Identification and Analysis of Stochastic Deception Attacks on Cyber Physical Systems

Soheila Barchinezhad[a], Mohammad Sayad Haghighi[a,*], Vicenc Puig[b]

[a]*School of Electrical and Computer Engineering, College of Engineering, University of Tehran, 1439957131, Iran.*
[b]*Institut de Robotica i Informatica Industrial (IRI), Universitst Politechnica de Catalunya (UPC), Barcelona, 08028, Spain.*

Abstract

Cyber Physical Systems (CPSs) refer to control systems which are composed of sensors, actuators, computers and network components. These systems are vulnerable to unforeseen failures and external malicious attacks. In this paper, we analyze the stability of CPSs under stochastic deception attacks. To this end: i) we propose a statistical framework for detection of deception attacks in CPSs; ii) identify the place of such attacks by taking advantage of a novel cryptographic adversarial model; iii) as an extra effort, put together an intelligent deception attack; and finally, iv) based on the real-time data and characterizing the imposed deception attacks by an Intrusion Detection System (IDS), we analyze the effect of both intelligent and blind/random deception attacks on the stability of CPSs. We do this through a Markov chain modeling and subsequently extract the sufficient stability conditions. We validate our findings by illustrative examples at the end of the paper. Our results show that proposed IDS can detect deception attacks with low false positive and negative rates in real time. The results also confirm the validity of the theoretically-predicted stability conditions.

Keywords: Cyber-Physical Systems, Deception Attack, Markov Jump Model, Stability Analysis, Intrusion Detection System, Switching System.

1. Introduction

Cyber-physical systems (CPS) consist of both computational/control centers and physical processes, connected by communication networks such as Internet. Using the communication networks, make these systems suffer from specific vulnerabilities which do not affect classical control systems. Therefore, the security of such real-time systems requires a comprehensive perception of network security [1], control theory, and system dynamics [2] which in turn necessitates the development of appropriate information security and control system analysis techniques. Recently, the cyber attacks on CPSs, especially those employed in critical infrastructures, have drawn the attention of researchers to the risks in

*Corresponding author
Email addresses: s.barchinezhad@ut.ac.ir (Soheila Barchinezhad), sayad@ut.ac.ir (Mohammad Sayad Haghighi), Vicenc.puig@upc.edu (Vicenc Puig)

this area [3–6]. In many cases, classical information security services such as confidentiality and data integrity are used for the protection of CPSs in critical infrastructure [7, 8]. Data confidentiality can be achieved by encryption algorithms and a key shared between sender and receiver. Data integrity service can also be provided by using keys and one-way hash functions which produce fixed-size hash codes. Although these security services boost the security of CPSs, they are ineffective against some types of cyber attacks.

Control data and sensory data are sometimes transmitted over networks. By finding the encryption key, an attacker can compromise the forward or backward link [9] (the data stream from controller to physical system or vice versa) and change the content of them through deception attacks [10]. Deception attacks consist of malicious interventions that are done in control loop to reduce the efficiency of physical process or destabilize the system by violating data integrity. In a more complex attack, that is introduced in [11], as controlled data injection, the attacker drops the original data and sends bogus data to the destination. In order to have more destructive effect, the attacker might inject data stochastically. This randomness causes many switchings between attack and non-attack modes which deprives the chance of compensation from the controller and potentially makes the system unstable. In this taxonomy, sometimes the bogus data injection in the closed-loop is done without having prior knowledge about the CPS, and sometimes, in order to achieve a desired behavior on the physical system, they are generated intelligently by gathering information about the system behavior and design. Study of deception attacks in CPSs has received increasing attention in the last few years. For instance, a covert deception attack for service degradation is proposed in [11], which is designed based on the intelligence information gathered by system identification attack. The authors of [12] study three kinds of stealthy deception attacks based on the attackers' ability in compromising the sensors and/or actuators of the system. Furthermore, [13, 14] focus on analyzing system's response during false data injection attacks on CPSs.

In order to promote the security of CPSs, using another safeguarding approach (i.e. intrusion detection system (IDS)) is essential. IDSs are a complement to security services including confidentiality and message integrity. Employment of IDS for CPSs has gained considerable attention since attacks can cause CPS failures [15–17]. The authors of [18] propose a specification-based anomaly detection framework using the information provided by some components in power systems. In [19] and [20], rule-based intrusion detection algorithms are developed, and [21] proposes an artificial intelligence approach for detection, estimation and compensation of attacks in nonlinear CPSs. In contrast to these model-based methods, some works introduce real time threshold-based approaches that have less computational complexity. For instance, in [22], a χ^2 failure detector is used for intrusion detection. The detector is instructed to trigger an alarm when a quadratic distance measure exceeds from a pre-defined threshold. It is assumed that this distance measure has the χ^2 distribution. In the same fashion, the authors of [23] use a dynamic procedure for change detection. They introduce the cumulative sum method, which employs the distance measure history. The alarm is triggered for the sample times in which the cumulative sum goes beyond the threshold. Another threshold-based detection framework is proposed in [24]. This is done by monitoring the expected alarm rate associated with consecutive changes to deal with hidden sensor at-

tacks. This method uses a pseudo-window to monitor the estimated alarm rate at run-time. In addition to these, paper [25] and its multi-sensor version [3] utilize a secure data transmission module using the pseudo-random number as a watermarking in the backward link, so that it can help the χ^2 detector to recognize undetected linear deception attacks. The advantage of these two works compared to the above-mentioned threshold-based ones is that they work well in recovering attacked data, but the disadvantage of all the mentioned methods is that they can not bring us the correct statistics of the stochastic attacks.

Cyber attacks that stochastically exploit the system vulnerability, will turn the closed loop control system into a stochastically switched one that has both normal and attacked modes. Sometimes switching between some modes makes the system unstable. Hence, stability analysis of CPSs under switching attacks is one of the popular subjects in the security analysis of CPSs. The authors of [26] are concerned with the secure and stable control problem for a class of discrete-time stochastic nonlinear systems under deception attacks. Stability of CPSs under DoS attacks is investigated in [27]. In this work, necessary and sufficient conditions for the stability of closedloop systems that are under attack are provided. In [28], by introducing a time-variant Lyapunov-Krasovskii function, a sufficient condition is derived to design controllers so that exponential stability is guaranteed under cyber attacks. Moreover, reference [29] shows that linear stochastic systems are exponentially almost surely stable (EAS-stable) if and only if their transition matrix is averagely contractive over a finite time interval.

To simulate mode mutation in a system with random changing, Markov Jump Systems (MJS) are commonly used for modelling [30–32], since MJSs can represent physical systems with abrupt changes in both structure and parameters [33, 34]. A sliding mode control problem for a class of CPSs is investigated in [35]. This study models system as a Markovian jump CPS. It is assumed that the control input data are transmitted via a communication network in which adversaries may inject false data into the control signals in a stochastic manner. In [36], an event-triggered control problem for the networked Markovian jump systems subject to deception attacks is studied. In this paper, it is believed that deception attacks cause asynchronous switching. Based on the variations between two consecutive states, a switching-like event-triggered scheme is formed and sufficient conditions are derived to ensure the mean-square exponential stability for such systems.

Although there are many studies focusing on attack analysis in CPSs [37–39], there are not many works that identify attacks in real time and predict their effect on system stability. In an intrusion case, only detecting the abnormality is not enough and determining whether the attack needs a reaction or compensation is also important. A prominent security issue that requires response in CPSs is instability. Stochastic deception attacks are complex, and further research in this area is needed. Determining the rate and the place of attack are helpful in analyzing the effect of attack on system stability. Motivated by the above, this study tends to use classical network security and control theories to identify stochastic cyber attacks and analyze their effects on a dynamic real-time CPS. In this paper, we first propose an anomaly-based IDS which tries to model the process or plant normal behavior locally, and compare the model and original plant outputs (by residual error and applying detection thresholds). It also determines the location of attack by using an adversarial model based on

cryptographic tools such as hash functions and message authentication codes. The purpose of this IDS is to detect as many intrusions as possible with minimum number of false alarms and to spot their locations in the shortest possible time. The proposed IDS provides information about the attacks probabilities (or the stationary distributions) and the place of them, that are useful in analyzing attack effects on CPSs. We assume that the attacker, on the forward or backward communication channel, stochastically changes the original data and sends fake data towards the destination instead. In fact, a specific share of received data at the destination is the attacker's data that are either generated by knowing the state of the system or without knowing that. In the former case, the false data injected by attacker are generated without prior knowledge of the CPS model/state and they are practically random. In the latter one, the forged data are generated intelligently in an optimal way by gathering information about the system. In this case, the attacker uses a control law that acts against the CPS controller and tries to maximize a quadratic cost function.

To study the effect of attacks on system stability, we suggest modeling the system under attack as a Markov jump linear system whose stationary distribution is provided by the proposed IDS. Afterwards, we find sufficient stability conditions in order to determine the boundaries for the probabilities of attacks (or the volume of attacks) under which the attacked system remains stable. To this end we consider the evolution of system states over many transitions. To evaluate the proposed scheme, we conduct a set of simulations in Matlab/Simulink for some examples. The simulation results approve the validity of the theoretical analyses. The paper contributions can be summarized as follows:

1) First, we describe and model CPSs under stochastic attacks. In the context of CPSs and attack scenarios, our system is modelled using a Markov jump system representation, and for the prototypical deception attacks as a special case. Additionally, in this section we describe the network transmission mechanism.

2) Then, an intrusion detection algorithm with a method to identify the attack location is proposed. This algorithm is capable of detecting deception attacks in real time with low false positive and negative rates.

3) Next, we find two sufficient stability conditions in the process of analyzing the effect of detected attack on system stability. These conditions depend on the attack information obtained by the intrusion detection algorithm (i.e. the probability and location of attack) and check whether attack will make the system unstable or not. As an additional effort, an intelligent attack is designed whose objective is to find an optimal control law that aims at destabilizing the system in a short period.

The rest of the paper is organized as follows. The model description and problem setup are given in Section 2. In Section 3, the proposed intrusion detection and identification method is discussed. Sufficient conditions for system stability are derived in Section 4. In order to check the validity of the theoretical developments, an application with a DC motor as the plant is studied in Section 5 . Finally, some concluding remarks are made in Section 6.

2. Problem Setup

In this section, we present a mathematical model to describe the dynamics of a CPS subject to cyber attacks and explain network data transmission

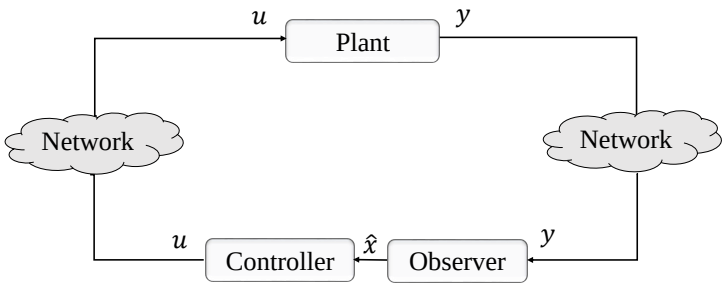

Figure 1: A simple form of closed loop CPS.

mechanism. We consider a discrete-time linear CPS under deception attacks. Deception attacks on CPSs may take place on the controller and/or on its communication channel. Assuming the controller is secure, these attacks affect the forward and/or the backward data channels.

2.1. Normal system dynamics

In this study, we use a linear time-invariant (LTI) discrete control system model. Such a simplified model is useful for studying fault and attack detection as well as stability analysis. Model of a sample linear discrete time system in benign mode with state-feedback controller is presented in Fig. 1, is given as,

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) \\ y(k) = Cx(k) \end{cases} \tag{1}$$

where $x(k)$, $u(k)$ and $y(k)$ are the system state, control signal and measured output for the kth sample, respectively. The matrices A, B and C are the state, input and output matrices. Since in our scenarios only the measured outputs are sent to the controller, an observer is included in the loop together with the controller such that using the measured outputs, it can estimate the system state. This information is required for the state feedback controller to generate the control actions.

$$u(k) = \mu(\hat{x}(k)) \tag{2}$$

In the above, μ is the mapping function of the state feedback controller, and $\hat{x}(k)$ is the state estimated by the observer. In our case, an observer with Luenberger structure is used [40]. It can make the estimated error approach zero asymptotically.

$$e(k) = x(k) - \hat{x}(k) \tag{3}$$

By considering that the error converges to zero, we will have $\hat{x}(k) \approx x(k)$. Hence, equations (1) and (2) can be rewritten as:

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) \\ u(k) = \mu(x(k)) \end{cases} \tag{4}$$

2.2. System dynamics under cyber attack

Stochastic cyber attacks make a CPS behave as a switching system. A switching system is a dynamic system consisting of a number of modes as well as a switching signal $r(k)$ that determines which mode is activated at any time. In our system, one of the modes is related to the normal state, and the other is created by an attack on the communication channel. Even if both modes or subsystems are stable, switching between them could still cause instability. It means that the stability of subsystems does not necessarily reflect the stability of the whole system. Consider the attack occurs along the forward channel. If the adversary discards the original packets and sends new ones, then, the system equation will be:

$$\begin{cases} x(k+1) = Ax(k) + Bu'(k) \\ u'(k) = \begin{cases} u(k) = \mu(x(k)) & if \;\; r(k) = 0 \\ u_a(k) = \delta(k) & if \;\; r(k) = 1 \end{cases} \end{cases} \tag{5}$$

In the above, $u(k)$ is the control signal sent by the controller. In this paper, a state feedback controller is employed (as function $\mu(.)$ to calculate the control signal. $u'(k)$ is the control signal received by the physical system and $\delta(k)$ denotes an alternative data generation function used by the attacker to generate the new control data to be sent to the plant. $r(k) \in \{0, 1\}$ is the switching signal and indicates the status of the packet/signal delivery. In case $r(k) = 0$, the system is normal and the plant receives the data sent by the controller, otherwise the received data are generated by the attacker.

Since Markov jump systems can effectively model physical systems with abrupt mutations, our above switching system can be described as a discrete-time Markov jump linear system (MJLS).

$$x(k+1) = A_{r(k)}x(k) \tag{6}$$

where the process $r(k)$ is a finite-state Markov stochastic jump taking values from a finite set $S = \{0, ..., d\}$. Transition probabilities of that is $Pr\{r(k+1) = j|r(k) = i\} = q_{ij}$, where q_{ij} is the transition rate from mode i at sample time k to mode j at sample time $k+1$. Let $q_{ii} = -\sum_{j=0, j\neq i}^{d} q_{ij}$ and $Q = [q_{ij}]$, by an initial probability distribution $\pi_0 = [\pi_{00}, \pi_{01}, ..., \pi_{0d}]^T$, where $\pi_{0i} = Pr\{r(0) = i\}$, then, the probability distribution $\pi_k = [Pr\{r(k) = 0\}, ..., Pr\{r(k) = d\}]^T$ satisfies the differential equation $\pi_{k+1} = Q^T\pi_k$. Assuming that the Markov chain $r(k)$ is irreducible aperiodic or ergodic, there is a unique invariant distribution $\bar{\pi} = [\bar{\pi}_0 \; \bar{\pi}_1 \; ... \; \bar{\pi}_d]^T$ which π_k converges to for any π_0 and vector $\bar{\pi}$ is called the steady state probability of the ergodic form process $r(k)$. In our case, the value of d is 1, i.e. $r(k)$ switches between two modes 1 (attacked) and 0 (non-attacked).

2.3. Secure transition mechanism

The most important security services in data transition in CPS are data integrity and confidentiality. A secure transmission mechanism, which guarantees message integrity by using hash functions and confidentiality by using encryption algorithms, is shown in Fig. 2. As shown in this figure, a sender provides a plain text message (M_p) as the input to the hash function that generates a unique hash code which is concatenated with M_p. Then, it can be encrypted

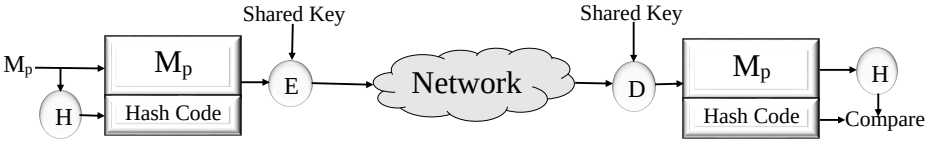

Figure 2: A secure transmission mechanism.

by encryption algorithms and shared keys, and sent to the destination. At the destination, that incoming data packet is first decrypted and the hash function is applied to the message again in order to calculate the digestion. If the result is equal to the appended hash code of incoming data packet, the message is considered authentic. Otherwise, the message is discarded.

3. The Proposed Intrusion Detection and Identification Method

In this section, the proposed methods for detecting cyber attacks and identifying their location are presented.

3.1. Intrusion detection

The solution for intrusion detection that is proposed here is based on placing an identifier, using the mathematical model of plant, on the controller side. We first consider a secure phase for the closed-loop system during which the dynamics of the normal system is identified and modeled. In fact, it is assumed that the system starts from a secure phase, let us say $[0 - T]$, during which no attack happens. By the end of this phase, we have gathered enough information for intrusion detection. After this phase, we monitor the system behavior and, if the residual, i.e., the difference between the model output and the received output by IDS, exceeds a threshold, an anomaly is detected. This threshold can be set by analyzing the collected information during the secure phase. The proposed scheme can be seen in Fig. 3 and Algorithm I. In the algorithm, we compare the statistics of two time-series; the first one is related to received output by IDS (Y) and the other one is related to model output (Y_M). As a result of this comparison, the residual vector can be generated. Assuming that the vector of residuals has a Gaussian distribution, detection thresholds are defined as $\mu \pm J\sigma$ where J and σ are the Gaussian mean and standard deviation of residuals, respectively, and J is a constant which depends on system sensitivity. J should be chosen based on false positive and false negative values.

3.2. Identifying the Location of Intrusion

Cyber attacks may occur in the forward or backward communication channel. Attacks in both channels may have adverse effects on the output of physical system and it is important to determine in which link they have occurred. Identifying the location of attacks is useful for compensation and control processes, as well as analyzing the effects of attacks on physical system. This is achievable by an adversarial model based on cryptography, hash functions, and an additional sensor. We use an additional sensor on the physical system side as shown in Fig. 3. Suppose all data streams (forward and backward channel data,

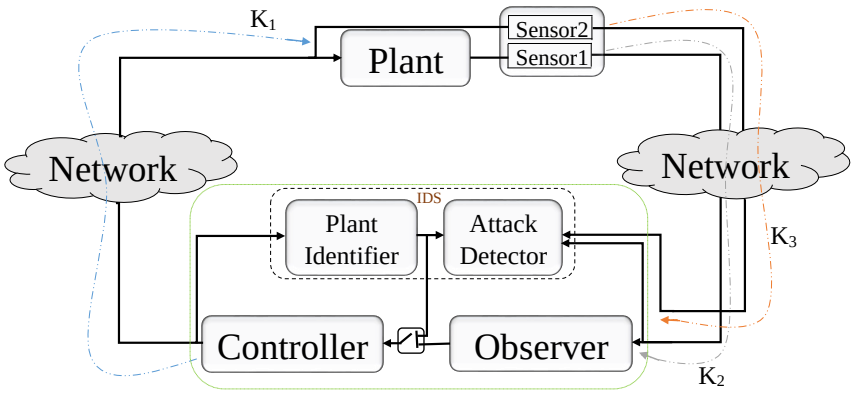

Figure 3: Proposed adversarial model schema to intrusion detection and identification.

as well as the second sensor data) are protected with different encryption keys such that integrity and confidentiality services are provided. The attacker who is going to inject false data in the channels, must compromise the keys then. We assume that the attacker can compromise at most one key in our adversarial model. It can be shown that it is possible to determine the location of the attack precisely if this assumption holds. It is reasonable to assume that an attacker acquires only one key at a time, because if the attacker obtains two different keys simultaneously, he/she can intercept the controller and the physical system and pretend to each one that he/she is the other. The schema of the adversarial model is shown in Fig. 3, and its functionality in Algorithm 1.

Remember the secure transmission mechanism of Fig. 2. The controller and the plant use this mechanism in forward and backward communication links. They use three shared (secret) keys to protect their data streams (i.e. forward and backward streams plus the information submitted by sensor 2). In Fig. 3, the control data stream in the forward channel is protected using the key K_1, and the measured output in the backward channel with the key K_2. In our model, the received data hash code is recorded by sensor 2 at the plant side and is sent to the controller to be used in the intrusion identification process. This data is protected by K_3, as shown in Fig. 3. By considering these, we can now explain the proposed attack place identification process.

Controller and plant send and receive data through a communication network. As we can see in Algorithm 1, they have a data sender and a data receiver unit. The sender unit calculates the hash code of data, concatenates the data with its hash code, encrypts the result with the corresponding encryption key, and sends the data packet to the destination. Then, the receiver unit receives the data packet (which can be the original data packet or one that is generated by the attacker).

Algorithm 1. Intrusion detection and identification.

*****Secure transmission in proposed adversarial model*****

Controller data sender:

1. $hash_1 \leftarrow HashFunction(u(k))$
2. $Buffer(hash_1)$
3. $M_{control} \leftarrow Encrypt(u(k) \| hash_1, K_1)$
4. $Send(M_{control})$

Plant data receiver:

5. $Receive(M'_{control})$
6. $(u'(k) \| hash'_1) \leftarrow Decrypt(M'_{control}, K_1)$

Plant data sender:

7. $hash_2 \leftarrow HashFunction(y(k))$
8. $M_{output} \leftarrow Encrypt(y(k) \| hash_2, K_2)$
9. $M_h \leftarrow Encrypt(hash'_1, K_3)$
10. $Send(M_{output}), \quad Send(M_h)$

Controller data receiver:

11. $Receive(M'_{output}), \quad Receive(M'_h)$
12. $(y'(k) \| hash'_2) \leftarrow Decrypt(M'_{output}, K_2)$
13. $hash''_1 \leftarrow Decrypt(M'_h, K_3)$

*****Intrusion detection and intrusion place identification*****

1. $R \leftarrow [y(1) - y_M(1), y(2) - y_M(2), ..., y(T) - y_M(T)]$
2. $\mu \leftarrow mean(R), \quad \sigma \leftarrow std(R)$
3. $\lambda_{high} \leftarrow \mu + J\sigma \ , \quad \lambda_{low} \leftarrow \mu - J\sigma$
4. $if \ (y'(k) - y_M(k)) > \lambda_{high} \quad OR \quad (y'(k) - y_M(k)) < \lambda_{low}$
5. $\quad if \quad hash_1 \neq hash''_1$
6. $\quad\quad r(k) = 1, \quad AttackPlace \leftarrow ForwardLink$
7. $\quad else$
8. $\quad\quad r(k) = 1, \quad AttackPlace \leftarrow BackwardLink$
9. $\quad end \ if$
10. $else$
11. $r(k) = 0$
12. $end \ if$

Notations:

$\| :$ *Concatenation*, $\quad$ Std : *Standard deviation*

In Algorithm 1, the controller calculates the hash code of the control signal ($hash_1$) and buffers this code to be used in the intrusion detection process. It encrypts ($u(k)||hash_1$) with K_1 and sends it out as a packet (which we refer to as $M_{control}$). The plant receives the control packet $M'_{control}$ (which may be different from $M_{control}$ due to an attack). It decrypts the packet and uses the embedded control signal. In the sender unit of the plant, the measured plant output is hashed, encrypted and sent back to the controller. Moreover, the hash code of the received control packet is recorded by sensor 2, encrypted with K_3 and sent to the controller. In the IDS, if the difference between the received plant output and the model output exceeds a threshold, the presence of intrusion in forward or backward channel is detected. If the hash code received by the controller is different from the buffered one, it can be deduced that the attack has happened in the forward link and accordingly, K_1 has been compromised, otherwise it has happened in the backward channel and K_2 has been compromised. In the case that the residual of the received plant output and the model output is within the thresholds, but the hash code received by the controller is different from the buffered one, K_3 must have been compromised. The other possibility could be that two or three keys have been compromised simultaneously, but since we have assumed only one key can be compromised at a time, this option is ruled out.

4. Stability Analysis

Stochastic cyber attacks make a CPS behave as a switching system, with a switching signal that is continually switched between different modes. Our goal is to analyze system stability under these switches and propose sufficient conditions for stability of a system under stochastic deception attacks. As we mentioned before, in the present taxonomy, sometimes injected false data are generated without a previous knowledge of the CPS model and they are random. On the other hand, sometimes injected packets are generated intelligently after gathering information about the system. We extract sufficient stability conditions in presence of both of these two switching systems.

Assumptions

1) In both attack cases, we assume that attack occurrence (i.e. the replacement of packets) follows a Bernoulli distribution with a probability of ρ. ρ can be estimated by using the proposed intrusion detection and identification schema.

2) Random attack is applied as a uniformly distributed multiplicative coefficient.

3) Intelligent attack employs as an optimal control law based on Linear Quadratic Regulator (LQR) [41] to destabilize the system (presumably in a short period).

Definition 1. According to [30, 31], the system is EAS-stable if and only if $\bar{\lambda} = E(\lambda) < 0$ where, $\lambda = \lim_{k\to\infty} sup\frac{1}{k} ln\|\phi(0,k)\|$ is the top Lyapunov exponent [42], $\phi(t_2, t_1)$ is the state transfer matrix over the interval $[t_1, t_2]$, E and $\|.\|$ denotes expectation and norm respectively.

Proposition 1. System (6) is AS-stable if, for some integer $m > 0$ we have,

$$E_{\bar{\pi}}[ln\|\phi(m,0)\|] < 0 \tag{7}$$

in which $E_{\bar{\pi}}$ is the expectation operator with respect to $\bar{\pi}$, as a steady state probability of $r(k)$.

Proof: See the Appendix.

In this section, using the above definition as well as the proposition, we will introduce two new sufficient conditions for AS-stable CPSs under stochastic deception attacks. Since ergodicity conditions for an attack with Bernoulli distribution is satisfied, it has a steady state probability and we can consider Proposition 1 to find sufficient the stability conditions for CPSs under both random/blind and intelligent attacks. The key idea here is to consider the evolution of the state $x(k)$ over m transitions, and to investigate that the system is converging, averagely over that time interval.

4.1. Stability condition for random deception attack

Proposition 2. The closed loop system of (5), with multiplicative uniform random attack in the interval $[a, b]$ on forward channel is AS-stable if there exists some finite integer value like $m > 0$, such that the following condition holds:

$$(1-\rho)m.ln(\|A_0\|) + \rho m.\frac{1}{b-a}\int_a^b ln(\|A - BK_c i\|)di < 0 \tag{8}$$

where a and b are the minimum and maximum values that can be multiplied by the control signal and results a value in the acceptable rang of plat input.

Proof: The state feedback controller of (5) can be formulated as follows:

$$u'(k) = \begin{cases} u(k) = -K_c x(k) & if \;\; r(k) = 0 \\ u_a(k) = \delta(k) & if \;\; r(k) = 1 \end{cases} \tag{9}$$

We assumed the attack occurs in the forward link. The attacker generates and sends random data as control signals. In this case, the original control data $u(k) = -K_c x(k)$ is the only information that the attacker has about the closed loop system, so he/she uses this information. Therefore, the switching system can be modelled as follows:

$$x(k+1) = \begin{cases} (A - BK_c)x(k) & if \;\; r(k) = 0 \\ (A - BK_c f(k))x(k) & if \;\; r(k) = 1 \end{cases} \tag{10}$$

where $f(.)$ is a random data generator and the switching signal $r(k)$ shows the existence or absence of attack. By considering the probability of attack ρ and the state transfer matrix property (i.e. $\phi(t2, t0) = \phi(t2, t1)\phi(t1, t0)$ [43]), for the system of (10), we have:

$$\|\phi(m, 0)\| = \|\phi(m, m-1)\phi(m-1, m-2)...\phi(1, 0)\| \tag{11}$$

$A_0 = A - BK_c$ is the system state transfer matrix in the benign mode and $A_1 = A - BK_c f(k)$ is the system state transfer matrix in the attack mode. From (11) and since A_0 is a constant matrix, by some mathematical simplifications it can be concluded that:

$$E(ln\|\phi(m, 0)\|) \leq ln\|A_0\|^{(1-\rho)m} + E(ln\|A_1\|^{\rho m}) \tag{12}$$

Since $f(k)$ produces a random value for each k, A_1 will be time-variant. In the case that function $f(.)$ produces uniform random numbers in $[a, b]$, we have,

$$E(ln\|A_1\|^{\rho m}) = E(ln\|A\_BK_c i\|^{\rho m}) = \rho m. \int_a^b ln(\|A\_BK_c i\|)p(i)di$$

$$= \rho m. \frac{1}{b-a} \int_a^b ln(\|A\_BK_c i\|)di \tag{13}$$

Now according to Proposition 1, a sufficient condition for stability of system (10) with uniform random attack is:

$$E[ln\|\phi(m, 0)\|] \leq (1-\rho)m.ln\|A_0\| + \rho m. \frac{1}{b-a} \int_a^b ln(\|A\_BK_c i\|)di < 0 \tag{14}$$

The proof of Proposition 2 is completed.□

4.2. Stability condition for intelligent deception attack

When the attacker is able to identify the controller and know the state of the system, forged packets can be generated intelligently. If the controller's K_c is designed to decrease a Lyapunov function and keep the system stable, the attacker uses an optimal K_a to maximize another Lyapunov function which works in his/her favor.

$$x(k+1) = \begin{cases} (A - BK_c)x(k) & if \;\; r(k) = 0 \\ (A - BK_a(k))x(k) & if \;\; r(k) = 1 \end{cases} \tag{15}$$

Designing intelligent attack: According to the assumptions, the attacker's objective in designing the intelligent attack is to find an optimal control law that maximizes a cost function. This problem can be considered as designing an optimal controller aims at minimizing the cost function named G that can be expressed as a negative of a quadratic cost function ($min_{u(k)}[G(x(0))]$). This kind of controller is known as LQR.

$$\begin{aligned} G(x(0)) = E[-(\sum_{k=1}^{N-1}(x^T(k+1)(Q_{r(k)}(k+1))x(k+1) + u^T(k)(R_{r(k)}(k))u(k)) \\ + x^T(N)(P_{r(N)}(N))x(N))] = E[\sum_{k=1}^{N-1}(x^T(k+1)(-Q_{r(k)}(k+1))x(k+1) \\ + u^T(k)(-R_{r(k)}(k))u(k)) + x^T(N)(-P_{r(N)}(N))x(N)] \end{aligned} \tag{16}$$

where $-Q_{r(k)}(k), -R_{r(k)}(k)$ and $-P_{r(N)}(N)$ are positive semi-definite matrices for each mode $r(k) \in S$ and each sample time $k = 0, 1, ..., N-1$, in addition for each $j \in S$, the following condition is met: $p_{ij} = Pr(r(k+1) = i | r(k) = j)$, $\sum_{i=1}^{M} p_{ij} = 1$,

$$-R_j(k) + B^T \sum_{i=1}^{M} p_{ij}(-Q_i(k+1))B \geq 0 \tag{17}$$

In particular, this condition is satisfied if $-R_j(k)$ and $-Q_j(k)$ are positive definite and positive semi-definite respectively. To minimization process, these matrices are given and have appropriate dimensions.

The solution of this jump linear quadratic control problem, can be calculated using dynamic programming [44]. To this end let $V_{r(k)}(x(k))$ be the expected cost to go from state $x(k)$ with mode $r(k)$ to state $x(k+1)$:

$$\begin{aligned} V_{r(k)}(x(k)) = min_{u(k)} E(x^T(k+1)(-Q_{r(k)}(k+1))x(k+1) \\ + u^T(k)(-R_{r(k)}(k))u(k) + V_{r(k)}(x(k+1))) \end{aligned} \tag{18}$$

and

$$V_{r(N)}(x(N)) = x^T(N)(-P_{r(N)}(N))x(N) \tag{19}$$

The iterative relationship (18) can be recursively solved by going backward from finite time N. In this way the optimal control law for minimizing (16) is given by [44]:

$$u_j(k) = -K_j(k)x(k) \tag{20}$$

where the optimal gain is:

$$K_j(k) = [-R_j(k) + B^T(\sum_{i=1}^{M} p_{ij}(-Q_i(k+1) - P_i(k+1)))B]^{-1}$$

$$.B^T \sum_{i=1}^{M} p_{ij}(-Q_i(k+1) - P_i(k+1))A \tag{21}$$

Which the set of symmetric positive semi-definite matrices $\{P_j(k) : j \in M\}$ satisfies the following set of M coupled differential equations are given by the dynamic programming algorithm.

$$P_j(k) = A^T[\sum_{i=1}^{M} p_{ij}(-Q_i(k+1) - P_i(k+1))][A - BK_j(k)] \tag{22}$$

$K_j(k)$ and $P_j(k)$ can be computed recursively. Then the optimal value of the cost function G is given by $G(x(0)) = x^T(0)(P_{r(0)}(0))x(0)$. For the case intelligent deception attack, in the attack mode, we have $K_a(k) = K_1(0)$.

Stability analysis of intelligent attack: For stability analysis in the intelligent attack case, in (11) the sequence of modes is important. Hence, the probability and the transfer matrix of the sequences are needed in calculation of $E[ln(\|\phi(m,0)\|)]$. For this purpose, we should consider the tree of mode occurrences and compute the probability and transition matrix in each branch. For example, for $M = 3$ (Fig. 4), the probability and its transition matrix is as follows:

Branch 1: $p_1 = (1-\rho)^3, \quad \phi_1(3,0) = (A - BK_c)^3$
Branch 2: $p_2 = \rho(1-\rho)^2, \quad \phi_2(3,0) = (A - BK_c)^2(A - BK_a(3))$
.
.
.
Branch 8: $p_8 = \rho^3, \qquad \phi_8(3,0) = (A - BK_a(1))(A - BK_a(2))(A - BK_a(3))$

Therefore,

$$E[ln\|\phi(3,0)\|] = \rho_1.ln\|\phi_1(3,0)\| + \rho_2.ln\|\phi_2(3,0)\| + ... + \rho_8.ln\|\phi_8(3,0)\| \tag{23}$$

Then, Proposition 1 can be investigated for this case to find the sufficient condition for system stability based on ρ.

Proposition 3. System (5) with intelligent deception attack in forward channel is AS-stable if there exists some finite integer value $m > 0$ such that the following condition holds,

$$E[ln\|\phi(m,0)\|] = \rho_1.ln\|\phi_1(m,0)\| + \rho_2.ln\|\phi_2(m,0)\| + ... + \rho_{2^m}.ln\|\phi_{2^m}(m,0)\| < 0 \tag{24}$$

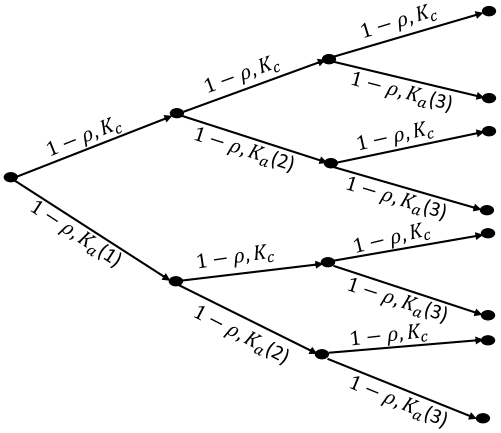

Figure 4: Tree of mode sequences for three time steps.

4.3. Stability Analysis when attack occurs in the backward link

Suppose that the explained deception attacks occur in backward link. In this case, the data receiver is the controller that unlike the physical system, is able to detect the attacked packets using the IDS engine. It discards non-benign packet and uses local model output. The dynamics of the switching system will be based on the following equations:

$$x(k+1) = Ax(k) + Bu(k)$$

$$u(k) = \begin{cases} \mu(x(k)) & if \quad r(k) = 0 \\ \mu(x_l(k)) & if \quad r(k) = 1 \end{cases} \tag{25}$$

If the local model produces outputs with high accuracy such that $x(k)$ and $x_l(k)$ remain close, stability of the switching system will be equivalent to that of the system in normal state.

5. Simulations and Results

In this section, an example is used to illustrate the detection and the effect of deception attacks on system stability. Some intuitions on intrusion detection and system stability are confirmed by numerical tests. The example plant is a second order one (DC motor). Its state space model formulation is

$$\frac{d}{dt}\begin{bmatrix}\dot{\theta}\\ i\end{bmatrix} = \begin{bmatrix}-\frac{b}{Z} & \frac{k}{Z}\\ -\frac{k}{L} & -\frac{R}{L}\end{bmatrix}\begin{bmatrix}\dot{\theta}\\ i\end{bmatrix} + \begin{bmatrix}0\\ \frac{1}{L}\end{bmatrix} V, \quad y = \begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}\dot{\theta}\\ i\end{bmatrix} \tag{26}$$

where $\dot{\theta}$ is the rotational speed of the shaft and i is the armature current. The system parameters are $Z = 0.1\ Kg.m^2$; $b = 0.09\ N.m.s$; $k = 0.01\ N.m.Amp$; $R = 1\ Ohm$; $L = 0.5\ H$. To control the speed of the DC motor under normal conditions, the state feedback controller $K_c = [0.04, -0.55]$ is used such that the

system output tracks a unit step function. We also assume the system is safe in the first 10 seconds of operation ($T = 10s$). The attacker intrudes after that. The sample time in simulations is set to 0.001s.

5.1. Intrusion detection system

The differences in attacks, their places (forward or backward channel) and their probabilities provide evidences to evaluate the reliability of the proposed intrusion detection method. To this end, we use different attack scenarios. We conduct experiments on random and intelligent deception attacks in forward channel, and random attack in backward channel.

The main purpose of an IDS is to detect as many intrusions as possible with minimum number of false alarms in the shortest possible time. Our IDS works in real time and uses statistical measures. The main problem in statistical methods is determining the detection thresholds. In our algorithm, parameter J extends the thresholds and should be tuned for each system individually. As we mentioned in Section 3, J should be chosen in a way that yields low false positive (FP) and false negative (FN) values, and eventuate a balance between them. A series of experiments was conducted to analyze the effect of changing this parameter on the test system. We can find the optimal threshold by some experiments. Because of using an extra sensor in the proposed adversarial model, if the intrusion happens in the forward link, the rate of FP is zero. Thus, the criteria that is important in selecting parameter J for proposed IDS is the rate of FN in the case that attack happens in forward or backward link and only the rate of FP when attacker is on the backward link.

We simulated the intrusion detection algorithm while sweeping the parameter J under both intelligent and random deception attacks with different probabilities. The average values of normalized FN and FP in 11 runs for different attack scenarios with different probabilities of intrusion are shown in Fig. 5. Each simulation was run for 150 seconds and for the intrusion probabilities of $0, 0.0001, 0.0002, ..., 0.001$. In these experiments, FN and FP are normalized by the number of attacked samples and the number of normal samples, respectively. It is clearly seen in the figure that the FN rate goes up with the increase of J, and FP goes down at the same time. Therefore, using larger thresholds might let small intrusions slip away. Let us suppose that the cost of false negative error is equal to that of false positive error. As we can see in Fig. 5, by using $J = 3.2$ to set the thresholds, we have balanced and negligible FN and FP rates under different attack scenarios. For FN, the average of three charts at the bottom is concerned.

To evaluate the intrusion detection system in over/under estimation of stationary distribution of MJLS, we use the metrics precision, recall and F-measure, which are defined as,

$$precision = \frac{TP}{TP + FP},\ recall = \frac{TP}{TP + FN},\ F = 2 \times \frac{precision \times recall}{precision + recall} \quad (27)$$

Since the costs of FP and FN are assumed to be equal, precision, recall and F-measure are all important. Over/under estimation of stationary distribution of MJLS affects the accuracy of the stability thresholds found by the proposed propositions.

Now, we examine the proposed IDS with the chosen value of J under different attack scenarios and attack probabilities.

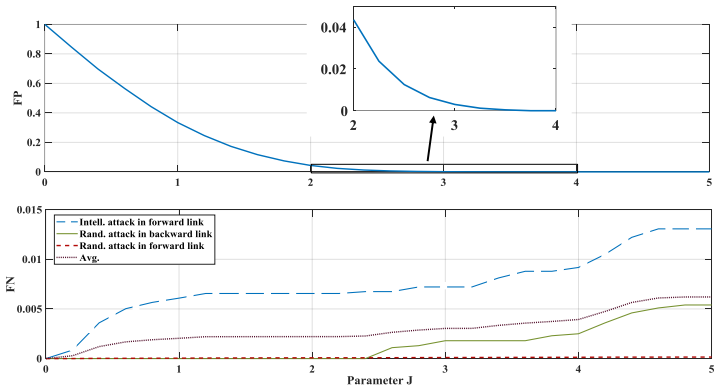

Figure 5: (Top) Mean rate of FP for random deception attack in the backward link, (Bottom) Mean rate of FN for random deception attack in the forward and backward links, and intelligent deception attack in the forward link with different values of J.

1) Scenario I: Random attack in forward channel

Consider the system (26) with uniform random attacks in the interval $[a, b]$ on the forward channel. Let $a = 0$ and $b = 40$ be the minimum and maximum values that can be multiplied by the control signal and get accepted at the DC motor. As an experimental setting, the number of intrusions changes by changing the probability of attack and then precision, recall and F-measure are calculated for the proposed IDS. The results of the experiments are listed in Table 1.

2) Scenario II: Intelligent attack in forward channel

Now, we examine the proposed intrusion detection algorithm by considering an intelligent attacker in the forward channel. First, we need to determine the parameters of LQR. We considered the results of Bryson's rule [45] as the starting point for a trial-and-error iterative process, therefore the following parameters were chosen for the system of (26),

$$R_0(k) = \left[-\tfrac{1}{9}\right], \quad Q_0(k) = \begin{bmatrix} -1e+08 & 0 \\ 0 & -\frac{1}{9000} \end{bmatrix}$$

$$R_1(k) = 0.4 \times \left[-\tfrac{1}{360}\right], \quad Q_1(k) = \begin{bmatrix} -\frac{1e+06}{45} & 0 \\ 0 & -\frac{1}{360} \end{bmatrix} \tag{28}$$

The attack occurrence follows the Bernoulli distribution. We set $p_{00} = 1 - \rho$, $p_{01} = \rho$, $p_{10} = 1 - \rho$, $p_{11} = \rho$.

We examined the proposed intrusion detection algorithm for intelligent attacks with these parameters and reported the results in Table 2 for different attack probabilities.

3) Scenario III: Attack in backward channel

Based on what was explained in Section 3, the attacker can obtain the secrete key of forward or backward channel. Suppose that the attacker compromises the backward channel key and sends its generated packet. Since the packet receiver is the controller, it can detect the malicious packets using the IDS. It discards the abnormal data and uses the local model output. There are no differences

Table 1: Detection results for proposed IDS by considering intelligent deception attack in forward channel.

Prob.	Precision	Recall	F-measure
0.01	1	0.9965	0.9983
0.02	1	0.9972	0.9986
0.03	1	0.9976	0.9988
0.04	1	0.9988	0.9994
0.05	1	0.9988	0.9994

Table 2: Detection results for proposed IDS by considering intelligent deception attack in forward channel.

Prob.	Precision	Recall	F-measure
0.0002	1	0.9111	0.9535
0.0004	1	0.9344	0.9661
0.0006	1	0.9895	0.9947
0.0008	1	0.9920	0.9960
0.001	1	0.9936	0.9968

between the malicious data generated randomly or intelligently here. Similar to the previous scenarios, we examine the proposed intrusion detection algorithm under different probabilities of intrusion. The results can be seen in Table 3.

4) Discussion

In this section, we confirm that our statistical intrusion detection and identification algorithm achieves successful and meaningful results. In a CPS, real-time intrusion detection is an important requirement that increases the robustness of the whole system and makes the system respond to intrusions more quickly. Our proposed algorithm is a real-time threshold-based IDS that detects intrusion and its place almost immediately. This approach, in addition to reducing the detection time, reduces the computational complexity of model learning in comparison with the model-based intrusion detection methods. We also analyze the security results, in term of precision, recall and F-measure. As we can observe in Table 1, Table 2, and Table 3, by adjusting proper detection thresholds the mentioned metrics in most cases goes ups to 0.99. First point which can be seen in Table 1 and Table 2 is that the proposed algorithm reduces the value of FP to zero, in the case that attack happens in the forward link,

Table 3: Detection results for proposed IDS by considering random deception attack in backward channel.

Prob.	Precision	Recall	F-measure
0.05	0.9924	0.9974	0.9949
0.1	0.9964	0.9968	0.9966
0.15	0.9978	0.9969	0.9973
0.2	0.9985	0.9969	0.9977
0.25	0.9989	0.9970	0.9979

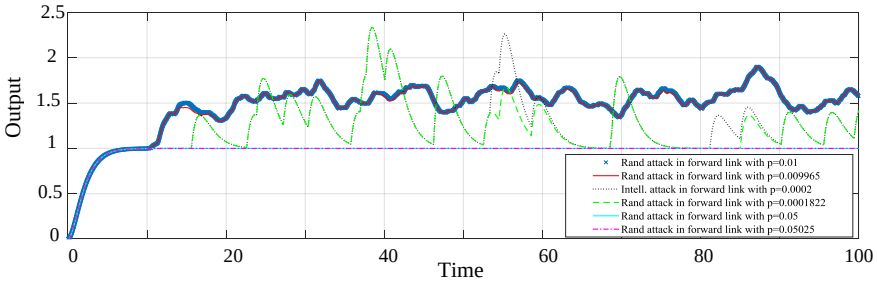

Figure 6: System output under different attack scenarios with determined probabilities and estimated ones by IDS.

thus it has the highest precision (*i.e.*1). Another thing that can be understood from the results is that, in the cases the attack probability is small, attacker tries to influence the control system and decrease its performance in a gentler way. Therefore, in such cases, most IDSs may classify some attacked data as normal and have nonzero FN and accordingly, low recall and F-measure rates. By increasing the attack probability, FN decreases and these two measures go up. In any case, the obtained results are satisfactory for our stability analysis purposes. For instance, in our system, according to Table 1 and for the random attack in the forward link with a probability of 0.01, the value of recall is equal to 0.9965. In fact, the IDS detects the presence of an attack at an incidence rate of 0.009965 that is so close to 0.01. As we can see in Fig. 6, the system outputs for these two rates of attack are close to each other. The results of mentioned example and two other cases (intelligent attack in the forward link with $\rho = 0.0002$ and random attack in the backward link with $\rho = 0.05$) are shown in Fig. 6. According to the relevant precision and recall, the probabilities for these two cases are estimated as $\rho = 0.0001822$ and $\rho = 0.05025$, respectively. Therefore, these results show that the proposed IDS gives good estimations of stationary distribution of switching signal, and the stability/instability of the system with the estimated attack probability can be checked with high confidence.

The results of our IDS are also comparable with other threshold-based intrusion detection methods. We examine the detection methods of [22], [23] and [24] in the presence of a random attack at $\rho = 0.0008$ and compare the triggered attack alarms of theirs with the result of our method. The results are reported in Fig. 7. As one can see in the figure, the intrusion detection methods of [22], [23] and [24], sense the presence of attack in most cases, but they fall short in the estimation of attack rate. The effects of attack remain for a while after its occurrence, and because of this phenomenon, these methods can not detect whether the data is still plagued by attack or not. Our proposed method, by using the adversarial model, decreases the misdiagnoses.

5.2. Stability analysis

Now we are going to discuss the stability of system under the three attack scenarios of the previous section. In order to identify the probability of attack (ρ) that satisfies sufficient stability conditions (discussed in Proposition 2 and 3), numerical examples are given. We illustrate the effect of stochastic switching signal on the stability of the system of (26). First, we implement the theoretical

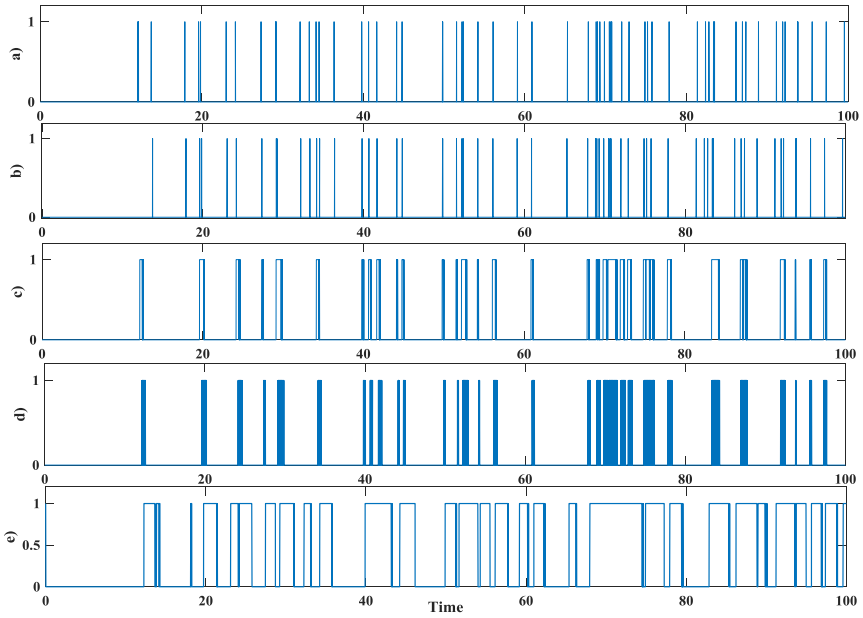

Figure 7: Comparison of different threshold-based intrusion detection methods for the system of Eq. (26) under a random attack with $\rho = 0.0008$: a) the attack signal (1: attack, 0: no attack), b) the triggered alarms by our proposed IDS, c) the triggered alarms by the intrusion detection method of [22], d) the triggered alarms by the intrusion detection method of [23], e) the triggered alarms by the intrusion detection method of [24].

part in Matlab and calculate the range of attack probabilities under which the system can stay stable. Then, we will evaluate the results by implementing the system in Simulink environment. These results show whether the attack is going to destabilize the system or not.

1) Scenario I: Random attack in forward channel

Consider the first attack scenario of Section 5.1. Results of running the Matlab code for this attack show that if the attacker changes the data in the forward channel by a probability of $\rho \leq 0.043$, the sufficient stability condition of Proposition 2 is satisfied and for $\rho > 0.043$, stability of system will not be guaranteed. To verify this theoretical result, the control system, is also implemented in the Simulink environment. The results of Simulink runs for the attack probabilities of $0, 0.025, 0.03, 0.035, 0.04, 0.045$ and 0.05, are shown in Fig. 8. Results confirm that the system under this attack with $\rho \leq 0.043$ remains stable.

2) Scenario II: Intelligent attack in forward channel

In the second scenario, the attacker replaces the data in the forward channel intelligently. These new data are based on an optimal control method that is designed through identifying the closed loop system. The LQR parameters have the same values as the ones presented in (28). Proposition 3 was implemented in Matlab in order to find the probabilities of attack (ρ) under which the system remains stable. By using the initial conditions of (28), for $\rho \leq 0.0004$, the sufficient condition will be satisfied. To validate this result, the system was examined for the attack probabilities of $0, 0.0002, 0.0003, 0.0004, 0.0007$, and 0.0009. The outputs of the system are depicted in Fig. 9, which validate the results of the theoretical part.

3) Scenario III: Attack in backward channel

Imagine that the attack occurs in the backward link. We examine this scenario by random attack with gain in range $[0, 45]$ which is equivalent to the random attack of forward channel with range $[0, 40]$. In this case, the controller detects the attacked data using the IDS, discards them and uses the local model output. As we said before, if the local model produces the output with high accuracy, stability of the attacked system will be equivalent to that of the normal system. Here, we examine the stability of system with attack in the backward channel with probabilities of 0.01 and 0.1. The results can be seen in Fig. 10. Considering the compensation mechanism that the controller uses, for any injected data and at any probability, the system should remain stable.

6. Conclusion

This paper studied intrusion detection and stability of CPSs under some kind of deception attack in which the attacker manipulates the original data stream by injecting false data. In this case, some of received packets at the destination are the original data packets and some are generated by the attacker, either randomly or intelligently. These attacks make the system behave as a switching system. The switching signal stochastically switches between normal and attacked modes and is modeled using a Markovian jump modelling approach. The contributions of this work were in two parts. We first proposed an adversarial model along with an intrusion detection and identification system that detects the presence of any deception attack in the closed loop and

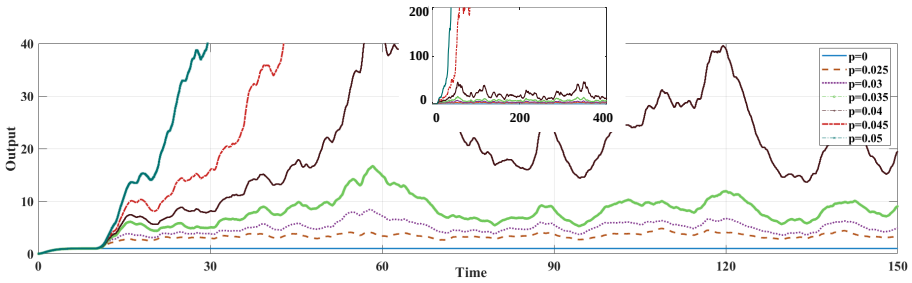

Figure 8: System output under random attack in forward link with the $\rho = 0, \rho = 0.025, \rho = 0.03, \rho = 0.035, \rho = 0.04, \rho = 0.045$, and $\rho = 0.05$.

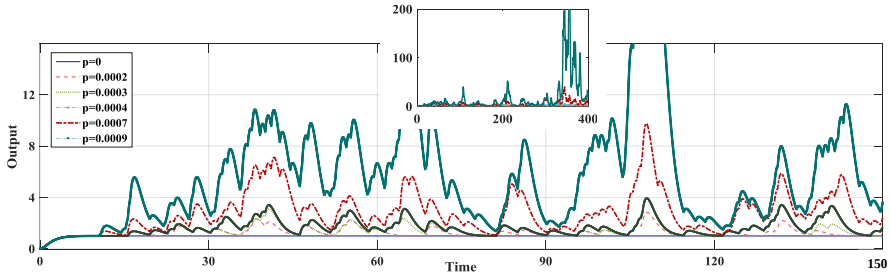

Figure 9: System output under intelligent deception attack with $\rho = 0, \rho = 0.0002, \rho = 0.0003, \rho = 0.0004, \rho = 0.0007$, and $\rho = 0.0009$.

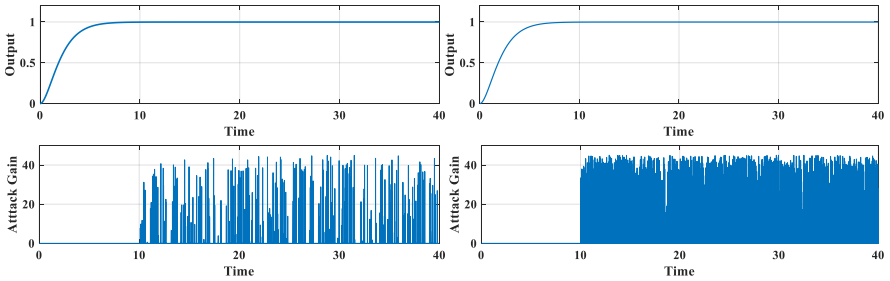

Figure 10: System output and attack gain under random attack in backward link when the attack probability is $\rho = 0.01$ (left figures) and $\rho = 0.1$ (right figures).

spots its location. This is a real time operation. The proposed schema brings us the rate and the place of attack, which are helpful to predict whether an attack will destabilize system or not. Then, two sufficient conditions for system stability under random and intelligent deception attacks were derived. Some numerical examples were used to validate the developed theory and illustrate its applications. The results were consistent with our intuitions and confirmed the validity of theoretical findings. Future research could be directed toward the development of stability conditions for nonlinear systems as well as randomly switched time-delay systems.

APPENDIX

Proof of Proposition 1. According to definition 1, $\lambda = lim_{k\to\infty} sup \frac{1}{k} ln\|\phi(k,0)\|$ is the top Lyapunov exponent of discrete time linear system. The mentioned system is AS-stable if and only if $\bar{\lambda} = E[\lambda] < 0$.

Consider

$$\bar{\lambda} = lim_{k\to\infty} \frac{1}{k} E_{\bar{\pi}}[ln\|\phi(k,0)\|] \tag{29}$$

Now, similar to the proof of the theorem 3.1 of the reference [29], let $k = Nm + h, 0 \leq h \leq m-1$ and

$$\frac{1}{k} E_{\bar{\pi}}[ln\|\phi(k,0)\|] = \frac{1}{Nm+h} E_{\bar{\pi}}[ln\|\phi(k,0)\|] \leq \frac{1}{Nm+h} E_{\bar{\pi}}[ln\|\phi(h,0)\|]$$

$$+\frac{1}{Nm+h} E_{\bar{\pi}}[\sum_{j=1}^{N} ln\|\phi(jm+h,(j-1)m+h)\|] \tag{30}$$

$$= \frac{1}{Nm+h} E_{\bar{\pi}}[ln\|\phi(h,0)\|] + \frac{N\bar{\alpha}}{Nm+h}$$

where $\bar{\alpha} = E_{\bar{\pi}}[ln\|\phi(m,0)\|] < 0$. Then, taking the limit as $N \to \infty$, it is immediate to see that

$$\bar{\lambda} \leq \frac{\bar{\alpha}}{m} < 0 \tag{31}$$

By increasing m, a sequence of sufficient conditions is obtained.

References

[1] M. Sharma, H. Elmiligi, F. Gebali, Network security and privacy evaluation scheme for cyber physical systems (cps), in: Handbook of Big Data Privacy, Springer, 2020, pp. 191–217.

[2] T. Glad, L. Ljung, Control theory, CRC press, 2018.

[3] C. Wang, J. Huang, D. Wang, F. Li, A secure strategy for a cyber physical system with multi-sensor under linear deception attack, Journal of the Franklin Institute 358 (13) (2021) 6666–6683.

[4] R. Qi, S. Ji, J. Shen, P. Vijayakumar, N. Kumar, Security preservation in industrial medical cps using chebyshev map: An ai approach, Future Generation Computer Systems 122 (2021) 52–62.

[5] S. M. Dibaji, A. Hussain, H. Ishii, A tutorial on security and privacy challenges in cps, Security and Resilience of Control Systems (2022) 121–146.

[6] S. Amin, X. Litrico, S. Sastry, A. M. Bayen, Cyber security of water scada systems part i: Analysis and experimentation of stealthy deception attacks, IEEE Transactions on Control Systems Technology 21 (5) (2012) 1963–1970.

[7] J. Xu, L. Wei, W. Wu, A. Wang, Y. Zhang, F. Zhou, Privacy-preserving data integrity verification by using lightweight streaming authenticated data structures for healthcare cyber–physical system, Future Generation Computer Systems 108 (2020) 1287–1296.

[8] Z.-H. Pang, G. Liu, Z. Dong, Secure networked control systems under denial of service attacks, IFAC Proceedings Volumes 44 (1) (2011) 8908–8913.

[9] L. Guo, H. Yu, F. Hao, Optimal allocation of false data injection attacks for networked control systems with two communication channels, IEEE Transactions on Control of Network Systems 8 (1) (2020) 2–14.

[10] Z.-H. Pang, G.-P. Liu, Design and implementation of secure networked predictive control systems under deception attacks, IEEE Transactions on Control Systems Technology 20 (5) (2011) 1334–1342.

[11] A. O. de Sá, L. F. R. da Costa Carmo, R. C. Machado, Covert attacks in cyber-physical control systems, IEEE Transactions on Industrial Informatics 13 (4) (2017) 1641–1651.

[12] C. Kwon, W. Liu, I. Hwang, Security analysis for cyber-physical systems against stealthy deception attacks, in: 2013 American control conference, IEEE, 2013, pp. 3344–3349.

[13] A. Tahoun, M. Arafa, Cooperative control for cyber–physical multi-agent networked control systems with unknown false data-injection and replay cyber-attacks, ISA transactions 110 (2021) 1–14.

[14] Z. Gu, X. Zhou, T. Zhang, F. Yang, M. Shen, Event-triggered filter design for nonlinear cyber–physical systems subject to deception attacks, ISA transactions 104 (2020) 130–137.

[15] M. Sayad Haghighi, F. Farivar, A. Jolfaei, M. H. Tadayon, Intelligent robust control for cyber-physical systems of rotary gantry type under denial of service attack, The Journal of Supercomputing 76 (4) (2020) 3063–3085.

[16] F. Farivar, M. S. Haghighi, S. Barchinezhad, A. Jolfaei, Detection and compensation of covert service-degrading intrusions in cyber physical systems through intelligent adaptive control, in: 2019 IEEE International Conference on Industrial Technology (ICIT), IEEE, 2019, pp. 1143–1148.

[17] S. Barchinezhad, M. S. Haghighi, Compensation of linear attacks to cyber physical systems through arx system identification, arXiv preprint arXiv:2002.05798.

[18] S. Pan, T. H. Morris, U. Adhikari, A specification-based intrusion detection framework for cyber-physical environment in electric power system., Int. J. Netw. Secur. 17 (2) (2015) 174–188.

[19] M. H. Monzer, K. Beydoun, J.-M. Flaus, Model based rules generation for intrusion detection system for industrial systems, in: 2019 International Conference on Control, Automation and Diagnosis (ICCAD), IEEE, 2019, pp. 1–6.

[20] M. S. Haghighi, F. Farivar, A. Jolfaei, A machine learning-based approach to build zero false-positive ipss for industrial iot and cps with a case study on power grids security, IEEE Transactions on Industry Applications.

[21] F. Farivar, M. S. Haghighi, A. Jolfaei, M. Alazab, Artificial intelligence for detection, estimation, and compensation of malicious attacks in nonlinear cyber-physical systems and industrial iot, IEEE transactions on industrial informatics 16 (4) (2019) 2716–2725.

[22] Y. Mo, E. Garone, A. Casavola, B. Sinopoli, False data injection attacks against state estimation in wireless sensor networks, in: 49th IEEE Conference on Decision and Control (CDC), IEEE, 2010, pp. 5967–5972.

[23] C. Murguia, J. Ruths, On model-based detectors for linear time-invariant stochastic systems under sensor attacks, IET Control Theory & Applications 13 (8) (2019) 1051–1061.

[24] P. J. Bonczek, N. Bezzo, Memoryless cumulative sign detector for stealthy cps sensor attacks, IFAC-PapersOnLine 53 (2) (2020) 838–844.

[25] J. Huang, D. W. Ho, F. Li, W. Yang, Y. Tang, Secure remote state estimation against linear man-in-the-middle attacks using watermarking, Automatica 121 (2020) 109182.

[26] D. Ding, Z. Wang, Q.-L. Han, G. Wei, Security control for discrete-time stochastic nonlinear systems subject to deception attacks, IEEE Transactions on Systems, Man, and Cybernetics: Systems 48 (5) (2016) 779–789.

[27] A.-Y. Lu, G.-H. Yang, Stability analysis for cyber-physical systems under denial-of-service attacks, IEEE Transactions on Cybernetics 51 (11) (2020) 5304–5313.

[28] X. Chen, Y. Wang, S. Hu, Event-based robust stabilization of uncertain networked control systems under quantization and denial-of-service attacks, Information Sciences 459 (2018) 369–386.

[29] P. Bolzern, P. Colaneri, G. De Nicolao, Almost sure stability of stochastic linear systems with ergodic parameters: An average contractivity criterion, in: Proceedings of the 45th IEEE Conference on Decision and Control, IEEE, 2006, pp. 950–954.

[30] P. Bolzern, P. Colaneri, G. De Nicolao, On almost sure stability of continuous-time markov jump linear systems, Automatica 42 (6) (2006) 983–988.

[31] Y. Song, H. Dong, T. Yang, M. Fei, Almost sure stability of discrete-time markov jump linear systems, IET Control Theory & Applications 8 (11) (2014) 901–906.

[32] D. J. Antunes, H. Qu, Frequency-domain analysis of networked control systems modeled by markov jump linear systems, IEEE Transactions on Control of Network Systems 8 (2) (2021) 906–916.

[33] J. Gao, Z. Zhao, J. Wang, T. Tan, M. Ma, Event-triggered output feedback control for discrete markov jump systems under deception attack, Journal of the Franklin Institute 357 (11) (2020) 6435–6452.

[34] B. Chen, Y. Niu, Y. Zou, Security control for markov jump system with adversarial attacks and unknown transition rates via adaptive sliding mode technique, Journal of the Franklin Institute 356 (6) (2019) 3333–3352.

[35] Z. Cao, Y. Niu, J. Song, Finite-time sliding-mode control of markovian jump cyber-physical systems against randomly occurring injection attacks, IEEE Transactions on Automatic Control 65 (3) (2019) 1264–1271.

[36] J. Lian, Y. Han, Switching-like event-triggered control for networked markovian jump systems under deception attack, IEEE Transactions on Circuits and Systems II: Express Briefs 68 (10) (2021) 3271–3275.

[37] S. Thakur, A. Chakraborty, R. De, N. Kumar, R. Sarkar, Intrusion detection in cyber-physical systems using a generic and domain specific deep autoencoder model, Computers & Electrical Engineering 91 (2021) 107044.

[38] H. Ge, D. Yue, X. Xie, S. Deng, C. Dou, A unified modeling of mutisources cyber-attacks with uncertainties for cps security control, Journal of the Franklin Institute 358 (1) (2021) 89–113.

[39] B. Xie, C. Peng, M. Yang, X. Kong, T. Zhang, A novel trust-based false data detection method for power systems under false data injection attacks, Journal of the Franklin Institute 358 (1) (2021) 56–73.

[40] M. S. Mahmoud, Observer-based control design: Basics, progress, and outlook, in: New Trends in Observer-Based Control, Elsevier, 2019, pp. 143–208.

[41] H. J. Chizeck, A. S. Willsky, D. Castanon, Discrete-time markovian-jump linear quadratic optimal control, International Journal of Control 43 (1) (1986) 213–231.

[42] A. Czornik, A. Nawrat, M. Niezabitowski, Lyapunov exponents for discrete time-varying systems, in: Advanced Technologies for Intelligent Systems of National Border Security, Springer, 2013, pp. 29–44.

[43] C. Chen, Linear system theory and design, 1999, Type for questions for University Exams Question (1)-Eight short answer question of five marks with two questions from each of four modules Question (2-5)-Two questions A & B of 15.

[44] D. Bertsekas, Dynamic programming and optimal control (athena scientific, 1995), Vol. II.

[45] J. P. Hespanha, Linear systems theory, Princeton university press, 2018.