Towards the Adaptability of Traffic-Based IoT Security Management Systems to the Device Behavior Evolutions

Chenxin Duan[1,2], Jia Li[3], Dongqi Han[1,2], Linna Fan[1,2], Shize Zhang[1,2], Jiahai Yang[1,2(✉)], and Zhiliang Wang[1,2]

[1] Institute for Network Sciences and Cyberspace, Tsinghua University, Beijing, China
{dcx19,handq19,fln19,zsz16}@mails.tsinghua.edu.cn
[2] Beijing National Research Center for Information Science and Technology, Beijing, China
{yang,wzl}@cernet.edu.cn
[3] National Computer Network Emergency Response Technical Team/Coordination Center of China, Beijing, China
lijia@cert.org.cn

Abstract. Different kinds of Internet-of-Things (IoT) devices have been widely deployed in recent years, bringing great convenience as well as security threats. Given the grim situation of IoT security, various traffic-based security management systems specially designed for IoT systems have also been developed, such as device identification systems and anomaly detection systems. A lot of such systems are trained and evaluated on datasets collected only in short time periods and lack long-term evaluation. Intuitively, the communication behaviors and traffic profile of IoT devices may keep evolving due to factors like software or firmware update and the changes of user habits. It remains to be evaluated whether these IoT security management systems can adapt well to the device behavior evolutions, which matters a lot to the real-world performance. In this paper, we give a systematic discussion about the adaptability of IoT security management systems. We summarize the factors that may cause changes on the traffic profiles of IoT devices and how they can influence the long-term performance of IoT security management systems. We hope our work can serve as a base for further study on the building of adaptive systems for the security of IoT devices.

Keywords: Internet-of-Things · Security management · Adaptability · Device behavior evolutions

1 Introduction

The proliferation of different kinds of Internet-of-Things (IoT) devices has facilitated many aspects of people's daily life, such as smart home, smart city and

This work is supported by the National Key Research and Development Program of China (No. 2017YFB0803004).

B. Chen and X. Huang (Eds.): AC3 2021, LNICST 386, pp. 203–209, 2021.
https://doi.org/10.1007/978-3-030-80851-8_15

industrial control. However, the deployment of these mini devices also poses new challenges to cyber security. For example, Mirai, a large-scale botnet which is mainly composed of compromised IoT devices, caused great damages by launching powerful distributed denial-of-service (DDoS) attacks [2]. It is also demonstrated that such IoT-based botnets are becoming more and more resilient and can even bring down the power grid [7,13]. Given the severe security threats brought by IoT devices, many systems to perform identification [6,9,10,12], monitor [5,8,14,16] and anomaly detection [11,15] to IoT devices by traffic modeling and analysis have been developed to help network managers ensure that the deployed IoT devices are functioning as expected. We refer to these systems as IoT security management systems. Considering IoT devices usually have only very simple functionalities and constrained communication and computing capacities, many IoT security management systems depend on the characterization of traffic generated by IoT devices and the construction of models that depict the normal traffic profiles of IoT devices. Such IoT security management systems are shown to achieve excellent performance and can realize the expected security management purposes.

However, for the evaluations of these IoT security management systems, both of the training and test processes are usually based on the traffic datasets collected in short time periods. Because it is impractical to collect traffic traces spanning long time periods for the training and test of these IoT security management systems in both laboratory and real-world environments. Nevertheless, many IoT devices interact with changing environments. The changes of weather or seasons will also influence the user activities and then change the way they use IoT devices. Moreover, technical issues like software or firmware updates and configurable properties may also make the traffic profiles of IoT devices change dramaticly. Thus, a problem of adaptability, whether existing IoT security management systems can keep their high performance in the long-term running, arises because the characteristics of traffic generated by the same IoT devices may keep evolving in the process of uses.

Adaptability is of great importance for practical IoT security management systems to be deployed in real-world scenarios, without which, the systems will be very fragile and the behavior evolutions of IoT devices can easily jeopardize their performance and then compromise the network management. However, it seems that previous works on IoT security have ignored the evaluation of the system adaptability to the device behavior evolutions and this property needs more attention in the future development of IoT security management systems. In this paper, we try to give a prelimilary but systematic investigation about the problem of adaptability by discussion on the following research questions (RQ):

- RQ1: How device behavior evolutions caused by different factors will affect the traffic profiles of IoT devices?
- RQ2: Can current IoT security management systems based on the traffic profiles adapt well to the device behavior evolutions?
- RQ3: What are the possible solutions to improve the adaptability of IoT security management systems or build self-adaptive systems?

We hope our work can serve as a base for further studies on the building of adaptive systems for the security of IoT devices. The remainder of this paper is organized as follows: Sect. 2 gives a review about the existing IoT security management systems based on the traffic features or profiles; Sect. 3 summarizes and categorizes different factors that may cause evolutions on the generated traffic of IoT devices; Sect. 4 discusses the possible solutions to build adaptive security management systems for IoT devices; Sect. 5 concludes the work.

2 IoT Security Management Systems

Existing IoT security management systems based on the traffic characteristics can be divided into 3 classes: device identification, event fingerprinting and intrusion detection. We will review these 3 classes of works and the traffic features they tend to use respectively.

Device Identification: A single IoT device has only very simple functionality and there are usually various different types of IoT devices in the environments equipped with IoT systems. Thus, knowing what IoT devices are connecting to the network is the first step to enable further management policies towards the IoT devices. Besides, device identification systems can also help with the discovery of unauthorized or vulnerable devices. Some works utilize identifiable fields in the traffic to recognize IoT devices, such as destination IP addresses, domain name queries and certificates [6]. Others leverage features in different resolutions, including packet level, flow level and session level, with machine learning models to classify traffic generated by different IoT devices [9,10,12]. Many device identification methods extract feature vectors in terms of traffic traces collected in short durations, ranging from minutes to hours. Both of the selected features and the durations of instances to be classified may influence the adaptability of the device identification methods. What's more, in the evaluations of these works, the training and test datasets are usually derived from traffic generated by the same devices in different periods, completely ignoring the impact of the deployment environment and user habits, which may matter much to the real-world scenarios, especially in long terms.

Event Fingerprinting: It is presented that the events happening on IoT devices that trigger the changes of their working status (trigger events) often correspond to some fixed traffic patterns, based on which, event fingerprinting systems can be developed to help network managers monitor the working status of deployed IoT devices [8,14]. With the persistent awareness of the working status of all the IoT devices, network managers can even monitor the semantic contexts of the IoT devices and detect context-relevant anomalies like event spoofing and device failure [5,16]. Unlike device identification, event fingerprinting systems tend to use very simple features, short sequences of packet lengths and directions, to detect the events happening on IoT devices by pattern matching [5,8,14,16]. Intuitively, although such simple signatures can be very precise, they are also very fragile at the same time, because the most minor update of

the firmware or changes on distinct device configuration parameters (*e.g.*, credentials and device IDs) could destroy the signatures and it will take much effort to maintain the signatures if the changes are frequent.

Intrusion Detection: Intrusion detection systems (IDS) aim to detect all kinds of attack traffic sent by malicious attackers to compromise IoT devices. Considering the simple functionalities of IoT devices, anomaly-based IDSes, which build the profile of the normal traffic generated by IoT devices and regard any traffic deviated from the profile as intrusions, become popular in the networks serving IoT devices [11,15]. However, the traffic generated by some IoT devices may change dramaticly according to the user activities. For example, surveillance cameras are silent when users are at home and generate much large-volume traffic when the users leave their houses; the using frequency of air-conditioners can vary a lot in different seasons, which means that the communication behaviors of IoT devices can drift to deviate from the known normal profiles by themselves, not just intrusions. Without taking these factors into consideration, the long-term performance of IDSes for IoT devices may be quite questionable.

3 IoT Device Behavior Evolutions

In this section, we summarize the factors that may incur evolutions in the traffic generated by IoT devices and their possible implications on the IoT security management systems. Technical issues are the most common reasons that cause IoT devices to behave differently. Regular software or firmware updates are the most common issues that change the communication patterns of IoT devices. The IP addresses of cloud servers and domain names queried by the devices, coming from content delivery networks (CDN), can change frequently with the updates and then compromise device identification systems based on these fields. In addition, many IoT devices support multiple users, configurable credentials or parameters and user-defined trigger-action rules, which can all be reflected in the traffic characteristics. And the impact of these changes on systems based on only simple traffic features, like exact packet lengths, will be catastrophic.

Besides technical issues, the ways in which users interact with the IoT devices will also often bring evolutions to device behaviors. The functionalities of most IoT devices are closely tied with people's living and producing activities, which naturally change with the seasons. The evolutions caused by human activities, such as daily routines, lifestyles and producing plans, are mainly reflected in the using frequencies and durations of different IoT devices, which can reshape many spatial-temporal characteristics of traffic generated by the devices. For intrusion or anomaly detection systems, the changes of user habits must be taken into consideration, otherwise, by regarding traffic traces collected only in some periods as the whole normal profiles, a storm of false alerts will be generated once immense changes of user habits take place.

Having said the behavior evolutions of IoT devices, there may also exist some stable traffic features, like protocols and ports used by the devices, because the hardware and core functionalities of IoT devices can hardly change after

coming into uses. However, intuitively, these relatively stable features are not distinguishable enough to achieve the goals of IoT security management. Thus, the adaptability of IoT security management systems to the device behavior evolutions should get more attention. Both of the empirical evaluations and the enhancement of the adaptability of IoT security management systems are in highly demand in the future works.

4 Possible Solutions

In this section, we briefly discuss the possible solutions to cope with the IoT device behavior evolutions. A prelimilary method is to enrich the training datasets so that the datasets can cover all the possible traffic patterns that IoT devices may exhibit as much as possible. However, on one hand, this increases the overhead to prepare the data and makes the performance of systems sensitive to the training datasets. On the other hand, this method may only apply to the evolutions caused by user activities and the changes incurred by device software or firmware updates are usually unpredictable and remain unsolved. What's more, for anomaly detection systems, current user habits are also important factors to determine whether the ongoing device behavior is abnormal. It may hinder the system performance to simply add the training data without proper contexts.

Another possible solution is to enable lifelong learning and detection for the systems, which is a hot topic in current anomaly detection community [3]. Lifelong learning means that when the system makes some mistakes, it can be trained increamentally by directly learning from the administrators' feedback to avoid making the same mistakes again. This method is a kind of remedy afterwards with lots of human intervention and cannot eliminate the performance loss caused by evolutions proactively. Additionally, it is also challenging to make the system get aware of the different contexts between the newly coming feedback and the previously learned model that should perhaps be forgotten sometimes.

Machine learning methods are widely used in existing IoT security management systems. Nevertheless, similar problems, called *concept drift*, have already been investigated in machine learning community [1,4]. Concept drift means that the statistical properties of the target variable, which the model is trying to predict, change over time in unforeseen ways. While in IoT scenarios, many security management systems are trying to fit some mapping relationships with the traffic features, which also keep changing over time due to the device behavior evolutions. Therefore, adaptive learning methods that update predictive models online during their operation to react to concept drifts, may provide inspirations to deal with the IoT device behavior evolutions. However, a lot of future work is still needed to dive deep into the regularities of IoT device behavior evolutions and combine general adaptive learning algorithms with the IoT security management objectives.

5 Conclusion

In this paper, we focus on the adaptability of IoT security management systems to device behavior evolutions. We give systematic summary and analyses on the existing IoT security management systems, different kinds of evolutions happening on IoT devices and the possible solutions to build adaptive systems for the security of IoT devices. We find that the adaptability of IoT security management systems did not get enough attention despite its great importance and we hope our work can serve as the base of further study on adaptive IoT security management systems.

References

1. CADE: Detecting and explaining concept drift samples for security applications. In: 30th USENIX Security Symposium (USENIX Security 2021). USENIX Association, Vancouver, B.C. (2021). https://www.usenix.org/conference/usenixsecurity21/presentation/yang
2. Antonakakis, M., et al.: Understanding the mirai botnet. In: 26th USENIX security symposium (USENIX Security 2017), pp. 1093–1110 (2017)
3. Du, M., Chen, Z., Liu, C., Oak, R., Song, D.: Lifelong anomaly detection through unlearning. In: Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security, pp. 1283–1297 (2019)
4. Gama, J.A., Žliobaitundefined, I., Bifet, A., Pechenizkiy, M., Bouchachia, A.: A survey on concept drift adaptation. ACM Comput. Surv. **46**(4) (2014). https://doi.org/10.1145/2523813
5. Gu, T., Fang, Z., Abhishek, A., Fu, H., Hu, P., Mohapatra, P.: IoTGaze: IoT security enforcement via wireless context analysis. In: IEEE INFOCOM 2020-IEEE Conference on Computer Communications, pp. 884–893. IEEE (2020)
6. Guo, H., Heidemann, J.: Detecting IoT devices in the internet. IEEE/ACM Trans. Netw. **28**(5), 2323–2336 (2020)
7. Herwig, S., Harvey, K., Hughey, G., Roberts, R., Levin, D.: Measurement and analysis of Hajime, a peer-to-peer IoT botnet. In: NDSS (2019)
8. Junges, P., François, J., Festor, O.: Passive inference of user actions through IoT gateway encrypted traffic analysis. In: 2019 IFIP/IEEE Symposium on Integrated Network and Service Management (IM), pp. 7–12 (2019)
9. Ma, X., Qu, J., Li, J., Lui, J.C., Li, Z., Guan, X.: Pinpointing hidden IoT devices via spatial-temporal traffic fingerprinting. In: IEEE INFOCOM 2020-IEEE Conference on Computer Communications, pp. 894–903. IEEE (2020)
10. Marchal, S., Miettinen, M., Nguyen, T.D., Sadeghi, A.R., Asokan, N.: AuDI: toward autonomous IoT device-type identification using periodic communication. IEEE J. Sel. Areas Commun. **37**(6), 1402–1412 (2019)
11. Mirsky, Y., Doitshman, T., Elovici, Y., Shabtai, A.: Kitsune: an ensemble of autoencoders for online network intrusion detection. In: Network and Distributed System Security Symposium, NDSS (2018)
12. Sivanathan, A., et al.: Classifying IoT devices in smart environments using network traffic characteristics. IEEE Trans. Mob. Comput. **18**(8), 1745–1759 (2018)
13. Soltan, S., Mittal, P., Poor, H.V.: BlackIoT: IoT botnet of high wattage devices can disrupt the power grid. In: 27th USENIX Security Symposium (USENIX Security 2018), pp. 15–32 (2018)

14. Trimananda, R., Varmarken, J., Markopoulou, A., Demsky, B.: Packet-level signatures for smart home devices. In: Network and Distributed System Security Symposium, NDSS (2020)
15. Wan, Y., Xu, K., Xue, G., Wang, F.: IoTArgos: a multi-layer security monitoring system for internet-of-things in smart homes. In: IEEE INFOCOM 2020-IEEE Conference on Computer Communications, pp. 874–883. IEEE (2020)
16. Zhang, W., Meng, Y., Liu, Y., Zhang, X., Zhang, Y., Zhu, H.: HoMonit: monitoring smart home apps from encrypted traffic. In: Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, pp. 1074–1088 (2018)