NETWORK SCIENCE CTA

Social Network Signal Processing for Real-time Situation Understanding

Tarek Abdelzaher

University of Illinois at Urbana Champaign

ILLINOIS UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

ARL

IBM Research

NSF

An Internet Minute
Today: Growing Social Network *Data* Size!
→ Novel Social Data Management Services!

60 seconds on the Internet

A Deluge of Information!!!

Goal

Tools that empower the analyst to quickly understand ongoing events

Goal: Supporting (Social) Data-to-Decision Pipelines

Events

Traffic jams

Disasters

Civil unrest

News and Public Sources
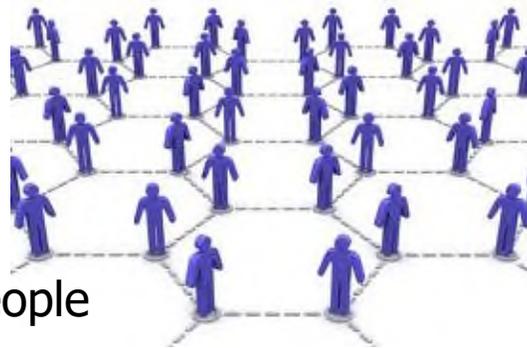
People

Sensors

Autonomic Service Stack

Data

Decision Support
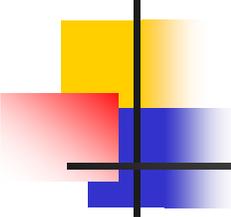
Building Blocks

- Physical event detection, localization and tracking from social media feeds (Twitter + Instagram)

- Fact-finding (separating true and false media posts)

- Polarization analysis (separating different views in a conflict, isolating the community adopting each view, and distilling more neutral narratives)

- Anomaly detection and explanation
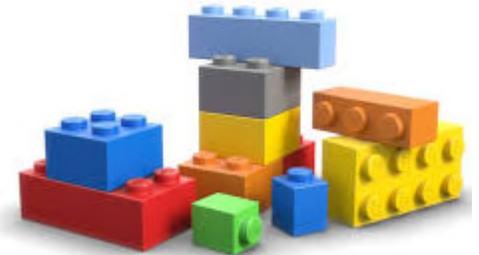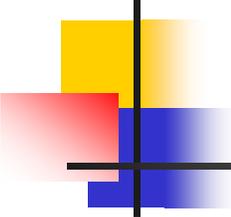
Putting Components Together

Assemble any way you like.

Example:

Putting Components Together

Assemble any way you like.

Example:

Add annotations to each tweet
(tweet polarity, source polarity)

Add annotations to each tweet
(belief: true/false, source reliability)

Short
microblogs
(tweets)

Event Detection, Localization, Tracking

Fact-finding

Polarization Analysis

Anomaly Explanation

Add annotations to each tweet
(event name, location, etc)
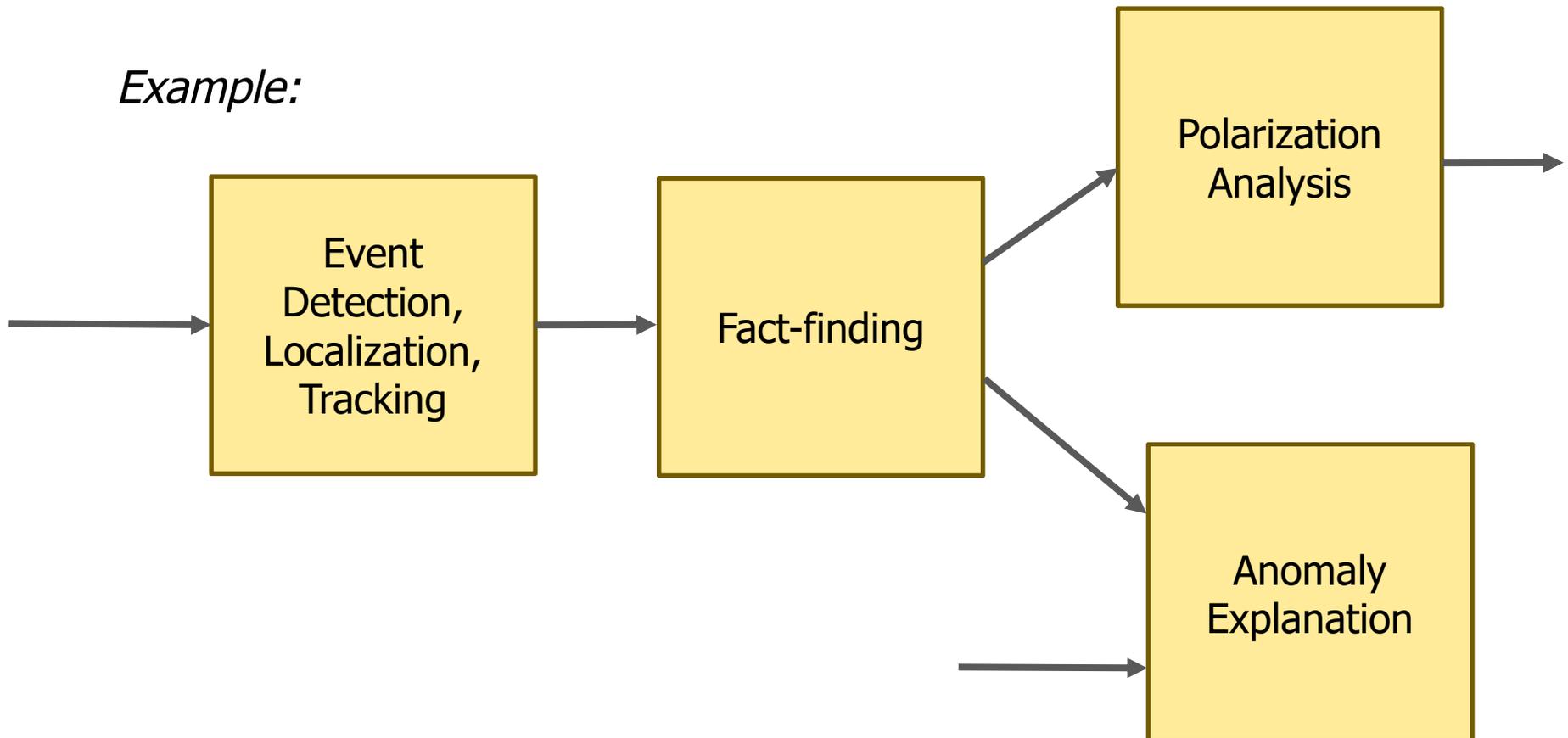
Anomaly (cue)
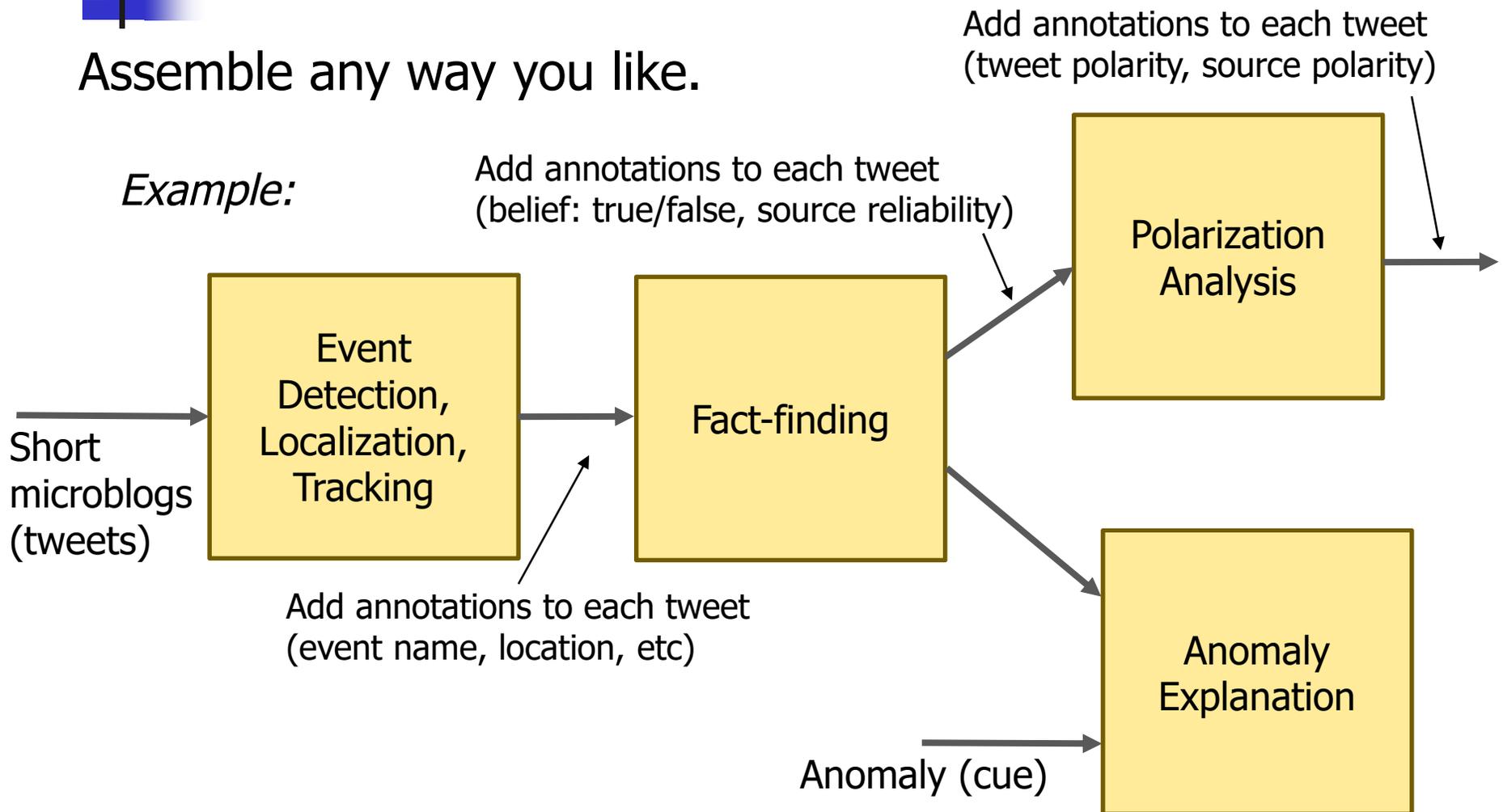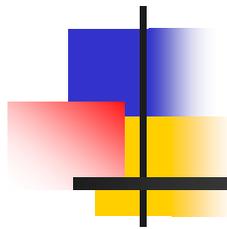
Capability #1

Event Detection, Localization, Tracking
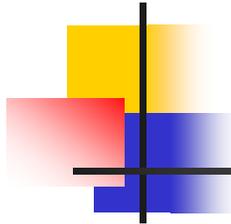
Insight

- An analogy between social media and physical media: both propagate signals that offer information on physical events

→ Build tools for "social sensing"

The Social Signal: An Analogy

Physical target

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

Received signature (energy in
multiple signal frequency bands)

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)
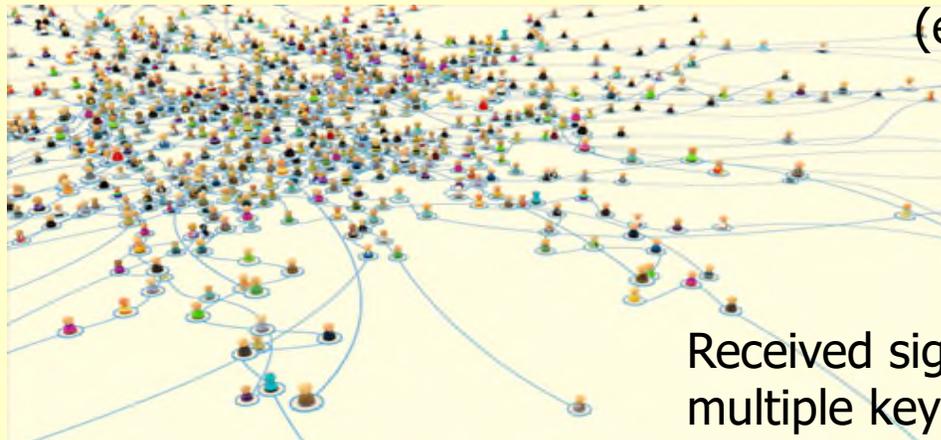
Received signature (energy in
multiple keyword frequency bands)
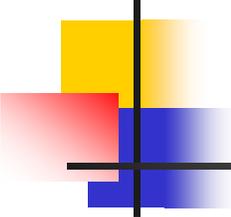
Demultiplexing
(New Protests, 6/13/2017)

- "Protests" 6/13/2017
  - For Trump's birthday, an army of protesters will be delivering the president a plane ticket to Russia… https://t.co/cB3GS8xHvj
  - Women Dressed As Handmaids Descend On Ohio Statehouse To Protest Anti-Abortion Law https://t.co/guuLo8DnKk
  - किसान आन्दोलन में किसानो में ही पड़ी फूट.. एक टूटे धड़े ने कहा कि - "ये आन्दोलन नहीं राजनीति" https://t.co/pKDtzVKU7A
  - Trump Supporters Plan Protest at CNN Headquarters in Atlanta to Decry Fake News - Breitbart https://t.co/SA4IbL6jvM via @BreitbartNews
  - FCT Indigenes Protest Unclad In Abuja Over Marginalization (Photos) >> https://t.co/vQ3yX6He7B
  - Parishioners protest priest reassignments across Catholic Diocese of Memphis. https://t.co/JC11n80trB
  - Thousands of people in more than 100 cities in Russia turned out to protest official corruption https://t.co/gOJKZSDBQh

Instagram

TVERSKOY DISTRICT, MOSCOW, RUSSIA

#ОННАМНЕДИМОН #IG_PHOTOOFTHEDAY #PEOPLE #MOSCOW #VSCO_GOOD #ИНСТАГРАМНЕДЕЛИ #РОССИЯ #F4F #FOLLOW #RUSSIA

Event Timeline
Moscow, Protest

06-12-2017 20:16:24

• #Navalny says officials prevented companies from renting him audio/visual equip for protest rally so he's moving protest to Tverskaya st.

• 1/1 Activist Navalny said protest meeting in Moscow today at 14.00 will be moved from outside the centre tot the main street, Tverskaya >

• #Navalny protest march on Tverskaya - will be hard to know how many turn out since they could be mixed in w/others celebrating Russia Day.

• Tomorrow on Tverskaya St, where Navalny has moved his protest, Moscow is planning a historical reenactment festival. 150k people expected.

• Navalny just moved his protest tomorrow to the Moscow's main Tverskaya street, citing govt efforts to block preparation of the rally.

06-13-2017 05:17:16

Liveuamap: .navalny was detained in Moscow, prior to mass-protest on Tverskaya street https://t.co/mhfUiFuf6t https://t.co/rcIzM6zMHE vi...

Tverskaya street in downtown #Moscow now where @Navalny plans to hold anti-corruption #protest today: https://t.co/KJLvt36o4v

• Tverskaya street in Moscow where #Navalny protest meant to take place absolutely blockaded by viking ship and group... https://t.co/nFQNbpwFVS

• Police in Moscow say they'll use face-recognition technology on street where Alexsey Navalny plans his main anti-corruption protest...

06-13-2017 09:17:25

• Police in Moscow say they'll use face-recognition technology on street where Alexsey Navalny plans his main anti-corruption protest...

• Police urge people to leave site of unsanctioned protest on Tverskaya Street in downtown Moscow... https://t.co/ItxDCy0hyp

#Protest in central #Moscow right now: Tverskaya street from above #RussiaProtests https://t.co/m5DipjUGrB

• Police urge people to leave site of unsanctioned protest on Tverskaya Street in downtown Moscow... https://t.co/K4ciARfypK

#Protest in central #Moscow right now: Tverskaya street from above #RussiaProtests https://t.co/m5DipjUGrB

06-13-2017 10:17:29

• On a day of mass arrests for "unsanctioned protest," girls make signs from Russia's constitution. (Photos via... https://t.co/8uzjTb18EE

Mass arrests reported as thousands across Russia protest corruption, confront Kremlin https://t.co/zO332tvHoB https://t.co/OOoamphsMo

06-13-2017 18:19:06

• Revealing! Western #MSM & pundits reporting non-stop on #Navalny 'protest' but no coverage of big celebrations... https://t.co/HoHFvkIQa4

• Globalist darling #Navalny organized protest 1) to ruin Russia Day festivities 2) to create photo-ops for Western p... https://t.co/bBrVlZiPlZ

• Moscow corrs for Western pubs conveniently leaving out fact that Navalny moved his protest to the site of June 12 holiday celebrations.

• Navalny protests: Western media don't care about facts when cheering their Russian opposition hero https://t.co/j6HUv3LEmu

• Revealing! Western #MSM & pundits reporting non-stop on #Navalny 'protest' but no coverage of big celebrations... https://t.co/HoHFvkIQa4

Events and Signal Processing: The Lexical Frequency Domain

- *"Portability"* → Language Independence
- *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?

Events and Signals:
A Data Association Problem

Sparsely populated feature space

Easy to associate data with events

Feature axis

Densely populated feature space

Hard to associate data with events
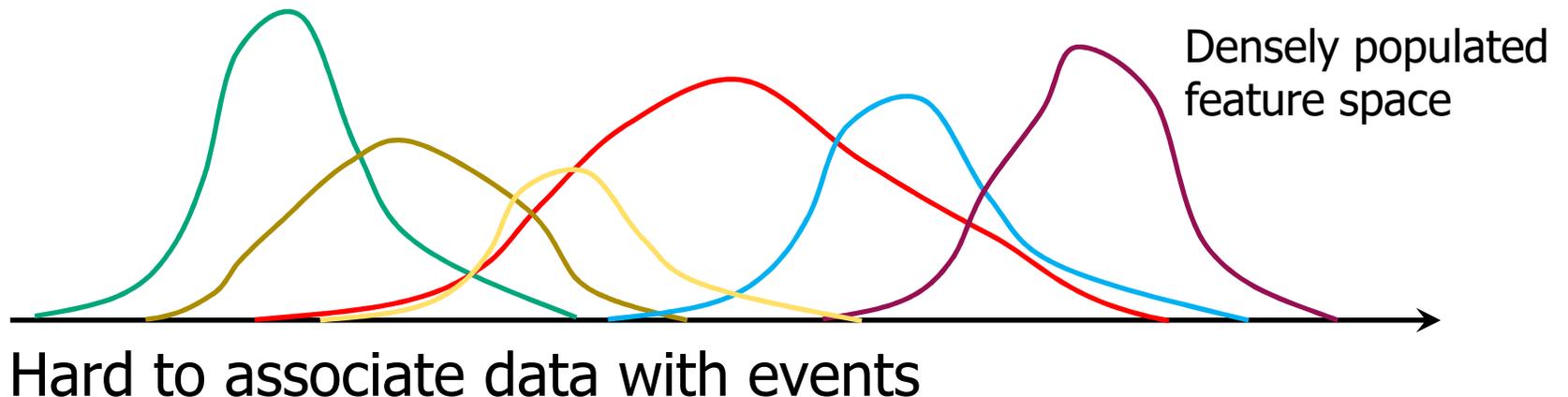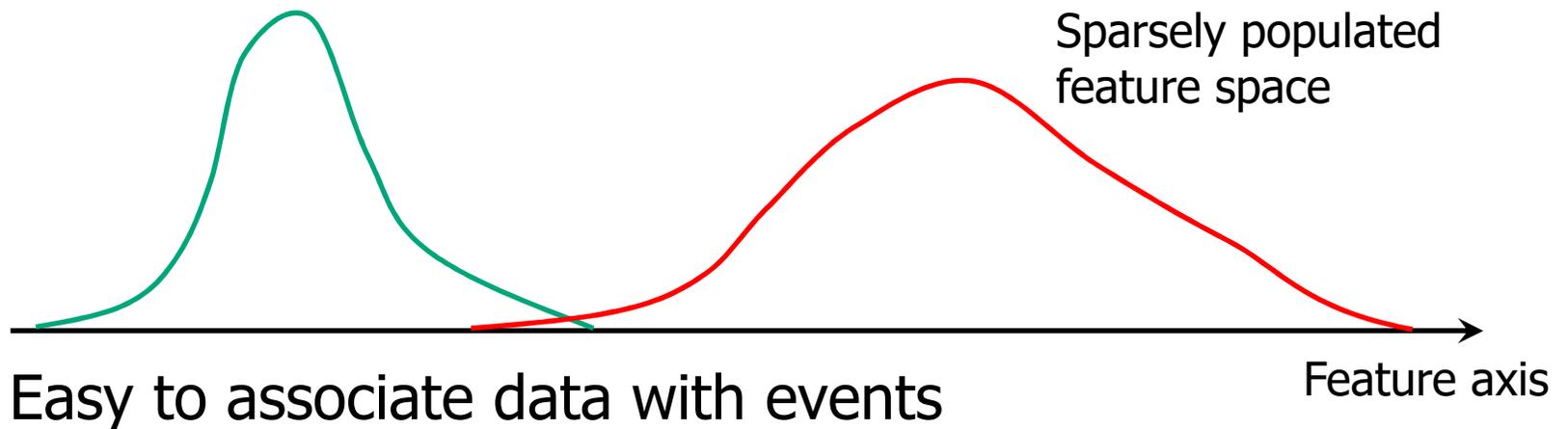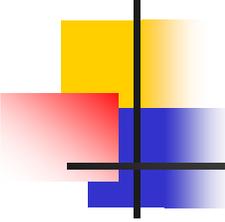
A Signal Sparsity Observation

- Most languages have about 10,000 frequent words.
- Number of "events" in a Twitter data trace may be in the 100s or 1000s
- The space of 1-word event signatures is not "sparse enough"
- Consider 2-word event signatures
  - There are at least 100,000,000 possible signatures
- The space of 2-word signatures is vastly sparse:
  - Different events → Different signatures (assuming independent keywords)
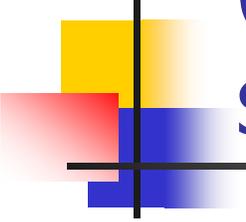
Event Detection, Consolidation, and Tracking:
## Signal Processing Questions

- 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)

Tweet cluster

Time
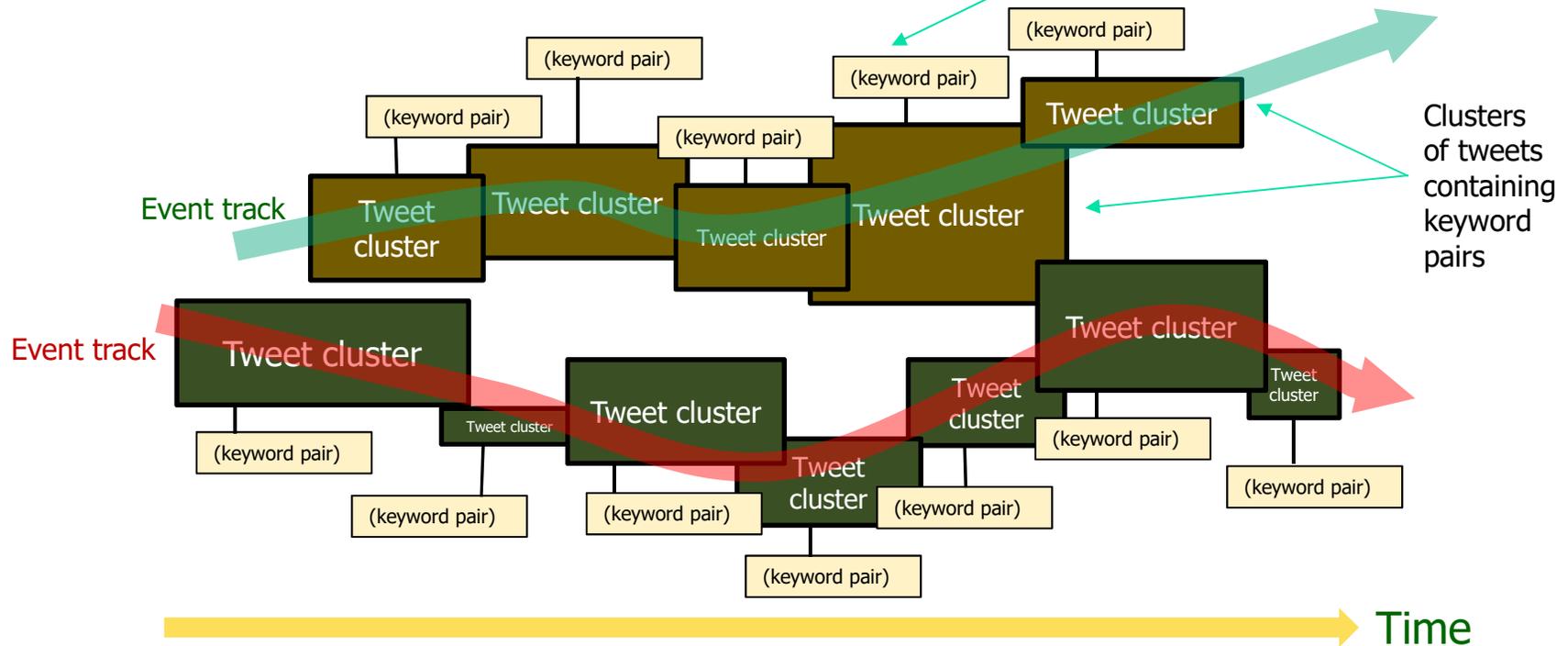
Distance Metrics
(For Merging Event Data Clusters)

- Cosine similarity
- Term frequency difference
- Jaccard distance
- KL divergence

Lexical frequency domain signal

Distance?

Lexical frequency domain signal

(keyword pair)

(keyword pair)

(keyword pair)

(keyword pair)

Tweet cluster

Event track

Tweet cluster

Tweet cluster

Tweet cluster

Tweet cluster

Clusters of tweets containing keyword pairs

Distance Metrics
## (For Merging Event Data Clusters)

Challenge: Adaptive data consolidation

- Cosine similarity
- Term frequency difference
- Jaccard distance
- KL divergence

Distance?
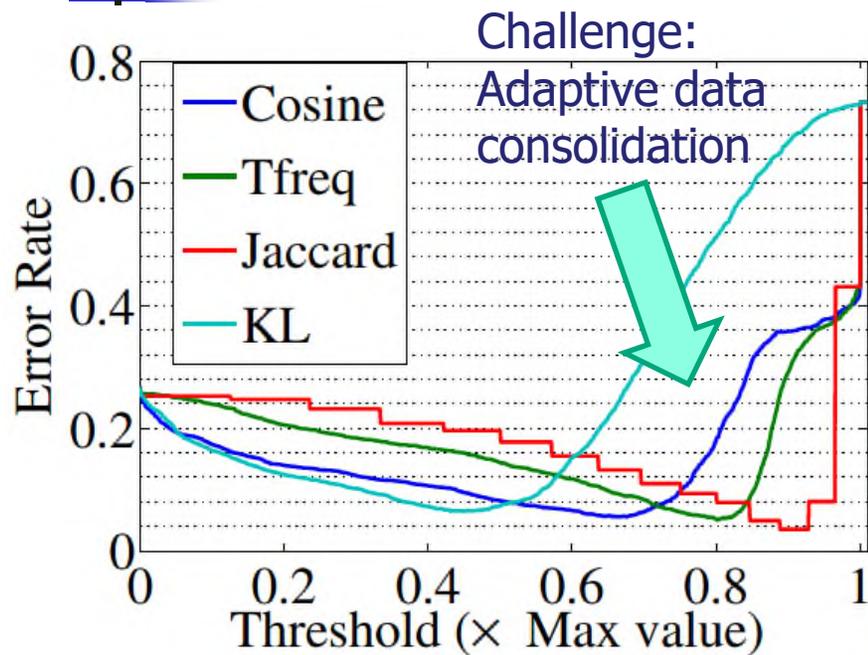
Clusters of tweets containing keyword pairs

Event track

(keyword pair)

Tweet cluster

Event Tracks

Recognizing Distinct Event Tracks

- Contribution: Efficient algorithms that "demultiplex" Twitter feed into sub-streams associated with different events in a class (e.g., different concurrent shootings or different concurrent protests)

Protest Name	Tweets
Bangladesh protests	Religion Bangladesh Braces for Protests After Islamist's Execution: senior official of the largest Islamis... http://t.co/NTBmIWsTme
	World News: Bangladesh braced for protests after Islamist leader's execution: Bangladeshi security personnel s... http://t.co/UPKiaFFHtW
	Bangladesh braces for protests after Jamaat leaders execution: Bangladesh braced for protests and fresh violen... http://t.co/3dcPFqKAQE
Brazil protests	Protests across Brazil seek ouster of president http://t.co/YmXZnsxbAQ
	FollowMePlease Brazil braces for nationwide protests, as groups seeking impeachment of presiden... http://t.co/5T150D0zlL BrinaldyHere
	Fresh anti-government protests in Brazil: Brazil on Sunday braced for more huge demonstrations against governm... http://t.co/8nJmX56MUm
Turkey protests	DTN Turkey: Turkey protests to Pope Francis after he brands Armenian killings 'genocide': Pontiff's run-in wit... http://t.co/dKx5iCwP9w
	Pope refers to Armenian genocide; Turkey protests. 24 April is 100th anniversary of start of the Armenian genocide http://t.co/ZFOCZnC411
	Telegraph: Turkey protests to Pope Francis after he brands Armenian killings 'genocide' http://t.co/YnJew4foN1 http://t.co/O6d5oYwG2p

Accuracy

- Accuracy in data association (four Twitter streams)
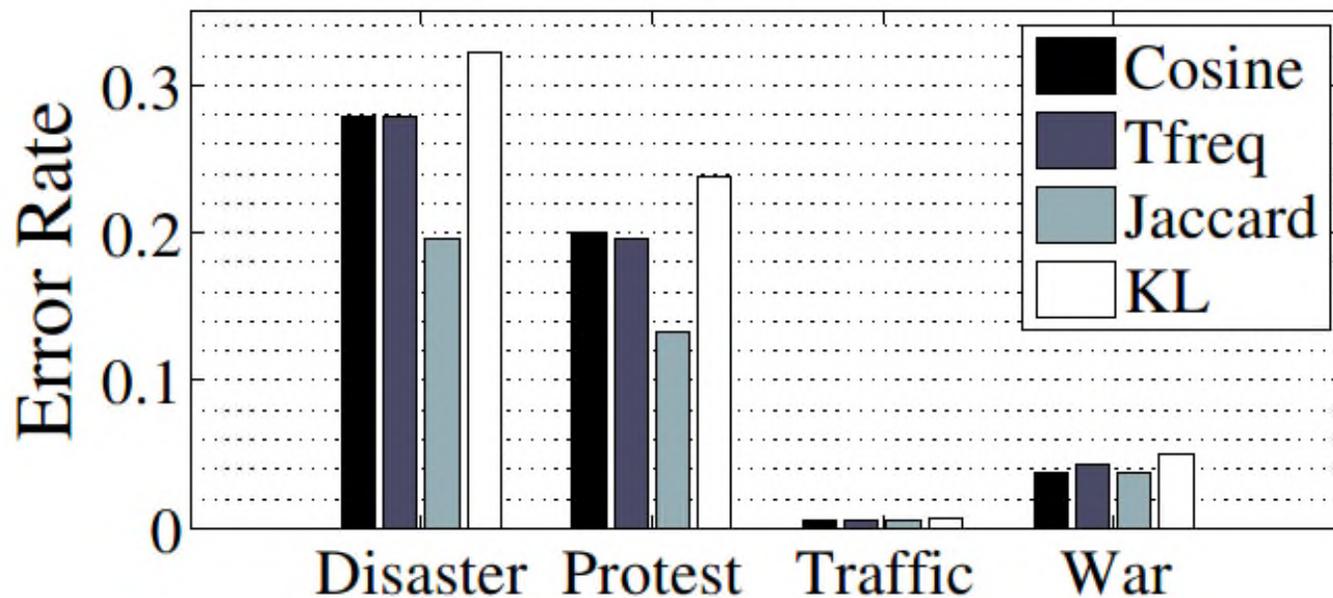
	Start Time (UTC)	End Time (UTC)	# Tweets Collected	# Tweets after Filtered	# Events Detected	Locations Identified
Disaster	Apr 19 19:41:08 2015	May 02 16:58:47 2015	319, 181	143, 448	121	104 countries or cities
Protest	Apr 01 06:03:51 2015	May 04 06:50:40 2015	329, 336	137, 432	178	129 countries or cities
Traffic	Apr 19 12:12:50 2015	May 02 17:05:59 2015	480, 717	140, 527	392	307 countries or cities
War	Apr 19 19:45:17 2015	May 02 17:03:21 2015	391, 711	196, 704	103	95 countries or cities
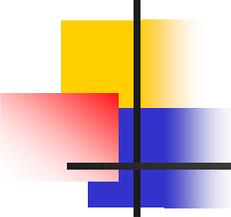
Precision Comparison

EVENT DETECTION PRECISION COMPARISON

| Algorithm	Traffic	Disaster	Protest	War
StoryLine	**72.55%**	**76.92%**	80.95%	**88.24%**
ET	57.14%	36.36%	**86.36%**	61.90%
TopicModel	55.10%	60.87%	65.22%	69.57%

Demultiplexing Quality

- A distribution (pie chart) of the number of different events per "demultiplexed" stream.
 - Ideally, must be 1, 100% of the time.

Chicago traffic dataset
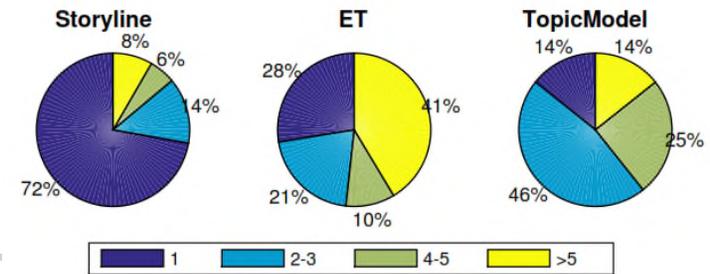
Storyline — ET — TopicModel

Legend: 1 | 2-3 | 4-5 | >5

Disaster dataset

Storyline — ET — TopicModel

Protest dataset

Storyline — ET — TopicModel

War dataset

Storyline — ET — TopicModel
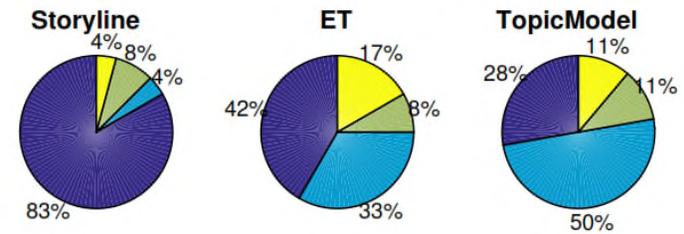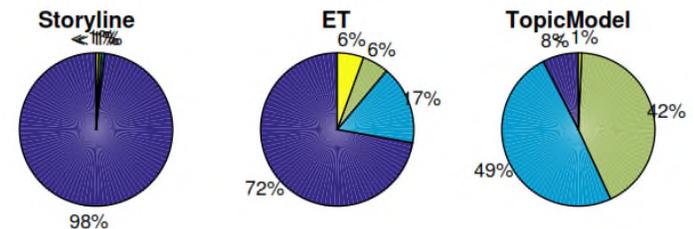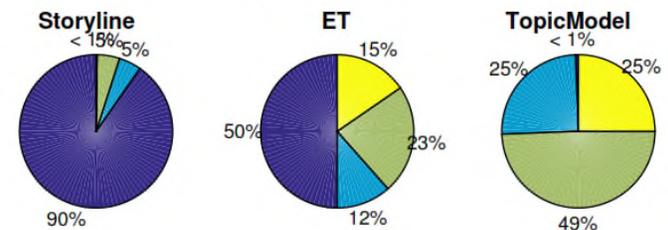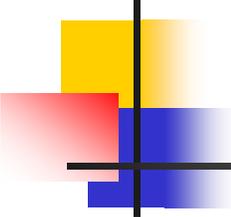
Contribution to the State of the Art in Event Detection

- Many techniques focus on detection, not demultiplexing (can detect a protest, but can't differentiate well text describing two different protests)
- Some do demultiplexing based on location information (e.g., geo-tags)
  - Most tweets (about 98%) are not geo-tagged.
- Others extract the location information from text (need training to understand location words)
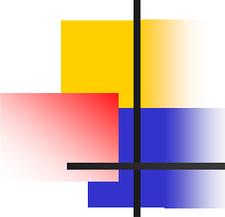
A Multilingual Approach

- Το γαλλικό πλήρωμα αναγκάστηκε σε προσγείωση στην Αθήνα, στην πορεία τους προς τη Μόσχα

- フランスの乗組員は、モスクワへ向かう途中、アテネの緊急着陸を余儀なくされた

- واضطر الطاقم الفرنسية إلى الهبوط اضطراريا في أثينا في طريقها إلى موسكو

- Iफ्रेंच चालक दल मास्को के लिए अपने रास्ते पर एथेंस में एक आपात लैंडिंग करने के लिए मजबूर किया गया
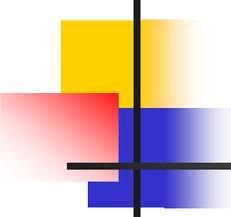
Contribution to the State of the Art in Event Detection

- Our demultiplexing technique
  - No location info (e.g., geotags)
  - No language understanding
 - No prior training

  Suitable for IoT applications
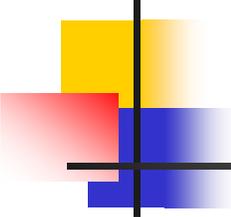
On False Positives

- Some "tweet buckets" created by the demultiplexing algorithm contain unrelated tweets that coincidentally share the same (common) keyword pairs.

- How to automatically detect and eliminate such theme-less buckets?

Removing False Positives

Beta Expectation (Empirical) CDF
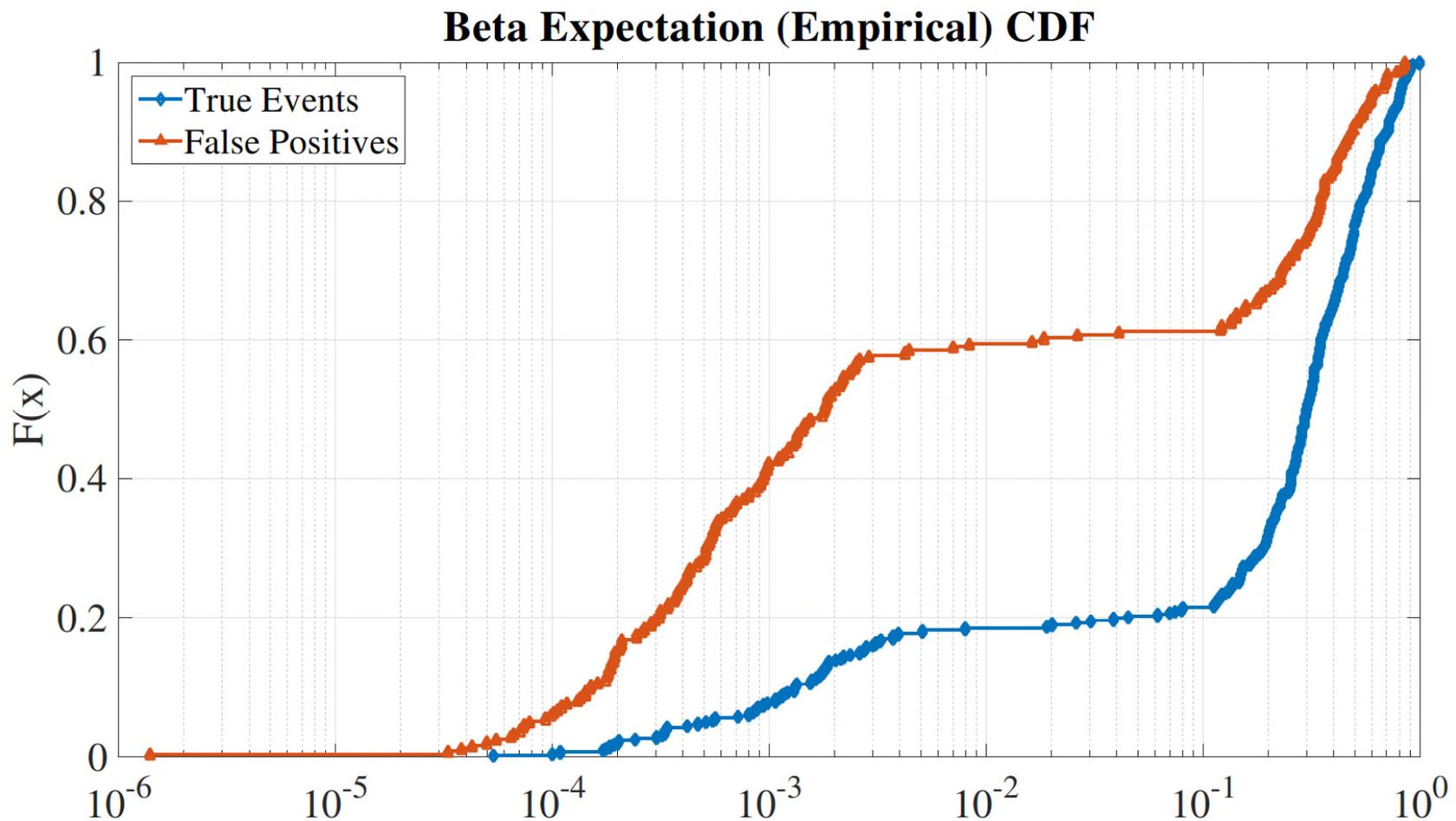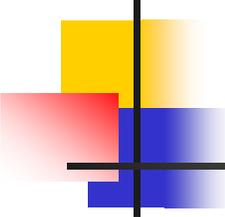
Detection, Localization and Tracking with Instagram

- *Motivation:* Why use Instagram for Event Localization?
  - Unlike text that can be written from anywhere, pictures of an event are generally taken at the event location.
  - Almost 25% of Instagram images are geotagged (but only 2% of tweets).
- *Analogy:* An Instagram picture = "binary sensor" signifying event "detection"
- *Method:* Leverage prior sensor network literature on localization and tracking with binary sensors

Feasibility of Instagram-based Localization

- Example: Tracking "LA Marathon" (2015)
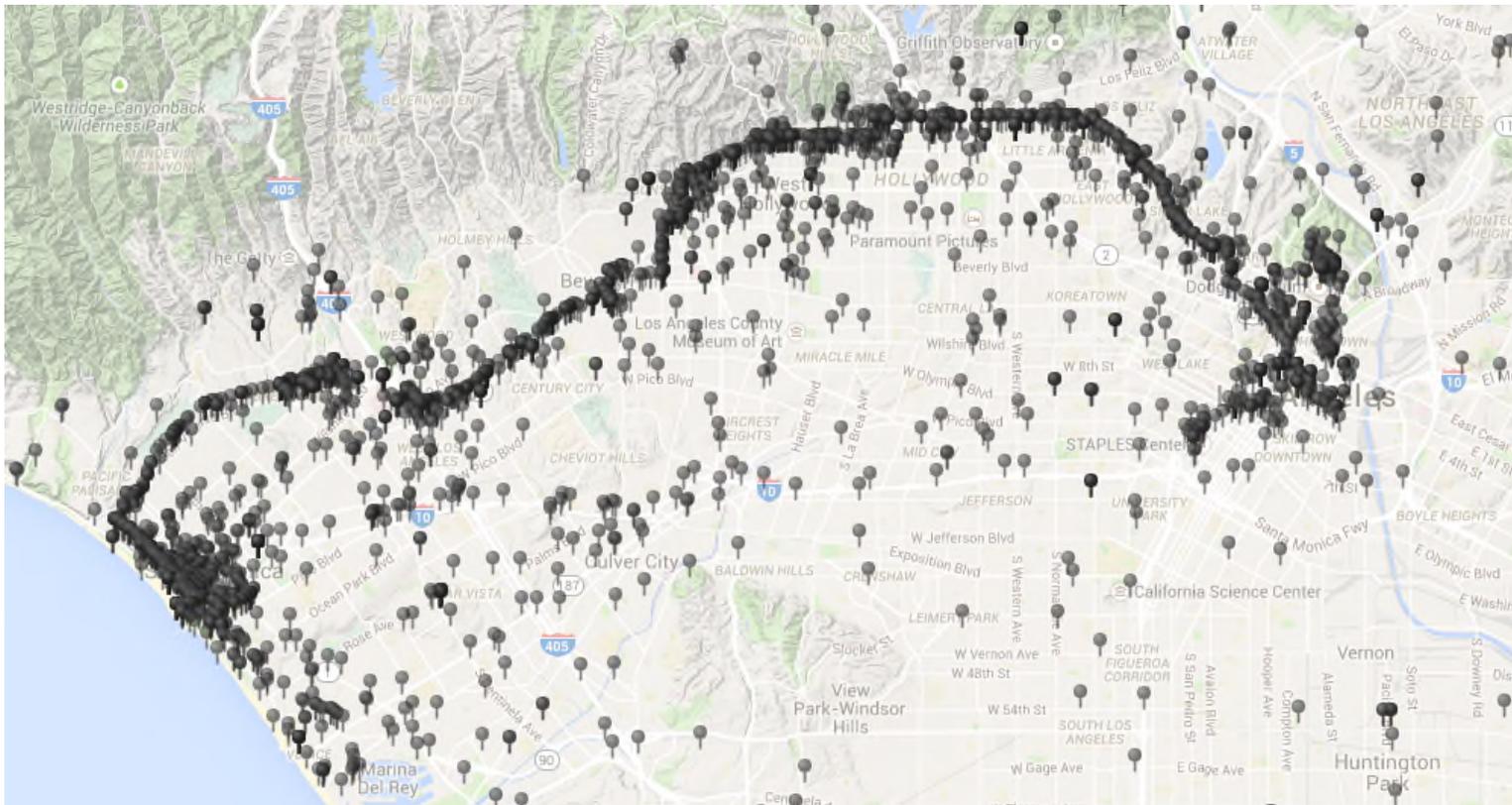
Instagram

- Tracking "LA Marathon"
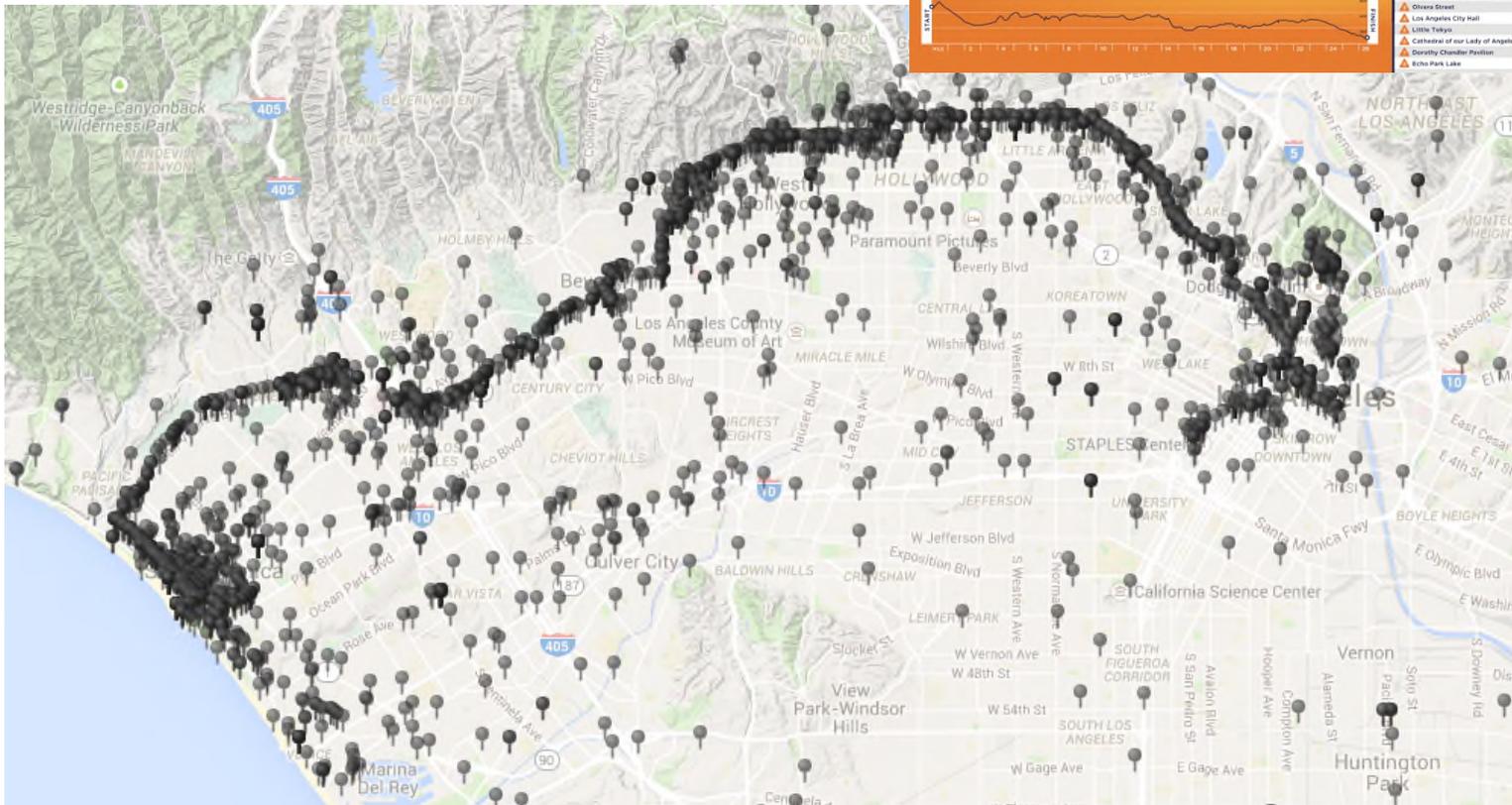
Feasibility of Instagram-based Tracking

- **Tracking "LA Marathon": Early Stage**
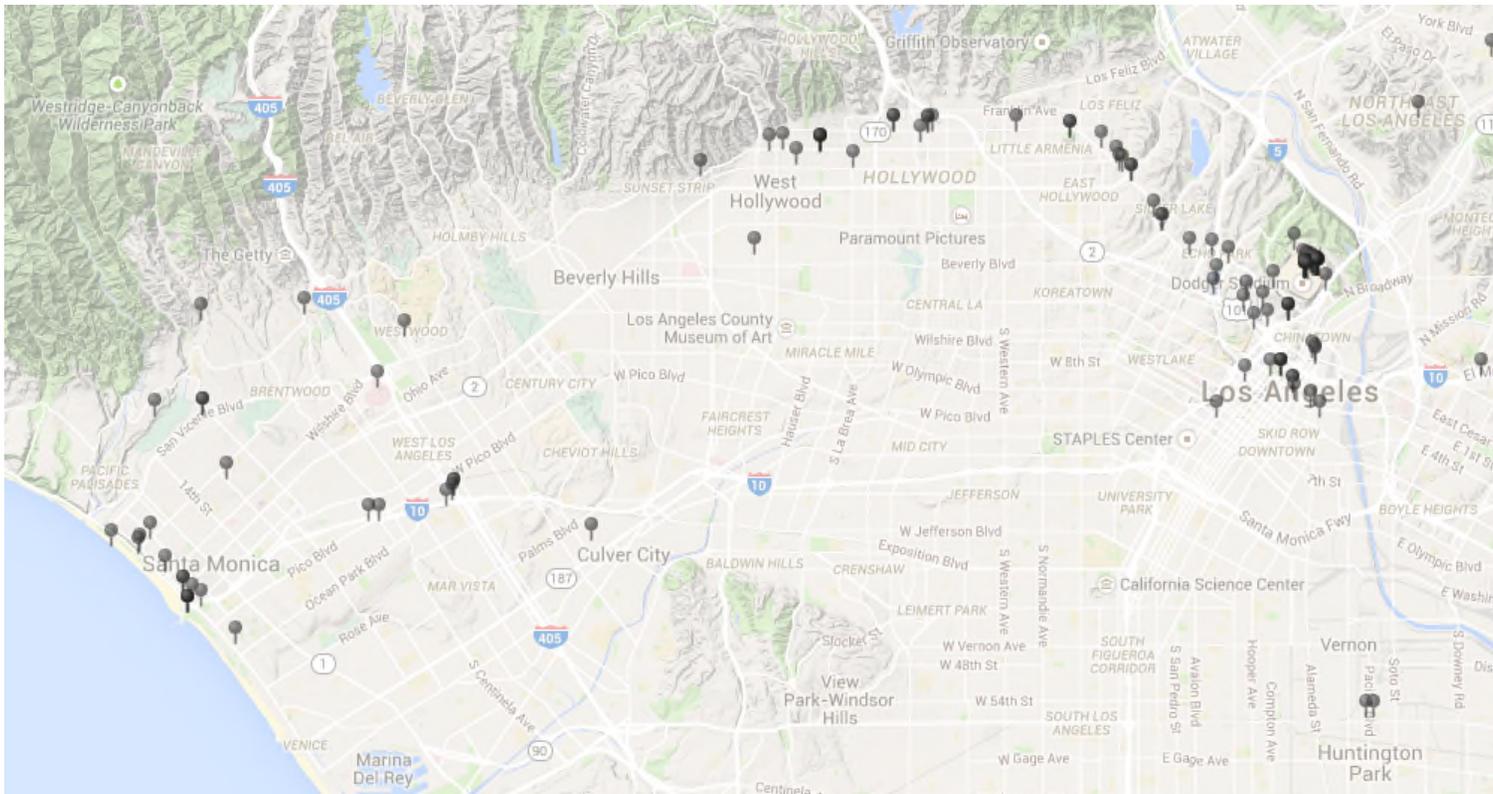
Feasibility of Instagram-based Tracking

- Tracking "LA Marathon": Middle
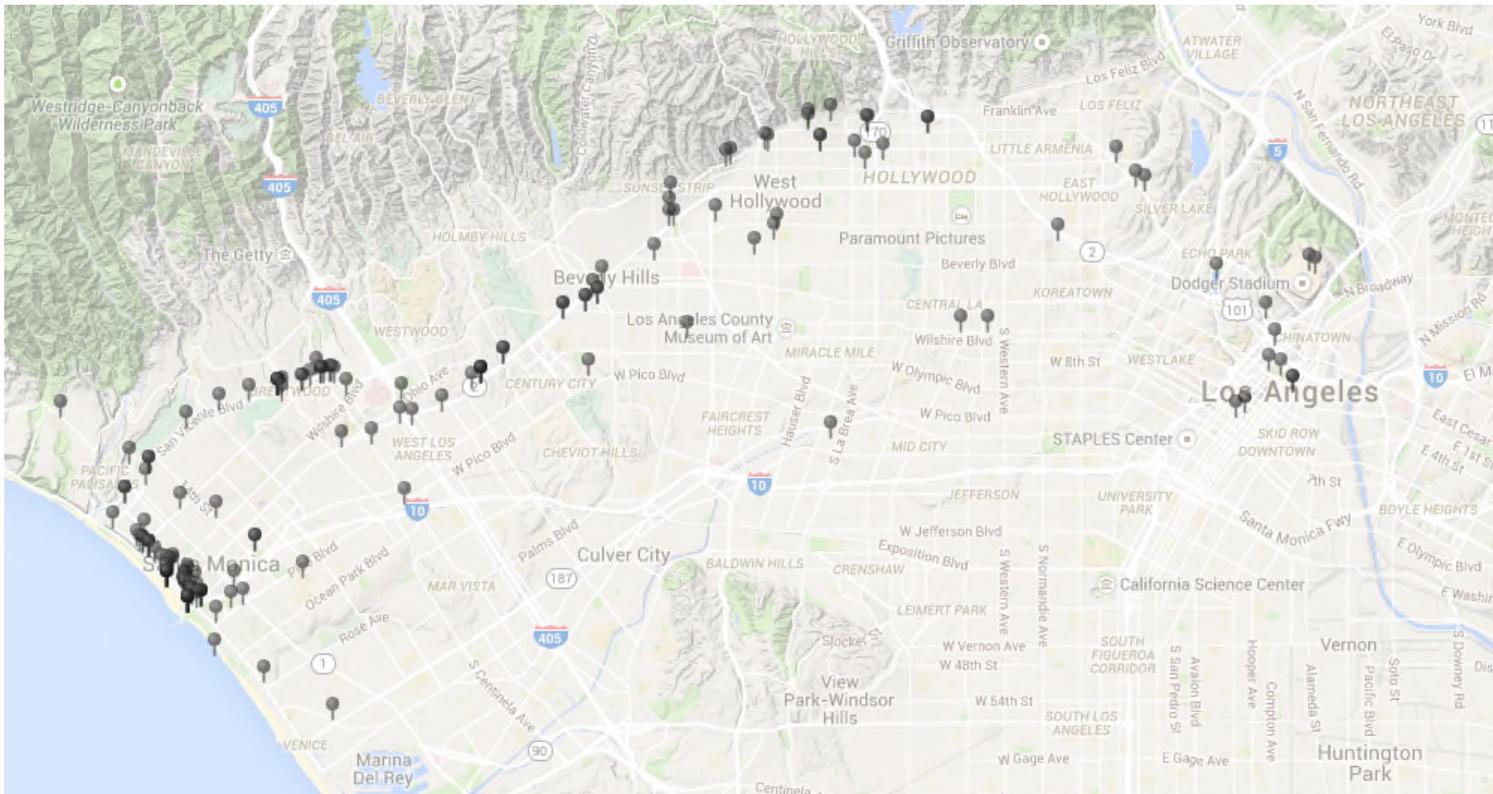
Feasibility of Instagram-based Tracking

- Tracking "LA Marathon": Late Stage
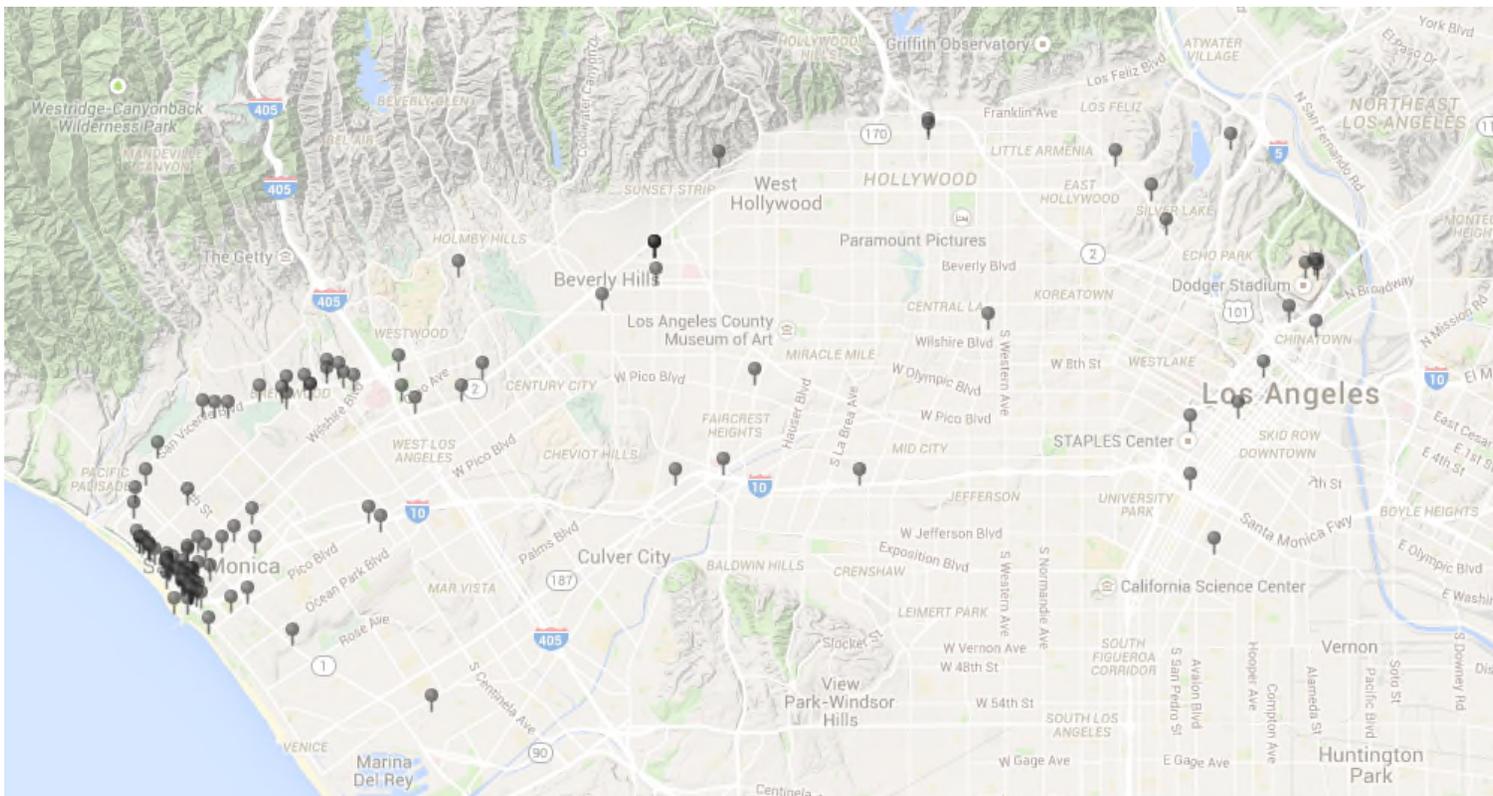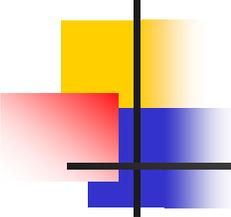
On False Positives

- Some picture clusters come from "echo" locations (e.g., large groups watching a concert event on a TV screen in bar). How to eliminate them?

- Main idea: eliminate smaller picture clusters that are similar in tags to a larger cluster
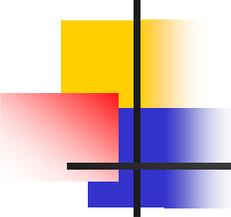
Event Detection Results

RECALL

| Dataset	Our Localization Algorithm	Tag Similarity Localization				Geo Event Detection [23]				Points of Interest [5]
		$X = 20\%$	$X = 40\%$	$X = 60\%$	$X = 80\%$	$\theta = 0.2$	$\theta = 0.4$	$\theta = 0.6$	$\theta = 0.8$	
Taylor Swift	28/28	24/28	25/28	26/28	26/28	28/28	28/28	28/28	28/28	27/28
Maroon V	17/17	13/17	15/17	17/17	17/17	17/17	17/17	17/17	17/17	17/17
Marathon	5/5	3/5	4/5	4/5	4/5	5/5	5/5	5/5	5/5	5/5
Tornado	3/3	2/3	3/3	3/3	3/3	3/3	3/3	3/3	3/3	3/3

FALSE POSITIVES

Dataset	Our Localization Algorithm	Tag Similarity Localization				Geo Event Detection [23]				Points of Interest [5]
		$X = 20\%$	$X = 40\%$	$X = 60\%$	$X = 80\%$	$\theta = 0.2$	$\theta = 0.4$	$\theta = 0.6$	$\theta = 0.8$	
Taylor Swift	2	18	10	5	4	35	16	9	9	26
Maroon V	0	5	4	4	2	19	8	8	8	14
Marathon	0	16	10	7	6	17	11	11	11	15
Tornado	1	3	3	3	2	6	6	6	6	19

40

Architecture for Twitter and Instagram Fusion

- Fusion of Twitter and Instagram Data
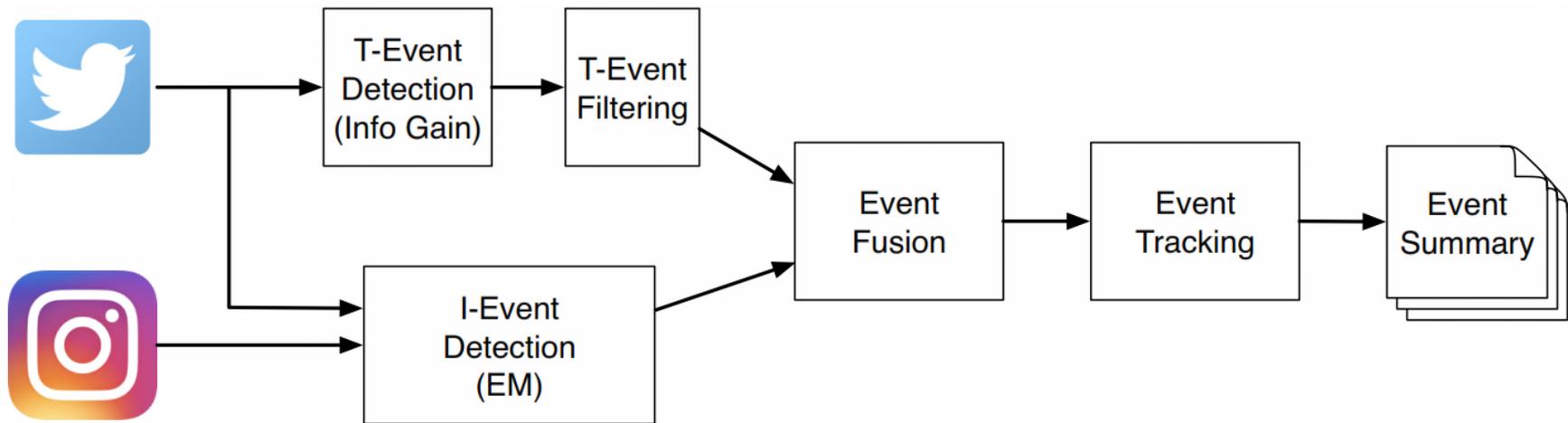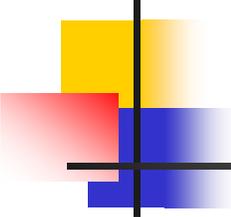
Evaluation

- Protest data set collected from Twitter and Instagram

Week Index	#Tweets	#Instagram posts
Feb 2016 Week 1	77001	1377
Feb 2016 Week 2	78334	1424
Feb 2016 Week 3	75639	1489
Feb 2016 Week 4	64669	1398

Evaluation

- Fusion offers a better trade-off between precision and recall

Algorithm	Total# events	Precision	Recall
Instagram	54	**87.037%**	47
Twitter	174	63.218%	110
Fusion	211	70.616%	**149**
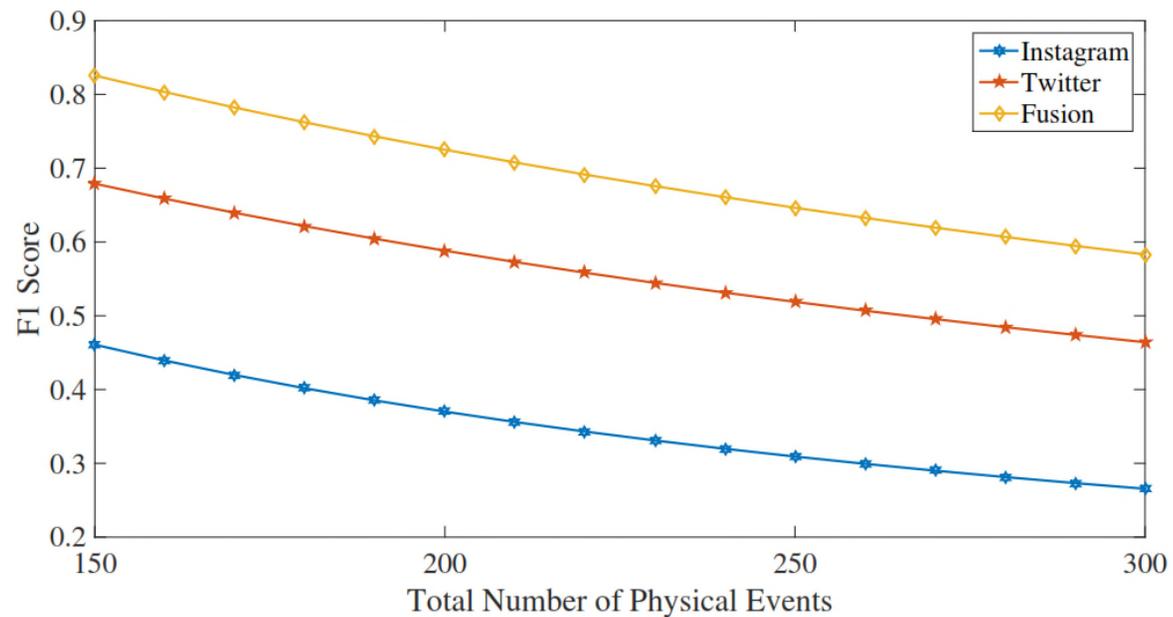
Example: Tracking the Nepal Earthquake

- A data association problem for a "moving" target

Date	New Detected Keywords	Sample Tweets
04/25/2015	nepal, earthquake, death	Powerful magnitude-7.8 earthquake that rocked Nepal triggered an avalanche on Mount Everest http://t.co/MULEuWhx3Q http://t.co/QeRKg8QgYp
		RT @BBCBreaking: At least 876 killed in Nepal #earthquake; deaths also reported in India, Tibet & Bangladesh http://t.co/3BTo9l1QZ4 http://Ž026
04/26/2015	help, nepalearthquake	RT @cnnbrk: At least 2,263 people have died in Nepal from massive #NepalEarthquake and aftershocks, official says. http://t.co/hCyjO7YyS7
		Anyone with information about my son Joseph Patrick please help #NepalEarthquake #Pray4Joe http://t.co/X2Kn7mOtRO http://t.co2026
04/27/2015	surges, devastation, drone, thankyoupm, donations	Nepal #earthquake: Death toll surges to 3,218; four aftershocks felt in last 12 hours http://t.co/Njvru9k2kQ
		@cnni: New drone footage shows the extent of devastation from the #NepalEarthquake: http://t.co/7PiPjayQZ1https://t.co/phIGRkYoZQ
		#ThankYouPM for massive rescue and relief operation by India in Nepal after #earthquake
		Nepal Earthquake: Facebook to Match Donations Made for Victims http://t.co/aLooadYNxj Free Submission http://t.co/J90dT2qnXb
04/28/2015	salute2indianforces, koirala, sanjay, humanitarian	Thank you very much Indin Forces for being with us.It means alot.... #Salute2IndianForces
		CNN's Dr. Sanjay Gupta performs surgery on girl in Nepal: CNN's Dr. Sanjay Gupta performed a life-saving... http://t.co/4EtmH28EwC #tcot
		Live: Nepal earthquake kills 4,352, PM Sushil Koirala says death toll could reach 10,000: A high-intensity ear... http://t.co/A68VtR6hWK
04/29/2015	survivor,hours, hospital, field, miracle	Nepal earthquake survivor drank urine while trapped for 82 hours http://t.co/v9DHM5Jhnf #worldnews
		That is amazing, Nepal Army rescued a 4-month kid alive after 22 hours! ::http://t.co/KzJPJeZDCx https://t.co/HvTkvS0Ba0 via @sharethis
		RT @haaretzcom: Nepal earthquake updates / Israeli field hospital opens, to treat 200 people per day http://t.co/PMwRlRT6YO http://t.co/s9i
04/30/2015	pakistan, serves, masala, teenage, lydia	Pakistan serves 'beef masala' to earthquake-hit Nepal via /r/worldnews http://t.co/GoFJO09mJP
		Teenage boy pulled out of rubble alive five days after Nepal earthquake http://t.co/0kiAigYE7M #telegraph #news
		Lydia Ko donating earnings to Nepal relief effort: The 18-year-old Ko, ranked No. 1 in the world, successfully... http://t.co/2nquCITqJa

Example: Jat Delhi Protest

Window ID	Tweets Sample	Bucket Type
1	VIDEO: Delhi water supplies sabotaged by protest No water left in Delhi due to Jat protest, schools closed, rationing begins due to CASTE…	T-bucket
2	Water rationed as India caste protest toll rises	T-bucket
3	Caste violence …violent protest had briefly shut down the water supply in New Delhi. …that Jat started protest, they destroyed munak canal, no water for Delhi	T-bucket
4	Heaven help us all. Upper caste protest in Delhi leads to death and destruction…	T-bucket
5	India Sends Soldiers To Area Of Caste Protest, water cut off in New Delhi Caste Protests Near Delhi Close Roads and Restrict #Water Supply: Though the Indian Army	T-bucket
6	RT @fakingnews: Fed up Delhi youth start a protest against protests	T-bucket
7	More than 10 million people in #India's capital are without water despite the army regaining control of its key water source after protest	T-bucket
8	Haryana State in India Proposes New Caste Status in Bid to Quell Protests Jats protest leaves millions in New Delhi without water	Fusion bucket
9	Deal reached to end Jat protests in India's Haryana state; roadblocks to be cleared, protest leader and police say	T-bucket

The Social Signal Layer

Frequency counts

Event Detection and Tracking

Event Data (Tweets Associated with the Event)

Keywords

Event Marathon

Event Protest

Event Flash mob

Social Sensing Signal, *Signal (k)*

Event Map
Events and Trajectories

Social Medium
Observers of Physical Events

Event Timeline
Moscow, Protest

06-13-2017 09:17:25

- Police in Moscow say they'll use face-recognition technology on street where Alexsey Navalny plans his main anti-corruption protest...
- Police urge people to leave site of unsanctioned protest on Tverskaya Street in downtown Moscow... https://t.co/ltxDCy0hyp

#Protest in central #Moscow right now: Tverskaya street from above #RussiaProtests https://t.co/m5DipjUGrB

- Police urge people to leave site of unsanctioned protest on Tverskaya Street in downtown Moscow... https://t.co/K4ciARfypK

#Protest in central #Moscow right now: Tverskaya street from above #RussiaProtests https://t.co/m5DipjUGrB

06-12-2017 20:16:24

- #Navalny says officials prevented companies from renting him audio/visual equip for protest rally so he's moving protest to Tverskaya st.
- 1/1 Activist Navalny said protest meeting in Moscow today at 14.00 will be moved from outside the centre tot the main street, Tverskaya >
- #Navalny protest march on Tverskaya - will be hard to know how many turn out since they could be mixed in w/others celebrating Russia Day.
- Tomorrow on Tverskaya St, where Navalny has moved his protest, Moscow is planning a historical reenactment festival. 150k people expected.
- Navalny just moved his protest tomorrow to the Moscow's main Tverskaya street, citing govt efforts to block preparation of the rally.

06-13-2017 05:17:16

Liveuamap: .navalny was detained in Moscow, prior to mass-protest on Tverskaya street https://t.co/mhfUiFuf6t https://t.co/rcIzM6zMHE vi...

Tverskaya street in downtown #Moscow now where @Navalny plans to hold anti-corruption #protest today: https://t.co/KJLvt36o4v

- Tverskaya street in Moscow where #Navalny protest meant to take place absolutely blockaded by viking ship and group... https://t.co/nFQNbpwFVS
- Police in Moscow say they'll use face-recognition technology on street where Alexsey Navalny plans his main anti-corruption protest...

06-13-2017 10:17:29

- On a day of mass arrests for "unsanctioned protest," girls make signs from Russia's constitution. (Photos via... https://t.co/8uzjTb18EE

Mass arrests reported as thousands across Russia protest corruption, confront Kremlin https://t.co/zO332tvHoB https://t.co/OOoamphsMo

06-13-2017 18:19:06

- Revealing! Western #MSM & pundits reporting non-stop on #Navalny 'protest' but no coverage of big celebrations... https://t.co/HoHFvkIQa4
- Globalist darling #Navalny organized protest 1) to ruin Russia Day festivities 2) to create photo-ops for Western p... https://t.co/bBrVIZiPlZ
- Moscow corrs for Western pubs conveniently leaving out fact that Navalny moved his protest to the site of June 12 holiday celebrations.
- Navalny protests: Western media don't care about facts when cheering their Russian opposition hero https://t.co/j6HUv3LEmu
- Revealing! Western #MSM & pundits reporting non-stop on #Navalny 'protest' but no coverage of big celebrations... https://t.co/HoHFvkIQa4

Event Timeline
Moscow, Protest

06-12-2017 20:16:24

- #Navalny says officials prevented companies from renting him audio/visual equip for protest rally so he's moving protest to Tverskaya st.
- 1/1 Activist Navalny said protest meeting in Moscow today at 14.00 will be moved from outside the centre tot the main street, Tverskaya >
- #Navalny protest march on Tverskaya - will be hard to know how many turn out since they could be mixed in w/others celebrating Russia Day.
- Tomorrow on Tverskaya St, where Navalny has moved his protest, Moscow is planning a historical reenactment festival. 150k people expected.
- Navalny just moved his protest tomorrow to the Moscow's main Tverskaya street, citing govt efforts to block preparation of the rally.

06-13-2017 05:17:16

Liveuamap: .navalny was detained in Moscow, prior to mass-protest on Tverskaya street https://t.co/mhfUiFuf6t https://t.co/rcIzM6zMHE vi...

Tverskaya street in downtown #Moscow now where @Navalny plans to hold anti-corruption #protest today: https://t.co/KJLvt36o4v

- Tverskaya street in Moscow where #Navalny protest meant to take place absolutely blockaded by viking ship and group... https://t.co/nFQNbpwFVS
- Police in Moscow say they'll use face-recognition technology on street where Alexsey Navalny plans his main anti-corruption protest...

06-13-2017 09:17:25

- Police in Moscow say they'll use face-recognition technology on street where Alexsey Navalny plans his main anti-corruption protest...
- Police urge people to leave site of unsanctioned protest on Tverskaya Street in downtown Moscow... https://t.co/ltxDCy0hyp

#Protest in central #Moscow right now: Tverskaya street from above #RussiaProtests https://t.co/m5DipjUGrB

- Police urge people to leave site of unsanctioned protest on Tverskaya Street in downtown Moscow... https://t.co/K4ciARfypK

#Protest in central #Moscow right now: Tverskaya street from above #RussiaProtests https://t.co/m5DipjUGrB

06-13-2017 10:17:29

- On a day of mass arrests for "unsanctioned protest," girls make signs from Russia's constitution. (Photos via... https://t.co/8uziT518EE

Mass arrests reported as thousands across Russia protest corruption, confront Kremlin https://t.co/zO332tvHoB https://t.co/OOoamphsMo

True or False?

06-13-2017 18:19:06

- Revealing! Western #MSM & pundits reporting non-stop on #Navalny 'protest' but no coverage of big celebrations... https://t.co/HoHFvkIQa4
- Globalist darling #Navalny organized protest 1) to ruin Russia Day festivities 2) to create photo-ops for Western p... https://t.co/bBrVIZiPlZ
- Moscow corrs for Western pubs conveniently leaving out fact that Navalny moved his protest to the site of June 12 holiday celebrations.
- Navalny protests: Western media don't care about facts when cheering their Russian opposition hero https://t.co/j6HUv3LEmu
- Revealing! Western #MSM & pundits reporting non-stop on #Navalny 'protest' but no coverage of big celebrations... https://t.co/HoHFvkIQa4

Capability #2

Veracity Analysis
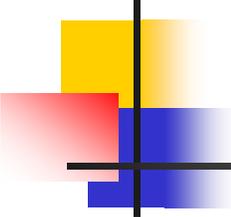
Insight

- Think of short text excerpts (e.g., tweets) as true/false. Jointly assess veracity of tweets and sources:
    - Tweets with more independent (and more truthful) sources are more likely to be true
    - Sources with a larger fraction of tweets deemed true are more likely to be reliable
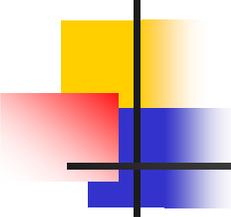
Tweets as Information "bits"

- ## A binary signal model

Each observation is a "bit" of information: Either true or false

Examples of Twitter "Sensing"

Crash blocking lanes on I-5S @ McBean Pkwy in Santa Clarita

BREAKING NEWS: Shots fired in Watertown; source says Boston Marathon terror bomb suspect has been pinned down

The police chief of Afghanistan's southern Kandahar province has died in a suicide attack on his headquarters.

Yonkers mayor has lifted his gas rationing order. Fill it up!
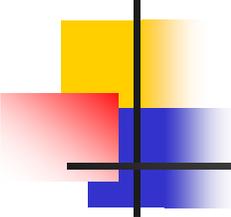
Noise on the Social Medium

Examples of Twitter "Noise"
Why are iphone chargers not called Apple Juice?
Nothing makes me happier than traffic in LA!

Logical "bit-channel" Layer: Data Cleaning
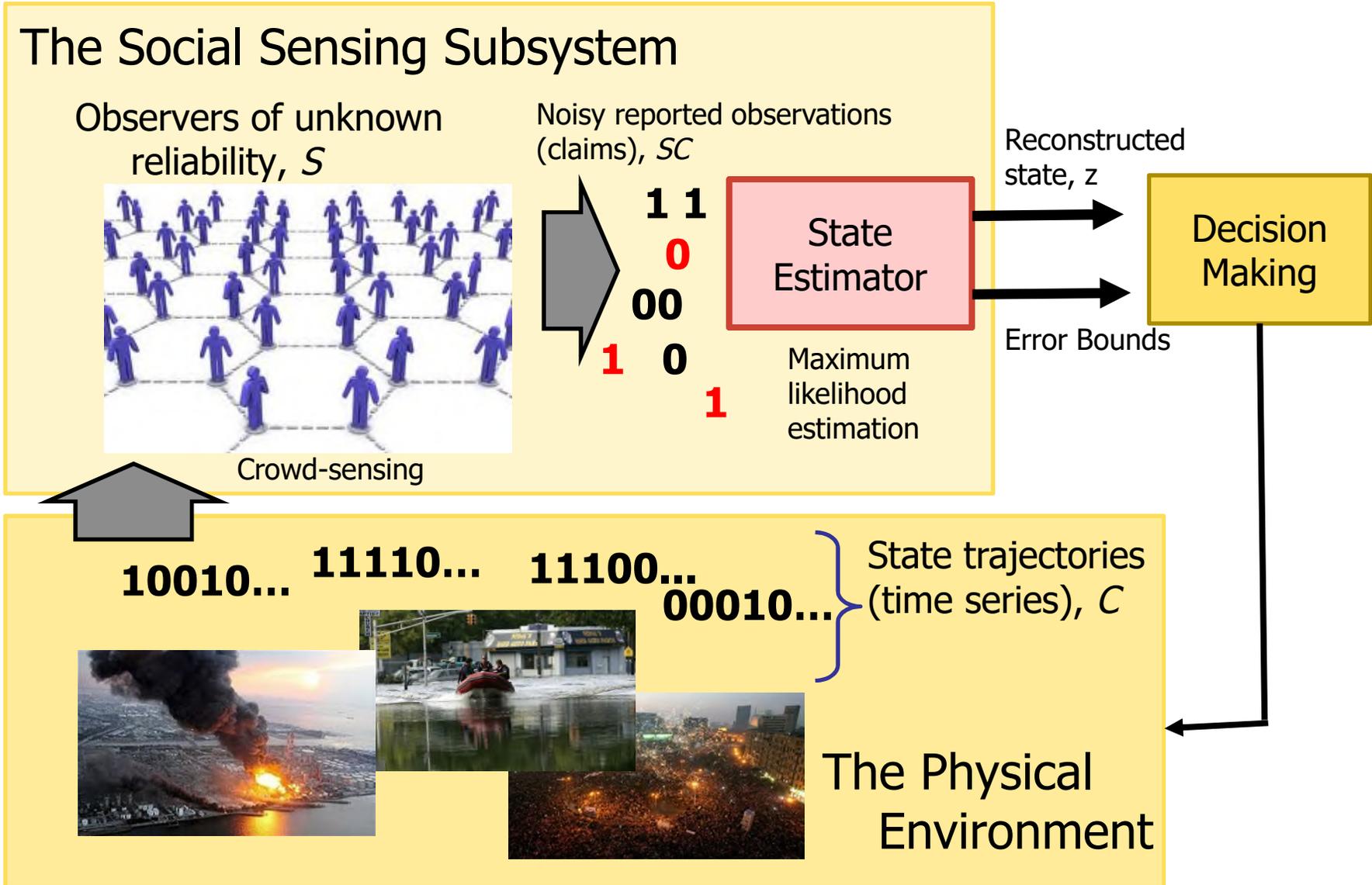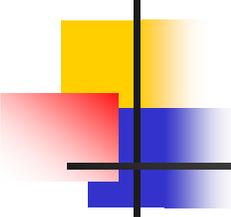
The Social Sensing Subsystem

Observers of unknown reliability, S

Noisy reported observations (claims), SC

Crowd-sensing

1 1
0
00
1 0
1

State Estimator

Maximum likelihood estimation

Reconstructed state, z

Error Bounds

Decision Making

10010... 11110... 11100...
00010...

State trajectories (time series), C

The Physical Environment

Social Networks as Noisy Channels

- The source noise model:
  - Source error distribution
    - Prob (correct), and
    - Prob (wrong) = 1 – Prob (correct)
  - The distribution is not known *a priori*
- Uncertain provenance
 - When John makes an observation it is impossible to tell if he is the original source
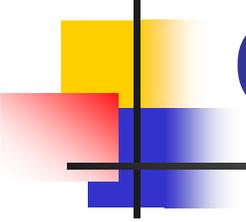
Social Networks as Noisy Channels

- The source noise model:
  - Source error distribution
    - Prob (correct), and
    - Prob (wrong) = 1 − Prob (correct)
  - The distribution is not known *a priori*
- Uncertain provenance
  - When John makes an observation it is impossible to tell if he is the original source

Source reliability
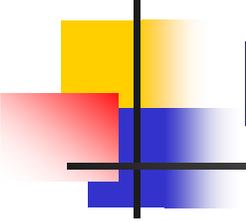
The Channel "Decoding" Problem

- Joint estimation of
  - Source reliability
  - True/false value of each observation (claim)
- Given
  - Who said what, and
  - Correlations between sources
    - The empirically observed probability that one source repeats another

Social Channel "Decoding"
## A Maximum Likelihood Estimation Problem


Boston Bombing

Hurricane Sandy
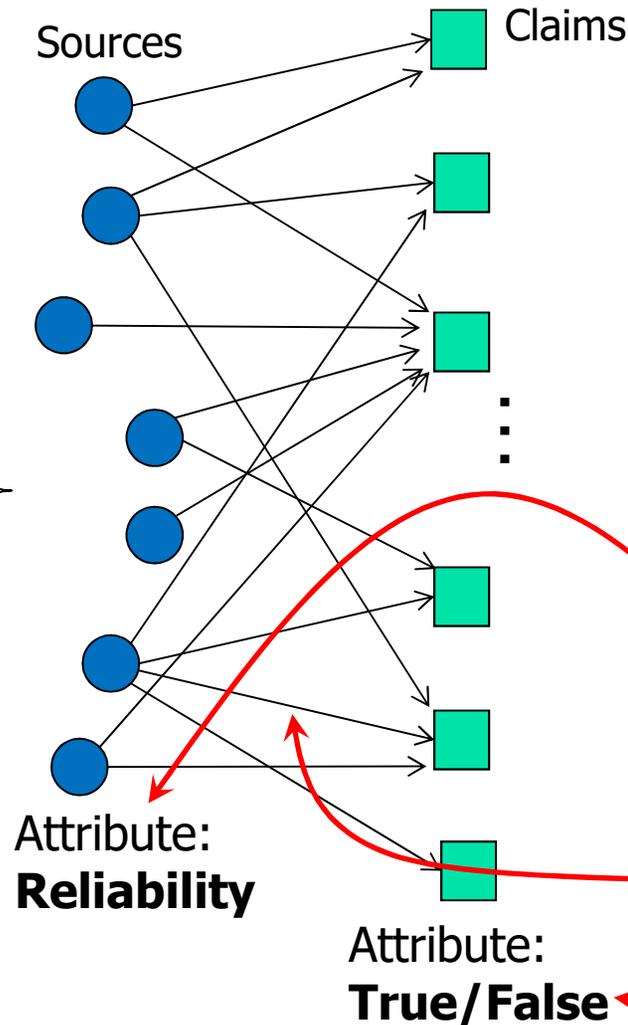
Egypt unrest

Sources

Claims

Attribute: **Reliability**

Attribute: **True/False**

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

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

Source Dependencies

Boston Bombing

Hurricane Sandy
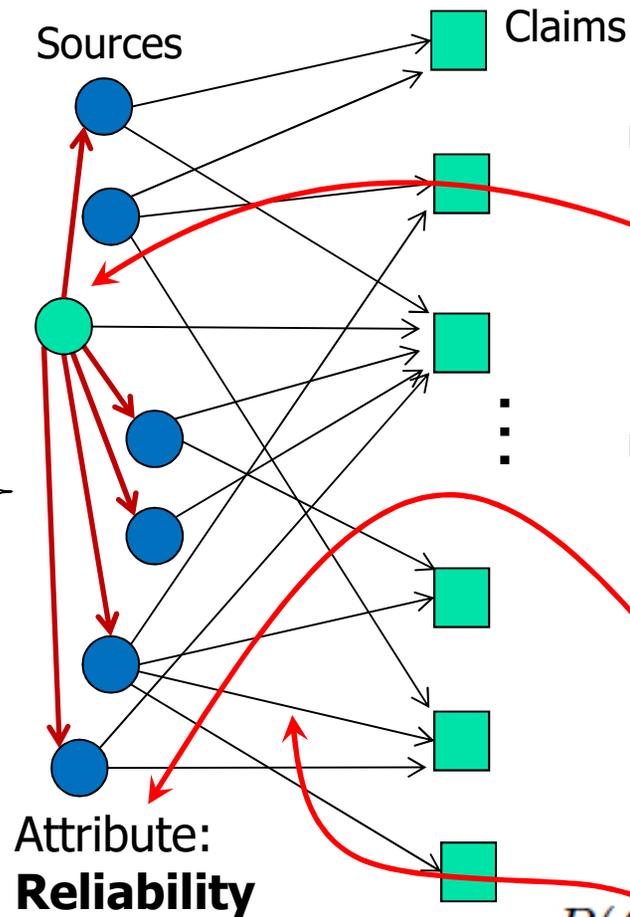
Egypt unrest

Sources

Claims

Attribute: **Reliability**

Attribute: **True/False**

- Joint estimation of
  - Source reliability
  - True/false value of each observation
- Given
  - Who said what, and
  - Correlations between sources

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

# Expectation Maximization

Likelihood Function Incorporating Source Dependency

$$P(SC, z|SD, \theta) = \prod_{j=1}^{N} P(z_j) \times$$

$$\{ \prod_{g \in M_j} P(S_g C_j|\theta, z_j) \prod_{i \in c_g} P(S_i C_j|S_g C_j) \}$$

$$P(z_j) = \begin{cases} d & z_j = 1 \\ (1-d) & z_j = 0 \end{cases}$$

$$P(S_g C_j|\theta, z_j) = \begin{cases} a_g & z_j = 1, S_g C_j = 1 \\ (1-a_g) & z_j = 1, S_g C_j = 0 \\ b_g & z_j = 0, S_g C_j = 1 \\ (1-b_g) & z_j = 0, S_g C_j = 0 \end{cases}$$

$$P(S_i C_j|S_g C_j) = \begin{cases} p_{ig} & S_g C_j = 1, S_i C_j = 1 \\ 1 - p_{ig} & S_g C_j = 1, S_i C_j = 0 \end{cases}$$

g

Claim

i
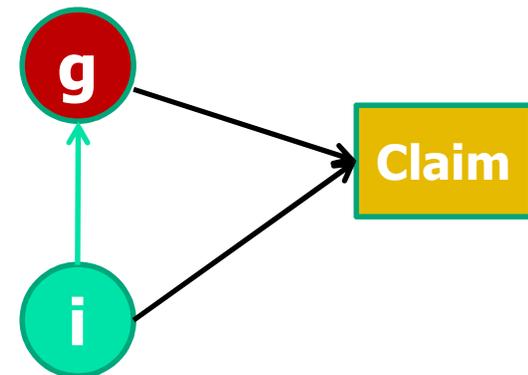
Dependent Sources

59

Expectation Maximization

E-Step

M-Step

$$Q\left(\theta|\theta^{(n)}\right) = \sum_{j=1}^{N} \left\{ Z(n,j) \times \left[\left\{ \sum_{g \in M_j} \left(\log P(S_g C_j|\theta, z_j) \right. \right. \right. \right.$$

$$+ \sum_{i \in c_g} \log P(S_i C_j|S_g C_j) \right) \right\} + \log d \right]$$

$$+ (1 - Z(n,j)) \times \left[\left\{ \sum_{g \in M_j} \left(\log P(S_g C_j|\theta, z_j) \right. \right. \right.$$

$$+ \sum_{i \in c_g} \log P(S_i C_j|S_g C_j) \right) \right\} + \log(1 - d) \right] \right\} \tag{10}$$

$$a_g^{(n+1)} = a_g^* = \frac{\sum_{j \in SJ_g} Z(n,j)}{\sum_{j=1}^{N} Z(n,j)}$$

$$a_i^{(n+1)} = a_i^* = \frac{\sum_{j \in \overline{SJ_g} \cap SJ_i} Z(n,j)}{\sum_{j \in \overline{SJ_g}} Z(n,j)} \quad \text{for } i \in c_g$$

$$b_g^{(n+1)} = b_g^* = \frac{\sum_{j \in SJ_g}(1 - Z(n,j))}{\sum_{j=1}^{N}(1 - Z(n,j))}$$

$$b_i^{(n+1)} = b_i^* = \frac{\sum_{j \in \overline{SJ_g} \cap SJ_i}(1 - Z(n,j))}{\sum_{j \in \overline{SJ_g}}(1 - Z(n,j))} \quad \text{for } i \in c_g$$

$$d^{(n+1)} = d^* = \frac{\sum_{j=1}^{N} Z(n,j)}{N}$$

A Controlled "Social Sensing" Experiment

Goal: Find Free Parking Lots on UIUC Campus

Y: Free Parking lot - no charge after 5pm
N: Not free parking lot

61

A Controlled "Social Sensing" Experiment

Goal: Find Free Parking Lots on UIUC Campus

Table I. Accuracy of Finding Free Parking Lots on Campus

Schemes	False Positives	False Negatives
EM-Conflict	**6.67%**	**10.87%**
EM-Regular	11.67%	17.39%
Average-Log	16.67%	19.57%
Truth-Finder	18.33%	15.22%
Voting	21.67%	23.91%

Experiment setup:
106 parking lots of interests, 46 indeed free
30 participants, 901 marks collected

Humans as (Noisy) Sensors

- 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
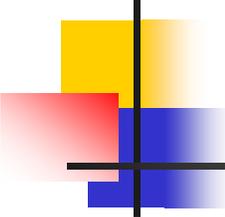
The Experiments

- Run the maximum likelihood estimator on Twitter data to determine the probability of correctness of different tweets
- Sort tweets by probability of correctness.
- Give the top N tweets to a human for "grading"
- Human must investigate each tweet to determine if it is true.
- Any tweet that cannot be shown to be true is considered "unconfirmed"
- Compare the percentages of unconfirmed tweets across different credibility estimation algorithms

Note: To remove bias, the grader was not told which algorithm "believed" which tweet.

# Evaluation: Sandy Trace
A Comparison of Confirmed True Tweets

RT: Retweet

FF: Follower-Followee

EC: Epidemic Cascade

65
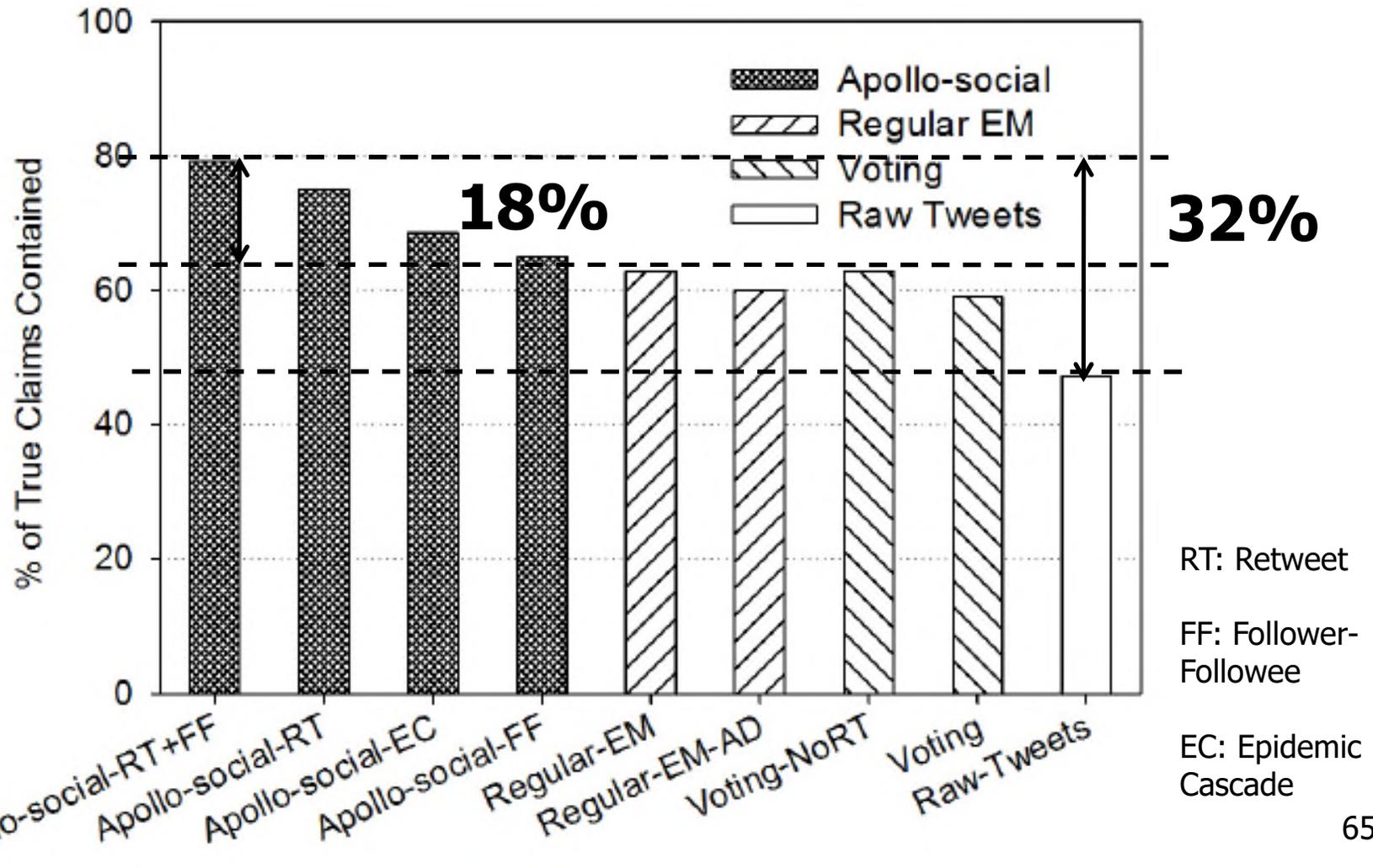
Evaluation: Irene Trace
A Comparison of Confirmed True Tweets

RT: Retweet

FF: Follower-Followee

EC: Epidemic Cascade
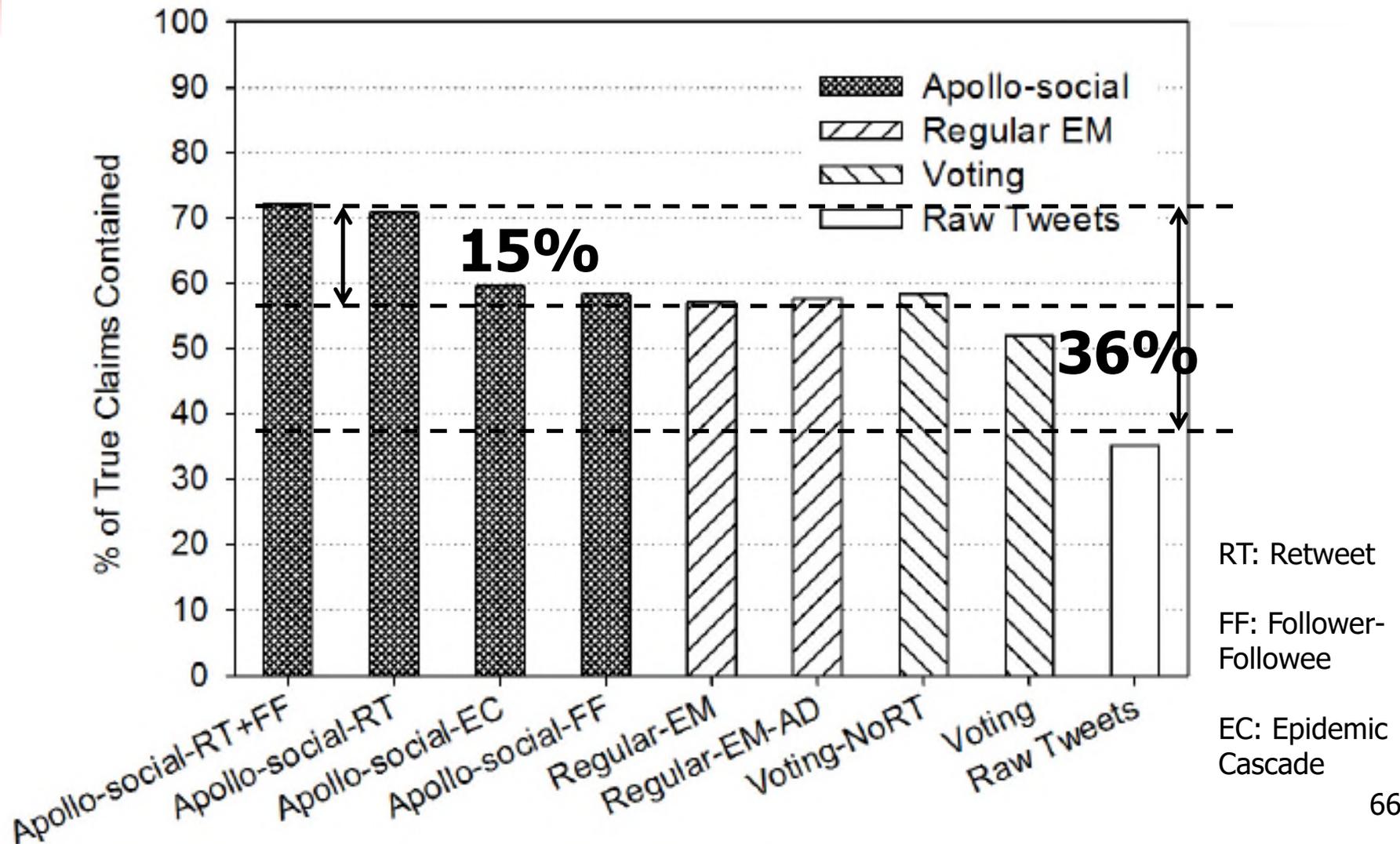
Evaluation: Egypt Trace
A Comparison of Confirmed True Tweets
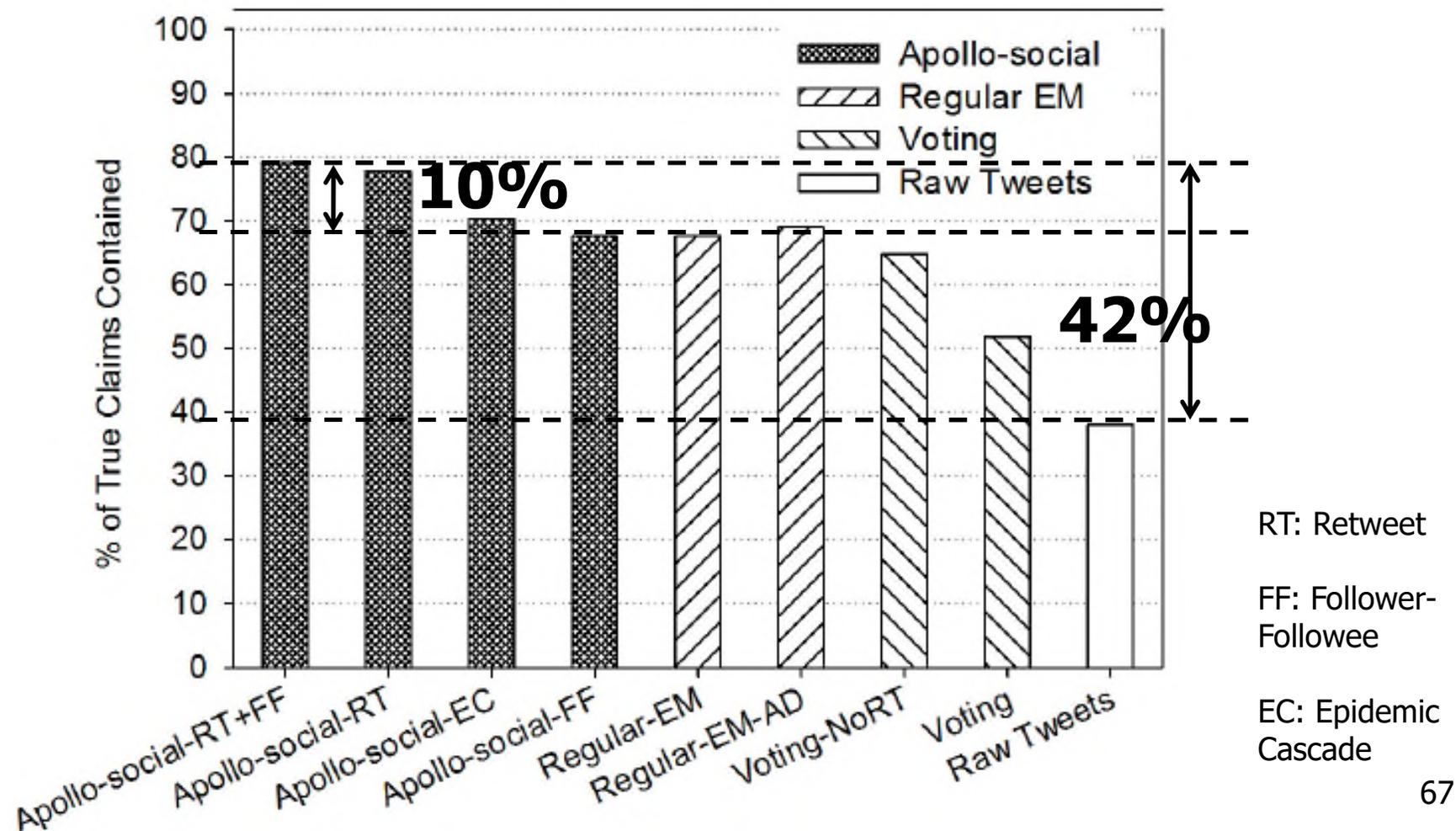

RT: Retweet

FF: Follower-Followee

EC: Epidemic Cascade

Real-time Estimation

- ## Offline "decoding":

Observations, probability of correctness, and source reliability estimates

Bulk Algorithm

Entire Data Set
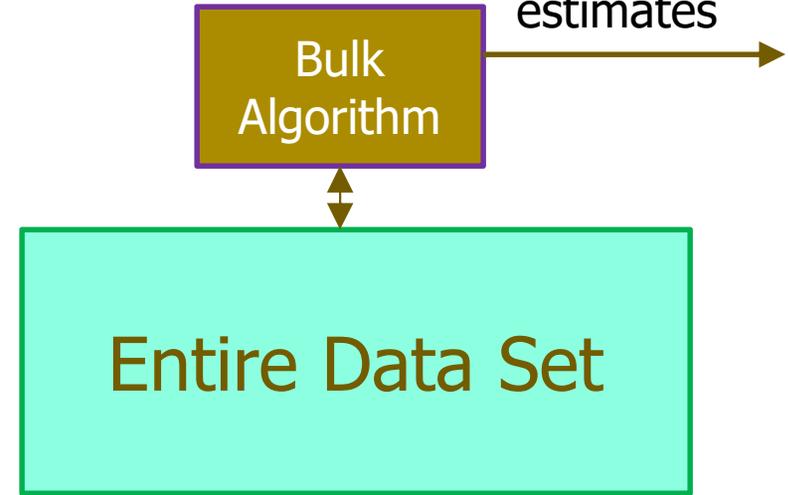
Adapt veracity and social network estimates dynamically

- ## Online "decoding":

Data Updates

Data Updates

Data Updates

Bulk Algorithm

Learned History

Observations, probability of correctness, and source reliability estimates
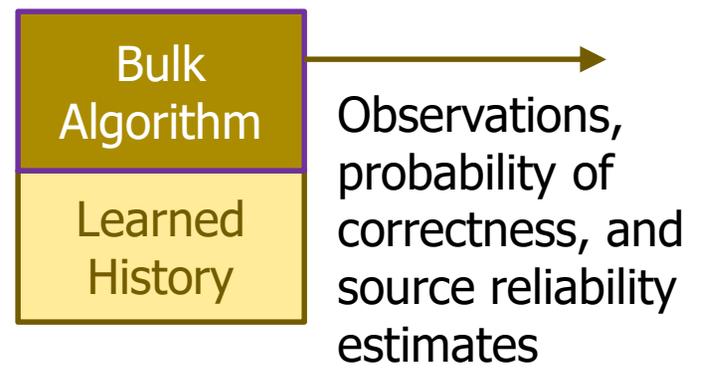
Beating the News

Example of Twitter "Signal"

#BREAKING Many feared dead including students as suicide attack hit #UN HQ in #Somalia's Garoowe city. More updates soon.

(4/20/2015, 1:45am EST)

Note: reported by Apollo 1 hour before first appearance in news.

(ABC @2:40am)(Fox @3:05am)(Yahoo @3:15am)(Aljazeera @3:30am)(BBC @ 3:40am)(Reuters @3:45am)(CNN @4:15am)...


1am EST   2am EST   3am EST   4am EST   Time

4/20/2015

@1:45am Declared "Likely True" by Apollo

@2:40am Reported by ABC News

@3:05am Reported by Fox News

@3:15am Reported by Yahoo

@3:30am Reported by Aljazeera

@3:40am Reported by BBC News

@3:45am Reported by Reuters

@4:15am Reported by CNN

Beating the News

Examples of Twitter "Signal"

Syrian rebels seize major army base in Deraa
Note: Reported by Apollo hours before first appearance in news.
(BBC @8am)(Daily Telegraph @9:30)(IB Times @11am)(Reuters @2pm)(Aljazeera @ 4pm)(The Guardian @5pm)...


Timeline (EST): 2am, 4am, 6am, 8am, 10am, 12pm, 2pm, 4pm, 6pm

Apollo "Likely True" reports

Channel News Asia

BBC

International Business Times

Reuters

Yahoo News

Aljazeera

The Guardian

New York Times

Daily Telegraph

#Syria #Daraa Rebels took control of Battalion 14 inside the Brigade 52 base
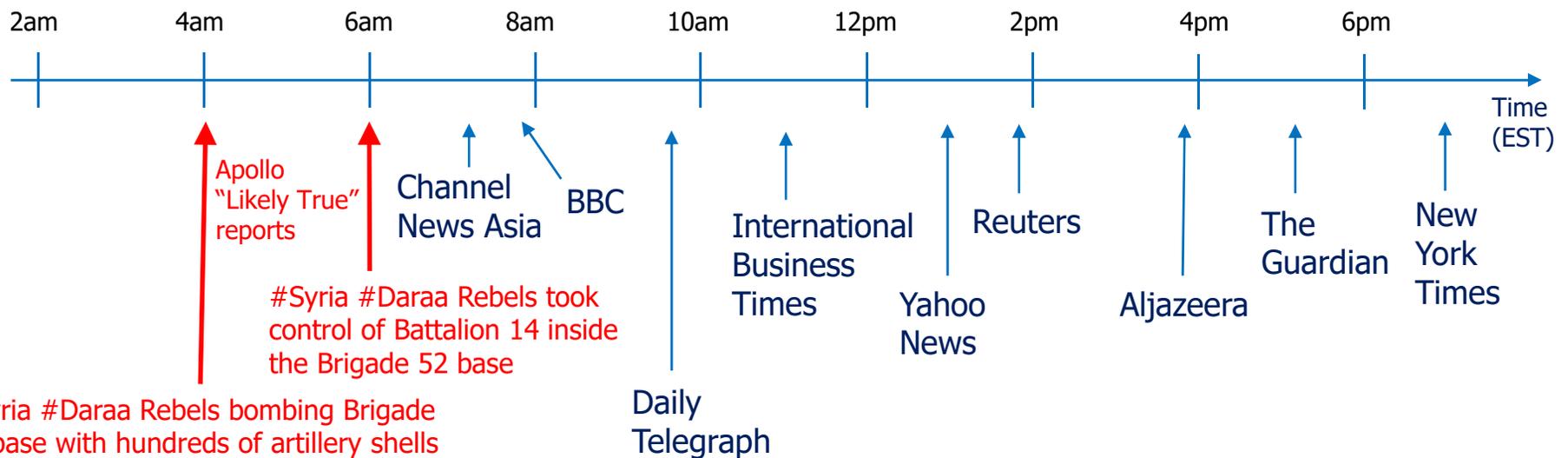
#Syria #Daraa Rebels bombing Brigade 52 base with hundreds of artillery shells & rockets -note base is 1200 Hecta big
http://t.co/y4k706jteX

A Study of Multiple Events
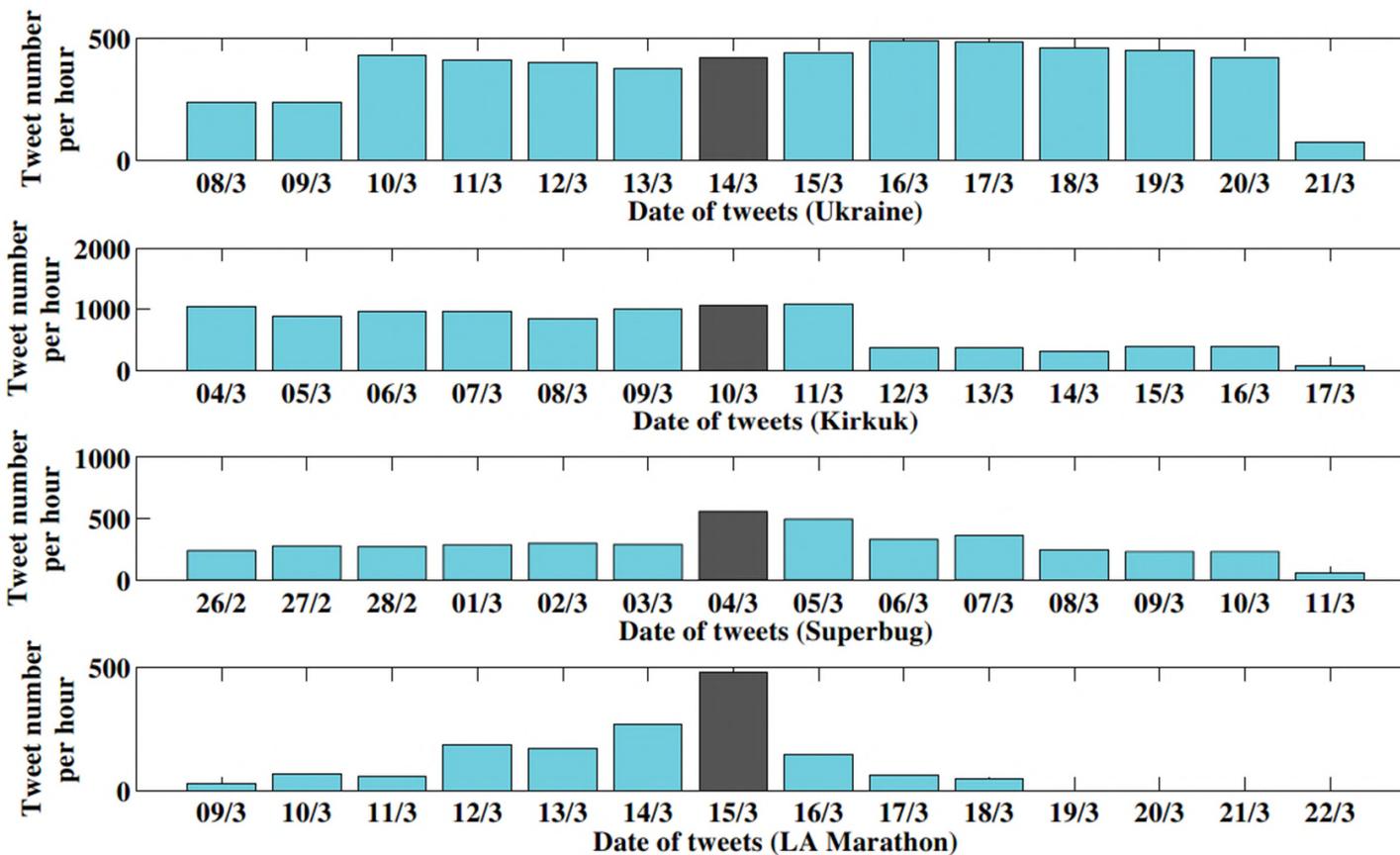
Accuracy Results of Real-time Estimation

- Improvement in *both* Accuracy and Speed

Event Timeline
Moscow, Protest

06-12-2017 20:16:24

- #Navalny says officials prevented companies from renting him audio/visual equip for protest rally so he's moving protest to Tverskaya st.
- 1/1 Activist Navalny said protest meeting in Moscow today at 14.00 will be moved from outside the centre tot the main street, Tverskaya >
- #Navalny protest march on Tverskaya - will be hard to know how many turn out since they could be mixed in w/others celebrating Russia Day.
- Tomorrow on Tverskaya St, where Navalny has moved his protest, Moscow is planning a historical reenactment festival. 150k people expected.
- Navalny just moved his protest tomorrow to the Moscow's main Tverskaya street, citing govt efforts to block preparation of the rally.

06-13-2017 05:17:16

Liveuamap: .navalny was detained in Moscow, prior to mass-protest on Tverskaya street https://t.co/mhfUiFuf6t https://t.co/rclzM6zMHE vi...

Tverskaya street in downtown #Moscow now where @Navalny plans to hold anti-corruption #protest today: https://t.co/KJLvt36o4v

- Tverskaya street in Moscow where #Navalny protest meant to take place absolutely blockaded by viking ship and group... https://t.co /nFQNbpwFVS
- Police in Moscow say they'll use face-recognition technology on street where Alexsey Navalny plans his main anti-corruption protest...

06-13-2017 09:17:25

- Police in Moscow say they'll use face-recognition technology on street where Alexsey Navalny plans his main anti-corruption protest...
- Police urge people to leave site of unsanctioned protest on Tverskaya Street in downtown Moscow... https://t.co/ltxDCy0hyp

#Protest in central #Moscow right now: Tverskaya street from above #RussiaProtests https://t.co /m5DipjUGrB

- Police urge people to leave site of unsanctioned protest on Tverskaya Street in downtown Moscow... https://t.co/K4ciARfypK

#Protest in central #Moscow right now: Tverskaya street from above #RussiaProtests https://t.co /m5DipjUGrB

06-13-2017 10:17:29

- On a day of mass arrests for "unsanctioned protest," girls make signs from Russia's constitution. (Photos via... https://t.co/8uziTb18EE

Mass arrests reported as thousands across Russia protest corruption, confront Kremlin https://t.co/zO332tvHoB https://t.co/OOoamphsMo

True or False?

06-13-2017 18:19:06

- Revealing! Western #MSM & pundits reporting non-stop on #Navalny 'protest' but no coverage of big celebrations... https://t.co/HoHFvkIQa4
- Globalist darling #Navalny organized protest 1) to ruin Russia Day festivities 2) to create photo-ops for Western p... https://t.co /bBrVIZiPlZ
- Moscow corrs for Western pubs conveniently leaving out fact that Navalny moved his protest to the site of June 12 holiday celebrations.
- Navalny protests: Western media don't care about facts when cheering their Russian opposition hero https://t.co/j6HUv3LEmu
- Revealing! Western #MSM & pundits reporting non-stop on #Navalny 'protest' but no coverage of big celebrations... https://t.co/HoHFvkIQa4

# Event Timeline
Moscow, Protest

06-12-2017 20:16:24

- #Navalny says officials prevented companies from renting him audio/visual equip for protest rally so he's moving protest to Tverskaya st.
- 1/1 Activist Navalny said protest meeting in Moscow today at 14.00 will be moved from outside the centre tot the main street, Tverskaya >
- #Navalny protest march on Tverskaya - will be hard to know how many turn out since they could be mixed in w/others celebrating Russia Day.
- Tomorrow on Tverskaya St, where Navalny has moved his protest, Moscow is planning a historical reenactment festival. 150k people expected.
- Navalny just moved his protest tomorrow to the Moscow's main Tverskaya street, citing govt efforts to block preparation of the rally.

06-13-2017 05:17:16

Liveuamap: .navalny was detained in Moscow, prior to mass-protest on Tverskaya street https://t.co/mhfUiFuf6t https://t.co/rcIzM6zMHE vi...

Tverskaya street in downtown #Moscow now where @Navalny plans to hold anti-corruption #protest today: https://t.co/KJLvt36o4v

- Tverskaya street in Moscow where #Navalny protest meant to take place absolutely blockaded by viking ship and group... https://t.co/nFQNbpwFVS
- Police in Moscow say they'll use face-recognition technology on street where Alexsey Navalny plans his main anti-corruption protest...

06-13-2017 09:17:25

- Police in Moscow say they'll use face-recognition technology on street where Alexsey Navalny plans his main anti-corruption protest...
- Police urge people to leave site of unsanctioned protest on Tverskaya Street in downtown Moscow... https://t.co/ItxDCy0hyp

#Protest in central #Moscow right now: Tverskaya street from above #RussiaProtests https://t.co/m5DipjUGrB

- Police urge people to leave site of unsanctioned protest on Tverskaya Street in downtown Moscow... https://t.co/K4ciARfypK

06-13-2017 10:17:29

Mass arrests reported as thousands across Russia protest corruption, confront Kremlin https://t.co/zO332tvHoB https://t.co/OOoamphsMo

06-13-2017 18:19:06

- Revealing! Western #MSM & pundits reporting non-stop on #Navalny 'protest' but no coverage of big celebrations... https://t.co/HoHFvkIQa4
- Globalist darling #Navalny organized protest 1) to ruin Russia Day festivities 2) to create photo-ops for Western p... https://t.co/bBrVIZiPIZ
- Moscow corrs for Western pubs conveniently leaving out fact that Navalny moved his protest to the site of June 12 holiday celebrations.
- Navalny protests: Western media don't care about facts when cheering their Russian opposition hero https://t.co/j6HUv3LEmu
- Revealing! Western #MSM & pundits reporting non-stop on #Navalny 'protest' but no coverage of big celebrations... https://t.co/HoHFvkIQa4

Apollo: True.
- Hundreds detained across Russia as thousands protest against Vladimir Putin https://t.co/Gc0Dkr33Kr

Event Timeline
Moscow, Protest

06-12-2017 20:16:24

• #Navalny says officials prevented companies from renting him audio/visual equip for protest rally so he's moving protest to Tverskaya st.

• 1/1 Activist Navalny said protest meeting in Moscow today at 14.00 will be moved from outside the centre tot the main street, Tverskaya >

• #Navalny protest march on Tverskaya - will be hard to know how many turn out since they could be mixed in w/others celebrating Russia Day.

• Tomorrow on Tverskaya St, where Navalny has moved his protest, Moscow is planning a historical reenactment festival. 150k people expected.

• Navalny just moved his protest tomorrow to the Moscow's main Tverskaya street, citing govt efforts to block preparation of the rally.

06-13-2017 05:17:14

Liveuamap: .navalny was detained in Moscow, prior to mass-protest on Tverskaya street https://t.co/mhfL... https://t.co/rcIzM...

...verskaya street in downtown #Moscow now where @Navalny plans to hold anti-corruption #protest today: https://t.co/KJLvt36o4v

• Tverskaya street in Moscow where #Navalny protest meant to take place absolutely blockaded by viking ship and group... https://t.co /nFQNbpwFVS

• Police in Moscow say they'll use face-recognition technology on street where Alexsey Navalny plans his main anti-corruption protest...

06-13-2017 09:17:25

• Police in Moscow say they'll use face-recognition technology on street where Alexsey Navalny plans his main anti-corruption protest...

• Police urge people to leave site of unsanctioned protest on Tverskaya Street in downtown Moscow... https://t.co/ltxDCy0hyp

#Protest in central #Moscow right now: Tverskaya street from above #RussiaProtests https://t.co /m5DipjUGrB

• Police urge people to leave site of unsanctioned protest on Tverskaya Street in downtown Moscow... https://t.co/K4ciARfypK

#Protest in c... #Mosc... T...

06-13-2017 10:17:29

...mass arrests for ...tioned protest," girls ...ke signs from Russia's constitution. (Photos via... https://t.co/8uzjTb18EE

Mass arrests reported as thousands across Russia protest corruption, confront Kremlin https://t.co/zO332tvHoB https://t.co/OOoamphsMo

06-13-2017 18:19:06

• Revealing! Western #MSM & pundits reporting non-stop on #Navalny 'protest' but no coverage of big celebrations... https://t.co/HoHFvkIQa4

• Globalist darling #Navalny organized protest 1) to ruin Russia Day festivities 2) to create photo-ops for Western p... https://t.co /bBrVlZiPlZ

• Moscow corrs for Western pubs conveniently leaving out fact that Navalny moved his protest to the site of June 12 holiday celebrations.

• Navalny protests: Western media don't care about facts when cheering their Russian opposition hero https://t.co/j6HUv3LEmu

• Revealing! Western #MSM & pundits reporting non-stop on #Navalny 'protest' but no coverage of big celebrations... https://t.co/HoHFvkIQa4

Can we automatically separate the different points of view?
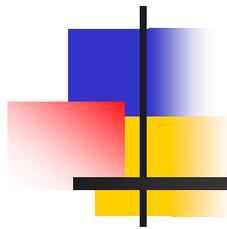
Capability #3

Analysis of Polarization
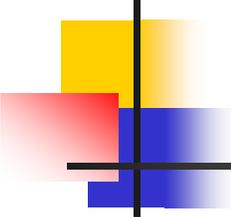
Insight

- People propagate content that agrees with their biases and beliefs (and filter out the rest)
  - Cluster content into different bins according to the way it propagates and you will isolate the different points of view!

Example: Polarization in Egypt
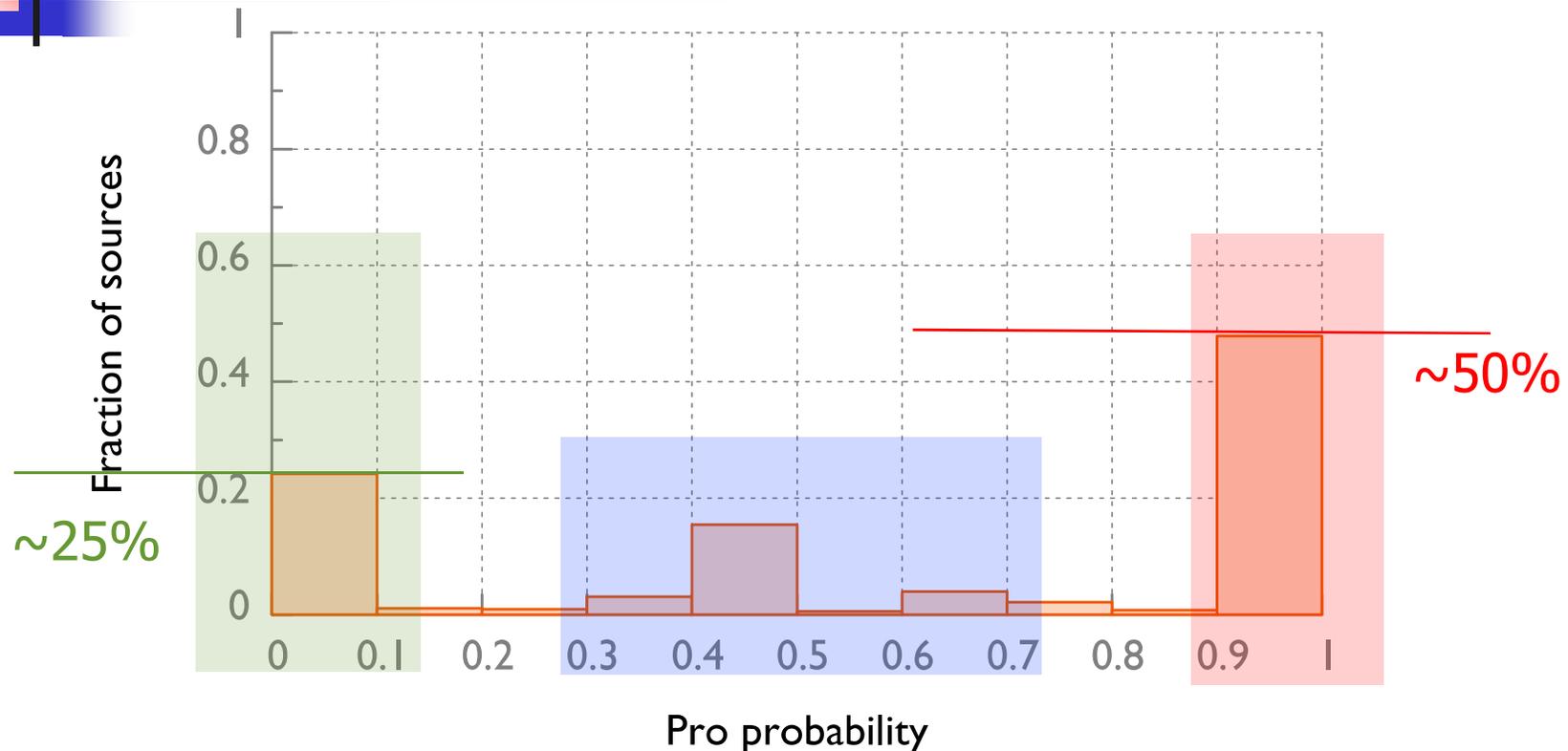
- Some users are highly polarized, and mostly forward tweets favoring camp they belong to
  - Pro-Morsi or Anti-Morsi

Egypt Social Network
of polarized sources

——— Anti network   ——— Pro network

Model of a Polarized Network
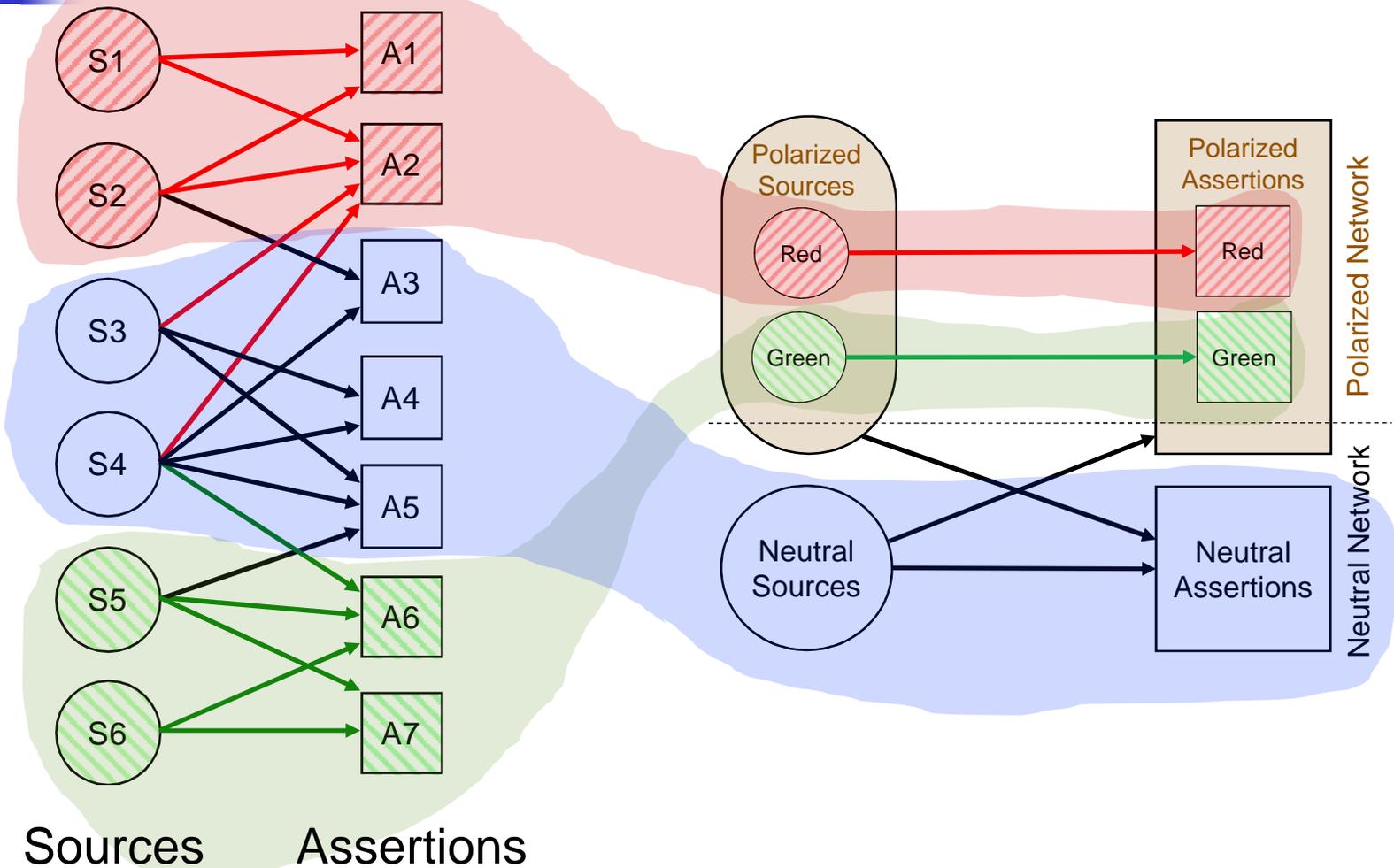
Matrix Factorization Approach

Minimize error of factorization

$$J = \sum_{(i,j) \in O} e_{ij}^2$$

$$= \sum_{(i,j) \in O} (a_{ij} - \sum_{q=1}^{k} u_{iq} v_{jq})^2$$

Account for the social network

Source i dependent on j

i's scores should be consistent with j's

Regularization

$$J = \sum_{(i,j) \in O} (a_{ij} - \sum_{q=1}^{k} u_{iq} v_{jq})^2$$
$$+ \sum_{i,j} \gamma t_{ij} ||\overline{u_i} - \overline{u_j}||^2 + \lambda(||U||_F^2 + ||V||_F^2)$$

Gradient descent

$$\frac{\partial J}{\partial u_{iq}} = 2 \sum_{j:(i,j) \in O} e_{ij}(-v_{jq})$$
$$+ 2 \sum_{m} (\gamma t_{im} + \gamma t_{mi})(u_{iq} - u_{mq}) + 2\lambda u_{iq}$$

$$\frac{\partial J}{\partial v_{jq}} = 2 \sum_{i:(i,j) \in O} e_{ij}(-u_{iq}) + 2\lambda v_{jq}$$

81

Quality of Classification
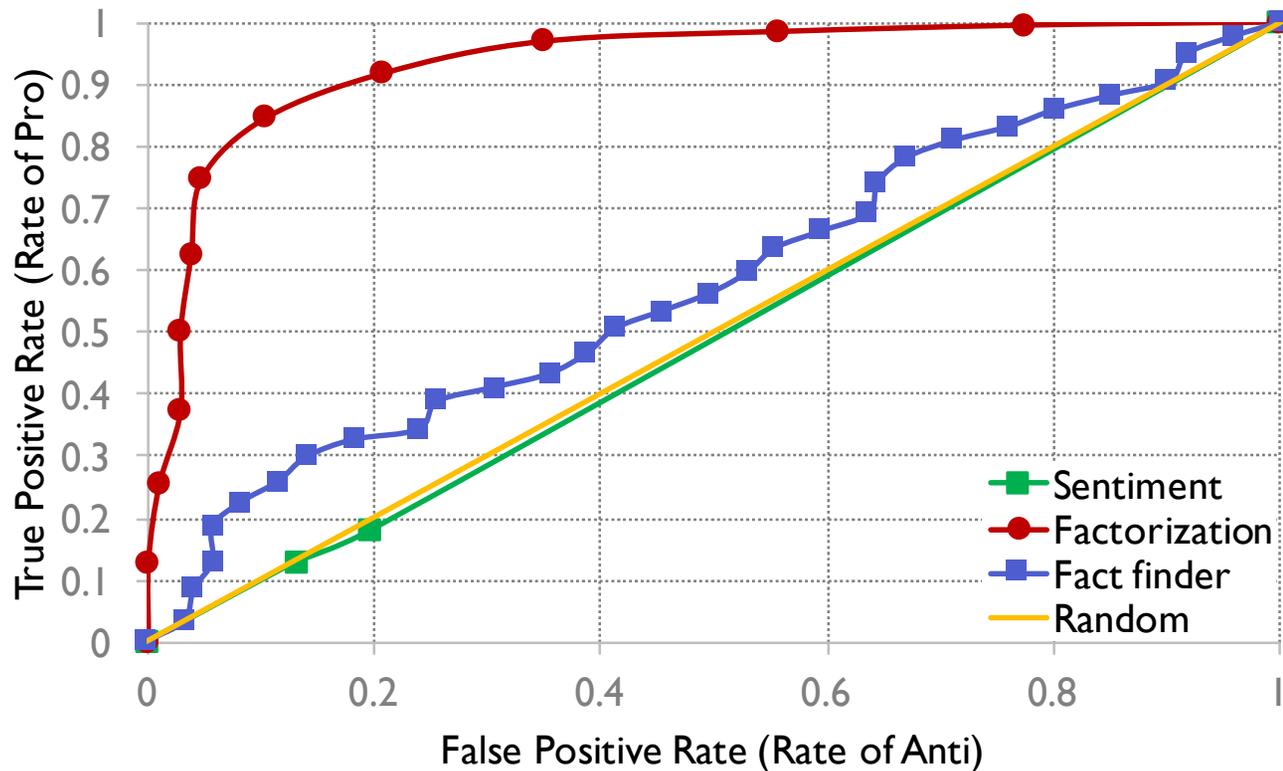
Receiver Operating Characteristics
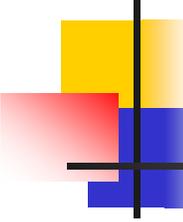
Pro or Anti?

Eurovision winner, Jamala, performed this song in Kiev a year ago. A year later it wins Eurovision.

Eurovision

| Pro Jamala	Anti Jamala
Incredible performance by #Jamala, giving Crimean Tatars, suffering persecution & abuse, reason to celebrate https://t.co/XWOZADrywH	@marcelsardo @jamala Next year #ESC2017 i represent Serbia,with song 1999, its not about @NATO bombing, its personal, and im hotter than her
Breaking: #Russia launches harassment campaign against #Jamala's @Twitter a/c. All known Kremlin trolls. @BBC_ua @AP https://t.co/btk9QyUpkH	Eurovision rules stipulate a song must be NEW (since Sept). @Jamala sang "1944" in May 2015 https://t.co/T8Pu21hPiM https://t.co/rwAFHHIllf
After deportation requiem, #Jamala explains how restrictions in/on #Crimea prevented parents fm joining her @ #ESC https://t.co/tMKIX1qNA9	Eliot, @EliotHiggins, Jamala openly performed that song at a concert in May 2015, it was published online. It's a rule violation #Eurovision
President awarded @jamala title of the People's Artist of Ukraine https://t.co/2df8J9zHP5	#Eurovision 10 mn 🎥 (English subs) by @anatoliisharii shows @Jamala singing 1944 song 5/2015 https://t.co/swA56P0MbF https://t.co/QzLPr0E43C
Russian coverage of Jamala's victory descends to the level of old Soviet anecdote https://t.co/HoNwJKdtCO via @EuromaidanPress	Jamala's "Ukrainian" Parents Prosper in Russian Crimea https://t.co/W5bZLj008H
Jamala's father: We do not talk with Russian journalists https://t.co/EEewVpZ9D2 https://t.co/51rcf8Je3K	"#CrimeanTatars" @Jamala now shows the worth of her patriotism and how much she REALLY cares https://t.co/KVnkhk3nGL https://t.co/MUNJh3AJH
Photo gallery: #Eurovision winner #Jamala arrives in #Kyiv https://t.co/MBKhuQ7W03 https://t.co/1eEHKAgsjB	#Ukraine's singer, Jamala, has performed her #Eurovision song a year ago, which is normally against the rules https://t.co/bxgeOK6X8D
Congratulations to Ukraine on winning #Eurovision 2016! @JAMALA wrote and composed her song '1944' by herself. ✊🏻 https://t.co/vZjYHvtoC	Crimean authorities have invited Jamala to an opening of monument dedicated to Crimean Tatar victims of deportation. https://t.co/oHaF7QGYQw
#Jamala sends everyone a postcard from home, #Ukraine https://t.co/l1eN6KPVfK	#BREAKING #Jamala is singing for #Russia-n agressors in #Sochi 2015 to earn money for #Eurovision #ESC2016 #Ukraine https://t.co/gnrH8lWLf4
Another thank you from #Jamala #Eurovision #CrimeaIsUkraine https://t.co/W8zOlDRnXn	So .@jamala wants to commemorate in Kiev not Crimea. Odd. In 1944 Crimea was (+ is now) part of Russia not Ukraine https://t.co/zTimmbDsKI
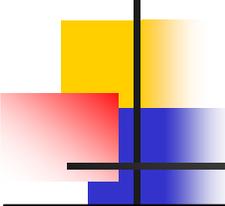
Donald Trump

| Pro Trump	Anti Trump
Retweet if you are 100 PERCENT voting for Donald Trump	Donald Trump said women should be punished for seeking an abortion. That's not a distraction—it's a disgrace. https://t.co/sbJ3opebyB
@realDonaldTrump Fugedaboudit!!! The woman in New York love Donald Trump!!! https://t.co/7yzgMHVzL4	At this point, Donald Trump has insulted the vast majority of Americans. The good news is, there's something we can all do about it: Vote.
Thank you, @NYPost! #Trump2016 https://t.co/KzGweIxaEo"	Read and sign this letter that people all over are signing to Donald Trump: https://t.co/S56QbW5K5C
The police are the most mistreated people in this country," Donald Trump #BlueLivesMatter #Trump2016 https://t.co/WfJvWUkMaB	We've earned more votes than any other candidate—Republican or Democrat. https://t.co/tRJNMj86AJ https://t.co/XJIt2bGevs
Nobody beats me on National Security. https://t.co/sCrj4Ha1I5	Study: Hillary Clinton, not Donald Trump, gets the most negative media coverage https://t.co/CyONOdTU0
Donald Trump #DonaldTrump #NewYork #NY #NYPrimary #RhodeIsland #Pennsylvania #NewYork4Trump #Delaware #CT #Maryland https://t.co/iYfo13Jjut	Donald Trump says wages are too high. (Yeah, you read that right.) https://t.co/up8ZI1WULC
Latinos For Donald Trump 2016 "Go Out & #VoteTrump" #LatinosForTrump #Hispanics4Trump #Trump2016 @realDonaldTrump https://t.co/eSNy170CXu	Happy to hear @realDonaldTrump accepted my challenge to debate one-on-one: https://t.co/mikc6fXZei
1987: Donald J. Trump Celebrated As Model Citizen in #NYC. Remember📺 TV without HD? #NYPrimary #MAGA #Trump2016 https://t.co/5ROvjhJyAK	I've released 9 years of tax returns. RT if you agree it's time for Donald Trump to release his! https://t.co/08whtFVC0r
The Post endorses Donald Trump https://t.co/bGIxG1DnZO https://t.co/1lC8E4Xi89	Donald Trump says wages are too high. Really? Hardworking Americans don't think so. https://t.co/5oEK9UhGI1 https://t.co/1z0tuCedJa
I'm a Veteran. I was born in Mexico, but I am here Legally! I am not racist! I support Donald Trump usus #Trump2016 https://t.co/pD076BcWU5	It's not just Trump: Every Republican presidential candidate has attacked women's health and rights. https://t.co/3TQdSvYTSs

Discussed So Far

Event Detection, Localization, Tracking

Fact-finding

Polarization Analysis
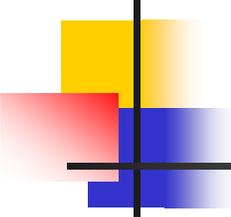
Anomaly Explanation
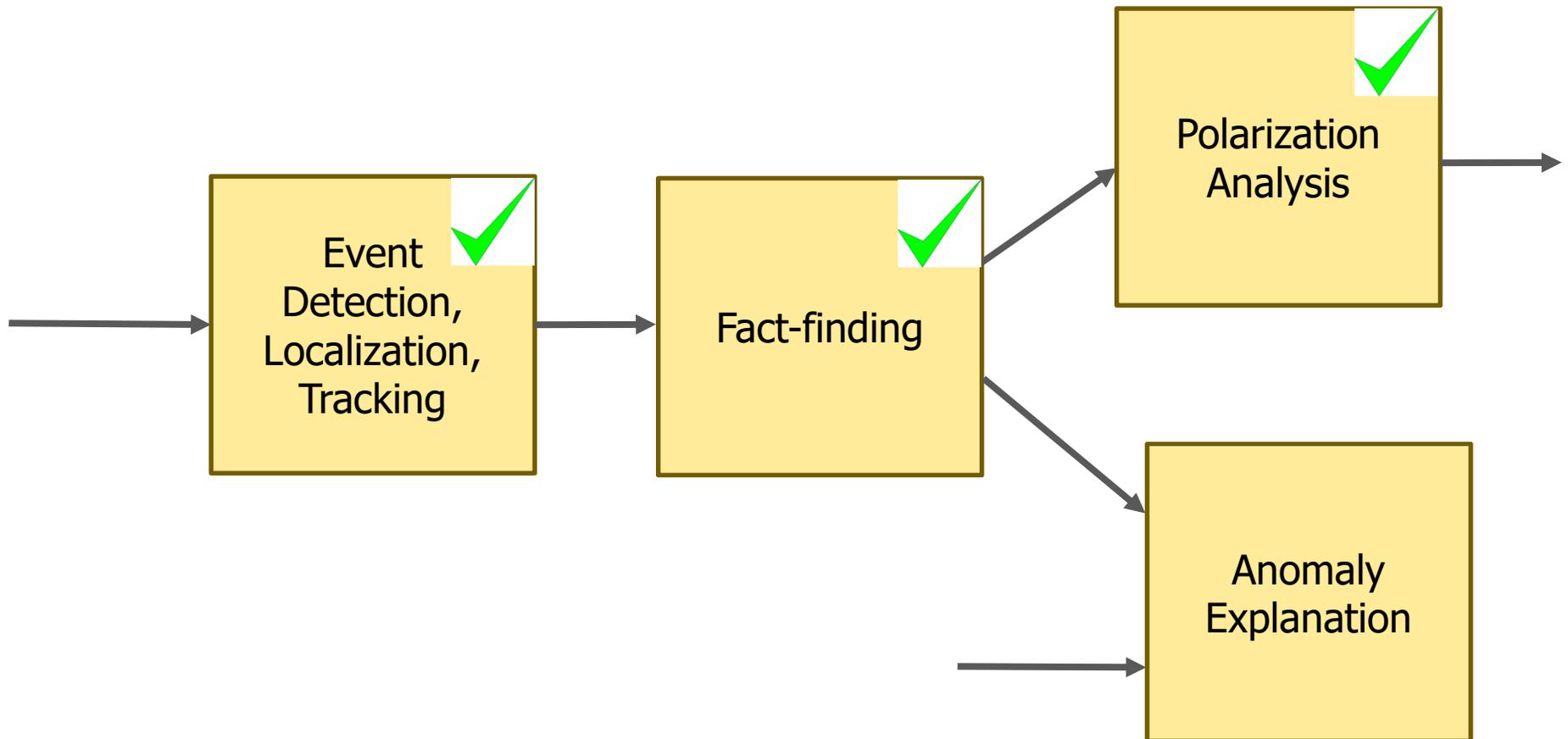
A Signal Processing Stack for Social Sensing

- **A layered social sensing architecture**

Physical World Events

| Application layer |
| Fusion layer ✔ |
| Logical bit-channel layer – "decoding the channel" |
| (Social) Signal layer – "tuning the radio" |
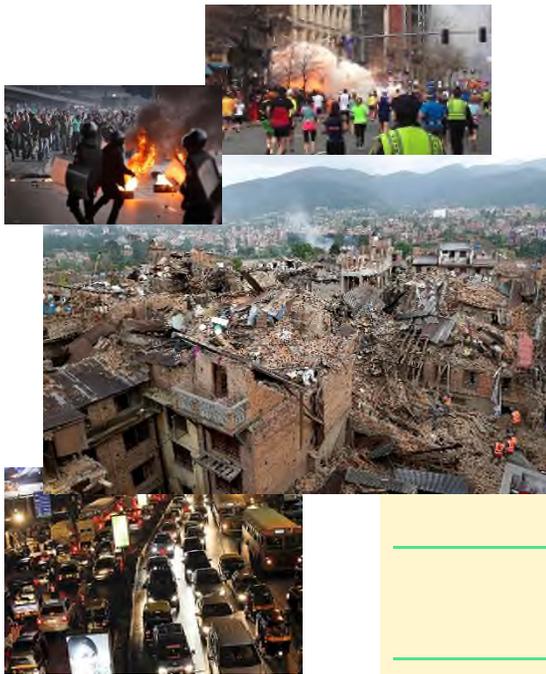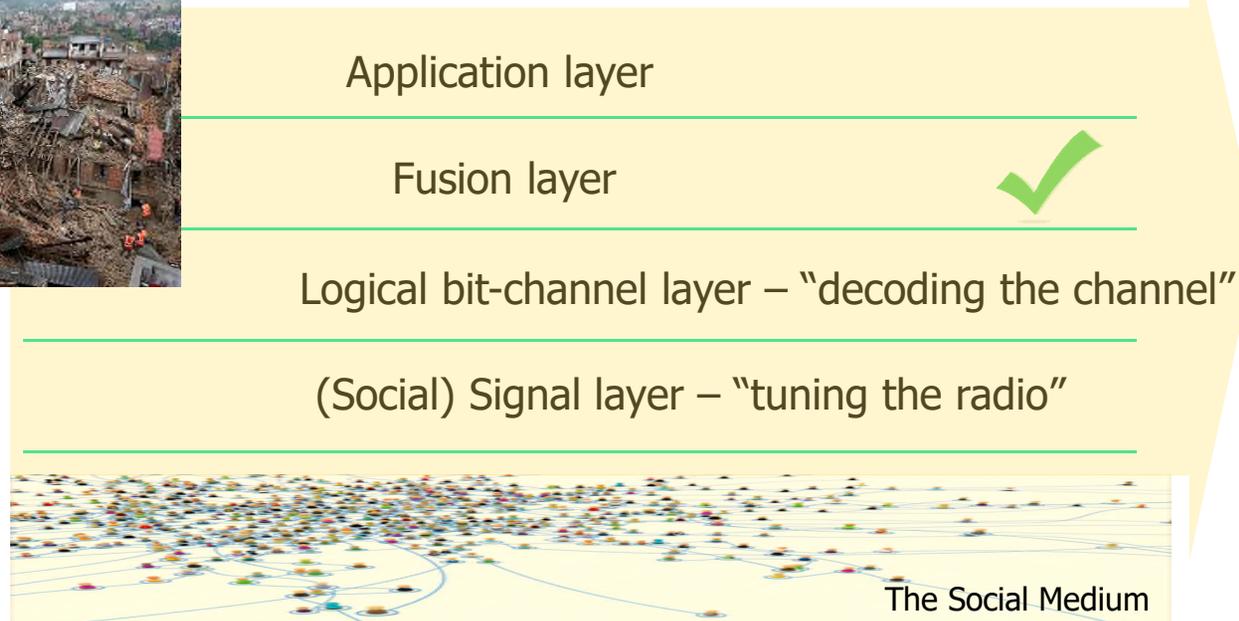
Receiver

The Social Medium
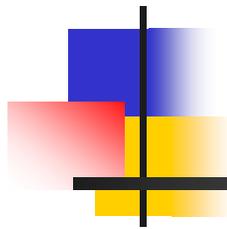
# Capability #4

Anomaly Explanation
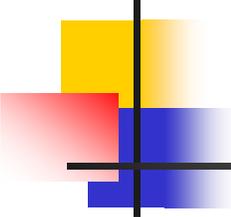
Insight

- Cued by anomaly seen by a sensor (or other source), find an event on social media that correlates with the anomaly's location and time. It will likely explain it!

A Different Use-case: Anomaly Explanation

- Explain anomalous traffic conditions on freeways in three California cities: Los Angeles, San Francisco, San Diego
  - Sensor data: Department of Transportation publishes freeway speed data (from thousands of sensors) once every 30 seconds
  - Twitter data: Collected California Twitter data with keyword "Traffic"

- Input: Location of sensor experiencing anomaly
- Output: Ranked list of tweets that comprise possible explanations.
  - Find events with specific sensor location references
 - Rank their tweets by information gain of underlying keyword pairs

Example 1: Forest Fire

Flow (Veh/5-Min)
12,720 Lane Points (98% Observed)
Mainline VDS 821709 - OAKIE FLATS - I15-N
Sun 08/11/2013 00:00:00 to Mon 08/19/2013 19:59:59
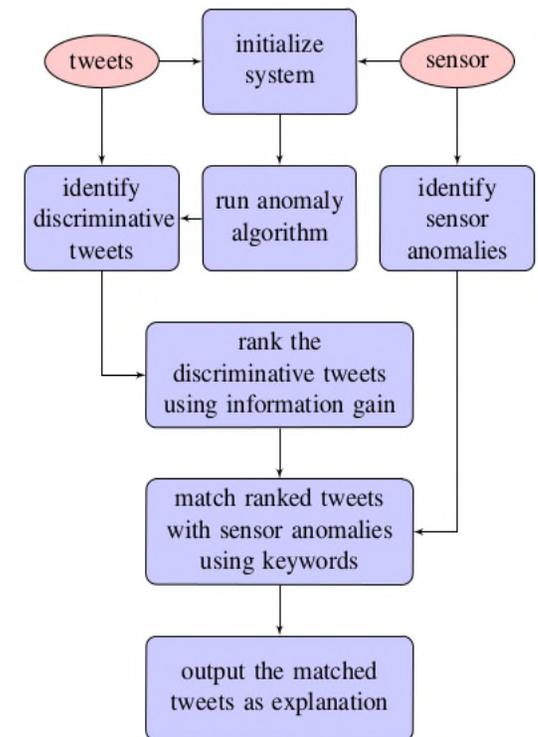
Time, Location:
6pm, Aug 17th, I-15 N, LA

**Event Signature:**
"Cajon", "Pass"

Possible Explanation (from Twitter):
Cleghorn Fire in **Cajon Pass** Snarls Traffic
on I-15: (KTLA) One northbound lane on
the 15 Freeway was reopened...
http:\\t.co\nieqh4nsMX
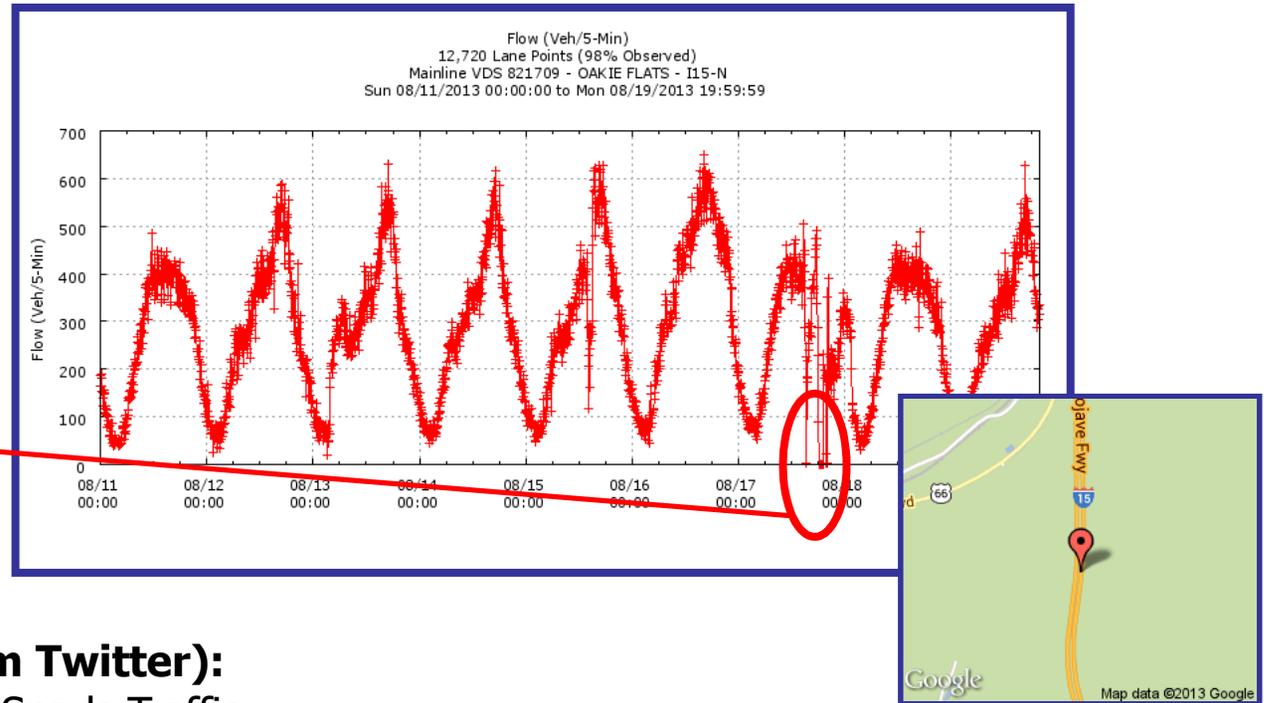
One northbound lane on the 15 Freeway was reopened Saturday evening in the Cajon Pass as firefighters continued to battle the so-called Cleghorn fire, authorities said. All southbound lanes on Interstate 15 were open, according to the U.S. Forest Service. State Road 138 remained closed from the 15 Freeway to Summit Valley Road.

URL:
http:\\t.co\nieqh4nsMX

Example 2: Pedestrian Death

Flow (Veh/5-Min)
2,544 Lane Points (90% Observed)
Mainline VDS 774358 - NORTH OF MILLER WAY - I710-S
Sun 08/18/2013 00:00:00 to Mon 08/26/2013 19:59:59
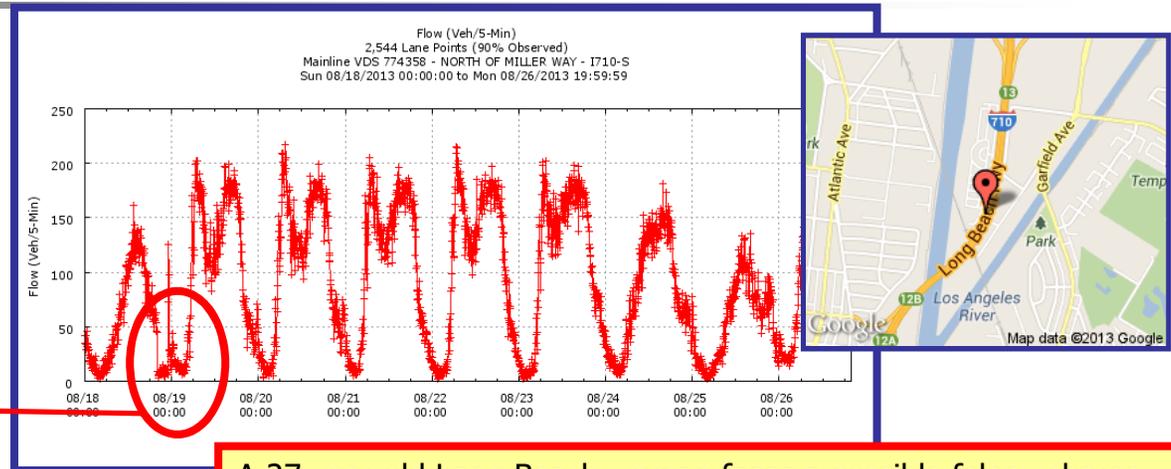
Time, Location:
8pm, Aug 18th, SB I-710, LA

Event Signature:
"Drunk", "Kills"

Possible Explanation (from Twitter):
27-Year-Old Drunk Driver Hits, Kills Man Trying To Stop Traffic On 710
http:\\t.co\rMoI7DxFH4

URL:
http:\\t.co\rMoI7DxFH4

A 27-year-old Long Beach woman faces a possible felony charge after fatally hitting a pedestrian on the 710 Freeway while she drove intoxicated.
Melanie Gosch struck a man in his 20s at about 8 p.m. on Sunday near Imperial Highway on the southbound 710 in South Gate, according to City News Service. The man, whose name has been withheld, was allegedly trying to stop traffic in the number three freeway lane.
Police responded to the scene after a caller reported a "long-haired man in dark clothing" on the freeway. Within a minute, the California Highway Patrol received a call that a pedestrian was lying in the freeway. The man was pronounced dead at the scene.
Gosch stopped her 2007 Nissan Sentra after striking the man, and she was arrested and booked on suspicion of causing injury or death while driving under the influence of alcohol or drugs.

Example 3: Vehicle Crash

Flow (Veh/5-Min)
9,264 Lane Points (79% Observed)
Mainline VDS 717254 - TOWNE - I10-E
Thu 08/15/2013 00:00:00 to Fri 08/23/2013 00:59:59

Time, Location:
8pm, Aug 19th, I-10E, LA

Event Signature:
"Crashes", "Divider"

Possible Explanation (from Twitter):
1 dead, 9 hurt in fiery **crashes** on 10 Fwy in Pomona; EB 10 closed, traffic allowed to pass along center **divider**
http:\\t.co\cCu8xLzx1U

URL:
http:\\t.co\cCu8xLzx1U
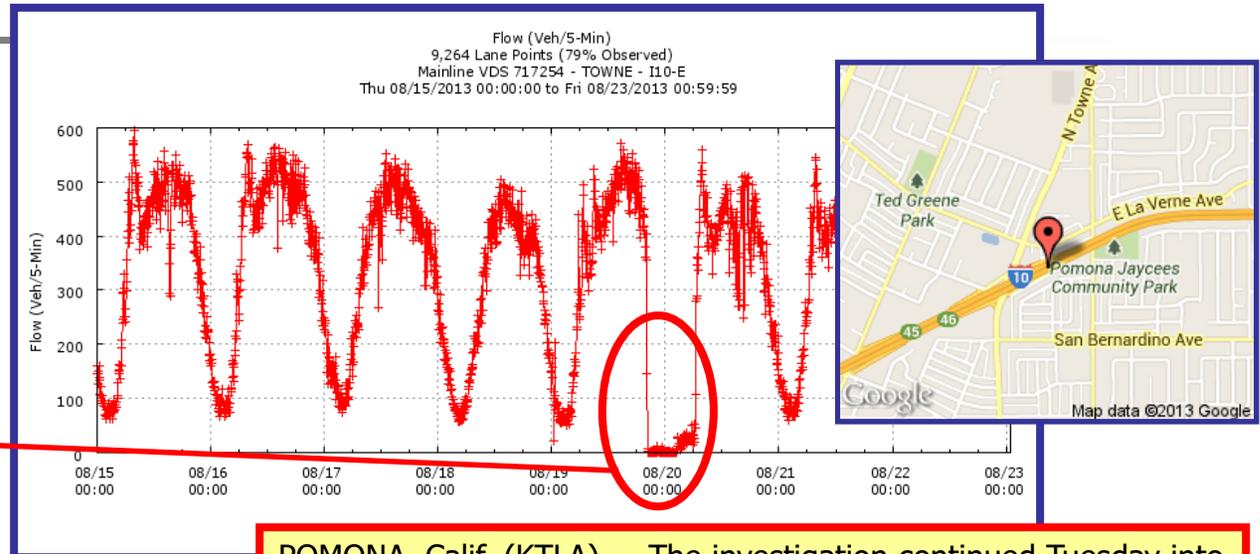
POMONA, Calif. (KTLA) — The investigation continued Tuesday into a pair of chain-reaction crashes on the 10 Freeway in Pomona that left one person dead and eight others injured. One killed, eight hurt in pair of chain-reaction crashes on 10 Freeway in Pomona. It all happened around 8 p.m. on Monday on the eastbound 10 Freeway near Towne Avenue, according to the California Highway Patrol. The first crash involved four cars and created a traffic back-up, CHP officials said. That's when a second crash occurred involving a big rig with a full tank of diesel fuel and three other vehicles. The fuel tank of the big rig ruptured, causing it to burst into flames, authorities said. The driver of the semi was able to get out safely. However, the driver of a red BMW that became trapped under the big rig was not able to escape. That person, who was not immediately identified, died at the scene.

Example 4: Vehicle Explosion

Time, Location:
8pm, Sept 1st , 405N, LA

Event Signature:
"Explosion", "lifeinLA"

Possible Explanation (from Twitter):
Heading to LAX. Traffic backed UP on 405
N due to an SUV explosion... #lifeinLA
#405parkinglot http://t.co/UBxUYMGrkS

URL:
http://t.co/UBxUYMGrkS
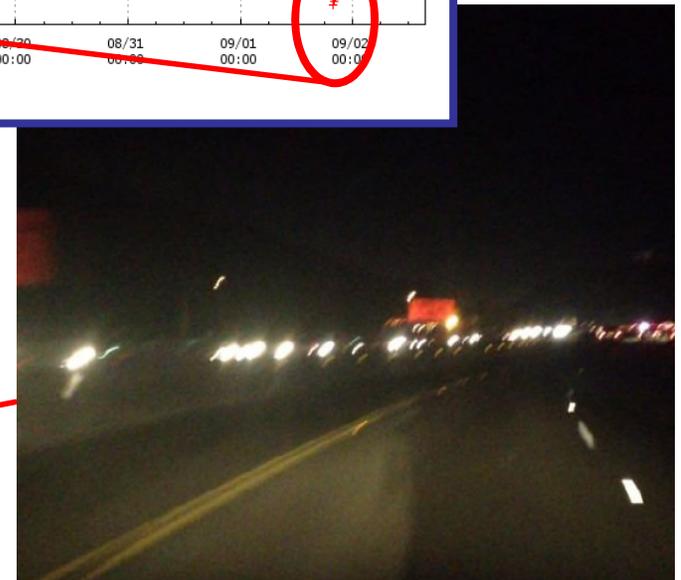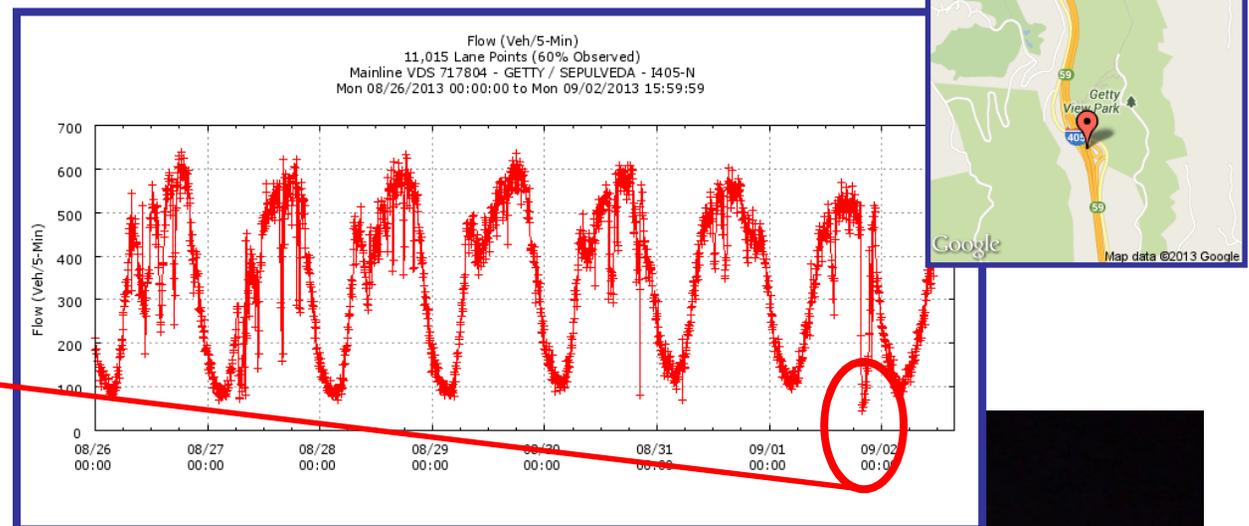
Results:
Anomaly Explanation Study

- Recall = Percentage of events explained in top 5 tweets.
    - Los Angeles: 83%
    - San Diego: 100%
    - San Francisco: 100%
  (for Hazard related events)
- Precision: Average rank of correct explanation of found traffic anomalies (after sorting and filtering).

	Information Gain (IG)	Spatial filter + IG	Credibility filter + IG	Spatial + Cred. + IG
Los Angeles	1.43	1.36	1.36	1.29
San Francisco	1.6	1.6	1.2	1
San Diego	1.14	1.07	1.21	1.2

Example: Anomaly Explanation
Fukushima leak (Feb 2014)

Peak observed in radiation sensor readings near the Fukushima Nuclear Plant.

Example: Anomaly Explanation
Fukushima leak (Feb 2014)

Peak observed in radiation sensor readings near the Fukushima Nuclear Plant.
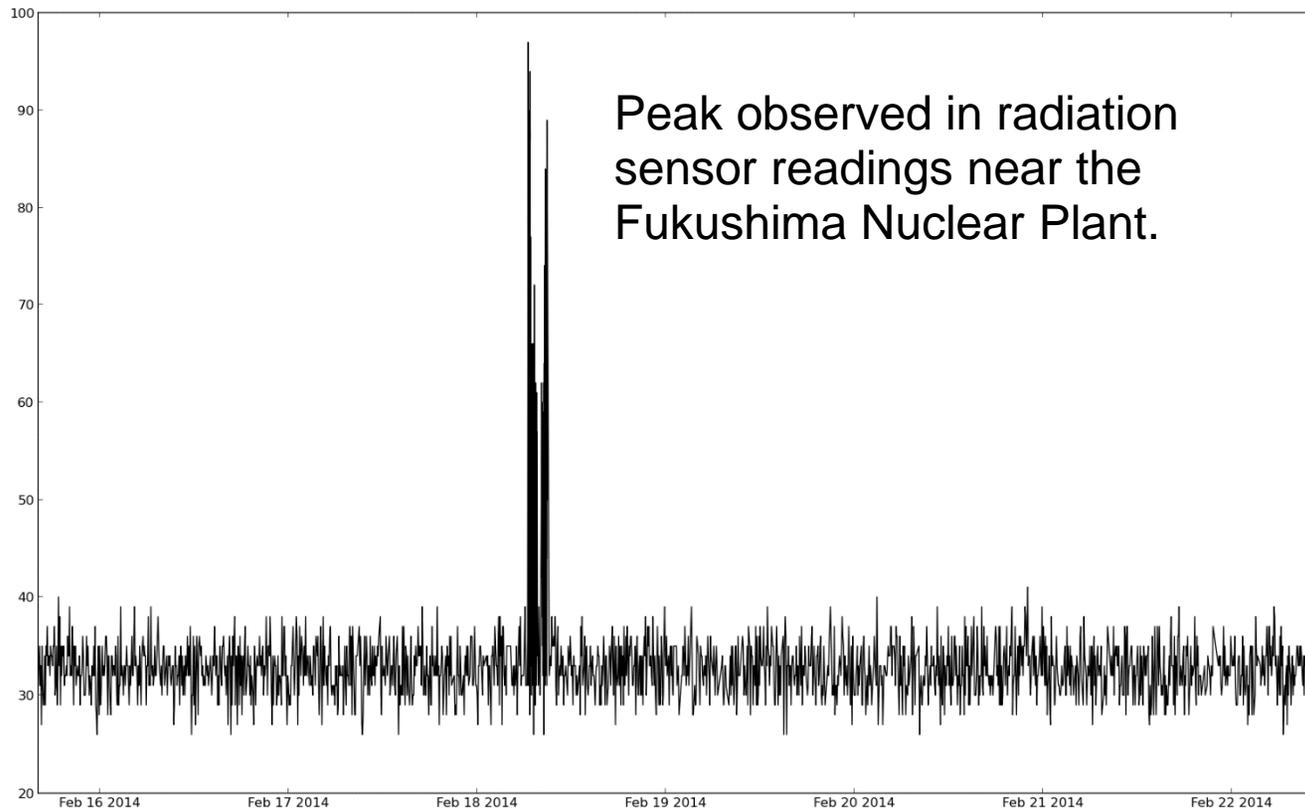
Feb 19, 2014

福島第一原子力発電所５号機タービン建屋地下１階における漏えい検知器の動作について | 東京電力 http://t.co/qbmpi49M2t

The operation of the leak detector in the first floor of Unit 5 the basement of the turbine building of Fukushima Daiichi Nuclear Power Station | TEPCO http://t.co/qbmpi49M2t
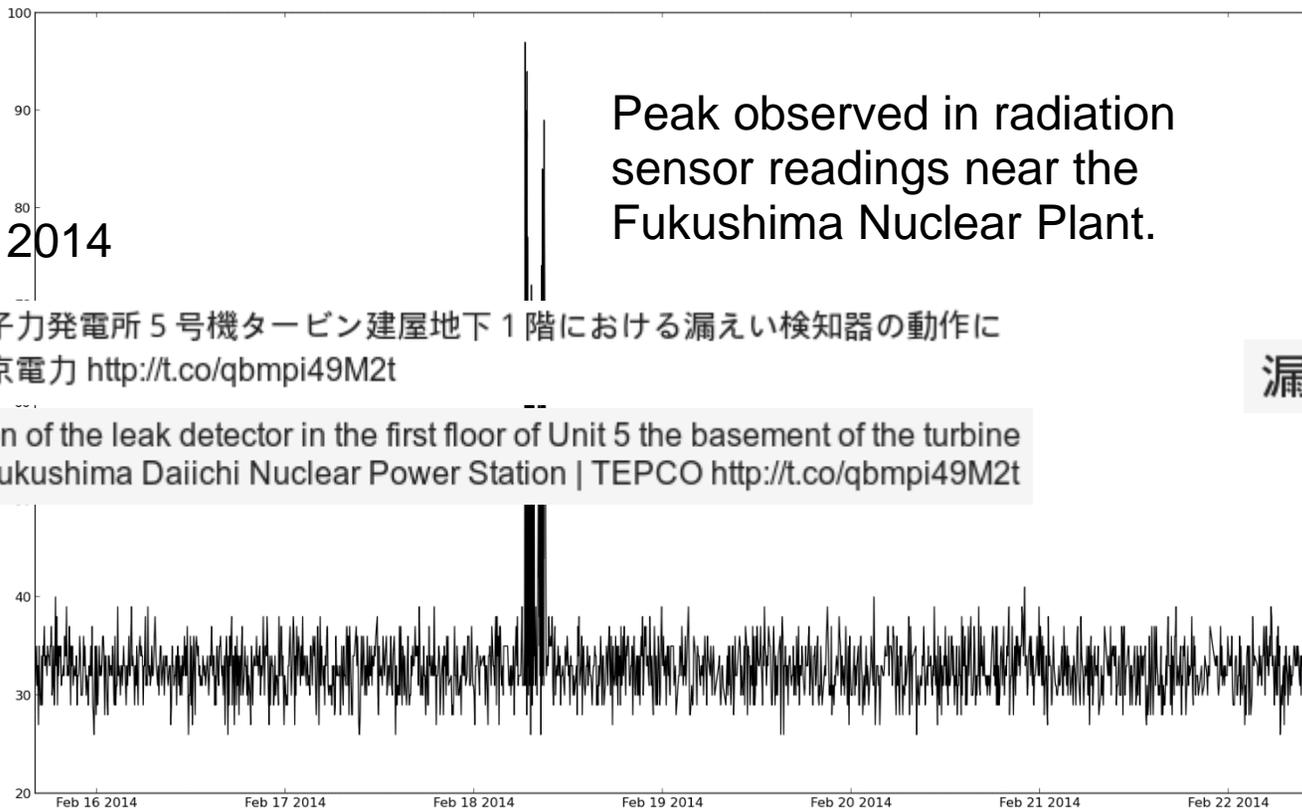
漏れ、建物
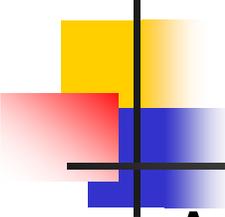
Conclusions

- An emerging computing area: Data services for decision support
  - Talk focused on social sensing
- Challenge: understand the social sensing modality (acoustic, magnetic, optical, and now social)
  - Physical events impact the social medium which responds with a signature
  - Analyze signature to measure the event
- Challenge: Accurately detect and track physical events (matching a user's interest)
- Challenge: Accurately assess data veracity
- Challenge: Automatically explain phenomena of interest
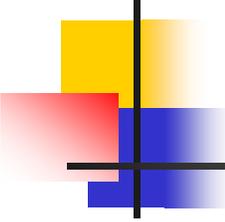
Tool Available Online

Link:

apollofactfinder.net

Book:
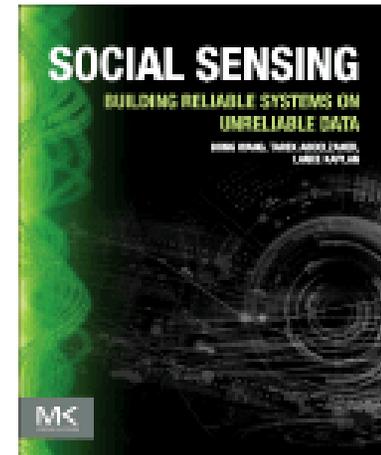
Social Sensing: Building Reliable Systems on Unreliable Data
Dong Wang, Tarek Abdelzaher, Lance Kaplan, *Morgan Kaufmann*, 1st
Edition, April, 2015