Attack Vectors and Advanced Persistent Threats

Sergio F. de Abreu, Stefan Kendzierskyj, and Hamid Jahankhani

Abstract Advanced Persistent Threats (APTs) are destructive and malicious cyber-attacks aimed at high profile, high value targets with clear objectives in mind with a range of desired outputs. In most cases, these threat groups are state sponsored which makes them extremely well financed, organised and resourced. The attack payloads range from data exfiltration and theft to the undermining of critical national infrastructure. These attacks differ from the typical cyberattacks in several different ways but a key differentiation is their patient "low and slow" approach to prevent detection. This approach, although slow, has been very successful and in many cases, detection is years after initial infection. Many of the attacks detected today, have been over a decade in the making. Most concerning is the fact that traditional defence mechanisms have been unsuccessful at detecting these attacks and so how successful will these methods be against a new generation of attacks? The earliest recording of an APT is probably "the cuckoo's egg". An attack in the 1980s in which a West German hacker infiltrated a series of computers in California and over time stole state secrets relating to the US "Star Wars" program. The hacker then sold the information to the Soviet KGB. Although at this point in time, cyber defence was not a government sponsored military department, it raised awareness of just how powerful this threat could be. Since then, worldwide attacks in the private and public sectors have grown exponentially and today, all governments have cyber warfare units.

Most APT attacks are state sponsored; however, this does not mean that attacks are limited to government entities. Far from it. These attacks affect individuals, companies, corporations and governments globally. Attacks can and do encompass a multitude of sophisticated techniques and affect not only the traditional LAN/WAN

S. F. de Abreu · H. Jahankhani
Northumbria University London, London, UK
e-mail: Hamid.jahankhani@northumbria.ac.uk

S. Kendzierskyj (✉)
Cyfortis, Worcester Park, Surrey, UK
e-mail: stefan@cyfortis.com

H. Jahankhani et al. (eds.), *Cyber Defence in the Age of AI, Smart Societies and Augmented Humanity*, Advanced Sciences and Technologies for Security Applications, https://doi.org/10.1007/978-3-030-35746-7_13

environments but could also contaminate new generation networks such as mobile 5G networks, vehicular ad hoc networks (VANET) and Internet of Things (IoT) to name but a few. Dealing with these attacks is challenging, most attacks take years to be discovered and traditional detection mechanisms have been woefully inadequate. The age of machine learning and artificial intelligence has brought significant improvement to the detection challenges faced. These fields allow us to look for far more than attack signatures and characteristics. They allow us to look for patterns of behaviour through massive data quantities at speeds previously unimaginable.

Keywords Advanced persistent threats · APTs · Malware · Machine learning · Artificial intelligence · Threat actors · Cyberattacks

1 Introduction

In June 2010, a cybersecurity researcher named Sergey Ulasen, discovered a malicious computer worm. This worm, codenamed Stuxnet, is thought to have been in development since at least early 2005 and is still regarded as one of the most sophisticated APTs ever seen. Stuxnet's purpose was to sabotage the Iranian nuclear program and reportedly ruined almost one fifth of Iran's nuclear centrifuges causing enough physical damage to the infrastructure to set the entire program back 4 years [6]. This malicious worm was part of what we now know and call an Advanced Persistent Threat (APT).

An APT can be described as a prolonged persistent cyber-attack in which access to a network is achieved but remains undetected over a long period of time. The attackers go to extraordinary lengths to avoid detection. The threat infiltrates the network of choice using a multitude of different attack vectors and once access is gained, advanced methods are used to avoid detection while increasing their foothold on the overall network. These attacks are then used to exfiltrate data, control systems and in some cases destroy infrastructure.

The complexity and cost of APTs suggests that in the vast majority of cases the attacks are specifically targeted, well-funded, resourced and patient which has led to a general consensus that they are state sponsored. According to a recent review of top threat actor groups and the countries they operate from [20], North Korea, Russia and Iran currently list in the top three.

It is widely accepted that the Stuxnet worm was part of an APT attack engineered by both American and Israeli intelligence, although this was never officially confirmed by either country the fact remains that this attack very successfully and significantly damaged the Iranian Nuclear program without the need for any physical military involvement.

Another APT codenamed Duqu, a derivative of Stuxnet suspected of either being created by the same organisations or at least a group with access to the original source code was discovered in 2011. This APT's payload was not to directly cause any damage but rather to gather information specifically around industrial control systems. One of the vital parts required in a sophisticated attack involving different phases of attack.

Traditional attacks tend to try achieve immediate and fast access to a target. The attack is carried out and once the objective is met, the attacker leaves with no clear plan or intention of returning. While APT's often use many of the same techniques to infiltrate a target network, their primary focus is to avoid all detection systems, gain a foothold and begin to spread across the network to ensure that if a compromised node is detected, they still have access to the network via one of the other infected nodes. This allows them to spread slowly and quietly ensuring that they go undetected while they go about their intended attack. A successful attack will not necessary mean that they will leave, if undetected, they will keep their foothold to either use at another point or even sell off to another adversary.

It is important to remember that the threat of APTs wouldn't be restricted to the traditional LAN/WAN network environment but could also be utilised on any type of network. This would include both Internet of Things (IoT), and Vehicular Ad Hoc networks (VANET) infrastructures posing a serious threat and risk to any network.

2 Advance Persistent Threats (APTs)

2.1 What Is an APT

An APT could be defined as a series of both basic and advanced malicious techniques and methods used in conjunction to build an attack which not only grants an attacker access to a victim network but expands and maintains access over a long term to ensure that as much valuable data and malicious damage can be done with the minimum chance of detection.

The attacks differentiate themselves from traditional threats in that:

- The attackers are highly organised, sophisticated, determined and operated by a well-resourced group.
- The targets are specific.
- The purpose is strategic.
- The approach is one of repeated attempts, stays low and slow, adapts to resist defences and is generally long term.

The National Institute of Standards and Technology (NIST) defines an APT as:

> An adversary that possesses sophisticated levels of expertise and significant resources which allow it to create opportunities to achieve its objectives by using multiple attack vectors (e.g., cyber, physical, and deception). These objectives typically include establishing and extending footholds within the information technology infrastructure of the targeted organizations for purposes of exfiltrating information, undermining or impeding critical aspects of a mission, program, or organization; or positioning itself to carry out these objectives in the future. 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. [17]

2.2 *The Actors*

The vast majority of APT attacks are state sponsored. Looking at currently identified and tracked APTs, their objectives and the groups known to have orchestrated them, and it quickly builds up a picture of the top 6 countries in which the actors operate from, namely:

- North Korea
- Russia
- Iran
- India
- Russia
- China

In a 2018 report by AlienVault [20], the top ten most reported active threat actor groups and their locations were as follows in Table 1:

The Lazarus group, also known to united states intelligence as "Hidden Cobra" is widely accepted to be sponsored and controlled by the North Korean government. This group's primary focus are attacks within the financial markets. One of their campaigns nicknamed "FASTCash" was responsible for large amounts of theft from ATMs in both Asia and Africa with an attack, which started in 2016, and is still ongoing. In 2018, the US department of homeland security (CISA) issued an alert to this effect. On the 10th of April 2019, CISA released another alert attributed to the Lazarus group [7]. This alert details a piece of malware which has the ability to connect to a command and control server in order to transfer stolen files from an infected network.

The Malware, known as "Hoplight" masks traffic between the victim and the remote server by acting as several proxy applications.

Table 1 10 most reported APTs

Rank	Advanced persistent threat	Location
1	Lazarus Group	North Korea
2	Sofacy	Russia
3	Muddy Water	Iran
4	Oil Rig	Iran
5	Patchwork	India
6	Energetic Bear	Russia
7	Kimsuky	North Korea
8	APT 15	China
9	Stone Panda	China
10	Turia	Russia

According to the alert, "The proxies have the ability to generate fake TLS handshake sessions using valid public SSL certificates, disguising network connections with remote malicious actors." [7]. North Korea's backing for the Lazarus group falls outside of the typical state sponsorship for the purpose of espionage and intellectual property theft. The objective of this group is purely financial gain, which when one looks at the severely isolated and cash starved state, it is clear why this group is so critical.

The Sofacy group also known as Fancy Bear is highly suspected of being sponsored by Russian military intelligence. In 2018 an indictment by Robert Mueller, the United States special council looking into Russian Interference in the United States 2016 presidential election, identified the Sofacy group as two GRU (Main Directorate of the General Staff of the Armed Forces of the Russian Federation) units knows as Unit 26165 and Unit 74455.

This group has been operating since around the mid 2000s and specifically targets government, military and security organisations.

One of the groups attributed attacks was an attack on German parliament in 2014. Specifically, the government's "Informationsverbund Berlin-Bonn" (IVBB) network, which is a separate and private network used by the Chancellery and Federal Ministries. Ironically, this network was setup separately from other public networks to ensure an added layer of security.

The Dutch Government also accused the group of data theft from the Organisation for the Prohibition of Chemical Weapons (OPCW) in The Hague and most recently and famously, this group has been specifically mentioned in ties to the 2016 American election meddling investigation. Their primary target is and has always been NATO member states.

Clear actor identification can be challenging. Various vendors and intelligence agencies often name the threat actors differently which can lead to some confusion within the market. Some naming conventions are designed to create a mythological or figurative emotion, others are just naming tags given for the sole purpose of identification, yet others are just named after specific malware that that was used in an attack. A further key reason for differences is that threat actors could occasionally join and then split up causing further confusion on the actual threat actor responsible.

An example of the varied naming conventions could be the APT group "Comment Crew" [10]. This Chinese group, attributed to the second Bureau of the People's Liberation Army (PLA) is named "Comment Panda" [5] by reseller Crowdstrike, "PLA Unit 61398" [8] by reseller IRL, "TG-8223" by Dell Secure Works "APT 1" [10] by reseller Mandiant and even "brown fox" by reseller iSight. These differences in naming can be confusing and there are calls for standardisation but it's just not that simple. There are technical and "people" reasons why certain vendors use certain naming conventions.

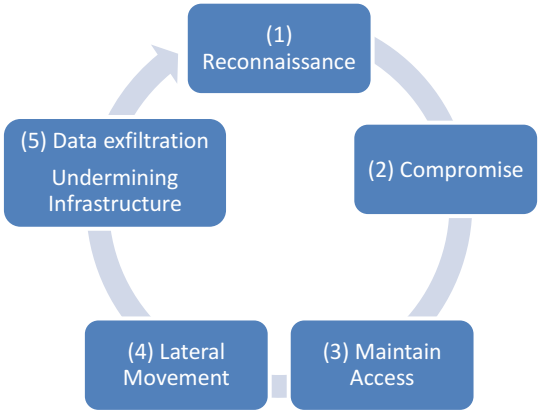

Fig. 1 Typical APT Lifecycle

2.3 APT Lifecycle

The typical APT lifecycle can be split into several different phases (see Fig. 1). Although various researchers break down the steps differently [4] ([22, 29], they all essentially break an attack down into five distinct steps.

Reconnaissance
Once attackers have identified the target and a strategy for attack, they need to research the target so that they are completely familiar with the people, systems and processes that are used. This reconnaissance would typically include both physical and passive cyberattacks in an effort to gather as much information as possible.

The people aspect of the reconnaissance would not necessarily only be staff but could include contractors, vendors and partners. These reconnaissance missions often employ large numbers of researchers and can involve a significant amount of time and cost and are almost always passive to ensure no red flags are raised. If the Stuxnet attack on the Iranian nuclear reactors is reviewed, it can be understood that the attackers had expert knowledge of the internal systems used and critically the Siemens programmable logic controllers (PLCs) used on the centrifuges within the facility. This is no small feat and would have involved significant research and knowledge.

With this knowledge, attackers would then need to identify an initial entry point to the network. This point would not only be the easiest path to entry but also the point where an attack would stand the best chance of going undetected. Wherever possible, multiple points would be targeted to ensure success.

Compromise
In this phase, the attacker crafts an attack with the sole purpose of infecting a victim's machine. This is commonly in the form of a socially engineered attack with spear phishing and watering hole attacks being the preferred route [22], but

could really be any available resource to the attacker. The attack could even come indirectly through a third party which is trusted by the victim.

Again, in the case of Stuxnet, it is suspected that an infected removable disk storage unit inadvertently plugged in by a staff member was used to distribute the attack. [6]. Analysis of Stuxnet shows us that four zero-day exploits were built into the malware. This is a massive number in comparison to all other APT attacks. The attacks are well crafted and designed to bypass traditional Intrusion Detection Systems (IDS) while the exploits used are often zero-day attacks that any proactive level of patching would not help to prevent [2].

Internal staff are often regarded as the most cost-effective way to infiltrate the network and this is seen by the amount of attacks targeting end users directly.

Maintaining Access

Maintaining access and lateral movement are really the two phases which set an APT apart from other typical opportunistic type attacks.

Once the attacker has managed to compromise an internal system, in almost all cases its vital that a back door is installed to continually maintain a level of access to the infrastructure. To do this, a Remote Access Trojan (RAT) is installed on the victim machine/s as described by [4, 22].

Once the attacker has created the backdoor to the network, they would then proceed to compromise several other machines thereby ensuring that access can be maintained even if one of the compromised systems are discovered or indeed just taken offline. The RAT will then make a connection to an external Command and Control server (CnC). This CnC server then dictates to the RAT what should be done on the victim machine/s. This would explain how [16, 22] the connection from the RAT to the CnC server will in almost all cases be initiated from the RAT outwards to the CnC. This is done to help hide the traffic and bypass typical security controls, as most networks are configured to be far more lenient on outgoing connections than incoming traffic.

Lateral Movement

APTs operate in a “low and slow” method, gaining access slowly and carefully and spreading their connectivity from within the network.

In this phase, the attacker would be able to perform internal scans to map out traffic routes and other hosts within the network segment. Details of the environment, systems, functions and processes are discovered, both hardware and software vulnerabilities, unprotected network resources and additional access points to the network are mapped. Although internal scans could be detected, the lateral movement is often not, due to the use of compromised valid credentials already obtained as detailed by [22]. Since an APT’s main goal is to gain access and remain in the network for an extended amount of time, every method and technique used is built around avoiding detection. One example of the techniques used in an attack is operation Aurora, also known as Hydraq or the Google hack attack. This attack originating in China [9], used an old obfuscation technique called spaghetti code to

help keep itself hidden from network protection mechanisms. This was originally recognised as an inefficient and unstructured way of coding which was highly discouraged but was used to great success when the coders were after exactly that effect.

Moving laterally within a network allows the attacker to access and infect further endpoints over time using the elevated privileges gained in earlier steps to access targeted data/systems.

Data Exfiltration
This is the final stage and the objective of the attack. However, this stage does not have to only be about data exfiltration; it could be about undermining critical aspects of the targeted infrastructure as described by [17]. Data exfiltration mentioned by [22] and collaborated by [16] and could be executed in many different ways:

- Encrypted or clear data could be exfiltrated to the CnC server(s). This could be done from one or multiple victims to either one or multiple CnC servers. The advantage of exfiltrating data in an encrypted format would make it even harder for intrusion detection and data loss prevention (DLP) systems to detect the data loss.
- Although data could possibly be exfiltrated all in one go but with the intention of longer-term access to the victim needing to be maintained, very low and slow levels of data leakage would help prevent being detected, successfully exfiltrating data and maintaining access for future use.
- Steganography is a technique that could be used to insert the data into an image which could be displayed as a .jpg file as was the case in the *Duqu* APT [34]. This would appear as normal day to day typical use by a user which would be very difficult to identify as anything malicious [14].
- Physical human intervention could be used to gather the exfiltrated data from a defined location. One way this could be accomplished would be a technique called “dead letter box”.

A recent example of successful data exfiltration is represented by the Equifax data leak in 2017 [12, 23] in which 147 million customers sensitive personal information was leaked.

3 Attack Examples

3.1 How Did They Do It?

Looking at two examples, Stuxnet and Lazarus Group, of well-known and successfully implemented APT attacks, we can analyse exactly how these attacks were carried out in each of the five phases to build a complete picture.

3.1.1 Stuxnet

One of the most sophisticated and precise APTs ever detected. This attack was very precisely aimed at Iran's Nuclear plant, Natanz (see Table 2 for attack phases and its descriptions).

Table 2 Stuxnet attack phases and descriptions

Attack phase	Description detail
Reconnaissance	The Stuxnet worm was targeted at very particular and specific Siemens Programmable Logic Controllers (PLCs). The worm was so well written, it required absolutely no intervention from any internal staff to work. A simple plugin to a USB drive was all that was necessary. To achieve this level of functionality the attacker would have to have detailed information of the network, infrastructure and centrifuges.
Compromise	The Natanz plant was air gapped from the internet. It was not possible to attack it directly from the internet however it is widely accepted that the Stuxnet worm was introduced into the plant via a USB key. It is not known whether this was done accidently by staff or deliberately.
Maintain access	Stuxnet was targeted directly at certain logic controllers controlling centrifuges within the plant. It was so specific that while it was programmed to spread from machine to machine, it was coded to search for certain hardware components and if they were not found, no action at all was taken. The worm would lie dormant taking no further action. Additionally, the worm was designed to self-destruct on the 24th of June 2012. In most cases, APT's establish a connection to the outside world by installing a remote access trojan (RAT) on the machine, however in the case of Stuxnet the attackers knew that it would not be possible for a RAT to communicate with the outside world once deployed so the worm had to be completely self-sufficient and run without waiting for any external instructions. An incredibly hard task to accomplish.
Lateral movement	This worm was specifically written to spread at a rapid pace using four in-built zero-day attacks to ensure that it would be able to achieve its target. Although traces of Stuxnet were found on systems all over the world, the biggest concentration of infections were all over Iran. It's important to consider that the worm would take absolutely no action on any machine that didn't have the correct Siemens controller software on it.
Data exfiltration undermining infrastructure	The payload was to destroy centrifuges in the plant. To achieve this, Stuxnet made the centrifuges spin dangerously fast for a short period of time but critically had already infected the monitoring systems within the plant to not detect this change. Although engineers could hear that the centrifuges were spinning dangerously high, the control systems indicated that all was within normal parameters. About a month later Stuxnet then slowed the centrifuges down dramatically for around 50 minutes, again with all control systems showing the plant running within perfectly normal operation parameters. The dangerous repetition of this caused over 1000 centrifuges (around 20%) at the plant to collapse.

3.1.2 Lazarus Group – Financial Threats

Founded in 2009, the Lazarus Group, a very active North Korean sponsored threat group best known for their attacks specifically targeted around financial gain. They attack the world cryptocurrency exchanges, financial institutions and banks Although this is not their only attack profile. Below is a high-level look at one of their most recent attacks on a Chilean organisation called Redbanc (see Table 3 for attack phases and its descriptions).

3.2 Detection Challenges

The sophisticated nature of APT's means there are significant challenges in detecting them. At every stage of their typical lifecycle, everything possible is done to avoid detection.

The reconnaissance is detailed, well-funded and passive to avoid any means of detection while the compromises take any and all approaches necessary from physical infiltration to cyber hacking. In most cases, multiple zero-day attacks are utilized to prevent being detected by traditional intrusion detection systems (IDS) [6], also rendering both system patching and signature based anti-virus and malware detection useless [18]. Messmer [19] and Kruegel [24] argue that even *Sandboxing*, an often used and preferred malware detection method can by bypassed by skilled and well-funded adversaries using methods such as, environment-specific-techniques, human-interaction-techniques, VMware-specific techniques, and configuration-specific-techniques. Using these detection avoidance techniques has led to a 200% rise in malware capable of evading detection [19]. The persistent nature of these attacks means that even in cases where a completely isolated system is enforced, the victim could still be physically compromised by being influenced into plugging a removable media drive into an internal system (USB drop attack) [30].

As previously discussed, maintaining access to the victim is a key aspect of the persistence of an APT. Data exfiltration or undermining the infrastructure can only happen when the correct targets are identified and compromised. This process can take a significant amount of time hence the need for access to be maintained. This is accomplished using external CnC servers which use various techniques to maintain access to the victims while avoiding detection. These methods as described by ([1, 6] include but are not limited to:

- Remote Access tools (RAT) which are often used in day to day business use and make use of a server and client agent.
- Social Networking sites that the victim's machine goes to which could put control information into blog posts and status messages
- TOR Anonymity Networks which by their very nature are designed to hide services and traffic.

Table 3 Lazarus group attack phases and descriptions

Attack phase	Description detail
Reconnaissance	Redbanc is a Chilean company whose business is responsible for all interconnectivity between the ATM infrastructure in the country. To gain access to the network, attackers created a front company and posted a job opening on LinkedIn for a developer position within the company. At that point, they were not sure who would apply for the job. An employee of Redbanc saw the posting and applied for the position.
Compromise	Once the employee had applied for the job, the group arranged a video conference interview over Skype™ and in that interview was asked to download and run a file that would help with the recruitment process seen below in Fig. 2 [11].
	The file appeared to generate a standard job application form, but this file called ApplicationPDF.exe was in fact a Microsoft Visual C#/ Basic .NET (v4.0.30319)-compiled executable file which infected the employee's computer with a piece of malware called PowerRatankba. This malware, allowed the attackers to gain information about:
	The hardware
	Operating system
	Running processes
	RPC and SMB file shares
	Computer name
	User name
	Proxy settings
	Through this compromise, the attackers were able to get further reconnaissance of the target and decide if the other stages of attack would be of value to them. The attackers clearly decided that this was a desirable target.
Maintain access	As well as feeding back information about the target computer, the malware constantly reports on the status of its own remote connection the attacker. The malware gives the attacker the ability to delete the malware from the victim machine, modify and replace ps1 and VBS files, send data to a chosen destination server and download an executable to run via PowerShell. This is archived through its support for several different commands [26].
Lateral movement	The ability to upload further executables from the attacker to the victim gives the attacker many different opportunities to not only maintain access but also spread infection through the network. With the reconnaissance information gained in step one, the attacker knows the machine type, operating system and running processes on a standard staff desktop thereby giving them vital information on the standard company installation profile. Information on running processes is extremely valuable as it allows the attacker to build a profile on any security measures and software running on the machines. This includes specific firewall and anti-virus tools.
	In the case of the attack on Redbanc, infection spread to a significant number of machines.
Data exfiltration	Exact financial losses are not clear as Redbanc has never released any information regarding this however, other attacks by the Lazarus group on ATM infrastructure in Asia and Africa are well documented.

(continued)

Table 3 (continued)

Attack phase	Description detail
Undermining infrastructure	The joint FBI, DHS and Treasury US-Cert technical alert report details the FASTCash scheme used against ATMS. "FASTCash" *schemes remotely compromise payment switch application servers within banks to facilitate fraudulent transactions. The U.S. Government assesses that HIDDEN COBRA actors will continue to use FASTCash tactics to target retail payment systems vulnerable to remote exploitation."*
	"According to a trusted partner's estimation, HIDDEN COBRA actors have stolen tens of millions of dollars. In one incident in 2017, HIDDEN COBRA actors enabled cash to be simultaneously withdrawn from ATMs located in over 30 different countries. In another incident in 2018, HIDDEN COBRA actors enabled cash to be simultaneously withdrawn from ATMs in 23 different countries."
	As previously mentioned, the US government defines the Lazarus group as Hidden Cobra.

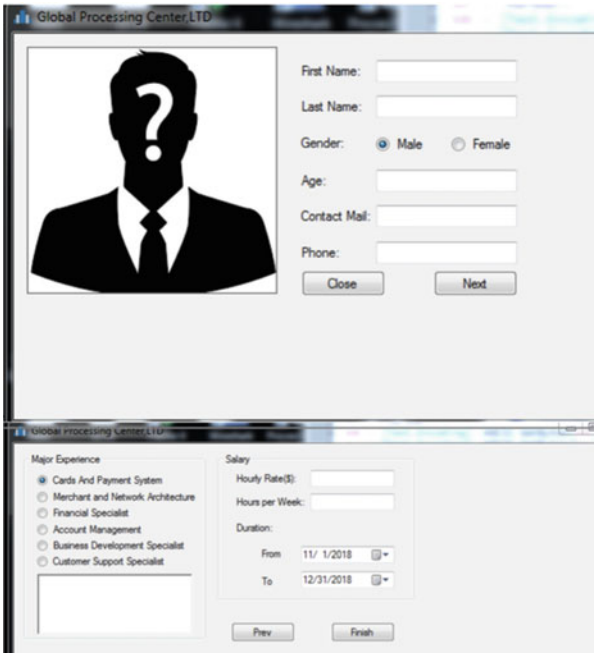

Fig. 2 Redbanc fake job application

The ability to move laterally is arguably the most dangerous phase of the attack and almost certainly the most time consuming. In this phase, the attackers remain undetected by often making use of built in Operating System (OS) features and utilities whose use cases would not look out of the ordinary to any security software. By using these in-built tools, internal reconnaissance would allow the adversary to obtain information about additional systems, network structure, network drives, security software used and network security detection systems. A key part of

this phase would be the ability to harvest user credentials, particularly those with elevated access rights. The use of authorised access credentials would generally not flag as suspicious to the internal systems unless accounts were used in multiple locations at the same time. Data exfiltration can be accomplished using low and slow techniques like DNS tunnelling as described by [28]. This technique when done slowly and making use of custom coding is very difficult to detect. Exfiltrated data is compressed to limit the size as much as possible. The data is additionally encrypted using SSL/TLS to restrict the type of scanning that can be performed masking the data and the communication channel. The use of TOR networks is often used to accomplish this.

There are three factors that any successful APT requires:

- The attacker must have the ability to execute their malicious code on a machine(s) within the target environment. This would include individual vehicles in VANET
- The attacker requires the ability to CnC the machine(s) on the target environment and this ability has to be maintained. There must be the ability to get messages in and out of the target network.
- Lateral movement requires that the attacker is visible. If they have valid network credentials, this is hard to detect but they will be visible.

3.3 How Do We Detect APT's Today

As discussed at length already, there are significant challenges with APT detection however significant research on this problem has been done and researches have discussed various different detections methods to deal with this issue.

3.3.1 Network Sensors

Bhatt et al. [3] argue that effective detection of APTs is only possible with network sensors which can detect all attack facets. Further to this [27] finds it is necessary to continuously monitor and analyse features of a TCP/IP connection. These include:

- Number of transferred packets
- Total count of the bytes exchanged
- Duration of TCP/IP connections
- Information on the number of packet flows

Bhatt et al. [3] suggests a method for detection is to install sensors in each layer of the network. All alerts and logs would then be collected and stored. Correlation of data for each layer could then be performed and this would assist in identifying attacks in progress. An issue highlighted with this approach is the sheer number of logs which are typically generated in all the layers of attack. Hale [13] and

MacDonald [21] point out that in a typical network of 100 hosts, one can expect around 100GB of logs and alarms a day. If we consider a typical network with varying node density, mobility and a constant increase in users, analysing this volume with current methods would be extremely challenging. Another proposed technique used to detect attacks is honeypots.

3.3.2 Honeypots

Jasek et al. [15] propose a system of detecting APTs using honeypots, a system or network of systems (honeynet) whose sole purpose it is to attract attackers and then record their activities. The proposal makes use of high and low interaction honeypots as well as separate honeypots on production systems. Jasek et al. [15] argues that traditional honeypots are limited in that they are passive and wait for the attacker. It proposes making use of an agent which is installed and directs the attacker to the honeypots. The engagement is a 5-step process as follows:

1. Connect the system of Honeypots to the production environment using low and high interactive honeypots and activated agents.
2. The attacker compromises a production client and, in their reconnaissance, discovers shared resources on other systems (honeypots)
3. The attacker gains access to the honeypot systems and compromises them.
4. The attacker collects data from the compromised systems and honeynets and sends the information out to the CnC server externally.
5. The administrator detects the compromise from the honeypot systems and the traffic outflow.

With the attacker activity logged and monitored, the administrator(s) is then fed this information. The administrator is then theoretically able to apply rules and procedures to defend against the attack on the production environment (Fig. 3).

While honeypots unquestionably increase our understanding of malicious network activity and provide an interesting option for detection of malicious activity, there are several issues that are raised with the use of honeypots. Questions around the legality and privacy of honeypots exist; collection and monitoring of user information, malicious or not could fall foul of privacy laws. Sokol et al. [32] highlights privacy issues within the European Union (EU) while [25] addresses the

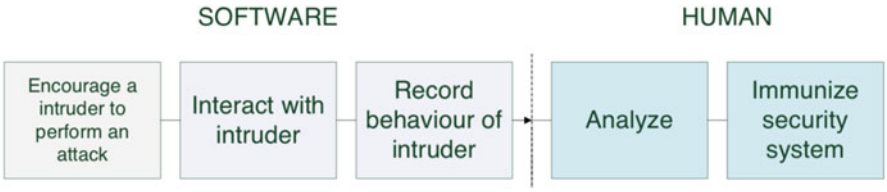

Fig. 3 Honeypot interaction model

same concerns from the legal jurisdiction of the United States of America (USA). There are also concerns around the risk of honeypots and how an attacker realising that a honeypot is being used could then compromise the honeypot in such a way as to attack, infiltrate or harm other systems or organisations [33]. Another prominent proposed detection method is that of machine learning (ML).

4 Machine Learning and Artificial Intelligence

4.1 Current Detection Methodologies

Typical security mechanisms do not adequately address APTs in in this new highly mobile, varied and complex ad-hoc type network world. It is impractical to think that human intervention and detection skills could solve the challenges presented in such a complex and completely ad-hoc network especially when one considers that in certain cases no input or information is available about the attack at all. In such cases unsupervised Machine Learning techniques (ML) are seen as a solution which could deal with this threat. Machine learning techniques can generally be split into two different approaches. Artificial Intelligence (AI) and Computational Intelligence (CI) [35] AI techniques have their roots in traditional methods like statistical modelling while CI techniques are most commonly based on nature-inspired methods that are used to deal with challenges that classic methods are unable to solve. CI methodologies include but are not limited to evolutionary computation (genetic algorithms), fuzzy logic, artificial neural networks (ANN), artificial immune systems (AIS) and swarm intelligence (SI). "*AI handles symbolic knowledge representation, while CI handles numeric representation of information*" [35]. Although it's not always easy to distinguish the boundary between these two broad categories. Hybrid methods are possible and sometimes proposed but generally speaking are used independently of each other.

Fractal dimension-based machine learning is one such possibility proposed by Siddiqui et al. [31]. The authors present a correlation algorithm which makes use of fractal dimensions to detect APT based anomalous traffic patterns with high accuracy and reliability using a feature vector obtained through the processing of TCP/IP session information.

The feature vector selected is based on two metrics:

- Total data packets transferred during a single TCP session
- The duration of a complete TCP session.

The researcher's analysis of TCP data concludes that APT traffic consists of a small count of data packets in a short or long-lived TCP session, whereas normal internet traffic exhibited patterns of a large amount of data packets in a short duration. This is consistent with the APT low and slow exfiltration method already discussed.

The basic requirement of the algorithm is an accurately labelled reference dataset of the features. Each data point is classified as anomalous by comparing the correlation fractal dimensions of the corresponding dataset.

The algorithm first calculates the correlation fractal dimension of the attack and normal reference datasets separately, and then forms a prototypical measure for each class. To classify new input samples, the methodology computes the correlation fractal dimension of the new samples with the reference data set and compares that, to the prototypical measures of the normal and attack data sets. The class for which there is a minimal change in the fractal dimension, indicates that, the point belongs to the particular class. This can also be regarded as finding the similarity index of the new sample and choosing the class to which the input is most similar. This methodology has proven more effective at reducing both false positives and false negatives.

Paredes-Oliva et al. [27] has proposed a novel scheme which also makes use of ML techniques to detect anomalies in traffic patterns. The authors make use of a combination of both frequent item-set mining and decision tree ML techniques to accomplish this and while not directly looking at APTs, such classification would detect anomalies which could then be classified as required. The authors argue that most anomaly detection systems differentiate between normal traffic and anomalies but they do not distinguish different anomaly types which is a key focus of the proposal. The authors first analyse a large set of flows for one or more flow features in common. This is called frequent item-set mining (FIM). An example of this would be a typical network scan; this will produce many separate flows with the same source IP address and destination port. After applying FIM, the result would be one frequent item set with two items: the scanner IP address and the scanned port number. The scheme then builds a decision tree to classify the FIMs as benign or anomalous. Once this process is complete, the anomalies could then be classified by specific type. Figure 4 visually illustrates this process.

Using this methodology, the authors were able to simultaneously monitor two high volume 10Gb/s links and maintain a classification accuracy of 98%.

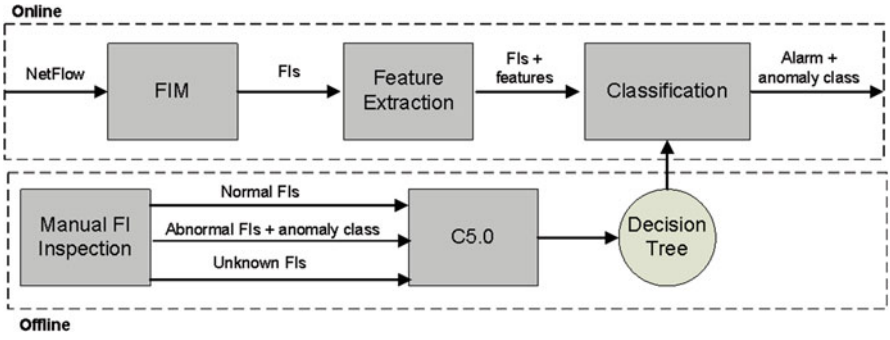

Fig. 4 Anomaly detection system overview [27]

This opens up the question of how does a machine learning classifier begin to identify an attack?

4.2 Attack Visualisation

If we take a standard dataset of benign network traffic and then randomly inject several APT attacks into it, we have the opportunity to analyse these flows and visualise just how the attacks integrate into the traffic.

Taking five separate attacks approximately 5 Mb in total size and injecting this into a 4.4GB standard benign network traffic dataset, we can extract each bidirectional data flow and analyse several attributes of the flows. Breaking these streams down results in 137 APT data streams amongst 7703 benign data streams. A total of 1.78% of the total data.

If we then extract some of the individual attributes of the streams such as:

- Flow duration
- Total forwarded packets (per flow)
- Total backward packets (per flow)
- Maximum forward packet length
- Minimum forward packet length
- Mean forward packet length
- Flow Bytes per second
- Flow packets per second
- Backward packets per second
- Standard packet length
- Down/Up ratio
- Average packet size
- Backward segment size average
- Average forward Bytes/b
- Label (Manually labelled as attack or benign).

It is then possible to view how these attributes are seen by a machine learning classifier. We do this by using WEKA, an application written by the university of Waikato which has built a collection of machine learning algorithms on a single platform to simplify the task of data mining using machine learning classifiers.

Figure 5 is how this data analysis displays in WEKA. The red dots are the benign data streams while the blue dots are the attack data sets. This very clearly highlights the characteristics of the typical low and slow APT data transmission. The duration of flows is much lower over the entire time period under analysis. This, as discussed, is one of the methods used by APTs to avoid detection by traditional intrusion detection systems.

A further illustration of this can be seen in Fig. 6 where average packet sizes are illustrated by grouping them by size over the same duration. A large percentage of the APTs are recorded in the lowest packet data size hidden amongst benign data

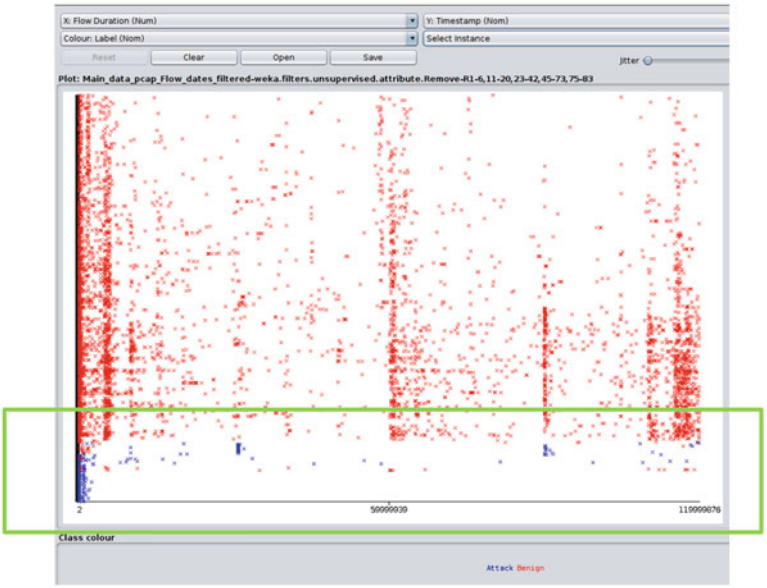

Fig. 5 Visual representation of flow duration typical of APTs

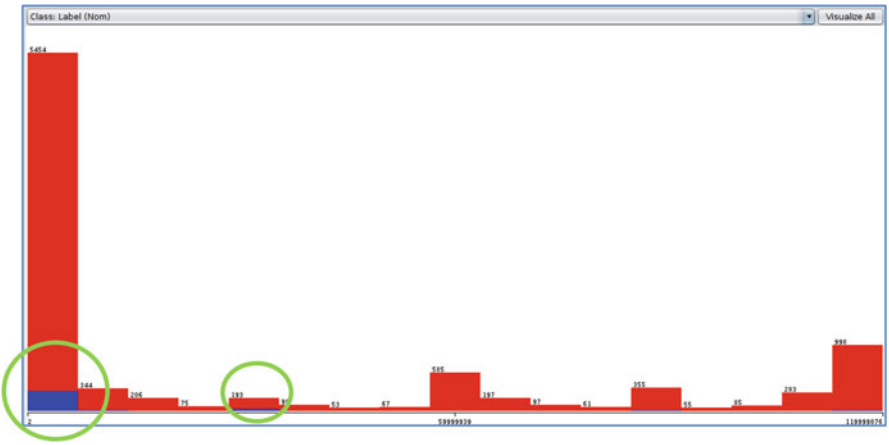

Fig. 6 Visual representation of packet size grouping

flows of the same nature. This grouped with the short flows shows just how data is transmitted, slowly over short periods and small sizes making it very difficult to detect.

4.3 Analysis

Although extremely challenging to detect, there are techniques which can be utilised that give a higher chance of detection. The attacks are sophisticated and well-crafted and often include components that traditional intrusion detection systems (IDS) do not detect.

Too many techniques are passive and look for particular signatures which only work when the attack types have been identified before. To add to this, the volume of data and logs created on a standard corporate LAN/WAN network is staggering. The ever-increasing quantity of data really does make detection a case of finding a needle in a haystack and a fact that attackers rely on.

One successful technique in this detection challenge is searching for suspicious behaviour but the key to this is that it has to be done in the absence of a baseline. One cannot simply analyse a network, assume it's clean and then create a benchmark based on that to analyse future traffic. Fundamentally, it can never be assumed that a network is clean and free from contamination. Applications vary greatly and there is a constant introduction of new and upgraded network components which create an ever-changing network traffic profile.

Honeypots, as mentioned earlier in this chapter, might help to detect an attack but this is a passive approach that doesn't allow for real time analysis and detection and can be extremely difficult to implement in a sophisticated network architecture. They do however help to build an overall knowledge of attacks which in turn helps to identify characteristics that attacks might have in common.

APTs use a combination of techniques and methodology to attack a victim and these will vary depending on who the victim is. Equally, successful defence against this type of adversary will require a combination of differing techniques. A one shoe fits all approach will not work and a consolidated approach will produce better results.

5 Conclusions

Advanced Persistent Threats are an attack type which cannot be underestimated and must be taken seriously. They are hard to detect, prevent, and if infected, to remove. No industry is immune from attack and APT is agnostic to any organisation type.

Reconnaissance of the target is detailed and effective and because most attacks are state sponsored, they are well funded and resourced. The attacks in themselves are specific, with clear objectives in mind.

Attacks are patient and run through several different phases from reconnaissance, compromise, lateral movement and eventually payload delivery. These attacks can take years to deliver their complete payload and all the while, the victim is completely unaware that they are infected. From intellectual property and financial theft to critical infrastructure destruction, the threat is real and applies to all

industries and network types and this ‘low and slow’ type attack is what makes this highly dangerous.

When considering the threats, landscape and attack types, attack consequences could be life threatening and devastating. An example of this could be a well-orchestrated attack on an autonomous vehicles VANET where a vehicle is taken over and maliciously used, but there are other attacks on VANET we could consider of a less severe nature where a vehicle could be infiltrated and the cars inbuilt microphone used for handsfree communication compromised, allowing the attacker to listen and record all conversations within the car over an extended period of time. This could be a source of invaluable information to the attacker.

Detection of these attacks using traditional techniques and intrusion detection systems is extremely challenging. A well-crafted attack making use of zero-day exploits used in conjunction with detailed knowledge of the target’s internal systems as in so many recorded cases can infect a network for years.

Real time Identification of suspicious behaviour in large data volumes can successfully be accomplished by systems which implement some form of machine learning classifiers. Human detection alone is impossible. While various detection methodologies have been researched, it is clear that the key lies in the accuracy of the detection and on how refined the classifiers are and how they are adapted to the data type. It is critical to keep false positive results as low as possible to avoid confusion. Artificial Intelligence might allow these classifiers to keep adapting and developing their algorithms as threats advance in this area and continued research in AI and ML may prove to provide beneficial outcomes.

References

1. Adair S, Deibert R, Rohozinski R, Villeneuve N, Walton G (2010) SHADOWS IN THE CLOUD: investigating cyber espionage 2.0|online safety & privacy|computer security. [online] Scribd. Available at https://www.scribd.com/doc/29435784/SHADOWS-IN-THE-CLOUD-Investigating-Cyber-Espionage-2-0#. Accessed 14 June 2018
2. Ben-Asher N, Gonzalez C (2015) Training for the unknown: the role of feedback and similarity in detecting zero-day attacks. Proc Manuf 3:1088–1095
3. Bhatt P, Yano E, Gustavsson P (2014) Towards a framework to detect multi-stage advanced persistent threats attacks. In: 2014 IEEE 8th international symposium on service oriented system engineering
4. Brewer R (2014) Advanced persistent threats: minimising the damage. Netw Secur 2014(4):5–9
5. Cdn0.vox-cdn.com (2014) crowdstrike-intelligence-report-putter-panda.original.pdf. [online]. Available at http://cdn0.vox-cdn.com/assets/4589853/crowdstrike-intelligence-report-putter-panda.original.pdf. Accessed 8 Sept 2019
6. Chen P, Desmet L, Huygens C (2014) A study on advanced persistent threats. In: Communications and multimedia security. Springer, Aveiro, pp 63–72
7. CISA Cyber Infrastructure (2019) MAR-10135536-8 – North Korean Trojan: HOPLIGHT|CISA. [online]. Available at https://www.us-cert.gov/ncas/analysis-reports/AR19-100A. Accessed 23 Sept 2019

8. Council on Foreign Relations (2019) Connect the dots on state-sponsored cyber incidents – PLA unit 61398. [online]. Available at https://www.cfr.org/interactive/cyber-operations/pla-unit-61398. Accessed 15 Sept 2019
9. Ferrer Z, Cebrian Ferrer M (2016) In-depth analysis of Hydraq – in-depth_analysis _of_hydraq_final_231538.pdf. [online] Paper.seebug.org. Available at https://paper.seebug.org/papers/APT/APT_CyberCriminal_Campagin/2010/in-depth_analysis_of_hydraq_final_231538.pdf. Accessed 6 Sept 2019
10. Fireeye Mandiant APT1 Report (2016) APT1: exposing one of China's cyber espionage units – mandiant-apt1-report. [online]. Available at https://www.fireeye.com/content/dam/fireeye-www/services/pdfs/mandiant-apt1-report.pdf. Accessed 18 Sept 2019
11. Flashpoint (2019) Flashpoint – disclosure of Chilean Redbanc intrusion leads to Lazarus Ties. [online]. Available at https://www.flashpoint-intel.com/blog/disclosure-chilean-redbanc-intrusion-lazarus-ties/. Accessed 6 Sept 2019
12. Gressin S (2017) The equifax data breach: what to do. [online] Consumer Information. Available at https://www.consumer.ftc.gov/blog/2017/09/equifax-data-breach-what-do. Accessed 7 Aug 2018
13. Hale B (n.d.) Estimating log generation for security information event and log management. [online] Content.solarwinds.com. Available at http://content.solarwinds.com/creative/pdf/Whitepapers/estimating_log_generation_white_paper.pdf. Accessed 9 June 2018
14. Hussain M, Wahab A, Idris Y, Ho A, Jung K (2018) Image steganography in spatial domain: a survey. Signal Process Image Commun 65:46–66
15. Jasek R, Kolarik M, Vymola T (2013) APT detection system using honeypots. [online] Pdfs.semanticscholar.org. Available at https://pdfs.semanticscholar.org/2f8e/f5890c39579bc9648158b710a1ef2b8366db.pdf. Accessed 12 July 2018
16. Jiang D, Omote K (2015) An approach to detect remote access Trojan in the early stage of communication. In: 2015 IEEE 29th international conference on advanced information networking and applications
17. Joint Task Force Transformation Initiative (2011) Managing information security risk. [online] Nvlpubs.nist.gov. Available at https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-39.pdf. Accessed 4 July 2018
18. Keragala D (2016) Detecting malware and sandbox evasion techniques. [online] Sans.org. Available at https://www.sans.org/reading-room/whitepapers/forensics/detecting-malware-sandbox-evasion-techniques-36667. Accessed 10 June 2018
19. Kruegel C (2015) Evasive malware exposed and deconstructed|USA 2015|RSA conference. [online] Rsaconference.com. Available at https://www.rsaconference.com/events/us15/agenda/sessions/2022/evasive-malware-exposed-and-deconstructed. Accessed 6 June 2018
20. LLC L (2018) Threat actors and exploits top ten lists of 2018|LIFARS, your cyber resiliency partner. [online] LIFARS, your cyber resiliency partner. Available at https://lifars.com/2018/11/threat-actors-exploits-top-ten-2018/. Accessed 19 Sept 2019
21. MacDonald N (2012) Information security is becoming a big data analytics problem. [online] Gartner.com. Available at https://www.gartner.com/id=1960615. Accessed 9 June 2018
22. Marchetti M, Pierazzi F, Colajanni M, Guido A (2016) Analysis of high volumes of network traffic for advanced persistent threat detection. Comput Netw 109:127–141
23. McCandless D (2018) World's biggest data breaches & hacks – information is beautiful. [online] information is beautiful. Available at http://www.informationisbeautiful.net/visualizations/worlds-biggest-data-breaches-hacks/. Accessed 10 Aug 2018
24. Messmer E (2013) Malware-detecting 'sandboxing' technology no silver bullet. [online] network world. Available at https://www.networkworld.com/article/2164758/network-security/malware-detecting%2D%2Dsandboxing%2D%2Dtechnology-no-silver-bullet.html. Accessed 17 June 2018
25. Mokube I, Adams M (2007) Proceedings of the 45th annual southeast regional conference. ACM, New York, pp 321–326

26. Paganini P (2019) Experts link attack on Chilean interbank network Redbanc NK Lazarus APT. [online] Security Affairs. Available at https://securityaffairs.co/wordpress/79929/breaking-news/chilean-research-redbank-lazarus.html. Accessed 6 Sept 2019
27. Paredes-Oliva I, Castell-Uroz I, Barlet-Ros P, Dimitropoulos X, Sole-Pareta J (2012) Practical anomaly detection based on classifying frequent traffic patterns. In: 2012 Proceedings IEEE INFOCOM workshops
28. Raman D, De Sutter B, Coppens B, Volckaert S, De Bosschere K, Danhieux P, Van Buggenhout E (2013) DNS tunneling for network penetration. In: Lecture notes in computer science. Springer, Cham, pp 65–77
29. Rashid P, Ramdhany D, Edwards M, Kibirige S, Babar D, Hutchison P, Chitchyan D (2014) Detecting and preventing data exfiltration. [online] seculanc_data_exfil_report. Available at https://www.lancaster.ac.uk/media/lancaster-university/content-assets/images/security-lancaster/seculanc_data_exfil_report.pdf. Accessed 10 June 2018
30. Scaife N, Carter H, Traynor P, Butler K (2016) CryptoLock (and Drop It): stopping ransomware attacks on user data. In: 2016 IEEE 36th international conference on distributed computing systems (ICDCS)
31. Siddiqui S, Khan M, Ferens K, Kinsner W (2016) Detecting advanced persistent threats using fractal dimension based machine learning classification. In: Proceedings of the 2016 ACM on international workshop on security and privacy analytics – IWSPA'16
32. Sokol P, Míšek J, Husák M (2017) Honeypots and honeynets: issues of privacy. EURASIP J Inf Secur 2017(1):1–9
33. Spitzner L (2002) Honeypots: tracking hackers. Addison-Wesley, Boston
34. Virvilis N, Gritzalis D (2013) The big four – what we did wrong in advanced persistent threat detection? In: 2013 international conference on availability, reliability and security
35. Zamani M, Movahedi M (2015) Machine learning techniques for intrusion detection. [online] Arxiv.org. Available at https://arxiv.org/pdf/1312.2177.pdf. Accessed 21 Dec 2017