Privacy Implication of Health Related Information in Android Apps

Jun Ma, Yuankun Li & Samuel Nzaramba

INTRODUCTION

Mobile applications collect and share enormous health related data, such as activity information, body information and some other kinds of health related information from users.

We implemented a static analysis tool to analyze the collection and sharing of health related information. We also used a wearable devices and a mobile phone to test if indeed the apps collected and shared health related information.

Related Tools

- Androguard:
  Decompile apps through reverse engineering
  Reverse most of source code of apps
- Apktool:
  Decompile apps through reverse engineering
  Reverse most of source code of apps
- Raccoon:
  Get .apk files of apps
- Test Devices:
  Moto Phone and Sony Smart Watch

METHODS

- Use Apktool and Androguard to reverse apps, find specific features of health apps codes
- Differentiate collection activities
  - Body Sensors Permission
  - Sensor call methods
  - Google Fit and other cloud service providers
- Differentiate sharing activities
  Google Fit and other cloud service providers
- Verify the accuracy of our methods
  - Manually analysis the activities of apps and then compare the results with the results from our methods
  - Calculating accuracy, precision, recall and F1 value
- Analysis
  - Download .apk files from Google play through raccoon
  - Apps came from recommended health & fit apps of Google Play

ACCURACY

We got recommended Health & Fitness apps from Google Play, first we determined if they collected and shared health related information by reading their description, secondly we installed them on real devices and tested whether they collected step count and heart rate information. Then, we analyzed the apps with our tool and used the results to calculate the accuracy, recall, precision and F1 value.

1) Check collection of step information

	No. of Apps detected by our tool as collecting steps	No. of Apps detected by our tool as not collecting steps
No. of Apps collect steps	15	5
No. of Apps don't collect steps	6	14

Accuracy	Recall	Precision	F1
0.725	0.75	0.7143	0.7317

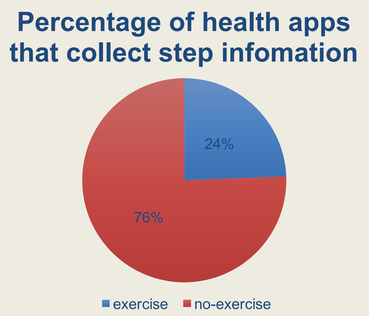
2) Check collection of heart rate information

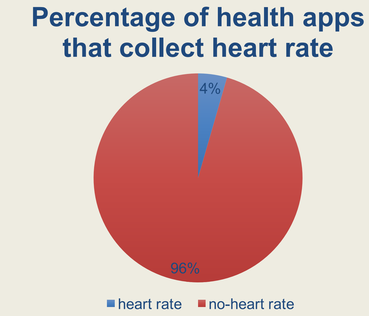
	No. of Apps detected by our tool as collecting heart rate	No. of Apps detected by our tool as not collecting heart rate
No. of Apps collect heart rate	5	0
No. of Apps collect heart rate	0	5

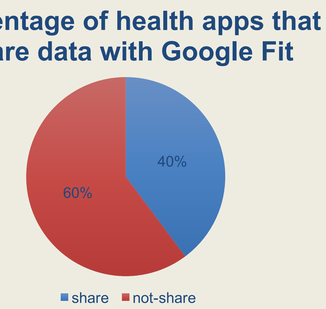
Accuracy	Recall	Precision	F1
1	1	1	1

RESULTS

We downloaded 201 apps from Google Play with Raccoon. And then ran experiments for 201 apps with our tool. All of those 201 apps came from recommended apps of Health & Fitness of Google Play.







DISCUSSION

- Limitation
  - Static analysis cannot fully reflect the actual behaviors of Apps
  - We cannot ensure the behaviors of apps without dynamic analysis
  - Developers might write useless codes
  - Some apps might collect health related information without standard sensors
  - Our methods cannot differentiate the categories of health related information being shared, and information being collected by Google fit.
- Future work
  From our work, we discovered that many applications collect and share health related information from users which may pose a potential privacy violation. We recommend more work to be done in identifying the laws and regulations that cover these activities and if they are followed by the app developers or data owners.