2015 IEEE 54th Annual Conference on Decision and Control (CDC)
December 15-18, 2015. Osaka, Japan

Event-based Attack Against Remote State Estimation

Yifei Qi[1] , Peng Cheng[1], Ling Shi[2] and Jiming Chen[1]

Abstract— **Security issues in Cyber-Physical Systems (CPS) have gained increasing attention in recent years due to the vulnerability of CPS to cyber attack. This paper focuses on designing an intelligent online attack, which can compromise a sensor, eavesdrop measurements and inject false feedback information, against remote state estimation. From the viewpoint of the attacker, we design an event-based attack strategy to degrade the estimation quality with arbitrary communication rate constraint. The approximate minimum mean-squared error estimation algorithm at the attacker side is derived under the Gaussian assumption. Furthermore, the relation between the attack threshold and the scheduling threshold is obtained in a closed form. Two examples are provided to demonstrate the main results.**

I. Introduction

Cyber-Physical Systems (CPS), which integrate information and physical elements, have attracted great research interest in the past decade [1] [2]. Driven by the integration of control, communication and computation, CPS are applied in many industrial fields such as manufacturing, energy, and transportation, to name a few [3] [4]. Since the Iran's nuclear facilities were attacked by the Stuxnet worm in 2010, which caused significant damage to the system, security issues in CPS have been investigated from different perspectives [5] [6] [7].

Remote state estimation (RSE) is a critical problem in CPS, since its accuracy is the precondition of the system performance. However, due to limited energy of sensors and wireless transmission between sensors and estimators, RSE is vulnerable to the increasing cyber attack. One fundamental issue is to investigate the influence of attack actions in RSE. Zhang et al. in [8] studied the optimal Denial of Service (DoS) attack strategy against state estimation. The authors considered an energy constrained attacker whose objective is to maximize the average estimation error covariance by jamming the transmission channel, proved that the consecutive jamming is the optimal strategy, and provided the performance reduction with a closed form. In [9], Li et al. considered the situation where both sensor and attacker have energy constraint and studied the interactive decision making process. They proved that the optimal strategies for both sides constitute a Nash equilibrium in a game-theoretical framework. In [10], Zhang et al. further investigated the stealthy deception attack to degrade the state estimation quality with communication rate constraint that is no less than $1/2$ and presented an analytical relation between the attack effect and the communication rate.

It should be pointed out that most existing works mainly assume that the sensor has enough computational capability, and the sensor is able to send the local estimation to the remote estimator. Therefore, it is beneficial to the attacker for designing and executing attack strategy. However, there are various practical application scenarios where the computational capability is highly limited, e.g., electronic healthcare system with body sensor [11], environment monitoring with temperature or humidity sensor [12], etc. Motivated by these observations, in this paper, we focus on designing an intelligent online attack strategy against remote state estimation, in which the sensors send measurements directly, with arbitrary communication rate constraint caused by the limited sensor's energy or limited communication bandwidth.

The objective of the attacker is to degrade the estimation quality by cooperatively eavesdropping the measurements, compromising and injecting false feedback data into the sensor node. Since the sensor has limited computation capacity and the remote estimator may detect attack behavior if the communication rate is changed, it is challenging to design the attack mechanism to deteriorate the estimation performance as much as possible under the communication rate constraint. Specifically, we are interested in design an event-based attack with proper trigger mechanism and attack threshold to degrade the state estimation quality of a linear system with Gaussian noises.

The major contributions of this paper can be summarized as follows:

1) To the best of our knowledge, it is the first work on designing the online attack against remote state estimation with arbitrary communication rate constraint.
2) We design an event-based attack mechanism which leverage real-time measurements and can be implemented through compromising and injecting false feedback data into the sensor node.
3) We derive the approximate minimum mean-squared error (MMSE) estimation algorithm at the attacker side under the Gaussian assumption. We further obtain an analytical relation between the attack threshold and the scheduling threshold, which guarantees invariability of the communication rate.

The remainder of the paper is organized as follows. Section II introduces the system architecture and problem

[1]Y. Qi, P. Cheng and J. Chen are with State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou, China yifeiqi@zju.edu.cn, pcheng@iipc.zju.edu.cn, jmchen@ieee.org

[2]L. Shi is with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong, China eesling@ust.hk

The work was partially supported by NSFC under grants U1401253 and 61429301, and National Program for Special Support of Top Notch Young Professionals. The work by L. Shi is supported by an HKUST Caltech Partnership FP004.

formulation. In Section III, we design the intelligent online attack mechanism against remote state estimation, and derive the approximate MMSE algorithm from the viewpoint of the attacker and the attack threshold. Some examples are presented in Section IV to illustrate our results. Conclusion is given in section V.

Notations: $\mathbb{R}^n$ is the n-dimensional Euclidian space. $\mathbb{S}_+^n$ is the set of $n \times n$ positive semi-definite matrices. When $X \in \mathbb{S}_+^n$, we simply write $X \geq 0$; when X is positive definite, we write $X > 0$. $f_X(x)$ denotes the probability density function (pdf) of the random variable x, and $f_{X|y}(x|y)$ represents the pdf of x conditional on y. $\mathcal{N}(\mu, \Sigma)$ denotes Gaussian distribution with mean μ and covariance matrix Σ. $\mathbb{E}[X]$ denotes the mathematical expectation of X and $\mathbb{E}[X|y]$ represents the expectation of X conditional on y. $\| z \|_\infty$ and $\| z \|_{-\infty}$ represent the maximal absolute value and minimal absolute value of all entries in z, respectively. For functions $f_1, f_2 : \mathbb{S}_+^n \to \mathbb{S}_+^n, f_1 \circ f_2$ is defined as $f_1 \circ f_2(X) \triangleq f_1(f_2(X))$.

II. Problem Formulation

A. System Model

Consider a discrete linear-invariant system (Fig.1)

$$x_{k+1} = Ax_k + \omega_k \tag{1}$$
$$y_k = Cx_k + \nu_k, \tag{2}$$

where $x_k \in \mathbb{R}^n$ is the state system, $y_k \in \mathbb{R}^m$ is the measurement obtained by the sensor, and A, C are known time-invariant real matrices. $\omega_k \in \mathbb{R}^n$ and $\nu_k \in \mathbb{R}^m$ are uncorrelated zero-mean Gaussian random noise with covariances $Q \geq 0$ and $R > 0$, respectively. The initial state x_0 is also a zero-mean Gaussian random vector which is uncorrelated with ω_k or ν_k and has covariance $P_0 \geq 0$. Assume the pair $(A, \sqrt{Q})$ is controllable and (C, A) is observable.

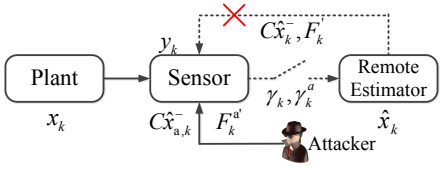

Fig. 1. System architecture

Denote $Y_k = \{y_0, y_1, \ldots, y_k\}$ as all the measurement data collected by the sensor from time 0 to time k. However, due to the energy constraint of the battery-powered sensor or communication bandwidth constraint, the measurements cannot be sent at each time slot, which leads to the sensor-to-estimator communication rate constraint. Thus, the sensor has to decide whether y_k shall be sent or not which is denoted as $\gamma_k = 1$ or 0, respectively. Then denote $\mathrm{I}_k = \{\gamma_0 y_0, \ldots, \gamma_k y_k\} \cup \{\gamma_0, \ldots, \gamma_k\}$ with $\mathrm{I}_{-1} = \emptyset$. The remote estimator shall estimate the system state by MMSE estimator including the a priori and a posteriori estimation which are defined as

$$\hat{x}_k^- \triangleq \mathbb{E}[x_k | \mathrm{I}_{k-1}] \qquad a\ priori, \tag{3}$$
$$\hat{x}_k \triangleq \mathbb{E}[x_k | \mathrm{I}_k] \qquad a\ posteriori. \tag{4}$$

The corresponding error and covariance are defined

$$e_k^- \triangleq x_k - \hat{x}_k^-, \qquad P_k^- \triangleq \mathbb{E}[e_k^- e_k^{-'} | \mathrm{I}_{k-1}], \tag{5}$$
$$e_k \triangleq x_k - \hat{x}_k, \qquad P_k \triangleq \mathbb{E}[e_k e_k' | \mathrm{I}_k]. \tag{6}$$

Further define the measurement innovation z_k as

$$z_k \triangleq y_k - \mathbb{E}[y_k | \mathrm{I}_{k-1}]. \tag{7}$$

To simplify the notation, we define the operators $h, \widetilde{g}_\lambda$ and $g_\lambda : \mathbb{S}_+^n \to \mathbb{S}_+^n$ as

$$h(X) \triangleq AXA' + Q, \tag{8}$$
$$\widetilde{g}_\lambda(X) \triangleq X - \lambda XC'[CXC' + R]^{-1}CX, \tag{9}$$
$$g_\lambda(X) \triangleq \widetilde{g}_\lambda \circ h(X), \tag{10}$$

where $\lambda \in [0, 1]$. While $\lambda = 1$, $\widetilde{g}_\lambda$ and g_λ will be written as $\widetilde{g}$ and g for brevity. Note that, $h(\cdot)$ and $\widetilde{g}$ are the Lyapunov operator and Riccati operator, respectively

B. Event-based Remote State Estimation

In this paper, we consider that the sensor has an average communication rate constraint defined as [13]

$$\gamma \triangleq \limsup_{T \to +\infty} \frac{1}{T+1} \sum_{k=0}^{T} \mathbb{E}[\gamma_k]. \tag{11}$$

While feedback is available from the estimator to the sensor, based on the fact that the innovation z_k is a Gaussian variable with zero mean and covariance $CP_k^-C' + R > 0$, Wu et al. in [13] proposed an event-based scheduling scheme for remote state estimation which can improve the estimation accuracy under the communication rate constraint.

The event-based sensor data scheduler is described as

$$\gamma_k = \begin{cases} 0, if \quad \| \epsilon_k \|_\infty \leq \delta, \\ 1, otherwise, \end{cases} \tag{12}$$

where δ is the event triggering threshold and

$$\epsilon_k \triangleq F_k' z_k, \ F_k = U_k \Lambda_k^{-\frac{1}{2}},$$

with $\Lambda_k = diag(\lambda_k^1, \ldots, \lambda_k^m) \in \mathbb{R}^{m \times m}$, $\lambda_k^1, \ldots, \lambda_k^m \in \mathbb{R}$ are the eigenvalues of $CP_k^-C' + R$, and U_k is a unitary matrix satisfying

$$U_k'(CP_k^-C' + R)U_k = \Lambda_k.$$

Then the approximate MMSE estimator and communication rate for remote estimator with (12) are summarized in the following lemma [13].

Lemma 2.1: Consider the remote state estimation with the event-based sensor scheduler (12). Under the Gaussian assumption

$$f_{x_k}(x | \mathrm{I}_{k-1}) = \mathcal{N}(\hat{x}_k^-, P_k^-),$$

the MMSE estimator is given recursively as follows:

1) Time update:

$$\begin{cases} \hat{x}_k^- = A\hat{x}_{k-1}, \\ P_k^- = h(P_{k-1}), \end{cases} \tag{13}$$

2) Measurement update:

$$\begin{cases} \hat{x}_k = \hat{x}_k^- + \gamma_k P_k^- C'[CP_k^- C' + R]^{-1} z_k, \\ P_k = \gamma_k \widetilde{g}(P_k^-) + (1-\gamma_k)\widetilde{g}_{\beta(\delta)}(P_k^-), \end{cases} \tag{14}$$

where

$$\beta(\delta) = \frac{2}{\sqrt{2\pi}} \delta e^{-\frac{\delta^2}{2}} [1 - 2Q(\delta)]^{-1},$$

and $Q(\cdot)$ is the standard $Q - function$ defined by

$$Q(\delta) \triangleq \int_{\delta}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.$$

Furthermore, the average sensor-to-estimator communication rate γ in (11) is given by

$$\gamma = 1 - [1 - 2Q(\delta)]^m. \tag{15}$$

Note that, the remote estimator can compute the average number of the received data, which is the unbiased estimation of γ. Consequently, the abnormal behavior may be detected if the communication rate is changed.

C. Event-based Attack Model

The energy-efficient state estimation framework proposed in [13] is beneficial for the energy-constrained cyber-physical systems such as Personal Area Network System [14], Wireless Identification and Sensing Platform [15], etc. However, as shown in [16], such systems are also vulnerable to be attacked.

In this paper, we consider an intelligent attacker that intrudes the sensor to learn and modify the transmission decisions, eavesdrops the measurements and injects the false feedback information to the sensor. The detailed capacities of the attacker and the model are described as follows.

1) The attacker is aware of the system parameters including A, C, Q, R, x_0, P_0 and the communication rate γ.
2) The attacker intrudes the sensor and modifies the program to change the transmission decision according to the following event-based attack model:

$$\gamma_k^a = \begin{cases} 0, & if\ \mathcal{M}(z_k^a) \geq \delta^a, \\ 1, & otherwise, \end{cases} \tag{16}$$

where γ_k^a and z_k^a are the transmission decision and the innovation under attack respectively, δ^a is the attack threshold and $\mathcal{M} : \mathbb{R}^m \to \mathbb{R}$ is a function which will be designed by the attacker.
3) Since the modified program is executed, the attacker eavesdrops the channel from the sensor to the estimator to obtain the measurements and runs its own estimation algorithm which will be derived in next section.
4) At each time slot, the attacker jams the feedback channel and injects the false feedback information which can be computed by the algorithm presented in next section.

Remark 2.1: The capacity 1) of the intelligent attacker could be realized by potentially performing system identification [7].

Remark 2.2: The attacker could exploit the sensor's unauthorized privileges to modify the program [17].

Remark 2.3: To degrade the estimation accuracy, intuitively and inversely, the measurements with small deviation from the last estimation should be transmitted. Therefore, the attack model (16) is leveraged by exploiting the innovation information to avoid violating the communication constraint.

Remark 2.4: The eavesdropping and estimation algorithm running in 3) and jamming and injecting in 4) can be realized by USRP [18] and here are due to the limitation of the computation in the sensor. It is worth noting that all of the attack actions in 3) and 4) can be implemented in the compromised sensor, if it has the ability to compute the matrix inversion.

D. Problem of Interest

For executing and evaluating the online attack by the model (16), the following four problems are considered and answered in this paper.

1) How to design the function $\mathcal{M}$ in event-based attack model (16)?
2) What is the MMSE estimation algorithm on the attacker side under the event-based attack with the given threshold δ^a ?
3) How to design δ^a to avoid being detected?

III. Event-based Attack Against Remote State Estimation

In this section, we present the details of the attack mechanism and derive the approximate MMSE estimator under the event-based attack.

A. Event-based Attack Mechanism

Denote $\mathrm{I}_k^a = \{\gamma_0^a y_0, \ldots, \gamma_k^a y_k\} \cup \{\gamma_0^a, \ldots, \gamma_k^a\}$ with $\mathrm{I}_{-1}^a = \emptyset$. Define $\hat{x}_{a,k}^-, \hat{x}_{a,k}, e_{a,k}^-, e_{a,k}, P_{a,k}^-, P_{a,k}$ and z_k^a by changing $\mathrm{I}_{k-1}, \mathrm{I}_k$ to $\mathrm{I}_{k-1}^a, \mathrm{I}_k^a$ in (3)-(7) as the a prior estimation and the a posteriori estimation and their error at the attacker side and the innovation at the sensor side under the online attack. According to section II-C, the attacker runs approximate MMSE estimation algorithm, which will be obtained in next part. Thus, the innovation under attack $z_k^a = y_k - \mathbb{E}(y_k|\mathrm{I}_{k-1}^a)$ is Gaussian random variable with zero mean and covariance $CP_{a,k}^- C' + R > 0$. Then there exists a unitary matrix $U_k^a \in \mathbb{R}^{m\times m}$ satisfying

$$U_k^{a'}(CP_{a,k}^- C' + R)U_k^a = \Lambda_{a,k}, \tag{17}$$

where $\Lambda_{a,k} = diag(\lambda_{a,k}^1, \ldots, \lambda_{a,k}^m) \in \mathbb{R}^{m\times m}$, and $\lambda_{a,k}^1, \ldots, \lambda_{a,k}^m \in \mathbb{R}$ are the eigenvalues of $CP_{a,k}^- C' + R$. Define $F_k^a \in \mathbb{R}^{m\times m}$ as $F_k^a = U_k^a \Lambda_{a,k}^{-\frac{1}{2}}$ and then $F_k^a F_k^{a'} = (CP_{a,k}^- C' + R)^{-1}$.

For ease of computing the attack threshold, the Mahalanobis transformation is first used in the mechanism $\mathcal{M}$ by pre-multiplying $F_k^{a'}$ to decorrelate the coordinates

of z_k^a. In the event-based scheduling mechanism (12), the measurements z_k is decided to be sent when the positive infinite norm of $F_k' z_k$ is larger than threshold which means the innovation is better for estimation. On the opposite, here the measurements will be sent while the negative infinite semi-norm of $F_k^{a'} z_k^a$ is smaller than the attack threshold. Mathematically, the attack mechanism $\mathcal{M}$ is defined as

$$\mathcal{M}(z_k^a) = \| F_k^{a'} z_k^a \|_{-\infty} \tag{18}$$

where $\| \cdot \|_{-\infty}$ is the negative infinite semi-norm of a vector.

B. The approximate MMSE estimator at the attacker side

In this part, we leverage an approximation technique which is widely used for nonlinear filtering in the literature, e.g., [13] [19], to assume that the conditional distribution of x_k given I_{k-1}^a is Gaussian, i.e.,

$$f_{x_k}(x|\mathrm{I}_{k-1}^a) = \mathcal{N}(\hat{x}_{a,k}^-, P_{a,k}^-). \tag{19}$$

Based on this assumption (19), the MMSE estimator with a simple form from the viewpoint of the attacker can be derived.

Before presenting the estimation and its proof, some preliminary results are stated. First, we prove the a priori statistical decorrelation property of $F_k^{a'} z_k^a$ as follows.

Lemma 3.1: Conditioning on I_{k-1}^a, $F_k^{a'} z_k^a$ is a Gaussian random variable with zero mean and covariance I_m.

Proof: See the Appendix. ∎

Then, we prove the a posteriori statistical properties of $F_k^{a'} z_k^a$ conditioned on $\hat{\mathrm{I}}_k$ which is defined as

$$\hat{\mathrm{I}}_k = \mathrm{I}_{k-1}^a \cup \{\gamma_k^a = 0\}.$$

Namely, we explore the properties of $F_k^{a'} z_k^a$ while the measurements y_k is not sent to the estimator/attacker. From the attack mechanism (18), $\gamma_k^a = 0$ means that

$$\mathcal{M}(z_k^a) = \| F_k^{a'} z_k^a \|_{-\infty} \geq \delta^a, \tag{20}$$

which is known at the attacker side. From Lemma 3.1, given I_{k-1}^a, the i-th and j-th elements of $F_k^{a'} z_k^a$, denoted as $\epsilon_{a,k}^i$ and $\epsilon_{a,k}^j$, are decorrelated with each other for arbitrary $i \neq j$. Therefore, we first investigate the a posteriori statistical properties of a one-dimensional Gaussian variable and prove the results in the following lemma.

Lemma 3.2: Let $x \in \mathbb{R}$ be a Gaussian random variable with zero mean and variance $\mathbb{E}[x^2] = \sigma^2$. Denoting $\Delta = \delta^a \sigma$, then $\mathbb{E}[x^2 | |x| \geq \Delta] = \sigma^2(1 + \overline{\beta}(\delta^a))$, where

$$\overline{\beta}(\delta^a) = \frac{1}{\sqrt{2\pi}} \cdot \delta^a \cdot e^{-\frac{(\delta^a)^2}{2}} \cdot [Q(\delta^a)]^{-1}. \tag{21}$$

Proof: Due to the space limits, the proof is omitted. ∎

Leveraging the results in Lemma 3.2, we have the following lemma.

Lemma 3.3: $\mathbb{E}[(F_k^{a'} z_k^a)(F_k^{a'} z_k^a)' | \hat{\mathrm{I}}_k^a] = [1 + \overline{\beta}(\delta^a)] I_m$.

Proof: See the Appendix. ∎

Furthermore, some equalities are required for deriving the approximate estimation and are given as follows.

Lemma 3.4: The following equalities hold.

$$\begin{aligned}
&\mathbb{E}[e_{a,k}^- z_k^{a'} | \hat{\mathrm{I}}_k^a] = L_k^a \mathbb{E}[z_k^a z_k^{a'} | \hat{\mathrm{I}}_k^a], \\
&\mathbb{E}[(e_{a,k}^- - L_k^a z_k^a) z_k^{a'} | \hat{\mathrm{I}}_k^a] = 0, \\
&\mathbb{E}[(e_{a,k}^- - L_k^a z_k^a)(e_{a,k}^- - L_k^a z_k^a)' | \mathrm{I}_{k-1}^a, z_k^a = z] = \widetilde{g}(P_{a,k}^-), \\
&\mathbb{E}[(e_{a,k}^- - L_k^a z_k^a)(e_{a,k}^- - L_k^a z_k^a)' | \hat{\mathrm{I}}_k^a] = \widetilde{g}(P_{a,k}^-),
\end{aligned}$$

where $L_k^a = P_{a,k}^- C' [C P_{a,k}^- C' + R]^{-1}$.

Proof: The proof is similar to that of Lemma 3.3 in [13] and is omitted here. ∎

Now we present the approximate MMSE estimation from the viewpoint of the attacker in the following theorem.

Theorem 3.1: Under the Gaussian assumption (19), the approximate MMSE estimator with the attack mechanism (18) is given recursively as follows:

1) Time update:

$$\begin{cases} \hat{x}_k^- = A \hat{x}_{k-1}, \\ P_{a,k}^- = h(P_{a,k-1}), \end{cases} \tag{22}$$

2) Measurement update:

$$\begin{cases} \hat{x}_k = \hat{x}_k^- + \gamma_k^a L_k^a z_k^a, \\ P_{a,k} = \gamma_k^a \widetilde{g}(P_{a,k}^-) + (1 - \gamma_k^a) \widetilde{g}_{[-\overline{\beta}(\delta^a)]}(P_{a,k}^-). \end{cases} \tag{23}$$

Proof: Based on the results in Lemma 3.1 and Lemma 3.4, it is easy to derive the time update and the measurement update when $\gamma_k^a = 1$. Thus, for saving space, we only prove the measurement update when $\gamma_k^a = 0$. Since the y_k is not transmitted, from the viewpoint of the attacker, the state is estimated as

$$\begin{aligned}
\hat{x}_{a,k} &= \mathbb{E}[x_k | \hat{\mathrm{I}}_k^a] = \mathbb{E}[\mathbb{E}[x_k | z_k^a = (F_k^{a'})^{-1}\epsilon, \mathrm{I}_{k-1}^a]] \\
&= \frac{1}{p_{\delta^a}} \int_\Omega \mathbb{E}[x_k | z_k^a = (F_k^{a'})^{-1}\epsilon, \mathrm{I}_{k-1}^a] f_{F_k^{a'} z_k^a}(\epsilon | \mathrm{I}_{k-1}^a) d\epsilon \\
&= \int_\Omega [\hat{x}_{a,k}^- + L_k^a (F_k^{a'})^{-1}\epsilon] \cdot p_{\delta^a}^{-1} \cdot f_{F_k^{a'} z_k^a}(\epsilon | \mathrm{I}_{k-1}^a) d\epsilon \\
&= \hat{x}_{a,k}^- \int_\Omega p_{\delta^a}^{-1} \cdot f_{F_k^{a'} z_k^a}(\epsilon | \mathrm{I}_{k-1}^a) d\epsilon \\
&\quad + L_k^a (F_k^{a'})^{-1} \int_\Omega p_{\delta^a}^{-1} \cdot \epsilon f_{F_k^{a'} z_k^a}(\epsilon | \mathrm{I}_{k-1}^a) d\epsilon,
\end{aligned}$$

where $p_{\delta^a} \triangleq Pr(\| F_k^{a'} z_k^a \|_{-\infty} \geq \delta^a | \mathrm{I}_{k-1}^a)$ and $\Omega = \{F_k^{a'} z_k^a \in \mathbb{R}^m : \| F_k^{a'} z_k^a \|_{-\infty} \geq \delta^a\}$. Note that, from Lemma 3.1, we have

$$f_{F_k^{a'} z_k^a}(\epsilon | \hat{\mathrm{I}}_k^a) = \begin{cases} \frac{f_{F_k^{a'} z_k^a}(\epsilon | \mathrm{I}_{k-1}^a)}{p_{\delta^a}}, & if \ \| F_k^{a'} z_k^a \|_{-\infty} \geq \delta^a, \\ 0, & otherwise, \end{cases}$$

which leads to $\int_\Omega p_{\delta^a}^{-1} \cdot f_{F_k^{a'} z_k^a}(\epsilon | \mathrm{I}_{k-1}^a) d\epsilon = 1$. Furthermore, $f_{F_k^{a'} z_k^a}(\epsilon | \mathrm{I}_{k-1}^a)$ is a even function, ϵ is a odd function and Ω is a symmetric subset with the center $[0, 0, \ldots, 0]$ in $\mathbb{R}^m$, which leads to $\int_\Omega p_{\delta^a}^{-1} \cdot \epsilon f_{F_k^{a'} z_k^a}(\epsilon | \mathrm{I}_{k-1}^a) d\epsilon = 0$. Therefore,

$$\hat{x}_{a,k} = \mathbb{E}[x_k | \hat{\mathrm{I}}_k^a] = \hat{x}_{a,k}^-.$$

Then the corresponding error covariance $P_{a,k}$ at the attacker's side can be computed as

$$\begin{aligned}
P_{a,k} &= \mathbb{E}\big[(x_k - \hat{x}_{a,k})(x_k - \hat{x}_{a,k})'|\hat{\mathrm{I}}_k^a\big] \\
&= \mathbb{E}\big[(x_k - \hat{x}_{a,k}^-)(x_k - \hat{x}_{a,k}^-)'|\hat{\mathrm{I}}_k^a\big] \\
&= \mathbb{E}\big[\{(e_{a,k}^- - L_k^a z_k^a) + L_k^a z_k^a\} \\
&\quad \cdot \{(e_{a,k}^- - L_k^a z_k^a) + L_k^a z_k^a\}'|\hat{\mathrm{I}}_k^a\big] \\
&= \mathbb{E}\big[(e_{a,k}^- - L_k^a z_k^a)(e_{a,k}^- - L_k^a z_k^a)' + L_k^a z_k^a z_k^{a'} L_k^{a'} \\
&\quad \cdot (e_{a,k}^- - L_k^a z_k^a) z_k^{a'} L_k^{a'} + L_k^a z_k^a (e_{a,k}^- - L_k^a z_k^a)'|\hat{\mathrm{I}}_k^a\big].
\end{aligned}$$

According to Lemma 3.3 and Lemma 3.4, we have

$$\begin{aligned}
P_{a,k} &= \widetilde{g}(P_{a,k}^-) + L_k^a \mathbb{E}[z_k^a z_k^{a'}|\hat{\mathrm{I}}_k^a] L_k^{a'} \\
&= \widetilde{g}(P_{a,k}^-) + [1 + \overline{\beta}(\delta^a)] L_k^a (F_k^a F_k^{a'})^{-1} L_k^{a'} \\
&= \widetilde{g}(P_{a,k}^-) + [1 + \overline{\beta}(\delta^a)] L_k^a (C P_{a,k}^- C' + R) L_k^{a'} \\
&= P_{a,k}^- - P_{a,k}^- C'[C P_{a,k}^- C' + R]^{-1} C P_{a,k}^- \\
&\quad + [1 + \overline{\beta}(\delta^a)] P_{a,k}^- C'[C P_{a,k}^- C' + R]^{-1} \\
&\quad \cdot (C P_{a,k}^- C' + R)(C P_{a,k}^- C' + R)^{-1} C P_{a,k}^- \\
&= P_{a,k}^- - P_{a,k}^- C'[C P_{a,k}^- C' + R]^{-1} C P_{a,k}^- \\
&\quad + [1 + \overline{\beta}(\delta^a)] P_{a,k}^- C'[C P_{a,k}^- C' + R]^{-1} C P_{a,k}^- \\
&= P_{a,k}^- + \overline{\beta}(\delta^a) P_{a,k}^- C'[C P_{a,k}^- C' + R]^{-1} C P_{a,k}^- \\
&= \widetilde{g}_{[-\overline{\beta}(\delta^a)]}(P_{a,k}^-),
\end{aligned}$$

which leads to (23). ∎

C. Design of the Attack Threshold

To avoid being detected by the estimator, the attack threshold should be designed properly, namely it must be computed to guarantee the communication rate. First, a basic property is derived.

Lemma 3.5: While $\delta^a > 0$, we have

$$Pr(\mathcal{M}(z_k^a) \geq \delta^a \big| \mathrm{I}_{k-1}^a) = [2Q(\delta^a)]^m. \tag{24}$$

Proof: The proof is similar to that of Lemma 3.3 in [13] and is omitted here. ∎

Then we show the relation between attack threshold and communication rate as well as the relation between attack threshold and scheduling threshold in the following theorem.

Theorem 3.2: For the approximate MMSE estimation under attack, the communication rate

$$\gamma^a \triangleq \limsup_{T\to+\infty} \frac{1}{T+1} \sum_{k=0}^{T} \mathbb{E}[\gamma_k^a]. \tag{25}$$

satisfies

$$\gamma^a = 1 - [2Q(\delta^a)]^m, \tag{26}$$

and to avoid being detected, the attack threshold δ^a must satisfies

$$Q(\delta^a) + Q(\delta) = \frac{1}{2}, \tag{27}$$

where δ is the event-based scheduling threshold without attack.

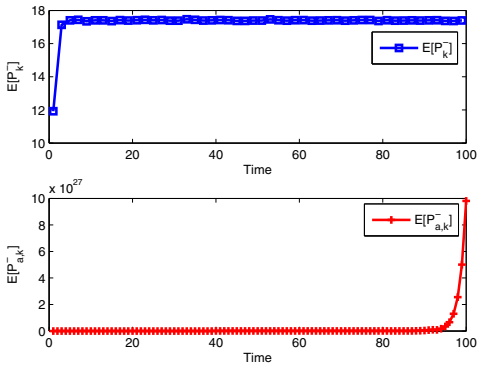

Fig. 2. Expectation of the prediction error covariance for Example 1 with and without attack (communication rate $\gamma = \gamma^a = 0.5$).

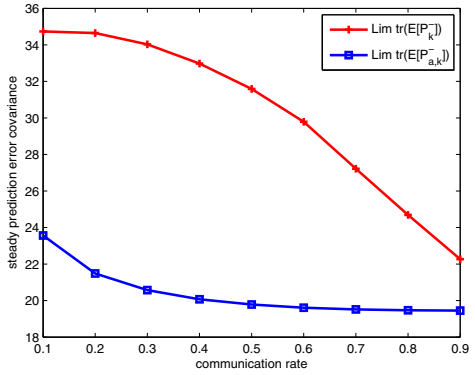

Fig. 3. Comparison of the steady prediction error covariance for Example 2 with and without attack .

Proof: Since γ_k^a is a random variable taking value only in $\{0, 1\}$, and according to Lemma 3.5, we have

$$\begin{aligned}
\gamma^a &= \limsup_{T\to+\infty} \frac{1}{T+1} \sum_{k=0}^{T} Pr(\gamma_k^a = 1|\mathrm{I}_{k-1}^a) \\
&= Pr(\| F_k^{a'} z_k^a \|_{-\infty} < \delta^a |\mathrm{I}_{k-1}^a) \\
&= 1 - Pr(\| F_k^{a'} z_k^a \|_{-\infty} \geq \delta^a |\mathrm{I}_{k-1}^a) \\
&= 1 - [2Q(\delta^a)]^m.
\end{aligned}$$

Furthermore, to avoid being detected, the communication rate under the attack should be equal to that without attack (15), namely

$$1 - [2Q(\delta^a)]^m = 1 - [1 - 2Q(\delta)]^m,$$

which leads to (27). ∎

Remark 3.1: For the effectiveness of the designed attack, we believe that the mean-squared stability condition and asymptotic performance of estimation become weakened under attack. However, due to the limited space, the theoretical analysis is omitted here.

IV. Examples

In this section, we leverage two examples to demonstrate the effectiveness of the proposed attack mechanism.

Example 1: Consider a unstable scalar system whose parameters are $A = 1.4, C = 1, Q = 8, R = 5$.

From the results in [20], we can obtain the critical communication rate without attack as $\gamma_c = 0.1198$, below which the estimation will diverge and above which the estimation will converge. In Fig.2, we fix the communication rate as $\gamma = \gamma^a = 0.5$, and plot the expected prediction error covariance $\mathbb{E}[P_k^-]$ when there is no attack and $\mathbb{E}[P_{a,k}^-]$ at the attacker side when estimation is under attack. It can be found that $\mathbb{E}[P_k^-]$ converges since $\gamma > \gamma_c$ and $\mathbb{E}[P_{a,k}^-]$ diverges, which shows that the mean-squared stability condition of estimation becomes weakened under the designed attack.

Example 2: Consider a two-dimensional stable system with the following parameters

$$A = \begin{bmatrix} 0.8 & 0.2 \\ 0 & 0.8 \end{bmatrix}, Q = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 \end{bmatrix}, R = 2.$$

For arbitrary communication rate, the estimation under attack will be stable, since system is stable. Then we compare the steady expected prediction error covariance $\lim_{k\to\infty} tr(\mathbb{E}[P_k^-])$ and $\lim_{k\to\infty} tr(\mathbb{E}[P_{a,k}^-])$ of the estimation without attack and under attack for the communication ranging from 0.1 to 0.9. The curves in Fig.3 show the estimation quality is always degraded for any communication rate, which verify the effectiveness of the designed attack mechanism.

V. Conclusion

In this paper, we investigate event-based attack design against remote state estimation with arbitrary communication rate constraint. We design an intelligent attack strategy that can degrade the estimation performance by leveraging the online measurements information. The approximate MMSE estimation algorithm is derived with a simple form at the attack side. The choice of attack threshold to avoid being detected is presented with an analytical form.

Appendix

Proof: [Proof of Lemma 3.1]

From the definition of z_k^a and the assumption (19), z_k^a is zero-mean Gaussian conditioned on I_{k-1}^a, and z_k^a is jointly Gaussian with x_k conditioned on I_{k-1}^a. Furthermore,

$$\begin{aligned}
&\mathbb{E}[z_k^a z_k^{a'} \big| \mathrm{I}_{k-1}^a] = \mathbb{E}[(Ce_{a,k}^- \nu_k)(Ce_{a,k}^- + \nu_k)' \big| \mathrm{I}_{k-1}^a] \\
&= C\mathbb{E}[e_{a,k}^- e_{a,k}^{-'} \big| \mathrm{I}_{k-1}^a]C' + R = CP_{a,k}^- C' + R, \\
&\mathbb{E}[(F_k^{a'} z_k^a)(F_k^{a'} z_k^a)' \big| \mathrm{I}_{k-1}^a] = F_k^{a'} \mathbb{E}[z_k^a z_k^{a'} \big| \mathrm{I}_{k-1}^a] F_k^a = I_m,
\end{aligned}$$

which completes the proof. ∎

Proof: [Proof of Lemma 3.3]

Given I_{k-1}^a, due to the decorraltion of $\epsilon_{a,k}^i$ and $\epsilon_{a,k}^j$, from Lemma 3.2, we have

$$\begin{aligned}
\mathbb{E}[(\epsilon_{a,k}^i)^2 \big| \hat{\mathrm{I}}_k^a] &= \mathbb{E}[(\epsilon_{a,k}^i)^2 \big| \mathrm{I}_{k-1}^a, \| F_k^{a'} z_k^a \|_{-\infty} \geq \delta^a] \\
&= \mathbb{E}[(\epsilon_{a,k}^i)^2 \big| \mathrm{I}_{k-1}^a, |\epsilon_{a,k}^i| \geq \delta^a] = 1 + \overline{\beta}(\delta^a),
\end{aligned}$$

and

$$\mathbb{E}[\epsilon_{a,k}^i \epsilon_{a,k}^i \big| \hat{\mathrm{I}}_k^a] = \mathbb{E}[\epsilon_{a,k}^i \epsilon_{a,k}^i \big| \mathrm{I}_{k-1}^a, |\epsilon_{a,k}^i| \geq \delta^a, |\epsilon_{a,k}^j| \geq \delta^a] = 0.$$

Thus,

$$\mathbb{E}[(F_k^{a'} z_k^a)(F_k^{a'} z_k^a)' \big| \hat{\mathrm{I}}_k^a] = [1 + \overline{\beta}(\delta^a)] I_m$$

which completes the proof. ∎

References

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

[2] X. Cao, P. Cheng, J. Chen, S. Ge, Y. Cheng, and Y. Sun, "Cognitive radio based state estimation in cyber-physical system," *IEEE Journal on Selected Areas on Communications*, vol. 32, no. 3, pp. 489–505, 2014.

[3] J. Chen, X. Cao, P. Cheng, Y. Xiao, and Y. Sun, "Distributed collaborative control for industrial automation with wireless sensor and actuator networks," *IEEE Transactions on Industrial Electronics*, vol. 57, no. 12, pp. 4219–4230, 2010.

[4] A. Teixeira, S. Amin, H. Sandberg, K. H. Johansson, and S. S. Sastry, "Cyber security analysis of state estimators in electric power systems," in *IEEE International Conference on Decision and Control*, Dec 2010, pp. 5991–5998.

[5] M. Pajic, J. Weimer, N. Bezzo, P. Tabuada, O. Sokolsky, I. Lee, and G. J. Pappas, "Robustness of attack-resilient estimators," in *Proceedings of International Conference of Cyberphysical Systems (ICCPS)*, 2014, pp. 163–174.

[6] H. Zhang, P. Cheng, L. Shi, and J. Chen, "Optimal dos attack scheduling in wireless networked control system," *IEEE Transactions on Control System Technology, DOI:10.1109/TCST.2015.2462741*, 2015.

[7] 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.

[8] H. Zhang, P. Cheng, L. Shi, and J. Chen, "Optimal Denial-of-Service attack scheduling with energy constraint," *IEEE Transactions on Automatic Control, DOI:10.1109/TAC.2015.2409905*, 2015.

[9] 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, DOI:10.1109/CYBER.2013.6705454*, 2015.

[10] H. Zhang, P. Cheng, J. Wu, L. Shi, and J. Chen, "Online deception attack against remote state estimation," in *Proceedings of World Congress of the International Federation of Automatic Control (IFAC)*, 2014, pp. 128–133.

[11] G. zhong Yang and M. Yacoub, *Body sensor networks*. Springer, 2006.

[12] A. Mainwaring, J. Polastre, R. Szewczyk, D. Culler, and J. Anderson, "Wireless sensor networks for habitat monitoring," in *Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications*, 2002, pp. 88–97.

[13] J. Wu, Q.-S. Jia, K. H. Johansson, and L. Shi, "Event-based sensor data scheduling:trade-off between communication rate and estimation quality," *IEEE Transactions on Automatic Control*, vol. 58, no. 4, pp. 1041–1046, 2013.

[14] E. Callaway, P. Gorday, L. Hester, J. A. Gutierreaz, M. Naeve, B. Heile, and V. Bahl, "Home networking with ieee 802.15.4: a developing standard for low-rate wireless personal area networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 70–77, 2002.

[15] J. R. Smith, A. Sample, P. Powledge, A. Mamishev, and S. Roy, "A wirelessly powered platform for sensing and computation," in *Proceedings of Ubicomp 2006: 8th International Conference on Ubiquitous Computing*, 2006, pp. 495–506.

[16] S. Lim, T. H. Oh, Y. B. Choi, and T. Lakshman, "Security issues on wireless body area network for remote healthcare monitoring," in *Proceedings of IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC)*, 2010, pp. 327–332.

[17] E. Bryant, M. Atallah, and M. Sytz, "A survey of anti-tamper technologies crosstalk," *The Journal of Defense Software Engineering*, vol. 17, no. 11, pp. 12–16, 2004.

[18] E. Bayraktaroglu, C. King, X. Liu, G. Noubir, R. Rajaraman, and B. Thapa, "Performance of ieee 802.11 under jamming," *Mobile Networks and Applications*, vol. 18, no. 5, pp. 678–696, 2013.

[19] A. Ribeiro, G. B. Giannakis, and S. I. Roumeliotis, "Soi-kf: Distributed kalman filterin with low-cost communications using the sign of innovations," *IEEE Transactions on Signal Processing*, vol. 54, no. 12, pp. 4782–4795, 2006.

[20] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1453–1464, 2004.