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
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

![Graph showing differences in usage patterns for various genres.]
Usage Patterns – Spatial Patterns (cntd.)

- a. aggregate (not normalized)
- b. aggregate normalized
- c. social networking
- d. education
- e. photography
- f. weather
Usage Patterns – Spatial Patterns (cntd.)

a. aggregate news apps
b. headquartered in ATLANTA\(^1\)
c. headquartered in NEW YORK\(^2\)
d. headquartered in WASHINGTON, D.C.\(^2\)
e. headquartered in CHICAGO\(^2\)
f. headquartered in VIRGINIA\(^1\)

\(^1\): 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

<table>
<thead>
<tr>
<th></th>
<th>books</th>
<th>business</th>
<th>education</th>
<th>entertainment</th>
<th>finance</th>
<th>games</th>
<th>healthcare</th>
<th>lifestyle</th>
<th>medical</th>
<th>music</th>
<th>navigation</th>
<th>news</th>
<th>photography</th>
<th>productivity</th>
<th>reference</th>
<th>social net</th>
<th>sports</th>
<th>travel</th>
<th>utilities</th>
<th>weather</th>
</tr>
</thead>
<tbody>
<tr>
<td>books</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>business</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>education</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>entertainment</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>26</td>
<td>0</td>
<td>26</td>
<td>1</td>
<td>16</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>16</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>8</td>
<td>4</td>
<td>4</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>finance</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>games</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>26</td>
<td>0</td>
<td>110</td>
<td>2</td>
<td>19</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>13</td>
<td>5</td>
<td>8</td>
<td>12</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>healthcare</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>8</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>lifestyle</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>16</td>
<td>0</td>
<td>19</td>
<td>2</td>
<td>30</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>6</td>
<td>4</td>
<td>10</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>medical</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>music</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>11</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>navigation</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>news</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>16</td>
<td>4</td>
<td>3</td>
<td>8</td>
<td>12</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>12</td>
<td>8</td>
<td>7</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>photography</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>productivity</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>reference</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>13</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>social networking</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>8</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>6</td>
<td>0</td>
<td>11</td>
<td>1</td>
<td>12</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>32</td>
<td>4</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>sports</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>13</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>travel</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>12</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>9</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>utilities</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>weather</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
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