EPL646 – Advanced Topics in Databases

Lecture 18

Mobile Crowdsourcing

Demetris Zeinalipour

http://www.cs.ucy.ac.cy/~dzeina/courses/epl646

Lecture Outline

- ***Mobile Crowdsourcing***
  - *Introduction, New Challenges (compared to Web CS)*
  - *Crowdsourcing in Developing Regions*
  - *Urban Sensing and Traffic Engineering*
  - *Content-based Multimedia Search & Q/A Systems*
  - *Spatial (Outdoor and Indoor) and Spatio-temporal Systems.*
  - *Testbeds and Infrastructures*
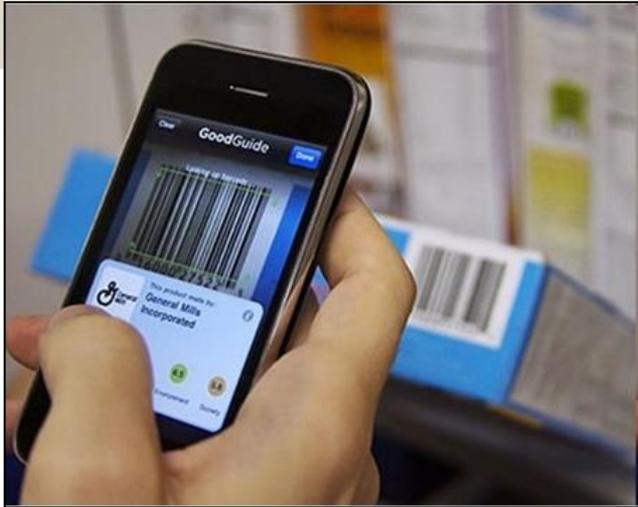
Smartphones

- **Smartphone:** a mobile device (phone, tablet, slate) that offers more *computing ability* than a basic **feature phone** (e.g., one running JavaME) and a "**dumb" phone**.
  - *Computing Ability: CPU*, Memory & Storage, Networking, Sensing.
- **Typical Features**
  - **CPU:** 1 Ghz Dual core / 1.5Ghz Quad-core (Qualcomm Snapdragon S4)
  - **RAM & Flash:** 1GB & 48GB / 2GB & 32GB
  - **Networking:** WiFi, 3G (Mbps) / 4G (100Mbps–1Gbps)
  - **Sensing:** Proximity, Ambient Light, Accelerometer, Microphone, Geographic Coordinates based on AGPS (fine), WiFi or Cellular Towers (coarse), Camera (13MB!)
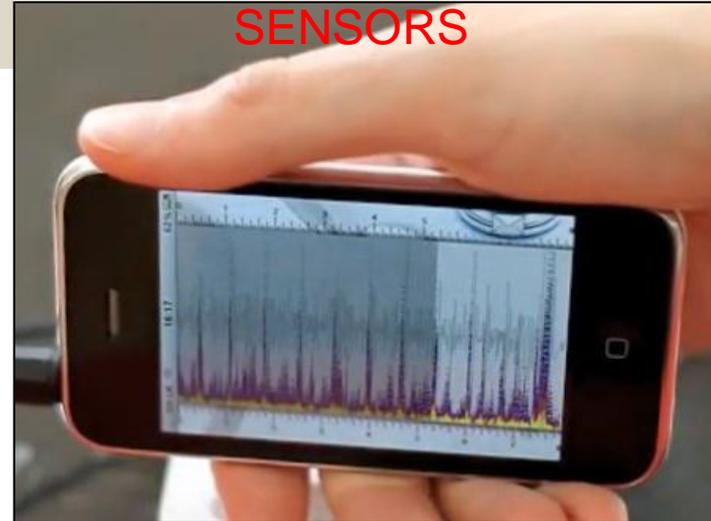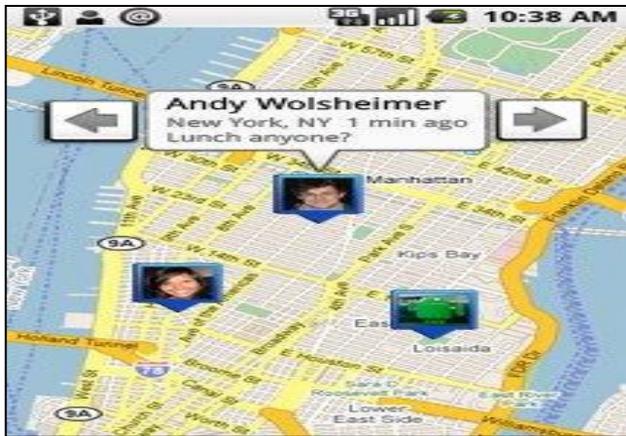
From Web to Mobile Crowdsourcing

- Obviously, all **Web Crowdsourcing** systems can be extended to the mobile world (with apps or HTML5)

- *However, this is **NOT** the **objective** of this second part.*

- *We strive to uncover the **intrinsic mobile & sensing characteristics** that make Smartphone Crowdsourcing a real new challenge*

    – *N Sensors => 2^N (powerset of N) crowdsourcing possibilities!*

Courtesy of: http://www.qualitativemind.com/trend-tamer/crowdsourcing/

From Web to Mobile Crowdsourcing

Camera: Find the right coupons on the right moment!

SENSORS

Microphone: Medical Stethoscope.

GPS/WIFI/Cell: Smartphone Social Networks
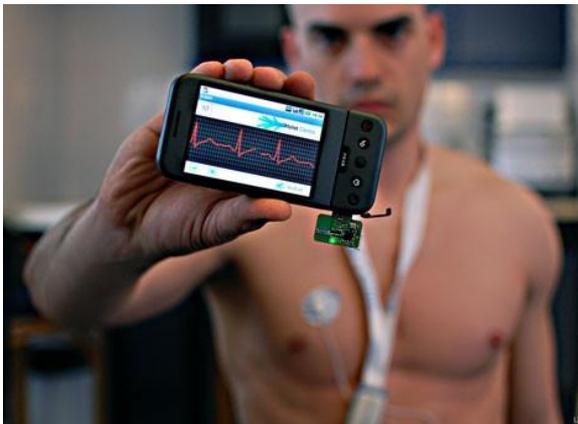

Compass / Accelerometer: Augmented Reality

ses - Demetris Zeinalipour (University of Cyprus)

From Web to Mobile Crowdsourcing

SENSORS

Nike+Apple

Movement Sensors for Athletes

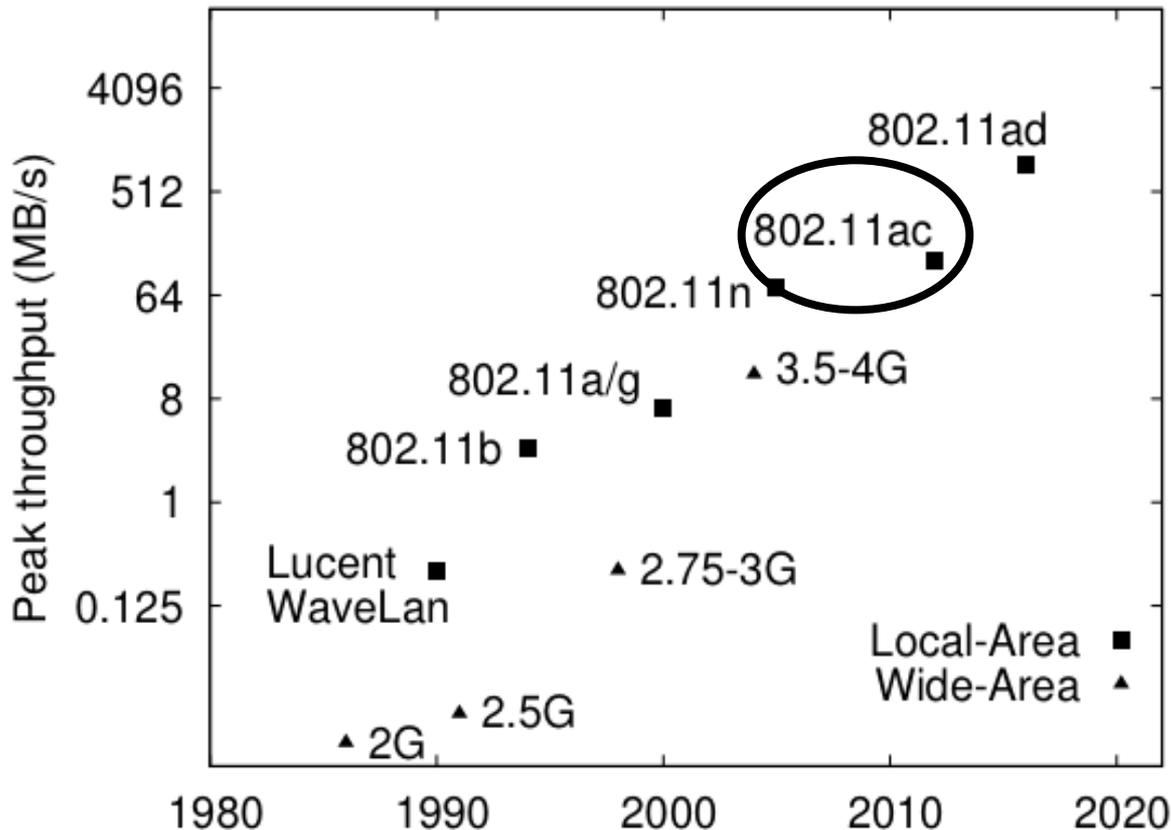Body Sensors: ECG, etc.

Urban Sensing: CO_2, etc.

From Web to Mobile Crowdsourcing

Wireless Data Transfer Rates BANDWIDTH

4G ITU peak rates:
- 100 Mbps (high mobility, such as trains and cars)
- 1Gbps (low mobility, such as pedestrians and stationary users)

*Plot Courtesy of H. Kim, N. Agrawal, and C. Ungureanu, "Revisiting Storage for Smartphones", The 10th USENIX Conference on File and Storage Technologies (FAST'12), San Jose, CA, February 2012. *** Best Paper Award ****

The Post PC era

Oct. 8, 2011. The Economist. "Beyond the PC"

Post-PC era

02/2012: Canalys

canalys
Insight. Innovation. Impact.
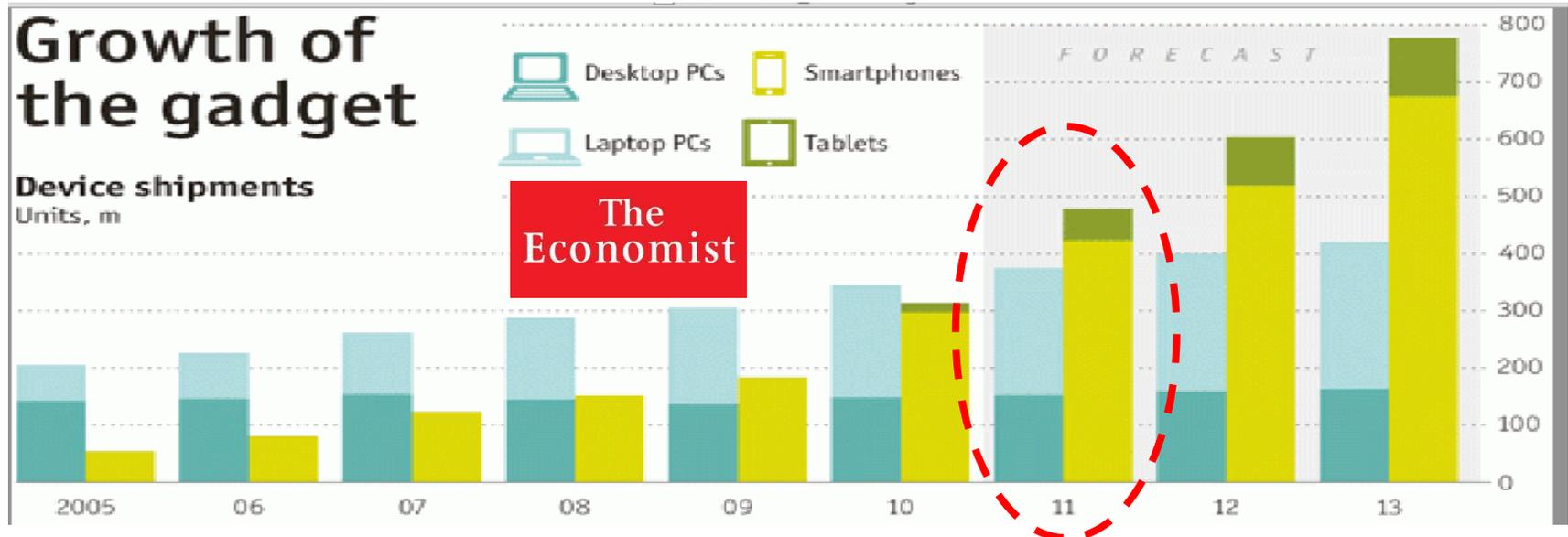
| Device	Ship. 2011	Annual growth
Smartphones	**487.7**	**62.7%**
Total PCs	**414.6**	**14.8%**
- Notebooks	209.6	7.5%
- Desktops	112.4	2.3%
- Tablets	63.2	274.2%
- Netbooks	29.4	-25.3%

The Smartphone Era

April 2013: Beginning of Smartphone Era!

Feature vs. Smartphone Phones

- The bulk of mobile phones are acquired in the **developing world** (e.g., China, India, Africa etc.)

- **In April, 2013**, for the **first time** in history the number of **Worldwide Smartphone** sales exceeded that of feature phones (according to IDC)
    - 51.6% were Smartphones (216M units)
    - 48.4% were Feature Phones (186M units)

- IDC attributes that to the fact that Chinese manufactures (ZTE, Huawei) started building smartphones for the wide markets.
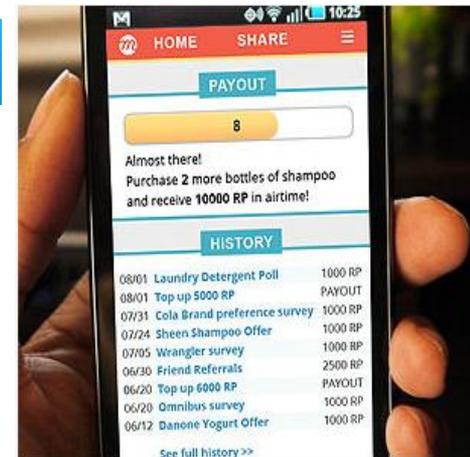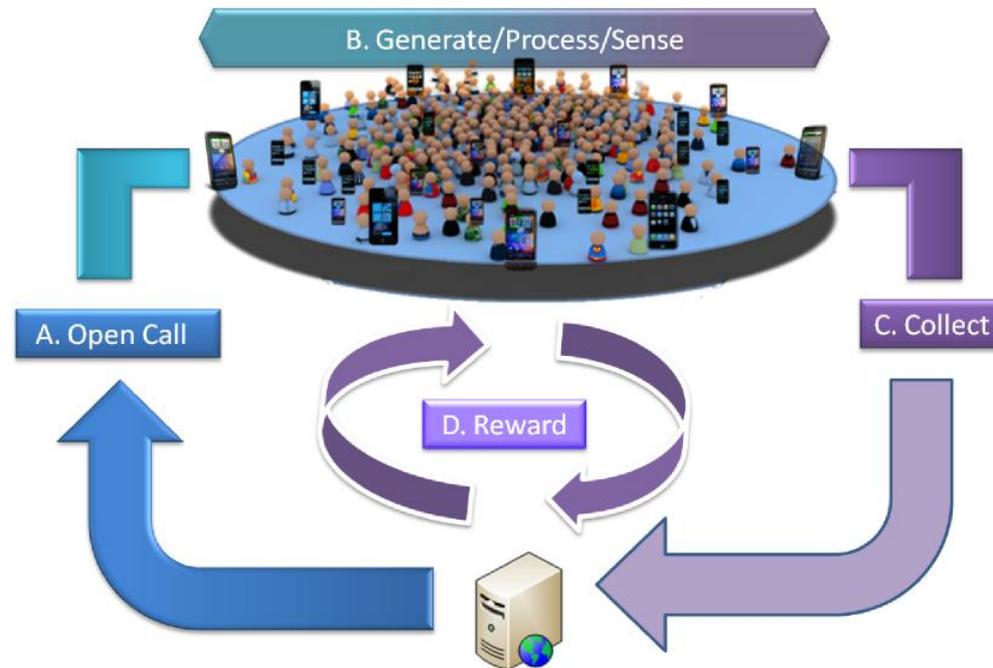
More Smartphones Were Shipped in Q1 2013 Than Feature Phones, An Industry First According to IDC, 25 Apr 2013, http://www.idc.com/getdoc.jsp?containerId=prUS24085413

Crowdsourcing in the Smartphone Era

- A smartphone crowd is constantly moving and sensing providing large amounts of opportunistic data that enables new services and applications

"Crowdsourcing with Smartphones", Georgios Chatzimiloudis, Andreas Konstantinidis, Christos Laoudias, Demetrios Zeinalipour-Yazti, IEEE Internet Computing (IC '12), Special Issue: Sep/Oct 2012 - Crowdsourcing, May 2012. IEEE Press, Volume 16, Pages: 36-44, 2012.

Mobile Crowdsourcing (MCS) Challenges

Inherited from Web Crowdsourcing

- Incentives

- Recruit

- Combine (data interpolation)

- Manage Abuse

- Scale/Manage

- Data Quality/Trustworthiness

- Disclosure Issues

Mobile Crowdsourcing (MCS) Challenges

Additional Challenges

1. Big Data
   - Volume and Velocity by sensor data

2. Typing and User Interfaces
   - Typing is cumbersome due to small form factor / display keyboard.
   - Scrolling & Crowded GUIs. Attention issues due to possible mobility.

3. (Location) Privacy
   - Coarse-grain (cell, wifi) vs. fine (gps)

4. Energy Consumption
   - Power Hungry (GPS, Brightness, etc.)

Mobile Crowdsourcing (MCS) Challenges

Additional Challenges

5. Calibration and Multi-device Issues

   – Different readings by different sensors (e.g., Wifi RSS, magnetic field, etc.)

   – Incomplete Data (Compressive Sensing reconstructing signal with missing values).

6. Connectivity Issues

   – Workforce might have intermittent connectivity.

7. Heterogeneous Clients hinders deployment

   – Different OSes, sensor, features, APIs, etc.

   – One supports active background tasks another OS doesn't, etc.

MCS in Developing Regions
Number of Mobiles in Use

Rank	Country or region	Number of mobile phones	Population	# of phones as % of population	Last updated date
-	World	Over 6 billion	7,012,000,000[1]	87	2011[2][3]
01	China	1,150,000,000[4]	1,349,585,838[5]	85.21	April 2013[6]
02	India	861,660,000	1,220,200,000[7]	70.42	28 Feb 2013[8]
03	United States	327,577,529	310,866,000[9]	103.9	June 2012[10]
04	Brazil	263,040,000	192,379,287[11]	136.7	February 2013[12]
05	Russia	256,117,000	142,905,200[11]	155.5	July 2011[13]
06	Indonesia	236,800,000	237,556,363	109.28	September 2012[11]
07	Pakistan	122,060,799	178,854,781[14]	68.83	Dec 2012[15][16]
08	Japan	121,246,700	127,628,095	95.1	June 2011[17]
09	Germany	107,000,000	81,882,342	130.1	2009[18]
10	Philippines	106,987,098	94,013,200	113.8	October 2011[19]
11	Nigeria	101,271,578	140,000,000	72.3	May 2012[20]
12	Bangladesh	98,470,000	148,090,000	68	July 2012[21]
13	Iran	96,165,000	73,973,000	130	February 2012[22]
14	Mexico	92,900,000	112,322,757	82.7	Dec. 2011[23]
15	Italy	88,580,000	60,090,400	147.4	Dec. 2008[24]
16	United Kingdom	75,750,000	61,612,300	122.9	Dec. 2008[25]
17	Vietnam	72,300,000	90,549,390	79	October 2010[26]
18	France	72,180,000	63,573,842	114.2	Dec. 2008[27]
19	Egypt	92,640,000	82,120,000	112.81	Egypt Ministry of Communications & IT, August 2012[28]
20	Thailand	69,000,000	65,001,021	105	2010[29]

36% World Population 2.5B

Mturk workers in India: 30% (educated with salary double the avg)

Wikipedia: http://en.wikipedia.org/wiki/List_of_countries_by_number_of_mobile_phones_in_use

MCS in Developing Regions

- **txtEagle (now JANA)** founded by Nathan Eagle (PhD, MIT, 2005) a first-of-a-kind mobile CS system:

  - **Requesters:** can assign small tasks (translation, transcription and surveys) on their mobile phones.

  - **Workers (**today 3.48 Billion Workers in 102 countries!**): :** rewarded with airtime on their mobile subscriber accounts or MPESA (mobile money described next).

txteagle: Mobile Crowdsourcing, Internationalization, Design and Global Development, LNCS Volume 5623, pp 447-456, 2009.

MCS in Developing Regions

- **Another app txtEagle SMS Bloodbank :**

  - **Idea:** to report blood levels of local hospitals centrally by nurses.

  - Initially, in the **absence** of an **incentive**, the system was a **complete failure.**

  - In **summer 2007**, **automatic airtime credit** was incorporated to award nurses for their contribution => then a **huge success!**

- **Other txtEagle SMS applications:**

  - Transcription mentioned previously (global market $18B in 2010)

  - Software Localization (60 local languages in Kenya, txtEagle generated a cookbook

  - Citizen Journalism, Sentiment Analysis, Surveys

MCS in Developing Regions

- **MClerk:** MCS (linguistic tasks) amongst low-income workers in developing countries.

  - 239 user study in India digitizing 25K words.

- **Challenges:**

  - Lack of skills (e.g., compose & write in English)

    - Authors focus on Digitization of local-language text (appropriate to local workers)

  - Assign tasks with images

    - Use protocol using sms.

- **Other projects:** http://samasource.org/

  Gupta et. al. mClerk: enabling mobile crowdsourcing in developing regions. In ACM CHI '12,

Urban Sensing

- Use sensors in urban environments in support of more classic **environmental sensing applications**.

- *"People sense and contribute data about their surroundings using mobile devices"* (Kanhere)

- Example Projects:
  - **Dartmouth | Metrosense:** SoundSense, CenceMe, Sensor Sharing, **BikeNet**, AnonySense, and Second Life Sensor.
  - *MIT | Cartel:* VTrack/CTrack, PotHole
  - **Harvard :** Citysense (grew out of MoteLab)
  - *UNSW: Noise (Earphone) & Air pollution (HazeWatch, CommonSense),*

Urban Sensing

- **Monitoring Urban Spaces**
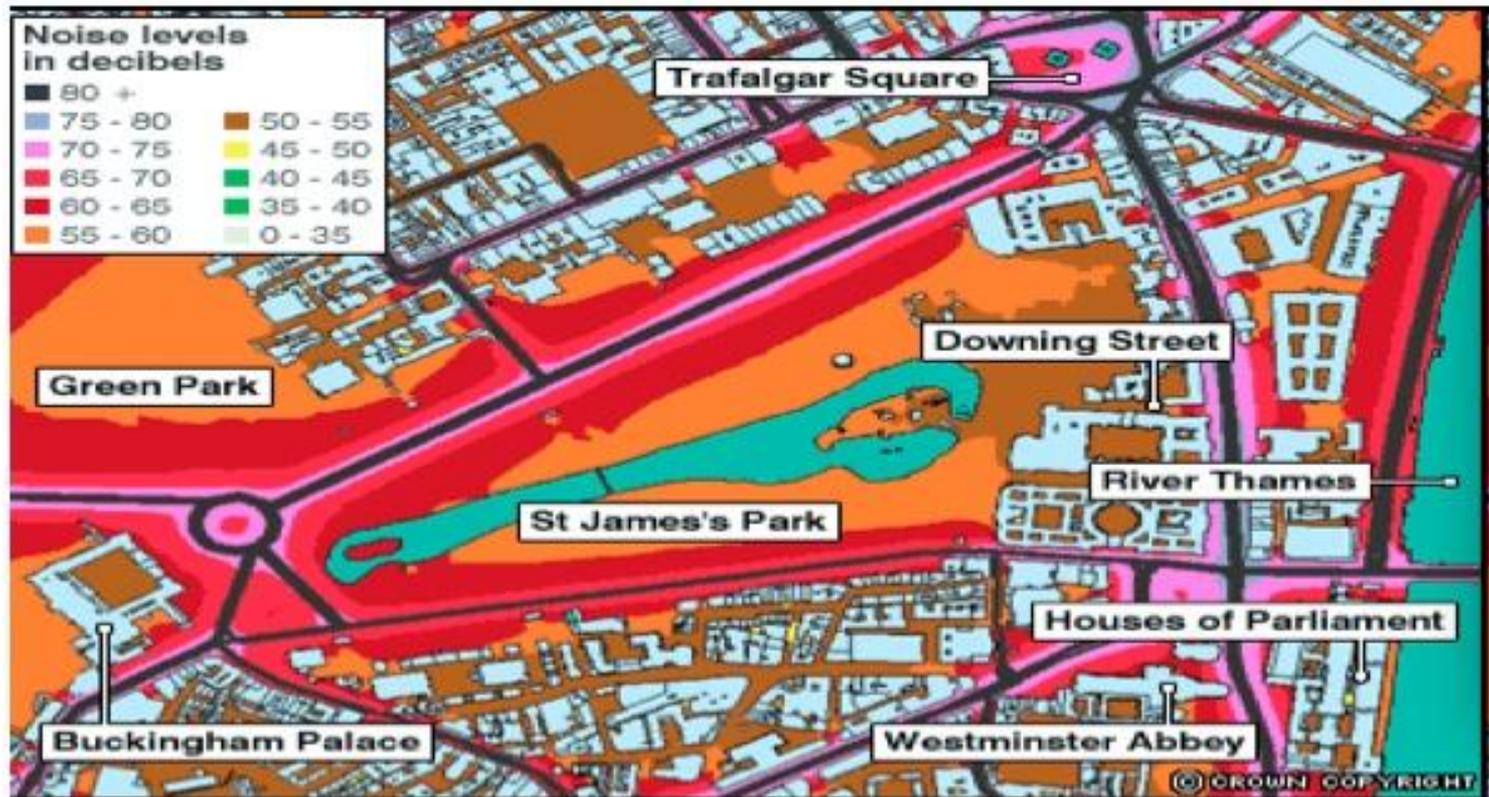
"Ear-Phone: An End-to-End Participatory Urban Noise Mapping System " Rajib Rana, Chun Tung Chou, Salil Kanhere, Nirupama Bulusu, and Wen Hu. In ACM/IEEE IPSN 10, SPOTS Track, Stockholm, Sweden, April 2010.

Urban Sensing

- **Useful Material on the Topic**
  - MDM'11 Tutorial by Salil Kanhere (Lulea, Sweden)*:*
    - *"Participatory Sensing: Crowdsourcing Data from Mobile Smartphones in Urban Spaces. Salil Kanhere*
  - MDM'13 Panel Discussion:
    - *Large-Scale Participatory Urban Sensing: A Fad or Reality?*

- ***Opportunistic (passive) or Participatory (active)***
  - *Monitoring Urban Spaces*
    - *Traffic, Road Quality, Air Quality, Noise Poll.*
  - *Commerce*
    - *Sharing Gas Pricing Information*
  - *Citizen Sensing*
    - *Reporting Suspicious Activity*
  - *Personal Sensing*
    - *Monitoring health, diet, carbon footprint*
  - *Social Sensing*
    - *Sensing meets mobile social networks*

Traffic Monitoring

Mapping Road Traffic with fixed cameras & sensors mounted on roadsides?

Traffic Monitoring

Mapping the Road traffic by collecting WiFi signals.

vTrack server

GPS positions/
AP observations

Users driving
with smartphones

Real time
routing updates

Received Signal Strength (RSS):
power present in WiFi radio signal
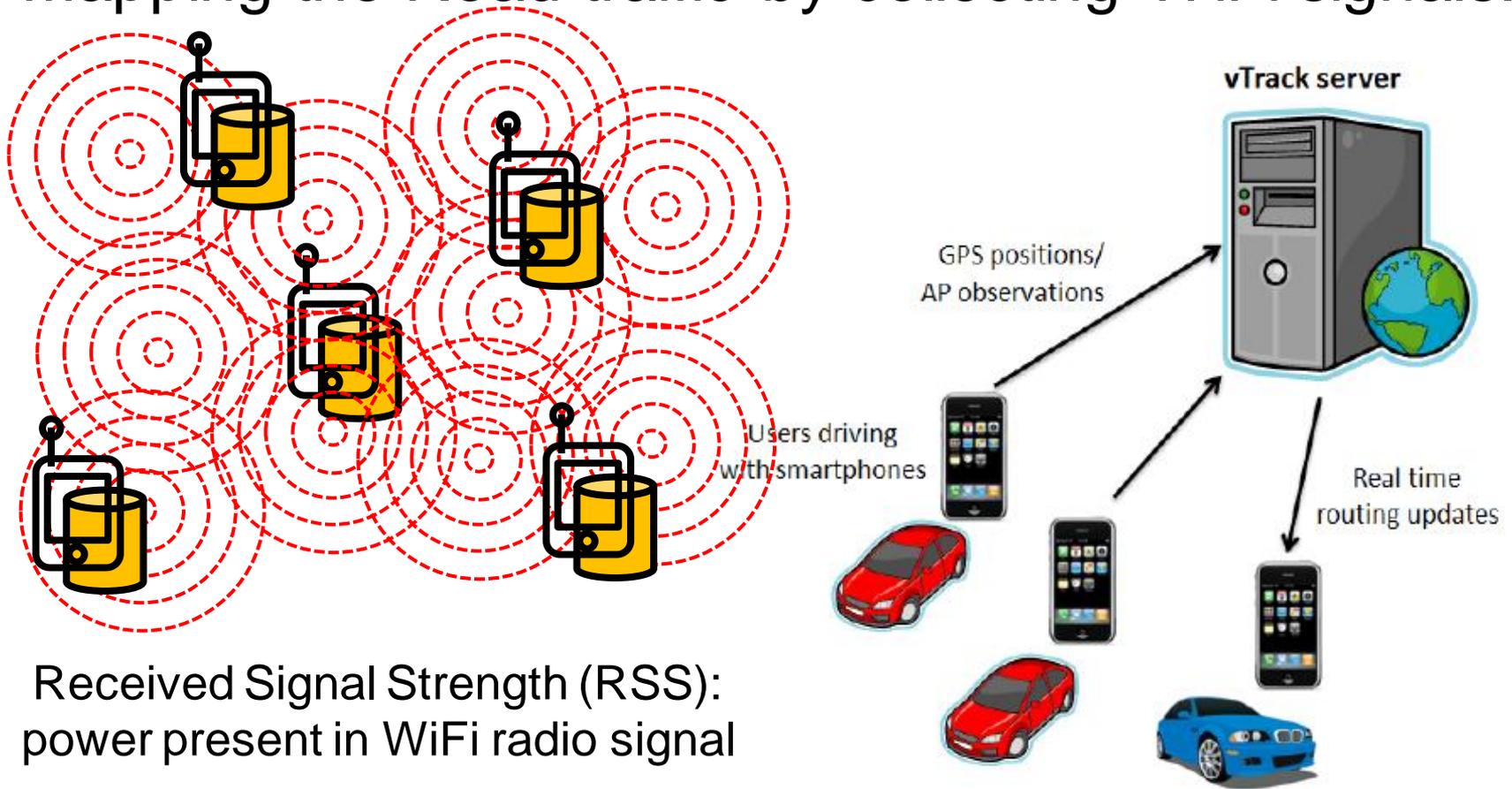
Graphics courtesy of: A .Thiagarajan et. al. "Vtrack: Accurate, Energy-Aware Road Traffic Delay
Estimation using Mobile Phones, In Sensys'09, pages 85-98. ACM, (Best Paper) MIT's CarTel Group

Traffic Monitoring

- Waze: Free GPS Navigation with Turn by Turn
  - Workers report their GPS location and events (gas prices, traffic jams, etc.)
  - Quite power hungry app…
  - Real-time updates to users
- Company Overview:
  - Israeli Mobile Crowdsourcing company founded in 2008, now 47M Users
  - **Facebook (wanted) & Google want to buy it for 1 Billion dollars!**
 - Waze Wins Best Overall App at Global Mobile Awards 2013!

Traffic Monitoring

Traffic Map for Riverside CA
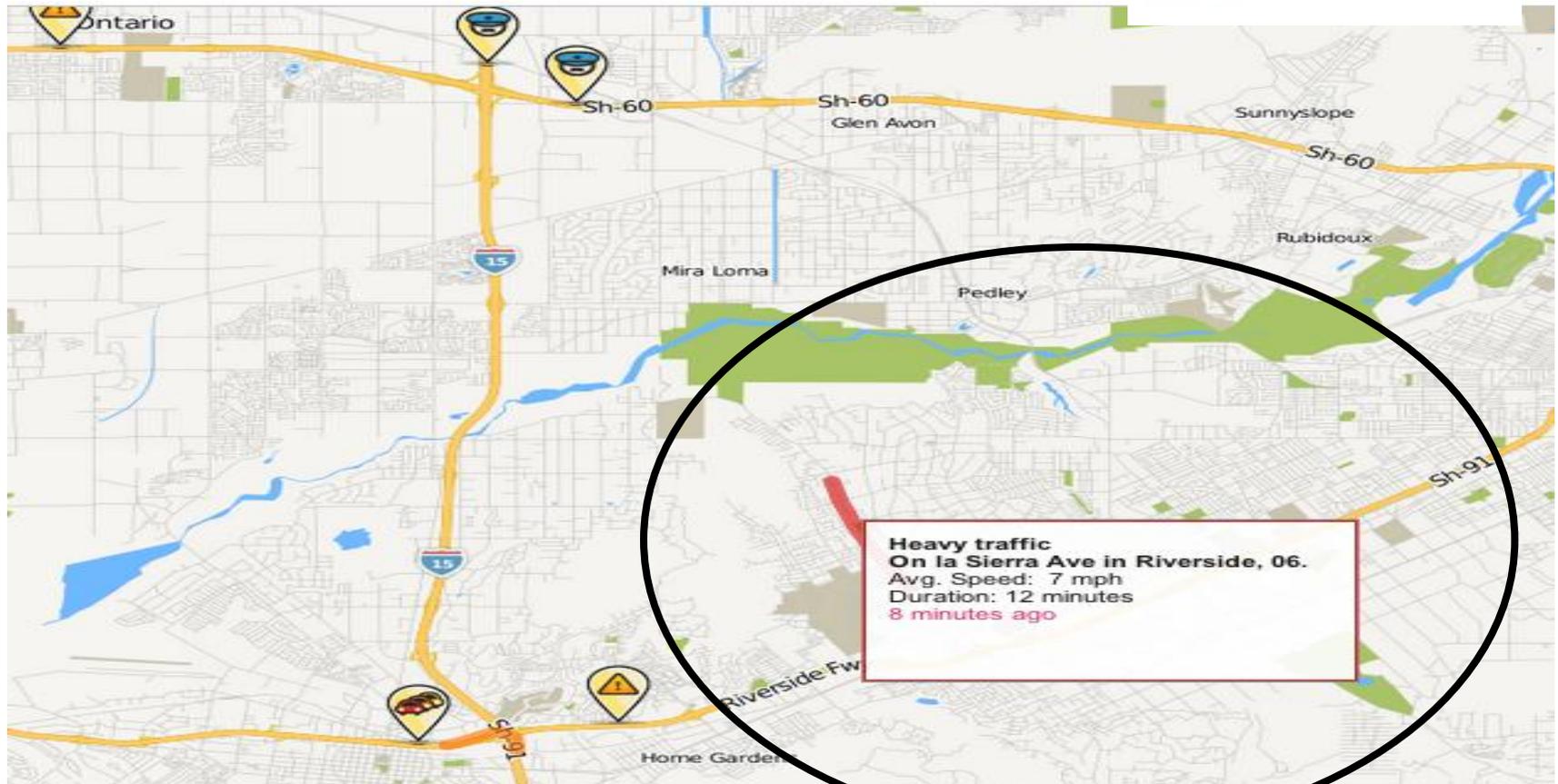
Content-based Multimedia Search

- CrowdSearch architecture similar to CrowdDB we saw earlier (engine filters, crowd fills the rest)

- Combines automated image search **filter (by engine)** with real-time human **validation** of search (by Mturk workers).

  - provides 95% precision, fast answers, with low cost (due to optimizations)
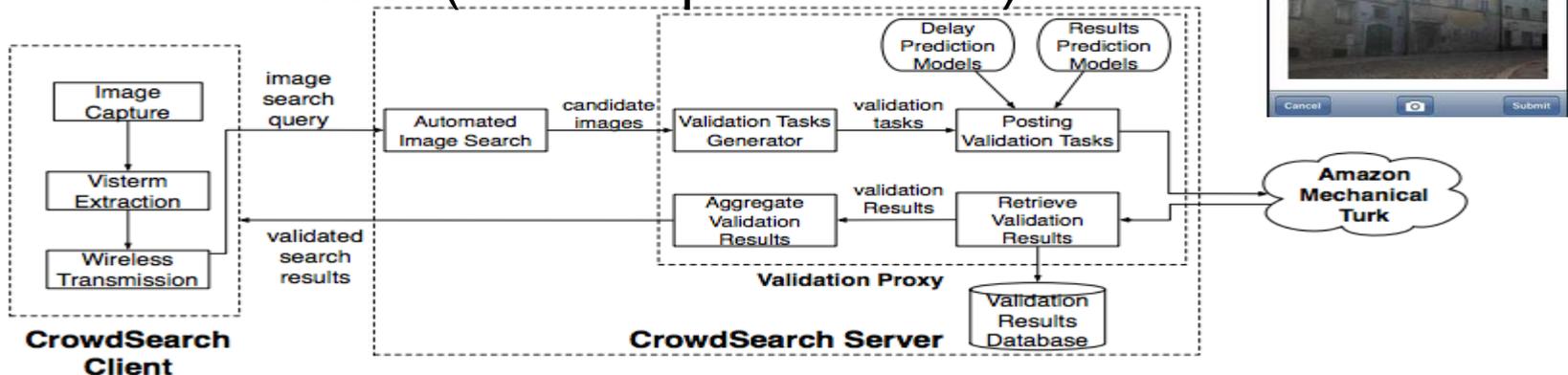
Yan, Kumar, Ganesan, CrowdSearch: Exploiting Crowds for Accurate Real-time Image Search on Mobile Phones, In MobiSys 2010.

Spatial / Outdoor

- Spatial MCS: Crowdsourcing a set of **spatial tasks** (i.e., related to a location) to be performed by **mobile workers***

- **GigWalk:** Perform "Gigs" (e.g., collect prices, check, inspect, collect, interview, touch, etc.) and earn money
  - *Price: Monetary (Paypal),*
  - **Workers:** 250K across US and Canada only, 90% Educated
 - Other: Scoopshot (photos & videos)

*adjusted from: Kazemi and Shahabi. GeoCrowd: enabling query answering with spatial crowdsourcing. In SIGSPATIAL '12

Spatial / Outdoor

- **Crowdfynder:** Still a Spatial MCS app but particularly targeted on Lost-N-Found "Fynding" tasks:
  - Lost, Found and Crime Alert (but without police notification)
  - Price associated with task.
- Other Lost-N-Found services:
  - Bikn (bee'-kin).
  - BeatTheBushes
 - NextDoor (Neigborhood Social Network)
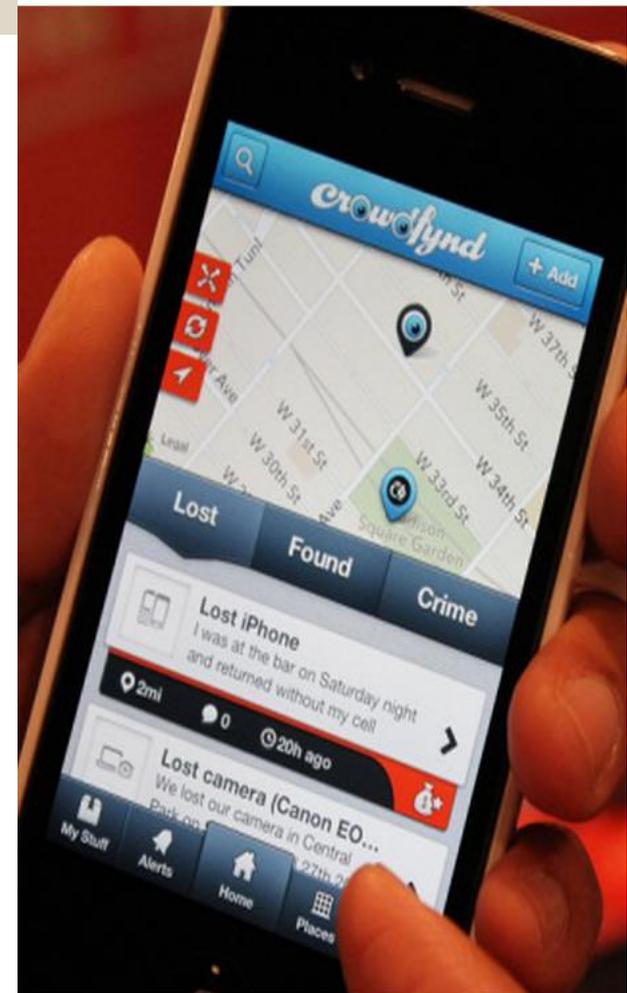
Spatial / Indoor

- Numerous ways to **localize** without power-hungry **GPS** but most of the solutions rely on additional hardware (RFID, sensor networks, etc.)

- Smartphones can nowadays localize **off-the-shelf** with onboard sensors and **WiFi signal fingerprints** (coined **Hybrid Localization**)

- **These solutions require that services acquire local data through Crowdsourcing (e.g., Google Indoor)**

  – Building RadioMaps, MagnetometerMap, etc.

  – **[Airplace]** "The Airplace Indoor Positioning Platform for Android Smartphones", C. Laoudias et. al., **Best Demo Award at IEEE MDM'12.**

  – **[HybridCywee]** "Demo: the airplace indoor positioning platform", C.-L. Li, C. Laoudias, G. Larkou, Y.-K. Tsai, D. Zeinalipour-Yazti and C. G. Panayiotou, in ACM Mobisys'13. Video at: http://youtu.be/DyvQLSuI00I

  – Wifislam.com (bought recently by Apple for 20M)

  – Google Indoor doing well with mapping but not

Spatial / Indoor

Cywee / Airplace
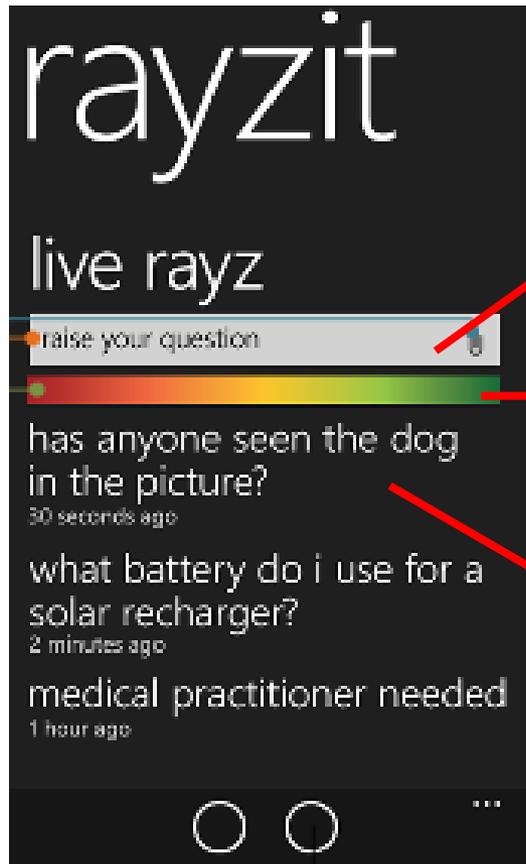http://youtu.be/DyvQLSuI00I
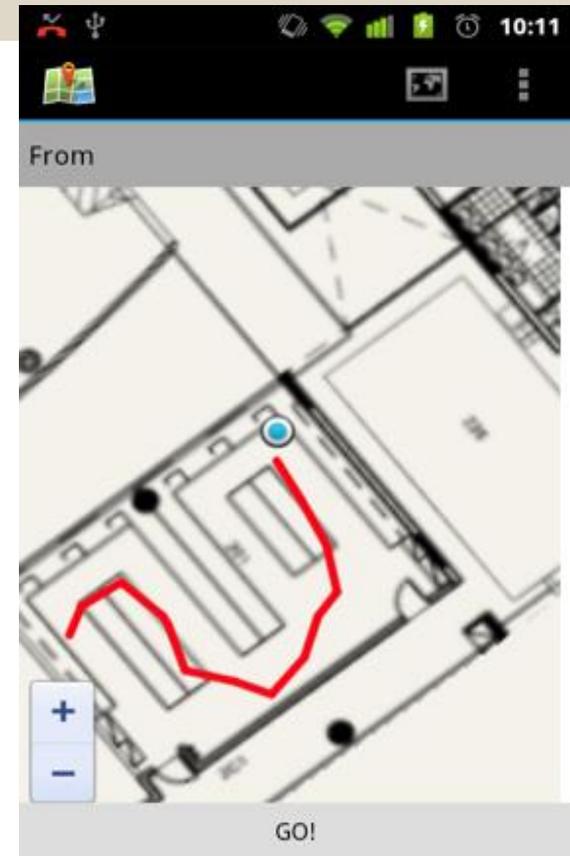
Anyplace
(Navigate) seamlessly indoor or outdoor
http://anyplace.cs.ucy.ac.cy/
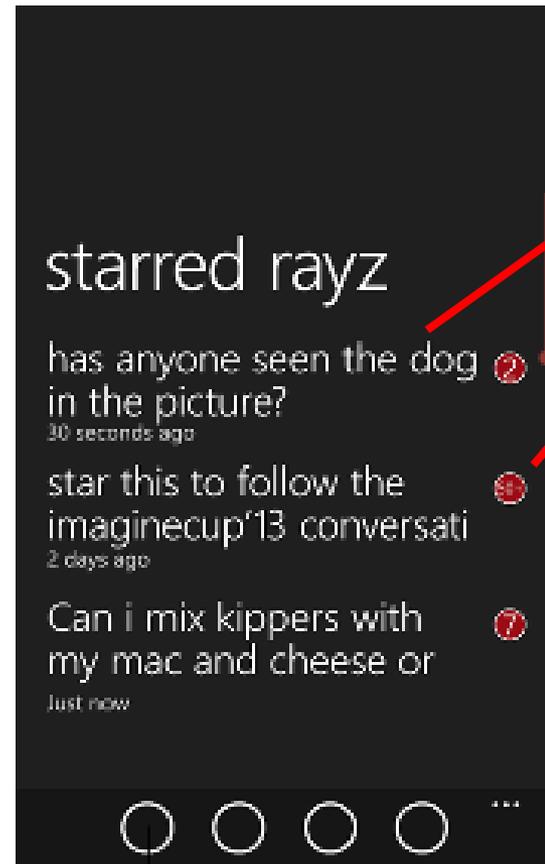
Proximity-based Q & A Systems

Rayzit: Mobile Crowdsourcing Q/A system (funded by alliance backed by Microsoft/Nokia)

Task propagated to KNN
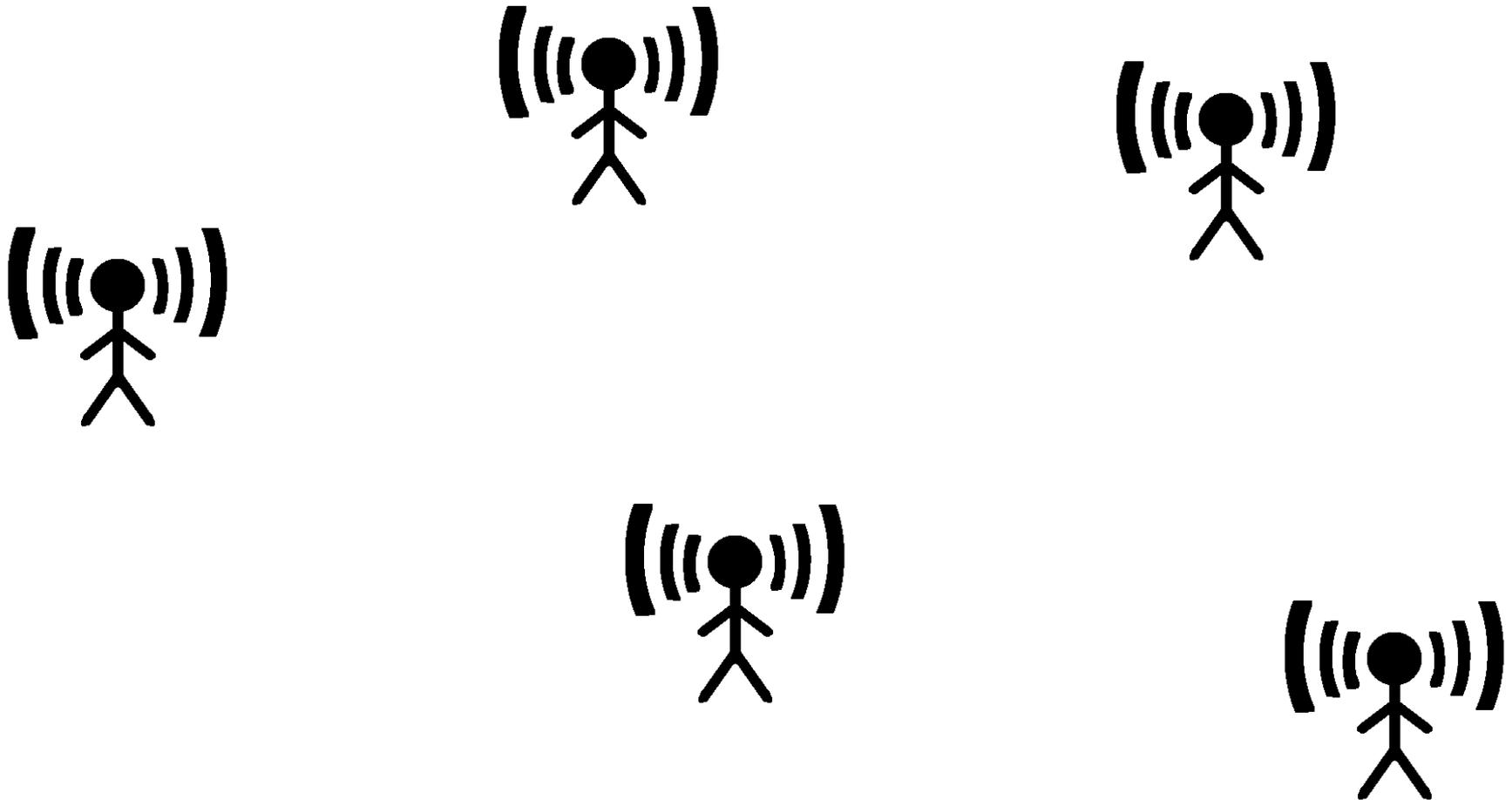
Abuse / Incentive

Engagement

Rerayz

Follow Conversations

Find 2 Closest Neighbors for <u>ALL</u> User
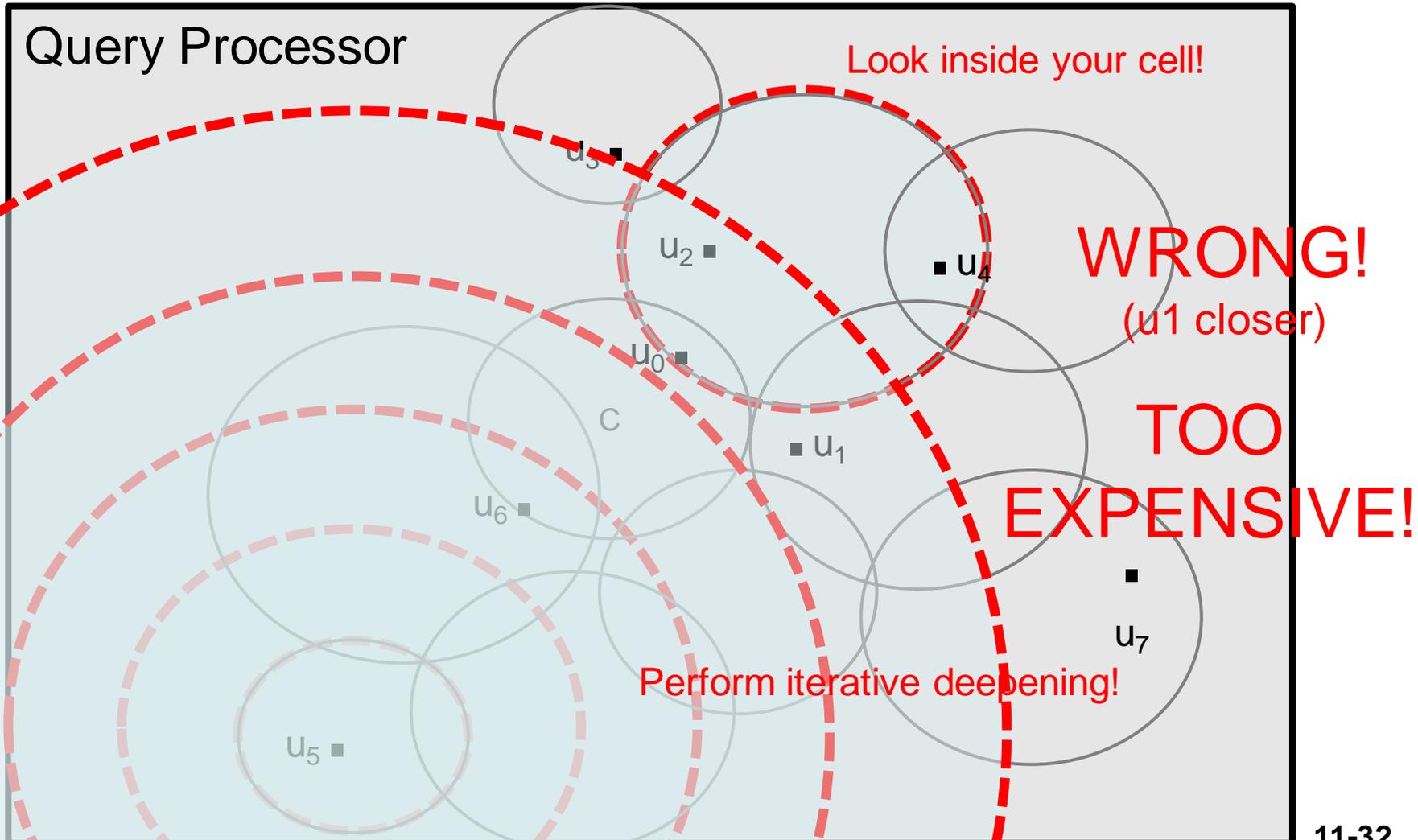
Proximity-based Q & A Systems

Find 2-NN for u_0 at timestep t.   For u_5?

Query Processor

Look inside your cell!

u_3

u_2

u_4

WRONG!
(u1 closer)

u_0

C

u_1

TOO
EXPENSIVE!

u_6

u_7

Perform iterative deepening!

u_5

Proximity-based Q & A Systems

- For every timestep:

1. Initialize a **k^+-heap** for every cell

2. Insert **every** user's location report to every **k^+-heap**
   - Notice that k^+-heap is a heap-based structure and most location reports will be dropped as a result of an insert operation

3. For every user scan the k^+-heap of its cell to find his k-NN

"Continuous all k-nearest neighbor querying in smartphone networks", Georgios Chatzimilioudis, Demetrios Zeinalipour-Yazti, Wang-Chien Lee, Marios D. Dikaiakos, In MDM'12.

Query Processor

u_3
u_2
u_4
u_0
u_1
u_6
u_5
u_7
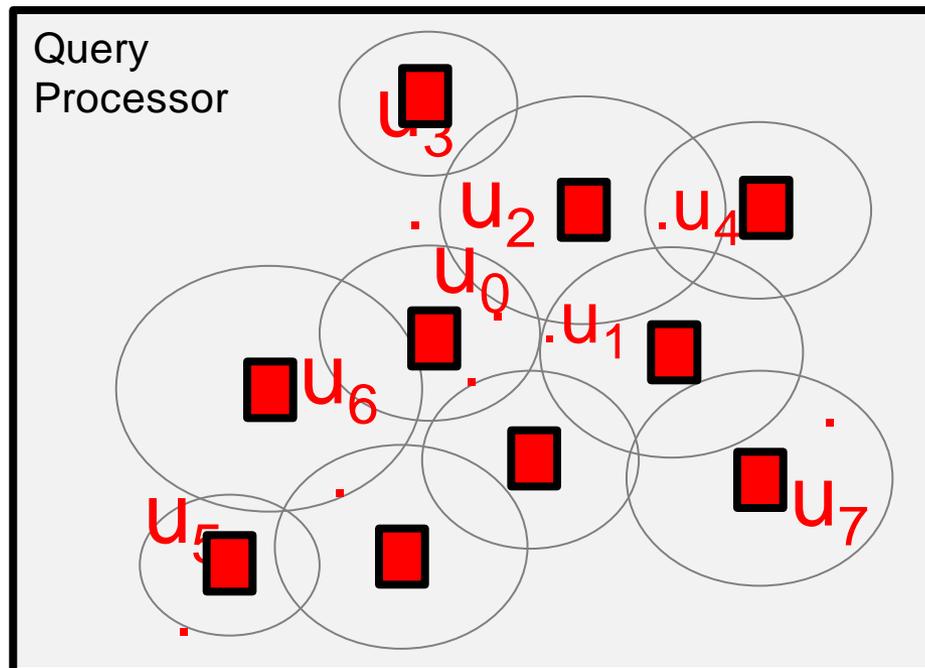
Testbeds

- **Smartphone Testbeds:** Allow the requestor to deploy the task (app, data collection, remote terminal etc.) directly on the end smartphone devices.

  - **[PhoneLab]** G. Challen et. al. "PhoneLab: A Large-Scale Participatory Smartphone Testbed", In USENIX NSDI'12 (poster).

  - **[SmartLab]** "Demo: a programming cloud of smartphones", A. Konstantinidis, C. Costa, G. Larkou, D. Zeinalipour-Yazti, In ACM Mobisys '12. **[By our Group]**

  - **[PRISM]** T. Das, P. Mohan, V.N. Padmanabhan, R. Ramjee, and A. Sharma, "PRISM: Platform for Remote Sensing using Smartphones", In ACM MobiSys'10.

  - **[CrowdLab]** E. Cuervo, P. Gilbert, B. Wu, and L.P. Cox, "CrowdLab: An Architecture for Volunteer Mobile Testbeds", In COMSNETS'11.

Testbeds / SmartLab

- Currently, there are no **testbeds** (like motelab, planetlab) for realistically prototyping Smartphone Network applications and protocols at a large scale.

- Currently applications are tested in emulators.
  - Sensors are **not emulated**. ☹
  - Reprogramming is difficult. ☹

- *SmartLab (*http://smartlab.cs.ucy.ac.cy/*) is* a first-of-a-kind programmable cloud of 40+ smartphones deployed at our department enabling a new line of systems-oriented research on smartphones.

"Crowdsourcing with Smartphones", Georgios Chatzimiloudis, Andreas Konstantinides, Christos Laoudias, Demetrios Zeinalipour-Yazti *IEEE Internet Computing* (IC '12), Special Issue: Sep/Oct 2012 - Crowdsourcing, May 2012. IEEE Press, 2012

"Demo: A Programming Cloud of Smartphones", A. Konstantinidis, C. Costa, G. Larkou and D. Zeinalipour-Yazti, "Demo at the 10th International Conference on Mobile Systems, Applications and Services" (Mobisys '12), Low Wood Bay, Lake District, UK, 2012.

Testbeds / SmartLab

SmartLab: Massive smartphone simulations with our first global open smartphone IaaS cloud – http://smartlab.cs.ucy.ac.cy/

Static Androids

Mobile Androids

TestBeds / SmartLab

http://smartlab.cs.ucy.ac.cy/

TestBeds / PhoneLab

- **PhoneLab**: a Participatory SmartPhone Sensing Testbed

- **200 Nexus S 4G phones** used by Students and Faculty Members at the Univ. of Buffalo

- Targeted for **Data Collection Scenarios** (not fine-grain access like SmartLab)

  – Each Data **Collecti**on task need to undergo an **Institutional Review Board process** (similar to other projects touching ethical issues)

  – **Data Collection:** Workers (Students) have to bring in their smartphones to have the app installed + data collected.

  – **Incentive:** Free Sprint Phone for 1st year. After that, only $44.23/month for an unlimited plan (claimed to be better than competition)

[PhoneLab] G. Challen et. al. "PhoneLab: A Large-Scale Participatory Smartphone Testbed", In USENIX NSDI'12 (poster).

Spatio-Temporal Systems

Incentive:

Contribute to the resoution of queries for Social Benefit (without revealing traces)

Applications

- **Intelligent Transportation Systems:** "F*ind whether a new bus route is similar to the trajectories of K other users.*"

- **Social Networks:** "*Find if there is an evening cycling route from MOMA to the Julliard*"

- GeoLife, GPS-Waypoints, Sharemyroutes, etc. offer centralized counterparts.

Spatio-Temporal Systems

- *Problem: Compare a query with all distributed trajectories and return the k most similar trajectories to the query.*

  - *Similarity between two objects A, B is associated with a distance function.*

Query

?

Distance

D = 7.3

D = 10.2

K

D = 11.8

D = 17

D = 22
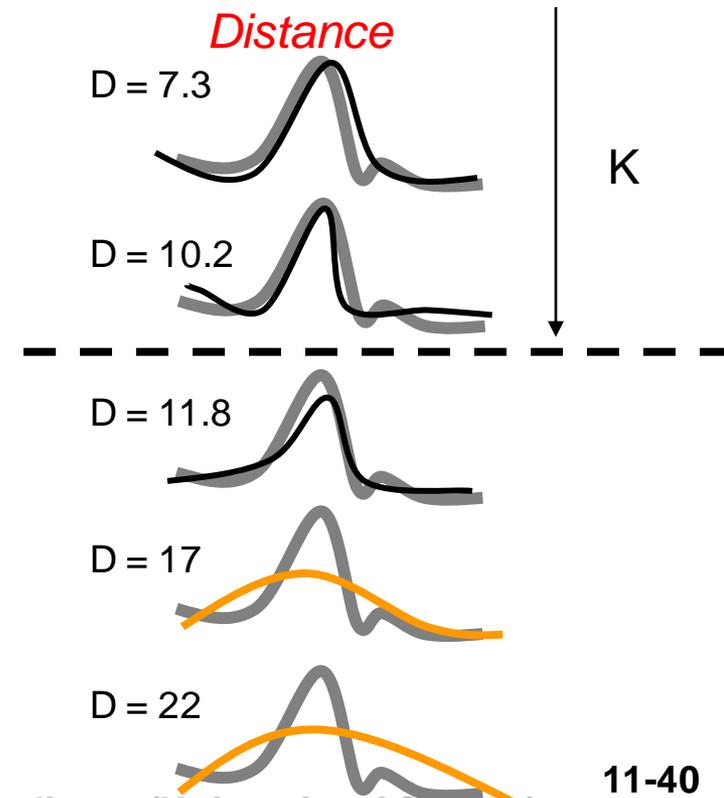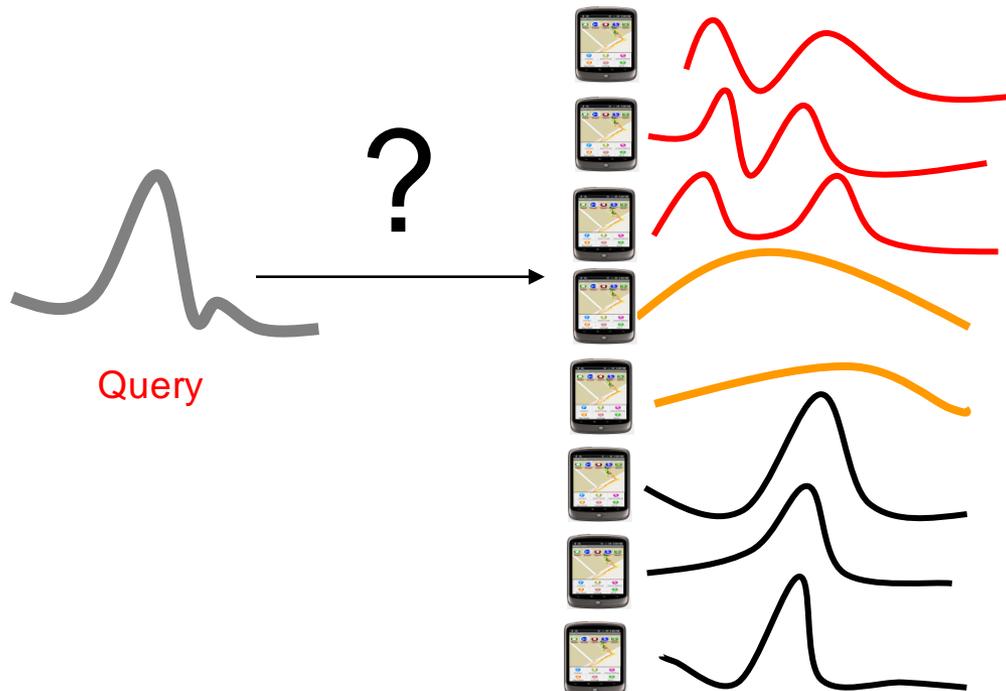
Spatio-Temporal Systems

- An intelligent top-K processing algorithm for identifying the K most similar trajectories to Q in a distributed environment.

  - Step A: Conduct an inexpensive linear-time $LCSS(MBE_Q, A_i)$ computation on the smartphones to approximate the answer.

  - *Step B: Exploit the approximation to identify the correct answer by iteratively asking specific nodes to conduct $LCSS(Q, A_i)$.*

• "Crowdsourced Trace Similarity with Smartphones", Demetrios Zeinalipour-Yazti et. al, In IEEE TKDE, Volume 25, Pages: 1240-1253, Los Alamitos, CA, USA, 2013.

• "SmartTrace: Finding similar trajectories in smartphone networks without disclosing the traces", Costa et al., IEEE ICDE'11.
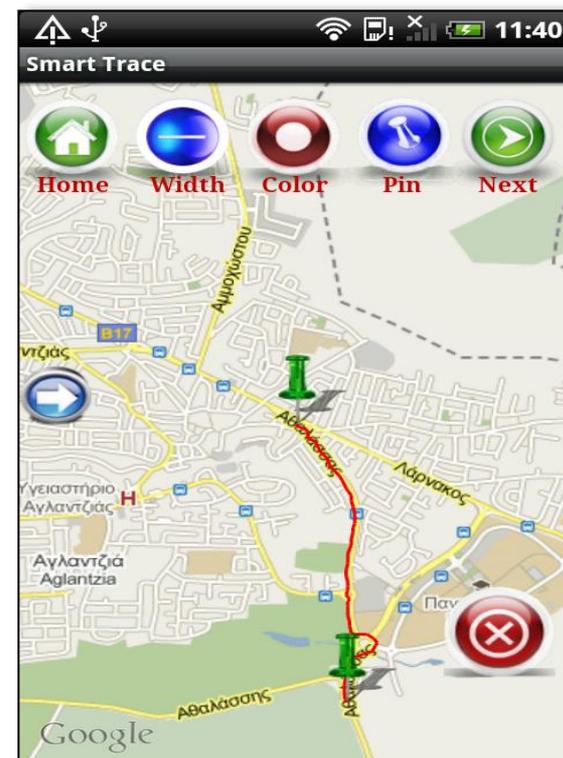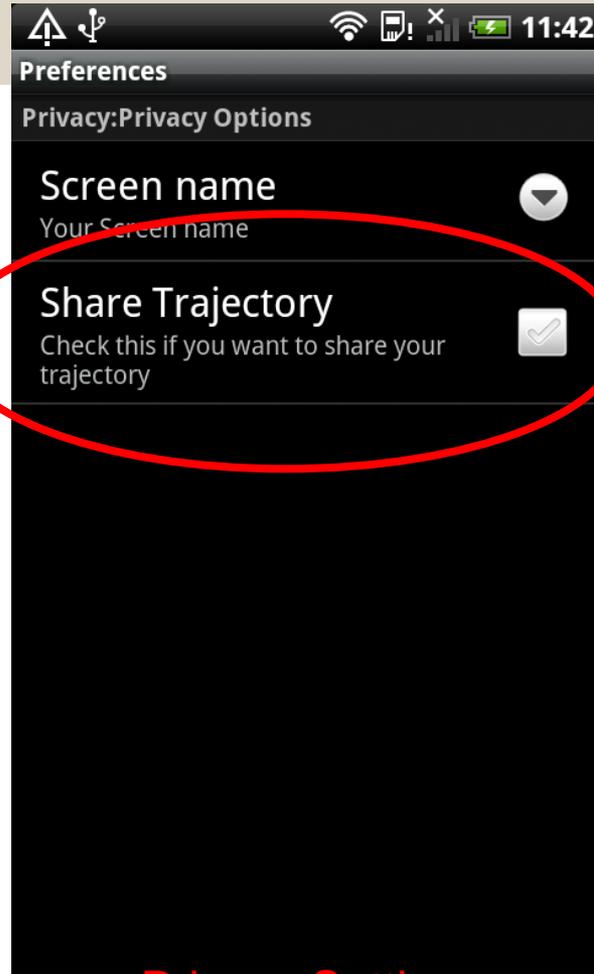
SmartTrace for Android!
http://smarttrace.cs.ucy.ac.cy/
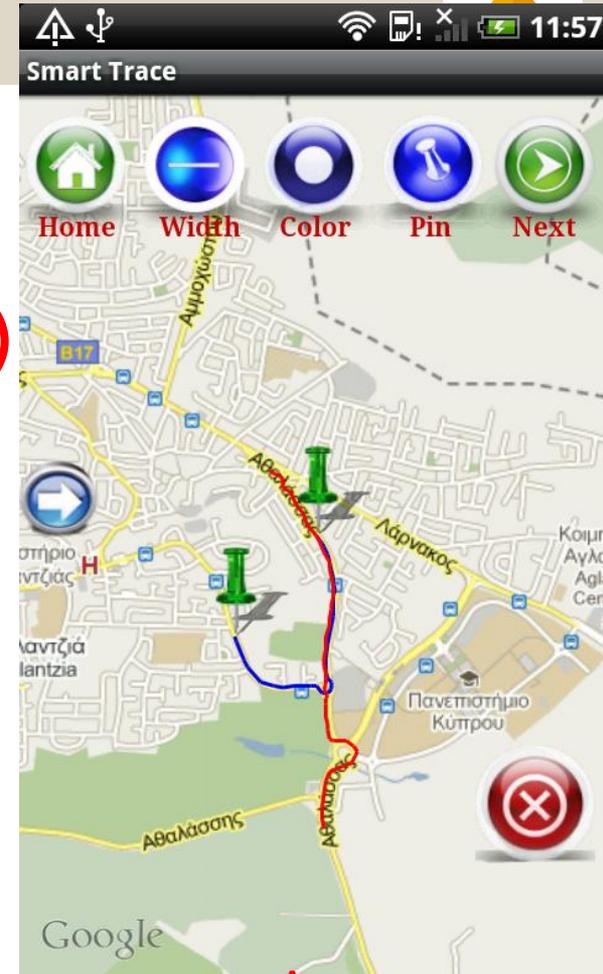

Query Q

Device B

Device C

EPL646: Advanced Topics in Databases - Demetris Zeinalipour (University of Cyprus)

Answer

Privacy Setting

Answer
With Trace