Identifying Unknown Android Malware with Feature Extractions and Classification Techniques

Ludovic Apvrille
ludovic.apvrille@telecom-paristech.fr

Axelle Apvrille
aapvrille@fortinet.com

TRUSTCOM’2015
Many Android Applications (and Malware!)

Application repositories

- Google Play: 1.6 million+
- F-Droid, APPSAPK, APKTOP, …
Many Android Applications (and Malware!)

Application repositories

- Google Play: 1.6 million+
- F-Droid, APPSAPK, APKTOP, . . .

Malware

- Aug. 2015. 2 millions+ malicious Android samples
- 2,000+ new malicious Android samples every day
Malware: Android Carbon 14 Dating ;)

Shortest detection delay for some samples by all AV vendors

<table>
<thead>
<tr>
<th>Name</th>
<th>Creation date</th>
<th>Detection date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Android/Wroba</td>
<td>June 16 2014</td>
<td>June 21 +5d</td>
</tr>
<tr>
<td>Android/Curesec</td>
<td>July 3 2014</td>
<td>July 11 +8d</td>
</tr>
<tr>
<td>Android/ScarePakage</td>
<td>July 13 2014</td>
<td>July 24 +11d</td>
</tr>
</tbody>
</table>
Malware: Android Carbon 14 Dating ;)

<table>
<thead>
<tr>
<th>Name</th>
<th>Creation date</th>
<th>Detection date</th>
<th>Delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>Android/Wroba</td>
<td>June 16 2014</td>
<td>June 21 +5d</td>
<td></td>
</tr>
<tr>
<td>Android/Curesec</td>
<td>July 3 2014</td>
<td>July 11 +8d</td>
<td></td>
</tr>
<tr>
<td>Android/ScarePakage</td>
<td>July 13 2014</td>
<td>July 24 +11d</td>
<td></td>
</tr>
<tr>
<td>Android/Ganlet</td>
<td>Nov 1 2013</td>
<td>May 15 2014 +6 months!!</td>
<td></td>
</tr>
</tbody>
</table>
What Are We Interested In?

KNOWN (and detected) MALWARE → NOTHING TO DO (BORING...)

MINOR VARIANT (usually detected) → NOT VERY INTERESTING FOR RESEARCH

MAJOR VARIANT

detected → AH HA?!

not detected

UNKNOWN MALWARE → TOP INTEREST

RESEARCH/PR INNOVATION
Problems with Manual Search

Too many apps and marketplaces to crawl
Waste time on clean apps
Even a team of 100 analysts is insufficient
Problems with Manual Search

Too many apps and marketplaces to crawl
Waste time on clean apps
Even a team of 100 analysts is insufficient

We need an automated system that helps identifying unknown malware with less effort

→ SherlockDroid
SherlockDroid to the Rescue!

Crawl Android marketplaces

Spot suspicious apps

Focus on major variants and unknown malware
SherlockDroid Architecture

**SherlockDroid Architecture**

- **Google Play**
- **APKTop**
- **SlideME**

**Crawling**

**Pre-Filtering**

**Database**

**Property Extractor - Droidlysis**

**Classification - Alligator**

**Yummy!**
Remarks on SherlockDroid

It is not an AV scanner because SherlockDroid does not handle known malware / minor variants
Remarks on SherlockDroid

It is not an AV scanner because SherlockDroid does not handle known malware / minor variants

We will miss some malware
We’re not (yet) perfect ;-)

8/20 Aug., 2015 Institut Mines-Telecom Identifying Unknown Android Malware
Remarks on SherlockDroid

It is not an AV scanner because SherlockDroid does not handle known malware / minor variants.

We will miss some malware

We’re not (yet) perfect ;-)  
But we would have missed them without SherlockDroid too
**DroidLysis - Extracting Properties/Features**

Static extraction of 289 properties

- **54 File-related properties**
  - Permissions, certificate, …

- **22 Resource properties**
  - Native code, resource risky calls (*su*, *mount*, etc.), Javascript, URLs, …

- **70 Dalvik code properties**
  - API usage, actions, intents, constants, implementation techniques (e.g., *junk bytecode injection*)

- **143 Third party kits properties**
  - Advertisements, statistics reporting, error reporting
**Alligator - Classification (1/2)**

### Usual classification approach

- **Data to be classified**
- **Classifier** 
  - (SVM or k-NN or ...)
- **Already classified data**
  1. 1
  2. 2
- **Classified data**

### Alligator

- **Data to be classified**
- **Classifier** 
  - (SVM and k-NN and ...)
- ** Already classified data**
  1. 1
  2. 2
- **Classified data**

*Alligator automatically combines classification algorithms in order to obtain better classification results*
Alligator - Classification (2/2)

Other capabilities

- Favor a cluster over another
- Forget/boost too abnormal elements

Other

- Shown to better classify than other classifiers (e.g., SVM) in various application domains (e.g., image classification)
- Free and open-source, easy to install and configure, scriptable

alligator.telecom-paristech.fr
SherlockDroid: Hall of “Fame”

- Android/MisoSMS.A!tr.spy
- Android/Odpa.A!tr.spy
- Adware/Geyser!Android
- Riskware/Flexion!Android
- Riskware/SmsControlSpy!Android
- Riskware/Zdchial!Android
- Riskware/SmsCred!Android
- Riskware/Blued!Android
- Riskware/SneakFont!Android

Descriptions: http://www.fortiguard.com/encyclopedia/
SherlockDroid: Unknown Malware Identified

Do you known any other framework who identified real unknown malware?
SherlockDroid: Unknown Malware Identified

Do you known any other framework who identified real unknown malware?

Answer: DroidRanger: 2
SherlockDroid: Unknown Malware Identified

Do you known any other framework who identified real unknown malware?

Answer: DroidRanger: 2

AAS, Andromaly, CopperDroid, Crowdroid, Drebin, MADAM, MAST, pBMDS, PUMA...
tested on artificial or known malware
Learning and Classification Results

Typical results we expect

▶ FP/FN shall be as low as possible (Obviously)
▶ FP shall be much lower than FN (Missing a malware is not a big deal w.r.t. wasting time on false alerts)

Samples

▶ Learning clusters: 500k samples used in the learning clusters
  ▶ ~ 487k malware, ~ 12k clean
  ▶ Gathered before June 2014

▶ Testing clusters: 1.5k clean and 3k malware gathered after Sept. 2014
# Learning and Classification Results (Cont.)

<table>
<thead>
<tr>
<th>Learning cluster size</th>
<th>Learning time</th>
<th>Classification time</th>
<th>FP</th>
<th>FN</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>480,000</td>
<td>∼ 11 hours</td>
<td>6 mn</td>
<td>1.78%</td>
<td>0.52%</td>
<td>0.93%</td>
</tr>
<tr>
<td>50,000</td>
<td>20 min</td>
<td>∼ 34 s</td>
<td>0.72%</td>
<td>2.1%</td>
<td>1.67%</td>
</tr>
</tbody>
</table>
Learning and Classification Results: Comparison

Alligator vs. SVM - Learning and classification time depending on the learning cluster size

- Alligator learning time
- Alligator classification time
- SVM learning time
- SVM Classification time

Number of samples in the learning clusters

Regular cluster size
Malware cluster size

Learning time (in seconds)
Classification time (in seconds)
Learning and Classification Results: Comparison (Cont.)

Alligator vs. SVM - False Positive and False Negative rates depending on learning cluster size

- Alligator FP
- Alligator FN
- Alligator FP+FN
- SVM FP
- SVM FN
- SVM FP+FN
What About the TrustCom Android App?

- Property extraction: 17 seconds
- Classification: 4 seconds
- Regular score: 77.0
- Malware score: 22.7
- → Classified as regular ;-)
Conclusion and Future Work

- SherlockDroid is operational, tested on a huge number of applications from various application markets
- 9 unknown malware identified
- For classification purpose, it relies on the Alligator meta-classifier

What’s next?

- Feature extraction: mix contextual information (e.g., call stack) and the related features
  - Sending an email is not the same if it is for a bug report or for a connection to a C&C server
- Differentiate between malware and Potentially Unwanted Applications
Thank You

Contact info

SherlockDroid: aapvrille at fortinet dot com
Alligator: ludovic dot aprville at telecom minus paristech dot com

References

Alligator Release: alligator.telecom-paristech.fr


Powerpoint slides? No way! This is \LaTeX- Beamer!