Multi-stage attack: Concepts, detection and defence

Delia Daniela Neagu
BSc (Hons) Computer Science for Cyber Security
University of Bradford
Bradford, West Yorkshire
ddneagu@bradford.ac.uk

***Abstract*— We live in a fast-paced world, surrounded by all kinds of technologies. However, unfortunately these technologies present weaknesses that exposes us to malicious attacks. The attackers are developing their skills and continuously find new ways of penetrating into systems where they have unauthorized access. These attacks can cause data breaches and losses and can even cause physical damage to the systems. By getting access to sensitive data, the attacker may steal the bank account details of an individual, their identity and even invade their privacy, which may cause a serious psychological impact. These attacks can cost companies huge financial losses, damage their reputation, and it might even place them in a situation where they have to face legal consequences for not being able to protect their customer's data. This report will discuss multi-stage attacks, which is a very dangerous type of attack, because it's not easily detectable. This attack, as it can be referred from the name, happens in multiple steps, where each step executed by its own does not seem malicious. However, all these phases executed together will cause great harm. This report will also present several proposed detection methods, as well as defence recommendations against this type of attack.**

Keywords- multi-stage attacks, cyber-attacks, stage, target, victim, attacker, network, system, machine, detection, method, malicious, prevention

I. INTRODUCTION

The evolution of innovative technologies unfortunately comes together with the development of cyber-attacks. The attackers are constantly finding new ways of infiltrating in unauthorized systems and develop more complex, intelligent and stealthier methods. Such types of attacks are the multi-stage attacks: they are complex in their phases, intelligent and provide a way to avoid detection over a prolonged period. Usually, the main target of multi-stage attacks are big companies, where the attacker gains access by first accessing an end-user system inside the company. To achieve that, they use different methods such as phishing, water holes or by infiltrating an infected device. By getting access to the end-user system, the attacker will further get access to the entire system.

II. CONCEPTS

A multi-stage attack consists of a series of steps through which the attacker can infiltrate through various networks and systems. Unlike a single stage attack, it is quite challenging to detect a multi-stage attack since each stage of the attack seems harmless as it's not typically malicious when executed alone.[5] However, when all these steps are executed together the overall attack becomes extremely dangerous and can cause hazards to both industries and individuals. Therefore, since the security systems fail to detect these types of attacks, the attacker will aim to maintain persistence in the compromised systems over an extended period, in a stealthy approach.

These attacks are possible because of several existing weaknesses in defense mechanisms. The anti-virus software is only capable to detect malware which has been registered in the database of signatures. Intrusion detection systems often give false negative or false positive alarms. The employees do not receive enough training and lack awareness when it comes to the security measures that they must implement constantly to keep the systems safe. Besides, often users are not changing the default settings that come with the software applications they are using which makes them vulnerable to attacks.

The multi-stage attack generally follows several steps. In the first stage of the attack, also known as "**Reconnaissance**", the actor will gather intelligence, conduct profound research about its target, and collect all information available from a vast range of resources available including both the public and private ones. Then different scanning procedures take place such as using fping to identify any active machines, nmap to detect the open ports, and OpenVAS to search for any vulnerabilities.[23] Based on the information gathered, the attacker will develop an attack plan using threat modelling. The purpose of the stage is to find the system's vulnerabilities.[1]

The second stage of the attack is "**Infection"**. The easiest way for the attacker to gain access to an organization' system is by first gaining access to internal user's account, usually an employee who is already working there. This can be done through different methods such as spear phishing or water holing attacks. In the Spear phishing case, the attacker will develop malicious code which will be used to exploit the weaknesses that were found. This malicious code will be contained in files such as pdfs, docs or pptx. Once the target falls victim to this type of attack, the malware will be installed through the accessed files. A water hole attack is an attack used to compromise websites. In the reconnaissance stage the attacker will identify which websites are mostly visited by the organizations. Then those websites will be injected with malicious code which will be downloaded in the background when the users interact with them.[1] The term "Watering hole" refers to the strategy predators use to hunt other animals by waiting nearby water holes for their prey. Similarly, the attackers wait for the victims to access the infected websites they usually visit.[22] Another method is to gain access through the internal system by infiltrating an infected device such as an USB or infected laptop.

The next stage after infection is "**Exploitation".** In this stage the actor may use a penetration testing framework such as Metasploit and execute a Brute Force attack to gain access to the victim's credentials.[21] Besides, the malicious code installed will exploit all the vulnerabilities of the system. After all these steps are completed, the attacker will connect to a "Command and control" server, to be able to control and

continue the attack.[8] This will give the attacker the opportunity to send commands remotely, control the functionalities of the system, and execute unauthorized operations.[1]

The following phase is "**Data Exfiltration".** Data exfiltration represents the process of transmitting data from the compromised machine to the attacker in a stealthy manner. The purpose of this stage is to steal sensitive information such as credentials, or bank account details.[1]

After all these stages are executed, the attacker will pursue to **maintain persistence**. By maintaining access to the end-user system, the actor will continue the attack and try gain further access to the other systems.[1]

III. Detection

Unfortunately, intrusion detection systems mostly fail to detect these types of attacks, since each step executed alone does not seem malicious.[3] Because of that reason, researchers have come up with several solutions to improve the network detection system and predict the multi-stage attacks.

One popular detection model is the **Hidden Markov Model (HMM)**. A key feature of the Markov Models is that unlike other detection models there's no need to gather extensive information before performing the prediction or detection methods. The Hidden Markov Model performs based on the DARPA 2000 dataset for the training model which analyses the information related to the HMM algorithms.[11]

However, when there is an extensive series of alerts, the Hidden Markov Model is facing difficulties when it comes to tackling the stage dependencies. To solve that problem, there has been proposed **a Sequence-to-Sequence** model for detecting a multi-stage attack. This approach involves: "encoding a series of alerts into a latent feature vector using a long-short term memory (LSTM) network and then decode this vector to a sequence of predicted attacking stages with another LSTM." This model provides a full spectrum of all the alerts detected on a long-term period.[9]

Another approach for early detection of multistage attacks proposes the use of **Machine Learning** in combination with MITRE Adversary Tactic Technique and Common knowledge (ATT&CK). This model has achieved 98% accuracy. This method involves 2 steps. Firstly, the run-time engine will detect any malicious content downloaded from the browser or from the launch of a new process. Then, the MITRE ATT&CK framework is used to detect and predict malicious activity.[13]

Another solution was proposed by Vinayakumar et al. which is based on machine learning and aims to detect the botnet within Internet of Things Network Systems. The detection is made by comparing normal and botnet behavior. First the model will examine the similarity score and then it will classify the normal and abnormal behaviors. [10]

To detect the multi-stage attacks in the Internet of Things industry, Xinghua Li et al. have proposed a machine learning model based on a "bidirectional long and short-term memory network with multi-feature layer (B-MLSTM)". This model is quite efficient since it has significant lower percentages compared to others existing models when it comes to false positives and false negatives alarms.[12]

Advanced Persistent Threat (APT) is a newer and stronger version of multi-stage attack. To perform this type of attack the attacker must maintain a persistent connection the Command-and-Control server. The security of these servers is maintained by Secure Sockets Layer (SSL) which is providing encryption for the communication traffic. This certificate raises great challenges in the detection process of malicious content. To solve this issue, I. Ghafir et al. proposed a detection solution which is based on a matching process of the SSL certificates found in the secure connection with blacklisted SSL certificates.[4]

Attack graphs are tools which are used to thoroughly analyse the network traffic and provide detection for the multi-stage attacks.[18] One example of attack graph detection method solution is discussed in [19], where the researchers rebuild the attack scenario based on the abnormal alerts identified in the IDS. A similar approach based as well on alert sets is also described in [15]. Another solution is proposed in [20], which in this case is using On-line Correlation Engine, where with the help of the Virtualization Engine, the organization will be able to monitor the presence and progression of an attack.

Power grids' networks have proven great capabilities in detecting abnormalities, which is why Omer Sen et al. have investigated an approach to detect multi-stage attacks in Smart Grids described in [16]. Also, H. Zhang et al. have presented a detection model based on alert graphs [17]. Ying-Mei Wang et al. have as well suggested a solution for identifying attack paths in networks.[15] Then, further research was taken to find methods to improve the 5 G network detection system [23].

Another method proposed to detect Multistage Advanced persistent Threats is a Game-Theoretic Approach for Dynamic Information Flow Traffic proposed by Shana Moothedath et al. This approach is based on information-flow graph, where the nodes are represented by objects and systems and the edges depict the interactions between them. The target of this method is to reduce the progression of the attack to a minimum level and increase its detection.[7]

IV. Defence

If this type of attack occurs, it is essential to come up with a strong response plan. The organization should have a clear understanding of what's the most critical data. They should use the Layered Security Architecture and store the essential data in the inner layers.[1]

To ensure the detection, the attacker should be forced to interact with the system multiple times. If the system is implemented in such a way that it will make it more difficult to gain access, the attacker will have to invest more time and

resources, which will cause more often interaction with the security system, that will increase the chances of the sensors to identify the attack. The most essential and valuable information will be contained in the inner layer.[8]

Accessing a layer should be possible by accessing the outermost layer first. For each layer the attacker will have to execute the intrusion kill chain. The weaknesses of each layer must be reduced to the most minimum level possible. Vulnerabilities found in each layer will give the attacker the opportunity to learn how to bypass the security of the next layer. These steps will ensure an effective layered model.[8] Each layer should be implemented with effective detection sensors which will perform based on rules defined by the patterns of the malicious behavior. The sensors will collect a series of alerts and logs which will be stored and correlated to identify the phases of the attack.[8]

This may be quite challenging to implement since the sensors will generate large amounts of data. To resolve this issue, there has been created a model using the Hadoop framework. Hadoop's Logging module contains different types of intrusion detection systems, firewalls, and logs. The setting for the configurations can be controlled by the administrator. The logs generated in the Logging Module will then be sent to the logging management module, where they can be accessed using Hive queries. The intelligence module contains detection algorithms. The malware analysis will present details about the malware effects. Finally, the Control module provide the administration with control over the logs.[8]

In addition to that, end user-systems must be protected by being regularly updated and patched. The plug-ins and adds-on must also be constantly updated. Also, the organization must ensure that the surfing browsers are using secure configuration settings. They should use anti-virus software which should also be updated regularly.[1]

The organizations should also ensure the network security system by using: IDS and IPS, Firewalls, email filtering, robust domain name systems, implementation of Honeypots, monitor the traffic closely, encrypt sensitive data, analyze server logs regularly, use virtual local area networks, and use strong security configurations for devices. [1]

Intrusion detection systems are crucial in ensuring the systems safety. However, as previously mentioned, it is difficult for them to detect multi-stage attacks since each stage executed alone does not seem malicious. Therefore, it is a necessity to continuously monitor the systems and verify whether there's any suspicious activity. Besides, it is essential to regularly update the systems, the software, and provide patches. They should also run regular security tests.[1]

Education and awareness are also vital especially when it comes to industries. Employees should receive trainings and understand the security dangers such as fallen victim to a phishing email, the importance of keeping their credentials secret and safe, and the dangers of using any unknown external devices.[1]

V. Real Life Examples

Havex in a Trojan virus that was first discovered in 2014. Its target are the industrial control systems (ICS), and its purpose is to collect information related to the network architectures or sensitive details. It is collecting it from a large number of websites on a world-wide scale where the attacks against the ICS have been successful. [6] This attack mainly affected energy, utility, and pharmaceutical industries and it was spread world-wide.

One common method to implement the Havex Trojan is through spear phishing email attack. The attacker creates the malicious code which is going to be integrated into the email attachment. When the victim opens the attachment, the malware will get installed. That will start the scanning procedures which will allow the attacker to find the machines vulnerabilities and then exploit them and collet information, which will be sent to the C2 server.[6]

The second method is to infect the website provided by the ICS. The scope of this attack is to transmit the malware through the HTPP protocol. If the infected website is visited by an administrator, the virus will get installed which will give the attacker high privileges and access to the users' sensitive data.[6]

Another intrusion method which is also the most common one, is by implementing malicious code on the provider's website. This type of intrusion is similar to the first 2 discussed previously, however the main difference consists in the information delivery. By using this method, the attacker aims to gain access to intranet of a company, which occurs when the administrators of the website are transferring files from the internet to the internal production system. [6]

The **Stuxnet** attack started in 2006 and it was not discovered until 4 years later, in 2010 when it caused physical damage to the centrifuges of the nuclear facility in Iran. The attacker has spent all those years acting from the shadow, collecting information, and working on creating a method to cause physical damage to the centrifuges.[6] Since the nuclear facilities were isolated at the time, and the connection to external network was less likely to be possible, it is believed that the attack occurred through the infiltration of an infected device, which was connected internally, allowing the virus to spread until it reached the attacker's C2 server. This is a real-life example of a 2-stage attack.[6]

Duqu is a malware similar to Stuxnet, which was first discovered in 2011. It is believed that they both were developed by the same people. The only difference between them however is that Duqu's purpose is to solely steal information, without causing any physical damage. [2]

Flame has been further investigated in 2012. It was first known by the name "sKyWIper", and it affected many areas. Its first interaction was in 2007 in Europe, then it affected the United Arab Emirates in the spring of 2008, and then it affected Iran in 2010. This malware acted as a proxy for the Microsoft Windows update, and its purpose was to steal the victim's information.[2]

VI. Conclusion

Multi-stage attacks are highly dangerous since they can cause data breaches and damage to the compromised systems. Each stage executed by itself does not appear malicious, however all the phases executed together can cause serious hazards to both individuals and companies. Therefore, the incapability of the intrusion detection systems to detect these types of attacks is raising a significant concern. Besides, the lack of detection it's giving the actor the opportunity to persist the attack over a prolonged period of time.

Detecting the attack in its early stages can prevent huge data loss and critical damage. Researchers have figure out several different detection methods which implement the use of machine learning, attack trees and others.

Defense methods are essential. The organization must have a strong recovery plan in case this happens. The systems must be protected and regularly updated and patched. The plug-ins and ads on should be updated as well. They should be using anti-virus software, intrusion detection systems, Firewalls, email filtering methods, robust domain name systems, and implementation of Honeypots. Besides that, they should always monitor the traffic closely and do regular check-ups, encrypt sensitive data, analyze server logs regularly, use virtual local area networks, and use strong security configurations for devices. It is also very important that the organizations educate their employees and offer them proper trainings, in the scope of making them aware of the potential dangers and the methods to prevent them.

References

[1] Aditya Sood, R. E. Targeted cyber attacks: Multi-staged attacks driven by exploits and malware. Elsevier Science, 2014

[2] Bencsáth, B. e. Duqu: Analysis, detection, and lessons learned. ACM European Workshop on System Security, EuroSec. Vol. 2012.

[3] Aparicio-Navarro, F.J., Kyriakopoulos, K.G., Ghafir, I., Lambotharan, S. and Chambers, J.A. Multi-Stage Attack Detection Using Contextual Information. IEEE Xplore1–9., 2018

[4] Ibrahim Ghafir, V. P. Malicious SSL Certificate Detection: A Step Towards AdvancedPersistent Threat Defence. *Proceedings of the International Conference on Future Networks and Distributed Systems.*, 2017

[5] M. Lefoane, I. G.-U. Multi-stage Attack Detection: Emerging Challenges for Wireless Networks," 2022 International Conference on Smart Applications. *Communications and Networking (SmartNets)*, (pp. pp. 01-05). Palapye, Botswana, 2022

[6] M. Li, Y. H. Research on Attack Mechanism of Network Intrusion in Industrial Control System. *2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)* (pp. 1904-1908). Chengdu, China: doi: 10.1109/IAEAC47372.2019.8997670, 2019

[7] Moothedath, S., Sahabandu, D., Allen, J., Clark, A., Bushnell, L., Lee, W., & P, R. . A Game-Theoretic Approach for Dynamic Information Flow Tracking to Detect Multistage Advanced Persistent Threats. *IEEE Transactions on Automatic Control*, 5248-5263. , 2020

[8] P. Bhatt, E. T. "Towards a Framework to Detect Multi-stage Advanced Persistent Threats Attacks,". *2014 IEEE 8th International Symposium on Service Oriented System Engineering, Oxford, UK*, pp. 390-395, 2014

[9] Peng Zhou, G. Z. Detecting multi-stage attacks using sequence-to-sequence model. *Computers & Security*, 2021

[10] R. Vinayakumar, M. A.-V. A Visualized Botnet Detection System Based Deep Learning for the Internet of Things Networks of Smart Cities . *IEEE Transactions on Industry Applications, vol. 56*, 4436-4456, July-Aug. 2020

[11] Timothy Chadza a b, K. G. Analysis of hidden Markov model learning algorithms for the detection and prediction of multi-stage network attacks. Future Generation Computer Systems, 636-649, 2020

[12] X. Li, M. X. Detection of Low-Frequency and Multi-Stage Attacks in Industrial Internet of Things. IEEE Transactions on Vehicular Technology, vol. 69, 8820-8831, 2020

[13] Y. S. Takey, S. G. Real Time early Multi Stage Attack Detection. 7th International Conference on Advanced Computing and Communication Systems (ICACCS) (pp. 283-290). Coimbatore, India: doi: 10.1109/ICACCS51430.2021.9441956, 2021

[14] Wang, Ying-Mei, et al. "An analysis approach for multi-stage network attacks." *2005 International Conference on Machine Learning and Cybernetics*. Vol. 7. IEEE, 2005.

[15] Steffen Haas and Mathias Fischer. GAC: graph-based alert correlation for the detection of distributed multi-step attacks. In Proceedings of the 33rd Annual ACM Symposium on Applied Computing (SAC '18). Association for Computing Machinery, New York, NY, USA, 979–988, 2018.

[16] Sen, Ömer, et al. "Towards an approach to contextual detection of multi-stage cyber attacks in smart grids." 2021 International Conference on Smart Energy Systems and Technologies (SEST). IEEE, 2021.

[17] H. Zhang et al., "A Multi-Step Attack Detection Model Based on Alerts of Smart Grid Monitoring System," in IEEE Access, vol. 8, pp. 1031-1047, 2020, doi: 10.1109/ACCESS.2019.2961517, 2019

[18] Bopche, G.S., Mehtre, B.M. Attack Graph Generation, Visualization and Analysis: Issues and Challenges. In: Mauri, J.L., Thampi, S.M., Rawat, D.B., Jin, D. (eds) Security in Computing and Communications. SSCC 2014. Communications in Computer and Information Science, vol 467. Springer, Berlin, Heidelberg, 2014

[19] Hu, Hao, et al. "Attack scenario reconstruction approach using attack graph and alert data mining." Journal of Information Security and Applications 54 (2020): 102522, 2020

[20] Angelini, Marco, et al. "An attack graph-based on-line multi-step attack detector." Proceedings of the 19th International Conference on Distributed Computing and Networking. 2018.

[21] Lefoane, Moemedi, et al. "Unsupervised learning for feature selection: A proposed solution for botnet detection in 5g networks." IEEE Transactions on Industrial Informatics 19.1 (2022): 921-929.

[22] Joey Allen, Z. Y. Mnemosyne: An Effective and Efficient Postmortem Watering Hole Attack Investigation System. In Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security (CCS '20). Association for Computing Machinery, (pp. 787–802). New York, NY, USA, 2020