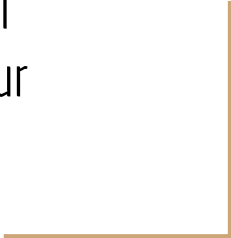




Detecting Malware in Android

Professor Fu
Hani Alshahrani
Harrison Mansour
Seaver Thorn



Outline

- Extra Feature Extraction
- Components
- Intents
- Network
- This week
- Questions

Feature Extraction

- Currently, we have ~590 useable features extracted from each app
 - From these features, our algorithms pick the top N features to input into our machine learning models.
 - It may be beneficial to expand the set of possible features in order to find an even more ideal set of top N features
- We are expanding our feature selection scripts via two methods:
 - AndroidManifest.xml file analysis
 - System network analysis

Manifest Extraction

- Every application that is downloaded on a device has an APK, which is a compiled folder containing key information about the application.
 - We modified DDefender to find, extract, and send this APK to our server.
- In its current state, the APK is unreadable. We feed it to Apktool, which is a reverse engineering tool that can decode the APK into nearly its original format.
- Now, we have easy access to the AndroidManifest.xml file, which contains key information about the application.

Manifest Analysis

- The AndroidManifest.xml file contains pertinent information about the application, including:
 - Declared permissions
 - A list of activities, services, receivers, and providers (these are the components)
 - Each component defines an “intent-filter” that lays out all of the relevant intents, which can be of type:
 - Action
 - Category
 - Extra
 - Developer-Defined
- We use minidom, which is an xml parser-scraper to extract relevant information. We then turn them into numerical features and add them to our database

Components

- A new feature we have defined focuses on how many of each component type (activity,service,receiver,provider) an app defines. This may be an interesting feature because it allows us to relate the size of an app (how many different activities or how many different providers that send information it has, for example)

Intents

- Intents are a way for different components within an app to communicate with each other. For example:
 - The startActivity intent allows you to start another activity.
- Action intents define the generic action to be performed (ex: view)
- Category intents specify additional information about the kind of component that should handle the event
- Extra intents specify intents that don't directly fall under Action or Category
- We ignore user-defined intents because they are likely to only show up in one application, and are thus useless to us.

Networks

- Many different ways to collect network information.
 - Tcpdump
 - Strace
 - VPN
 - Usage Statistics (Android 6.0+)
- Many upsides and downsides of these solutions.
- Tcpdump - cannot filter by process id.
- Strace - just records system call information of network calls.
- VPN - time consuming to implement into our app.
- Usage Statistics - only available in Android after version 6.0

Implementation of Networks

- Currently, tcpdump is implemented and works correctly.
- Cannot filter by apps, so the network information is from every app on the device.
- We plan on implementing the Usage statistics solution. We hope this to be done by the end of the week.
- Since this implementation is only available after Android 6.0 (Marshmallow), about 40% of current Android users will be able to use it according to current Android OS statistics collected in June 2017. This percentage will increase over time.

This Week

- Finalize new features.
- Run all these features through machine learning to see if it improves accuracy.
- Start drafting research paper.

Questions

