Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Online Deception Attack Against Remote State Estimation

Heng Zhang * **Peng Cheng** * **Junfeng Wu** ** **Ling Shi** **
Jiming Chen *

* *State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou, P.R. China (e-mail: ezhangheng@gmail.com, pcheng@iipc.zju.edu.cn, jmchen@ieee.org).*
** *Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong, P.R. China (e-mail: jfwu@ust.hk,eesling@ust.hk)*

Abstract: Security issue has become a new hotspot in cyber-physical systems (CPS) research field in recent years due to the vulnerability of CPS to security threats. This paper focuses on stealthy deception attack in remote state estimation, which is one typical attack in CPS. From the standpoint of deception attacker, we investigates how to design proper deception attack strategy to degrade the state estimation quality with communication rate constraint. We design an online attack strategy and prove that the proposed attack strategy can degrade the estimation quality. To study the effectiveness of the proposed strategy, we analyze the cost deviation, which depicts the difference between the estimation quality with and without the proposed attack strategy. Our results show that the cost deviation will be maximum when the communication rate is 0.75. A numerical example is presented to demonstrate the main results.

Keywords: Cyber-Physical Systems; Secure; State estimation.

1. INTRODUCTION

Cyber-physical systems (CPS), which smoothly integrate information and physical elements, have a large spectrum of applications, including smart grid (Bitar et al., 2011), smart building (Novak and Gerstinger, 2010), intelligent transportation (Qu et al., 2010), public health (Sarwate and Chaudhuri, 2013), etc. Due to its importance, it is of great research interests to investigate the vulnerability of CPS under various threats launched in either cyber or physical space (Zhang et al., 2014). A well-known example is the Stuxnet worm, which attacked Iran's nuclear facilities and resulted in more than 1000 centrifuges (10 percent) breakdown between November 2009 and late January 2010 (Wilson, 2013).

In this paper we focus on stealthy deception attack, which compromises sensor nodes, aiming at degrading the system performance without being detected (Cardenas et al., 2009). A typical deception attacker can capture the sensor nodes, exploit its unauthorized privileges to inject malicious code or modify the program, and then deteriorate the system performance stealthily (Bryant et al., 2004; Song et al., 2007; Kavitha and Sridharan, 2010).

One basic issue in CPS security is to study the consequence of attack actions (Shoukry et al., 2013). Zhang et al. (2013) has studied an optimal offline DoS attack strategy against state estimation, where, subject to an energy constraint in a finite time horizon, the attacker jams the transmission channel without being detected. In (Zhang et al., 2013), it was assumed that the sensor can always send the data to the estimator and every data can be received by the estimator with a certain probability. However, if the sensor has energy constraint or communication bandwidth constraint, it cannot send the data in every time slot. Therefore needs to design its transmission schedule to improve the estimation quality. Wu et al. (2013b) designed an online transmission schedule under communication rate constraint to minimize the state estimation error. It is interesting and challenging to design an attack strategy to maximize the attack effect under such communication rate constraint.

Since the remote estimator may detect the attack behavior if the communication rate constraint is violated, the basic research direction is whether and how the attacker can exploit the online information to degrade the system performance as much as possible under the communication rate constraint. Motivated by this, we focus on the online attack strategy design in order to degrade the estimation quality. Specifically, we consider deception attack strategy against state estimation of a linear system with Gaussian noises. In the viewpoint of attacker, we are interested in design proper online deception attack strategy to degrade the state estimation quality.

Our main contributions can be summarized as follows:

⋆ The work was supported in part by NSFC under Grants 61222305, 61290325, the SRFDP under Grant 20120101110139, NCET-11-0445, and National Program for Special Support of Top-Notch Young Professionals. The work by L. Shi is supported by an HKUST Caltech Partnership FP004.

(1) We propose an online attack strategy against state estimation and prove that the proposed strategy can degrade the estimation quality.
(2) We study the cost deviation under the proposed attack strategy, and prove that there exists a sensor-to-estimator rate to maximize this deviation.
(3) We obtain a closed-form expression of the sensor-to-estimator rate which maximizes the cost deviation.

The remainder of the paper is organized as follows: Section 2 formulates the problem. Section 3 proposes an online attack strategy and then evaluates the impact of this strategy. Section 4 illustrates the effectiveness of our proposed attack strategy. Section 5 concludes the whole paper.

Notations: $\mathbb{S}_+^n$ is the set of $n \times n$ positive semi-definite matrices. $\mathbb{R}^r$ is the r dimensional Euclidean space. $\mathbb{E}[X]$ and $\mathbb{D}[X]$ stand for the mean and variance of random variable X, respectively. $\mathbb{E}[X|Y]$ stands for the the mean of random variable X conditioned on Y. $\phi(\cdot)$ is the probability density function of Gaussian distribution $\mathcal{N}(0,1)$. $Tr(\cdot)$ is the trace operation of matrix. $\|\xi\|$ stands for Euclidean norm of a vector ξ. I_r represents $r \times r$ identity matrix. $\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_r)$ stands for the diagonal matrix with the diagonal elements $\lambda_i, i = 1, 2, \ldots, r$. $\mathrm{rank}(\cdot)$ is the rank of a matrix. $(\cdot)'$ is the transpose operation of a matrix.

2. PROBLEM FORMULATION

Consider the following linear system (Fig. 1)

$$\begin{aligned} x_{k+1} &= Ax_k + w_k, \\ y_k &= Cx_k + v_k, \end{aligned} \tag{1}$$

where $x_k \in \mathbb{R}^n$ is the state variable with $n \in \mathbb{N}$, $y_k \in \mathbb{R}^m$ is the measurement variable with $m \in \mathbb{N}$, $w_k \in \mathbb{R}^n$ is the process noise, $v_k \in \mathbb{R}^m$ is the measurement noise, w_k and v_k are uncorrelated zero mean Gaussian noises with covariance Σ_w and Σ_v, respectively. The pair $(A, \Sigma_w^{\frac{1}{2}})$ is stabilizable and (A, C) is assumed to be detectable .

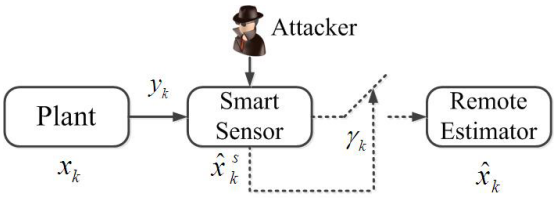

Fig. 1. System architecture

2.1 System architecture

The sensors, which have sufficient computational capability to estimate the system state x_k after reading the measurement y_k, are referred to as smart sensors. We assume a smart sensor is used. Its local estimate is calculated by a Kalman filter, i.e., $\widehat{x}_k^s = \mathbb{E}[x_k|y_0, \ldots, y_k]$. Sensor's estimation error is defined as $e_k^s = x_k - \widehat{x}_k^s$, and its error covariance matrix is defined as $P_k^s = \mathbb{E}[(e_k^s)(e_k^s)'|y_0, \ldots, y_k]$.

It is assumed that $\widehat{x}_0^s = 0$ and $P_0^s = \Pi_0 \geq 0$. From (Anderson and Moore, 1981), one can see that error covariance matrix P_k^s converges to a steady-state value P exponentially. We shall ignore the transient period and assume that $\Pi_0 = P$.

The sensor then decides whether or not to send this state estimate to the remote estimator. We denote $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_N)$ as the sensor's decision vector in a finite time horizon $[1, N]$, i.e., $\gamma_k = 1$ if the sensor sends $\widehat{x}_k^s$, and $\gamma_k = 0$ otherwise.

Denote the data set at remote estimator as $\mathcal{D}(\gamma)$. Then, its state estimate and corresponding error covariance are given by

$$\hat{x}_k(\gamma) = \mathbb{E}[x_k|\mathcal{D}(\gamma)]$$

and

$$P_k(\gamma) = \mathbb{E}[(x_k - \hat{x}_k)(x_k - \hat{x}_k)'|\mathcal{D}(\gamma)].$$

For simplicity, we write $\hat{x}_k(\gamma)$ as $\hat{x}_k$, etc., when the schedule γ is given.

From (Shi et al., 2011b), it can be seen that

$$\widehat{x}_k = \begin{cases} \widehat{x}_k^s, & \text{if } \gamma_k = 1; \\ A\widehat{x}_{k-1}, & \text{otherwise.} \end{cases} \tag{2}$$

The estimation quality is measured by the cost

$$J(\gamma) = \limsup_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} Tr\{\mathbb{E}[P_k]\}.$$

From the sensor's point of view, it aims to find transmission strategy which minimizes J for a given average sensor-to-estimator communication rate [1] $\overline{\gamma}$, where

$$\overline{\gamma} = \limsup_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \mathbb{E}(\gamma_k).$$

It is assumed that the sensor runs an online scheduler θ^s (cf.(Wu et al., 2013a)) in the sensor as follows:

$$\gamma_k^s = \begin{cases} 0, \text{ if } k \text{ is even and } \|E'\epsilon_k\| < \delta; \\ 1, \text{ otherwise.} \end{cases} \tag{3}$$

where $\epsilon_k = \widehat{x}_k^s - A\widehat{x}_{k-1}$, and δ is event-triggering threshold which is determined by the given average sensor-to-estimator communication rate $\overline{\gamma} \in [\frac{1}{2}, 1]$. The matrix E will be defined shortly.

This schedule can improve the estimation quality with the sensor-to-estimator communication rate constraint (cf. (Wu et al., 2013b)).

2.2 The objective of attacker

In our scenario, the deception attacker intrudes the sensor, stealing the compromised sensor's codes to learn its online transmission strategy. It then tampers the sensor's program by implanting its designed codes to the sensor.

Here we assume that the estimator knows that the sensor's online schedule is of the form given by (3). It means that the attacker can only recompose the transmission schedule in even time since no transmission at odd time can be easily detected by estimator. The estimator can estimate the communication rate from its prior knowledge

[1] The communication rate is defined due to sensor's energy constraint or the limitation of communication bandwidth (cf.(Wu et al., 2013a)).

about sensor's energy budget and communication channel's bandwidth constraint.

From the viewpoint of attacker, it aims to design a proper attack mechanism against remote state estimation without being detected by estimator. Specifically, the attacker modifies the transmission schedule as follows:

$$\gamma_k^a = \begin{cases} 0, \text{ if } k \text{ is even and } \|E'\epsilon_k\| \geq \delta^a; \\ 1, \text{ otherwise.} \end{cases} \quad (4)$$

Here ϵ_k is the online information in the sensor at time k, and δ^a is a new event-triggering threshold tampered by the attacker. The basic idea of our proposed attack strategy is that if the difference of $\widehat{x}_k^s$ and $A\widehat{x}_{k-1}$ is big in even time k, i.e., $\|E'\epsilon_k\| \geq \delta^a$, the attacker will decide not to send the data $\widehat{x}_k^s$ to the estimator.

This paper aims to answer the following two questions:

(1) How to design the parameter δ^a to avoid being detected by the remote estimator?
(2) Whether the online attack strategy (4) with designed parameter δ^a can degrade the estimation quality?

3. ATTACK STRATEGY DESIGN AND ANALYSIS

In this section, we first present the online attack strategy and then prove the proposed strategy can degrade the estimation quality. We also study the deviation of the estimation quality under the proposed attack strategy. Our result shows that there exists a sensor-to-estimator communication rate such that the deviation reaches maximum.

3.1 Preliminaries

Before studying our proposed attack strategy, we present some preliminaries in this subsection.

We define the functions $h : \mathbb{S}_+^n \to \mathbb{S}_+^n$ as

$$h(X) = AXA' + \Sigma_w.$$

Define a matrix H in terms of steady-state error covariance P as

$$H \triangleq h(P) - P,$$

and the rank of matrix H as $r \triangleq \text{rank}(H)$. Then one can see that $H \geq 0$ from (Shi et al., 2011a).

One can see that there exists an orthonormal matrix F such that

$$H = F \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} F',$$

where $\Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_r)$ and the diagonal elements $\lambda_i, i = 1, 2, \ldots, r$ are the positive eigenvalues of H. Let

$$E = F \begin{bmatrix} \Lambda^{-\frac{1}{2}} & 0 \\ 0 & I_{n-r} \end{bmatrix}.$$

Then one has

$$E'HE = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}.$$

Let $\xi \in \mathbb{R}^r$ be a random variable following Gaussian distribution $\mathcal{N}(0, I_r)$. Define

$$\rho(\delta) \triangleq \Pr(\|\xi\| \leq \delta), \quad (5)$$

$$\widehat{\Gamma}(\delta) \triangleq \mathbb{E}[\xi\xi' \mid \|\xi\| \leq \delta], \quad (6)$$

$$\Gamma(\delta) \triangleq \Lambda^{\frac{1}{2}} \widehat{\Gamma}(\delta) \Lambda^{\frac{1}{2}}. \quad (7)$$

Lemma 3.1. $\Gamma(\delta)$ has following property:

$$\Gamma(\delta) < \Lambda.$$

Proof. See Appendix.

From (Wu et al., 2013b), one can see that $\epsilon_k = \widehat{x}_k^s - A\widehat{x}_{k-1}$ follows Gaussian distribution $\mathcal{N}(0, H)$. Thus $E'\epsilon_k$ follows Gaussian distribution $\mathcal{N}(0, E'HE)$.

3.2 Attack strategy design

Note that in order not to be detected, the threshold δ^a needs to be designed to follow the sensor-to-estimator communication rate $\overline{\gamma}$.

The following theorem presents the relation between communication rate and event-triggering threshold δ^a, which can help the deception attacker to implement the attack strategy without being detected by estimator.

Theorem 3.1. The communication rate $\overline{\gamma}$ and threshold δ^a satisfy

$$\overline{\gamma} = \frac{1}{2} + \frac{1}{2}\rho(\delta^a). \quad (8)$$

Furthermore the threshold δ and δ^a satisfy

$$\rho(\delta) + \rho(\delta^a) = 1. \quad (9)$$

Proof. The communication rate under attack is given by

$$\begin{aligned} \overline{\gamma}(\delta^a) &= \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} \mathbb{E}[\gamma_k^a] \\ &= \lim_{N\to\infty} \frac{1}{N} \Big(\sum_{k \text{ is odd}} \mathbb{E}[\gamma_k^a] + \sum_{k \text{ is even}} \mathbb{E}[\gamma_k^a] \Big) \\ &= \frac{1}{2} + \lim_{N\to\infty} \frac{1}{N} \sum_{k \text{ is even}} \Pr(\|E'\epsilon_k\| \leq \delta^a) \\ &= \frac{1}{2} + \frac{1}{2}\rho(\delta^a). \end{aligned}$$

Let $\overline{\gamma}(\delta) = \overline{\gamma}(\delta^a)$, then from (A.1) in Appendix one can see that

$$1 - \frac{1}{2}\rho(\delta) = \frac{1}{2} + \frac{1}{2}\rho(\delta^a)$$

which leads to (9).

From Theorem 3.1, it can be seen that the threshold δ^a is determined by the communication rate $\overline{\gamma}$. When the attacker obtains the value δ from sensor's original codes, the threshold δ^a can be computed by (9) which answers the first question in Section 2.

In what follows, in order to show that the proposed attack strategy is able to degrade the system performance, we first present the following lemma which provides the analytical expression of system cost J under the attack schedule θ^a.

Theorem 3.2. Under attack schedule θ^a, the cost $J(\theta^a)$ is given by

$$J(\theta^a) = Tr\left[P + \frac{1}{2} F \begin{bmatrix} \Lambda - \rho(\delta^a)\Gamma(\delta^a) & 0 \\ 0 & 0 \end{bmatrix} F' \right]. \quad (10)$$

Proof. When k is even and $E'\epsilon_k > \delta^a$, we have

$$\begin{aligned}P_k^a &= \mathbb{E}[(x_k-\widehat{x}_k)(x_k-\widehat{x}_k)' \mid \|E'\epsilon_k\| > \delta^a] \\ &= \mathbb{E}[(e_k^s)(e_k^s)' \mid \|E'\epsilon_k\| > \delta^a] + \mathbb{E}[\epsilon_k\epsilon_k' \mid \|E'\epsilon_k\| > \delta^a] \\ &= P + \mathbb{E}[\epsilon_k\epsilon_k' \mid \|E'\epsilon_k\| > \delta^a],\end{aligned}$$

where

$$\begin{aligned}&\mathbb{E}[\epsilon_k\epsilon_k' \mid \|E'\epsilon_k\| > \delta^a] \\ &= \frac{1}{1-\rho(\delta^a)}\Big\{H - \rho(\delta^a)\mathbb{E}[\epsilon_k\epsilon_k' \mid \|E'\epsilon_k\| \le \delta^a]\Big\} \\ &= \frac{1}{1-\rho(\delta^a)} F \begin{bmatrix}\Lambda - \rho(\delta^a)\Gamma(\delta^a) & 0 \\ 0 & 0\end{bmatrix} F'.\end{aligned}$$

Therefore

$$\begin{aligned}J(\theta^a) &= \lim_{N\to\infty}\frac{1}{N}\sum_{k=0}^{N-1} Tr[P_k^a] \\ &= \lim_{N\to\infty}\frac{1}{N}\sum_{k=0}^{N-1}[1-Pr(\gamma_k^a=0)]Tr[P_k^a] \\ &\quad + \lim_{N\to\infty}\frac{1}{N}\sum_{k=0}^{N-1} Pr(\gamma_k^a=0)Tr[P_k^a] \\ &= Tr\Big[\Big(\frac{1}{2}+\frac{1}{2}\rho(\delta^a)\Big)P\Big] + Tr\Big[\Big(\frac{1}{2}-\frac{1}{2}\rho(\delta^a)\Big)\Big(P \\ &\quad + \frac{1}{1-\rho(\delta^a)}F\begin{bmatrix}\Lambda-\rho(\delta^a)\Gamma(\delta^a) & 0\\ 0 & 0\end{bmatrix}F'\Big)\Big] \\ &= Tr\left[P + \frac{1}{2}F\begin{bmatrix}\Lambda-\rho(\delta^a)\Gamma(\delta^a) & 0\\ 0 & 0\end{bmatrix}F'\right].\end{aligned}$$

Theorem 3.3. Comparing the cost J under schedules θ^a and θ^s with the same sensor-to-estimator communication rate, one has

$$J(\theta^a) > J(\theta^s).$$

Proof. From Property A.1 in Appendix and Theorem 3.2, it can be seen that

$$\begin{aligned}&J(\theta^a) - J(\theta^s) \\ &= \frac{1}{2}Tr\left[F\begin{bmatrix}\Lambda-\rho(\delta^a)\Gamma(\delta^a)-\rho(\delta)\Gamma(\delta) & 0\\ 0 & 0\end{bmatrix}F'\right] \quad (11) \\ &> 0.\end{aligned}$$

The last inequality is true since

$$\rho(\delta^a)\Gamma(\delta^a) + \rho(\delta)\Gamma(\delta) < \Lambda[\rho(\delta^a)+\rho(\delta)] = \Lambda.$$

From Theorem 3.3, one can see that our proposed attack strategy can degrade the estimation quality which gives the answer to the second question in Section 2.

3.3 Effectiveness of attack strategy

It can be found that the threshold δ^a is determined by the communication rate $\overline{\gamma}$. Thus another problem is how to drive the optimal communication rate which maximizes the cost deviation under deception attack.

In order to study the effectiveness of our proposed strategy, we define the cost deviation under proposed attack strategy as

$$\Delta J(\delta^a) \triangleq J(\theta^a) - J(\theta^s). \quad (12)$$

In this subsection, we first focus on the problem how does attack threshold δ^a impact the cost deviation $\Delta J(\delta^a)$ and then find out the optimal communication rate such that $\Delta J(\delta^a)$ reaches its maximum.

Property 3.1. Cost deviation $\Delta J(\delta^a)$ has the following properties:

(1) $\Delta J(\delta^a)$ is a continuous function of δ^a in $[0,+\infty)$.
(2) $\Delta J(0) = \Delta J(+\infty) = 0$.

Proof. From (11) and (12), one has

$$\begin{aligned}\Delta J(\delta^a) &= \frac{1}{2}Tr\left[\begin{bmatrix}\Lambda-\rho(\delta^a)\Gamma(\delta^a)-\rho(\delta)\Gamma(\delta) & 0\\ 0 & 0\end{bmatrix}F'F\right] \\ &= \frac{1}{2}Tr\begin{bmatrix}\Lambda-\rho(\delta^a)\Gamma(\delta^a)-\rho(\delta)\Gamma(\delta) & 0\\ 0 & 0\end{bmatrix} \\ &= \frac{1}{2}Tr\Big[\Lambda-\rho(\delta^a)\Gamma(\delta^a)-\rho(\delta)\Gamma(\delta)\Big],\end{aligned}$$

where $\rho(\delta)+\rho(\delta^a)=1$. Then one can easily confirm that statements (1) and (2) are true.

Theorem 3.4. There exists $\delta^{a*} \in (0,+\infty)$ such that

$$\Delta J(\delta^{a*}) = \max_{\delta^a\in[0,+\infty)} \Delta J(\delta^a).$$

Proof. Using Property 3.1 and Lemma C.2 in Appendix, one can obtain this result directly.

Although from Theorem 3.4, the maximum of deviation $\Delta J(\delta^a)$ exists, how to find out this maximum value and the corresponding variable δ^{a*} is still difficult, due to the implicit expression of $\Delta J(\delta^a)$. Hereafter, we will show how to calculate the explicit expression of ΔJ in terms of δ^a.

From (7), it can be seen that

$$\begin{aligned}Tr\Big(\Gamma(\delta^a)\Big) &= Tr\Big(\Lambda^{\frac{1}{2}}\widehat{\Gamma}(\delta^a)\Lambda^{\frac{1}{2}}\Big) = Tr\Big(\widehat{\Gamma}(\delta^a)\Lambda\Big) \\ &= \sum_{i=1}^{r}\lambda_i\widehat{\Gamma}_{i,i}(\delta^a),\end{aligned}$$

where $\widehat{\Gamma}_{i,i}(\delta^a)$ is the (i,i)-th element of matrix $\widehat{\Gamma}(\delta^a)$. From (6), one can see that $\widehat{\Gamma}_{i,i}(\delta^a)$ can be calculated as

$$\begin{aligned}&\widehat{\Gamma}_{i,i}(\delta^a) \\ &= \mathbb{E}[\xi_i^2 \mid \|\xi\| \le \delta^a] \\ &= \frac{1}{r}\mathbb{E}[\xi_1^2+\xi_2^2+\cdots+\xi_r^2 \mid \|\xi\| \le \delta^a] \\ &= \frac{1}{r\rho(\delta^a)}\int_{\sum_{i=1}^r u_i^2 \le (\delta^a)^2}\Big(\sum_{i=1}^r u_i^2\Big)\cdot\prod_{i=1}^r\phi(u_i)du_1\cdots du_r,\end{aligned}$$

where $\mathbf{G}(r) = \int_0^\infty u^r e^{-u}\frac{du}{u}$ is the Gamma function (cf. (Lanczos, 1964)). From polar coordinates transformation

$$\begin{cases}u_1 = \nu\sin\theta_1\cdots\sin\theta_{r-2}\cos\theta_{r-1}, \\ u_2 = \nu\sin\theta_1\cdots\sin\theta_{r-2}\sin\theta_{r-1}, \\ \cdots \\ u_{r-1} = \nu\sin\theta_1\cos\theta_2, \\ u_r = \nu\cos\theta_1,\end{cases}$$

where $(\nu,\theta_1,\ldots,\theta_{r-1},\theta_r) \in [0,\delta^a]\times[0,\pi]\times\cdots\times[0,\pi]\times[0,2\pi]$, one can see that

$$\begin{aligned}\widehat{\Gamma}_{i,i}(\delta^a) = &\frac{1}{r\rho(\delta^a)}\int_0^{\delta^a}\int_0^{\pi}\cdots\int_0^{\pi}\int_0^{2\pi}\nu^2\\ &\cdot\phi(\nu\sin\theta_1\cdots\sin\theta_{r-2}\cos\theta_{r-1})\\ &\cdot\phi(\nu\sin\theta_1\cdots\sin\theta_{r-2}\sin\theta_{r-1})\\ &\cdots\phi(\nu\cos\theta_1)\nu^{r-1}\prod_{k=1}^{r-2}(\sin\theta_k)^{r-k-1}\\ &\cdot d\theta_1\cdots d\theta_{r-1}d\nu,\\ =&\frac{1}{r\mathbf{G}(\frac{r}{2})\rho(\delta^a)}\int_0^{\delta^a}\nu^{r+1}e^{-\frac{\nu^2}{2}}d\nu,\end{aligned}$$

Then one has

$$\begin{aligned}\frac{d}{d\delta^a}\Big[Tr\Big(\rho(\delta^a)\Gamma(\delta^a)\Big)\Big] &= \frac{d}{d\delta^a}\Big[\frac{\sum_{i=1}^r\lambda_i}{r\mathbf{G}(\frac{r}{2})}\int_0^{\delta^a}\nu^{r+1}e^{-\frac{\nu^2}{2}}d\nu\Big]\\ &= \frac{\sum_{i=1}^r\lambda_i}{r\mathbf{G}(\frac{r}{2})}(\delta^a)^{r+1}e^{-\frac{(\delta^a)^2}{2}}.\end{aligned}$$

From Lagrange multiplier method, we define

$$L(\delta,\delta^a,\mu)\triangleq\Delta J(\delta,\delta^a)-\mu(\rho(\delta)+\rho(\delta^a)-1).$$

Let

$$\begin{cases}\dfrac{\partial L}{\partial\delta} = \dfrac{\sum_{i=1}^r\lambda_i}{r\mathbf{G}(\frac{r}{2})}\delta^{r+1}e^{-\frac{\delta^2}{2}}-\mu\dfrac{2^{-\frac{r}{2}+1}}{\mathbf{G}(\frac{r}{2})}\delta^{2r-1}e^{-\frac{\delta^2}{2}}=0,\\ \dfrac{\partial L}{\partial\delta^a} = \dfrac{\sum_{i=1}^r\lambda_i}{r\mathbf{G}(\frac{r}{2})}(\delta^a)^{r+1}e^{-\frac{(\delta^a)^2}{2}}-\mu\dfrac{2^{-\frac{r}{2}+1}}{\mathbf{G}(\frac{r}{2})}(\delta^a)^{2r-1}e^{-\frac{(\delta^a)^2}{2}}\\ \qquad=0,\\ \dfrac{\partial L}{\partial\mu} = \rho(\delta)+\rho(\delta^a)-1=0.\end{cases}\tag{13}$$

Then one can obtain that the solution to (13) is

$$\delta=\delta^a.\tag{14}$$

Theorem 3.5. ΔJ will reach maximum when $\overline{\gamma}=\frac{3}{4}$.

Proof. From (14) one can see that ΔJ will reach maximum when $\delta=\delta^a$. Thus $\rho(\delta)=\rho(\delta^a)$. From (8) and (9) one has $\overline{\gamma}=\frac{3}{4}$.

4. EXAMPLE

Consider system (1) with

$$A=\begin{bmatrix}1.2 & 0.1\\ 0 & 1\end{bmatrix}, C=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}, \Sigma_w=\begin{bmatrix}1 & 0\\ 0 & 2\end{bmatrix}, \Sigma_v=0.5C.$$

By running a Kalman filter at the sensor, one can obtain the steady-state error covariance P as

$$P=\begin{bmatrix}0.3776 & 0.0020\\ 0.0020 & 0.4143\end{bmatrix}.$$

Then one obtains $r=\mathrm{rank}(H)=\mathrm{rank}(h(P)-P)=2$.

The effectiveness of attack strategy θ^a is evaluated by simulation in this section. From Fig.2, one can see that both $J(\theta^a)$, the cost under attack θ^a, and $J(\theta^s)$, the cost without attack, decrease with the increase of sensor-to-estimator communication rate $\overline{\gamma}$. The cost under attack is always larger than that without being attacked which verifies that the proposed attack strategy can degrade the estimation quality. The deviation ΔJ is firstly increasing and then descending as communication rate $\overline{\gamma}$ increases. From Fig.2, one can also see that ΔJ reaches maximum when $\overline{\gamma}=0.75$. It means that when the communication rate is less than 0.75, a larger communication rate will result in more obvious cost deviation, and when the communication rate is more than 0.75, larger communication rate will result in the decrease of cost deviation.

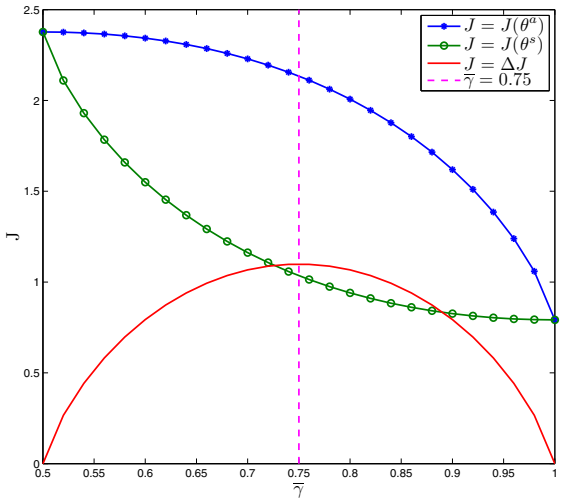

Fig. 2. An example to illustrate the effectiveness of proposed attack strategy.

5. CONCLUSION

In this paper, we investigate online deception attack strategy against remote state estimation with sensor-to-estimator communication rate constraint. We design a new deception attack strategy which can degrade the estimation quality without being detected by the estimator. In order to study the effectiveness of proposed attack strategy, we define the cost deviation as the difference between the cost under attack and that without attack. Our results show that the cost deviation varies with communication rate and reaches the maximum when communication rate is 0.75.

REFERENCES

Anderson, B. and Moore, J. (1981). Detectability and stabilizability of time-varying discrete-time linear systems. *SIAM Journal on Control and Optimization*, 19(1), 20–32.

Bitar, E., Khargonekar, P.P., and Poolla, K. (2011). Systems and control opportunities in the integration of renewable energy into the smart grid. In *Proceedings of World Congress of the International Federation of Automatic Control*, volume 18, 4927–4932.

Bryant, E., Atallah, M., and Stytz, M. (2004). A survey of anti-tamper technologies. crosstalk. *The Journal of Defense Software Engineering*, 17(11), 12–16.

Cardenas, A.A., Amin, S., Sinopoli, B., Giani, A., Perrig, A., and Sastry, S. (2009). Challenges for securing cyber physical systems. In *Proceedings of Workshop on future directions in cyber-physical systems security*.

Kavitha, T. and Sridharan, D. (2010). Security vulnerabilities in wireless sensor networks: A survey. *Journal of information Assurance and Security*, 5(1), 31–44.

Lanczos, C. (1964). A precision approximation of the gamma function. *Journal of the Society for Industrial & Applied Mathematics, Series B: Numerical Analysis*, 1(1), 86–96.

Novak, T. and Gerstinger, A. (2010). Safety and security critical services in building automation and control systems. *IEEE Transactions on Industrial Electronics*, 57(11), 3614–3621.

Qu, F., Wang, F., and Yang, L. (2010). Intelligent transportation spaces: vehicles, traffic, communications, and beyond. *IEEE Communications Magazine*, 48(11), 136 –142.

Sarwate, A. and Chaudhuri, K. (2013). Signal processing and machine learning with differential privacy: Algorithms and challenges for continuous data. *IEEE Signal Processing Magazine*, 30(5), 86–94.

Shi, L., Cheng, P., and Chen, J. (2011a). Sensor data scheduling for optimal state estimation with communication energy constraint. *Automatica*, 47(8), 1693 – 1698.

Shi, L., Johansson, K.H., and Qiu, L. (2011b). Time and event-based sensor scheduling for networks with limited communication resources. In *Proceedings of World Congress of the International Federation of Automatic Control*, volume 18, 13263–13268.

Shoukry, Y., Araujo, J., Tabuada, P., Srivastava, M., and Johansson, K.H. (2013). Minimax control for cyber-physical systems under network packet scheduling attacks. In *Proceedings of the 2nd ACM international conference on High confidence networked systems*, 93–100.

Song, H., Xie, L., Zhu, S., and Cao, G. (2007). Sensor node compromise detection: the location perspective. In *Proceedings of the 2007 international conference on Wireless communications and mobile computing*, 242–247.

Wilson, C. (2013). Cybersecurity and cyber weapons: Is nonproliferation possible? In *Cyber Security*, 11–24. Springer.

Wu, J., Jia, Q., Johansson, K.H., and Shi, L. (2013a). Event-based sensor data scheduling: Trade-off between communication rate and estimation quality. *IEEE Transactions on Automatic Control*, 58(4), 1041–1046.

Wu, J., Yuan, Y., Zhang, H., and Shi, L. (2013b). How can online schedules improve communication and estimation tradeoff? *IEEE Transactions on Signal Processing*, 61(7), 1625–1631.

Zhang, H., Cheng, P., Shi, L., and Chen, J. (2013). Optimal dos attack policy against remote state estimation. In *Proceedings of IEEE Conference on Decision and Control*, 5444–5449.

Zhang, H., Cheng, P., Shi, L., and Chen, J. (2014). Optimal denial-of-service attack scheduling against linear quadratic gaussian control. In *Proceedings of American Control Conference*, to appear.

Appendix A. ONLINE SCHEDULE WITHOUT BEING ATTACKED

From (Wu et al., 2013b), we have the following result for online schedule without being attacked:

Property A.1. Consider the sensor's schedule when the sensor-to-estimator communication constraint $\overline{\gamma} \in [\frac{1}{2}, 1]$ is given. Under sensor's online schedule θ^s (without attack), the expected sensor communication rate is given by

$$\overline{\gamma}(\delta) = 1 - \frac{1}{2}\rho(\delta), \tag{A.1}$$

and the corresponding cost $J(\theta^s)$ is given by

$$J(\theta^s) = Tr\left[P + \frac{1}{2}\rho(\delta)F\begin{bmatrix}\Gamma(\delta) & 0\\ 0 & 0\end{bmatrix}F'\right]. \tag{A.2}$$

Appendix B. PROOF OF LEMMA 3.1

Proof. [Proof of Lemma 3.1] From Lemma 3.3 (Wu et al., 2013b), we have

$$\widehat{\Gamma}(\delta_1) < \lim_{\delta_2 \to \infty} \widehat{\Gamma}(\delta_2) = I.$$

Then it leads to

$$\Gamma(\delta) = \Lambda^{\frac{1}{2}}\widehat{\Gamma}(\delta)\Lambda^{\frac{1}{2}} < \Lambda.$$

Appendix C. PRELIMINARIES FOR PROVING THEOREM 3.4

In order to prove Theorem 3.4, the following lemmas are needed.

Lemma C.1. (Weierstrass Extreme Value Theorem). If a real-valued function f is continuous in $[a, b]$, then f must attain a maximum and a minimum, i.e., there exist $c, d \in [a, b]$ such that

$$f(c) = \max_{x\in[a,b]} f(x), f(d) = \min_{x\in[a,b]} f(x).$$

Lemma C.2. If a real-valued function f is continuous in $[0, +\infty)$, and satisfies $f(0) = f(+\infty) = 0$ and $f(x) > 0$ for all $x > 0$, then f must attain a maximum in $(0, +\infty)$.

Proof. [Proof of Lemma C.2] The proof process consists of three steps:

Step 1. Consider the existence of maximum of $f(x)$ for $x \in [0, 1]$. From Lemma C.1, one can see that there exists $a_1 \in [0, 1]$ such that

$$f(a_1) = \max_{x\in[0,1]} f(x).$$

Since $f(0) = 0$ and $f(x) > 0$ for all $x > 0$, it can be seen that $a_1 \in (0, 1]$.

Step 2. Consider the existence of maximum of $f(x)$ for $x \in [1, +\infty)$. Let $x = \frac{1}{y}$ for $x \in [1, +\infty)$ and

$$g(y) = \begin{cases} f(\frac{1}{y}) = f(x), & y \in (0, 1];\\ 0, & y = 0.\end{cases}$$

One can see that $g(y)$ is continuous in $[0, 1]$ with $g(0) = 0$ and $g(y) > 0$ for $y \in (0, 1]$. Then from Step 1 it can be seen that there must exist $y_1 \in (0, 1]$ such that $g(y_1) = \max_{y\in[0,1]} g(y)$. Thus there exists $a_2 = \frac{1}{y_1} \in [1, +\infty)$ such that

$$f(a_2) = \max_{x\in[1,+\infty)} f(x).$$

Step 3. Consider the problem in $[0, +\infty)$. Integrating Step 1 and Step 2, one can see

$$\max_{x\in[0,+\infty)} f(x) = \max\{f(a_1), f(a_2)\},$$

which completes the proof.