International Journal of Engineering & Technology, 7 (4) (2018) 5618-5621

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET
doi: 10.14419/ijet.v7i4.23675
Research paper

An integrated multi layers approach for detecting unknown malware behaviours

Humam Imad Wajeeh Al-Shahwani [1]*, Warusia Mohamed Yassin [1], Zaheera Zainalabidin [1], Mohammed Rasheed [2]

[1] Universiti Teknikal Malaysia Melaka, faculty of communication and information, 76100 Hang Tuah Jaya, Durian Tunggal, Melaka, Malaysia
[2] Universiti Teknikal Malaysia Melaka, Faculty of Electrical Engineering, Industrial Power, 76100 Hang Tuah Jaya, Durian Tunggal, Melaka, Malaysia
**Corresponding author E-mail: humam_alshahwany@yahoo.com*

Abstract

Malware represents one of the dangerous threats to computer security. Dynamic analysis has difficulties in detecting unknown malware. This paper developed an integrated multi – layer detection approach to provide more accuracy in detecting malware. User interface integrated with Virus Total was designed as a first layer which represented a warning system for malware infection, Malware data base within malware samples as a second layer, Cuckoo as a third layer, Bull guard as a fourth layer and IDA pro as a fifth layer. The results showed that the use of fifth layers was better than the use of a single detector without merging. For example, the efficiency of the proposed approach is 100% compared with 18% and 63% of Virus Total and Bellegarde respectively.

Keywords: *Registry; Virus Total; Bullugruad; IDA Pro; Cuckoo and Multi-Layer Malware Detection.*

1. Introduction

Malware refers to malicious software. It is written to cause damage to data, devices or persons. Malware is classified into many types such as viruses, Trojans, spyware and worms. Malware writers such as the developers of the BLACK HAT software sell their skills and abilities to whoever can pay more for their services like criminal organizations for the digital world businesses or government intelligent agencies that tend to get blocked data from computers, networks or mobile devices to reach their goals [1], [6],[7 - -10].

Malware represents one of the most dangerous threats to computer security with samples reaching to more than 140 million populations in 2015. Signature matching remains the core defense against malware. Defense against malware attack becomes difficult with static analysis because malware possesses evasion techniques such as polymorphism, obfuscation and encryption [2], [3], [11 - 13], [20].

An efficient way to defend against malware is through Dynamic analysis known as the sandbox which utilizes string matching approaches. A popular way to execute this method is by analyzing the binary system of the behaviors in order to detect and determine the existence of malware in a controlled environment. In this environment, the binary behaviors of the samples are viewed in depth in order to be classified into practical malware threat or malware families or none.

The strength of this sand box is that it is designed to assume that those behaviors are randomized, and it is difficult to discover its signals. Then, their interactions with the hardware and the operating system and with the system resources are analyzed. The Dynamic analysis in general often focus on the system calls which is the only way for the application to interact with the hardware used in this analysis to examine unknown samples. These samples are analyzed, and their effects are put in the table of victors to compare these effects with the malware behaviors. Any matching occurrences will be classified as malwares. The Dynamic analysis can fix many evasion attacks such as shadow attacks and induction attacks. Furthermore, the dynamic analysis uses based-signal method to detect malware where signals are produced from many types of malware actions such as Ransom war and adware which execution requires interactions with a visible resource at operating system level and the interactions produce signals that are difficult to hide [2], [4], [5], [14 - 16] [22]. The current available dynamic analysis has difficulties in detecting unknown malwares. Further analyses against malware behaviors are lacking, causing challenges in detecting their patterns more correctly. This paper satisfied the requirement for determining more accurate malware detection to vary important parts of windows 7 registry which are Hash-key user and Hash-key local machine. The paper integrated a multilayer detection system from malware in order to provide an.

2. Research methods

The proposed method aimed to detect malware that infects the files of two registries from the registry of the windows (Hash Key User and Hash Key Local Machine). The method consisted of five sequence integrated layers; the other contents of the proposed method are preparing the required data to feed into the fives- layers properly and recording the final results of those fives layers properly as shown in Figure 1. This method has nine steps; extracting the Hash key user (HKU) and Hash key local machine (HKLM) files from the registry of windows and putting those files in a folder; checking those files with Virus Total, as a first layer of detection; converting the extracted files into a binary form; extracting the header of each converted file with the header of mal-

ware samples that are stored in malware database as a second layer of detection.; considering the existing files in the malware database as a detected malware; skipping the files that are considered as a detected malware for the next security procedures; applying the other files to the layers 3,4 and 5; checking the results of the previous three layer detection, if any one of them is one and finally if any of the output from the 3,4 and 5 layers is one ,the file is considered a malware as shown in Table 1.

Table 1 describes how the detected file can be considered as a malware or not, where each one of cuckoo, IDA Pro and Bellegarde has two states; one if the tool considers the file as a malware or zero if the tool decides the file is not a malware. If the product and operation are equal to zero, then, when all the three tools decide the file is not malware, it could be classified as a non-malware, otherwise, if the product and operation are equal to one, the file is considered as a malware.

Table 1: The Design Outputs

And Operation Output	Cuckoo	Blugrad	Ida
1	1	1	1
0	0	0	0
0	1	0	0
0	0	1	0
0	0	0	1
0	1	1	0
0	0	1	1
0	1	0	1

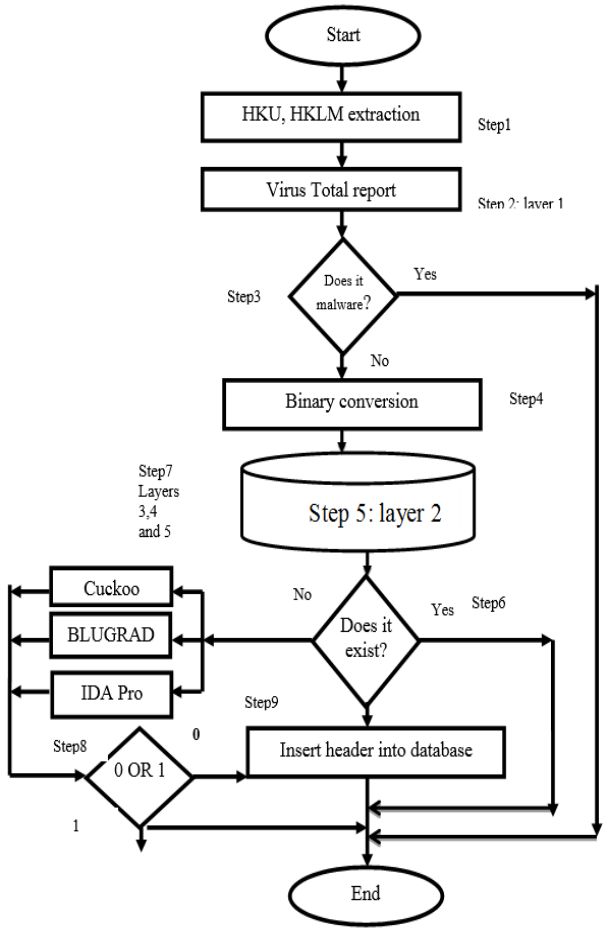

Fig. 1: Proposed Method Step by Step.

The proposed method aimed to detect malware that infect two registries from the registry of the windows (Hash Key User and Hash Key Local Machine). The method consisted of five sequence integrated layers; the other contents of the proposed method are preparing the required data to feed into the fives- layers properly, and to take out the final results of those fives layers properly.

This method has nine steps, in the first step the method extract the Hash key user (HKU) and Hash key local machine (HKLM) files from the registry of windows, and put those files in a folder. In the second step the method check those files with virus total, as a first layer of detection. In the third step the method convert the extracted files into binary form. In fourth step the method extract the header of each converted file with the header of malware samples that are stored in malware database, as a second layer of detection. In the fifth step the method will consider the existing files in the malware database as a detected malware. In the sixth step the method will skip the files that consider as a detected malware from the next security procedures. In the seventh step the method applying the other files to the layers 3,4 and five. In the eight step the method will check the result of the previous three layers detection, if any one of them is one. In the ninth step if any one of the output of the 3,4 and 5 layers is one, the file will consider as a malware.

Step1: extract (HKU) and (HKLM) files.
Step2: check the extracted files with virus total (first layer).
Step3: Convert the extracted files into binary form.
Step4: Compare the header of each converted file with the headers stored at the database (layer 2).
Step5: If the file exists in the malware database, then, it will be considered as a detected malware.
Step6: If a file is detected as a malware, it will not proceed through the other detecting tools.
Step7:. Apply cuckoo, bullugruad and IDA pro on the file as the layer 3,4 and 5.
Step8: check whether the output of any one from 3,4 and 5 layers is one.
Step9: If any one of cuckoo, bullugraud, and IDA Pro detects the file as a malware then it will be considered as a malware.

As a final result of the proposed method output describe how the detected file can be considered as a malware or not, where each one of cuckoo, IDA Pro and Bullugraud had two states; one if the tool considered the file as a malware or zero if the tool decided the file was not a malware. If the product and operation were equal to zero, then, when all the three tools decided the file was not malware then it could be classified as a non-malware, otherwise, if the product and operation were equal to one, the file was considered as malware.

3. Results and analysis

3.1. Subsection

Table 2 describes the detection of malware samples through Bull Guard tool. The zero state means undetected malware whereas the one state means detected malware which can consist of virus, Trojan and backdoor. The Bullguard caught 7 out of 11 malware samples.

The efficiency of Bullugruard.

TPR=TP/TP+TN*100% =

TPR= 7/7+4 *100% = 7/11 *100% = 63%.

Table 2: Describes the Detection of Malware Samples by Bull Guard Tool

Malware	Type	Tool	Detection
Vsd33.exe	Virus	Bullguard	0
Hxdofena.exe	Generic	Bullguard	0
Sbplus.exe	Virus	Bullguard	1
Hxdefloor.exe	Backdoor	Bullguard	1
File-loopad.exe	Trojan	Bullguard	1
Pc-setup.exe	Trojan	Bullguard	1
Aig32.exe	Trojan	Bullguard	1
Bdcli100.exe	Trojan	Bullguard	1
Hxdef100.exe	Backdoor	Bullguard	1
Av-test-tcp.exe	Virus	Bullguard	0
Rdrbs100.exe	backdoor	Bullguard	0

3.2. Subsection2

Table 3 describes malware detection through Virus Total tool. The zero state means undetected malware whereas the one state means detected malware from all types of virus. The tool detected 2 out of 11 samples.
The efficiency of Virus Total is:

TPR= TP/TP+TN * 100%.

TPR=2/2+9 *100% =

TPR=2/11*100% = 18%.

Table 3: Describe the Detection of Malware by Virus Total Tool

Malware	Type	Tool	Detection
Vsd33.exe	Virus	Virus total	1
Hxdofena.exe	Generic	Virus total	0
Sbplus.exe	Virus	Virus total	0
Hxdefloor.exe	Backdoor	Virus total	0
File-loopad.exe	Trojan	Virus total	0
Pc-setup.exe	Trojan	Virus total	0
Aig32.exe	Trojan	Virus total	0
Bdcli100.exe	Trojan	Virus total	0
Hxdef100.exe	Backdoor	Virus total	0
Av-test-tcp.exe	Virus	Virus total	1
Rdrbs100.exe	backdoor	Virus total	0

3.3. Subsection 3

Table 4 describes the detection of malware by the integrated approach. The zero state means undetected malware whereas the one state means detected malware. The approach detected viruses like backdoor, Trojan and generic. The approach detected 11 out of 11 samples.
TPR=11/11 *100% = 100% is the efficiency of the integrated multi-layer approach. The efficiency of the integrated multi-layer approach in comparison with single detectors that are used without merging showed that the multi-layer approach was much better and more efficient for detecting malware behaviors.

Table 4: Describes the Detecting of Malware by the Integrated Approach

Malware	Type	Tool	Detection
Vsd33.exe	Virus	Integrated approach	1
Hxdofena.exe	Generic	Integrated approach	1
Sbplus.exe	Virus	Integrated approach	1
Hxdefloor.exe	Backdoor	Integrated approach	1
File-loopad.exe	Trojan	Integrated approach	1
Pc-setup.exe	Trojan	Integrated approach	1
Aig32.exe	Trojan	Integrated approach	1
Bdcli100.exe	Trojan	Integrated approach	1
Hxdef100.exe	Backdoor	Integrated approach	1
Av-test-tcp.exe	Virus	Integrated approach	1
Rdrbs100.exe	backdoor	Integrated approach	1

4. Conclusion

Malware is one of the major security threats that can damage computer operation. Malware writers try to avoid detection using several techniques such as polymorphic, metamorphic and also hiding technique. In order to overcome that issue, an integrated multi-layer approach is proposed for detecting unknown malwares. The end results show that the proposed approach is more efficient than the use of single detector to detect malware for the efficiency of proposed approach is 100% compared with 18% and 63% for Virus Total and Bellegarde respectively.

Acknowledgement

I would like to thank for faculty of communication and information (FTMK), Universiti Teknikal Malaysia Melaka (UTeM).

References

[1] Stamatatos, E., 2009. A Survey of Modern Authorship Attribution Methods. Journal of the American Society for Information Science and Technology, 60(3), pp.538–556.

[2] V. Surducan and E. Surducan, "Low-Cost Microwave Power Generator for Scientific and Medical Use [Application Notes]," in IEEE Microwave Magazine, vol. 14, no. 4, pp. 124-130, June 2013. doi: 10.1109/MMM.2013.2248651

[3] Cho JH, Chang SA, Kwon HS, Choi YH, KoSH, Moon SD, Yoo SJ, Song KH, Son HS, Kim HS, Lee WC, Cha BY, Son HY & Yoon KH (2006), Long-term effect of the internet-based glucose monitoring system on HbA1c Reduction and glucose stability: a 30-month follow-up study for diabetes management with a ubiquitous medical care system. *Diabetes Care* 29, 2625–2631. https://doi.org/10.2337/dc05-2371.

[4] Fauci AS, Braunwald E, Kasper DL & Hauser SL (2008), Principles of Harrison's Internal Medicine, Vol. 9, 17thedn. *McGraw-Hill*, New York, NY, pp.2275–2304.

[5] Kim HS & Jeong HS (2007), A nurse short message service by cellular phone in type-2 diabetic patients for six months. *Journal of Clinical Nursing* 16, 1082–1087. https://doi.org/10.1111/j.1365-2702.2007.01698.x.

[6] Lee JR, Kim SA, Yoo JW & Kang YK (2007), The present status of diabetes education and the role recognition as a diabetes educator of nurses in korea. Diabetes Research and Clinical Practice 77, 199–204. https://doi.org/10.1016/j.diabres.2007.01.057.

[7] McMahon GT, Gomes HE, Hohne SH, Hu TM, Levine BA & Conlin PR (2005), Web-based care management in patients with poorly controlled diabetes. *Diabetes Care* 28, 1624–1629. https://doi.org/10.2337/diacare.28.7.1624.

[8] Thakurdesai PA, Kole PL & Pareek RP (2004), Evaluation of the quality and contents of diabetes mellitus patient education on Internet. *Patient Education and Counseling* 53, 309–313. https://doi.org/10.1016/j.pec.2003.04.001.

[9] Stiborek, J., Pevný, T. & Rehák, M., 2018. Multiple instance learning for malware classification. Expert Systems with Applications, 93, pp.346–357.

[10] WAGNER, M. et al., 2017. A Knowledge-Assisted Visual Malware Analysis System:Design, Validation, and Reflection of KAMAS. Elsevier computers and security, pp.1–15.

[11] Stamatatos, E., 2009. A Survey of Modern Authorship Attribution Methods. Journal of the American Society for Information Science and Technology, 60(3), pp.538–556.

[12] Benzmüller, R., 2017. malware-trends-2017. Available at: https://www.gdatasoftware.com/blog/2017/04/29666-malware-trends-2017 04/10/2017.

[13] Z. et al., 2013. A survey on heuristic malware detection techniques. IKT 2013 - 2013 5th Conference on Information and Knowledge Technology, (May), pp.113–120

[14] Bazrafshan, Z. et al., 2013. A survey on heuristic malware detection techniques. IKT 2013 - 2013 5th Conference on Information and Knowledge Technology, (May), pp.113–12

[15] Hang, H. et al., 2016. "Infect-me-not": A User-centric and Site-centric Study of web-based malware. , pp.234–2

[16] Chaczko, Z. & Ahmad, F., 2009. "Wireless Sensor Network Based System for Fire Endangered Areas." In Third International Conference on Information Technology and Applications. 2 (4–7). pp. 203–207.

[17] Bidoki, S.M., Jalili, S. & Tajoddin, A., 2017. PbMMD: A novel policy based multi-process malware detection. Engineering Applications of Artificial Intelligence, 60(August 2016), pp.57–70. Available at: http://dx.doi.org/10.1016/j.engappai.2016.12.008

[18] Tanaka, Y., Akiyama, M. & Goto, A., 2017. Analysis of malware download sites by focusing on time series variation of malware. Journal of Computational Science, 22, pp.301–313. Available at: https://doi.org/10.1016/j.jocs.2017.05.027.

[19] Maestre Vidal, J., Sandoval Orozco, A.L. & García Villalba, L.J., 2017. Alert correlation framework for malware detection by anomaly-based packet payload analysis. Journal of Network and Computer Applications, 97(February), pp.11–22. Available at: http://dx.doi.org/10.1016/j.jnca.2017.08.010.

[20] Ceron, J.M., Margi, C.B. & Granville, L.Z., 2017. MARS: From traffic containment to network reconfiguration in malware-analysis systems. Computer Networks, 129, pp.261–272

[21] Stamatatos, E., 2009. A Survey of Modern Authorship Attribution Methods. Journal of the American Society for Information Science and Technology, 60(3), pp.538–556.

[22] H. S. Abbas, S. A. Bakar, M. Ahmadi, and Z. Haron, "Experimental studies on corrugated steel-concrete composite slab," vol. 67, pp. 225–233, 2015.