AI in mobile protection

kaspersky

Victor Chebyshev,
Lead Security Researcher,
Global Research & Analysis
Team, Kaspersky

In **2009** there were **3** Android mobile threats per month on average

I thought that I could handle it by myself using a simple signature-based engine

What happened in 2010?

20,255 classified samples per month

“Well, it is still possible to handle it with the same engine, but I have to work on Saturdays." - I thought

What did mobile malware look like in 2011?


```
int v8 = v7 + 1;
Log.v("MessageSender", v9.getNumber(v8) + " " + v9.getPrefix(v8) + " " + v9.getSuffix(v8));
PendingIntent v4 = PendingIntent.getActivity(context, 0, new Intent(context, Context.class),
        0);
SmsManager v0 = SmsManager.getDefault();
String v3 = v9.getPrefix(v8) + " " + GlobalConfig.getInstance().getValue("id") + " " + "none"
         + " " + "android" + " " + v9.getSuffix(v8);
StringBuilder v6 = new StringBuilder();
v6.append("number:").append(v9.getNumber(v8)).append(v3);
Log.v("debug", v6.toString());
v0.sendTextMessage(v9.getNumber(v8), v2, v3, v4, ((PendingIntent)v2));
++v7;
goto label_22;
```

So, there were simple SMS Trojans, they changed only numbers and text messages, however the code remained almost the same

What happened in 2012?

467,515 classified samples per month

Malware became complex, harder to analyze

We used:

- Four analysts
- Heuristics
- Statistics

But we were still unable to handle that avalanche

Fttkit

- Obfuscation as a service
- Trojan-Dropper
- Carry financial threats (banking Trojans)
- 366,636 unique files total and growing

So, we created two technologies

Sandbox technology

Machine learning engine

Spoiler: I hoped that sandbox would show its power. But ML won.

Machine learning is based on two aspects

Huge marked-up data that we collected since 2009

Expert knowledge of the most common threat features

Up to **33%** of new mobile malware files per month were classified by machine learning

Time to chill?
Not for **Kaspersky**

Mobile ML online

Idea: to effectively protect users using ML we must operate on user's device

Problems: battery and storage limitations

Kaspersky creed – not to downgrade performance

Mobile ML online

We placed the model into the cloud service

Mobile solution calculates features and send them to the cloud

Cloud responds with reputation score

ML detected the most often threats to mobile devices

Top 3 mobile malware programs	
The following malware rankings omit riskware, such as Risk Tool and AdWare	
Verdict	Share of users, attacked by this type of malware, %
DangerousObject.Multi.Generic	36.95
Trojan.AndroidOS.Boogr.gsh	9.54
DangerousObject.AndroidOS.GenericML	6.63

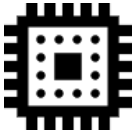
Source: Mobile malware evolution report 2020: https://securelist.com/mobile-malware-evolution-2020/101029/

Machine learning is a complex story:

You must hire a senior experienced mobile virus analyst

You must have a huge relevant virus database

You must hire a senior experienced data scientist



You must have enough computing power to retrain the model almost every day

It's worth it!

We plan even more:

Machine learning against mobile web threats

New techniques, e.g. dynamic data processing

Thank you!

kaspersky