Capturing Urban Context: Its Limitations and Possibilities

Hideyuki Tokuda
Keio University

hxt@sfc.keio.ac.jp
http://www.ht.sfc.keio.ac.jp/~hxt
Keio University

- Private comprehensive education institution
- 10 undergraduate faculties, 14 graduate schools and over 20 research centers
- 6 campuses across the greater Tokyo area
- University hospital, schools from elementary to high school levels

Mita Campus
Hiyoshi Campus
Yagami Campus

Shinanomachi Campus
Shonan Fujisawa Campus
Shiba-Kyoritsu Campus
Hide Tokuda Lab., Keio University

- Smart Spaces
- Ubiquitous Service Platform (HW/SW)
- HCI
- Sensors and Dependable Ubiquitous Nodes
- MANET and Heterogeneous MANET

MANET, Heterogeneous MANET
Ubiquitous Network Browser

SS Lab., Smart Living, uPlatea

Service Platform: Smart Furniture, uTexture

uPhoto, @Reader, uTexture, InfoRod
photo-based Interaction, Texture-based Interaction
Multi-display Interaction
Recovering from 3.11 Disaster
Thank you for Supporting and Praying for Japan
Recovering from 3.11 Disaster
Recovering from 3.11 Disaster
(by ABC News)
Recovering from 3.11 Disaster

(by ABC News)
Urban Context Capturing for Disaster Recovery
Honda & Google Collaboration
Passage Route Map (by HONDA)
Yamamoto-town’s (山元町) Geiger Counter Map by a community bus (by Ubiteq)
Outline

• A bit of History
  • March 11 Disaster and Urban Context

• Ubiquitous Services
  • What are Ubiquitous Services

• Capturing Urban Context
  • Limitations and Possibilities

• Place-triggered Geotagged Tweets Analysis
  • Case Study

• Summary
What are Ubiquitous Services
Ubiquitous Services

• Service type: any3 vs. only3
  • At anytime, anywhere, for anyone
  • Only now, only here, only for me/us

• Ubiquitous Services
  • Context-aware Services
  • Context-aware Health Care
  • Context-aware Information Services
    • Presence Service for your friends (Real-Space SNS)
    • Push-type information service
  • Mobile e-Commerce with RFID tags
  • and more…
Classification of Ubiquitous Services

Cyber space

Real space

Body
Room
Bldg
Outdoor (on)
Outdoor (off)

CPC in Small
CPC in Large

14
Cyber-Physical Coupling

Coupling = Sensing + Processing + Actuation
Outline

• A bit of History
  • March 11 Disaster and Urban Context

• Ubiquitous Services
  • What are Ubiquitous Services

• Capturing Urban Context
  • Limitations and Possibilities

• Place-triggered Geotagged Tweets Analysis
  • Case Study

• Summary
Urban Context Capturing
Limitations: Useful and Harmful
Context-awareness in Ubiquitous Services

- **Personal Context**
  - e.g. sleeping, eating, standing, running, walking, moving, stopping, ... etc.

- **Group Context**
  - e.g. group meeting, discussion, sports, ad hoc chatting, lecture, ... etc.

- **Urban Context**
  - e.g. City-wide context
    - blackout area, rain, hot spots, traffic jam, train accident, social events, ... etc.

- **Nation-wide Context**
  - e.g. population distribution, power distribution, ... etc
My Sports Pals (www.mysportspals.com)
SkyHook’s SpotRank
(http://www.skyhookwireless.com)

SpotRank In Action
NTT DoCoMo (Mobile Space Statistics(2010))
Mobile Space Statistics (2010)
Limitations: Useful and Harmful

- Anonymity Set and Privacy Enhancement
  - Visualization Problem: Density vs. Actual Data
    - Sport Pals: No cycling path
  - Small Anonymity Set Problem
- Data Accuracy
  - Mobile Statistics/Skyhook
- Real-Time Sensing/Processing/Actuation
  - Mobile Statistics
- Target Users
  - City Planner vs. Individual
Urban Context Capturing Possibilities
Weather News: Hybrid Sensing Model

Defense forces for Guerrilla Thunderstorm
Weather News: Better Prediction

・昨年の「ゲリラ雷雨メール」実績

<table>
<thead>
<tr>
<th>各都府県</th>
<th>ゲリラ雷雨 発生数</th>
<th>ゲリラ雷雨 捕捉率</th>
<th>ゲリラ雷雨メール 事前送信時間</th>
</tr>
</thead>
<tbody>
<tr>
<td>東京都</td>
<td>172回</td>
<td>76.7%</td>
<td>平均38分前</td>
</tr>
<tr>
<td>大阪府</td>
<td>128回</td>
<td>62.5%</td>
<td>平均8分前</td>
</tr>
<tr>
<td>愛知県</td>
<td>172回</td>
<td>63.9%</td>
<td>平均19分前</td>
</tr>
</tbody>
</table>
Electricity Consumption: Improving Citizen’s awareness
ユビキタス技術で環境を感じよう！

Airy Notes

Visualization of Shinjuku City Park with Airy Notes
Heathlandscape (www.healthlandscape.org)
Possibilities: Big Potentials

- Improved Data Accuracy and Prediction
  - Use of Physical Sensors with Human Sensors
  - Hybrid Sensing Model with Crowdsourcing

- Human as a Sensor
  - Crowdsourcing with Gamification

- Tweet as a Sensor
  - Geo-tagged Tweets

- Real-Time Dynamic Event Analysis
  - Prescheduled event vs. Dynamic Event

- Open data as a Sensor
Outline

• A bit of History
  • March 11 Disaster and Urban Context

• Ubiquitous Services
  • What are Ubiquitous Services

• Capturing Urban Context
  • Limitations and Possibilities

• Place-triggered Geotagged Tweets Analysis
  • Case Study

• Summary
Detection, Classification and Visualization of Place-triggered Geotagged Tweets

- Shinya Hiruta (1)
- Takuro Yonezawa (1)
- Marko Jurmu (1,2)
- Hideyuki Tokuda (1)

- ¹Keio University, ²University of Oulu
Real world event

- Structured as a collection of descriptive attributes
  - e.g. Place, Time, Content, ...
    - “Baseball game will be held at PNC park from 6:00 PM”

- However, attributes are often dynamic
  - e.g. Baseball game that gets postponed because of rain
  - e.g. A traffic accident occurring on a way and causing traffic congestion

- LBSN are suitable for extraction of dynamic information
Motivation:
Geotagged tweets are not always useful for real world event detection!

Content is related to the location

Pumpkin spice lattes at Starbucks. So good!
I’m at Convention Center
I want to watch today’s animation on TV!
It’s about to rain
@_BaracObama How are you today?
I love Justin Bieber!

Content is NOT related to the location

Useful Tweets

Unuseful Tweets
Place-triggered Geotagged Tweets

• Definition
  • Tweets that have both:
    • Geotag metadata
    • Content relevant to the associated location

• Research Goal
  • Detection
  • Classification
  • Application
Detecting Place-triggered Geotagged Tweets

Without our system  With our system

From: 2012–04–29 00:00:00  To: 2012–04–29 04:00:00

From: 2012–04–29 00:00:00  To: 2012–04–29 04:00:00
Detecting Place-triggered Geotagged Tweets

Without our system

With our system

From: 2012–04–29 04:00:00 To: 2012–04–29 08:00:00
Detecting Place-triggered Geotagged Tweets

Without our system

With our system
Detecting Place-triggered Geotagged Tweets

Without our system

With our system

From: 2012–04–29 12:00:00 To: 2012–04–29 16:00:00
Detecting Place-triggered Geotagged Tweets

Without our system

With our system

From: 2012–04–29 16:00:00 To: 2012–04–29 20:00:00
Detecting Place-triggered Geotagged Tweets

Without our system

With our system

Noise
Related Work

• Earthquake shakes twitter users: Real-time event detection by social sensors.
  • T. Sakaki, M. Okazaki, and Y. Matsuo.

• Measuring geographical regularities of crowd behaviors for twitter-based geo-social event detection.
  • R. Lee and K. Sumiya.
Comparison with Related Work

Existing Research

Specific Real World Event

- e

Top-down process

Tweet

Tweet

Our Approach

Various Real World Event

- x
- y
- z

Bottom-up process

Place-triggered

Non Place-triggered

Tweet

Tweet

Tweet

Tweet
Preliminary Survey

- Geotagged tweets in Twitter around Japan
- Period: From 2011-11-21 to 2011-12-31
- Number of sample: 2,000
- Classified these tweets to certain types based on their content

Most of the tweets (42.5%) were classified as noise
Classification of the Place-triggered Geotagged Tweets

• Classified to Five types:
  • Report of whereabouts
    • A tweet that user refers to his/her current location
  • Food
    • A tweet where user shares information regarding current food or drink
  • Weather
    • A tweet about weather of the location
  • Back at home
    • A tweet where user reports the fact that he/she is back at home
  • Earthquake
    • A tweet in which user reports the feeling of the earthquake
Approach

• How do we detect Place-triggered Geotagged Tweets?
• We started with straightforward approach
  • Report of whereabouts
    • Detecting checkin activity
      (Foursquare, Loctouch, Imakoko-now)
  • Food, Weather, Back at home and Earthquake
    • Naive keyword matching method with dictionary
      • We assume that people tend to classify tweets mainly by distinctive keywords
Design and Implementation

Crawling Module → Repository → Analysis Module

Social Media

Geotagged Tweets

<location, text, time>

Database

<location, text, time, type>

Query → Analysis Result

Visualize Applications
Interactive Visualization of Place-triggered Geotagged Tweets

- Filtering by date/time
- Animation view
- Each tweet is color by type
- Filtering by types of tweet
- Plotting area
場所誘因型位置情報付き発言の検出と可視化
Detection and Visualization of Place-triggered Geotagged Tweets

Top | Demo | 解説

時間帯
12:02 ~ 13:02

表示時間
60分

日付選択
March 2012
March 14, 2012 with Food Filter

Detection and Visualization of Place-triggered Geotagged Tweets

soba eater (@ soba Sato) http://twitter.com/soha_NzvRp0

日付選択

Microsoft Research Asia 2012
Detection and Visualization of Place-triggered Geotagged Tweets
March 14, 2012 with Earthquake Filter

場所誘因型位置情報付き発言の検出と可視化
Detection and Visualization of Place-triggered Geotagged Tweets

Top | Demo | 解説

優美ちゃん無事?! 曇れた〜(>__<)

2012年3月14日 21:11:55 JST

Microsoft Research Asia Faculty Summit 2012
Evaluation

• Methodology
  • Creating Ground-truth
    • Asked 18 third party people to classify tweets
      • 12 men in their 20s
      • 2 men in their 30s
      • 5 women in their 20s
  • Dataset
    • Geotagged tweets nearby Japan
    • Period: From 2012-01-01 to 2012-03-31
    • Total amount: 4,524,257
    • Each participants reviewed 500 tweets which were randomly sampled from the dataset
Evaluation Results

<table>
<thead>
<tr>
<th>Type of Tweets</th>
<th>Precision</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Report of whereabouts</td>
<td>93.18%</td>
<td>77.16%</td>
<td>84.42%</td>
</tr>
<tr>
<td>Food</td>
<td>53.6%</td>
<td>17.8%</td>
<td>26.7%</td>
</tr>
<tr>
<td>Weather</td>
<td>57%</td>
<td>21%</td>
<td>30%</td>
</tr>
<tr>
<td>Back at Home</td>
<td>54%</td>
<td>23%</td>
<td>32%</td>
</tr>
<tr>
<td>Earthquake</td>
<td>76%</td>
<td>66%</td>
<td>71%</td>
</tr>
</tbody>
</table>

Table 1. Classification result by the system

<table>
<thead>
<tr>
<th></th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRUE</td>
<td>40.09%</td>
<td>15.84%</td>
</tr>
<tr>
<td>FALSE</td>
<td>2.18%</td>
<td>41.89%</td>
</tr>
</tbody>
</table>

Table 2. Accuracy rate of detecting place-triggered geotagged tweets
Future Work

• Expanding the classification
  • Expand to other countries
  • More complete categories

• Improving detection accuracy
  • Linguistic analysis, slang

• Discovering real events
  • Automatic event detection
  • Temporal-spacial analysis should be investigated
Conclusion

- Capturing Urban Context
  - Limitations: Useful and harmful
    - Anonymity Set and Privacy Enhancement
    - Visualization Problem
    - Small Anonymity Set Problem
  - Possibilities
    - Hybrid Sensing Model
    - Crowdsourcing and Gamification
    - Real-Time Dynamic Event Analysis

- Place-triggered Geotagged Tweets Analysis
  - Detecting Five types of the place-triggered geotagged tweets
    - Report of whereabouts, Food, Weather, Back at home and Earthquake
    - Showed that the system can detect place-triggered geotagged tweets with an overall accuracy of 82%
Thank you!
http://www.ht.sfc.keio.ac.jp/