Cyber-Physical Attacks Detection in Networked Control Systems with Limited Communication Bandwidth

Cyber-Physical Attacks Detection in Networked Control Systems with Limited Communication Bandwidth

Eman Mousavinejad[1], Fuwen Yang[2], Qing-Long Han[3], and Ljubo Vlacic[4]

***Abstract*— This paper is concerned with cyber-physical attack detection problem in networked control systems subject to limited communication bandwidth. This constraint arises when an attack detection system is located at a remote site and so the required signals, measurement output and control signals, need to be transmitted over a digital communication channel. Therefore, data before being sent to the remote site must be encoded and converted from analog signals to digital signals by using quantizer. A quantizer maps the amount of information from a continuous space to a finite set which is compatible with the limited communication bandwidth. Considering the quantized measurement output, a detection algorithm by means of a set-membership filtering approach will be proposed. The algorithm consists of a prediction ellipsoid set and an estimation ellipsoid set updated with the quantized measurement output. The detection method depends on the existence of intersection between two sets computed by the filter. Simulation results for some possible physical and cyber attacks are provided to demonstrate the effectiveness of the proposed method.**

***Index Terms*— Cyber-physical attack detection, Networked control systems, Set-membership filtering, quantization.**

I. INTRODUCTION

A networked control system (NCS) is such a system whose components, i.e. controllers, sensors, and actuators, are implemented distributively. In an NCS, information among system components is exchanged via a shared communication network. As a result, an NCS has several advantages over traditional control systems such as low installation and maintenance costs, elimination of unnecessary wiring, high reliability, and more flexibility in upgrading the system with new components [1].

However, the use of communication networks usually introduces some challenging issues. To mention a few, the growing trend towards using pervasive computing systems, such as the Internet and wireless communication, leads NCSs to become dramatically vulnerable to attacks that violate both their physical infrastructure and the communication layer [2]. Due to the crucial role of NCSs in industrial control systems, security of NCSs needs to be properly addressed to ensure that the system is operating in a safe manner.

Furthermore, in NCSs some communication issues are induced by the network due to wireless connections. Generally speaking, three common induced communication issues include network-induced delay, packet dropout and quantization effect [1]. As this paper is concerned with the problem of attack detection using quantized measurement output, quantization effects on the performance of the proposed attack detection system is the only network-induced communication issue considered in this study.

Signal quantization is a natural phenomenon in digital control systems and it is inevitable in NCSs due to the widely use of digital computing technology. In NCSs, sensing data must be encoded by transforming analog signals to digital signals before being transmitted to a remote site via a network [3]. As design of the controller is beyond the scope of this paper, the controller is considered to be at the same site as the plant and, thus, providing the physical plant with a known control signal. As quantization may result in limited communication bandwidth and unpredictable behaviour, quantization may have some effects on the performance of the attack detection system considered to be in the remote site of an NCS in this paper.

Quantized state estimation in NCSs has received considerable attention. The set-valued state estimation problem is discussed for continuous-time systems in [4]. They considered that the observation must be transmitted via a communication channel with a limited capacity. The proposed state estimation utilizes uniform quantization of the state of the system. In [5], a reset state estimator is presented to improve the position estimation for motion control systems with measurement output quantization. The reset estimator updates the estimated state using the exact information obtained from the uniform-typed quantizer. In [6], the state estimation problem is investigated for linear discrete-time systems using quantized measurements by a logarithmic quantizer. The problem of state estimation in quantized feedback control systems is studied in [7]. They proposed a new state estimation method for a discrete-time system in which the measurement output is uniformly quantized. In [8], a reliable tracking controller is proposed for discrete time-varying systems subject to time-delays, quantization effects and parameter uncertainties. Set-membership filtering method is used to estimate the state of the system in which the measurement output is quantized based on the logarithmic quantizer. However, up to now, the attack detection problem in NCSs with limited communication bandwidth has not yet received adequate attention. The impact of the communication capacity constraint on the attack detection performance should be further discussed.

Motivated by the discussion above, it is the objective of this paper to design an algorithm by means of set-membership filtering techniques to detect cyber-physical

[1]E. Mousavinejad, [2]F. Yang, and [4]L. Vlacic are with the Griffith School of Engineering, Griffith University, Gold Coast, QLD 4222, Australia eman.mousavinejad@griffithuni.edu.au, fuwen.yang@griffith.edu.au, l.vlacic@griffith.edu.au

[3]Q.-L. Han is with the School of Software and Electrical Engineering, Swinburne University of Technology, John Street, Hawthorn, Melbourne, VIC 3122, Australia qhan@swin.edu.au

attacks into NCSs subject to limited communication bandwidth. First, a recursive convex optimization algorithm that provides a region of state estimate where the true state resides despite the presence of limited communication bandwidth is proposed. This algorithm estimates the state in two steps: a prediction step and a measurement update step. Then, two sub-algorithms are provided to detect attacks in NCSs subject to quantized measurement output caused by the limited communication bandwidth. The idea to detect attacks introduced in the sub-algorithms depends on the existence of intersection between the prediction ellipsoid set and the estimation ellipsoid set updated with the quantized measurement output. If there exists no intersection between the prediction ellipsoid set and the estimation ellipsoid set updated with the previous quantized measurement output, control signal is compromised by attacks; the sensor (measurement output) signal is violated by attacks if there is no intersection between the prediction ellipsoid set and the estimation ellipsoid set updated at the current time instant. The so-called "zoom-in/zoom-out" uniform quantized strategy is employed to model the quantizer.

II. Problem Statement and Formulation

let us consider the framework of an NCS as shown in Fig.1. The concept depicted on Fig.1 is equally applicable to both open and closed control system configurations. Consequently, Fig.1 is only focused onto the link between a sensor, the controllers output and the attack detection system. In this framework, the attack detection system is located at the remote site. Therefore, the measurement output and the control signal are both sent to the attack detection system via a communication network. The propagation of the output of the attack detection system throughout the control system is not considered in this paper. Instead, it has just been considered as an alarm signal. It is assumed that the communication network transmitting data to the attack detection system has a limited communication bandwidth. However, design of the controller is not discussed in this paper and the control signal is considered as a constant-value signal, therefore, there is no need for the control signal to be quantized.

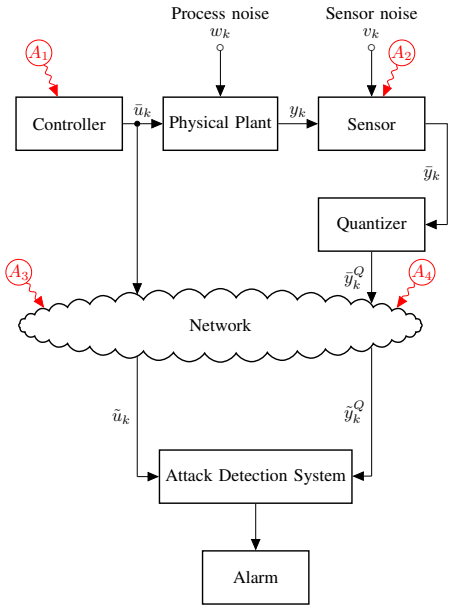


Fig. 1. NCS framework with quantized measurement output showing some possible attack points into NCS.

Attacks on an NCS can be practically considered either attacks directly on physical components (physical attack), i.e. sensors and controllers, or attacks on the communication network that transmits data to receivers located at the remote site (cyber attack). If a physical attack is applied into controllers and/or sensors (attack points A_1 and/or A_2), the control signal, u_k, and/or quantized measurement output, y_k^Q are changed into $\bar{u}_k$ and/or $\bar{y}_k^Q$, respectively. If the communication network is under cyber attacks (attacks points A_3 and/or A_4), the control signal and/or quantized measurement output are changed into $\tilde{u}_k$ and/or $\tilde{y}_k^Q$, respectively. If there is no attack on the NCS, then $\tilde{u}_k = \bar{u}_k = u_k$ and $\tilde{y}_k^Q = \bar{y}_k^Q = y_k^Q$.

A physical plant considered in the proposed framework, Fig.1, can be developed in a mathematical way as a discrete time-varying system in the form of

$$x_{k+1} = A_k x_k + F_k u_k + B_k w_k \tag{1}$$

$$y_k = C_k x_k + D_k v_k \tag{2}$$

where $x_k \in \mathbb{R}^n$ is the system state; $u_k \in \mathbb{R}^l$ is the known deterministic input; $y_k \in \mathbb{R}^m$ is the measurement output; A_k, B_k, C_k, D_k, and F_k are known time-varying matrices with appropriate dimensions; $w_k \in \mathbb{R}^r$ and $v_k \in \mathbb{R}^p$ are the known-but-bounded process and measurement noises, respectively (totally known as system noises), which are assumed to belong to the following specified ellipsoid sets

$$\mathcal{W}_k := \{w_k : w_k^T Q_k^{-1} w_k \leq 1\} \tag{3}$$

$$\mathcal{V}_k := \{v_k : v_k^T R_k^{-1} v_k \leq 1\} \tag{4}$$

where $Q_k = Q_k^T \succ 0$ and $R_k = R_k^T \succ 0$ are known matrices with compatible dimensions.

The initial state x_0 is assumed to belong to a given ellipsoid set $\mathcal{X}_{0|0}(\hat{x}_{0|0}, E_{0|0})$

$$\mathcal{X}_{0|0} := \{x_0 : (x_0 - \hat{x}_{0|0})^T P_{0|0}^{-1} (x_0 - \hat{x}_{0|0}) \leq 1\} \tag{5}$$

where $\hat{x}_{0|0}$ is an estimate of x_0, which is assumed to be given, $P_{0|0} = P_{0|0}^T \succ 0$ is a known matrix and $P_{0|0} = E_{0|0} E_{0|0}^T$.

Due to the limitation of the communication bandwidth, a quantizer is considered at the plant site to transmit a finite number of codewords. In this paper, the "zoom-in/zoom-out" uniform quantized strategy is adopted in the coding/decoding communication [9]. The dynamics of the coder/decoder is considered as

$$\xi_{k+1} = \xi_k + f_k \alpha_k \tag{6}$$

where the quantizer function is

$$\alpha_k = Q\left(\frac{y_k - \xi_k}{f_k}\right) \tag{7}$$

The coder/decoder state ξ_{k+1} contains the quantized measurement output y_k^Q. So,

$$y_k^Q = \xi_k + f_k \alpha_k \tag{8}$$

Therefore, the quantized measurement output satisfies the recursive relation

$$y_{k+1}^Q = \xi_{k+1} + f_{k+1}\alpha_{k+1} = y_k^Q + f_{k+1} Q(\frac{y_{k+1} - y_k^Q}{f_{k+1}}) \tag{9}$$

The function $Q(\cdot)$ is defined by

$$Q(\varepsilon) = \begin{cases} \ell, & \text{if } (2\ell-1)/2 < \varepsilon < (2\ell+1)/2 \\ L, & \text{if } \varepsilon \geq (2L-1)/2 \\ -Q(-\varepsilon), & \text{if } \varepsilon \leq -1/2 \end{cases} \tag{10}$$

where $\{\ell \in \mathbb{N} : 0 \leq \ell \leq L-1\}$, and

$$f_{k+1} = \begin{cases} k_{in} f_k, & \text{if } |\alpha_k| < L \\ k_{out} f_k, & \text{if } |\alpha_k| = L \end{cases} \tag{11}$$

where $k_{in} \in (0,1)$ is the zoom-in parameter and $k_{out} \in (1, +\infty)$ is the zoom-out parameter, L is the number of finite levels.

Remark 1: Since, the error between the current measurement output and the previous quantized measurement output, i.e. $y_{k+1} - y_k^Q$, is much less than the value of the current measurement output y_{k+1}, a quantization of the error can reduce the transmitted rate. Therefore, the quantization strategy in (9) is a quantization of the error, rather than the current measurement output. The zoom in-zoom out quantizer can adjust the transmitted rate at a certain range by changing the scaling parameters k_{in} and k_{out}.

Remark 2: Assuming

$$\Delta_k = \frac{(y_{k+1} - y_k^Q)}{f_{k+1}} - Q(\frac{y_{k+1} - y_k^Q}{f_{k+1}}) \tag{12}$$

Then, if k_{in} and k_{out} in (11) are appropriately scaled, then the quantizer in (10) will never be saturated. For the unsaturated quantizer,

$$\|\Delta_k\| \leq \frac{1}{2} \tag{13}$$

In this paper, a strategy presented to detect cyber-physical attacks is based on an ellipsoidal set-membership filtering approach. This approach involves two ellipsoid sets: a prediction ellipsoid set and an estimation ellipsoid set updated with the quantized measurement output. An attack can be detected if there exists no intersection between the two sets.

This strategy will be refereed as a detection of cyber-physical attacks that compromise sensor measurement and control signal data in an NCS with limited communication bandwidth using a quantized set-membership filtering method.

III. Attack Detection Using Quantized Set-Membership Filtering

In this section, a quantized set-membership filtering method is designed for the proposed cyber-physical attack detection problem.

A. Prediction Ellipsoid Set

The prediction filter is considered in the form of

$$\hat{x}_{k+1|k} = G_k \hat{x}_{k|k} + F_k u_k \tag{14}$$

where $\hat{x}_{k|k}$ is an estimation of the state x_k; G_k is the filter parameter to be determined.

In the first step, a previous state estimation ellipsoid at instant k is assumed to be known and belong to a set $\mathcal{X}_{k|k}(\hat{x}_{k|k}, E_{k|k})$ with the centre $\hat{x}_{k|k}$ and the shape matrix $E_{k|k}$, i.e.

$$x_k = \hat{x}_{k|k} + E_{k|k} z \tag{15}$$

where $E_{k|k}E_{k|k}^T = P_{k|k}$, and for some z such that $\|z\| \leq 1$.

In this step, our goal is to obtain the prediction ellipsoid set $\mathcal{X}_{k+1|k}$:

$$\begin{aligned} \mathcal{X}_{k+1|k} := \{x_{k+1} : \\ (x_{k+1} - \hat{x}_{k+1|k})^T P_{k+1|k}^{-1} (x_{k+1} - \hat{x}_{k+1|k}) \leq 1\} \end{aligned} \tag{16}$$

Theorem 1: For the system (1)-(2), suppose that the state x_k belongs to its state estimation ellipsoid (15). Then the one-step ahead state x_{k+1} resides in its state prediction ellipsoid (16) for any value of the system noises belong to their specified sets, if there exist $P_{k+1|k} > 0$, G_k, $\tau_1 \geq 0$, and $\tau_2 \geq 0$ such that the following recursive matrix inequality holds.

$$\begin{bmatrix} -P_{k+1|k} & \Pi_1(\hat{x}_{k|k}) \\ \Pi_1^T(\hat{x}_{k|k}) & -\Theta_1(\tau_1, \tau_2) \end{bmatrix} \leq 0 \tag{17}$$

where

$$\begin{aligned} \Pi_1(\hat{x}_{k|k}) &= \begin{bmatrix} (A_k - G_k)\hat{x}_{k|k} & A_k E_{k|k} & B_k \end{bmatrix} \\ \Theta_1(\tau_1, \tau_2) &= \text{diag}(1 - \tau_1 - \tau_2, \tau_1 I, \tau_2 Q_k^{-1}) \end{aligned}$$

Corollary 1: Consider the system (1)-(2) and the prediction filter (14). As the size of the ellipsoid can be measured by means of the sum of squared semiaxes lengths, which is given by Tr(P)[1], the convex optimization approach is applied to determine an optimal ellipsoid. $P_{k+1|k}$ is obtained by solving the following optimization problem:

$$\begin{aligned} &\underset{P_{k+1|k}>0, G_k, \tau_1 \geq 0, \tau_2 \geq 0}{\text{minimize}} \quad \text{Tr}(P_{k+1|k}) \\ &\text{subject to (17)} \end{aligned} \tag{18}$$

B. Estimation Ellipsoid Set Updated with Current Quantized Measurement

A filter based on the current quantized measurement is considered for the system (1)-(2), which is in the form of

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + L_{k+1}(y_{k+1}^Q - \hat{y}_{k+1|k}) \tag{19}$$

where L_{k+1} is the filter parameter to be determined.

[1]Tr(P) denotes trace of P.

At the measurement update step, the prediction ellipsoid set $\mathcal{X}_{k+1|k}$ is given by (16) for the state x_{k+1}, which can be written as

$$x_{k+1} = \hat{x}_{k+1|k} + E_{k+1|k} z \tag{20}$$

where $E_{k+1|k} E_{k+1|k}^T = P_{k+1|k}$, and for some z such that $\|z\| \leq 1$.

Our objective is to update this prediction set with the one yielding from the current quantized measurement y_{k+1}^Q. Thus, the updated ellipsoid set should satisfy the condition

$$(x_{k+1} - \hat{x}_{k+1|k+1})^T P_{k+1|k+1}^{-1} (x_{k+1} - \hat{x}_{k+1|k+1}) \leq 1 \tag{21}$$

whenever the equality (output constraint)

$$\begin{aligned} y_{k+1}^Q = C_{k+1}\hat{x}_{k+1|k} + C_{k+1}E_{k+1|k}z \\ + D_{k+1}v_{k+1} - f_{k+1}\Delta_k \end{aligned} \tag{22}$$

holds for some $\|z\| \leq 1$, $v_{k+1} \in \mathcal{V}_{k+1}$ and $\|\Delta_k\| \leq \frac{1}{2}$.

Theorem 2: For the system (1)-(2), if the state x_{k+1} belongs to its state prediction ellipsoid (20), then the updated state x_{k+1} resides in its state estimation ellipsoid (21), if there exist $P_{k+1|k+1} > 0$, L_{k+1}, $\tau_3 \geq 0$, $\tau_4 \geq 0$, $\tau_5 \geq 0$, and N_{k+1} such that the following recursive matrix inequality holds.

$$\begin{bmatrix} -P_{k+1|k+1} & \Pi_2 \\ \Pi_2^T & \Theta_3(\tau_3, \tau_4, \tau_5, \hat{x}_{k+1|k}, N_{k+1}) \end{bmatrix} \leq 0 \tag{23}$$

where

$$\Pi_2 = \begin{bmatrix} 0 & (I - L_{k+1}C_{k+1})E_{k+1|k} & -L_{k+1}D_{k+1} \\ & L_{k+1}f_{k+1} \end{bmatrix}$$

$$\begin{aligned} \Theta_3(\tau_3, \tau_4, \tau_5, \hat{x}_{k+1|k}, N_{k+1}) = -\Theta_2(\tau_3, \tau_4, \tau_5) \\ + N_{k+1}^T \Pi_y(\hat{x}_{k+1|k}) \\ + \Pi_y^T(\hat{x}_{k+1|k}) N_{k+1} \end{aligned}$$

$$\Pi_y(\hat{x}_{k+1|k}) = \begin{bmatrix} C_{k+1}\hat{x}_{k+1|k} - y_{k+1}^Q & C_{k+1}E_{k+1|k} \\ D_{k+1} & -f_{k+1} \end{bmatrix}$$

$$\begin{aligned} \Theta_2(\tau_3, \tau_4, \tau_5) = \ \text{diag}(1 - \tau_3 - \tau_4 - \frac{1}{2}\tau_5, \tau_3 I, \\ \tau_4 R_{k+1}^{-1}, \tau_5 I) \end{aligned}$$

Corollary 2: Consider the system (1)-(2) and filter (19). Applying the convex optimization approach, an optimal ellipsoid can be determined. $P_{k+1|k+1}$ is obtained by solving the following optimization problem:

$$\begin{aligned} \underset{P_{k+1|k+1}>0, L_{k+1}, N_{k+1}, \tau_3 \geq 0, \tau_4 \geq 0, \tau_5 \geq 0}{\text{minimize}} \quad & \text{Tr}(P_{k+1|k+1}) \\ \text{subject to (23)} \end{aligned} \tag{24}$$

C. Attach Detection Algorithm

The algorithm that recursively computes the prediction and estimation ellipsoid sets is summarized as Algorithm 1. Based on the intersection between these two ellipsoid sets, two sub-algorithms are defined to detect attacks that violate the control signal and the sensor measurement.

Algorithm 1 :Recursive State Estimation

1. Initialization:
Given an initial ellipsoid $\mathcal{X}_{0|0}(\hat{x}_{0|0}, E_{0|0})$, the current value of input u_0, recursive times N, and set $k = 0$. Let $\hat{x} \leftarrow x_{0|0}$, $E \leftarrow E_{0|0}$, $u \leftarrow u_0$.

2. Prediction:
1) Calculate $P_{k+1|k}$ and G_k by solving the optimization problem (18). Obtain the matrix $E_{k+1|k}$ by means of the Cholesky factorization of $P_{k+1|k}$: $P_{k+1|k} = E_{k+1|k} E_{k+1|k}^T$.
2) Calculate the centre of the prediction ellipsoid $\hat{x}_{k+1|k}$ by (14).

Sub-Algorithm 1a: Diagnosis Attack on Controller Data

1) If $\mathcal{X}_{k|k} \bigcap \mathcal{X}_{k+1|k} \neq \varnothing$.
There is no attack
2) If $\mathcal{X}_{k|k} \bigcap \mathcal{X}_{k+1|k} = \varnothing$.
Indicate attack, and then
$\mathcal{X}_{k+1|k} \leftarrow \mathcal{X}_{k|k}$

3. Measurement Update:
1) Calculate the quantized measurement output, y_{k+1}^Q, from (9).
2) Calculate $P_{k+1|k+1}$ and L_{k+1} by solving the optimization problem (24). Obtain the new $E_{k+1|k+1}$ by the Cholesky factorization of $P_{k+1|k+1}$.
3) Calculate the centre of the updated estimation ellipsoid $\hat{x}_{k+1|k+1}$ by (19).

Sub-Algorithm 1b: Diagnosis Attack on Sensor Data

1) If $\mathcal{X}_{k+1|k+1} \bigcap \mathcal{X}_{k+1|k} \neq \varnothing$.
There is no attack
2) If $\mathcal{X}_{k+1|k+1} \bigcap \mathcal{X}_{k+1|k} = \varnothing$.
Indicate attack, and then
$\mathcal{X}_{k+1|k+1} \leftarrow \mathcal{X}_{k+1|k}$

4. Loop
If $k == N$ then Exit, Else $k \leftarrow k + 1$ and Goto *Prediction* step

IV. An Illustrative Example

In this section, a maneuvering target tracking system with a constant-acceleration input model [10] is considered as

$$x_{k+1} = \begin{bmatrix} 0.8 & 0.6 \\ 0 & 0.8 \end{bmatrix} x_k + \begin{bmatrix} 0.18 \\ 0.6 \end{bmatrix} u_k + \begin{bmatrix} 2 \\ 1 \end{bmatrix} w_k$$

where the components of the state $x_k = \begin{bmatrix} x_1(k) & x_2(k) \end{bmatrix}^T$ are the position and the velocity; u_k is the commanded acceleration; w_k represents a process noise which belongs to a specified ellipsoid set.

Since the position is measured by radar, the measurement is described as

$$y_k = \begin{bmatrix} 1 & 0 \end{bmatrix} x_k + v_k$$

and v_k is the radar measurement noise which is confined to a specified ellipsoid set.

In the simulation, the control signal is considered as the constant commanded acceleration u_k which is set to 2 m/s^2. w_k and v_k are chosen as $0.5\sin(2k)$ and $0.2\sin(20k)$, respectively. $Q_k = 0.3$ and $R_k = 0.1$. The initial state is set as $x_0 = \begin{bmatrix} 3 & 2 \end{bmatrix}^T$, which belongs to the ellipsoid $\mathcal{X}_{0|0} := \{x_0 : (x_0 - \hat{x}_{0|0})^T P_{0|0}^{-1} (x_0 - \hat{x}_{0|0}) \leq 1\}$, where $\hat{x}_{0|0} = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$, and $P_{0|0} = \left[\begin{smallmatrix} 50 & 0 \\ 0 & 50 \end{smallmatrix}\right]$. The quantizer parameters k_{in} and k_{out} are set to 0.7 and 11, respectively; the initial scale factor is considered as $f_0 = 8$, and the quantization level is $L = 16$.

To perform a successful attack by an attacker, it is assumed that an attacker has knowledge of the accurate values of the

control signal u_k and the sensor measurement output y_k in real time, and he is also capable of violating the integrity of the sensor measurement and control signal and modifying the true values of y_k and u_k to arbitrary ones.

A. Physical Attacks on Controller and Sensor

A physical attack can be directly applied to physical components and perturb the system dynamics. An attacker can access to the plant site and compromise the controller and the sensor to introduce fake control and measurement output commands. The physical attack applied on the controller (attack point A_1 shown in Fig.1), which changes F_k in (1), can be modeled as

$$\bar{F}_k = F_k + A_1$$

and the physical attack applied on the sensor (attack point A_2 shown in Fig.1), which changes C_k in (2), can be modeled as

$$\bar{C}_k = C_k + A_2$$

B. Bias Injection Attack on Control Signal and Replay Attack on Sensor Data

The replay attack consists of two phases: (1) the attacker starts to record sequences of data from sensors' communication channels without entering any input to the system. (2) the attacker replays the recorded data to the system [11].

An attacker can record quantized sensor measurement data from k_i till k_r with the window size of $\tau = k_r - k_i$ in the first phase. Then, in the second phase, the attacker replays the recorded data to the system from $k = k_r + 1$ till the end on the attack at $k = k_f$. This attack (attack point A_4 shown in Fig.1) can be modeled as

$$A_4 = (y^Q_{k-\tau} - y^Q_k)$$

Thus, the quantized sensor measurement output affected by the attack is

$$\tilde{y}^Q_k = y^Q_k + A_4$$

In the bias injection attack, the attacker injects a constant bias in the system [11]. The bias injection attack on the control signal (attack point A_3 shown in Fig.1) can be modeled as

$$A_3 = \delta_u$$

where δ_u is a constant value injected by the attacker. Therefore, the control signal's data affected by the attack is

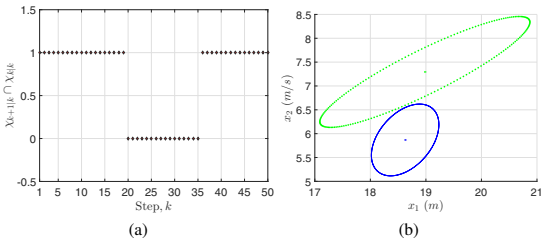
$$\tilde{u}_k = u_k + A_3$$

V. Results and Discussion

The simulation results are obtained under Matlab 8.6 with YALMIP and the solver Sdpt 3 during 50 sampling steps. The simulated attacks are applied from step $k = 20$ to step $k = 35$.

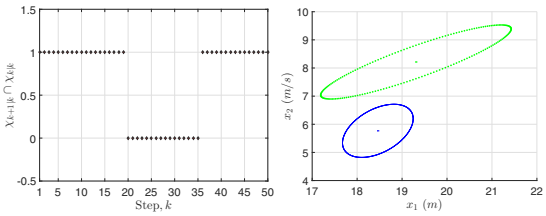
A. Physical and Bias Injection Attacks on Controller

In this section, first, the attacker applies a physical attack on the controller to provide the plant with a fake control command, i.e. the attack point A_1. Second, the control signal is violated via applying a bias injection attack into the communication network between the controller and the attack detection, i.e. the attack point A_3. In the simulation, the attack vectors are modeled as $A_1 = \begin{bmatrix}0.1 & 0.7\end{bmatrix}^T$ and $A_3 = 4$.



(a) (b)

Fig. 2. Controller physical attack. (a) intersection between $\chi_{k+1|k}$ and $\chi_{k|k}$ (b) $x_1 - x_2$ phase-plane at step 20 ($\chi_{k+1|k}$ green, dotted line, $\chi_{k|k}$ blue, solid line).



(a) (b)

Fig. 3. Controller bias injection attack. (a) intersection between $\chi_{k+1|k}$ and $\chi_{k|k}$ (b) $x_1 - x_2$ phase-plane at step 20.

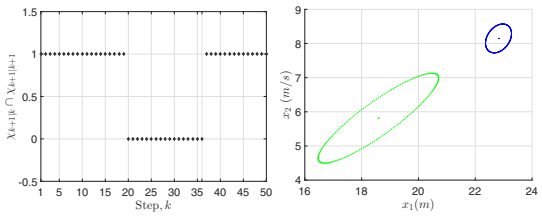
Figs. 2a and 3a show the sequences of the intersection between the prediction ellipsoid set and the estimation ellipsoid set updated with the previous quantized measurement output[2]. As the physical and bias injection attacks start at $k = 20$, the centre of the prediction ellipsoid set (14) is affected by the attacks; however the previous updated estimation ellipsoid set is not affected by the attacks as it is calculated with the prediction ellipsoid set at $k = 19$. Therefore, it is expected from the *sub-algorithm 1a* that there must exist no intersection between these two sets, i.e. $\chi_{21|20} \cap \chi_{20|20} = \varnothing$, as illustrated in Figs. 2b and 3b.

B. Physical and Replay Attacks on Sensor

In this case, first, the attacker applies a physical attack on the sensor to change the quantized measurement output received by the attack detection, i.e. the attack point A_2. Also, the measurement output is compromised via injecting a replay attack into the communication network between the sensor and the attack detection, i.e. the attack point A_4. In the simulation, the attack vector that represents A_2 is modeled as $A_2 = \begin{bmatrix}2 & 0\end{bmatrix}$.

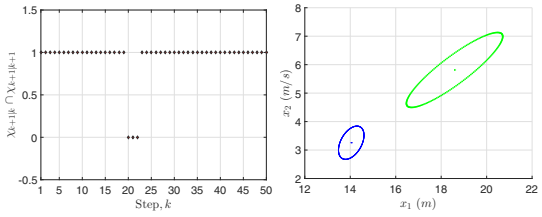
Figs. 4a and 5a demonstrate the sequences of the intersection between the prediction ellipsoid set and the estimation ellipsoid set updated with the current quantized measurement output. As the physical and replay attacks start at $k = 20$,

[2]"1" indicates there exists intersection, "0" indicates there is no intersection.



(a) (b)

Fig. 4. Sensor physical attack. (a) intersection $\chi_{k+1|k}$ and $\chi_{k+1|k+1}$ (b) $x_1 - x_2$ phase-plane at step 20 ($\chi_{k+1|k}$ green, dotted line, $\chi_{k+1|k+1}$ blue, solid line).



(a) (b)

Fig. 5. Sensor replay attack. (a) intersection $\chi_{k+1|k}$ and $\chi_{k+1|k+1}$ (b) $x_1 - x_2$ phase-plane at step 20.

the prediction ellipsoid set is calculated from the quantized measurement output obtained at $k = 19$ when there is no attack. However, the estimation ellipsoid set is updated with the current quantized measurement output at $k = 20$. Therefore, it is expected from the *sub-algorithm 1b* that there must be no intersection between these two sets, i.e. $\chi_{20|19} \cap \chi_{20|20} = \varnothing$, as illustrated in Figs. 4b and 5b[3].

C. Quantization Effects

Since the proposed attack detection algorithm relies on the existence of intersection between the prediction ellipsoid set and the estimation ellipsoid set updated with the quantized measurement output, quantizing the measurement output affects the performance of the attack detection system. The proposed attack detection system can be sensitive to the quantization parameters, i.e. L, k_{in}, and k_{out}. Consider the physical attack on the sensor discussed in Section V-B. If the quantization error increases through decreasing the number of quantization levels from $L = 16$ to $L = 3$, the attack detection system is no longer able to detect the attack at its time of occurrence, $k = 20$, as shown in Fig. 6 and therefore, there is some delays to identify the attack compared with the results shown in Fig. 4a. Moreover, the proposed set-membership filter cannot recover the state of the system after the attack finishes. Therefore, there are some false alarms after step $k = 36$ produced by the attack detection system.

VI. Conclusion

This paper has presented a cyber attack detection algorithm based on the ellipsoidal set-membership filtering approach in NCSs subject to the limited communication bandwidth. The algorithm consists of a prediction ellipsoid set and an estimation ellipsoid set updated with the quantized measurement output. The attack detection method relies on the intersection of the two sets. In addition, it is investigated that the uniform quantizer in the system makes the attack detection system sensitive to the quantization levels and zoom-in/zoom-out scaling parameters.

Future work would focus on considering logarithmic quantization of the measurement output. Moreover, this study can be further improved through considering a robust and resilient controller in the proposed NCS framework that uses information from the attack detection system output to mitigate effects of cyber-physical attacks on the state of the system.

[3]the updated estimation ellipsoid at step 20 is magnified with the ratio of 10^7.

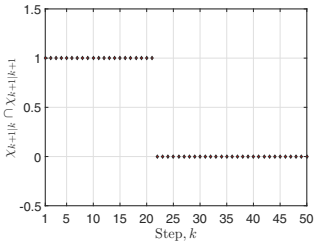


Fig. 6. Intersection $\chi_{k+1|k}$ and $\chi_{k+1|k+1}$, sensor physical attack, $L = 3$

References

[1] X. M. Zhang, Q. L. Han, and X. Yu, "Survey on recent advances in networked control systems," *IEEE Trans. Ind. Informat.*, vol. 12, no. 5, pp. 1740–1752, Oct 2016.

[2] A. A. Cárdenas, S. Amin, and S. Sastry, "Research challenges for the security of control systems," in *Proc. The 3rd Conference on Hot Topics in Security*, San José, CA, Jul. 2008, pp. 1–6.

[3] L. Zhang, H. Gao, and O. Kaynak, "Network-induced constraints in networked control systems - a survey," *IEEE Trans. Ind. Informat.*, vol. 9, no. 1, pp. 403–416, Feb 2013.

[4] A. V. Savkin and I. R. Petersen, "Set-valued state estimation via a limited capacity communication channel," *IEEE Trans. Autom. Control*, vol. 48, no. 4, pp. 676–680, April 2003.

[5] J. Zheng and M. Fu, "A reset state estimator for linear systems to suppress sensor quantization effects," *IFAC Proceedings Volumes*, vol. 41, no. 2, pp. 9254 – 9259, 2008, 17th IFAC World Congress.

[6] Z. Duan, V. P. Jilkov, and X. R. Li, "State estimation with quantized measurements: Approximate mmse approach," in *2008 11th International Conference on Information Fusion*, June 2008, pp. 1–6.

[7] T. Ohtsuka, T. Zanma, and K. Liu, "State estimation in quantized feedback control system," in *2014 IEEE 13th International Workshop on Advanced Motion Control (AMC)*, March 2014, pp. 746–751.

[8] S. Liu, G. Wei, Y. Song, and Y. Liu, "Error-constrained reliable tracking control for discrete time-varying systems subject to quantization effects," *Neurocomputing*, vol. 174, pp. 897 – 905, 2016.

[9] R. Carli, F. Bullo, and S. Zampieri, "Quantized average consensus via dynamic coding/decoding schemes," in *2008 47th IEEE Conference on Decision and Control*, Dec 2008, pp. 4916–4921.

[10] X. R. Li and V. P. Jilkov, "Survey of maneuvering target tracking. part i. dynamic models," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1333–1364, Oct 2003

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

[12] A. A. Kurzhanskiy and P. Varaiya, "Ellipsoidal toolbox," EECS Department, University of California, Berkeley, Tech. Rep., May 2006.