Poster Abstract: Towards identifying IoT traffic anomalies on the Home Gateway

Eman Maali, David Boyle, Hamed Haddadi
Imperial College London

ABSTRACT

The number of IoT devices continues to grow despite the alarming rate of identification of security and privacy issues. There is widespread concern that development of IoT devices is performed without sufficient attention paid to security and privacy issues. Consequently, networks have a higher probability of incorporating vulnerable IoT devices that may be easy to compromise to launch cyber attacks. Inclusion of IoT devices paves the way for a new category of anomalies to be introduced to networks. Traditional anomaly detection techniques (*e.g.,* semi-supervised and signature-based methods), however, are likely inefficient in detecting IoT-based anomalies. This is because these techniques require static signatures of known attacks, specialized hardware, or full packet inspection. They are also expensive, and may be inaccurate or unscalable. Vulnerable IoT devices can be used to perform destructive attacks or invade privacy. The ability to find anomalies in IoT traffic has the potential to assist with early detection and deployment of countermeasures to thwart such attacks. Thus, new techniques for detecting infected IoT devices are needed to mitigate the associated security and privacy risks. In this research, we investigate the possibility to identify IoT traffic using a combination of behavioural profile, predefined blocklist and device fingerprint. Such a system may be able to detect anomalous and/or malicious devices and/or traffic reliably and quickly. Initial results show that for our implementation of such a system, IoT traffic can be identified using device behaviour profile, fingerprint, and contacted destinations. This work takes the first step towards designing and evaluating *iDetector*, a framework that can detect anomalous behaviour within IoT networks. In our experiments, *iDetector* was able to correctly identify 80-90% of all captured traffic traversing a home gateway.

CCS CONCEPTS

• **Computer systems organization** → **Embedded and cyber-physical systems**; • **Security and Privacy** → Anomaly detection and mitigation; • **Networks** → Network Measurement.

KEYWORDS

Internet of Things, IoT, anomaly detection, traffic analysis

SenSys '20, November 16–19, 2020, Virtual Event, Japan

ACM ISBN 978-1-4503-7590-0/20/11…$15.00
https://doi.org/10.1145/1122445.1122456

ACM Reference Format:
Eman Maali, David Boyle, Hamed Haddadi. 2020. Poster Abstract: Towards identifying IoT traffic anomalies on the Home Gateway. In *The 18th ACM Conference on Embedded Networked Sensor Systems (SenSys '20), November 16–19, 2020, Virtual Event, Japan.* ACM, Yokohama,Japan, 2 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

The number of IoT devices connected to the Internet is steadily increasing [2]. As IoT devices are mostly used in their default settings (among other reasons), they are vulnerable to a variety of security and privacy risks [2]. Manufacturers often prioritize functionality, so IoT devices may be shipped without sufficient protection [9]. This, coupled with typical resource constraints, makes applying traditional anomaly detection techniques like semi-supervised and signature-based methods inefficient. Traditional anomaly detection techniques require static signatures of known attacks, specialized hardware, or full packet inspection. These techniques are thus expensive, and may be inaccurate or unscalable [1]. Vulnerable IoT devices can be used to perform destructive attacks (*e.g.,* Mirai, Stuxnet) or pose privacy risks to users [3]. Thus, it is important to detect anomalous behaviour, which can help with early detection and mitigation of the associated risks. There are several challenges in the field of anomaly detection within IoT networks, starting from defining what normal behaviour is to scalability and datasets. The main motivation of this research is to understand these challenges. Hence, we try to understand behavioural profiles of IoT devices and propose a framework for real-time anomaly detection on the home gateway. Our first results show that it is possible to identify normal traffic. The normal traffic pattern can be inferred from the device behavioural profile, contacted destinations, and device fingerprint.

2 SYSTEM DESIGN

This section gives an overview of the design and implementation of our *iDetector* system. First, we analyzed traffic from two testbeds and a collected dataset from earlier works, specifically Mon(IoT)r [8] and IoT-inspector [5]. These sources were used to build a library that includes all possible normal behaviours of the IoT devices.

Figure 1 shows the architecture of the proposed framework. The framework is proposed to be deployed on home gateways. Traffic is passed through the system to allow normal activity and deny/isolate anomalous behaviour(s). *iDetector* consists of three main stages: (1) Identification, (2) Anomaly detection, (3) System Update. The first step in finding anomalies is to set an identification model to identify what is normal. Several identification approaches, e.g. [4], have been proposed, such as using behavioural profile or device fingerprint, yielding promising results. Such approaches, however, are based on input from users that may be subject to several problems, *e.g.,* spoofing. A possible solution is a hybrid

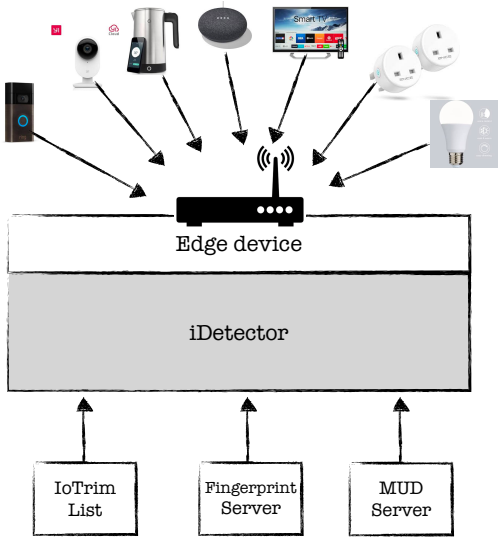

Figure 1: *iDetector* Architecture

approach, which combines behavioural profile, which defines the intended communication patterns for the device when it connects to the network such as Manufacture Usage Description (MUD) [6]; device fingerprint, which is the traffic-based representation of the device; and, predefined blocklist, which defines non-functional destinations for each device [7]. In case of failure of one of the elements in the approach, the others may handle the identification process. Algorithm 1 provides an overview of the identification process flow.

Algorithm 1 IoT Traffic Identification Algorithm

Require: traffic to be analysed
1: Initialize Extract Domain, fingerprint, behaviour profile
2: **for each** traffic **do**
3: check for functional destination
4: check for valid behavioural pattern
5: check for valid fingerprint
6: **end for**

Three outputs are expected from the identification stage. The behaviour profile and the non-functional blocklist is a binary classification, with two possible outcomes exist or non-exist, while the device fingerprint will compute a matching score (*e.g.*, cosine similarity). Then, if the matching score is greater than a certain threshold, the fingerprint matches that device. Each element (behaviour profile, predefined blocklist, and device fingerprint) in the identification stage will be weighed according to its importance. The greatest sum of the attained votes will win. The final stage is updating the framework This must be done often, because behaviour-profile, non-functional blocklist, and fingerprint might be updated with some regularity.

3 IMPLEMENTATION & RESULTS

For the first evaluation, experiments were conducted with *iDetector* deployed on a Raspberry Pi 4 configured as an access point. The experiments were centered around a single professional working from home who has four deployed IoT devices (TP-Link bulb, YI camera, Google Nest mini, Magic Home strip), performs on average 20 activities per day, generates in total 99,423 packets and lasts for two days. For the purposes of experimentation, the following assumptions were made: there is one behaviour-profile for each device, one non-functional block-list, and the captured traffic captures all possible normal behaviours of the four devices. It is important to evaluate how well the identification phase identifies a packet. Thus, the received traffic is compared with the normal behaviour dataset. The comparison result from the identification stage is used to calculate the number of packets that are identified as normal when they are indeed normal. *iDetector* was able to detect normal packets when they are indeed normal, with no false positives and accuracy of 97.8%. The ARP, broadcast, ICMP port unreachable and local network packets (these packets formed 10% of total captured traffic) were filtered out before calculating the percentage of the identified traffic. The percentage of the identified traffic as normal compared with the total number of captured traffic during the experiments was between 80-90%. The remaining traffic consisted of packets that had never been seen before, which may have anomalous and/or malicious traffic. This result is promising; however, the model needs further testing, evaluation, and subjugation to anomalies.

4 CONCLUSION & FUTURE WORK

In this poster, we have presented *iDetector*, a novel framework for real-time traffic anomaly detection on home gateways. *iDetector* identifies normal behaviour and identifies traffic that may be anomalous. *iDetector* uses a fingerprint, behavioural profile, and functional blocklist to identify the traffic. We showed that *iDetector* was able to correctly identify up to 90% of IoT network traffic. Given the positive findings of the first experiment, only four devices at the time of writing, we are in the process of expanding our analysis to a much larger dataset.

REFERENCES

[1] Ross Anderson. 2008. *Security engineering.* John Wiley & Sons.
[2] Cisco. [n.d.]. *Cisco 2018 Annual Cybersecurity Report. Technical Report.*
[3] Daniel J Dubois, Roman Kolcun, Anna Maria Mandalari, Muhammad Talha Paracha, David Choffnes, and Hamed Haddadi. 2020. When Speakers Are All Ears: Characterizing Misactivations of IoT Smart Speakers. *Proceedings on Privacy Enhancing Technologies* 2020, 4 (2020), 255–276.
[4] Ayyoob Hamza, Hassan Habibi Gharakheili, Theophilus A Benson, and Vijay Sivaraman. 2019. Detecting volumetric attacks on loT devices via SDN-based monitoring of MUD activity. In *Proceedings of the 2019 ACM Symposium on SDN Research.* 36–48.
[5] Danny Yuxing Huang, Noah Apthorpe, Frank Li, Gunes Acar, and Nick Feamster. 2020. Iot inspector: Crowdsourcing labeled network traffic from smart home devices at scale. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 2 (2020), 1–21.
[6] E. Lear, R. Droms, and D. Romascanu. 2019. Manufacturer Usage Description Specification. RFC 8520 (Proposed Standard). https://doi.org/10.17487/RFC8520
[7] Anna Maria Mandalari, Roman Kolcun, Hamed Haddadi, Daniel J Dubois, and David Choffnes. 2020. Towards Automatic Identification and Blocking of Non-Critical IoT Traffic Destinations. *IEEE Security and Privacy Workshops on Technology and Consumer Protection* (2020).
[8] Jingjing Ren, Daniel J. Dubois, David Choffnes, Anna Maria Mandalari, Roman Kolcun, and Hamed Haddadi. 2019. Information Exposure for Consumer IoT Devices: A Multidimensional, Network-Informed Measurement Approach. In *Proc. of the Internet Measurement Conference (IMC).*
[9] Andrew Tannenbaum. [n.d.]. *Why Do IoT Companies Keep Building Devices with Huge Security Flaws?* https://hbr.org/2017/04/why-do-iot-companies-keep-building-devices-with-huge-security-flaws