Advanced Persistent Threat Detection: A Survey

Adam Khalid
School of Computing
Faculty of Engineering
Universiti Teknologi Malaysia
81310 Johor Bahru, Johor, Malaysia
adam.khalid@mnu.edu.mv

Anazida Zainal
Cyber Threat Intelligence Lab,
Information Assurance & Security
Research Group (IASRG)
School of Computing, Faculty of Engineering
Universiti Teknologi Malaysia
81310 Johor Bahru, Johor, Malaysia
anazida@utm.my

Mohd Aizaini Maarof
Cyber Threat Intelligence Lab,
Information Assurance & Security
Research Group (IASRG)
School of Computing, Faculty of Engineering
Universiti Teknologi Malaysia
81310 Johor Bahru, Johor, Malaysia
aizaini@utm.my

Fuad A. Ghaleb
Cyber Threat Intelligence Lab,
Information Assurance & Security
Research Group (IASRG)
School of Computing, Faculty of Engineering
Universiti Teknologi Malaysia
81310 Johor Bahru, Johor, Malaysia
abdulgaleel@utm.my

2021 3rd International Cyber Resilience Conference (CRC) | 978-1-6654-1844-7/20/$31.00 ©2021 IEEE | DOI: 10.1109/CRC50527.2021.9392626

***Abstract*—Advanced Persistent Threat is a very sophisticated targeted attack aimed at organizations. Several approaches have been proposed to detect APT. This paper defines an APT as an attack that has certain objectives to be achieved, and are performed by well-funded organizations, and is long term campaign. In this paper we have identified APT as a threat that follows a kill chain process. Intrusion detection and intrusion detection methods are summarized in this paper. Detection of an APT is a challenge. In this paper various detection methods used by researchers and the challenges in detecting APT is highlighted.**

***Keywords*—Advanced Persistent Threat, anomaly, signatures, machine learning, targeted attacks, intrusion detection**

I. Introduction

With the ease of access to the internet and increasing online services more and more networks are getting connected, and are using these online services. This open nature of the internet and the ease of access has increased exposure to cyber threats [1]. Today cyber threats are a major concern and are compromising the availability, integrity and confidentiality of data in computer systems [2].

Several cyber-attacks have been reported recently and the losses due to these cyber-attacks are innumerable. In 2018 cyber attackers attacked Facebook and leaked about 50 million users' personal data [3]. In June 2017 attackers attacked a dozen nuclear power stations in the US causing huge losses [3]. In 2014 attackers earned 3 million USD through ransomware attacks [4]. In 2018 the United States alone got losses of 57 million USD due to cybercrime [5]. This figure is expected to rise and is expected to reach 3 billion in the year 2021 [6].

Cyber-attacks can be categorized into two groups; namely targeted and untargeted attacks. In an untargeted attack, the attacker indiscriminately tries to attack as many targets as possible. The attacker does not care who the victim is. The machines with vulnerabilities will be the likely victims. In a targeted attack, the attacker will pick a target and try to launch the attack on that particular target.

Today, a new and more sophisticated form of a targeted attack is becoming prominent. This new form of attack is known as an Advanced Persistent Threat acronym APT [7]. Today APT attacks have become a real threat to governments, businesses, research institutes, etc: The acronym APT stands for the following [8]:

- *Advanced:* refers to the stealthy nature of the attack. The attackers are data focused and will constantly change their attack patterns. APT attackers will keep on trying until the attack becomes successful.
- *Persistent:* APT attackers will follow a low and slow approach to penetrate a system. The main aim of the attacker will be to penetrate and maintain a long-term presence in the system and stay hidden in the network until the attacker's objectives are fulfilled.
- *Threat:* The actors in an APT attack will select a specific organization and will do whatever within their means to achieve their goals. APT attacks are well funded by organizations, sometimes governments [9].

The main goal of this study is to explicitly study various techniques and solutions researched to detect APT. An APT is a six stage attack process. Several research has been done to detect APT but only a few detects the stages of APT and correlate them. In this study various detection strategies used to detect APT is explored and their main strength and weaknesses are highlighted. In this paper Intrusion detection is defined and explores two different intrusion detection strategies commonly used, namely signature based and anomaly-based intrusion detection. Research shows that signature-based intrusion detection is not effective in detecting an APT attack. Researchers have used different methods to detect APT. This paper reveals that current detection approaches have several drawbacks and APT detection is still a challenge. The main challenges faced by researchers are also discussed in the paper.

The rest of this paper is organized as follows, Section 2 defines APT, Section 3 describes an APT life cycle, and Section iv explore intrusion detection systems, and section v explores the state of the detection techniques for an APT attack and paper concludes by a conclusion in Section vi.

II. APT Attacks Definition

National Institute of Technology defines an APT using three principles [10]. An APT attacker (i) pursues its

978-1-6654-1

978-1-6654-1844-7/20/$31.00 ©2021 IEEE

objectives repeatedly over an extended period of time, (ii) will adapt to the defender's effort to resist it, and (iii) are determined to maintain the level of interaction needed to execute its objectives [10]. This definition differentiates between an APT attack and a traditional attack. The man differences between an APT and a highlighted in Table I.

TABLE 1 DIFFERENCES BETWEEN AN APT ATTCK AND A TRADITIONAL MALWARE ATTACK

Feature	APT attacks	Common malware attack
Definition	Organized and highly sophisticated and well planned.	A malicious software developed, sometimes to show abilities
Attacker	An organization, sometimes governments	A cracker and a hacker involved in criminal activities
Target	Government organizations, computer-controlled systems, research institutes	Random victim selected by the attacker
Purpose	Destruction, steal sensitive information from the organization.	Personal recognition
Attack Life Cycle	Stay as long in the system until the attacker's goals are fulfilled	Ends when it is detected by the security team

APT attacks have some unique characteristics specific to an attack. These specific characteristics are as follows:

1) *Specific Objectives and Targets*: APT attackers will have a specific target and objectives. The targets are chosen by the funding organization. The attackers will keep on trying until the attack goals are achieved.
2) *Highly Organized and well resourced*: APT attacks are generally funded by organizations. Sometimes governments. The attackers are well funded and will have all the tools and funds required to launch the attack. This financial support and equipment make APT attacks difficult to detect.
3) *A long-Term Campaign:* An APT attack is a long-term campaign. The attackers will keep on trying until the attack becomes successful.
4) *Stealthy and evasive techniques:* APT attackers will have the ability to stay inside the victim's machine, interacting minimally with just enough resources to achieve their objectives. APT attackers commonly use zero-day attacks, to avoid signature-based detection, and encryption to obfuscate network traffic.

III. APT ATTACK LIFE CYCLE

APT attackers follows a kill chain model, consisting of six phases [11]. Figure 1 shows the shows the six phases of the APT kill chain process, commonly used by attackers to launch an attack. This kill chain model corresponds to the model proposed by Bhat [12]. The different phases in the kill chain process are as follows:

1) *Reconnaissance:* In this stage the attackers gather information from the organization. The information collected are the technical information about the organization, its network [13]. Attackers often use social engineering to obtain the required information, sometimes they use Open Source Intelligence Tools to collect information. Once the required information is obtained; the attackers will then process and analyse the collected data, using big data or data mining techniques. Once the information is analysed the attackers will develop the attack plan.

2) *Weaponize:* In this stage the attacker will develop malicious code to explore vulnerabilities, attach the malicious code to pdf, doc, and ppt.

3) *Delivery:* In this stage the malware is delivered into the victim's machine. Two types of delivery methods are used. Direct and indirect delivery. Some of the direct delivery methods are highlighted in Table II

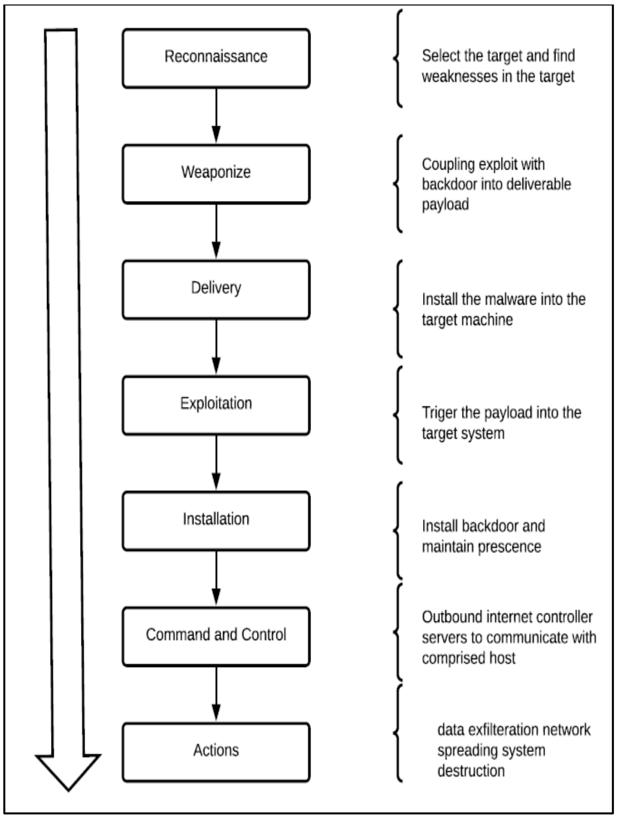
Reconnaissance
Select the target and find weaknesses in the target
Weaponize
Coupling exploit with backdoor into deliverable payload
Delivery
Install the malware into the target machine
Exploitation
Triger the payload into the target system
Installation
Install backdoor and maintain prescence
Command and Control
Outbound internet controller servers to communicate with comprised host
Actions
data exfilteration network spreading system destruction

Fig. 1 APT Kill Chain Process

4) *Exploitation:* Once the malware is installed the next stage is the exploit stage. In this stage the malware and its payload will get executed. Once the exploit is installed and triggered, the malware will become active and will start to communicate with the command-and-control server.

5) *Installation:* In this stage remote access trojans are installed which allows the attacker to maintain its presence in the target environment.

6) *Command and Control:* APT attackers once penetrated the network, will create a communication channel to control its malware, and will continue

communicating with the victim's servers and machines

7) *Actions:* The attacker achieves their objectives by performing data exfiltration.

TABLE II APT DIRECT DELIVERY METHODS

Method	Description
Spear phishing email	a fraudulent email send to selected staff in the organization. The staff are selected on the basis of the information collected in the reconnaissance stage. This email will contain malicious links and malicious attachments
Drive by downloads	The victim is encourages to visit a website, where there is a hidden Iframe, which will redirect to a malicious domain. This domain will run a browser exploit pack. This pack will download malware directly to the victims machine
Watering hole attack	Based on the information obtained in the reconnaissance stage the attacker will infect some of the websites the victim visits frequently. Once the website is visited malware will get downloaded directly
Zero day attacks	zero day attacks are security flows which are patched by software. Attackers use this vulnerability to launch the attack
Attacks on servers	The attackers will infiltrate the servers. The attacker will then use this server to infiltrate the network
Storage media	Attackers often use storage media to gain access to the system. Once access is obtained they will then execute the malware directly

IV. INTRUSION DETECTION SYSTEMS

An intrusion is defined as an attempt that comprises the integrity, availability and confidentiality of data in a computer system [14]. A software that monitors events in a network and analyses them for malicious activities is known as an intrusion detection system. Intrusion detection systems can either be signature based or anomaly based. The pros and cons of each of this detection methodology is tabulated in Table III

Signature Based Detection

In signature based detection, the signatures of all the known attacks are stored in a database of signatures, and when traffic arrives, the signatures of the traffic are matched with that of the signatures in the database. If a match occurs then the traffic is considered as malicious [14]. Figure 2 shows the general architecture for signature based detection. This method is easy to deploy and gives high accuracy for existing and already known attacks, whereas fail to detect new attacks.

APT attackers use a new attack technique for every attempt, and very often they use zero-day attacks. Due to this new attack methods developed for every attack, signature based detection alone cannot be used to detect APT attacks.

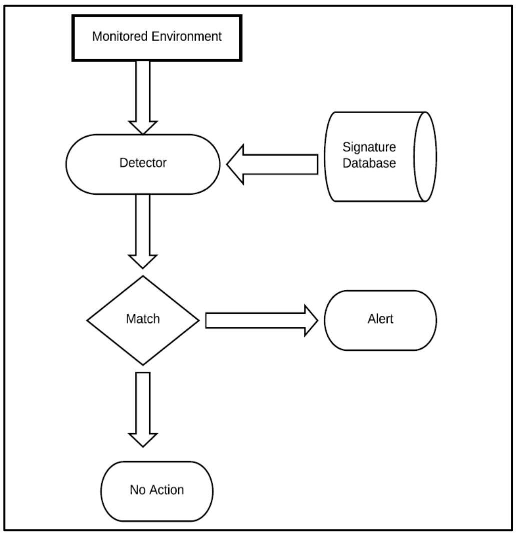

Fig. 2 Signature Based Detection Methodology [15]

Anomaly Based Detection

An anomaly is defined as a deviation from the normal behaviour. Once a malware enters the system the system will deviate from the normal behaviour, and this abnormality can be tracked [16]. In anomaly based detection, a detector examines the events occurring in the network against a baseline profile. The events occurring in the network are matched against a baseline profile. If the event matches with that of the baseline profile then the event is considered normal. If the events does not match then detector will check whether the events are within a threshold range. If the events are within the threshold range then profile is updated and if events are outside the threshold range then the events are considered anomalous. Figure 3 shows the basic architecture of an anomaly based detector. Table 3 compares the two different approaches.

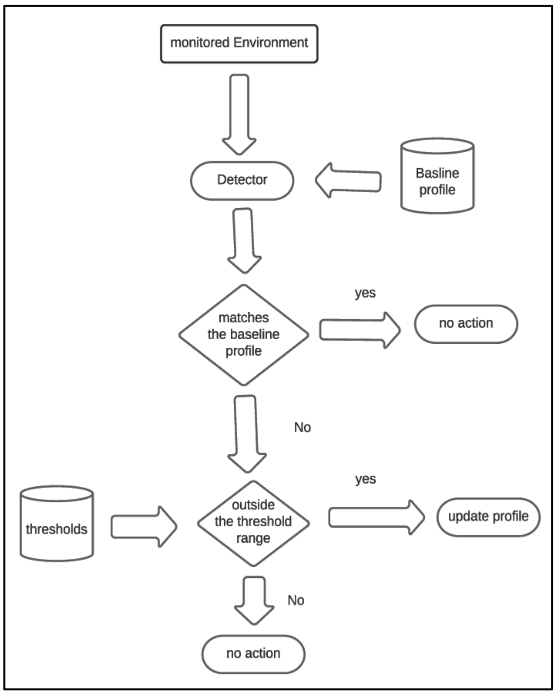

Fig. 3 Anomaly based Detection [15]

Anomaly based detection is the heart of any intrusion detection system. Anomaly based detection can operate on three modes; supervised, semi supervised and unsupervised modes. The different approaches to anomaly-based detection model is shown in Figure 4.

In supervised mode of detection a labelled data set is required [17]. A predictive model for normal and abnormal data are constructed. The unseen data are then compared to determine which class it belongs to. This method also has problems. One major issue with this approach is that anomalous data is much less than normal data. This gives an imbalanced distribution of normal and abnormal data [17]. Secondly obtaining malicious data is a challenge. To overcome this challenge researchers, inject artificial anomalies into normal data, to obtain a labelled malicious training data set. Labelling the data set is the main bottleneck in supervised learning. Supervised learning methods gives accurate results, however, fails to detect new attacks.

In semi supervised mode of detection only one data set containing the labelled data for normal data is required. Labelled data for anomalous data is not required. This makes the semi supervised model of detection widely applicable [17]. In semi supervised mode, a model representing the normal class is constructed. Any event that deviates from this normal class is considered as anomalous.

TABLE III COMPARISON OF INTRUSION DETECTION METHODOLOGIES

Detection Method	Advantages	Disadvantages
Signature Based	i) Very effective in detecting already existing attacks. ii) Very fast in detecting existing attacks. iii) Simple design	i) Needs to update the signature database regularly. ii) Unable to detect new and unseen attacks. iii) Cannot detect zero-day attacks. iv) Not suitable for detecting multistep attacks.
Anomaly based detection	i) Can be used to detect new and unseen attacks. ii) Can be combined with signature-based detection	i) Building a normal profile is complex. ii) Needs training iii) Generates unclassified alerts

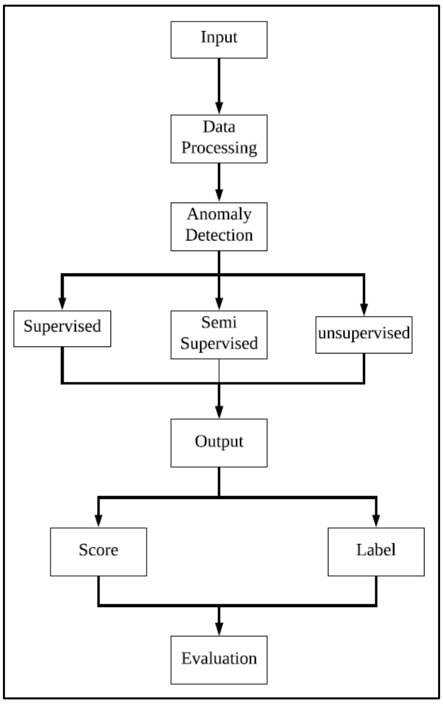

Fig. 4 Anomaly Detection Engine

In unsupervised mode of detection, no training data is required. Unsupervised mode of detection assumes that normal instances occur more frequently than anomalies in the test data. If this assumption fails, then the technique will give a high false alarm rate.

The output of anomaly-based detection will be either scores or labels. Scores assign a value to each instance. A threshold value is selected, based on the selected threshold anomalies are classified. In the labelling technique the outputs are either classified as either normal or anomalous [18]. Labelling method is technically more efficient since providing an anomaly score for each instance is complex.

Anomaly detection models are build using machine learning models. Machine learning models are build using two approaches: shallow learning and deep learning. Shallow learners build a predictive model based on the features whereas deep learners can extract better representations from the raw data and create more accurate results. Deep learners consist of several layers. At each layer a better representation of the features is obtained. Machine Learning and deep learning algorithms commonly used in intrusion detection are shown in the Figure 5.

V. APT DETECTION STATE OF THE ART

Several research works have been in detecting and preventing APT attacks. In this section we will explore some of the detection techniques and their drawbacks.

Friedburg in 2015 used machine learning approach to detect APT [19]. Friedberg's method consists of collecting system logs and extracting patterns of strings from these logs. Hypothesis are generated from these log lines. A hypothesis that hold for a long time are converted to rules.

Rules are generated and Machine Learning is applied on these rules to detect anomalies. Freiburg's method applies mainly to computer controlled devices. This method cannot detect zero-day attacks and also detect only one stage of an APT attack.

In 2014 Wang proposed a network gene-based approach [20] Network genes are the digital elements extracted by network protocols, reverse analysis and their combined sequences. Wang defined three levels of network genes, namely, messages, protocols, and operations.

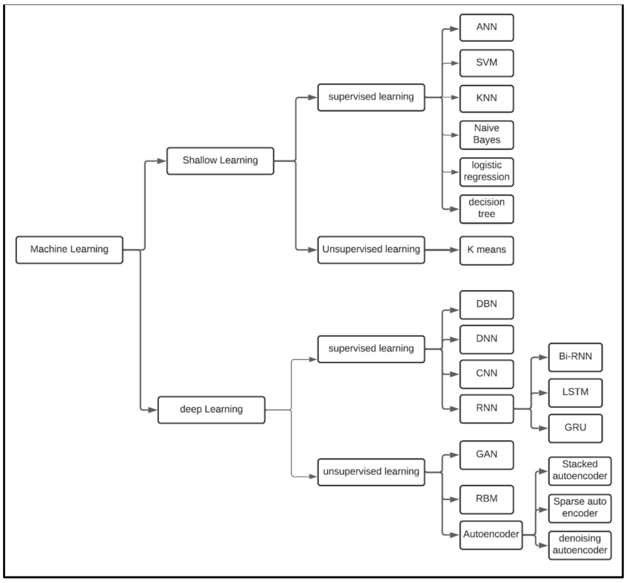

Fig. 5 Taxonomy of Machine Learning Algorithms

Combinations of all the three genes form the network genome of the application. From the network genes a network gene pool is created, and from this pool a blacklist and a whitelist are created. This method works by creating a blacklist and whitelist. New and unseen attacks such as zero-day attacks cannot be detected by this method [20].

Bhat and Gustavsson developed a framework to detect APT [12]. They used a layered defence system. Prevention and detection mechanisms are developed for each layer. The main aim is to make the penetration process more difficult for the attacker. The main drawback of this approach is that the system cannot correlate between attacks.

A distributed framework architecture was proposed by [21]. The main aim is to develop an intelligent system that can detect an APT attack. The proposed architecture runs in three phases. In the first phase the intrusion is detected using a Trusted Platform Module (TPM); which is a hardware device embedded in the mother board of the system. This TPM is designed to detect node to node communication.

The first step of this framework is to collect and analyse network traffic and study the possible strategies that can be used by an APT attacker to launch an attack. Four different detection methods are used to perform this task. All the four are autonomous and independent. The outputs of these are then fed into the correlation phase. In this phase all the outputs of each phase are taken as inputs to the correlation module and correlated individually, according to the rules specified by the administrator. After correlation the outputs are then taken into the voting phase. The result is determined by the voting phase.

The main problem with this approach is that the detection is done through hardware. The hardware device TPM is programmed based on previous detections. The main drawback of this approach is that the method can only detect known attacks. In this approach four there are four detection modules which does the same thing, and this will be a waste of resources.

A framework to detect APT was proposed by Ghaffir [22]. Ghaffir's proposed framework consists of three phases, threat detection, alert correlation, and alert prediction. In the threat detection stage eight modules are developed to detect threats. All these modules detect threats by creating a blacklist of threats. There are eight detection modules developed. Each module is independent from other modules. Ghaffir's framework can detect and correlate APT. The main problems with this method are that detection is based on blacklists, which means attacks such as zero-day attacks cannot be detected.

Wang developed a detection system that can be access to C&C domains [20]. Wang used an assumption that illegal access to C&C domains independent, while legal access to domains is correlated. This property was used, an analysed using machine learning techniques. The data was tested on a public data set and achieved significant results. This method also can only detect the C&C stage of an APT attack.

Chandra proposes detection based on spear phishing emails [23]. This approach uses statistical analysis to filter spam emails. Emails are split into tokens and specific words are searched by the detection algorithm to separate legitimate emails from spam emails. This method also detects one step in an APT attack. Emails that does not include any of the tokens will not be detected.

Sexton proposes an APT attack as a five stage process namely; delivery, exploit, install, command and control and actions [24]. In this model there is no specified way to move from one phase to the other. Within each phase there are several event types. Events in each phase are then combined. Combining is done by giving a score for each event type. An anomaly score for each host, and for each cluster with the same type of events. Events with anomaly score greater than a threshold are considered as APT attacks. The main drawback of this system is that the system will require expert knowledge to set up and maintain.

An APT detector SPunGe is proposed by Baduzzi [25] SPunGe gathers data from the host side and detects attacks on the host side. SpunGe detects targeted attacks through behaviour clustering, and location industry URL. Spunge determines the host distance and request distance and groups processed requests

VI. MAIN CHALLENGES IN APT DETECTION

Detecting and defending against an APT attack is a challenge to the research community. In this section we will highlight some of the challenges in detecting APT.

1) *Determined and Powerful Attackers:* The deterministic nature and the strength of the attackers causes a challenge in detecting APT. The system might have a strong defence mechanism but for the attacker it all boils down to building complex tools that can bypass this defence mechanism. For an APT attacker resources are plentiful, and this will enable to develop new tools and malware to achieve their goals.

2) *Duration of the attacks:* APT attacks span a long duration of time. Detecting a sophisticated attack is a challenge, correlating attacks that span a long period is bigger challenge. In detecting an APT, attack the state of the machines with suspicious behaviour needs to be tracked and stored. This anomalous behaviour needs to be correlated with

further incidents. For a large network this is a challenge.

3) *Internal Employees:* In the reconnaissance stage the attackers gather information about the organization. This is done through social engineering. In the APT kill chain process people are considered to be the weakest point in the process [14]. To prevent this clear security policies must be implemented in the organization.

4) *Powerful Resources:* APT attackers are sponsored most of the time. The determination, skills of the attackers, and the sponsoring makes detection a challenge.

5) *Infrastructure:* Today cloud computing systems are also growing rapidly. Evaluating vulnerability to these cloud computing systems is a challenge

CONCLUSION

Advanced Persistent threats are sophisticated attacks. APT attackers are well funded and the attackers equipped and technically knowledgeable. The funds, the quality and determination of the attackers makes detection of an APT difficult. Normal intrusion detection are of two types misuse and anomaly based detection. In misuse based detection the detector compares with that of signatures stored in a database of signatures, if there is a match then it is termed as a malicious. This method gives accurate results for known attacks but fail to detect unknown attacks. The approach is to use anomaly based machine learning detection to detect APT. This approach also gives high false positives and negatives. Several approaches are used to detect APT. But most the research done detects only one aspect of an APT attack. An APT is a six stage attack. All the stages needs to be detected and correlated. Detecting all these stages an correlating them is still an open research problem.

REFERENCES

[1] J. Liu et al., "ANID-SEoKELM: Adaptive network intrusion detection based on selective ensemble of kernel ELMs with random features," Knowledge-based systems, vol. 177, pp. 104-116, 2019.

[2] M. Jouini and L. B. A. Rabai, "A security framework for secure cloud computing environments," in Cloud security: Concepts, methodologies, tools, and applications: IGI Global, 2019, pp. 249-263.

[3] C. Analytica, "Facebook: The Scandal and the Fallout So Far," The New York Times, April, vol. 4, 2018.

[4] S. Ghosh and S. Sampalli, "A survey of security in SCADA networks: Current issues and future challenges," IEEE Access, vol. 7, pp. 135812-135831, 2019.

[5] P. Bell, "Cyber Threat Report 03 January 2019," 2019.

[6] S. Homayoun, A. Dehghantanha, M. Ahmadzadeh, S. Hashemi, and R. Khayami, "Know abnormal, find evil: frequent pattern mining for ransomware threat hunting and intelligence," IEEE transactions on emerging topics in computing, vol. 8, no. 2, pp. 341-351, 2017.

[7] A. Beuhring and K. Salous, "Beyond blacklisting: Cyberdefense in the era of advanced persistent threats," IEEE Security & Privacy, vol. 12, no. 5, pp. 90-93, 2014.

[8] I. Jeun, Y. Lee, and D. Won, "A practical study on advanced persistent threats," in Computer applications for security, control and system engineering: Springer, 2012, pp. 144-152.

[9] S. Quintero-Bonilla and A. Martín del Rey, "A New Proposal on the Advanced Persistent Threat: A Survey," Applied Sciences, vol. 10, no. 11, p. 3874, 2020.

[10] R. Kissel, Glossary of key information security terms. Diane Publishing, 2011.

[11] E. M. Hutchins, M. J. Cloppert, and R. M. Amin, "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.

[12] P. Bhatt, E. T. Yano, and P. Gustavsson, "Towards a framework to detect multi-stage advanced persistent threats attacks," in 2014 IEEE 8th international symposium on service oriented system engineering, 2014, pp. 390-395: IEEE.

[13] P. Chen, L. Desmet, and C. Huygens, "A study on advanced persistent threats," in IFIP International Conference on Communications and Multimedia Security, 2014, pp. 63-72: Springer.

[14] T. Hamed, J. B. Ernst, and S. C. Kremer, "A survey and taxonomy of classifiers of intrusion detection systems," in Computer and network security essentials: Springer, 2018, pp. 21-39.

[15] D. Mudzingwa and R. Agrawal, "A study of methodologies used in intrusion detection and prevention systems (IDPS)," in 2012 Proceedings of IEEE Southeastcon, 2012, pp. 1-6: IEEE.

[16] M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, "Network anomaly detection: methods, systems and tools," Ieee communications surveys & tutorials, vol. 16, no. 1, pp. 303-336, 2013.

[17] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," ACM computing surveys (CSUR), vol. 41, no. 3, pp. 1-58, 2009.

[18] M. Ahmed, A. N. Mahmood, and J. Hu, "A survey of network anomaly detection techniques," Journal of Network and Computer Applications, vol. 60, pp. 19-31, 2016.

[19] I. Friedberg, F. Skopik, G. Settanni, and R. Fiedler, "Combating advanced persistent threats: From network event correlation to incident detection," Computers & Security, vol. 48, pp. 35-57, 2015.

[20] 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 2016 IEEE International Conference on Communications (ICC), 2016, pp. 1-6: IEEE.

[21] P. K. Sharma, S. Y. Moon, D. Moon, and J. H. Park, "DFA-AD: a distributed framework architecture for the detection of advanced persistent threats," Cluster Computing, vol. 20, no. 1, pp. 597-609, 2017.

[22] I. Ghafir et al., "Detection of advanced persistent threat using machine-learning correlation analysis," Future Generation Computer Systems, vol. 89, pp. 349-359, 2018.

[23] J. V. Chandra, N. Challa, and M. A. Hussain, "Data and information storage security from advanced persistent attack in cloud computing," International Journal of Applied Engineering Research, vol. 9, no. 20, pp. 7755-7768, 2014.

[24] J. Sexton, C. Storlie, and J. Neil, "Attack chain detection," Statistical Analysis and Data Mining: The ASA Data Science Journal, vol. 8, no. 5-6, pp. 353-363, 2015.

[25] M. Balduzzi, V. Ciangaglini, and R. McArdle, "Targeted attacks detection with spunge," in 2013 Eleventh Annual Conference on Privacy, Security and Trust, 2013, pp. 185-194: IEEE.