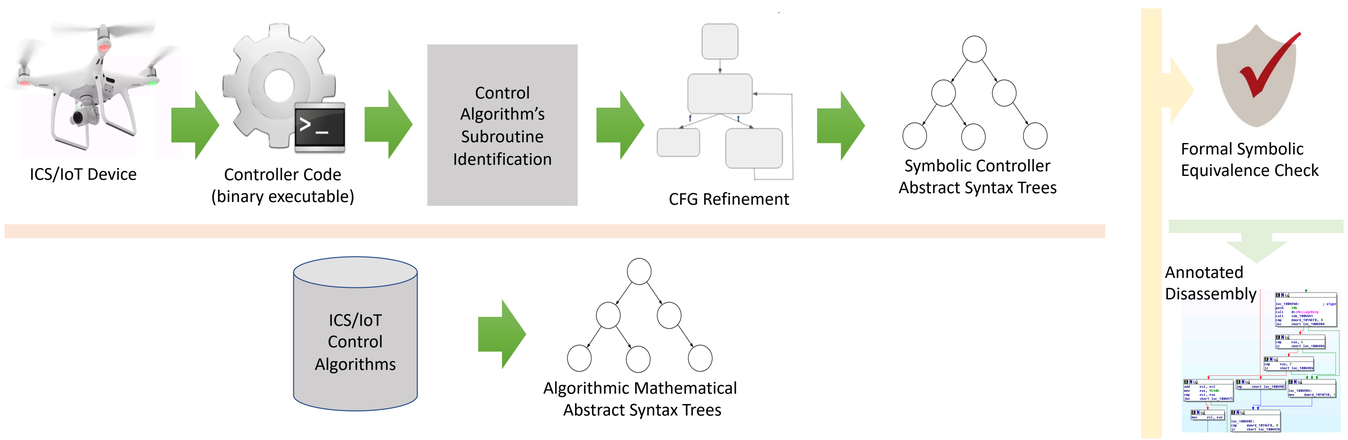
A **domain-specific** reverse engineering framework to extract **high-level algorithmic** control- and data-flow **semantics** from embedded controller binary executables

Towards Robust Semantic Reverse Engineering of Control System Binaries

Pengfei Sun, Luis Garcia, Saman Zonouz
Shape Security, UCLA and Rutgers University

Mismo Overview



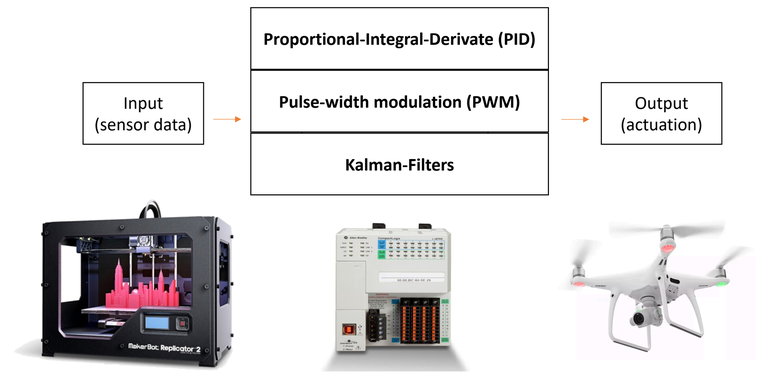
Main Idea

- A general framework to extract semantic information of an embedded firmware binaries with respect to its associated high-level control algorithm.
- Using dynamic binary analysis and symbolic comparison of the mathematical and binary expressions to fill the semantic gap between high-level algorithm descriptions and low-level stripped binary segments

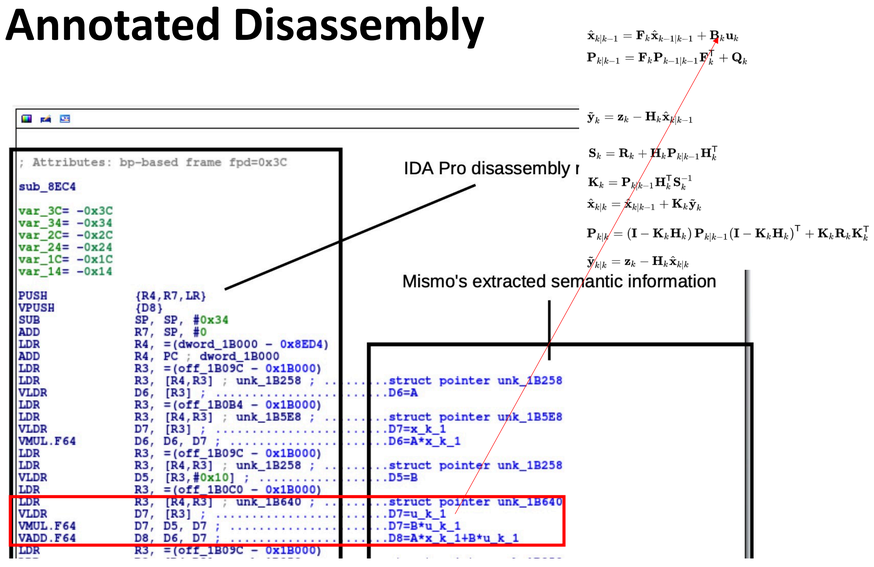
Potential Use-cases

- Binary vulnerability assessment
- Memory forensics analysis
- Sensitive code and data segment protection
- Correct algorithm implementation verification
  - ➢ Discovered a zero-day vulnerability in the Linux kernel controllers versions 3.13 and above.
- Binary-level software similarity measures

Domain Knowledge (CPS)



Annotated Disassembly



Future Work

- Robustness to obfuscation
  - ➢ Obfuscated through techniques such as neural-network approximation.
- Larger control algorithm datasets
  - ➢ Collect more control algorithms
  - ➢ Distinguish different control algorithms

UCLA SHAPE

Take a picture to download the full paper