International Journal of Advanced Smart Convergence Vol.12 No.1 59-63 (2023)
http://dx.doi.org/10.7236/IJASC.2023.12.1.59

IJASC 23-1-7

Malware Detector Classification Based on the SPRT in IoT

Jun-Won Ho

Professor, Department of Information Security, Seoul Women's University, South Korea
jwho@swu.ac.kr

Abstract

We create a malware detector classification method with using the Sequential Probability Ratio Test (SPRT) in IoT. More specifically, we adapt the SPRT to classify malware detectors into two categories of basic and advanced in line with malware detection capability. We perform evaluation of our scheme through simulation. Our simulation results show that the number of advanced detectors is changed in line with threshold for fraction of advanced malware information, which is used to judge advanced detectors in the SPRT.

Keywords: *Sequential Probability Ratio Test (SPRT), Malware Detector Classification, IoT*

1. Introduction

Nodes with sufficient computing and communication resources can act as malware detectors in IoT. As far as malware detection is concerned, we can consider basic malware that can be discerned with relatively low computing and communication resources and advanced malware that can be identified with relatively high computing and communication resources. Under this consideration, we can classify malware detectors as basic and advanced detectors in line with detection capability. In particular, we apply the Sequential Probability Ratio Test (SPRT) [4] to this malware detector classification with the threshold for fraction of advanced malware information. We discern that our simulation results are changed in line with threshold for fraction of advanced malware information.

2. Related Work

We introduce a couple of research work have been proposed in the field of malware. In [1], the relevant work to zero-day malware detection is investigated. Static malware analysis work for IoT is proposed in [2]. In [3], IoT malware detection method rooted on Markov chain behavioral model is developed . Evasive malware detection method rooted on bare-metal analysis is devised in [5].

3. Malware Detector Classification with Using the SPRT

For malware detector classification, we leverage the intuition that the more advanced malware information

Manuscript Received: December. 24, 2022 / Revised: December. 27, 2022 / Accepted: December. 29, 2022
Corresponding Author: jwho@swu.ac.kr
Tel: +82-2-970-5607, Fax: +82-2-970-5981
Professor, Department of Information Security, Seoul Women's University, South Korea

is detected by malware detector the higher likelihood of being advanced malware detector is given to it.

We define the number of malware information required for the SPRT execution as the number of malware information detected by malware detector that is needed for the SPRT to be initiated. For instance, if the number of malware information required for the SPRT execution is set to 10, the SPRT is performed each time 10 malware information is detected by malware detector.

Moreover, we define the fraction of advanced malware information as the number of advanced malware information detected by malware detector over the total number of malware information detected by malware detector, where the total number of malware information is sum of the number of basic malware information detected by malware detector and the number of advanced malware information detected by malware detector. The fraction of advanced malware information is computed each time the number of malware information required for the SPRT execution is met. We also configure threshold for the fraction of advanced malware information, which is a threshold value used for decision process in the SPRT. We assume that a central entity runs the SPRT for each malware detector. Furthermore, each sample is assumed to be judged as Bernoulli random variable which is independent and identically distributed, where EI is a success probability in Bernoulli distribution.

In the SPRT, we define a null hypothesis as a hypothesis that malware detector is basic malware detector. We also define an alternate hypothesis as a hypothesis that malware detector is advanced malware detector. The specific procedure of the SPRT is as follows: Variable DS is made use of counting the number of samples in the SPRT and Variable DT is made use of counting the number of samples with alternate hypothesis type in the SPRT. Both DS and DT are initialized to 0. Note that Q (resp. R) is a user-set false-positive rate (resp. user-set false-negative rate). Both EI_0 and EI_1 are pre-configured parameters such that $EI_0 < EI_1$. The case that EI is greater than or equal to EI_1 will lead to the higher likelihood at which the SPRT selects an alternate hypothesis. On the other hand, the case that EI is smaller than or equal to EI_0 will lead to the higher likelihood at which the SPRT selects a null hypothesis.

Each time the number of malware information needed for the SPRT execution is satisfied, the fraction of advanced malware information is computed. The fraction of advanced malware information is thought of as a sample of the SPRT. If the fraction of advanced malware information is larger than or equal to threshold for the fraction of advanced malware information, doSPRT(1) procedure is executed. Otherwise, doSPRT(0) procedure is executed.

The specific procedure of doSPRT(type) is defined as:

DS = DS+1;
If type == 1, then DT = DT+1;

JK and PK are variables acting as decision threshold for null hypothesis and alternate hypothesis, respectively. JK and PK are computed as follows:

$$JK = \frac{\ln\frac{R}{1-Q} + DS\ln\frac{1-EI_0}{1-EI_1}}{\ln\frac{EI_1}{EI_0} - \ln\frac{1-EI_1}{1-EI_0}} \quad , \quad PK = \frac{\ln\frac{1-R}{Q} + DS\ln\frac{1-EI_0}{1-EI_1}}{\ln\frac{EI_1}{EI_0} - \ln\frac{1-EI_1}{1-EI_0}}$$

If DT <= JK, then DT=DS=0; Accept null hypothesis;
If DT >= PK, then DT=DS=0; Accept alternate hypothesis; return 1;
return 0;

4. Performance Evaluation

For the evaluation of our devised method, we implement a basic simulation program pondering the following case: It is assumed that each malware detector randomly selects the number of malware information between minimum number of malware information and maximum number of malware information, determines whether each chosen malware information is advanced or not with a randomly chosen probability, and performs the SPRT whenever the number of malware information checked in this advanced malware information decision process is equal to the number of malware information required for the SPRT execution.

We set Q=R=0.01. We think over two configurations of (EI_0, EI_1)= (0.3, 0.7), (0.1, 0.9). We configure the number of detectors to 100. We also set the minimum number of malware information to 1000 and the maximum number of malware information to 10000, the number of malware information required for the SPRT execution to 10. Furthermore, we think over a configuration set of threshold for fraction of advanced malware information, (0.5, 0.6, 0.7, 0.8, 0.9). We also think over the incurrence of error that basic (resp. advanced) malware information is misidentified as advanced (resp. basic) malware information and that error probability is configured to 0.01.

Our simulation is reiterated 100 times and we exhibit average results of our evaluation results. As shown in Figures 1,2, we see that the number of advanced malware detectors decays as threshold for fraction of advanced malware information rises. As displayed in Figure 3,4, however, we discern that the number of advanced malware information per advanced malware detector increases as threshold for fraction of advanced malware information rises. This means that the higher threshold for fraction of advanced malware information contributes to delay in the end of the SPRT with an alternate hypothesis acceptance, incurring the larger number of advanced malware information per advanced malware detector.

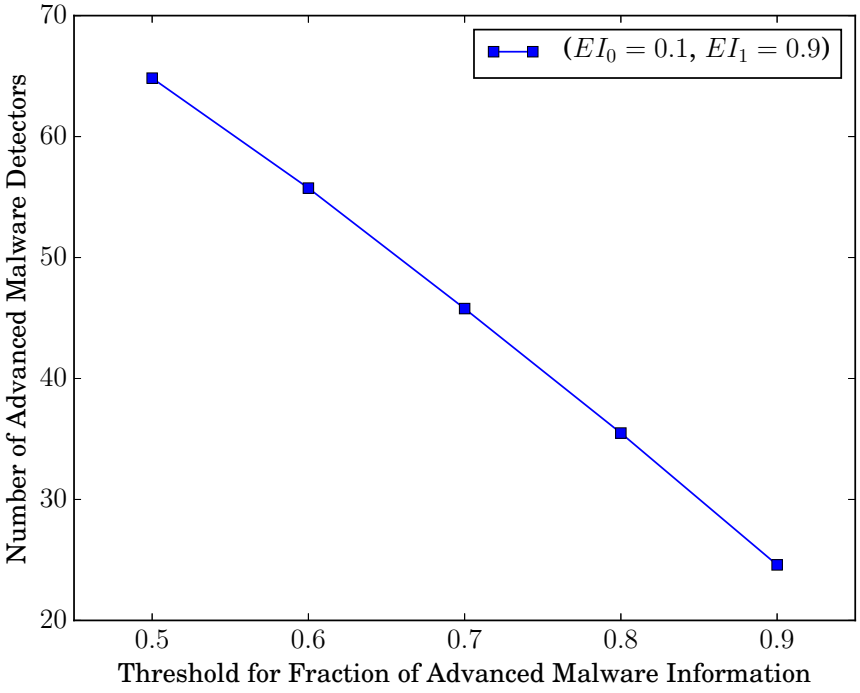

Figure 1. Effect of threshold for fraction of advanced malware information on the number of advanced malware detectors when EI_0=0.1, EI_1=0.9.

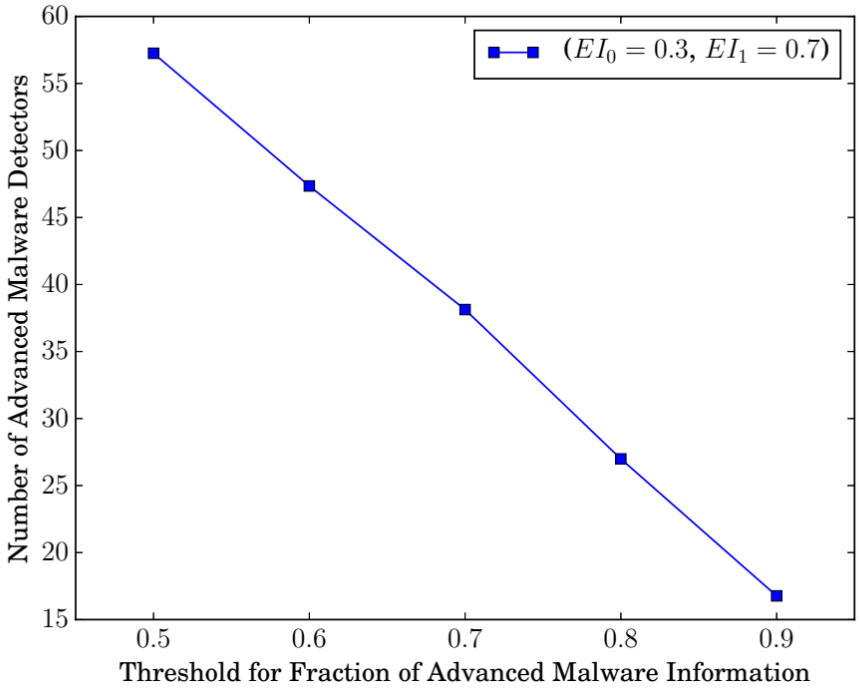

Figure 2. Effect of threshold for fraction of advanced malware information on the number of advanced malware detectors when EI_0=0.3, EI_1=0.7.

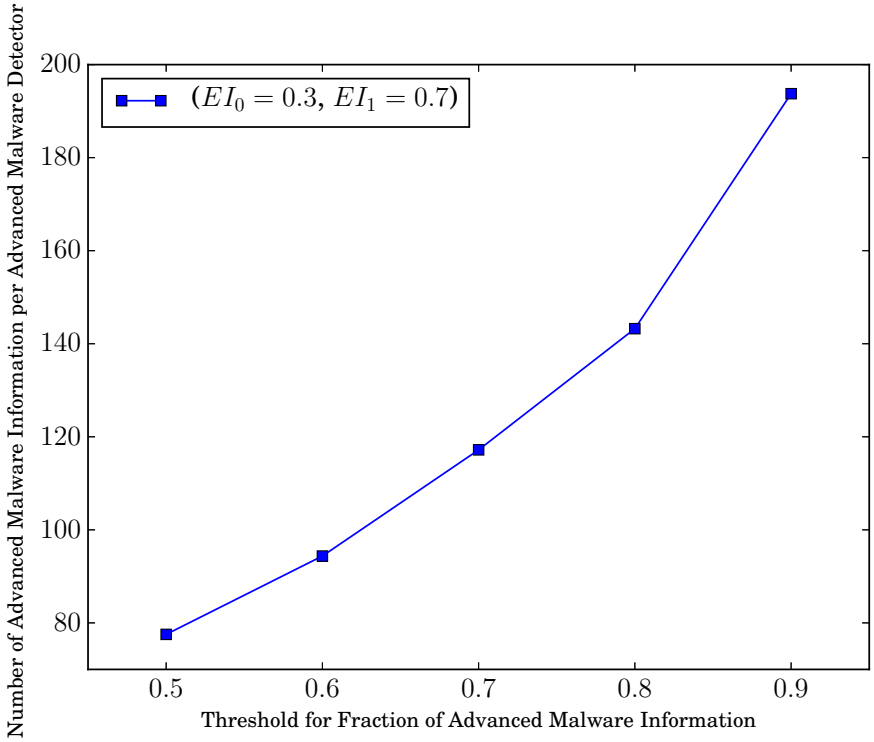

Figure 3. Effect of threshold for fraction of advanced malware information on the number of advanced malware information per advanced malware detector when EI_0=0.3, EI_1=0.7.

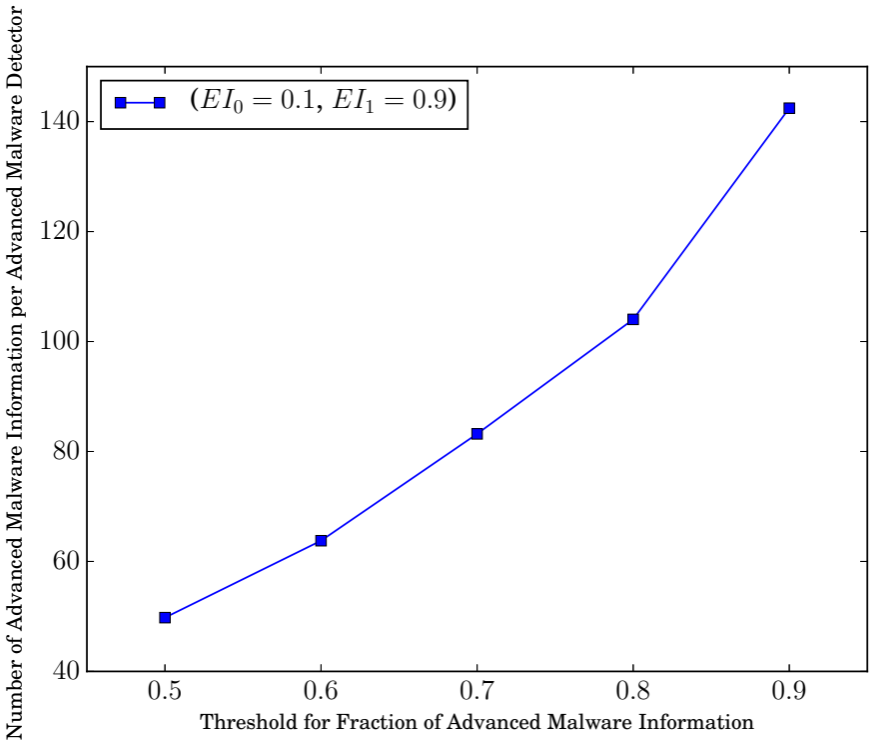

Figure 4. Effect of threshold for fraction of advanced malware information on the number of advanced malware information per advanced malware detector when EI_0=0.1, EI_1=0.9.

5. Conclusion

We create a method classifying malware detectors into basic and advanced detectors by utilizing the SPRT. Moreover, we evaluate our devised method through simulation. From our evaluation results, we recognize that the number of advanced malware detectors decreases and the number of advance malware information per advanced malware detector increases as threshold for fraction of advanced malware information rises.

Acknowledgement

This work was supported by a research grant from Seoul Women's University(2023-0001).

References

[1] E. Gandotra, D. Bansal and S. Sofat, "Zero-day malware detection," *2016 Sixth International Symposium on Embedded Computing and System Design (ISED)*, Patna, India, 2016, pp. 171-175, DOI: https://doi.org /10.1109/ISED.2016.7977076.

[2] A. Ravi and V. Chaturvedi, "Static Malware Analysis using ELF features for Linux based IoT devices," *2022 35th International Conference on VLSI Design and 2022 21st International Conference on Embedded Systems (VLSID)*, Bangalore, India, 2022, pp. 114-119, DOI: https://10.1109/VLSID2022.2022.00033.

[3] M. Ficco, "Detecting IoT Malware by Markov Chain Behavioral Models," *2019 IEEE International Conference on Cloud Engineering (IC2E)*, Prague, Czech Republic, 2019, pp. 229-234, DOI: https://10.1109/IC2E.2019.00037.

[4] A. Wald. Sequential Analysis, Dover, 2004.

[5] D. Kirat, G. Vigna, C. Kruegel. BareCloud: Bare-metal Analysis-based Evasive Malware Detection. In Usenix Security, 2014.