



Privacy-preserving Data Sharing on Sensor Map

Presented by:

Tarek Abdelzaher

Department of Computer Science

University of Illinois at Urbana-Champaign

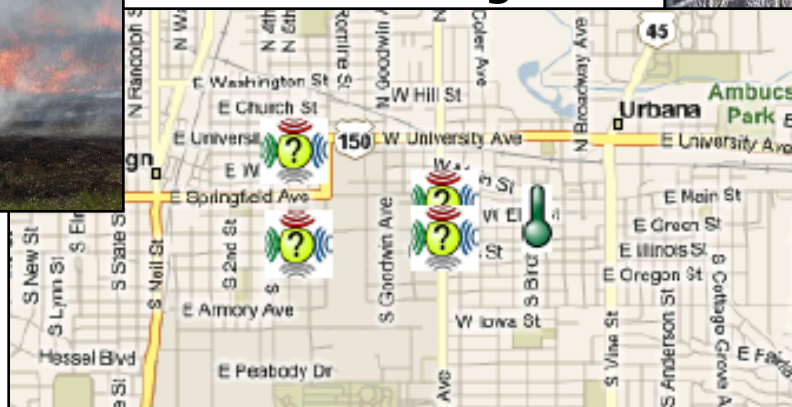
In Collaboration with Microsoft Research

Sensor Map

A Microsoft Research Project

- A browser for the physical world
 - Anyone can register “sensors”
 - Anyone can browse sensor data

Wild fire detection
and monitoring



Acoustic
sensing

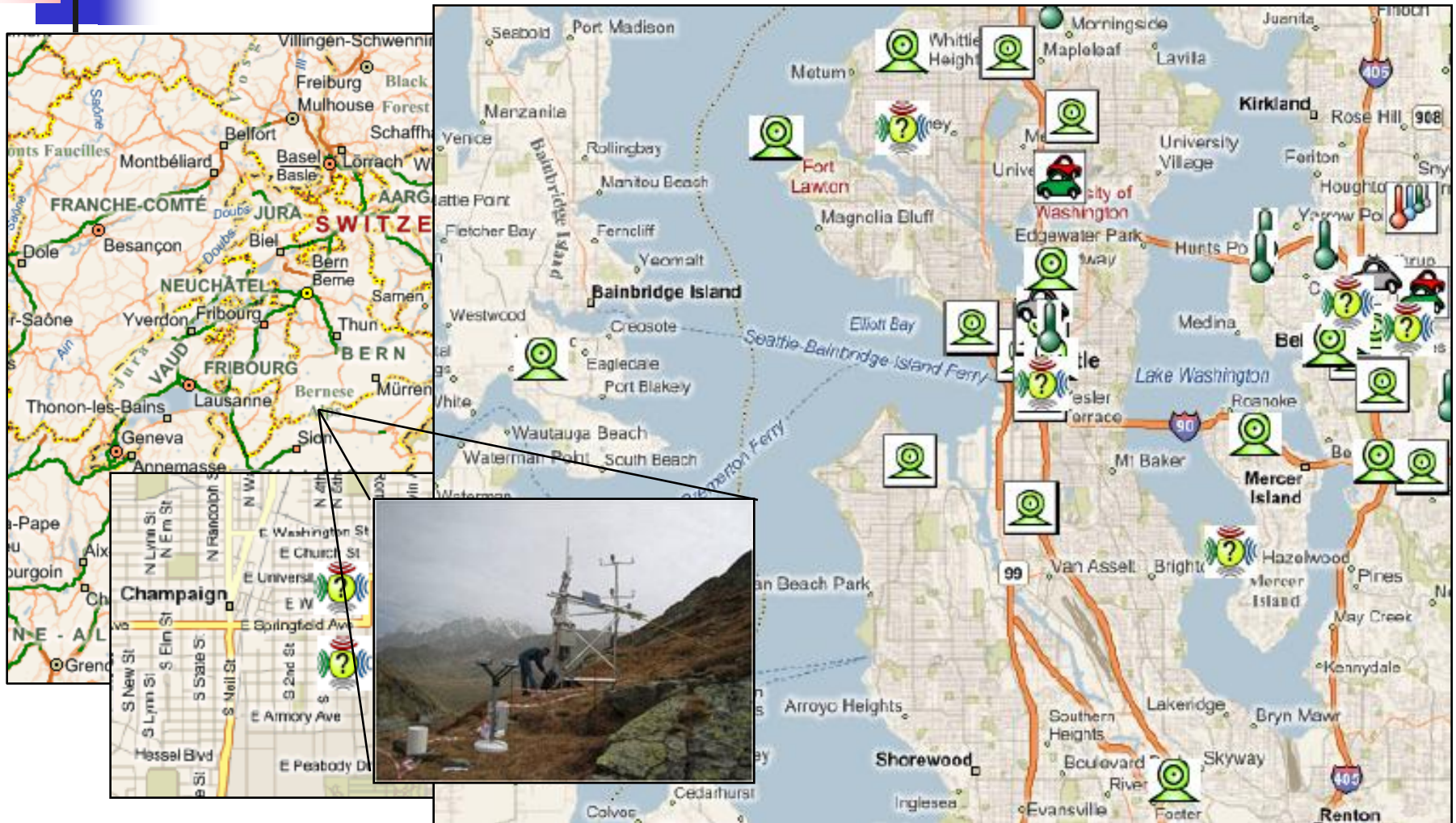


World-wide Sensor Web

Body sensor networks

Browsing the Physical World in Real Time - Feng Zhao

<http://atom.research.microsoft.com/sensormap/>





UIUC Sensor Map Focus: Participatory Sensing

- WWW → a gathering place for special interest social groups
- SensorMap → a gathering place for special interest data pools
 - Joggers sharing route and fitness data
 - Drivers sharing GPS/speed data
 - Environmentalists sharing CO2 measurements
 - Biologists sharing instrumented animal biometrics
 - Chronic sinus patients sharing allergen counts

A Sensor Web is a Social Network

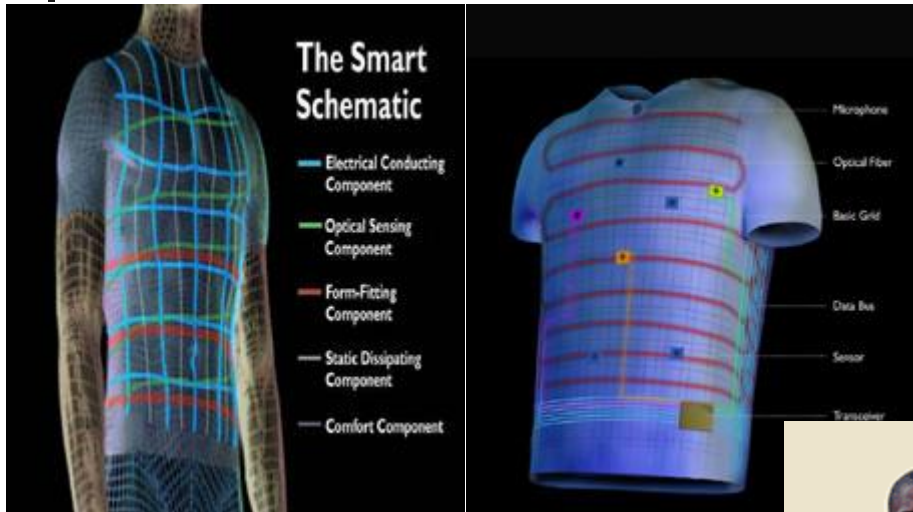
Sponsored by
Microsoft Research
and NSF

People with Sensors

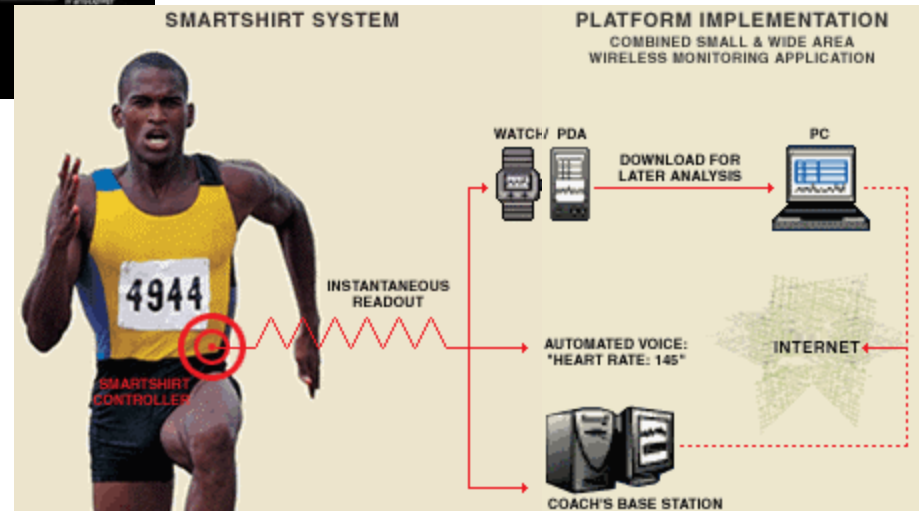
The master controller



Nike -iPod



<http://www.sensatex.com>



Spot

Wii



Technology introduced by 2007

Health and Wellness

Health Vault

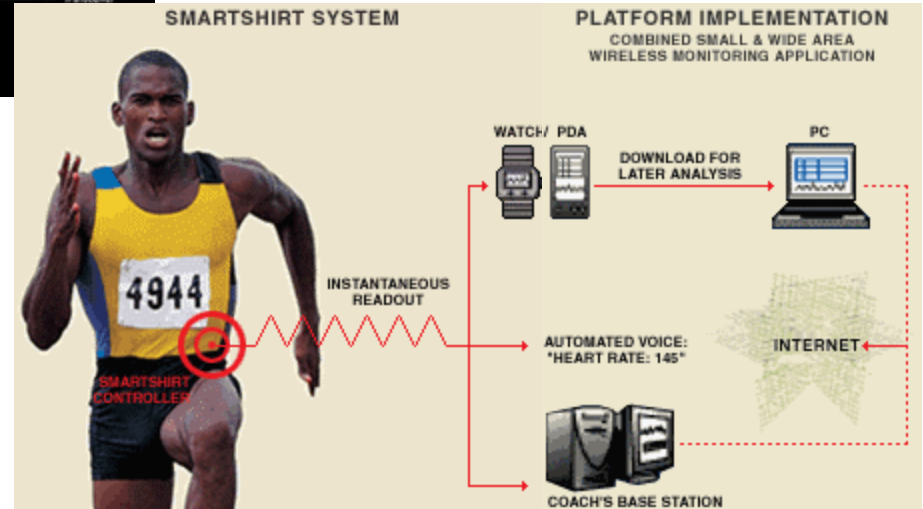
- The master controller



Nike -iPod



<http://www.sensatex.com>



Spot

Wii



Sports and Entertainment

The master controller



Nike iPod



The Smart Schematic

- Electrical Conducting Component
- Optical Component

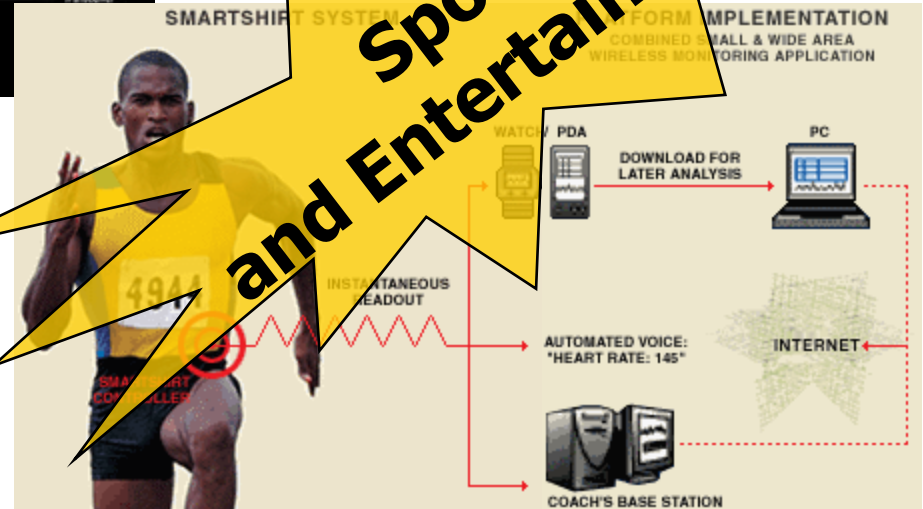
Health and Wellness

<http://www.sensatex.com>



Spot

Wii



Multiplayer Games (beyond 2nd life)

The master controller



The Smart Schematic

- Electrical Conducting Composite
- Optical Composite

Health and Wellness



<http://www.sensatex.com>

SMARTSHIRT SYSTEM FORM IMPLEMENTATION
COMBINED SMALL & WIDE AREA WIRELESS MONITORING APPLICATION

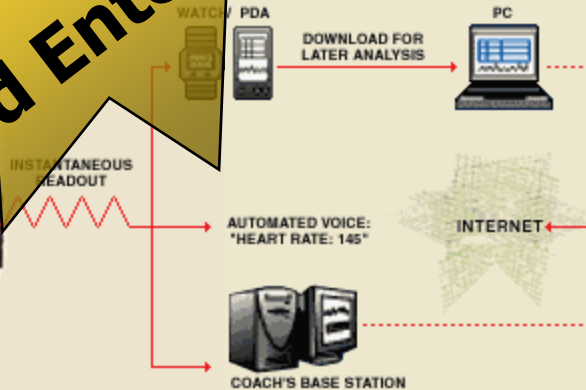


Sports and Entertainment



Spot

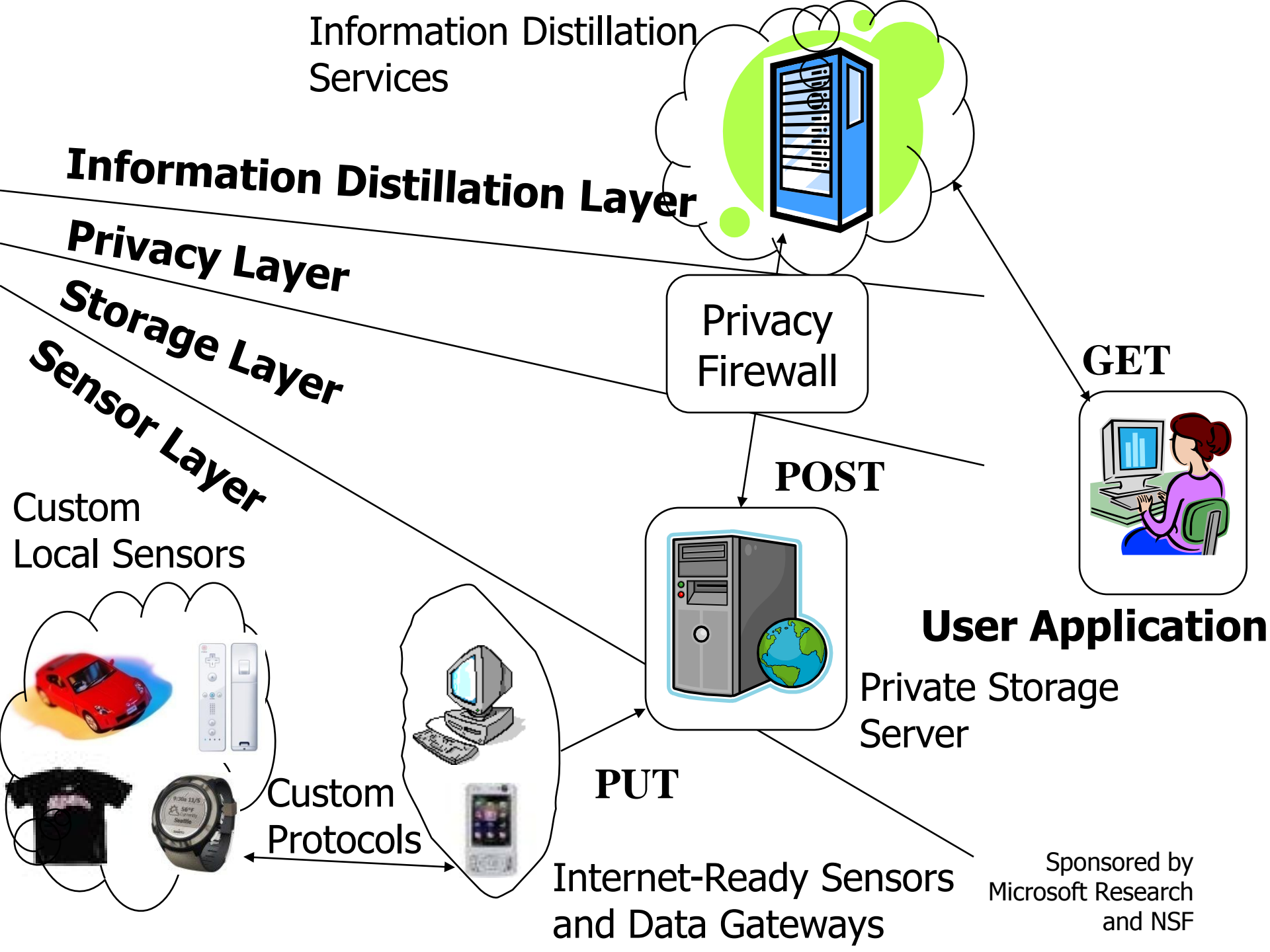
Wii Games

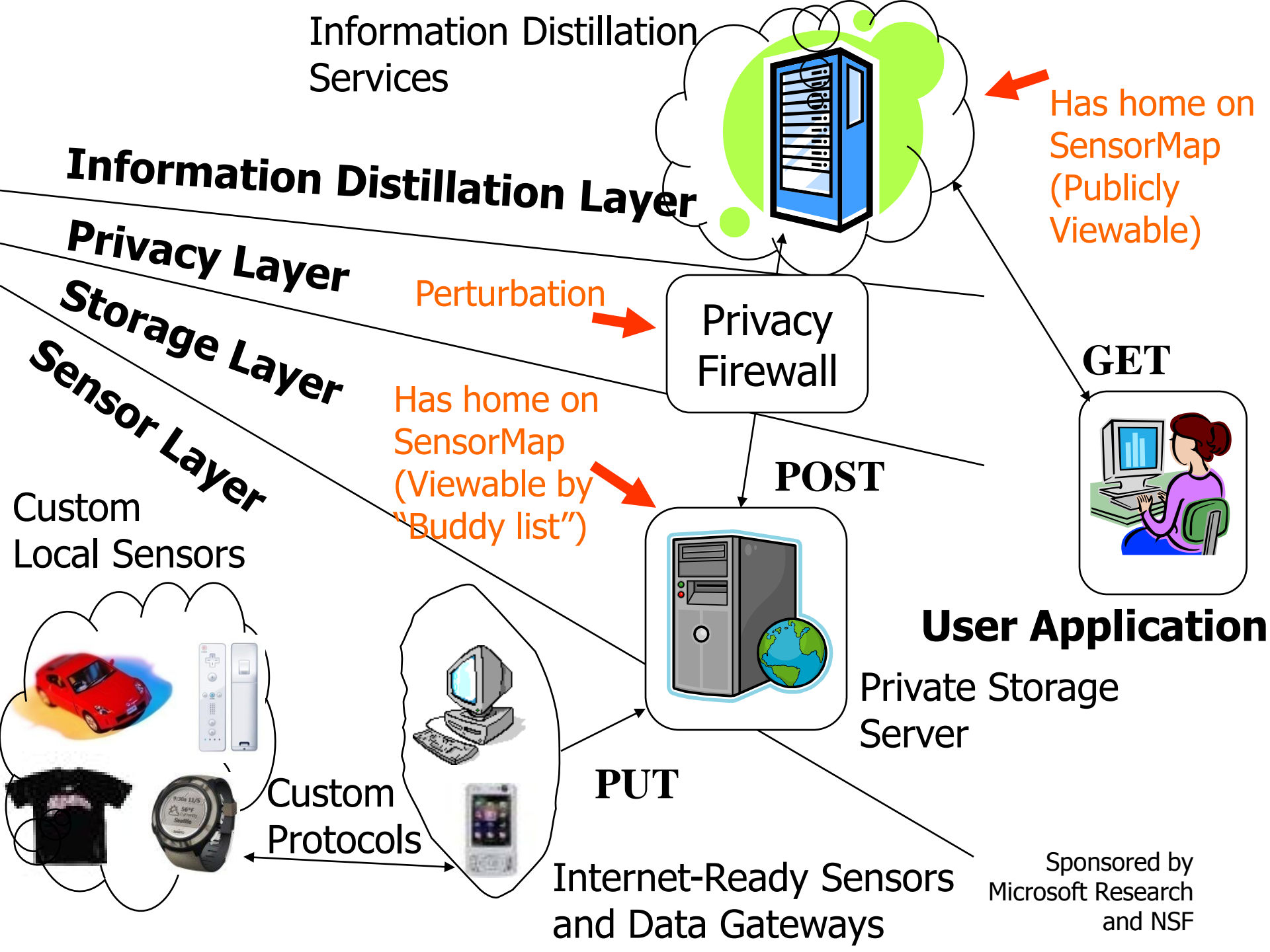


Cars with Sensors



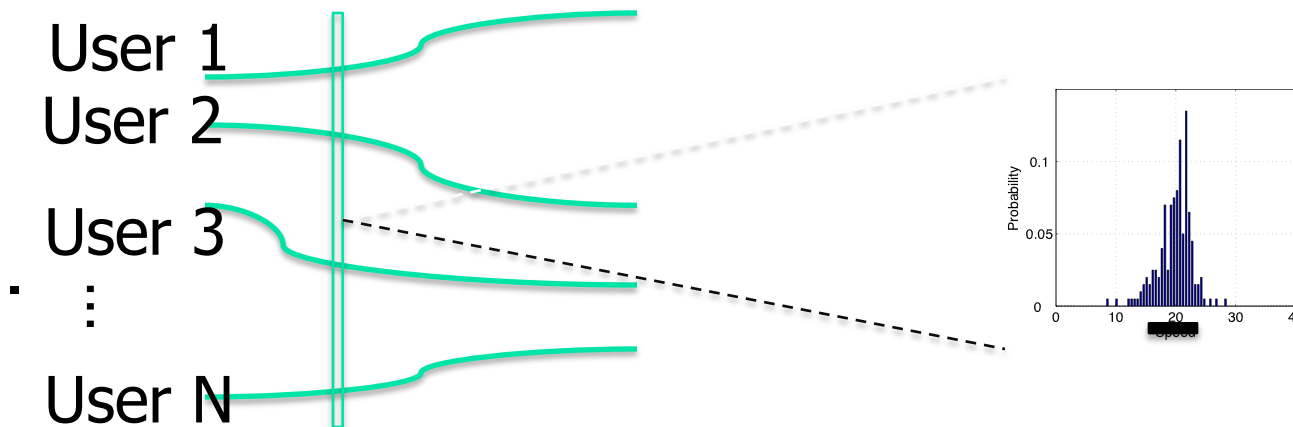
- GPS measurements offer insights on traffic conditions





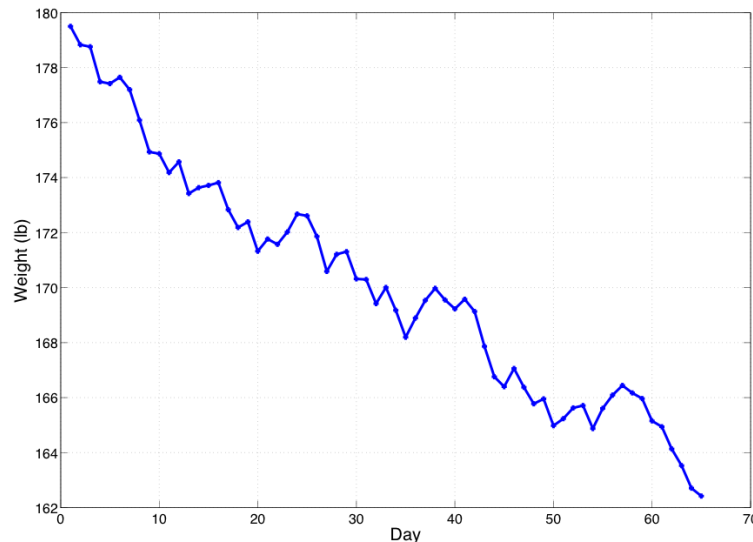
Data Sharing: A Privacy Problem Formulation

- Users have **time-series** sensor data
 - Examples – weight of an user on diet measured once a day, speed of an user measured periodically on a given route
- Goal : Let users to “lie” about their data, yet allow computing accurate distribution over the community at any point

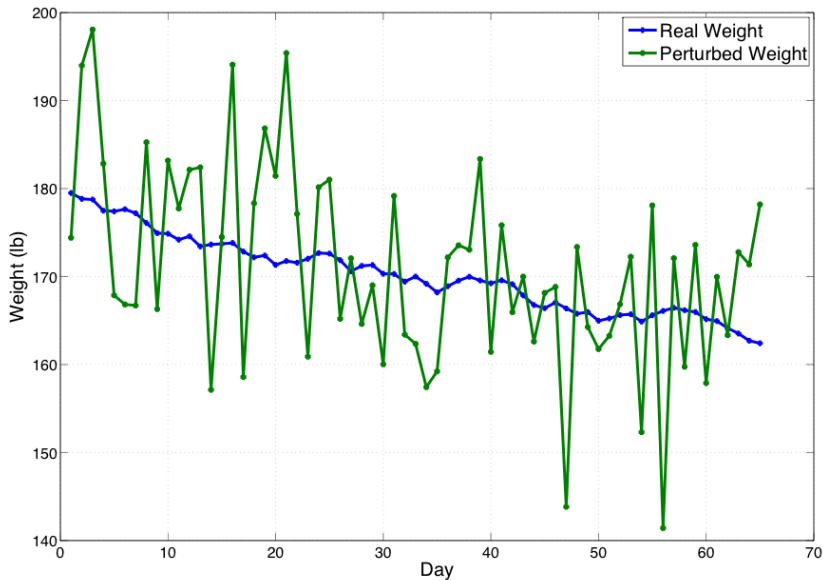


An Example

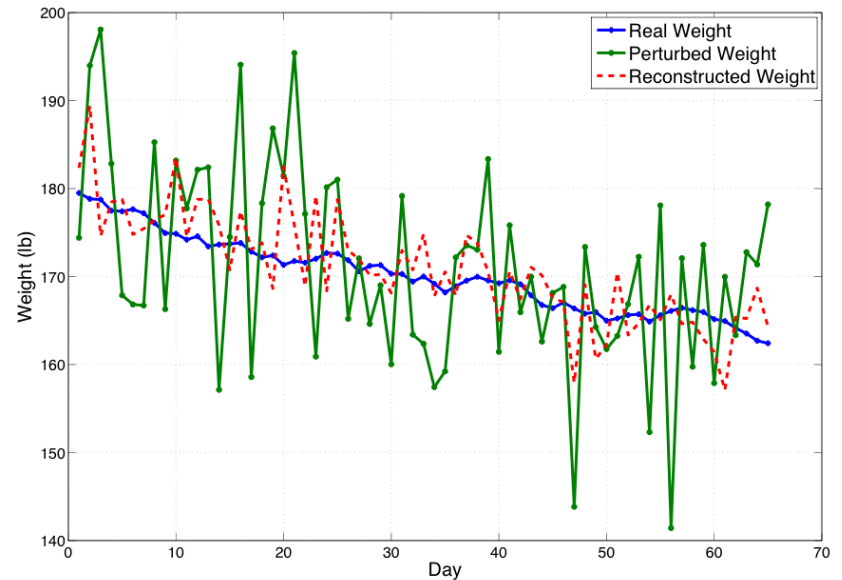
- Dieters want to share weight information to find efficacy of the given diet, without revealing their true weight, average, trend (loss or gain of weight), etc...



Perturb data? Add Noise?

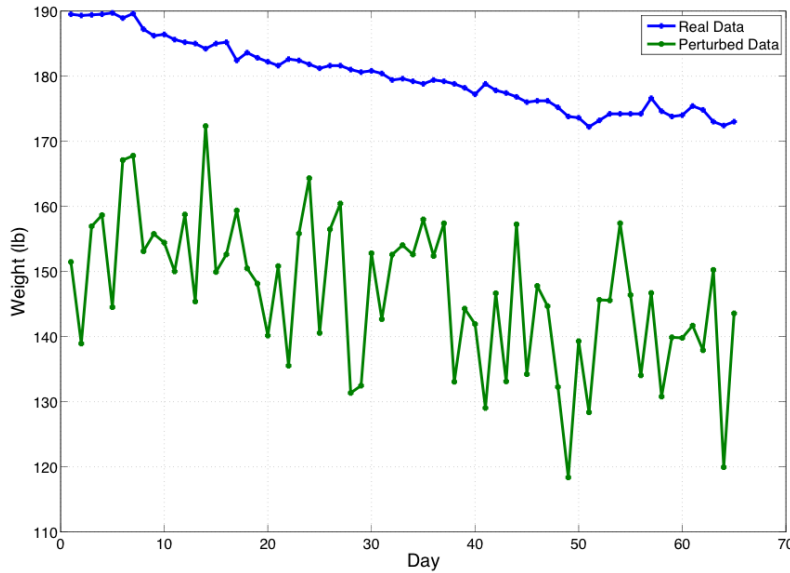


Weight curve perturbed by adding independent random noise (Gaussian in this case)

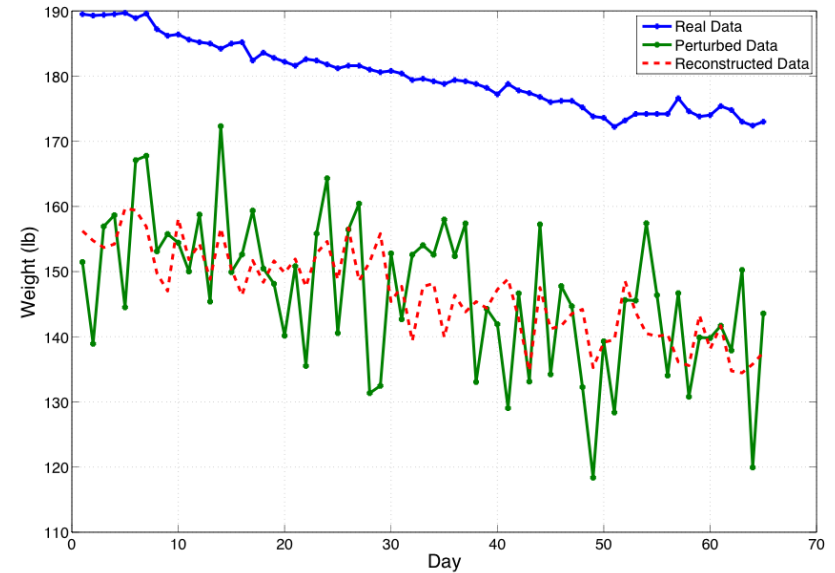


Estimation using PCA to breach privacy of user

Add Noise and Random Offset?



Weight curve perturbed by adding independent random Gaussian noise and a random offset



Estimation using PCA to estimate the data of the user



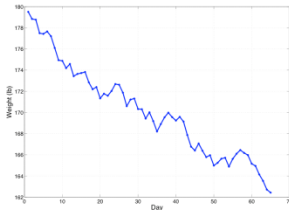
Challenge

- Develop perturbation that preserves privacy of individuals
 - Cannot infer individuals' data without large error
 - Reconstruction of community distribution can be achieved within proven accuracy bounds
 - Perturbation can be applied by non-expert users

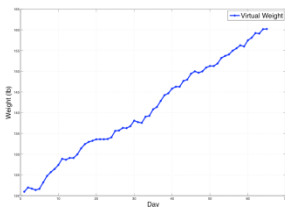
General Approach



Real user

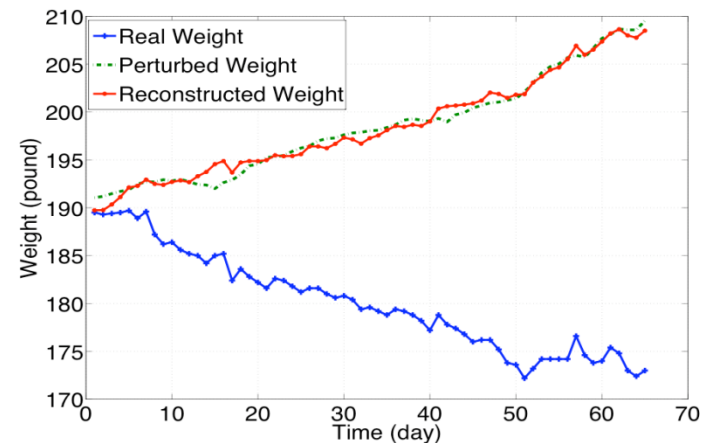


Virtual user



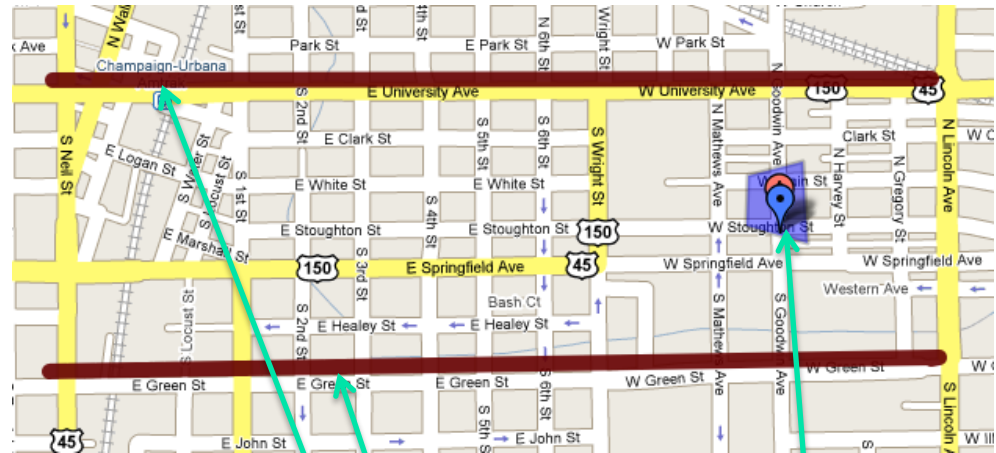
Add virtual user curve to real curve

Perturbed data curve



Traffic Analyzer

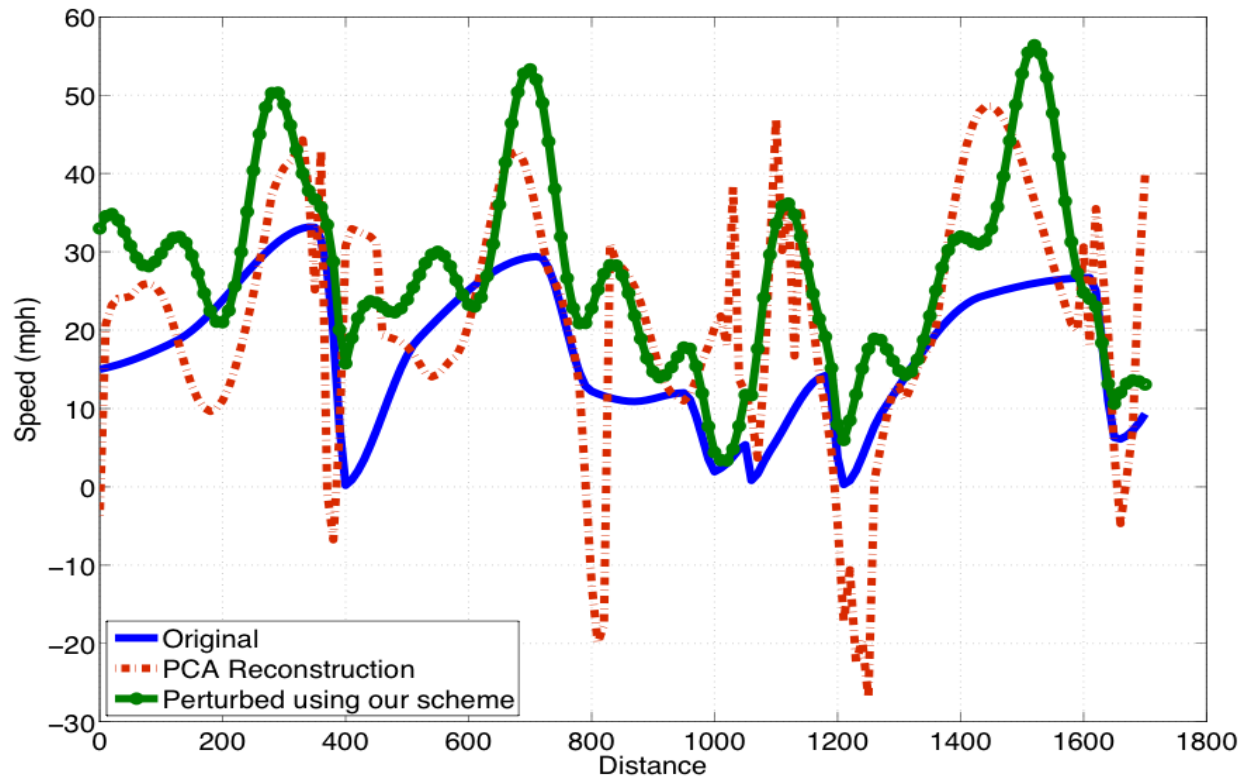
- Users share perturbed speed data with aggregation server
- Server combine perturbed speed data and use de-convolution with noise model to compute original speed distribution
- Garmin GPS used for data collection
- Results are from real data collection in Urbana-Champaign in 2008



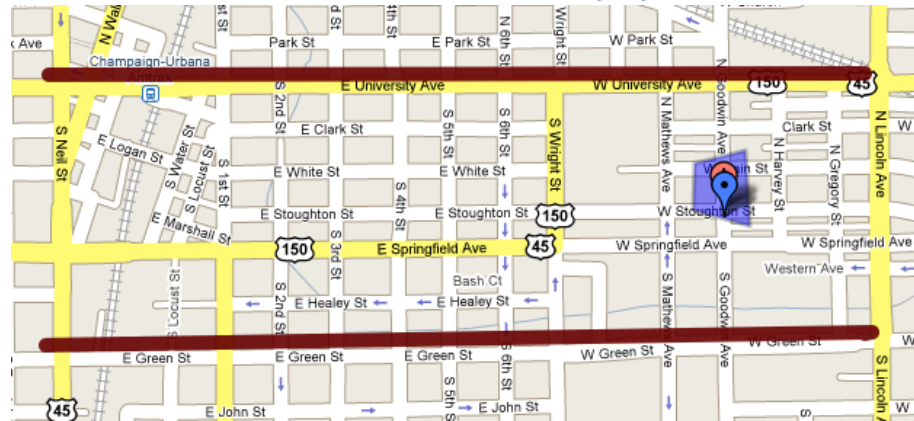
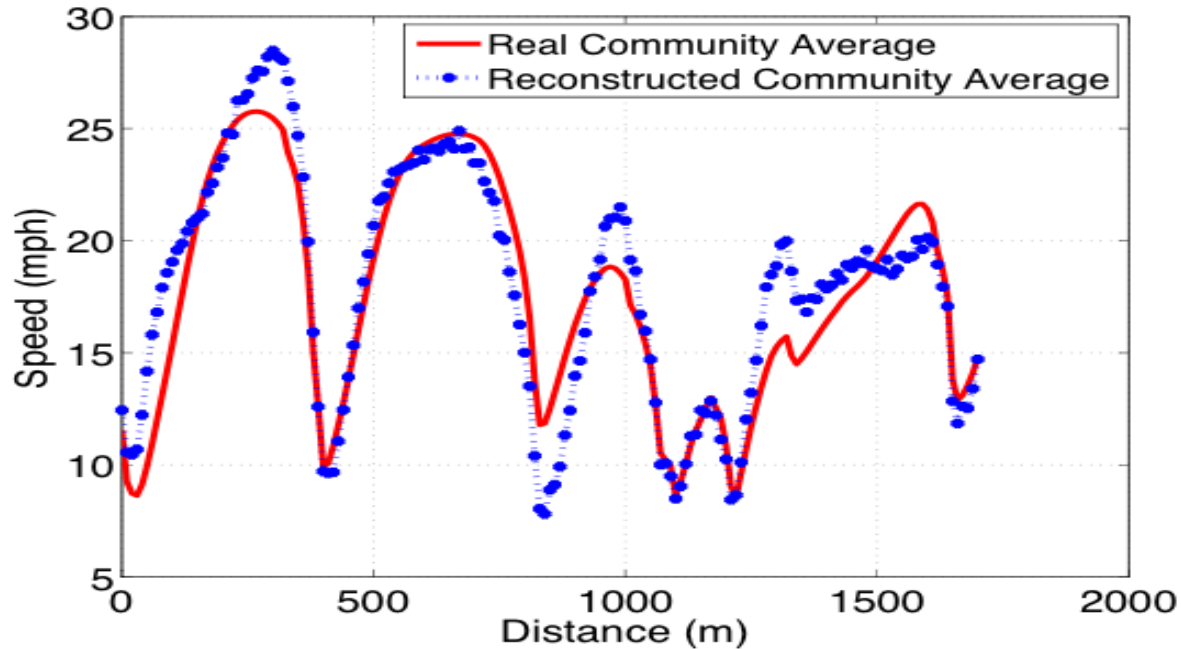
Roads for which we want to estimate average speed

Dept. of Computer Science

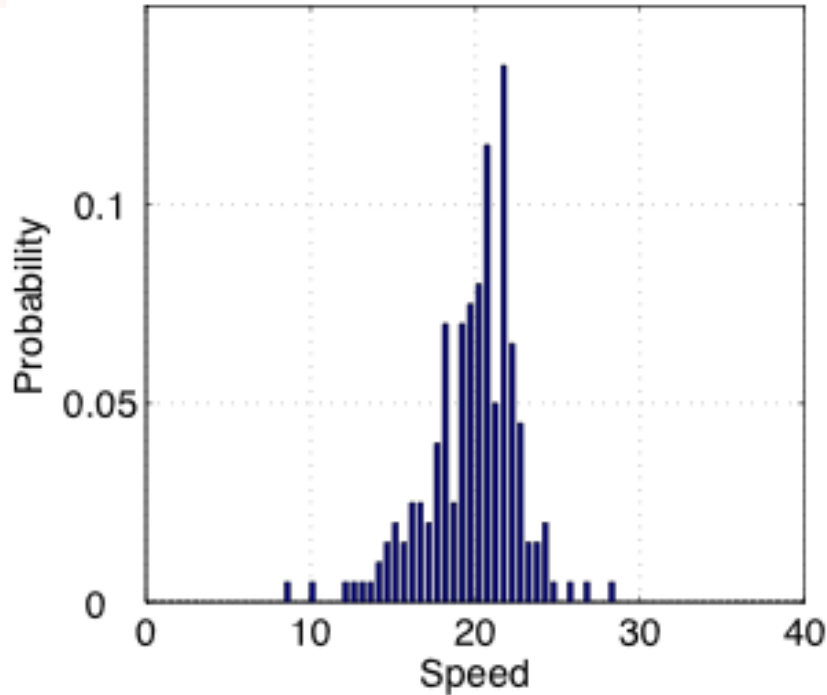
Perturbation



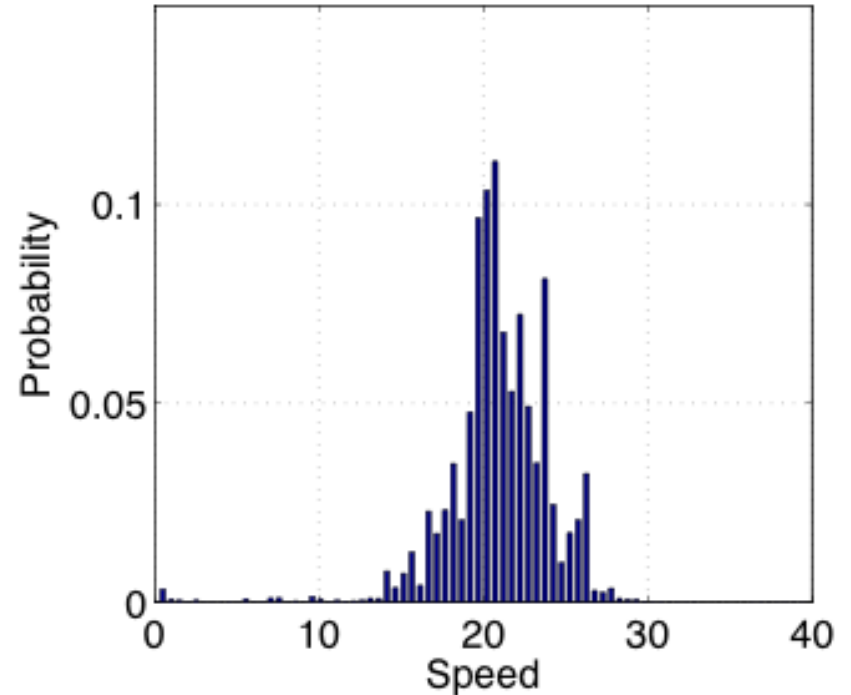
Reconstruction of Average Speed



Reconstruction of Community Speed Distribution



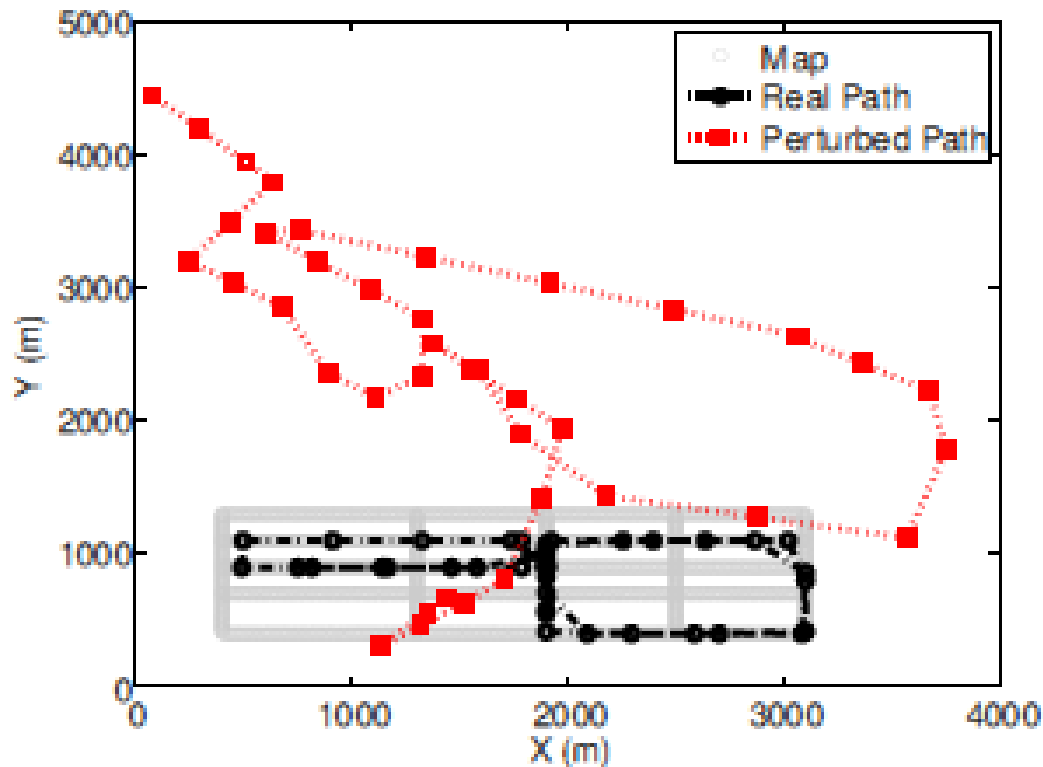
Real community distribution of speed



Reconstructed community distribution of speed

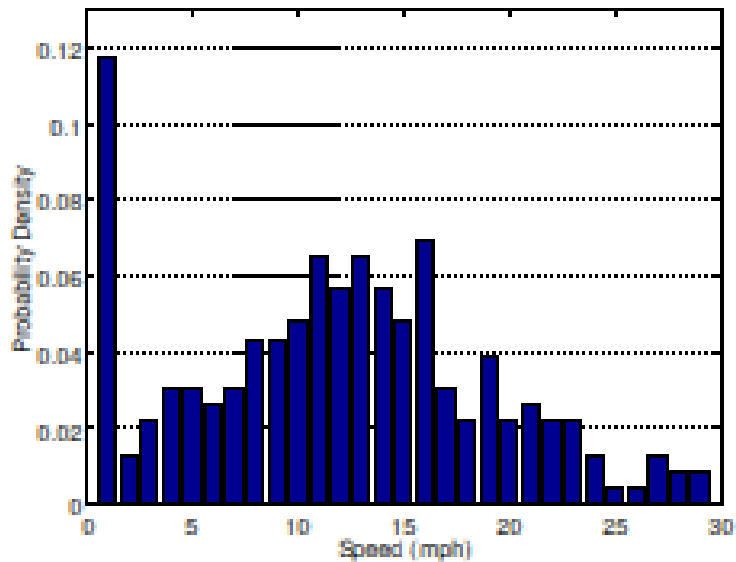
Perturbing Speed and Location

- Clients lie about both location and speed

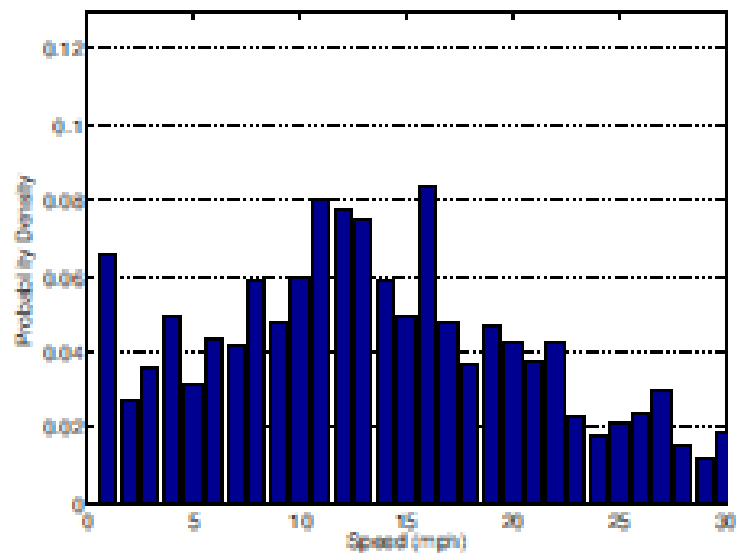


Reconstruction Accuracy

- Real versus reconstructed speed



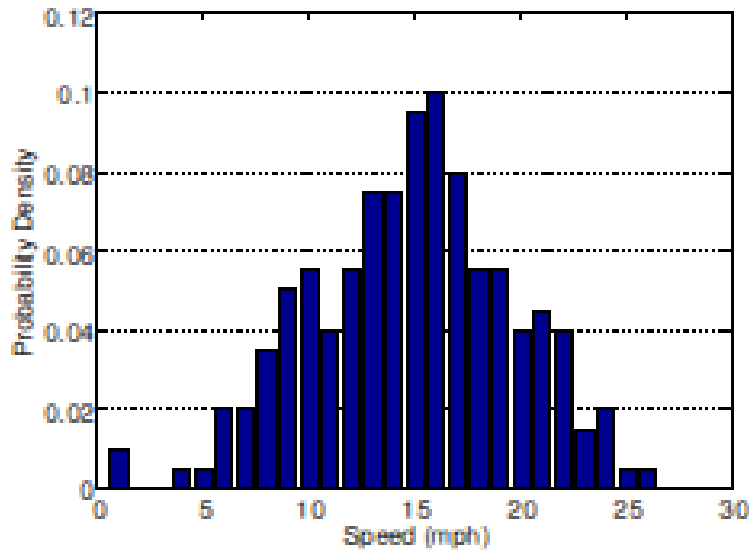
Real community distribution of speed



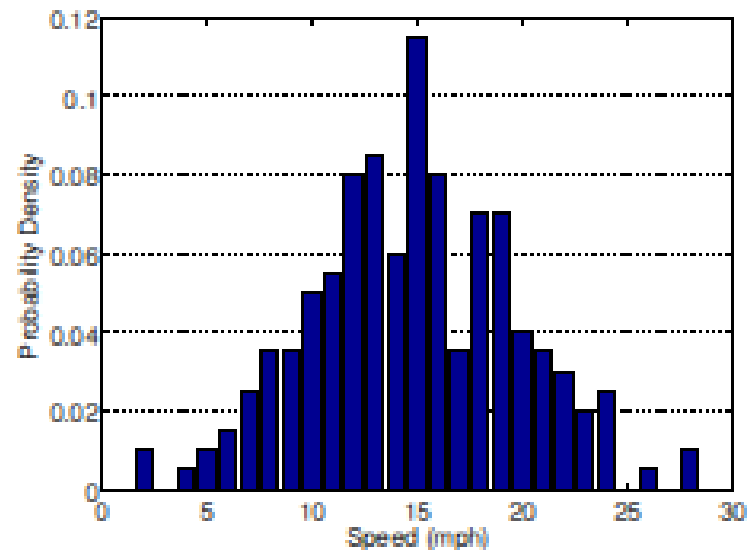
Reconstructed community distribution of speed

More on Reconstruction Accuracy

- Real versus reconstructed speed on Washington St., Champaign



Real community distribution of speed



Reconstructed community distribution of speed



How Many are Speeding?

- Real versus estimated percentage of speeding vehicles on different streets (from data of users who “lie” about both speed and location)

Street	Real % Speeding	Estimated % Speeding
University Ave	15.6%	17.8%
Neil Street	21.4%	23.7%
Washington Street	0.5%	0.15%
Elm Street	6.9%	8.6%



Sensor Map Demo

- For participatory sensing data in Urbana Champaign:

<http://atom.research.microsoft.com/sensormap/>



Other Ongoing Work with Microsoft Research

- Sneak peak:
 - Energy efficient communication on the MSR Sensornet Research Platform (mPlatform)
 - New mote platform featuring 802.11 and 802.15.4 radios
- Idea:
 - 802.15.4 is a low-power radio that is slow
 - 802.11 uses more power but is disproportionately faster (hence, less energy per bit)
 - Idea: Can we reduce overall communication energy consumption by adding 802.11 to 802.15.4?