System Identification Attacks, Model-based Offensives and Countermeasures in Networked Control Systems*

Alan Oliveira de Sá[1,2] (author), Luiz F. R. da C. Carmo[1,3](advisor), Raphael C. S. Machado[3,4](co-advisor)

[1]Post-graduate Program in Informatics, Federal University of Rio de Janeiro - RJ - Brazil

[2]Admiral Wandenkolk Instruction Center, Brazilian Navy – Rio de Janeiro, RJ – Brazil

[3]National Institute of Metrology, Quality and Technology (Inmetro) – Duque de Caxias, RJ – Brazil

[4]Fluminense Federal University – Niterói, RJ – Brazil

alan.oliveira.sa@gmail.com, {lfrust,rcmachado}@inmetro.gov.br

Abstract. *Networked Control Systems (NCS) are widely used in Industry 4.0 and to control critical infrastructures. However, at the same time that they provide several advantages, they are prone to cyberattacks. This work investigates new classes of threats in NCSs (System Identification attacks and covert/model-based offensives), and proposes novel countermeasures to mitigate them. The results indicate that the study on the new attacks introduced in this work and the countermeasures herein proposed contribute to the cybersecurity of NCSs.*

Resumo. *Sistemas de Controle em Rede, ou Networked Control Systems (NCS), são amplamente utilizados na Indústria 4.0 e no controle de infraestruturas críticas. No entanto, ao mesmo tempo em que oferecem diversas vantagens, os NCSs são propensos a ataques cibernéticos. Este trabalho investiga novas classes de ameaças em NCSs (ataques de Identificação de Sistemas e ofensivas furtivas/baseadas em modelos) e propõe novas contramedidas para mitigá-las. Os resultados indicam que o estudo dos novos ataques apresentados neste trabalho e as contramedidas aqui propostas contribuem para a segurança cibernética dos NCSs.*

1. Introduction

The integration of systems used to control physical processes via communication networks aims to provide better operational and management capabilities, as well as reduce costs. Motivated by these advantages, there is a trend to have an increasing number of industrial processes and critical infrastructures driven by Networked Control Systems (NCS) – which, not by chance, are widely used in Industry 4.0. As detailed in Figure 1, an NCS consists of a controller, which runs a control function $C(z)$, a physical plant, described by its transfer function $G(z)$, and a communication network that interconnects both devices through a forward stream and a feedback stream.

At the same time it brings several advantages, the integration of controllers and physical plants through communication networks also exposes such control systems to new threats, typical of the cyber domain. In fact, the literature [McLaughlin et al. 2016]

*The complete thesis is available at:
https://drive.google.com/file/d/12DXmrS-51CqhWFHjhAnuyJHQ7GWPZP-6/view?usp=sharing

reports the execution of real cyber attacks against physical plants since 1982, affecting a wide variety of targets, such as a diesel generator, a gas pipeline, a steel plant, and even a uranium enrichment plant (which occurred in the well known case of the Stuxnet worm).

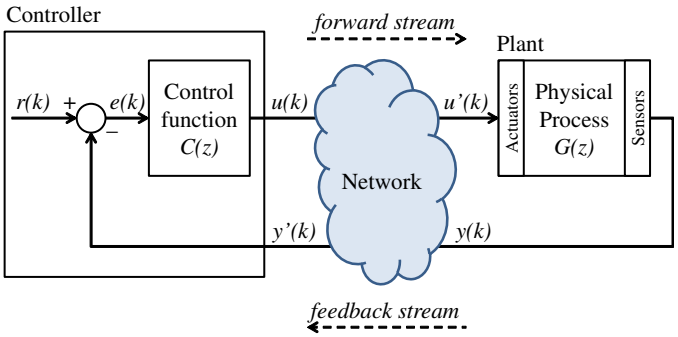


Figure 1. Networked Control Systems (NCS) [de Sá et al. 2017c].

One possible way for an attacker to impair an NCS is by interfering on its communication process (by inducing jitter, causing data loss, or injecting false data in the NCS links). Note that, although some new industrial communication protocols were developed including security features, there are protocols widely used in industry that still lack security mechanisms [Collantes and Padilla 2015] – such as the Profinet, MODBUS/TCP, and Ethernet/IP. The main issue of these protocols is the lack of encryption and authentication, which makes a plethora of plants vulnerable to cyber threats. Moreover, even when the NCS uses secure protocols, the possibility of an attacker overcoming security mechanisms must be considered (through social engineering attacks, for instance). Therefore, given the feasibility of occurring cyber attacks against physical systems, as evidenced by the literature, studies have been conduced aiming to characterize vulnerabilities and propose security solutions for NCSs. In this context, this thesis investigates a class of covert/model-based attacks that use not evident nor trivial mechanisms to lead the plant to harmful conditions when the NCS communication is prone malicious interferences. This work reveals novel techniques that can be used to build such kind of attacks (for the sake of awareness), and proposes countermeasures to mitigate them. In summary, the main contributions of this work are listed below:

1. It proposes two System Identification attacks (the Passive and the Active System Identification attacks), which are studied and evaluated as attack tools to learn the NCS models and, thus, support the design of covert/model-based offensives;
2. It proposes a novel model-based attack that operates by causing controlled data loss in the NCS links;
3. It evaluates the joint operation of System Identification attacks with three model-based offensives: the novel controlled data loss attack; and two other attacks that operate by injecting false data into the system. The effectiveness and accuracy of the attacks are evaluated considering an example of a common industrial device – a DC motor – and a nuclear critical infrastructure – a large Pressurized Heavy Water Reactor (PHWR);
4. It introduces a taxonomy to support the discussion regarding the relationship between System Identification attacks and covert/model-based attacks in NCSs;
5. It proposes two countermeasures that contribute to the security of NCSs in case of failure or absence of other conventional security mechanisms – such as encryption, authentication, and network segmentation.

This research resulted in the publication of papers in five international journals [de Sá et al. 2017c, Ferrari et al. 2020, de Sá et al. 2020, de Sá et al. 2017b, de Sá et al. 2018b], six conference papers [de Sá et al. 2016, de Sá et al. 2017a, de Sá et al. 2017d, de Sá et al. 2018a, de Sá et al. 2019b, de Sá et al. 2019a], an award, and an international research partnership, as detailed in the remainder of this paper. It is worth emphasizing that the applications of NCSs can range from cooperative control of vehicles using mobile networks to wired NCS intended to control devices in Industry 4.0, water canal systems or even large Pressurized Heavy Water Reactors (PHWR). It includes a vast number of potential – sometimes critical – targets that can suffer from the attacks herein studied, as well as benefit from the countermeasures herein proposed.

2. Motivation

As discussed in Section 1 and reported in the literature, the network communication in NCSs makes the controlled physical processes prone to be impaired by cyberattacks. In addition to the exploitation of vulnerabilities present in real NCSs [Collantes and Padilla 2015], studies indicate that advanced techniques can be used to produce covert and effective attacks in such systems [Smith 2011, Smith 2015]. Due to the wide application of NCSs in industries and critical infrastructures, cyberattacks against these systems may result in significant impacts to the society. So, this work is motivated by the the need to better understand the advanced attack techniques that can be perpetrated against NCSs and propose countermeasures to protect these critical systems.

3. Objectives

The first objective (O1) of this work is to investigate disclosure attacks – particularly system identification attacks – as a tool to gather and learn information from the plant and control algorithms, as well as the role of these attacks in the design of covert/model-based offensives against NCSs. The second objective (O2) is to study new covert/model-based attacks in NCSs. With the lessons learned from the first two objectives, the third objective (O3) is to develop countermeasures to mitigate system identification attacks and model-based offensives in NCSs, while ensuring adequate levels of plant control.

4. Results

This section briefly describes the results obtained for each objective aimed in this research. The objectives O1, O2, and O3 are addressed in Sections 4.1, 4.2 and 4.3, respectively.

4.1. System Identification attacks

From the point of view of control theory, the literature indicates that covert/model-based attacks can be planned based on an accurate knowledge about the NCS models [Teixeira et al. 2015, Smith 2015]. However, despite the importance of model knowledge for this set of covert/model-based offensives, the literature does not explore attacks intended to reveal/learn the NCS models.

To fill this gap, in this work, two System Identification attacks are proposed, studied and evaluated as an attack tool to support the design of covert/model-based attacks. As demonstrated in this work, the system identification – *i.e.* the action of building mathematical models of dynamic systems – can be considered a key step for the execution of covert/model-based attacks against NCSs. The attacks herein proposed are: the Passive

System Identification attack [de Sá et al. 2017c]; and the Active System Identification attack [de Sá et al. 2017a, de Sá et al. 2017b]. Both attacks are designed based on bio-inspired metaheuristics (the Backtracking Search Optimization and the Particle Swarm Optimization), and assessed in scenarios with and without data loss and noise.

To learn the NCS models, the Passive System Identification attack [de Sá et al. 2017c] analyzes signals that typically flow through the NCS links, as a result of its normal operation. The attacker first eavesdrops the input and output signals of the attacked NCS device (a controller or a plant) and, then, runs a system identification algorithm to learn the device's model – whose coefficients are unknown. As explained in [de Sá et al. 2017c], it is possible to establish an analogy between the Passive System Identification attack and the Known Plaintext cryptanalytic attack. In such analogy, the captured input and output signals of the attacked device correspond to the plaintext and ciphertext, respectively, the generic model (a generic transfer function) of the device corresponds to the encryption algorithm, and the coefficients of the actual model (intended to be found) correspond to the secret key.

The Active System Identification attack [de Sá et al. 2017a, de Sá et al. 2017b], in turn, is an alternative to the Passive System Identification attack when the attacker cannot wait for the occurrence of an event that produces the signals needed for the identification process. In this case, the attacker injects an attack signal $a(k)$ into the system, in order to estimate the NCS models based on the system response to such signal. Similarly, as explained in [de Sá et al. 2017b], an analogy can be established between the Active System Identification attack and the Chosen Plaintext cryptanalytic attack, wherein the injected signal $a(k)$ corresponds to the chosen plaintext, the response of the system to $a(k)$ represents the ciphertext, the generic model of the attacked system corresponds to the encryption algorithm, and the coefficients of the actual model (intended to be found) correspond to the secret key.

To support the discussion regarding the relationship between System Identification attacks and covert/model-based attacks, this work introduces a taxonomy that formalizes of a number of concepts related to covertness and intelligence in the context of NCS security [de Sá et al. 2017b, de Sá et al. 2017c]. This taxonomy also sets the requirements for the attacks discussed in this work, which helps on the development of layered defense strategies against System Identification attacks and covert/model-based offensives. Moreover, regarding covert attacks in NCS, this taxonomy also dismembers the concept of covertness in two different domains: the cyber domain; and the physical domain.

4.2. Model-based Offensives

This thesis evaluates the effectiveness and accuracy of the joint operation of System Identification attacks and model-based offensives against NCSs. Three model-based offensives are addressed:

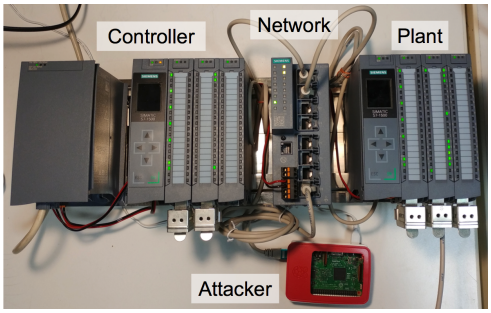
- The novel Controlled Data Loss attack [Ferrari et al. 2020] proposed in this thesis;
- The Controlled Data Injection attack [de Sá et al. 2017c] characterized in this research;
- The Covert Misappropriation attack proposed in [Smith 2015];

The attacks addressed in this thesis are evaluated through simulations considering a common industrial device (a DC motor) [de Sá et al. 2016, de Sá et al. 2017c,

Ferrari et al. 2020], as well as a nuclear critical infrastructure (a Pressurized Heavy Water Reactor) [de Sá et al. 2018a].

In the Controlled Data Injection and Covert Misappropriation attacks, the attacker acts as a Man-in-the-Middle (MitM) to inject false data in the NCS links. The injected data is computed based on the models previously learned through a System Identification attack in order to produce accurate and harmful effects in the plant. The injected false data is also computed to make the attacks covert: physically covert in the case of the Controlled Data Injection attack [de Sá et al. 2017c]; and cybernetically covert (from controller perspective) in the case of the Covert Misappropriation attack [Smith 2015, de Sá et al. 2018a].

The novel Controlled Data Loss attack, in turn, uses the models learned through a System Identification attack to smartly decide which packets the NCS must lose – due to an attacker interference – to produce harmful effects on the plant. The attack uses a bio-inspired metaheuristic to optimize its activity, ensuring accuracy and (at the same time) making the data loss limited to a reduced number of packets. The proposed approach avoids the indiscriminate loss of packets (which could facilitate the disclosure of the attack). It is noteworthy that this attack strategy, formalized in Sections 3.2.3 and 5.2 of the thesis, had its feasibility demonstrated in the Live Demo Track[1] of the 2019 IEEE International Workshop on Metrology for Industry 4.0 and IoT, and is published in [Ferrari et al. 2020][2]. In both cases (paper and demo), the attack was carried out in the real setup shown in Figure 2, composed by commercially available industrial hardware.



Controller:
Siemens SIMATIC S7-1500 PLC
Plant:
Siemens SIMATIC S7-1500 PLC
Network equipment:
Industrial switch Siemens Scalance XB208
Network Protocol:
PROFINET
Attacker device:
PC Engines APU2C4 board

Figure 2. Real setup used to demonstrate the Controlled Data Loss attack.

4.3. Countermeasures

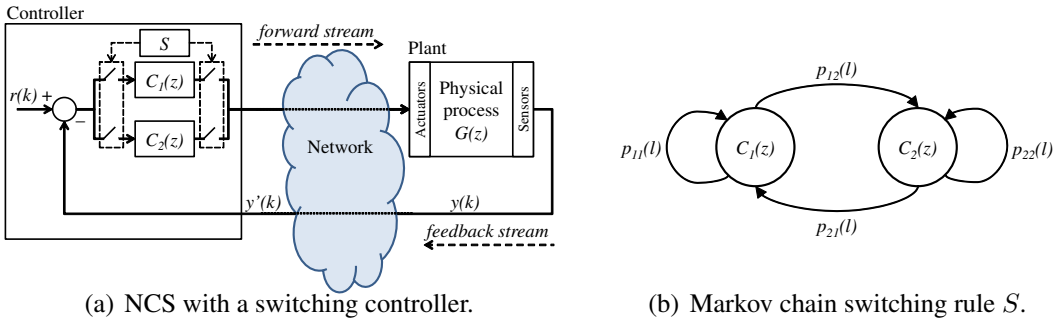
The analysis of system identification processes as feasible attacks led to the development of a countermeasure intended to inhibit the identification task, in case of failure of other conventional security mechanisms – such as encryption, network segmentation and firewall policies. In this sense, another contribution of this work is the proposal of a switching controller design [de Sá et al. 2017d, de Sá et al. 2018b], shown in Figure 3, to hinder the System Identification attacks proposed in this work – and, therefore, dissuade covert/model-based attacks. The switching controller, shown in Figure 3(a), consists of a set of control functions $C_i(z)$, $i \in \mathcal{I} = \{1, ..., N\}$, that are switched among N states (here, for the sake of simplicity, $N = 2$) by a switching rule S, to perform the control of a plant $P(z)$. The switching rule S is described as a Markov chain, shown in Figure 3(b),

[1] A video of the demonstration is available at https://www.youtube.com/watch?v=zPhrq52jQL8.

[2] Paper published in collaboration with researchers from the University of Brescia (Italy) and the University of Lisbon (Portugal).

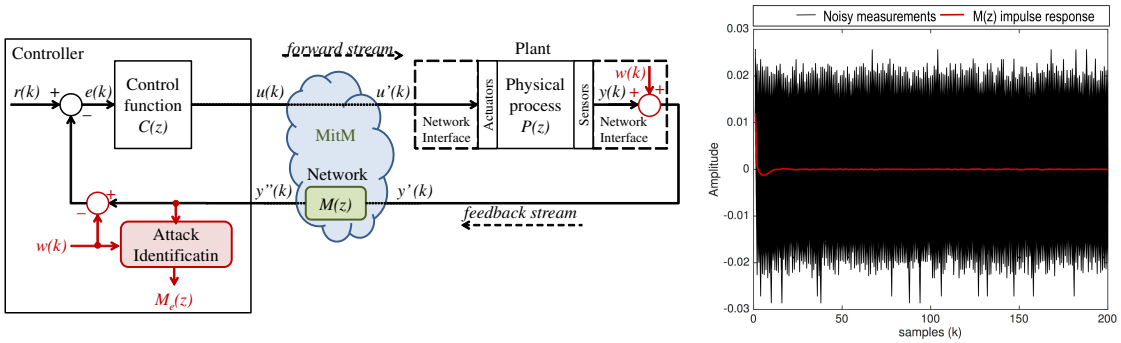
wherein the probabilities $p_{11}(l)$, $p_{12}(l)$. $p_{21}(l)$ and $p_{12}(l)$ are taken from probability density functions designed to hinder the identification process and provide system stability. The results indicate that, with this countermeasure, the models obtained by the attacker are imprecise/ambiguous in such a way that, with the obtained information, the attacker may hesitate in launching covert/model-based attacks. At the same time the countermeasure provides an adequate control performance.

Last, this work also proposes the countermeasure shown in Figure 4(a) to detect/identify the linear time-invariant (LTI) functions executed by a Man-in-the-Middle (MitM) during controlled data injection attacks in NCSs. It consists of a link monitoring strategy [de Sá et al. 2019a, de Sá et al. 2020], which uses white gaussian noise $w(k)$ to excite possible attack functions $M(z)$, in order to obtain the information necessary to identify the attack. To increase the accuracy when identifying the attack using white gaussian noise, this work also proposes a Noise Impulse Integration (NII) technique [de Sá et al. 2019b, de Sá et al. 2020], which is developed inspired by the pulse integration process of radar systems. As shown in Figure 4(b), it is demonstrated that the NII technique is able to reveal the impulse response of the attack function $M(z)$ (in red) based on the signals produced by the white gaussian noise injected in the NCS (in black). The results show that NII effectively contributes to enhance the accuracy of the countermeasure.



(a) NCS with a switching controller. (b) Markov chain switching rule S.

Figure 3. Countermeasure to hinder System Identification attacks.



(a) Architecture of the countermeasure. (b) Result of the NII technique.

Figure 4. Countermeasure to identify controlled data injection attacks in NCSs.

5. Scientific Production: Publications, Award and Partnership

To date, this research resulted in the publication of papers in journals/conferences, an award, and an international research partnership, as follows:

- **Papers:** The results of this research are published in five international journals, whose classifications and number of citations (from Google Scholar on September

11, 2020) are shown in Table 1. Also, six papers on this research were published in Conferences, Symposiums and Workshops [de Sá et al. 2016, de Sá et al. 2017a, de Sá et al. 2017d, de Sá et al. 2018a, de Sá et al. 2019b, de Sá et al. 2019a].

Table 1. Papers published in journals

Paper	Classification	Citations*
[de Sá et al. 2017c]	Qualis A1 / Impact Factor JCR 7.377	50
[Ferrari et al. 2020]	Qualis A1 / Impact Factor JCR 7.515	0
[de Sá et al. 2020]	Qualis A1 / Impact Factor JCR 3.031	1
[de Sá et al. 2017b]	Qualis A2 / Impact Factor JCR 2.39	7
[de Sá et al. 2018b]	Qualis B1	11

- **Award:** This work was awarded with the first place in the Student Contest of the 2018 IEEE International Workshop on Metrology for Industry 4.0 and IoT (Brescia, Italy), with the poster: "Covert Attacks and Challenges for Metrology in Industrial Control Systems".
- **Partnership:** The partial results of this work, presented in the 2018 IEEE International Workshop on Metrology for Industry 4.0 and IoT (Brescia, Italy), motivated an international partnership between the National Institute of Metrology, Quality and Technology (Inmetro, Brazil) and the University of Brescia (Italy). The advisor and co-advisor of this thesis are affiliated to Inmetro, which is interested in this research (such as the University of Brescia). The partnership is formalized through a Memorandum of Understanding signed by both institutions in 2018.

6. Conclusions

This thesis has resulted in contributions that encompass novel System Identification attacks and model-based offensives in NCS. It introduces a new taxonomy that supports the discussion on the relationship between these classes of attack. Finally, it proposes countermeasures to improve the cybersecurity of NCSs against these kinds of threat. The results of this research are published in relevant journals, as well as conferences/symposiums/workshops, which we consider important to share the outcomes and promote the cybersecurity of NCSs – our main goal. As discussed, it encompasses a wide number of potential – sometimes critical – targets that can suffer from the attacks herein studied, as well as benefit from the countermeasures proposed in this work.

References

Collantes, M. H. and Padilla, A. L. (2015). Protocols and network security in ics infrastructures. Technical report, Spanish National Institute for Cyber-security (INCIBE.

de Sá, A. O., Carmo, L. F. R. d. C., and Machado, R. C. S. (2018a). Evaluation on passive system identification and covert misappropriation attacks in large pressurized heavy water reactors. In *2018 Workshop on Metrology for Industry 4.0 and IoT*, pages 203–208. IEEE.

de Sá, A. O., Carmo, L. F. R. d. C., and Machado, R. C. S. (2019a). Countermeasure for identification of controlled data injection attacks in networked control systems. In *2019 II Workshop on Metrology for Industry 4.0 and IoT (MetroInd4. 0&IoT)*, pages 455–459. IEEE.

de Sá, A. O., Casimiro, A., Machado, R., da C, C., and Luiz, F. (2020). Identification of data injection attacks in networked control systems using noise impulse integration. *Sensors*, 20(3):792.

de Sá, A. O., Casimiro, A., Machado, R. C. S., and da Costa Carmo, L. F. R. (2019b). Bio-inspired system identification attacks in noisy networked control systems. In *11th EAI International Conference on Bio-inspired Information and Communications Technologies (BICT)*, pages 1–11, Pittsburgh, USA. Springer.

de Sá, A. O., da Costa Carmo, L. F. R., and Machado, R. C. S. (2016). Ataques furtivos em sistemas de controle físicos cibernéticos. In *Anais do XVI Simpósio Brasileiro em Segurança da Informação e de Sistemas Computacionais (SBSeg)*, pages 128–141.

de Sá, A. O., da Costa Carmo, L. F. R., and Machado, R. C. S. (2017a). Bio-inspired active attack for identification of networked control systems. In *10th EAI International Conference on Bio-inspired Information and Communications Technologies (BICT)*, pages 1–8, New Jersey, USA. ACM.

de Sá, A. O., da Costa Carmo, L. F. R., and Machado, R. C. S. (2017b). Bio-inspired active system identification: a cyber-physical intelligence attack in networked control systems. *Mobile Networks and Applications*, pages 1–14.

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

de Sá, A. O., da Costa Carmo, L. F. R., and Machado, R. C. S. (2017d). Use of switching controllers for mitigation of active identification attacks in networked control systems. In *2017 IEEE Cyber Science and Technology Congress (CyberSciTech2017)*, pages 1–6, Orlando, FL, USA. IEEE.

de Sá, A. O., da Costa Carmo, L. F. R., and Machado, R. C. S. (2018b). A controller design for mitigation of passive system identification attacks in networked control systems. *Journal of Internet Services and Applications*, 9(1):1–19.

Ferrari, P., Sisinni, E., Bellagente, P., Rinaldi, S., Pasetti, M., de Sá, A. O., Machado, R. C. S., d. C. Carmo, L. F. R., and Casimiro, A. (2020). Model-based stealth attack to networked control system based on real-time ethernet. *IEEE Transactions on Industrial Electronics*, pages 1–12. Accepted to be published in the next issue. Available as early access with doi: 10.1109/TIE.2020.3001850.

McLaughlin, S., Konstantinou, C., Wang, X., Davi, L., Sadeghi, A.-R., Maniatakos, M., and Karri, R. (2016). The cybersecurity landscape in industrial control systems. *Proceedings of the IEEE*, 104(5):1039–1057.

Smith, R. (2011). A decoupled feedback structure for covertly appropriating networked control systems. In *Proceedings of the 18th IFAC World Congress 2011*, volume 18, Milano, Italy. IFAC-PapersOnLine.

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

Teixeira, A., Shames, I., Sandberg, H., and Johansson, K. H. (2015). A secure control framework for resource-limited adversaries. *Automatica*, 51:135–148.