

Machine Learning and APTs

Pedro Ramos Brandão

Full Professor – ISTECS Lisbon

Gabriel Pereira Matos

Computer Science MSc Student

I. Abstract

APTs, also known as Advanced Persistent Threats, are a type of cyberattack characterized by slow and stealthy methods of attack. As one of the most worrying attack methods today, it's important to understand what they are and how they work. At the moment, there are already some techniques for detecting APTs through the training and learning method known as Machine Learning. This article introduces the definitions of APTs and machine learning clarifies the operation of APTs, and introduces and discusses some techniques for APTs detection.

Keywords: Advanced Persistent Threats, Cybersecurity, Machine Learning

II. Introduction

The security of the armed forces and well-known organizations has become a priority for all organizations due to the strong emphasis given to information security by information security investigators throughout the world. But this isn't enough, for we are introduced daily to new forms of malware, and new means of attack. To showcase their abilities, attackers target organizations for financial gain or, simply, to damage the organization's reputation. They don't tend to hide their actions in the course of all the attacks. Some attacks are characterized by their slow and hidden movement because the attacker intends to steal data without getting caught. These attacks are known as Advanced Persistent Threats (APT). The attackers that use such means may use familiar methods of

entering the network of the target, but the tools at their disposal are not common. As the name, APT, suggests, those tools need to be advanced to allow the attacker to be more persistent for long periods of time. They stay hidden, slowly expanding their foothold, from one system to another inside the network of the organization, obtaining useful information as they move and export said information to their command and control center in a strategic manner. The APTs are typically executed by well-funded attackers, with the necessary resources to execute the attack, as the funding entity requests it. The attack is only terminated when it is detected or when the funding entity acquires the data it needs. Regardless of the outcome, the damage would be substantial and, sometimes, beyond repair, which is more common in the following case, when the APT attack isn't detected until all the data of the organization has fallen into the wrong hands.

The only question that the victim of an APT attack has is, why it wasn't able to detect the attack.

These types of threats are normally invisible to a firewall or an ordinary antivirus. To ensure that the system has a high-security level mechanism, we can fight these threats using the Machine Learning (ML) method. Yet, it's required that the ML algorithms used in the detection of malicious content in the systems or networks are updated. If the ML is taught and trained correctly, it will be able to detect and predict several threats in a short period of time, with a small rate of false positives and a big rate of precision. The objective of this article is to

discover if it is possible to utilize ML to detect and fight APT attacks.

III. Methodology

To answer the question, “Is it possible to detect and fight APTs with the assistance of Machine Learning?” Several articles about Machine Learning, cyberattacks, and information security risk were reviewed in a literature review. Not being well versed in the topics was the reason for the preference in obtaining the most clarifying articles through the usage of a search engine, in this particular case, Google Scholar. This article is the result of that search and acquisition of knowledge. The article is structured to introduce APTs, what these are, how they operate, and their phases of operation. A brief description of Machine Learning is also introduced and after, it's worked out and discussed some existing techniques given the abilities of Machine Learning in detecting and fighting Advanced Persistent Threats (APTs).

IV. APTs

A. What are APTs?

Advanced Persistent Threats, better known as APT, according to the following authors Alshamrani et al. [1], are different attacks made by the typical hacker. APTs are normally made by groups of experienced and specialized hackers, funded by organizations and governments. Initially, a military term that refers to attacks against the United Nations (UN), it was adapted to the context of information security. APTs are defined by the combination of three words, which are Advanced, Persistent, and Threats. More specifically:

Advanced: the attackers that use APTs, can develop limitless advanced tools through the combination of several offensive strategies and attack in different stages.

Persistent: the attackers that use APTs are extremely persistent to achieve their goals, and for that, they plan evasive techniques to avoid being detected by security systems and digital security specialists. This is normally done through a persistent strategy with the mindset and approach that the APTs are “slow, possess low immediate impact, and are hidden”.

Threat: the attackers that use APTs focus precisely on a specific organization to achieve their goal. Normally, they have the

potential to jeopardize an information system through destruction, exposure, modification of data, and/or denial of service. The attacker's objective normally involves obtaining access to the information of the attack's target, filtering all the information about the target that was removed from the compromised systems. Seeing as the attack can access the target's system it must be “behind the scenes”, which means in a concealed fashion, in order to not be discovered and be able to retrieve information that the attackers are looking for, and then send to the entity that funded the attack. The APT attacks normally involve several compromised knots, not just one, in contrast to the regular attacks. According to the National Institute of Standards and Technologies of the United States of America [2], a group of attackers that uses APTs: pursues its goals repetitively throughout a long period of time; it adapts to the efforts coming from the defending side that resists the attack; is committed to guaranteeing that the level of necessary interaction to achieve its goal is kept. With consideration of these characteristics of the APTs, it's hard to detect and analyze attackers capable of “cheating” the existing security systems.

Aforementioned, the attackers have to go through numerous steps to achieve their goals. These stages involve establishing a foothold in the target's network, exploring the internal network to search for system flaws, and laterally moving from one system to another until the main target system is reached. Given that you are in the network, it's possible to obtain more privileges to achieve the targeted system that contains the sensitive information and, finally, retrieve and send the information to the attacker's command and control center.

After a filtration process of the data, the attackers can decide if they continue to retrieve information as the new data is introduced or, simply, leave the system. This choice depends on the requirements of the entity that funded the attack.

Table I demonstrates the key differences between the “traditional” attacks and the APT-type attacks. This table is the result of a summary made by the authors, Chen et al. [3].

B. How does an APT act?

As stated before, APTs are well-planned, structured, and organized attacks, in order to increase the probability of the attack being

successful. These have to go through several stages to achieve their goal. To better explain these stages, it will the 6 stages model, introduced by [3], and one more stage called “Covering the Trails”. The 6 stages model is based on the concept of a “chain of destruction via intrusion” presented by [4]. The application of a model of a chain of destruction helps the comprehension of the practices of the players of the threats in each level, and also supplies the directives for the defense against the APT attacks.

The stages of a typical APT may be visualized in Figure I.

1. Recognition and Armament

To initiate the attacks, the retrieval of information is a fundamental step in the process. The more the attackers comprehend the target, the more successful the attack will be. The attackers identify and study the organization, retrieving the maximum amount of information possible about the technical environment and the “key” personnel of that organization. The attackers retrieve relative data about individual employees, such as

	execution	and long-duration protections
--	-----------	-------------------------------

Table I: A comparison between traditional attacks and APT attacks

social life, habits, and websites that they frequently visit, as well as details from the infrastructure of subjacent information technologies, such as the category of used switches, routers, used antivirus tools, firewall, web servers, open doors, among others.

This information not only would allow a network foothold but in addition, it would be used to penetrate the target’s network deeper. Open Source Intelligence Tools (OSINT) and social engineering techniques are used to retrieve this information. The social engineering, which involves the psychological manipulation of people to reach objectives that may, or may not, be in the target’s best interest, is one of the techniques utilized to retrieve information[3].

	Traditional Attack	APT Attack
Attacker	Probably only one person	Highly organized, sophisticated, determined, and well funded group
Target	Not specified, probably individual systems	Organizations, commercial companies, and specific governmental institutions
Purpose	Economic Benefits and abilities showcase	Competitive and strategic advantages
Approach	Single execution, “Smash and Grab”, short time	Repetitive tries, remain hidden and slow, adjusting to the defense

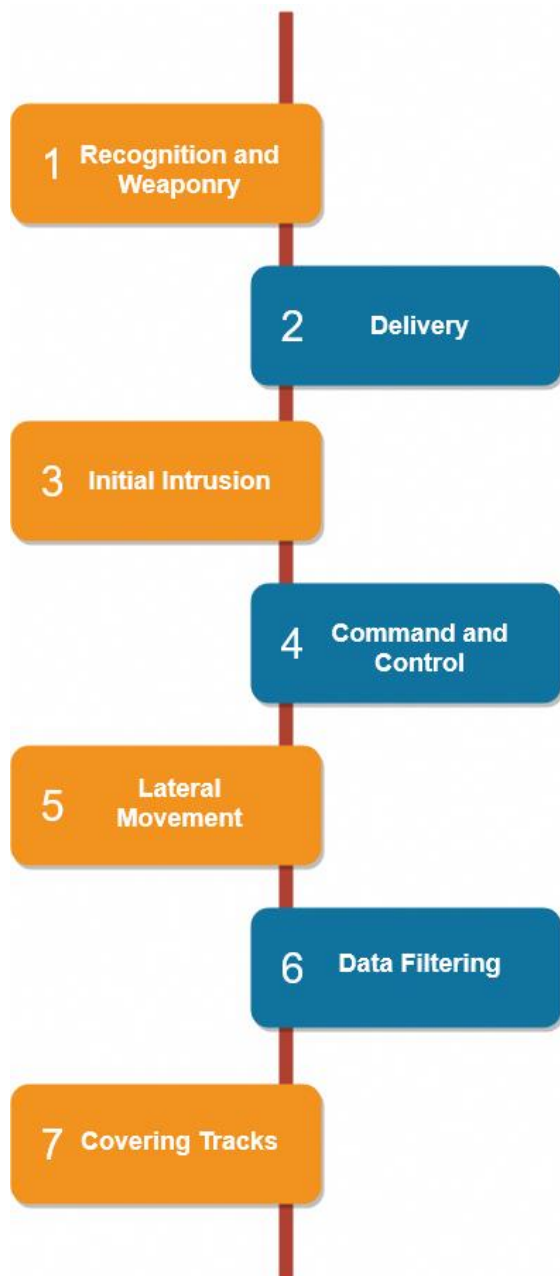


Figure I: stages of a typical APT attack

In cyberattacks and cybercrime, it's frequently utilized to obtain sensitive information, or make it so that the target executes certain actions (for example, executing malware). OSINT is a form of information gathering from publicly available sources, and nowadays, it typically refers to information aggregation about a subject coming from paid or free sources on the Internet. The information can be retrieved via OSINT, from the personal profile of an employee to an organization's hardware and software configurations. To produce actionable intelligence, the attackers can utilize data

prospection and great analytical data techniques to process the retrieved data.

Based on the gathered intel, the attackers will build an attack plan and will prepare the necessary tools. The attackers generally prepare several tools for different attacks, to adapt their tactics in case of failure. Besides that, these information gatherings look for publicly available repositories, such as WHOIS and BGP, to search domain and routing information, as well as websites on the network that may have been visualized, and have high-risk vulnerabilities, such as cross-site scripting (XSS) and SQL injections[5].

2. Delivery

The attackers deliver their exploits to the targets in this stage, the usage of known vulnerabilities is one of the sources that the attackers utilize to execute advanced and persistent threats. The known weaknesses are generally exposed and may be obtained from well-known vulnerabilities databases, such as Common Vulnerabilities and Exposures List (CVE), Open Source Vulnerability Database (OSVDB)[6], and NIST National Vulnerability Database (NVD)[7]. Each vulnerability is identified using a unique CVE-ID. Beyond that, in some cases, the attackers may share and gather useful information from vulnerabilities found in dark-web and deep-web forums[8]. As for the study related to the paper[9], the majority of APT attacks were based on known vulnerabilities. Therefore, applying security updates is essential, when such are available. Direct and indirect delivery are two categories of mechanisms for the delivery of exploits. For direct delivery, the attackers utilize several social engineering techniques, such as spear-phishing, to send exploits to their targets. The indirect delivery is less perceivable. The attackers compromise first a third party whom the target trusts, and after utilize that third party to serve explorations. It's possible to find a software/hardware supplier utilized in the target organization or a legitimate website that is frequently visited by the concerned people. Up next, some exploit delivery mechanisms are described.

Spear-phishing: Spear-phishing is a way of phishing where fraudulent e-mails only aim at affecting a small group of selected recipients. The Spear-phishing attacks, specifically under the shape of a Business e-

mail Compromise, are favored by attackers instead of old mass e-mail attacks with set phrases, according to a Symantec report. The attackers use social engineering to obtain information about the organization and then, they send malicious software e-mails. These fraudulent e-mails are ingeniously elaborated, enough to intrigue the selected recipients to open up the sent attachments. Employees that don't possess knowledge of the malware's existence can make the organization's network risk the download of an "apparently" harmless attachment that contains a vulnerability exploit, or click on a link that sends the recipient to a malicious website that serves drive-by-download exploits[10]. When executed, this malware can try giving an exploit to known or unknown vulnerabilities, to establish a foothold in the organization's network. In APT-type attacks, malicious attachments are frequently utilized instead of malicious links, since people are more prone to sharing files (for example, reports, commercial documents, and resumes) by electronic mail in the business or governmental environment.

Zero-Day Vulnerability: a zero-day vulnerability is a software bug that the software manufacturer doesn't have knowledge of or has the notion of, but couldn't correct it before the attackers could utilize it. The operative system versions, executed patches, and installed software components in those systems are some of the information that the attackers gather from the organization. Then, they proceed to the identification of any vulnerabilities in those versions that could be utilized to obtain an entrance to the target's network. However, in accordance with the related study in the paper [9], only some APT attacks were accomplished by zero-day vulnerabilities. The majority of APT attacks were based on known exploits.

Watering Hole Attack: the concept of Watering Hole Attack is like a predator waiting in a water hole in a desert, since the predator knows that the victims will have to go to the water hole. In the case of a Watering Hole Attack, the attackers can identify third-party websites that are frequently visited by the selected people, and then, the attacker can begin trying to inject malicious content into a vulnerable website. After sending e-mails with attachments or malicious links that go to these malicious software websites, the attackers patiently wait for the malware to work in the organization's network, which would open

doors for their entry into the organization's system. Here, the challenge for an APT attacker is to have the malware working without being detected by the antivirus tools, and by the detection and intrusion prevention systems. Eventually, the delivery is accomplished when the infected Internet pages are viewed by the victims [11]. The use of Water Hole Attacks has been seen in several APT attacks [12], [13].

Given that the attackers obtained control of the system through the execution of malware that explores vulnerabilities in the system, they stay in hiding to continue undetected for the next stage. Also in this regard, the APT attackers look to build a communication channel of Command and Control (C&C) after infiltrating the chosen network, to implement subsequent attacks. The majority of malware resorts to the Domain Names System (DNS) to locate the domain names servers and compromised devices, so that the APT attackers can establish a long-term link to the victims' devices to steal confidential data.

3. Initial Intrusion

When the attacker obtains first non-authorized access to the computer or network of the target, the initial intrusion occurs. Although the attackers can obtain access credentials through social engineering, and simply, utilize them to gain "legitimate" access, the typical form of intrusion is through the use of malicious code that explores a vulnerability in the target's computer. After the delivery of the malicious code in the delivery stage, the attackers obtain access to the target's computer when they successfully execute the exploit. The attacks when the attackers concentrate the vulnerabilities in programs such as Adobe PDF, Adobe Flash, and Microsoft Office are known as APTs. Even though several APT attacks [14], [15] have exploited the zero-day vulnerability for the initial intrusion stage, many APT attacks also use the oldest vulnerabilities of non-corrected applications. In an attack, the initial intrusion is the most important stage, as the attackers establish a support foothold in the network with the information obtained in the other stages. The installment of a backdoor is, generally, the result of a successful intrusion. The possibility of the defenders detecting the persistent advanced threat in an initial stage may exist, for

network traffic is produced, and the evidence is left in the victims' computers.

4. Command and Control

When an attacker establishes a backdoor, there must be an open communication channel between the server itself and the victims' devices. This is known as C&C (or C2) and it's a fundamental element of the network throughout the attacks and can be utilized to further attack the network. The communication of the C&C applies to the main services of the network, such as the Hypertext Transport Protocol (HTTP) or Internet Relay Chat (a protocol that allows the sending of messaging in real-time in an only-text environment). The C&C traffic based on HTTP is preferable to others, because, first and foremost, the C&C traffic based on HTTP is certified as "legitimate" in most companies, and secondly, other C&C protocols such as P2P and IRC traffic have distinct network characteristics, such as doors and package content, being easily identifiable and blockable [9], [16]. To avoid detection, the attackers trust even more various legitimate services and publicly available tools that anyone can use.

Social Media: the attackers register on several social media websites, and then insert control information in blogs' publications or status messages [17].

Tor anonymity network: The Tor anonymity network includes configured servers to only receive entrance links by Tor. These servers are called 'hidden services'. By allocating C2 servers as 'hidden services' on the network, Tor makes them harder to identify, place on the blacklist, or eliminate the server.

Remote Access Tools (RATs): although they are frequently utilized for legitimate remote administration, the Remote Access Tools are frequently associated with cyberattacks [11], [18]. A RAT consists of two components: a server that resides on the victims' endpoint, and an installed client in the attacker's machine. To function, the component 'server' needs to be delivered first to the target's machine. This is frequently achieved by spear-phishing type e-mails.

5. Lateral Movement

After obtaining access to the target's system, an attacker can spread to other systems

in the target's internal environment. The attackers utilize various techniques to access other hosts through a compromised system, to obtain access to sensitive resources. The stolen credentials to legitimate users are frequently utilized during this stage. This includes the placement of malware and other tools in different machines, the compromised components of the system, and its concealment. Sometimes, this stage involves obtaining more access privileges, and other times involves obtaining keyloggers' passwords. At times, this can be done using pass-the-hash techniques and vulnerability exploits. The environment of the targeted system is what determines the chosen method. The attackers want to expand their foothold to other systems inside this environment, to find the data they want to filter, in this stage. Therefore, as long as the attacker has reached this advanced stage, it's really hard to remove the attacker completely from the environment [9]. The dumping of credentials is the process of obtaining the login and password information from accounts coming from the operative system and software. There exist credentials that can be used to access restricted information. The attackers that use APT can take advantage of the valid credentials and move stealthily inside an environment. Suitable applications that can download this information from a system are the main method of hashes and password retrieval. The Mimikatz is one of the most used password dump tools because it can dump clear text passwords, and it also has more characteristics compared to other tools in the market. Another tool used by attackers to gather valid credentials is the Credentials Editor from Windows. Although there exist different techniques for the dumping of credentials from Windows, the most common method is to extract and analyze parts from the process of the Windows Local Security Authority (LSA) [9]. These tools are used not only by security professionals, but by their opponents as well.

6. Data Filtering

Data filtering is a critical step for attackers, to obtain sensitive data and strategic benefits. When the attackers find the data that they are looking for, they try to contact the control center and export them to a remote site. The data is generally sent to an internal processing server, where they are compacted and then sent to other sites under the control of

the attackers. To avoid a transmission of information from being detected, the attackers that use APTs are usually known for using secure protocols like SSL/TLS, or by taking advantage of the anonymity given by the Tor network [17]. Given that the majority of the systems of detection and intrusion prevention don't use an outgress filter, the filtering of the data can pass unnoticed. If the organization has a defense methodology, the attacker could divide the data into blocks to be sent to the server.

7. Covering Tracks

The objective of an APT attack is not to simply steal the data of an organization, but to keep doing it until the attack is canceled by the attack's sponsor. The sponsor can keep up with the attack and the gathering of data while they can, or the gathering of data as these are being obtained. Whatever the case, the attackers should have to cover their tracks so that the target has no knowledge of who is attacking. If the need doesn't arise for permanent and continued monitoring, the tools will normally be removed to cover their tracks. This is frequently done by the establishment of an open door, that would make for an easier return.

VI. Machine Learning

According to authors Alshamrani et al. [1], Machine Learning (ML) is the capacity of which a machine can modify the result of a situation, or behavior on the basis of knowledge, or observation. The learning algorithms of machines utilize data to learn new patterns and identify patterns previously unknown. These algorithms can be used to analyze data immediately as they are gathered, therefore allowing more precise results than the previous methods. ML can be applied in different cases, depending on the wanted result that is previously known. In ML, the models can interact with the environment to learn with the gathered data for different means. A learning system can give three different ways of knowledge, depending on the nature of learning. Identifying what constitutes regular and irregular behavior is the most important part of detection based on anomalies. There exist several works that aim at identifying the difference between these. Some utilize a group of practice data that represents the normal behavior of a system or network, and

meanwhile, there are others that utilize models in learning about the system to identify abnormal behaviors.

VII. Development of the article and Discussion

With the use of Machine Learning, it's possible to teach machines to detect and, possibly, fight APTs. In reality, presently, there exist some detection techniques, as can be observed in table II, as a result of the work accomplished by Rajalakshmi et al. [19].

Each one of these techniques has its own advantages and disadvantages. There are countless sublevel tests that consume an enormous amount of time. But the number 1 technique has an enormous detection precision rate, being of 85% [20]. Technique 2 can analyze the malware content present in the IoT devices, utilizing a recurring neural network [21].

This classifies the malware based on the operational codes, although, the operational codes require storage. The Life pattern characteristic in Technique 3 can increase the precision of threat detection [22], but the groups of data must be trained with different metrics. Technique 4 involves the usage of a specific search engine to apply a Hadoop platform to investigate APT victims [23]. This approach is more effective than the approaches based on the signature, but it's not perfect. The Naïve Bayes and SVM methods in Deep Belief Network are utilized in technique 5 [24]. The detection of malware is highly predictable, but the detection process takes time to detect the threats and malware.

Technique 6 is used to detect ransomware, based on the analysis of the signature pattern of the active ransomware. This is obtained with the usage of the Perceptron multilayers algorithm [25], but it's hard to implement. Technique 7 allows the detection of attacks with targets, such as APTs that are hard to detect with conventional methods, with the usage of SPuNge [26]. Even though methodical, Technique 7 is a time-consuming process. Technique 8 can detect malware in 5 to 10 seconds, but has a low precision rate [27].

As previously clarified, the APT type attacks differ from traditional attacks, only don't they possess different approaches, but the goals and reasons for the attack are also distinct. These attacks are a time-consuming process,

with different stages and different demands by stage. The hard part of detecting an APT-type attack is knowing which stage the attacker is currently on. The ideal would be to have software that could identify and document the current stage that is occurring to the target of the attack, to know what is occurring in the systems of the target of the attack. Many of these techniques cannot possess by themselves an enormous impact but, if these are used simultaneously, or even combined, it's possible to increase the precision as well as reduce the necessary time to detect. Also, the ideal would be to implement these techniques to act in real-time in the networks, like a normal antivirus or firewall.

Techniques of APT detection	
1	Detection of threats through the correlation analysis by Machine Learning
2	An RNN approach for the "hunting" of malware threats
3	System of detection of intrusions, resulting from anomalies
4	Search engine to discover malware and/or threats
5	Use of the Deep Learning algorithm to detect Malware
6	Frequent model of exploration for Malware and threat persecution
7	Detection of target-specific attacks, with the use of SPuNge
8	Prediction of the initial stage of malware recurring to RNN

Table II: techniques of detection of APTs, relying on Machine Learning

VIII. Reflection of the work

When I started formulating the idea for this article, the first thing that came to mind was that I did not know enough about APT and/or Machine Learning, since these weren't topics that I have never been interested in. Still, during

my research, I discovered these to be extremely fascinating subjects. In the case of Machine Learning, to be able to teach a machine, or robot, to do things is very fascinating, to be able to evolve and gain knowledge as a typical human child would. In the case of APTs, I have never heard about this category of attacks until my professor mentioned them. The initial idea that I had about these attacks was very different from that of reality, for example, initially, I thought that it was a typical cyberattack, but with the use of very advanced and sophisticated tools, which is very wrong, since it's almost the opposite, it's about a hidden and slow attack that isn't detected and does the most damage possible, with very specific goals. This article was a good way of challenging me to obtain more knowledge about these topics and create a desire to know more about them, in a way that in my free time I will continue searching and consulting, and informing myself about APTs, Machine Learning, and their junction.

IX. Conclusion

APTs (Advanced Persistent Threat) are one of the biggest existing cybersecurity threats presently. Different from traditional attacks, the APTs attacks act in a slow and hidden way, to be able to realize the objective that took to the attack existing, normally what the funding entity of the attack will obtain with this attack. These attacks go through different stages, specifically: recognition and weaponry, delivery, initial intrusion, command and control, lateral movement, data filtering, and covering tracks, to be accomplished, so it's possible to use Machine Learning to teach machines to recognize these stages through data, patterns, and metrics. There exist several techniques for the detection of these attacks from which 8 were mentioned in this article, if these techniques are used simultaneous or even still combined amongst themselves to create "the ultimate technique", it will be possible to be able to fight and reduce the risk that APTs are.

X. References

[1]A. Alshamrani, S. Myneni, A. Chowdhary, and D. Huang, "A survey on Advanced persistent threats: Techniques, solutions,

- challenges, and research opportunities,” *IEEE Communications Surveys & Tutorials*, vol. 21, no. 2, pp. 1851–1877, Jan. 2019.
- [2]R. S. Ross, “Managing information security risk: Organization, mission, and information system view,” *Special Publication (NIST SP)-800-39*, 2011.
- [3]P. Chen, L. Desmet, and C. Huygens, “A study on advanced persistent threats,” in *IFIP International Conference on Communications and Multimedia Security*. Springer, 2014, pp. 63–72.
- [4]E. M. Hutchins, M. J. Cloppert, R. M. Amin, and others, “Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains,” *Leading Issues in Information Warfare & Security Research*, vol. 1, no. 1, p. 80, 2011.
- [5]A. K. Sood and R. J. Enbody, “Targeted cyberattacks: a superset of advanced persistent threats,” *IEEE security & privacy*, vol. 11, no. 1, pp. 54–61, 2013.
- [6]O. S. V. D. (OSVDB), “Open source vulnerability database (osvdb),” 2012.
- [7]P. Mell, K. Scarfone, and S. Romanosky, “Common vulnerability scoring system,” *IEEE Security & Privacy*, vol. 4, no. 6, 2006.
- [8]M. Motoyama, D. McCoy, K. Levchenko, S. Savage, and G. M. Voelker, “An analysis of underground forums,” in *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference*. ACM, 2011, pp. 71–80.
- [9]M. Ussath, D. Jaeger, F. Cheng, and C. Meinel, “Advanced persistent threats: Behind the scenes,” in *Information Science and Systems (CISS), 2016 Annual Conference on*. IEEE, 2016, pp. 181–186.
- [10]A. TrendLabsSM, “Spear-Phishing Email: Most Favored APT Attack Bait”, 2012.
- [11]G. O’Gorman and G. McDonald, “The elderwood project”. *Symantec Corporation*, 2012.
- [12]W. Gragido, “Lions at the watering hole: The voho affair,” *RSA blog*, vol. 20, 2012.
- [13]D. Kindlund, D. Caselden, X. Chen, N. Moran, and M. Scott, “Operation SnowMan: DeputyDog Actor Compromises US Veterans of Foreign Wars Website,” *FireEye*, 13-Feb-2014. [Online]. Available: <https://www.fireeye.com/blog/threat-research/2014/02/operation-snowman-deputydog-actor-compromises-us-veterans-of-foreign-wars-website.html>. [Accessed: 13-Jul-2022].
- [14]S. McClure *et al.*, “Protecting your critical assets-lessons learned from operation aurora,” *Tech. Rep.*, 2010.
- [15]RSA FraudAction Research Labs, “The anatomy of the RSA attack,” *RSA blog*, 01-Apr-2011. [Online]. Available: <http://blogs.rsa.com/anatomy-of-an-attack/>. [Accessed: 13-Jul-2022].
- [16]X. Wang, K. Zheng, X. Niu, B. Wu, and C. Wu, “Detection of command and control in advanced persistent threat based on independent access,” in *Communications (ICC), 2016 IEEE International Conference on*. IEEE, 2016, pp. 1–6.
- [17]B. Harris, “Shadows in the cloud: An investigation of cyber espionage 2.0,” *GovTech*, 02-Aug-2010. [Online]. Available: <https://www.govtech.com/dc/articles/shadows-in-the-cloud-an-investigation.html>. [Accessed: 13-Jul-2022].
- [18]M. Z. Rafique, P. Chen, C. Huygens, and W. Joosen, “Evolutionary algorithms for classification of malware families through different network behaviors,” in *Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation*, 2014, pp. 1167–1174.
- [19]E. Rajalakshmi, N. Asik Ibrahim, and V. Subramaniaswamy, “A survey of machine learning techniques used to combat against the advanced persistent threat,” *Applications and Techniques in Information Security*, pp. 159–172, Nov. 2019.
- [20]I. Ghafir, M. Hammoudeh, V. Prenosil, L. Han, R. Hegarty, K. Rabie, and F. J. Aparicio-Navarro, “Detection of advanced persistent threat using machine-learning correlation analysis,” *Future Generation Computer Systems*, vol. 89, pp. 349–359, Jul. 2018.
- [21]H. HaddadPajouh, A. Dehghantanha, R. Khayami, and K.-K. R. Choo, “A deep recurrent neural network based approach for internet of things malware threat hunting,” *Future Generation Computer Systems*, vol. 85, pp. 88–96, Mar. 2018.
- [22]F. J. Aparicio-Navarro, K. G. Kyriakopoulos, Y. Gong, D. J. Parish, and J. A. Chambers, “Using Pattern-of-Life as Contextual Information for Anomaly-Based Intrusion Detection Systems,” *IEEE Access*, vol. 5, pp. 22177–22193, 2017, doi: 10.1109/ACCESS.2017.2762162.

- [23]S.-T. Liu, Y.-M. Chen, and S.-J. Lin, "A novel search engine to uncover potential victims for apt investigations," in *IFIP International Conference on Network and Parallel Computing*, 2013, pp. 405–416.
- [24]G. E. Hinton, "Deep belief networks," *Scholarpedia*, vol. 4, no. 5, p. 5947, 2009.
- [25]M. Moradi and M. Zulkernine, "A neural network based system for intrusion detection and classification of attacks," in *Proceedings of the IEEE international conference on advances in intelligent systems-theory and applications*, 2004, pp. 15–18.
- [26]M. Balduzzi, V. Ciangolini, and R. McArdle, "Targeted attacks detection with sponge," in *2013 Eleventh Annual Conference on Privacy, Security and Trust*, 2013, pp. 185–194.
- [27]A. Azmoodeh, A. Dehghantanha, and K.-K. R. Choo, "Robust malware detection for internet of (battlefield) things devices using deep eigenspace learning," *IEEE transactions on sustainable computing*, vol. 4, no. 1, pp. 88–95, 2018.