A Study on Security Framework Against Advanced Persistent Threat

Qingyun Zhang,Huan Li and Jinsong Hu
Department of Network
Electrical Engineering Institution
Hefei, Anhui, China
1070553997 @qq.com

Abstract—**Advanced Persistent Threat (APT) and traditional cyber attacks are different in kinds of aspects, which make the traditional defense is difficult to detect APT and protect the network. Therefore, an APT detection framework based on OpenIOC is established for the characteristics of APT system attack. Firstly, real-time attack data related to APT from massive fragmented threat data is output. Secondly, the data is transformed into real-time OpenIOC threat information by IOC which is used to measure the similarity with the history APT organization's OpenIOC threat information. Finally, the relationship between real-time attack characteristics and APT organization is analyzed, and history APT organization's attack behavior will be discovered.**

Advanced Persistent Threat; OpenIOC; threat sequence;TTPs; similarity;IOC

I. INTRODUCTION

The term APT Attack was first proposed by USAF (United States Air Forces) in 2006.According to "SP800-53 management information security risk" published by the US National Institute of Standards and Technology (NIST) in 2013, in the definition of APT: proficient technology attackers use a variety of attack programs (network, physical and fraud) with rich resources to achieve the purpose of attack. In recent years, the security situation of cyberspace has become more and more serious, and advanced threat has become the most threatening cyber attack. Almost all important industries, such as government, finance, electric power and education, are threatened by APT attack[1-3].

Therefore, it is necessary to study the security defense means to deal with the APT attack effectively. APT attacks are different from traditional network attacks. They have the characteristics of pertinence, continuity, advancement, phase, sharing, indirection, so the traditional attack detection technology is facing great challenges of APT attacks.

At present, many scholars have proposed different defense detection framework for APT attacks, which laid the foundation for the development of APT attack detection technology [4-5].

Paper [6] propose a comprehensive defending framework and a detecting framework based on intelligent feedback. Paper [7] based on the in-depth analysis of the denotation and connotation of threat, this paper explores defense models to threat in details and proposes a theoretic security and defense framework to deal with APT: abnormal discovery, so as to solve the problem of threats detection.

However, due to the advanced nature and sharing of APT attacks, there is similarity between the APT attacks issued by the same APT organization, and threat sharing is an important support for the discovery of APT attacks. Therefore, the effective use of the deep information about the APT organizations reflected by the known APT attacks can provide a basis for detecting the attacks committed by the same APT organization. However, the detection framework of the above methods do not involve such a function

Therefore, an APT detection framework based on OpenIOC is established. Firstly, real-time attack data related to APT from massive fragmented threat data is output. Secondly, the data is transformed into real-time OpenIOC threat information by IOC which is used to measure the similarity with the history APT organization's OpenIOC threat information. Finally, the relationship between real-time attack characteristics and APT organization is analyzed, and history APT organization's attack behavior will be discovered.

II. SECURITY FRAMEWORK AGAINST ADVANCED PERSISTENT THREAT

A. An security framework aginst APT based on OpenIOC

The overall testing process is shown in Figure 1. The framework testing process shown below is divided into three stages: TTPs eigenvalue acquisition, TTPs feature sequence modeling and TTPs feature sequence similarity calculation.

The relevant concepts involved are as follows.

Concept 1(Real-time data) Real-time data is a collection of communication data and host data acquired in real-time monitors. The communication data intercepted from the gateway including packet data and message data. The host data obtained from the internal network including the system running log, the system running status record, the application software log and so on.

Concept 2(History-threat-data) History-threat-data is a collection of threat data temporarily stored in history that has not been determined as an APT attack.

978-1-5090-3025-5/17/$31.00 ©2017 IEEE

Concept 3(TTPs) TTPs (technique, tool and producer) are the key information in the threat information produced by IOC including kinds of network attack method. Such as method of recon (port scan, probing…), method of deliver (phishing…), method of C&C (DGA, R2L....), method of execute (Dos, R2L, Worms ...).

Concept 4(OpenIOC) OpenIOC stands for Open Indicators of Compromise. OpenIOC is an extensible XML schema for the description of technical characteristics that identify a known threat, an attacker's methodology, or other evidence of compromise. OpenIOC was created by MANDIANT, a security company headquartered in Alexandria, Virginia.

Concept 5(IOC) IOC stands for Indicators of Compromise. The term may be used to refer to specific artifacts left by an intrusion, or greater sets of information that allow for the detection of intrusions or other activities conducted by attackers. The term is also used as the name for a file in the OpenIOC format that contains a set of data. The file extension for such files is .ioc.

Concept 6(IKC-APT) The IKC-APT model is an APT model based on the kill chain (IKC).

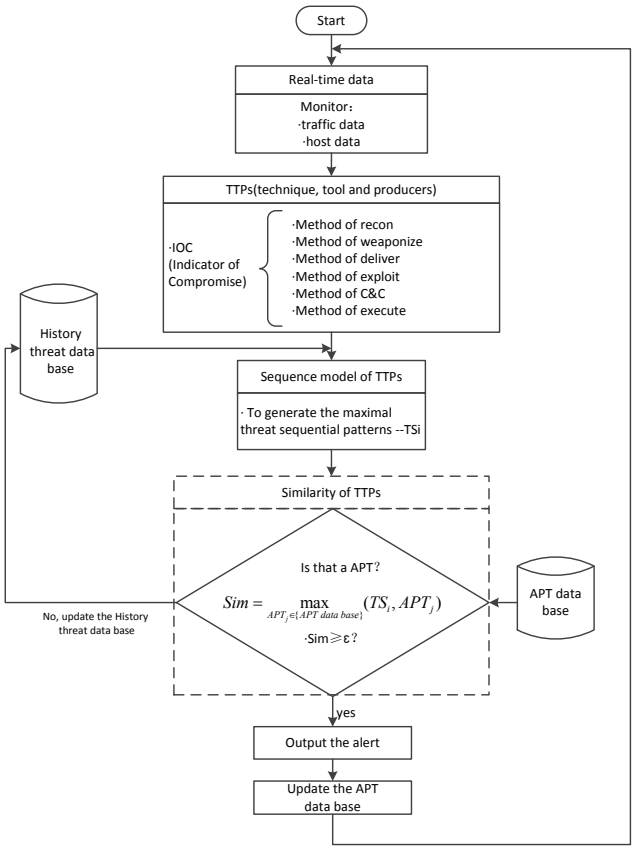

Figure 1 The security framework testing process

The framework detection process is as follow.

Step1 Input data. Input the real-time data.

Step2 Get TTPS. Extracting TTPs from IOC.

Step3 Threat sequence modeling. The TTPS is serialized based on the kill chain (IKC) to generate a threat sequence.

Step4 Similarity analysis. The similarity between the threat sequence generated in real time and the threat sequence of the known APT is analyzed.

Step5 Warning and update.

B. TTPs acquisition

This stage mainly uses various network data collection devices to obtain real time data. And then using IOC to dig out TTPs, for further analysis of the higher level to provide the basis for analysis.

TTPs will be a comprehensive description of the security events, and sub-means, technology, process three dimensions of analysis, including the malicious attack behavior, tools, victim goals, weaknesses, effects and consequences, etc.

C. TTPs sequence modeling

This stage transforms the massive TTPs obtained from the previous stage into the threat sequence based on the IKC-APT model.

Intrusion kill chain (IKC) [8] was originally used in the combat process to refer to "discovery-location-tracking-aiming-attack-evaluation". In March 2011, three security researchers at Lock Martin presented the IKC (intrusion kill chain) model for the first time at the ICIW conference.

From the point of view of intrusion detection, the model decomposes the attacker's attack process into 7 steps: information collection - assembly - release - use - implantation - command - control - continuous attack.

APT attackers generally have a clear purpose of attack. In order to achieve the purpose, attackers need to implement different attack steps. Although these attacks steps belong to different types of attacks, there is a causal relationship between each other, and the order of occurrence is also logical. Therefore, APT attacks have a feature of phase. Using IKC to describe it can quickly help us understand the latest attack scenarios and how to effectively carry out and defense.

APT attack model based on IKC is produced, which divides APT process into six stages as shown in Figure 2.

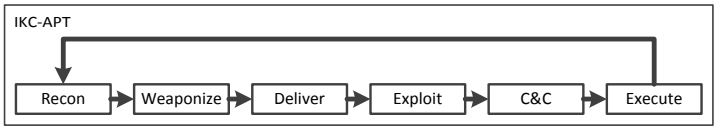

Figure 2 IKC-APT model

The threat sequence pattern is an implementation of the IKC-APT model.

(1) Recon

Through the analysis of the published APT report, it is found that the attacker collects the information of the target user and the target host in the target scanning phase to collect the social engineering information of the target user and the

environmental information of the target host in order to design the attack scheme.

The collected information includes scanning host information, network information, application information, disk information, process information, user behavior characteristics, user information (account name, password) and so on.

2) Weaponize

The attacker uses the public RAT (Remote Access Trojan) to add functionality, or independently develop RAT, based on the target scan results.

(3) Deliver

Malicious code is delivered to the target host via payload delivery methods such as harpoon mail, phishing sites, mobile media, and so on. It can be divided into two ways: direct introduction and indirect import.

Among them, indirect introduction can be done by setting up phishing sites, targeting target users, or by using devices such as U disks containing malicious programs to make the target host stallion or use social engineering methods for password guessing and so on.

The specific way of load delivery is determined by the specific characteristics of the target host or target user information collected in the first stage. There are also the use of social networking and instant messaging tools to send malicious code and other new delivery methods.

(4) Exploit

Attackers often make use of known or unknown system vulnerabilities and application vulnerabilities to realize the operation of codes and the promotion of their rights, so as to achieve covert dissemination and control of the system within the network.

(5) C&C

Through the last stage of the installation of the remote access to the Trojan or back door to maintain a sustained access to the target system, the attacker has been in the target user's computer installed in the corresponding malicious programs or start the remote control,

This time already has a target host File access. Through the communication with the RAT, update the malicious code itself or the configuration file; accept the relevant control instructions; and return to steal the data information.

(6) Execute

After establishing a secure communication channel with the target host, you can steal the secret information by scanning the file system, or use the virus, Trojan horse to destroy the target system, to achieve the purpose of target paralysis.

The TTPs is a 7-tuple, including the ID, time of occurrence, the IKC phase, the source IP address, the destination IP address, the source port number, the destination port number, and the protocol type.

$$T\_TTPs = \{ID, Time, IKC\_ID, SrcIP, DestIP, SrcPort, DestPort, Prot\}$$

The threat sequence (TS) can be expressed as $< a_1, a_2, a_3, ..., a_n >$. The TS is sorted by time which is used to describe the APT attacker's behavior cycle. The TS's construction method is shown in Figure 3.

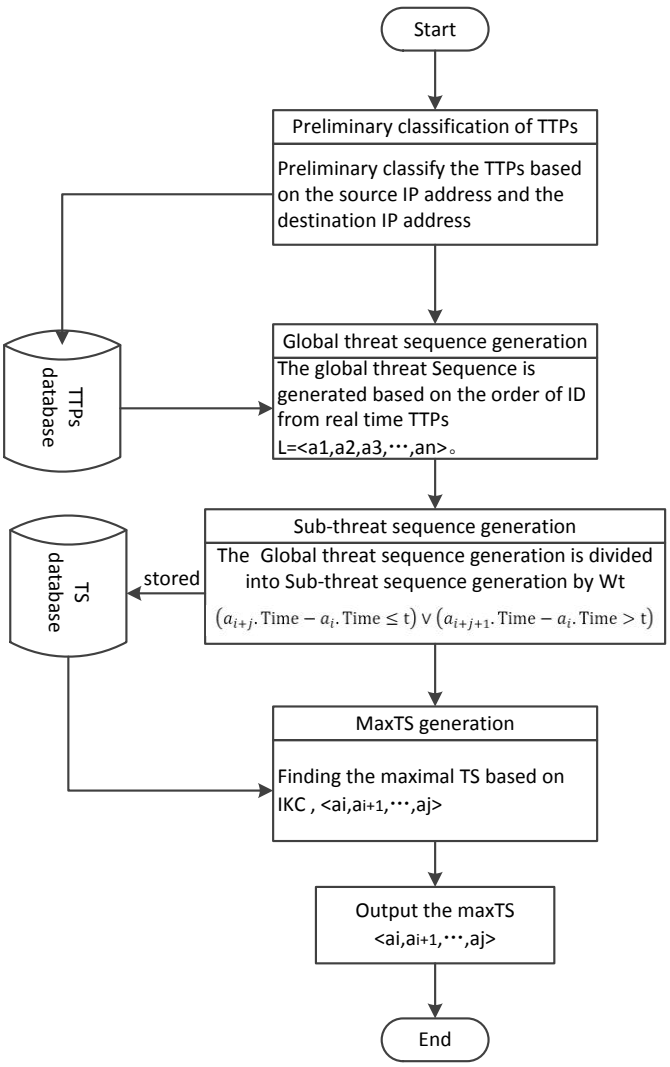

Figure 3 The TS's construction method process

Based on the above method, the transformation of the TTPs to the threat sequence can be realized.

D. Threat sequence similarity calculation

This phase analyzes the similarity between the real-time TS generated by the previous stage and the known APT TS in the APT database, and determines whether the attack behavior of the known APT has occurred.

According to paper [9] and based on the APT characteristics, a similarity calculation method of APT based on TTPs similarity is proposed. The similarity of the method is composed of two parts: one is the attribute similarity. By

calculating the distance D_{A-B} between the real time TS and the known APT's TS, the smaller the distance value, the higher the similarity between the real time TS and the known APT's TS.

$$D_{A-B} = \overline{\omega_{ioc}} D_{HM}(A,B)$$

And another part is the similarity degree of traffic -- S. S is calculated by calculating the similarity determined by the C&C communication information.

$$S = \sigma S_C + (1-\sigma) S_{CF}$$

Finally, the similarity of two parts is linearly fitted, and the total similarity Sim_{A-B} is calculated.

$$Sim_{A-B} = \alpha \frac{1}{\overline{\omega_{ioc}} D_{HM}(A,B)} + \beta(\sigma S_C + (1-\sigma) S_{CF})$$

III. COMPARATIVE ANALYSIS

Compared with the existing security framework, the security framework based on OpenIOC against APT of the main advantages are as follows.

(1) Aiming at the APT attack chain, this paper proposes an APT modeling method based on IKC, which can describe the causal relationship between APT attack scene and mining APT attacker behavior.

(2) Because the traditional defense means can't fully share the threat of intelligence, threat information exchange method based on OpenIOC is proposed.

(3) Because the traditional defensive means always ignore the characteristics of APT organization behavior, the similarity calculation method based on TTPs is proposed to realize the discovery and early warning of known APT organization attack behavior.

(5) On the discovery of the attack can be timely warning, and feedback updates, and then guide the next round of testing process.

However, since the framework is based on OpenIOC threat information from known APT principals, the framework can not detect the attack behavior of unknown APT principals. The next step should be to think about how to combine threat information and discover unknown APT attacks.

IV. CONCLUSION

The use of threat intelligence to identify APT attacks is the development trend of APT attack defense means. Aiming at the shortcomings of the traditional defense methods in the analysis of threat intelligence and APT attack behavior. IOC, attack sequence pattern and similarity calculation are merged to form the framework of APT detection system based on OpenIOC. The framework extracts the APT-related threat data through the IOC, integrates with the IKC-APT model to form the attack sequence, and finally finds the known APT organization's attack behavior by similarity calculation.

REFERENCES

[1] Mandiant. M-Trends-2016[M]. 2016.

[2] Xiaomei L. Research on Prevention Solution of Advanced Persistent Threat[Z]. Singapore,Singapore: 20144.

[3] Cole E. Advanced Persistent Threat: Understanding the Danger and How to Protect Your Organization[M]. Syngress Publishing, 2012.

[4] Yutong W, Chaowen C, Zengbang M. An Advanced Persistent Threats Awareness Technology Based on "Condensed matter" [Z]. Xi'an,China: 20156.

[5] Bin D, Wentao Z, Jianglong S. Modeling Analysis of Advanced Persistent Threat-Based on UML[Z].Guilin,Guangxi: 20155.

[6] Vukalović J, Delija D. Advanced Persistent Threats - detection and defense[C]. 2015.

[7] Du Yuejin, Z Lidong, L Yue, et al. Security Architecture to Deal with APT Attacks: Abnormal Discovery[J]. Journal of Computer Research and Development. 2014, 51(7): 1633-1645.(In Chinese)

[8] Hutchins E, Cloppert M, Amin R. Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains[C]. 2011.

[9] H Rong, S Huo, C Hu, et al. User similarity-based collaborative filtering recommendation algorithm[J]. Journal on Communications.2014(02): 16-24.(In Chinese)