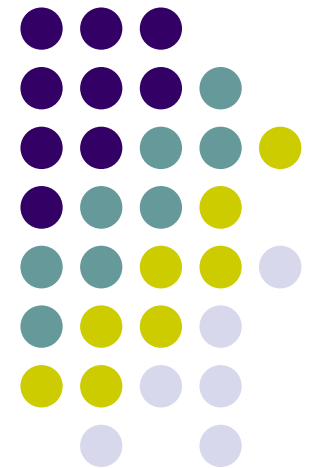


Advanced Computer Graphics

CS 525M: ProfileDroid: Multi-layer Profiling of Android Applications

Cheng Cheng

*Computer Science Dept.
Worcester Polytechnic Institute (WPI)*





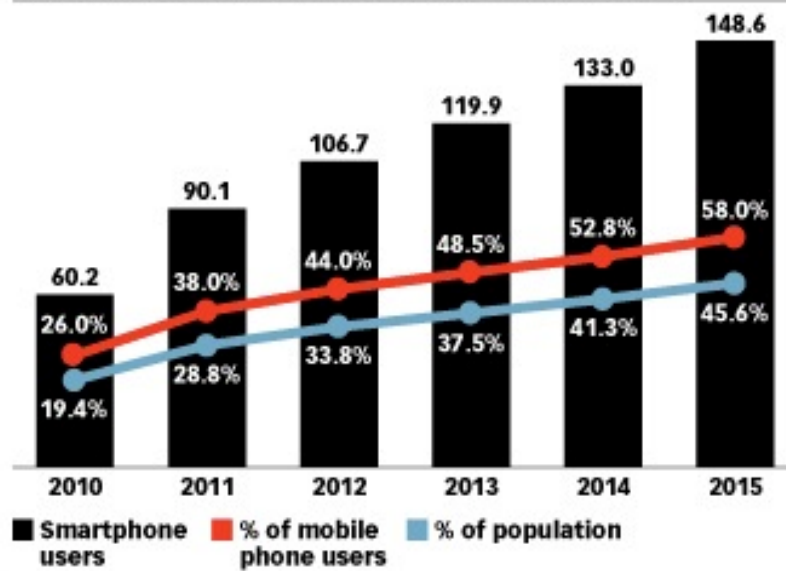
Motivation

More and more people use smartphones

Android is an very important platform

US Smartphone Users and Penetration, 2010-2015

millions, % of mobile phone users and % of population



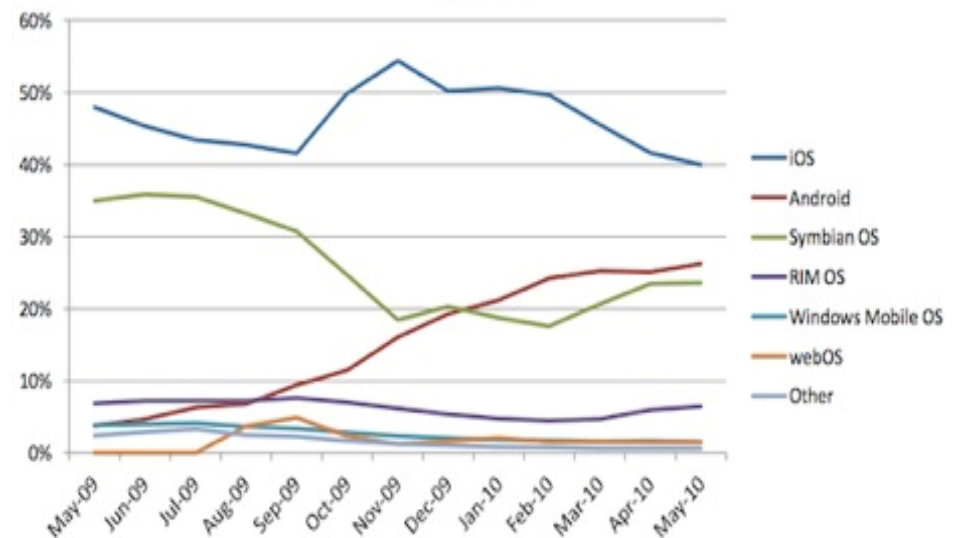
Source: eMarketer, Aug 2011

130909

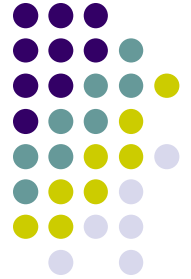
www.eMarketer.com

Operating System Share

Worldwide



Motivation



The image shows a screenshot of the Google Play Store interface. The top navigation bar includes 'SHOP', 'MY MUSIC', 'MY BOOKS', 'MY MAGAZINES', and 'MY APPS'. The main content area displays a grid of app cards, each with a rating, an 'INSTALL' button, and a brief description. Overlaid on this interface are several speech bubbles containing the following questions:

- my battery?
- Will this app leak my photos?
- Which radio is best for me?
- Will this app tell my friends that I'm a moron?

The app cards visible include:

- Facebook for Android (4 stars)
- Google Search (7 stars)
- Pandora® internet radio (10 stars)
- Slacker Radio (17 stars)
- Netflix (16 stars)
- Adobe Reader (19 stars)
- YouTube (5 stars)
- Google+ (8 stars)
- Twitter (14 stars)
- Pinterest (13 stars)
- Skype - free IM & video calls (20 stars)
- Angry Birds (21 stars)
- Instagram (12 stars)
- Google Play Movies & TV (15 stars)
- The Moron Test: Section 1 (18 stars, 34,090 reviews)



Related Work

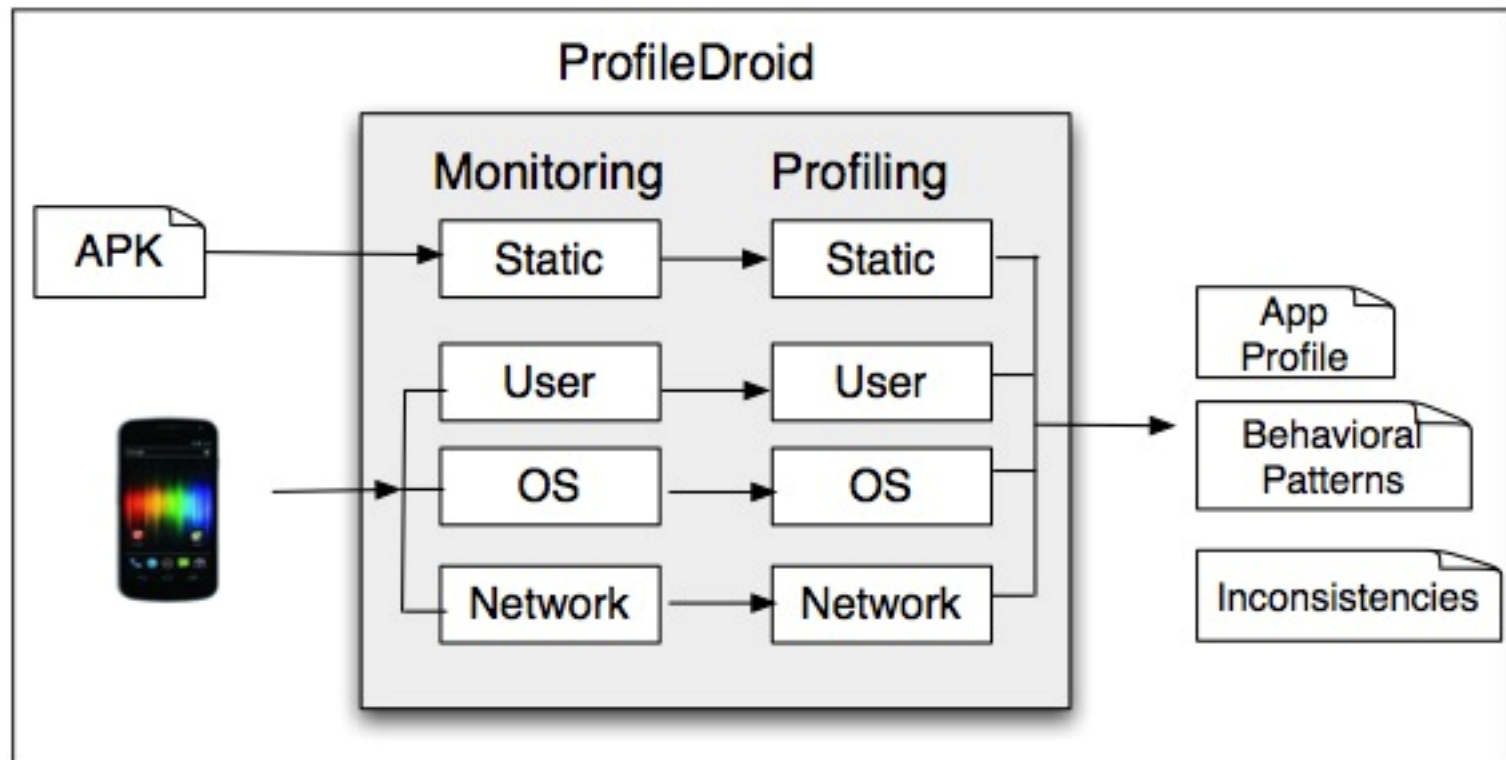
- Smartphone Measurements and Profiling
 - do not analyze the Android apps themselves.
- Android Security Related Work.
 - Static Layer do not include Intent Usage
 - Profiles the app do not from multiple layers
 - Profile the network layer was not with a more fine-grained granularity

Approach



- Four different layers:
 - (a) static, or app specification
 - (b) user interaction
 - (c) operating system
 - (d) network approach
- For each layer,
 - the monitoring component runs on the Android device
 - The profiling part runs on the connected computer.

Approach





Experiment

- Capture-and-replay

Round1:

- Each user ran each app one time for 5 minutes
- Capture the interaction using event logging

Round2:

- Using replay tools, replay back 5 times in the morning and 5 times at night. (10 runs each per user per app)

Round3:

- Apply the logs for different experiments.

Test Apps



App name	Category
Dictionary.com, Dictionary.com-\$\$	Reference
Tiny Flashlight	Tools
Zedge	Personalization
Weather Bug, Weather Bug-\$\$	Weather
Advanced Task Killer, Advanced Task Killer-\$\$	Productivity
Flixster	Entertainment
Picsay, Picsay-\$\$	Photography
ESPN	Sports
Gasbuddy	Travel
Pandora	Music & Audio
Shazam, Shazam-\$\$	Music & Audio
Youtube	Media & Video
Amazon	Shopping
Facebook	Social
Dolphin, Dolphin-\$\$	Communication (Browsers)
Angry Birds, Angry Birds-\$\$	Games
Craigslist	Business
CNN	News & Magazines
Instant Heart Rate, Instant Heart Rate-\$\$	Health & Fitness

Table 1: The test apps; app-\$\$ represents the paid version of an app.



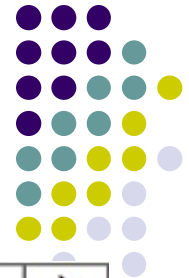
Static Layer (Layer 1)

- Analyze the APK (Android application package) file
- Use apktool to unpack the APK file to extract relevant data.
 - Focus on the Manifest.xml file
 - Bytecode files contained in smali folder.



Static Layer (Layer 1)

- Permissions (shown at install)
 - Internet
 - GPS
 - Camera, Microphone, Bluetooth, Telephony
- Intent Usage (not shown at install)
 - Resource use without permission via deputy apps



Static Layer (Layer 1)

- Result:

App	Internet	GPS	Camera	Microphone	Bluetooth	Telephony
Dictionary.com	✓			I		I
Dictionary.com-\$\$	✓			I		I
Tiny Flashlight	✓		✓			
Zedge	✓					
Weather Bug	✓	✓				
Weather Bug-\$\$	✓	✓				
Advanced Task Killer	✓					
Advanced Task Killer-\$\$	✓					
Flixster	✓	✓				
Picsay	✓					
Picsay-\$\$	✓					
ESPN	✓					
Gasbuddy	✓	✓				
Pandora	✓				✓	
Shazam	✓	✓		✓		
Shazam-\$\$	✓	✓		✓		
YouTube	✓					
Amazon	✓		✓			
Facebook	✓	✓	I			✓

Table 2: Profiling results of *static* layer; '✓' represents use via permissions, while 'I' via intents.

Craigslist	✓					
CNN	✓		✓			
Instant Heart Rate	✓		✓		I	I
Instant Heart Rate-\$\$	✓		✓		I	I



User Layer (Layer 2)

- Focus on user-generated events
- Events result from interaction between the user and the Android device while running the app.
- Use combination of the
 - Logcat: capture the system debug output and log messages from the app.
 - Getevent(read /dev/input/event*): collect the user input events

User Layer (Layer 2)

- Focus on
 - TouchScreen
 - Accelerometer
 - Proximity sensor.





User Layer Result(Layer 2)

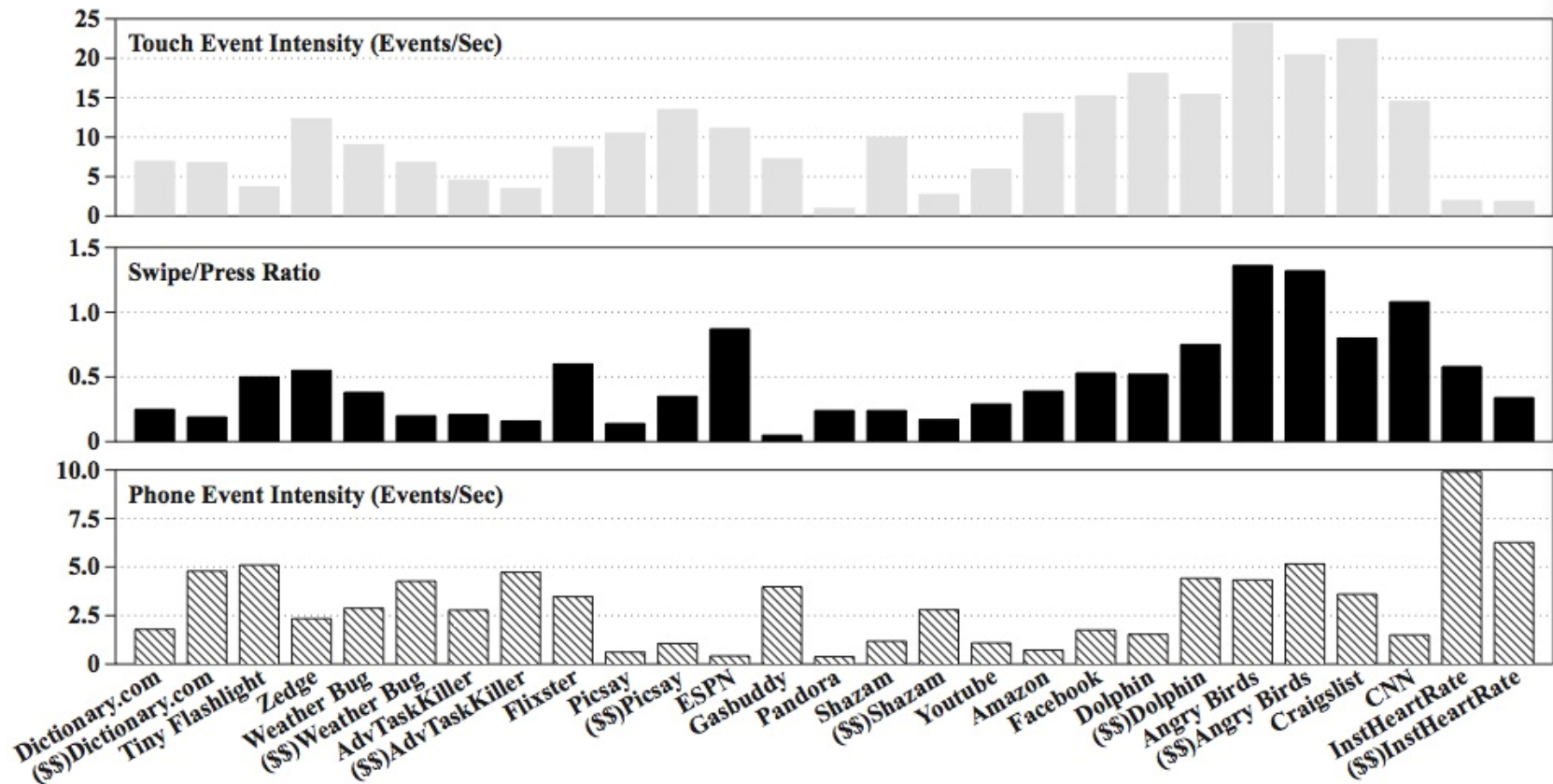


Figure 2: Profiling results of *user* layer; note that scales are different.

Operating System Layer (Layer 3)



- Monitor system calls
- Strace: collect system calls invoked by the app
- Classify system calls into four categories:
 - Filesystem
 - Network
 - VM/IPC
 - Enforces isolation
 - Overhead: scheduling, idling, IPC
 - miscellaneous

Operating System Layer(Layer 3)



App	Syscall intensity (calls/sec.)	FS (%)	NET (%)	VM&IPC (%)	MISC (%)
Dictionary.com	1025.64	3.54	1.88	67.52	27.06
Dictionary.com-\$\$	492.90	7.81	4.91	69.48	17.80
Tiny Flashlight	435.61	1.23	0.32	77.30	21.15
Zedge	668.46	4.17	2.25	75.54	18.04
Weather Bug	1728.13	2.19	0.98	67.94	28.89
Weather Bug-\$\$	492.17	1.07	1.78	75.58	21.57
AdvTaskKiller	75.06	3.30	0.01	65.95	30.74
AdvTaskKiller-\$\$	30.46	7.19	0.00	63.77	29.04
Flixster	325.34	2.66	3.20	71.37	22.77
Picsay	319.45	2.06	0.01	75.12	22.81
Picsay-\$\$	346.93	2.43	0.16	74.37	23.04
ESPN	1030.16	2.49	2.07	87.09	8.35
Gasbuddy	1216.74	1.12	0.32	74.48	24.08
Pandora	286.67	2.92	2.25	70.31	24.52
Shazam	769.54	6.44	2.64	72.16	18.76
Shazam-\$\$	525.47	6.28	1.40	74.31	18.01
YouTube	246.78	0.80	0.58	77.90	20.72
Amazon	692.83	0.42	6.33	76.80	16.45
Facebook	1030.74	3.99	2.98	72.02	21.01
Dolphin	850.94	5.20	1.70	71.91	21.19
Dolphin-\$\$	605.63	9.05	3.44	68.45	19.07
Angry Birds	1047.19	0.74	0.36	82.21	16.69
Angry Birds-\$\$	741.28	0.14	0.04	85.60	14.22
Craigslist	827.86	5.00	2.47	73.81	18.72
CNN	418.26	7.68	5.55	71.47	15.30
InstHeartRate	944.27	7.70	1.73	75.48	15.09
InstHeartRate-\$\$	919.18	12.25	0.14	72.52	15.09

Table 3: Profiling results: *operating system* layer.



Network Layer (Layer 4)

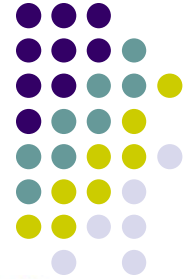
- Analyze network traffic by logging the data packets.
- Tcpdump: collect all network traffic on the device.



Network Layer Result (Layer 4)

App	Traffic intensity (bytes/sec.)	Traffic In/Out (ratio)	Origin (%)	CDN+Cloud (%)	Google (%)	Third party (%)	Traffic sources	HTTP/HTTPS split (%)
Dictionary.com	1450.07	1.94	-	35.36	64.64	-	8	100/-
Dictionary.com-\$\$	488.73	1.97	0.02	1.78	98.20	-	3	100/-
Tiny Flashlight	134.26	2.49	-	-	99.79	0.21	4	100/-
Zedge	15424.08	10.68	-	96.84	3.16	-	4	100/-
Weather Bug	3808.08	5.05	-	75.82	16.12	8.06	13	100/-
Weather Bug-\$\$	2420.46	8.28	-	82.77	6.13	11.10	5	100/-
AdvTaskKiller	25.74	0.94	-	-	100.00	-	1	91.96/8.04
AdvTaskKiller-\$\$	-	-	-	-	-	-	0	-/-
Flixster	23507.39	20.60	2.34	96.90	0.54	0.22	10	100/-
Picsay	4.80	0.34	-	48.93	51.07	-	2	100/-
Picsay-\$\$	320.48	11.80	-	99.85	0.15	-	2	100/-
ESPN	4120.74	4.65	-	47.96	10.09	41.95	5	100/-
Gasbuddy	5504.78	10.44	6.17	11.23	81.37	1.23	6	100/-
Pandora	24393.31	28.07	97.56	0.91	1.51	0.02	11	99.85/0.15
Shazam	4091.29	3.71	32.77	38.12	15.77	13.34	13	100/-
Shazam-\$\$	1506.19	3.09	44.60	55.36	0.04	-	4	100/-
YouTube	109655.23	34.44	96.47	-	3.53	-	2	100/-
Amazon	7757.60	8.17	95.02	4.98	-	-	4	99.34/0.66
Facebook	4606.34	1.45	67.55	32.45	-	-	3	22.74/77.26
Dolphin	7486.28	5.92	44.55	0.05	8.60	46.80	22	99.86/0.14
Dolphin-\$\$	3692.73	6.05	80.30	1.10	5.80	12.80	9	99.89/0.11
Angry Birds	501.57	0.78	-	73.31	10.61	16.08	8	100/-
Angry Birds-\$\$	36.07	1.10	-	88.72	5.79	5.49	4	100/-
Craigslist	7657.10	9.64	99.97	-	-	0.03	10	100/-
CNN	2992.76	5.66	65.25	34.75	-	-	2	100/-
InstHeartRate	573.51	2.29	-	4.18	85.97	9.85	3	86.27/13.73
InstHeartRate-\$\$	6.09	0.31	-	8.82	90.00	1.18	2	20.11/79.89

Table 4: Profiling results of *network* layer; ‘-’ represents no traffic.



ProfileDroid: Profiling apps

- Extract information from each layer in isolation or in combination with other layers.

App	Static (# of func.)	User (events/ sec.)	OS (syscall/ sec.)	Network (bytes/ sec.)
Dictionary.com	L	M	H	M
Dictionary.com-\$\$	L	M	M	M
Tiny Flashlight	M	L	M	L
Zedge	L	M	M	H
Weather Bug	M	M	H	M
Weather Bug-\$\$	M	M	M	M
AdvTaskKiller	L	M	L	L
AdvTaskKiller-\$\$	L	M	L	L
Flixster	M	M	L	H
Picsay	L	M	L	L
Picsay-\$\$	L	M	M	M
ESPN	L	M	H	M
Gasbuddy	M	M	H	M
Pandora	M	L	L	H
Shazam	H	L	M	M
Shazam-\$\$	H	L	H	M
YouTube	L	M	M	H
Amazon	M	M	M	H
Facebook	H	H	H	M
Dolphin	M	H	M	H
Dolphin-\$\$	M	H	M	M
Angry Birds	L	H	M	M
Angry Birds-\$\$	L	H	H	L
Craigslist	L	H	H	H
CNN	M	M	M	M
InstHeartRate	M	L	H	M
InstHeartRate-\$\$	M	L	H	L

Table 5: Thumbnails of multi-layer intensity in the *H-M-L* model (*H*:high, *M*:medium, *L*:low).



Result

- **Free apps are not as free as we might think**
 - 50—100% higher system call intensity
 - Dramatically higher network traffic (usually ads&tracking)
 - Bad for your dataplan, your battery life, and your privacy
- **VM-based isolation comes at a cost**
 - 64—87% of system calls are due to VM and IPC



Result

- **Apps talk to many servers spread across many top-level domains**
 - AngryBirds\$\$: 4 domains, AngryBirds free: 8 domains
 - Weatherbug: 13 domains, Shazam: 13 domains
- **Most network traffic is not encrypted**
- **Google traffic is predominant**
 - Except for Amazon and Facebook which have 0 (zero) Google traffic



Future Work

- **Expand study to include more apps**
- **User profiles**
 - Study the variance across users
- **Fully automate process**
- **Profiler as an app to run on the device**
- **• Provide summary of usage on close**



References

- <http://www.sigmobile.org/mobicom/2012/slides/Gomez.pdf>