“Community Imaging” techniques:

Understanding people through their use of technology

Prof. Vassilis Kostakos
Director, Center for Ubiquitous Computing
University of Oulu

24 August 2016
Summer School on Smart Cities, Vienna, Austria
Advanced evaluation techniques

• How to measure performance of
  – large numbers of people
  – from remote settings (e.g. online, distributed)
  – using large datasets

• We want to determine
  – What happened / what did people do?
  – What does it mean?
  – What does it tell us about people?

• “Community Imaging”
  – Understanding people through their use of technology.
Smartphones

Online behaviour

Human Computation & Crowdsourcing
Do phones create habits?

We built an app and deployed to an appstore. The app collects data.
Differences in launching

Y-axis is LOGARITHMIC!
Histograms per user

Device Revisititation Curve for User cb980575

Frequency

0 0.05 0.1 0.15 0.2 0.25

1min 2min 4min 8min 16min 1/2hr 1hr 2hr 4hr 8hr 16hr 1.3days 2.6days 5.2days >10days
Patterns emerge (users)
Histograms per app

Revisititation Curve for Whatsapp

Frequency

1min 2min 4min 8min 16min 1/2hr 1hr 2hr 4hr 8hr 16hr 1.3days 2.6days 5.2days >10days
Patterns emerge (apps)

<table>
<thead>
<tr>
<th>Cluster Label</th>
<th>Description</th>
<th>Centroid Revisitation Curve</th>
<th>Example Apps</th>
<th>Cluster Size (# Apps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>Fast</td>
<td></td>
<td>Google Play Store, Facebook Messenger, InoReader, Chrome Beta, BlackBerry Messenger, Reddit, Okcupid</td>
<td>32 (13%)</td>
</tr>
<tr>
<td>F2</td>
<td>Fast</td>
<td></td>
<td>Chrome, Whatsapp, Facebook, Google Hangouts, SMS/MMS, Viber, Youtube, Contacts+, Google Maps, Firefox, Spotify, Skype, Snapchat, Xperia Conversations, Line, Reddit News, Telegram Messenger, Music, Falcon Pro</td>
<td>82 (33%)</td>
</tr>
<tr>
<td>M1</td>
<td>Medium</td>
<td></td>
<td>Phone, Gmail, Contacts, Email, Dialer, Clash of Clans, Instagram, Outlook, Yahoo Mail, Opera Browser</td>
<td>47 (19%)</td>
</tr>
<tr>
<td>S1</td>
<td>Slow</td>
<td></td>
<td>Gallery 3D, Calendar, Camera, Twitter, Calculator, Clean Master (Speed Booster) Runkeeper Pro, Flipboard, Google Play Services, Mobile Bank, Mobile Weather, Flickr, Google Doc Editor, Tumblr, Quick Office, Google Translate</td>
<td>30 (12%)</td>
</tr>
<tr>
<td>S2</td>
<td>Slow</td>
<td></td>
<td>Settings, Desk Clock, Organiser, Tinder, Plants vs. Zombies 2, Clash of Lords, Titanium Backup, Hot or Not, Control Panel, Candy Crush Saga, Castle Clash</td>
<td>40 (16%)</td>
</tr>
<tr>
<td>H1</td>
<td>Hybrid</td>
<td></td>
<td>Evernote, Google+, Google Docs, MusicBox, Adobe Reader, 9gag, Video Player, Meo Remote, Waze, Dictionary, Opera Mini</td>
<td>21 (8%)</td>
</tr>
</tbody>
</table>
Compare to Web revisitaton (2000’s)

<table>
<thead>
<tr>
<th>Cluster Group</th>
<th>Centroid Curve</th>
<th>Description</th>
<th>Corresponding cluster group descriptions from Adar et al. [1]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fast (F1,F2)</td>
<td></td>
<td>Instant</td>
<td>Hub &amp; Spoke, Shopping &amp; Reference, Auto refresh, Fast</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Messaging,</td>
<td>monitoring, Pornography &amp; Spam.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Browser, Social Media</td>
<td></td>
</tr>
<tr>
<td>Medium (M1)</td>
<td></td>
<td>Email and Phone Communication</td>
<td>Popular homepages, Communication, .edu domain, browser homepages.</td>
</tr>
<tr>
<td>Slow (S1,S2)</td>
<td></td>
<td>Utilities, Multimedia, Health and Fitness, Games, Dating, Phone Settings</td>
<td>Entry pages, Weekend activity, Search engines used for Revisitation, Child-oriented content, Software updates</td>
</tr>
<tr>
<td>Hybrid (H1)</td>
<td></td>
<td>Documents, Notes, Video, Satnav</td>
<td>Popular but infrequently used, Entertainment &amp; Hobbies, Combined Fast &amp; Slow.</td>
</tr>
</tbody>
</table>
Similar patterns.
What does this mean?

Do phones create habits?

Technology A

Technology B

Behaviour
Online behaviour is a proxy

Can online behaviour be used as a proxy for studying urban mobility?
WiFi access points
Across oulu – 3 years

Amount of pedestrians detected across the city

Upward trend. Needs normalisation.
Notice that the y-axis is now 0~1
Each location has a pattern
Google web searches also have a pattern.

Dataset:

- Frequency: 31, 78, 65, ..., 83
“Google Trends” lets us match the patterns
Which keyword patterns match to a location patterns?

<table>
<thead>
<tr>
<th>University</th>
<th>r-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helka (Helsinki university library)</td>
<td>0.914</td>
</tr>
<tr>
<td>google scholar</td>
<td>0.895</td>
</tr>
<tr>
<td>scholar</td>
<td>0.894</td>
</tr>
<tr>
<td>tutkimus (research)</td>
<td>0.893</td>
</tr>
<tr>
<td>learning</td>
<td>0.887</td>
</tr>
<tr>
<td>optima oulu (student environment)</td>
<td>0.878</td>
</tr>
<tr>
<td>funktio (function)</td>
<td>0.876</td>
</tr>
<tr>
<td>luento (lecture)</td>
<td>0.874</td>
</tr>
<tr>
<td>development</td>
<td>0.872</td>
</tr>
<tr>
<td>nelli (university e-library portal)</td>
<td>0.871</td>
</tr>
</tbody>
</table>
Which keyword patterns match to a location patterns?

<table>
<thead>
<tr>
<th>Ice hockey hall</th>
<th>r-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>sjl (Finnish ice hockey union)</td>
<td>0.726</td>
</tr>
<tr>
<td>finnhockey</td>
<td>0.724</td>
</tr>
<tr>
<td>keilahalli (bowling hall)</td>
<td>0.722</td>
</tr>
<tr>
<td>kiekko kaleva (hockey in local newspaper)</td>
<td>0.716</td>
</tr>
<tr>
<td>sm-liiga nelonen (tv-channel with hockey)</td>
<td>0.715</td>
</tr>
<tr>
<td>nelonen sm (same as above)</td>
<td>0.713</td>
</tr>
<tr>
<td>jääkiekkoliitto (ice hockey union)</td>
<td>0.711</td>
</tr>
<tr>
<td>nelonen sm liiga (tv-channel, ice hockey)</td>
<td>0.707</td>
</tr>
<tr>
<td>lihapata (meat stew)</td>
<td>0.699</td>
</tr>
<tr>
<td>finhockey</td>
<td>0.699</td>
</tr>
</tbody>
</table>
Which keyword patterns match to a location patterns?

<table>
<thead>
<tr>
<th>High school</th>
<th>r-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>wilma kempele (student environment)</td>
<td>0.795</td>
</tr>
<tr>
<td>wilma kiiminki (student environment)</td>
<td>0.791</td>
</tr>
<tr>
<td>helmi (student management system)</td>
<td>0.781</td>
</tr>
<tr>
<td>wilma oulu (student environment)</td>
<td>0.771</td>
</tr>
<tr>
<td>edu (news and info on education)</td>
<td>0.763</td>
</tr>
<tr>
<td>wilma kuusamo (student environment)</td>
<td>0.758</td>
</tr>
<tr>
<td>pedanet (support for online learning)</td>
<td>0.747</td>
</tr>
<tr>
<td>wilma raahe (student environment)</td>
<td>0.744</td>
</tr>
<tr>
<td>wilma kemi (student environment)</td>
<td>0.736</td>
</tr>
<tr>
<td>varoitusmerkit (warning signs)</td>
<td>0.726</td>
</tr>
</tbody>
</table>
Similar patterns.
What does this mean?
Human Computation & Crowdsourcing

Can we use humans to solve problems that are hard for computers?

<table>
<thead>
<tr>
<th>Feature</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reliable memory</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Understands context</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Reliable calculations</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Fast calculations</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Sensory translation</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Can do research</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>
Offer decision-support to humans (not answers)
Which car should I buy?

<table>
<thead>
<tr>
<th>Price</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Volvo</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Things to evaluate

• Is a solution “good”? (BMW, Mercedes, Donkey)
• Is a criterion “good”? (Price, Safety, Swag)
• Is the decision-support system “good”? 
Available solutions in order of your preferences:

1. United Kingdom
   - UK
2. United States
   - US
3. Russia
   - Russia
4. India
   - India

Economic equality
This basically means the gap between rich and poor! If people are economically equal, choose 10 (very equal). If you think people are not equal at all economically, choose 0!

Weather
Generally, how pleasant is the weather. Choose 0 for a really bad weather, 10 for great weather all year round!
Smart City research

• Brings together
  – Large numbers of people/users
  – Large volumes of data
  – Large number of disciplines

• Provides a natural
  – Metaphor for computation (4th wave of computing)
  – Application domain for interesting problems/solutions
The end

Prof. Vassilis Kostakos
Director, Center for Ubiquitous Computing
University of Oulu

Vassilis.kostakos@oulu.fi
http://ubicomp.oulu.fi