Available online at www.sciencedirect.com

ScienceDirect

IFAC Papers Online CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 50-1 (2017) 387–392

A Game Theoretic Approach to Design a Resilient Controller For a Nonlinear Discrete System

Hossein Salehghaffari * **Prashanth Krishnamurthy** * **Farshad Khorrami** *

* *Control/Robotics Research Laboratory (CRRL), Department of Electrical and Computer Engineering, NYU Tandon School of Engineering, NY 11201, USA*
e-mails: {h.saleh,prashanth.krishnamurthy,khorrami}@nyu.edu

Abstract: In this paper, the problem of designing a cyber-resilient controller is studied for a class of nonlinear discrete systems subject to actuator attacks. We develop a coupled control design approach, which incorporates interaction between the intrusion detection system (IDS) and the attacker, with the controller design in the physical layer. In the considered cyber-attack scenario, the attacker attempts to first bypass the IDS to spoof the system actuator and then deteriorate the controller performance in the physical layer. The optimal attack strategy to deceive the IDS is obtained based on a zero-sum game in the detection layer. In the physical layer, it is assumed that the system has a secure compensator along with the actuator to deal with the injected attack values. Therefore, the controller and the compensator are designed based on the results of the game in the detection layer, and stochastic stability analysis and Lyapunov function methods are used to prove boundedness of the system state in the probabilistic sense. Finally, through numerical analysis of a representative example, the proposed design procedure is illustrated and its efficacy in maintaining robust stability of the cyber-physical system under actuator attacks is demonstrated.

Keywords: Cyber-physical system, resilient control, zero-sum game, stochastic stability, attack strategy, control systems security.

1. INTRODUCTION

While increasing connectivity of embedded controller devices in many current infrastructures such as power grids and transportation systems has enhanced on-demand reconfiguration/reprogrammability functionalities and reduced operator workload, this increasing connectivity has also made these cyber-physical systems (CPS) vulnerable to cyber-attacks. In particular, a crucial class of attacks that are relevant to CPS are process-aware attacks (Khorrami et al. (2016)) that attempt to utilize knowledge of the system dynamics or characteristics of its hardware/software components to hamper performance or stability of the system. An attack on control systems can be considered as an unexpected change in run-time signals from/to control systems components such as sensors and actuators. In one point of view, this type of attack is similar to disturbances, although many assumptions about disturbances in control theory may not hold in the case of such an attack. The term "resilient control" is used for a control system which is robust to cyber-physical attacks, and the resiliency of the controller relates to its capability in retaining state awareness and acceptable operational normalcy in the presence of unexpected threats (Rieger (2010)).

In the literature, multiple approaches to cope with various types of attacks have been considered. To model denial of service (DoS) attacks on control systems, Liang et al. (2009); Lin et al. (2009); Wang et al. (2009, 2007); Zhao et al. (2009) have considered the robust control problem for systems with random communication packet losses. Wang et al. (2007) utilized Bernoulli random variables to model packet losses in communication channels from sensors to controller and controller to actuators. Thereafter, the problem of output feedback control design for the discrete linear time invariant system has been solved based on a stochastic definition of the Lyapunov function and the concept of mean square stability in the probabilistic sense. Finally, the controller gain is designed based on the solvability of a certain LMI condition. Lin et al. (2009) uses the concept of exponential mean square stability to design an observer-based control for continuous-time systems with random sensor delays.

Mo and Sinopoli (2009) analyze the effects of a replay attack on the control system where the attacker records the sensor readings for a certain amount of time and repeats the readings while launching the attack on the system. Mo and Sinopoli (2009) propose a feasibility condition for replay attack in a discrete linear system equipped with a specific failure detector. In Miao and Zhu (2014), the problem of secure control design is considered as a hybrid stochastic game. The game model is used for designing defense policies against various types of attacks. Based

2405-8963 © 2017, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved.
Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2017.08.178

on the saddle-point solution of the game, a switching mechanism is designed to choose the appropriate controller among a set of controllers.

On the other hand, to increase robustness of systems to various attack types, several papers have considered the attack detection problem along with the controller design problem (Linda et al. (2011); Pasqualetti et al. (2012, 2013)). The problem of machine learning based attack detection has been analyzed in Keliris et al. (2016). Utilizing support vector machines (SVMs), Keliris et al. (2016) demonstrated a process-aware defense and mitigation strategy. A model based attack detection and mitigation for power systems has been proposed in Sridhar and Govindarasu (2014). The impact of a data integrity attack on power system frequency is analyzed, and then a general framework to design an attack resilient control for power systems as a combination of the attack detection and mitigation mechanism is proposed. A mathematical framework for attack detection and security in water systems was developed in Eliades and Polycarpou (2010). Additionally, to increase the efficiency of the detection system, the sensor placement problem is solved by utilizing various optimization approaches. Yuan et al. (2013) has integrated the IDS configuration problem with the control design problem to build a resilient control against DoS attack. The paper utilizes a reinforcement algorithm to compute the optimal IDS configuration and control laws.

In this paper, we consider attack detection and controller design as a coupled problem. The problem of data integrity attack on the actuator of a nonlinear discrete system in strict feedback form is investigated in this paper. A Bernoulli random variable is utilized to model actuator spoofing, where the expected value of the Bernoulli variable specifies the fraction of time that the actuator is under attack on average. It is assumed that the attacker is intelligent, and has complete information about the IDS vulnerabilities, cost of the detection, and risk of the attack on each subsystem of the IDS. With such information, an adversary can attempt to optimally bypass the IDS. In the absence of such information, the attacker strategy is sub-optimal. However, the proposed control design is still applicable to offer resiliency against actuator attacks. Motivated by the literature (Alpcan and Basar (2004)), a zero-sum stationary game is proposed to capture various security aspects of interaction between the attacker and the IDS. The game result determines the optimal value injected by the attacker into the system. It is assumed that the system has a secure compensator along with the actuator. In the physical layer, actuator spoofing is modeled by a Bernoulli random variable. Thereafter, the controller and the compensator are designed based on the backstepping approach and the stochastic Lyapunov function approach.

The organization of the paper is as follows: A game theoretic approach to analyze the defender and attacker strategies is developed in Section 2. In Section 3, the actuator attack model is formulated and relevant aspects of stochastic control theory are briefly summarized. The control design problem and stability analysis are addressed in Section 4. Section 5 provides simulation results of the proposed attack resilient control design. In the last section, concluding remarks and directions for future works are discussed.

2. CYBER LAYER GAME

Game theory is a very effective tool to address problems where multiple players with different objectives compete and interact with each other. In this paper, we use a game theoretic approach to model interaction between the attacker and the defender in a class of control systems. The attacker goal is to degrade the control performance of the system by manipulating run-time signals such as sensor and actuator values. In the cyber layer, the IDS is responsible for detecting abnormality in the system by monitoring states of the system.

To model a cross-layer attack for the integrated cyber-physical system, a continuous zero-sum game is introduced to capture various security aspects of the game played by the attacker and the IDS. The result of the cyber layer game is used to find the best attack strategy in the physical layer. The architecture of the proposed cyber-physical resilient control system is shown in Fig. 1. In this section, we propose a general framework to demonstrate the competition between the attacker and the IDS. This framework can be utilized to obtain the optimal attack strategy for multiple types of attack including sensor and actuator attacks.

In this paper, it is assumed that the attacker chooses his attack value from a set of unknown probability distributions which model additive perturbations introduced into the system dynamics with known expected values, and the attack set is defined as the expected values of these probability distribution functions. The attacker launches his attack from the attack set $A = \{a_1, a_2, \dots, a_m\}$ where $a_i, i = 1, 2, \dots, m$, is the expected value of an unknown probability distribution function. It is assumed that the IDS uses a machine learning approach to detect the attack. Therefore, its detection capability depends on the data set used for training the classifier. The set $D = \{d_1, d_2, \dots, d_p\}$ is considered to be the set of all data sets available to the IDS. The IDS can choose different subsets of the data set to detect various kinds of attacks. So, the IDS can maximally have $2^{|D|}$ different configurations where $|D|$ is cardinality of the data set D. In this paper, we assume that the IDS utilizes n different configurations to deal with different attack types. $F_i \in D, i = 1, 2, \dots, n$, denote the IDS configurations to deal with different attack types. The performance of the IDS depends on its configuration. For example, assume that configurations F_1 with data set $D_{F_1} = \{d_1, d_2\}$ and F_2 with data set $D_{F_1} = \{d_1, d_2, d_3\}$ are used to detect attack types a_1 and a_2. Consider that the attack a_1 is fully detectable with configuration F_1 and F_2. Also, consider that the attack a_2 is partially detectable with F_1 configuration but fully detectable with F_2 configuration. Consequently, to detect attack a_1, F_1 has a superior performance because F_1 uses fewer data sets than F_2. On the other hand, a_2 is not fully detectable by F_1. This provides an example to illustrate the importance of using different configurations for the IDS in different situations. The utilization of different configurations of data sets is especially relevant under computation/memory constraints that are always present,