Android - Security Investments and Management

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

October 7, 2019

1 Introduction

In this paper, we explore the security issue of malware publishing on application markets from the perspective of security strategies for mitigating the issue. We first identify the different actors that play a role in this security issue, including the problem owner. We define possible strategies for each actor, based on their motivations, and then select one actor and strategy for a more in-depth analysis. In this analysis, we seek to find out whether the selected strategy is viable investment for the actor. To do so, we aggregate the possible costs and benefits of this investment and calculate the return of security investment (**ROSI**).

2 Problem owner

In our previous report, we defined metrics to measure the threat of malwares on Android application stores. Malware are developed by malicious developers, who embed it in an application and subsequently upload it on the app stores. The users, who potentially download these applications, are the victims of malware.

Application stores depend on their users for their business. Some, like the Google Play Store, take a percentage of every transaction that happen on the store (like buying an application). But since a majority of the applications on these stores are free, these stores mostly earn money by advertising applications. For example, developers can pay to have their applications appear first in the user searches on to be highlighted on the front page.

As the Google Play Store is the official application store for Android and because the data is more detailed than the data about some other application stores, we will mostly focus on this store.

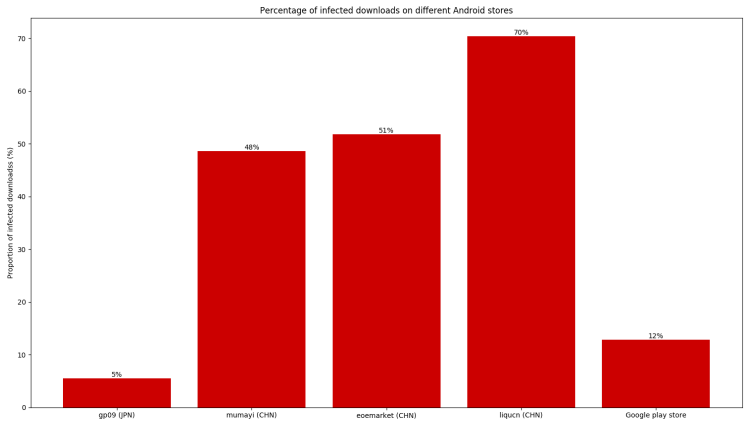

Figure 1: Proportions of malwares in application downloads

3 Security Performance

In the previous paper, we defined multiple metrics to measure the security performance of the different application stores. In figure 1, we looked at the proportion of downloaded infected applications to measure the risk of downloading a malware depending on the application store used.

In Figure 2, one can see that the markets which had the higher proportion of malware in downloads in figure 1 are also the ones which have the higher proportion of malware in the available applications. These measures are of course linked together (the more malware there is, the higher the chance users will download it).

These metrics reveal that the stores do not have the same controls quality, both in mitigating the risk and preventing it. Therefore, there are significant differences in security performance between application stores.

4 Risk strategies

The metrics presented in Figures 1 and 2 reveal that the stores do not have the same risk strategies against malware. Most choose risk acceptance as their strategy (i.e. they ignore the issue at hand). The Google Play store chooses to reduce the risk by analysing each uploaded or updated application for potential malware. They also promote very popular, useful and safe applications by pushing them to the top of search results ('Editor's Choice'). These measures do not reduce the security issue completely, but it is the most optimal solution out of ignoring or avoiding the issue (e.g. by removing categories which have a high proportion of malware), or even transferring the risk by being insured for

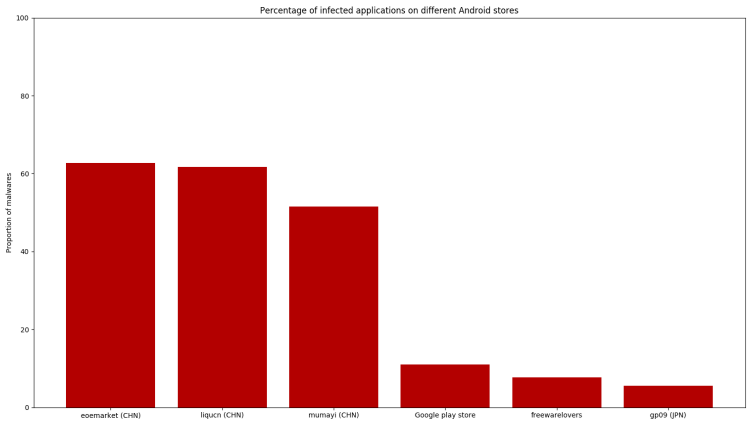

Figure 2: Proportions of malwares in application stores

potential lawsuits.

As shown in Figure 3, the proportions of malware are not equally distributed in the store categories. This metric can also help to refine the controls to analyse the applications that are most likely to be or contain malware.

The Google Play Store can mitigate the risk of having malware applications in the store by :

- Having stronger security checks before the application can be uploaded on the Play Store (especially in the most targeted categories, such as *Media and Video* or *Games*).
- Reducing the impact of an uploaded malware application by not promoting it and removing it shortly after its upload.

5 Actors

Malware developers are the main actors aside from the application store itself, because they are the threat source.

Application store markets are important actors. They have assets such as their reputation, user base and services (application storage) to protect. If an application store has a low reputation due to a relatively high amount of infected applications, their users will move to a store with a better reputation. Therefore, not only does it matter how severe the issue is in a single store, but also how the stores compare against one another in terms of security.

Users are also actors of this security issue. They are the primary targets of the attackers and they dominantly influence the consequences of an untreated

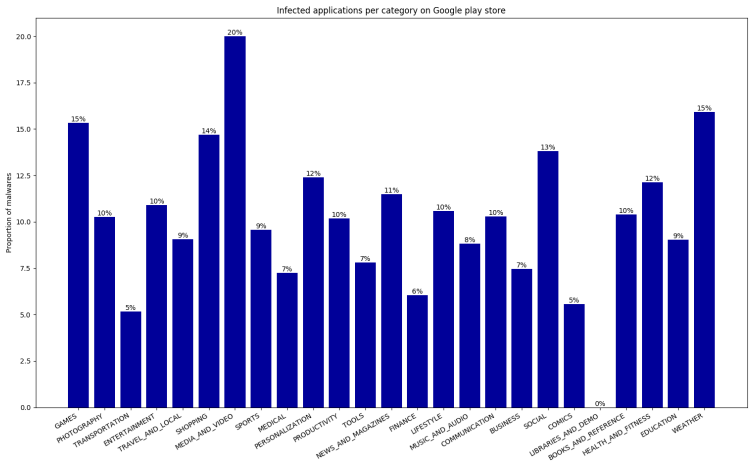

Figure 3: Proportion of infected applications per category on Google Play store

security issue. If the users are faced with malware-infected applications from an application store, they may be discouraged from visiting that particular store again. The loss of revenue and popularity, as discussed in the first assignment, weighs heavily on the severity of the security issue.

Antivirus developers are actors since they work on the malware detection. They influence the effectiveness of mitigating this risk, whether it is by improving their detection rates or by providing cheaper controls.

Regulators heavily influence this security issue, as the application stores must follow regulations or face serious fines and other penalties. However, this actor acts independently and works on already existing security issues. It therefore cannot directly influence (upcoming) security issues, as opposed to the stores themselves.

6 Choosing a strategy

The different actors here can have different strategies, because they all have different goals and assets to protect:

- Malware developers want to maintain or increase the number of infections.
- The stores want to protect their reputation and critical assets such as the availability of the network.
- A user wants to protect the integrity and availability of their phone and the confidentiality of their data.
- Antivirus developers want to increase their market share and maintain their (product) reputation.

- Regulators want to protect the market and users, without restricting it to a high degree. They also want to maintain the trust the users have in them.

For malware developers, in order for them to remain relevant in their field, their only option is to tackle the risk and find a solution. If they were to accept/ignore the risk, the (perceived) quality would degrade over time and their reputation would suffer. If they were to avoid the risk, they would not be able to continue their work, as the risk is directly affected by the criticality of their asset.

Application stores suffer from the same choices, albeit slightly less threatening. They could avoid the issue by limiting the amount of categories they have applications for. Avoiding the issue results in less costs, but would also result in less traffic. They could also implement the product of anti-virus developers (or design their own given enough budget) into their own system to analyse applications. This would have a higher cost associated to it than by avoiding the issue, but it would result in a higher quality of service, thus increasing popularity and market share.

One possible strategy for the users could be risk avoidance: if one store (or one category of a specific store) contains too much malware, the user could decide to boycott this store or this specific category by simply using another store. This strategy has the added benefit that the costs of this strategy for the user is near zero, yet it would have a relatively considerable impact on the store itself if multiple users did this.

Another strategy for the users could be to mitigate the risk by using an anti-virus application to detect the malware application after its installation. The success of such a strategy is however strongly correlated with the efficiency of the anti-virus. And we see in the dataset that anti-viruses don't always agree on which application is a malware. Users can also be careful about the permissions they give to each application, if they are well enough informed to detect incoherent permissions (for example, an image editing application which asks for the permission to access the user's contacts).

7 ROSI calculation

7.1 Benefits

7.1.1 Main benefits for Google Play

We have the price of each application defined in the dataset of the Play Store. We assume that there is no other transaction (no 'in app purchases') that the Play Store could charge and get benefit on. So each time a user buys an application, the Play Store gets 30% of the price [6]. Google Play also takes 15% of the price of in-app purchases. Therefore, we could compute the benefits on the Google Play store by multiplying the number of downloads for an application by 30% of this application price. Note that the dataset does not contain all the

available applications, but that we will assume that the distribution for prices and malware are the same in the dataset and in the store. This method could give interesting results, but all the dataset applications are free. Therefore, the dataset will not give any useful information about this type of benefits.

For each different developer, there is also a $25 fee. Because a developer can upload many apps, we need to compute the number of different developers in the dataset. But this is a one time fee, so this won't be present in the costs.

Another estimated benefit is prevented impact/loss. We established that users/traffic are the main source of income. If one of the devices of a user has been infected by malware by the Google Play Store, the user might be hesitant to use that store again. Therefore, the prevented loss equals the value of a user.

We compute this estimated loss on the dataset and assume that the ROSI based on the information from the dataset is the same as the one on the whole Google Play Store.

7.1.2 Probability of detecting new malware

In order to compute the benefits from this stronger controls, we need to evaluate the probability of detecting a malware in the Play Store.

A solution to detect malware more efficiently on Google Play could be to develop an antivirus software to analyse applications. Since Google only gives few information about Google Play Protect, we do not know how efficient it is. To estimate the possible efficiency of a new antivirus software, we will analyse the performances of a well-known antivirus which was used to analyse applications in the dataset: *McAfee*.

In the example of the McAfee antivirus, we compute the probability of detection of a malware as follow : if an application is detected as a malware by 10% or more of the antivirus programs in the dataset, then we consider it a virus. The overall probability of detecting a virus is given by the sum of malicious applications detected by McAfee normalized by the total amount of malware detected by 10% or more antivirus programs.

We obtain a probability $p_i = 70\%$.

7.1.3 Estimated loss per paying user

We make the following assumption: A user who downloaded malware will never buy something in the Play Store again. This quantifies the loss of reputation from the Play Store due to malware. Since our dataset does not include enough information to quantify impact this way, we use external information to estimate the loss per user:

On average, according to SensorTower [4], one US user spends 30$ (per device) each year in the Play Store. However, the report *The State of In-App Spending* by AppsFlyer [1] indicates that this value is not the same all over the world. However, the values in this report indicate that Americans are close to the global average.

For the sake of simplicity, let us assume that each user only has one device. Since Google will take between 15% and 30% of these $30 depending on the type of purchases (applications or in-app purchases), we will consider the lower bound and take 15% of this amount. This means a loss of $l = 30\$ \cdot 15\% = 4.50\$$ for the store per downloaded malware due to the user leaving.

7.1.4 Total of the estimated benefits

The number of malware downloads can be computed from the dataset. In the dataset, there are $m = 984$ malicious apps, which have an average of $d_m = 6,867,081$ downloads. However, according to the report by AppsFlyer [1], only 5% of users are paying users. Assuming that paying and non-paying users are equally affected by malware (which might not be the case since malware is almost always free), we can assume that 5% of these downloads affected paying users and caused a loss for Google. The total benefits per year of a solution that eliminates p_i=70% of the malware can be given as:

$$\begin{aligned} \text{benefits} &= l \cdot 0.05 \cdot d_m \cdot m \cdot p_i \\ &= \$4.50 \cdot 0.05 \cdot 6,867,081 \cdot 984 \cdot 0.7 \\ &= \$1,064,260,213 \end{aligned}$$

7.2 Costs

If Google chooses to implement stronger controls on uploaded applications, it may have to face these additional costs:

- the cost of developing controls (paying developers to code a tool that has a better detection of malware applications);
- the losses of the malware applications that have not been uploaded;
- the losses due to unforeseen bugs: false positives. When a non-malicious application is refused by the store, the store loses the 15% fee on every hypothetical transaction (including the price of the application).

7.2.1 Cost of developing an antivirus tool

For computing this, we can find out what the budget for the anti-virus is in an anti-virus company, assuming that the cost of developing an anti-virus tool would be the same if Google decided to develop its own antivirus software.

We choose the example of the McAfee company, which in 2019 had 7000 employees according to the fact sheet they published in August 2019 [5]. The website ComputerCareers [2] gives an approximation of the average wage for a computer engineer: $102,450 per year.

This gives us a cost of development equal to **$693,381,600 per year** for the employees salaries, which is a very high estimation seeing as Google does not

need the employees from the HR/Finance/Support division of McAfee; sadly we did not have more precise values to provide a better estimation.

In addition, we need to consider the cost of the buildings for the 7,000 employees working on stronger controls: according to Market Watch [3], the annual office price per employee in San Francisco is 13,032$. This incurs an additional cost of **$91,224,000**.

7.2.2 Cost of a malware that has not been uploaded

The dataset gives us a list of malware applications (we consider as malware every application that is detected as a malware by at least one anti virus) and their cost on the Google Play Store.

On the dataset, all the applications are free. It would be logical for a malware to be free, since it will spread more in average. So we can consider that this loss is equal to zero or negligible compared to the cost of development.

7.2.3 Cost of the false positives

In order to compute the false positives and false negatives of this new anti virus solution for the PlayStore, we can again look at an existing company and compute the False Positive Rate (FPR) and True Positive Rate (TPR).

False positive ratio with McAfee is 0.07% on the dataset which means that McAfee tagged as malicious 0.07% of applications that were considered safe by other antiviruses. In the dataset, there were 7984 safe applications, thus we get an amount of:

$$\begin{aligned} \#FP &= \#apps_{safe} * FP_{rate} \\ &= 7984 * 0.0007 \\ &= 5.59 \end{aligned}$$

About five to six maliciously-marked application is actually benign. According to [1], users spend on average $0.50 per app. Google keeps the initial investment the developer had to pay to put his app on the application store, so if we assume there to be 6 maliciously-marked applications that are actually benign, there is a cost of $3 per month, which is $36 a year. This value is insignificant compared to the development costs.

7.2.4 Total cost of the solution

The total cost per year is $693,381,600 + $91,224,000 + $36 = $784,605,636.

7.3 ROSI value

We find a final ROSI of :

$$\text{ROSI} = \frac{1,064,260,213 - 784,605,636}{784,605,636} = 0.3564 = 35.6\%$$

The ROSI is 35% per year, and therefore, the investment could be considered attractive. However, this value mostly depends on the assumptions we made in the costs and benefits estimation and might be over-estimated. The assumptions we made were heavily favored towards a positive ROSI. For instance, the assumption that a single downloaded application infected by malware on a single device would result in a user indefinitely boycotting Google Play is very weak. The user would maybe stop using its infected phone, but could buy another one and come back to Google Play after some time.

ROSI does not take the time factor into account. We assume that our anti-virus software is instantly developed and ready for use, yet the development process for a proper anti-virus software might take longer than a year. The invested costs before the product is finished could also have been invested in something else with a positive return rate.

The time factor is also hard to take into account because developing an antivirus is not a one time investment. Once it is developed, an antivirus software also requires maintenance, such as keeping the malware database up-to-date.

However, the cost of development has probably been overestimated : Google doesn't need as much employees as an antivirus company such as McAfee since it's not developing a software to sell but an internal to issues to the PlayStore. The development of controls to avoid malicious applications still remain very interesting for the PlayStore, because they are developing a such tool since 2017, called Play Protect.

8 Conclusion

In this paper we explored various strategies that can be used for mitigating the security issue of malware distribution in app-stores. We identified various actors that play a role in this security issue, such as the app-stores themselves, their clients, malicious developers, and regulators, and we offered a few examples of security strategies based on the motivations of each actor.

The main part of our exploration lies in the evaluation of the returned value of a security investment. For this, we chose the Google Play Store as our main actor, and the development of an antivirus software as the security investment. We evaluated the various costs and benefits related to this investment, making some assumptions where data was insufficient, and obtained a **ROSI** value of 35%. It is worth noting that this value is somewhat optimistic, because of some of our assumptions (e.g. the assumption that each malware download will result in a lost user, or the antivirus will be developed in a year or less, or ignoring the maintenance costs for the antivirus).

9 References

References

[1] AppsFlyer. *The State of In-App Spending: Global & Regional Benchmarks*. 2016. URL: https://www.appsflyer.com/resources/state-of-in-app-spending-global-regional-benchmarks/.

[2] *Computer Engineer Salary*. URL: https://www.computercareers.org/computer-engineer-salary/.

[3] Sally French. *Here's how much your company pays to rent office space*. May 2015. URL: https://www.marketwatch.com/story/heres-how-much-your-company-pays-to-rent-office-space-2015-05-27.

[4] Ruika Lin. *U.S. Android Users Spent an Average of $30 on Google Play Apps in 2016*. June 2017. URL: https://sensortower.com/blog/revenue-per-android-device-2016.

[5] McAfee. *Corporate fact sheet; PDF*. Aug. 2019. URL: https://www.mcafee.com/enterprise/en-us/assets/fact-sheets/fs-mcafee-fact-sheet.pdf.

[6] *Transaction fees for merchants*. URL: https://support.google.com/paymentscenter/answer/7159343.