ICIC Express Letters
Volume **14**, Number **8**, August **2020**

ICIC International ©2020 ISSN 1881-803X
pp. 821–827

A STUDY ON EFFECTIVE OUTER DOCUMENT FLOW BASED ON DAST FOR APT RESPONSE

Jong Pil Kim[1], Onechul Na[2], Harang Yu[2], Giwan Hong[2]
and Hangbae Chang[3]

[1]Softcamp Co., Ltd.
17, Pangyo-ro 228beon-gil, Bundang-gu, Seongnam-si, Gyeonggi-do 13487, Korea
jpkim@softcamp.co.kr

[2]Department of Security Convergence
Graduate School
[3]Department of Industrial Security
Chung-Ang University
84 Heukseok-ro, Dongjak-gu, Seoul 06974, Korea
{ nastop; hryu356; ghd9201; hbchang }@cau.ac.kr

Received January 2020; accepted April 2020

Abstract. *The Advanced Persistent Threat (APT) is a method of social engineering which refers to a meticulous attack by utilizing a precise attack target and sophisticated program and is also rapidly increasing in continuous progress and development of commercial detection technology. Accordingly, in this research, Document Attachment Sanitization Technology (DAST) based harmless technology through contents reconstitution is established and implemented by analyzing the types of malicious code, attack method, etc. and furthermore, precedent researches which analyze these were investigated. S-ince the existing technology can only detect simple signature comparison or detection of known malicious code, the technology proposed in this study is able to actively and quickly respond to new malicious code such as concealment, bypass, and latency. DAST is a technology with new perspective of interpretation and a differentiated access method, which is applicable on documentation which inflows from exterior to interior from various types of channels and uses technology which reconstructs documentation after extracting safe contents and therefore able to preemptively and actively respond to APT attack. This study suggested a solution that can improve limitation of existing detection technology; however, it has a limitation of detection by damage on the original document during the process of documentation reconstitution and non-supporting extension, and therefore requires to conduct an additional study in the future time.*

Keywords: Advanced Persistent Threat (APT), Document Attachment Sanitization Technology (DAST), Malicious code, Detection technology, Spreading course of malicious code

1. **Introduction.** As the penetration rate of the Internet and the rapid development of the Internet infrastructure are threatened by the user's computer environment, these threats become more intelligent and diversified over time. High-performance computers are used as a specific zombie computer for botnets or are used as a propagation medium for infected computers.

The recent occurrence of serious security incidents that affect the overall life of the country, such as the recent cyber attacks of 6.20 and cyber terrorism of 3.25, has heightened awareness of the importance of security. The government has imposed various security technologies such as firewall, Intrusion Prevention Systems (IPS), vaccine as well as network separation to major institutions such as public institutions and financial institutions by strengthening the information protection standards and guidelines [1].

DOI: 10.24507/icicel.14.08.821

However, even these efforts are causing serious security incidents such as the outflow of public corporations and the leaking of personal information in financial institutions, and they are damaging by bypassing existing security technologies and bypassing countermeasures. This is attributed to Advanced Persistent Threat (APT) attacks.

The purpose of the APT is to suspend the core information and telecommunication facilities of a specific company or organization and to acquire key information. The attacker can use the information infrastructure such as Information Technology (IT) infrastructure, work environment, employee information to collect various information. On the basis of this, zero-day attacks and social engineering techniques are used to collect and exploit vulnerabilities possessed by attack targets and execute attacks [2]. The characteristics of APT attacks are that it is difficult to detect and respond to attacks because they use very sophisticated programs that define clear attack targets and collect various information over a long period of time to exploit vulnerabilities [3,4].

In order to cope with such advanced and intelligent APT attacks, a solution has been developed that incorporates a variety of security technologies such as network traffic analysis, behavior based and reputation analysis using a virtual environment, and is now being released as a defense solution for APT attack. However, it is difficult to anticipate and preemptively respond to changing APT attacks because they are all based on expert analysis and are not known or constantly advanced like zero-day attacks [5,6].

In this paper, we propose a document based on Document Attachment Sanitization Technology (DAST) for external inflow documents that is not malicious by judging signatures and behavior-based analysis against constantly changing external attacks and blocking inflows. We want to study the base technology that harmlessly reconfigures and protects efficiently.

2. Recent Study.

2.1. Malicious code and attack technique.

Malware is an abbreviation of malicious software, which refers to malicious software and refers to any software that is maliciously designed and can adversely affect the computer. Such software may be installed or executed without the users' order or approval and may perform malicious actions such as degradation of system performance or leakage of personal information. Domestically, malware is collectively referred to as "malicious code". Computer viruses, worms, Trojan horses, spyware, and rootkits all belong to malicious codes [7].

The attack technology of malicious code has changed from the early self-motivation to the aggressive type with the purpose and the target in recent years and has been developed with various techniques for concealment, detour, and camouflage. Recent attacks have been increasing in number of attack types based on social engineering attack techniques in the same way as zero-day vulnerability [7].

Malware attacks using social engineering attack techniques mainly take advantage of e-mail or vulnerabilities, and these types of malicious codes include Conficker, Waledac, IRCBot, etc. Conficker is a malicious code that mainly propagates through Windows security vulnerabilities, removable storage media such as Universal Serial Bus (USB), shared folders with weak passwords, and accesses specific websites from an infected platform [8]. Waledac attack technology is a malicious code that spreads via e-mail and increases network traffic by sending large amounts of spam from an infected system [8]. IRCBot is a malicious code that spreads via Windows vulnerabilities, network share folders, removable storage media, and changes the system date. By changing the date of the system, normal service cannot be provided, and a backdoor is installed by connecting to a specific IRC server [8].

The attack techniques of malicious code were found to be attacked by using various technologies from the outside as seen in the above investigation, and it was investigated that they were intelligent not to be distinguished from each other visually through cloaking, camouflage and detour.

2.2. **Spreading course of malicious code and type of attack.** Recently, malicious code has been developed to infect more computers by using various methods of spreading course. According to the study of Korea Internet & Security Agency (KISA), the path and type of spreading course of malicious code is defined as shown in Figure 1, and is typified by storage medium, e-mail, network, download, instant messenger program and security vulnerability [9]. In particular, there are two ways to spread malicious code: passive spreading course and automatic spreading course [7]. Specifically, a storage medium is a medium for storing digital data such as text, audio, and images, and is mainly used to store data or to transfer data from one computer to another. Examples of the storage medium include a Compact Disk (CD), a floppy disk, and a removable storage device. The network is steadily expanding into our lives right now. We have access to the network at least once a day and are using multiple services. However, an unspecified anonymous majority network is becoming one of the main causes of malware spreading. Spoofing, blogging, bulletin board, and Social Network Service (SNS) are examples of breaking down malicious code using a network. Currently, Internet users can download various programs according to their needs in various ways. Especially executable files and double-extension files are especially important. The executable file is usually a file with a file extension of exe. The extensions are MOV, SCR, AVI, PIF, etc., which are video files and picture files but actually executable files. Since 1999, e-mail has been used as a medium to spread malicious code to attackers. Early viruses, such as I-Worm/Happy99, which used e-mail as a propagation method, simply sent malicious code to send an e-mail. Recently,

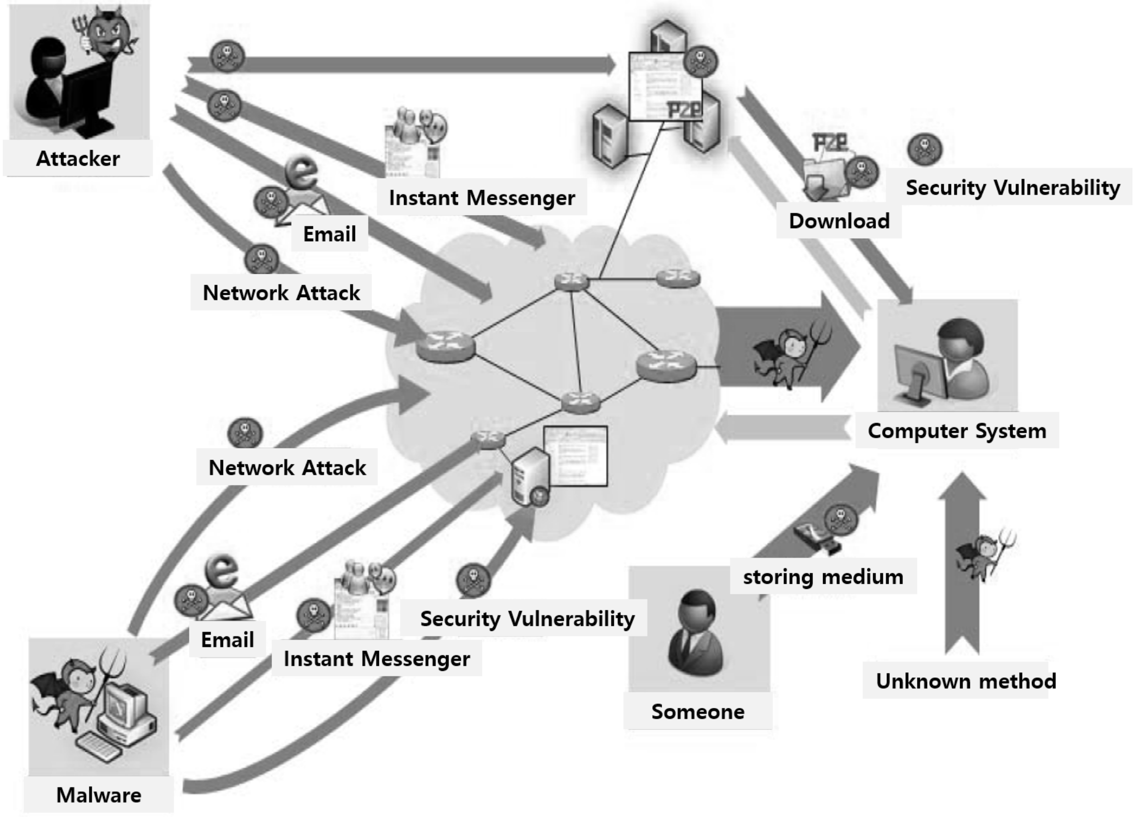

FIGURE 1. Types of spreading course of malicious code

however, this technique has become so popular that people have been particularly careful when using e-mail. As a result, when malicious code became difficult to distribute in the early days, attackers used social engineering techniques to embed irritating and interesting titles or content in e-mails, opening up users' mail or running attached files. The instant messenger program is a program used by two or more users to send and receive text in real time over the Internet. An instant messaging program differs from mail in that it can send and receive text instantly.

2.3. **Detection method of malicious code.** Recent methods for generating malicious code, scripts such as Visual Basic (VB), mIRC script, and Java script, are usable [10]. These scripts are mainly propagated through e-mail. In order to detect and block these scripts in advance, signature scanning, application conversion, and static analysis are used. Malicious code analysis technology is divided into host-based analysis technology and network-based analysis technology, and it is classified into behavior characteristics of analysis technology, signature-based, and behavior-based [8,10,11].

Host-based malicious code analysis technology is a method of analyzing malicious code of a file or system using a malicious code detection program installed in each host. In contrast, network-based malicious code analysis is a method of analyzing malicious code by collecting and analyzing the network traffic that is transmitted from each end of the network to each host. The signature-based malicious code analysis technique is a technique for analyzing whether a specific or unique part of a file is in conformity with a pattern of known malicious code. The behavior-based malicious code detection technique analyzes various behaviors occurring in the system to detect malicious code suspicious files [12].

2.4. **Differentiation from existing research.** The market size for APT attacks is increasing. According to the Gartner report, the APT market is expected to grow from about \$ 200 million in 2012 to \$ 1.17 billion in 2017. In addition, related technologies such as detecting malicious code, malware, and spam e-mail are analyzed and provided as a virtual environment in relation to cyber attack defense, and many of them have already been utilized and patented. However, most of them are malicious code, malware, spam mail and so on.

On the other hand, the underlying technology of this study protects information assets from APT attacks by monitoring and blocking the access of local system resource (important file, registry, communication, etc.) by the file without any analysis of external inflow file and it is a completely different study from the existing detection methods investigated through previous studies. In addition, if the existing analysis was a one-time analysis of a file or application, this study is not a one-time analysis but a response to fundamental APT.

Significant downside of existing signature-based and behavior-based analysis is that it needs a periodical update in order to detect new malicious code or behavior, since it is based on blacklist method and previously analyzed exemplary regarding a influx file. However, DAST fundamentally reorganizes structural form of documentation so that information which includes malicious code can be removed and analyzing technology cannot be detoured. Also, it has a significant differentiation in the fact that it is a technology which does not need a periodical pattern update unlike existing method and can ultimately prevent an influx of documented malicious code.

3. **Establishment of DAST.** The DAST, which is a core technology of this research, aims to systematically design and implement access procedures to physically monitor visitors in relation to documents that are imported inside, and by utilizing DAST, intends to analyze and handle documentation by influx routines, form of influx file.

The DAST concept is designed based on the access procedure of physical security and has the same structure as the procedure for outsiders to go in and it is easy to compare

Table 1. DAST research differentiation

Division	*Signature-based*	*Behavior-based*
Strength	High detection performance for malicious code already known by signature method	– Easy to build and manage – Network traffic analysis/detection (some available)
Disadvantages	– Unprotected against unknown attacks such as zero-day – If the inspection (analysis) passes and it flows into the inside, it cannot be controlled and controlled separately	– Need to collect excessive activity logs for analysis – Defenseless when bypassing virtual environment – Separate management/control of files passed analysis – Difficult to detect malicious code that is brought in through USB other than communication – No solution for proactively infected malicious code in case of internal inflow
Differentiation of this study	– Preemptive and active defensive technologies that complement the weaknesses of the two bases – Threats removal technique of external inflow file, not analysis/detection method of malicious code – File structure improvement method that does not need pattern update – Analysis of behavior by removal of hidden elements (malicious code) by extracting visible contents – Shipment response technology due to pre-removal of necessary elements – Document reconstruction technology by removing document hidden attachment – Maintain the format of the original extension – Discard and block files with malicious code – Rapid response and analysis with minimal resources – Post-analysis (after inflow) defense continuous maintenance structure	

and understand. Visitors need to apply for a visit and visit permission at the meeting room to get inside. After receiving a pass, the visitor gets to go inside and inspects the bags and the carrying items, and carries out security procedures to prevent USB memory, notebook, and other storage devices from being brought in or to obtain prior permission. Even now, various security measures are conducted by prohibiting the use of smartphone camera by attaching a sticker [13].

As shown in Figure 2, the DAST structure also has the identical structure which regards the visitor as an external document and even when an external document is imported, the same procedure as the access security procedure is performed, which is referred to as document sanitization. The role of an entrance gate is connected to an interlocking module Application Programming Interface (API) between systems transmitted from the outside to the inside, and document evasion processes such as a pass granting procedure and a belongings inspection are performed.

Stage 1 is a process of delivering an external document to interior system. If vaccine exists, external document is preferentially transmitted to vaccine server and existing signature-based or behavior-based malicious code detection analysis is conducted. In

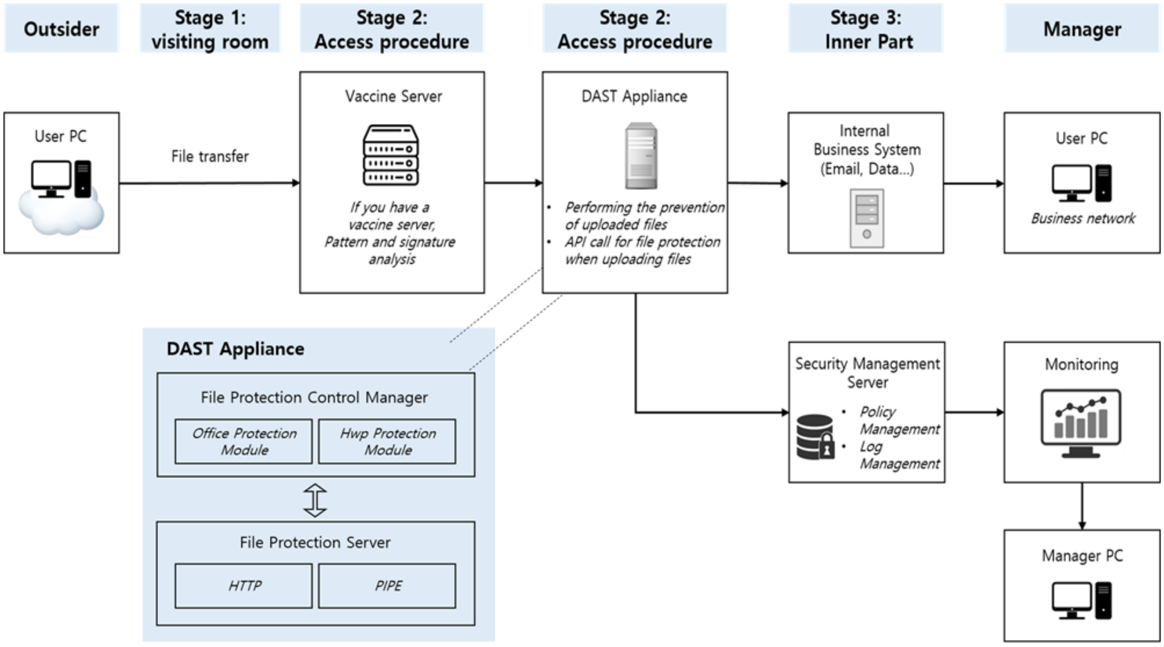
Outsider
Stage 1: visiting room
Stage 2: Access procedure
Stage 2: Access procedure
Stage 3: Inner Part
Manager
User PC
File transfer
Vaccine Server
If you have a vaccine server, Pattern and signature analysis
DAST Appliance
• Performing the prevention of uploaded files
• API call for file protection when uploading files
Internal Business System (Email, Data...)
User PC
Business network
DAST Appliance
File Protection Control Manager
Office Protection Module
Hwp Protection Module
File Protection Server
HTTP
PIPE
Security Management Server
• Policy Management
• Log Management
Monitoring
Manager PC

FIGURE 2. Concept of combination of access security and DAST

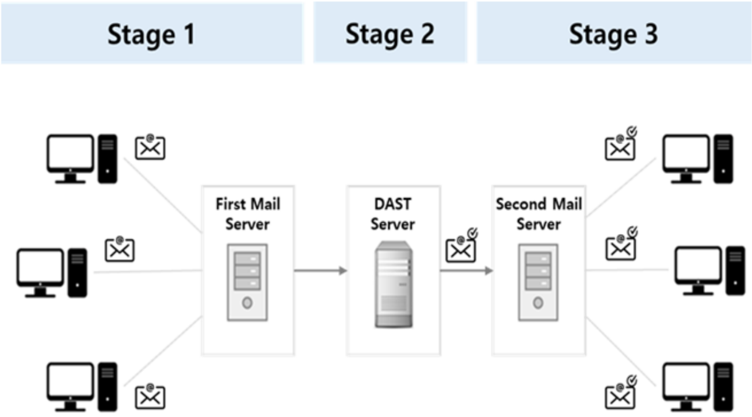

FIGURE 3. Configuration of correlation between DAST server (sanitization server) and e-mail

Stage 2, external document is transmitted to DAST application server and fundamentally reconstruct a structural form of document in order to remove concealment information which can include malicious code of document such as belonging inspection. Reconstructed file is in the status granted to pass. Lastly in Stage 3, reconstructed file is completely delivered to interior system.

In order to facilitate the understanding of this study, concept of combination of access security and DAST in correlation with e-mail is illustrated in Figure 3. It is composed in conjunction with e-mail to sanitize various viruses and malicious codes which is included in e-mail and attached files. It is composed with Stage 1, where DAST server proactively receives electronic document. Stage 2 consists of removing a macro setting within DAST server and in Stage 3, DAST server which macro setting is deleted is sent to interior system.

4. **Conclusions.**

In this study, DAST is a preemptive countermeasure against intelligent cyber attacks. Through this study, we tried to reconstruct the document by extracting only the safe contents when importing the document inflowed from outside and also tried to improve the limitation of the existing researches by removing the hidden code by applying it to the documents of various paths from the outside.

In order to design the DAST basic structure based on the DAST concept, a DAST server and an e-mail are linked to form a three-stage system. Next, we will construct the entire DAST structure for each module and describe the operation method for the detailed structure in detail. In addition, DAST test version that can be applied in real enterprise environment is implemented and carry out an experiment of bringing in the malicious code from outside to inside in order to verify the reliability of DAST technology.

Acknowledgment. This paper was supported by Korea Institute for Advancement of Technology (KIAT) Grant Funded by the Korea Government (MOTIE) (P0008703, The Competency Development Program for Industry Specialist).

REFERENCES

[1] Financial Security Institute, *Guidelines for Separating Financial Computer Network*, 2013.

[2] Y. Li, T. Zhang, X. Li and T. Li, A model of APT attack defense based on cyber threat detection, *China Cyber Security Annual Conference*, Singapore, pp.122-135, 2018.

[3] J. Chen, C. Su, K. H. Yeh, and M. Yung, *Special Issue on Advanced Persistent Threat*, 2018.

[4] M. Nicho and S. N. Khan, A decision matrix model to identify and evaluate APT vulnerabilities at the user plane, *The 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pp.1155-1160, 2018.

[5] Birdman, The evolution of Windows spyware techniques, *HIT2005*, 2005.

[6] A. Sung, J. Zu, P. Chavez and S. Mukkamala, Static analyzer for vicious executavles (SAVE), *The 20th Annual Computer Security Applications Conference*, pp.326-334, 2004.

[7] Korea Internet & Security Agency, *A Study on the Prediction Method for Similar Types and Variants of Malware*, 2008.

[8] C. Im, J. Oh and H. Jeong, Study of technical trends and analysis method of recent malware, *The Korean Institute of Information Scientists and Engineers*, vol.28, no.11, pp.117-126, 2010.

[9] Korea Internet & Security Agency, *Cyber Trend Threat Report*, 2016.

[10] C. Woo and K. Ha, A development of malware detection tool based on signature patterns, *The Korean Society of Computer and Information*, vol.10, no.6, pp.127-135, 2008.

[11] V. Ganapathy, S. A. Seshia, S. Jha, T. W. Reps and R. E. Bryant, Automatic discovery of API-level exploits, *Proc. of the 27th International Conference on Software Engineering*, pp.312-321, 2005.

[12] T.-T. Teoh, Y.-Y. Nguwi, Y. Elovici, W.-L. Ng and S.-Y. Thiang, Analyst intuition inspired neural network based cyber security anomaly detection, *International Journal of Innovative Computing, Information and Control*, vol.14, no.1, pp.379-386, 2018.

[13] Korea Communications Commission, *IT Subcontractor Security Control Guide*, 2011.