Detection of Attacks

Securing critical information infrastructures

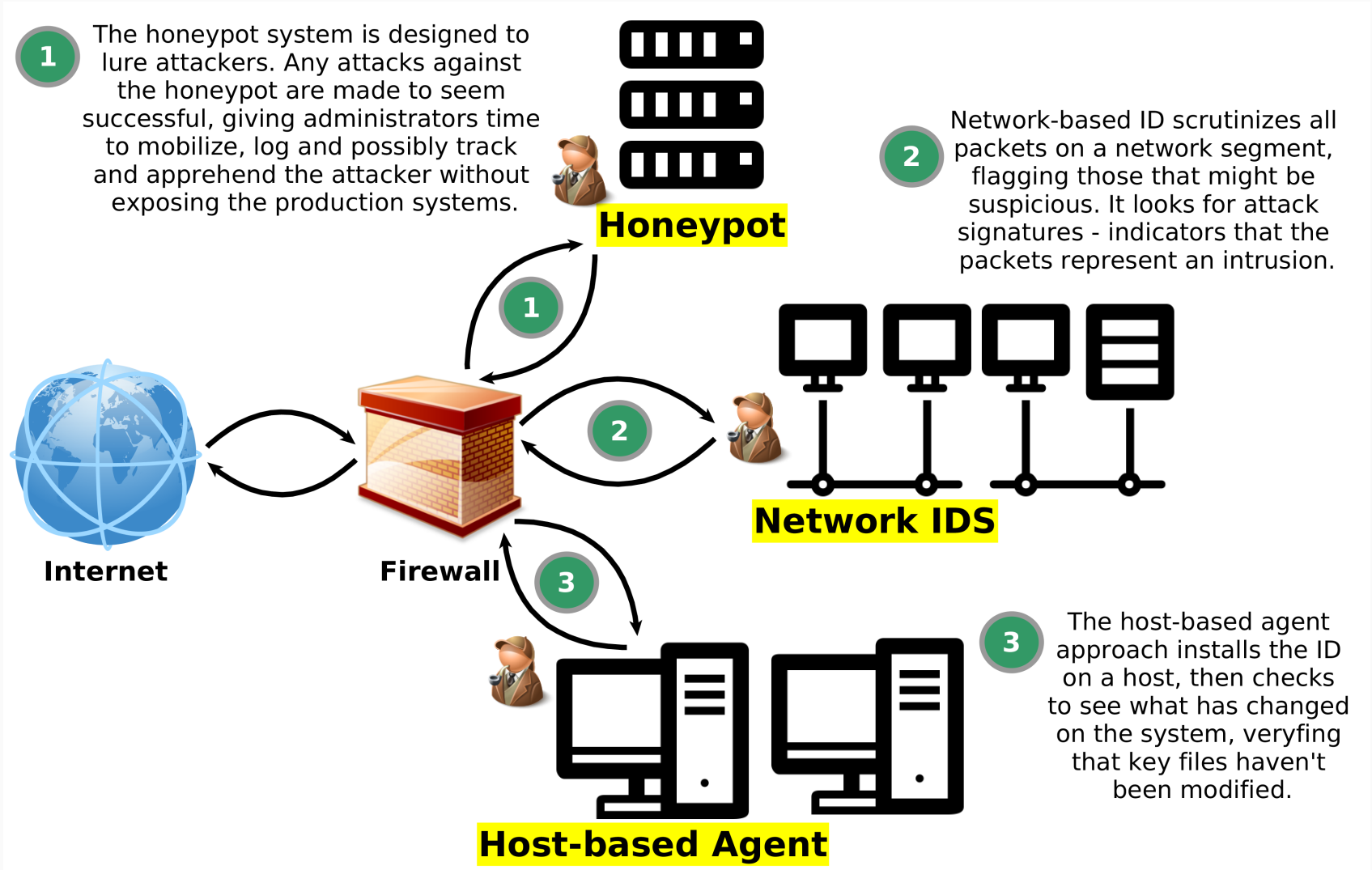
Jonathan Jogenfors
Leonardo Iwaya

2016-09-26

Why detect attacks?

- Model of information security: Prevention, Detection, Reaction (PRD)
  - Prevention: Difficult because attacker has the advantage. Large attack surface.
  - Reaction: Too late!
  - Detection: What our papers are all about

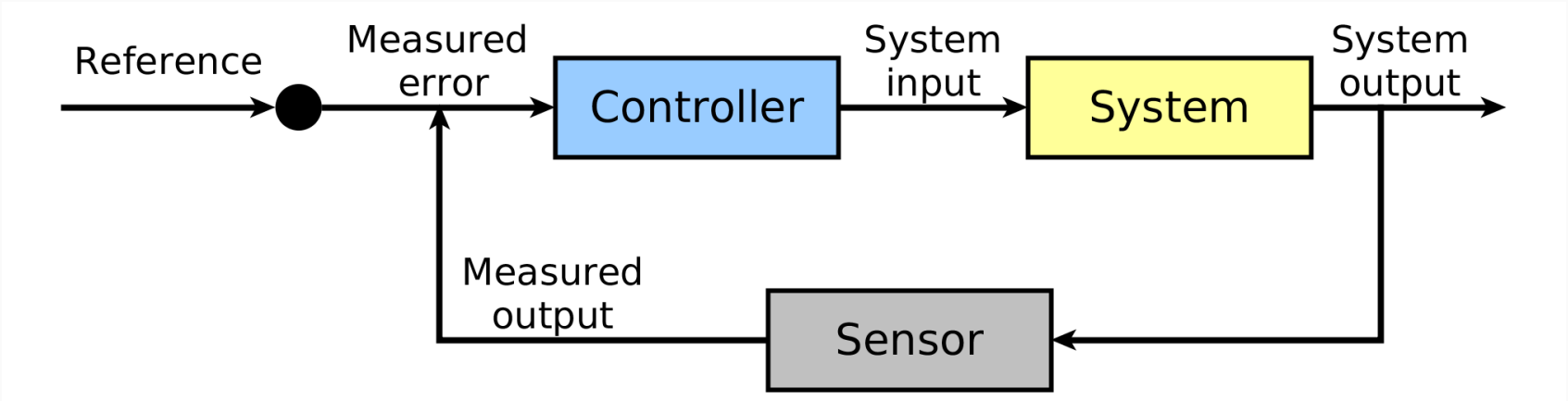
Intrusion Detection Systems & Honeypots



Source: http://www.computerworld.com/article/2592425/lan-wan/intrusion-detection.html

Control Theory & Cybernetics

Control systems: measure, compare, compute and correct.



Feedback loop to control the behavior of a system by comparing its output to a desired value, and applying the difference as an error signal to dynamically change the output so it is closer to the desired output.

Papers

Motivation

- Papers [1] and [2] were **selected** based on their **relevance** to the theme of *"attack detection methods"* for *"critical infrastructures"*.
- Paper [3] was in the course's reading list.

Paper	Year	CI Sub-area	Citations	Journal IF
Pasqualetti et al [1]	2013	Attack Detection	210	2.777
Genge et al [2]	2015	Attack Prevention & Detection	12	1.351
Vasilomanolakis et al [3]	2016	Attack Detection	1	n.a

Attack detection and identification in cyber-physical systems (2013)

Model systems as a formal linear system:

$$\begin{aligned} E\dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{aligned} \tag{1}$$

The paper is highly theoretic but delivers a number of *fundamental* results on attack detections

Attack detection and identification in cyber-physical systems (2013)

Assumptions:

- Determinant is nonzero:

$$|sE - A| \neq 0 \tag{2}$$

- Initial condition is consistent
- Smooth input signal (no impulses)

Attack detection and identification in cyber-physical systems (2013)

- Attack monitor uses the output signal $y(x, u, t)$ to detect attacks on the input $u(t)$.
- Attack is detectable (weaker)
- Attack is identifiable (stronger)

Attack detection and identification in cyber-physical systems (2013)

Some of the results on detectability:

Theorem

A nonzero attack is undetectable iff

$$y(x_1, u_K, t) = y(x_2, 0, t) \tag{3}$$

Theorem

A nonzero attack is unidentifiable iff

$$y(x_1, u_K, t) = y(x_2, u_R, t) \tag{4}$$

for some attack R with $|R| \leq |K|$ and $R \neq K$.

Attack detection and identification in cyber-physical systems (2013)

The paper continues to design attack detection systems

- Centralized system
- Distributed system

as well as some results on the "Attack Identification Problem".

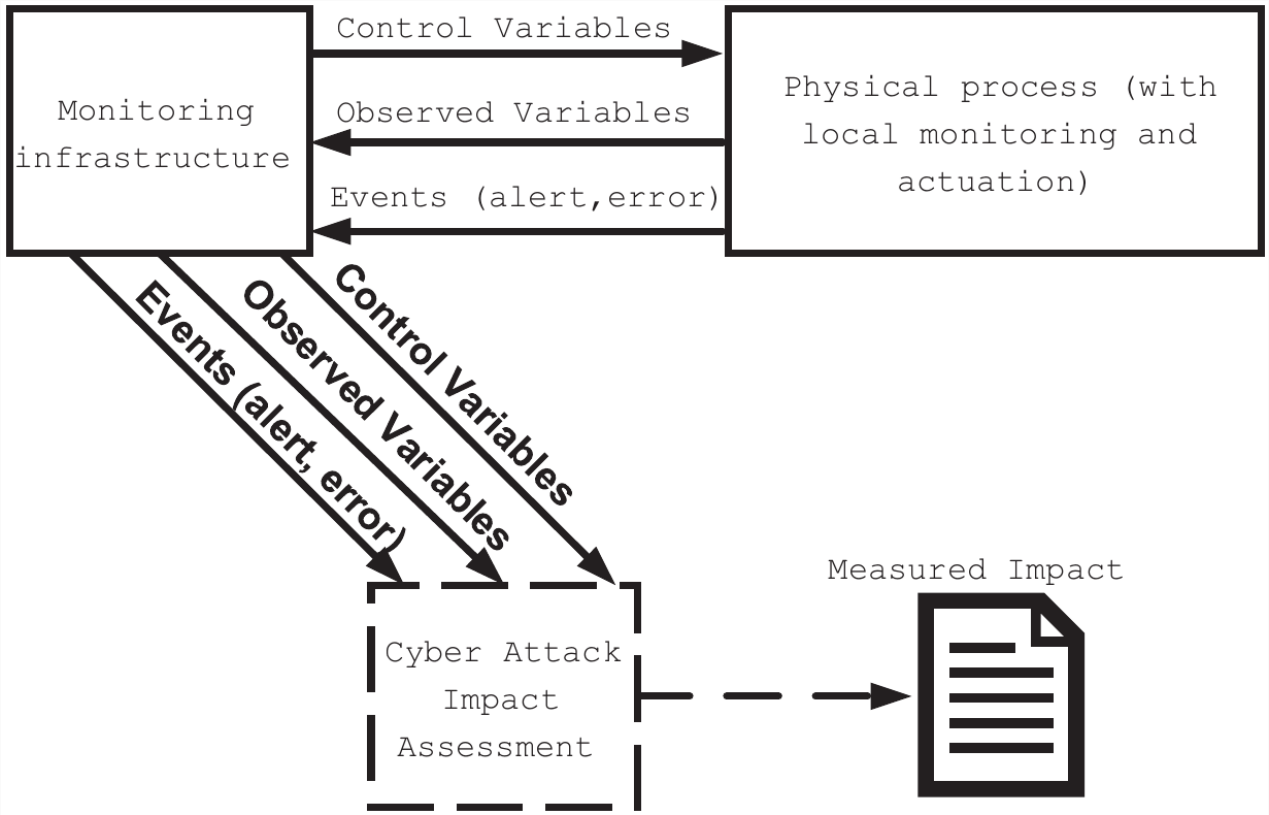
A system dynamics approach for assessing the impact of cyber attacks on CI (2015)

Aim & Contribution

- To **identify** and **rank** assets in complex, large-scale and heterogeneous CIs.
- **Cyber Attack Impact Assessment** (CAIA) methodology that helps system admins to understand:
    1. How cyber attacks affect the normal functioning of physical processes?
    2. What cyber assets would cause the most negative impact if compromised?

A system dynamics approach for assessing the impact of cyber attacks on CI (2015)

CAIA Methodology



A system dynamics approach for assessing the impact of cyber attacks on CI (2015)

Experiments & Comparisons

- First, the **basic functioning** of CAIA is demonstrated using IEEE 14-bus electric grid model.
- Second, CAIA's **scalability** is proven by using attack scenarios in the context of IEEE 300-bus electric grid model.
- Third, CAIA's **cross-sector applicability** is evaluated using Tennessee Eastman chemical process system.
- The methodology was also **compared** with other approaches (i.e., graph-theoretic and electrical centrality metric techniques).

A system dynamics approach for assessing the impact of cyber attacks on CI (2015)

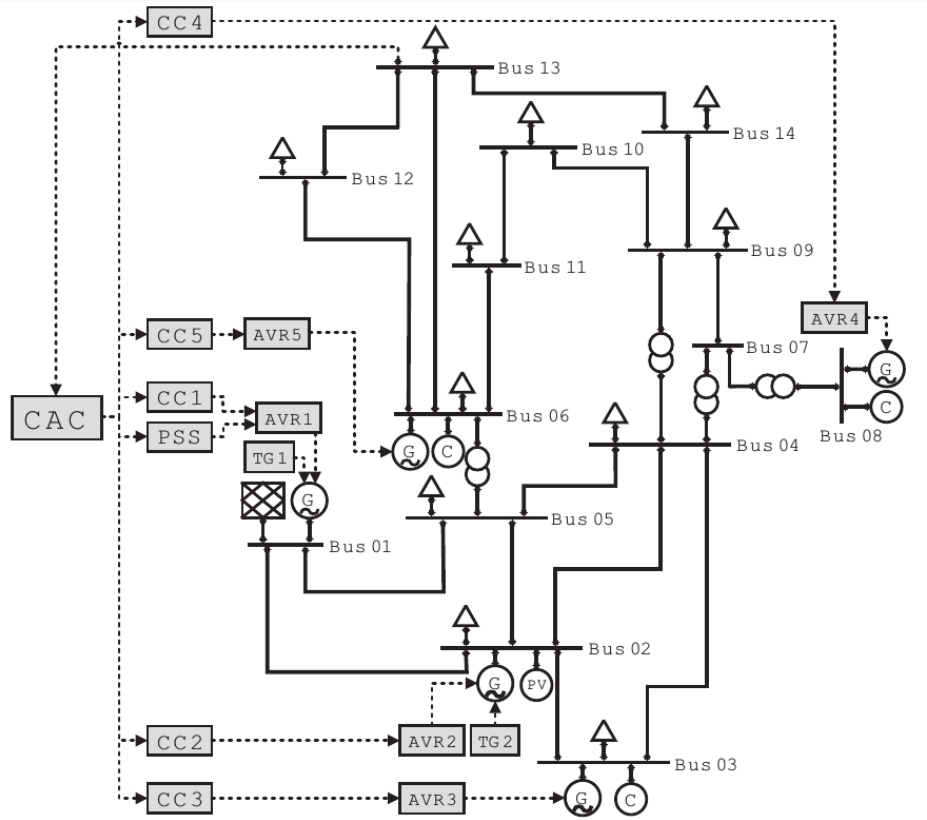


Fig. 5 – IEEE 14-bus model and its associated controllers.

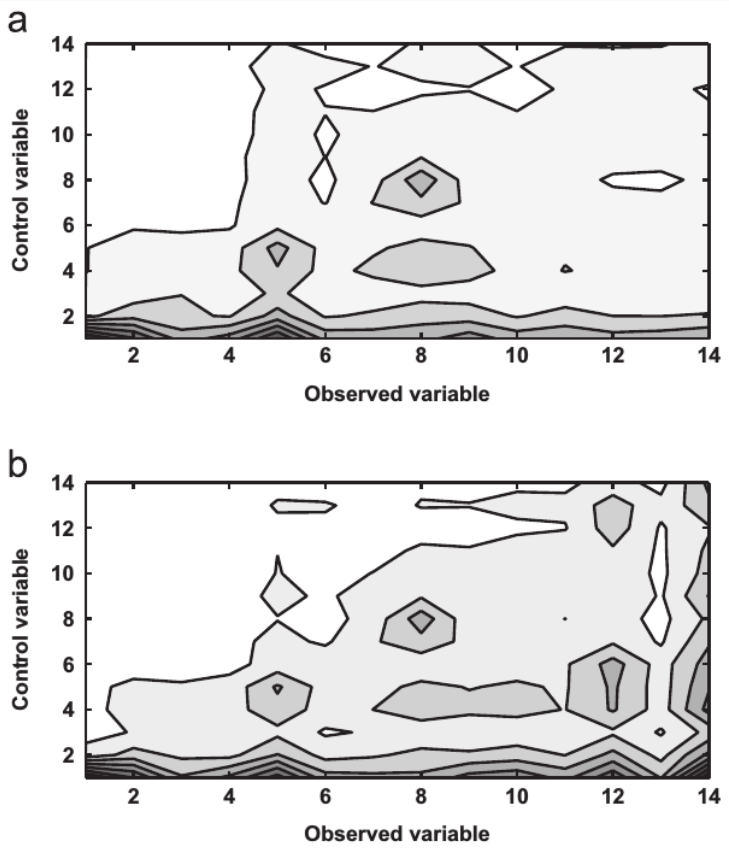


Fig. 7 – Effects of observed variable weights on the impact matrix for the IEEE 14-bus model. (a) Equal weights for all observed variables and (b) increased weights for observed variables (bus line voltage levels) 10, 12 and 14.

A system dynamics approach for assessing the impact of cyber attacks on CI (2015)

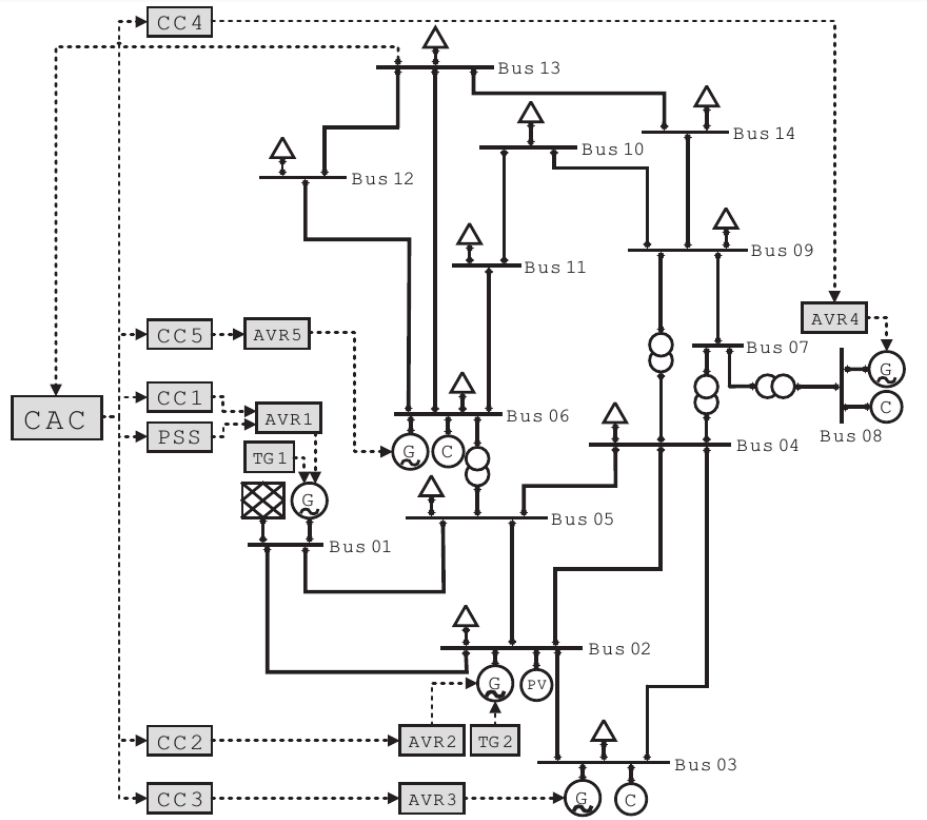


Fig. 5 – IEEE 14-bus model and its associated controllers.

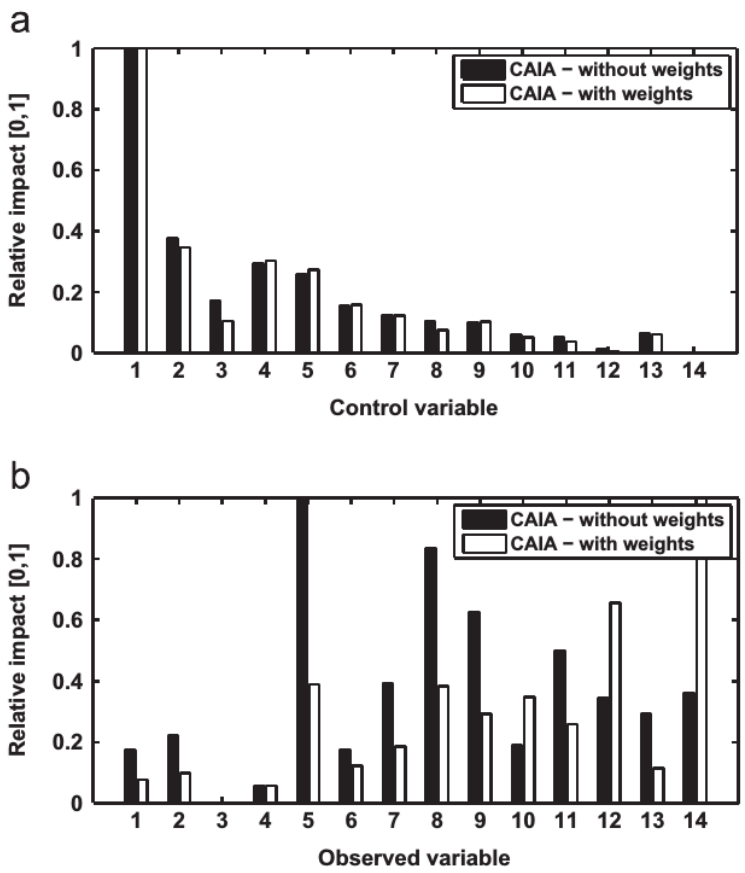


Fig. 8 – Effects of observed variable weights on impact rankings for the IEEE 14-bus model. (a) Impacts on control variables and (b) impacts on observed variables.

A system dynamics approach for assessing the impact of cyber attacks on CI (2015)

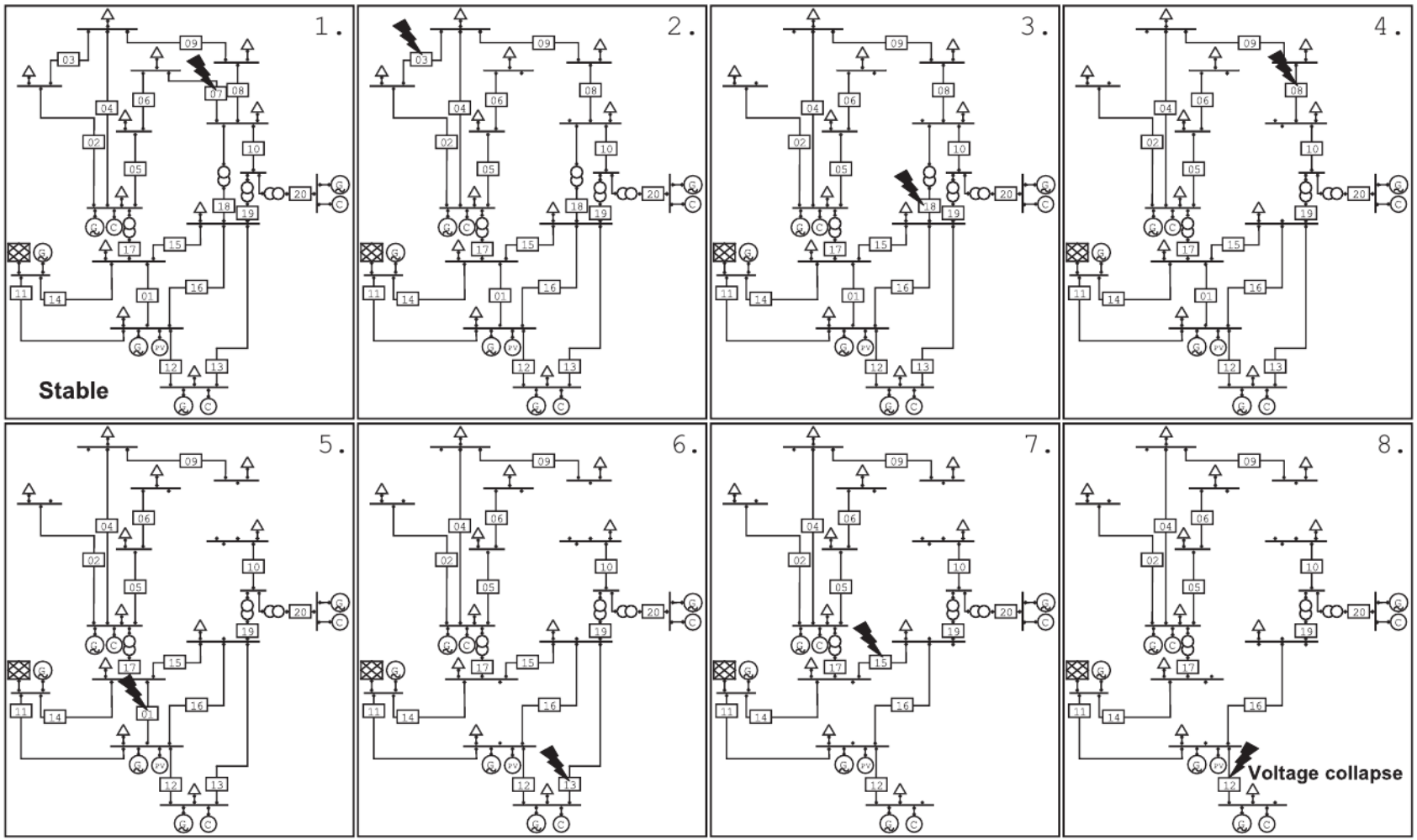


Fig. 20 – Stealthy cyber attack sequence that disconnects substation lines.

A system dynamics approach for assessing the impact of cyber attacks on CI (2015)

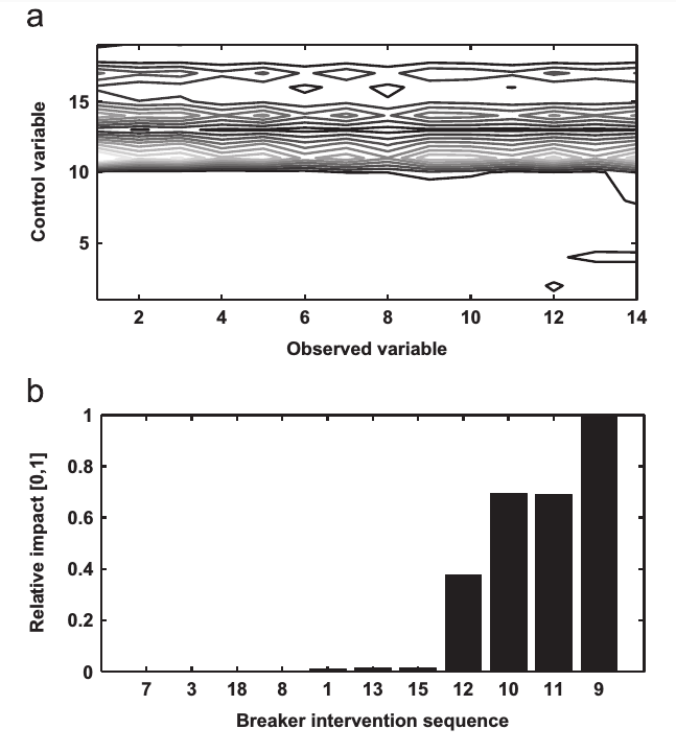


Fig. 19 – Stealthy cyber attack on the IEEE 14-bus model line breakers. (a) CAIA impact matrix and (b) ordered impact ranking of breakers.

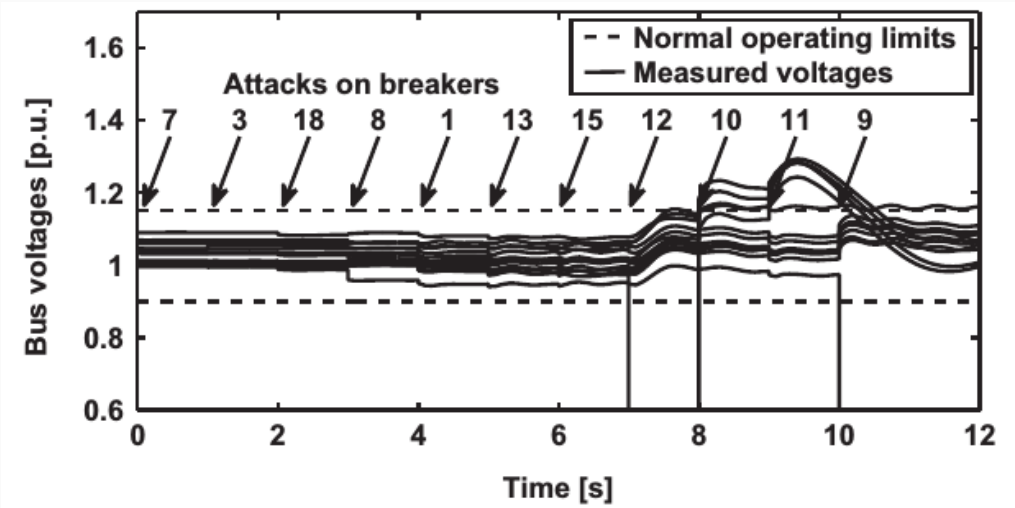


Fig. 21 – Operator's view of the stealthy attack sequence.

A system dynamics approach for assessing the impact of cyber attacks on CI (2015)

Limitations

- CAIA helps to identify and rank assets given specific interventions (e.g., an attack)
- Which interventions are relevant to test (?), and, how to protect the assets after generating the impact matrix (?) are open questions; out of the paper's scope.
- Obvious Note: the knowledge of impact matrices would be definitely valuable to attackers(!); as any risk assessment information.
- Seems hard to reproduce since no detailed information is given about the simulations; plus, no source code.

Multi-stage Attack Detection and Signature Generation with ICS Honeypots (2016)

Aim & Contribution

- HosTaGe: honeypot for detecting **multi-stage attacks** in ICS networks.
- Honeypot extension with capabilities of ICS protocols, i.e., Modbus, S7, SNMP, HTTP, Telnet, SMB and SMTP.
- Basic functions:
  1. notify the network administrators;
  2. produce an attack signature;
  3. forward the signature to the internal IDSs.

Multi-stage Attack Detection and Signature Generation with ICS Honeypots (2016)

Expermients & Comparisons

- HosTaGe was **compared** with "CONPOT ICS/SCADA Honeypot" [1]
- Criteria:
  1. ability to not be evade (i.e., be perceived by attackers);
  2. ability to detect multi-stage attacks;
  3. ability to generate valid signatures for Bro IDS [2].

[1] http://conpot.org/
[2] https://www.bro.org/

Multi-stage Attack Detection and Signature Generation with ICS Honeypots (2016)

Formal Model - Extended Finite State Machine (EFSM)

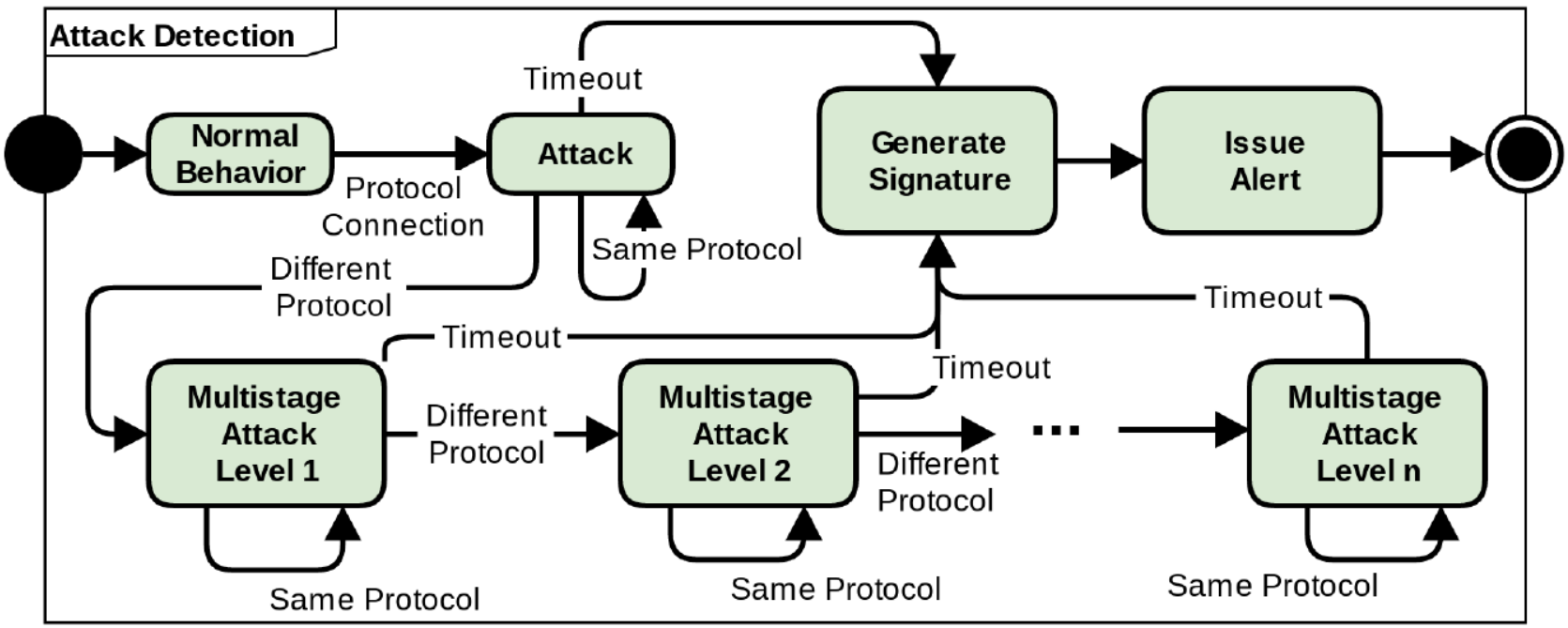


Fig. 1. EFSM of the attack detection and signature generation mechanism.

Multi-stage Attack Detection and Signature Generation with ICS Honeypots (2016)

Formal Model - Extended Finite State Machine (EFSM)

- Detection Mechanism
  1. Single-Protocol Level Detection (SPLD)
  2. Multi-Stage Level Detection (MSLD)
  3. Payload Level Detection (PLD)
- **Time window** (*tw*) determines whether an attack should be mapped as SPLD or MSLD

Multi-stage Attack Detection and Signature Generation with ICS Honeypots (2016)

Example - Signature Generation

- Automatically generate signature for well-known Metasploit script[3] for Modbus services identification.

Listing 1. Modbus attack signature generated by *HosTaGe*

```
signature modbus-signature{
    ip-proto == tcp
    dst-port == 502
    payload
        /\x21\x00\x00\x00\x00\x06\x01\x04\x00
    event "Modbus attack"
    }
```

[3] No further information given by the authors...

Comparison - Honeypot x CONPOT

- Controlled environment, no firewalls, 8 to 12 weeks, probing by Shodan[4].

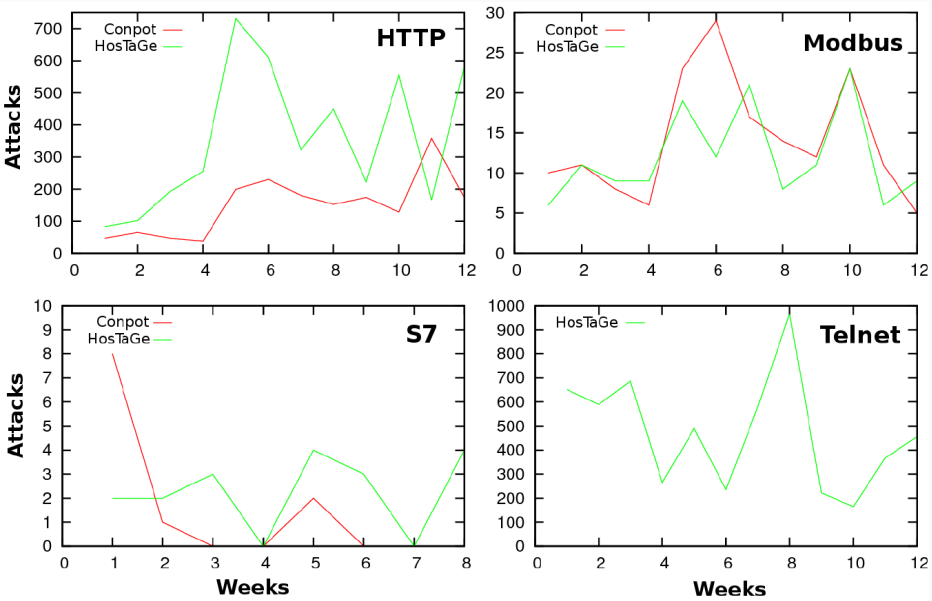


Fig. 3. Comparison of attacks on *HosTaGe* and Conpot for HTTP, Modbus, *S7* and Telnet. Note, that Conpot does not support the Telnet protocol.

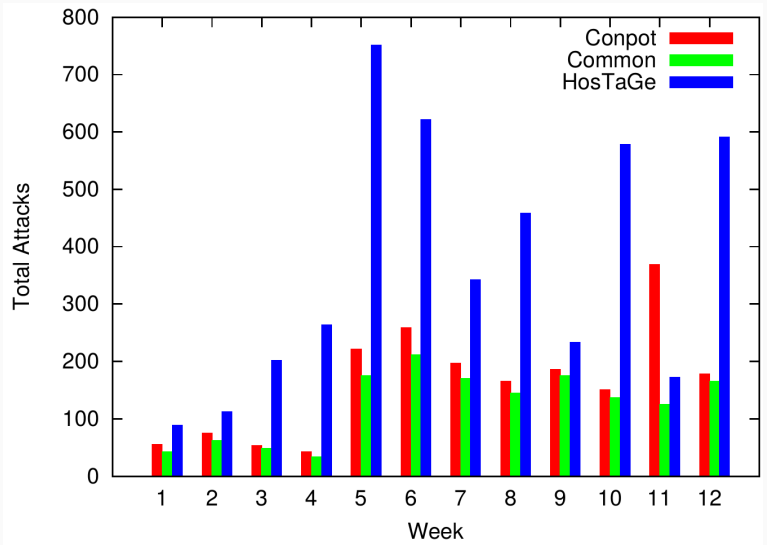


Fig. 4. Comparison of unique and common malicious IP addresses targeting *HosTaGe* and Conpot

[4] https://www.shodan.io/

Multi-stage Attack Detection and Signature Generation with ICS Honeypots (2016)

Limitations

- The evaluation of multi-stage signature generation was rather shallow.
- Shodan's probes were not explained in details, i.e., how Shodan detect a honeypot?

More info about HosTaGe can be found at Darmstad's research group website[5].

[5] https://www.tk.informatik.tu-darmstadt.de/de/research/secure-smart-infrastructures/hostage/

Q&A and Discussions

Debate Suggestions

- **Attacker models** [1, 2] – strong assumptions; absolute knowledge and control.
- **Study validation** – enough tests; data sources; experiment description.
- **Reproducibility** – enough information; open source; plant models.
- **Overall critics** about the papers – readability; depth; contribution.

Question for paper "Attack detection and identification in cyberphysical systems (Pasqualetti et al, 2013)"

Regarding the fundamental limitations derived in the paper, which of the following statements is TRUE:

(a) An unidentifiable attack is also undetectable

(b) The derived limitations are probabilistic in nature

(c) The monitors in this paper are assumed to have no false alarms

(d) An attack can remain undetected if it excites a nonzero system response

Question for paper "Attack detection and identification in cyberphysical systems (Pasqualetti et al, 2013)"

Regarding the fundamental limitations derived in the paper, which of the following statements is FALSE:

(a) A (nonzero) attack is undetectable if the measurements due to the attack coincide with the measurements due to some nominal operationg condition

(b) Due to fundamental limitations in attack detection methods, the model in the paper only allows centralized attack detection systems to be analyzed

(c) The attack signal is assumed to be smooth (i.e. no impluse inputs)

(d) The paper considers colluding, omniscient attackers with unlimited computation capabilities

Question for paper "A system dynamics approach for assessing the impact of cyber attacks on critical infrastructures (Vasilomanolakis et al, 2013)"

Regarding the Cyber Attack Impact Assessment (CAIA) methodology proposed in the paper, which of the following statements is FALSE:

(a) CAIA helps system administrators to analyze how cyber attacks affect the normal functioning of physical processes.

(b) The proposed approach computes the covariances of observable variables before and after an specific intervention in control variables.

(c) CAIA methodology is mainly inspired in graph-theoretical and electric centrality metric approaches.

(d) CAIA helps to identify and rank assets in the context of critical infrastructures.

Question for paper "A system dynamics approach for assessing the impact of cyber attacks on critical infrastructures (Vasilomanolakis et al, 2013)"

The proposed methodology (CAIA) was validated by various experiments. Which of the following statements is FALSE:

(a) The conducted experiments were able to demonstrate CAIA's efficiency, scalability and cross-sector applicability.

(b) Final results demonstrate that CAIA methodology is only suitable for electric grid models, such as IEEE 14-bus and 300-bus.

(c) To demonstrate CAIA's cross-sector applicability, the authors used the Tennessee Eastman chemical plant model.

(d) The authors described how an attacker can use CAIA results to plan and execute a stealthy cyber attack, in which multiple low-impact variables are affected simultaneously to cause severe infrastructure degradation.

Question for paper "Multi-stage Attack Detection and Signature Generation with ICS Honeypots (2016)"

Regarding the concept of honeypot addressed in the paper, which of the following statements is FALSE:

(a) Honeypots exhibit a high rate of false positives.

(b) Honeypots are systems whose only value is to be probed, attacked and compromised.

(c) Honeypots are used to attract malicious users and study their activities.

(d) One essential requirement for honeypots is their ability to remain undetected.

Question for paper "Multi-stage Attack Detection and Signature Generation with ICS Honeypots (2016)"

Regarding the comparative study between HosTaGe and Conpot, which of the following statements is FALSE:

(a) Overall, HosTage was able to detect more attacks than Conpot.

(b) The analysis of multi-stage attacks was performed only for HosTaGe.

(c) HosTaGe honeypot presented better evasion (ability to remain undetected) capabilities than CONPOT.

(d) Conpot is not an ICS-specific honeypot and therefore it supports a smaller number of protocols.

References

F. Pasqualetti, F. Dörfler, and F. Bullo. "Attack Detection and Identification in Cyber-Physical Systems". In: *IEEE Transactions on Automatic Control* 58.11 (Nov. 2013), pp. 2715–2729. ISSN: 0018-9286. DOI: 10.1109/TAC.2013.2266831.

B. Genge, I. Kiss, and P. Haller. "A System Dynamics Approach for Assessing the Impact of Cyber Attacks on Critical Infrastructures". In: *International Journal of Critical Infrastructure Protection* 10 (Sept. 2015), pp. 3–17. ISSN: 18745482. DOI: 10.1016/j.ijcip.2015.04.001.

References II


E. Vasilomanolakis, S. Srinivasa, C. G. Cordero, and M. Muhlhauser. “Multi-Stage Attack Detection and Signature Generation with ICS Honeypots”. In: IEEE, Apr. 2016, pp. 1227–1232. ISBN: 978-1-5090-0223-8. DOI: 10.1109/NOMS.2016.7502992.