

Identifying Diverse Usage Behaviors of Smartphone Apps

Presented by
Besim Avci

Outline

- ▶ Introduction
- ▶ Dataset
- ▶ Usage Patterns
 - ▶ Metrics
 - ▶ Popular apps
 - ▶ Spatial patterns
 - ▶ User patterns
 - ▶ Temporal patterns
 - ▶ Device patterns
- ▶ Implications
- ▶ Conclusion

Introduction

- ▶ 150K apps in Android
- ▶ 350K apps in AppStore
- ▶ Appearance of new products;
 - ▶ Tablets,
 - ▶ Smartwatches etc.
- ▶ Usage patterns has significant diversity in terms of:
 - ▶ Locality, time of the day, etc.

Introduction (cntd.)

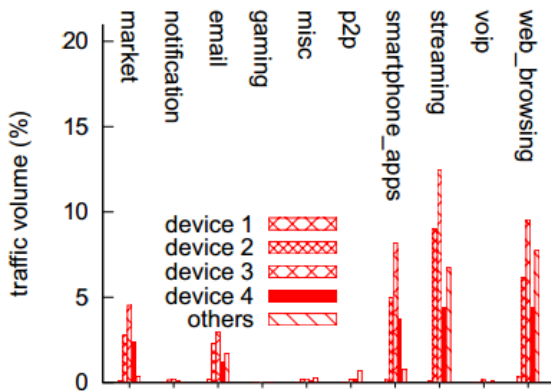
- ▶ IP-level networking traces.
 - ▶ Signatures based on HTTP headers
 - ▶ User-Agent header for app name extraction
- ▶ One week in August 2010
- ▶ Patterns are mined depending on different features such as locality, user mobility etc.

Dataset

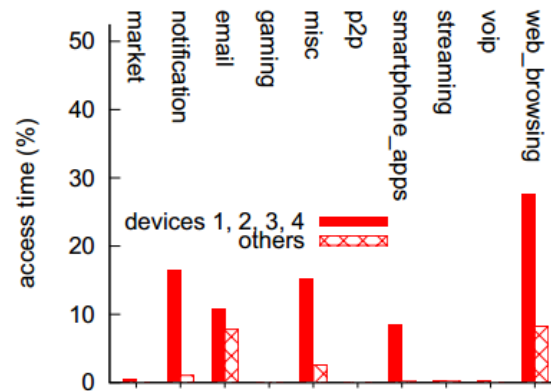
- ▶ Data from/to PDP context tunnels.
- ▶ Nationwide view of the traffic.
- ▶ Uniform sample of devices.
 - ▶ Uniformly random (reservoir sampling)
- ▶ Data is solely based on smartphone apps.
 - ▶ Youtube, browser traffic omitted.
- ▶ Four main features:
 - ▶ Traffic volume,
 - ▶ Access time,
 - ▶ Unique subscribers,
 - ▶ Locations.

Dataset (cntd.)

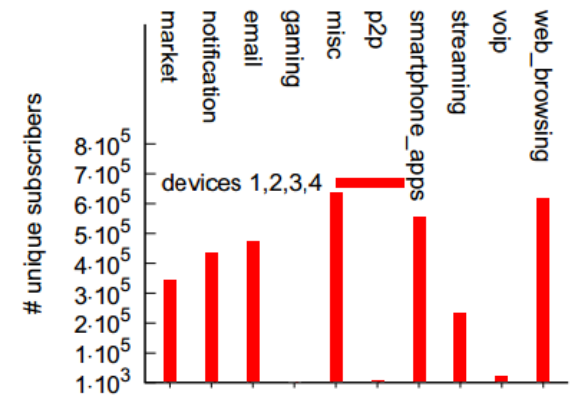
- ▶ 600K unique subscribers
- ▶ 22K distinct apps



a. distribution of traffic volume



b. distribution of access time

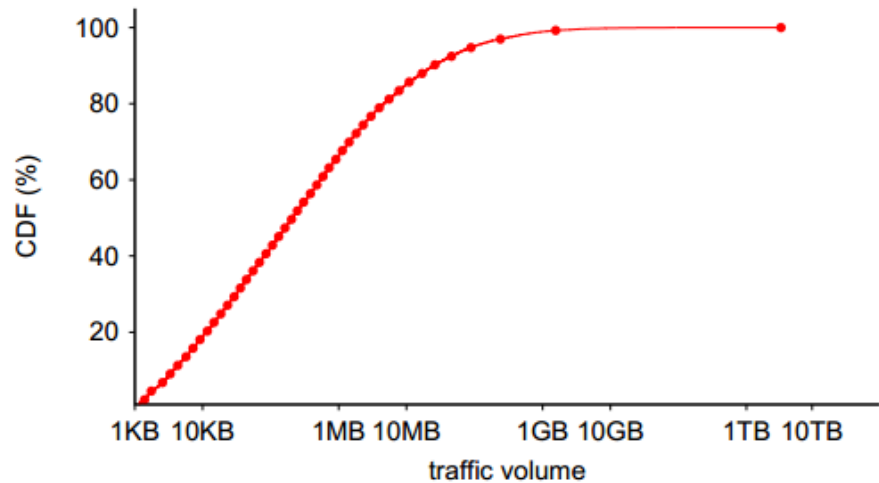


c. distribution of subscribers

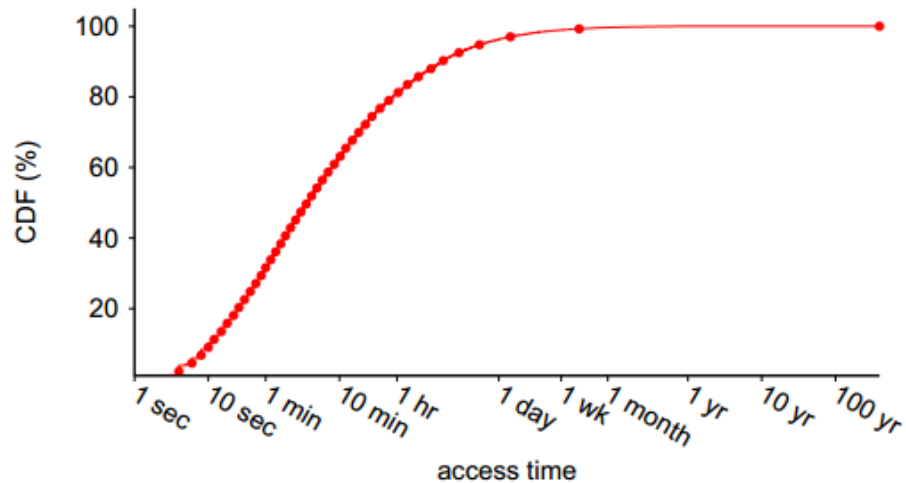
Usage Patterns – Metrics

- ▶ **Traffic volume;**
 - ▶ Number of bytes consumed by subs.
- ▶ **Number of unique subscribers.**
- ▶ **Network access;**
 - ▶ Total duration summed across all IP flows.

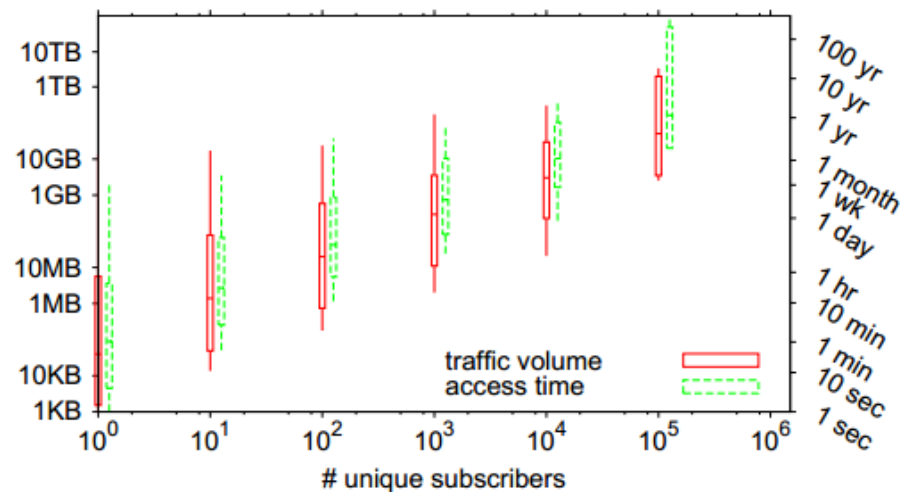
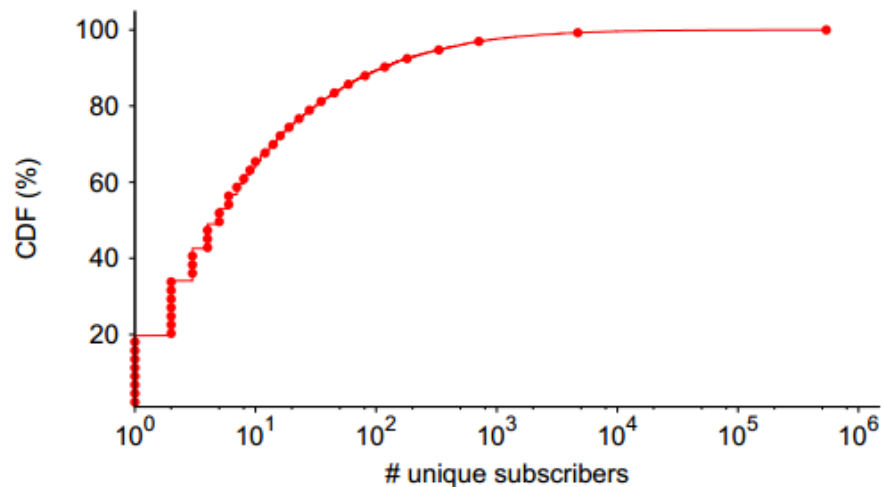
Metrics



a. traffic volume of smartphone apps



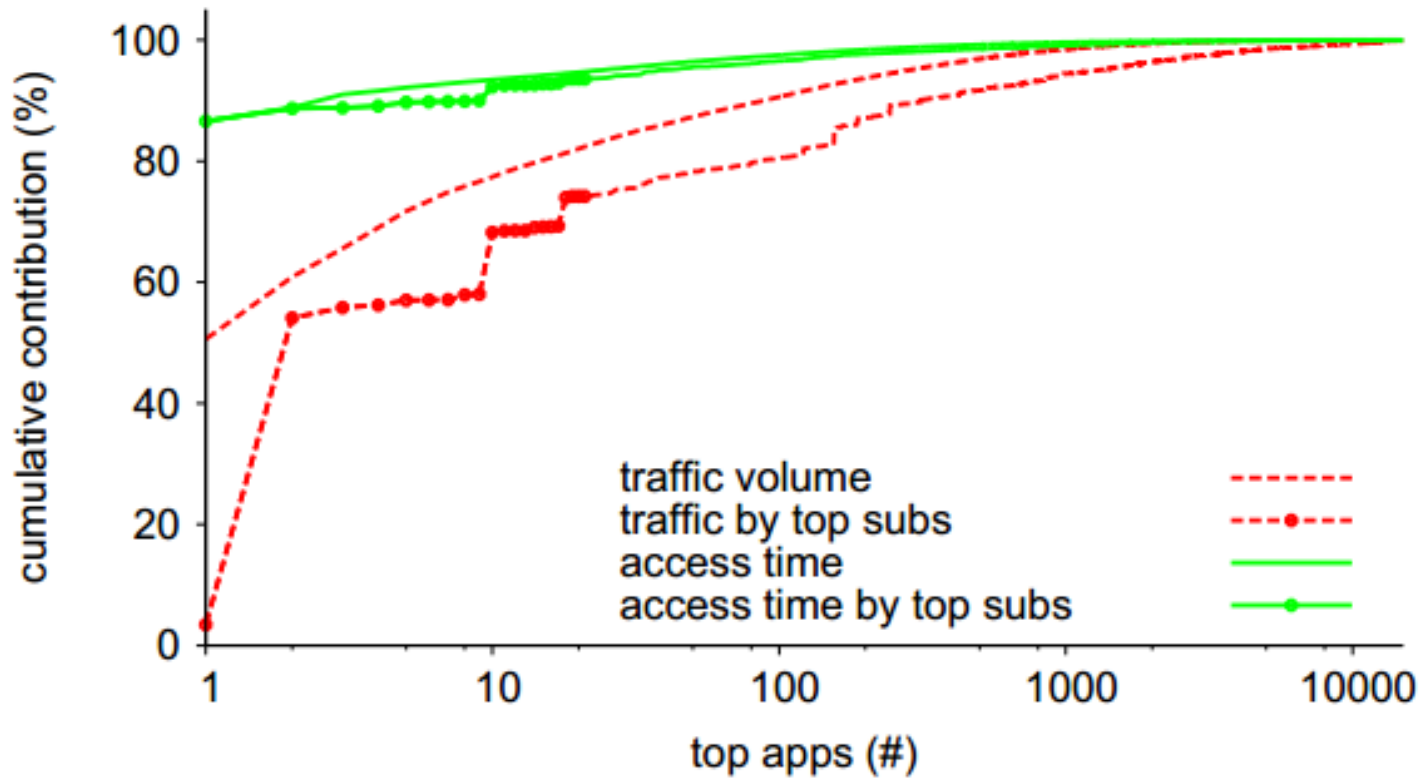
b. access time of smartphone apps



Usage Patterns – Popular apps

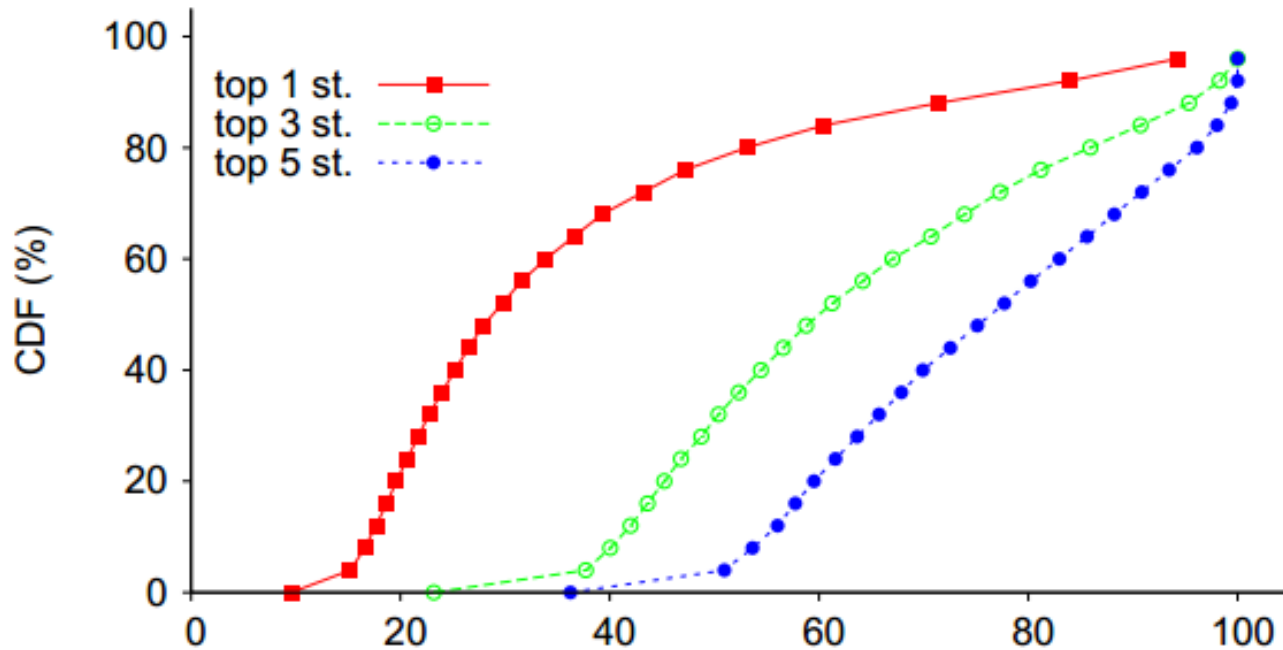
- ▶ %60 of apps have no more than 10 distinct subs.
- ▶ Eliminating unpopular apps while maintaining reliability.
- ▶ Top 1000 apps are attributed to:
 - ▶ %90 of total volume and access time.
- ▶ 1000th popular app has 471 unique subs.

Usage Patterns – Popular apps



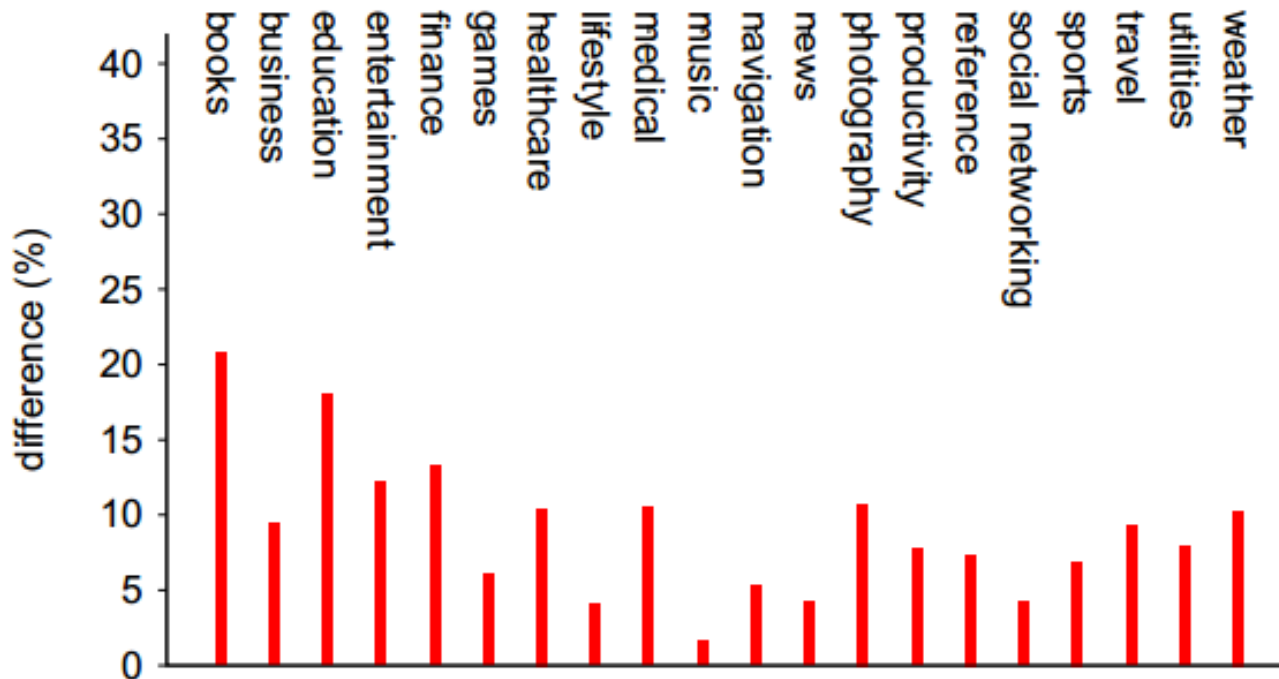
Usage Patterns – Spatial Patterns

- ▶ Total traffic is divided into states

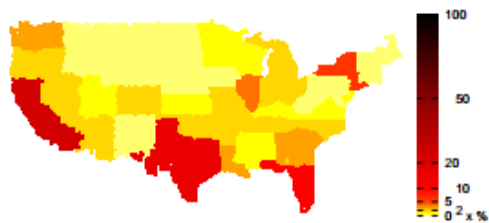


Usage Patterns – Spatial Patterns (cntd.)

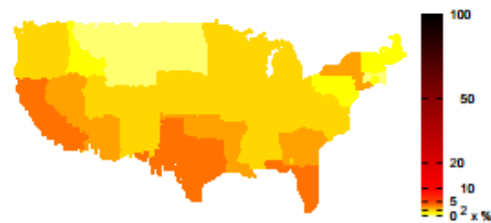
- ▶ Difference in the geographic usage of different genres



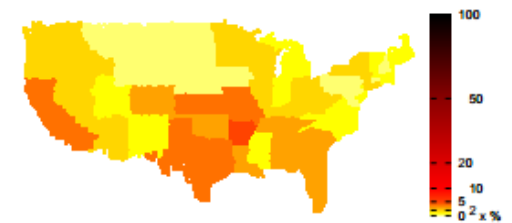
Usage Patterns – Spatial Patterns (cntd.)



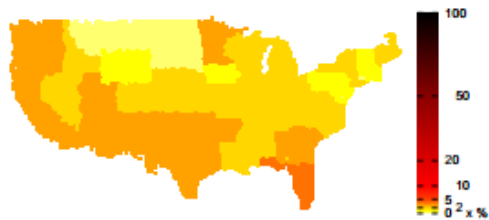
a. aggregate (not normalized)



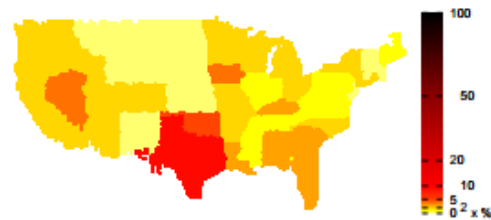
c. social networking



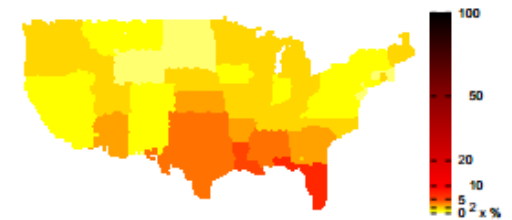
e. photography



b. aggregate normalized

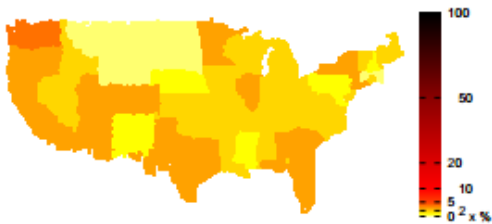


d. education

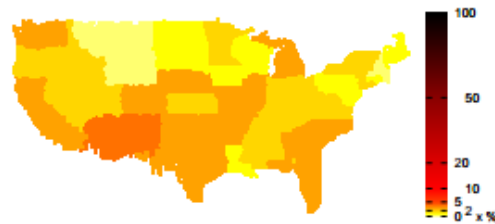


f. weather

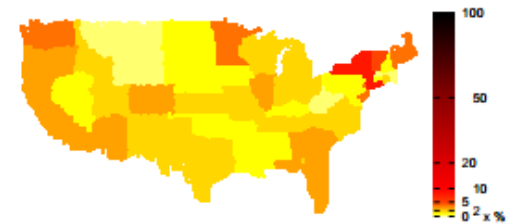
Usage Patterns – Spatial Patterns (cntd.)



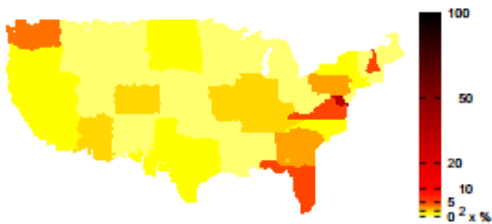
a. aggregate news apps



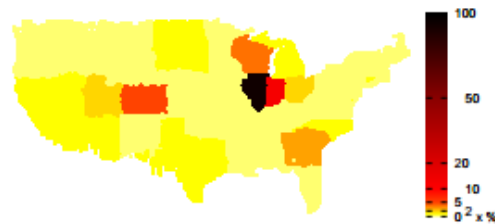
b. headquartered in ATLANTA¹



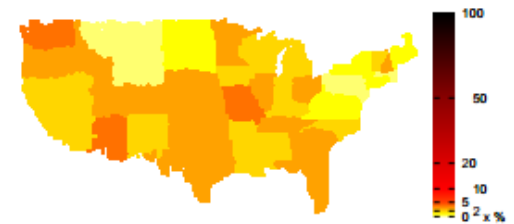
c. headquartered in NEW YORK²



d. headquartered in WASHINGTON, D.C.²



e. headquartered in CHICAGO²



f. headquartered in VIRGINIA¹

¹: names of these apps do not have any location indicated.

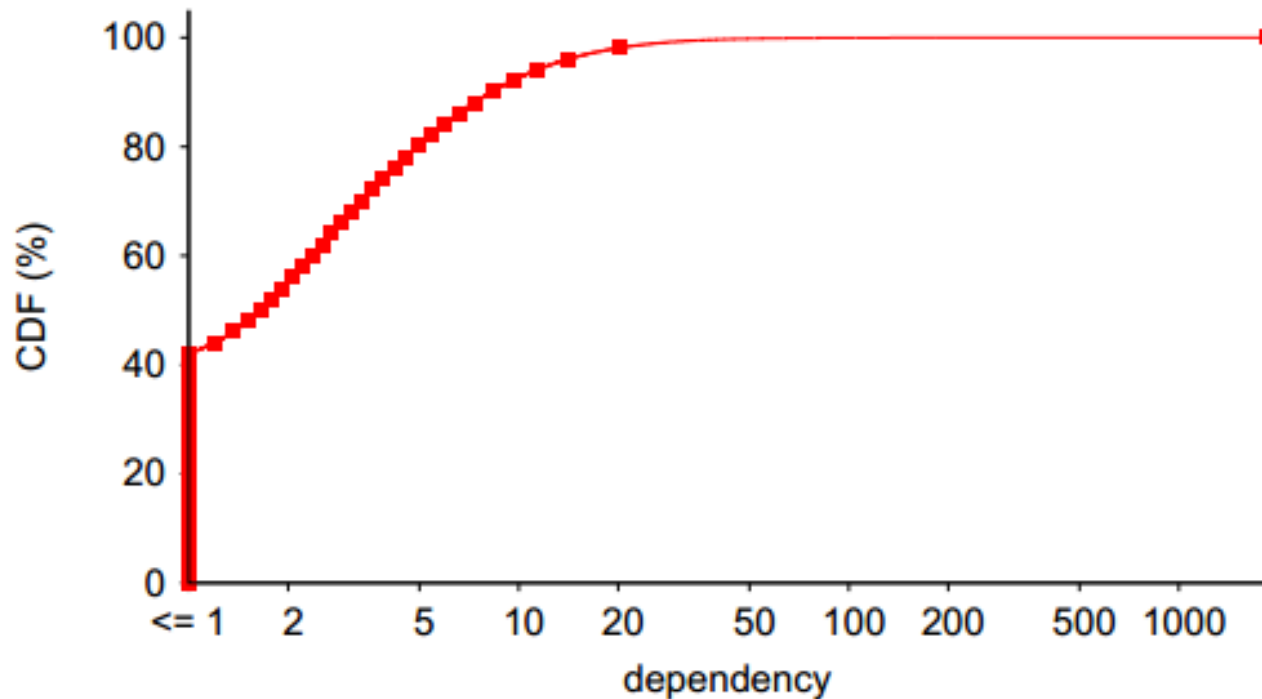
²: names of these apps include locations.

Usage Patterns – User Patterns

- ▶ Identifies correlation between apps.
- ▶ *Jaccard Similarity Coefficient* is used;
 - ▶ $\text{joint}(a,b) / \text{union}(a,b)$ for all 1000 popular apps.
- ▶ Not a significant correlation observed.

Usage Patterns – User Patterns

- ▶ Dependency ratio: $P(A.B)/P(A)P(B)$



- ▶ 254 pairs exceeding 100 ratio.

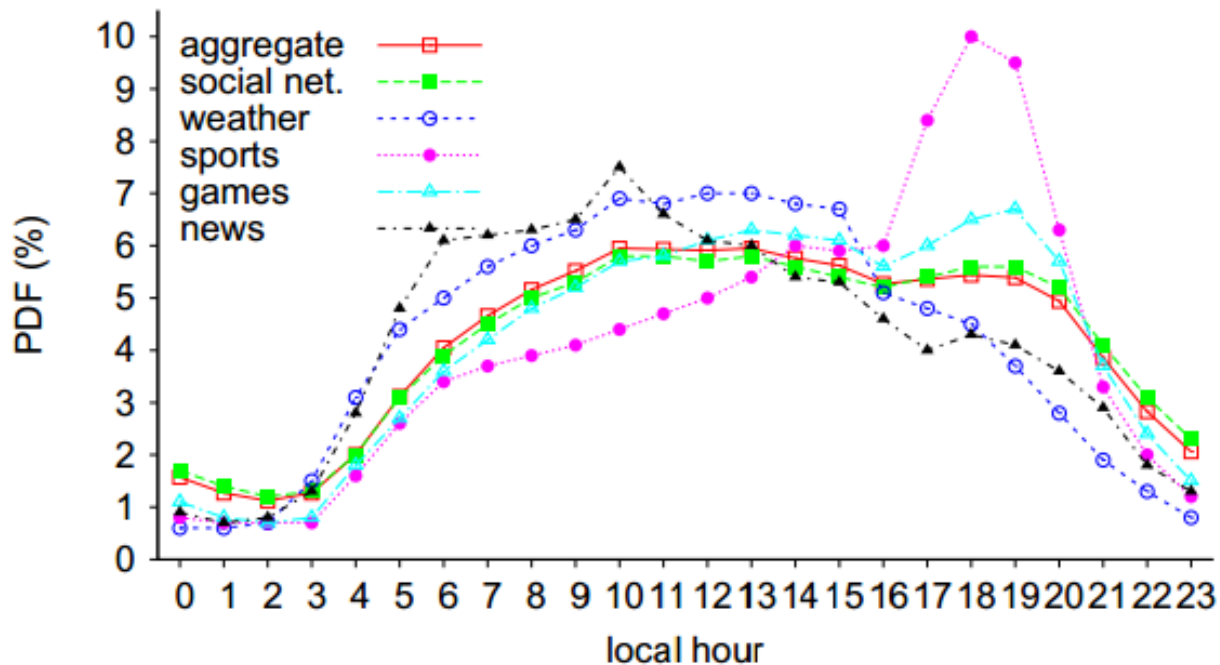
Usage Patterns – User Patterns

	books	business	education	entertainment	finance	games	healthcare	lifestyle	medical	music	navigation	news	photography	productivity	reference	social net.	sports	travel	utilities	weather
books	0	0	0	2	0	3	0	1	0	0	0	1	0	0	0	0	0	1	0	0
business	0	0	0	0	0	0	0	4	0	0	0	2	0	0	0	0	0	0	1	0
education	0	0	5	2	0	2	0	0	0	0	2	3	0	0	0	1	0	0	0	0
entertainment	2	0	2	26	0	26	1	16	0	5	1	16	4	3	3	8	4	4	7	0
finance	0	0	0	0	2	0	1	0	0	0	0	4	0	0	0	0	0	1	1	0
games	3	0	2	26	0	110	2	19	1	5	1	3	2	1	13	5	8	12	1	1
healthcare	0	0	0	1	1	2	1	2	1	0	1	8	1	0	0	1	0	0	0	0
lifestyle	1	4	0	16	0	19	2	30	0	5	0	12	0	1	7	6	4	10	6	0
medical	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0
music	0	0	0	5	0	5	0	5	0	6	0	5	0	1	1	11	3	0	3	0
navigation	0	0	2	1	0	1	1	0	0	0	2	1	0	1	0	1	0	4	0	1
news	1	2	3	16	4	3	8	12	1	5	1	77	1	3	2	12	8	7	4	4
photography	0	0	0	4	0	2	1	0	0	0	0	1	1	1	0	0	0	0	1	0
productivity	0	0	0	3	0	1	0	1	0	1	1	3	1	0	0	9	1	0	0	0
reference	0	0	0	3	0	13	0	7	0	1	0	2	0	0	1	0	1	2	0	1
social networking	0	0	1	8	0	5	1	6	0	11	1	12	0	9	0	32	4	0	3	0
sports	0	0	0	4	0	8	0	4	0	3	0	8	0	1	1	4	13	1	0	1
travel	1	0	0	4	1	12	0	10	0	0	4	7	0	0	2	0	1	9	2	0
utilities	0	1	0	7	1	1	0	6	0	3	0	4	1	0	0	3	0	2	2	0
weather	0	0	0	0	0	1	0	0	0	0	1	4	0	0	1	0	1	0	0	1



Usage Patterns – Temporal Patterns

- ▶ Apps show different usage patterns over the day.



Usage Patterns – Device Patterns

- ▶ **Intuition: faster devices enable;**
 - ▶ Longer sessions,
 - ▶ Faster downloads,
 - ▶ More interactivity.
- ▶ **Device 2 and 4;**
 - ▶ Roughly same access time.
 - ▶ 2 has lower traffic volume.
- ▶ **Device 3 and 4 are almost equal.**

Implications

- ▶ **Content providers**
 - ▶ Local app usage => local content distribution
- ▶ **Context-awareness**
 - ▶ App recommendation.
 - ▶ Ad personalization.
- ▶ **Network providers**
 - ▶ Network providers can leverage on temporal patterns
- ▶ **OS vendors and app designers**
 - ▶ Prefetching

Conclusion

- ▶ Comprehensive diverse usage pattern analysis.
- ▶ %20 of apps are local.
- ▶ Slight correlation between different app usage.
- ▶ Diurnal patterns are different throughout the day.
- ▶ Social networking and gaming apps are related with user mobility.

Questions / Discussion
