

# Understanding Mobile App Usage Patterns Using In-App Advertisements

Alok Tongaonkar<sup>1</sup>, Shuaifu Dai<sup>2,3</sup>, Antonio Nucci<sup>1</sup>, Dawn Song<sup>3</sup>

<sup>1</sup> Narus Inc, USA

<sup>2</sup> Peking University, China

<sup>3</sup> University of California, Berkeley, USA



# Motivation

- Mobile Device vs PC
  - 488 million vs 415 million (2011)
- Applications of understanding app behavior
  - Network Management
  - Security
  - Market Analysis

# New Challenges

- Explosive growth rate of apps
  - 700,000 apps in Google Play (Oct. 2012)
- Bring Your Own Device(BYOD)
  - Network admins have no control over apps on personal device in enterprises
- Network operator needs to be aware of all apps

# State of the Art Techniques

- User-Agent
  - [Xu, Q., etc.] Identifying diverse usage behaviors of smartphone apps. IMC, 2011.
  - Not strictly enforced, even worse for Android apps
- Host
  - [Falaki, H., etc.] A first look at traffic on smartphones. IMC, 2010
  - Not unique
- Manually running apps
  - [Wei, X., etc.] Profiledroid: Multi-layer profiling of android applications. MobiCom, 2012
  - Require human labor

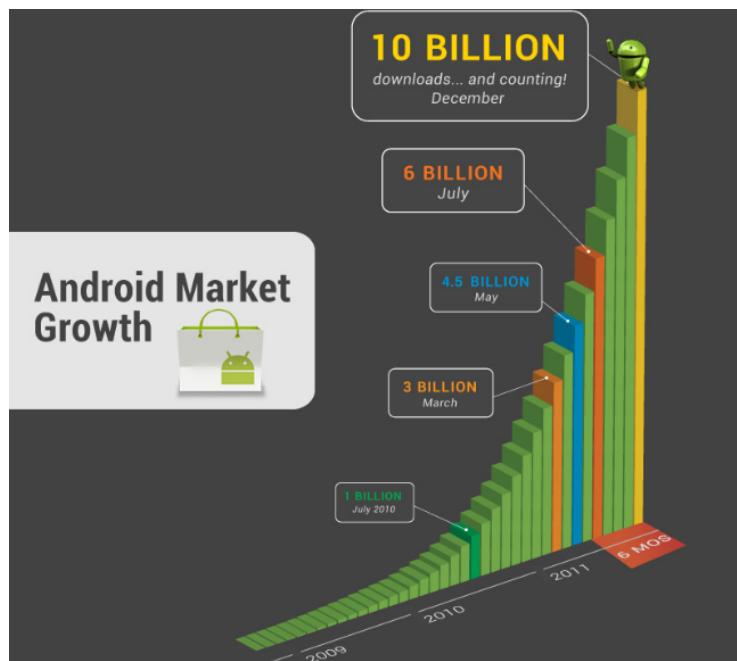
# Key Insight

- Mobile traffic classification\*
  - 1. Origin: e.g. app provider
  - 2. CDN+Cloud: e.g. Amazon AWS
  - 3. Third party: e.g. ads & analytics
- Key Idea: Use ads for app identification and analysis
- Observation regarding ads traffic:
  - Most Apps contains ads
  - Ads contains unique identifier
  - Most of the ads will auto-loaded when the application starts

\* Wei, X. etc., Profiledroid: Multi-layer profiling of android applications. MobiCom'2012

# Android Platform

- More difficult to identify apps on Android platform
- Growth rate of apps on Android is tremendous



# Ad Information in Android Apps

## ■ In manifest file for apps

## ■ In traffic

GET /getInfo.php?appid=523e4ae0705248b0b2b770a91d33d1c6&appver=300&client=2

### (a) HTTP Traffic of AdWhirl

GET /mads/gma?preqs=2&...&u\_w=320&msid=com.portugalemgrande.LiveClock&...

#### (b) HTTP Traffic of Google Ads

# Objective

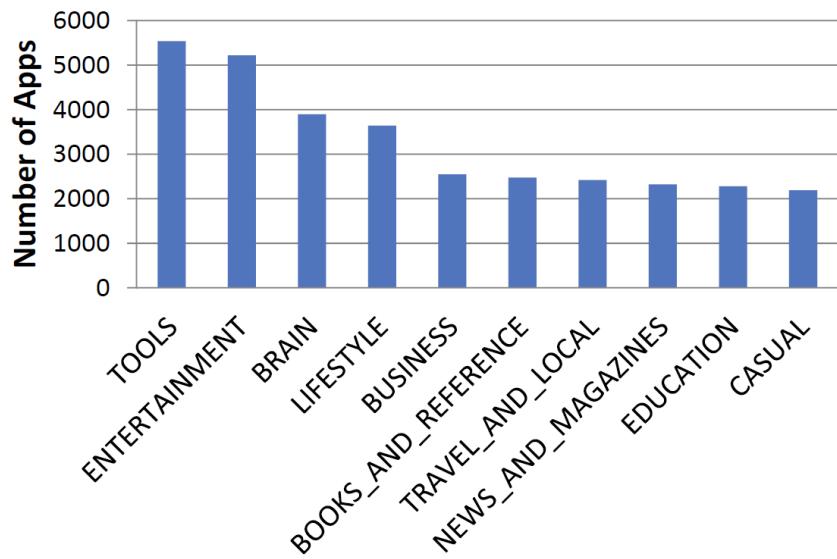
- Use ads information to:
  - Analyze apps in Google Play
  - Analyze app usage in real traffic
- Goal: To show the kinds of analysis enabled by ads

# Methodology for Google Play Analysis

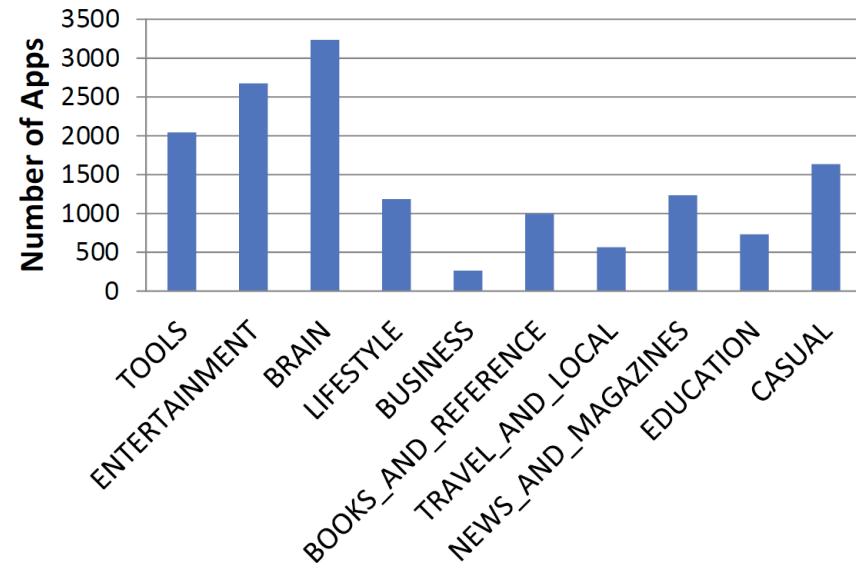
- Download apps, reverse engineering manifest files, query category, etc.
- Data set
  - 55k free apps from Google Play
    - 46K ask Internet Access Permission (**84%**)
      - 19K contains TOP 30 Ads Libraries\* (**41%**)
  - 2 days traffic from Tier-2 cellular service provider

\* Grace, M.C, etc., Unsafe exposure analysis of mobile in-app advertisement. WISEC'2012

# Top 10 Categories for Apps

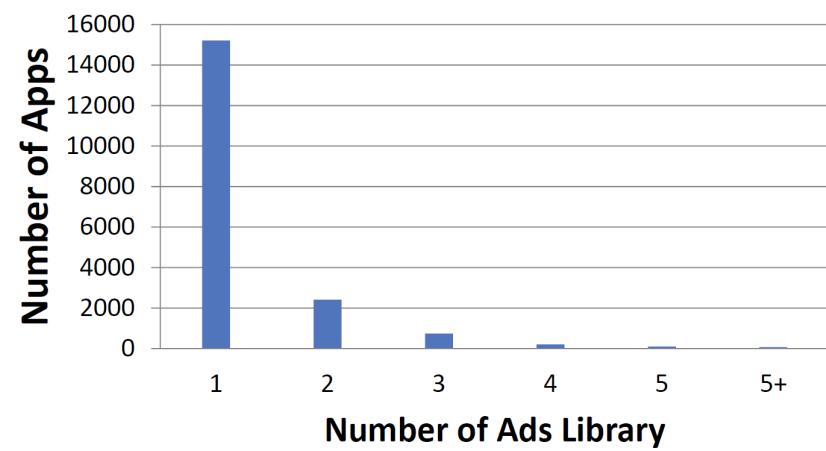


(a) All Apps

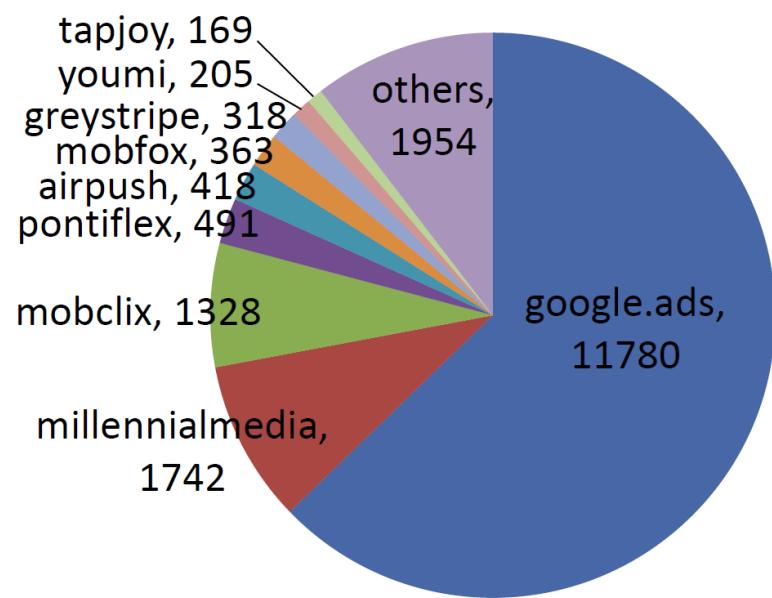


(b) Apps Containing Ads Library

# Ads Library Info

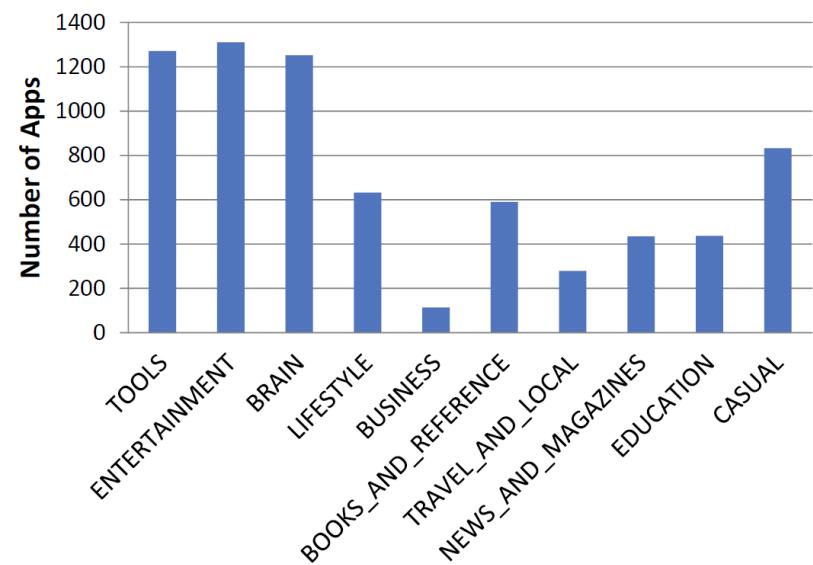


(a) Distribution of Ad Libraries in Apps

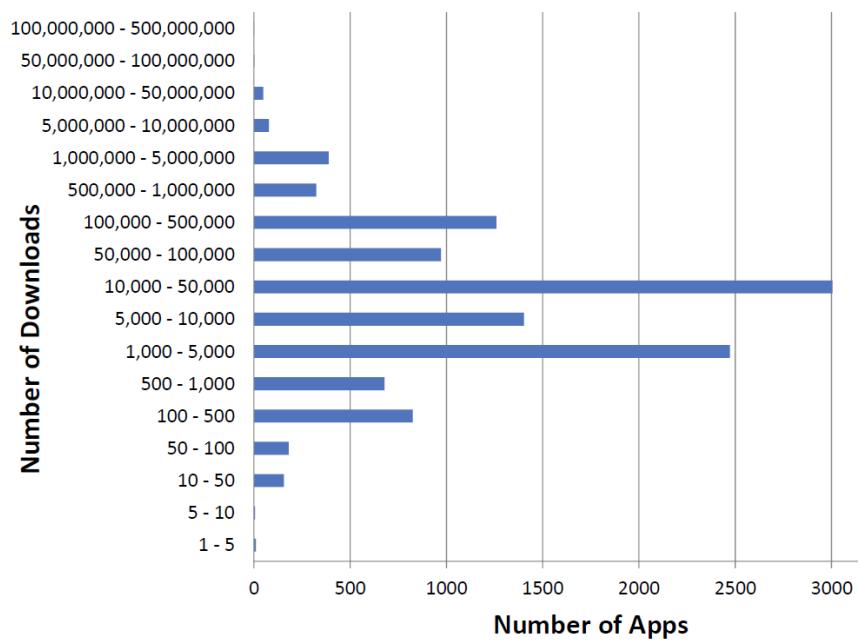


(b) Number of Apps Per Ad Library

# Google Ads Information

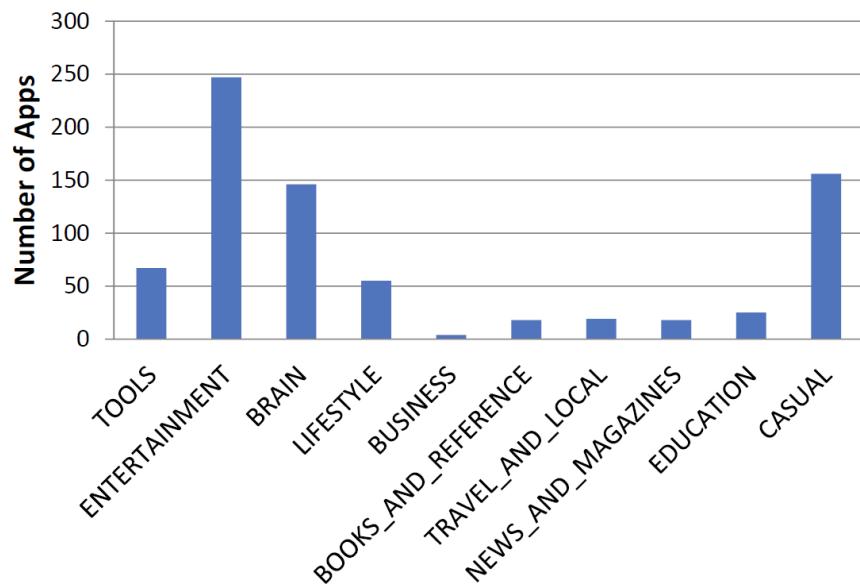


(a) Distribution of Categories

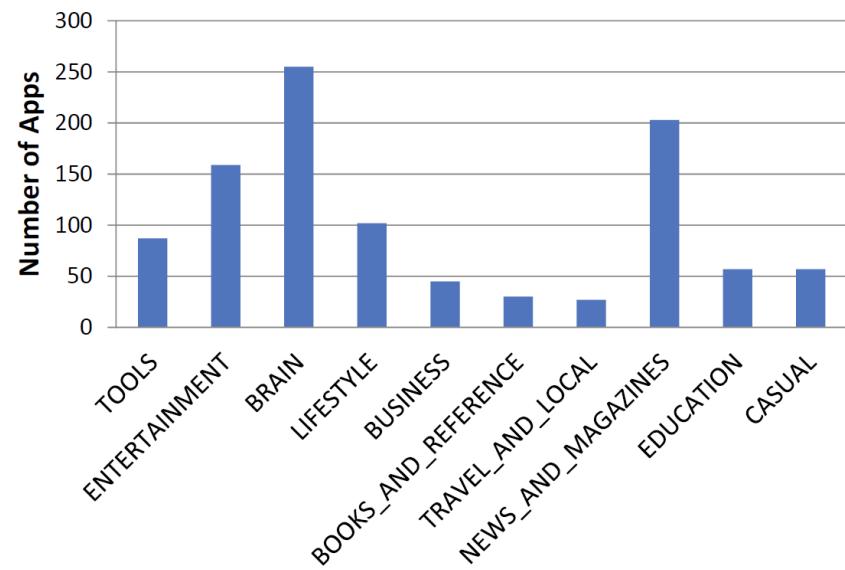


(b) Number of Downloads Per App

# Categories of Other Ad library

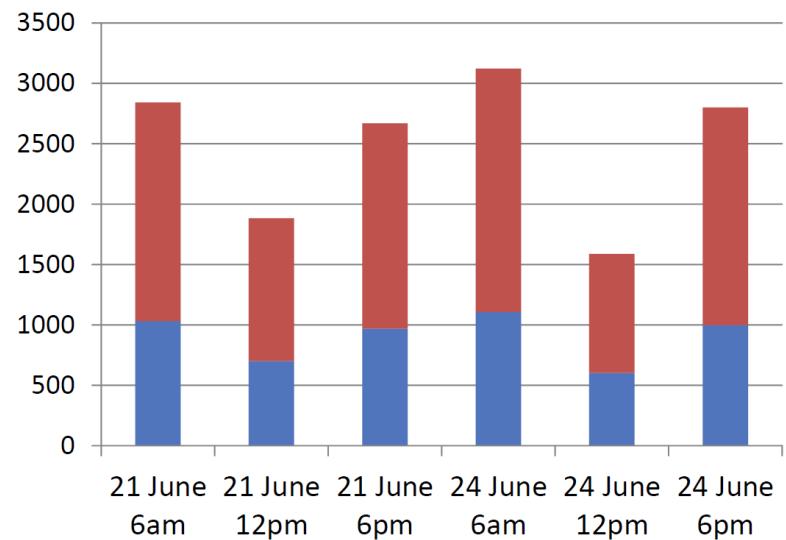


(a) Mobclix

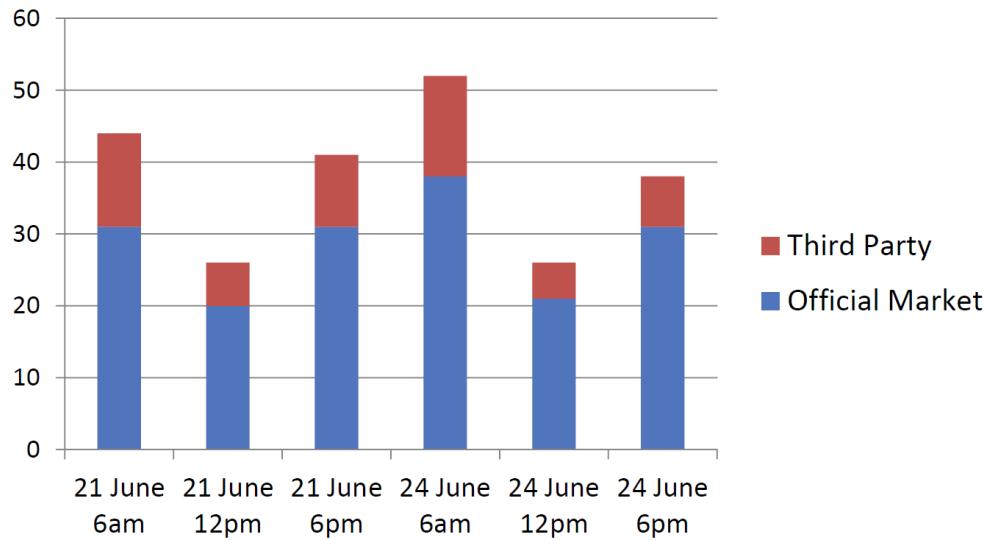


(b) Millenial Media

# Apps in Network Traffic

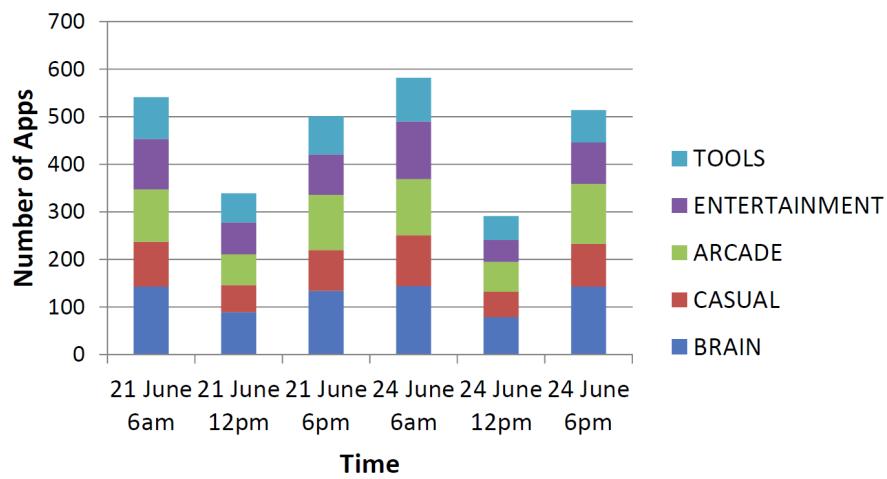


(a) Google Ads

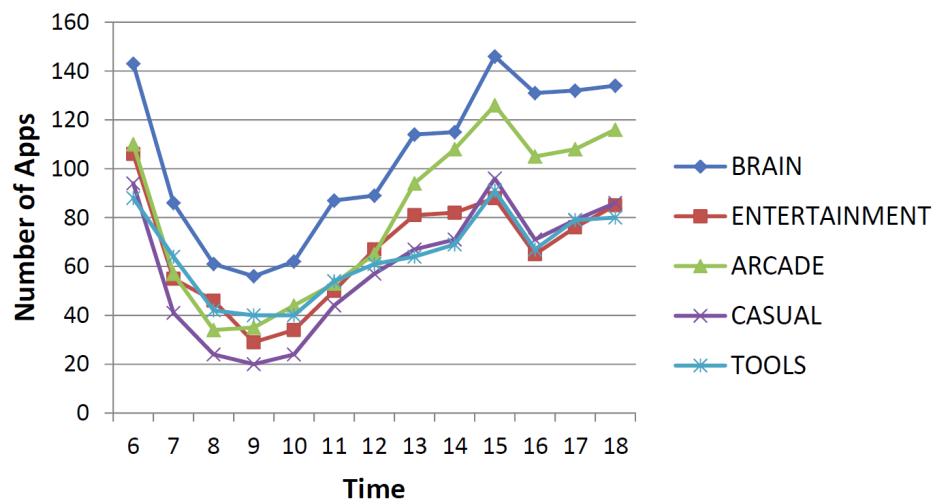


(b) Smaato

# Apps Containing Google Ads in Network Traffic



(a) Patterns Over 2 Days



(b) 12 Hours Pattern (21 June)

# Future Work

- Analyze ads containing developer ids
- Look at ads in paid apps
- Do more detailed analysis of the market as well as larger traffic sets
- Create large database of identifiers

# Our Related Work

- Mobile traffic classification
  - 1. Origin: e.g. app provider
  - 2. CDN+Cloud: e.g. Amazon AWS
  - 3. Third-party: e.g. ads & analytics
- We have developed techniques for identifying Origin+CDN+Cloud traffic in Infocom'13 paper

# Conclusion

- We presented a new direction for analyzing usage behavior of mobile apps based on ad flows.
- This technique provides a new perspective and it can be combine with other analysis methods.

**Thanks!**

**Q&A?**