The 4th Unit Project

Required if you are taking class for 4 credits

Offered for extra credit (5%) if you are taking class for 3 credits and *cannot* take it for 4 credits

No more than 2 people per project

Project Ideas (Set #1): Intelligent/Learning IoT Systems

- Goal: Endow everyday devices with (some aspect of) human-like behaviors/interactions!

- Examples:
  - A voice activated command interface (without cloud support)
  - A fridge that understands its content (and gives eating/cooking suggestions)
  - A dashcam that understands your emotional state so your car can do "small talk" (e.g., "Wow! You sure look happy. What happened?" or "You seem upset. Let's go eat pizza!")
  - A doorknob that detects stress, fever, etc, when you open the door.
  - A floor sensor that monitors your activity
  - A fitbit that detects the most suspenseful scenes in a scary movie you are watching

Intelligent/Learning IoT Systems

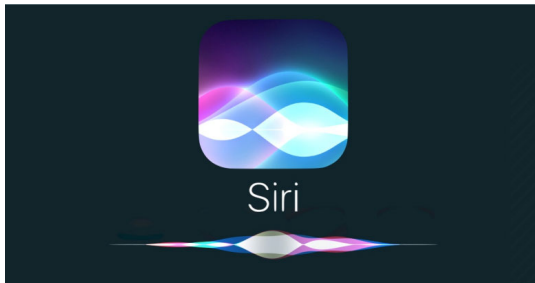Empower everyday objects with human-like interaction modalities

Embedded & Mobile Devices

Smart City Infrastructure

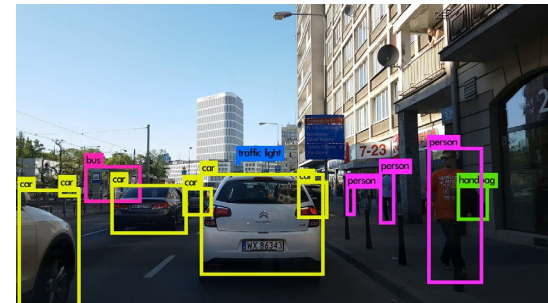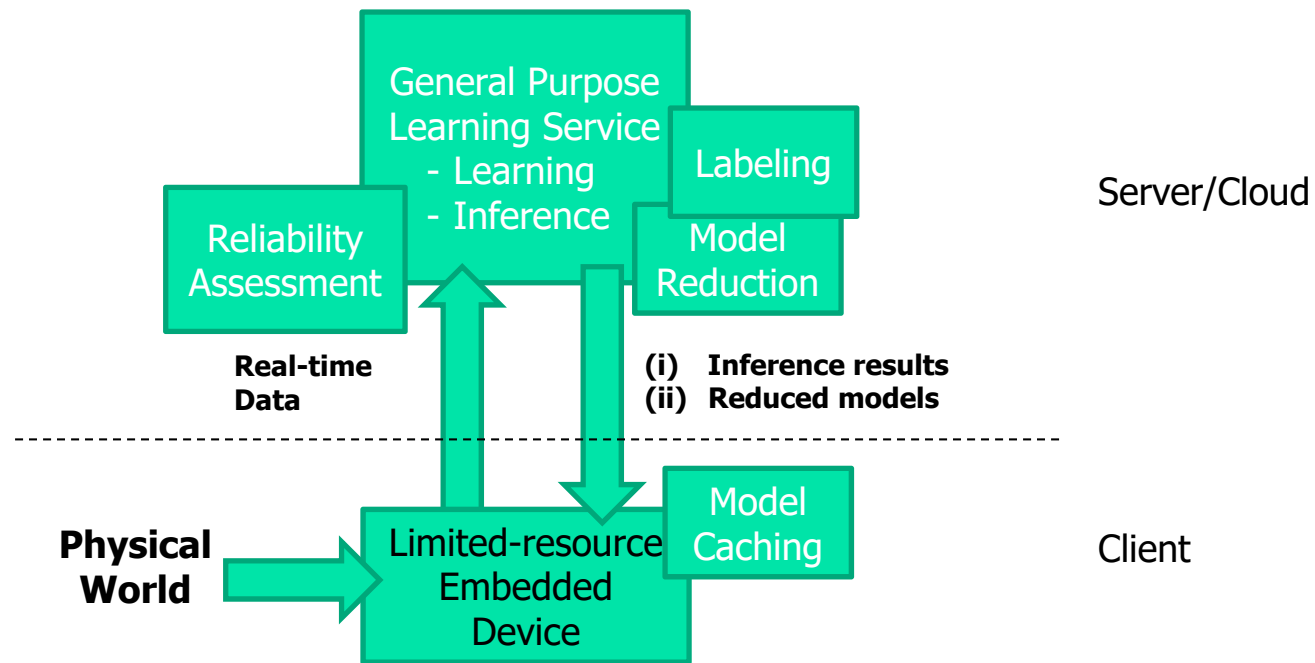# Deep Learning: Towards Human-like Interactions with the Physical Environment
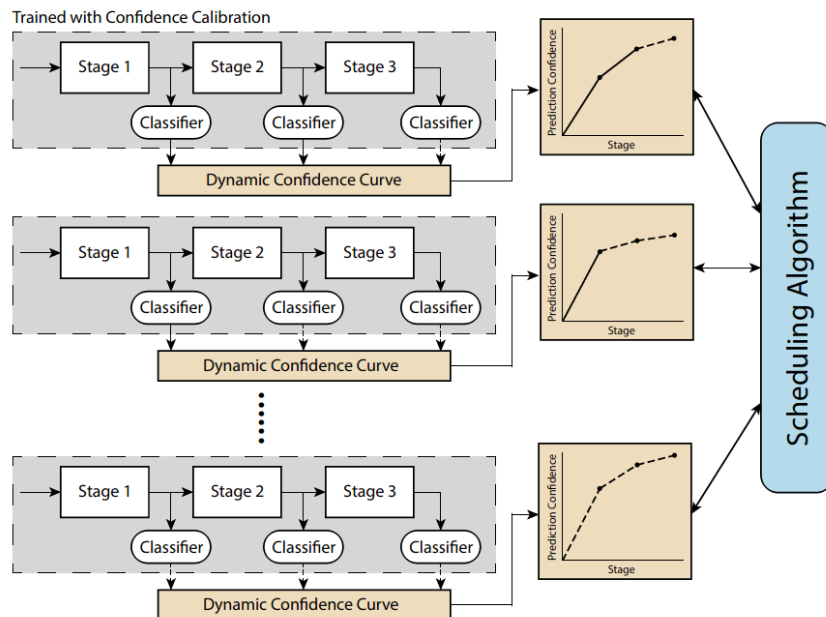
Speech Recognition

Context/Activity Recognition

Vision/Object Detection

The "Intelligence as a Service" Model

General Purpose
Learning Service
- Learning
- Inference

Labeling

Model
Reduction

Reliability
Assessment

Server/Cloud

**Real-time
Data**

**(i)   Inference results
(ii) Reduced models**

Model
Caching

**Physical
World**

Limited-resource
Embedded
Device

Client

Server-side Real-time Scheduling of Inference Tasks

Challenge: Different data inputs offer different degrees of complexity. Some are easily recognizable patterns, but others are not.

Idea:
- Break execution into stages
- Use the confidence estimates to predict utility from executing the next stage of each task
- Scheduler executes the task (stage) with the highest marginal utility

# Project Ideas (Set #2): Disaster Response Services

- Develop information services for disaster response using a combination of sensor data and crowdsourcing
  - Evacuation route planning
  - Damage assessment
  - Survivor search

Damage Assessment and Route Planning

- Route reconnaissance for evacuation, logistics/resupply planning, humanitarian aid distribution, …
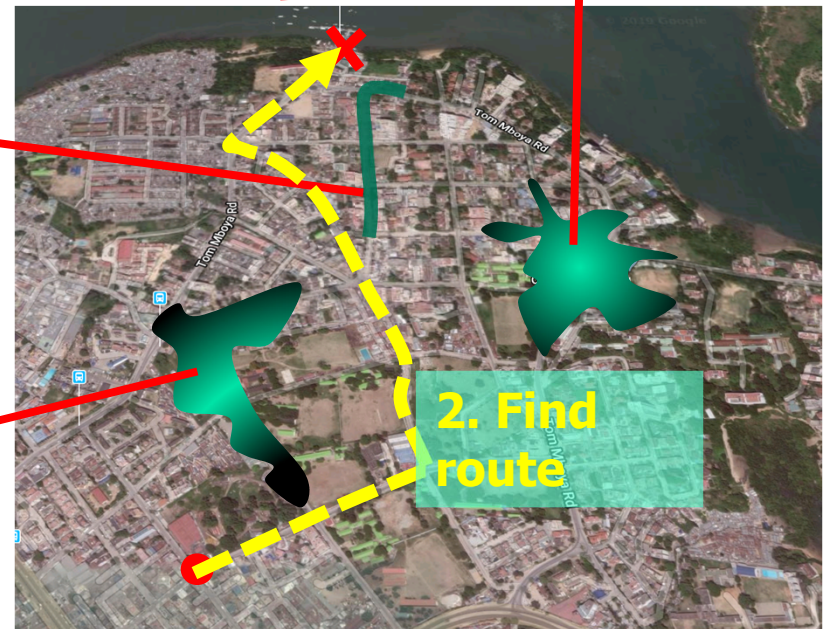
1. Use Images from selected sensors to find "safe routes"
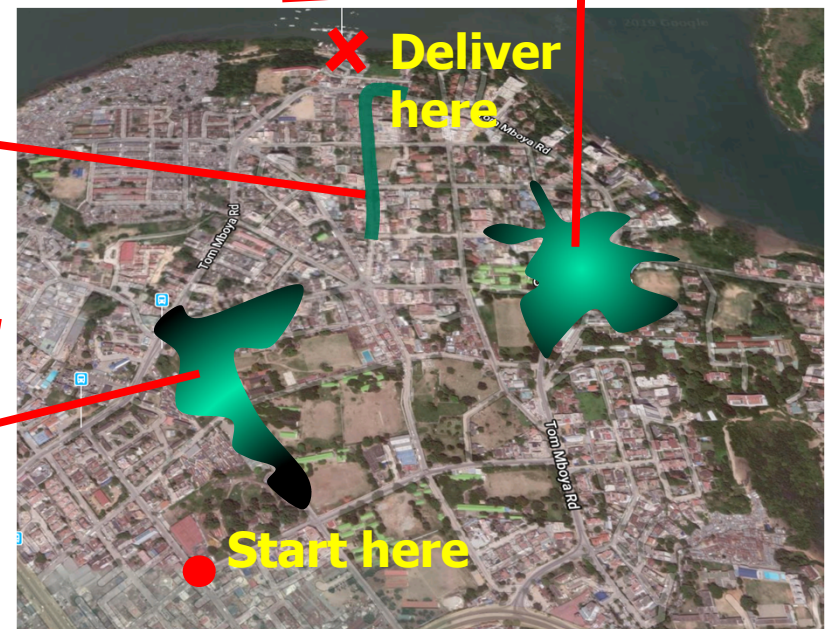
Deliver here

Start here

Damage Assessment and Route Planning

- Route reconnaissance for evacuation, logistics/resupply planning, humanitarian aid distribution, …
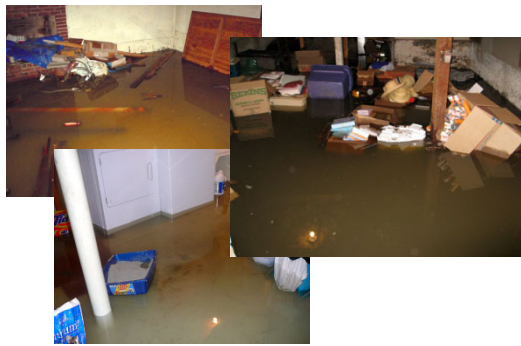
2. Find route

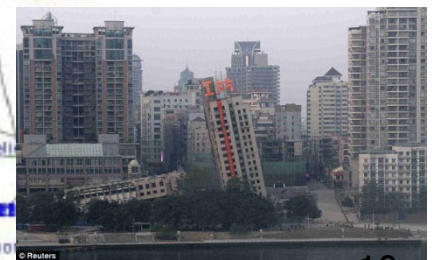# Challenge: Data Selection to Maximize Coverage

Fire on 6th and Main.

Collapse on Park Ave.
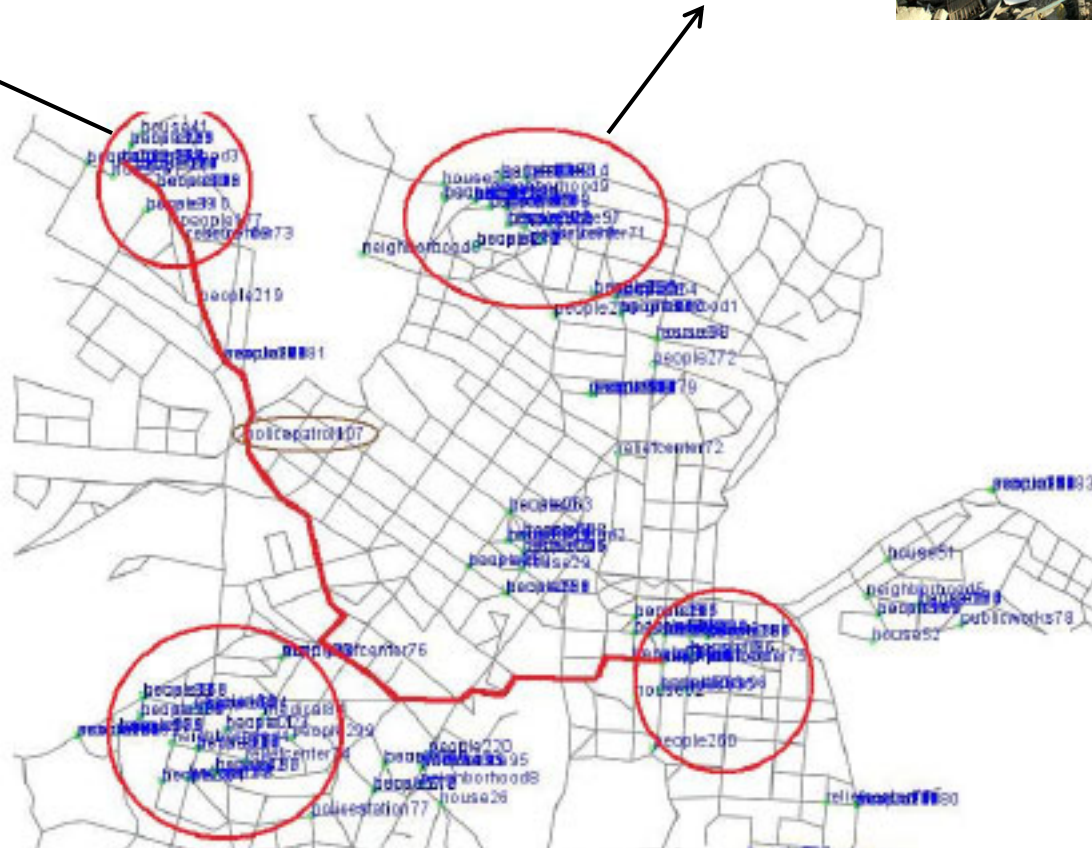
Flooding on State St.

Structural damage on Pier Square

Example of Bad Coverage

Fire on 6th and Main.

Collapse on Park Ave.

An Example of Poor Data Selection (Low Coverage)

11

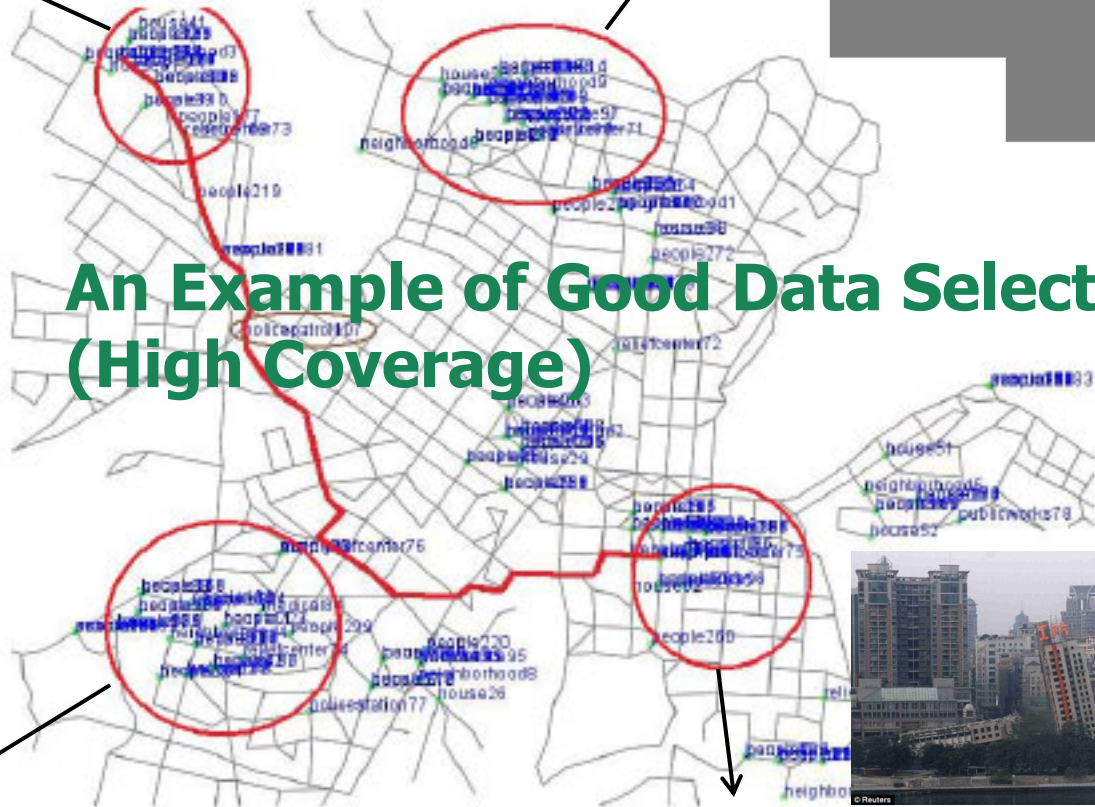# Example of
Good Coverage

Fire on 6th and Main.

Collapse on Park Ave.
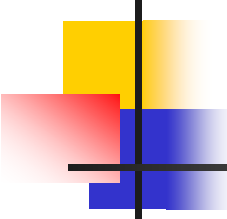
An Example of Good Data Selection (High Coverage)

Flooding on State St.

Structural damage on Pier Square

A Real-time Scheduling Approach: Coverage-maximizing Priorities

- Implement coverage-maximizing in-network prioritization for real-time data forwarding and storage

  - Objects are forwarded/dropped in a priority order aimed to maximize coverage of delivered content while meeting timing constraints

    - Objects similar to previously forwarded ones get lower priority

  - Challenge: Forwarding and dropping must be made aware of the degree of semantic redundancy (i.e., similarity) between objects

Project Ideas (Set #3):
Real-time Exploitation of Social Media as Sensor Networks

Physical World

Civil Unrest

Hurricanes

Man-made disasters

People

Sensors

→ **Information**

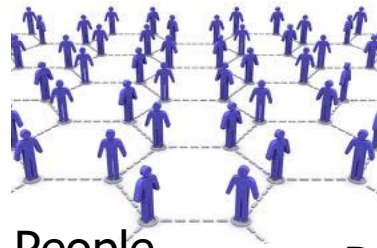# Real-time Exploitation of Social Media as Sensor Networks

Physical World

Civil Unrest

Hurricanes

Man-made disasters

People

Sensors

Data Mining/Machine Learning/Estimation

Estimated State

There exists a *unique "ground truth"* state (vector) is being estimated

As opposed to: opinion mining, sentiment analysis, statistical correlation mining, ...
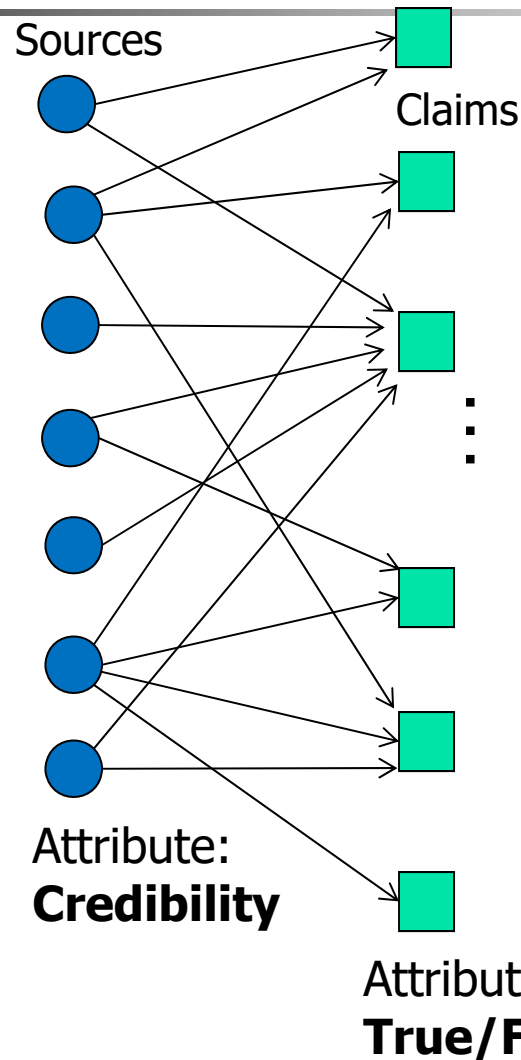
15

Social Media Misinformation Detection

Events

Civil Unrest

Hurricanes

Man-made disasters

Sources

Claims

Attribute: **Credibility**

Attribute: **True/False**
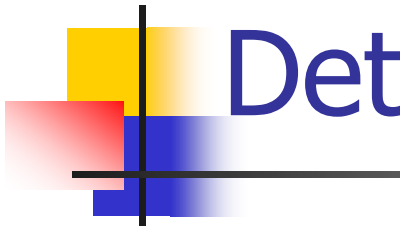
Maximum Likelihood Estimation
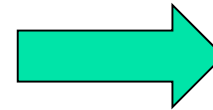
Clean Event Summary

- Credibility of sources
- Correctness of claims
- Confidence intervals

- Formulate the fact-finding problem as one of maximum likelihood estimation
- Solve it using the *Expectation Maximization* (EM) algorithm
- Compute a bound on estimation accuracy (using the Cramer Rao Bound)

16

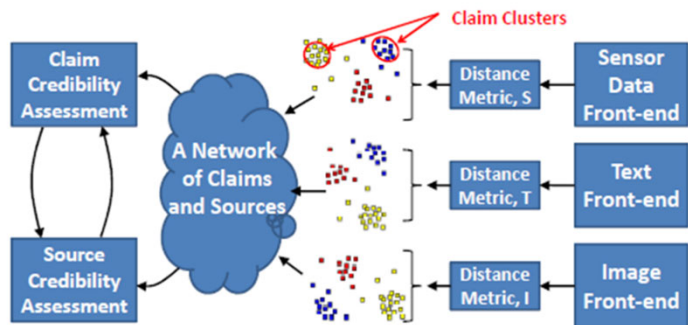# Social Media Misinformation Detection

Events

Civil Unrest

Hurricanes

Man-made disasters

Sources

Claims
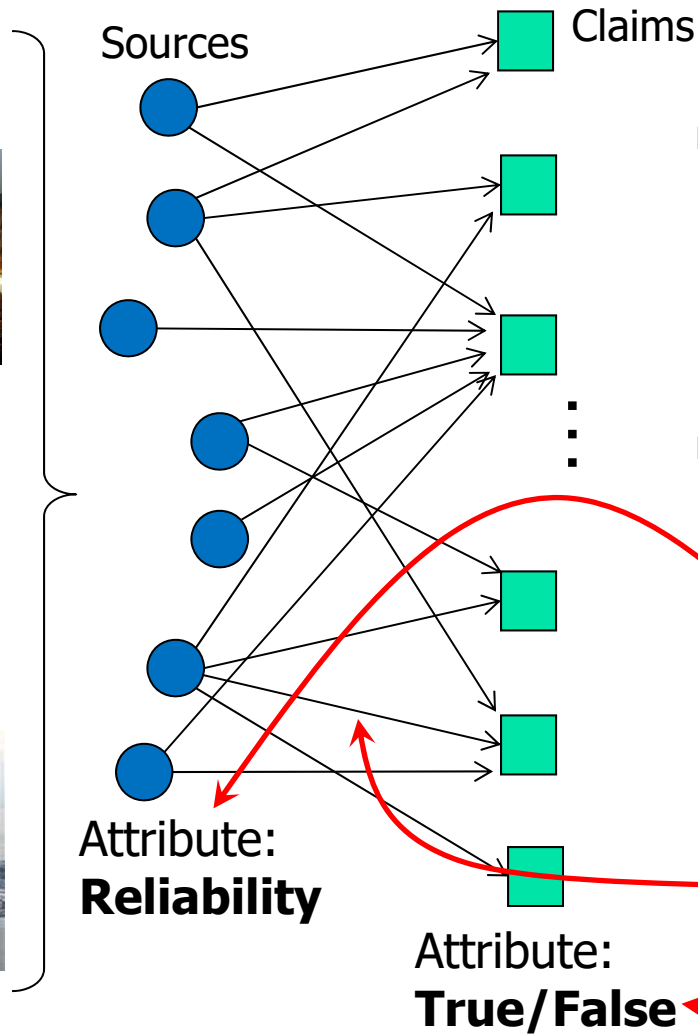
- Joint estimation of
  - Source reliability
  - True/false value of each observation
- Given
  - Who said what

Attribute: **Reliability**

Attribute: **True/False**

$$P(SC|\theta) = \sum_z P(SC, z|\theta)$$

Social Media Misinformation Detection
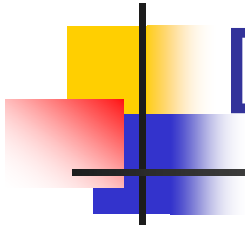
Create new task

Keyword 1	Syria	or
Keyword 2	Chemical	or
Keyword 3	Weapons	or from
Latitude	35.14051	Longitude 37.10675 Radius (miles) 100 Hide Map

Crawl with Search API   Enter your short name here ...   Create Task

18

Social Media Misinformation Detection

- Example of tweets collected in the aftermath of the Syrian chemical weapons attack in August 2013.

- Tweets were crawled for ten days after the event using the keywords "Syria", "attack", "dead"

- Table shows results of maximum likelihood estimation, automatically separating tweets into "socially corroborated" and "not corroborated".

Triage Result: Recommended for Viewing	Triage Result: Dismissed/Unimportant
Medecins Sans Frontieres says it treated about 3,600 patients with 'neurotoxic symptoms' in Syria, of whom 355 died http://t.co/eHWY77jdS0	So sad. All but one of the activists who filmed the chemical attack in Syria died of toxins: http://t.co/7Xc9u8achL
Weapons expert says #Syria footage of alleged chemical attack "difficult to fake" http://t.co/zfDMujaCTV	Saudis offer Russia secret oil deal if it drops Syria via @Telegraph http://t.co/iOutxSiaRs
U.N. experts in Syria to visit site of poison gas attack http://t.co/jol8OlFxnf via @reuters #PJNET	Putin Orders Massive Strike Against Saudi Arabia If West Attacks Syria http://t.co/SFLJ9ghwbt
Syria Gas Attack: 'My Eyes Were On Fire' http://t.co/z76MiHj0Em	Miley Cyrus twerks meanwhile in other news the U.S.A. might declare war on Syria....
Long-term nerve damage feared after Syria chemical attack http://t.co/8vw7BiOxQR	I posted a new photo to Facebook http://t.co/FRWBFC0vKb
Syrian official blames rebels for deadly attack http://t.co/76ncmy4eqb	Two Minds on Syria http://t.co/ogDjKFH7Rs via @NewYorker
Assad regime responsible for Syrian chemical attack, says UK government http://t.co/pMZ5z7CsNZ	We may be going to war in Syria, and somehow Miley Cyrus Is trending on twitter
US forces move closer to Syria as options weighed: WASHINGTON (AP) — U.S. naval forces are moving closer to Sy... http://t.co/F6UAAXLa2M	Syrian Chemical Weapons Attack Carried Out by Rebels, Says UN (UPDATE) http://t.co/lN4CkUePUj #Syria http://t.co/tTorVFUfZF
400 tonnes of arms sent into #Syria through Turkey to boost Syria rebels after CW attack in Damascus --> http://t.co/KLwESYChCc	For those in the US, please text SYRIA to 864233 to donate $10 via @unicefusa http://t.co/YMXnrk1jcb #childrenofsyria
UN Syria team departs hotel as Assad denies attack http://t.co/O3SqPoiq0x	Attack! http://t.co/wY5KKm7R3s
Vehicle of @UN #Syria #ChemicalWeapons team hit by sniper fire. Team replacing vehicle & then returning to area.	A fathers last words to his dead daughters killed by Bashar al-Assad & his supporter army with chemical weapon attack http://t.co/DN25pLfCq8
International weapons experts leave Syria, U.S. prepares attack. More @ http://t.co/4Z62RhQKOE	What the media isn't telling you about the Syrian chemical attack http://t.co/LQ479S1Tiv
Military strike on Syria would cause retaliatory attack on Israel, Iran declares http://t.co/M950o5VcgW	France on the phone. Apparently they surrendered to #Syria weeks ago.
Asia markets fall on Syria concerns: Asian stocks fall, extending a global market sell-off sparked by growing ... http://t.co/06A9h2xCnJ	Poll: Do you think the chemical attack in #Syria could have been a false flag attack to push for war? RT for yes. Favourite for no
UK Prime Minister Cameron loses Syria war vote (from @AP) http://t.co/UlFF1wY9gx	Lebanon was once part of Syria and will forever be with Syria. #PrayForSyria #PrayForLebanon

19

Extensions:

- The current estimation framework makes simplifying assumptions on sources and observations (e.g., independence)
  - How to detect copying/influence?
  - How to account for source non-independence due to information dissemination?
  - How to account for physical relations between observations?
  - How to include inference and other logical relations when some observations imply others?
  - How to separate "opinions" from ground-truthable facts?
  - How to de-bias observations?
  - How to detect degree of "polarization" among sources?
  - How to compute fundamental error bounds?
  - How to influence sources such as error bound is reduced?

20

Project Ideas (Set #4) Event Detection

Physical target

Response of physical propagation medium
(e.g., acoustic, vibration, optical, …)

Received signature (energy in
multiple signal frequency bands)

Event Detection: An Analogy

Physical target

Response of physical propagation medium
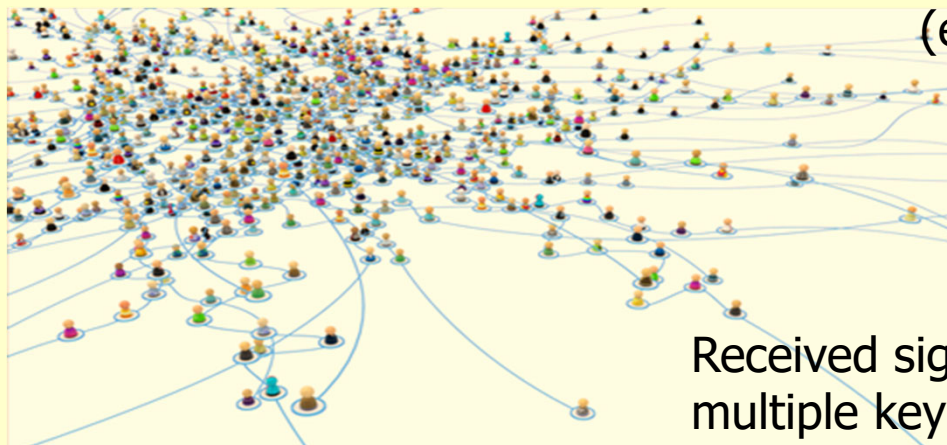(e.g., acoustic, vibration, optical, …)

Received signature (energy in multiple signal frequency bands)

Physical event

Response of social propagation medium
(e.g., tweets)

Received signature (energy in multiple keyword frequency bands)

Event Detection:
The Demultiplexing Challenge

- A world of "protest" – this morning:
  - Angry French farmers and 1,000 tractors head for Paris protest. Photo @MartinBureau1 #AFP http://t.co/j5DdveSHZh
  - VIDEO: Tractor protest descends on Paris: French farmers protesting about high taxes have taken a convoy of tr... http://t.co/hKievMFpq3
  - WATCH LIVE: Farmers on tractors gather in Paris streets https://t.co/peTOvKrIAF http://t.co/3vDK6qc060
  - MORE: Police detained refugees who lay on train tracks in protest at being taken to a camp,This is 2015 not 1940's http://t.co/TbQrwWBWrH
  - RIGHT NOW: Activists & giant polar bear protest Arctic oil outside Shell London HQ http://t.co/1Ae9mgc1ZF #ArcticRoar http://t.co/5tJaKv0mHZ
  - Underwater sculptures emerge from Thames in climate change protest http://t.co/mg6RiURn6t
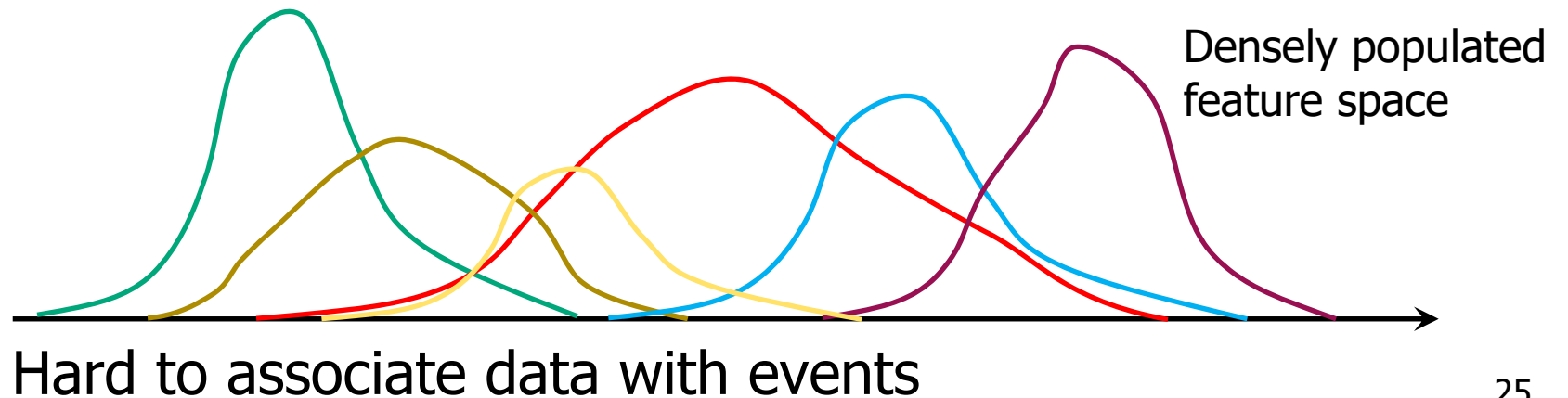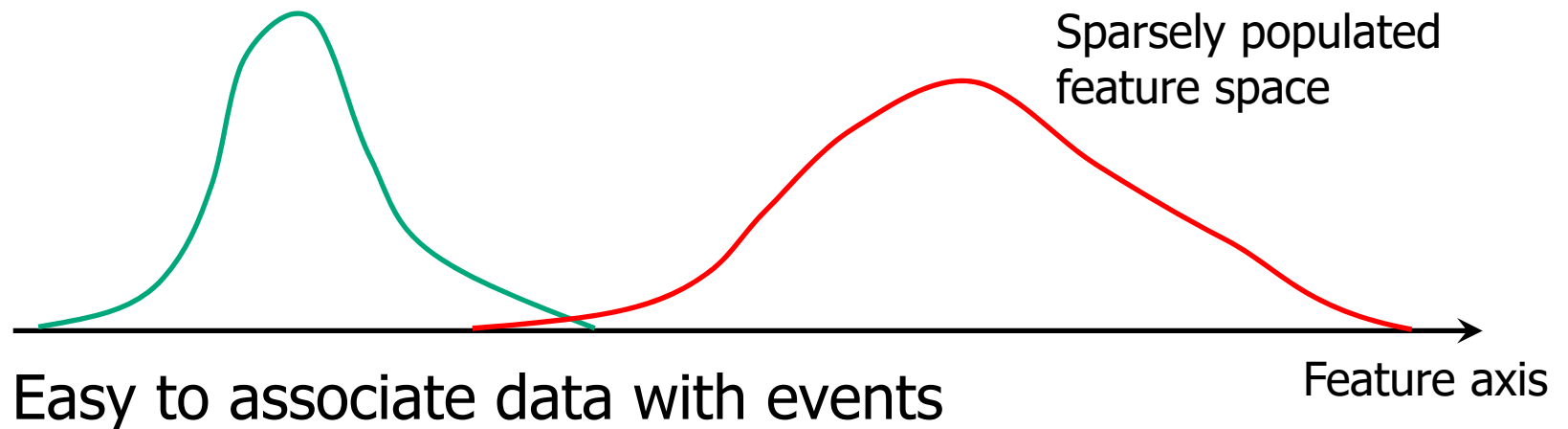
Events and Information Signals

- *Observation:* Targets can be recognized using frequency domain signatures

- *Question:* Can we detect and track events using "frequency domain" signatures only?

  - At first glance: text has complex semantics, so the ordering of keywords has great impact on meaning

    - "John killed Mary" versus "Mary killed John"

  - Do we need natural language processing to identify and track distinct events?

24

Events and Information Signals

Sparsely populated feature space

Easy to associate data with events

Feature axis

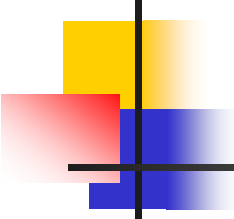Densely populated feature space

Hard to associate data with events

25

Events and Information Signals

- Most languages have about 10,000 frequent words.
- Consider a 2-word event signature
  - There are at least 100,000,000 possible signatures
- Number of "events" in a Twitter data trace may be in the 100s or 1000s
- The space of keyword signatures is vastly sparse:
  - Different events → Different signatures (assuming independent keywords)

Event Detection, Consolidation, and Tracking:
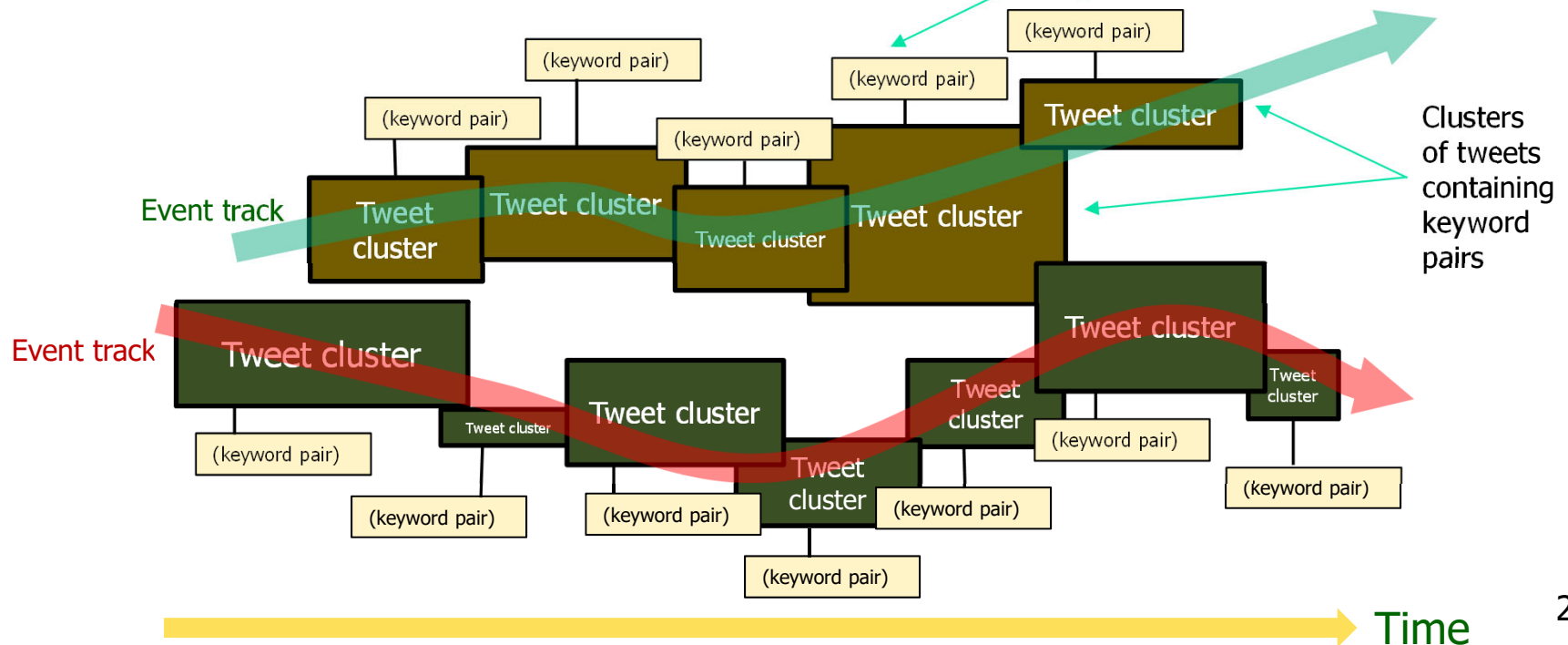## Signal Processing Questions

- How to detect new event signatures?
  - Find high-information-gain signatures (new spikes in the frequency spectrum)
  - Bin tweets that contain a new signature into a cluster
  - Determine if this cluster is of a new event or not using frequency domain distance (note: some events will have more than one signature)
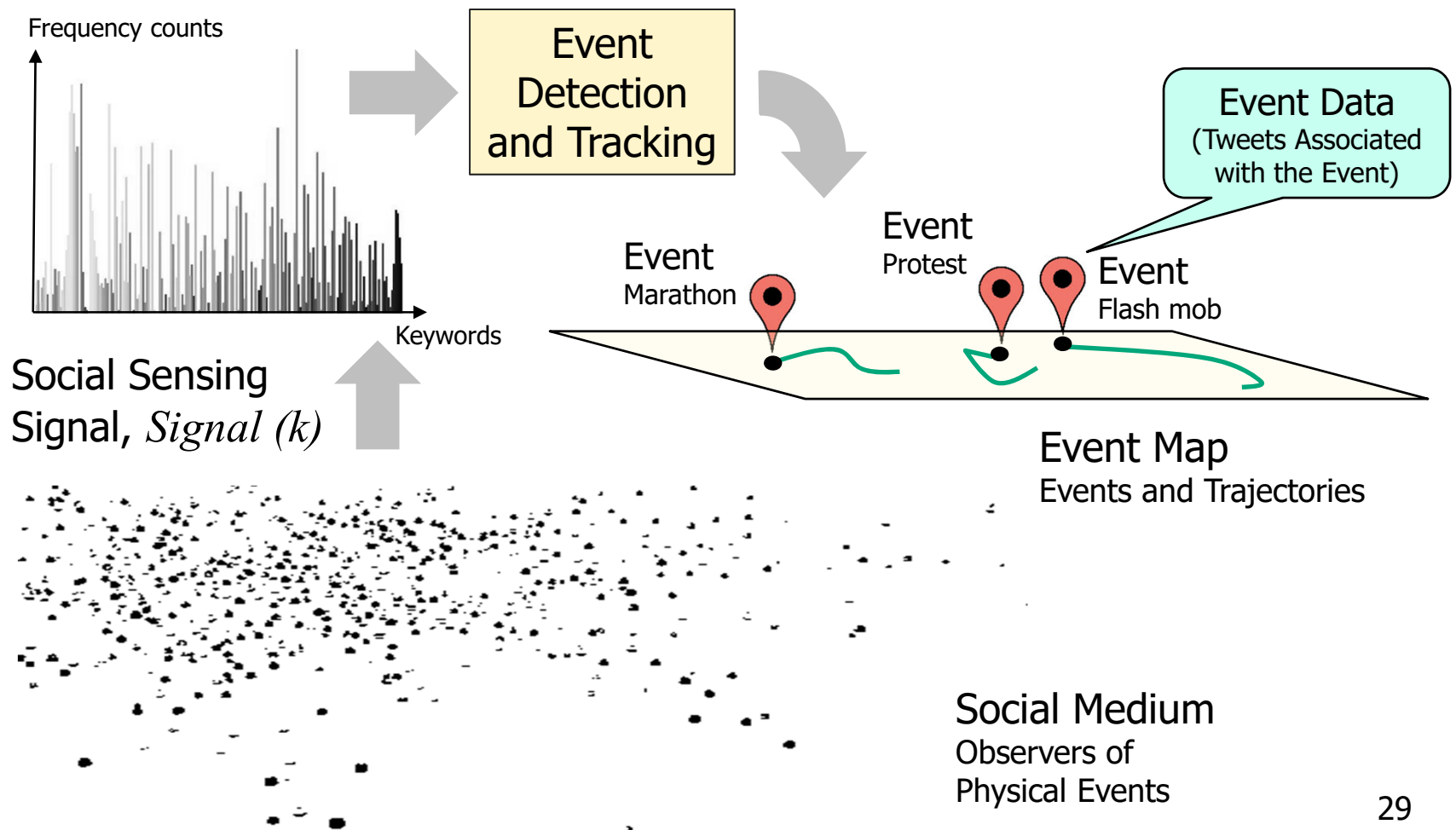
Event Detection, Consolidation and Tracking

Three key ideas:
1. Use information gain to detect new keyword pairs (event signatures)
2. Each pair gives rise to a cluster of tweets (that contain the pair)
3. Merge clusters with similar keyword distributions

Automatically detected high-information-gain keyword pairs

Clusters of tweets containing keyword pairs

Event track

Event track

(keyword pair)

(keyword pair)

(keyword pair)

(keyword pair)

(keyword pair)

(keyword pair)

(keyword pair)

(keyword pair)

(keyword pair)

(keyword pair)

(keyword pair)

(keyword pair)

(keyword pair)

(keyword pair)

Tweet cluster

Tweet cluster

Tweet cluster

Tweet cluster

Tweet cluster

Tweet cluster

Tweet cluster

Tweet cluster

Tweet cluster

Tweet cluster

Tweet cluster

Tweet cluster

Tweet cluster

Time

The Social Signal Layer

Frequency counts

Event Detection and Tracking

Event Data (Tweets Associated with the Event)

Event Marathon

Event Protest

Event Flash mob

Keywords

Social Sensing Signal, $Signal\ (k)$

Event Map
Events and Trajectories

Social Medium
Observers of
Physical Events

29

Event Localization with Instagram

- Taking a picture requires being on location
- There is a substantial overlap between Twitter users and Instagram users
  - Implication: Many shared hashtags/labels
- "Demultiplex" events on Twitter, identify relevant keywords/hashtags, search Instagram, find location!

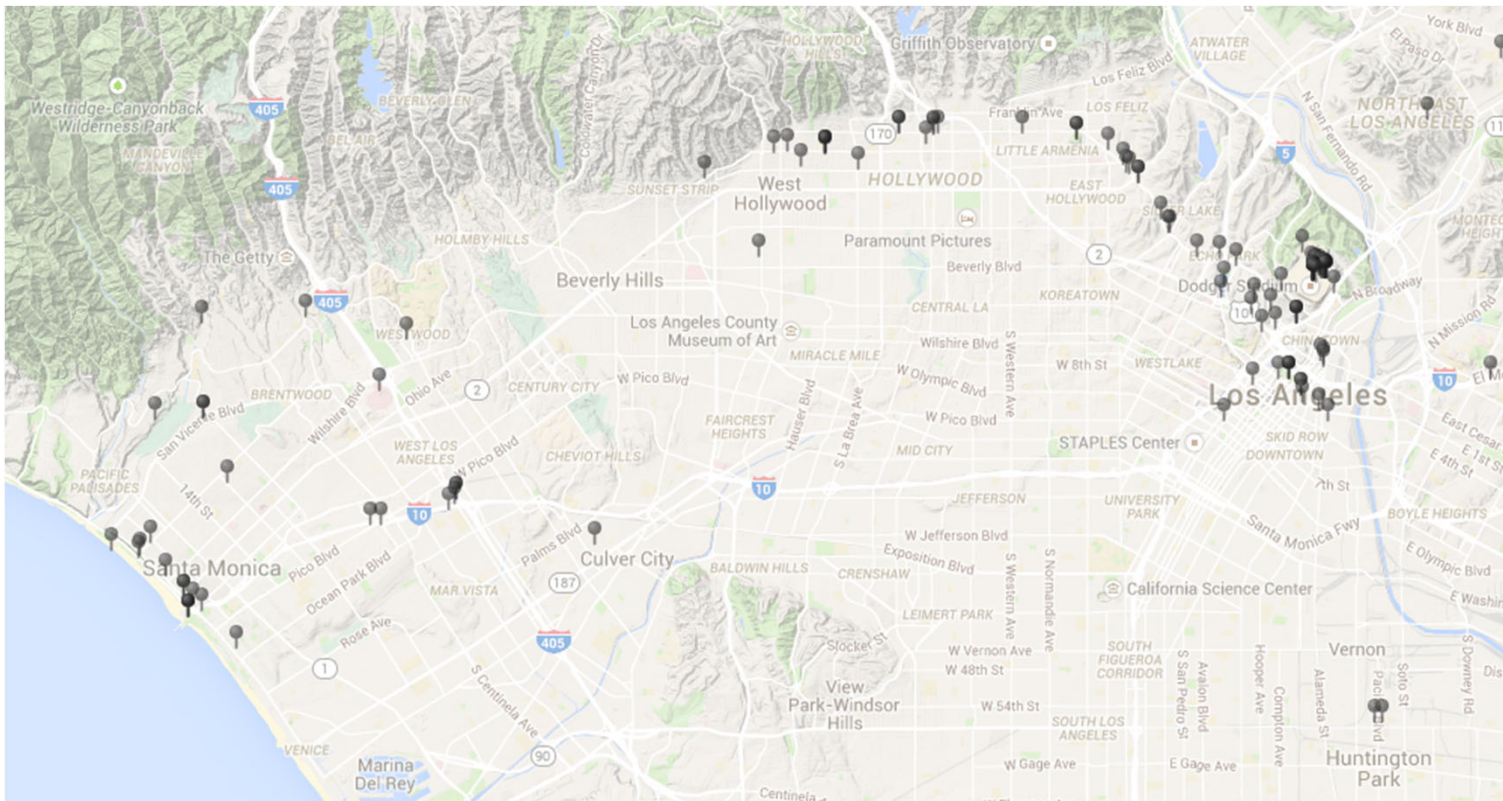# Instagram Localization

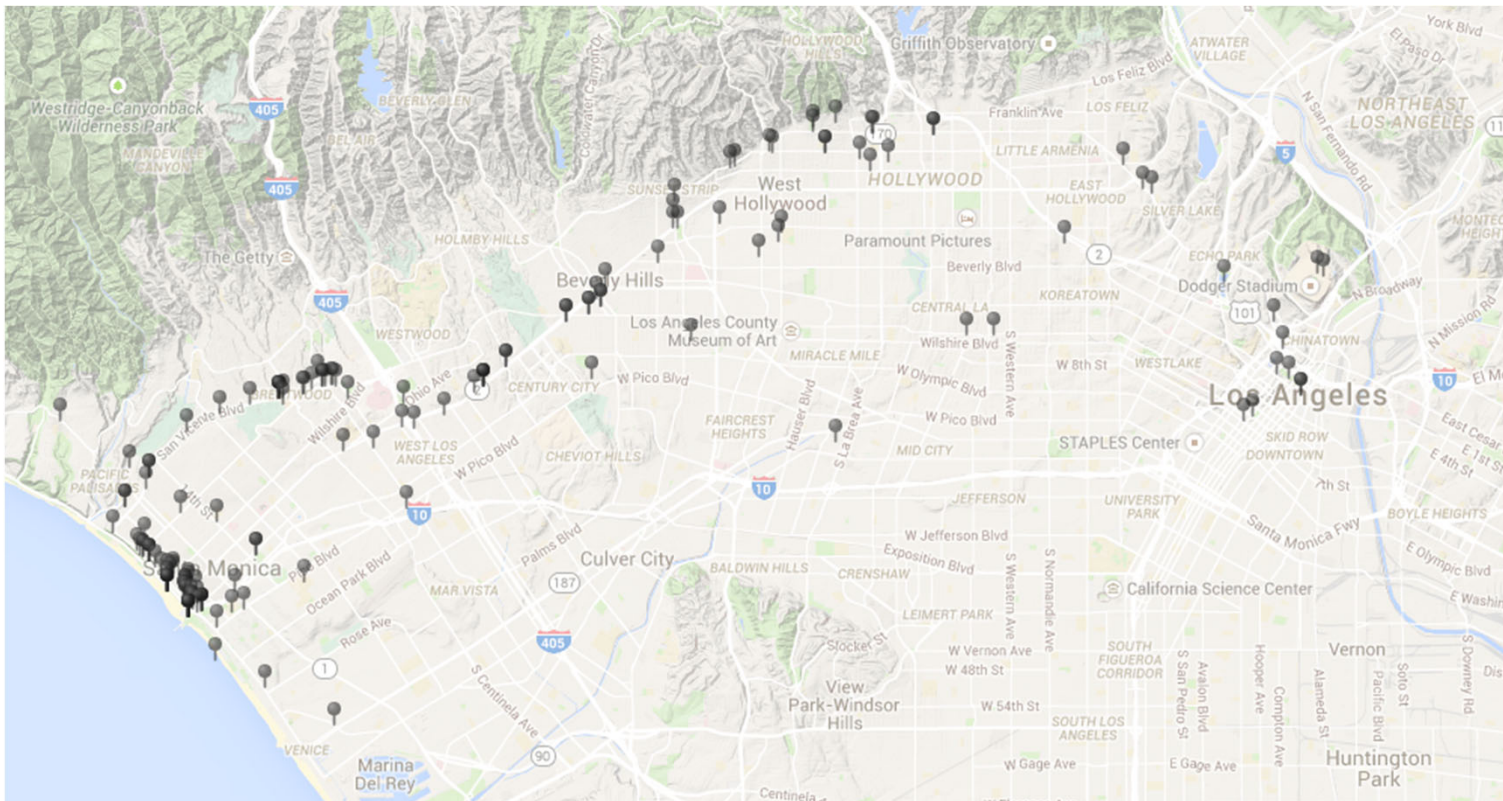- Tracking "LA Marathon"
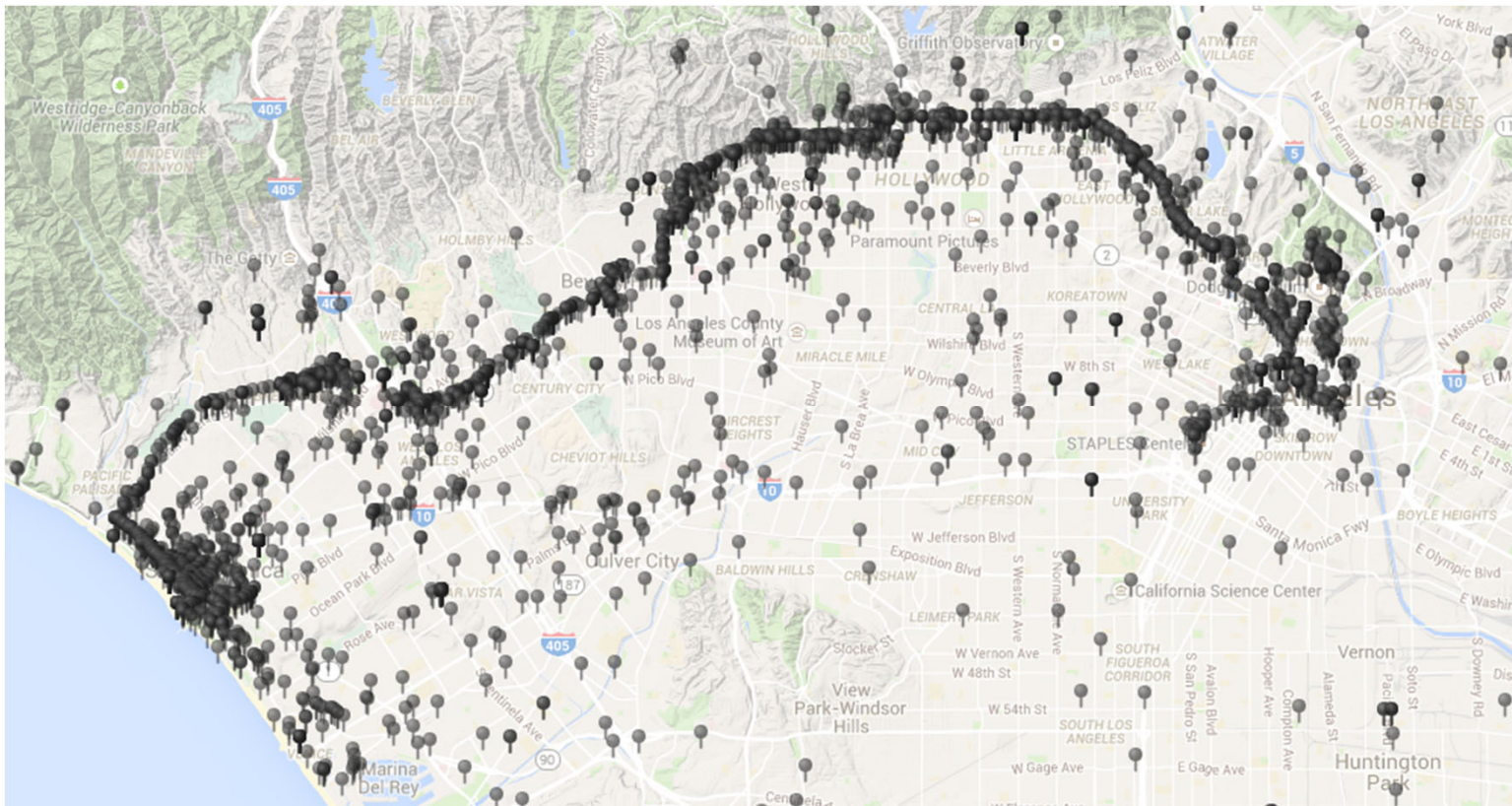
Instagram

- Tracking "LA Marathon"
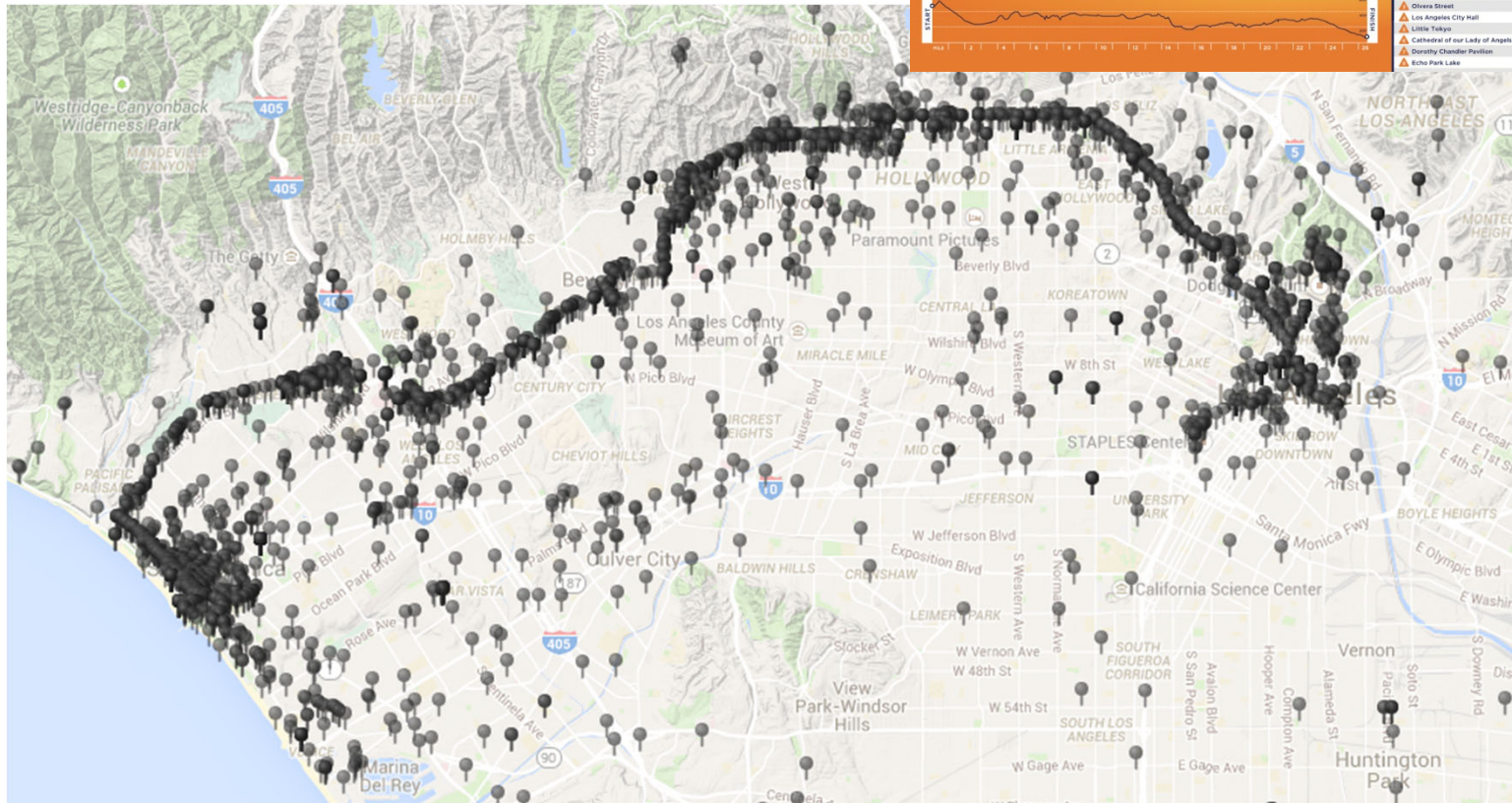
Instagram Tracking
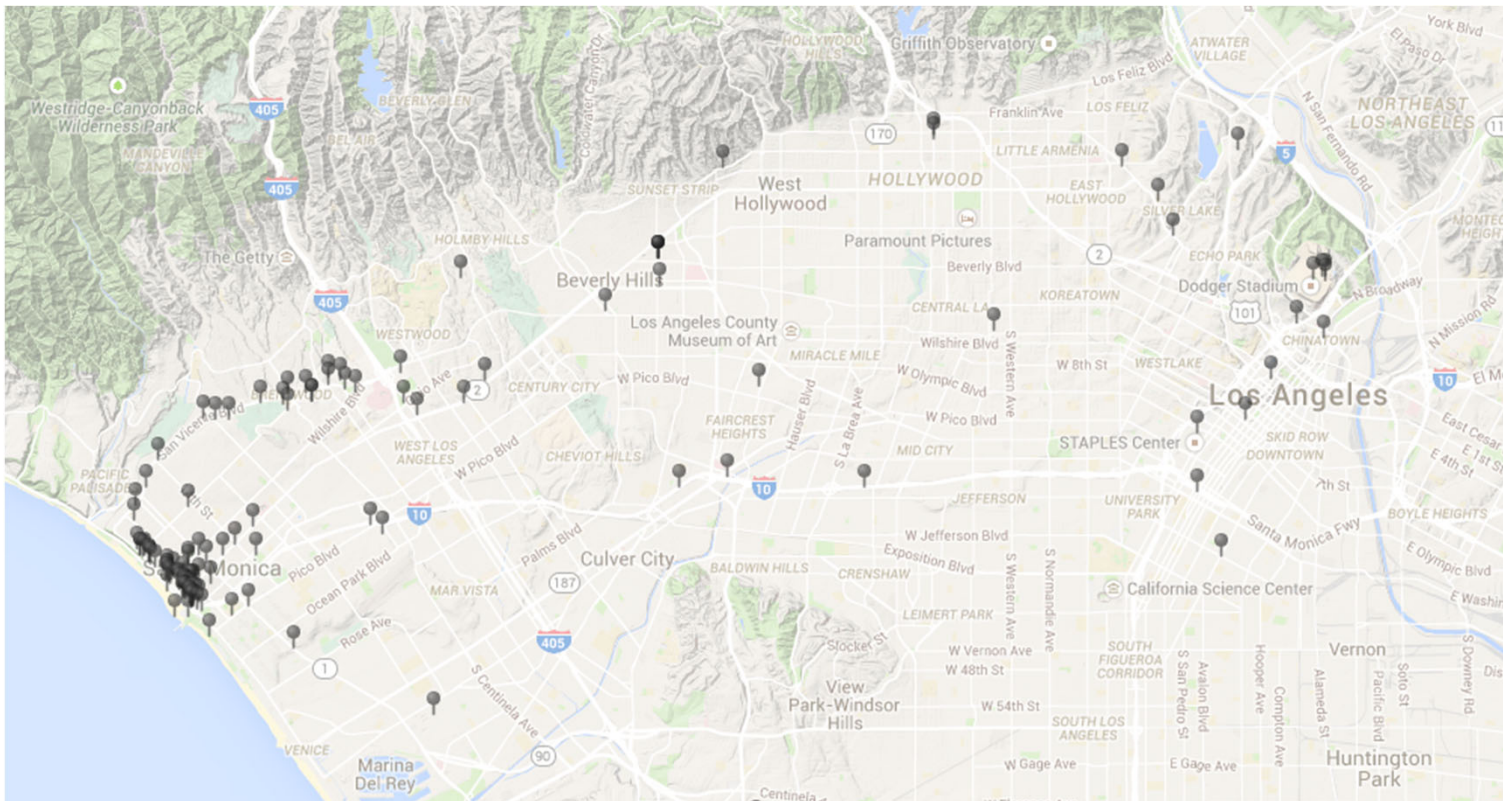
- Tracking "LA Marathon": Early Stage

Instagram Tracking

- Tracking "LA Marathon": Middle

Instagram Tracking

- Tracking "LA Marathon": Late Stage

Project Ideas (Set #5)
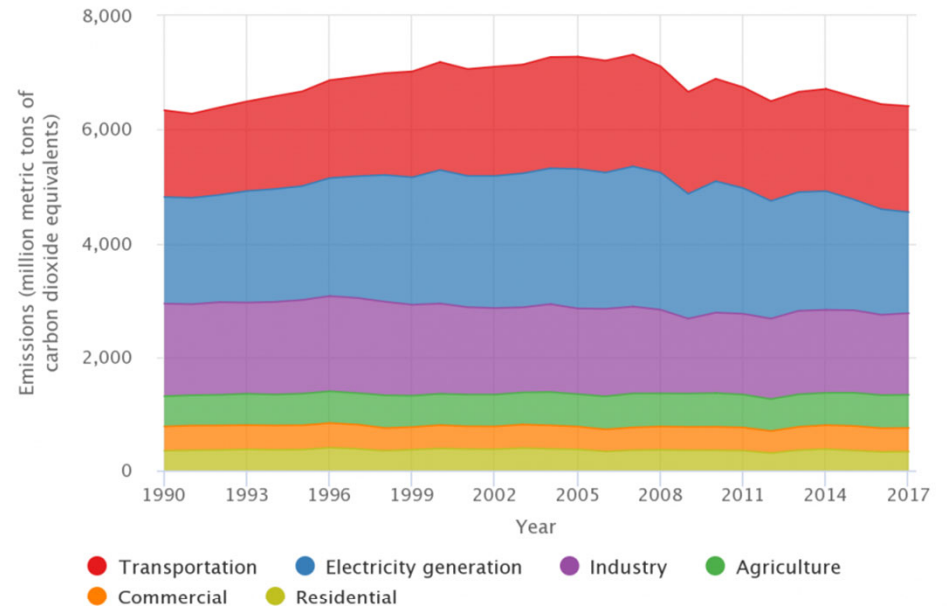## Transportation Energy Minimization

- US is 5% of world's population but 21% of GHG emissions

- The transportation sector is one of the largest sources of GHG emissions in the US
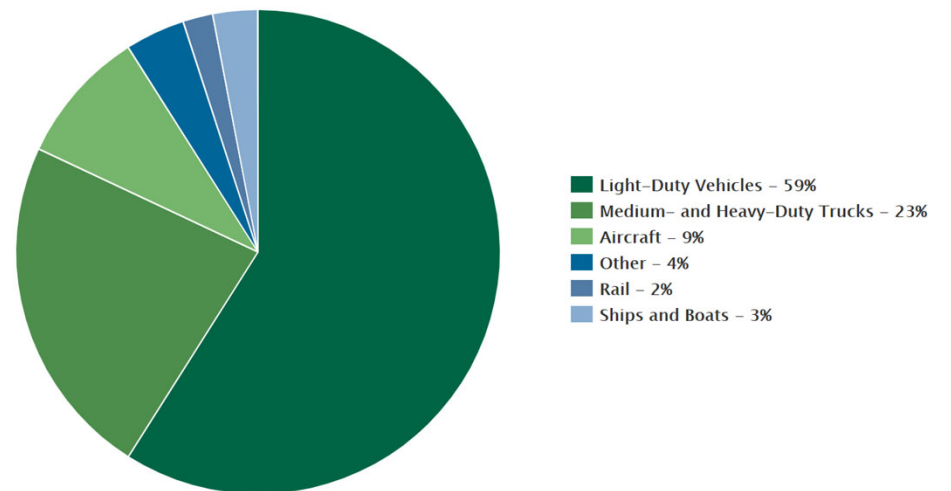
- How to reduce energy & emissions?

Trends

- Carbon emissions by sector

- Transportation emissions by vehicle type

U.S. Greenhouse Gas Emissions by Economic Sector, 1990–2017

Source: U.S. EPA's Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2017.
https://www.epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks

2017 U.S. Transportation Sector GHG Emissions by Source

- Light–Duty Vehicles – 59%
- Medium– and Heavy–Duty Trucks – 23%
- Aircraft – 9%
- Other – 4%
- Rail – 2%
- Ships and Boats – 3%

Class Project Idea

- Improve fuel-efficiency of transportation via "green" navigation
  - Measure fuel-efficiency of vehicles
  - Model fuel-consumption as a function of driver characteristics, road characteristics (average speed, speed variability, waiting time, slope, etc), and vehicle characteristics
  - Compute least-energy routes for a given vehicle and driver
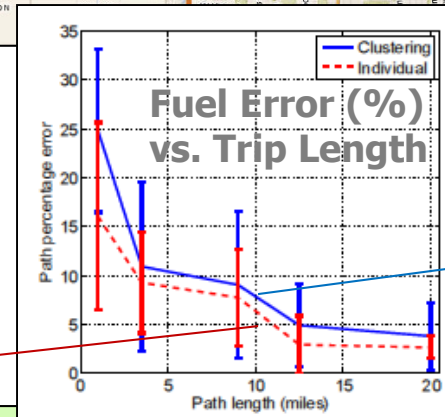
Green GPS

Saves 6% over shortest path and 13% over fastest path
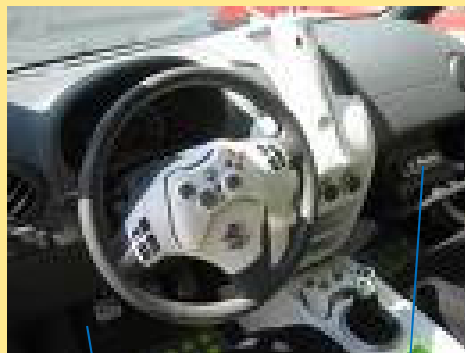
Shortest and fastest

Green GPS
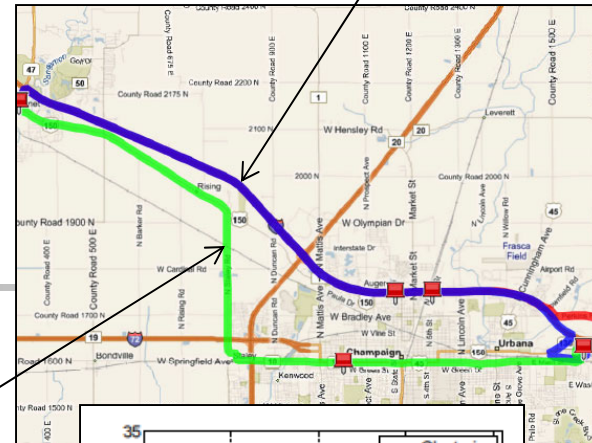
Most fuel-efficient

Subscribers

OBDII-WiFi Adaptor ($50) + GPS Phone

Subscribers:
Premium service
High savings

Fuel Error (%) vs. Trip Length

Open access:
Standard service
Average savings
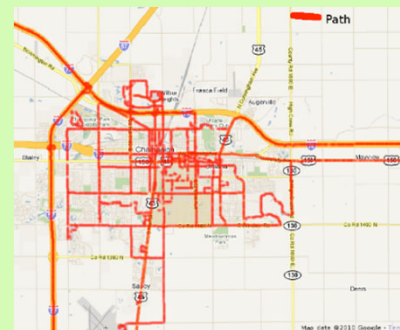
Fuel Data + Physical Models

$$F_{engine} = \frac{\Gamma(\omega) G g_k}{r}$$

F_{engine}

F_{air}

$$F_{air} = \frac{1}{2} c_d A \rho v^2$$

F_g

$F_{friction}$

$$F_{friction} = c_{rr} mg\cos(\theta)$$

$$F_g^g = mg\sin(\theta)$$

θ

$$F_{car} = F_{engine} - F_{friction} - F_{air} - F_g$$

Server

39

A Modeling Challenge
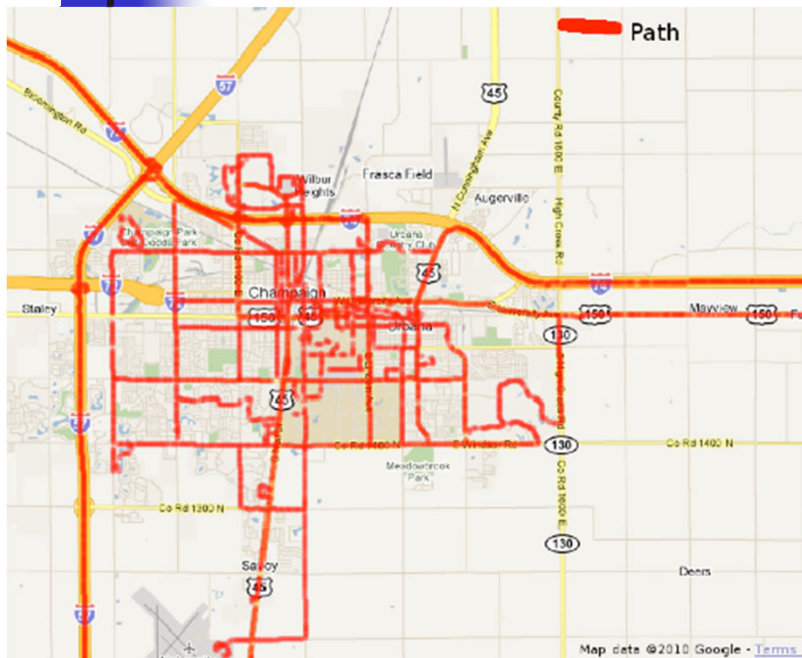
Fuel consumption of a few cars driven on a few roads by a few driver

Predict fuel consumption of any car on any road by any driver

Fuel Savings Evaluation

- How efficient is the fuel-efficient route?

Car Details	Landmarks	Route	Savings %
Honda Accord 2001	H1 to Mall	Shortest	31.4
	H1 to Gym	Shortest	19.7
Ford Taurus 2001	H2 to Restaurant	Shortest	26
Toyota Celica 2001	H2 to Work	Fastest	10.1
Nissan Sentra 2009	H3 to CUPHD	Fastest	8.4
Honda Civic 2002	Grad to Work	Fastest	18.7

Average fuel savings across 5 cars

Project Ideas (Set #6) Buildings and Smart Spaces

- On average, Americans spend about 90 percent or more of their time indoors

- Buildings accounted for 38.9% of total U.S. energy consumption in 2005

- Buildings accounted for 72% of total U.S. electricity consumption in 2006

- The average household spends at least $2,000 a year on energy

- Out of the total energy consumption in an average household, 50% goes to space heating, 27% to run appliances, 19% to heat water and 4% goes to air conditioning.

Related Class Projects

- Build smart services that improve residential energy consumption
  - Energy consumption modeling
  - Smart lighting
  - Smart door/window control

Testbed

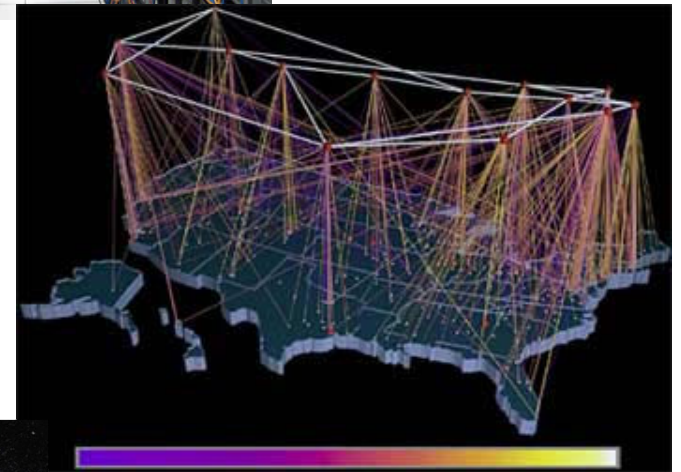|	Knob	Sensor
CPU	Frequency	Utilization, Frequency, Temperature
MEM		Utilization,
NIC		Received / Sent packets/bytes
PDU		Power consumption of each individual machine
CRAC	set point*	Input and outlet temperature, Set point

44

Failures in Complex Systems
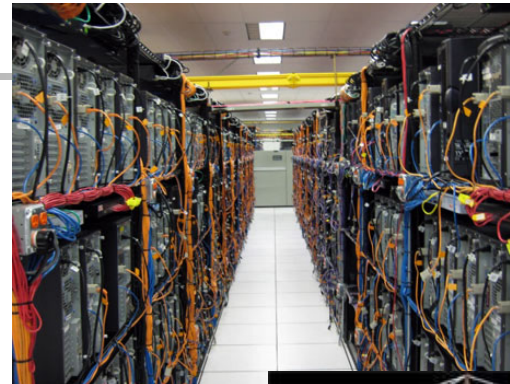
When systems fail, a common goal is:

Localize and fix the root cause!

Failures in Complex Systems
Another Thought

Individual software components are easy to "debug"

- Therefore, they are typically built reliably

Systems do not fail because of "bugs" localized to single components

- Systems fail because of unexpected interactions between many *individually well-behaved* components
- No single component is to blame
- No predicate over current state explains failure
- Unexpected *sequences of events* lead to problems