This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2016.2570003, IEEE Transactions on Control of Network Systems

Optimal Linear Cyber-Attack on Remote State Estimation

Ziyang Guo, Dawei Shi, Karl Henrik Johansson, Ling Shi

***Abstract*—Recent years have witnessed the surge of interest of security issues in cyber-physical systems. In this paper, we consider malicious cyber attacks in a remote state estimation application where a smart sensor node transmits data to a remote estimator equipped with a false data detector. It is assumed that all the sensor data can be observed and modified by the malicious attacker and a residue-based detection algorithm is used at the remote side to detect data anomalies. We propose a linear deception attack strategy and present the corresponding feasibility constraint which guarantees that the attacker is able to successfully inject false data without being detected. The evolution of the estimation error covariance at the remote estimator is derived and the degradation of system performance under the proposed linear attack policy is analyzed. Furthermore, we obtain a closed-form expression of the optimal attack strategy among all linear attacks. Comparison of attack strategies through simulated examples are provided to illustrate the theoretical results.**

***Index Terms*—Cyber-Physical Systems, Deception Attack, Security, Remote State Estimation.**

I. Introduction

CYBER-Physical Systems (CPS) are systems that smoothly integrate sensing, communication, control, computation and physical processes [1]. CPS applications range from large-scale industrial applications to critical infrastructures including chemical processes, smart grids, mine monitoring, intelligent transportation, precision agriculture, civil engineering, aerospace, etc. [2]–[4].

The rapid growth of CPS and its safety-critical applications have generated a surge of interest in CPS security in recent years [5]. Since the measurement and control data in CPS are commonly transmitted through unprotected communication networks, such systems are vulnerable to cyber threats. Any successful CPS attack may lead to a variety of severe consequences, including customer information leakage, damages to national economy, destruction of infrastructure, and even endangering of human lives [6], [7].

Z. Guo and L. Shi are with Department of Electronic and Computer Engineering, the Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong (e-mail: zguoae@ust.hk, eesling@ust.hk).

D. Shi is with State Key Laboratory of Intelligent Control and Decision of Complex Systems, School of Automation, Beijing Institute of Technology, Beijing, 100081, China (e-mail: dawei.shi@outlook.com).

K. H. Johansson is with ACCESS Linnaeus Centre and Department of Automatic Control, School of Electrical Engineering, KTH Royal Institute of Technology, Stockholm, Sweden (e-mail: kallej@kth.se).

The work by Z. Guo and L. Shi is supported by a Hong Kong RGC GRF grant 16209114. The work by D. Shi is supported by Natural Science Foundation of China (61503027). The work by K. H. Johansson is supported by the Knut and Alice Wallenberg Foundation and the Swedish Research Council.

The cyber-physical attack space can be divided according to the adversary's system knowledge, disclosure resources and disruption resources. Attack models, such as Denial-of-Service (DoS), replay, false data injection and zero dynamic attacks were analyzed in [8]. Cardenas et al. [9] studied cyber attacks compromising measurement and actuator data integrity and availability. They considered two types of CPS attacks: DoS and deception attacks. The DoS attack, which jams the communication channels and prevents the exchange of information containing both sensor measurements and control inputs, was further analyzed for a resource-constrained attacker in [10], [11]. Moreover, a game-theoretic approach was utilized to provide an effective framework to handle security and privacy issues in communication networks in [12]. With energy constraints on both the sensor and the attacker, Li et al. [13] studied the interactive decision-making process of when to send and when to attack using a zero-sum game. They proved that the optimal strategies for both sides constitute a Nash equilibrium. Agah et al. [14] formulated a repeated game between the intrusion detector and the sensor nodes to study the prevention of DoS attack in wireless sensor networks. A framework to enforce cooperation among sensor nodes and punishment for non-cooperative behavior was proposed.

The deception attacks, which affect the integrity of data by modifying its content, have recently received attention. The replay attack is a special type of deception attack where the attacker does not have any system knowledge but is able to access, record, and replay the sensor data. Mo et al. [15], [16] studied the feasibility of the replay attack on a control system equipped with a bad-data detector and proposed a countermeasure to detect the existence of such an attack. Miao et al. [17] proposed a zero-sum stochastic game framework to balance the tradeoff between the control performance and the system security. Another type of deception attack with perfect system knowledge, false-data injection attack, was initially proposed for power networks [18]. Sandberg et al. [19] analyzed the minimum number of sensors required for a stealthy attack and proposed the concept of measurement security metric. A more general framework for security indices was provided in [20]. Furthermore, the consequence of the false-data injection attack and the reachable state estimation error have been analyzed in [21]. Besides the aforementioned studies where the models used are static, data injection attacks on dynamic control systems have also been considered. A covert data attack, which misleads the control center to remove useful measurements, was proposed and analyzed in [22]. Pasqualetti et al. [23] studied the set of undetectable false-data injection attacks for omniscient attackers who have full system

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2016.2570003, IEEE Transactions on Control of Network Systems

information but only compromise a part of existing sensors and actuators. A unified framework and advanced monitoring procedures to detect components malfunction or measurements corruption were also proposed. Further results on different formulations of integrity attack and secure estimation problems were investigated in [24], [25].

In this paper, we consider deception attacks in a remote state estimation scenario. We study the optimal linear deception attack on the sensor data without being detected by a false data detector at the remote state estimator. The motivation of the current work is three-fold:

1) A deception attack is subtler and may cause more severe consequences compared with many other attacks.
2) Existing models of deception attack are quite simple, many focusing on static parameter estimation [18]–[21]. The need for analyzing potential consequences of attacks on a dynamic system is important.
3) To propose effective countermeasures, one needs to understand what the worst attack might be.

The main contributions of this paper are summarized as follows:

1) We propose a novel type of linear attack strategy and present the corresponding feasibility constraint, which guarantees the attacker to successfully inject false data and remain undetected by the false data detector at the same time.
2) We compute the evolution of the estimation error covariance at the remote estimator and analyze the degradation of system performance under various linear attack strategies (**Theorem 1**).
3) We derive a closed-form expression of the optimal linear attack strategy which yields the largest error covariance (**Theorem 2**).

The remainder of the paper is organized as follows. Section II presents the problem formulation and revisits some preliminaries of the Kalman filter and the false data detector. Section III proposes a new type of deception attack strategy and states the feasibility constraint. Section IV illustrates the degradation of system performance and derives the optimal strategy among all linear attacks. Simulation results are provided in Section V. Some concluding remarks are given in the end.

Notations: $\mathbb{N}$ and $\mathbb{R}$ denote the sets of natural numbers and real numbers, respectively. $\mathbb{R}^n$ is the n-dimensional Euclidean space. $\mathbb{S}_+^n$ and $\mathbb{S}_{++}^n$ are the sets of $n \times n$ positive semi-definite and positive definite matrices. When $X \in \mathbb{S}_+^n$, we simply write $X \geq 0$ (or $X > 0$ if $X \in \mathbb{S}_{++}^n$). $X \geq Y$ if $X - Y \in \mathbb{S}_+^n$. $\mathcal{N}(\mu, \Sigma)$ denotes Gaussian distribution with mean μ and covariance matrix Σ. The superscript ^{T} stands for transposition. $\mathrm{tr}(\cdot)$ refers to the trace of a matrix. $\mathbb{E}[\cdot]$ denotes the expectation of a random variable. $\Pr\{\cdot\}$ denotes the probability of an event. For functions f, f_1, f_2 with appropriate domain, $f_1 \circ f_2(x)$ stands for the function composition $f_1(f_2(x))$, and $f^n(x) \triangleq f(f^{n-1}(x))$.

II. Problem Setup

The system architecture of cyber attacks in a remote state estimation application considered in this paper is shown in Fig. 1. It consists of six main components, namely the process, smart sensor, attacker, remote estimator, false data detector, and wireless network. The smart sensor performs local estimation based on the process measurements and transmits data packet to the remote estimator through a wireless network where a malicious attacker may intercept and modify the transmitted data. A false data detector at remote side monitors the system behavior and identifies the existence of the attacker. The detailed models are described in the following.

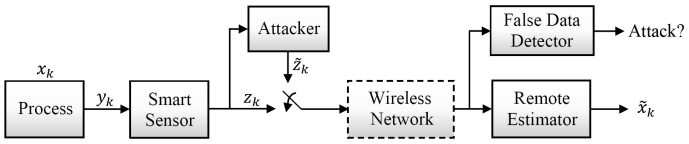


Fig. 1. System architecture. The attacker is able to intercept and modify sensor data, which affects the remote estimation performance despite the false data detector.

A. Process Model

Consider a discrete-time linear time-invariant process:

$$x_{k+1} = Ax_k + w_k, \tag{1}$$

$$y_k = Cx_k + v_k, \tag{2}$$

where $k \in \mathbb{N}$ is the time index, $x_k \in \mathbb{R}^n$ the vector of system states, $y_k \in \mathbb{R}^m$ the vector of sensor measurements, $w_k \in \mathbb{R}^n$ and $v_k \in \mathbb{R}^m$ are zero-mean i.i.d. Gaussian noises with covariances $Q \geq 0$ and $R > 0$, respectively. The initial state x_0 is zero-mean Gaussian with covariance matrix $\Pi_0 \geq 0$, and is independent of w_k and v_k for all $k \geq 0$. The pair (A, C) is detectable and $(A, \sqrt{Q})$ is stabilizable.

B. Smart Sensor and Remote Estimator

The concept of smart sensors refers to sensors that provide extra functions beyond those necessary for generating the measured quantity. The functions included might be signal processing, decision-making and alarm functions, which can be used to improve system performance [26], [27]. Thus, we assume that the smart sensor first locally processes the raw measurement data and transmits its innovation to the remote estimator in this work. To estimate the system state, the following standard Kalman filter is adopted by the remote estimator:

$$\hat{x}_k^- = A\hat{x}_{k-1}, \tag{3}$$

$$P_k^- = AP_{k-1}A^T + Q, \tag{4}$$

$$K_k = P_k^- C^T (CP_k^- C^T + R)^{-1}, \tag{5}$$

$$\hat{x}_k = \hat{x}_k^- + K_k z_k, \tag{6}$$

$$P_k = (I - K_k C)P_k^-, \tag{7}$$

where z_k is the local innovation transmitted to the remote estimator with

$$z_k = y_k - C\hat{x}_k^-, \tag{8}$$

$\hat{x}_k^-$ and $\hat{x}_k$ are the *a priori* and the *a posteriori* Minimum Mean Squared Error (MMSE) estimates of the state x_k at the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2016.2570003, IEEE Transactions on Control of Network Systems

remote estimator, and P_k^- and P_k are the corresponding error covariances. The recursion starts from $\hat{x}_0^- = 0$ and $P_0^- = \Pi_0 \geq 0$.

For notational brevity, we also define the Lyapunov and Riccati operators $h, \tilde{g} : \mathbb{S}_+^n \to \mathbb{S}_+^n$ as:

$$h(X) \triangleq AXA^T + Q, \tag{9}$$

$$\tilde{g}(X) \triangleq X - XC^T(CXC^T + R)^{-1}CX. \tag{10}$$

It is well known that the gain and the error covariance of the Kalman filter converge from any initial condition [28]. Hence, we denote the steady-state value of the *a priori* estimation error covariance as

$$\overline{P} = \lim_{k\to\infty} P_k^-,$$

where $\overline{P}$ is the unique positive semi-definite solution of $h \circ \tilde{g}(X) = X$.

To simplify our subsequent discussions, we assume that the Kalman filter at the remote estimator starts from the steady state, i.e., $\Pi_0 = \overline{P}$, which results in a steady-state Kalman filter with fixed gain

$$K = \overline{P}C^T(C\overline{P}C^T + R)^{-1}. \tag{11}$$

Remark 1 *Using the smart sensor instead of the conventional sensor not only improves measurement accuracy, but also reduces the computations at the remote estimator and improve communication efficiency [29]. Another reason why sending the innovation z_k rather than the measurement y_k or the local estimate $\hat{x}_k$ is that the innovation z_k will approach a steady-state distribution that can be easily checked by a false data detector. If y_k or $\hat{x}_k$ is sent instead, it is difficult to find an appropriate detector which can detect potential malicious attacks.*

C. False Data Detector

The innovation sequence z_k sent by the smart sensor is a white Gaussian process with zero mean and covariance $\mathscr{P}$, where $\mathscr{P} = C\overline{P}C^T + R$ [28]. The false data detector at the remote estimator side monitors the system behavior and detects cyber attacks by checking the statistical characteristic of the arriving innovation sequence. The mean and covariance of the innovations are used to diagnose the existence of potential cyber attacks.

The χ^2 detector is a residue-based detector widely used to reveal system anomalies [30], [31]. The detector makes a decision based on the sum of squared residues z_k which is normalized by the steady-state innovation covariance matrix $\mathscr{P}$. At time slot k, we suppose the detection criterion is given in the following form:

$$g_k = \sum_{i=k-J+1}^{k} z_i^T \mathscr{P}^{-1} z_i \underset{H_1}{\overset{H_0}{\lessgtr}} \delta, \tag{12}$$

where J is the window size of detection, δ is the threshold, the null hypotheses H_0 means that the system is operating normally, while the alternative hypotheses H_1 means that the system is under attack. The left hand side of (12) satisfies the χ^2 distribution with mJ degrees of freedom. Thus, it is easy to calculate the false alarm rate from the χ^2 distribution. If g_k is greater than the threshold, the detector triggers an alarm.

D. Problems of Interest

Based on the model of the process, the smart sensor, and the false data detector, the main problems we are interested in consist of the following:

1) What are the possible attack strategies under which the attacker remains undetectable to the false data detector?
2) What is the corresponding estimation error at the remote estimator under such an attack?
3) Does there exist an optimal attack strategy that renders maximum estimation error?

The detailed mathematical formulations and solutions to these problems will be introduced in the following two sections.

III. Linear Attack Strategy

In this section, we consider the existence of a malicious agent who intentionally launches cyber attacks to degrade the system performance. We will first define the attack policy and then analyze the feasibility constraint needed for such attack from being detected by the false data detector.

A. Linear Deception Attack

Similar to the attack models in existing works [32], [33] and the man-in-the-middle attack where the attacker has knowledge of all relevant messages passing between the two victims and can inject new ones [34], [35], we suppose that the attacker is able to intercept and modify the transmitted data. At each time k, the attack strategy is defined as

$$\tilde{z}_k = f_k(z_k) + b_k,$$

where z_k is the currently intercepted innovation, $\tilde{z}_k$ the innovation modified by the attacker, f_k an arbitrary function, $b_k \sim \mathcal{N}(0, \mathscr{L})$ an i.i.d. Gaussian random variable which is independent of z_k.

In this paper, we focus on the subset of all linear attack strategies where f_k is a linear transformation of the innovation z_k. We shall consider the general nonlinear attack strategies in the future work. The proposed linear attack strategy is defined as

$$\tilde{z}_k = T_k z_k + b_k, \tag{13}$$

where $T_k \in \mathbb{R}^{m\times m}$ is an arbitrary matrix. Since $z_k \sim \mathcal{N}(0, \mathscr{P})$, where $\mathscr{P} = C\overline{P}C^T + R$, it is easy to see that $\tilde{z}_k$ is also an i.i.d. Gaussian random variable with zero mean and variance $T_k\mathscr{P}T_k^T + \mathscr{L}$.

According to the detection criterion (12) of the false data detector, the detection rate of the proposed linear attack (13) is the same as without attack *if the modified innovation $\tilde{z}_k$ preserves the same statistical characteristic as z_k*. In other words, to bypass the false data detector, $\tilde{z}_k$ is supposed to satisfy the Gaussian distribution $\mathcal{N}(0, \mathscr{P})$, i.e., have zero mean and covariance $\mathscr{P}$. Hence,

$$T_k\mathscr{P}T_k^T + \mathscr{L} = \mathscr{P}.$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2016.2570003, IEEE Transactions on Control of Network Systems

Consequently, it must hold that

$$\mathscr{P} - T_k \mathscr{P} T_k^T = \mathscr{L} \geq 0. \tag{14}$$

Remark 2 *In principle, the attacker has knowledge of all the past innovations such that it can design the linear attack strategy in the form of* $\tilde{z}_k = f(z_0, z_1, \ldots, z_k) + b_k = \sum_{i=0}^{k} T_i z_i + b_k$. *However, in order to bypass the false data detector, the modified innovation* $\tilde{z}_k$ *needs to preserve the same Gaussian distribution* $\mathcal{N}(0, \mathscr{P})$ *as* z_k. *The zero mean condition forces* $T_0 = T_1 = \cdots = T_{k-1} = 0$ *when choosing* $\tilde{z}_k$. *Hence, it is reasonable that the proposed linear attack strategy* $\tilde{z}_k$ *only depends on current innovation* z_k.

B. Attacker with Perfect System Information

In the case that the attacker has perfect system information, it can accurately calculate $\overline{P}$ by solving the Riccati equation $h \circ \tilde{g}(X) = X$ based on the system parameters A, C, Q, R. Then for each time slot k, the attack strategy that remains undetected by the false data detector can be easily generated by firstly selecting any matrix $T_k \in \mathbb{R}^{m \times m}$ which satisfies $\mathscr{P} - T_k \mathscr{P} T_k^T \geq 0$, and then selecting $\mathscr{L} = \mathscr{P} - T_k \mathscr{P} T_k^T \geq 0$.

C. Attacker with No System Information

In the case that the attacker does not have any system information, it needs to estimate the mean and the variance of the innovation z_k before launching the attack. In statistics, interval estimation is the use of sample data to calculate an interval of probable values of an unknown population parameter [36]. Suppose $\{X_1, X_2, \ldots, X_n\}$ is an independent sample from a normally distributed population with mean μ and variance σ^2. Let

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i,$$
$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2,$$

where $\bar{X}$ is the sample mean, and S^2 the sample variance. In order to estimate the mean of the population, one has

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1),$$

where $t(n-1)$ represents a Student's $t-$distribution with $n-1$ degrees of freedom. For a given confidence level $1 - \alpha$, we obtain

$$\Pr\left\{ \bar{X} - t_{\frac{\alpha}{2}}(n-1) \frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{\frac{\alpha}{2}}(n-1) \frac{S}{\sqrt{n}} \right\} = 1 - \alpha, \tag{15}$$

which means that the value of the estimated parameter μ falls into the confidence interval with probability $1 - \alpha$.

Similarly, to estimate the variance of the population, we have

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1),$$

where $\chi^2(n-1)$ represents a chi-squared distribution with $n-1$ degrees of freedom. For a given confidence level $1 - \beta$, the probability that the estimated value is between two stochastic endpoints is shown as follows:

$$\Pr\left\{ \frac{(n-1)S^2}{\chi^2_{\frac{\beta}{2}}(n-1)} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\frac{\beta}{2}}(n-1)} \right\} = 1 - \beta. \tag{16}$$

From (15) and (16), it is not difficult to see when n is sufficiently large, a small confidence interval with a high confidence level can be obtained. Therefore, even though the attacker may not know anything about the system parameters, it can still launch the proposed linear attack without being detected after a certain period of time, during which it can successfully estimate the mean and the variance of the innovation z_k.

Based on the attack strategy, the problem we are interested in is to find the largest degradation of the system performance at the remote estimator under the proposed linear attack. To quantify the estimation performance, we define $\tilde{x}_k^-$ and $\tilde{x}_k$ as the *a priori* and the *a posteriori* MMSE estimates of the state x_k at the remote estimator when the system is under attack, and $\tilde{P}_k^-$ and $\tilde{P}_k$ as the corresponding error covariances. Note that the linear attack is able to start at any time $k \in \mathbb{N}$. Without loss of generality, we assume that the attack starts at $k = 1$ and investigate the error covariance iteration at the remote side. Due to the assumption that the remote estimator starts from the steady state, we obtain the initial conditions $\tilde{x}_0^- = \hat{x}_0^-$ and $\mathbb{E}[(x_0 - \tilde{x}_0^-)(x_0 - \tilde{x}_0^-)'] = \mathbb{E}[(x_0 - \hat{x}_0^-)(x_0 - \hat{x}_0^-)'] = \overline{P}$.

IV. Performance Analysis

We consider the system under the linear attack (13) with the feasibility constraint (14) in this section. First, we derive the evolution of the estimation error covariance at the remote estimator during an attack, which quantifies the system performance degradation. Second, we derive the error covariance iteration and optimal attack strategy for scalar systems. Then, we formulate the problem of finding the optimal attacker as a convex optimization problem and apply semi-definite programming (SDP) to find a numerical solution. Finally, we prove that the optimal attack strategy is obtained when $T_k = -I$, i.e., when the attacker flips the sign of all the innovation.

A. Error Covariance Evolution

Consider the process (1)–(2) under the proposed linear attack $\tilde{z}_k = T_k z_k + b_k$. The state estimate of the remote estimator follows

$$\tilde{x}_k^- = A \tilde{x}_{k-1}, \tag{17}$$
$$\tilde{x}_k = \tilde{x}_k^- + K \tilde{z}_k, \tag{18}$$

where the fixed gain K is given in (11).

Since the false data detector cannot detect any anomaly if the linear attack strategy (13) satisfies the feasibility constraint (14), the state estimate $\tilde{x}_k$ produced by the remote estimator will deviate from the true system state. The following theorem

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2016.2570003, IEEE Transactions on Control of Network Systems

summarizes the evolution of the estimation error covariance under such an attack.

Theorem 1 *For the system in Fig. 1, the linear attack* (13) *satisfying the feasibility constraint* (14) *remains undetected by the false data detector. Moreover, the estimation error covariance at the remote estimator follows the recursion*

$$\tilde{P}_k = A\tilde{P}_{k-1}A' + Q + \overline{P}C^T(\Sigma - T_k^T\Sigma - \Sigma T_k)C\overline{P}, \quad (19)$$

where $\Sigma = (C\overline{P}C^T + R)^{-1}$.

Proof: The stealthiness of the proposed linear attack strategy follows from the previous reasoning.

According to the process model (1)–(2) and the iteration of state estimate (17)–(18), one has

$$\begin{aligned} x_k - \tilde{x}_k^- &= A(x_{k-1} - \tilde{x}_{k-1}) + w_{k-1}, \\ x_k - \tilde{x}_k &= x_k - \tilde{x}_k^- - K\tilde{z}_k, \end{aligned}$$

from which the error covariance at the remote estimator side can be obtained as

$$\begin{aligned} \tilde{P}_k^- &= \mathbb{E}[(x_k - \tilde{x}_k^-)(x_k - \tilde{x}_k^-)^T] \\ &= A\tilde{P}_{k-1}A^T + Q, \\ \tilde{P}_k &= \mathbb{E}[(x_k - \tilde{x}_k)(x_k - \tilde{x}_k)^T] \\ &= \tilde{P}_k^- + K(C\overline{P}C^T + R)K^T \\ &\quad - \mathbb{E}[(x_k - \tilde{x}_k^-)\tilde{z}_k^T K^T] - \mathbb{E}[K\tilde{z}_k(x_k - \tilde{x}_k^-)^T]. \end{aligned} \quad (20)$$

To calculate the last two terms of (20), we first evaluate

$$\begin{aligned} & x_k - \tilde{x}_k^- \\ &= Ax_{k-1} + w_{k-1} - A(\tilde{x}_{k-1}^- + K\tilde{z}_{k-1}) \\ &= A^k x_0 + \sum_{i=0}^{k-1} A^i w_{k-1-i} - A^k\tilde{x}_0^- - \sum_{i=0}^{k-1} A^{i+1}K\tilde{z}_{k-1-i} \\ &= A^k(x_0 - \hat{x}_0^-) + \sum_{i=0}^{k-1} A^i w_{k-1-i} - \sum_{i=0}^{k-1} A^{i+1}K\tilde{z}_{k-1-i}, \end{aligned} \quad (21)$$

where the last equality follows from the steady-state assumption $\tilde{x}_0^- = \hat{x}_0^-$. Since $\tilde{z}_k$ is an i.i.d. Gaussian random variable, we obtain that $\mathbb{E}[\tilde{z}_i\tilde{z}_j^T] = 0$, $\forall i \neq j$. Thus, we are only concerned of the correlation between the first two terms of (21) and $\tilde{z}_k$. Then, based on

$$\begin{aligned} & x_k - \hat{x}_k^- \\ &= Ax_{k-1} + w_{k-1} - A(\hat{x}_{k-1}^- + Kz_{k-1}) \\ &= Ax_{k-1} + w_{k-1} - A[\hat{x}_{k-1}^- + K(C(x_{k-1} - \hat{x}_{k-1}^-) + v_{k-1}))] \\ &= A(I - KC)(x_{k-1} - \hat{x}_{k-1}^-) + w_{k-1} - AKv_{k-1}, \end{aligned} \quad (22)$$

we can further represent $\tilde{z}_k$ in the form of

$$\begin{aligned} \tilde{z}_k &= T_k z_k + b_k \\ &= T_k C(x_k - \hat{x}_k^-) + T_k v_k + b_k \\ &= T_k CA(I - KC)(x_{k-1} - \hat{x}_{k-1}^-) + T_k C w_{k-1} \\ &\quad - T_k CAKv_{k-1} + T_k v_k + b_k \\ &= T_k C[A(I - KC)]^k(x_0 - \hat{x}_0^-) \\ &\quad + \sum_{i=0}^{k-1} T_k C[A(I - KC)]^i w_{k-1-i} + V, \end{aligned} \quad (23)$$

where $V = T_k v_k + b_k - \sum_{i=0}^{k-1} T_k C[A(I - KC)]^i AKv_{k-1-i}$ is independent of the first two terms of (21). It now follows that the second last term of (20) can be written as

$$\begin{aligned} & \mathbb{E}\left[(x_k - \tilde{x}_k^-)\tilde{z}_k^T K^T\right] \\ &= \mathbb{E}\left[\left\{A^k(x_0 - \hat{x}_0^-) + \sum_{i=0}^{k-1} A^i w_{k-1-i}\right\}\left\{T_k C\left[A(I - KC)\right]^k\right.\right. \\ &\quad \left.\left.(x_0 - \hat{x}_0^-) + \sum_{i=0}^{k-1} T_k C\left[A(I - KC)\right]^i w_{k-1-i}\right\}^T K^T\right] \\ &= \left\{A^k\mathbb{E}\left[(x_0 - \hat{x}_0^-)(x_0 - \hat{x}_0^-)^T\right]\left[(I - KC)^T A^T\right]^k\right. \\ &\quad \left. + \sum_{i=0}^{k-1} A^i\mathbb{E}\left[w_{k-1-i}w_{k-1-i}^T\right]\left[(I - KC)^T A^T\right]^i\right\}C^T T_k^T K^T \\ &= \left\{A^k\overline{P}\left[(I - KC)^T A^T\right]^k\right. \\ &\quad \left. + \sum_{i=0}^{k-1} A^i Q\left[(I - KC)^T A^T\right]^i\right\}C^T T_k^T K^T \\ &= \overline{P}C^T T_k^T K^T, \end{aligned} \quad (24)$$

where the last equality is due to the fact that $\overline{P}$ is the unique positive semi-definite fixed point of $h \circ \tilde{g}$, i.e.,

$$\begin{aligned} \overline{P} &= (h \circ \tilde{g})^n(\overline{P}) \\ &= [A(I - KC)]^n\overline{P}(A^T)^n + \sum_{i=0}^{n-1}[A(I - KC)]^i Q(A^T)^i \\ &= A^n\overline{P}[(I - KC)^T A^T]^n + \sum_{i=0}^{n-1} A^i Q[(I - KC)^T A^T]^i. \end{aligned}$$

Similarly, we obtain

$$\mathbb{E}\left[K\tilde{z}_k(x_k - \tilde{x}_k^-)^T\right] = KT_k C\overline{P}. \quad (25)$$

Substituting (24) and (25) into (20), the error covariance at the remote estimator is given by

$$\begin{aligned} \tilde{P}_k &= \tilde{P}_k^- + \overline{P}C^T(C\overline{P}C^T + R)^{-1}C\overline{P} \\ &\quad - \overline{P}C^T T_k^T(C\overline{P}C^T + R)^{-1}C\overline{P} \\ &\quad - \overline{P}C^T(C\overline{P}C^T + R)^{-1}T_k C \\ &= A\tilde{P}_{k-1}A^T + Q + \overline{P}C^T(\Sigma - T_k^T\Sigma - \Sigma T_k)C\overline{P}, \end{aligned}$$

where $\Sigma = (C\overline{P}C^T + R)^{-1} > 0$. ■

Remark 3 *The obtained iteration of the remote estimation error covariance* (19) *when the system is under linear attack is based on the steady-state assumption, i.e.,* $\tilde{x}_0^- = \hat{x}_0^-$ *and* $\tilde{P}_0^- = \mathbb{E}[(x_0 - \hat{x}_0^-)(x_0 - \hat{x}_0^-)'] = \overline{P}$. *Otherwise, the error covariance recursion at the remote estimator depends on system initial state when the attack begins. However, the analytical method and the obtained result are the same.*

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2016.2570003, IEEE Transactions on Control of Network Systems

B. Scalar Systems

For processes with scalar outputs ($m = 1$), T_k is a scalar, so the linear attack strategy and the corresponding feasibility constraint become

$$\tilde{z}_k = T_k z_k + b_k, \tag{26}$$
$$\mathscr{L} = \mathscr{P} - T_k^2 \mathscr{P} \geq 0. \tag{27}$$

Hence, the parameters of the linear attack strategy should be chosen as $T_k \in [-1, 1]$ and $\mathscr{L} \in [0, \mathscr{P}]$. According to (19), the error covariance at the remote estimator is

$$\tilde{P}_k = A\tilde{P}_{k-1}A^T + Q + (1 - 2T_k)\triangle, \tag{28}$$

where $\triangle = \overline{P}C^T(C\overline{P}C^T + R)^{-1}C\overline{P}$. Then the optimal linear attack strategy which yields the largest error covariance is obtained when $T_k = -1$ and $\mathscr{L} = 0$, i.e., when $\tilde{z}_k = -z_k$.

It is worth noticing that there are some interesting special cases for the attack strategy (26):

1) $\tilde{z}_k = z_k$: When $T_k = 1$, $b_k = 0$, the error covariance is recursively given as

$$\tilde{P}_k = A\tilde{P}_{k-1}A^T + Q - \triangle$$

with the initial state $\tilde{P}_0 = (I - KC)\overline{P} = \overline{P} - \Delta$. It produces the same result with the steady-state Kalman filter, which can be represented as $P_k = (I - KC)(AP_{k-1}A^T + Q) = (I - KC)\overline{P} = \overline{P} - \Delta$.

2) $\tilde{z}_k = -z_k$: When $T_k = -1$, $b_k = 0$, the attacker launches the attack $\tilde{z}_k = -z_k$ and the corresponding error covariance is

$$\tilde{P}_k = A\tilde{P}_{k-1}A^T + Q + 3\triangle.$$

3) $\tilde{z}_k \sim \mathcal{N}(0, C\overline{P}C^T + R)$: When $T_k = 0$, $b_k \sim \mathcal{N}(0, C\overline{P}C^T + R)$, the attacker generates i.i.d. Gaussian noise as an attack and the error covariance is

$$\tilde{P}_k = A\tilde{P}_{k-1}A^T + Q + \triangle.$$

All the estimation error covariance iterates above converge if the system is stable.

In the case that the remote estimator uses no data to update its estimate, the error covariance is

$$\tilde{P}_k = A\tilde{P}_{k-1}A^T + Q.$$

Hence, the optimal attack strategy $T_k = -1$ degrades the system performance by adding $3\triangle$ to the iteration of the error covariance.

The optimal attack strategy for the multiple-output case ($m > 1$) is more difficult to derive. The scalar case suggests the conjecture $T_k = -I$. This motivates our investigations in the following two subsections, where we show that it is indeed true.

C. Numerical Solution of the Optimal Attack Strategy

The remote estimation error covariance under the linear attack $\tilde{z}_k = T_k z_k + b_k$ can be represented as $P_k(T_k)$ according to (19), based on which we define the optimal attack strategy T_k^* as the one that yields the largest estimation error covariance, i.e., for any other T_k, $\tilde{P}_k(T_k^*) - \tilde{P}_k(T_k)$ is positive semi-definite. Then the following proposition uses the trace of the error covariance matrix as a metric to quantify the attack effect formulates the problem of finding an optimal attack strategy as a convex optimization problem, from which a numerical solution can be easily obtained.

Proposition 1 *The optimal attack strategy is given by the solution of the convex optimization problem*

$$\begin{aligned} \min_{T_k \in \mathbb{R}^{m\times m}} \quad & \operatorname{tr}(C\overline{P}\overline{P}C^T\Sigma T_k) \\ s.t. \quad & \begin{bmatrix} \mathscr{P} & T_k \\ T_k^T & \mathscr{P}^{-1} \end{bmatrix} \geq 0. \end{aligned} \tag{29}$$

Proof: According to the iteration of the error covariance (19) under linear deception attack, we obtain that

$$\begin{aligned} \operatorname{tr}(\tilde{P}_k) &= \operatorname{tr}(A\tilde{P}_{k-1}A^T + Q + \overline{P}C^T(\Sigma - T_k^T\Sigma - \Sigma T_k)C\overline{P}) \\ &= \operatorname{tr}(A^k\overline{P}[A^T]^k) + \sum_{i=0}^{k-1} \operatorname{tr}(A^iQ[A^T]^i) \\ &\quad + \sum_{i=0}^{k} \operatorname{tr}(A^i(\overline{P}C^T(\Sigma - T_i^T\Sigma - \Sigma T_i)C\overline{P})[A^T]^i). \end{aligned}$$

It can be observed that at time k, for any given system, maximizing the trace of the error covariance matrix is equivalent to maximizing the trace of the last term of (19). Since $\operatorname{tr}(A + B) = \operatorname{tr}(A) + \operatorname{tr}(B)$, $\operatorname{tr}(ABCD) = \operatorname{tr}(DABC)$ and $\operatorname{tr}(X^TY) = \operatorname{tr}(XY^T)$, one has

$$\begin{aligned} &\operatorname{tr}(\overline{P}C^T[\Sigma - T_k^T\Sigma - \Sigma T_k]C\overline{P}) \\ &= \operatorname{tr}(\overline{P}C^T\Sigma C\overline{P}) - \operatorname{tr}(\overline{P}C^TT_k^T\Sigma C\overline{P}) - \operatorname{tr}(\overline{P}C^T\Sigma T_kC\overline{P}) \\ &= \operatorname{tr}(\overline{P}C^T\Sigma C\overline{P}) - \operatorname{tr}(\Sigma C\overline{P}\overline{P}C^TT_k^T) - \operatorname{tr}(C\overline{P}\overline{P}C^T\Sigma T_k) \\ &= \operatorname{tr}(\overline{P}C^T\Sigma C\overline{P}) - 2\operatorname{tr}(C\overline{P}\overline{P}C^T\Sigma T_k), \end{aligned} \tag{30}$$

where Σ and P are semi-definite matrices.

Ignoring the constant term of (30), the problem of finding the optimal attack strategy is equivalent to solving the optimization problem

$$\begin{aligned} \max_{T_k \in \mathbb{R}^{m\times m}} \quad & -\operatorname{tr}(C\overline{P}\overline{P}C^T\Sigma T_k) \\ s.t. \quad & T_k\mathscr{P}T_k^T - \mathscr{P} \leq 0. \end{aligned}$$

To solve this optimization problem, we use Schur complement to change the constraint to a linear matrix inequality:

$$\begin{aligned} \min_{T_k \in \mathbb{R}^{m\times m}} \quad & \operatorname{tr}(C\overline{P}\overline{P}C^T\Sigma T_k) \\ s.t. \quad & \begin{bmatrix} \mathscr{P} & T_k \\ T_k^T & \mathscr{P}^{-1} \end{bmatrix} \geq 0. \end{aligned}$$

■

Remark 4 *Using the CVX toolbox [37] in MATLAB to solve the optimization problem given in* (29)*, we can find a numerical solution based on SDP.*

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2016.2570003, IEEE Transactions on Control of Network Systems

D. Optimal Attack Strategy

Based on the conjecture of the optimal attack strategy from the scalar case and the numerical solutions, we aim to find out a closed-loop expression of the optimal T_k. The main result is summarized in the following theorem.

Theorem 2 *For the system in Fig. 1 with the linear attack* (13)*, $T_k = -I$ and $b_k = 0$ is the optimal linear attack strategy in the sense that it yields the largest estimation error covariance.*

Proof: The iteration of estimation error covariance at the remote estimator under the linear attack $\tilde{z}_k = T_k z_k + b_k$ is given by (19). Obviously, the optimal attack strategy which maximizes $\tilde{P}_k$ is equivalent to the strategy which maximizes $\overline{P}C^T(\Sigma - T_k^T\Sigma - \Sigma T_k)C\overline{P}$. We then derive the optimal attack strategy based on the correspondence between the optimal attack and the optimal estimate.

According to the attack strategy (13) and the iteration equation 18, one has

$$\begin{aligned}\tilde{x}_k &= \tilde{x}_k^- + K\tilde{z}_k \\ &= \tilde{x}_k^- + K(T_k z_k + b_k) \\ &= \tilde{x}_k^- + K[T_k + b_k(z_k^T z_k)^{-1} z_k^T] z_k \\ &= \tilde{x}_k^- + \tilde{K}_k z_k, \end{aligned} \tag{31}$$

where $\tilde{K}_k = K[T_k + b_k(z_k^T z_k)^{-1} z_k^T]$. Note that the state estimate at time k is a linear combination of all the past innovations z_i, $i \in \{1, 2, \ldots, k\}$. Due to the orthogonality between z_i and z_j, i.e., $\mathbb{E}[z_i z_j] = 0$, $\forall i \neq j$, whether there exists malicious attacks during the past time instants or not, the optimal state estimate at time k which minimizes the remote estimation error covariance is obtained when $\tilde{K}_k = K$. This corresponds to the estimation error covariance when $T_k = I$ and $b_k = 0$, i.e.,

$$\tilde{P}_k = A\tilde{P}_{k-1}A^T + Q - \overline{P}C^T\Sigma C\overline{P}.$$

It other words, $T_k = I$ yields the smallest error covariance at time k among all attacks given by (13).

Hence, if we denote $T_{k1} = I$ and note that

$$\Sigma - T_{k1}^T\Sigma - \Sigma T_{k1} = -\Sigma,$$

for any $T_{k2} = T_{k1} + M$, where M is an arbitrary matrix satisfying the constraint

$$T_{k2}\mathscr{P}T_{k2}^T = (I + M)\mathscr{P}(I + M)^T \leq \mathscr{P}, \tag{32}$$

the following inequality

$$\begin{aligned}&\overline{P}C^T[\Sigma - T_{k2}^T\Sigma - \Sigma T_{k2} - (\Sigma - T_{k1}^T\Sigma - \Sigma T_{k1})]C\overline{P} \\ &= \overline{P}C^T[\Sigma - (I + M)^T\Sigma - \Sigma(I + M) + \Sigma]C\overline{P} \\ &= \overline{P}C^T[-M^T\Sigma - \Sigma M]C\overline{P} \\ &\geq 0 \end{aligned} \tag{33}$$

must be true since $T_k = I$ is the optimal estimate which yields the smallest error covariance.

Now we obtain a one-to-one correspondence in finding the optimal attack strategy. Denote $T_{k3} = -I$ and note that

$$\Sigma - T_{k3}^T\Sigma - \Sigma T_{k3} = 3\Sigma.$$

For any $T_{k4} = T_{k3} - M$, where M is an arbitrary matrix satisfying the constraint

$$\begin{aligned}T_{k4}\mathscr{P}T_{k4}^T &= (-I - M)\mathscr{P}(-I - M)^T \\ &= (I + M)\mathscr{P}(I + M)^T \leq \mathscr{P}, \end{aligned} \tag{34}$$

we obtain that

$$\begin{aligned}&\overline{P}C^T[\Sigma - T_{k3}^T\Sigma - \Sigma T_{k3} - (\Sigma - T_{k4}^T\Sigma - \Sigma T_{k4})]C\overline{P} \\ &= \overline{P}C^T[3\Sigma - \Sigma + (-I - M)^T\Sigma + \Sigma(-I - M)]C\overline{P} \\ &= \overline{P}C^T[-M^T\Sigma - \Sigma M]C\overline{P} \\ &\geq 0 \end{aligned} \tag{35}$$

from (33). This means that for any T_k matrix different from $-I$, the difference between the two error covariances is always positive semi-definite. Therefore, $T_k = -I$ is the optimal attack that yields the largest error covariance. ■

V. Simulation Examples

To demonstrate the analytical results, we present some numerical simulations in this section. We compare the attack strategies $\tilde{z}_k = -z_k$ and $\tilde{z}_k \sim \mathcal{N}(0, C\overline{P}C^T + R)$ with the extreme case where the remote estimator does not use any data to update its state estimate. We consider a stable process with parameters $A = 0.8, C = 1.2, Q = 1, R = 1$ and an unstable process with parameters $A = 1.02, C = 1.2, Q = 1, R = 1$.

A. Stable Process under Linear Attack

For the stable process, the simulation results of the remote state estimate and estimation error covariance under different attacks are shown in Fig. 2 and Fig. 3, respectively. During time periods $[0, 20]$, $[40, 60]$ and $[80, 100]$, the remote estimator runs a Kalman filter and enters the steady state. The attacker uses the false data $\tilde{z}_k = -z_k$ during the time period $[60, 80]$ and randomly generates an i.i.d. Gaussian innovation $\tilde{z}_k \sim \mathcal{N}(0, CPC^T + R)$ during the time period $[100, 120]$ to launch cyber attack. Since the false data detector cannot successfully detect the existence of the attack, the remote estimator still uses the Kalman filter to update its state estimate and the corresponding error covariance, which are the green dash-dot lines in Fig. 2 and Fig. 3. However, the real state estimate deviates from the true state and leads to large estimation error, which can be seen by the red dashed lines in Fig. 2 and Fig. 3. To compare different attack effects, the extreme case where the remote estimator uses no data to update its state estimate is shown in Fig. 2 and Fig. 3 during the time period $[20, 40]$. It can be observed that the error covariance under the $\tilde{z}_k = -z_k$ attack is larger than that under the $\tilde{z}_k \sim \mathcal{N}(0, CPC^T + R)$ attack, and the latter is larger than using no data. Hence, using bad data is worse than using no data and $T_k = -I$ is the optimal linear deception attack strategy. Moreover, the error covariance converges when the process is stable.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2016.2570003, IEEE Transactions on Control of Network Systems

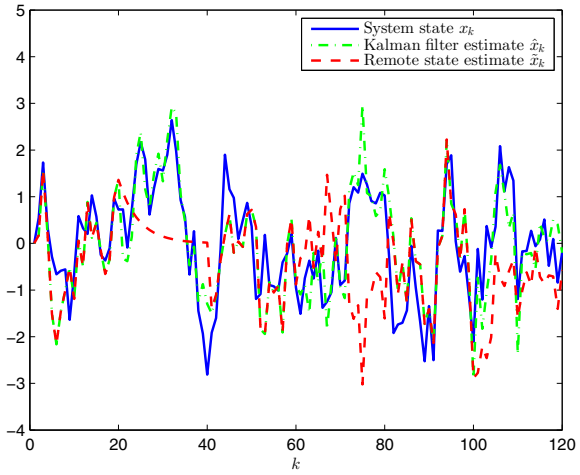


Fig. 2. Remote state estimate for stable process.

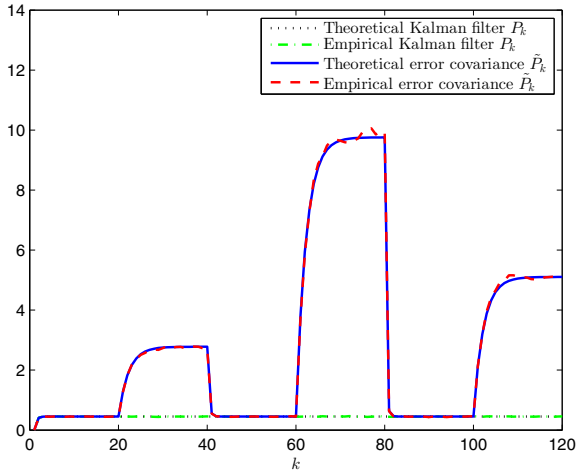


Fig. 3. Remote estimation error covariance for stable process.

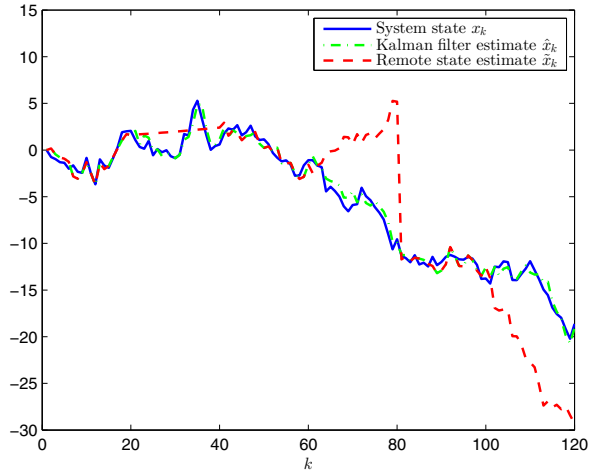


Fig. 4. Remote state estimate for unstable process.

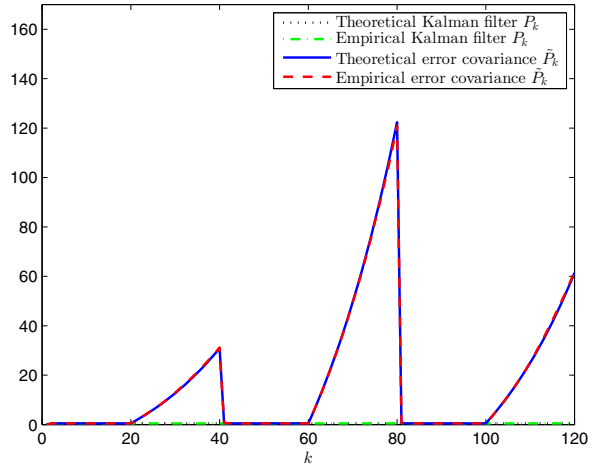


Fig. 5. Remote estimation error covariance for unstable process.

B. Unstable Process under Linear Attack

Fig. 4 and Fig. 5 represent the state estimate and the error covariance of the remote estimator for the unstable process. The attacker launches a cyber attack using $\tilde{z}_k = -z_k$ and $\tilde{z}_k \sim \mathcal{N}(0, CPC^T + R)$ at $k = 60$ and $k = 100$, respectively. The remote estimator believes that it acts as a Kalman filter and tracks the system state with a small error because the false data detector cannot detect any system anomaly. In reality, however, the real state estimate is quite different from the true state, which leads to the divergence of the error covariance, which are shown by the red dashed lines in Fig. 4 and Fig. 5. Compared with the case where the remote estimator uses no data to update its state estimate during the time period $[20, 40]$, all the error covariances diverge exponentially fast for the unstable process.

VI. Conclusion

In this paper, we proposed a novel linear attack strategy on remote state estimation and analyzed the corresponding feasibility constraint to ensure that the attack can successfully bypass a χ^2 false data detector. We investigated the evolution of the remote estimation error covariance under the attack and analyzed the degradation of system performance. Furthermore, we proved that $T_k = -I$ is the optimal among all linear attacks. Simulation and comparison were provided to demonstrate the analytical results. Future work includes the analysis of system performance under other types of attack strategies and the development of detection criterion to prevent these attacks.

References

[1] R. Poovendran, K. Sampigethaya, S. K. S. Gupta, I. Lee, K. V. Prasad, D. Corman, and J. Paunicka, "Special issue on cyber-physical systems," in *Proceedings of the IEEE*, vol. 100, no. 1, 2012, pp. 1–12.

[2] S. H. Ahmed, G. Kim, and D. Kim, "Cyber physical system: Architecture, applications and research challenges," in *Wireless Days, IFIP*, 2013, pp. 1–5.

[3] K. D. Kim and P. R. Kumar, "Cyber-physical systems: A perspective at the centennial," *Proceedings of the IEEE*, vol. 100, no. Special Centennial Issue, pp. 1287–1308, 2012.

[4] E. A. Lee, "Cyber physical systems: Design challenges," in *11th IEEE International Symposium on Object Oriented Real-Time Distributed Computing*, 2008, pp. 363–369.

[5] H. Sandberg, S. Amin, and K. H. Johansson, "Cyberphysical security in networked control systems: An introduction to the issue," *IEEE Control Systems Magazine*, vol. 35, no. 1, pp. 20–23, 2015.

[6] Y. Mo, T. H. Kim, K. Brancik, D. Dickinson, H. Lee, A. Perrig, and B. Sinopoli, "Cyber-physical security of a smart grid infrastructure," *Proceedings of the IEEE*, vol. 100, no. 1, pp. 195–209, 2012.

[7] A. A. Cardenas, S. Amin, and S. Sastry, "Research challenges for the security of control systems." in *HotSec: Procceedings of 3rd Conference in Hot Toptics in Secuity*. Berkeley, CA, USA, 2008, pp. 1–6.

[8] A. Teixeira, D. Pérez, H. Sandberg, and K. H. Johansson, "Attack models and scenarios for networked control systems," in *Proceedings of the 1st International Conference on High Confidence Networked Systems*. ACM, 2012, pp. 55–64.

[9] A. A. Cardenas, S. Amin, and S. Sastry, "Secure control: Towards survivable cyber-physical systems," *28th International Conference on Distributed Computing Workshops*, pp. 495–500, 2008.

[10] S. Amin, A. A. Cárdenas, and S. S. Sastry, "Safe and secure networked control systems under denial-of-service attacks," in *Hybrid Systems: Computation and Control*. Springer, 2009, pp. 31–45.

[11] A. Gupta, C. Langbort, and T. Basar, "Optimal control in the presence of an intelligent jammer with limited actions," in *49th IEEE Conference on Decision and Control*, 2010, pp. 1096–1101.

[12] M. H. Manshaei, Q. Zhu, T. Alpcan, T. Bacşar, and J. P. Hubaux, "Game theory meets network security and privacy," *ACM Computing Surveys*, vol. 45, no. 3, pp. 25:1–39, 2013.

[13] Y. Li, L. Shi, P. Cheng, J. Chen, and D. E. Quevedo, "Jamming attacks on remote state estimation in cyber-physical systems: A game-theoretic approach," *IEEE Transactions on Automatic Control*, vol. 60, no. 10, pp. 2831–2836, 2015.

[14] A. Agah and S. K. Das, "Preventing dos attacks in wireless sensor networks: A repeated game theory approach," *International Journal of Network Security*, vol. 5, no. 2, pp. 145–153, 2007.

[15] Y. Mo and B. Sinopoli, "Secure control against replay attacks," in *47th Annual Allerton Conference on Communication, Control, and Computing*, 2009, pp. 911–918.

[16] ——, "Integrity attacks on cyber-physical systems," in *Proceedings of the 1st International Conference on High Confidence Networked Systems*. ACM, 2012, pp. 47–54.

[17] F. Miao, M. Pajic, and G. J. Pappas, "Stochastic game approach for replay attack detection," in *52nd Annual Conference on Decision and Control*, 2013, pp. 1854–1859.

[18] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Transactions on Information and System Security*, vol. 14, no. 1, pp. 13–24, 2011.

[19] H. Sandberg, A. Teixeira, and K. H. Johansson, "On security indices for state estimators in power networks," in *First Workshop on Secure Control Systems, Stockholm, Sweden*, 2010.

[20] A. Teixeira, K. C. Sou, H. Sandberg, and K. H. Johansson, "Secure control systems: A quantitative risk management approach," *IEEE Control Systems Magazine*, vol. 35, no. 1, pp. 24–45, 2015.

[21] L. Xie, Y. Mo, and B. Sinopoli, "Integrity data attacks in power market operations," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 659–666, 2011.

[22] J. Kim, L. Tong, and R. J. Thomas, "Subspace methods for data attack on state estimation: A data driven approach," *IEEE Transactions on Signal Processing*, vol. 63, no. 5, pp. 1102–1114, 2014.

[23] F. Pasqualetti, F. Dorfler, and F. Bullo, "Cyber-physical attacks in power networks: Models, fundamental limitations and monitor design," in *50th IEEE Conference on Decision and Control and European Control Conference*, 2011, pp. 2195–2201.

[24] D. Shi, T. Chen, and M. Darouach, "Event-based state estimation of linear dynamic systems with unknown exogenous inputs," *Automatica*, vol. 69, pp. 275–288, 2016.

[25] D. Shi, R. J. Elliott, and T. Chen, "On finite-state stochastic modeling and secure estimation of cyber-physical systems," *IEEE Transactions on Automatic Control*, vol. PP, no. 99, p. 1, 2016.

[26] F. L. Lewis *et al.*, "Wireless sensor networks," *Smart environments: technologies, protocols, and applications*, pp. 11–46, 2004.

[27] R. Frank, *Understanding smart sensors*. Artech House, 2013.

[28] B. D. O. Anderson and J. B. Moore, *Optimal filtering*. Dover Publications, Mineola, N.Y, 2005.

[29] J. Favenne, "Smart sensors in industry," *Journal of Physics E: Scientific Instruments*, vol. 20, no. 9, pp. 1087–1090, 1987.

[30] R. K. Mehra and J. Peschon, "An innovations approach to fault detection and diagnosis in dynamic systems," *Automatica*, vol. 7, no. 5, pp. 637–640, 1971.

[31] A. S. Willsky, "A survey of design methods for failure detection in dynamic systems," *Automatica*, vol. 12, no. 6, pp. 601–611, 1976.

[32] Y. Mo, S. Weerakkody, and B. Sinopoli, "Physical authentication of control systems: designing watermarked control inputs to detect counterfeit sensor outputs," *IEEE Control Systems Magazine*, vol. 35, no. 1, pp. 93–109, 2015.

[33] R. S. Smith, "Covert misappropriation of networked control systems: Presenting a feedback structure," *IEEE Control Systems Magazine*, vol. 35, no. 1, pp. 82–92, 2015.

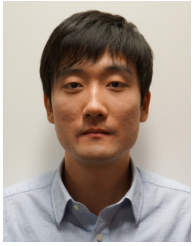
[34] U. Meyer and S. Wetzel, "A man-in-the-middle attack on UMTS," in *Proceedings of the 3rd ACM workshop on Wireless security*, 2004, pp. 90–97.

[35] F. Callegati, W. Cerroni, and M. Ramilli, "Man-in-the-middle attack to the HTTPS protocol," *IEEE Security and Privacy Magazine*, no. 1, pp. 78–81, 2009.

[36] D. Cox and D. Hinkley, *Theoretical statistics*. New York: Chapman and Hall, distributed in U.S. by Halsted Press, 1979.

[37] M. Grant, S. Boyd, V. Blondel, and H. Kimura, *CVX: Matlab Software for Disciplined Convex Programming, version 2.0*, 2011.

Ziyang Guo was born in Henan, China, in 1992. She received the B.Eng. degree (Honors) in College of Control Science and Engineering from Zhejiang University, Hangzhou, China, in 2014. She is currently pursuing the Ph.D. degree in Electronic and Computer Engineering at the Hong Kong University of Science and Technology, Hong Kong. Her research interests include cyber-physical system security, state estimation and wireless sensor network.



Dawei Shi received his B.Eng. degree in Electrical Engineering and Automation from the Beijing Institute of Technology in 2008. He received his Ph.D. degree in Control Systems from the University of Alberta in 2014. Since December 2014, he has been appointed as an Associate Professor at the School of Automation, Beijing Institute of Technology, China. His research interests include event-based control and estimation, robust model predictive control and tuning, and wireless sensor networks. He is a reviewer for a number of international journals, including IEEE Transactions on Automatic Control, Automatica, and Systems & Control Letters. In 2009, he received the Best Student Paper Award in IEEE International Conference on Automation and Logistics.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2016.2570003, IEEE Transactions on Control of Network Systems

Karl H. Johansson n is Director of the ACCESS Linnaeus Centre and Professor at the School of Electrical Engineering, KTH Royal Institute of Technology, Sweden. He is a Wallenberg Scholar and has held a Senior Researcher Position with the Swedish Research Council. He also heads the Stockholm Strategic Research Area ICT The Next Generation. He received M.Sc. and Ph.D. degrees in Electrical Engineering from Lund University. He has held visiting positions at UC Berkeley, California Institute of Technology, Nanyang Technological University, and Institute of Advanced Studies, Hong Kong University of Science and Technology. His research interests are in networked control systems, cyber-physical systems, and applications in transportation, energy, and automation systems. He has been a member of the IEEE Control Systems Society Board of Governors and the Chair of the IFAC Technical Committee on Networked Systems. He has been on the Editorial Boards of several journals, including Automatica, IEEE Transactions on Automatic Control, and IET Control Theory and Applications. He is currently a Senior Editor of IEEE Transactions on Control of Network Systems and Associate Editor of European Journal of Control. He has been Guest Editor for a special issue of IEEE Transactions on Automatic Control on cyberphysical systems and one of IEEE Control Systems Magazine on cyberphysical security. He was the General Chair of the ACM/IEEE CyberPhysical Systems Week 2010 in Stockholm and IPC Chair of many conferences. He has served on the Executive Committees of several European research projects in the area of networked embedded systems. He received the Best Paper Award of the IEEE International Conference on Mobile Ad-hoc and Sensor Systems in 2009 and the Best Theory Paper Award of the World Congress on Intelligent Control and Automation in 2014. In 2009 he was awarded Wallenberg Scholar, as one of the first ten scholars from all sciences, by the Knut and Alice Wallenberg Foundation. He was awarded Future Research Leader from the Swedish Foundation for Strategic Research in 2005. He received the triennial Young Author Prize from IFAC in 1996 and the Peccei Award from the International Institute of System Analysis, Austria, in 1993. He received Young Researcher Awards from Scania in 1996 and from Ericsson in 1998 and 1999. He is a Fellow of the IEEE.

Ling Shi received the B.S. degree in electrical and electronic engineering from Hong Kong University of Science and Technology, Kowloon, Hong Kong, in 2002 and the Ph.D. degree in control and dynamical systems from California Institute of Technology, Pasadena, CA, USA, in 2008. He is currently an associate professor at the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology. His research interests include networked control systems, wireless sensor networks, event-based state estimation and sensor scheduling, and smart energy systems. He has been serving as a subject editor for International Journal of Robust and Nonlinear Control from 2015. He also served as an associate editor for a special issue on Secure Control of Cyber Physical Systems in the IEEE Transactions on Control of Network Systems in 2015-2016.