

Earthquake Shakes Twitter User: Real-Time Event Detection By Social Sensors

Wang Yingze

CS 2310 Seminar

Twitter

❖ Twitter

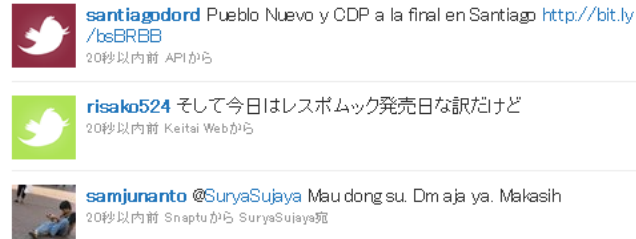
- is one of the most popular microblogging services
- Short message within 140 characters
- Real-time nature



Real-time Nature of Microblogging

social events
parties
baseball games
presidential campaign

disastrous events
storms
fires
traffic jams
heavy rain-falls
earthquakes



- There is a large number of tweets, which results in many reports related to events
- We can know how other users are doing in real-time
- We can know what happens around other users in real-time.

The motivation

- ❖ Adam Ostrow, an Editor in Chief wrote the possibility to detect earthquakes from tweets in his blog.

Japan Earthquake Shakes Twitter Users ... And Beyonce:

Earthquakes are one thing you can bet on being covered on Twitter first, because, quite frankly, if the ground is shaking, you're going to tweet about it before it even registers with the USGS and long before it gets reported by the media.*

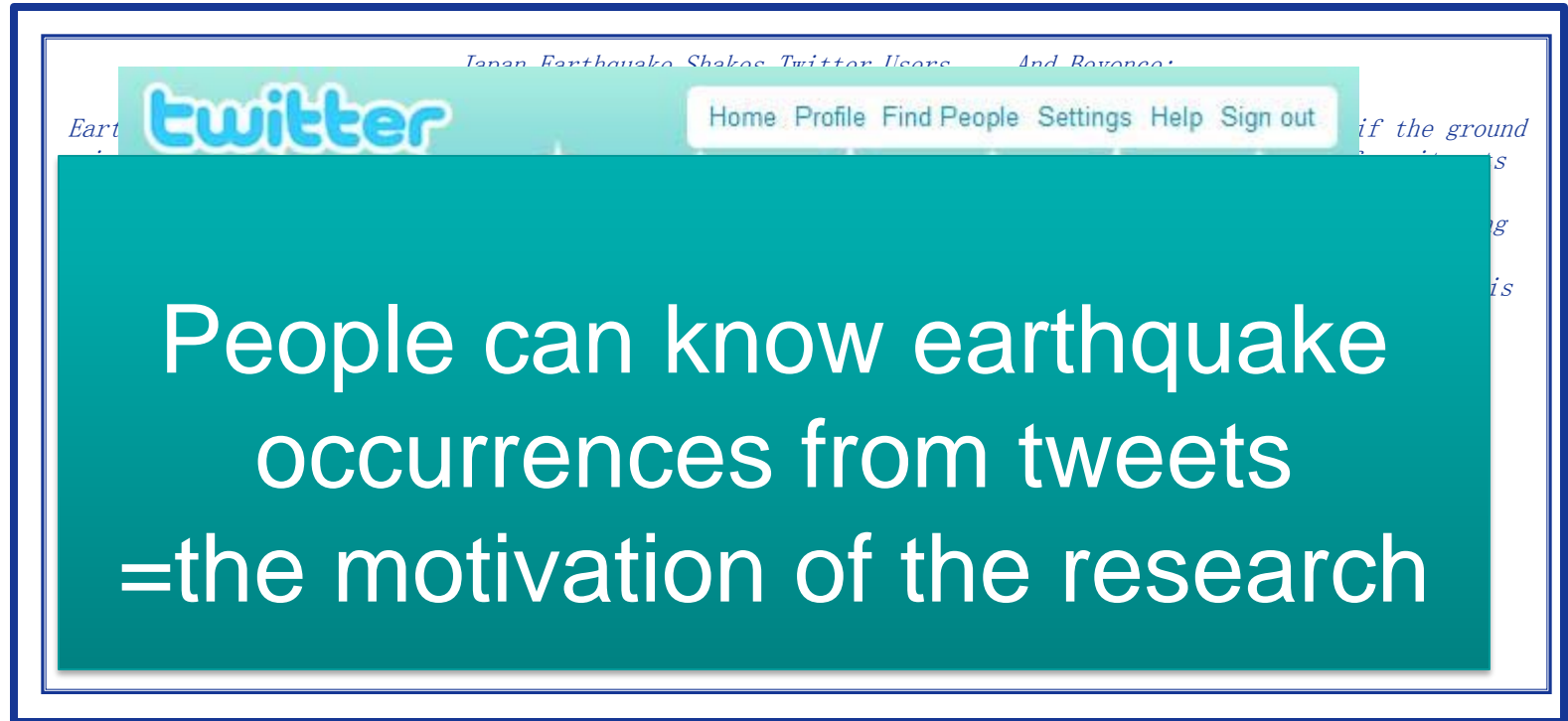
That seems to be the case again today, as the third earthquake in a week has hit Japan and its surrounding islands, about an hour ago.

The first user we can find that tweeted about it was Ricardo Duran of Scottsdale, AZ, who, judging from his Twitter feed, has been traveling the world, arriving in Japan yesterday.

*USGS : United States Geological Survey

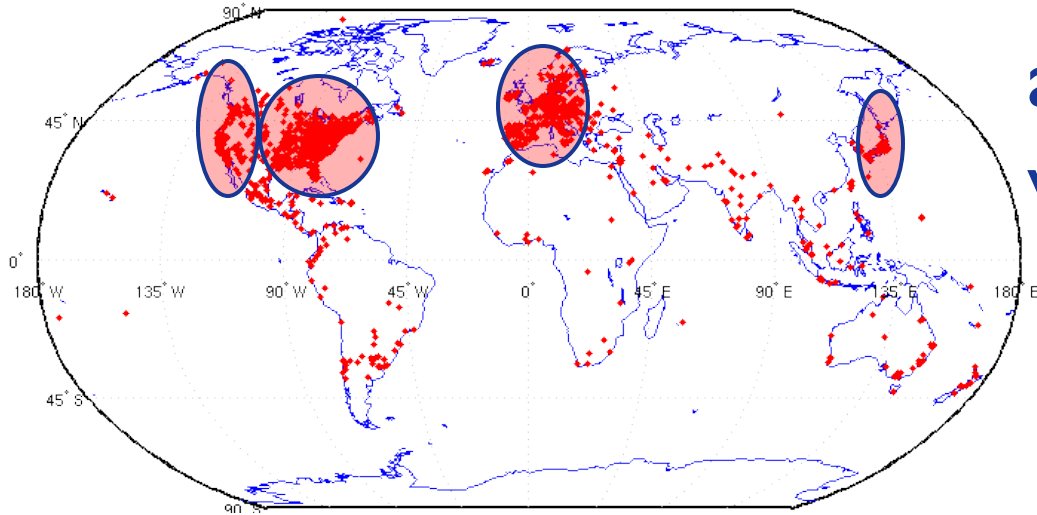
The motivation

- ❖ Adam Ostrow, an Editor in Chief wrote the possibility to detect earthquakes from tweets in his blog.

A screenshot of a Twitter page is shown within a blue-bordered frame. The page header includes the Twitter logo and navigation links: Home, Profile, Find People, Settings, Help, and Sign out. A teal rectangular box is overlaid on the page, containing white text. The text reads: "People can know earthquake occurrences from tweets =the motivation of the research".

People can know earthquake occurrences from tweets =the motivation of the research

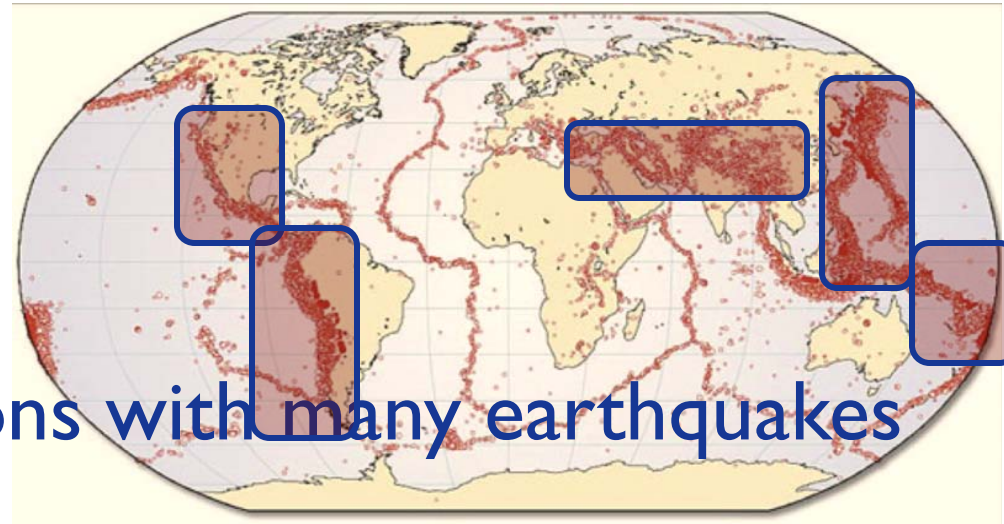
Twitter and Earthquakes in Japan



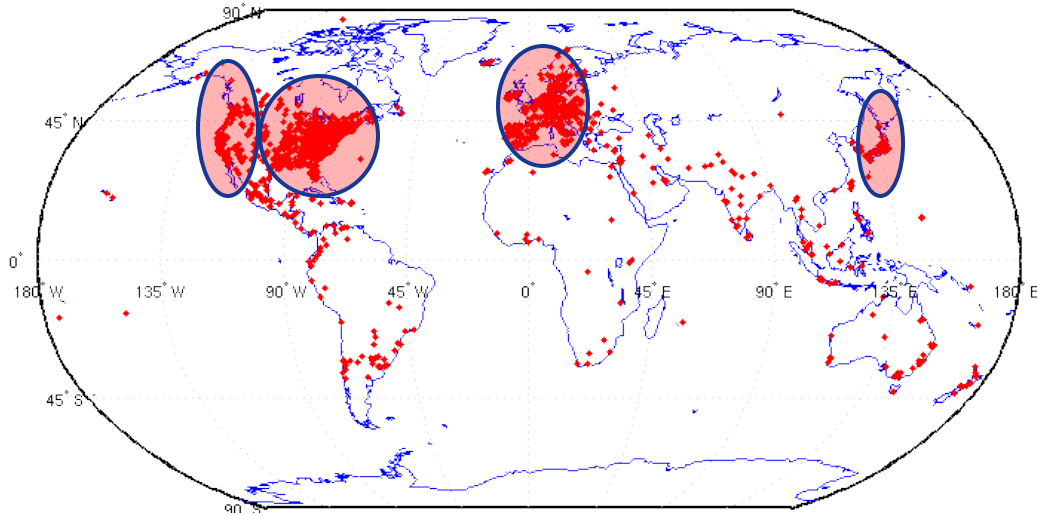
a map of Twitter user world wide

a map of earthquake occurrences world wide

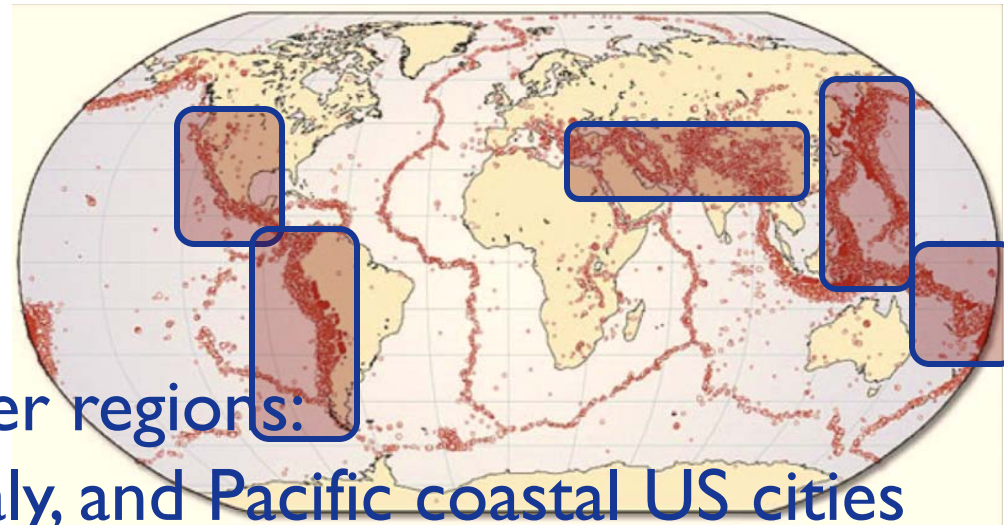
The intersection is regions with many earthquakes and large twitter users.



Twitter and Earthquakes in Japan



0



Other regions:

Indonesia, Turkey, Iran, Italy, and Pacific coastal US cities

The Goals

- ❖ Propose an algorithm to detect a target event
 - do semantic analysis on Tweet
 - ▶ to obtain tweets on the target event precisely
 - regard Twitter user as a sensor
 - ▶ to detect the target event
 - ▶ to estimate location of the target

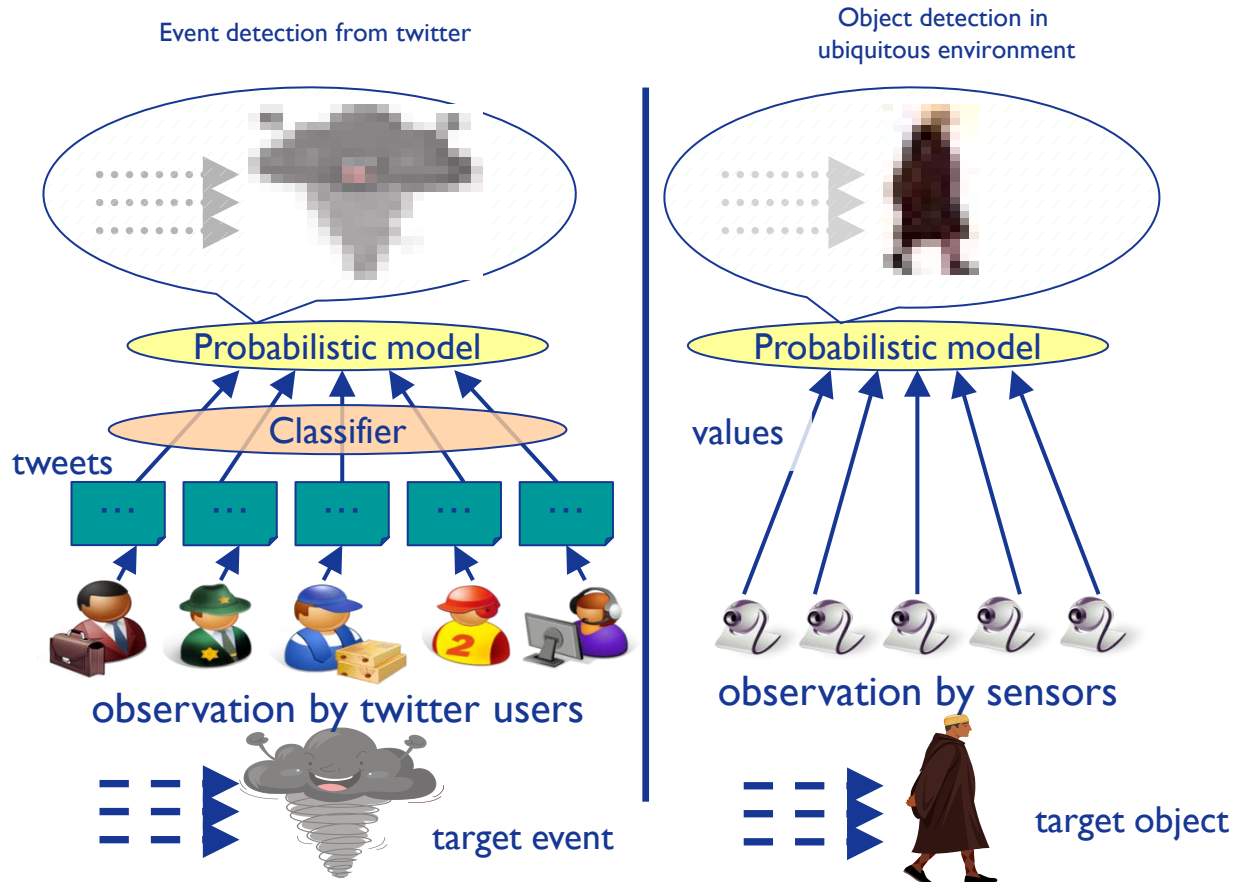
- ❖ Propose Earthquake Reporting System using Japanese tweets

Assumption: Tweet as a Sensory Value

- ❖ Two assumptions to apply methods for observation by sensors
 - Assumption 1: Each Twitter user is regarded as a sensor
 - ▶ a tweet → a sensor reading
 - ▶ a sensor detects a target event and makes a report probabilistically
 - ▶ Example:
 - ▶ make a tweet about an earthquake occurrence
 - ▶ “earthquake sensor” return a positive value
 - Assumption 2: Each tweet is associated with a time and location
 - ▶ a time : post time
 - ▶ location : GPS data or location information in user’s profile

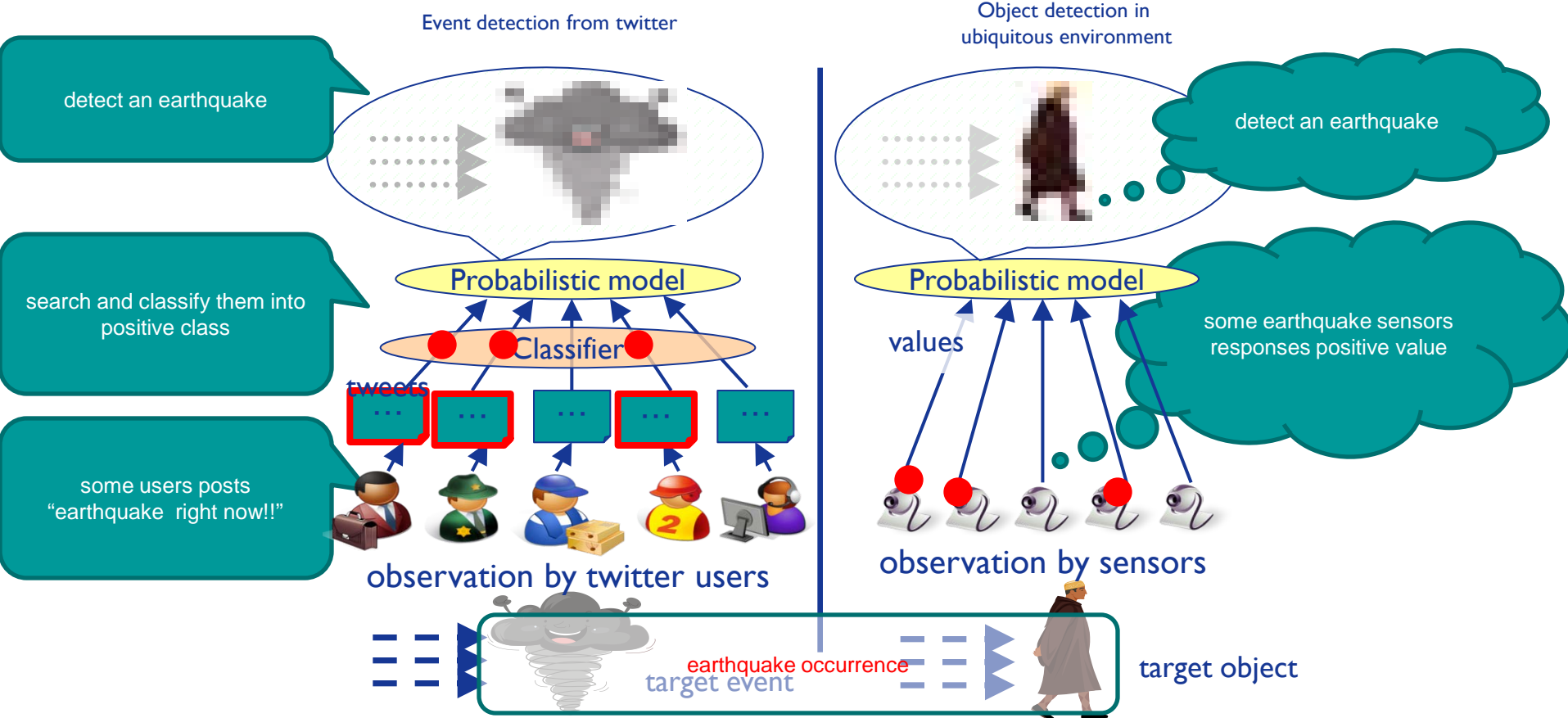
Processing time information and location information, we can detect target events and estimate location of target events

Tweet as a Sensory Value



the correspondence between **tweets processing** and **sensory data detection**

Tweet as a Sensory Value (Cont's)



We can apply methods for sensory data detection to tweets processing

Semantic Analysis on Tweet

- ❖ Search tweets including keywords related to a target event
 - Example: In the case of earthquakes
 - ▶ “shaking”, “earthquake”

- ❖ Classify tweets into a positive class or a negative class
 - Example:
 - ▶ “Earthquake right now!!” ---positive
 - ▶ “Someone is shaking hands with my boss” --- negative

Semantic Analysis on Tweet (Cont's)

❖ Create classifier for tweets

- use Support Vector Machine(SVM)

❖ Features (Example: I am in Japan, earthquake right now!)

- Statistical features (7 words, the 5th word)

the number of words in a tweet message and the position of the query within a tweet

- Keyword features (I, am, in, Japan, earthquake, right, now)

the words in a tweet and the number of each word in a tweet.

- Word context features (Japan, right)

the words before and after the query word

Probabilistic Model

❖ Why we need probabilistic models?

- Sensor values are noisy and sometimes sensors work incorrectly
- We cannot judge whether a target event occurred or not from one tweet
- We have to calculate the probability of an event occurrence from a series of data

❖ Paper propose probabilistic models for

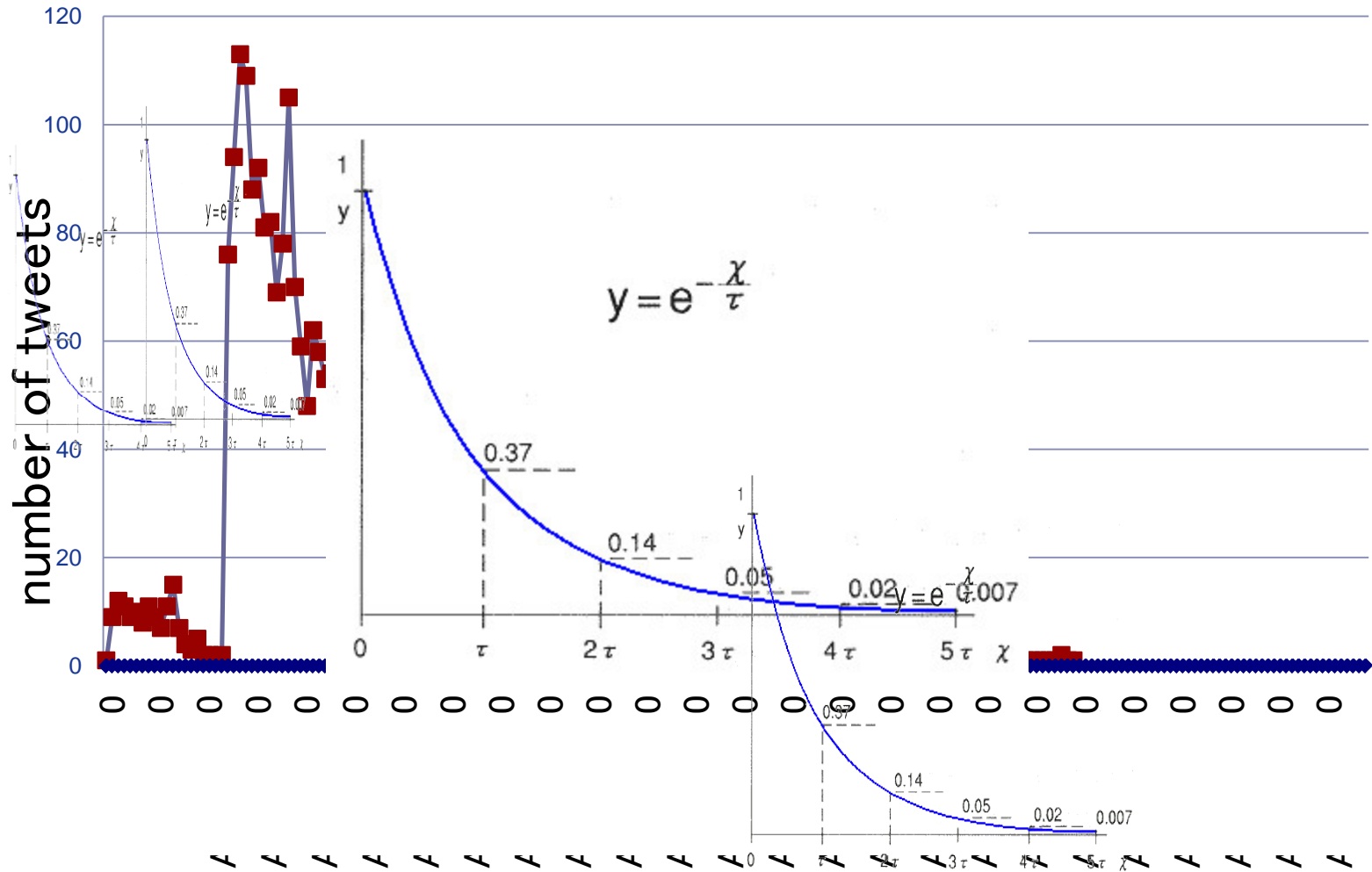
- event detection from time-series data
- location estimation from a series of spatial information

Temporal Model

- ❖ Calculate the probability of an event occurrence from multiple sensor values
- ❖ Examine the actual time-series data to create a temporal model

$$f(t; \lambda) = \lambda e^{-\lambda t} (t > 0, \lambda > 0) \quad \lambda = 0.34$$

Temporal Model (Cont's)



Spatial Model

- ❖ Calculate the probability distribution of location of a target
- ❖ Apply Bayes filters to this problem which are often used in location estimation by sensors
 - Kalman Filters
 - Particle Filters

Earthquake Reporting System

❖ Toretter

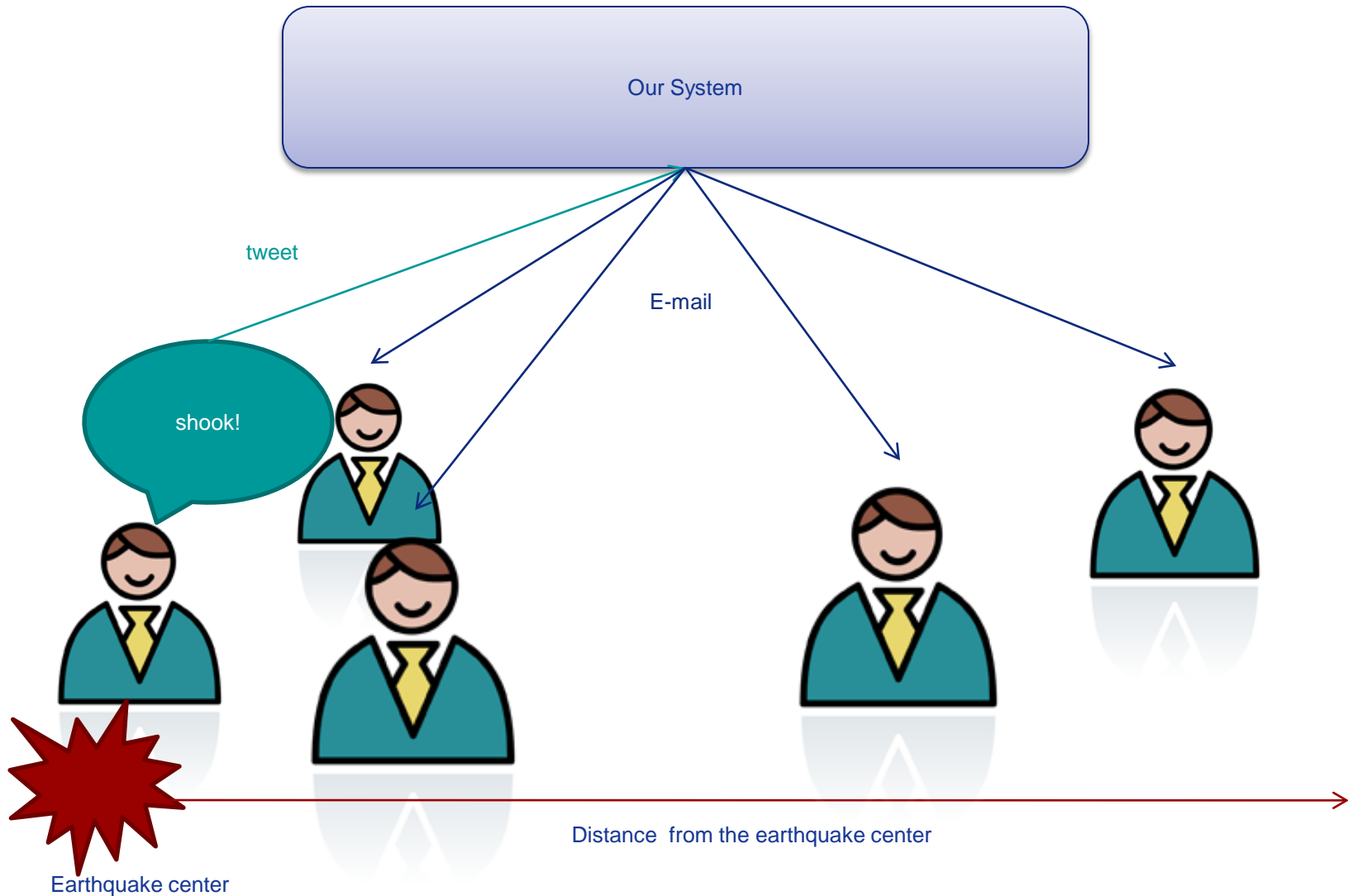
- Earthquake reporting system using the event detection algorithm
- All users can see the detection of past earthquakes
- Registered users can receive e-mails of earthquake detection reports

Dear Alice,

We have just detected an earthquake around Chiba. Please take care.

Toretter Alert System

Earthquake information system



Screenshot of Toretter.com



Published	Location	Title	Screen_name	URL
2009-08-11 05:08:57	Saitama, Japan	地震おおいわー	tondol	http://twitter
2009-08-11 05:08:56	unknown	地震。	tiroly	http://twitter
2009-08-11 05:08:53	iPhone: 35.509506,139.615601	揺れたね	Hakkan	http://twitter
2009-08-11 05:08:53	Mie Prefecture	すごい地震だ [mb]	narude531 masu	http://twitter
2009-08-11 05:08:52	Kawasaki city	地震だ！！	yaketasamma	http://twitter
2009-08-11 05:08:52	unknown	地震こわいですかんべん	wzcc	http://twitter
2009-08-11 05:08:52	Kansai	あら、地震？	Haru_iro	http://twitter
2009-08-11 05:08:52	Sakado, Saitama, Japan	地震だ	d_wackys	http://twitter
2009-08-11 05:08:51	unknown	愛知も揺れたw	edomain	http://twitter
2009-08-11 05:08:51	unknown	また地震 長いな...	laukaz	http://twitter
2009-08-11 05:08:51	JP	地震なう	echomutti	http://twitter

Callout boxes for English translations:

- Lots of earthquakes
- Earthquake.
- It shook.
- Terrible earthquake.
- Earthquake!!
- Earthquake! My gosh!
- Oh, earthquake?
- I feel earthquake!
- Shook Aichi
- Earthquake again. This is a long one...
- Earthquake now.

Results of Earthquake Detection

Date	Magnitude	Location	Time	E-mail sent time	time gap [sec]	# tweets within 10 minutes	Announce of JMA
Aug. 18	4.5	Tochigi	6:58:55	7:00:30	95	35	7:08
Aug. 18	3.1	Suruga-wan	19:22:48	19:23:14	26	17	19:28
Aug. 21	4.1	Chiba	8:51:16	8:51:35	19	52	8:56
Aug. 25	4.3	Uraga-oki	2:22:49	2:23:21	31	23	2:27
Aug.25	3.5	Fukushima	2:21:15	22:22:29	73	13	22:26
Aug. 27	3.9	Wakayama	17:47:30	17:48:11	41	16	1:7:53
Aug. 27	2.8	Suruga-wan	20:26:23	20:26:45	22	14	20:31
Ag. 31	4.5	Fukushima	00:45:54	00:46:24	30	32	00:51
Sep. 2	3.3	Suruga-wan	13:04:45	13:05:04	19	18	13:10
Sep. 2	3.6	Bungo-suido	17:37:53	17:38:27	34	3	17:43

In all cases, system sent E-mails before announces of JMA.

Conclusions

- ❖ Investigated **the real-time nature of Twitter** for event detection
 - Semantic analyses were applied to tweets classification
 - We consider each Twitter user as a sensor and set a problem to detect an event based on sensory observations
 - Location estimation methods such as Kaman filters and particle filters are used to estimate locations of events
- ❖ Developed **an earthquake reporting system**, which is a novel approach to notify people promptly of an earthquake event

Reference

- [1] Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo, “Earthquake shakes Twitter users: real-time event detection by social sensors”, In *Proceedings of the 19th international conference on World wide web (WWW '10)*. ACM, 2010, New York, NY, USA, 851-860.
- [2] Makoto Okazaki and Yutaka Matsuo, “Semantic twitter: analyzing tweets for real-time event notification”, *Recent Trends and Developments in Social Software (2010)*, Volume: 6045, Publisher: Springer Berlin / Heidelberg, Pages: 63-74

Thank You!

Q & A