To: Bonnie Hsia, Professor

From: Partha Sarathi Ghosh

Subject: Analysis of a Professional Journal Article

Date: December 13, 2019

Purpose

The purpose of this memo is to analyze the journal article, *The Evolution of Android Malware and Android Analysis Techniques*, from the technical journal ACM Computing Surveys (CSUR) by (Tam, Kimberly and Feizollah, Ali and Anuar, Nor Badrul and Salleh, Rosli and Cavallaro, Lorenzo, February 2017) for its readability and pseudoscience.

Summary

The article *The Evolution of Android Malware and Android Analysis Techniques* [Tam et al., 2017] from the technical journal CSUR is reviewed to determine how well it is written as a technical document and if the authors have used proven scientific methods for the technical document. It is concluded that the article proves its technical content with ample evidences of the use of scientific method or design model.

Journal and Article

The title of the article being analyzed is *The Evolution of Android Malware and Analysis Technique* [Tam et al., 2017]. The article has been published in the peer reviewed journal CSUR, Volume 49 Issue 4, February 2017, Article No. 76. CSUR is published quarterly and is available both online and in print. CSUR is edited by Albert Zoyama [Zomaya, 2017], who is the director of the Australian Research Council Professorial Fellow Chair of High Performance Computing and Networking, at Centre for Distributed and High Performance Computing in School of Computer Science. On its website ACM states, ”Computing Surveys publishes comprehensive, readable tutorials and survey papers that give guided tours through the literature and explain topics to those who seek to learn the basics of areas outside their specialties” [ACM, 2019]. It is to be noted that CSUR does not publish ”new” research [ACM, 2019]. The target audiences for publication from ACM are professionals from computer and software industry and academics.

The Authors

There are five authors of this paper.

- Kimberly Tam is a researcher in mobile security with Information Security Group (ISG), Royal Holloway, University of London. Miss Tam has a Ph.D. in Information Security and is the author of three other papers in the area of android malware [Dash et al., 2016], [Tam, 2016] and [Tam et al., 2015]
- Ali Feizollah is a research fellow in the Deputy Vice Chancellor(Research & Innovation) Office at University of Malaya, Kuala Lumpur, Malaysia [Feizollah, 2019]. Mr. Feizollah has published 16 articles in the area of computer security. ”Evaluation of machine learning classifiers for mobile malware detection” is his most widely cited article [Narudin et al., 2016].
- Dr. Nor Badrul Anuar is an associate professor at the Department of Computer, System and Technology, University of Malaya. Dr. Anuar has published over hundred technical articles in the area of computer security. His article ”The rise of big data on cloud computing: Review and open research issues” [brahim Abaker Targio Hashem et al., 2015] has been cited in over 1700 different technical papers on cloud computing and computer security.

- Dr. Rosli Salleh is an associate professor at the Department of Computer, System and Technology, University of Malaya. His research area has been in mobile malware. He has published over 25 articles in this area. Once of his key work has been in the area of wireless security titled "Overview of Security Issues in Wireless Sensor Networks" [Modares et al., 2011], which points out the fragility of wireless sensors' security mechanisms.
- Lorenzo Cavallaro is a professor of computer science and is the chair of Cybersecurity at Kings college, London. He has published over 90 articles in the area of computer security [Scholar,]. "Your botnet is my botnet: Analysis of a botnet takeover" is his most cited publication (over 750 citation) [Stone-Gross et al., 2009].

This technical article [Tam et al., 2017] has been refereed, cited by 18 papers and downloaded 2832 times as of . These facts about the paper indicate that this is a quality technical paper.

Introduction

This is an article which chronicles the evolution of the Android Operating System (OS). The authors describe proliferation of Android mobile devices in our lives. This has also caused the mobile malware in Android OS to grow at a rapid rate. The authors use statistics from the year 2010 to 2015 to establish these facts by comparing Android usage with the other mobile operating systems. They define the scope of their research in these words - *"Unlike previous works, this article is not a general study on mobile . . . but instead focuses on Android-related analysis techniques systematically and in detail."* [Tam et al., 2017]
The article walks the path of the evolution of Android malware using the method of "*classification* and not *partition*" [Markel and Selber, 2018] to identify the types of mobile malware. However, the techniques used to develop the malware for Personal Computers (PC) are different from those of the mobile devices. The technique of *"comparision and contrast"*, a technique described by Merkel in [Markel and Selber, 2018] is used to explain the inherent hardware architectural differences between these two computing devices. The authors organize the information on solving the problem of mobile malware using *problems-methods-solution* approach as described by Merkel in [Markel and Selber, 2018].
This paper writes in detail about the different methods of malware detection used. The authors describe the static and dynamic analysis methods and then explain how a hybrid approach could be more beneficial for malware detection. The authors say that machine learning methods, coupled with hybrid mechanism is a possible solution for malware detection in the current Android OS landscape.
After reading this article I have concluded that the proliferation of the malware in Android OS has been because of the rapid growth of the Android OS. The method to mitigate the growth of malware is to analyze the malware with a hybrid approach of static analysis, dynamic analysis, and machine learning. The authors of this technical paper hypothesize a solution to detect malware in a constantly changing threat-landscape for Android OS.

Visual Aids

The authors substantiate their arguments with relevant statistics using numerous tables and graphs conforming the American Psychological Association (APA) style of citation. In one of the graphs, the smart phone operating systems' market share data serves well as a visual aid for the reader. These graphs are important tools for conveying complex data to the readers. Since the authors traverse the evolution of malware proliferation in Android OS from the year 2010 to 2015, there are graphs to show progression of a single data annually, like mobile usage among people and worldwide smartphone market shares. Another good example of visual aid is the description of Android Linux kernel as shown in Figure 2. Any reader would appreciate the clarity of the system blocks in this architectural diagram.
In Figure 3, the authors' citations overwhelm the Venn diagram. The Venn diagram tries to explain the different static analysis methods used in Android OS. The authors could have removed the citations to a separate table to make the figure cleaner. This Venn diagram should have had only the static analysis

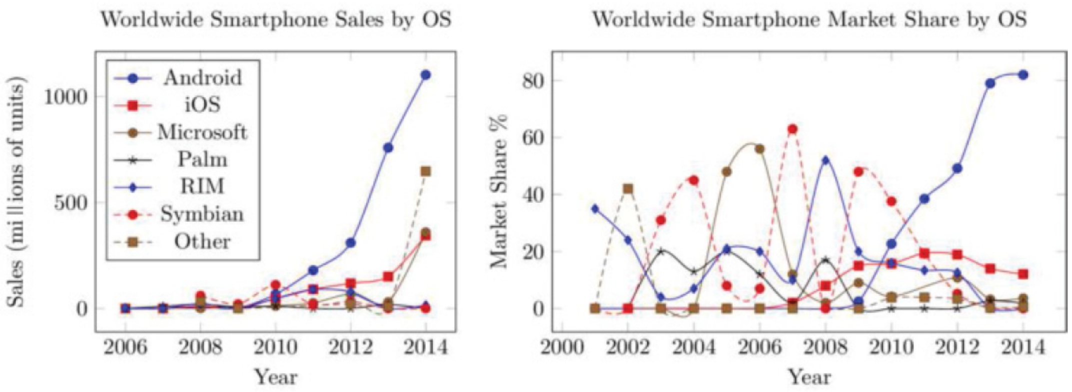

Figure 1: Smartphone market share, [Tam et al., 2017, fig 1]

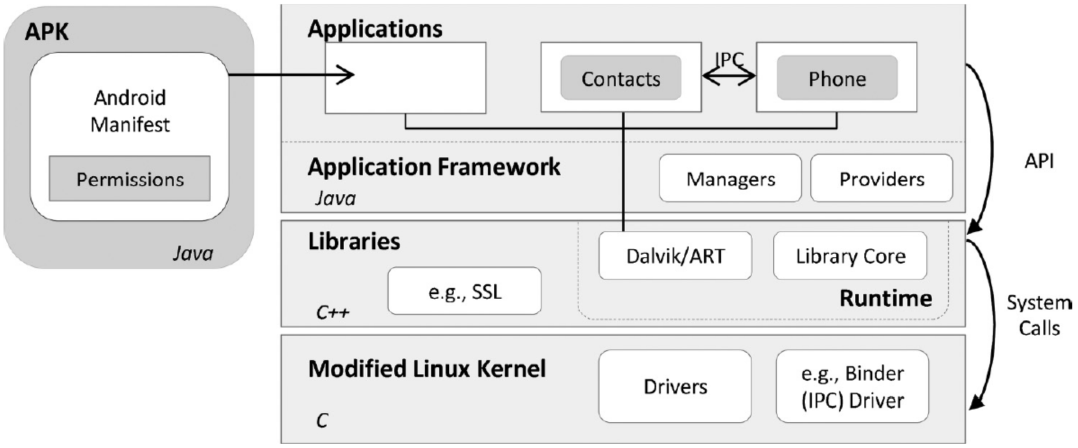

Figure 2: Overview of Android Operating System Architecture [Tam et al., 2017, fig 2]

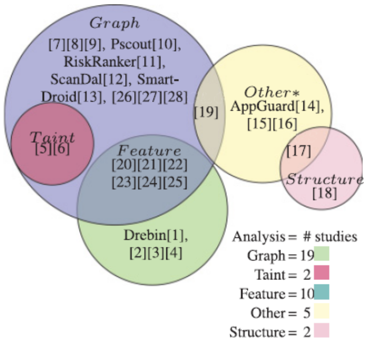

Figure 3: Venn diagram of static analysis methods [Tam et al., 2017, fig 3 b]

methods and descriptions. Thus, in this particular case, the authors' purpose of using the Venn diagram is defeated.

Readability of the Article

The readability of this article is analyzed and assessed using the sections “Characteristics of a Technical Document” and “Measures of Excellence in Technical Documents” from [**?**].

Characteristics of Technical Writing

The five authors of this article have worked collaboratively using the methods of *classification*, *comparison and contrast*, and *problems-methods-solution* [Markel and Selber, 2018] to organize the information in this technical paper. The authors use *cause and effect* pattern as described in [Markel and Selber, 2018] to depict the “time vs mobile OS usage” and “time vs malware proliferation” graphs. Statistical information used by the authors demonstrate that they have been able to monitor the best information in the area for Android OS and malware analysis.

Measures of Excellence in Technical Writing

The authors’ research has been thorough. They have been *honest* in their approach to the problem by citing statistical data. The usage of diagrams to explain architecture brings out *clarity*. *Accuracy, comprehensiveness and conciseness* have been the key focus for the authors and they are consistent with the *problems-methods-solution* [Markel and Selber, 2018] approach in the paper. The document is written following the APA style, which provides a *professional appearance* to the document. The document is *usable* for future studies because of the data that the authors use to substantiate their arguments on malware solutions. In terms of *correctness*, the authors have reviewed the document well as there are hardly any linguistic errors.

Scientific Validity of the Article

The authors of this paper describe Android architecture with a few diagrams. This is important for the reader who may not be familiar with the Android architecture. It must be noted here that though Android OS has a Linux kernel, the application architecture is different. This is because most of the mobile devices that have Android OS use the ARM hardware architecture, smaller power requirements, different storage hardware, etc.

The usage of tables to illustrate different multi variable comparison is very apt. This helps to read the document more effectively and identify the points of differentiation with greater clarity. The authors use graphs to demonstrate single variable changes over the years, for example, Fig 1 [Tam et al., 2017]. The explanations of the dynamic analysis and static analysis are written with great detail. Other mobile operating systems are detailed in this paper. I think this is redundant because the focus of the paper is on Android OS malware. It was unnecessary to explain any of the other mobile operating systems as the security features of those devices are not in the scope of this research. In this paper, the authors should have provided details on Virtual Machine (VM) and Just In Time (JIT) compilation, as there are variants of implementations. VM and JIT are integral parts to devise any security mechanism in Android OS.

Scientific method:

The authors have researched extensively over 100 technical papers. To substantiate their argument, the authors have compared papers and their approaches on malware detection. This technical paper is full of references of statistics like, growth percentage of malware year over year, usage of different mobile operating systems, financial incentives of a malware, etc. The authors also look at the types of static analysis methods and dynamic analysis methods. It looks into different techniques of dynamic analysis. In this paper, the authors conclude with empirical evidence that a hybrid approach would be the way for the future of malware analysis and detection. The finding of the researchers are significant because the authors realize that mobile device ecosystem is dynamic and only a hybrid approach will lead to an effective solution for malware detection. In the Taxonomy of Mobile Malware Analysis [Tam et al., 2017] section of the technical paper, the authors detail the ineffectiveness of malware detection, using binary signature analysis. The malware could evade detection by rearranging the text or the data section of malware binary. This creates a challenge and opportunity to the researchers to look for newer techniques of malware detection. The authors look into different techniques used in static code analysis, dynamic code analysis and hybrid approach at length.

To explain the dynamic analysis method, the authors research into the Application Programming Interface (API) and its interaction with the kernel, to build up the security hole context. Since dynamic analysis can expose the executable part of the source code (for malware analysis), the authors describe different input techniques that could be used to create a graph to explore the different paths of code execution. Dynamic analysis needs to be done at different architectural layers (application, kernel, device drivers, hardware interactions, etc.).

The authors also look into the evolution of the evasion techniques used by the malware. One of the simple techniques that the malware uses is to run the application at odd hours when the users are not using the devices. The authors identify that permission infrastructure is the weak point in the Android architecture. In this paper, the authors try to pivot on this architectural issue and try to devise different malware analysis and detection methods.

Proposed Solutions

The number of different mobile devices that are available complicates the problem of malware detection. It's hard to perform the analysis of the malware in all the target (available) mobile devices. The authors *propose* a hybrid approach that combines static and dynamic analysis methods with use of VM to simulate the multitude of target android devices for malware detection. The VM can be simulated to operate like a real hardware with certain changes in the software.

In this paper, the authors detail another hybrid approach in which static and dynamic analysis methods along with machine learning (ML) could be used for analysis of the malware. This method seems to be superior because it can scale to support the huge number of different hardware Android OS.

As alternative approaches, the authors explore other techniques of identifying malware like, using network traffic analysis, code coverage of API interfaces, monitoring system calls, creation of application dependency graphs for all the applications running, information-flow among the application and operating system infrastructure, inter-process communication analysis, hardware analysis, and application metadata analysis. The authors also mention that detection of malware after malware analysis is a *classification* problem. In situations where malware cannot be identified using binary classification, different attributes of the malware need to be analyzed, thereby translating the analysis problem to a multiple class *stochastic* problem.

Scope of Future Work

The authors state that the future of the Android malware research is in pursuing the path of hybrid approach that uses static analysis, dynamic analysis, and ML along with the use of virtual machines to simulate target hardware. The advances in parallel computing would allow future researchers to follow the hybrid approach of malware analysis. The authors highlight the advantages of performing code coverage analysis on the applications using the hybrid mechanism. Code coverage method can present different branches of execution, which would help in identifying the weak links in an application. Since malware has started evading detection in virtualised environment, one area of research would be to use virtualization to detect malware.

Unanswered Questions

One of the primary reasons for the proliferation of the Android malware is the failure of App store to check for malicious applications. In this paper, the authors should have detailed the mechanism of application upload and application distribution in App store. A section should have been written about this in the technical paper. The authors do not describe the installation process of application in Android OS. Details on the tools that had been used in this research is missing. The authors do not rule out or talk about the flaws in ARM architecture, which could be the reason for malware proliferation in Android.

Conclusion

The authors of this technical paper mention that they studied a wide range of Android malware analysis and detection frameworks. The technical paper created the context by describing the threat-landscape for mobile devices, followed by the evaluation of different malware detection techniques and solutions. The authors think that they have discussed threats and solutions for android devices and that they have been able to provide a superior solution than those available.

The authors acknowledge UK EPSRC grant EP/L022710/1 and Ministry of Science, Technology and Innovation, under Grant eScienceFund 01-01-03-SF0914 for funding their research for this paper.

Usefulness to your major

The San Jose State University graduate program in software engineering focuses on cybersecurity. Mobile security is a relevant topic in cybersecurity. In this technical article, the evolution of malware in Android OS is highlighted and a solution to mitigate the proliferation of malware is proposed. Reading this technical paper made me aware of the hybrid method of malware elimination in Android OS.

References

[ACM, 2019] ACM (2019). Acm csur editorial charter. https://csur.acm.org/editorial_charter.cfm#editorial-charter. (Accessed on 12/11/2019).

[brahim Abaker Targio Hashem et al., 2015] brahim Abaker Targio Hashem, IcbrarYaqoob, BadrulAnuar, N., okhtar, S. M., and AbdullahGani, S. K. (2015). The rise of big data on cloud computing: Review and open research issues - sciencedirect. https://www.sciencedirect.com/science/article/abs/pii/S0306437914001288?via%3Dihub. (Accessed on 12/11/2019).

[Dash et al., 2016] Dash, S., Suarez-Tangil, G., Khan, S., Tam, K., Ahmadi, M., Kinder, J., and Cavallaro, L. (2016). Droidscribe: Classifying android malware based on runtime behavior. In *Security and Privacy Workshops (SPW), 2016 IEEE*, pages 252–261. IEEE.

[Feizollah, 2019] Feizollah (2019). Umexpert - dr. ali feizollah. https://umexpert.um.edu.my/ali-feizollah.html. (Accessed on 12/11/2019).

[Markel and Selber, 2018] Markel, M. H. and Selber, S. A. (2018). *Organizing Your Informations*. Bedford/St. Martins.

[Modares et al., 2011] Modares, H., Salleh, R., and Moravejosharieh, A. (2011). Overview of security issues in wireless sensor networks. In *2011 Third International Conference on Computational Intelligence, Modelling Simulation*, pages 308–311.

[Narudin et al., 2016] Narudin, F. A., Feizollah, A., Anuar, N. B., and Gani, A. (2016). Evaluation of machine learning classifiers for mobile malware detection. *Soft Computing*, 20(1):343–357.

[Scholar,] Scholar, G. Lorenzo cavallaro - google scholar citations. https://scholar.google.com/citations?user=oWT7fIYAAAAJ&hl=en.

[Stone-Gross et al., 2009] Stone-Gross, B., Cova, M., Cavallaro, L., Gilbert, B., Szydlowski, M., Kemmerer, R., Kruegel, C., and Vigna, G. (2009). Your botnet is my botnet: Analysis of a botnet takeover. In *Proceedings of the 16th ACM Conference on Computer and Communications Security*, CCS '09, pages 635–647, New York, NY, USA. ACM.

[Tam, 2016] Tam, K. (2016). *Analysis and Classification of Android Malware*. PhD thesis, Royal Holloway, University of London.

[Tam et al., 2017] Tam, K., Feizollah, A., Anuar, N. B., Salleh, R., and Cavallaro, L. (2017). The evolution of android malware and android analysis techniques. *ACM Comput. Surv.*, 49(4):76:1–76:41.

[Tam et al., 2015] Tam, K., Khan, S., Fattori, A., and Cavallaro, L. (2015). Copperdroid: Automatic reconstruction of android malware behaviors. In *NDSS Symposium 2015*, pages 1–15.

[Zomaya, 2017] Zomaya (2017). Professor albert zomaya. https://sydney.edu.au/engineering/about/our-people/academic-staff/albert-zomaya.html. (Accessed on 12/11/2019).