Situational Awareness: Detecting Critical Dependencies and Devices in a Network

Martin Laštovička(✉) and Pavel Čeleda

Institute of Computer Science and Faculty of Informatics,
Masaryk University, Brno, Czech Republic
lastovicka@ics.muni.cz, celeda@ics.muni.cz

Abstract. Large-scale networks consisting of thousands of connected devices are like a living organism, constantly changing and evolving. It is very difficult for a human administrator to orient in such environment and to react to emerging security threats. With such motivation, this PhD proposal aims to find new methods for automatic identification of devices, the services they provide, their dependencies and importance. The main focus of the proposal is to find novel approaches to building cyber situational awareness in an unknown network for the purpose of computer security incident response. Our research is at the initial phase and will contribute to a PhD thesis in four years.

Keywords: Situational awareness · Cybersecurity · Device importance evaluation · Threat impact estimation · Graph theory · Network monitoring

1 Introduction

The impacts of cyber threats became more serious with organisations increasing dependency on computer infrastructure. To defend against such threats, system administrators must build situational awareness which allows them to understand and orient in the complex networks [6]. The aim of this PhD thesis is to find new ways to automatically build situational awareness to help administrators understand possible impacts of a cyber threat.

Situational awareness means the knowledge and understanding of the current situation. It is possible for a system administrator to know what is going on in a small network, but with the growing number of connected devices, this becomes more and more difficult. A basic solution is to manually create a list of all devices in the network. But it is impossible to maintain such list throughout time and keep it updated with the dynamic changes of the network. Moreover, the trend of nowadays networks, containing mobile devices or IoT (Internet of Things), and cloud environments goes directly against the idea of device list and makes it useless in practice An automated approach is needed to deal with the constantly changing environment [6].

D. Tuncer et al. (Eds.): AIMS 2017, LNCS 10356, pp. 173–178, 2017.
DOI: 10.1007/978-3-319-60774-0_17

The current approach for device and service identification focus on very specific networks, e.g., industrial control systems, or selected subset of services [1,2] which reduces their value in modern networks described above. Manual evaluation by security expert is still prevalent in the field of dependency detection and importance estimation. These risk assessment methods are not automated [5] or need active cooperation of the devices [8].

In our work, we intend to find new methods of building situational awareness based on data from network monitoring that will not depend on a specific type of network. We will define a computer network model containing information about devices and services, their dependencies and importance for the organisation. The importance of a device can then be expressed as how the device outage or compromise would impact other devices and goals of the host organisation. The nature of continuous information gathering from the network also overcomes the ever-changing nature of large networks and allows us to evaluate the data throughout time.

2 Research Questions

This research aims to discover new ways of threat impact estimation with respect to current situation, devices and services. To achieve this goal we attempt to answer following research questions:

1. **How can device and its services be identified in a complex network using passive network monitoring?**
   Many devices are not willing (end-user devices) or not able to (IoT) provide information about themselves in large networks. But every device communication over network could be analysed [3] and used to identify the type of the device, its operating system and provided services. However, current trends in modern networks, e.g., encrypted communication, port obfuscation, high transfer rate, make such identification hard. We plan to investigate those issues and propose methods to handle them.
2. **How can device dependencies be detected in a network?**
   To understand the situation in a network, it is not enough to know only what a device is and what services it provides. It is important to know which devices it depends on and how many devices depend on it. To answer this research question, we will study relationships between devices in internal network and propose new methods for their detection from network monitoring data.
3. **How can device importance be estimated from the perspective of reaction to cyber threats?**
   The importance of a particular device for organisation mission differs according to the provided services and the number of clients depending on the device. We plan to take these factors into account to build a model for importance estimation and we will find new ways of automatic importance evaluation based on traffic monitoring.

3 Proposed Approach

Our first step towards the building of situational awareness will be the definition of a network model. The natural representation of a computer network is a graph, where each node stands for a device in the network. Edges between nodes represent device communication, while another type of edge can represent dependencies or the presence of a cyber threat. This model allows us to separate the mostly static nature of what device is from its dynamic behaviour on the network.

3.1 Identification of Devices and Services

The knowledge of what a device is and what services it provides is a fundamental part of understanding the network. The goal of this part is to research methods of processing network traffic data to identify the type of the device (server, workstation, mobile, IoT), its operating system (Windows, Linux) and its services (web, mail, database).

Easiest way to determine a device type and services is to simply ask it. To do it in an organised way, many Service Discovery Protocols have been implemented [7] and deployed. They build a directory of all devices and their services, as an example, we can name well-known protocols such as BitTorrent or UPnP. However, this approach require active cooperation of the devices and hence we will not focus on them. Another way is to use active scanning. Our plan is to focus on passive methods only, yet we can use outputs of network scanning projects, e.g., Shodan, Censys, as a verification or an enhancement of our methods.

To achieve passive classification described above a sophisticated method must be used. Simple methods using protocol and port numbers currently falls short in classifying services with a dynamic port assignment or port obfuscation, e.g., hiding behind TCP port 80 [10]. To overcome these issues, more characteristics need to be taken into consideration.

The current trend of traffic encryption makes the analysis of its content hard, but on the other hand, it opens new ways of host identification. A client needs to send a lot of data to establish encrypted communication. For example, supported ciphersuites can be used to identify communicating clients during TLS (Transport Layer Security) handshake [4]. Similarly, we plan to investigate other properties of encrypted communication to identify the client device.

The most promising service identification method nowadays is the use of machine learning algorithms to classify the network traffic. Current methods perform well in a controlled environment where every application is known in advance, but cannot efficiently handle unknown traffic. Zhang et al. [11] presented an iterative method to improve identification accuracy, yet this field is still not fully explored. The two challenges we plan to address are the accuracy of identification in real network and performance of such algorithms when processing large amounts of data continuously coming from the monitored network.

3.2 Dependency Detection and Importance Estimation of a Device

The problem of asset criticality evaluation is known as vital for proper decision making during cyber-attacks but is difficult to achieve [5]. Research in this area is mainly focused on finding ways how a group of security experts can determine criticality by following prepared guidelines just like in risk assessment. But this approach is very time-consuming and cannot be repeated very often which leads to the data being outdated.

On the contrary, automatic evaluation is able to run continuously and can provide results when needed. We are aware that some important services or dependencies can be discovered only during exceptional operations or back-up servers become active only after failure of the main one. Automatic detections can still provide good staring point for risk assessment and save resources. Moreover, automatic system can identify operations that the administrators do not know about as presented in [9]. We propose three components to combine in order to estimate the device importance:

1. **Traffic Statistics** – Analysis of ongoing traffic in the network can point out the most used services in the terms of connected clients and data transfer volume. We will link these volumetric statistics to identified services to give them the dynamic context for importance evaluation, e.g., heavily loaded web server will be set as more important than another one scarcely visited. Our research will focus on real-time statistics computations so that it will be possible to dynamically adjust the evaluation as the network usage changes in time.
2. **Dependency detection** – Based on the identification of device type and traffic statistics, the basic dependency between client and server will be modelled. Using graph centrality algorithms we can then estimate the servers importance and the impact of its outage as the number of affected clients weighted by their own criticality. More complex dependencies can be discovered by clique detection. Dependencies forming a clique between servers can indicate strong relationship and exploitation of one will affect the whole group. The first steps towards automatic dependency detection using graph algorithms were made in [9], but they rely on active probing (i.e., Nagios system) to discover effects of service failure and backup detections, whereas we plan to achieve the same with passive network monitoring.
3. **Attacks Statistics** – Network attack is a manifestation of a cyber threat. The understanding of attack targets and discovery of most attacked devices should lead to raising the protection level of those devices. Our assumption is that parts of critical infrastructure will be targeted by attackers more often than user stations. Moreover, the type of the attack should differ and these differences could help to identify the most important devices. However, such assumption needs to be carefully verified before using in the criticality calculations. For example, attackers could target the most vulnerable device instead of critical infrastructure. In that case, such observation should be used as an advisory for the administrator rather than for criticality estimation.

4 Conclusion

In this research, we focus on building situational awareness from passive network observation without the necessity of active device probing. From those data, we intend to determine what a device is, what services it provides, what are its dependencies and how important it is for the network. Our methods will evaluate the situation continuously in order to follow changes in network and will be designed to be autonomic to minimise the need for human administrator assistance. Achieving our goals will help system administrators to better understand the situation in their network and to perceive the possible impacts of cyber threats.

Acknowledgement. This research was supported by the Security Research Programme of the Czech Republic 2015–2020 (BV III/1 VS) granted by the Ministry of the Interior of the Czech Republic under No. VI20172020070 Research of Tools for Cyber Situation Awareness and Decision Support of CSIRT Teams in the Protection of Critical Infrastructure.

Martin Laštovička is Brno Ph.D. Talent Scholarship Holder – Funded by the Brno City Municipality.

References

1. Callado, A., Kamienski, C., Szabó, G., Gero, B.P., Kelner, J., Fernandes, S., Sadok, D.: A survey on internet traffic identification. IEEE Commun. Surv. Tutorials **11**(3), 37–52 (2009)
2. Franke, U., Brynielsson, J.: Cyber situational awareness - a systematic review of the literature. Comput. Secur. **46**, 18–31 (2014)
3. Hofstede, R., Čeleda, P., Trammell, B., Drago, I., Sadre, R., Sperotto, A., Pras, A.: Flow monitoring explained: from packet capture to data analysis with NetFlow and IPFIX. IEEE Commun. Surv. Tutorials **16**(4), 2037–2064 (2014, Fourthquarter)
4. Husák, M., Čermák, M., Jirsík, T., Čeleda, P.: HTTPS traffic analysis and client identification using passive SSL/TLS fingerprinting. EURASIP J. Inf. Secur. **2016**(6), 1–14 (2016)
5. Kim, A., Kang, M.H.: Determining asset criticality for cyber defense. Technical report, Naval Research Lab, Washington DC (2011)
6. Kott, A., Wang, C., Erbacher, R.F.: Cyber Defense and Situational Awareness. Springer, Heidelberg (2014). ISBN: 978-3-319-11390-6
7. Meshkova, E., Riihijärvi, J., Petrova, M., Mähönen, P.: A survey on resource discovery mechanisms, peer-to-peer and service discovery frameworks. Comput. Netw. **52**(11), 2097–2128 (2008)
8. Weintraub, E., Cohen, Y.: Continuous monitoring system based on systems's environment. In: Proceedings of the Conference on Digital Forensics, Security and Law, p. 151. Association of Digital Forensics, Security and Law (2015)
9. Zand, A., Houmansadr, A., Vigna, G., Kemmerer, R., Kruegel, C.: Know your achilles' heel: automatic detection of network critical services. In: Proceedings of the 31st Annual Computer Security Applications Conference, pp. 41–50. ACM (2015)

10. Zander, S., Nguyen, T., Armitage, G.: Automated traffic classification and application identification using machine learning. In: The IEEE Conference on Local Computer Networks 30th Anniversary (LCN 2005), pp. 250–257. IEEE (2005)
11. Zhang, J., Chen, C., Xiang, Y., Zhou, W.: Robust network traffic identification with unknown applications. In: Proceedings of the 8th ACM SIGSAC Symposium on Information, Computer and Communications Security, pp. 405–414. ACM (2013)