Asian Journal of Computer Science and Technology
ISSN: 2249-0701 (P) Vol.7 No.S1, 2018, pp.90-95

DOI: https://doi.org/10.51983/ajcst-2018.7.S1.1797

# A Detailed Study on Advanced Persistent Threats: A Sophisticated Threat

**A. Sarkunavathi[1] and V. Srinivasan[2]**
[1]Research Scholar, [2]Professor and Head
[1&2]Department of Information Technology, Annamalai University, Annamalai Nagar, Tamil Nadu, India
E-Mail: sarkuna.bala@gmail.com, profvsau@gmail.com
(Received 1 October 2018; Revised 15 October 2018; Accepted 31 October 2018; Available online 7 November 2018)

***Abstract* - In the present world computer networks are used to store sensitive information and to provide services for organizations and society. The growth of internet and the increased use of computers in society along with smart devices lead to the increase in cyber crimes and persistent attacks. The most complex and advanced attacks are targeted attacks which are specifically aimed at companies or governments to accomplish the predetermined goals such as economic advantages, strategic benefits, getting control of sensitive information. Hackers try to access sensitive data from cyber space and there by become as advanced malware developers for the security systems. One type of such attack is Advanced Persistent Threats (APT) which targets the governmental institutions, military, multinational enterprises, financial industry, manufacturing and banks. The approach that is followed by the attackers are repeated attempts using different methods such as , stealth approach, adapting to the existing defense mechanisms, stealthily infiltrating the network to avoid any suspicions like involving in sleep modes before commencing any attack. The effects of these attacks are exfiltaration of key intelligence property, stoppage of fundamental services, and destruction of critical infrastructure. This paper is about the detailed study of Advanced Persistent threats to provide an idea about the advanced attacks.**
***Keywords:*** **Advanced Persistent Threats, Attacks, Effects**

## I. INTRODUCTION

Nowadays, the daily operation of public and private sectors ranging from government and military organisations to large enterprises and financial institutions depends largely on computers and networks. This dependency paves way to high range of cyber attacks. Traditional cyber attacks include computer viruses, worms, and spyware. Conventional cyber defense measures including firewall and intrusion detection becomes ineffective for these cyber attacks. Advanced Persistent Threats (APTs) are the most dangerous malwares of the present and near future. It can be defined as a cyber network attack in which an unauthorized person gains access to a network and stays there undetected for a long period of time.

The main objective of this type of attack is to ex-filtrate data rather than causing damage to the organizations [1]. It looks for specific information from a specific target, and will span over a long period of time using multiple targeting methods, tools and techniques. APT uses the strategy of slow and steady multiple attacks and it doesn't care how long it might take to accomplish its goal. APT attacks are mostly carried out by groups because individuals do not possess the ability to attack highly secured systems in the targeted organizations. Individuals usually choose easily vulnerable targets as they do not have sufficient money and infrastructure to carry out large-scale attacks.

From the year 2006 to 2009, knowledge about APT was with the military and intelligence services. The situation changed in the year December 2009 with the Google Aurora attack. According to Google, this attack originated in China, targeted Google and several other large American companies (Adobe Systems, Juniper Networks, Rack space, etc.), and resulted in the theft of Google Intellectual Property (IP). This single event made the world to know about this dangerous threat. Since Operation Aurora, several other commercial enterprises including Morgan Stanley, RSA Security, and the World Bank have suffered APT attacks of their own.

The existing security systems are able to detect only the common types of threats, but they are inefficient to identify APTs because an APT attacker always impersonate normal behaviour of the users and infiltrates the host. It is a fact that no two APTs are same, and the traditional detections mechanism fails at this point. Also different combination of attack techniques and ex-filtration methods are used by some well-organized cyber-attack organizations and are called as APT groups.

It takes several months to years to detect an APT attack, because these attacks are of an intelligent design and have the characteristics of being persistent. [3] The APT cases and their related attack detection delay: Flame (six years: 2006- 2012), Red October (five years: May 2007, October 2012), Stuxnet (one year: June 2009-June 2010), Duqu (one year: November 2010 –September 2011). In several cases, attacks are not detected by organizations. [2] When detected it's found that 69% of revealed intrusion cases analysed during 2014, are by external entities or by attackers themselves.

The objective of this paper is to study about the above said challenges, as well as the emerging challenges of the APT

life cycle and also to know how security experts focus on understanding the tactics and the techniques carried out by the attackers and detection mechanisms.

## II. ADVANCED PERSISTENT THREATS

### A. APT-Definition

The term Advanced Persistent Threat (APT) was first coined by cyber security analysts at the United States Air Force in 2006. This term Advanced Persistent Threat can be defined as:

1. *Advanced*: Attacks that are coordinated by a group of people with advanced resources and knowledge and are often well funded.
2. *Persistent*: The attackers repeatedly try to compromise the victim's system until they get access to sensitive data.
3. *Threat*: Possible source of harm or danger that has the potential to exploit vulnerabilities, and make security breach to the organization.

APTs share many characteristics with traditional "low and slow" hacking techniques which remain invisible to the detection system by compromising the network. The significant features of the APT threats that are defined by security experts are:

1. *Targeted:* APTs targets the organizations with the purpose of stealing specific data or causing damage.
2. *Persistent:* It goes through many phases over a long period of time to steal the data.
3. *Evasive*: They evade traditional security products by gaining privileged access of the hosts within the targeted network.
4. *Complex:* They use complex attack methods leading to multiple vulnerabilities in the targeted systems.

### B. Evolution of APT

A new class of cyber threat called Advanced Persistent Threat (APT) was firstly used in 2006 [3] and became well known in 2010 because Google was attacked by operation Aurora [3]. In 2011, APT was formally defined by National Institute of Standards and Technology (NIST) [3] as "The advanced persistent threat: (i) pursues its objectives repeatedly over an extended period of time; (ii) adapts to defenders' efforts to resist it; and (iii) is determined to maintain the level of interaction needed to execute its objectives."

## III. APT PHASES

The APT process includes several phases which may take place over a period of several months. The five main phases are reconnaissance; compromise (Infiltration); maintaining access; lateral movement; data ex-filtration (Stealing Data).

*A. Reconnaissance:* This phase is all about knowing the target by both social and technological aspects. Social engineering is one set of attack vector that APT groups may use targeting the human as the weakest link of the system. This includes attack like phishing, calling support to get information or change passwords, to sell data or information about the multinational firms such as IT hardware, security application, employees personal data etc;

*B. Compromise (Infiltration)*: An APT attacker infiltrates the system, possibly through social engineering strategies that exploit information gathered during the reconnaissance phase. They exploit the weakness in the infrastructure and gain access to the target network. This can be performed by two ways namely direct and indirect ways. In direct way the attacker compromise any third party working at the organization in order to use him as a loop hole to steal the data. In indirect way they employ techniques such as Spear Phishing, Watering Hole attack, Zero day virus and install a RAT (Remote Access Trojan, or Remote Administration Tool).The common approaches are [8]; by using email link that appears to be from reliable sources leading to vulnerability of sensitive data. The user visits the link website which contains a malicious JavaScript payload; which contain a built-in a zero-day Internet Explorer exploit. The common direct approach is when the storage media like USB is attached to a system. Once an infected USB is attached to a window based system, malware would auto-execute without user interaction, utilizing zero-day vulnerability (in some cases using a modified autorun.inf technique).The RAT embedded in a Trojan horse are then often used to take control of a machine.

*C. Maintaining Access*: An attacker uses the RAT (Remote Access Toolkit) to communicate with an external Command and Control (C&C) server. An internal host of the organization initiates this communication, because outgoing traffic passes more easily through firewalls.

*D. Lateral Movement*: The attackers typically use legitimate computer features to move within the network undetected. This takes place after the initial breach and the establishment of command-and-control links back to the attacker.

*E. Data Ex-filtration*: Acquiring the appropriate right to access particular data. Once identified, infiltrators can deploy malware extraction tools to steal desired data. In some cases root kits [6] can be secretly installed on targeted systems and network access points to monitor or capture data and commands as they stream over the network. Usually this means creating "white noise attacks" to cover cyber attackers who want to mask their intentions. They also mask their entry point, leaving it open for further attacks. Being persistent is also a key feature for this step to be successful.

## IV. TRADITIONAL ATTACK VS. APT ATTACK

The following Table I illustrates the difference between the traditional attack and APT attack

TABLE I DIFFERENCE BETWEEN TRADITIONAL AND APT ATTACK

| | Traditional Attack | APT Attack |
|---|---|---|
| Reason | Personal or financial benefits, show off | Economic advantages, strategic benefits, stealing sensitive information |
| Target | Undetermined | Governmental institutions, multinational enterprises and banks. |
| Approach | Aggressive, very rapid, smash and grab, tactic based on a very limited time based attack. | Repeated attempts using number of methods, stealth approach, adapts to resist defenses, very slow to avoid any suspicions may involve sleep modes before commencing any attack. |
| Attacker | Usually One person | Highly organized, sophisticated , Determined, highly skilled and no shortage of resources |

*1. Customized Attacks:* Techniques such as intrusion and specifically designed tools are used by APTs for targeted attack. In an APT campaign attackers target zero-day vulnerabilities in the software and plant highly complex root kits, worms, and viruses. APT's are customized to launch multiple attacks on the target simultaneously and take over the whole system. These APT's are so intelligent they trick the target, making the target think that attack has been receded which in reality remains secretly active on targets network.

*2. Covertness Attacks:* To avoid easy detection as in targeted attacks, APT attacker's activities are usually slow and mostly go undetected. The main goal of the attackers would be to stay calm and establish a persistent connection into the target network.

*3. Highly Aggressive Attacks*: In most cases the motive of APT attack is obscure. The reason for obscurity and secrecy is due to the requirements of activities like international espionage, intelligence and information gathering from sensitive organizations like military, political and economic organizations, there by interrupting the operations of the target, by damaging the infrastructure.

*4. Highly Aggressive Attacks*: In most cases the motive of APT attack is obscure. The reason for obscurity and secrecy is due to the requirements of activities like international espionage, intelligence and information gathering from sensitive organizations like military, political and economic organizations, there by interrupting the operations of the target, by damaging the infrastructure.

*5. Specific Targets:* Information Technology has ushered an era of globalized society and its markets. This in turn triggered the need for national and international security because APT attackers can originate from anywhere around the globe. This makes an easy escape for the APT attackers due to localization of legal structures and institutions. Statistics shows majority of the APT attacks are targeted towards government organizations due to inefficient international cyber crime laws.

## V. APT PROPAGATION METHODS

Most APT attacks simulates the normal user behavior and outperforms the installed security measures, they normally follow the slow and steady" policy by compromising the target system. The APTs propagate by many means once they compromise the system by social engineering methods, the attackers monitor the command and control infrastructure which is the indication of successful breach. In most cases of the APT attack the antivirus is not an obstacle. The attackers also use the pre-installed administrative tools and legitimate credentials for the lateral movement which is difficult to detect. The attackers use the evasion technique by which they hide the malicious code which is polymorphic and customized or dynamically modified.

The attackers try to obtain the administrative privileges by installing additional malware or by creating backdoors and install the root kits. The malwares are mostly dynamic, that is the malware signature differs in each attack and they mostly doesn't match with the existing malware database. The malwares can also be with mutation codes.

These attacks even try to modify the event and audit log entries, make the system enter into sleeping state for some time to hack the sensitive data .While gathering the data, they encrypt the network traffic by itself to conceal their true identity, the data is masked to resemble like normal traffic and slowly extract the data from the organization. The encryption and applicative hidden tunnels are used to hide the attacker action by http and https command, fully automated browser, web session to send information and receive instruction from attackers by highlighting the use of cloud services in order to make exfiltration stealthier and harder to notice.

The Command and Control(C&C) helps in routing back to the attacker. They also delete the traces of their own files on the system after successful exfiltration of data.

At the initial state the attackers performs no action they stay observed and evaluate for days or months to initialize the attack. They mostly remain outside the organization perimeter and uses partner or subcontractor infrastructure to lighten the control access to the network. The initial compromise stage is mostly high spear phishing mails, and in lateral movement phase they use the standard OS tools and HTTP and HTTPs commands.

The tools that are used to access the remote system by the attackers are PSExec, Window Management Instruction, Remote desktop protocol and Power shell, often these tools are used for remote administration of the system, which now are used by the attackers to access the remote system. They use custom protocols in which the protocols use encryption or obfuscation technique to hide from protection. The most common propagation methods, information and techniques followed by APT attackers to invade and compromise the network are described below

The attributes that are present in most of the logs collected from the APT activity indicates the presence of APT attack:

1. API resolving
2. File Access
3. File read
4. Registry Access
5. Registry read
6. Random logon

The above attributes is missing in the normal network behavior, but in the APT breached network all these attributes are present. The data attributes that are to be focused in the network to monitor the APT attacks are:

1. Bytes Received
2. Bytes Sent
3. Packet Received
4. Packet Sent
5. Source Port
6. Destination Port
7. Destination IP

*A. Encryption*

Encryption of the network traffic and applicative hidden tunnels hide the attackers from the security mechanisms. This can be performed by HTTP and HTTPs commands, during the web session to send information and receive instruction from attackers by highlighting the cloud services to make the exfiltration stealthier and harder to notice.

*B. C&C Channel:*

The Command and Control channel between the infected machine and the attacker is responsible for sending commands and transferring data between them. The domain name is used to locate the command and control which is performed by the Trojan backdoor and other remote access tools. The malwares that are used for this purpose are Ghost, PC Share, and Poison Ivy. They hide the real attack source by using the servers that are controlled and managed in different countries as proxies. Since the domain names are flexible to change the IP address of the malware C&C servers, they help the attacker to hide the true attack source behind the proxy server. The attacker's hardcode the IP of the C&C server into malware binary which cause some kind of failure that cannot be recovered. If C&C server goes down compromised machine will be out of the attacker's control.

*C. DGA (Domain Generation Algorithm):*

DNS behavioral feature of APT malware are very different from malicious flux service and DGA [11] [12]. It has short life time feature that is DGA domain are used only for a short duration that are extracted from domains generated by DGA. The crafted malware of APT don't use the malicious flux service or DGA domain

*D. Malicious DNS*

The features we extracted from big data for detection consist of malicious DNS features and network traffic features. By studying the DNS traffic, we achieved to extract distinguishable DNS features that are able to define the APT malware. The Domain named features are containing famous domain names, containing some particular names, and containing Phishing names.

The following are the characteristics of the outgoing traffic to discover the malicious activities in the APT attacks:

*1. The Packet Transfer Rate and Quantity:* In a normal network the users mostly send small data such as requesting for a particular file or page to the receiver, whereas in the case of APT the infected machine generates large amount of outbound data which is encrypted and masked, that is unusual in a network. They even adjust the monitoring engine to decrease the number of false positive it might produce. Also there will be large outgoing traffic in regular periods of time.

*2. Malicious Sites:* The outgoing traffic trying to communicate with the malicious site, routing the data to the particular destination IP address with no association with the organization.

*3. Domain Name Space and Cache:* For the command and control creation, the APT registers new domain names, frequently communicate with the unknown new domains and use dynamic DNS to route the traffic between the infected machines which is difficult to track .Whenever the user makes the requests for a particular data, they use URL or hyperlink which in turn makes the DNS resolver in DNS cache to find the matching address, whereas in the case of APT they use pure IP address connection instead of the domain names.

The vulnerability is more in the wireless devices in the near future. From the hardware view of point the microchip which has some access point on purpose to use in post manufacturing tests is the loop hole for the attack. The sleep instructions are executed by the APT before any other activity so to avoid from any suspicion.

## VI. APT DETECTION SCHEMES

Because of the dynamic nature of the APT malwares, neither the antivirus software nor the IDS will have the APT signature in their database, so it is harder to detect the attacks. The following are the detection schemes that can be used to detect the attack.

1. Anomaly based detections schemes
2. Signature based detection schemes.

Signature based detection system are ineffective in detecting APT, they are not scalable to the ubiquitous nature of organization networks, signature lacks the ability to identify completely new attacks or even significant variants of the same attack, therefore, some novel approach is required for combating such attacks. Some of the detection and prevention methods are used to handle the APT attacks are discussed below.

The automatic malware monitoring tools used are to monitor the network and detect the threats in the network:

*A. OSSEC:* The open source host based intrusion detection system which collect, read and analyze the logs, email alerts etc; this operates in two modes namely local and agent/server. The Local monitors only one system whereas the agent/server collects and monitors the logs from multiple sources in the network. The Analysis uses UDP which checks with the preset or manually created rules and works in all Operating System.

*B. Snort:* An open source network intrusion detection and prevention tool which performs packet analysis and collect the logs from the network. They filter generated events to reduce false positives. This can be based on the rate of transfer, number of events as well as the completely hiding events. It works in three modes Sniffer, packet logger, and Network intrusion detection mode. The first two display or store the network packets; the later one performs traffic analysis with predefined set of rules.

*C. Sguil:* They are set of networking tools with Intrusion Detection System (IDS) and Snort which performs network connection security profiling, TCP connection analyzing, and stores the results from multiple tools in My SQL database and provide the GUI about the logs. Since this tool integrates the different sources and analyzes the solution, the detection rate is increased and reduces the work load.

*D. Splunk:* They process the machine generated big data which is being analyzed, indexed ,correlated ,visualized and finally provides the report about the attack. They are widely used in IOT, web analytics etc; It performs statistical analysis within the network from the indexed data. The Splunk can be used with the Machine learning toolkit which uses linear regression for checking the relationship between input and output. The SVM (Support Vector Machine) is used to separate the normal and malicious events and finally Random forest, a supervised machine learning method used as continuous variable decision tree algorithm to identify the exfiltration.

*E. Sandboxing*: It is threat detection and blocking based on the signature which is deployed in the entry point of organization network.

*F. HoneyPots:* The internal resources which are purposefully placed in certain zones to access the network collect the information about the methods used by the attackers to exfiltrate the data stop their spread and also identify the new exploits and threats [5]. The problem with this method is when the honeypots pay attention to any small technical details it can reveal the unreal nature of them to the skilled APT attackers. To overcome this problem the honeypot agents are used which directs the attacker to the system without indicating the presence of security measures in the network. They are used as traps for APT attackers.

*G. Ranking Method:* By analyzing the huge volume of the network traffic, find the weakest spot related to exfiltration and rank the malicious host and focus on small set of hosts, by extracting the features and compare them with other hosts [6].

*H. Detect APT malwares based on malicious DNS*: It is an anomaly based and signature based detection method which finds the malicious APT C&C domains [7]. A reputation engine which computes reputation score for every IP address, categorizes the property of the malware DNS. The malicious traffic will have low reputation rather than the normal traffics which indicate the APT attack. The IDNs used here reduce the volume of network traffic to be analyzed which detects the inbound and outbound traffic based on the signature (with known set of malwares) and anomaly (with unknown or new set of malwares).

*I. Passive DNS Analysis technique:* Attackers use the DNS name along with complex large distributed infrastructure which is normally good services which makes difficult to identify the domain names of the malicious servers [8]. This method finds the APT attackers who identify the malicious domain which helps in extenuating any threat.

*J. SIEM (Security Information and Event Management):* It is the collection of real time analysis of security events which can detect both the internal and external threats [4]. They collect information from various components in the network such as [9] firewalls, antivirus, IDS, IPS, OS, workstation and network devices. The logs are collected, stored, monitored and correlated. The network generates several hundred thousand of logs and events per second this can be stored in QRadarIBM, QILabQradar, and NetIQ Security Manager. The Apache Drill (or) Dremel analyse the streaming data in real time.

*K. Rule Generated Alerts in Splunk SIEM [10]:* It uses RETE algorithm which formulates the rule and store the event attributes in the database. It collects heterogeneous logs such as application log, windows event log, system logs. An alert is triggered when the rule and attack pattern matches.

*L. NGFw (Next Generation Firewalls):* They combine the features of firewalls (packet filtering, website blocking, virtual IP address) with IPS, application control and context protection. The antivirus software used here scans all the packets, also the compressed and encrypted files. The IPS performs in depth monitoring of network traffic equipped with zero day threat minimizing mechanism. The IDS uses exploit signatures whereas the IPS performs detection based on anomaly statistics and vulnerability signature.

## VII. CONCLUSION

In this paper, we discussed the nature of an advanced persistent threat and its impact on the organization. The different phases along with the propagation methods were also studied. The user's must be educated about the good password practices and organization have to maintain strict access and usage policies with endpoint security to all host, by implementing the Network access control and blocking high risk applications. We need various intrusion detection tools at the host level and the network level and correlate them to discover the security breaches covering the whole organization. Since the single technical solution may not be sufficient to detect the threats. The identified indicators can be analyzed and correlated. Intrusion detection system can be chosen wisely based on the type of network traffic in the organization. A comprehensive approach towards the organization's security can protect the organization against multi-vectored attacks.

## REFERENCES


[1] V. N. Harikrishnan, and T. Gireesh Kumar, "Advanced Persistent Threat Analysis using Splunk", Vol. 118, No. 20, pp. 3761-3768, 2018

[2] Mandian - FireEye Inc, "M-Trends 2015 A View from the Front Lines", *Tech. Rep.*, 2014.

[3] N. Virvilis and D. Gritzalis, "The Big Four - What We Did Wrong in Advanced Persistent Threat Detection?" I*n Availability, Reliability and Security (ARES), Eighth International Conference on, Sept 2013*, pp. 248–254, 2013.

[4] Splunk Inc., "Splunk for Security: Supporting a Big Data Approach for Security Intelligence", [Online] Available at: http://www.splunk.com/web_assets/pdfs/secure/Splunk_for_Security.pdf, 2014

[5] Jasek, R.O.M.A.N., M.A.R.T.I.N. Kolarik, and T.O.M.A.S. Vymola, "APT detection system using honeypots", *Proceedings of the 13th International Conference on Applied Informatics and Communications (AIC'13), WSEAS Press*. 2013.

[6] Marchetti, and Mirco, *et al.*, "Analysis of high volumes of network traffic for Advanced Persistent Threat detection", *Computer Networks*, Vol. 109, pp. 127-141, 2016.

[7] Zhao, and Guodong, *et al.*, "Detecting APT malware infections based on malicious DNS and traffic analysis", *IEEE Access*, Vol. 3, pp.1132-1142, 2015

[8] Bilge, and Leyla, *et al.*, "Exposure: Finding Malicious Domains Using Passive DNS Analysis", *NDSS*. 2011.

[9] Anastasov, Igor, and DancoDavcev, "SIEM implementation for global and distributed environments", *Computer Applications and Information Systems (WCCAIS), 2014 World Congress on. IEEE*, 2014.

[10] Raja, M. Siva Niranjan, and A. R. Vasudevan, "Rule Generation for TCP SYN Flood attack in SIEM Environment", *Procedia Computer Science*, Vol. 115, pp. 580-587, 2017.

[11] L. Bilge, E. Kirda, C. Kruegel, and M. Balduzzi, "EXPOSURE: Finding malicious domains using passive DNS analysis", *in Proc. NDSS*, 2011.

[12] E. Stalmans and B. Irwin, "A framework for DNS based detection and mitigation of malware infections on a network", *in Proc Inf. Secur.South Africa (ISSA)*, pp. 1- 8, Aug. 2011.