Nuclear Engineering and Design 331 (2018) 204–210

Contents lists available at ScienceDirect

Nuclear Engineering and Design

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

Uncovering cyber-threats to nuclear system sensing and observability

Lee T. Maccarone, Christopher J. D'Angelo, Daniel G. Cole*

Department of Mechanical Engineering and Materials Science, University of Pittsburgh, 3700 O'Hara Street, Pittsburgh, PA, United States

ARTICLE INFO

Keywords:
Vulnerability analysis
Cybersecurity
Cyber-attack
Observability
Kalman decomposition

ABSTRACT

Cyber-physical systems are engineered systems that integrate physical processes and computational resources. But, by integrating cyber and physical worlds, the physical assets are vulnerable to cyber-attack. Two things are of importance for the security of cyber-physical assets: access to control inputs by the attacker, and the ability of an attacker to mask inputs. This combination of attacker control and masking measurements can allow an attacker to cause significant damage to a system while remaining undetected. By masking certain measurement signals, an attacker may affect the observability of the system and create a condition where part of the state space is unobservable, meaning that it is impossible to reconstruct those states. This is called an observability attack.

This paper presents a technique for analyzing observability attacks. How an attacker can design an attack to maximize the impact on the unobservable states while minimizing the possibility of detection is discussed. Criteria for maintaining a stealthy attack are given, and a design method is provided. For a nuclear balance of plant system, combinations of sensor omissions are analyzed to find an observability attack with maximum impact and minimum detection. An appropriate attack input signal is created, an attack is simulated, and the system response is shown.

1. Introduction

Cyber-physical systems (CPS) are engineered systems that integrate seamlessly physical processes, computational resources, and the communication networks that connect the two (Lee, 2008). Nuclear power plants are an example of CPS that integrate the reactor core, reactor coolant system, engineered safeguards, instrumentation and control (I&C), plant control systems, and plant networks. As nuclear I&C continues to transition to digital, it is important to protect the physical assets of the plant, like the core, that could be vulnerable to a cyber-attack. This paper develops tools for analyzing the vulnerability of nuclear systems to a certain type of cyber-attack.

An attack on a CPS can have severe physical consequences such as damaging equipment and processes. A well-known instance of a successful cyber-attack is Stuxnet (Falliere et al., 2011). Stuxnet targeted Iran's nuclear program, specifically centrifuges used to enrich uranium, and infected, by modifying code on programmable logic controllers, the industrial control systems used to control the centrifuges (Falliere et al., 2011). These modifications were designed both to hide critical measurements by masking sensor signals and to be stealthy by causing deviations in standard control signals (centrifuge variable-frequency drives) so that the accumulated effects would only become detectable long after infection (Karnouskos, 2011; Fidler, 2011). By causing them to operate erratically and at unsafe speeds, Stuxnet damaged the centrifuges unbeknownst to the operators (Greengard, 2010).

While policies and procedures exist to reduce the likelihood of a cyber-event, attacks like Stuxnet show that cyber-attacks can cross air gaps and that other vulnerabilities may exist that give attackers access to system inputs and outputs. For example, viruses and malicious code can infect critical digital assets when vendors connect to plant equipment using vendor diagnostic equipment. Also, malicious firmware, having infected equipment somewhere up the supply chain, can be present in digital I&C. Scanning device firmware is very difficult, with few, if any, commercial solutions to protect critical digital assets from such threats (Skorobogatov and Woods, 2012).

The Stuxnet example highlights two things that are of importance for the security of cyber-physical assets: First, an attacker can gain access to a control input to the system in question. Using this control input, the attacker can create inputs that drive the physical system to unsafe conditions. In a nuclear power plant, the attacker might cause temperatures, pressures, power, or levels to exceed limits. When these limits are exceeded, operators rely upon alarms and trip signals to be triggered to ensure the integrity of plant systems. Second, if the attacker can mask signals, either by limiting their availability or by replaying nominal signals, they can hide abnormal conditions from the operator. This combination of attacker control and masking measurements can

* Corresponding author.
E-mail addresses: ltm10@pitt.edu (L.T. Maccarone), cjd66@pitt.edu (C.J. D'Angelo), dgcole@pitt.edu (D.G. Cole).

https://doi.org/10.1016/j.nucengdes.2018.02.028
Received 7 September 2017; Received in revised form 8 January 2018; Accepted 19 February 2018
Available online 20 March 2018

allow the attacker to cause significant damage to a system while remaining undetected.

State observers are a powerful tool for monitoring system behavior and detecting cyber-threats even when an attacker has masked signals. When a system is observable, one can use the available measurements and control inputs with a system model to reconstruct mathematical estimates of internal (unmeasured or hidden) state variables. Previously hidden signals can be monitored, and if they exceed limits, mitigating actions can be taken by an operator. The ability to construct an observer depends upon whether the system is observable, a system characteristic discussed later in the paper. The observability of a system depends upon the dynamics of the system, how its states interact, and how those states map to the measured outputs. One problem is that by masking certain measurement signals, an attacker may affect the observability of the system. It may be possible for the attacker to create a condition where part of the state space is unobservable, meaning it is impossible to reconstruct those states. If that is the case, the attacker may be able to steer the unobservable states to unsafe conditions and maintain stealth by limiting the response of the observable states.

In this paper, we present a technique for analyzing an existing CPS for such *observability attacks*. In addition, we demonstrate how an attacker can design an attack to maximize the impact on the unobservable states while minimizing the possibility of detection. We consider the scenario where an attacker has access to the system's inputs, and can eliminate one or more measurements. Combinations of sensor omissions are analyzed to find an observability attack with maximum impact and minimum detection. Unobservable subspaces, those that the attacker can control but that cannot be observed, are identified by applying a Kalman decomposition. Given this information, an appropriate attack input signal is created, an attack is simulated, and the system response is shown.

The layout of this paper is as follows: Section 2 contains the theoretical foundation for an unobservable attack, followed by Section 3, which describes how an observability attack would be designed and gives criteria for a stealthy attack. Section 4 describes the model of the power conversion cycle for a nuclear power plant. The results of the system decomposition and physical consequences of the observability attack are provided in Section 5. Finally, a summary and conclusion are presented in Section 6.

2. Attacker controllability and attack observability

When considering the cybersecurity of a nuclear system two questions may be asked: When does an attacker have sufficient control to affect the state of the system? When an attacker masks measurements, how does this affect the ability to monitor the system and its hidden states? The answer to these questions is tied to the system theory concepts of controllability and observability.

The system under attack, G, is described by the model,

$$G:\left\{\begin{array}{ll}\dot{x} & = Ax+Bu\\ y & = Cx\end{array}\right. \sim \left[\begin{array}{c|c}A & B\\ \hline C & \end{array}\right] \tag{1}$$

where $x\in\mathbb{R}^n$ is the state variable, $u\in\mathbb{R}^q$ is the input to which the attacker has access, and $y\in\mathbb{R}^p$ is the measured output. The matrix $A\in\mathbb{R}^{n\times n}$ is the dynamics matrix. The matrix $B\in\mathbb{R}^{n\times q}$ is the input matrix, and $C\in\mathbb{R}^{p\times n}$ is the output matrix; these matrices describe how inputs enter the system and how measurements relate to the internal state variables. It should be noted that a system representation is not unique and can be transformed, using any non-singular matrix T, to another equivalent representation with the relationship, $x = Tz$, so that

$$\left[\begin{array}{c|c}A & B\\ \hline C & \end{array}\right] \sim \left[\begin{array}{c|c}T^{-1}AT & T^{-1}B\\ \hline CT & \end{array}\right] \tag{2}$$

In our scenario, we assume the attacker has access to possibly multiple control inputs that can steer the system to any desired state—it is controllable—and that in its nominal, uncompromised condition all of the internal states can be reconstructed—it is observable.

A system is controllable if it is possible to find some input, u, that can steer the state, x, to any desired value in finite time. Testing for controllability is straightforward:

Test 1. (Controllability) *A system with representation given in Eq. (1)is controllable if and only if the matrix*

$$\mathscr{C} = [B \quad AB \quad A^2B \quad \cdots \quad A^{n-1}B] \tag{3}$$

is full row rank.

Similarly, a system is observable if the state, x, can be determined from the observation of y in finite time. The test for observability is similar to the one for controllability:

Test 2. (Observability) *A system with representation given in Eq. (1)is observable if and only if the matrix*

$$\mathscr{O} = [C' \quad (CA)' \quad (CA^2)' \quad \cdots \quad (CA^{n-1})'] \tag{4}$$

is full row rank.

The tests for controllability and observability both depend on the rank of a matrix. Rank is defined as the number of linearly independent rows (columns) in a matrix. While this definition is theoretically pleasing, it is difficult in practical situations when numerical issues must be considered or the elements of the matrix have finite precision. A more workable test for rank compares the singular values of the matrix to some positive tolerance (Golub and Van Loan, 1996).

Test 3. (Matrix Rank) *The rank, r, of a matrix, X, can be determined from the singular values, σ_i, of X according to the inequalities*

$$\sigma_1 \geqslant \ldots\geqslant\sigma_r > \delta \geqslant \sigma_{r+1} \geqslant \ldots\geqslant\sigma_n \tag{5}$$

A matrix is full row rank if r is equal to the number of rows in the matrix.

The tolerance, δ, is defined to be consistent with the precision of the problem, ε. This precision would be the precision of the data in matrix X or machine precision for data with infinite accuracy. For a matrix X, the tolerance δ is defined by

$$\delta = \|X\|_\infty\varepsilon \tag{6}$$

where the matrix -norm is the maximum absolute row sum of the matrix,

$$\|X\|_\infty = \max_{1\leqslant i\leqslant n}\sum_{j=1}^{n}|x_{ij}|. \tag{7}$$

The primary advantage of this rank test is that computing singular values is straightforward and many implementations of the singular value decomposition exist.

What is important about these tests, for our discussion, is that the masking of signals amounts to the elimination of rows in C, which directly affects the rank test in Test 2. This result is not surprising—even using model knowledge of the system, eliminating measurements may make it impossible to reconstruct the state of the system. In the context of a cyber-attack, this could have huge implications because the attacker could create an unobservable system and then specifically attack the unobservable states. There would be no way, even using observers or virtual sensing, to determine the values of those states.

Knowing the structure of the state space provides insight to how the states map to the measurements. It also allows us to describe how an attacker might construct an attack. A Kalman decomposition is a representation of the system that makes clear the (un) controllable and (un) observable parts of the system. There exists a transformation, T, that transforms the system representation to the following Kalman decomposition:

$$\left[\begin{array}{c|c}\hat{A} & \hat{B}\\ \hline \hat{C} & \end{array}\right] = \left[\begin{array}{c|c}T^{-1}AT & T^{-1}B\\ \hline CT & \end{array}\right] = \left[\begin{array}{cc|c}A_1 & A_{12} & B_1\\ 0 & A_2 & B_2\\ \hline 0 & C_2 & \end{array}\right] \tag{8}$$