Microsoft Research

Microsoft Research Asia
Faculty Summit 2012

Capturing Urban Context: Its Limitations and Possibilities

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

Microsoft Research Asia
Faculty Summit 2012

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

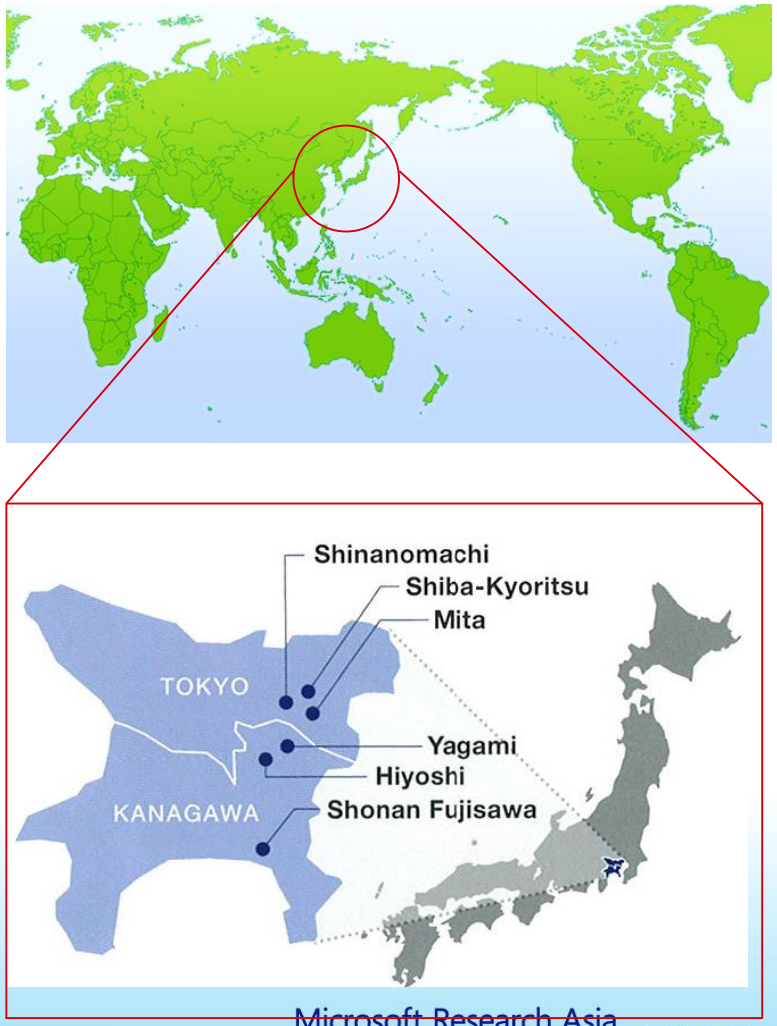

Hiyoshi Campus


Yagami Campus


Shinanomachi Campus


Shonan Fujisawa Campus

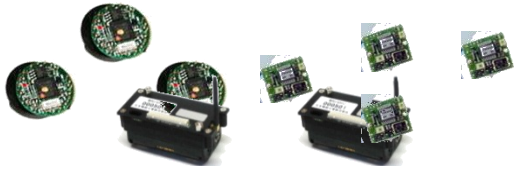

Shiba-Kyoritsu Campus

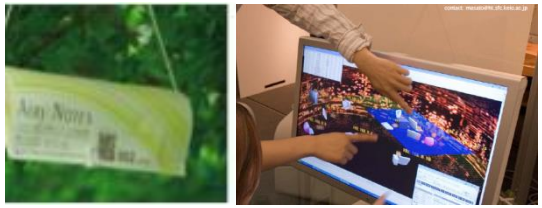


Hide Tokuda Lab., Keio University

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

SS Lab., Smart Living, uPlatea

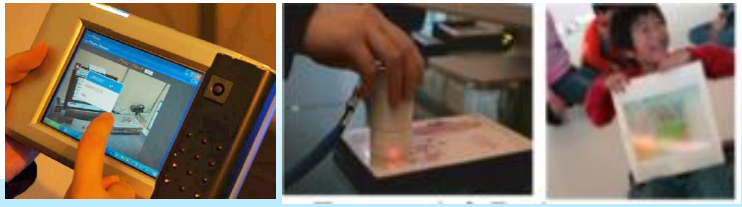




MANET, Heterogeneous MANET
Ubiquitous Network Browser

Service Platform: Smart Furniture, uTexture

Dependable Ubiquitous Nodes
Environmental Sensor Nodes



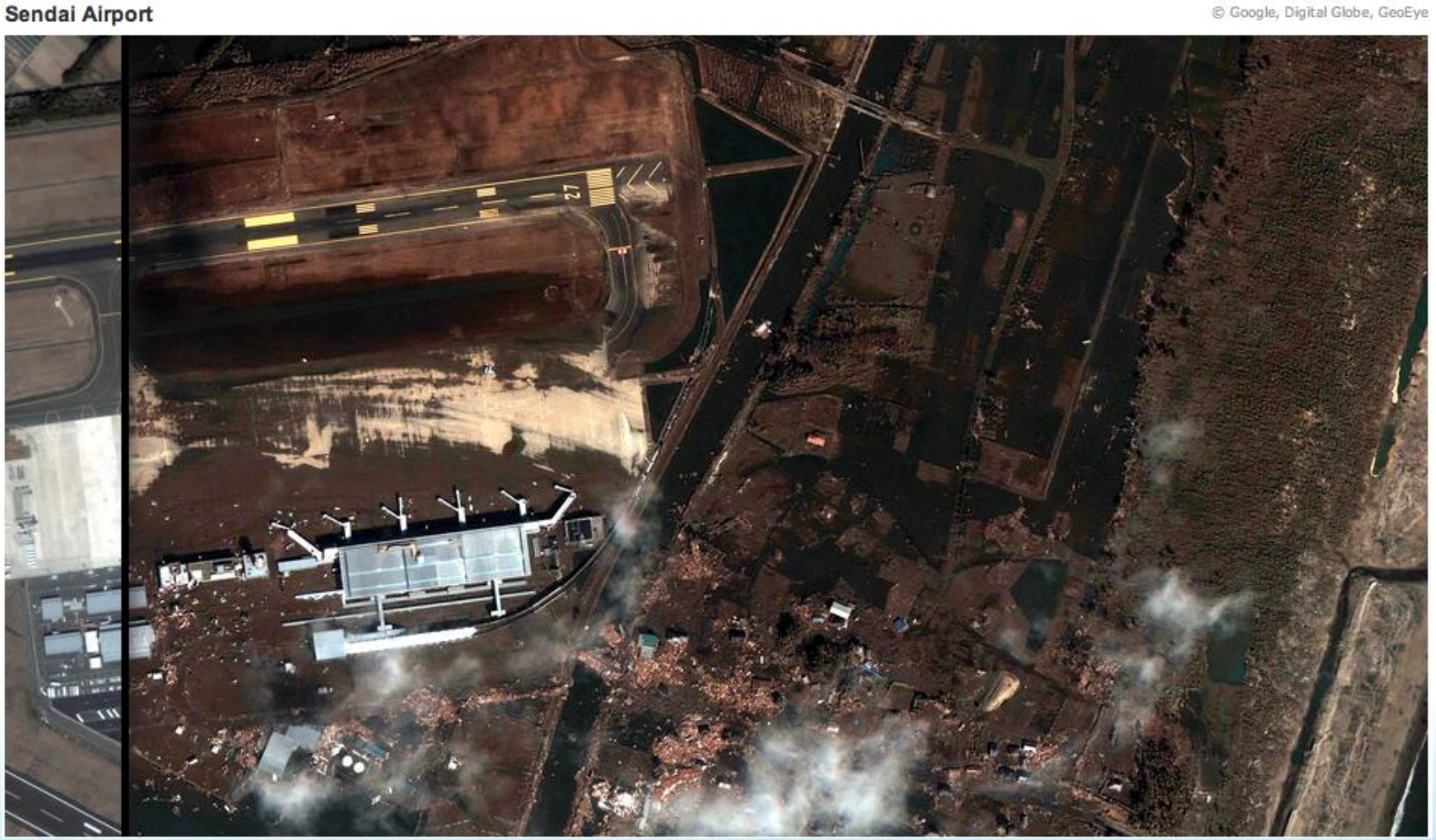
uPhoto, @Reader, uTexture, InfoRod
photo-based Interaction, Gesture-based Interaction
Multi-display Interaction

Recovering from 3.11 Disaster

Thank you for Supporting and Praying for Japan

Recovering from 3.11 Disaster

Recovering from 3.11Disaster (by ABC News)

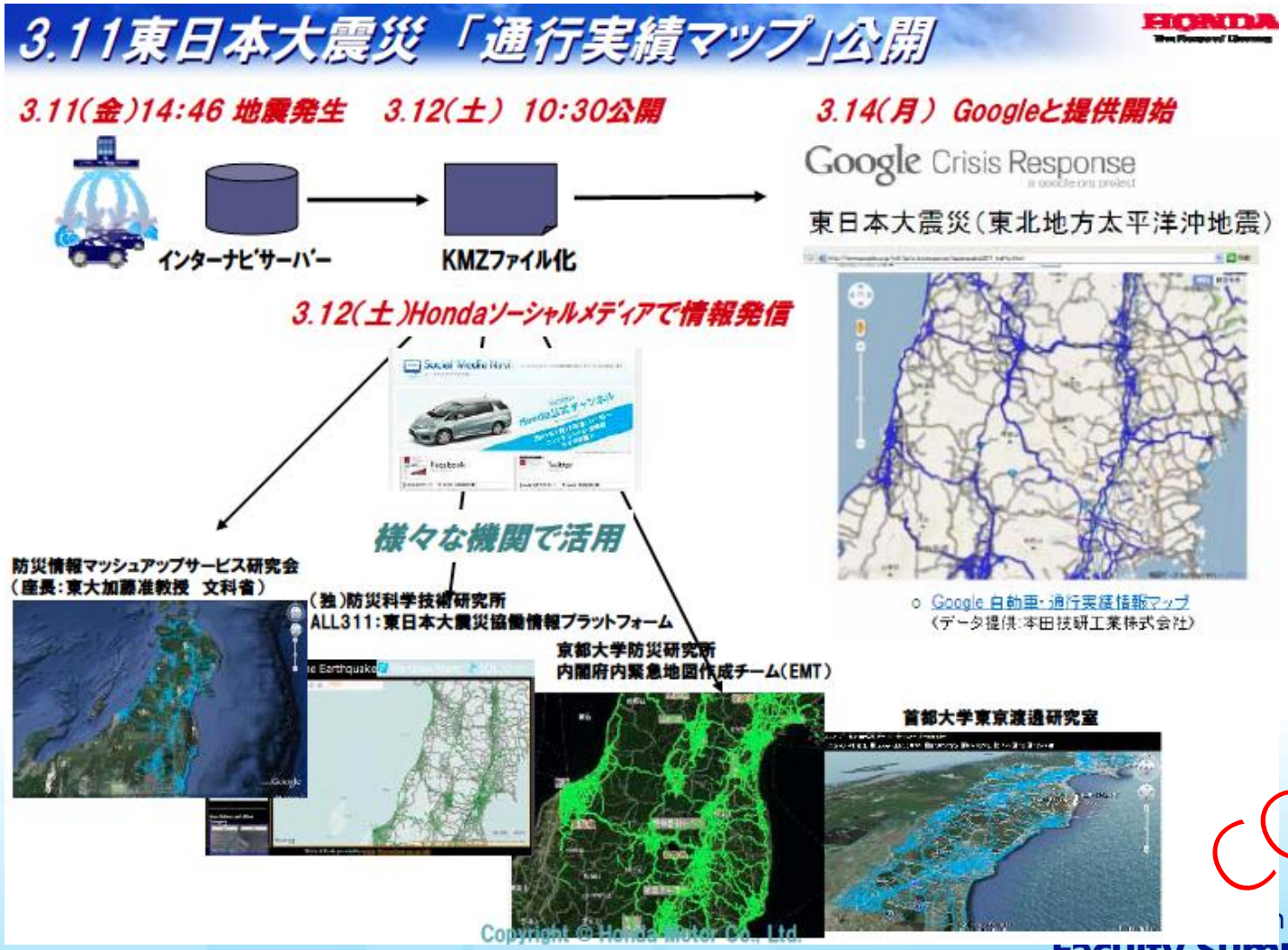


Recovering from 3.11Disaster (by ABC News)

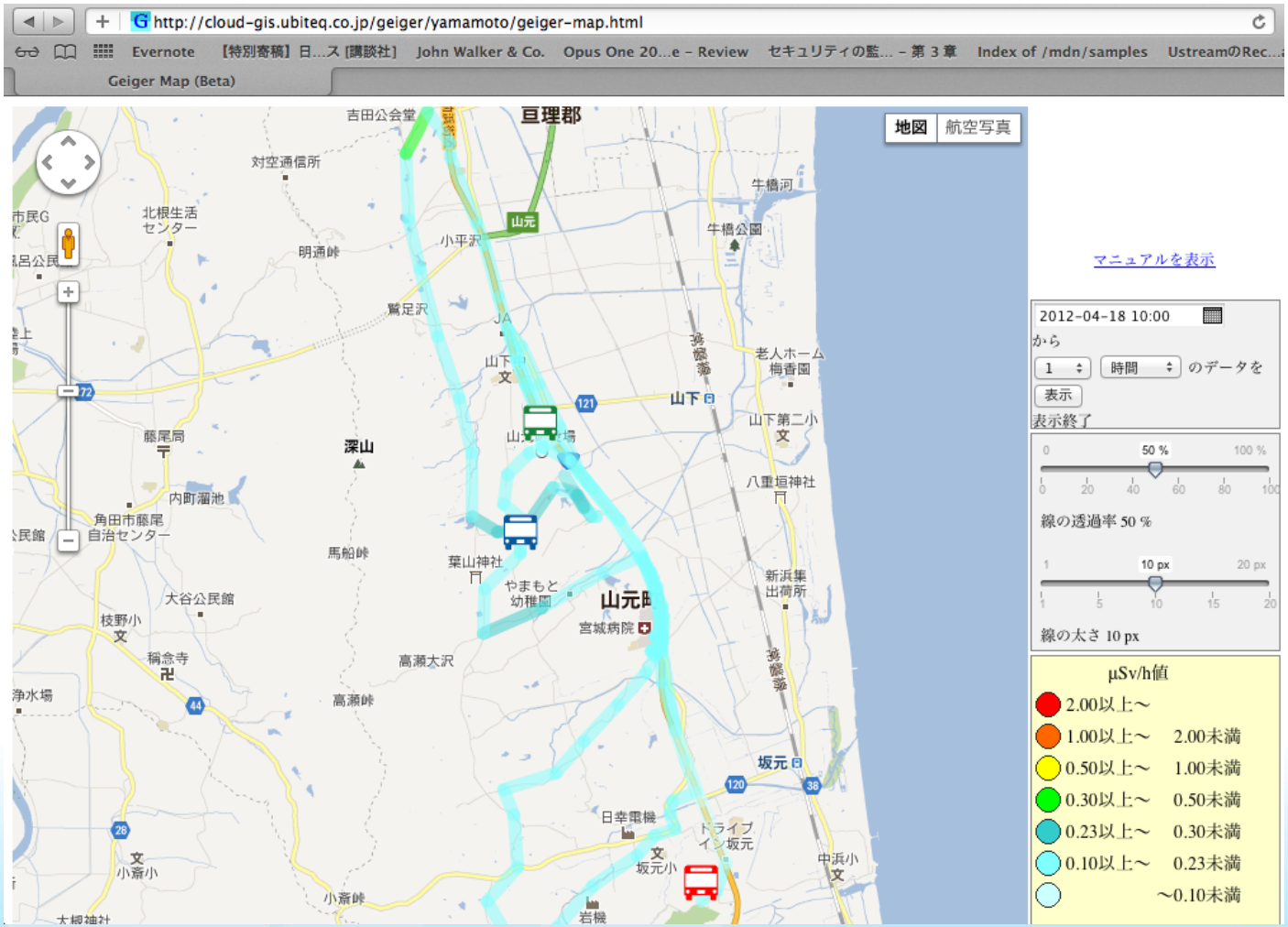
Urban Context Capturing

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

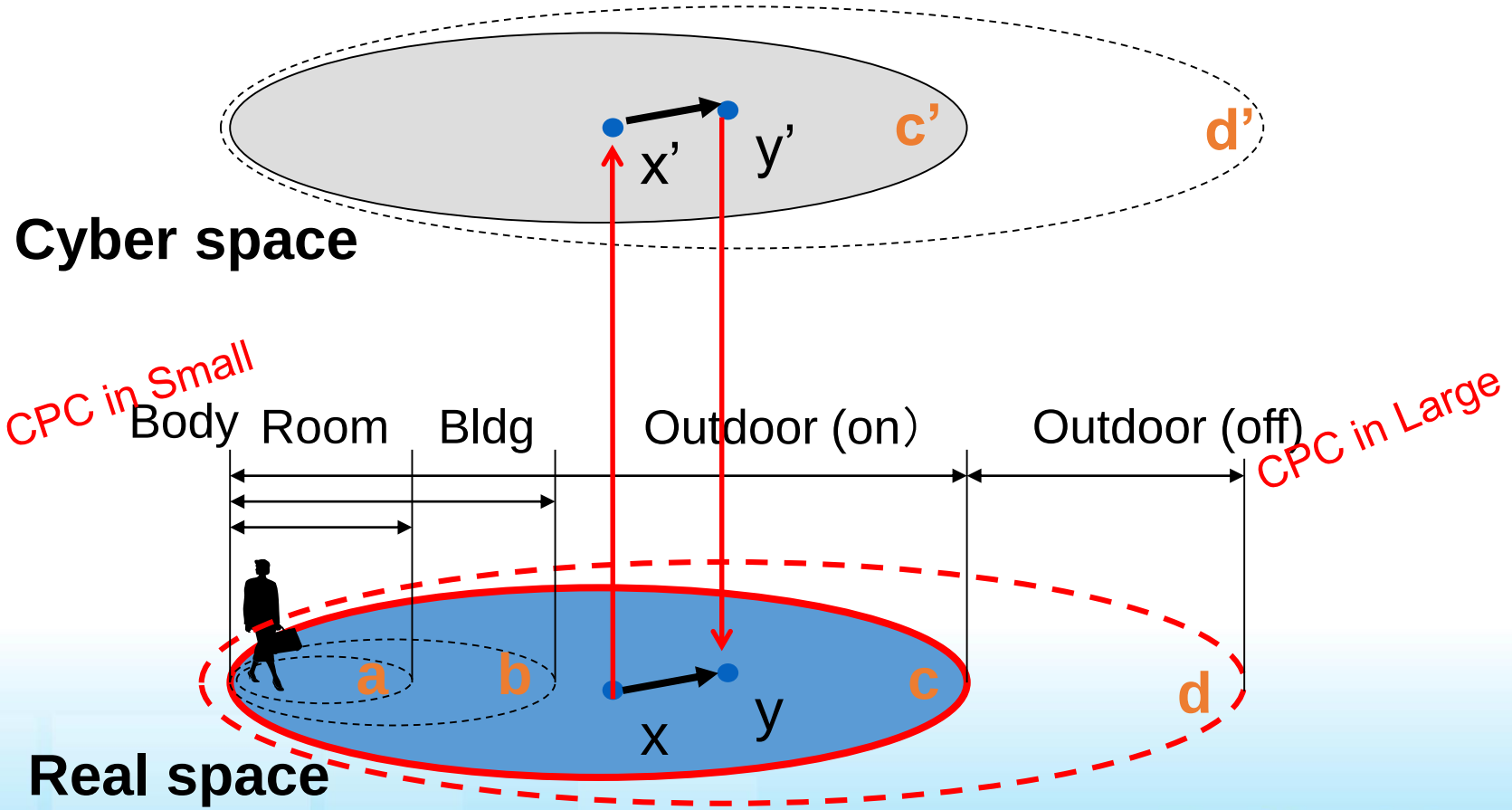
- A bit of History
  - March 11 Disaster and Urban Context
- Ubiquitous Services
  - What are Ubiquitous Services
- Capturing Urban Context
  - Limitations and Possibilities
- Place-triggered Geotagged Tweets Analysis
  - Case Study
- Summary

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-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

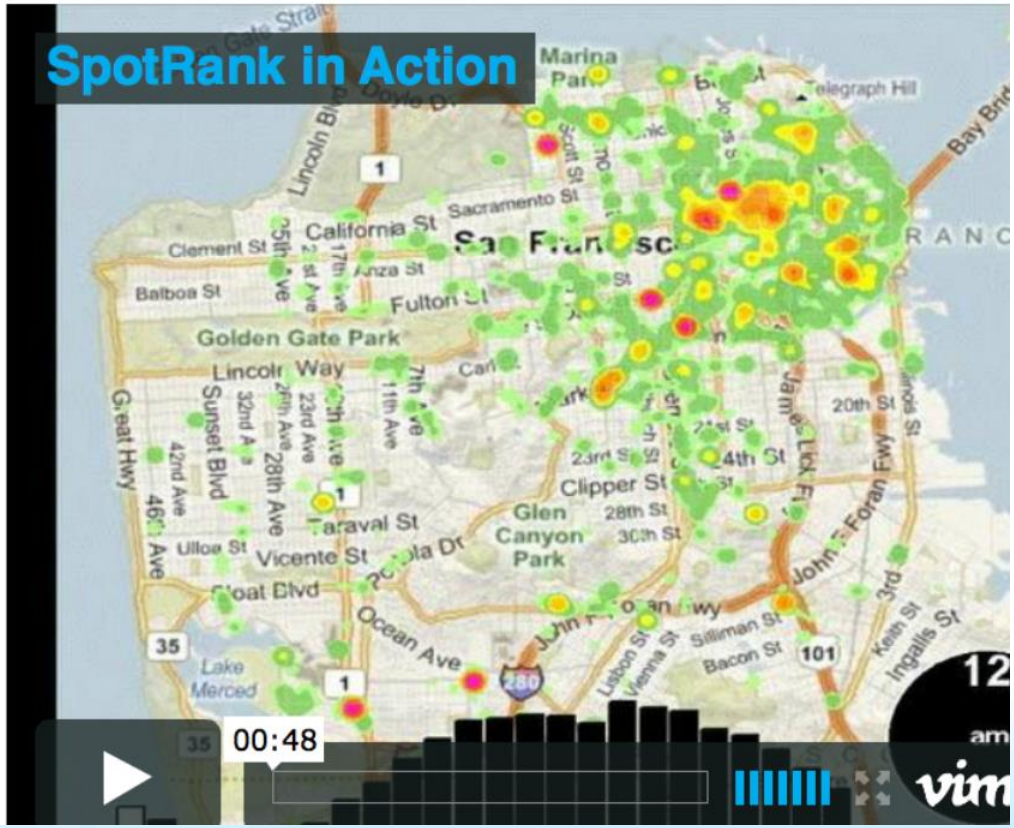
Context-awareness in Ubiquitous Services

- 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)

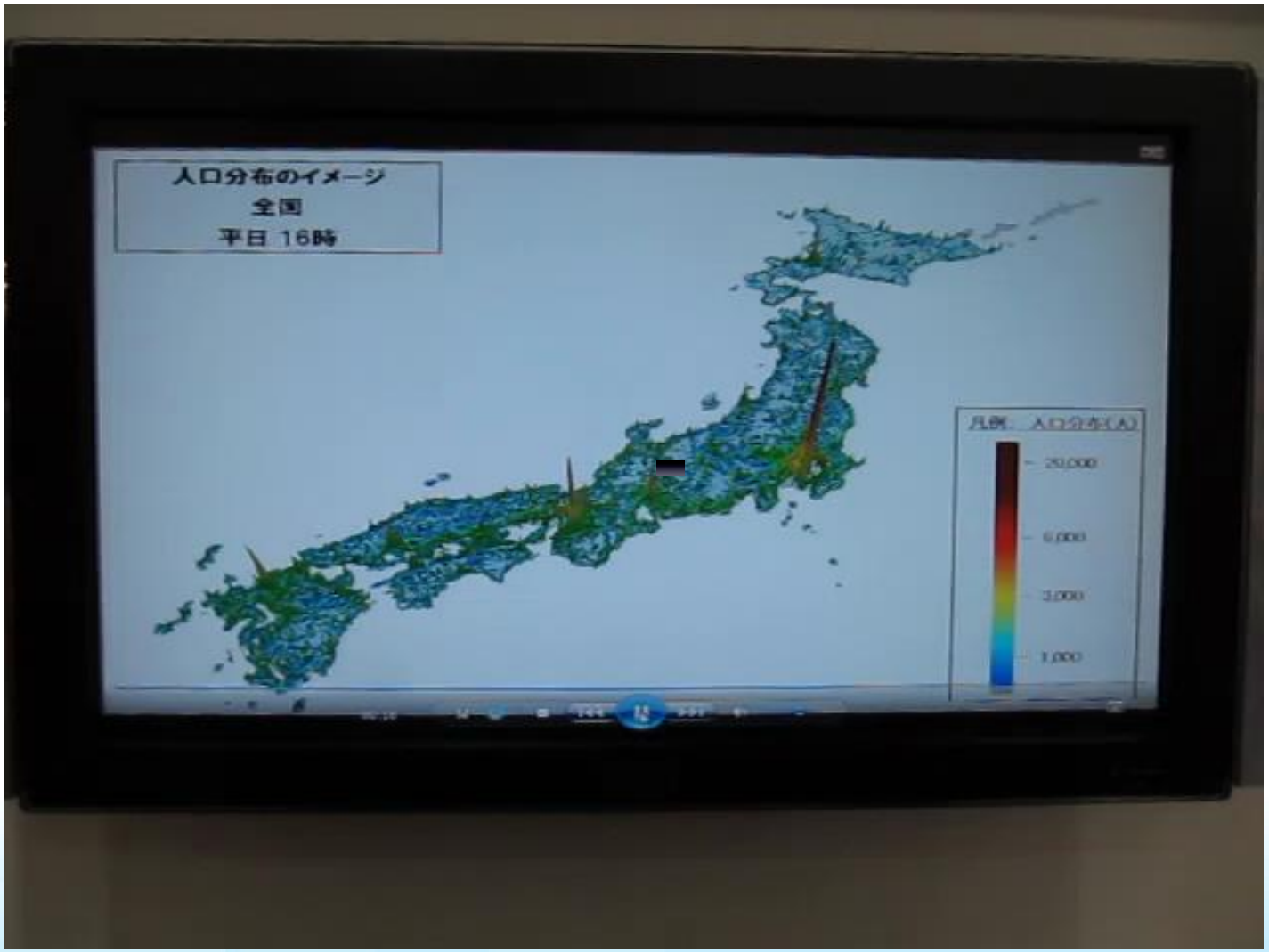
SkyHook's SpotRank
(http://www.skyhookwireless.com)

SpotRank In Action



NTT DoCoMo (Mobile Space Statistics(2010))

Mobile Space Statistics (2010)



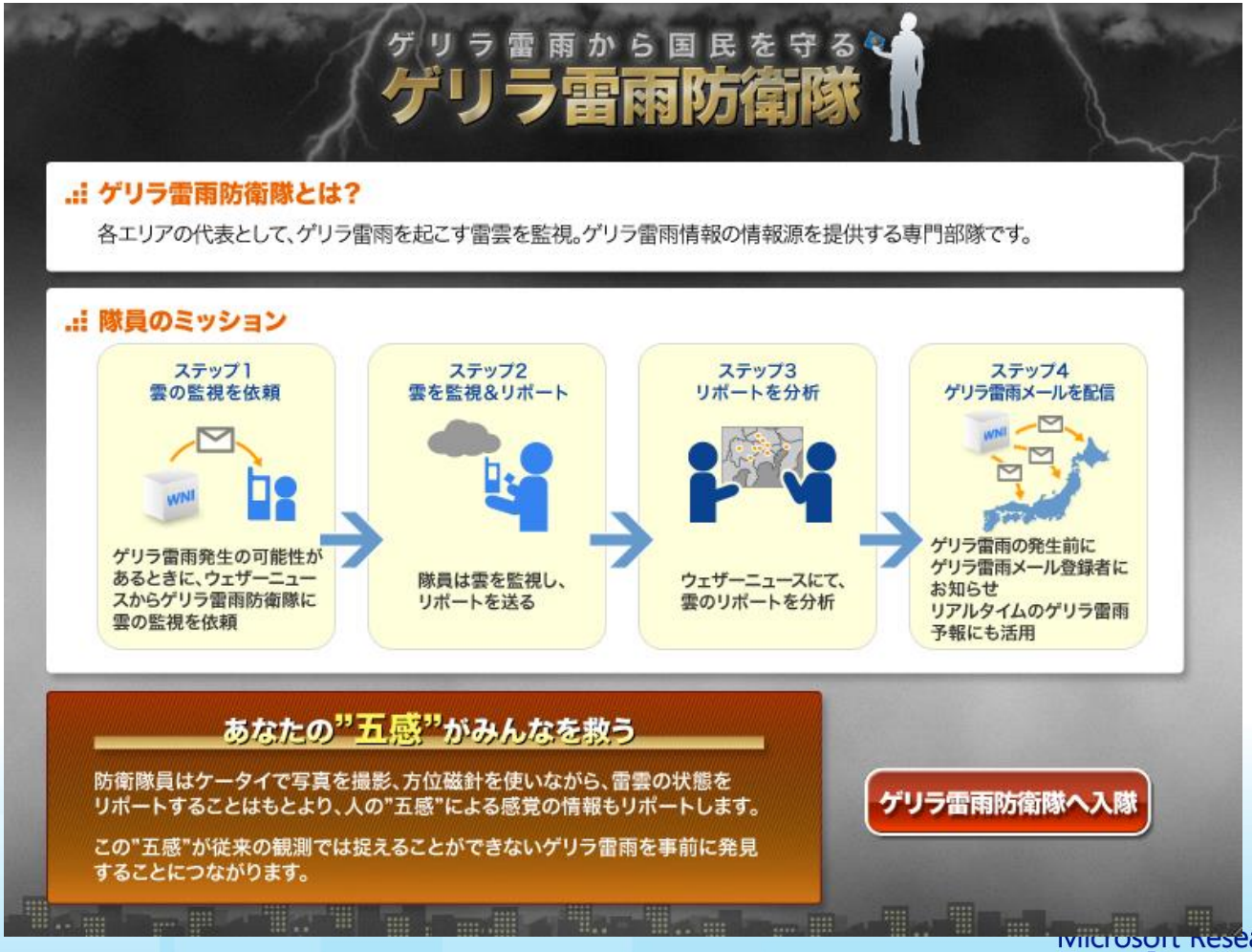
Limitations: Useful and Harmful

- 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

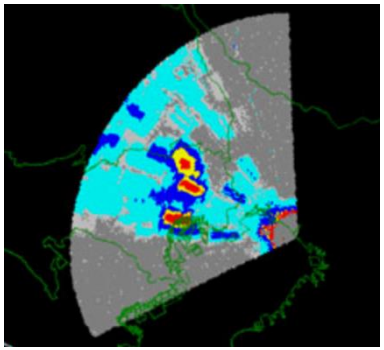
Urban Context Capturing
Possibilities

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

各都府県	ゲリラ雷雨 発生数	ゲリラ雷雨 捕捉率	ゲリラ雷雨メール 事前送信時間
東京都	172回	76.7%	平均38分前
大阪府	128回	62.5%	平均8分前
愛知県	172回	63.9%	平均19分前

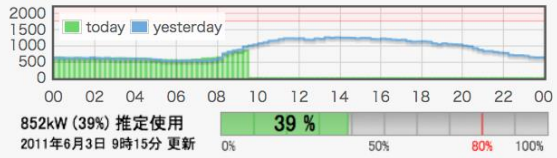
Electricity Consumption: Improving Citizen's awareness

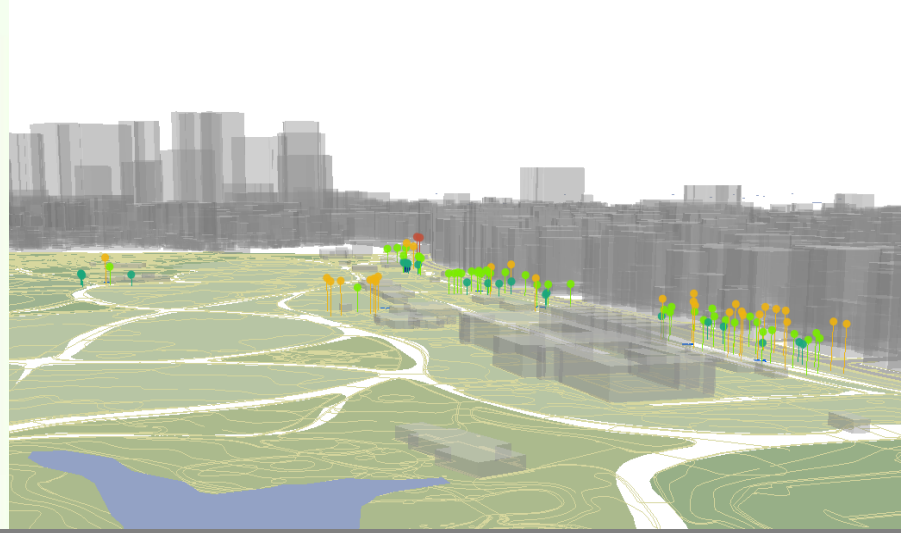
projects

LINKが取り組んでいるプロジェクトです。

SFC Power Awareness Campus Project

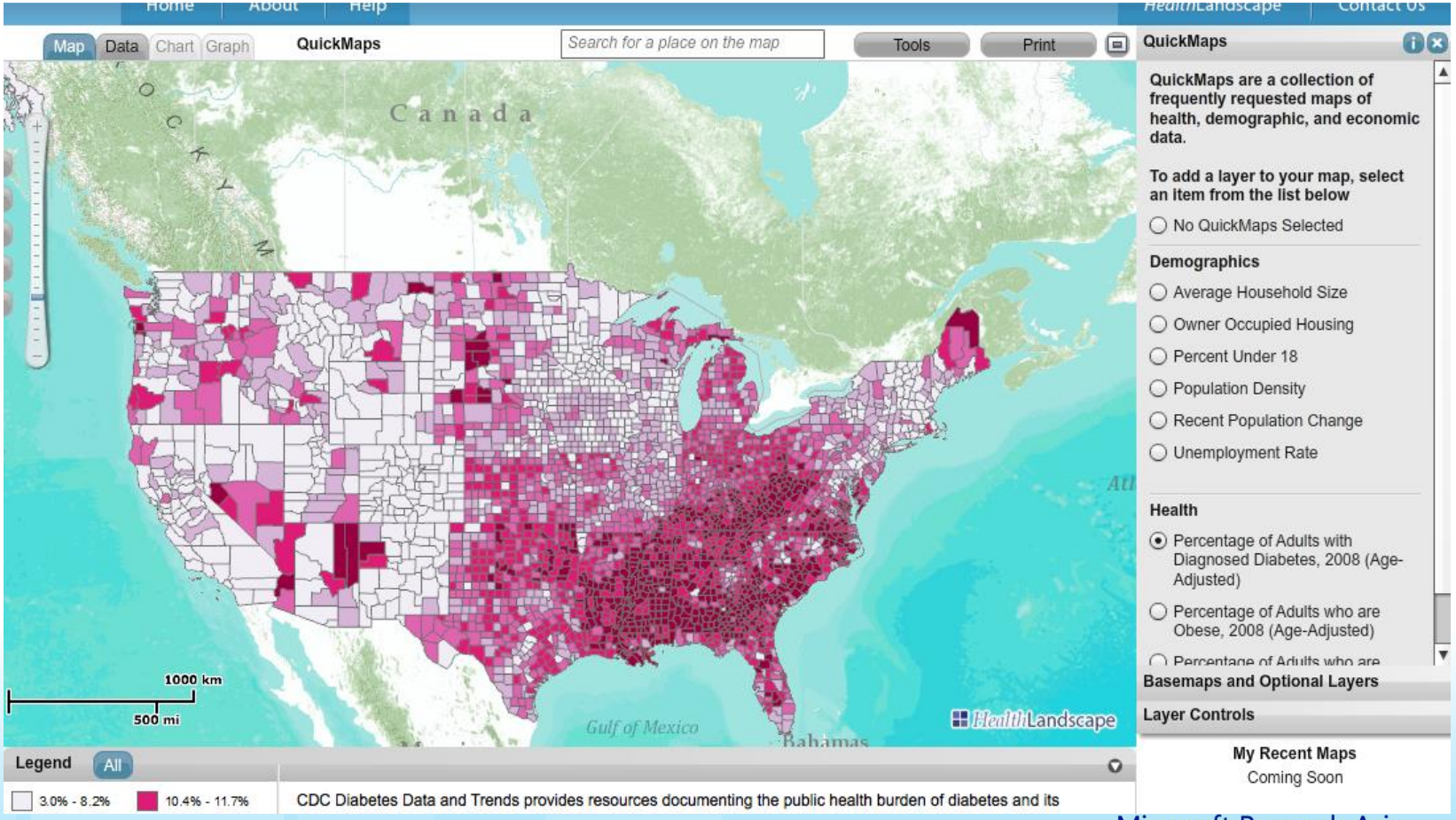
SFCの消費電力をモニタリングし、可視化することで、電力予測、節電通知、意識向上など、SFC全体の電力awarenessの向上に貢献しています。





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]

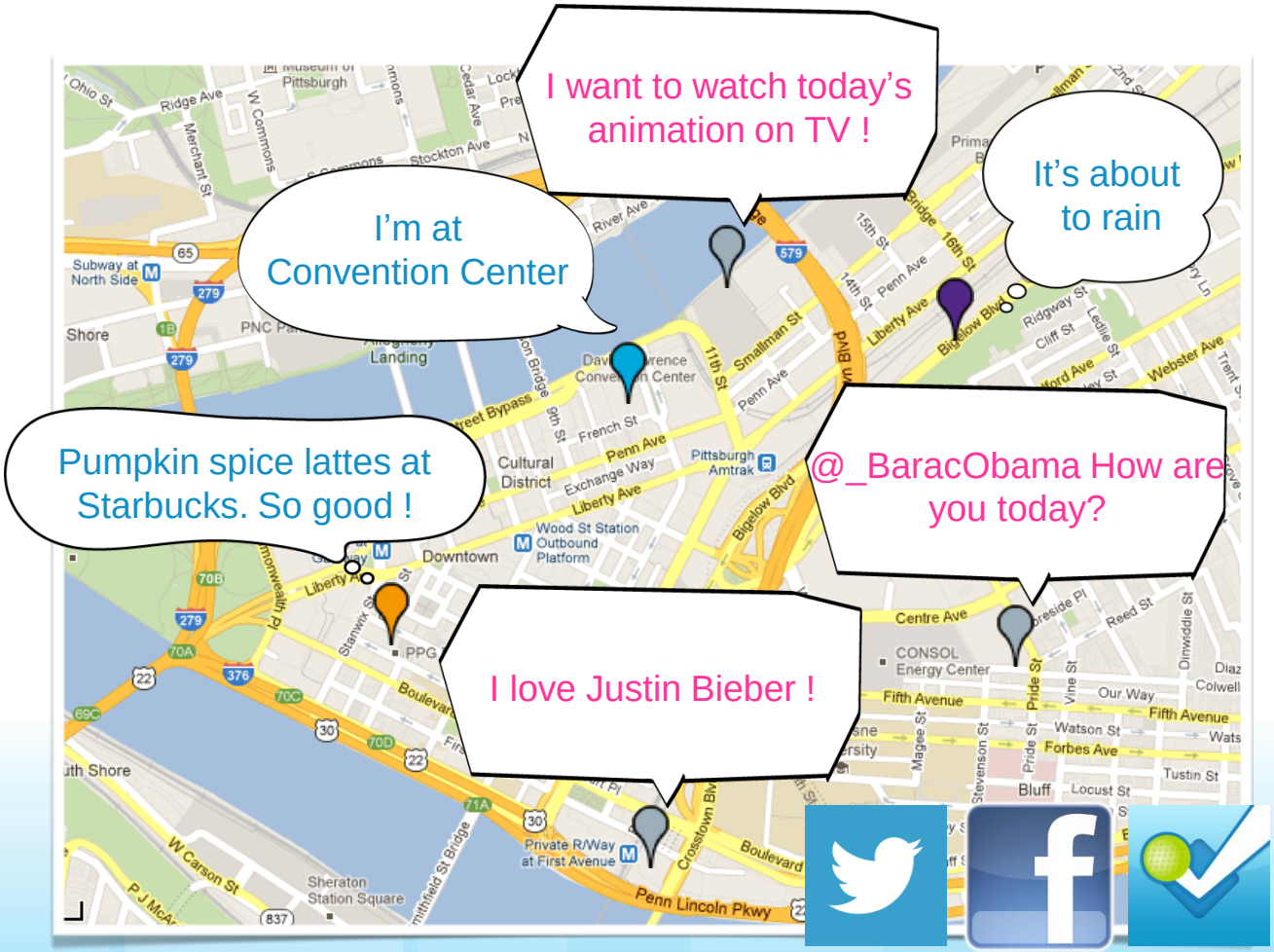
- [1] Keio University, [2] University of Oulu

Background: Real World Event Detection with Location-Based Social Networks

- 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!



Useful Tweets

Content is related to the location

Unuseful Tweets

Content is NOT related to the location

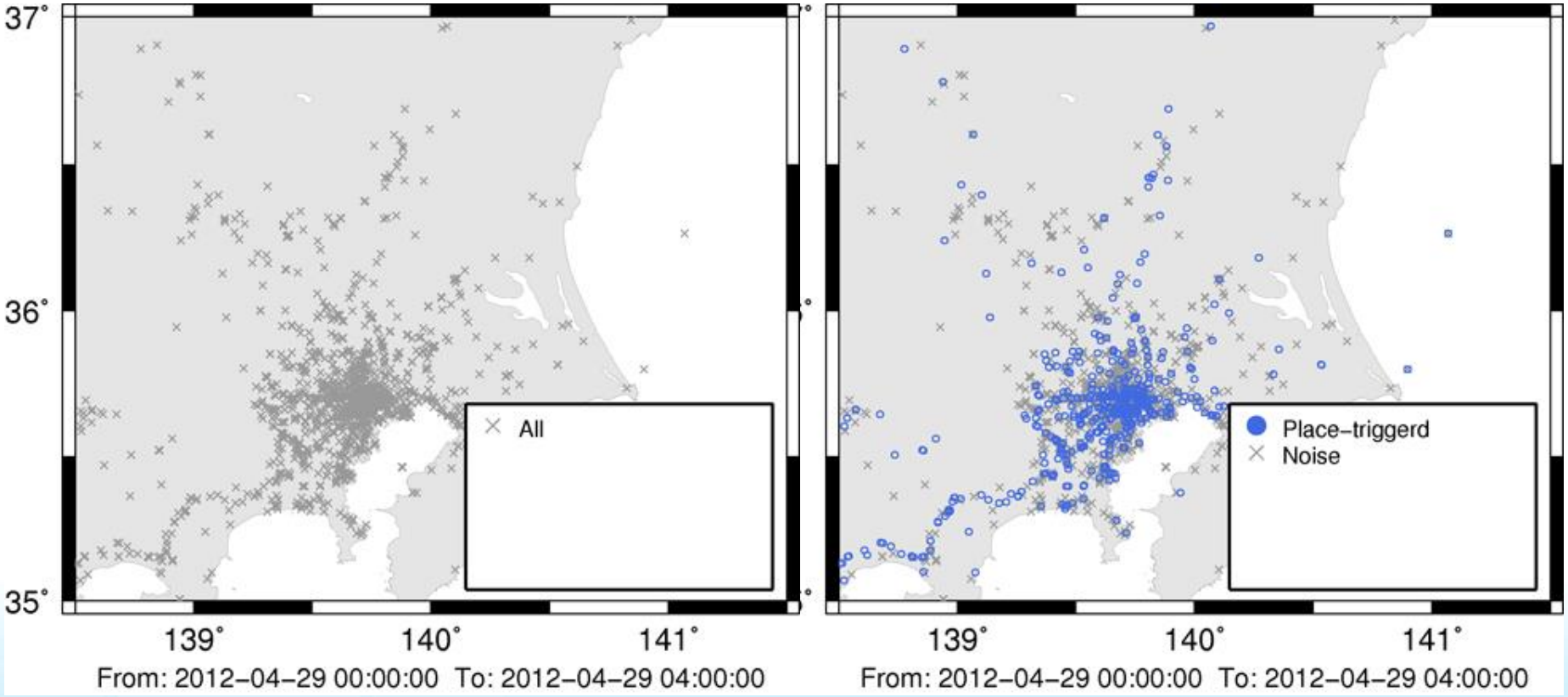
Place-triggered Geotagged Tweets

- 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

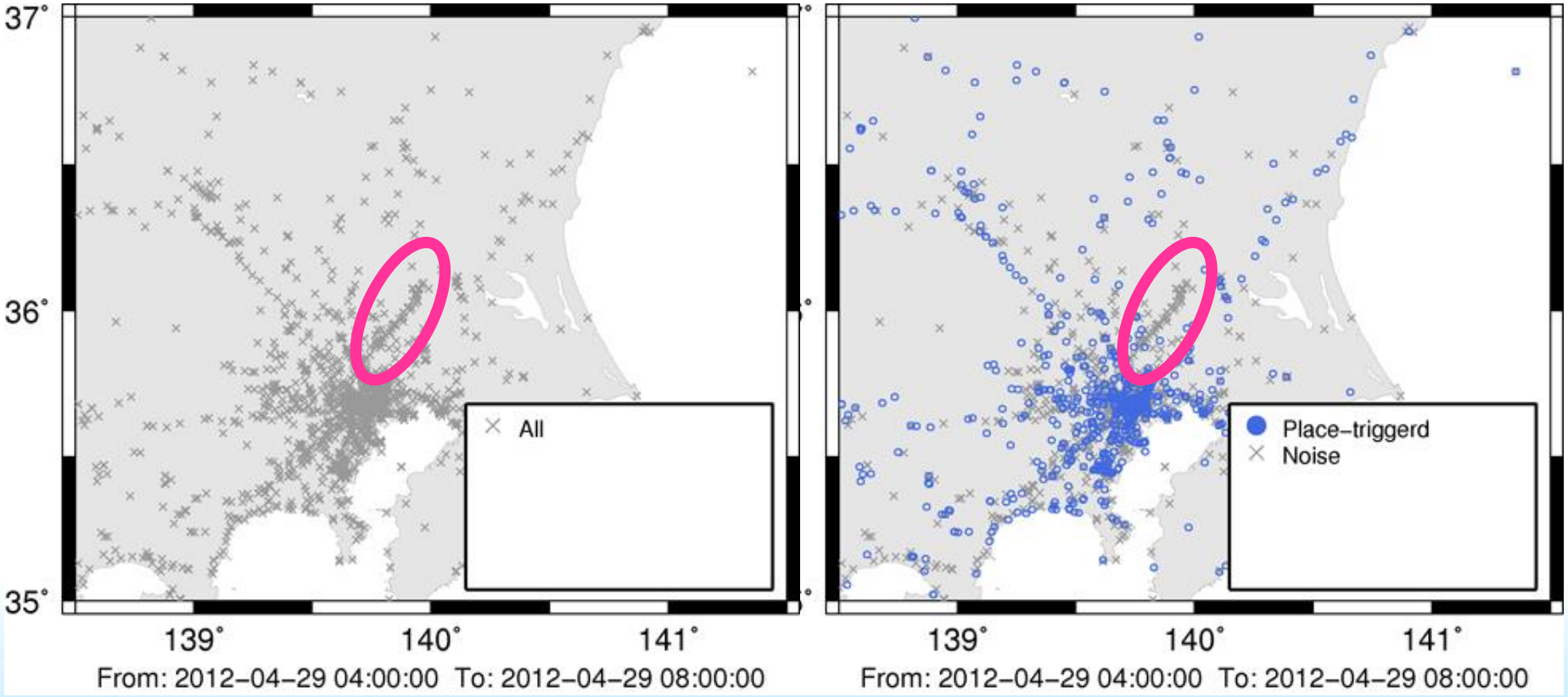
With our system



Detecting Place-triggered Geotagged Tweets

Without our system

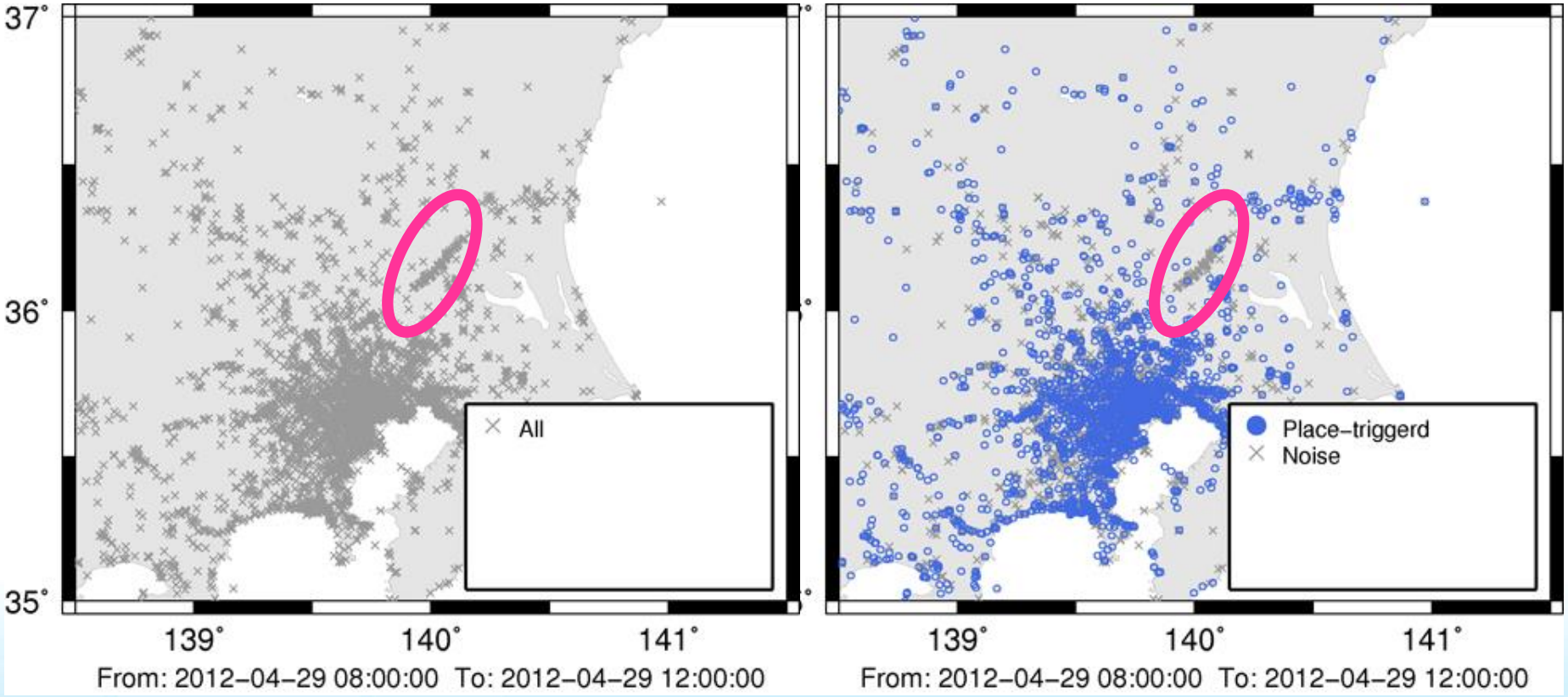
With our system



Detecting Place-triggered Geotagged Tweets

Without our system

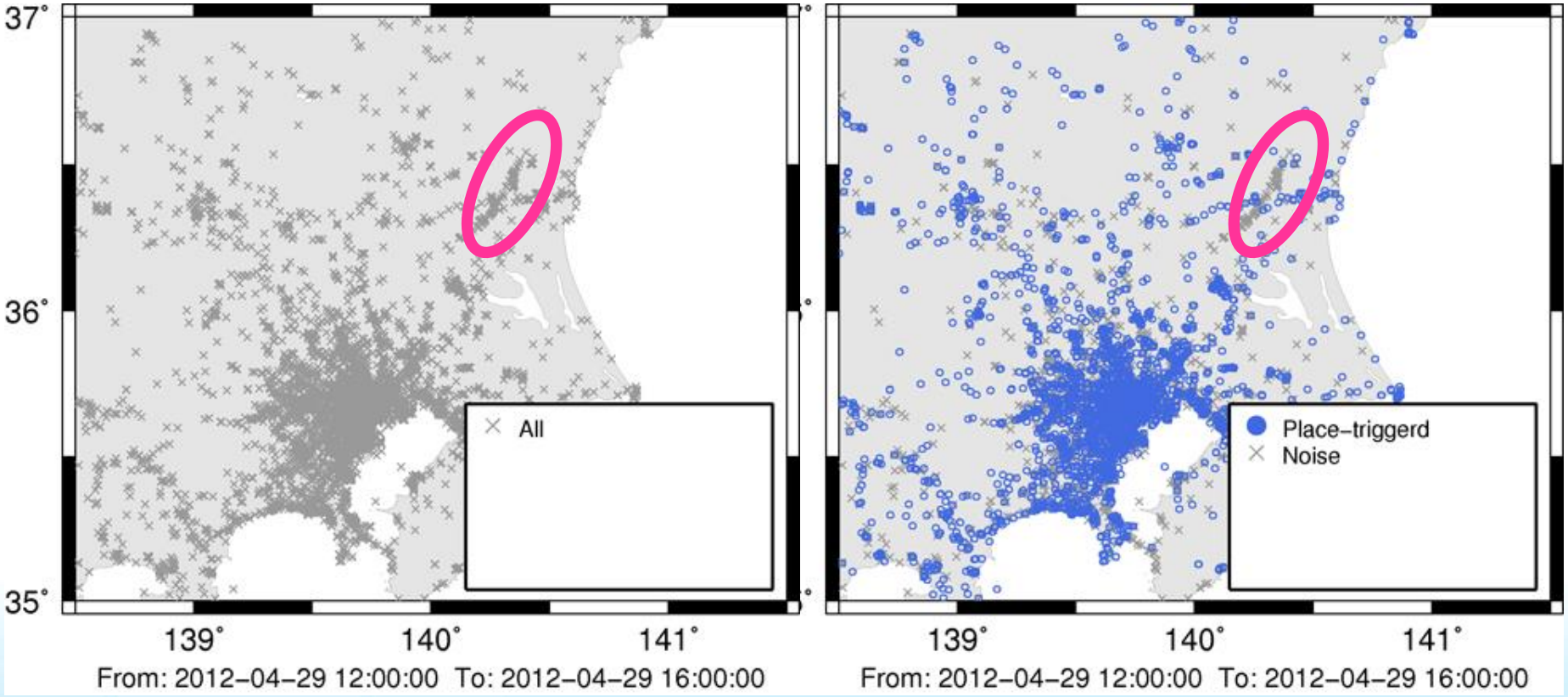
With our system



Detecting Place-triggered Geotagged Tweets

Without our system

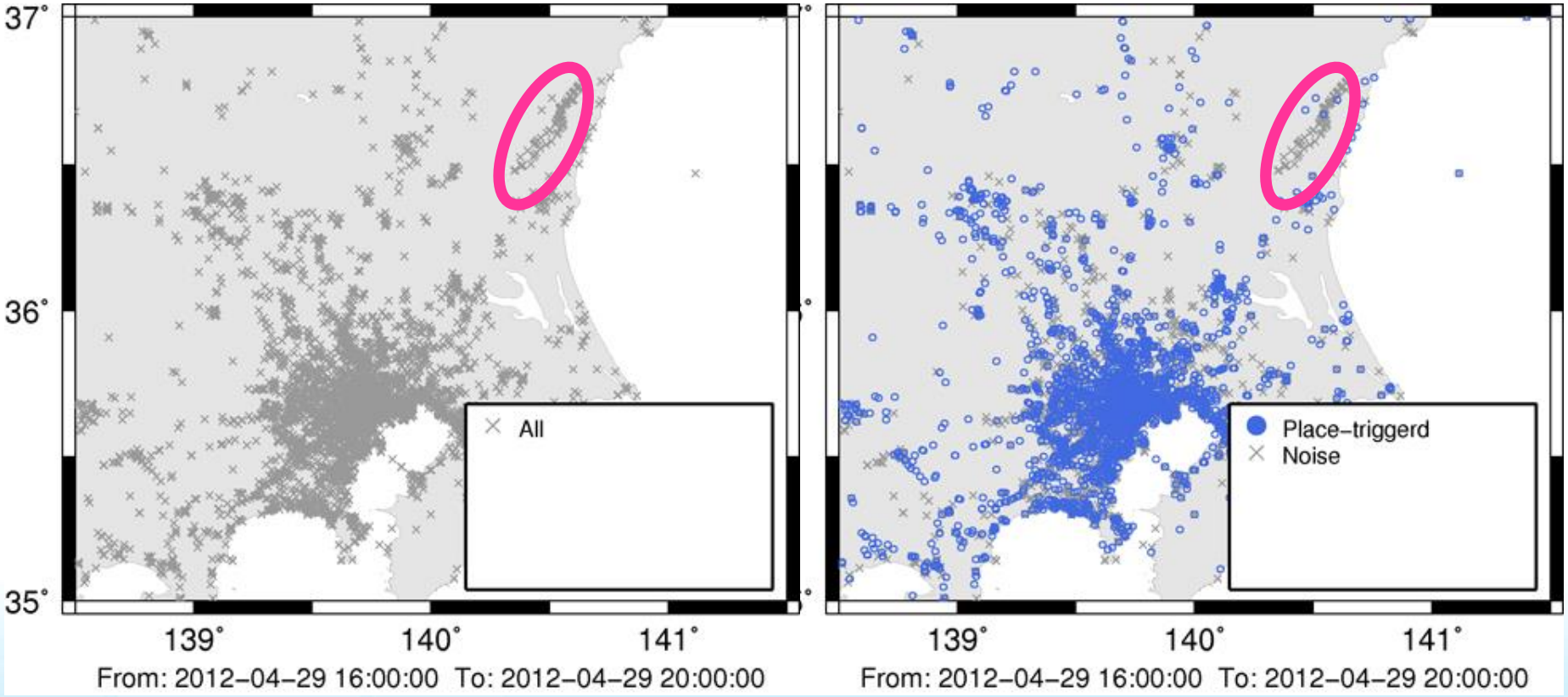
With our system



Detecting Place-triggered Geotagged Tweets

Without our system

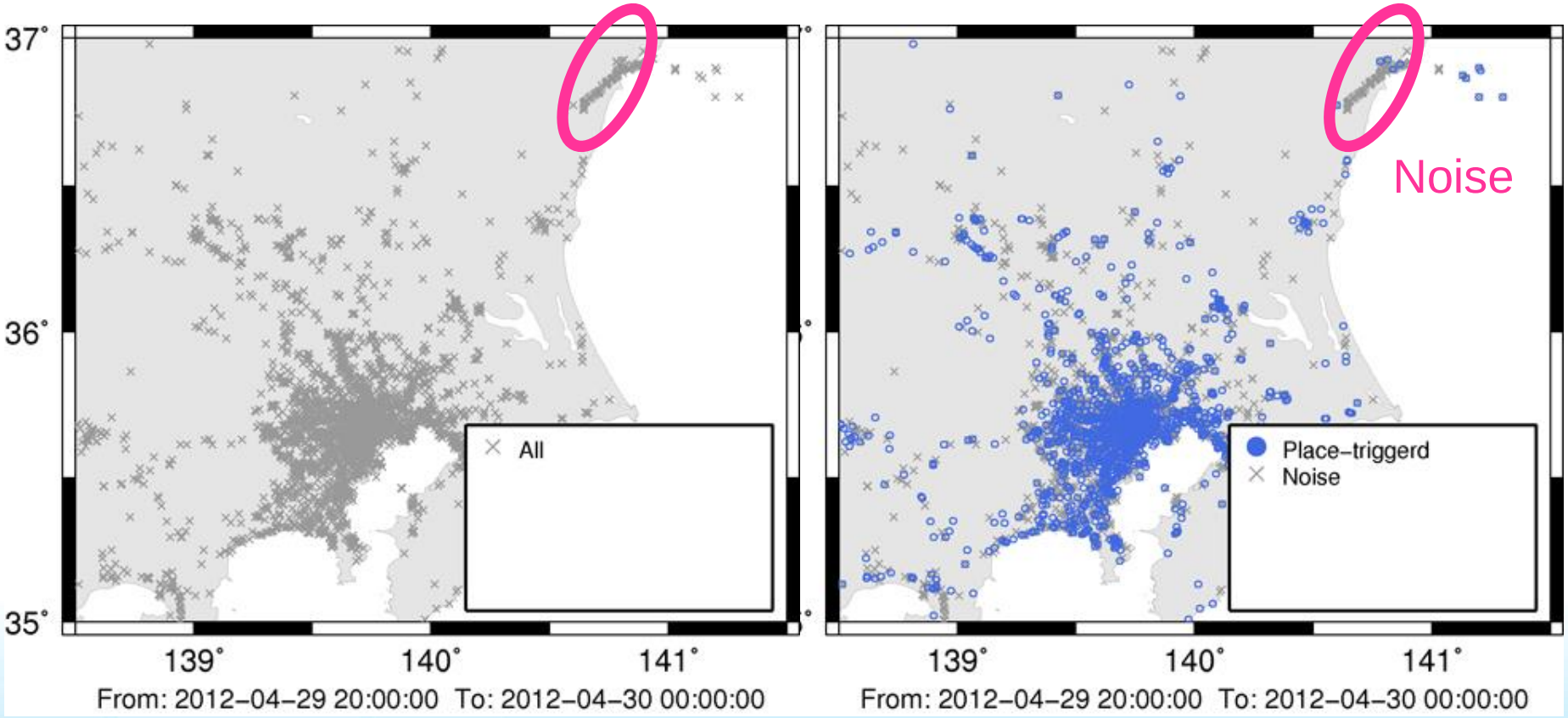
With our system



Detecting Place-triggered Geotagged Tweets

Without our system

With our system



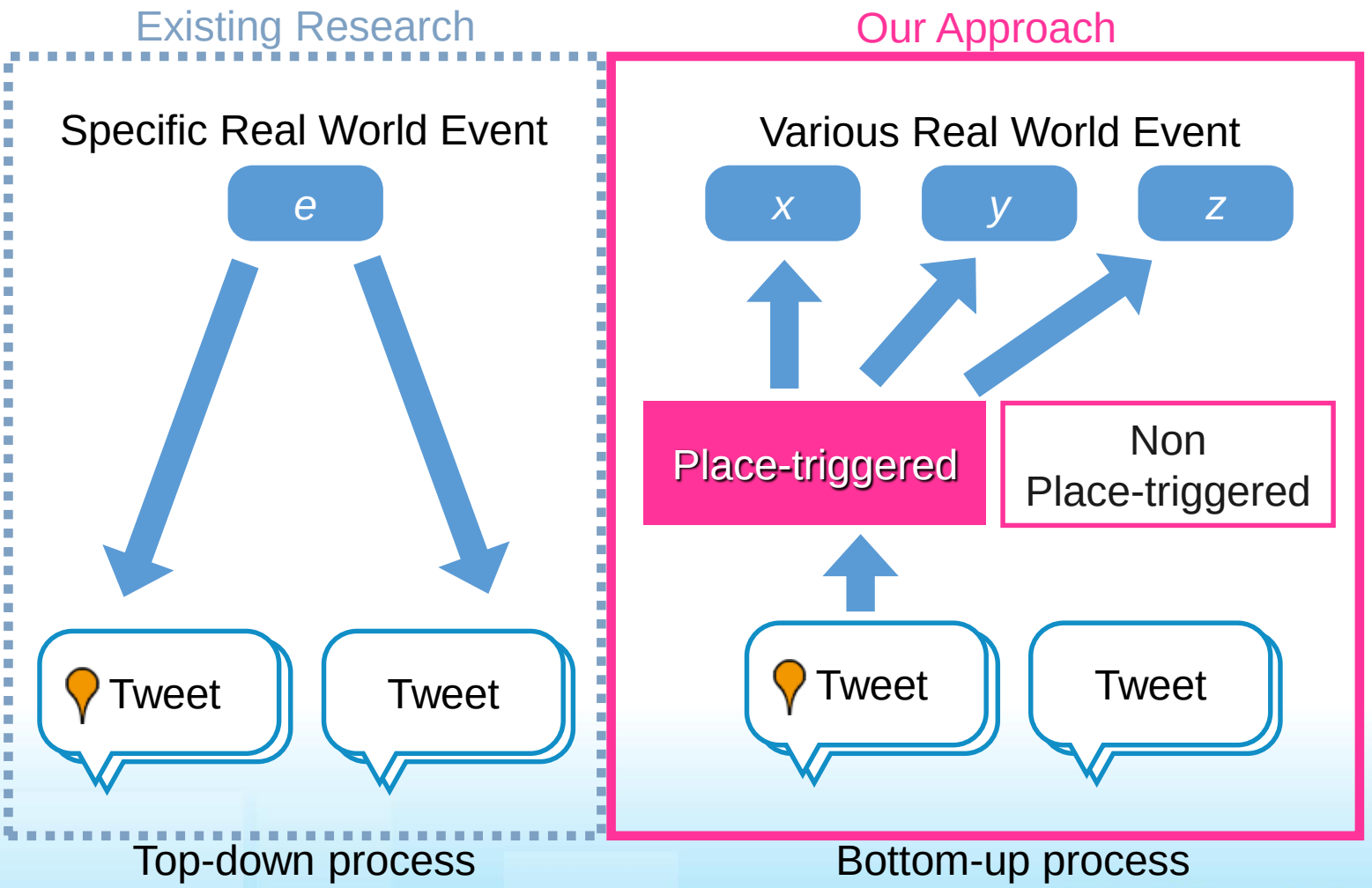
Related Work

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

    In Proceedings of the 19th International Conference on World Wide Web, pages 851–860, 2010.
- Measuring geographical regularities of crowd behaviors for twitter-based geo-social event detection.
  - R. Lee and K. Sumiya.

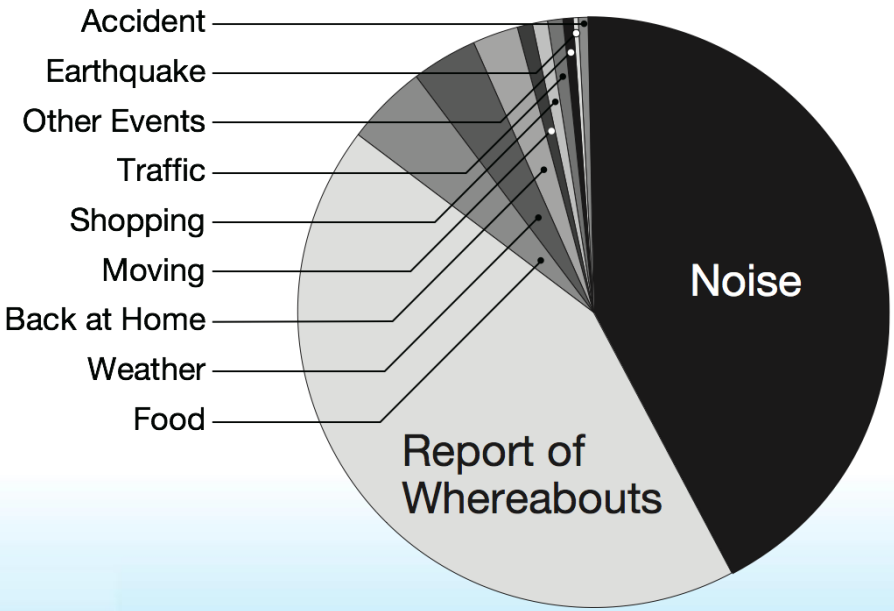
    In Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks, pages 1–10, 2010.

Comparison with Related Work



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

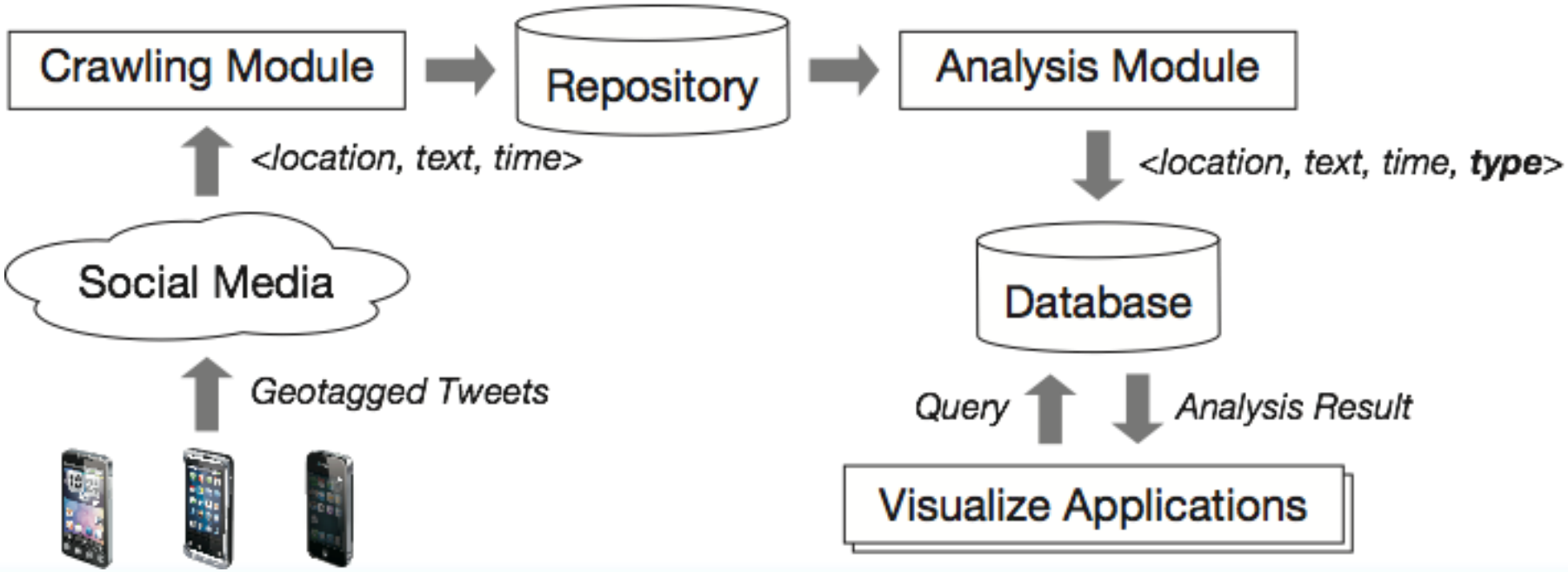
Classification of the Place-triggered Geotagged Tweets

- 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

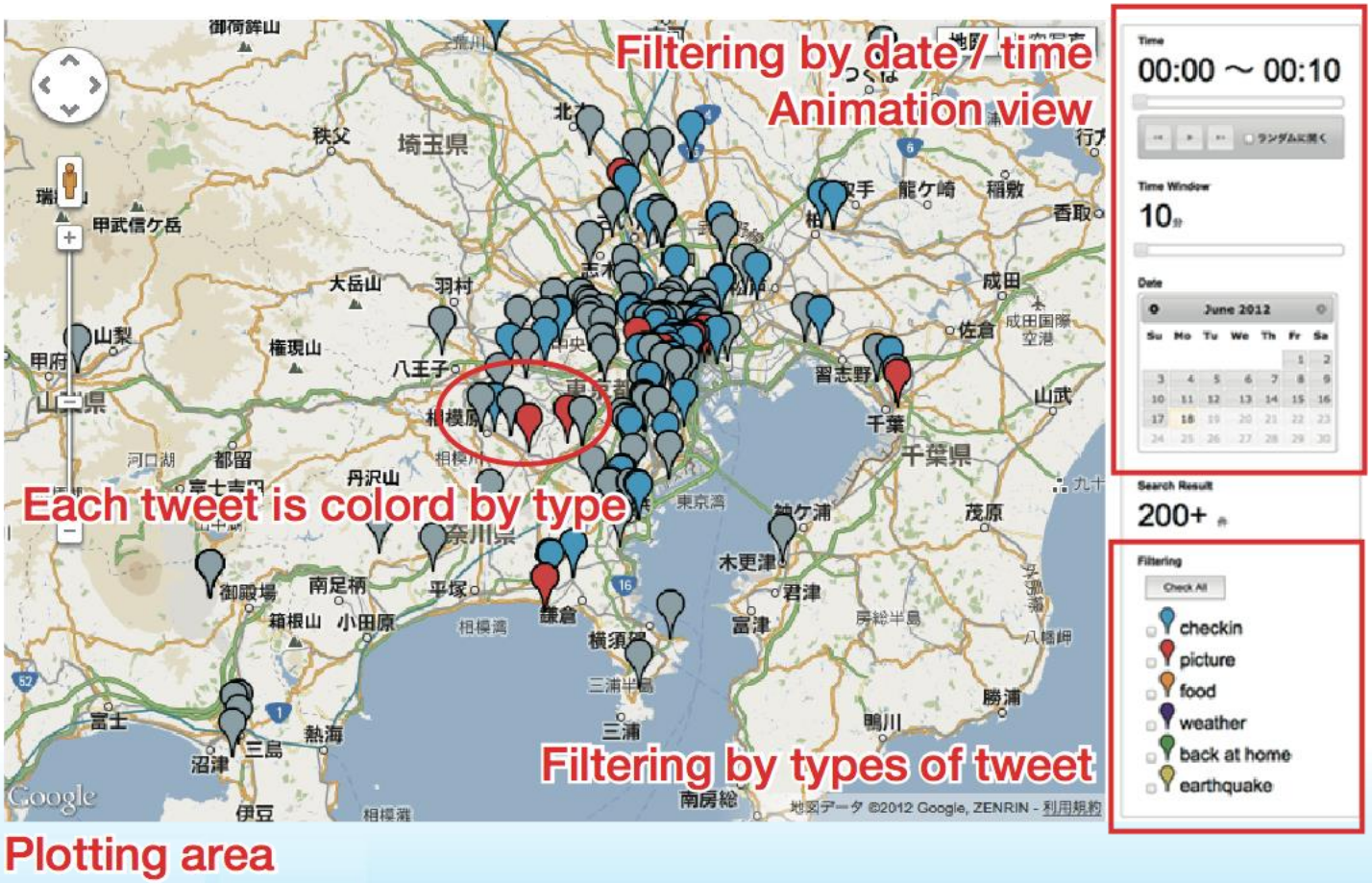
Approach

- 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



Interactive Visualization of Place-triggered Geotagged Tweets



March 14, 2012 without Food Filter

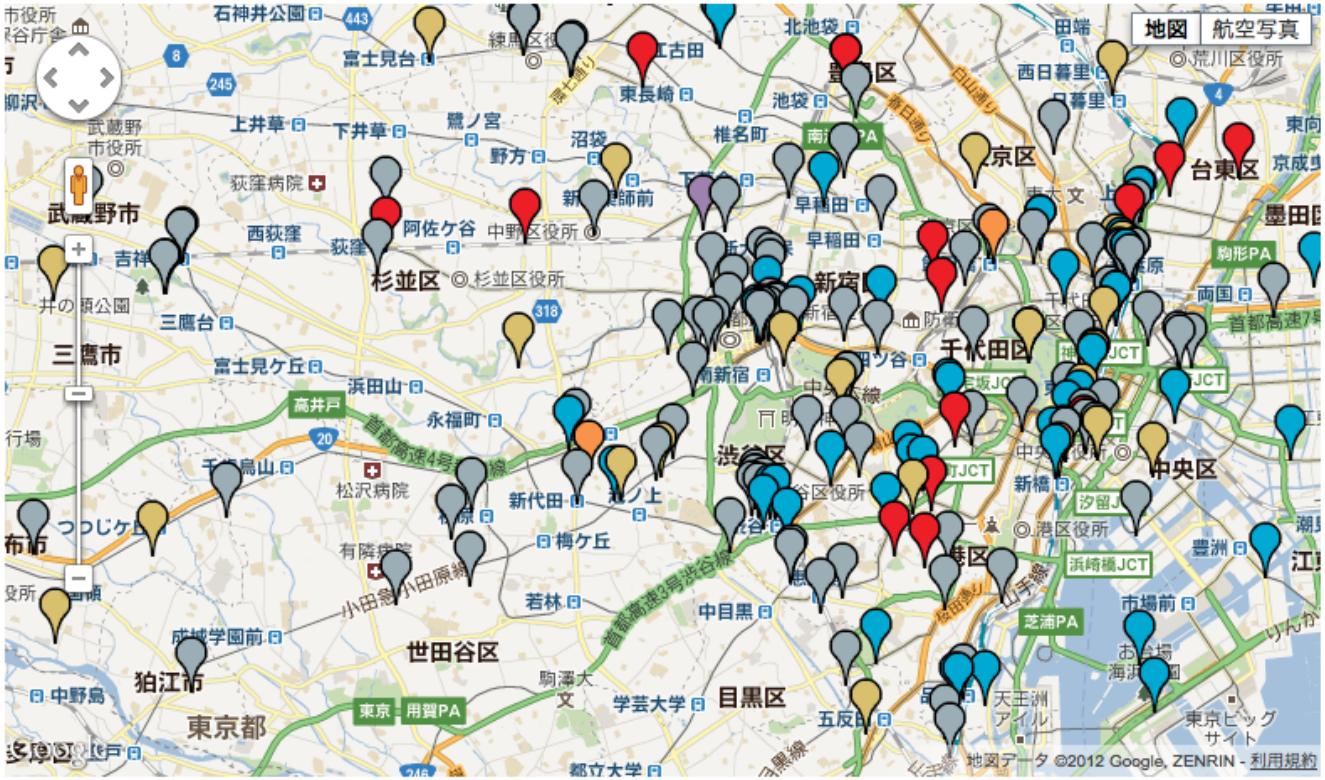
March 14, 2012 with Food Filter

March 14, 2012 without 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



時間帯

20:30 ～ 21:30

ランダムに開く

表示時間

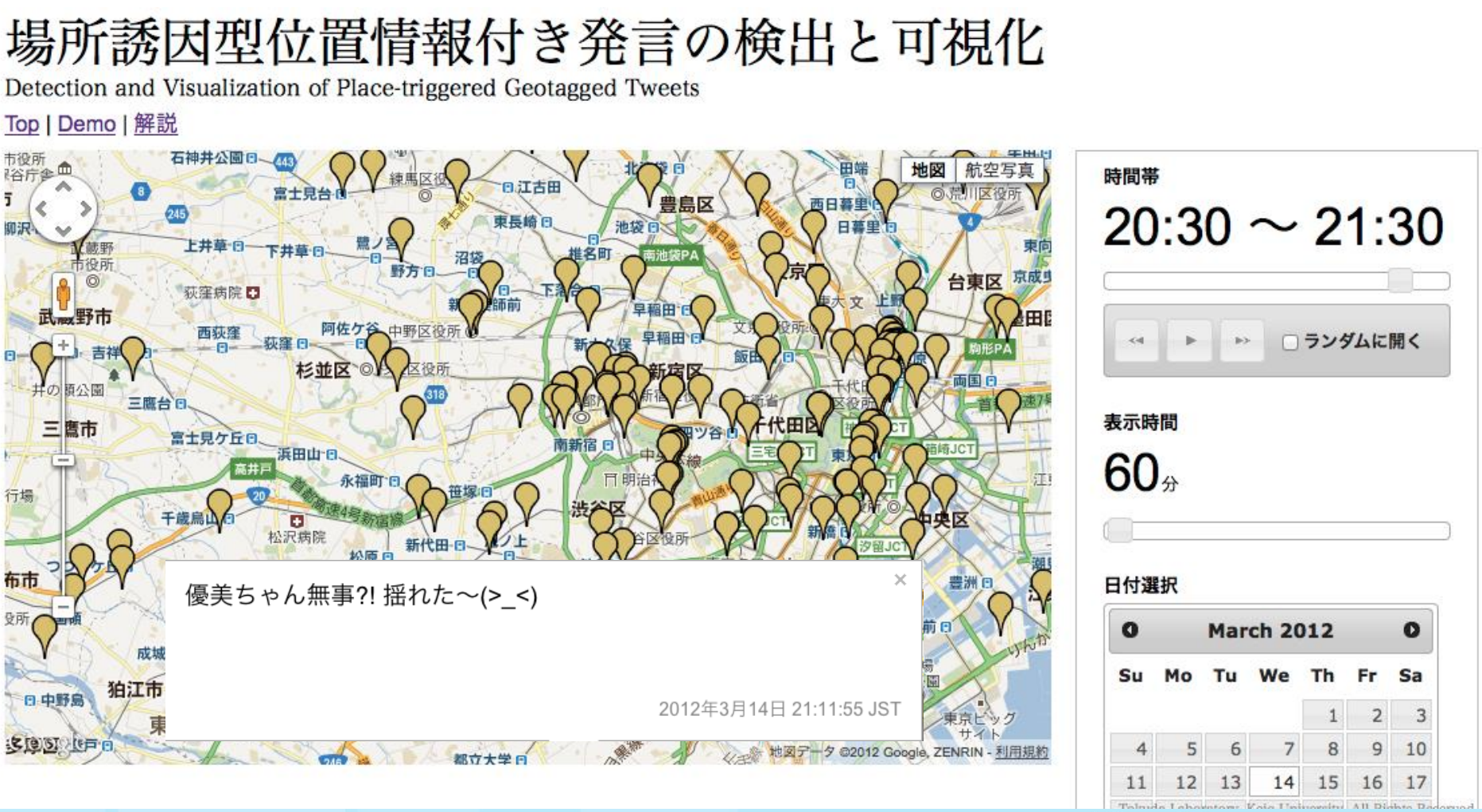
60分

日付選択

March 2012

Su	Mo	Tu	We	Th	Fr	Sa
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17

March 14, 2012 with Earthquake Filter



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

* Harmonic mean

Type of Tweets	Precision	Recall	F-measure
Report of whereabouts	93.18%	77.16%	84.42%
Food	53.6%	17.8%	26.7%
Weather	57%	21%	30%
Back at Home	54%	23%	32%
Earthquake	76%	66%	71%

Table 1. Classification result by the system

	Positive	Negative
TRUE	40.09%	False Negative 15.84%
FALSE	False Positive 2.18%	41.89%

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%

Microsoft Research

Thank you!

http://www.ht.sfc.keio.ac.jp/

Microsoft Research Asia
Faculty Summit 2012