[9053] Characterizing Building Automation System Attacks and Attackers

PRELIMINARY DECISION: accept

Summary of Reviews

- Review 1: 3 (5)
- Review 2: 2 (3)
- Review 3: 0 (3)
- Review 4: 0 (2)

Reviews

Review 1

TOTAL SCORE: 3

Overall evaluation: 3 (strong accept)

Reviewer's confidence: 5 (expert)

The paper describes a characterisation of attackers of building automation systems, and convincingly argues that this characterisation is different from the existing Cyber Physical Systems attacker characterisation. The paper briefly describes some cases and how the information from these cases leads to the different attacker characterisations.

The authors will also share the dataset of documented BAS attacks, which will be a good contribution to the field.

Some minor remarks to improve the paper:

* Firstly/Secondly -> First/Second

* In section III.B you first use the value 'null' to describe missing information, and then describe '0 (i.e. null)' as absence of a dimension. This confused me, I would suggest using the same wording to describe both values.

* Table I has a field containing '\' in the Honesty row

* The inclusion of the example in Figure 2 confused me, I would delay describing those examples to later.

* The placement of Table V is very far from the reference and explanation of the table, could this be improved?

One question to the authors, the Mirai botnet actors are described and categorised. Initially this was a single actor group, but soon after this botnet got serious attention, the source code was made public. Has the categorisation been made before or after this leak? Does that leak change the categorisation?

Review 2

TOTAL SCORE: 2

Overall evaluation: 2 (accept)

Reviewer's confidence: 3 (medium)

This is an interesting paper, focusing on characterizing building automation system attacks (BAS) and attackers. The main aim of the paper is to create and analyze an attacker model for BAS.

The authors provide a description of existing relevant work. Also, the method is being described sufficiently. Firstly, the authors collected a database of 26 attacks involving building automation systems. The database is built from publicly disclosed security incidents involving Building Automation Systems (BAS). Secondly, they created and analysed the attacker model for BAS based on empirical observations and highlighted core differences and similarities between BAS and cyber-physical-system (CPS) attackers. The methodology follows certain steps of attack collection, characterisation and evaluation. Also, the taxonomy and methodology presented by Rocchetto and Tippenhauer is followed. The model is then tested by applying in 22 security incidents.

It is interesting to see that the authors discuss profiling details of different types of threat actors, such as the insider, the cybercriminal or the terrorist, trying to identify their characteristics, motivations and practises. The model developed

The results are quite interesting and well presented. The authors discuss the findings, and it is quite interesting to see that these attacks do not require advanced resources or knowledge, something that the authors predict will lead to more attacks of this type.

Review 3

TOTAL SCORE: 0

Overall evaluation: 0 (borderline paper)

Reviewer's confidence: 3 (medium)

The article presents an adaption of a well-known framework from attacker characterization to the particular case of attacks against building automation systems (BAS). The article applies this framework to 26 cases as identified from public sources.

The article is motivated under the assumption that the attackers that go against BAS and CPS differ substantially. While this claim is not substantiated with evidence, the article does not clarify these differences. In fact, my main concern is the lack of definition of what types of systems the article considers BAS and in particular a BAS attack.

The identified 26 attacks against BAS are gathered by desk research. While the authors describe three different phases to gather these articles, the final set of search keywords is not clearly stated in the paper. While this should be easily fixable, my main concerns arise from looking at characteristics of the identified attacks against BAS. A great majority of these seems to be emerging from IoT botnets and hence not specific to attacks against BAS. The paper goes in depth into the analysis of the Mirai botnet (and successors) whose authors have been convicted. How do the authors distinguish IoT-botnet attacks from BAS attacks? Considering attackers behind Mirai to have any specific incentive to attack BAS sounds like a misattribution of the attacker's motivation. Hence, this seems like a major issue regarding the scope of the paper that also hinders the motivation of the authors to define a framework specific for BAS.

Similarly, the values added in Table II seem to be chosen based on the criteria of the person using the framework, and hence, it seems that such vectors cannot be assessed objectively. How would the investigator attributing these values distinguish between a value 1 to a value 2?

On the other hand, the limitation acknowledges in the paper regarding the modelling purely based on publicly available attacks also puts at risk the resulting framework. By the own nature of attacks against BAS and CPS, these are not always disclosed, hence the derived model might be biased towards a particular type of attacker profiles whose attacks reached the general public.

In short, I believe the paper should clarify the definition of BAS and BAS attackers. The current set of attacks selected from the literature contained also consumer IoT-based botnet attacks that could be hardly related to BAS systems. The paper is building the framework based on these attacks, hence, it is crucial that these set of attacks are directly related to BAS and not just collateral non-target specific attacks.

Review 4

TOTAL SCORE: 0

Overall evaluation: 0 (borderline paper)

Reviewer's confidence: 2 (low)

This paper presents an attacker model for Building Automation Systems (BAS). The paper analyzes BAS security incidents and derives an attacker model from these incidents, follow an existing model for Cyber-Physical Systems by Rocchetto and Tippenhauer.

I find the analysis of existing attacks according to the dimensions presented in the paper interesting and useful. However, there are some important parts of the paper that I do not understand. Most importantly, I'm confused by Table 5. Especially, what do the cycles mean?

Furthermore, the paper does not define CPS, ICS and BAS. Thus, footnote 2 states: "The term Industrial Control System (ICS) will be used almost interchangeably with CPS in this work". If CPS is the same as ICT, then I do not wonder that attacker model for BAS is different from attacker model for CPS/ICS. Actually, I do not need a paper to convince me about this.

On the other hand, p.3 states. "From a technological perspective, BASs are much more "open" and interconnected than other CPSs (e.g., an ICS)." From this I derive that BASs are subset of CPSs. I think that the paper should be much more clear about the definitions of the respective systems.

The paper contains some claims that are not supported by citations or data. This support should either be provided, or the claims toned down or removed:

p. 2: "Even though the study was published in 2016, more recent papers do not provide additional contributions with respect to the characterization of CPS attackers." – Please cite these more recent papers.

p.2: "In fact, we are not aware of an attack targeting ICS where the attacker tampered a device controlling the physical world when the goal was not connected to the physical world (e.g., only to gain a foothold on the IT network of the target)." – This should be either supported by an analysis, or removed.

p. 4: "This process is repeated until we reach an agreement ratio higher than 70%." – Why is 70% a magic threshold? How was the agreement calculated? There are some established inter-rater agreement calculation methods, such as Cohen's Kappa, with well-established guidelines for "good" agreement. In any case, the 70% threshold should be justified.

Small remarks:

p. 5: "The thresholds used to map attacker traits to numeric values are described in Table I." – Why are these numeric values needed? Here they surface for the first time, and therefore, should be introduced and explained.