3rd International Conference on Automation, Control, Engineering and Computer Science (ACECS'16)

Proceedings of Engineering & Technology (PET)
pp. 572-578

Copyright IPCO-2016

A Model-Based Stealthy Attacks Detection Scheme for Networked Control Systems

Taouba Rhouma[1], Karim Chabir[2], Mohamed Naceur Abdelkrim[3]

University of Gabes
Research Unit of Modeling, Analysis and Control Systems MACS 06/UR/11-12
National Engineering School of Gabes, 6029 Gabes, Tunisia.
[1]taouba.rhouma@gmail.com
[2]karim.chabir@yahoo.fr
[3]naceur.abdelkrim@enig.rnu.tn

***Abstract*— since most cyber attacks happen in stealthy ways, it is difficult to detect them. In this paper we are focused on the problem of stealthy attacks detection for networked control systems (NCS). The first part of this paper describes a deception attack strategy that a malicious agent located inside the network of a NCS can use to act on the physical world while remaining undetectable. The second part shows how to reveal the presence of the adversary by using a secure communication channel for the transmission of critical measurements. The results of the detection scheme are illustrated through numerical examples.**

***Keywords*— Networked control systems, Cyber Physical Systems, Kalman filtering, Linear Quadratic Gaussian control, Anomaly detector**.

I. Introduction

With the rapid advancements of technology and novel control strategies, Networked Control Systems (NCSs) have been at the core of several infrastructure systems and industrial plants. Transport systems, electrical power systems, chemical processes, water and gas distribution networks, manufacturing and transportation networks can be considered as examples of application areas of Cyber-physical systems (CPS). CPS is an integration of communication capabilities, computational resources and physical processes. In general, CPS consists of a group of agents including sensors, actuators, communication network. Such systems are often considered as large scale distributed physical processes and can be monitored by using can be monitored and controlled using a supervisory control and data acquisition (SCADA) software. The design of control systems taking into account the effects of packet losses and packet delays in NCS have been discussed in [1]. Besides several network-induced effects such as time-delays and packet losses, NCSs become vulnerable to Cyber physical attacks incorporating Cyber and physical activities into a malicious attack. Recently, a sharp rise in the number of Cyber attacks has been reported. Consequently, many researchers have shown a great concern for the analysis of vulnerabilities of CPS integrating physical processes, computational resources, and communication capabilities to external attacks (see [2]). For instance, in [3] Denial of Service (DoS) attacks against a networked control system are defined when the adversary prevents the controller from receiving sensor measurement or the plant from receiving control law. In [4] and [5] deception attacks (also called false data injection attacks) are introduced when the adversary sends false information on sensors or actuators. Replay attacks are discussed in [6] when the adversary generates artificial measurement delays. The effects of covert attacks against control systems are investigated in [7] and [8] when the adversary takes the control of the plant. Direct physical attacks on the plant (including sensors and actuators) close to traditional faults are taken into account by Fault Detection and Isolation (FDI) techniques as discussed in [9]-[13].

The detection problem of coordinated attacks in CPS seems closely related to the detection problem of multiple components, sensors or actuators faults, but there exists a significant difference. Multiple faults are considered as a phenomenon which occurs randomly on actuators, sensors or communication channels while a coordinated attack is an intentional action designed by adversaries to remain undetectable from traditional model-based FDI schemes. In this new context, it is necessary to design a new generation of FDI schemes having the ability to detect the presence of coordinated attacks. Kalman filtering has played a significant role in systems theory and has been produced on wide variety of application domains. Over the last three decades, there has been an increasing research interest in the problem of optimal state filtering in the presence of random loss of observations represented by Markovian or Bernoulli processes, see for examples [14]-[18]. In order to detect the cyber attacks, the monitoring system compare a sequence of the compromised data to the expected output of the healthy system. Since the attacker has a good knowledge of the system, he can design an advanced deception attack that could not be detected by the monitoring system. If the system does not know that there is a cyber attack, it cannot protect itself against the attack.

The first part of this paper briefly explains how a false data injection on the control signal generated by a Linear Quadratic Gaussian (LQG) controller can be designed to act on the state variables of the NCS and remain quasi undetectable to any passive detector applied on the innovation sequence of the Kalman filter. In order to monitor the occurrence of such attack, the second part proposes to destroy the stealthy strategy of the attacker by using a secure communication channel for the transmission of critical measurements. The monitoring of the plant is realized by a chi-squared detector applied on the innovation sequence of the Kalman filter with intermittent measurements.

The paper is organized as follows: Section 2 presents a stealthy attack strategy close to covert attack that a malicious agent can use to successfully realize his attack without being detected. Section 3 presents the proposed stealthy attack detection scheme. Several numerical examples are presented in section 4 before to conclude in section 5.

II. PROBLEM STATEMENT

Consider the NCS of Fig. 1:

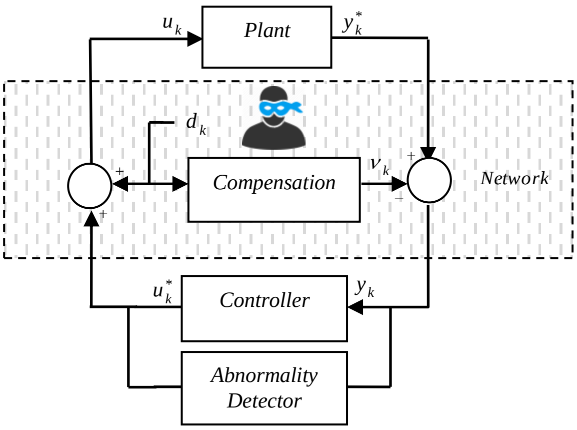

Fig.1. NCS subject to covert attack.

The plant is represented by a linear discrete-time stochastic system

$$x_{k+1} = Ax_k + Bu_k + w_k \tag{1.a}$$

$$y_k^* = Cx_k + \varepsilon_k \tag{1.b}$$

where $x_k \in \Re^n$, $u_k \in \Re^q$ and $y_k^* \in \Re^m$ are the state, input and measurement vectors and where $w_k \in \Re^n$ and $\varepsilon_k \in \Re^m$ are zero mean uncorrelated Gaussian random sequences with

$$E\left\{\begin{bmatrix} w_k \\ \varepsilon_k \end{bmatrix}\begin{bmatrix} w_j \\ \varepsilon_j \end{bmatrix}^T\right\} = \begin{bmatrix} W & 0 \\ 0 & V \end{bmatrix}\delta_{k,j} \text{ with } W \geq 0 \tag{1.c}$$

The initial state x_0 , assumed to be uncorrelated with w_k and v_k , is a Gaussian random variable with $E\{x_0\} = \bar{x}_0$ and $P_0 = E\{(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T\} \geq 0$. The pair (A, C) is detectable and (A, B) is stabilizable. Designed under the assumption that $u_k = u_k^*$ and $y_k^* = y_k$ $\forall k \geq 0$, with Q , R are positive semi definite matrices, the LQG control law solution to

$$J = \min \lim_{T\to\infty} E\left\{\frac{1}{T}\left[\sum_{k=0}^{T-1} x_k^T Q x_{k_k} + u_k^T R u_k\right]\right\} \tag{2.a}$$

is generated by

$$u_k^* = -L\hat{x}_{k/k} \tag{2.b}$$

with

$$L = (B^T SB + R)^{-1} B^T SA \tag{2.c}$$

$$S = A^T SA + Q - A^T SB(B^T SB + R)^{-1} B^T SA \tag{2.d}$$

where $\hat{x}_{k/k}$ is the state estimation given by the Kalman filter taking the following recursive form:

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + K_k(y_k - C\hat{x}_{k/k-1}) \tag{3.a}$$

$$P_{k/k} = (I - K_k C)P_{k/k-1}(I - K_k C)^T + K_k V K_k^T \tag{3.b}$$

$$K_k = P_{k/k-1} C^T (CP_{k/k-1} C^T + V)^{-1} \tag{3.c}$$

$$\hat{x}_{k+1/k} = A\hat{x}_{k/k} + Bu_k \tag{3.d}$$

$$P_{k+1/k} = AP_{k/k} A^T + W \tag{3.e}$$

Initialized with $\hat{x}_{0/-1} = x_0$ and $P_{0/-1} = P_0$, the monitoring of the plant is realized by a chi-squared detector applied on the innovation sequence $\gamma_k = y_k - C\hat{x}_{k/k-1}$ of the Kalman filter

$$T_k = \sum_{j=k-(N-1)}^{k} \gamma_j^T (CP_{j/j-1} C^T + V)^{-1} \gamma_j \quad \underset{\textit{Normality}}{\overset{\textit{Abnormality}}{\gtrless}} \quad \mu \tag{4}$$

where μ is the threshold level, $\alpha = \int_0^{\mu} \chi(0, l)dx$ the desired rate of false alarm computed from the central chi-squared distribution $\chi(0, l)$ with $l = mN$ degrees of freedom.

In deception attacks, the adversary attempts to prevent the actuator or the sensor from receiving an integrity data. His goal is to modify the control action or the sensor measurements from their real values by sending false information from controllers or sensors. The false information can be a wrong sender identity, an incorrect sensor measurement or control input, an incorrect time when a measurement was observed. The attacker can also inject a bias data that cannot be detected in the system and launch these attacks by obtaining the secret keys or by compromising some controllers or sensors.

In our work, we assume that the malicious agent acts on the plant by adding the false data d_k on the control signal u_k^* and by subtracting the false data v_k from the sensor's readings y_k^*.

Assume that to compute the appropriate attack policy the attacker has access to the detailed model of the system.

To remain undetectable to the abnormality detector, the attacker can simultaneously trigger at time r the false data injection $\{d_r, d_{r+1}, \ldots, d_k\}$ $\forall k \geq r$ and the compensation sequence $\{v_r, v_{r+1}, \ldots, v_k\}$ $\forall k \geq r$ computed as

$$\Delta x_{k+1} = A\Delta x_k + Bd_k \tag{5.a}$$

$$\nu_k = C\Delta x_k \tag{5.b}$$

with $\Delta x_r = 0$ where $\Delta x_k \quad \forall k \geq r$ represents the additive consequence of the attack on the state of the plant.

Contrary to a zero dynamic attack conditioned by the existence of invariant zeros, the false data injection $\{d_r, d_{r+1}, .., d_k\}$ $\forall k \geq r$ and its compensation sequence $\{\nu_r, \nu_{r+1}, .., \nu_k\}$ $\forall k \geq r$ computed by (2) do not need particular structural properties of the plant and can be designed by the adversary even if $\ker(C) = 0$, for example when the plant is over-measured with $m \geq n$.

III. Detection of stealthy deception attack

It has been shown in section 2 how a false data injection on the control signals can be designed to act on the state variables of the NCS while remaining undetectable to any passive detector applied on the innovation sequence of the Kalman filter. In this section, our objective is to create a detection strategy to reveal the presence of the adversary by using a secure communication channel for the transmission of critical measurements.

Consider the NCS of Fig.2 subject to covert attack and packet dropouts:

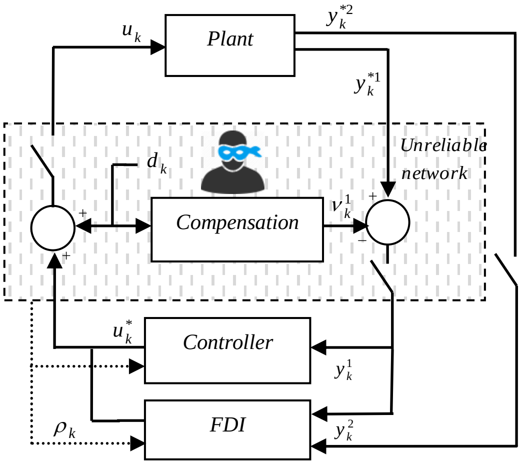

Fig. 2. NCS subject to covert attack and packet dropouts.

where the plant is represented by

$$x_{k+1} = Ax_k + Bu_k + w_k \tag{7.a}$$

$$y_k^{*1} = C_1 x_k + \varepsilon_k^1 \tag{7.b}$$

$$y_k^{*2} = C_2 x_k + \varepsilon_k^2 \tag{7.c}$$

with $x_k \in \Re^n$, $u_k \in \Re^q$, $y_k^{*1} \in \Re^{m1}$ and $y_k^{*2} \in \Re^{m2}$ are the state, the input and the measurements vectors, and where $w_k \in \Re^n$, $\varepsilon_k^1 \in \Re^{m1}$ and $\varepsilon_k^2 \in \Re^{m2}$. y_k^{*2} is the secure communication channel designed for the transmission of critical measurements so that the attacker cannot affect them.

The nominal system model of the plant viewed by the bloc FDI is described as follows

$$x_{k+1} = Ax_k + \rho_k^1 Bu_k + w_k \tag{8.a}$$

$$y_k^1 = \rho_k^1 (C_1 x_k + \varepsilon_k^1) \tag{8.b}$$

$$y_k^2 = \rho_k^2 (C_2 x_k + \varepsilon_k^2) \tag{8.c}$$

where the binary variables $\{\rho_k^1, \rho_k^2\} \in \{0,1\}$, assume to follow a random Bernoulli processes with $\lambda_1 = \Pr[\rho_k^1 = 1]$ and $\lambda_2 = \Pr[\rho_k^2 = 1]$, represents the acknowledgement signal indicating the status of reception/delivery with $\rho_k^i = 1$ when the control signal y_k^{*i} transmitted by the sensor is received by the plant or $\rho_k^i = 0$ when y_k^{*i} is lost on the unreliable network with $i \in \{1, 2\}$.

When the attacker knows (8.a) and (8.b) (but not (8.c)), its stealthy strategy is given by

$$\Delta x_{k+1} = A\Delta x_k + \rho_k^1 Bd_k \tag{9.a}$$

$$\nu_k^1 = C_1 \Delta x_k \tag{9.b}$$

Designed on (8.a), (8.b) and (8.c), the Kalman filter with intermittent measurements equations are described as follows

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + K_k (Q_k y_k - C_k \hat{x}_{k/k-1}) \tag{10.a}$$

$$P_{k/k} = (I - K_k C_k) P_{k/k-1} (I - K_k C_k)^T + K_k Q_k V Q_k^T K_k^T \tag{10.b}$$

$$K_k = P_{k/k-1} C_k^T (C_k P_{k/k-1} C_k^T + Q_k V Q_k^T)^{-1} \tag{10.c}$$

$$\hat{x}_{k+1/k} = A\hat{x}_{k/k} + \rho_k^1 Bu_k \tag{10.d}$$

$$P_{k+1/k} = AP_{k/k} A^T + W \tag{10.e}$$

with

$$C_k = Q_k \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} \tag{11.a}$$

$$Q_k = \rho_k^1 \rho_k^2 \begin{bmatrix} I_{m_1} & 0 \\ 0 & I_{m_2} \end{bmatrix} + \rho_k^1 (1 - \rho_k^2) \begin{bmatrix} I_{m_1} & 0 \end{bmatrix} + (1 - \rho_k^1) \rho_k^2 \begin{bmatrix} 0 & I_{m_2} \end{bmatrix} + (1 - \rho_k^2)(1 - \rho_k^1) 0 \tag{11.b}$$

where

$$\gamma_k = Q_k y_k - C_k \hat{x}_{k/k-1} = Q_k (y_k - C\hat{x}_{k/k-1}) \tag{12}$$

is the innovation sequence given information about the attack at the current time if and only if $\rho_k^2 = 1$.

We propose to detect stealthy attacks by applying on (12) the chi-squared detector

$$T_k = \sum_{j=k-(N-1)}^{k} \gamma_j^T (C_j P_{j/j-1} C_j^T + Q_j V Q_j^T)^{-1} \gamma_j \underset{\text{No attack}}{\overset{\text{Attack}}{\underset{<}{\geq}}} \mu_k \tag{13}$$

where the threshold detection level μ_k depends on N and the binary sequence $\left\{\rho_j^1, \rho_j^2\right\}$ for $j \in [k-N \quad k]$.

The proof that the attacker can perform his malicious act while forcing the system out of its safe operating region without being detected from any anomaly detector are established via a the first simulation example given in the following section. Another illustrative example will be given to prove that the proposed detection attack strategy works very well with such attacks.

IV. Illustrative examples

After having modeled the NCS under false data injection attacks and provided the detection strategy, two simulation examples are given in this section to demonstrate the effectiveness of the obtained results. We illustrate firstly how the attacker can successfully realise his malicious act while remaining undetectable from passive detectors. We then apply the proposed detection scheme of section 3 to detect the presence of such attacks.

To illustrate the effect of this attack, consider first the NCS of Fig.1 described by:

$$A = \begin{bmatrix} 0.7 & 0.4 & 0.1 & 0 \\ 0 & 0.8 & 0.3 & 0.6 \\ 0 & 0 & 0.9 & 0.2 \\ 0 & 0 & 0 & 0.8 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (14)$$

The stealthy constant bias injection attack $d_k = [20 \quad 30]^T$ $\forall k \geq r$ and its compensation sequence computed by (5.b) are triggered at time $r = 100$.

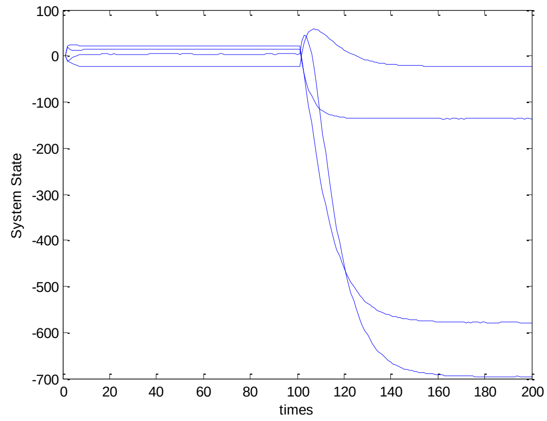

Fig.3: State x_k .

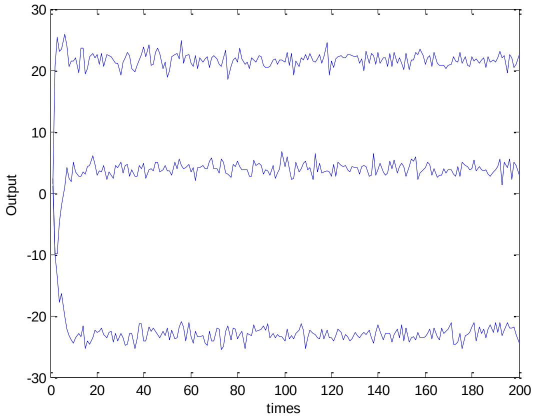

Fig.4: Measurement y_k .

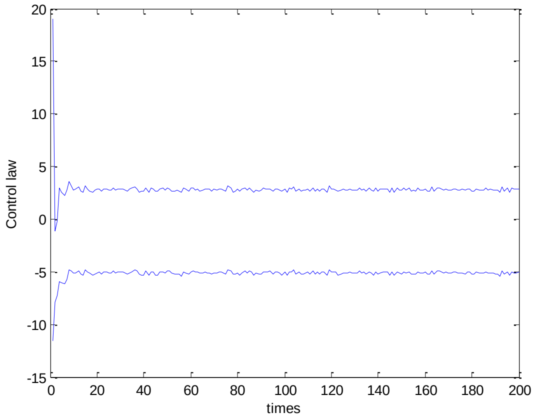

Fig.5: Control law u_k^* .

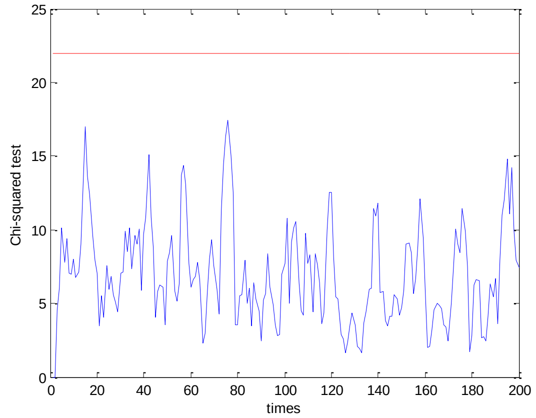

Fig.6: Abnormality detector.

Noting that the attack occurs at the time $r = 100$, The Fig. 3 shows the consequences of the attack on the system's state, which is invisible on the measurement given by Fig.4. One can see in the next figure Fig.5 that despite the presence of the attack, the control law is not altered. The Fig.6 shows that the attack cannot be detected using the detection variable. This demonstrates that an attacker located inside the network of a

NCS can provide malicious consequences on the system's state using a stealthy deception strategy without being detected.

To prove the usefulness of the proposed detection strategy, we consider NCS subject to covert attack and packet dropouts of Fig.2 where the plant is described by:

$$A=\begin{bmatrix}0.6 & 0.34 & 0.35 & 0\\ 0 & 0.8 & 0 & 0.37\\ 0 & 0 & 0.5 & 0\\ 0 & 0 & 0 & 0.9\end{bmatrix},\quad B=\begin{bmatrix}1 & 0 & 0\\ 1 & 0 & 1\\ 0 & 0 & 2\\ 0 & 1 & 1\end{bmatrix},\quad C=\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\end{bmatrix} \qquad (15)$$

The stealthy constant bias injection attack $d_k=[-10 \quad 20 \quad -30]^T \quad \forall k \geq r$ and its compensation sequence computed by (9.b) are triggered at time $r=80$. The occurrence rate of data losses fixed at $\lambda_1=\lambda_2=0.7$. The random binary variable ρ_k^1 and ρ_k^2 are plotted on Fig.7 and Fig.8, respectively. Consequences of the attack on the state variables, output measurements and control signal are plotted on Fig.9, Fig.10 and Fig.11, respectively. As we can see on Fig.12, this strategy allow detecting the presence of the stealthy attack under $\rho_k=1$ when the detection variable T_k exceeds the threshold levels of significance values.

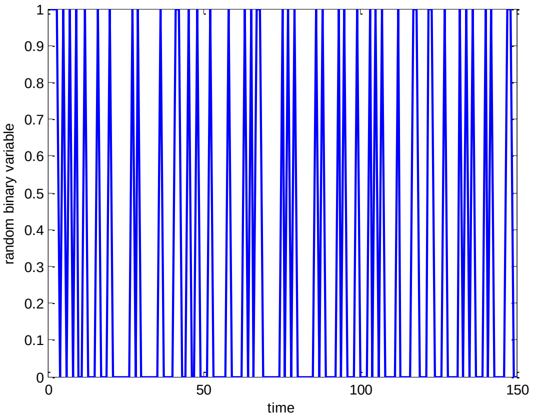

Fig.7: The random binary variable ρ_k^1 .

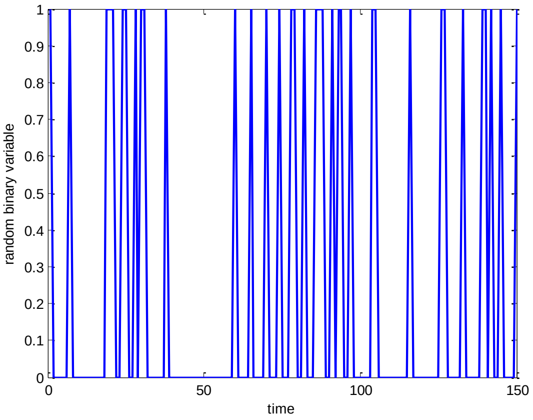

Fig.8: The random binary variable ρ_k^2 .

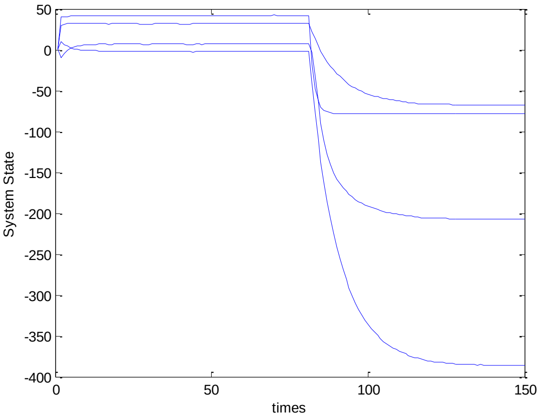

Fig.9: State x_k .

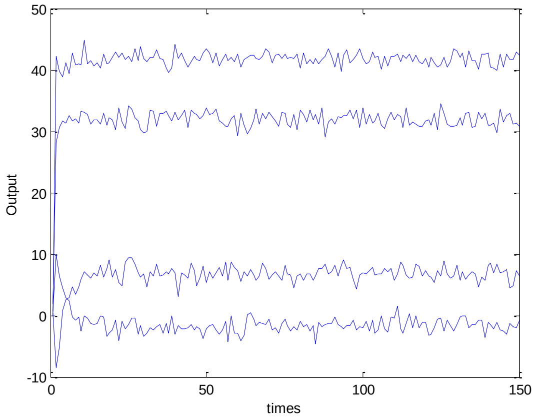

Fig.10: Measurement y_k .

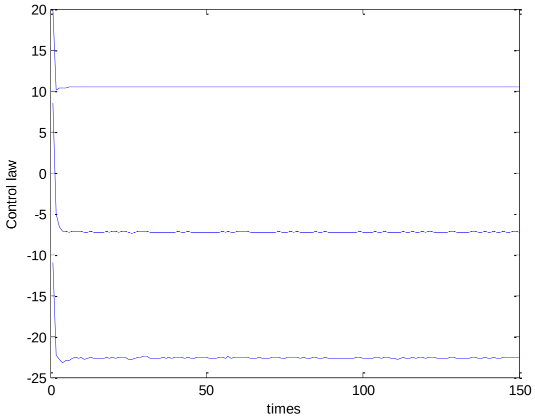

Fig.11: Control law u_k^* .

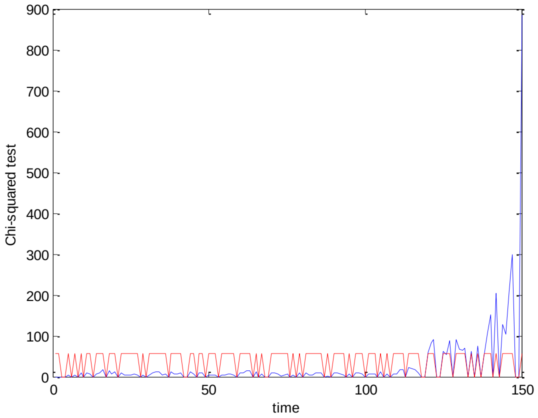

Fig.12: The detection variable T_k and its threshold level μ_k

Detection and isolation of multiple attacks allowing the design of a resilient controller in the presence of actuator redundancy is open question yet. Future works will concern the design of distributed resilient controllers for large scale NCS decomposed into subsystems as explained in [19]. The Kalman filter with intermittent unknown input, designed in [20] will help the design of distributed resilient controllers under denial of service attacks on information exchanged between subsystems.

V. Conclusion

The problem of stealthy attacks detection in networked control systems (NCSs) was tackled in this paper. First, we have described how an adversary could generate malicious attacks on the system by using a deception attack strategy without being detected. Then we have provided a technique that a defender can use to detect the presence of the attacker by using the chi-squared detector applied on the innovation sequence of the Kalman filter with intermittent measurements. Two numerical simulation examples are given to prove the effectiveness of the obtained results.

References

[1] J.P. Hespanha, P. Naghshtabrizi and Y. Xu , “Survey of recent results in networked control systems,” *Proceedings of IEEE*, vol. 95, pp. 138–162, 2007.

[2] A. Cardenas, S. Amin and S. Sastry, “Secure control: Towards survivable cyber-physical systems,” *First International Workshop on Cyber-Physical System*, Beijing, China, pp. 495-500, 2008.

[3] S. Amin, A. Cardenas and S. Sastry, “Safe and secure networked control systems under denial-of-service attacks,” Hybrid Systems: Computation and Control, vol. 5469, pp. 31–45, 2009.

[4] Y. Liu, M.K. Reiter and P. Ning, “False data injection attacks against state estimation in electric power grids,” ACM Conference on Computer and Communications Security, Chicago, IL, USA, pp. 21–32, 2009.

[5] A. Teixeira, H. Sandberg and K.H. Johansson, “Networked control system under cyber attacks with applications to power networks,” American Control Conference, Baltimore, pp. 3690-3696, 2010.

[6] Y. Mo and B. Sinopoli, “Secure control against replay attacks,” Allerton Conf. on Communications, Control and Computing, Monticello, IL, USA, pp. 911–918, 2010.

[7] R. Smith, “A decoupled feedback structure for covertly appropriating network control systems,” IFAC World Congress, Milan, Italy, pp. 90–95, 2011.

[8] A. Teixeira, Iman Shames, Henrik Sandberg and K.H. Johansson, “Revealing stealthy attacks in control systems,” 50th Annual Allerton Conference on Communication, Control, and Computing, 2012.

[9] R.K. Mehra and J. Peshon, “An innovation approach to fault detection and diagnosis in dynamic systems,” *Automatica*, vol. 7, pp. 637-640, 1971.

[10] M. Basseville and I. Nikiforov, *Detection of abrupt changes: Theory and application*, Prentice Hall, Thomas Kailath Editor, 1994.

[11] J. Chen and R.J. Patton, *Robust Model based fault diagnosis for dynamic systems*, Kluwer Academic Publishers, 1999.

[12] F. Gustafsson, “Statistical signal processing approaches to fault detection,” *Annual Reviews in Control*, vol. 31, pp. 41-54, 2007.

[13] J.Y. Keller and D. Sauter, “Restricted diagonal detection filter and updating strategy for multiple fault detection and isolation,” *International Journal of Adaptive Control and Signal processing*, vol. 25, pp. 68-87, 2011.

[14] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. Jordan and S. Sastry, “Kalman filtering with intermittent observations,” *IEEE transactions on Automatic Control*, vol. 49, pp. 1453-1464, 2004.

[15] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla and S. Sastry, “Foundations of control and estimation over lossy networks,” *Proceedings of IEEE*, vol. 95, pp. 163–187, 2007.

[16] J.Y. Keller and D. Sauter, “Kalman filter for discrete-time stochastic linear systems subject to intermittent unknown inputs,” To appear in IEEE TAC.

[17] D. Sauter and J.Y. Keller, “Intermittent bias and state filtering for linear systems subject to deception attacks in communication networks,” Submitted to the IEEE SMC conference on System, Man and Cybernetics, 2013.

[18] X. Liu and A. Goldsmith, “Kalman Filtering with Partial Observation Losses”, 43rd IEEE Conference on Decision and Control, 2004.

[19] D. Sauter, T. Boukhobza and F. Hamelin, “Decentralized and autonomous design for FDI/FTC of networked control systems,” IFAC Symposium SAFEPROCESS, pp. 163-168, Beijing, China, 2006.

[20] J.Y. Keller and D. Sauter, “Kalman filter for discrete-time stochastic linear systems subject to intermittent unknown inputs,” IEEE transactions on Automatic Control, vol. 58, pp. 1882-1887, 2013.