Poster: APT Detection through Sensitive File Access Monitoring

Wenjia Song
Virginia Tech
wenjia7@vt.edu

Danfeng (Daphne) Yao
Virginia Tech
danfeng@vt.edu

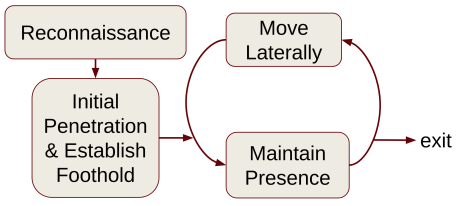
***Abstract*—Advanced persistent threats (APTs) have become more destructive in recent years, impacting a wide range of organizations, from government agencies to critical infrastructures, and the daily life of millions of people. There have been many security works focusing on this topic and various solutions have been proposed. However, challenges still exist for the accurate detection of new APT variants. Innovative and strong defenses need to be invented. To achieve this, we plan to systematically analyze recent APT attacks and the current practice of system monitoring of sensitive information, which could lead us to new and secure solutions.**

I. Introduction

Ransomware and zero-day supply-chain trojan, a new form of attack, have posted unprecedented threats to various organizations. $350 worth of cryptocurrency ransom was collected by hackers in 2020[1]. $4.4 million ransom was paid by Colonial Pipeline Co. in May 2021 and $11 million was paid to the hackers in June 2021 by the meat supplier JBS USA. Zero-day supply-chain trojan leverages software updates from established vendors. The SolarWinds and Kaseya hacks are examples of this new form, threatening the data safety of hundreds of companies. Both of these attacks could be variants of advanced persistent threats (APTs).

As the name implies, advanced persistent threats (APTs) are often achieved by groups of sophisticated hackers. They try to stay present in the system as long as they can and threaten the safety of sensitive data and critical components of targeted organizations [1]. Well-known and documented examples of APTs include APT-28[2] and APT-38[3]. APTs usually consist of multiple key stages [1], [2]. The life cycle of APTs is shown in Figure 1.

Fig. 1. Life cycle of APTs.



[1]https://blog.chainalysis.com/reports/ransomware-ecosystem-crypto-crime-2021/

[2]https://www.fireeye.com/content/dam/fireeye-www/global/en/current-threats/pdfs/rpt-apt28.pdf

[3]https://content.fireeye.com/apt/rpt-apt38

TABLE I
Summary of existing APT detection works categorized by the attack stage they focus on and the approach they use

	Rule-based Approach	Machine learning-based Approach
Detection of initial penetration	Mohammad et al. [3] Chandra et al. [4]	APTGuard [5] Kumar and Somani [6]
Detection of malicious movement	HOLMES [2] Poirot [7]	MLAPT [8] UNICORN [9]

- **Reconnaissance.** This is the beginning of the attack. The more the hackers know about the victim, the more likely they will succeed.
- **Initial penetration and establish foothold.** This stage represents the successful entry of the attack. The hackers may also set as many backdoors as possible to stay inside the system.
- **Maintain presence.** Once inside the victim system, the malware wants to remain present and undetected for as long as possible.
- **Move laterally.** To collect critical information, the malware laterally moves within the victim system.
- **Complete mission.** When the mission is completed, the hackers may clear their trace and exit the target system. Otherwise, they would stay in and repeat the early stages.

There have been many research studies on the detection of APTs. However, challenges still exist. Innovative and strong security defenses are needed to ensure the safety of sensitive data.

II. Related Work

Various APT detection tools have been proposed in recent years. They can be roughly categorized by the attack stage they focus on (e.g., initial penetration or latent movement) and the approach they use for detection (e.g., rule-based or machine learning-based), as shown in Table 1.

To prevent the malware from entering the victim machine, Mohammad et al. [3] propose a feature-based classification of phishing websites. Chandra et al. [4] use a mathematical model to filter spam emails. APTGuard [5] is a tool to detect spear phishing URLs using decision tree and neural network. Additionally, Kumar and Somani [6] train various machine

learning models to detect malware before installation using static, dynamic, and origin information of executables. On the other hand, researchers have also done works on the detection of malicious movements of APTs after the malware get inside the victim systems. HOLMES [2] finds the behavior pattern connections of key stages of APTs. POIROT [7] analyzes the semantic of causality of system alters for anomalous events. For machine learning-based approaches, MLAPT [8] is a multi-phased machine learning detection framework using network traffic. Another tool, named UNICORN [9], clusters provenance graph sketches and dynamically detects anomaly system changes.

III. Challenges

Despite the research advances, there are still several challenges for detecting sophisticated APTs.

- **Hard to catch new variants.** APTs constantly change and update to disguise themselves from detection. The supply-chain trojan is an example of their new form of entry. Behavior patterns summarized from historical data may be outdated and unable to identify newer versions.
- **Hard to generalize across different hosts and operating systems.** Many anomaly detection models are trained on normal system behaviors. However, normal behavior may vary from system to system. To generalize, a large amount of normal data needs to be collected for each system.
- **False alters and missed detections.** Due to the large amount of system and network data and the lack of clear distinction between malicious and benign behavior, many machine learning-based approaches tend to produce false positives and false negatives.

IV. Ongoing Work

Although APTs change their form, structure, and behavior to avoid detection, their goal remains the same. Therefore, new approaches focusing on monitoring and protecting sensitive information, which is the target of attacks, are needed. In order to develop such a security defense, we need to understand the recent APT attacks and how the current systems monitor sensitive information. The major tasks of our next steps include:

- Analyze the current practice of how critical and sensitive information is monitored by the system. Different auditing policies generate logs with different types of events. A key step is to find the optimal configuration of security auditing.
- Analyze the recent APT attacks from various sources. The resources include executable malware samples (e.g., real-world APT malware samples summarized in [6]), APT execution logs (e.g., APT-EXE[4]), datasets containing APT traces (e.g., DARPA dataset, KDD-99[5] and its derivative NSL-KDD[6], DAPT 2020 [7]), and descriptions and documentation of various attacks (e.g., APTNotes[8], MITRE ATT&CK[9]).
- Fine-grain security logs and focus on sensitive information access-related events. Systems generate an overwhelming amount of event logs that contain noise. Focusing on a smaller set of relevant events can potentially help reduce false positives. As shown in Figure 2, events 5379 (i.e., Credential Manager credentials were read) and 4656 (i.e., a handle to an object was requested) are examples of sensitive information-related events that should be handled carefully.

Fig. 2. Example of Windows security event log.

Keywords	Event ID	Task Category
Audit Success	5379	User Account Management
Audit Success	4672	Special Logon
Audit Success	4624	Logon
Audit Success	4658	Other Object Access Events
Audit Success	4656	SAM

Acknowledgment

This work has been supported by the Office of Naval Research under Grant No. N00014-22-1-2057.

References

[1] A. Alshamrani, S. Myneni, A. Chowdhary and D. Huang, A Survey on Advanced Persistent Threats: Techniques, Solutions, Challenges, and Research Opportunities, in IEEE Communications Surveys & Tutorials, vol. 21, no. 2, pp. 1851-1877, Secondquarter 2019.

[2] S. M. Milajerdi, R. Gjomemo, and B. Eshete, et al., HOLMES: Real-time apt detection through correlation of suspicious information flows, in 2019 IEEE Symposium on Security and Privacy (SP), 2019, pp. 1137–1152.

[3] R.M. Mohammad, F. Thabtah, and L. McCluskey, Intelligent rule-based phishing websites classification. IET Information Security, 2014. 8(3): p. 153-160.

[4] J. V. Chandra, N. Challa, S. K. Pasupuleti, A practical approach to e-mail 900 spam filters to protect data from advanced persistent threat, in: Circuit, Power and Computing Technologies (ICCPCT), 2016 International Conference on, IEEE, 2016, pp. 1–5.

[5] B.L.J. Chuan, M.M. Singh, A.R.M. Shariff, APTGuard : Advanced Persistent Threat (APT) Detections and Predictions using Android Smartphone. In: Alfred R., Lim Y., Ibrahim A., Anthony P. (eds) Computational Science and Technology. Lecture Notes in Electrical Engineering, vol 481. Springer, Singapore, 2019.

[6] T. Kumar, G. Somani, Origin Information Assisted Hybrid Analysis to Detect APT Malware. In: Tripathy S., Shyamasundar R.K., Ranjan R. (eds) Information Systems Security. ICISS 2021. Lecture Notes in Computer Science, vol 13146. Springer, Cham.

[7] J. Yang, Q. Zhang, X. Jiang, S. Chen and F. Yang, Poirot: Causal Correlation Aided Semantic Analysis for Advanced Persistent Threat Detection,"in IEEE Transactions on Dependable and Secure Computing, 2021.

[8] I. Ghafir, M. Hammoudeh, V. Prenosil, L. Han, R. Hegarty, K. Rabie, and F. J. Aparicio-Navarro, Detection of advanced persistent threat using machine-learning correlation analysis, Future Generation Computer Systems, vol. 89, pp. 349–359, 2018.

[9] X. Han, T. Pasquier, and A. Bates, et al., UNICORN: Runtime provenance-based detector for advanced persistent threats, in Network and Distributed Systems Security (NDSS) Symposium 2020, Jan 2020.

[4]https://github.com/aptresearch/datasets

[5]http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html.

[6]http://www.unb.ca/cic/datasets/nsl.html.

[7]https://gitlab.thothlab.org/Advanced-Persistent-Threat/apt-2020/-/tree/master

[8]https://github.com/kbandla/APTnotes

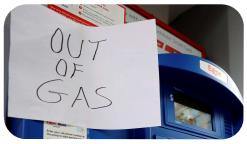
[9]https://attack.mitre.org/

Poster: APT Detection through Sensitive File Access Monitoring

Wenjia Song, Danfeng (Daphne) Yao
Department of Computer Science, Virginia Tech, VA, USA
{wenjia7, danfeng} @vt.edu

1. Motivation

The scope and impact of APT attacks, including ransomware and supply-chain trojans, are unprecedented. They affected a wide range of organizations and daily life of millions of people.

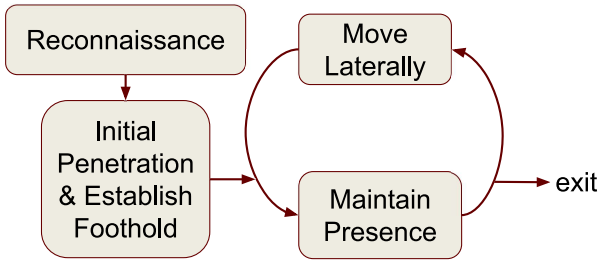


$350M ransom was paid in 2020.
$4.4M and **$11M** were paid by Colonial Pipeline Co. and JBS USA, respectively, in ransom attacks in 2021 .

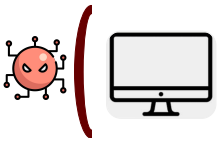
100-280 companies install a trojanized version of the SolarWinds Orion. **800-1500** companies were compromise in the Kaseya hack in July 2021.

2. APT

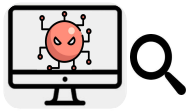
APTs stands for **Advanced Persistent Threats.** They are usually **highly sophisticated** attacks aiming for **long-term** access.



3.Existing Solutions



Focus on detecting **initial penetration**



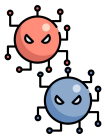
Focus on detecting **malicious movement**

Rule-based approaches

Machine learning-based approaches

Rule-based approaches	Machine learning-based approaches
Mohammad et al.	APTGuard
Chandra et al.	Kumar and Somani
HOLMES	MLAPT
Poirot	UNICORE

4. Challenges



Hard to detect variants.
APT always change its form to disguise itself, e.g., the new supply-chain trojan.

Hard to generalize.
Different hosts and systems may different normal behaviors.

False alerts and missed detections.
Due to the large amount of system and network data and lack of clear distinction between malicious and benign behavior.

5. Ongoing Work

The APT malware may change its **form, structure, and behavior** to disguise itself, but its **goal** remains the same.

Task 1: Analyze system monitoring on sensitive information (target of the attacks).

Task 2: Analyze recent attacks from different sources in various formats.

APT-EXE | **DAPT 2020** | **MITRE ATT&CK** | **DARPA** | **KDD-99 & NSL-KDD** | **APTNotes**

Task 3: Fine grain the system logs and focus on sensitive file and critical component access to reduce false positives.

Keywords	Event ID	Task Category	
Audit Success	5379	User Account Management	A read operation is performed on stored credentials.
Audit Success	5379	User Account Management	
Audit Success	4672	Special Logon	
Audit Success	4624	Logon	
Audit Success	4658	Other Object Access Events	
Audit Success	4658	Other Object Access Events	
Audit Success	4656	SAM	An access is requested to an object (security account manager).
Audit Success	4656	SAM	

This work has been supported by the Office of Naval Research under Grant No. N00014-22-1-2057.

VIRGINIA TECH

Yao Group on Cyber Security
https://yaogroup.cs.vt.edu/

NDSS SYMPOSIUM