UrbanMatch – linking and improving Smart Cities Data

Irene Celino, Simone Contessa, Marta Corubolo, Daniele Dell’Aglio, Emanuele Della Valle, Stefano Fumeo and Thorsten Krüger

CEFRIEL – Politecnico di Milano – SIEMENS
Three converging trends

Urban Computing and Location-based Services
- citizens as sensors, check-in logging, mobile apps

Linked Data and Semantic Web
- open/gov data, structured data, social networks, tourism data and recommendations

Urban Games
- collecting data, cleaning data, engaging the user, supporting the user while entertaining him/her

Games with a Purpose and Crowdsourcing
Research hypothesis

- Urban Games:
  - to **consume**, **create** and **assess the quality** of Smart Cities-related **Linked Data**
  - via a **Human Computation** approach
  - for users in **mobility** with **smart phone** devices

- Traditional Human Computation approaches are based on users' **domain knowledge**…
  …while Urban Games are based on and aim at exploiting "on site" users' **experience knowledge**

- **UrbanMatch** is our first prototype of Urban Games: a GWAP for iPhone aimed to create links between urban POIs and their most representative photos
UrbanMatch: input and output data

Places & POIs from OpenStreetMap

Manual Selection of linked photos

Trusted sources

UrbanMatch server

Uncertain sources

UrbanMatch clients

Trusted links:
<POI> foaf:depiction <photo>
UrbanMatch: gameplay

Video at: http://youtu.be/FYBXxwAL0NY
UrbanMatch: achieving the purpose

Trusted links are constituted by correct POI-photo couples (i.e. the correlation between the photo and the depicted POI was verified)

Candidate links are those returned by the uncertain sources (e.g. Duomo-photoXYZ with photoXYZ retrieved via a Flickr API geo-search for Duomo)

Players’ coupling decisions are collected and then post-processed

When a POI-photo link is assessed to be correct, it becomes trusted

Incorrect link discarded

When a POI-photo link is assessed to be incorrect, it is discarded and never again proposed to players
UrbanMatch: links and confidence value

Trusted sources
- manually selected photos

Uncertain sources
- Wikimedia Commons
- flickr

Candidate links
- POI ↔ POI
- ~ 200
- ~ 37,000

Confidence value
- initial value
- value changes on players' choices
- trustable when > upper threshold
- incorrect when < lower threshold

60%
40%
UrbanMatch evaluation (1/2)

Data Quality Metrics

- **Evaluation data**
  - 54 unique players, 290 games (781 game levels)
  - 2006 input links, 1284 assessed (trustable/incorrect)
  - upper threshold 70%, lower threshold 20%

- **Completeness**
  - def: assessed links (trusted+incorrect) / all input links
  - intuitively: *game ability to assess input links*
  - globally from 1.54% (only manually-checked links) to **4.98%**
  - trusted links have tripled (incremented by **322%**)

- **Accuracy**
  - def: correctly assessed links (w/ FP+FN) / all assessed links
  - intuitively: *game ability to correctly assess input links*
  - **99.4%** (only 4 False Positive and 8 False Negative links)
UrbanMatch evaluation (2/2)

User-based evaluation: actual "engagement"

- Based on game evaluation literature and integrated with our research-specific questions
- Findings:
Conclusions and next steps

- Evaluation results seem to prove our research hypothesis
  - *Purpose "hiding"* is key to GWAP success
  - Publication of generated links as **Linked Open Data**
  - Further evaluation on **GWAP-specific metrics**

- UrbanMatch Milano soon to be followed by its "brother" app **UrbanMatch Munich 😊**
  - **Comparative evaluation** of UrbanMatch vs. manual assessment via domain experts

- A more complex Urban Game under development to prove our research hypothesis on different kinds of Smart Cities-related Linked Data
  - **Urbanopoly** game coming soon from PlanetData project...
Thanks for your attention! Any question?

Irene Celino – CEFRIEL, ICT Institute Politecnico di Milano
email: Irene.Celino@cefriel.it – web: http://swa.cefriel.it