A taxonomy of malicious traffic for intrusion detection systems

Hanan Hindy
Elike Hodo
Ethan Bayne
Amar Seeam
Robert Atkinson
Xavier Bellekens

This is the accepted version of a paper presented at the Cyber Science 2018: Security, Safety and Survivability in an era of constant, contemporary and complex Physical and Cyber Attacks, 11-12 June 2018, Glasgow, UK which will be published by IEEE

A Taxonomy of Malicious Traffic for Intrusion Detection Systems

Hanan Hindy*, Elike Hodo‡ Ethan Bayne*,
Amar Seeam†, Robert Atkinson‡ and Xavier Bellekens*
*Division of Cyber Security, Abertay University, Dundee Scotland
†Department of Computer Science, Middlesex University (Mauritius)
‡EEE Department, University of Strathclyde, Glasgow, Scotland

***Abstract*—With the increasing number of network threats it is essential to have a knowledge of existing and new network threats in order to design better intrusion detection systems. In this paper we propose a taxonomy for classifying network attacks in a consistent way, allowing security researchers to focus their efforts on creating accurate intrusion detection systems and targeted datasets.**

I. Introduction

Cyber-security is defined as the discipline concerned with protecting networks, computer devices, programs and data from different forms of attacks. Cyber-attacks could cause data loss, allow attackers to access confidential information or affect system or service availability. Research in this domain focuses on detecting the attacks and on preventing and predicting them. Although research in this field started in the early fifties, it is an evolving domain [1].

The importance of research in this domain grow with the increasing prevalence of Internet of Things (IoT) systems to aggregate, transfer and send data to and from sensors. It is predicted by CISCO that there will be 50 billion devices connected to the Internet by 2020 [2].

Attacks are becoming more complex, targeting a wide range of devices (industrial, personal, etc) [3]. Some attacks focus on obtaining information without causing any damage to a system, whilst some attacks cause damage either by manipulating information or masquerading the attack to access privileges maintain an access. Moreover, users are becoming more aware of attacks and as a result systems are now designed to be more secure [4]. Therefore, the current iteration of strategic analysis tools and methods will no more fit the need for threats detection and prevention. The taxonomy of malicious traffic presented in this paper can help researchers and software engineers to design up-to-date detection tools and to find ways to prevent and predict these attacks.

The remainder of the paper is organised as follows, Section II highlights the need for new taxonomy focused on network threats. Section III provides an insight on related work, while in Section IV the taxonomy is described and analysed, finally, the paper summarises with the conclusion in Section V.

II. Problem Statement

Current rule based intrusion detection system only consider a subset of known attacks to defend large enterprise networks. Rule based intrusion detection systems rely highly on prior detection of attacks and regularly updated rules [5]. While these systems provide a first step to security, they do not enable the detection of unknown attacks. Machine learning intrusion detection systems on the other hand allow the detection of unknown threats, however these system rely heavily on existing training datasets, that may be outdated and leave out a number of recent threats [6]. This problem underpins the need for a taxonomy of attacks allowing researchers and engineers to build better datasets [7].

III. Related Work

The increasing number of threats has led to advances in cyber-security. These systems however have numerous drawbacks leading to systems being compromised. To this end, researchers have published a number of taxonomies with the aim to increase the overall efficacy of threat detection systems.

Zhu *et al.* [8] provide a taxonomy detailing the different flaws and attacks against industrial SCADA systems. The attacks are classified using the TCP/IP stack. Chakrabarti *et al.* highlight the threats the Internet infrastructure is facing [9] the attacks are illustrated through different scenarios. Hoque *et al.* [10] provides a taxonomy of tools and systems against network attacks, detailing the numerous tools used to identify network flaws. The tools are described and then classified with their pros and cons described.

While these taxonomies provide information on network attacks, they focus on specific systems and tools. The taxonomy presented in this paper focuses primarily on threats that can be detected via an intrusion detection systems and is aimed at researchers building datasets to ensure that relevant attacks are included, hence, increasing the efficiency of future intrusion detection systems that incorporate machine learning.

IV. Taxonomy of Malicious Traffic

The aim of this manuscript is to help researchers and engineers to develop techniques to detect current and new malicious traffic occurring on the network.The taxonomy

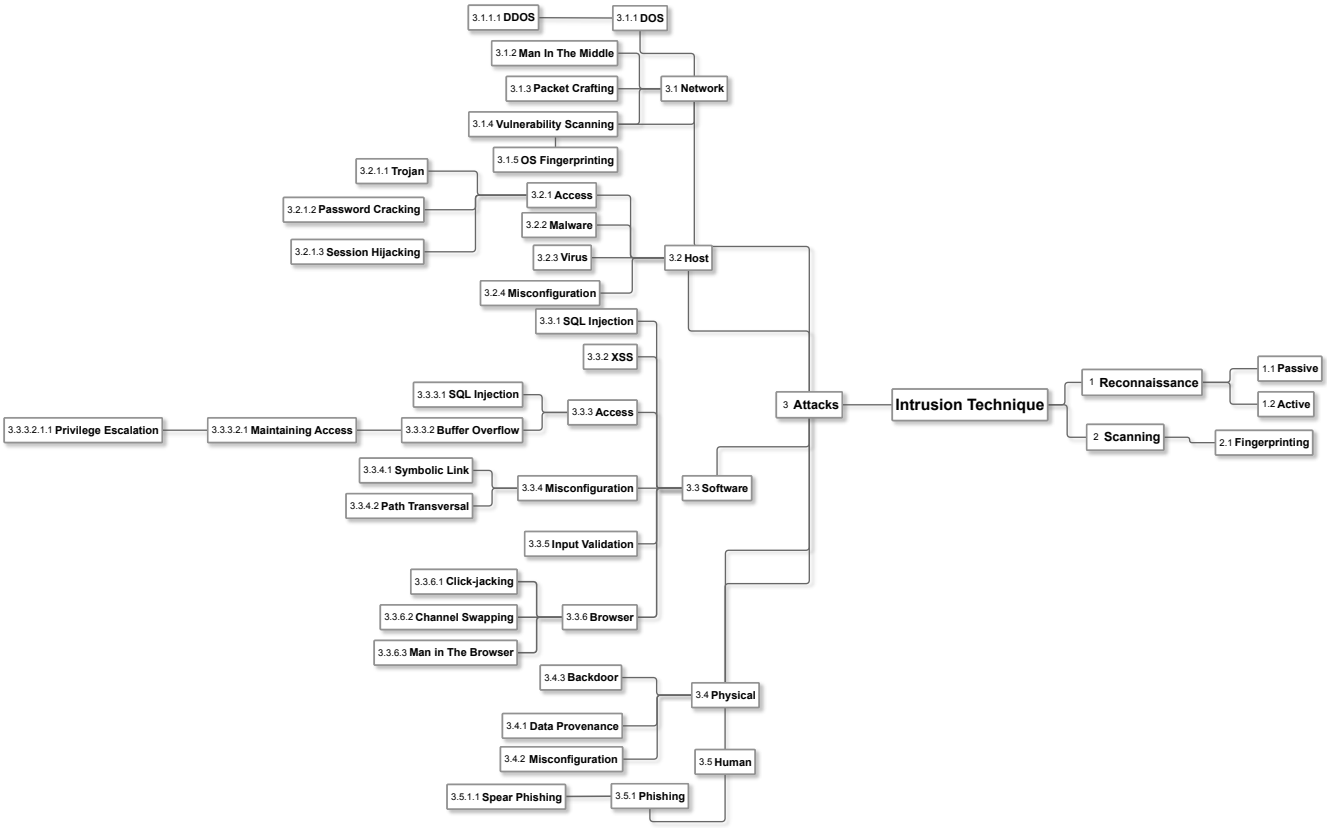

Fig. 1. Malicious Traffic Taxonomy

presented in this work is composed of three control stages (CS) representing attack steps taken by malicious users. Each stage is then further composed of sub-stages that can be either executed in parallel or subsequently during an attack. An example of multiple sub-steps being subsequently being executed is a malware taking advantage of a physical host executing a buffer overflow in order to provide access to a malicious user.

The taxonomy provided is designed to provide the end-user with singular sub-components that can be added together to represent a more complex attack.

Control Stage 1 : Reconnaissance The reconnaissance stage is often the initial and the most important stage [11]. It is used by malicious users to gather data on the target. The data gathered by the reconnaissance inform both CS2 and CS3. This stage enables the malicious user to learn critical information about the target, through various passive and active reconnaissance and obtain specific data such as

- IP Address Range
- DNS Records
- Mail Servers

Control Stage 2: Scanning Malicious users can gather critical information about the network target by mapping the system and the network. Important network components such as firewalls, routers, etc, can be discovered. This task is often realised by using a port scanner. The port scanner is designed to try and target opened ports. Scanning, is an active step, to identify network services running on a host; it also allows the malicious user to test the network and the firewall for the different security policies in place. This stage informs CS3.

Control Stage 3 : Attacks Whilst CS1 and CS2 can be operated with relative stealth on the network, the attack is essentially critical for the malicious user as both CS1 and CS2 inform the choice of target (e.g. which server represents the best target), which of exploit should be used to obtain the best result, etc. Figure 1 provides a taxonomy of malicious attacks that can be detected by analysing network traffic using an intrusion detection system or deep packet inspection. The attacks are classified in five different subsections described hereafter;

TABLE I
PORT SCANNING RESPONSES

Category	Description
Opened	The target responded on a that the port, implying that a particular service is listening
Closed	The target responded that all connections are denied on that port
Filtered / Blocked	the target did not reply

A. Network

Network nodes are vulnerable to a large range of attacks. Figure 1 (3.1) depicts common attacks on a network ecosystem. DoS and DDoS attacks (3.1.1) and (3.1.1.1) can be characterised by the large number of packets or requests received by target with an intent to render it's main function unusable (e.g. Sending a high number of HTTP request on a web server in an attempt to overwhelm the server and discard legitimate connections) [12]. The Man in The Middle attack (3.1.2) (e.g. ARP flooding attack). This attack attempts to map the wrong MAC address with an IP address, allowing the malicious user to redirect incoming and outgoing traffic and snoop on the network conversation [13]. The packet crafting attacks (3.1.3) highlights the possibility for an attacker to replay or craft a new packet to bypass a firewall, test the TCP/IP stack of network components, or replay a packet to gain access to a network component [14]. The vulnerability Scanning (3.1.4) is composed of all the methods used by malicious attackers to scan a network for vulnerabilities with known scanners such as Nessus, NMAP, OpenVAS, MBSA etc. Vulnerability scanners can allow malicious user to detect vulnerabilities and how to exploit them [15].

B. Host

As shown in Figure 1 (3.2), Host based threats encompass infection of a host or its access through malicious intent. Gaining Access (3.2.1) describes multiple methods to gain illegal access using trojans (3.2.1.1) or backdoors for remote access [16], password cracking (3.2.1.2) using tools such as john the ripper to gain access to the target [17], or session hijacking (3.2.1.3) to access the session of a legitimate user using a replay attack [18]. Host based (3.2) attacks also describe malware attacks (3.2.2) describing adware, sypware and worms, as well as viruses (3.2.3) which is in itself a malware, however has the ability to infect other hosts on the network during execution.

C. Software

Figure 1 provides an overview of software based attacks (3.3) that can be detected in network traffic. (3.3.3.1) SQL Injection are commonly used to maliciously acquire data from a database [19], (3.3.3.2) Buffer Overflow can be used to access restricted data in memory, they can also be used to gain root access on a computer, or for privilege escalation (3.3.3.2.1) in order to maintain access on the target host [20]. Software can also be subject to misconfiguration such as the symlink attack (3.3.4.1) on web servers or path transversal attacks (3.3.4.2) that are used to access folders and files outside the web root folder. Software can also be vulnerable to input validation (3.3.5) attacks [21]. This attack requires the malicious user to test different input fields to obtain error messages and gain access either to a system, or to execute code to obtain data.

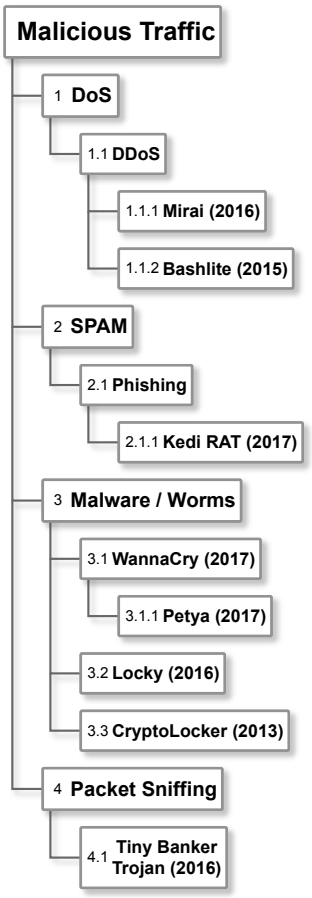

Fig. 2. Malicious Traffic Including Exploit and Worm Names

D. Physical

Network nodes are vulnerable to large range of physical attacks, a number of these attacks are highlighted in Figure 1 (3.4). Physical attacks include physical back-doors added in foundries that can be enabled remotely or through a specific combination of events (3.4.3) [22]. Back-doors can also enable remote-access hence be detected in network traffic. The data obtained through the sensor should also be verified through data provenance (3.4.1), a malicious user could spoof sensors data, making end-users and operators believe in an event which is not occurring [23]. Physical devices can also be prone to misconfiguration leading to remote access, and data leaks (3.4.2).

E. Human

The Human factor also plays a decisive role in network security, without appropriate training and awareness [24]. Humans can also be seen as the weak link, hence, when designing intrusion detection systems it is also important to identify attacks against the end-users and expert-users [25]. As shown in Figure 1 (3.5.1) phishing and spear phishing attacks. These attacks are designed to trick the user to provide

credentials or to allow a malicious user to access data without its consent. Some of these attacks may be targeted to have a higher probability of success.

F. Attack Examples

This paper, aims at providing a simple and concise taxonomy of network attacks, and to this end, Figure 2 provides an example or recent attacks classified through our malicious taxonomy. These attacks include the Mirai botnet (1.1.1) that infected numerous IoT devices to launch distributed denial of service attacks against numerous Internet provided and services [26]. It also includes the Tiny Banker Trojan (4.1), which was active from 2012 to 2016 targeting financial institutions stealing data from the users [27]. Whilst these attacks are reported in the media, they are often excluded from training datasets.

V. Conclusion

In this paper a taxonomy of network threats for intrusion detection systems is presented. The taxonomy is divided into three control stages in order to describe more complex attack processes. The aim of this work is to create a taxonomy with the ability to inform researchers developing both intrusion detection systems and training datasets in order to increase the detection accuracy and decrease the false positive rate. With the increasing number of connected systems and networks the taxonomy aims at facilitating the design of future defense mechanisms as well as robust systems.

References

[1] E. Tikk-Ringas, *Evolution of the Cyber Domain: The Implications for National and Global Security*. Routledge, for the International Institute for Strategic Studies, 2015.

[2] D. Evans, "The internet of things: How the next evolution of the internet is changing everything," *CISCO white paper*, vol. 1, no. 2011, pp. 1–11, 2011.

[3] X. Bellekens, A. Seeam, K. Nieradzinska, C. Tachtatzis, A. Cleary, R. Atkinson, and I. Andonovic, "Cyber-physical-security model for safety-critical iot infrastructures," in *Wireless World Research Forum Meeting*, vol. 35, 2015.

[4] K. M. Krister, "Automated analyses of malicious code," Master's thesis, Institutt for datateknikk og informasjonsvitenskap, 2009.

[5] X. J. Bellekens, C. Tachtatzis, R. C. Atkinson, C. Renfrew, and T. Kirkham, "Glop: Enabling massively parallel incident response through gpu log processing," in *Proceedings of the 7th International Conference on Security of Information and Networks*, p. 295, ACM, 2014.

[6] E. Hodo, X. Bellekens, E. Iorkyase, A. Hamilton, C. Tachtatzis, and R. Atkinson, "Machine learning approach for detection of nontor traffic," *Journal of Cyber Security and Mobility*, vol. 6, no. 2, pp. 171–194, 2017.

[7] E. Hodo, X. Bellekens, A. Hamilton, C. Tachtatzis, and R. Atkinson, "Shallow and deep networks intrusion detection system: A taxonomy and survey," *arXiv preprint arXiv:1701.02145*, 2017.

[8] B. Zhu, A. Joseph, and S. Sastry, "A taxonomy of cyber attacks on scada systems," in *Internet of things (iThings/CPSCom), 2011 international conference on and 4th international conference on cyber, physical and social computing*, pp. 380–388, IEEE, 2011.

[9] A. Chakrabarti and G. Manimaran, "Internet infrastructure security: A taxonomy," *IEEE network*, vol. 16, no. 6, pp. 13–21, 2002.

[10] N. Hoque, M. H. Bhuyan, R. C. Baishya, D. Bhattacharyya, and J. K. Kalita, "Network attacks: Taxonomy, tools and systems," *Journal of Network and Computer Applications*, vol. 40, pp. 307–324, 2014.

[11] T. Wrightson, *Advanced persistent threat hacking: the art and science of hacking any organization*. McGraw-Hill Education Group, 2014.

[12] J. Mirkovic and P. Reiher, "A taxonomy of ddos attack and ddos defense mechanisms," *ACM SIGCOMM Computer Communication Review*, vol. 34, no. 2, pp. 39–53, 2004.

[13] F. Callegati, W. Cerroni, and M. Ramilli, "Man-in-the-middle attack to the https protocol," *IEEE Security & Privacy*, vol. 7, no. 1, pp. 78–81, 2009.

[14] D. Papp, Z. Ma, and L. Buttyan, "Embedded systems security: Threats, vulnerabilities, and attack taxonomy," in *Privacy, Security and Trust (PST), 2015 13th Annual Conference on*, pp. 145–152, IEEE, 2015.

[15] H. Holm, "Performance of automated network vulnerability scanning at remediating security issues," *Computers & Security*, vol. 31, no. 2, pp. 164–175, 2012.

[16] T. Boraten and A. Kodi, "Mitigation of hardware trojan based denial-of-service attack for secure nocs," *Journal of Parallel and Distributed Computing*, vol. 111, pp. 24–38, 2018.

[17] R. Mahey, N. Singh, C. Kumar, N. Bhagwat, and P. Verma, "Graphical password using an intuitive approach," in *International Conference on Intelligent Computing and Applications*, pp. 153–161, Springer, 2018.

[18] H. Shahriar and H. M. Haddad, "Fuzzy rule-based vulnerability assessment framework for web applications," in *Application Development and Design: Concepts, Methodologies, Tools, and Applications*, pp. 778–797, IGI Global, 2018.

[19] W. G. Halfond, J. Viegas, and A. Orso, "A classification of sql-injection attacks and countermeasures," in *Proceedings of the IEEE International Symposium on Secure Software Engineering*, vol. 1, pp. 13–15, IEEE, 2006.

[20] A. Pasupulati, J. Coit, K. Levitt, S. F. Wu, S. Li, J. Kuo, and K.-P. Fan, "Buttercup: On network-based detection of polymorphic buffer overflow vulnerabilities," in *Network Operations and Management Symposium, 2004. NOMS 2004. IEEE/IFIP*, vol. 1, pp. 235–248, IEEE, 2004.

[21] S. Simmons, D. Edwards, N. Wilde, J. Just, and M. Satyanarayana, "Preventing unauthorized islanding: cyber-threat analysis," in *System of Systems Engineering, 2006 IEEE/SMC International Conference on*, pp. 5–pp, IEEE, 2006.

[22] F. Koushanfar, "Trusting the open latent ic backdoors," in *Proceedings of the sixth ACM workshop on Scalable trusted computing*, pp. 1–2, ACM, 2011.

[23] K. Xu, H. Xiong, C. Wu, D. Stefan, and D. Yao, "Data-provenance verification for secure hosts," *IEEE Transactions on Dependable and Secure Computing*, vol. 9, no. 2, pp. 173–183, 2012.

[24] X. Bellekens, K. Nieradzinska, A. Bellekens, P. Seeam, A. Hamilton, and A. Seeam, "A study on situational awareness security and privacy of wearable health monitoring devices," *Int. J. Cyber Situat. Aware*, vol. 1, pp. 1–25, 2016.

[25] P. Legg, "Visual analytics for non-expert users in cyber situation awareness," *International Journal on Cyber Situational Awareness*, vol. 1, no. 1, 2016.

[26] C. Kolias, G. Kambourakis, A. Stavrou, and J. Voas, "Ddos in the iot: Mirai and other botnets," *Computer*, vol. 50, no. 7, pp. 80–84, 2017.

[27] P.-F. Marteau, "Sequence covering for efficient host-based intrusion detection," *arXiv preprint arXiv:1712.02084*, 2017.