Poster: Towards Robust Semantic Reverse Engineering of Control System Binaries

Pengfei Sun
Shape Security
psun@shapesecurity.com

Luis Garcia
University of California, Los Angeles
garcialuis@ucla.edu

Saman Zonouz
Rutgers University
saman.zonouz@rutgers.edu

1 Poster Abstract

In recent years, cyber-physical Internet-of-things (IoT) have received considerable attention due to security concerns originated by the trend to connect those critical platforms to the Internet [8]. Critical infrastructures connected to and controlled by Cyber-physical systems (CPS) substantiate these security concerns. Control algorithms in cyber-physical IoT platforms act as functional guarantees for the entire cyber-physical system [1, 6, 7]. As indicated by the past attacks [3–5, 9], adversaries are often attracted to vulnerabilities that directly affect the core embedded controller algorithm implementations. The recent exponential growth of major cyber-physical IoT attacks indicate the insufficiency of existing security analysis solutions to protect controller software in aforementioned cyber-physical platforms. A common feature in most of the past attacks has been the adversaries' focus on affecting the controller software behavior.

On the protection side, however, the state-of-the-art reverse engineering and vulnerability assessment tools are unable to extract and leverage precise, domain-specific semantics of low-level embedded binary modules. Therefore, they fail to reason about the impact of a particular vulnerability to the overall system. Our previous work, MISMO [10], provided a reverse engineering framework that extracts algorithm-level semantics from stripped embedded software binary implementations of IoT and cyber-physical control algorithms. MISMO utilizes dynamic binary analysis and comparison of mathematical expressions to recover a particular algorithm implementation's cyber-physical execution semantics . The overview of MISMO is presented in Figure 1. MISMO performs dynamic binary analysis to locate the target subroutines of the executable that implements the control algorithm. The arithmetic operations of the execution paths are analyzed symbolically to build a binary-level abstract syntax tree (AST) for the corresponding output values. The generated AST subtrees are recursively compared to and matched with the algorithm-level AST subtrees of the control theoretic expressions. Consequently, our solution fills the semantic gap between the low-level binary executables and high-level algorithmic descriptions with regards to control and data flows.

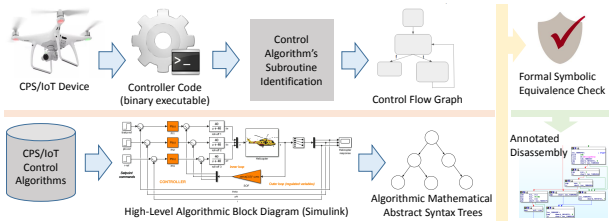

Figure 1: Overview of MISMO Framework.

1.1 Future Work

However, there still remain several research challenges in the context of semantic reverse engineering. In the poster, we will detail the MISMO framework and discuss how MISMO scales in the context of larger datasets as well as the efficacy of MISMO in the context of obfuscation techniques.

Robustness to obfuscation. We will identify the research challenges with respect to control system algorithms that are obfuscated through techniques such as neural-network approximation [2]. We hypothesize that characteristics of the dynamic behavior will be invariant across obfuscation techniques.

Larger control algorithm datasets. In our previous work, we only searched Commercial-off-the-shelf (COTS) binaries using a small database of control algorithms. We propose a larger case study in a variety of applications that utilize a larger set of algorithms. For instance, a drone may have several control system algorithm implementations for its internal control loop that may also be compositional. A much more rigorous evaluation would show the efficacy of MISMO in the real world.

References

[1] Alvaro A Cardenas, Saurabh Amin, and Shankar Sastry. Secure control: Towards survivable cyber-physical

systems. In *2008 The 28th International Conference on Distributed Computing Systems Workshops*, pages 495–500. IEEE, 2008.

[2] Hadi Esmaeilzadeh, Adrian Sampson, Luis Ceze, and Doug Burger. Neural acceleration for general-purpose approximate programs. In *Proceedings of the 2012 45th Annual IEEE/ACM International Symposium on Microarchitecture*, pages 449–460. IEEE Computer Society, 2012.

[3] F-Secure Labs. BLACKENERGY and QUEDAGH: The convergence of crimeware and APT attacks, 2016.

[4] Nicolas Falliere, Liam O Murchu, and Eric Chien. W32. stuxnet dossier. *White paper, Symantec Corp., Security Response*, 5(6):29, 2011.

[5] Luis Garcia and Saman A Zonouz. Hey, my malware knows physics! attacking plcs with physical model aware rootkit. 2017.

[6] Cheolhyeon Kwon, Weiyi Liu, and Inseok Hwang. Security analysis for cyber-physical systems against stealthy deception attacks. In *American Control Conference (ACC), 2013*, pages 3344–3349. IEEE, 2013.

[7] Yilin Mo, Emanuele Garone, Alessandro Casavola, and Bruno Sinopoli. False data injection attacks against state estimation in wireless sensor networks. In *Decision and Control (CDC), 2010 49th IEEE Conference on*, pages 5967–5972. IEEE, 2010.

[8] European Network and Information Security Agency (ENISA). Protecting industrial control systems – recommendations for Europe and Member States. https://www.enisa.europa.eu/, 2011.

[9] Devendra Shelar, Pengfei Sun, Saurabh Amin, and Saman Zonouz. Compromising security of economic dispatch in power system operations. In *Dependable Systems and Networks (DSN), 2017 47th Annual IEEE/IFIP International Conference on*, pages 531–542. IEEE, 2017.

[10] Pengfei Sun, Luis Garcia, and Saman Zonouz. Tell me more than just assembly! reversing cyber-physical execution semantics of embedded iot controller software binaries. In *Dependable Systems and Networks (DSN), 2019 49th Annual IEEE/IFIP International Conference on*. IEEE, 2019.