INL/CON-19-54494-Revision-1

A Covert System Identification Attack on Constant Setpoint Control Systems

Tyler Phillips, Hoda Mehrpouyan, John Gardner, Stephen J Reese

January 2020

The INL is a U.S. Department of Energy National Laboratory
operated by Battelle Energy Alliance

INL/CON-19-54494-Revision-1

A Covert System Identification Attack on Constant Setpoint Control Systems

Tyler Phillips, Hoda Mehrpouyan, John Gardner, Stephen J Reese

January 2020

Idaho National Laboratory
Idaho Falls, Idaho 83415

http://www.inl.gov

Prepared for the
U.S. Department of Energy
Office of Nuclear Energy
Under DOE Idaho Operations Office
Contract DE-AC07-05ID14517

A Covert System Identification Attack on Constant Setpoint Control Systems

Tyler Phillips Hoda Mehrpouyan
Computer Science Department
Boise State University
Boise, ID
tylerphillips1, hodamehrpouyan@boisestate.edu

John Gardner
Mechanical & Biomedical Engineering
Boise State University
Boise, ID
jgardner@boisestate.edu

Stephen J Reese
Idaho National Laboratory
Idaho Falls, ID
stephen.reese@inl.gov

***Abstract*—Industrial Control Systems (ICS) are the brain and backbone of nation's critical infrastructure such as nuclear power, water treatment, and petrochemical plants. In order to increase interoperability, real-time availability of data, and flexibility, information/communication technologies are adopted in this domain. While these information technologies have been effective, they are integrated into operational technologies without the necessary security defense. Designing an effective, layered security defense is not possible unless security threats are identified through a structural analysis of the ICS.**

For that reason, this paper provides an attacker's point of view on the reconnaissance effort necessary to gather details of the system dynamics - which are required for the development of sophisticated attacks. We present a reconnaissance approach which uses the system's I/O data to infer the dynamic model of the system. In this effort, we propose a novel cyber-attack which targets the controller proportional-integral-derivative gain values in a constant setpoint control system. Our findings will help researchers design more secure control systems.

***Index Terms*—cybersecurity, control systems, system identification, covert attack**

I. Introduction

Technology advancements and investments in smart manufacturing have resulted in the integration of digital instrumentation and computational control through communication networks. Smart manufacturing not only results in processes which are more responsive, precise, reliable, and efficient, they also provide better operational and management capabilities through factory and supply chain visibility [3]. Although, this transformation has many advantages, it has resulted in systems that are traditionally configured to operate in an air gap environment (i.e. a server cluster without access to the internet) to be exposed to new threats which originate in the cyber domain [18], [19], [22], [25], [29]. The perceived threat of a large impact cyber-attack on control systems proved to be a reality in 2010 with the launch of the Stuxnet worm [9], prompting plant owners, engineers, technicians, and researchers to feel the need to design and develop algorithms, tools, and techniques to protect the security of control systems. There are core features that separate the security of control systems from that of the traditional information technology (IT) domain. The fact that operational technologies (OT) and process control systems comprise proprietary hardware, software, and communication protocols, presents a new set of opportunities that require detection and protection techniques beyond what IT security can offer. Security technologies in the IT domain aim at protection of data and software by not allowing access from unauthorized users. The integration of IT security in control systems has lead to a false sense of security, as no amount of perimeter hardening can guarantee restriction of access by an attacker [10], [20], [27]. To address this issue, researchers have put forth efforts in physics-based detection methods to identify irregularities in the physics of the system [6].

In order to design and develop appropriate detection and protection techniques, researchers first turned their focus on constructing attack models [1], [15], [16], [23], [24], [26]. However, there is a lack of research on how the attackers are able to gain specific system knowledge that is required to carry out a successful attack. In most research studies, it is assumed that the reconnaissance efforts have been already carried out and the dynamics of the systems are known to the attackers. While attacks on cyber-physical systems (CPS) and industrial control systems (ICS) can have devastating impacts on human lives and the environment, it is not easy for attackers to inflict their desired effects on a targeted system. Krotofil and Larsen [8] outline five questions that an attacker should be able to answer to successfully complete the stages of an ICS kill chain[1]: (I) *Access*: How to utilize traditional IT network hacking, (II) *Discovery*: How to discover the system configuration and dynamics, (III) *Control*: What system parameters can be modified and in what degree these changes can be implemented so they are not detected, (IV) *Damage*: How can the attack scenario cause the greatest damage, (V) *Cleanup*: How to stay undetected after the attack is completed.

This paper provides a perspective from an attacker's point of view on the reconnaissance effort necessary to gather details of the system dynamics - which are required for the development of sophisticated attacks. Our findings will help researchers to design a more secure control system. We present a reconnaissance approach which is based on a data-driven

Identify applicable funding agency here. If none, delete this.

[1]https://www.sans.org/reading-room/whitepapers/ICS/industrial-control-system-cyber-kill-chain-36297

technique using the system's input/output (I/O) data to infer the dynamic model of the system. This process is known as system identification. We propose a novel cyber-attack which targets the controller Proportional-Integral-Derivative (PID) gain values in a constant setpoint control system. Accurately identifying the dynamic model of constant setpoint control systems is challenging, because there is little excitation of the system variables, i.e. the signal-to-noise ratio of the dynamic characteristics of the system are too low. Thus, the intent of our PID attack is to initiate excitation in the data so the dynamic characteristics of the system are present in the data, leading to more accurate system models. Additionally, we demonstrate the covertness of our attack in regards to physics-based detection algorithms.

The rest of this paper is organized as follows: Section II introduces related works and gives necessary background information. In Section III, we present our proposed system identification attack and analyze the accuracy and covertness of the attack. Finally, conclusions and possible directions of future work are covered in Section IV.

II. Background and Related Work

Numerous research studies have investigated security issues of control systems; however, research communities (i.e. control engineers and cybersecurity experts) often work independent of one another in the areas of *"cyber"* and *"physical"* and do not consider the overlap of the two domains. A good example of this is when in the control area mathematical models are constructed from the observed data to discover the dynamic models of the system. This process is known as *system identification* [11], which could be utilized by an attacker to learn about the system dynamics, and as a result, carry out more targeted attacks. However, not all system identification approaches used by the control community, such as an impulse-response, could be used by attackers for the discovery of the system dynamics, because it might raise an alarm by physics-based detection algorithms. In order to provide more details on the proposed approach, we first study control systems and the specific architecture that is the focus of this paper.

A control system is composed of four general components; the plant or physical system, sensors which measure the physical state of the plant, the controller which calculates control commands to send the actuators, and the actuators which make the physical changes to the plant. A continuous feedback-loop design, depicted in Fig. 1, is the general landscape used for continuous control of a system. Here, controlled variables such as pressure, temperature, or flow rate are measured using sensors, $y(t)$, and new control commands, $u(t)$, are sent to actuators based on the calculated error, $e(t)$, from their desired setpoint. In this work, we consider constant setpoint systems, i.e. the desired setpoint does not change in time. New control commands are calculated using the error between the setpoint and sensor measurements using the PID algorithm. The PID is the most commonly applied algorithm in practice today [28]. It calculates the control command sent to plant actuators using

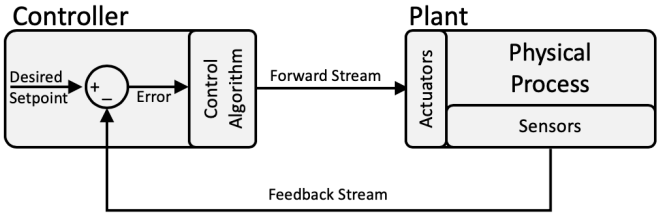

Fig. 1. Block diagram of a feedback-loop industrial control system. The controller calculates control commands using the setpoint and feedback data from the sensor measurements to control actuators.

three terms; proportion, integral, and derivative, hence the name. Mathematically this is given as

$$u(t) = K_p e(t) + K_i \int_0^t e(t)dt + K_d \frac{de(t)}{dt} \tag{1}$$

here, K_p, K_i, and K_d are the gain values of the proportional, integral, and derivative terms, respectively. In practice on a controller such as a programmable logic controller, a discrete form of the PID is used, given as

$$u_k = K_p e_k + K_i \sum_{n=1}^{k} e_n + K_d \left[e_k - e_{k-1}\right] \tag{2}$$

Discrete PID control is usually implemented using the so-called velocity form

$$u_k = u_{k-1} + K_p[e_k - e_{k-1}] + K_i e_k + K_d[e_k - 2e_{k-1} + e_{k-2}] \tag{3}$$

which is obtained by subtracting u_{k-1} from u_k. The obvious advantage of the velocity form is that there is no need to keep track of the sum of the errors.

In these types of controllers, it is possible to target and alter the PID gain values on the controller. This attack can influence the I/O data and result in a more accurate dynamic model of the physical system.

A. System Models and System Identification

System models are a representation of real-world phenomena where the essential aspects of a system are described by mathematical equations [21]. Historically, system modeling has been based on physical laws to derive system model. For example, mechanical systems follow Newton's and Hooke's laws, electrical systems follow Ohm's and Kirchoff's laws, and thermodynamics follow the ideal gas law and entropy.

As complexity of systems has increased, the models which describe their dynamics have become extremely complex as well. Researchers could rely on abstraction or simplification of the model, however, this could result in loss of information about physical phenomena that might be crucial for system discovery and analysis. In these cases, control engineers generally rely on system identification methods which construct the mathematical models using the systems I/O data. One of these mathematical models is the transfer function, which is the ratio of the output of a system to the input in the Laplace

domain. The mathematical formula for the transfer function H is given as

$$H(s) = \frac{N(s)}{D(s)} \tag{4}$$

Here, N and D are polynomials with unknown parameters in the frequency domain, (s).

To estimate the polynomial coefficients of the transfer function we apply the MATLAB [13] discrete-time transfer function estimation algorithm, *tfest*. This algorithm applies an estimated output-error polynomial model represented as

$$y(t) = \frac{B(q)}{F(q)}u(t - n_k) + e(t) \tag{5}$$

where $y(t)$ is the output, $u(t)$ is the input, n_k is the system delay, and $e(t)$ is the error. $B(q)$ and $F(q)$ are polynomials with respect to the backward shift operator, q^{-1}, and defined as follows

$$B(q) = b_1 + b_2 q^{-1} + \cdots + b_{nb} q^{-nb+1} \tag{6}$$

and

$$F(q) = 1 + f_1 q^{-1} + \cdots + f_{nf} q^{-nf} \tag{7}$$

In this algorithm, the polynomial coefficients are initialized using ARX, followed by nonlinear least squares search-based updates to minimize a weighted prediction error norm.

The objective for control engineers and attackers is to estimate the unknown parameters of the system model as accurately as possible. Based on this accuracy, control engineers can optimize system performance, whereas attackers can better design attacks which are more likely to remain covert and reach their goals.

B. Attack Scenarios for Control Systems

In order to compromise the control system, an attacker could affect its forward and feedback streams by attacking any of its components (i.e. controller, sensor, and actuator) or its communication system. Long et al. [12] and Farooqui et al. [5] provide examples of such attack models. In [12], the communication network of the control system is arbitrarily flooded, causing jitter and packet loss in the communication links. Whereas [5] uses false signals that are randomly generated and transmitted to the controller and actuator to impact the overall system. In these tactics, the system may become unstable leading to unpredictable behavior which is easier to identify using physics-based detection.

To this point, Teixeira et al. [26] investigated the attack models demonstrated in [1], [4], [23] and concluded that the design of a successful covert attack requires a high level of knowledge about the dynamics of the system. For example, the design and development of the man-in-the-middle attacks carried out in [23], [24] was based on the assumption that the dynamics of the control systems are known to the attack model. Based on that, the proper values were computed and injected into the feedback stream to remain covert. Hence, these attacks require an inside knowledge of the dynamics of the system and is limited to, and dependent on, inside attackers.

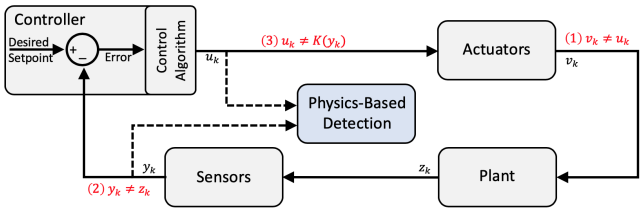

Fig. 2. General block diagram of an ICS continuous feedback-loop indicating where cyber-attacks can target and compromise the system: (1) actuators, (2) sensors, or (3) controller. Here, altered data from the attack is highlighted in red which the physics-based detection intends to detect.

To overcome this limitation, de Sá et al. [2], [4] designed and developed cyber-attack techniques known as *cyber-physical intelligence attacks* to acquire the system knowledge necessary to model covert and controlled attacks. In their earlier work, de Sá et al. [4] carried out a passive system identification attack and eavesdropped on the forward and feedback data streams to estimate the system model's *transfer function*. However, since the effectiveness of the passive attack depends on the occurrence of events or excitation of the system variables, the authors introduced an active system identification attack [2]. In the second attempt, they tailor signals to insert into the communication channel and observe the resulting response. While, the active system identification resulted in a faster discovery of system dynamics, there is a higher probability of getting detected by an anomaly detection algorithm.

In this paper, we further investigate an active system identification attack by altering the control command PID gain calculation. We will demonstrate that while the attack is still effective, it is much more difficult to detect our proposed approach. In order to prove the covertness of the proposed attack, we need to be able to pass the intrusion detection systems (IDS) used by these type of control systems. In the next section, we will discuss the different types of attacks that occur in control systems and the IDS approaches that are utilized to detect these attacks.

C. Physics-Based Anomaly Detection

The types of attacks that occur in a control system are depicted in Fig. 2 and summarized as follows:

1) When an actuator or forward stream is compromised, the actuation, v_k, to the plant is different than the intended action by the controller, $v_k \neq u_k$. This false actuation will in turn affect the measured variables of the plant.
2) When a sensor or feedback stream is compromised, the controller logic will accept incorrect input which is different than the real state of the plant, $y_k \neq z_k$.
3) When the controller is compromised, it will generate a control command that does not satisfy the intended logic of the controller, $u_k \neq K(y_k)$, where K is the control logic and a function of the sensor measurements, y_k.

In order to detect the above attacks, hardware- and/or software-based intrusion detection systems are designed and developed to monitor network and system activities to detect malicious acts [7]. An attack's ability to elude detection by the IDS determines its covertness. Covertness can be analyzed in the traditional IT domain as well as the physical domain; in this work we are interested in the latter.

Physics-based detection focuses on the problem of using real-time measurements to detect attacks. Two popular methods are anomaly-based and safety limit detection. Anomaly-based detection relies on the fact that physical processes must follow immutable laws of physics. In general, detection is done through the use of mathematical models of the system to predict the expected measurement, $\hat{y}_k$, using the current control commands, u_k, and previous sensor measurement, y_{k-1}.

The anomaly detection test itself uses a time series of residual values, r_k. The residual is the difference between the measured and predicted values, given as

$$r_k = |y_k - \hat{y}_k| \tag{8}$$

The residuals are then used in either a stateless or stateful anomaly test. A stateless test raises an alarm every time a residual value reaches a threshold value, $r_k \geq \tau$, shown by Fig. 3. In a stateful test the historical changes of the residual are kept as an additional statistic denoted as S_k, to generate an alert if $S_k \geq \tau$. There are many ways to keep track of the residual for a stateful test, such as taking an average over a time-window, an exponential weighted moving average, or using change detection statistics such as the non-parametric cumulative sum statistic.

On the other hand, safety limit detection is based on the normal operating range of the system variables. In this case an alarm is raised if the sensor measurement, y_k, exceeds lower or upper limits, given as

$$y_k < y_{k_{\text{min}}} \tag{9}$$

and

$$y_k > y_{k_{\text{max}}} \tag{10}$$

From the standpoint of a cyber-attacker the measure of covertness in regards to physics-based detection is important. Remaining covert is often necessary in order to be successful in reaching their attack goals.

III. Proposed System Identification Attack

In this section, we present the proposed covert active system identification attack approach along with the running example of an inverted pendulum. We begin by deriving the *transfer function* of the inverted pendulum which mathematically describes the behavior of the system. The transfer function is then used to perform simulations which model the behavior of the system under normal and attack scenarios. We then analyze the effectiveness and covertness of attacks which alter the derivative gain value.

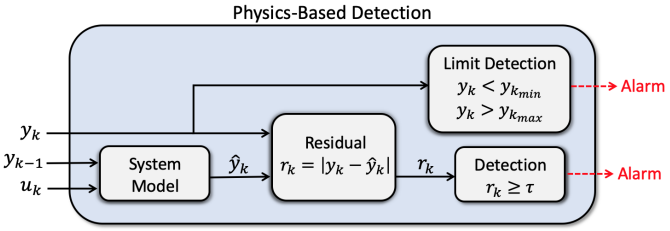

Fig. 3. Physics-based detection where the residual between the measured value, y_k, and model prediction, $\hat{y}_k$, is used for an alarm if exceeding a given threshold, τ.

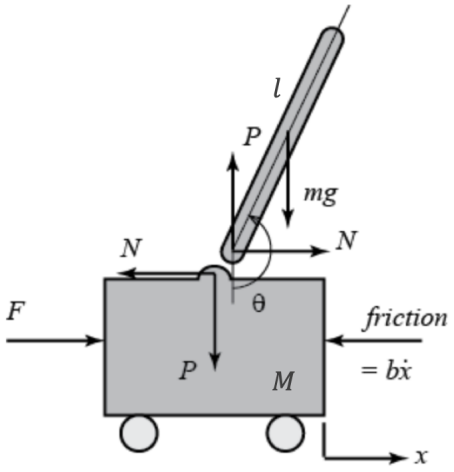

Fig. 4. Schematic showing the forces acting on an inverted pendulum attached to a cart. Image adapted from Messner and Tilbury [17].

A. Inverted Pendulum as a Target System

In this example, the control system objective is to keep the pendulum at the vertical position, i.e. a constant setpoint. To accomplish this objective, a PID controller is used to apply an input force to the cart on which the inverted pendulum is mounted. The input and system forces are shown in Fig. 4. Using the system forces, the transfer function can be derived. We present a partial derivation, for the full derivation the reader is referred to Messner and Tilbury [17]. First, the horizontal forces acting on the cart lead to the following

$$M\ddot{x} + b\dot{x} + N = F \tag{11}$$

and summing the forces on the pendulum results in

$$N = m\ddot{x} + ml\ddot{\theta}\cos\theta - ml\dot{\theta}^2\sin\theta \tag{12}$$

From here, substitution gives the first governing equation

$$(M+m)\ddot{x} + b\dot{x} + ml\ddot{\theta}\cos\theta - ml\dot{\theta}^2\sin\theta = F \tag{13}$$

We get the second governing equation by summing the forces perpendicular to the pendulum at the axis, giving

$$(I + ml^2)\ddot{\theta} + mgl\sin\theta = -ml\ddot{x}\cos\theta \tag{14}$$

where I is the moment of inertia of the pendulum. To linearize the governing equations we assume that the pendulum only has

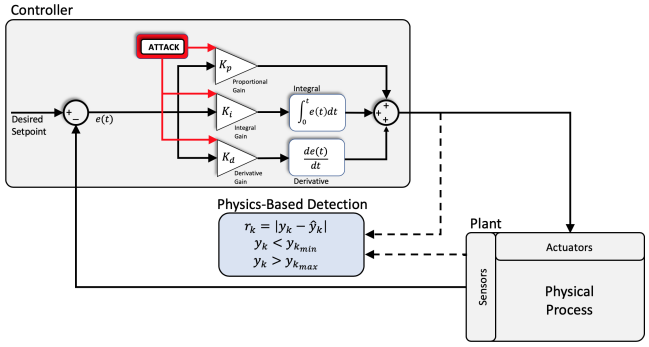

Fig. 5. Block diagram of a continuous feedback-loop control system showing where our system identification attack targets the Proportional-integral-derivative (PID) gain calculation.

small deviations from the vertical position and use the small angle approximation. This lead to a set of linearized governing equations

$$(I + ml^2)\ddot{\phi} - mgl\phi = ml\ddot{x} \tag{15}$$

and

$$(M + m)\ddot{x} + b\dot{x} - ml\ddot{\phi} = F \tag{16}$$

To obtain the transfer function of the linearized governing equations, we first take the Laplace transform and assume zero initial conditions. The resulting Laplace transforms are given as

$$(I + ml^2)\Phi(s)s^2 - mgl\Phi(s) = mlX(s)s^2 \tag{17}$$

and

$$(M + m)X(s)s^2 + bX(s)s - ml\Phi(s)s^2 = U(s) \tag{18}$$

In this study, we are concerned with the output of the angle, $\Phi(s)$, and its relation to the force input, $U(s)$. We eliminate $X(s)$ from (17) and (18) by solving for $X(s)$ and then using substitution. The transfer function of the pendulum angle becomes

$$P_{\text{pend}}(s) = \frac{\frac{ml}{q}s^2}{s^3 + \frac{b(I+ml^2)}{q}s^2 - \frac{(M+m)mgl}{q}s - \frac{bmgl}{q}} \tag{19}$$

where

$$q = [(M + m)(I + ml^2) - (ml)^2] \tag{20}$$

The linearized transfer function of the inverted pendulum is used for the simulations carried out in this work.

B. Active System Identification Attack

The goal of a system identification attack is to increase the accuracy of the transfer function estimation - which represents the dynamics of the system. The accuracy of system identification algorithms increases for data types that have high signal-to-noise ratios for the dynamic characteristics of the system, i.e. variable excitation. Thus, in order to force the excitation in the system, we employ a novel attack which briefly targets the PID gain values, K_p, K_i, and K_d, as depicted by Fig. 5.

The performance of the attack is evaluated through a set of simulations performed in Simulink [14]. Simulink is a graphical programming environment for modeling, simulating, and analyzing multi-domain dynamic systems. We utilize its environment to compute the control command, u_k, using the angle measurement, y_k, in a simulated environment, shown in Fig. 6.

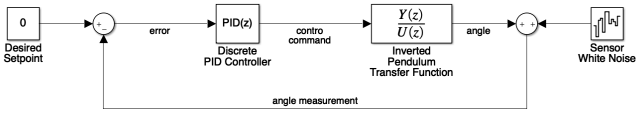

Fig. 6. Simulink block diagram of an inverted pendulum transfer function controlled using a discrete PID controller.

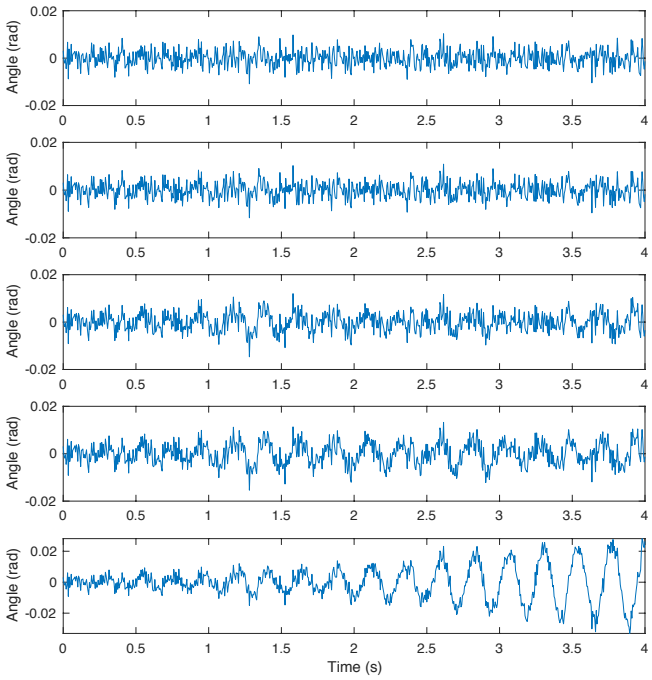

Fig. 7. Results of the inverted pendulum angle under different derivative attack scenarios. From top to bottom the derivative gain is 5 (no attack), 3.2, 1.05, 0.67, and 0.27, respectively.

In this paper, we run 100 trial simulations under normal and different attack scenarios. The normal operation is based on "tuned" PID gain values of 162, 124, and 5, respectively. The different attack scenarios in this work reduces the derivative gain in intervals of 20%, i.e. the derivative gain value is 5, 4, 3.2, 2.56, etc. In order to emulate a real-world control system, we apply a discrete PID controller running at 200 Hertz, add white noise to the sensor measurements, and run the simulation for 4 seconds of real-time. The resulting angle measurements under different derivative values is shown in Fig. 7. Hence, it is clear that the derivative gain attack can force excitation of the system variable, forcing the pendulum to oscillate about its setpoint.

To evaluate the accuracy of the estimated transfer functions we calculate the normalized root mean square error (NRMSE) measure of the goodness of the fit. The NRMSE is a fitness

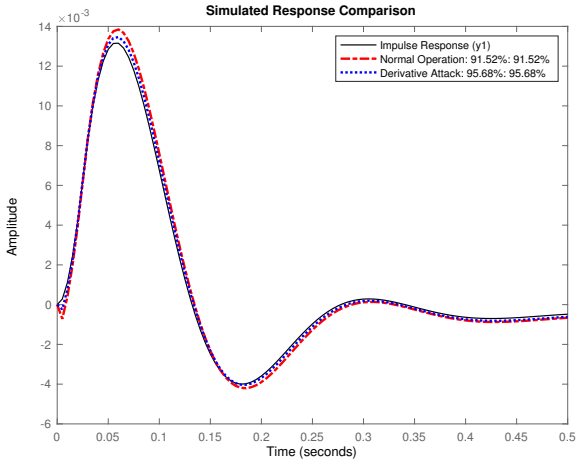

Fig. 8. Normalized root mean square error of an estimated transfer function. Here the derivative of the attack estimation has been reduced from 5 (no attack) to 1.05.

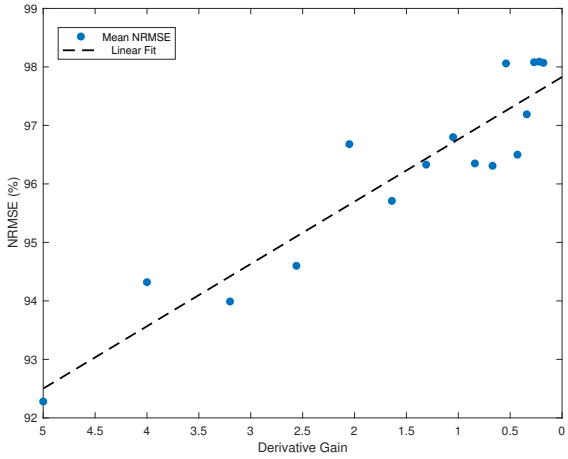

Fig. 9. Normalized root mean square error results under different derivative attack scenarios. Each data point represent the mean of 100 simulations.

value indicator of how well an estimated model matches validation data, given mathematically as

$$\text{NRMSE} = 100\left(1 - \frac{\| y - \hat{y} \|}{\| y - \text{mean}(y) \|}\right) \tag{21}$$

where y is the validation data and $\hat{y}$ is the estimation model. The validation data used in this analysis is generated using an impulse response simulation using the linearized transfer function given in (19). To make a direct comparison, the sensor noise is removed from these simulations. It can be seen in Fig. 8 that the estimated transfer function has an increased NRMSE when we employ our attack and the derivative term is decreased. The NRMSE mean value under the different attack scenarios is presented in Fig. 9. It is demonstrated that the transfer function estimation in general increases as the derivative term is reduced and excitation of the angle measurement is increased. However, we must also consider the covertness of these attacks, which is analyzed in the next section.

C. Covertness to Physics-based Detection

The measure of covertness in regards to physics-based detection is analyzed in both anomaly and limits detection. In an anomaly detection test the time-series residual values are calculated based on control commands that are sent to the actuator and the resulting angle measurement based on the previous measurement (Fig. 3). In our attack, we do not inject false data into the control loop, i.e. the actuator and anomaly detection algorithm receive the same control command. Therefore, the residual values are calculated to be the sensor noise when system disturbances are not present. Thus, we argue that our attack is covert to anomaly based detection statistics.

On the other hand, the attack effects the angle measurements and limit based detection can potentially identify the attack. Therefore, the covertness depends on the amount of excitation we force during the attack and the allowable limits that are set. To determine the limits in this study, we assume the angle data follows a Gaussian or normal distribution and select a limit which would give a false positive alarm once per year. This is calculated using the approximate expected frequency equation where the frequency of occurrence is 1 in

$$\frac{1}{1 - \text{erf}\left(\frac{k\sigma}{\sqrt{2}}\right)} \tag{22}$$

Here, erf is the error function and k is the number of standard deviations, σ. Using the frequency of our controller we get $k \sim 6.4$. Since the angle mean is 0 and our limits are symmetric, the detection limit for an alarm is calculated as

$$|y_k| > 0.022 \tag{23}$$

Therefore, we infer the covertness of the attack with a comparison of the absolute maximum deviation from the setpoint under each derivative attack scenario. It can be seen in Fig. 10 that the absolute maximum deviation from the setpoint slowly increases until the derivative is reduced to below a value of ~ 0.6, where the system becomes unstable. We argue that the attack would likely remain covert until the unstable region is reached.

IV. Conclusion

At present time, we propose an attack which targets the PID gain values of a controller. The intent of the attack is to force excitation in system parameters in order to increase the accuracy of data-driven system identification in constant setpoint control systems. The effectiveness of the attack is analyzed with the use of simulations, and we demonstrate that the estimated system model's accuracy increases as we reduce the derivative gain value. Additionally, the PID attack is covert in regards to physics-based anomaly detection by virtue of not injecting false data into the system. However, physics-based limits detection can potentially detect our attack if the altered PID gains force too much excitation into the system.

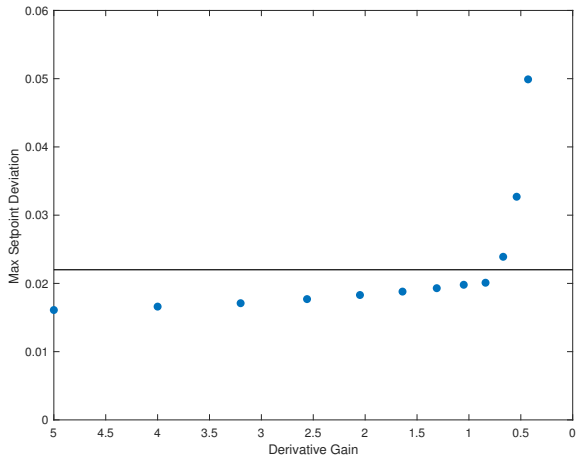

Fig. 10. Absolute maximum deviation of the angle measurement from the setpoint under normal (Derivative = 5) and different derivative attack scenarios. Here the horizontal line represents the maximum value of once a year deviation under normal operation.

Currently, we manually change the derivative gain and check the results of the estimated model. Since the actual system model is unknown from an attackers standpoint, our future work includes the implementation of an algorithm which watches the parameter deviations from the setpoint and alters the PID gains in order to maximize excitation while staying covert to limits detection. Additionally, we plan to investigate the effectiveness of our PID attack in a real-world simulated environment such as the Tennessee Eastman Process. However, we foremost encourage the development of new identification techniques in order to identify attacks of this nature.

ACKNOWLEDGMENT

This work was supported by National Science Foundation Computer and Information Science and Engineering (CISE), award number 1846493 of the Secure and Trustworthy Cyberspace (SaTC) program: Formal TOols foR SafEty aNd Security of Industrial Control Systems (FORENSICS).

REFERENCES

[1] S. Amin, X. Litrico, S. Sastry, and A. M. Bayen. Cyber security of water scada systemspart i: Analysis and experimentation of stealthy deception attacks. *IEEE Transactions on Control Systems Technology*, 21(5):1963–1970, Sep. 2013.

[2] Alan Oliveira de Sá, Luiz F. R. da C. Carmo, and Raphael C. S. Machado. Bio-inspired active system identification: a cyber-physical intelligence attack in networked control systems. *Mobile Networks and Applications*, Oct 2017.

[3] Alan Oliveira de Sá, Luiz Fernando Rust da Costa Carmo, and Raphael Machado. A controller design for mitigation of passive system identification attacks in networked control systems. *Journal of Internet Services and Applications*, 9:1–19, 2017.

[4] Alan Oliveira de Sá, Luiz Fernando Rust da Costa Carmo, and Raphael Machado. Covert attacks in cyber-physical control systems. *IEEE Transactions on Industrial Informatics*, 13:1641–1651, 2017.

[5] A. A. Farooqui, S. S. H. Zaidi, A. Y. Memon, and S. Qazi. Cyber security backdrop: A scada testbed. In *2014 IEEE Computers, Communications and IT Applications Conference*, pages 98–103, Oct 2014.

[6] Jairo Giraldo, David Urbina, Alvaro Cardenas, Junia Valente, Mustafa Faisal, Justin Ruths, Nils Ole Tippenhauer, Henrik Sandberg, and Richard Candell. A survey of physics-based attack detection in cyber-physical systems. *ACM Comput. Surv.*, 51(4):76:1–76:36, July 2018.

[7] Shijoe Jose, D. Malathi, Bharath Reddy, and Dorathi Jayaseeli. A survey on anomaly based host intrusion detection system. *Journal of Physics: Conference Series*, 1000:012049, apr 2018.

[8] Marina Krotofil and Jason Larsen. Rocking the pocket book: Hacking chemical plants for competition and extortion. Report P-41, DEFCON, 2015.

[9] Ralph Langner. Stuxnet: Dissecting a cyberwarfare weapon. *IEEE Security and Privacy*, 9(3):49–51, May 2011.

[10] L. Lerner. *Trustworthy Embedded Computing for Cyber-Physical Control*. PhD thesis, Virginia Tech, 2015.

[11] Lennart Ljung. System identification. *Wiley Encyclopedia of Electrical and Electronics Engineering*, 2001.

[12] M. Long, Chwan-Hwa Wu, and J. Y. Hung. Denial of service attacks on network-based control systems: impact and mitigation. *IEEE Transactions on Industrial Informatics*, 1(2):85–96, May 2005.

[13] The Mathworks, Inc., Natick, Massachusetts. *MATLAB version 9.5 (R2018b)*, 2018.

[14] The Mathworks, Inc., Natick, Massachusetts. *Simulink version 9.2 (R2018b)*, 2018.

[15] Hoda Mehrpouyan, Dimitra Giannakopoulou, Irem Y Tumer, Chris Hoyle, and Guillaume Brat. Combination of compositional verification and model checking for safety assessment of complex engineered systems. In *ASME 2014 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. American Society of Mechanical Engineers Digital Collection, 2015.

[16] Hoda Mehrpouyan, Irem Y Tumer, Chris Hoyle, Dimitra Giannakopoulou, and Guillaume Brat. Formal verification of complex systems based on sysml functional requirements. Technical report, Columbus State University Columbus United States, 2014.

[17] Bill Messner and Dawn Tilbury. Control tutorials for matlab and simulink (CTMS): Inverted pendulum system modeling. http://ctms.engin.umich.edu/CTMS/index.php?aux=Home.

[18] N. Nicolaou, D. G. Eliades, C. Panayiotou, and M. M. Polycarpou. Reducing vulnerability to cyber-physical attacks in water distribution networks. In *2018 International Workshop on Cyber-physical Systems for Smart Water Networks (CySWater)*, pages 16–19, April 2018.

[19] Luciana Obregon. Secure Architecture for Industrial Control Systems. Technical report, SANS Institute, 2015.

[20] US Department of Homeland Security. Recommended practice: Improving industrial control system cybersecurity with defense-in-depth strategies: Industrial control systems. Technical report, Homeland Security Cyber Emergency Response Team, 2016.

[21] Babatunde A. Ogunnaike and W. Harmond Ray. *Process Dynamics, Modeling, and Control*. Oxford University Press, New York, 1994.

[22] Farhad Rasapour and Hoda Mehrpouyan. Misusing sensory channel to attack industrial control systems. In *Proceedings of the Eighth ACM Conference on Data and Application Security and Privacy*, pages 158–160. ACM, 2018.

[23] R. S. Smith. Covert misappropriation of networked control systems: Presenting a feedback structure. *IEEE Control Systems Magazine*, 35(1):82–92, Feb 2015.

[24] Roy S. Smith. A decoupled feedback structure for covertly appropriating networked control systems. *IFAC Proceedings Volumes*, 44(1):90 – 95, 2011. 18th IFAC World Congress.

[25] Logan D. Sturm, Christopher B. Williams, Jamie A. Camelio, Jules White, and Robert Parker. Cyber-physical vulnerabilities in additive manufacturing systems: A case study attack on the .stl file with human subjects. *Journal of Manufacturing Systems*, 44:154 – 164, 2017.

[26] André Teixeira, Iman Shames, Henrik Sandberg, and Karl Henrik Johansson. A secure control framework for resource-limited adversaries. *Automatica*, 51(C):135–148, January 2015.

[27] US Department of Energy. 21 steps to improve cyber SCADA security. Technical report, DOE, 2005.

[28] Antonio Visioli. *Practical PID Control*. Springer. Advances in Industrial Control, London, 2006.

[29] Yinan Wang, Gangfeng Yan, and Ronghao Zheng. Vulnerability assessment of electrical cyber-physical systems against cyber attacks. *Applied Sciences*, 8:768, 05 2018.