Automatica 87 (2018) 176–183

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

State estimation under false data injection attacks: Security analysis and system protection☆

Liang Hu [a], Zidong Wang [a,1], Qing-Long Han [b], Xiaohui Liu [a]

[a] Department of Computer Science, Brunel University London, Uxbridge, Middlesex, UB8 3PH, UK
[b] School of Software and Electrical Engineering, Swinburne University of Technology, Melbourne, VIC 3122, Australia

ARTICLE INFO

Article history:
Received 11 September 2016
Received in revised form 19 May 2017
Accepted 5 September 2017

Keywords:
False data injection attacks
State estimation
Security analysis
Networked control systems

ABSTRACT

The security issue in the state estimation problem is investigated for a networked control system (NCS). The communication channels between the sensors and the remote estimator in the NCS are vulnerable to attacks from malicious adversaries. The false data injection attacks are considered. The aim of this paper to find the so-called *insecurity* conditions under which the estimation system is insecure in the sense that there exist malicious attacks that can bypass the anomaly detector but still lead to unbounded estimation errors. In particular, a *new* necessary and sufficient condition for the insecurity is derived in the case that all communication channels are compromised by the adversary. Moreover, a specific algorithm is proposed for generating attacks with which the estimation system is insecure. Furthermore, for the insecure system, a system protection scheme through which only a few (rather than all) communication channels require protection against false data injection attacks is proposed. A simulation example is utilized to demonstrate the effectiveness of the proposed conditions/algorithms in the secure estimation problem for a flight vehicle.

1. Introduction

The use of communication networks in networked control systems (NCSs) makes the system vulnerable to cyber-attacks, and the possible malicious attacks on NCSs may cause negative impact on the economy, the environment and the national security. The first-ever cyber-attack in real-world control systems was reported in 2010 (Chen, 2010). Since then, the cybersecurity of NCSs has been a hot topic of research that stirs considerable interest. In general, two kinds of attacks have been studied in NCSs (Teixeira, Shames, Sandberg, & Johansson, 2015). One is the denial-of-service (DoS) attack that violates data *availability* through blocking information flows between different components of NCSs, and the other is the deception attack that violates data *integrity* through modifying data packets. Compared with DoS attacks, deception attacks are more difficult to detect because the adversary could keep the deception attacks stealthy to the anomaly detector in NCSs through deliberately designing the attack sequences.

The deception attacks have been first considered in Liu, Ning, and Reiter (2009) for the state estimation problems of power systems modelled by *static* system models. The minimum number of comprised sensors that needed to launch deception attacks has been investigated in Teixeira, Amin, Sandberg, Johansson, and Sastry (2010). As for *dynamic* systems, when the system model is unknown to the adversary, a specific type of deception attack called replay attack has been investigated in Mo, Weerakkody, and Sinopoli (2015) and Zhu and Martínez (2014). In the case that the dynamic model is known to the adversary, another type of deception attack, namely, false data injection attack, has recently been put forward. For deterministic systems without stochastic noises, fundamental issues such as detectability and identifiability for false data injection attacks have been analysed in Chong, Wakaiki, and Hespanha (2015), Fawzi, Tabuada, and Diggavi (2014), Pasqualetti, Dorfler, and Bullo (2013) and efficient control/estimation algorithms have been developed against false data injection attacks (Shoukry & Tabuada, 2016). In Pang and Liu (2012), a data encryption scheme (together with time-stamp techniques) has been adopted to detect the deception attacks and compensate the side-effects.

As is well known, stochastic models have come to play a more and more important role in characterizing noisy phenomena from

☆ This work was supported in part by the Royal Society of the UK, the National Natural Science Foundation of China under Grant 61329301, the Australian Research Council Discovery Project under Grant DP160103567, and the Alexander von Humboldt Foundation of Germany. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Antonio Vicino under the direction of Editor Torsten Söderström.

E-mail addresses: l.hu@lboro.ac.uk (L. Hu), Zidong.Wang@brunel.ac.uk (Z. Wang), qhan@swin.edu.au (Q. Han), XiaoHui.Liu@brunel.ac.uk (X. Liu).
[1] Fax: +44 1895 251686.

https://doi.org/10.1016/j.automatica.2017.09.028
0005-1098/

real-world systems. Accordingly, it is of practical significance to investigate the cybersecurity of stochastic dynamic systems. As pointed in Kosut (2013), the detection task of malicious behaviours for stochastic systems (with external noises) is more difficult than that for deterministic (without stochastic noises) due to the fact that the injected attack by the adversary could be mistaken as a type of noises by the protection devices. Based on the setting that the smart sensors send innovation information (rather than measurements) to the remote estimator, explicit forms of optimal attacks on remote state estimation have been presented in Guo, Shi, Johansson, and Shi (2017), while several attack detection methods for multi-sensor remote estimation have been proposed in Li, Shi, and Chen (in press). A secure state estimation algorithm has been proposed in Mishra, Shoukry, Karamchandani, Diggavi, and Tabuada (2017) for stochastic dynamic systems where a key assumption of sparse observability has been made. While the results reported in Mishra et al. (2017) are indeed interesting, it is quite possible that the adversary attacks at a large number of (or even all) sensors, in which case the system cannot be guaranteed to be "sparsely observable". Motivated by the above observation, we aim to investigate the case where the attacker could inject false data into measurements from any sensor and, accordingly, the results obtained would constitute one of the main contributions of our paper.

In this paper, we focus on the remote state estimation problem for a class of *stochastic* systems under possible false data injection attacks where a χ^2 detector is employed to monitor the state estimates. Note that false data injection attacks have been considered in Kwon, Liu, and Hwang (2013), Mo, Garone, Casavola, and Sinopoli (2010) and Mo and Sinopoli (2010, 2016) for state estimation problems of stochastic systems equipped with χ^2 detectors. In particular, an approximation method has been proposed in Mo et al. (2010) and Mo and Sinopoli (2016) to analyse the cybersecurity of the system by calculating the estimation error bound caused by the malicious attacks, and some *insecurity* conditions have been derived in Kwon et al. (2013) and Mo and Sinopoli (2010) to determine whether or not there exist malicious attacks which can cause unbounded estimation error for the state estimation system. Nevertheless, a thorough investigation reveals that (1) there is still room to improve the existing insecurity conditions; and (2) there is also an engineering need to develop a system protection scheme by using only necessary number of communication channels requiring protection against cyber-attacks.

In this paper, we aim to propose new insecurity conditions for the state estimation problem under false data injection attacks. Specifically, in the case when all communication channels are compromised by the adversary, we propose a *new* necessary and sufficient condition under which the system is insecure in the sense that the estimation error caused by attacks is unbounded. Such a new condition is shown to be concise that simplifies the existing results. In the case when only parts of the communication channels are compromised by the adversary, a sufficient condition is proposed as well. Furthermore, we propose a criterion which determines a sufficient number of communication channels that require protection. According to the criterion, only necessary number of (rather than all) communication channels need to be protected in order to make the overall system secure against the attacks.

The contributions of the paper are summarized as follows: (1) *new security criteria are proposed for state estimation systems under false data injection attacks and, in the case that all communication channels are compromised by the adversary, our criteria are shown to be necessary and sufficient that simplify the existing ones*; (2) *an effective protection scheme is proposed for the system which is insecure under false data injection attacks*; *and* (3) *the developed criteria are applied to security analysis and system protection in the state estimation system of a flight vehicle*.

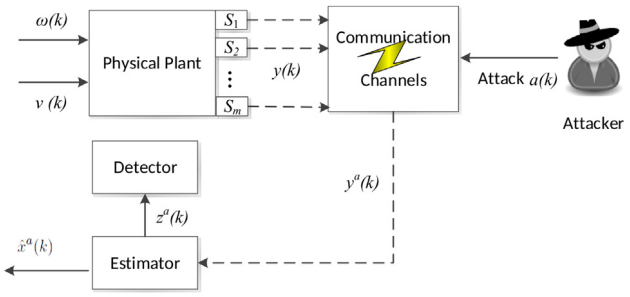

Fig. 1. Diagram of state estimation under cyber-attacks.

Notation: $\mathbb{N}$, $\mathbb{R}$ and $\mathbb{C}$ denote, respectively, the set of non-negative integers, the set of all real numbers, and the set of all complex numbers. $\{x(k)\}$ denotes an infinite sequence $x(1), x(2), \ldots, x(k), \ldots$. $\mathbb{R}^{n\times m}$ ($\mathbb{C}^{n\times m}$) denotes the set of all $n \times m$ real (complex) matrices, and $\mathbb{R}^n$ denotes the n dimensional Euclidean space. For $\alpha \in \mathbb{C}$, $\mathrm{Re}(\alpha)$ and $|\alpha|$ denote its real part and its modulus, respectively. For $a \in \mathbb{R}^n$, $\|a\|$ denotes its l_2 norm. For a matrix $P \in \mathbb{R}^{n\times m}$, P^T, P^{-1}, $\mathrm{Tr}\{P\}$ and $\mathrm{Rk}\{P\}$ represent its transpose, inverse, trace, and rank, respectively. For square matrix A, $\det(A)$ stands for the determinant of A, and $\rho(A)$ stands for the spectral radius of A. $\mathrm{diag}\{\cdots\}$ and I denote a block-diagonal matrix and identity matrix of compatible dimension, respectively, and I_m (0_m) denotes the $m \times m$-dimensional identity (zeros) matrix. I_m^s denotes the sth column of $m \times m$-dimensional identity matrix I_m, e.g., $I_m^s = [\overbrace{0, \ldots, 0}^{s-1}, \underbrace{1, 0, \ldots, 0}_{m}]^T$.

2. Problem formulation

In this section, we describe the model of false data injection attack and analyse how the injected attacks affect the estimation system. The structure of the state estimation system under cyber-attacks is shown in Fig. 1. For presentation convenience, we first introduce the estimation system without cyber-attacks (i.e., $y^a(k) = y(k)$ in Fig. 1).

2.1. State estimation without cyber-attacks

Let the physical plant be given by

$$\mathcal{P}: \begin{cases} x(k+1) = Ax(k) + \omega(k) \\ y(k) = Cx(k) + v(k) \end{cases} \tag{1}$$

where $x(k) \in \mathbb{R}^n$, $y(k) = [y_1(k), \ldots, y_m(k)]^T \in \mathbb{R}^m$ are the system state and measurement output, respectively, and $y_i(k)$ is the output of the ith sensor (labelled as S_i in Fig. 1) at time instant k. The initial state $x(0)$ has mean $\bar{x}(0)$ and covariance $\Sigma(0)$, the process noise $\omega(k) \in \mathbb{R}^n$ and the measurement noise $v(k) \in \mathbb{R}^m$ are assumed to be mutually uncorrelated zero-mean random signals with known covariance matrices W and R, respectively. The process noise $\omega(k)$ represents the external disturbance on dynamic systems and the measurement noise $v(k)$ characterizes the error of sensor and/or measurement process, respectively.

The following time-invariant state estimator is proposed:

$$\mathcal{E}: \begin{cases} \hat{x}(k+1) = A\hat{x}(k) + Kz(k+1) \\ z(k+1) = y(k+1) - CA\hat{x}(k) \end{cases} \tag{2}$$