TaintDroid Malware Categorizer

Imoleayo Abel - (iabel1) Davis Ancona - (dancona1) William Schneider - (wschnei1)
Swarthmore College

October 2nd, 2013

Abstract

The goal of our project is to expand upon the the output of TaintDroid[3] by in essence, "putting a name to a face," in regards to the tainted data it identifies. By studying the types of malware present today, we hope to modify TaintDroid to determine the most likely type(s) of malware present in an application based on the tainted data being sent out of a device. This can be used as an application vetting process to determine whether or not a third party application is infected with malware. The potential drawbacks of this idea are that in its infancy it will be very susceptible to false positives. However, time permitting we hope to solve this by integrating permissions of the application into its malware analysis.

1 Motivation

As technology becomes more and more advanced we are becoming less and less reliable on stationary methods of computation which we have spent a great deal of time learning to secure. In today's day and age the aspect of mobility is now almost synonymous with the business and social world. Smart phones are used daily to transfer data via a wide variety of both native and third party applications. However, for despite their usage they have yet to reach the same level of security as desktop computing. For such a large amount of data transferred via mobile devices there is not nearly enough security, especially for the type of private data many people store on there phones such as gps location, banking information, contacts, emails, etc. In order to help address this, the makers of TaintDroid have developed a system which can identify the pieces of private data being sent from your mobile device and the IP where they are being sent. Our motivation in taking on this project is to further improve mobile security by expanding upon what TaintDroid has started and categorizing the tainted data which it identifies as being output. This will allow for the development of a better vetting process for untrusted applications, helping achieve a higher level of overall security, and hopefully paving the way for a new style of machine learning based security.

2 Background

TaintDroid is capable of tracking personal data when used by allegedly benign applications. Our project seeks to measure how malware operates by tracking which personal data it uses with a modified version of TaintDroid. Thus, our research will rely on the original TaintDroid paper to understand how it works and modify its code to suit our purposes for not only tracking data but recognizing and diagnosing malware infections. We will also draw from a paper from UC Berkeley entitled "A Survey of Mobile Malware in the Wild"[4] which talks about the different types of malware on mobile devices, how they operate, what steps are being taken to combat them, and how different incentives can add to the danger and prevalence of malware attacks. This will tell us more about what types of malware we can expect to find and their behavior. Another paper from Carnegie Mel-

lon entitled, "All Your Droid Are Belong To Us: A Survey of Current Android Attacks"[5] will tell us more about Android malware attacks that exist today as well as identify some Android vulnerabilities that they exploit.

3 Idea

On a high level, the basic idea of this project is to categorize as best as possible the potential harm a software application could pose to a mobile platform from examining the tainted data. In other words, we plan to conduct an extensive study of different kinds of malware and the nature of the data they use or act upon. With this knowledge, we then create a classification model that we hope will probablilistically predict the potential damage an application could cause on the integrity of the mobile device.

4 Goals for Milestones

4.1 Milestone 1

For milestone 1, we hope to have TaintDroid running on the Galaxy Nexus One obtained for this project. Specifically, we hope to have gained familiarity with and to have a fundamental understanding of the basic functionality of TaintDroid. At this point in the project, we would be able to reproduce the some scenarios in the TaintDroid paper i.e. we would be able to identify delicate information sent out by mobile applications with or without prior notification of the user.

In addition, an implicit goal of this milestone is to have a decent understanding of a reasonable portion of the TaintDroid source code – enough to be able to make appropriate modifications and or grab desired data and ouput from TantDroid.

4.2 Milestone 2

The second milestone for this project will be to develop and implement a categorization model from our findings about various mobile malware applications. We plan to conduct extensive research on mobile malware and the pieces of data they act upon. For example, given a piece of data, say geographic location or IMEI code of a phone, we want to be able to make decent predictions about possible compromise and exposed vulnerabilities based on the piece of data. To achieve this, we will be reading about different malware applications and their functionality.

4.3 Milestone 3

The final milestone for this project will be to successfully implement a mapping scheme that maps taint values from TaintDroid for a specific application to potential harm the application can cause to the host device using the categorization model from Milestone 2. This will complete the vetting system with which we hope to be able to determine if a running application on the host Android phone poses possible malicious threats.

References

[1] Michael Becher, Felix C. Freiling, Johannes Hoffmann, Thorsten Holz, Sebastian Uellenbeck, and Christopher Wolf. Mobile security catching up? revealing the nuts and bolts of the security of mobile devices. In *Proceedings of the 2011 IEEE Symposium on Security and Privacy*, SP '11, pages 96–111, Washington, DC, USA, 2011. IEEE Computer Society.

[2] Abhijit Bose, Xin Hu, Kang G. Shin, and Taejoon Park. Behavioral detection of malware on mobile handsets. In *Proceedings of the 6th international conference on Mobile systems, applications, and services*, MobiSys '08, pages 225–238, New York, NY, USA, 2008. ACM.

[3] William Enck, Peter Gilbert, Byung-gon Chun, Landon P. Cox, Jaeyeon Jung, Patrick McDaniel, and Anmol N. Sheth. Taintdoid: An information-flow tracking system for realtime

privacy monitoring on smartphones. In *Proceedings of the USENIX Stmposium on Operating Systems Design and Implementation*, OSDI '10, 2010.

[4] Adrienne Porter Felt, Matthew Finifter, Erika Chin, Steve Hanna, and David Wagner. A survey of mobile malware in the wild. In *Proceedings of the 1st ACM workshop on Security and privacy in smartphones and mobile devices*, SPSM '11, pages 3–14, New York, NY, USA, 2011. ACM.

[5] Timothy Vidas, Daniel Votipka, and Nicolas Christin. All your droid are belong to us: a survey of current android attacks. In *Proceedings of the 5th USENIX conference on Offensive technologies*, WOOT'11, pages 10–10, Berkeley, CA, USA, 2011. USENIX Association.