Evaluating Malware Mitigation by Android Market Operators

Yosuke KIKUCHI*, **Hiroshi MORI***, Hiroki NAKANO*,
Katsunari YOSHIOKA*, Tsutomu MATSUMOTO*, Michel VAN EETEN**

*Yokohama National University
**Delft University of Technology

9th USENIX Workshop on Cyber Security Experimentation and Test

YNU YOKOHAMA National University

TU Delft
Delft University of Technology

Motivation

- The number of Android apps is continuously increasing
- There are many third party app markets(e.g. Amazon App Store)
- At the same time, the number of Android malware are increasing[1]
- In this situation, we are interested in whether these market operators make effort for mitigating malware and if yes, how well they do



Agenda

- Motivation
- Agenda
- Method
- Result of Evaluation
- Conclusion
- Limitations

Method



1. Download apps and meta-data
(# of downloads and
date of last update)



What markets we have selected?

- We measured five markets
- Why did we select these markets?
  - We wanted to track survival period of apps, so we selected popular (at least once) and not too new markets

Market	Collected Apps	Collection Period
Google Play (GP_PO)	12280	2015/09/21 – 2015/09/28
Goolgle Play (GP_PD)	8966	2014/08/07 – 2014/11/03
Mumayi (MY)	19794	2015/07/14 – 2015/08/09
Liqucn (LQ)	6844	2015/09/02 – 2015/09/04
eoeMarket (EM)	10203	2015/09/15 – 2015/09/19
Freeware Lovers (FL)	3652	2015/06/20 – 2015/06/23

Data Collection
- Google Play -

- We collected apps from Google Play in **two ways**

1. Top popular apps for each categories(GP_PO)
   - e.g. games, tools and etc...
   - We could collect **12,280 apps**
   - **Note**: There is a bias, because these apps are popular and could have been more closely examined by market operator

2. We also use PlayDrone DataSet[3] (over 1.4M apps) that were collected for research purpose. And we used **8,966 apps** finally(GP_PD)
   - We selected 15,000 free apps from the dataset at random
   - Since Play Drone was collected in 2014, we then checked whether these apps still exist in the Play store at the time of the study
   - We picked up 8,966 apps that are still in Google Play
   - **Note**: Collection period of this dataset is older than our other app dataset

[3]JAKEJ. Playdrone apk's. https://archive.org/details/ playdrone- apks/

Data Colection
- Other Markets -

- We could collect apps by crawling web pages of apps in other markets
  - Mumayi(EM): **19,794 apps**
  - eoeMarket(EM): **10,203 apps**
  - Liqucn(LQ): **6,884 apps**
  - Freeware Lovers(FL): **3,652 apps**

Method



1. Download apps and meta-data
(# of downloads and
date of last update)



Method



2. Submit the apps to VirusTotal and log results
- Detection Ratio
- Labels

virustotal

1. Download apps and meta-data (# of downloads and date of last update)



Method

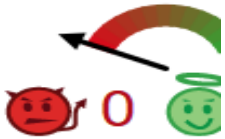
virustotal

SHA256: f49410d96e93822e44dfa1a45b55e250abca233a43f4afb0237a1a0f76b62e67

Detection Ratio: How many AV vendors flag the app as malicious?

Detection ratio: 43 / 55

Analysis date: 2016-07-25 07:53:40 UTC (1 week ago)



Analysis | File detail | Relationships | Additional information | Comments 0 | Votes

Antivirus	Result	Update
ALYac	Trojan.Android.DroidDream.D	20160725
AVG	Android	20160725
AVware	Exploit.AndroidOS.Lotoor	20160725
Ad-Aware	Android.Exploit.Exploid.B	20160725

Label: What kinds of malware(Note: sometimes adware is included)

Method



2. Submit the apps to VirusTotal and log results
- Detection Ratio
- Labels

virustotal

1. Download apps and meta-data (# of downloads and date of last update)



3. Evaluate with three metrics

Method
- Metrics for Evaluation -

- We calculate three security metrics for each market for evaluation:
  - **the malware presence ratio (MPR)**
    - percentage of all collected apps that are detected as malware
  - **the malware download ratio (MDR)**
    - percentage of all downloads of all collected apps that are detected as malware
  - **the survival period of malware**
    - how long apps detected as malware remain in the app store

Positive Detection Ratio

- We defined Positive Detection Ratio (PDR)
  - Ratio of number of AV products that flag the app “malicious” (=Detection Ratio)

$$PDR(S,a) = |\{av \in S | av(a) = malicious\}| / |S|$$

a : Android App
av: Anti Virus Product
S : Set of AV products with which VirusTotal used to scan the app

Positive Detection Ratio

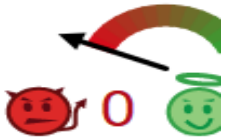
virustotal

SHA256: f49410d96e93822e44dfa1a45b55e250abca233a43f4afb0237a1a0f76b62e67

Detection Ratio: How many AV vendors flag the app as malicious?

Detection ratio: 43 / 55

Analysis date: 2016-07-25 07:53:40 UTC (1 week ago)



PDR = 78[%]

Analysis | File detail | Relationships | Additional information | Comments 0 | Votes

Antivirus	Result	Update
ALYac	Trojan.Android.DroidDream.D	20160725
AVG	Android	20160725
AVware	Exploit.AndroidOS.Lotoor	20160725
Ad-Aware	Android.Exploit.Exploid.B	20160725

S : Set o

Label: What kinds of malware (Note: sometimes adware is included)

Malware Presence Ratio

- We define Malware Presence Ratio (MPR)
  - This means percentage of all collected apps that are detected as malware
- How to judge malware or not?
  - We judge an app as malware if PDR of the app above m[%]

$$MPR(m,S,M) = |\{a \in M | PDR(S,a) > m\}| / |M|$$

a : Android App

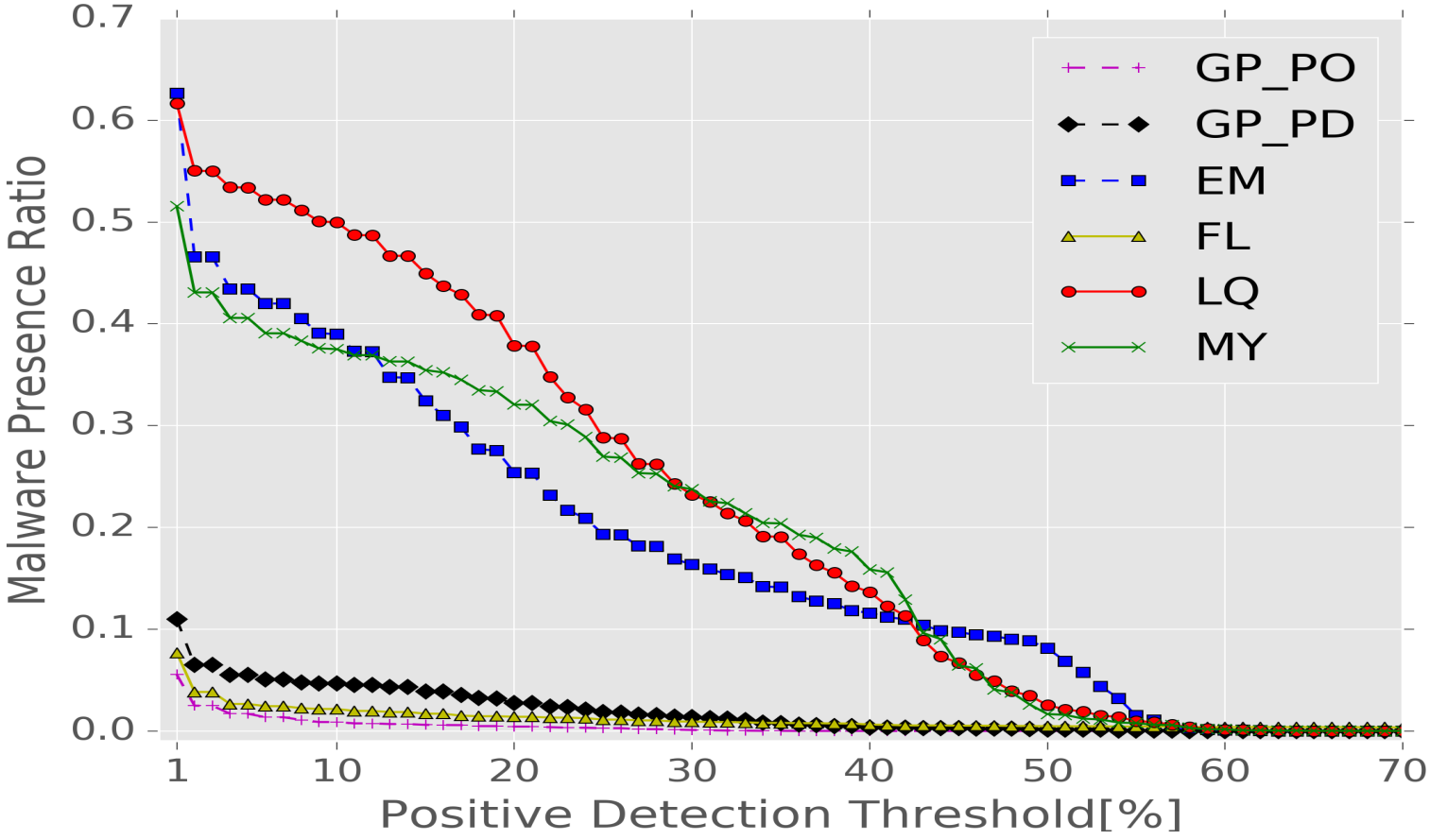
av: Anti Virus Product

S : Set of AV products with which VirusTotal used to scan the app

M: Set of Apps in a market

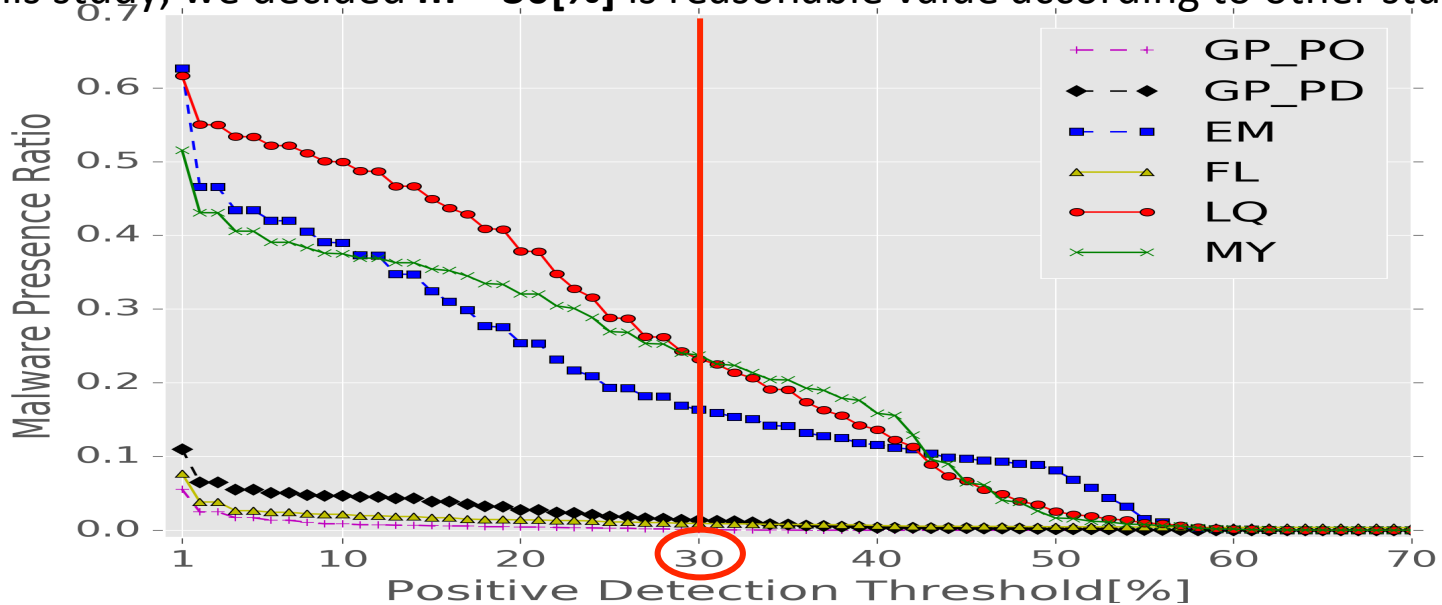
m: Threshold

Malware Presence Ratio



Malware Presence Ratio
- What is a reasonable value of “m”? -

MPR depend on “m” (Positive Detection Threshold), so this value is important.
If we set m to too high value (e.g. m= 70[%]), almost all apps are counted as not malware.
In this study, we decided **m = 30[%]** is reasonable value according to other studies[4][5]



[4] F-SECURE. Threat report h2 2013. https://www.f-secure. com/documents/996508/1030743/Threat_Report_H2_ 2013.pdf.
[5] ZHOU, Y., WANG, Z., ZHOU, W., AND JIANG, X. Hey, you, get off of my market: Detecting malicious apps in official and alternative android markets. In NDSS (2012).

Google Play is Really Safe?
- MPR for Google Play via Keyword Search -

- The MPR for GP_PO is low, but we have explored whether this risk was perhaps concentrated, and thus higher, in certain areas of the market.

- For this purpose, we collected another sample of apps, via keyword searches.

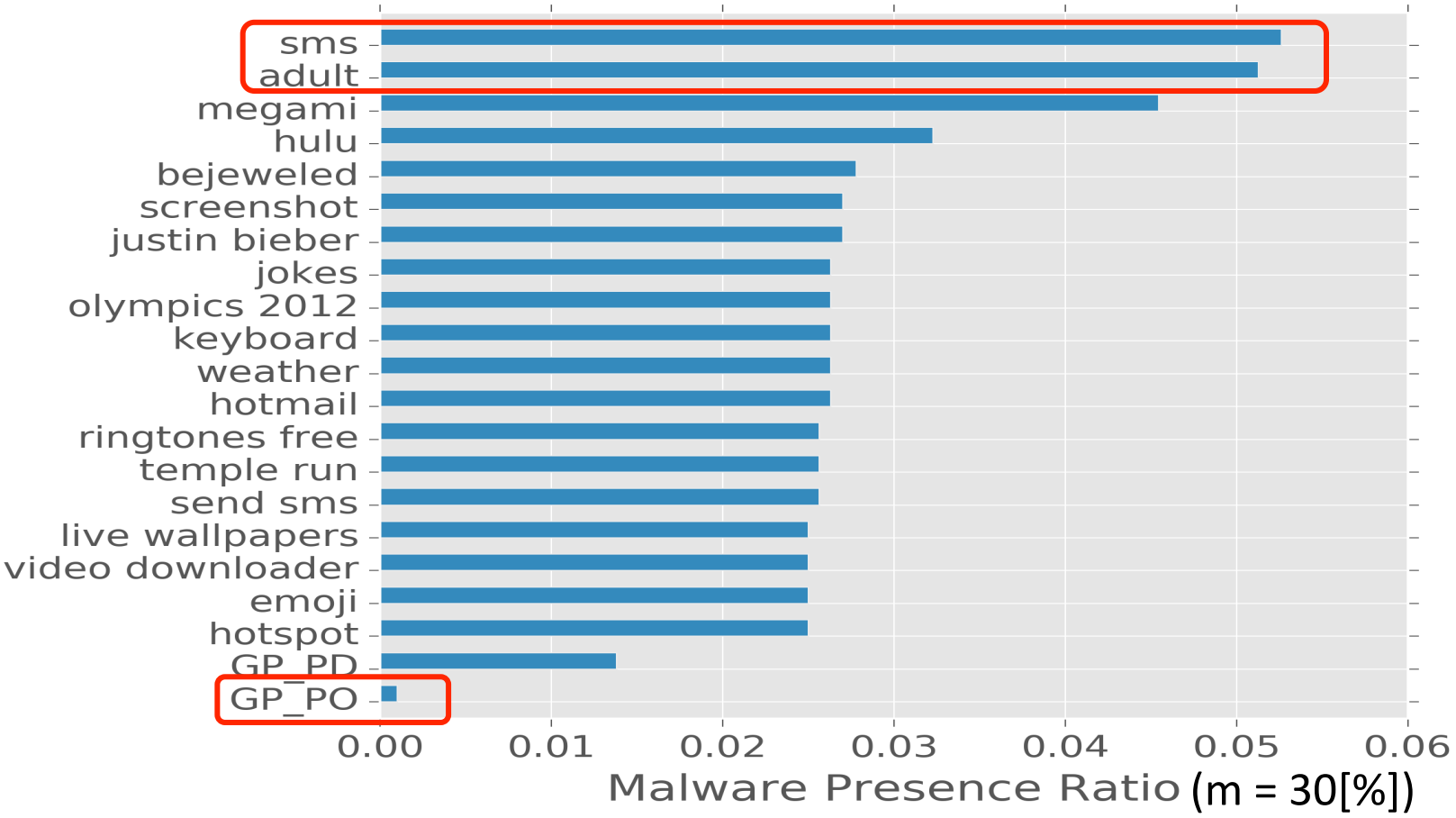
Google Play is Really Safe?
- MPR for Google Play via Keyword Search -

- We used 155 popular keywords listed on the keyword ranking sites [6-8], and 11 keywords in the security report by Symantec [9]

- We then collected the 40 apps that are shown on the first page of search results

[6] The most popular app store keywords from chomp & google play. http://searchengineland.com/the-most-popular-app-store-keywords-from-chomp-google-play-135744
[7] SIMILARWEBPRO. Google play top 20 popular keywords in apr 2015. http://www.similar- web.jp/blog/archives/1004
[8] App annie app store stats google play top app charts united states overall. https://www.appannie.com/apps/google-play/top/
[9] SYMANTEC. Malicious websites by search term. http://securityresponse.symantec.com/threatreport/topic.jsp?id=threat_activity_trends&aid= malicious_websites_by_search_term.

MPR for Google Play via Keyword Search



Google Play is Really Safe?
- MPR for Google Play via Keyword Search -

- We used 155 popular keywords listed on the keyword ranking sites [6-8], and 11 keywords in the security report by Symantec [9]

- We then collected the 40 apps that are shown on the first page of search results

- Keywords like ”sms” and ”adult” lead users to apps with an MPR that is more than 50 times higher than the rate for GP_PO.

Malware Download Ratio

- We define Malware Download Ratio(MDR)
  - This means percentage of all downloads of the collected apps that are detected as malware

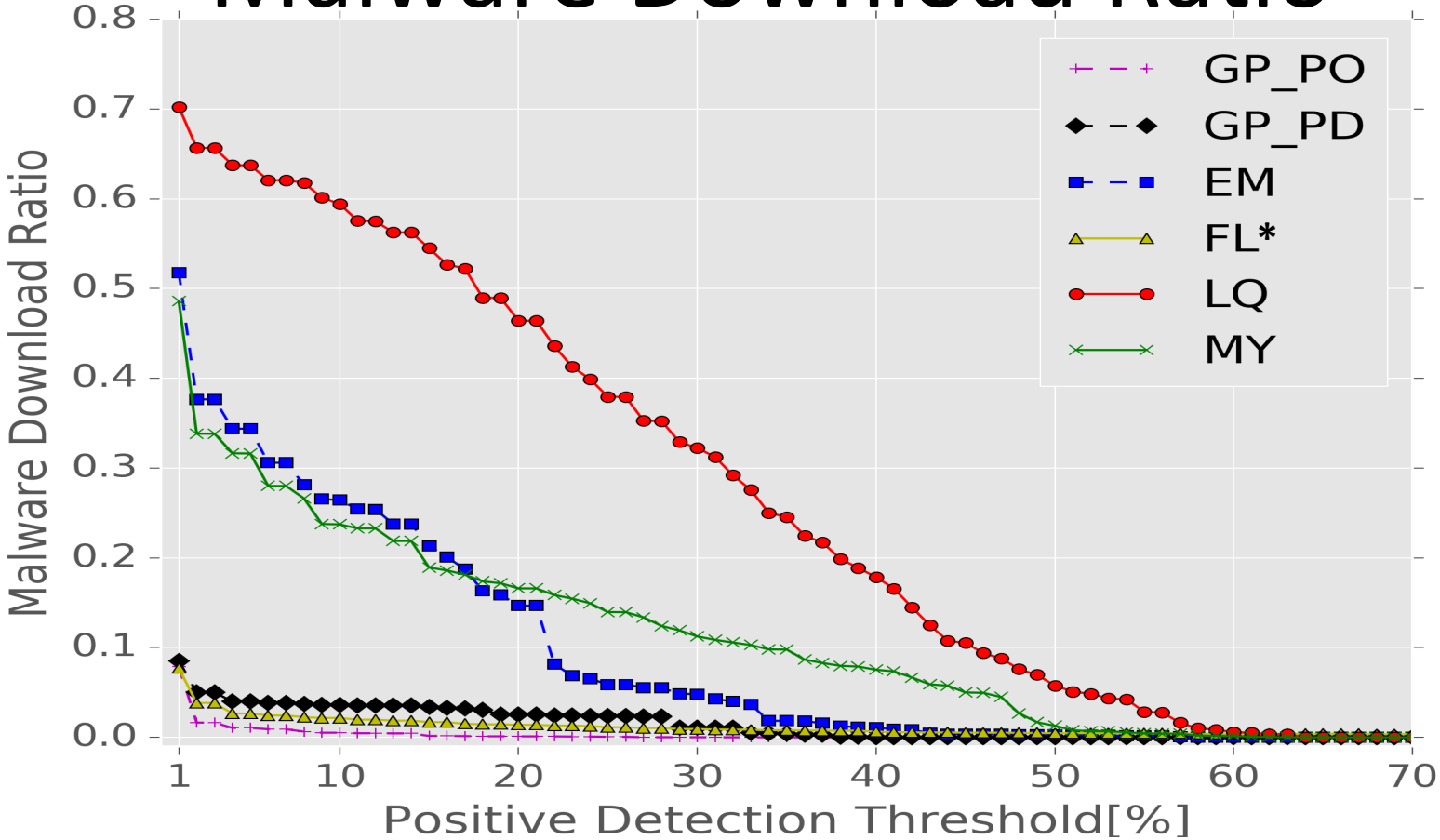
Total download number of all malware

$$MDR(m,S,M) = \frac{\sum_{a \in M'} DL(a)}{\sum_{a \in M} DL(a)}$$

Total download number of all apps

m: Threshold
S : Set of AV products VirusTotal used to scan
M: Set of Apps in a market
M': Set of Malware in M
a : Android App

Malware Download Ratio

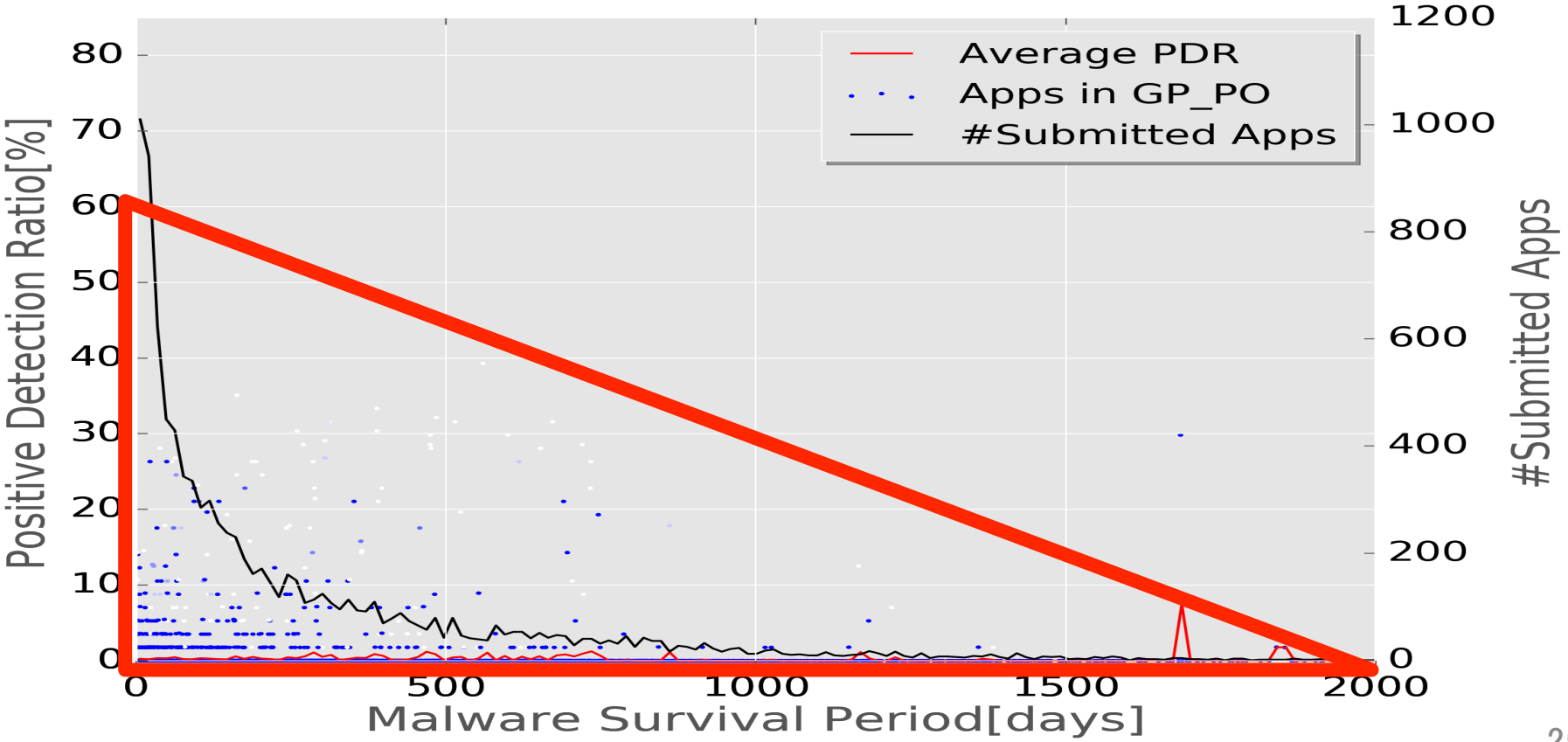


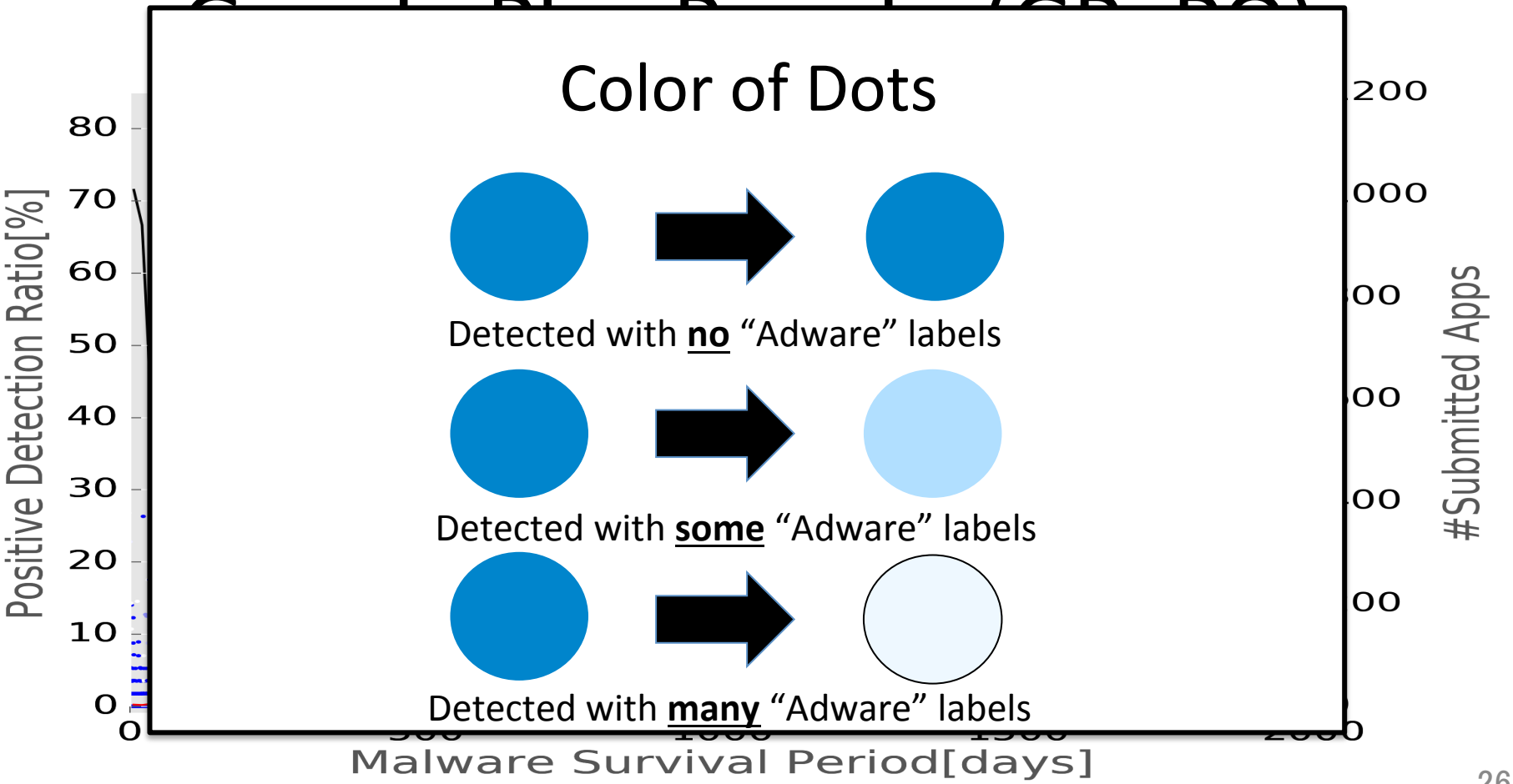
*In FL, the number of downloads is unpublished. For reference, we calculated the index on the assumption that each app was downloaded same times.

Malware Survival Period

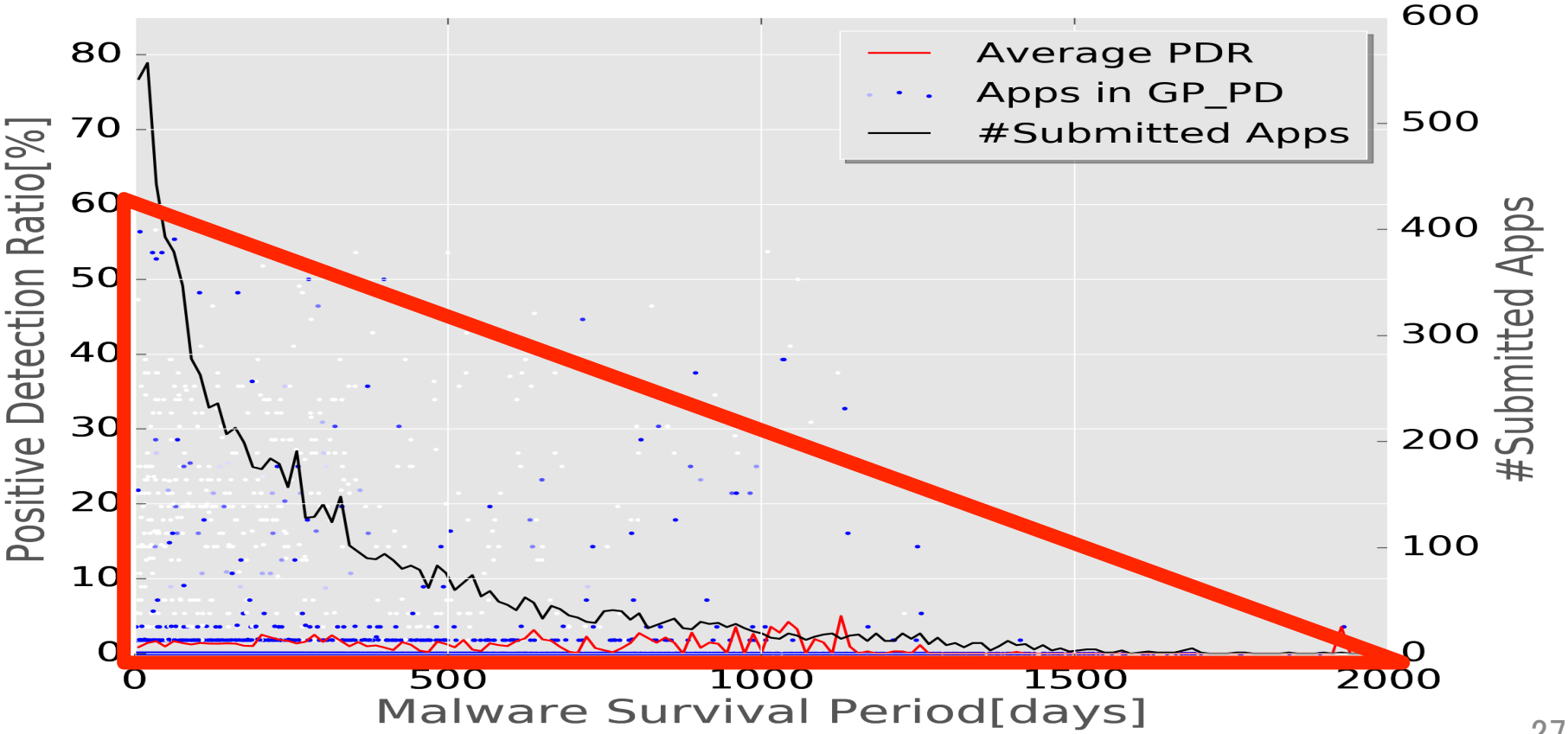
- We measure malware survival period
  - The survival period of an app is the time that an app has been present in an app store.
- To calculate a survival period, we count the number of days between the date when we downloaded the app and the date when it was last updated
- We expected that if market operators make effort to mitigate malware, high positive detection ratio (PDR) apps cannot live a long time

Google Play Popular(GP_PO)

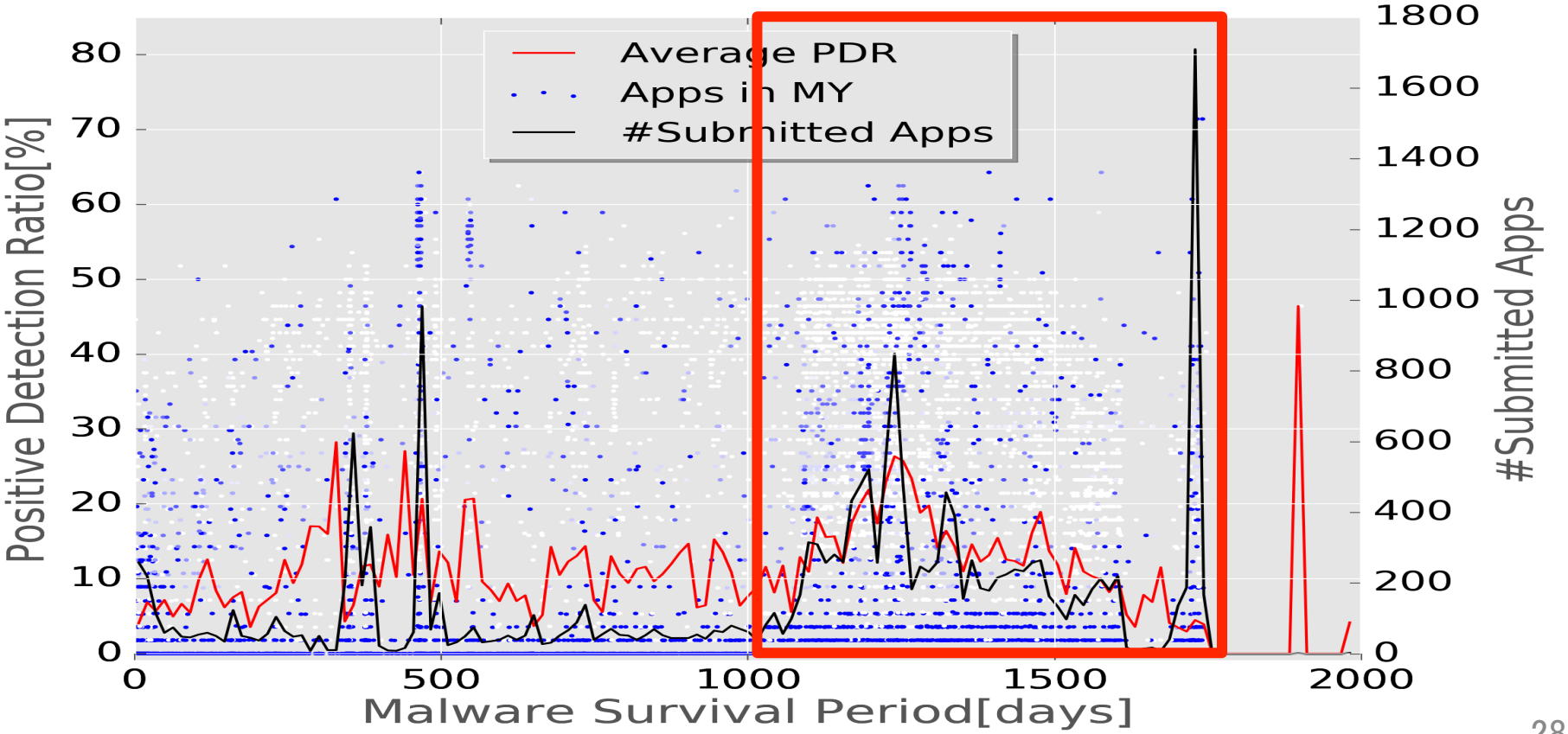



Color of Dots
Detected with no “Adware” labels
Detected with some “Adware” labels
Detected with many “Adware” labels
Positive Detection Ratio[%]
#Submitted Apps
Malware Survival Period[days]
0
10
20
30
40
50
60
70
80
200
000
00
00
00
00

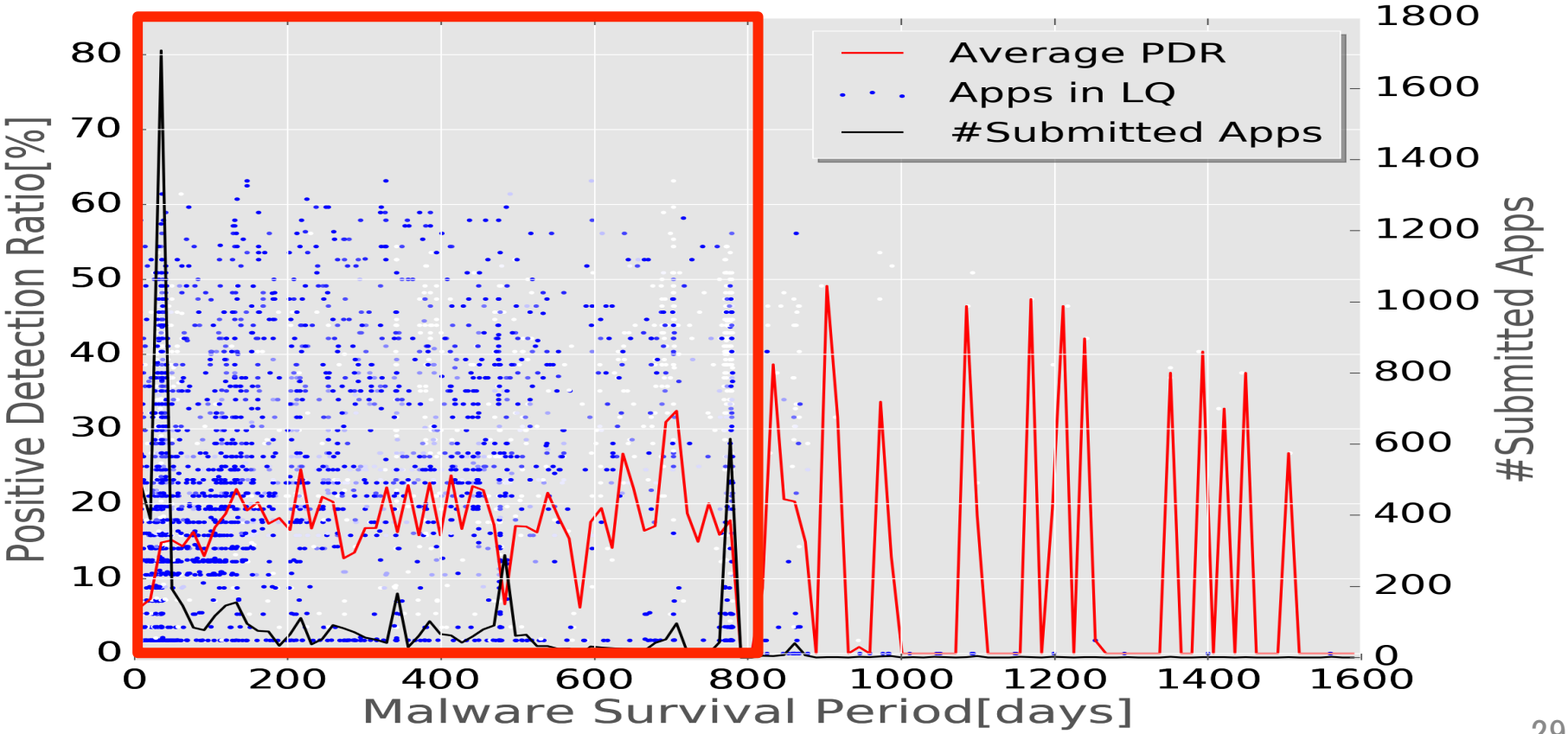
Google Play Play Drone(GP_PD)



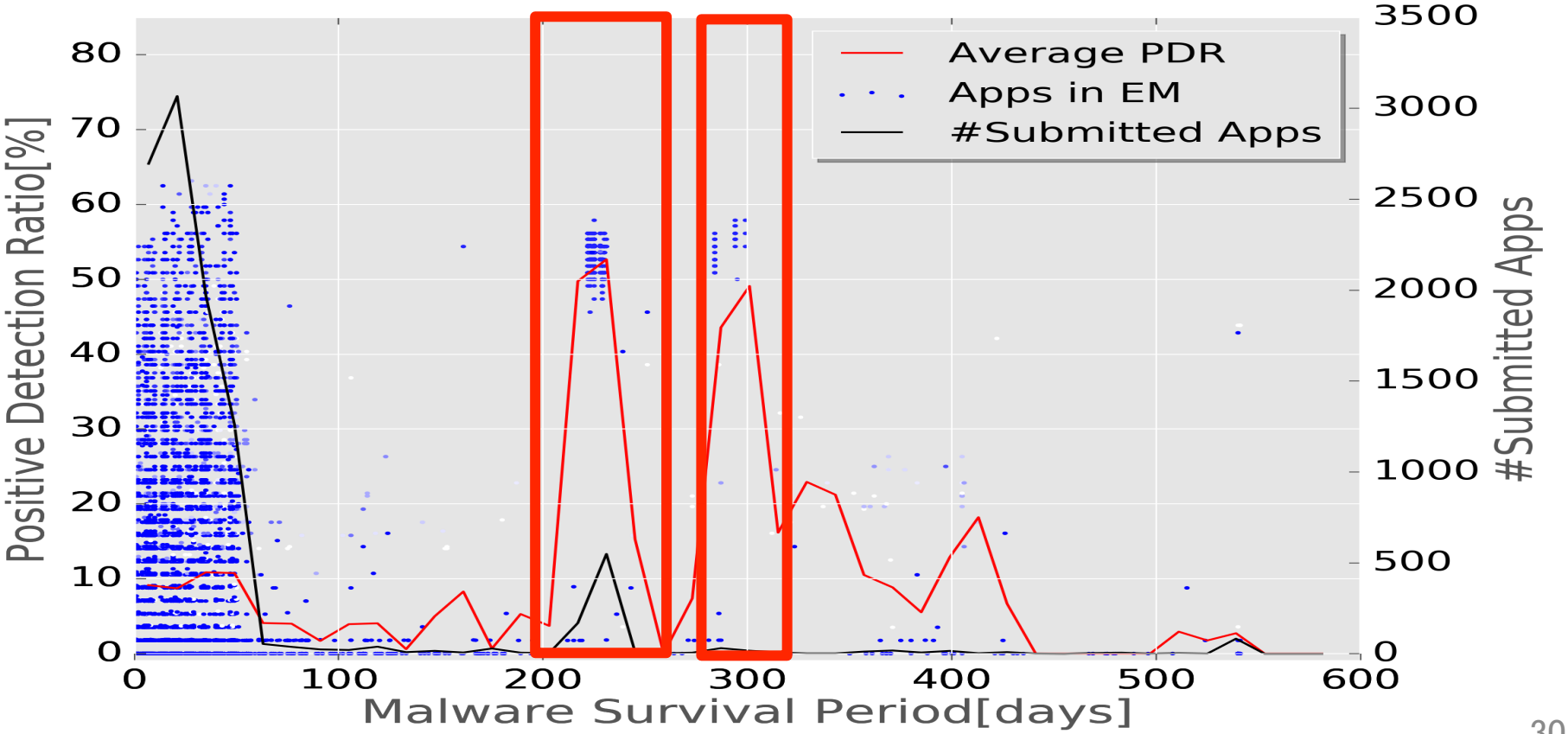
Mumayi(MY)


Average PDR
Apps in MY
#Submitted Apps
Positive Detection Ratio[%]
#Submitted Apps
Malware Survival Period[days]
0
10
20
30
40
50
60
70
80
0
200
400
600
800
1000
1200
1400
1600
1800
0
500
1000
1500
2000

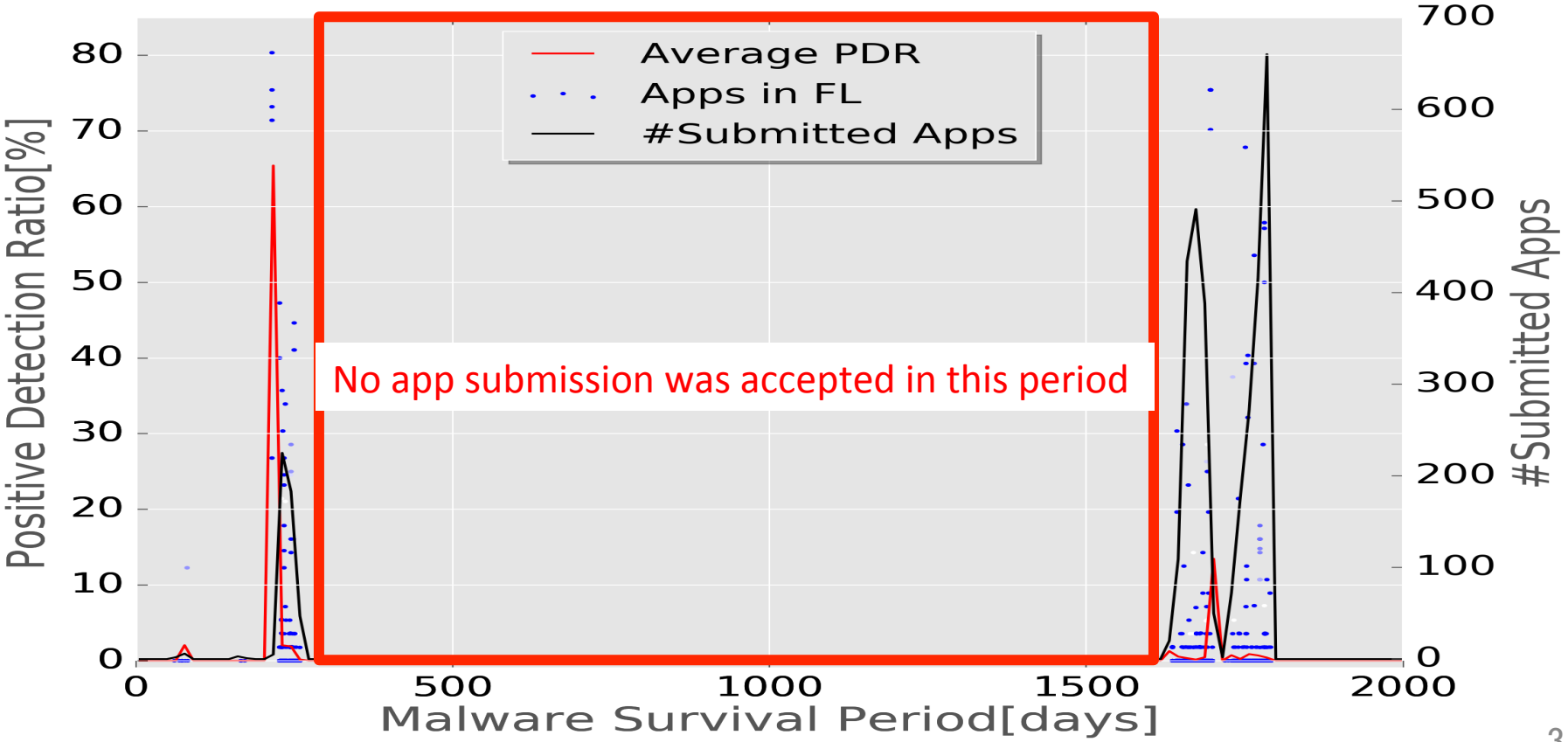
Liqucn(LQ)


Average PDR
Apps in LQ
#Submitted Apps
Positive Detection Ratio[%]
#Submitted Apps
Malware Survival Period[days]
0
10
20
30
40
50
60
70
80
0
200
400
600
800
1000
1200
1400
1600
1800
0
200
400
600
800
1000
1200
1400
1600

eoeMarket(EM)



Freeware Lovers(FL)



Conclusion

- At a PDR of 30%, around one in five apps are flagged as malicious in EM, MY and LQ
- MPR, MDR are relatively low for GP and FL but when users of Google Play search with specific keywords, they may have a 50 times higher probability of encountering malware that when selecting from the popular app list
- EM, MY, LQ and FL do not seem to act against malware even when PDRs are very high
- Google Play is the only market that seems to conduct active malware removal

Limitations

- Our results depend on VirusTotal (Anti Virus products)
- We simply count numbers of detection. We did not delve deeply into the labels
- There are some markets becoming popular in last few years that we didn't evaluate. (e.g. Baidu)

Thank You for Your Attention!