WM0824 - Economics of Security

Assignment Block 2
Android App Repositories

Group 5

Ana Guerra

Eren Celik

Luigi Tuttobene

Raditya A

1. What security issue does the data speak to?

Our data set shows two different markets in China (360 and Baidu) with application from two categories games and software. Each category has subcategories and in each subcategory there are different apps. The same includes metadata such as the size, package information, number of downloads, and the day when they were last updated.

Currently, the data set do not speak about security issues. It is needed to link the data set with malware information, or to evaluate each application to some malware/virus detector to analyze if the apps can be malicious.

2. What would be the ideal metrics for security decision makers?

We believe one of the objectives of security decision makers of apps marketplace is to realize a secure marketplace that is free from security harm. Therefore, an ideal security metric, at least, should be able to support this effort. However, it appears unlikely that a single and straightforward metric have the capability to reflect the overall security level of a particular apps marketplace. Our group suggests that more than a single metric is necessary.

The "percentage of applications which identified as malicious" could be one of the metrics. The reason is quite straightforward. The percentage of malicious apps can directly reflect the effectiveness rate of a marketplace's filtering procedure. By using this metric, the security decision makers in each marketplace will be able to measure the effectiveness of their efforts.

Another metric that could be useful is the "percentage of the most popular applications which identified as malicious". The difference from the first metric is that this metric populates only the most popular apps instead of populating the entire applications in the marketplace. One source from the Internet has shown that the top one percent of app publishers for a certain marketplace accounted for 70 percent of all downloads. Since the most popular apps constitute a much larger portion of the user's download, focusing cleansing efforts on these popular apps means that the security decision makers have a good chance of securing 70 percent of their users' downloads.

There are obviously more metrics that can be used to support the effort of securing a marketplace even further, such as survival period of malicious apps, malicious apps download ratio, and so on. The main point is that currently there is no metric that singlehandedly can measure the security level of any apps marketplace. However, combining several metrics together, a set of ideal metrics that complement each other might be achieved.

3. What are the metrics that exist in practice?

There are currently several metrics that exist in the market (Kikuchi et al., 2016):

- Infection rate of devices
- Number of app that are repackaged by third parties, sometimes for malicious purposes
- Relation between repacking of apps and malware
- Malware detection (using software such as VirusTotal and Androguard)
- Malware presence ratio (% of all collected apps that are detected as malware)

- Malware download ratio (% of all download of the collected apps belonging to apps that are detected as malware)
- Survival period of malware (how long apps detected as malicious remains in the app store)
- Infection rate (fraction of all apps in the market that are malware)

4. A definition of the metrics you can design from the dataset

The most obvious metric that we can produce from these datasets is the "percentage of the most popular applications which identified as malicious." Since the listed apps in the datasets are already a population of the most popular apps from their respective categories, the calculation of the metric is fairly straightforward.

With the information of our data base, we can evaluate each application using VirusTotal and/or Androguard to detect how many applications are detected as infected or malicious apps. Then we can compare this information between the two markets to see which market is more likely to have more malicious apps and in which category and subcategory. In addition, we can relate the number of infected app with the number of downloads to evaluate which apps are more likely to be targeted of malware.

5. An evaluation of the the metrics you have defined. This should include graphical representations of the metrics (e.g., histograms, scatter plots, time series, bar charts).

(*To be updated*)

Resources

Kikuchi, Y., Mori, H., Nakano, H., Yoshioka, K., Matsumoto, T., & Van Eeten, M. (2016). Evaluating Malware Mitigation by Android Market Operators. Proceedings of the 9th USENIX Conference on Cyber Security Experimentation and Test, 8.

Nelson, R. (2016, May 10). 94% of U.S. App Store Revenue Comes From the Top 1% of Monetizing Publishers. Retrieved September 25, 2016, from https://sensortower.com/blog/app-store-one-percent

Zhou, Y., & Jiang, X. (2012). Detecting Passive Content Leaks and Pollution in Android Applications. NDSS Symposium 2013, (October), 16. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.687.1281&rep=rep1&type=pdf