Android - Actors and Security Strategies

Alexandru Babeanu 4881133
Antoine d'Estalenx 5149371
Arjen Brussen 4974484
Irène Förstel 5149398

October 21, 2019

1 Introduction

In the previous reports, we described metrics about the security of mobile applications, by comparing the number of malwares and downloaded malware applications in the Play Store and in other stores (for example, Chinese stores).

We continue with three actors for this report: the Play Store itself, the customer who install applications on its phone, and the antivirus companies. We detail the role of each actor in the security issue, how it affects them and what strategies they can employ to solve it. We also explore their incentives for employing such strategies.

In later sections, we chose one of the metrics described in our first report, namely the proportion of malware withing an app-store, and explore its variation across multiple stores. We identify various factors that could explain this variance, then perform a statistical analysis to measure the impact some these factors have on our metric.

2 Strategies for three actors

2.1 The Play Store

The Play Store can, as described in the previous report, develop an anti-virus to better detect the malwares among the uploaded applications on the store (e.g. by improving the Google Play Protect tool, that since 2017 tries to detect and remove malwares on the platform using machine learning).

The main costs of such a tool would be the developing costs: Google needs to hire a lot of Computer Engineers/Scientists to develop a tool (i.e. a tool with a very low false positive rate, since the Play Store does not want to remove applications that are not a malware). For scale, an antivirus company such as McAfee works with more than 6,000 people and a smaller one like Avast with 1,700 employees worldwide, which means that Google would require a very

high initial investment, because of employment costs and the rent of offices and computer material.

The main benefit for the store would be the avoidance of reputation loss: when a customer downloads malware on the Play Store, the customer could blame the Play Store for letting a malware inside in the first place and leave. Even if a customer spends very little on the Play Store, the amount of downloads (even for the malware applications) makes it interesting for the store to mitigate the issue.

This control tool can be an externality for the users of the store: even if they don't pay for it, they will have a better security while downloading applications - which leads to a better security hygiene of the Android network in general.

2.2 The customer

The customers cannot directly act on the applications in the Play Store, but they can take countermeasures to avoid security problems. For example, by installing an antivirus on their phone. The antivirus will detect some (maybe not all) the malicious applications and avoid potential data leaks.

The cost of such a solution is the cost of the antivirus. However, plenty of free antivirus tools exist. Since it's difficult to assess the quality of an antivirus (because of the lack of available data and the difficulty to measure security), people will probably choose a free antivirus funded by an alternative such as ads. So the only cost remaining is the time it takes for a customer to choose an antivirus and install it.

The benefits of such an installation are difficult to measure. Often, the antivirus will not detect anything (since the customer is more likely to download non-malicious applications rather than infected applications). Most antivirus software use a great deal of processing power, negatively affecting battery life and performance of a device. When it does prevent the installation of a malicious application, the prevented loss could be anything. Loss of personal information, Loss of device performance, loss of credit or money. It could also be adware that hinders the usage of the device by continuous advertisements.

The uncertainty of how well an antivirus performs (i.e. its only benefit), especially with free solutions which - ironically - are often regarded as insecure [1], results in strong uncertainty of the benefit of such a solution. Customers, therefore, do not have many incentives to install antivirus applications on their phone, unless required to do so by, for example, their employer. Installing an antivirus is not a priority, because most customers have never experienced (the effects of) a serious data breach.

The hacked phone could be used as an bad externality, though. For example, an email hacked by a malicious application can send spam to other contacts of this email-address. However, since this does not directly concern the owner of the phone (which is not the one receiving spam), he could ignore that problem.

[1] https://www.comparitech.com/antivirus/android-antivirus-vulnerabilities/

2.3 The antivirus companies

The countermeasure for the issue designed by antivirus companies is obvious. More R&D on developing better algorithms for detecting malicious code is an effective long-term countermeasure.

The costs are directly tied to the product of these companies. Employee salaries, office rent and hardware are the most common costs. However, if we assume the company to have been operational for some time already, such as McAfee, then the costs are negligible as it requires no initial investment, but rather a continuous R&D cost.

The benefits are, once the product is released, initially very significant. However, over time the relative increase of quality in updates to the antivirus software decreases. Most malware will be detected by the antivirus and, therefore, the security issue reduces in risk.

Obviously, the antivirus companies have a high priority in developing this countermeasure, as it is their sole source of income.

The externality of this is that malware developers are less inclined to create 'lazy malware'-applications. With antiviruses, these developers must invest time and money in their applications in order to fool the antivirus software. Therefore, it would also result in a lower proportion of malware vs. non-malware applications, as less malware applications would be built.

3 Differences in the metrics

The main difference observed in the metrics is on the proportion of malicious applications in application stores which is represented in figure 1. This graph shows that all the stores don't have the same security controls. Furthermore, comparing the American store (the Google Play Store) and Chinese stores show that Chinese stores have far more malicious applications on it.

3.1 Factors

We have identified various factors that can influence our metric:

1. **The laws and regulations specific to the regions where the app-stores operate.** For example, some US states have regulations about data breach disclosures that force the companies suffering from a data breach to inform the customers and the authorities as soon as possible. Because this will result in loss of reputation, it is an incentive for the app markets to install better security controls.

2. **The openness of the application market.** Malicious developers are more likely to target app-stores with less security controls, lower publishing standards, and lower publishing prices.

3. **The size of the user-base of each app-store.** Our metric accounts for the size of the platform by normalizing the amount of malware based

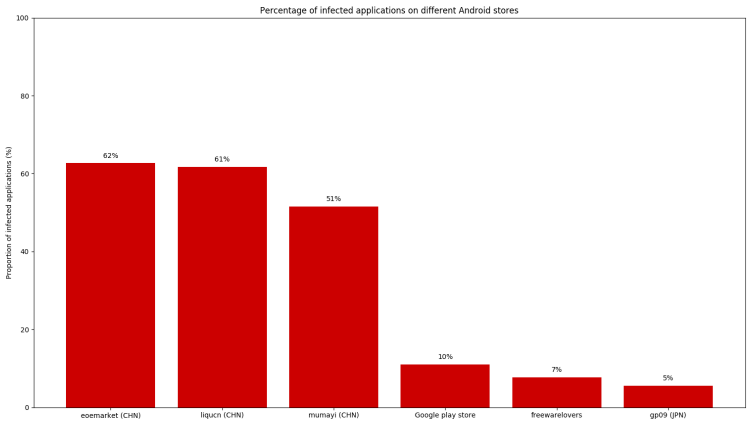

Figure 1: Percentage of infected applications on each store

on the total amount of applications, but the size of the platform by itself constitutes an incentive to the attackers. It is very likely that malicious developers will prioritize platforms with larger user-bases, since this translates to more potential victims.

4. **The types of applications available on each market.** If attackers have a preference for certain application categories, then stores that specialize in these categories will have higher malware ratios than stores that don't.

We will mainly focus on the last factor in the statistical analysis, first by looking at the proportion of games in each market and secondly by looking at the correlation between the popularity of categories in the market and the proportion of malware in these categories. Finally, we will discuss the possible effect of laws and regulations on the variance.

3.2 Statistical analysis

3.2.1 Proportion of games in a market

The datasets contain malware analyses of applications downloaded from six different application stores. Since the applications were also labelled with their categories, it is possible to look at popular categories, considering which have the most applications. In our first report [2] about security metrics, we had found that the "game" category was popular in application stores and also had a significant proportion of malware in it.

The idea is, therefore, to check if a higher proportion of games in a market could lead to a higher proportion of malware. One can note that Chinese mar-

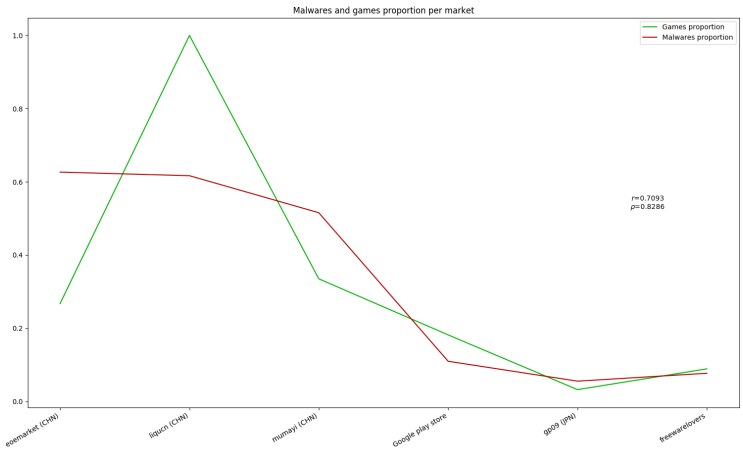

Figure 2: Correlation between games and malware proportions in markets

kets, that have a lot of malware also offer a lot of games relative to their size (note that Chinese market *liqucn* is only dedicated to games), while more secure markets, such as the Google Play store, have fewer games. This could be because games are really attractive for users and have more downloads than other applications, which makes them an attractive category for malware developers.

Figure 2 shows that there seem to be a small correlation (note the values of Pearson and Spearman correlation coefficients $r = 0.71$ and $\rho = 0.83$) between games and malware proportions in markets. A correlation higher than 0.8 is considered significant and indicates that there is a positive relation between the two variables, i.e. when one increases, so does the other. This does not necessarily imply that a higher ratio of games will cause a higher ratio of malware. While the reverse seems implausible, i.e. more malware causing more games, it is likely that the two factors share a common cause. For example, low standards for publishing, lax controls and low costs for publishing applications are just a few examples of characteristics that appeal to both malicious and non-malicious game developers, and could explain the increase in both variables.

A limitation in our analysis is the low amount of ata available. When considering the proportions at market-level, we end up with one data-point per market, which results in very few data-points. Obviously, this highly impacts the significance of our results.

3.2.2 Correlation category malware ratio

To confirm the assumption that the more popular a category is, the more malware it contains, we look at the popularity and malware for every category in a market.

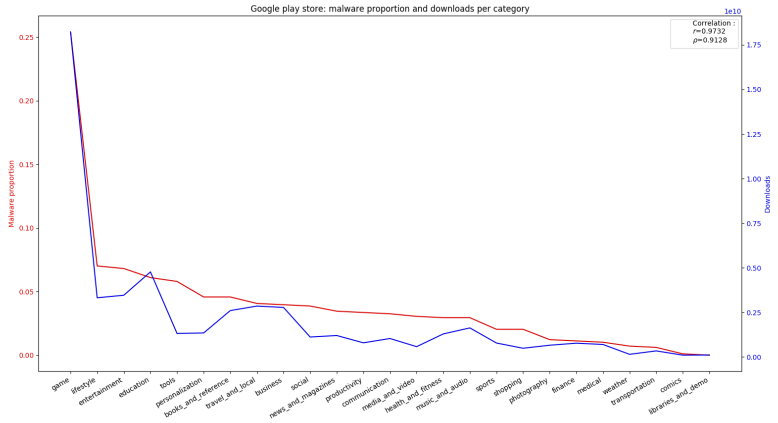

Figure 3: Correlation between downloads and malware proportion in Google Play store categories

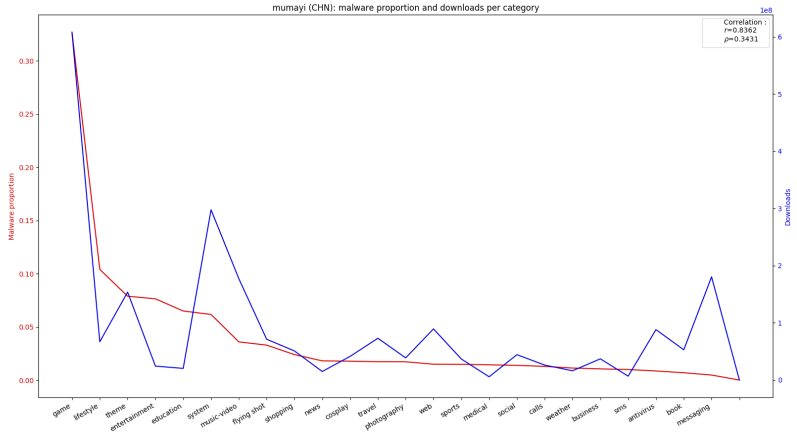

Figure 4: Correlation between downloads and malware proportion in mumayi categories

Figures 3 and 4 present the correlation between the number of downloads and the proportion of malware in each category of Google Play and Mumayi application stores. One can first notice that the game category is very popular in terms of downloads and also has a high malware proportion, as explained previously. But it can be noted that the total amount of downloads and the

malware proportion in each category seem to be correlated. This confirms the idea that malware developers may upload their malware in popular categories hoping that more users will download infected applications. Such a result could be helpful for application stores since they could refine their looking for infected application by focusing more on popular categories. Such a result echoes another article [3] in which researchers had also noted that some queries for popular keywords had very high malware rates in their results.

However, figure 4 (*mumayi* store) combined with the analysis on other stores shows that the evolution of both variables is not perfectly correlated. This suggests that there are multiple significant factors influencing the malware proportion in a market. In Figure 4, one can note that the *system* category has an important proportion of malware, even if it is not a very popular category. This might be because applications in this category require more permissions to fully customize the system and users often accept to give permissions to system applications thinking that these are required for customization. Since more permissions also means more attack vectors [1], attackers are likely to target this category more, even if it isn't so popular.

4 Regulatory variance

The data used in our research was gathered in 2016. In this year, no country had strong and fully-enveloping data protection laws in place. China, however, did not properly start researching cybersecurity law pertaining to other consumers than system and infrastructure until 2014 [4]. The release of its first official cybersecurity law was Jun 1 2017. Since the data was gathered earlier, it was still the wild west for most companies not related to infrastructure or critical systems.

This wild west time period for China resulted in a higher proportion of infected applications on each store, as can be seen in figure 1. It is, however, difficult to quantify how well a set of cybersecurity laws reduce the proportion of infected applications, thus changing the variance, without having data from after the release of the Chinese cybersecurity law.

Given that the new Chinese cybersecurity law is even more strict than the GDPR, we expect application stores in China to majorly overhaul their security policies as well in order to meet compliances. This should result in a lower proportion of malware comparable to other application stores in other countries.

5 Conclusion

We identified some factors that could influence the proportion of malware in a market, and more precisely inside a market: a statistical analysis shows that there is a correlation between the proportion of games in a market and the proportion of malware. This correlation could be caused by the popularity of game applications, which is appealing to attackers since they can reach a higher

amount of victims, but can also be caused by other factors related to the store itself. There also seems to be a correlation between the number of downloads in a single category and the proportion of malware in that same category: popular categories attract more malware.

Of course, other factors influence the distribution of malware across the markets and the distribution inside the categories as well: the cost of uploading an application (as developer), the permissions an application can reasonably obtain from their users, or the laws and regulations, depending on the country.

6 References

References

[1] Jagdish Prasad Achara, Mathieu Cunche, Vincent Roca, and Aurélien Francillon. "Short Paper: WifiLeaks: Underestimated Privacy Implications of the ACCESS_WIFI_STATE Android Permission". In: *Proceedings of the ACM Conference on Security and Privacy in Wireless and Mobile Networks (ACM WiSec)*. July 2014.

[2] Alexandru Babeanu, Antoine d'Estalenx, Irène Förstel, and Arjen Brussen. "Android - Security Metrics". In: (). URL: https://github.com/ArjenB96/EOCS_2019/blob/master/Block2/Android_Security_Metrics_FINAL.pdf.

[3] Yosuke Kikuchi, Hiroshi Mori, Hiroki Nakano, Katsunari yoshioka, Tsutomu Matsumoto, and Michel Van Eeten. "Evaluating Malware Mitigation by Android Market Operators". In: (). URL: https://www.usenix.org/system/files/conference/cset16/cset16-paper-kikuchi.pdf.

[4] KPMG. *Overview of Cybersecurity Law*. 2017. URL: https://assets.kpmg/content/dam/kpmg/cn/pdf/en/2017/02/overview-of-cybersecurity-law.pdf (visited on 10/21/2019).