twitter

Crowdsourced sensing & collaboration using Twitter

Murat Demirbas

SUNY Buffalo

Cellphones!

- 3-4B cellphone users worldwide
- 1.13 billion phones sold in 2009 (36 per sec) vs 0.3 billion PCs
- 174M were smart phones
- 15% (up from 12.8% in 2008)

Which is a better sensor?

Mote platform

- 4Mhz ATmega128 CPU, 8 bit microprocessor
- 8 Kb RAM, 128 Kb flash, 500Kb flash
- Chipcon CC2420 radio, 300Kbps
- Sensors: light, temperature, acceleration, magnetometer
- cost > $150

iPhone 3G

- 600mhz ARM processor, 256MB ram, 16GB flash: ~Pentium III
- WiFi, GSM, Bluetooth, USB
- Sensors: Camera, mic, GPS, compass, proximity, ambient light, ambient noise, 3D accelerometer, touchscreen, temperature
- + other integrated sensors

Go cellphones!

- Cared for by the user; battery life not a big problem
- Mobile; coverage is good
- Human intelligence comes free
- Impact: +10% mobile penetration => GDP up 0.8%
- Combine that with singlehop access to the cloud!

Status quo in cellphones

Each device connects to Internet to download/upload data and accomplish an individual task that does not require collaboration and coordination

What is missing?

An infrastructure to task/utilize these devices for collaboration and coordination

Any node should be able to search & aggregate the data published by other nodes in a region, as well as task nodes in the region to acquire the needed data

Our goal

We provide a crowdsourced sensing and collaboration service using Twitter to enable aggregation and sharing of data as well as tasking of other cellphones

Why Twitter?

- Open publish-subscribe system
  - different actors may integrate published data differently, and in unanticipated ways to offer new services
- Social networks aspect is useful for crowdsourced sensing and collaboration applications
- 105 million users, over 30 million users in US, 55 million tweets 600 million search queries everyday

Twitter is simple

- Each tweet has 140 char limit
- Twitter provides an open source Search API and a REST API (enables developers to access tweets, timelines, and user data)

Our contributions

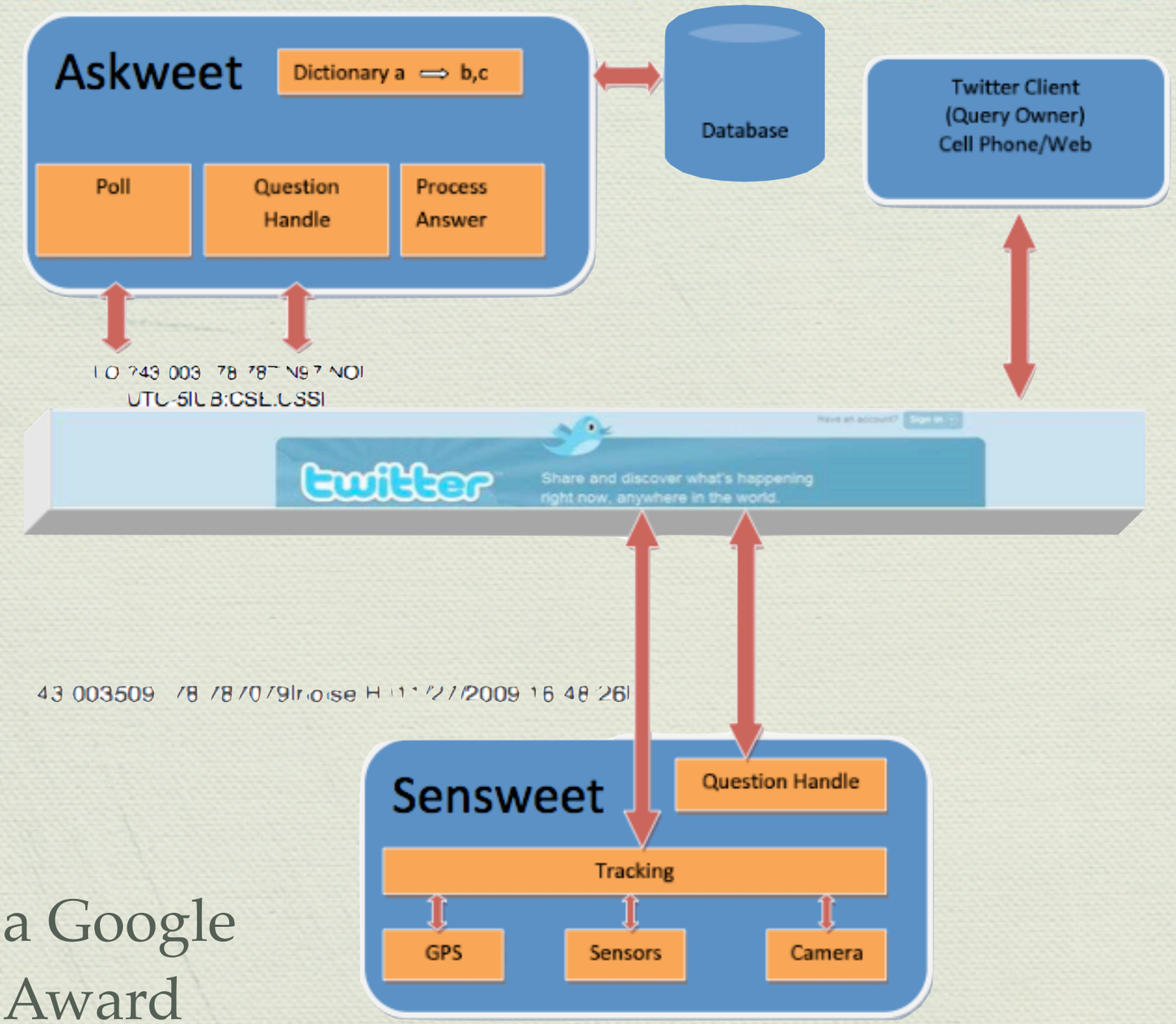
- Sensor & smartphone integration to Twitter
- Twitter bot architecture for crowdsourcing
- Experimental results: Rainradar, Local queries, NoiseMap
- Our current work: Upinion, CityPulse, Public health
- Future research directions

Sensor / phone integration

To search and process sensor values on Twitter, we need to agree on a standard for publishing these sensor readings

- Bio-code: Uses Twitter bio sections & allows users to search for the sensors they are looking for on-the-fly
- TweetML: Uses hashtags for searchability

Crowdsourcing architecture



Funded by a Google Research Award

Sensweet

Employs the smartphone's ability to work in the background without distracting the mobile user. Sensweet applications sense the surrounding environment and send these data to Twitter

Askweet

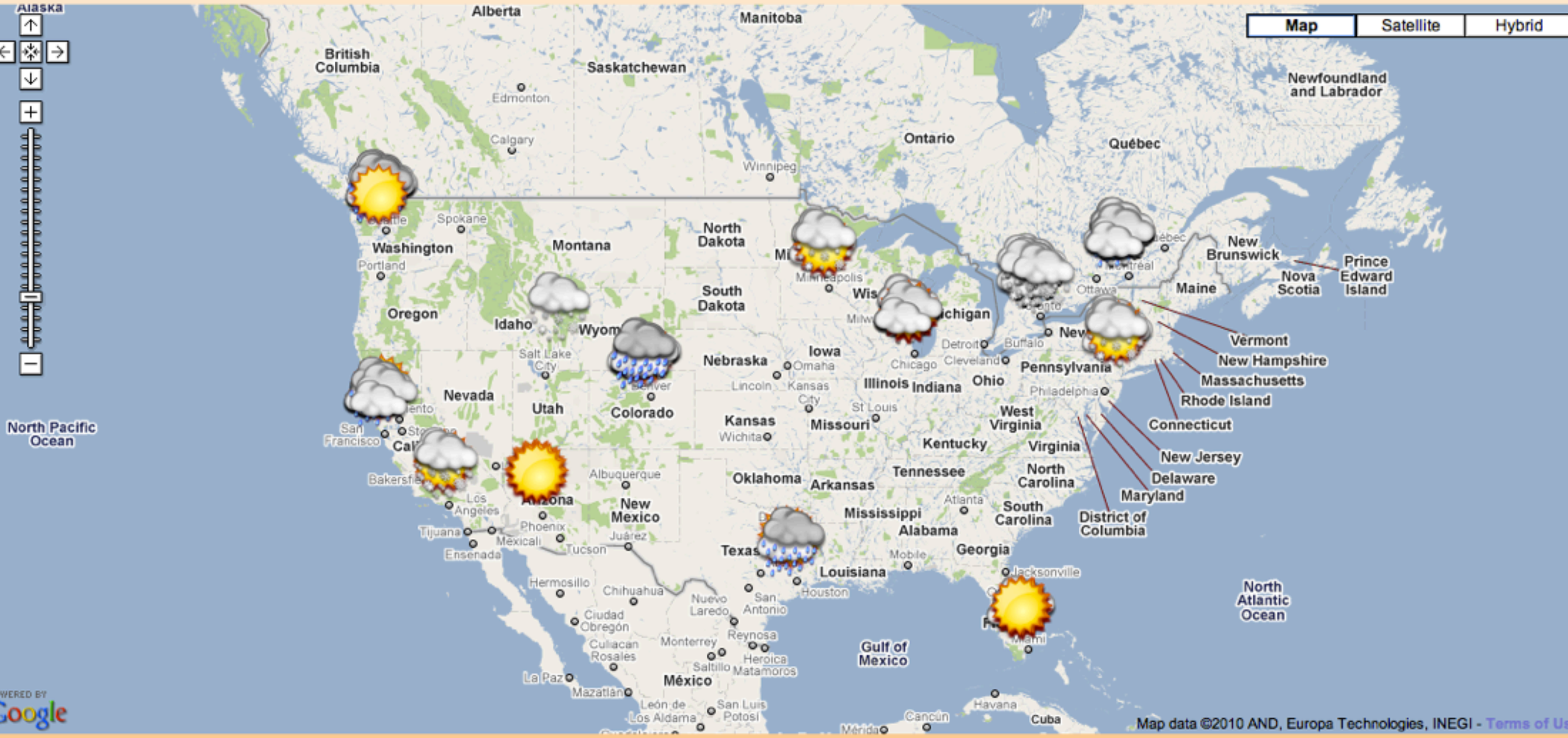
- Accepts a question from Twitter;
  - tries to answer the question using the data on Twitter, potentially data published by Sensweets
  - if that is not possible, Askweet finds experts on Twitter and forwards the question to these experts
- Parallelizable, easy to “cloudify” for elastic scalability

Crowdsourced weather

- Current weather, everybody on Twitter can be an expert
- Question to Askweet: “?Weather Loc:Buffalo,NY”
- Forwarded question:“How is the weather there now? reply 0 for sunny, 1 for cloudy, 2 for rainy, and 3 for snowy”

MAP WHAT IS RAINRADAR? CROWD-SOURCED SENSING PROJECT TE

http://ubicomp.cse.buffalo.edu/rainradar

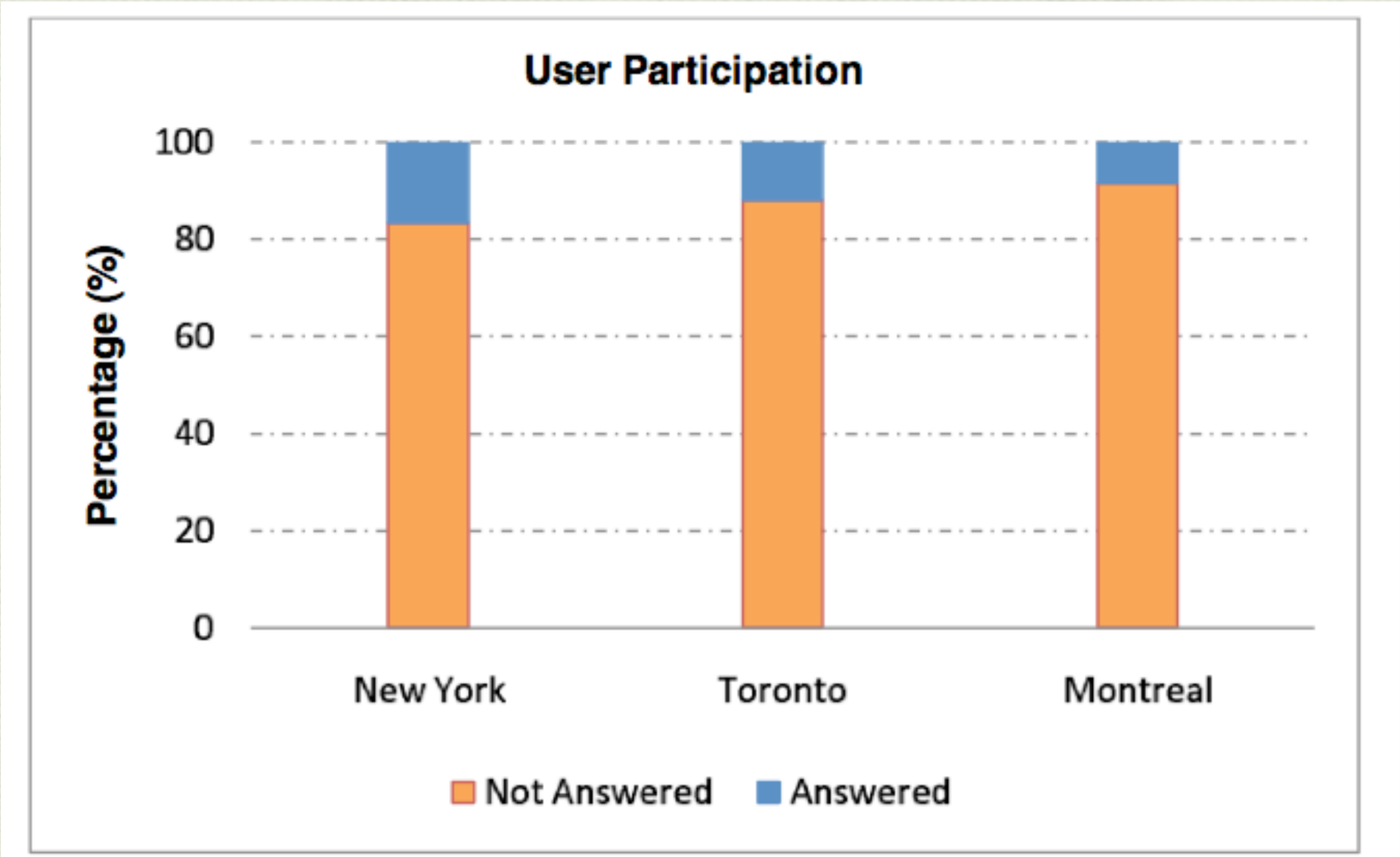


Select Date from here : Selected date is : 2010-01-05

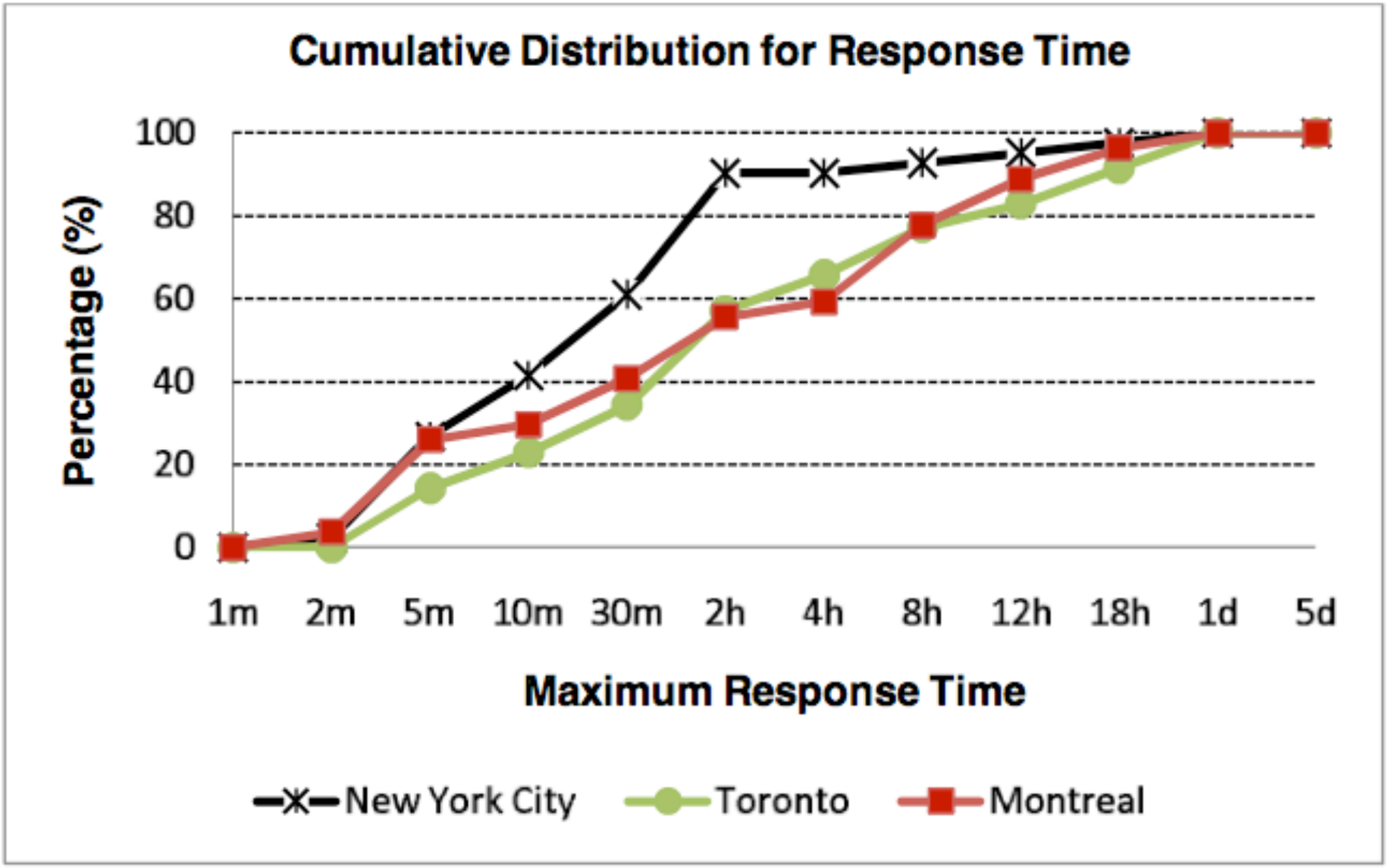
The map is configurable to show results from previous days, and also is zoomable to show fine-grain locations of the replies.

You can select dates between **December, 3 2009** and **January, 6 2010** to see the results!

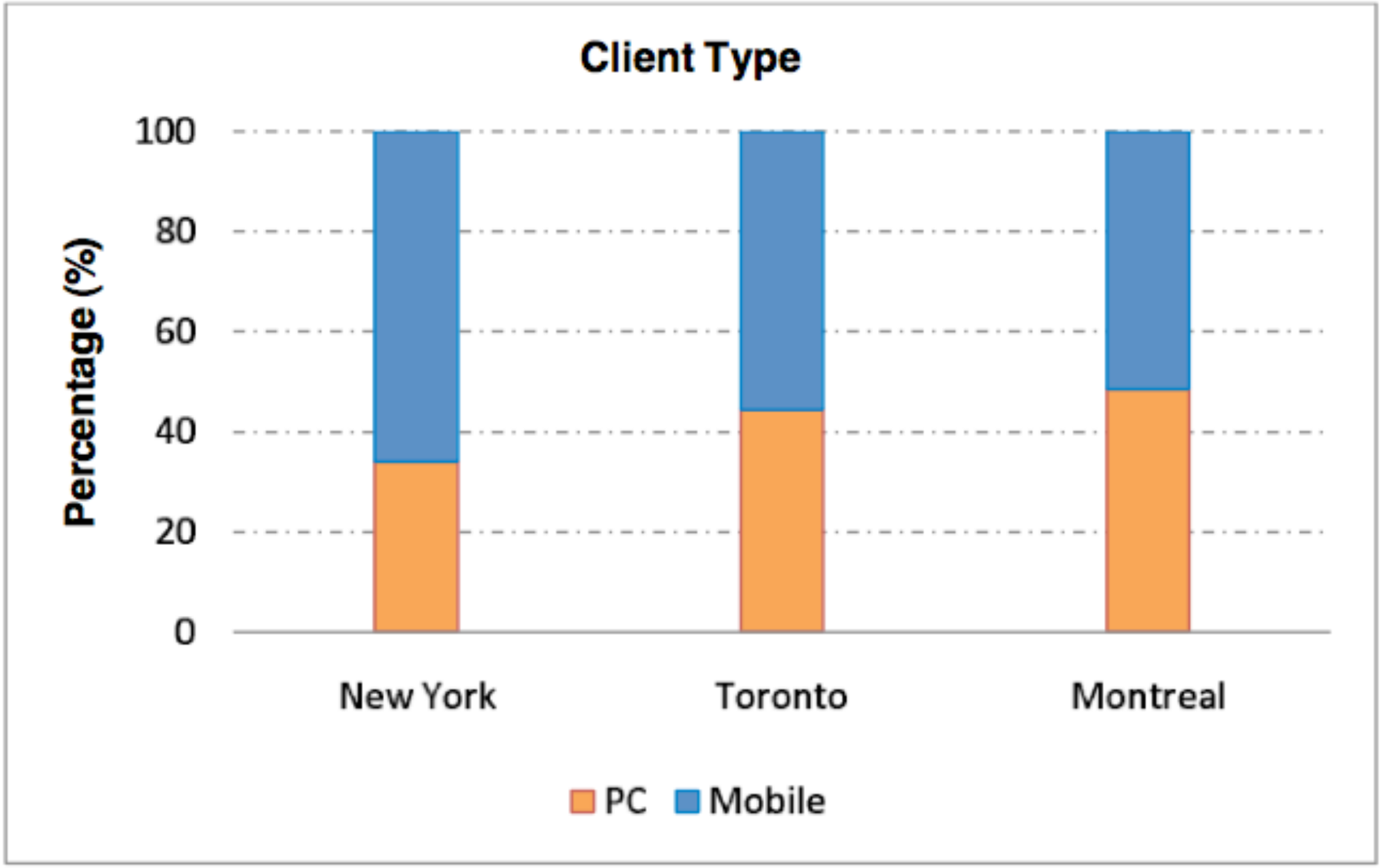
User participation



Response time



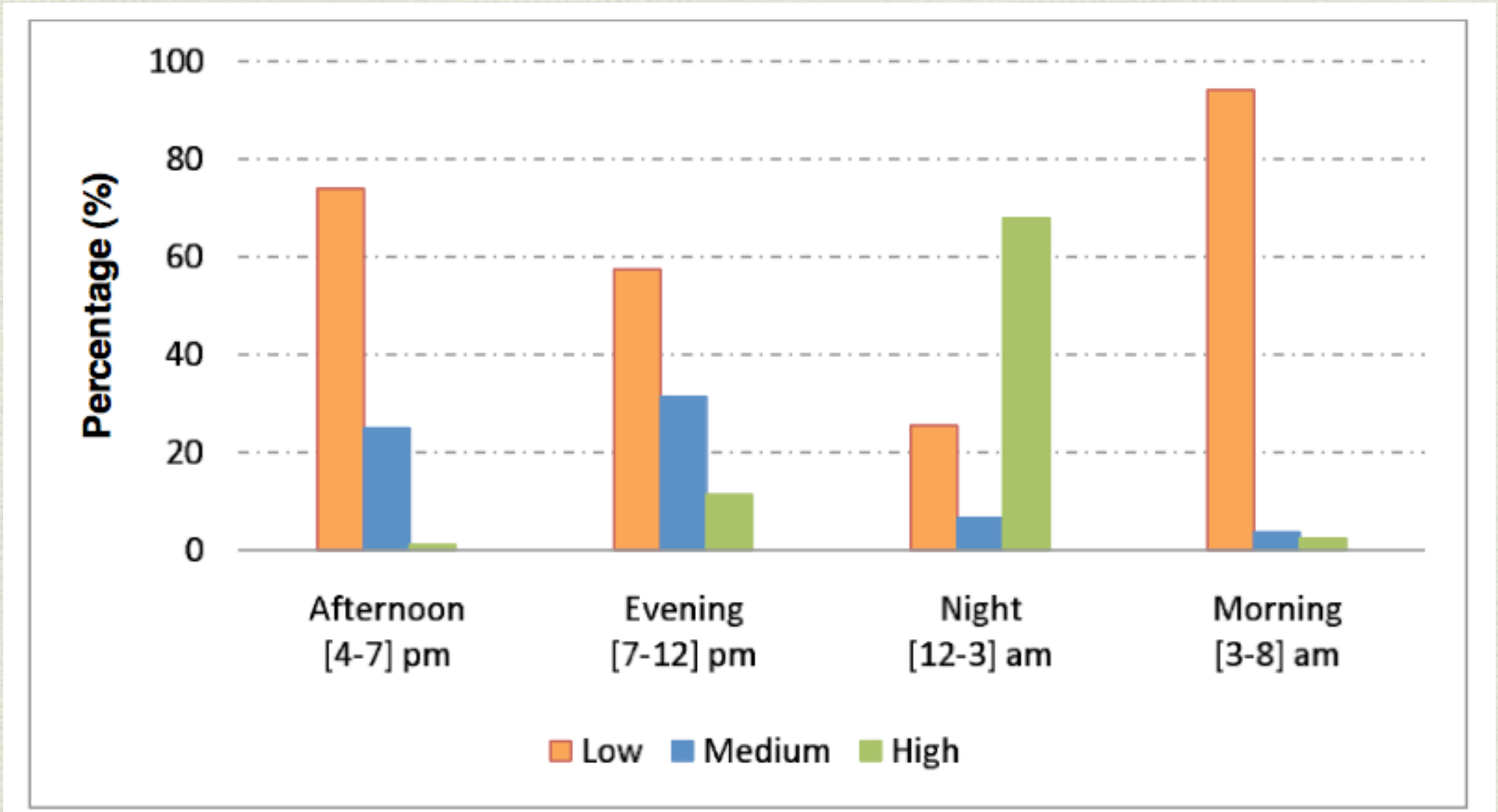
Smartphone ratio

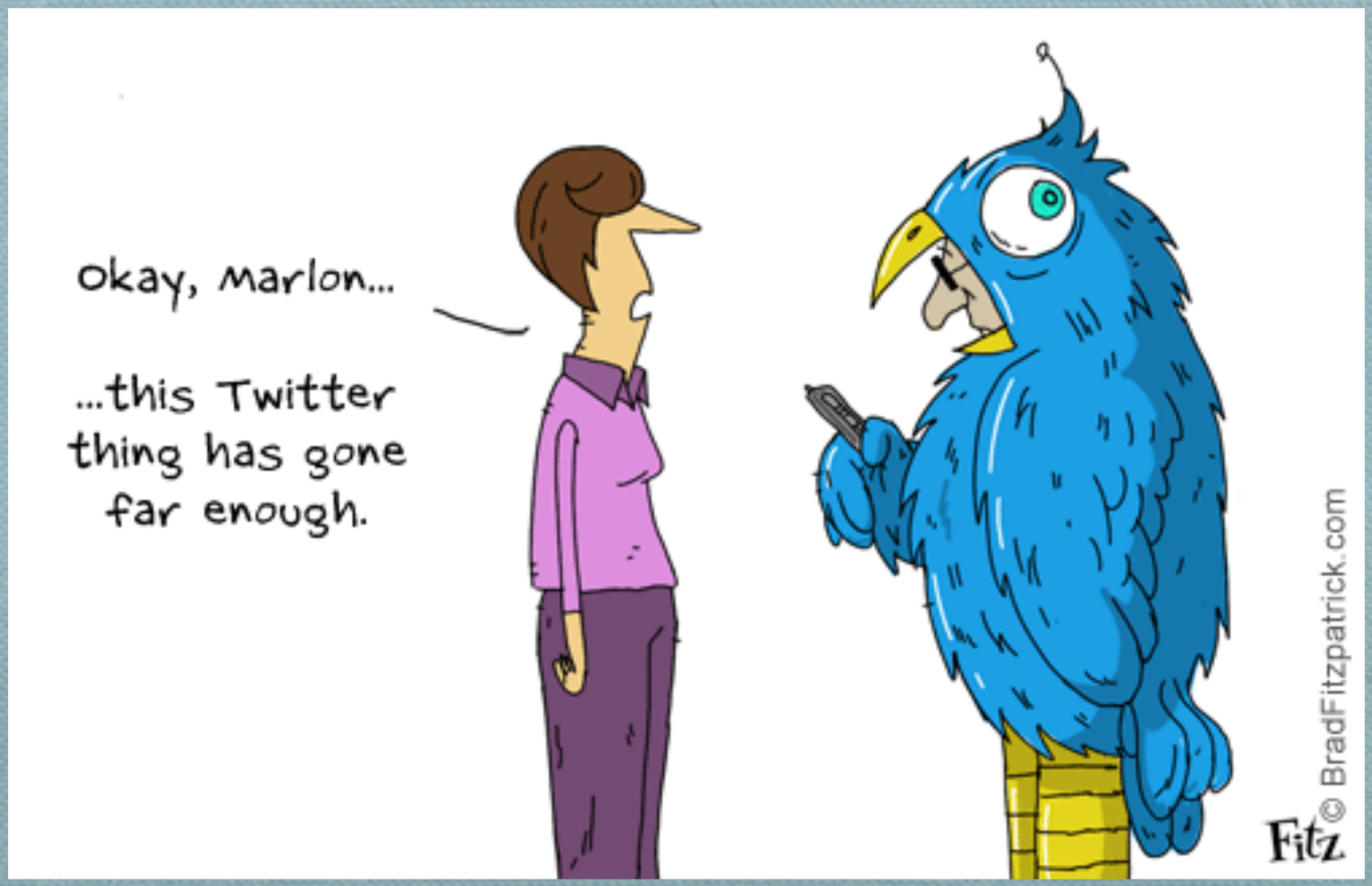


Noise map application

We implemented a Sensweet client for the Nokia N97 Smartphone series.The Sensweet client detects the noise level of the surrounding environment and forwards this data to Twitter using our TweetML format

Noise levels for a user





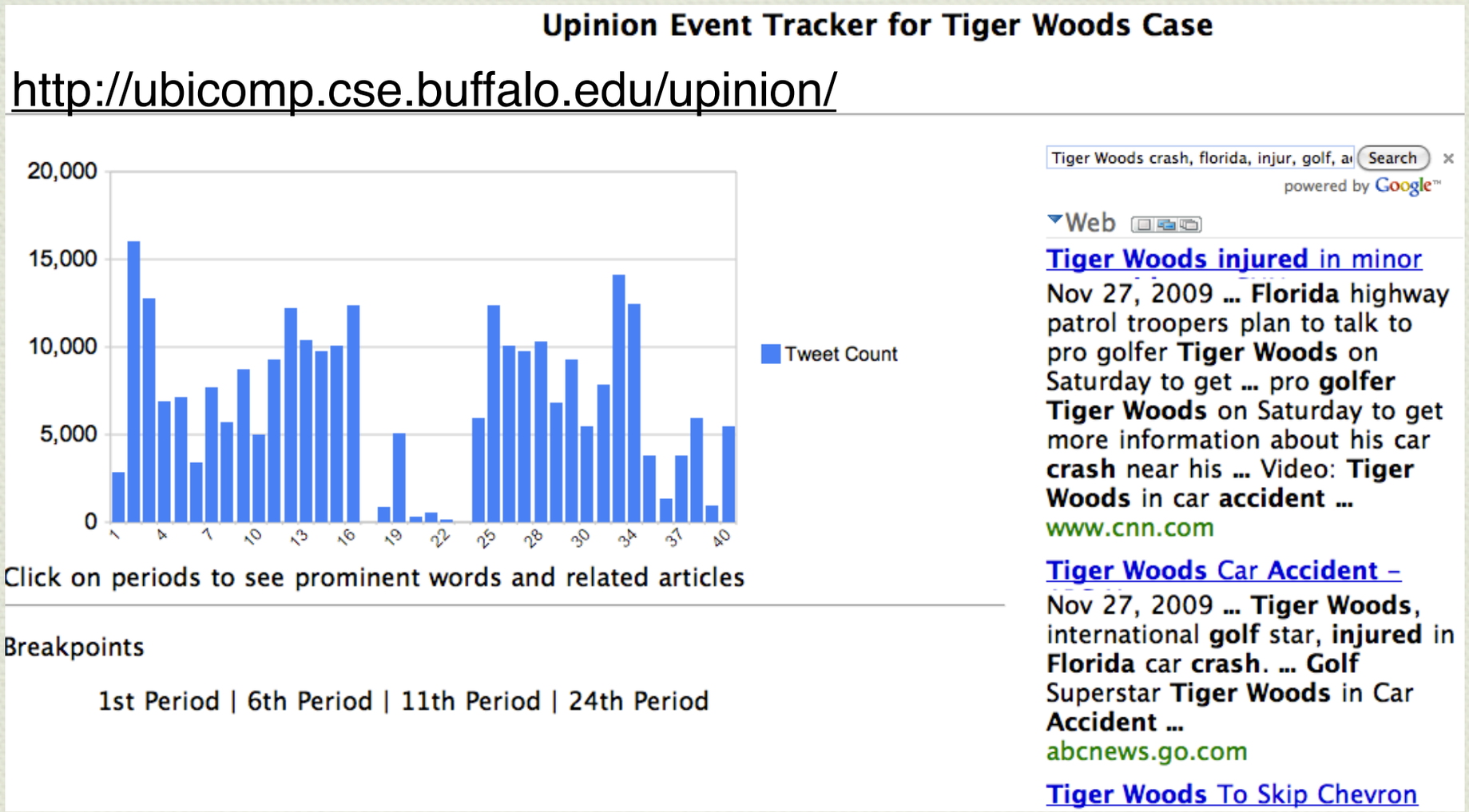
Our current work

Local querying via Twitter

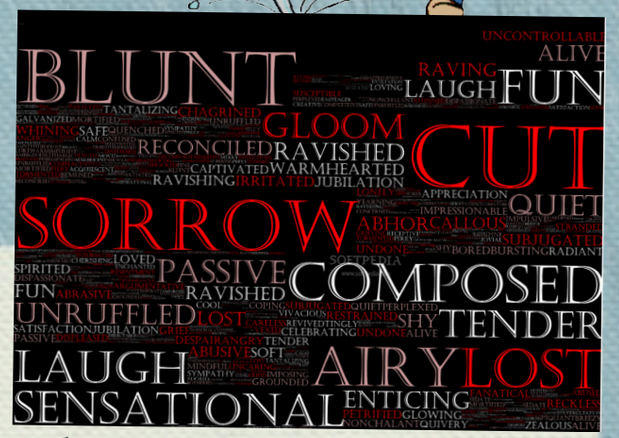
- We integrate mobility profiling to our Twitter crowdsourcing architecture to answer location-based queries more effectively
- We use category and location information provided by Foursquare to direct questions to mayors in queried locales
- We get better answers for food, nightlife, colleges
- We can answer 75% of both factual and non-factual queries, Google answers 78% answer rate of factual queries and 29% of non-factual queries

Upinion

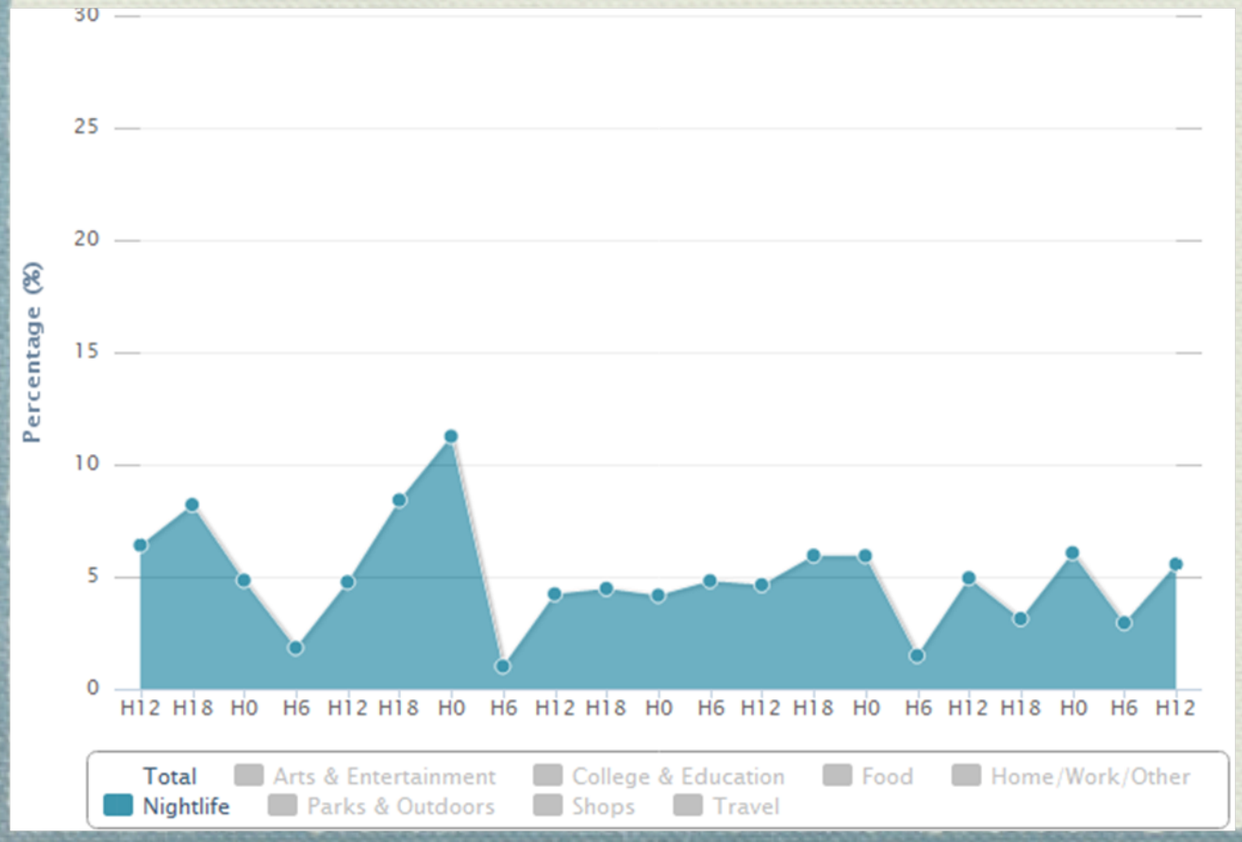
- Data mining over Twitter for identifying breakpoints in public opinion for a given keyword

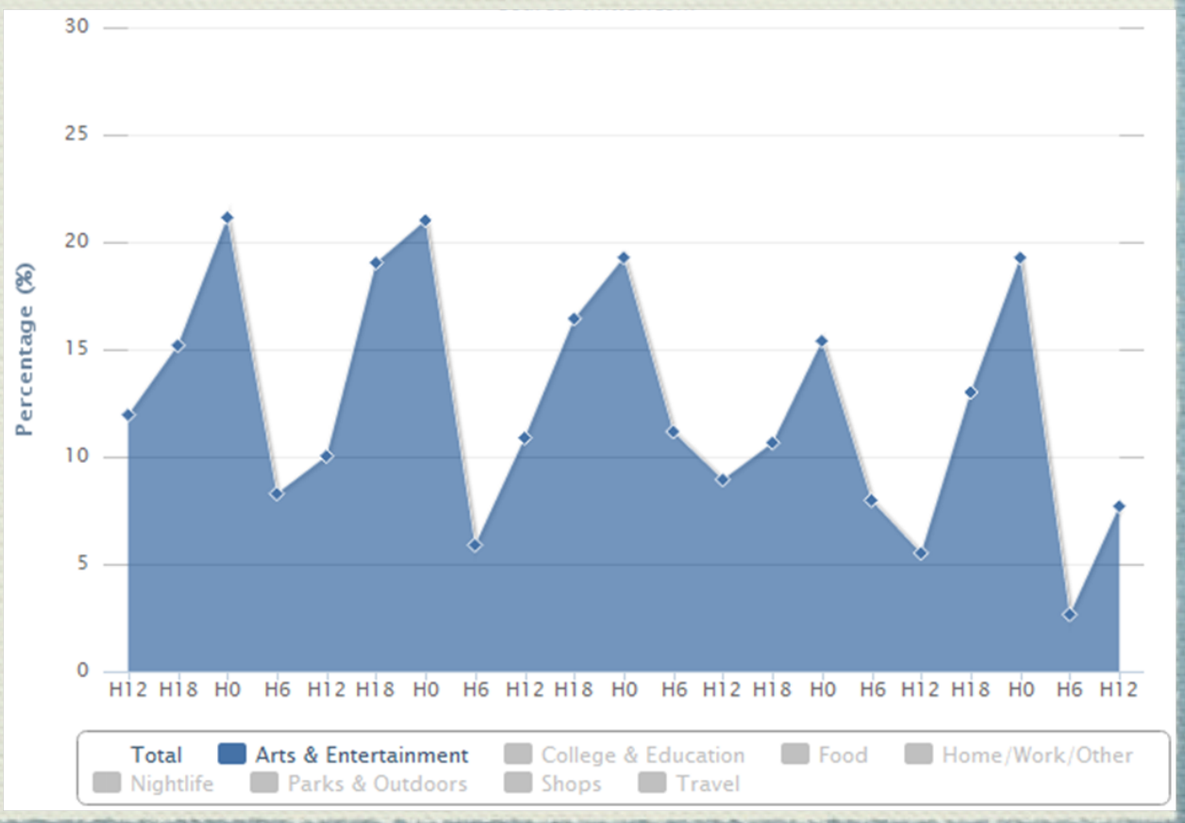


CityPulse



- We adopted the upinion framework to monitor changes in location related tweets in cities.





Air pollution exposure

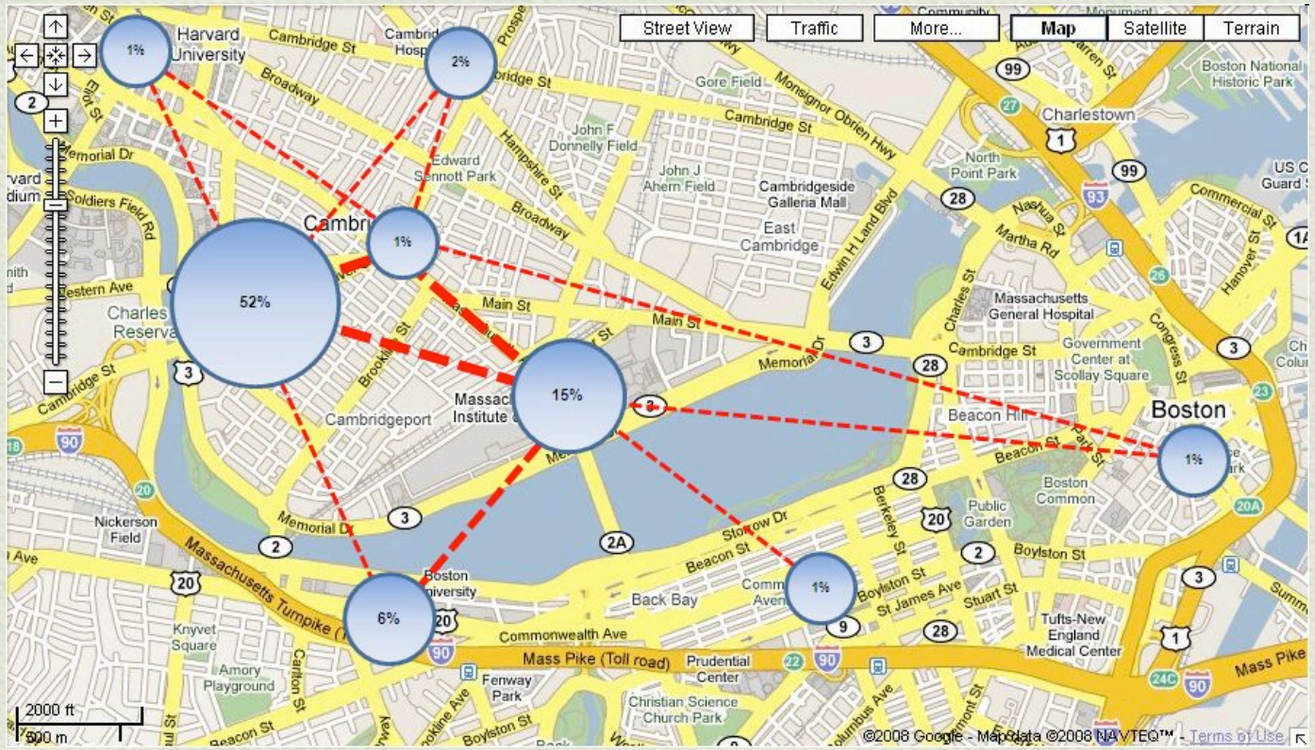
- 2 million premature deaths due to air pollutants yearly; cardiovascular & respiratory, asthma, low birth weight

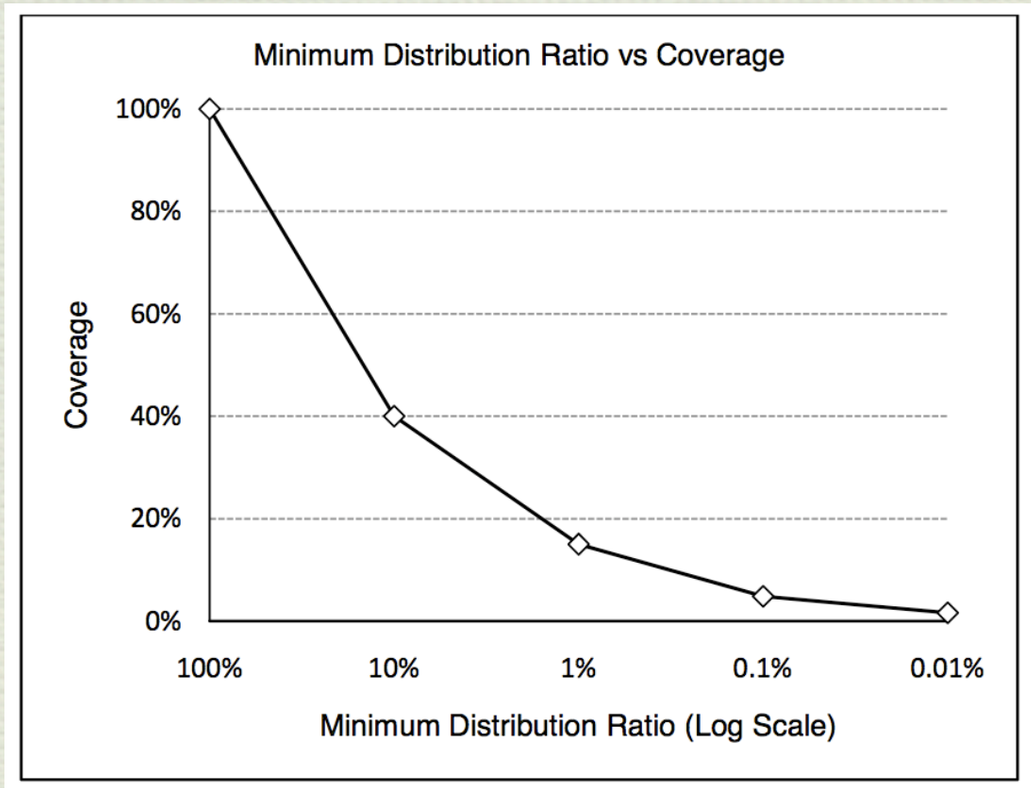
- Current approaches use questionnaires to determine spatiotemporal behavior (top-k locations), and use regression to calculate risks

- They ignore travel time and little time-spent location, but this leads to inaccurate results, using mobility logs and profiles yields more accurate results

Funded by: NIH/NIEHS

Long tail in distribution





- We spend about 85% of our time in our top 3 locations. The remaining 15% shows heavy tail distribution.

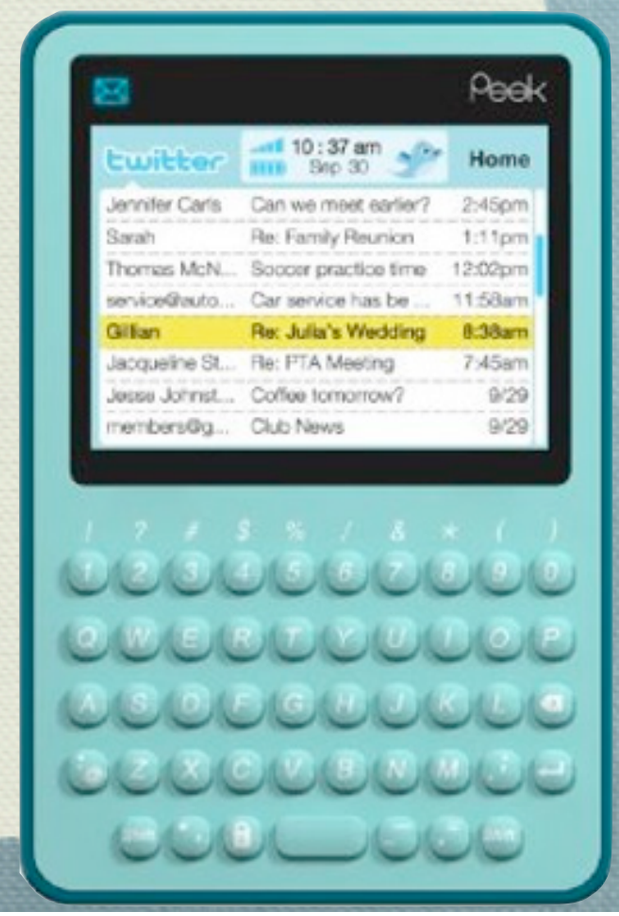
Future directions

Related work

- Participatory sensing, UCLA (air pollution monitoring, invasive plant species)
- People-centric sensing, Dartmouth Univ. (BikeNet, CenseMe, SoundSense)
- Crowdsourcing/coordination has not been the focus in these
- Crowdsearch distributed image search on phones, UMASS
- Microblogging + smartphones + collaboration, Duke

New WSN architectures

- TwitterPeek connects to Twitter from anywhere using GSM
  - Costs $199 for the lifetime of the device –without any bills
- TwitterPeek sensors can directly upload to Internet singlehop, and can be easily reconfigured over Twitter



DARPA's grand challenge

Find 10 balloons in US accurately and quickly

Exploring the roles the Internet and social networking play in the timely communication, wide-area team-building, and urgent mobilization required to solve broad-scope, time-critical problems

Spatiotemporal search

- What is the noise level at student union?
- What is the waiting time at the coffee shop?
- Whom can I share a ride to downtown?
- What could a guy like me do for a good time in this town?

These all require some crowdsourcing/tasking

Passive sensing to active

- We search for “noisy street”, “pollution”, “traffic accident”
- Then, when an event is detected, we direct questions to users in that area to get more information

Social collaboration

Requires back-and-forth interaction in contrast to the asymmetric one-shot interaction involved in crowd-sourcing. Examples of social collaboration applications include pick-up soccer games, arranged ride-sharing, community-organization events.

E.g., Gov2.0

E.g., http://www.redditisland.net/

PhoneCloud

- We are planning to build a 1000 phone reprogrammable testbed at UB
  - Geoffrey Challen, Murat Demirbas, Steve Ko, Tevfik Kosar
- Dense, controlled, yet realistic environment for testing and developing next generation collaborative smartphone apps and operating systems

Other open problems

- Security and trust: http://pleaserobme.com/
- Incentive mechanisms
- Sustainable deployments

Questions?