An Overview of Advanced Persistent Threat Detection Based on Machine Learning

Kai Xing, Aiping Li* and Rong Jiang
College of Computer, National University of Defense Technology Changsha, 410073, China

*Corresponding author. Email: xingkai18@nudt.edu.cn

Abstract. Cyberspace has been constantly threatened by attacks since its birth. With the development of high-tech and artificial intelligence, intelligent and efficient attack methods have emerged endlessly, and technological methods have been constantly renovated. In particular, Advanced Persistent Threat (APT) attacks are intensifying. How to effectively prevent this attack method has become the focus. With the advantages of machine learning, the thinking and technology of detection have made great progress. This article mainly discusses several innovative methods for detecting APT attacks based on machine learning, and looks forward to the future development direction.

Keywords: advanced persistent threat (APT), machine learning, cyber-kill-chain.

1. Introduction
Cyberspace has been hailed as the fifth battlefield [1], and governments of all countries have increasingly attached importance to the rights and interests of cyberspace. With the development of high-tech and artificial intelligence, people are always threatened by network attacks while enjoying the great convenience brought by the network. According to statistics from the National Internet Emergency Center (CNERT), only in January 2010, China received a total of 7,580 cybersecurity incidents reported at home and abroad [2]. There are endless forms of cyberattacks, and technical means are constantly being renovated, especially for artificial intelligence, the rapid development has given birth to intelligent and efficient means of network attacks. The "Stuxnet" virus attack on Iran’s nuclear facilities in 2010 is a masterpiece of APT. Therefore, APT has gradually received great attention from governments, international agencies and security companies. The academic community has also made many useful explorations on APT. Lots of detection methods based on machine learning have been proposed. Based on this, this article summarizes the current machine learning-based APT detection technology, and the follow-up structure is arranged as follows: Section 2 briefly introduces the background knowledge of APT, Section 3 states the detection technology based on machine learning; Section 4 summarizes and discuss the development direction of APT detection technology based on machine learning; the last is the future outlook.
2. Background knowledge

The term Advanced Persistent Threat (APT) originated from the warning of abandoning the Trojan horse to disclose sensitive information [3]. Although the name "APT" was not used at that time, "Advanced Persistent Threat" has been used. Until 2006, it was proposed by the U.S. Air Force Greg Rattray in a corrective manner [3]. As Iranian nuclear facilities were attacked by the "Stuxnet" virus in 2010 and Google was severely hacked, APT attacks are well known in the information security field [3]. With the rapid popularization of network technology, attacks such as this one continue to emerge, such as "Operation Aurora", "Black Energy" and so on, always threatening the security of cyberspace.

2.1. Basic characteristics of APT attacks

APT attacks become more frequent, three specific characteristics [3] of APT attacks can be clearly summarized through specific cases shown in Table 1:

(I) Advanced attack methods. Attackers often have all-round support and use advanced attack methods such as social engineering, zero-day vulnerabilities, covert communication, and machine learning to achieve their goals, but this does not mean that an attack is a single attack, often a combination of multiple methods use.

(II) Long duration. The attacker seeks long-term access control to the target, which is highly concealed and will not be easily exposed until the goal is reached, which can take several years.

(III) The degree of threat is high. APT attacks mainly target high-value government departments, infrastructure and large enterprises. Once these targets are overcome, the losses are often in the hundreds of millions.

| Organization   | Target                  | Tools                        |
|----------------|-------------------------|------------------------------|
| Hades          | Korea                   | Olympic Destroyer            |
| APT28 (Suspect)| Ukraine                 | VPNFilter                    |
| APT28          | North America, Europe   | Cannon, Zebrocy              |
| BlueMushroom    | China                   | PowerShell backdoor          |
| OceanLotus     | Southeast Asian, China  | Denis, Cobalt Strike         |
| BITTER         | China, Pakistan         | Unique backdoor procedures   |
| APT38          | Global, SWIFT           | Multiple homemade malicious programs |
| DarkHotel      | China                   | Plug-in Trojan Backdoor      |
| APT33 (Suspect)| Middle East, Europe     | Shamoon V3                   |

2.2. Cyber kill chain

The concept of "kill chain" was not originally used in cyberspace, but originated in the military field. It is a six-stage model describing the attack process, that is "discover, locate, track, target, strike, achieve" stages [3]. The earlier the link in the kill chain is used to stop the attack, the better the protection effect and the less harmful it is. Lockheed Martin proposed a similar attack kill chain (Cyber-Kill-Chain), which essentially divides the attack into 7 stages, namely "detection tracking, weapon construction, load delivery, vulnerability exploitation, installation implantation, command control, goal achievement" [4], as shown in Figure 1. This theory can be used not only for attacks but also for protection.
3. Detection technology based on machine learning

3.1. Detection of APT using machine learning classification based on fractal dimensions
Sana Siddiqui et al. [5] focus on classifying abnormal traffic patterns based on APT by processing the feature vectors obtained from TCP/IP session information to achieve the goals of high accuracy and reliability and then propose a new algorithm based on the correlation fractal dimension. The experimental results, compared with k-NN, show that the algorithm proposed has improved 8% in "accuracy", "sensitivity" and "specified degree", but the "accuracy" and "F-measure" have generally improved by more than 12%, indicating that the fractal method reduced false positives and false positives provide better performance. Similarly, the performance of the proposed algorithm is better than k-NN when the data set is highly imbalanced.

3.2. Detection of APT based on GAN-LSTM
The traditional APT detection methods are mainly aimed at comparing a single attack mode. It has obvious advantages in processing attacks with short time spans, but some attacks have long time spans. Traditional attacks cannot well reflect the timeliness of APT attacks. With fewer attack data samples and longer attack durations, the accuracy rate becomes very low. In response to this problem, Liu et al. [6] creatively proposes an APT attack detection method combining Generative Adversarial Networks (GAN) and Long Short-term Memory (LSTM). The detection method is determined by the attack data generation module and attack data. The module and timing processing module are composed, as shown in Figure 2.
Generate attack data based on GAN simulation, generate a large number of attack samples for the discrimination model, solve the problem of small samples, and improve the accuracy of the model; the memory unit and gate structure based on the LSTM model ensure the correlation and time in the APT attack sequence, the feature memory between large-sequence sequence fragments solves the problem of large time span. Generate simulated attack data through the generative model and then optimize the discriminative model, which improves the accuracy of the original discriminant model by 2.84 percentage points. Compared with the APT attack sequence detection method based on Recurrent Neural Network (RNN), the detection accuracy rate has been improved by 0.99%. The experimental results fully demonstrate that the APT attack detection algorithm based on GAN-LSTM can increase the sample capacity by introducing a generative model, thereby improving the accuracy rate of the discriminant model and reducing the false positive rate; meanwhile, compared with other time-series structures, using LSTM model to detect APT attack sequences has better accuracy and lower false positive rate.

3.3. Detection of APT use machine learning correlation analysis
Ibrahim Ghafir et al. [7] propose a new MLAPT system based on machine learning, as shown in Figure 3. The system can systematically accurately and quickly detect and predict APT attacks. MLAPT goes through three main stages: (1) threat detection, in which eight methods are developed to detect different technologies used in various APT steps; (2) early warning correlation, designing a correlation framework, linking the output of detection methods, and identifying Early warning that may be relevant and belong to a single APT scenario; (3) Attack prediction, based on the output of the correlation framework, proposes a prediction module based on machine learning, which is used by the network security team to determine the probability of early warning, thereby finding a complete APT attack. The experimental evaluation of MLAPT shows that the system can make early prediction of APT with a prediction accuracy of 84.8%.
3.4. Detection of APT based on a new Deep Learning stack

Tero Bodström et al. [8] propose a novel deep learning stack based on a theoretical approach for detecting APT attacks, as shown in Figure 4. In this approach, APT is considered a multi-media multi-stage attack with a continuous strategic campaign.

![Figure 3. The framework of the system.](image)

![Figure 4. Data flow through the detection process.](image)

To capture these attacks, the entire network flow, especially raw data, must be used as input to the detection process. By combining different types of custom DL methods, you can catch certain types of
exceptions and behaviors. In essence, the larger problem is decomposed into smaller tasks, trying to solve these problems in turn, and finally returning conclusive results.

3.5. Detection of APT attacks based on semi-supervised learning and complex network feature

Advanced Persistent Threat (APT) provides the most sophisticated types of attacks to modern networks. Attackers use sophisticated attack techniques to remotely control infected computers and steal sensitive information from organizations and governments. Due to the dynamic nature of the APT attack process, security products deployed on traditional defense-based enterprise networks often cannot detect APT infections. In order to overcome the limitations of the current attack network dynamics faced by APT research, Aaron Zimba et al.[9] propose an innovative APT attack detection model based on semi-supervised learning methods and complex network features, as shown in Figure 5. The entire target network is modeled as a small world network, while the evolving APT-Attack Network (APT-AN) is modeled as a scaleless network. The finite state machine is used to model the state transitions of the nodes in the time domain to characterize the state changes during the APT attack. The method proposed in the experiment effectively analyzes a large data set to reveal the characteristics of APT attacks between the command and control center and the victim host. The average detection accuracy of the APT test proposed by the author is 90.5%. The results show that the model can effectively detect suspicious hosts at different stages in the APT attack process.

![Figure 5. The framework of the system.](image)

4. Conclusions

This article mainly sorts out the representative APT detection technologies and ideas based on machine learning in recent years. Although not many, it can also cause deep thinking. The methods and ideas mentioned in this article are basically in-depth exploration of a certain problem using machine learning methods, and have proved their advantages through experiments, such as indicators, detection accuracy and false alarm rate. The shortcomings of each method are also summarized, as shown in Table 2.

| Methods/Ideas                              | Shortage                                      |
|-------------------------------------------|-----------------------------------------------|
| Classification based on fractal dimensions| Hard to deal with malicious data classification|
| GAN-LSTM                                  | Unable to deal with long sequence             |
| Machine Learning Correlation Analysis     | Lack of publicly available datasets           |
| A new Deep Learning stack                | Uncomputed time complexity                     |
| Semi-supervised Learning and complex network feature | Computational overhead                           |
5. Future outlook

The development of machine learning has brought great opportunities to cyberspace. The accompanying machine learning-based APT attacks have brought unprecedented challenges to cyberspace security. How to efficiently and accurately handle APT attacks has become a top priority of cybersecurity defense weight. Machine learning has its obvious advantages. How to solve the problems of computational overhead, small samples, and time span has a long way to go.

References

[1] Chen Rui-dong, Zhang Xiao-song, Niu Wei-na, and Lan Hao-yue, A Research on Architecture of APT Attack Detection and Countering Technology, Journal of University of Electronic Science and Technology of China. 48 (2019) 870-879.

[2] Information on https://www.cert.org.cn/publish/main/upload/File/2020-1.pdf.

[3] Information on https://www.freebuf.com/column/160412.html

[4] Information on https://www.sohu.com/a/204031269_490113

[5] Sana Siddiqui, Muhammad Salman Khan, Ken Ferens, Witold Kinsner, Detecting Advanced Persistent Threats using Fractal Dimension based Machine Learning Classification, Proceedings of the 2016 ACM on International Workshop on Security And Privacy Analytics, 2016, 64–69.

[6] Liu Hai bo, Wu Tian bo, Shen Jing and Shi Chang ting, Advanced Persistent Threat Detection Based on Generative Adversarial Networks and Long Short-term Memory, Computer Science.

[7] Ghafir I, Hammoudeh M, Prenosil V, Detection of advanced persistent threat using machine-learning correlation analysis, Future Generation Computer Systems. 89: 349-359.

[8] Tero Bodström, Timo Hämäläinen, A Novel Deep Learning Stack for APT Detection, Appl. Sci. 2019, 9(6), 1055.

[9] Aaron Zimba, Hongsong Chen, Zhaoshun Wang, Mumbi Chishimb, Modeling and detection of the multi-stages of Advanced Persistent Threats attacks based on semi-supervised learning and complex networks characteristics, Future Generation Computer Systems. 106 (2020) 501-517.