Journal of Process Control 68 (2018) 160–170

Contents lists available at ScienceDirect

Journal of Process Control

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

A mechanistic fault detection and isolation approach using Kalman filter to improve the security of cyber physical systems

Venkata Reddy Palleti[a,*], Tan Yu Chong[b], Lakshminarayanan Samavedham[b]

[a] *iTrust Centre for Research in Cyber Security, Singapore University of Technology and Design, Singapore*
[b] *Department of Chemical and Biomolecular Engineering, National University of Singapore, Singapore*

ARTICLE INFO

Article history:
Received 28 January 2018
Received in revised form 18 April 2018
Accepted 12 May 2018

Keywords:
Cyber-physical systems
Fault detection and isolation
Security
Cyber-attack

ABSTRACT

This paper seeks to understand and identify ways in which fault detection and isolation (FDI) methods can be utilized to enhance the cyber-security of cyber-physical systems (CPS). In this work, we have used state estimation to validate relations between process variables, termed *invariants*, thereby detect the onset of attacks. Multiple attack scenarios were studied and the proposed state estimation method was found to excel in detecting attacks conducted on sensors within the system; provided that not all sensors are compromised during the attack. However, the proposed methods for subsequent isolation and correction faced limitations due to delays or lack of information to pinpoint the attacked component. We demonstrated our proposed approach on a well-instrumented pilot scale water treatment plant equipped with controllers.

1. Introduction

Chemical processes involve precise unit operations to convert raw materials into valued products demanded by consumers [1]. Hence, disruptions to these operations will result in lost time and product quality degradation, which will compromise the profitability of the operation and at worst, cause safety incidents. It is therefore paramount to regulate the process, which is the primary objective of process control [2]. The study of process control involves understanding the dynamics of the unit operations and subsequent design of control strategies to maintain them at their designated states. The interactions between the industrial control system (ICS), which execute the control strategies, and unit operations, which are governed by physical laws, form a network of computing and physical system which is termed as the cyber physical system (CPS). They are considered as the new generation of embedded control systems, where sensor and actuator networks are embedded [3].

It is often believed that CPS is a well-isolated system that is not susceptible to external malicious attacks. However, the advent of Internet-of-Things and increasing Internet connectivity has greatly diminished the physical isolation [4], allowing malwares to establish attacks on CPS. Recent examples include Stuxnet, which damaged centrifuges in Iran's nuclear facilities [5] and BlackEnergy, which caused a blackout in Ukraine. Consequently, improving network security and traffic monitoring is required to combat cyber-attacks.

However, besides attempting to tackle cyber-attacks head-on, process control may help devise smarter control strategies that can aid in counteracting cyber-attacks. This paper explores the subject of conventional fault detection and isolation (FDI) methods and attempts to apply the mechanistic methods of FDI using Kalman filter and CUSUM test for online detection of cyber-attacks. Subsequent isolation and correction techniques are incorporated to observe the real-time functionality of the proposed FDI. A simulated environment was constructed based on a water treatment test bed at Singapore University of Technology and Design.

Modern plant management is designed such that plant level controllers communicate with higher level control system like manufacturing execution system (MES) and enterprise resource planning (ERP) that considers additional inputs from inventory management and economic factors to perform real-time optimization in order to achieve business objectives; the optimal results are returned to the plant level controllers as operating set points. Modern alarm management instruments offer wireless features that deliver real time updates of process performance to human machine interfaces (HMIs) or even to original equipment manufacturers, for the purpose of online troubleshooting. These added integration open up more entry points for malicious attacks and fur-

[*] Corresponding author.
E-mail address: venkata_palleti@sutd.edu.sg (V.R. Palleti).

https://doi.org/10.1016/j.jprocont.2018.05.005
0959-1524/© 2018 Elsevier Ltd. All rights reserved.

ther diminishes the isolation of the ICS, placing the CPS in danger of attack as well.

Many works on the subject have been largely centered on the cyber-aspect, which focuses on improving network security and monitoring of the data packets in the network traffic [6–9]. More efficient zoning and partitioning are proposed [10] to establish secured doorways where data transfer can be closely monitored at entry into the plant-level controls. Design and incorporation of honeypots [11] are used to reconnaissance upcoming attacks in order to alert operators and learn the nature of the attack or to act as decoys to minimise interactions that attackers have with the actual CPS infrastructure.

However, besides securing the borders of the CPS network, researchers in [12] assert that there is also an impetus to build smarter control units within the CPS perimeter that are capable of flagging attacks and stabilizing the system when under an attack. CPS needs to be complemented with a smarter control system that is able to inoculate itself against malicious attack, which calls for multi-disciplinary collaborations between information technology and process control. In particular, FDI has the potential to aid the development of smart control system due to the similar characteristics displayed by faulty states and attacked states. Both faulty and attacked states propagate error through the system which may lead to failure. However, the source of error is different. For faulty states, the source is usually physical, i.e. a drifting sensor, thus its effects can theoretically be predicted with a first principles model. For attacked states, it is usually due to a malicious attempt to shift the operating state and it is presumed that the attacker has the motivation to mask his attack. This introduces disparity between the observed and controlled state, thus the future states of the system may not change according to prescriptions of its first principle model. Thus attacks may be uncovered by observing the coherence between the observed and controlled state. FDI will further aid in identify the source of the attack and provide means to implement online corrections to mitigate the effect of the error source.

The work in this paper will seek to illustrate how FDI methods can contribute towards enhancing the cyber-security of CPSs. Specifically, state-estimation based FDI is examined and tested on a water treatment testbed.

2. Background

Recent trend in fault detection has been leaning towards the use of multivariate statistical techniques and machine learning approaches due to improved data collection and handling facilities. Furthermore, mechanistic models are time-consuming and difficult to build. In the context of data rich but information poor chemical processes, principal component analysis (PCA) and other techniques have shown to be able to model large amounts of noisy data obtained during operations and provide useful information for effective fault detection [13–15].

Further fault isolation is done using contribution plots [16], this technique works effectively for simple faults that has little propagation effects on neighboring variables and is widely favored when historical faulty data are not readily available. Furthermore, it was demonstrated on a simulation of a CSTR system with feedback control that with added learning data of a few common-caused faults, the multivariate statistical method can be very effective in isolating both simple and complex faults [17]. In [18], researchers discussed the problem of FDI of nonlinear process systems subjected to input constraints and sensor faults in the context of asynchronous measurements or complete sensor failure. They designed a fault tolerant controller that accounts for the process nonlinearity. Similarly, FDI is implemented on multi-input multi-output nonlinear system subject to fault in the control actuators and constrains on the manipulated inputs [19]. Further, [20] presented a nonlinear feed-forward/feedback based control model and FDI techniques on a high recovery reverse osmosis desalination process. The benefits of multivariate statistical methods in FDI are numerous and promising but at the time of writing, the authors are unaware of any publications that has demonstrated the use of PCA in detecting cyber-attacks on CPS. A driving cause may be the computation intensive nature of its algorithms hence ill-suited for online implementations [21].

In contrast, traditional fault detection techniques employ mechanistic models and state estimators [17]. State estimators like Kalman filters are computationally non-intensive and are well-suited for online implementations. Real time detection capabilities are highly desirable as cyber-attacks, if unnoticed, are capable of quickly driving processes out of control. Additionally, mechanistic models offer the benefit of illustrating causal relations between variables; though they may be difficult to obtain for complex processes. In [22], authors devised a model predictive control method that optimizes a finite time-horizon into the future to generate alarms that flag hazardous event in the future. Their method involves the use of first-principles model, state estimators and a set of operating constraints. Further, [23] also showed another form of fault isolation accomplished using a bank of auxiliary state observers that processes subsets of the process variables.

2.1. Development in process-oriented cyber attack prevention

In line with the notion of uncovering cyber-attacks from the process angle, [24] has proposed a systematic method to derive *invariants* for detecting cyber-attacks on a CPS. The invariant described is essentially a causal relation between variables that is established by physical laws or constraints of the operation. Examples are the material balance between inflows and outflows across a tee and process safety interlocks engineered to prevent tank overflows. They assert that these invariants must hold true for normal operating conditions hence any violations detected are signs of the system under an attack. Their work concerned with strategic manipulation of sensor measurements and strategic control of actuators. By doing so, an attacker can maintain the invariants derived using, for example, material balance, but cause the process to move into an abnormal state without a controller knowing that this has happened. The method was implemented on a water treatment test bed and has shown to be effective in uncovering a variety of attacks. However, the method fails to differentiate if the violations of the invariants were due to actual cyber-attacks or a plant/instrumentation fault developed over time. The method will also cease to detect violations if a complete compromise of all sensors is achieved by the attacker as the control system will lose observability of its process. The attacker can report normal operating data to the controlling units while manipulating actuators with false commands (Man-in-the-Middle attack). The latter scenario is assumed to be uncommon due to the practical difficulties in designing and launching such a precise and synchronised attack. The former scenario will require additional inputs from network monitoring or physical inspection of relevant instruments to distinguish the exact cause of failure and therefore define the subsequent course of action. Nonetheless, invariants-based detection provides useful clues towards a more robust cyber-attack prevention. At the time of writing, the authors recognise that a research team at the water treatment test bed is pursuing the identification of invariants in the system using PCA techniques in order to develop a full-scale detection scheme and test for the effectiveness of the scheme.

Recently, [25] has also investigated the effectiveness of yet another detection scheme built with Kalman filter and a chi-square detector on the same water treatment testbed system. Their work claimed that the state-estimator has limited use in cyber-attack