Study on Detection and Analysis of Zero-day Malicious Email and Software

ポムケオナ, サヌパーブ

https://doi.org/10.15017/4060190

出版情報：九州大学, 2019, 博士（情報科学）, 課程博士
バージョン：
権利関係：

（別紙様式 2）

氏　　名　：ポムケオナ　サヌパーブ

論 文 名　：Study on Detection and Analysis of Zero-day Malicious Email and Software
（メールとソフトウェアに含まれる未知の脅威の検出と解析に関する研究）

区　　分　：甲

論文内容の要旨

The purpose of our research is to protect users, organization’s network or businesses’ services from cyber threats which are increasing both quantity and diversity in recent years. There are many type of cyber threats such as malicious software (malware), phishing, denial-of-service (DoS), man-in-the-middle, SQL injection and so on. Moreover, in term of cyber-attack that makes those cyber threats more terrify is a zero-day attack. A zero-day attack is a threat that exploits an unknown computer security vulnerability which takes place before or on the first (or zeroth) day of an awareness of the exploit or bug. In addition, zero-day threat can be any kind of malware, phishing or attack which is new and hard to detect by using just cybersecurity systems, devices or anti-virus software. On the other hand, all cybersecurity reports said that the most common target of attack entry-points is by using malicious emails and APT threat attacks. Consequently, a Zero-day malware which is created by cyber deviants is a critical risk and menace because neither machines nor cyber security tools can easily detect them.

In first contribution, we develop a method to handle with suspicious emails which are the most common entry point of a zero-day malware attack in the real network environment. The method is a combination between URL investigation and file investigation together which cover a number of operations such as format file investigation, surface analysis, dynamic analysis, static analysis, etc. and then from the malicious behavior and their activities, we create a call graphs for each of them, make human being easier to compare, match or classify them with other malicious cases. we have completed investigate more than 100 suspicious emails characteristics, most of them are URSNIF (spyware), ransomware, downloader, phishing and spam mails. Compare with general Windows OS with Windows Defender, our method gives a better result of phishing and spam mails detection and increase a detection rate from 49.5% up to 89.9%. There is a huge better accuracy rate detection especially on phishing and spam mail using our method, however the manual job requires time to proceed. We have finished investigation, record details, behavior, category and solution more than 100 malicious emails and it can be used to find or help infected victims inside an organization’s network.

Next, because of everyday users have to face off spam, phishing or malicious emails and it would be a huge problem for whole organizations if only one user clicked on a single link from those. The difficult issue is how to classify and detect those malicious emails from the ordinary, especially spear phishing emails, which are designed for a particular target, or zero-day malicious emails that nobody has found before. In second research contribution, we proposed a method by using new features extracted from email and deep-learning approach to detect zero-day malicious spam (malspam). We have successfully extracted 27 features from email's header and body part, included machine translation detected, risk words detected and other features by using several APIs. We also use 4 different languages email dataset for more diversity and realistic purpose to build a words database and create features. Our experiment results show the accuracy rate of a zero-day malspam detection is about 78% and 92.8% for normal spam. Thus, we believe that the system still can be improved by adding more malicious spam dataset to train the system, as well as using a better accuracy translation API.

In the other hand, the traditional antivirus software is using virus definition to identify malware infection. In addition, antivirus needs to update the new virus definitions to guarantee its detection accuracy. However, due to the number of malware variants and new types of them are increase, it is very difficult to detect and respond them all. Moreover, there will be a serious incident if an unknown malware that did not correspond to the data definition had installed and expanded the infection without any notification. Therefore, in third research contribution, we proposed a malware detection method using an Execution Registry Access (ERA) with URSNIF malware which has the feature of adding values to the registry for automatically executed when the computer starts up, and the feature that created high failure rate of registry access. We applied the proposed method to 8 URSNIFs and showed the detected result that the failure rate of registry access was about 35.3% and the value of 6 that access to the specific registry was calculated. The result turned out that there was a possibility that it could be used as a new countermeasure method corresponding to the subspecies. Also, there was a possibility that it could be capable used for malware other than URSNIF. And also design an Active Method (AAM) for spyware detection which detects spyware using a bait function and achieved 88.16% ACC with 76.23% TPR and 23.68% FPR which is the unique function of detecting spyware after the infection completed.

From all above research contributions, we expect this research can be used to improve a performance of cybersecurity, protect, find or help infected victims inside an organization's network as well as keep the computer system or business service smoothly moving on.