Some Challenges in Using Social Media for Disaster Response

Huan Liu

Data Mining and Machine Learning Lab
Arizona State University

Disasters

A crisis beyond the coping capabilities of a community

Natural	Man-made
Hurricane Matthew - 2016	Boston Marathon Bombings - 2013
	
https://goo.gl/AXY1Wi	https://goo.gl/PLOJY2

Social Media and Disasters

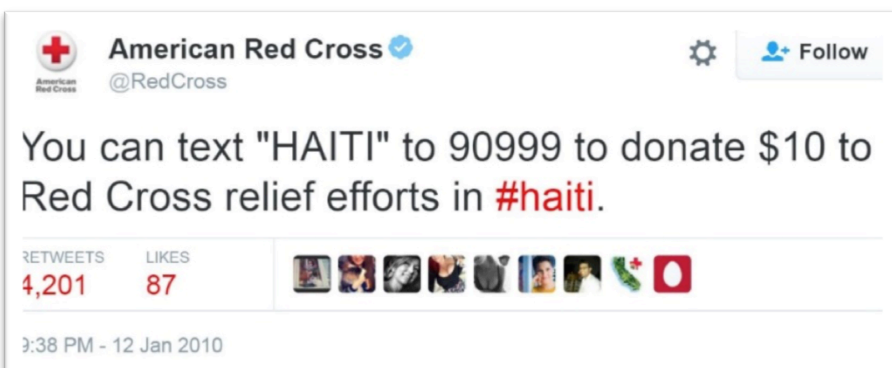
- Social media is a *new* and *important* source of dynamic information

Organizations

- Donations (Red Cross)
- Emergency phone numbers (FEMA, US State Dept, US Military)

Citizens

- 35% of citizens post requests for assistance on Twitter or Facebook*
- E.g., Hurricane Sandy



* Adam Crowe, "Disaster 2.0," CRC Press, 2012

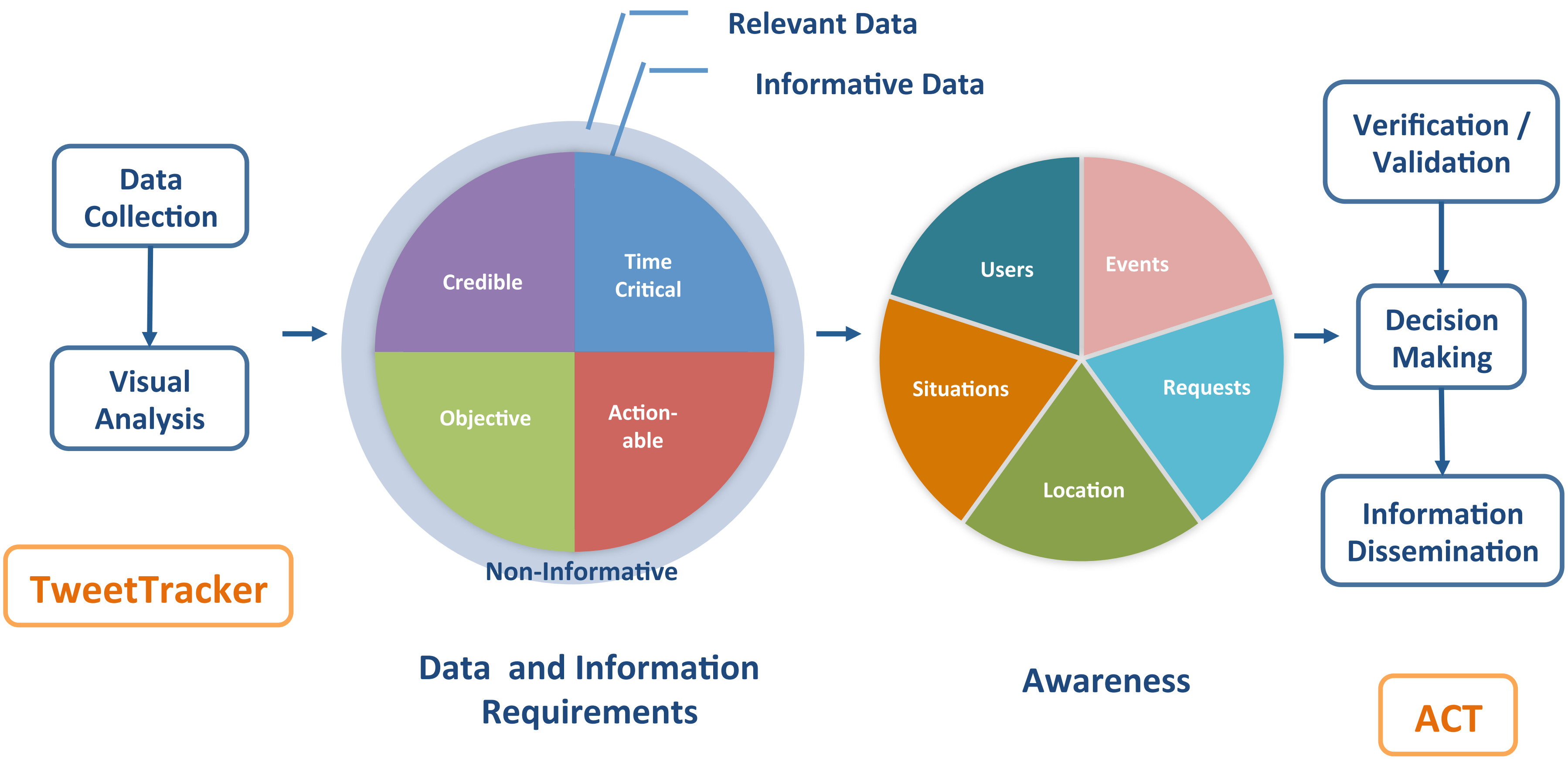
Social Media and Disasters

Citizens

- 35% of citizens post request for assistance on Twitter or Facebook*
- Social media as the primary source of information
  - London subway bombing*
  - Virginia Tech shootings**

* Adam Crowe, "Disaster 2.0," CRC Press, 2012
** http://articles.dailypress.com/2011-12-08/sports/dp-nws-tech-social-media-1209-20111208_1_twitter-and-facebook-twitter-website-power-of-social-media

A Process, Requirements, Tools



Data, Information, and Actions

Ensure Data to be Useful and Make Information Actionable

- Useful Data
  - Credible
  - Time-critical
  - Objective
- Actionable information requires Awareness
  - Events
  - Requests
  - Location
  - Situations
  - Users (responders, volunteers, victims, decision makers)

Credibility in Disaster Response

Definition

- Credible information is true or published by trustworthy sources

Motivations

- We want to learn about *people, true and relevant information*
- Social media is overwhelmed with bots & bot-generated content
- Users face the challenge of finding trustworthy sources of information
- False information or rumors can cause panic and distress

What we attempted

- <u>Bot Detection</u> [Morstatter ASONAM'16]
- <u>Trust Prediction</u> [Beigi SDM'16]
- <u>Rumor Detection</u> [Sampson CIKM'16]

Bot Detection

- Bots
  - Innocuous: relay information from official sources
  - Malicious: spread rumors and false information

- Goal: Remove bots from social media data with high Recall

- Challenges
  - Acquiring ground truth
  - Increasing Recall without significantly reducing Precision

Bots in Social Media

- Bots on Twitter:
  - Twitter claims 5% of 230M users are bots.
  - One study found 20M bot accounts = 9%**.
  - 24% of all tweets are generated by bots***.

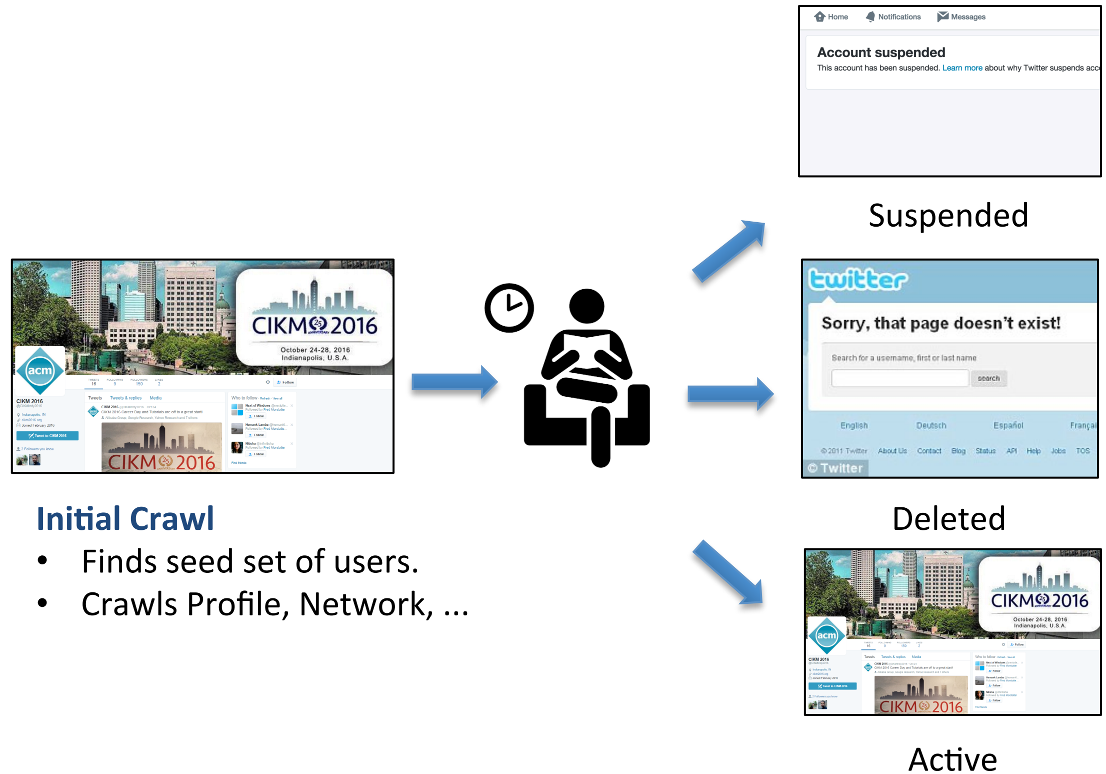
- 5-11% of Facebook accounts are fake****.

* http://blogs.wsj.com/digits/2014/03/21/new-report-spotlights-twitters-retention-problem/
** http://www.nbcnews.com/technology/1-10-twitter-accounts-fake-say-researchers-2D11655362
*** https://sysomos.com/inside-twitter/most-active-twitter-user-data
**** http://thenextweb.com/facebook/2014/02/03/facebook-estimates-5-5-11-2-accounts-fake/

Finding Ground Truth

Status on Twitter as a labeling mechanism

- Three states of a Twitter user:
  - Active
  - Suspended
  - Deleted
- **Idea:**
  - Use these states as labels
  - Two snapshots of each user is taken

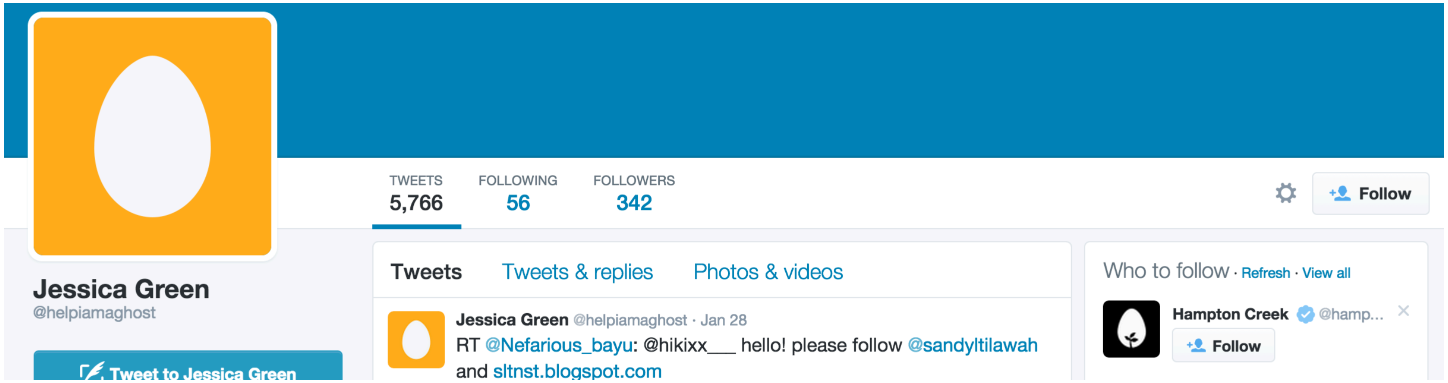


Initial Crawl

- Finds seed set of users.
- Crawls Profile, Network, ...

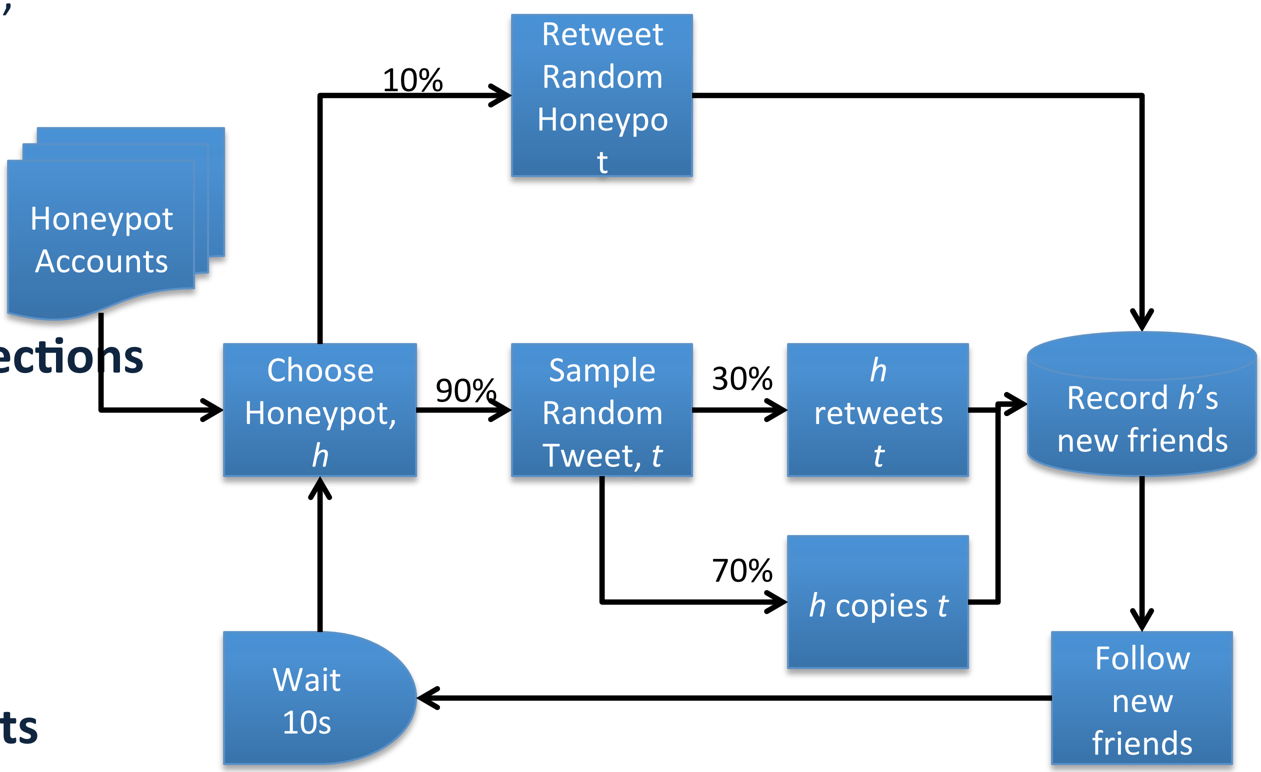
Ground Truth - Honeypots

- Act as obvious bot accounts
- Attract other bot accounts
- Bots are identified when they follow our account
- **Assumption:** Real users do not follow bots



Honeypots - Logic

- **Post "Luring" Content**
  - Post content that will be seen
  - trending topics, hashtags, "famous" tweets...

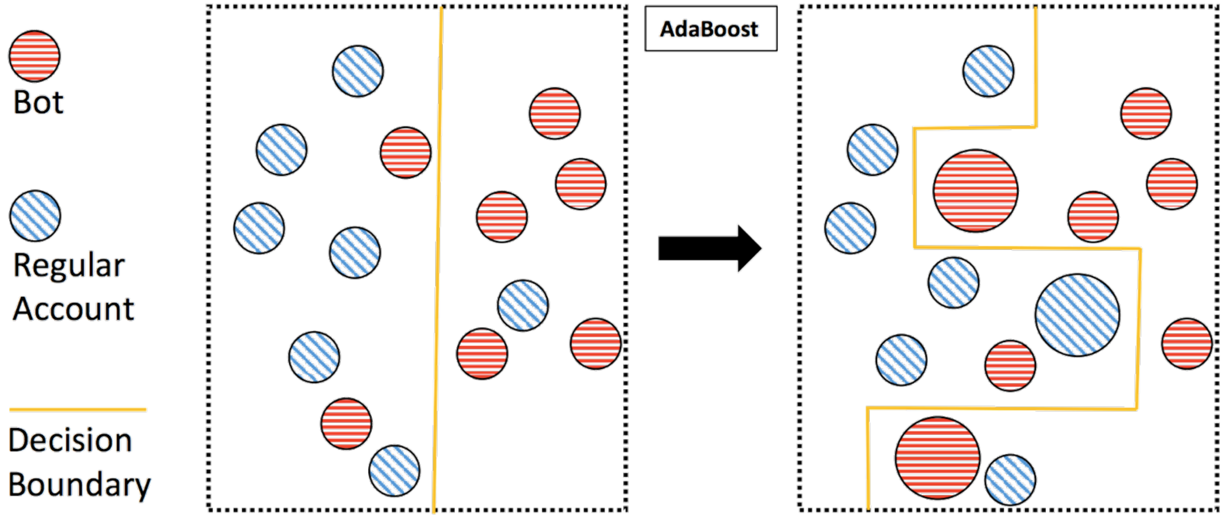


- **Maintain Network Connections**
  - "Follow back", Retweets
  - Fame begets fame
- **Promote Other Honeypots**
  - Retweet each other's tweets
  - Mention each other

AdaBoost Review

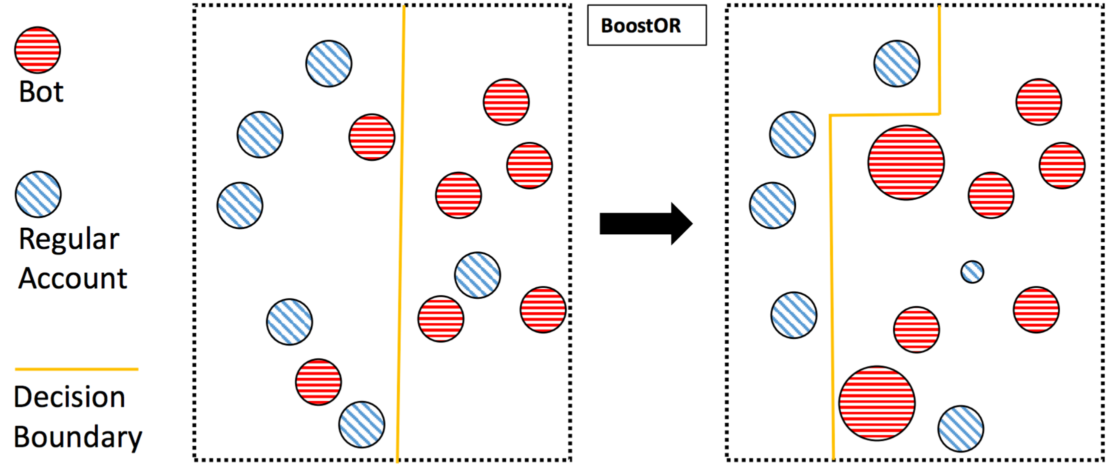
Two types of weights:

- Weights on weak learners.
- Weights on individual instances.



BoostOR

- Based on AdaBoost
- Try to increase Recall without drastic decrease in Precision
- Iteratively update the weight of instances:
  - Unchanged
    if correctly classified
  - Decreased
    if false negative
  - Increased
    if false positive



Trust Prediction

- Goal
  - Trust and distrust relations play an important role in helping online users collect reliable information
  - Finding trustworthy users and reliable information is of significant importance in aftermath of disasters
  - How to predict trust relations between users?

- Challenges
  - Trust relations are extremely sparse
  - Distrust relations are even much sparser than trust relations
  - What are strong indicators of trust/distrust?

Trust and Emotions

- According to psychologists, user's emotions are strong indicators of trust and distrust relations
- Emotional information is more available than trust/distrust
- There exists a correlation between emotions and trust/distrust relations

Modeling Emotional Information

- Users with positive (negative) emotions are more likely to establish trust (distrust) relations
- Users with high positive (negative) emotion strengths are more likely to establish trust (distrust)
- The Emotional Trust Distrust framework ETD
  - Low-rank matrix factorization
  - Emotional information regularization

Rumor Detection

- *Rumor*: unverified and relevant information that circulates in the context of ambiguity.

- Goal: detecting emerging rumors with minimum information as early as possible

- Challenges:
  - How to overcome the lack of information in a single tweet?
  - How to detect rumors in their formative stage?

Rumors on Social Media

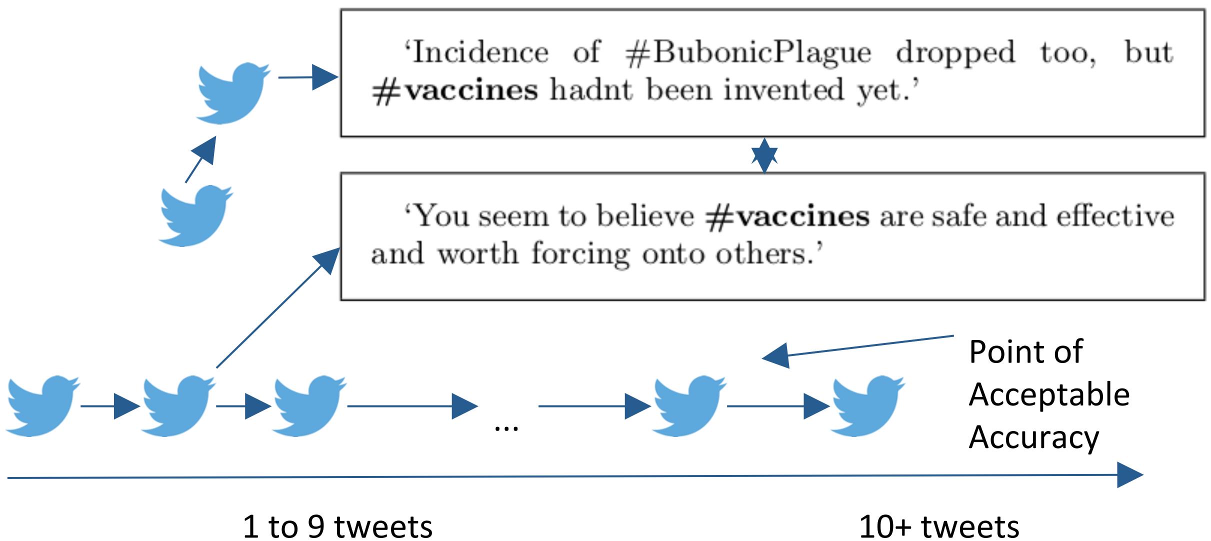
- False rumor propagated by AP hijacked account regarding a bombing in the White House
- Caused significant drop in Dow Jones Industrial Average (2013)

- More than ten thousand tweets contained fake images circulated after hurricane Sandy (2012)

Insufficient Information in a Single Tweet

- Not enough data in a single tweet
- Treat batches of tweets as "conversations"
  - Based on keyword similarities
  - Based on reply chains

- Aggregate conversations
  - Shared hashtags
  - Common links
  - Cosine similarity



Detection of Emerging Rumors

- Emergent detection - link the first tweet in a rumor with those already posted
- Standard rumor classifications are not effective for small conversations
  - Lack of network and statistical data
  - Data sparsity issues.
- Implicit linking works effectively for detecting small rumor cascades

Time-Critical Information in Disaster Response

Motivations

- Social media is used to request for immediate assistance during disasters
- Time-critical posts demand immediate attention
- Addressing these queries promptly can help in emergency response
- How can these posts be distinguished from others?

What we attempted

- Finding Time-Critical Responses [Ranganath ICDM'15]

Finding Time-Critical Responses

- Many questions asked during disasters should be immediately attended
- Many responders are busy
- How can we find a prompt responder who can provide a relevant answer?

- Challenges of Identifying Prompt Responders
  - How do we estimate the *reply time* of users to identify prompt responders?
  - Timeliness and relevance: how do we integrate timeliness with relevance to rank candidate respoders?

Information Seeking in Social Media

- Social media is used to request for help during disasters
- Addressing these queries promptly can help in emergency response

Identifying Candidate Responders

- Timeliness
  - The user can respond more quickly if she is available soon after the question is posted. It can be estimated using the previous posting times.
  - A user responds to questions faster if she has replied promptly to similar questions in the past.
- Relevance
  - Users whose previous content is similar to the question have higher relevance and their response is more likely to be a relevant answer.
- Timeliness and relevance are integrated by combining the ranking scores.

Finding *Objective* Posts in Disasters

- Objective posts convey a piece of unbiased factual information without opinions, chatters, or emotions
- Social media can contain opinions or intrinsic *sample bias*
- For sample bias, we ask
  - Is this sample representative of the whole population?
  - Can this sample be manipulated by malicious users?

What we attempted

- Twitter's Streaming API vs. Twitter's Firehose [Morstatter ICWSM'13]
- Bias in Twitter's Streaming API [Morstatter WWW'14]

Studying Bias in Social Media Data

- Twitter shares its data.
  - "Firehose" feed - 100% - costly
  - "Streaming API" feed - 1% - free
- We usually obtain data via sampling
  - Is the sampled data from the Streaming API representative of the true activity on Twitter's Firehose?

- Challenges
  - How to determine if the sample is biased when we do not have access to the whole data?
  - How to obtain an unbiased sample?

Twitter's Streaming API vs. Firehose

- Data from Firehose and Streaming API has been collected for specific period of time to perform analysis.
- More than 90% of all geotagged tweets are available via Streaming API and there is not significant difference in location distribution.
- Based on in-degree centrality and betweenness centrality in user-user retweet networks, the Streaming API finds ~50% of the key users.

Mitigating Bias in Twitter's Streaming API

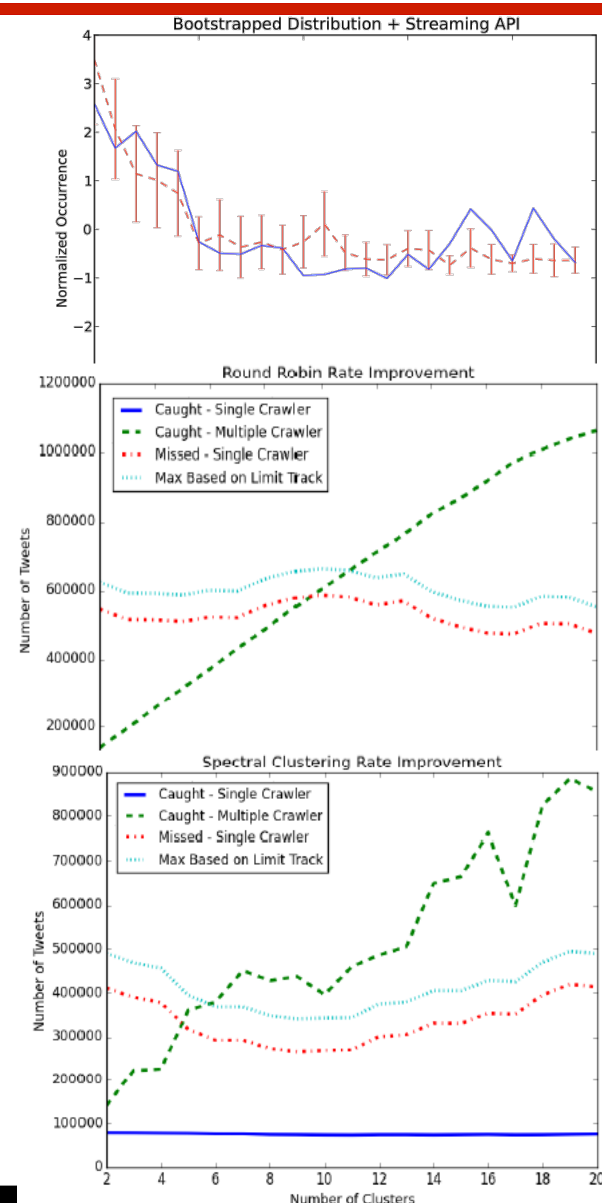
Can we find bias without the Firehose?

Estimating Bias from Streaming API:

- Obtain trend of hashtag from Sample API and Streaming API
- Bootstrap Sample API to obtain confidence intervals
- Mark regions where Streaming API is outside of confidence intervals

Mitigating Bias:

- Leverage multiple crawlers to maximize data for each query
- Round Robin Splitting



Users in Disaster Response

Motivations

- Relief campaigns require volunteers to assist with disaster relief
- Mixed with daily chatter, opinions, and sympathies on social media
- Which properties can be used to distinguish potential campaigners from others posting on the same topic

What we attempted

- <u>Detecting Advocates for Campaigns </u>[Ranganath WSDM'16]

Detecting Advocates

- Advocates are individuals who use social media to strategically advance their agenda for a given campaign
- Goal: identifying advocates from random users using social mobilization theories
- Challenges
  - Nuanced strategies to shape user opinions
  - Diverse strategies
    - Individual activities like constructing persuasive messages
    - Relational patterns like shared language and interactions
  - Scalability
    - Millions of users post on a given issue
    - Obtaining labels is hard

Detection based on Social Mobilization Theories

- Social mobilization theories
  - Message Strategies: Persuasive language, Topical focus
  - Propagation Strategies: Targeting popular users
  - Community Structure : Social and Interaction networks
- These features are used in a Logistic Regression classifier

Factors	Features	Elections	Gun Rights
Message Strategies	Persuasion	1.71**	1.04***
	Focus	-3.63*	-0.21**
Propagation Strategies	Targeting	2.15***	1.91*
	Hubs	2.91***	-0.31
	Authorities	6.39***	0.51*
Community Structure	Following	1.56***	0.90***
	Followers	2.04***	0.95***
	Interactions	1.44***	0.01**

Event Detection for Disaster Response

Definition of an Event

- Non-trivial incidents happening in a certain place at a certain time

Motivation

- Exogenous events cause changes in social media activities
- These changes are useful for early event detection
- Early detection can accelerate relief process

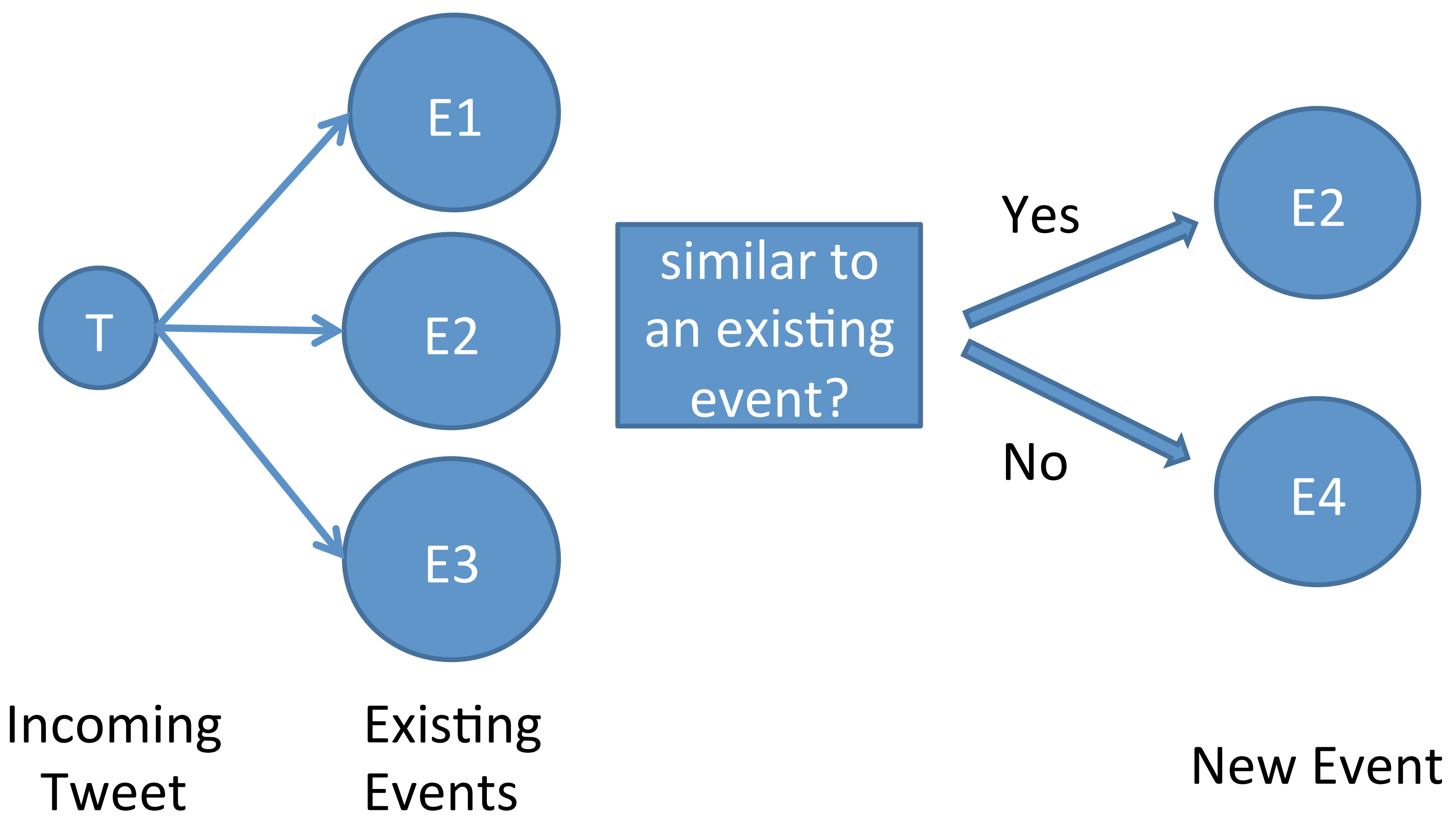
What we attempted

- Event Detection in Twitter Streams [Kumar ASONAM'15]
- Crisis-Mapping Using Language Distribution [Sampson ICDM'15, Demo]

Event Detection in Twitter Streams

- Goal: Given a stream of tweets T = t1, t2, t3, ...,
  detect events E = e1, e2, e3, ...
- Challenges
  - Informal language
  - Rapidly evolving events
  - User diversity
  - Handling streaming data
- Find events *E* such that
  - Constituent tweets are similar
  - User diversity is maintained
  - Distance between events is maximized

An Intuitive Clustering Approach for ED



Crisis-Mapping Using Language Distribution

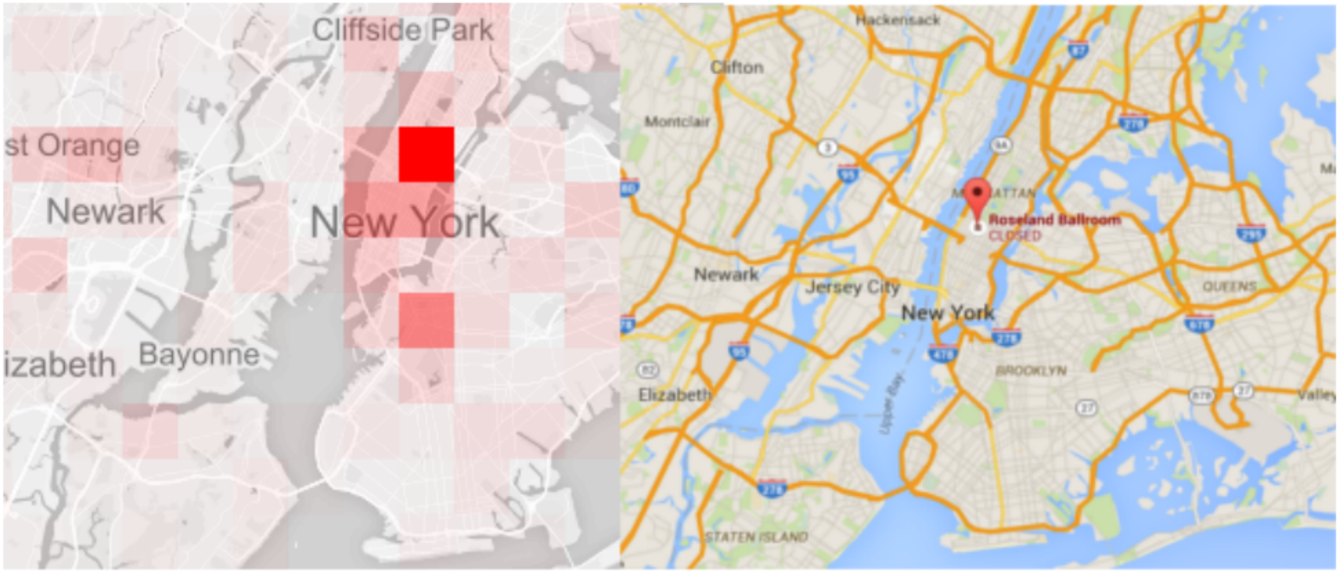
- Language distribution is probability distribution of words in the vocabulary

- Goal: Detecting events using changes in the regional language

- Challenges:
  - How to model the language of a region?
  - How to measure language change?
  - What amount of change can indicate the occurrence of an event?

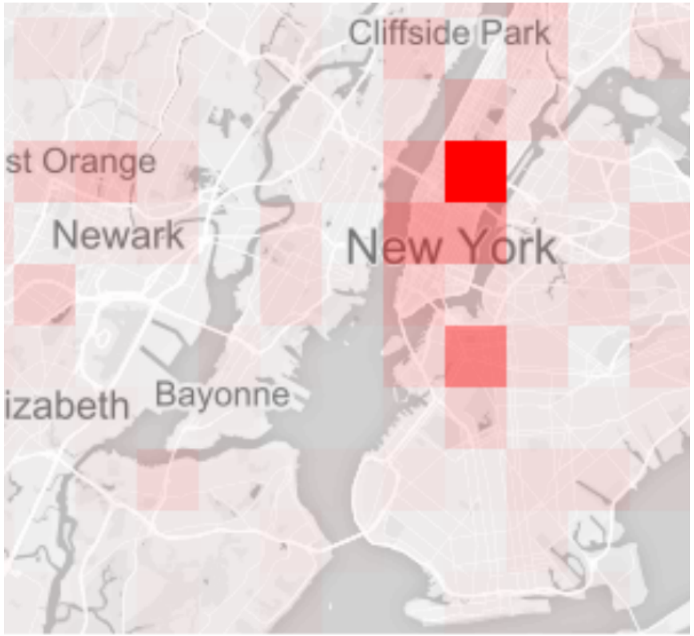
Macro-level Divergence Mapping

- Compare the distribution of a specific hour in a region to its expected distribution
- Spikes in the divergence of the language model indicate events
- Divergence is measured using Jensen-Shannon divergence



Micro-level Surprise

- Baseline: average of previous distributions for a region
- The language distribution in the same region for a desired period of time is calculated
- Subtraction of the two distributions shows the surprise level of each word in the vocabulary
- Provides insight for the detected event



An small event detected after Hurrican Sandy
Most surprising words: reseland, pick, ballroom, grace, okc, pic, times, show, num, #gracejones

Questions and Disaster Response

Motivations

- There are many social media sites
- Each site is designed for a specific purpose
- Difficult for users to choose the one to search for information

What we attempted

- Social Answer System [Dani ICDM'15, Demo]

Social Answer System

- Goal: Given a question and set of social media sites, rank the sites based on the probability of answering the question

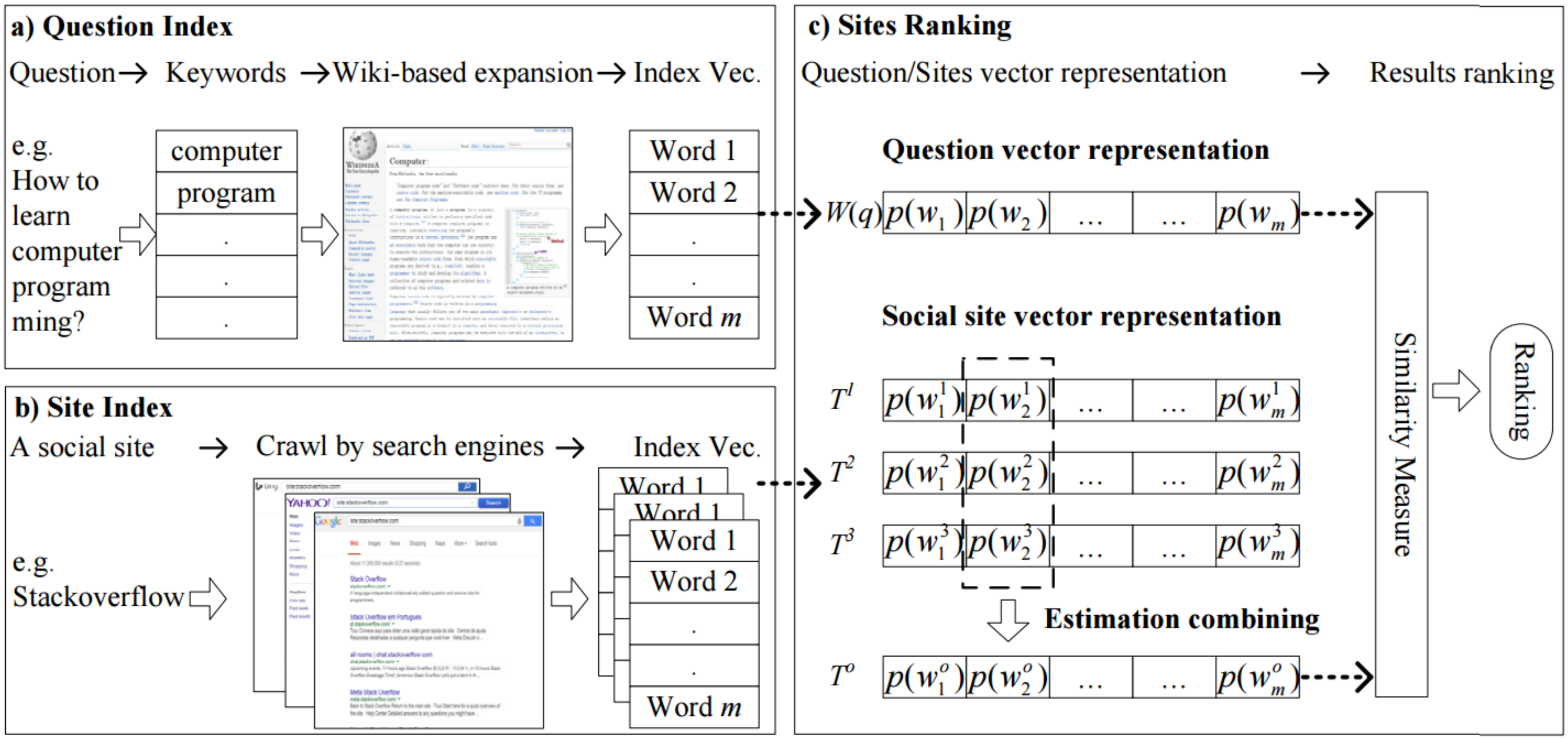
- Challenges
  - Understanding the topics of the question and topics discussed on social media sites
  - Finding similarity between the question and discussions

Social Answer System Overview

a) Query expansion using Wikipedia (question Index)

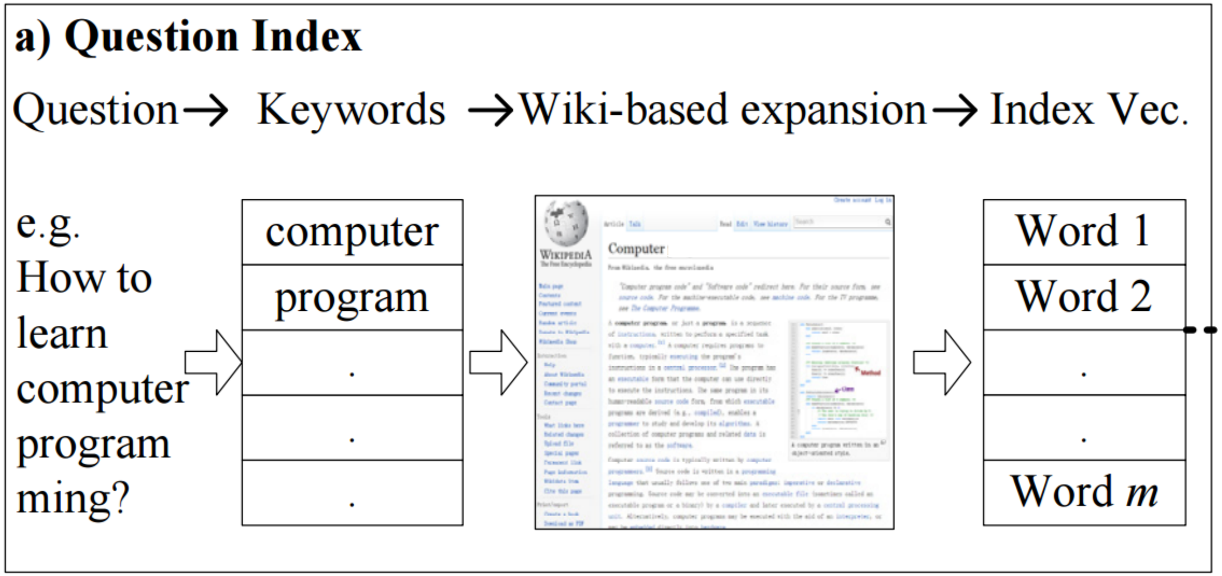
b) Search engine crawling (site index)

c) Question and search engine vector similarity site ranking)



Query Expansion Using Wikipedia

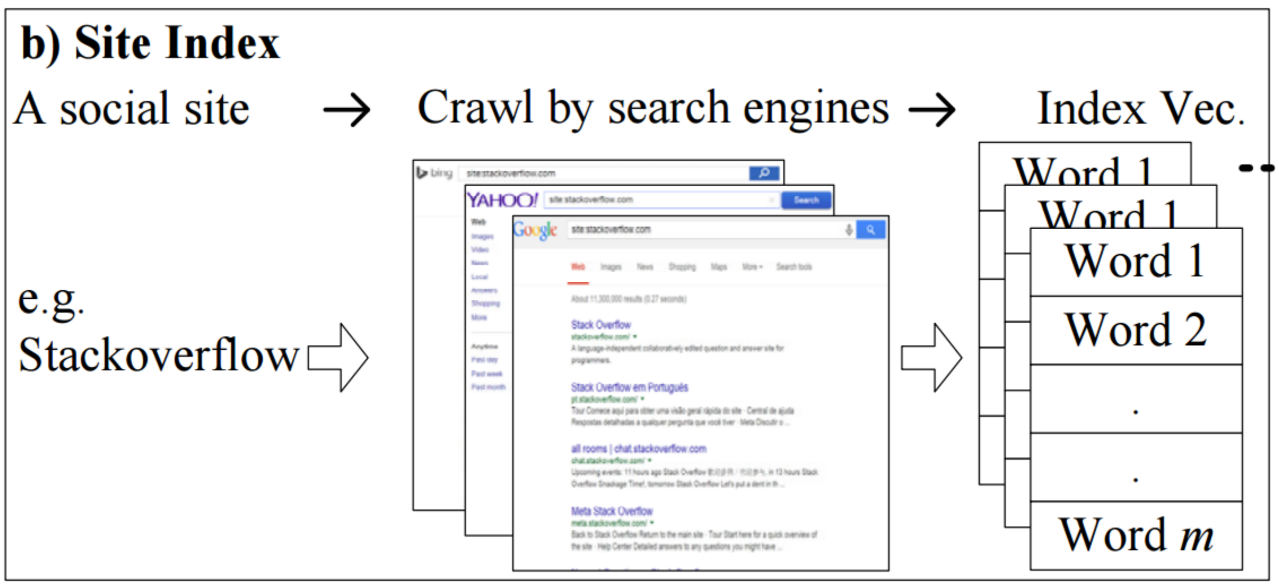
- Extract keywords from the question
- Obtain Wikipedia articles for each keyword
- Stem and remove stopwords from documents
- Send the aggregate of the documents as the new query



Search Engine Crawling

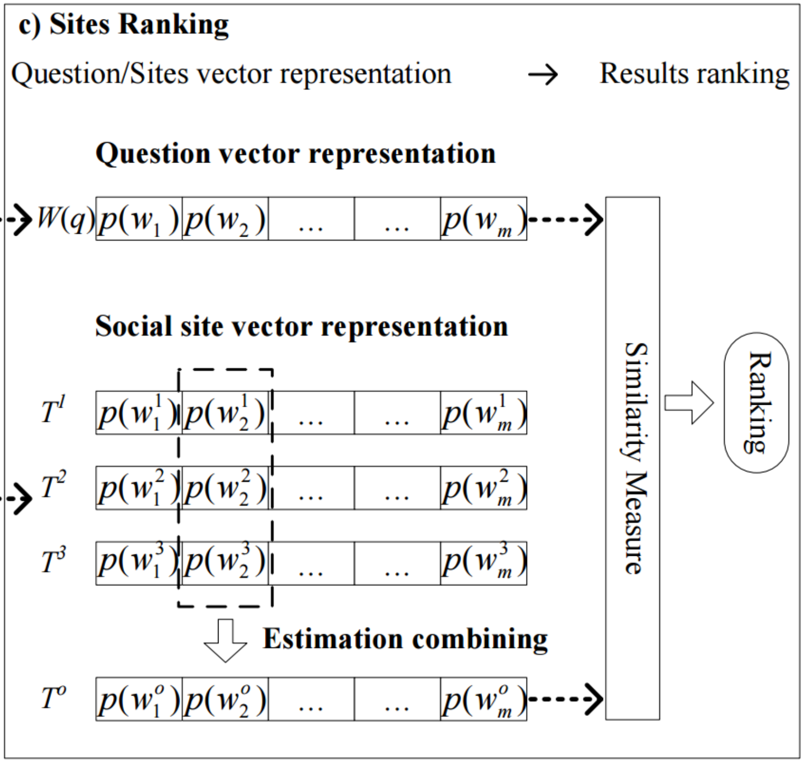
- 25 social media sites
- Find top ten pages of each site
- Using search results of Google, Yahoo, and Bing

Category	Site (Alexa rank)
Blogs	Wordpress(26), Blogger(53)
Microblogs	Twitter(7)
Social Networks	Facebook(2)
Professional Networks	Linkedin(12), Xing(908), Viadeo(1838)
Media Content Sharing	Youtube(3), Pinterest(25), Instagram(30) Tumblr(39), Imgur(49), Flickr(103)
Collaborative Knowledge Base	Wikipedia(6)
Social filter	Reddit(50), Yelp(125)
Collaborative Q&A	Answers(200), StackOverflow(45)



Question and Search Engine Vector Similarity

- Estimate the frequency of each question word in the content of sites
- Combine estimates from different search engines using conflict allocation algorithm
- Use cosine similarity measure to obtain the similarity score between query and each site



Examples

How can I perform XML Processing in Python? Go!

1. Stackoverflow , http://www.stackoverflow.com , 0.084778
2. Lastfm , http://www.last.fm , 0.055786
3. Wikipedia , http://en.wikipedia.org , 0.052185
4. Wordpress , http://www.wordpress.com , 0.048706
5. Tumblr , http://www.tumblr.com , 0.048645

How long to stay in Las Vegas? Go!

1. Tripadvisor , http://www.tripadvisor.com , 0.057861
2. Blogger , http://www.blogger.com , 0.045624
3. Answers , http://www.answers.com , 0.04364
4. Yelp , http://www.yelp.com , 0.03949
5. Viadeo , http://www.viadeo.com , 0.038574

Locations and Disaster Response

Motivations

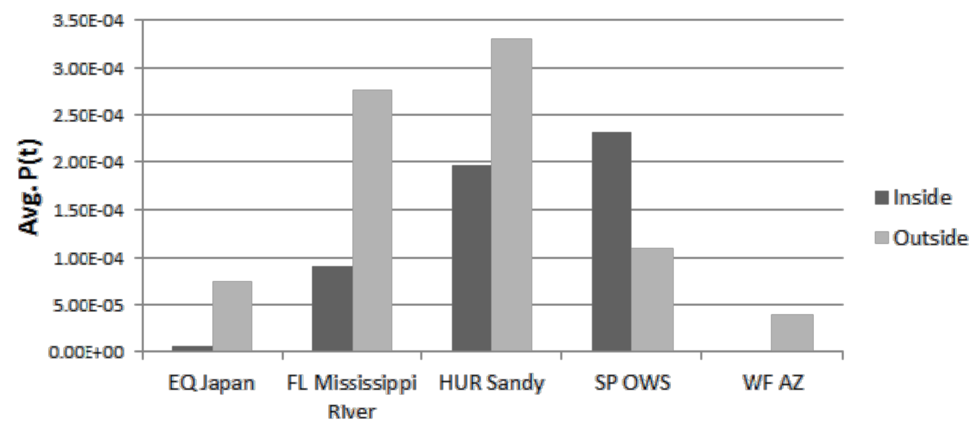
- Situational Awareness
  - Assess the impact of a crisis
  - Identify severely impacted regions
- Disaster response
  - Coordinate rescue efforts
  - Disseminate information to people
- Prioritizing information processing for faster response

What we attempted

- Identifying Tweets from Crisis Regions [Kumar HT'14]

Identifying Tweets from Crisis Regions

- Goal: identifying tweets generated from crisis regions based on the characteristics that discriminate them from the ones which are published elsewhere.



- Challenges:
  - Paucity of geotagged Tweets
  - limited history of users
  - User network is expensive to collect
  - Topic bias in crisis data

A Behavior-Differential Approach

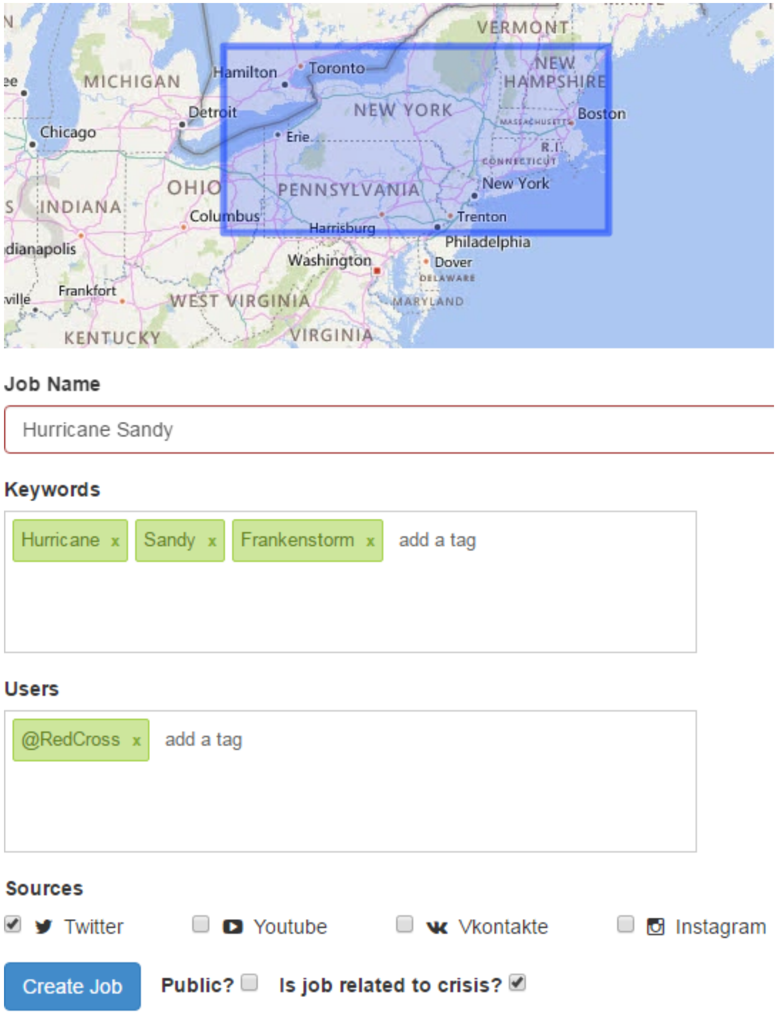
- Hypotheses: users in crisis region exhibit unique behaviors

- User behaviors in crisis regions
  - More likely to contain resources
  - Less likely to be retweets
  - Less likely to contain hashtags
  - Less likely to reference entities

- These characteristics can be harnessed to locate tweets

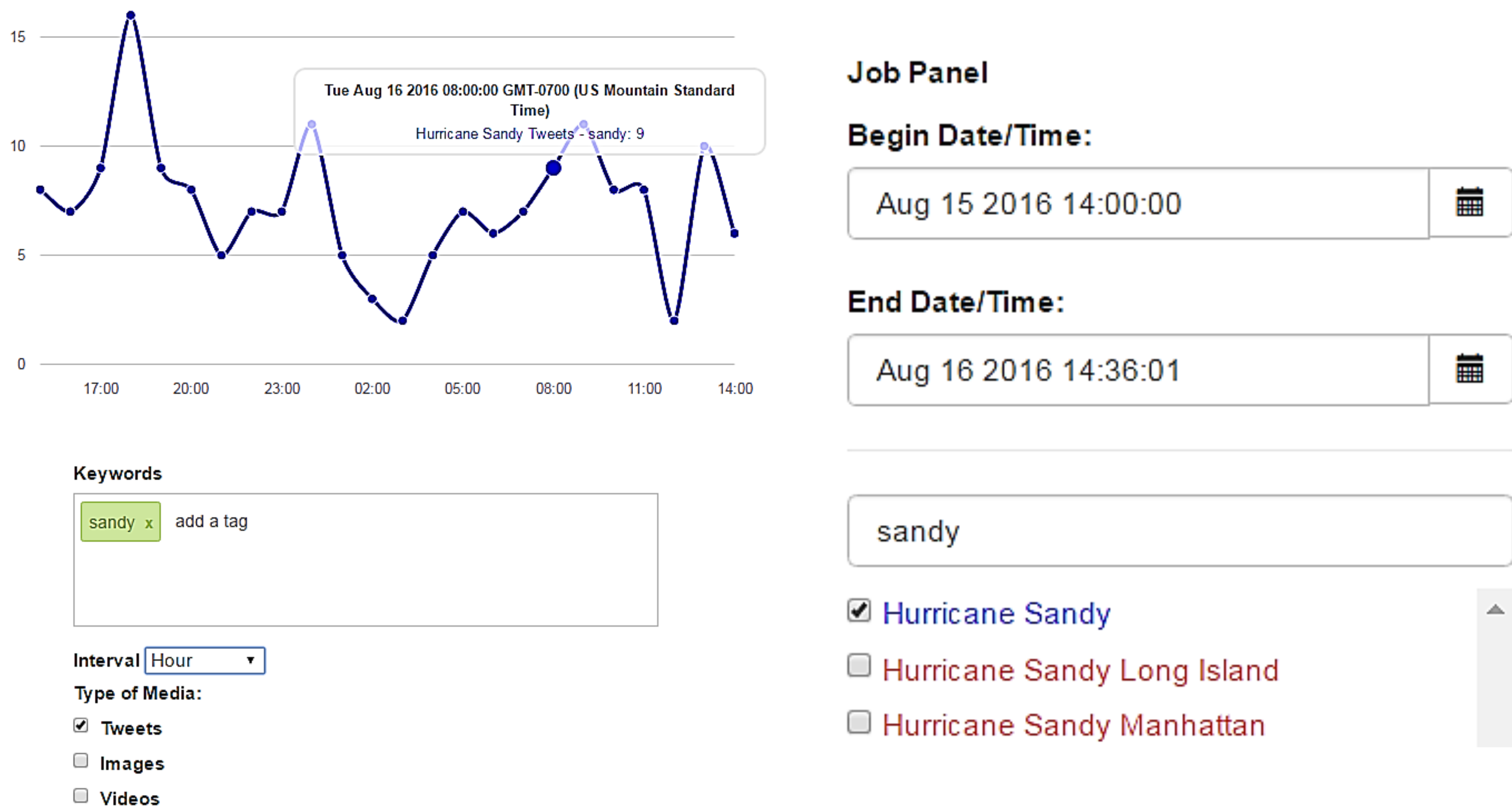
TweetTracker - Track

- Developed in DMML @ ASU
- Data collection
  - Keywords
  - hashtags
  - Location
  - Users
- Sources
  - Twitter
  - Instagram
  - YouTube
  - VK
- Collection type
  - Crisis-related
  - Normal



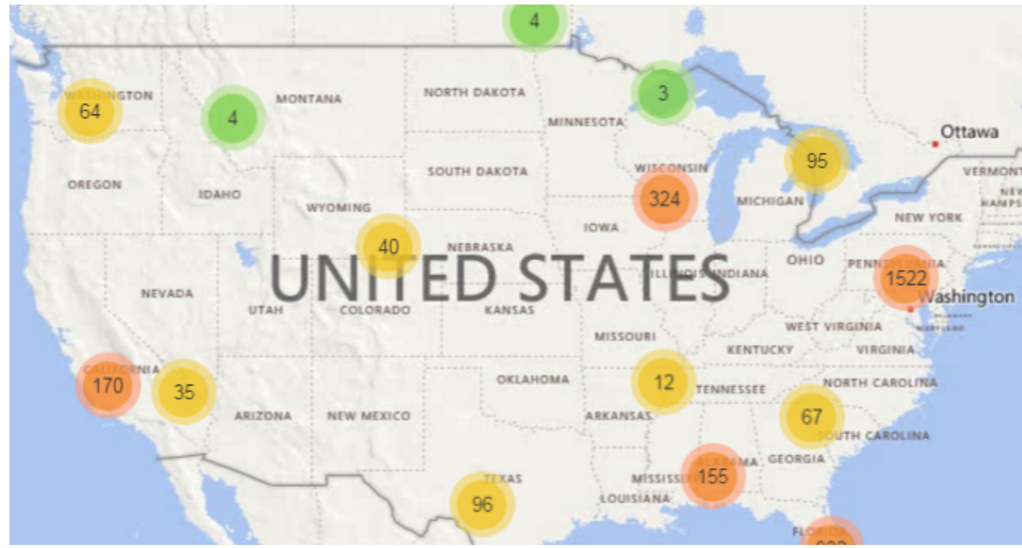
[Kumar ICWSM'11]

TweetTracker - Analyze



TweetTracker - Understand

- Location distribution
- Wordcloud
- Top links
- Top users
- Images
- Videos
- Top hashtags
- Most probable bots



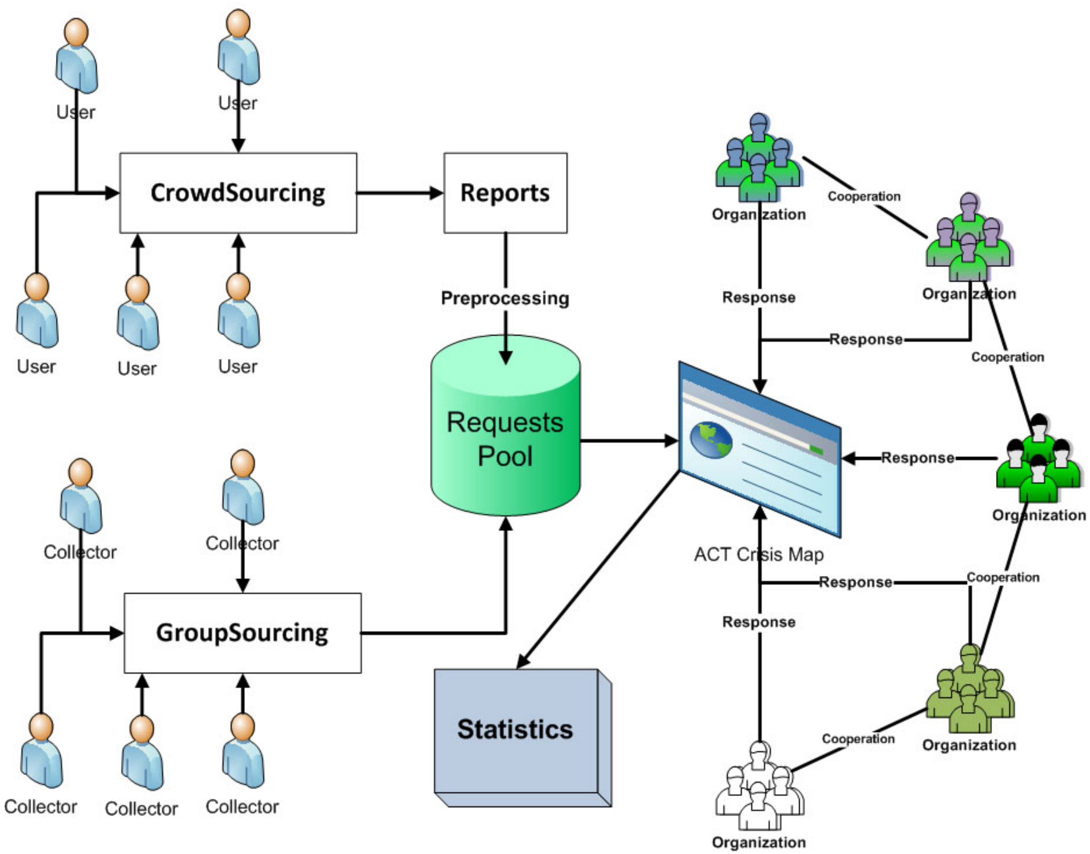
ACT Disaster Relief System Features

- Enables relief organizations to respond in an organized and collaborative manner
- Leverages the information available from a relief group to supplement the crowdsourcing information
- Provides helpful statistics
  - Organizational contribution during relief operations
  - Spatio-temporal distribution of requests
  - Request delivery status for evaluating relief progress

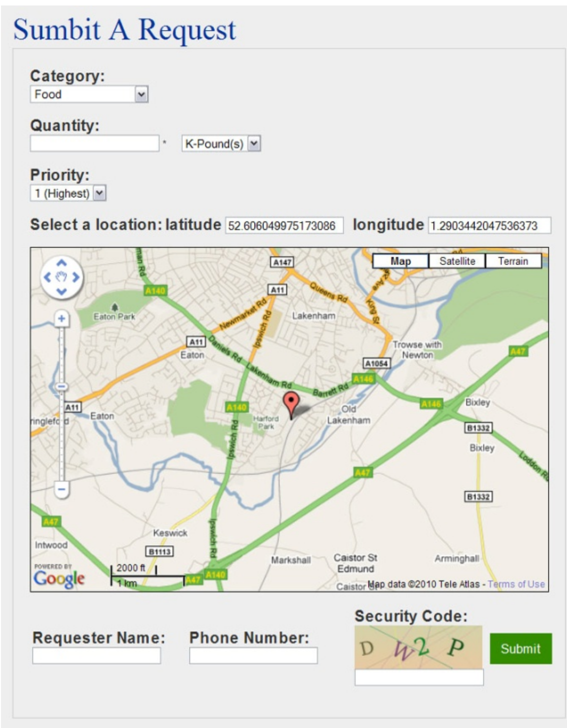
[Gao DTIC 11]

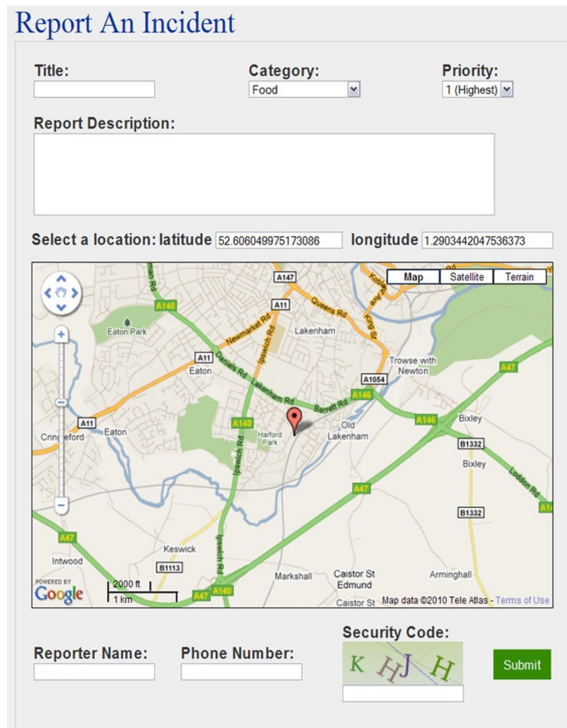
ACT Architecture

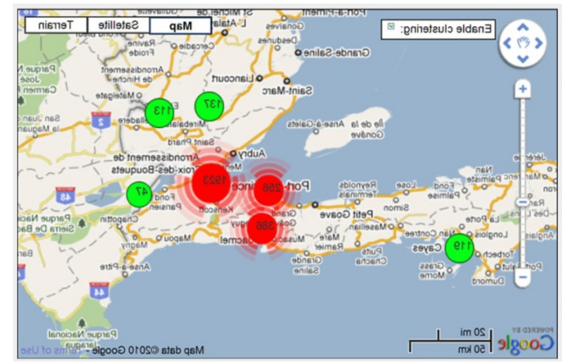
- System displays requests pool on crisis map
- Organizations respond to requests and coordinate with each other
- A statistics module helps track relief progress

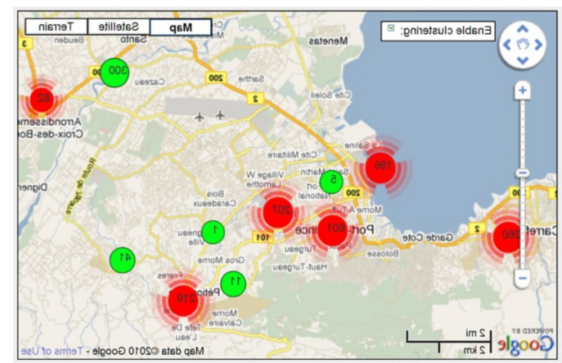


ACT Modules - Crowdsourcing

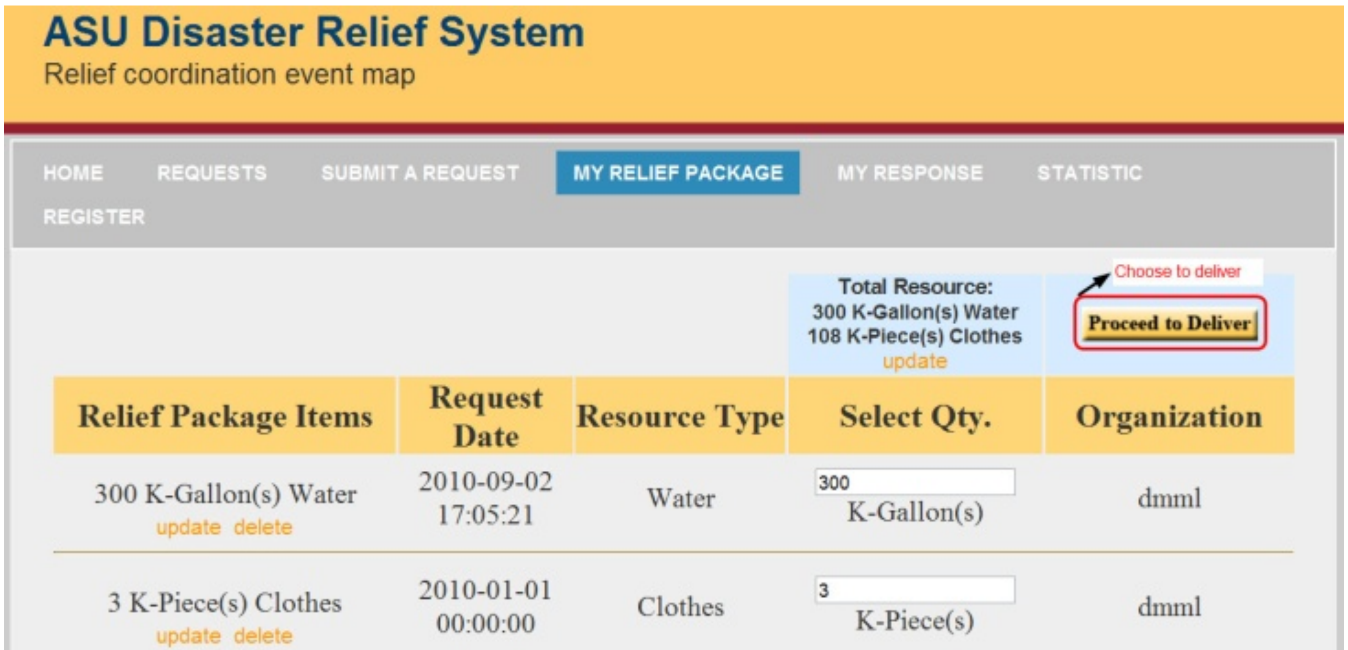


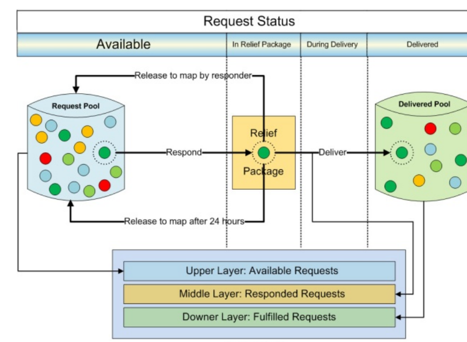






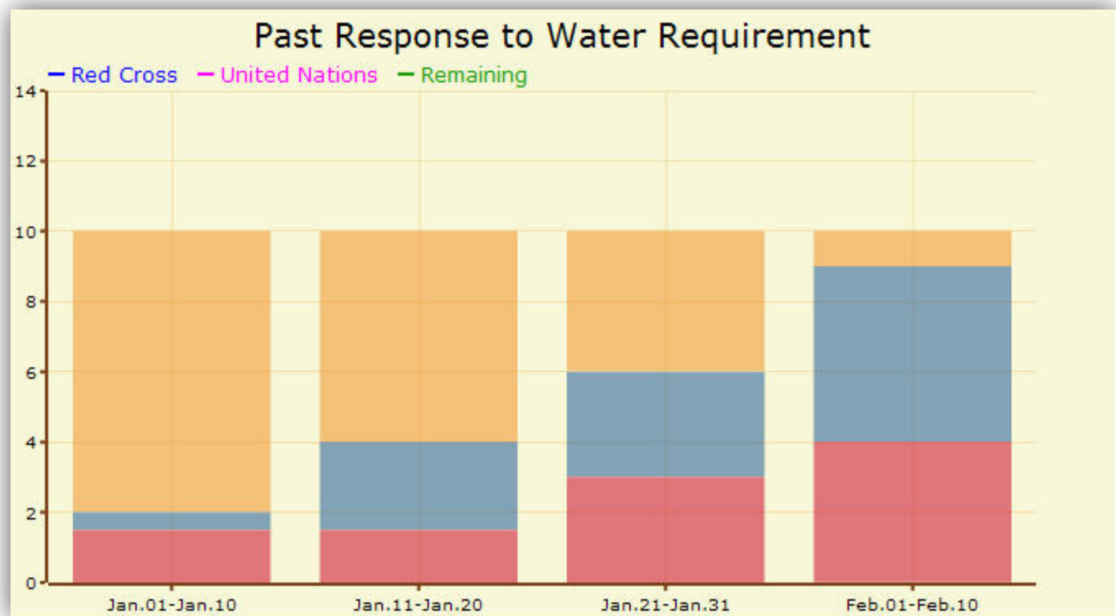
ACT Modules - Response & Coordination

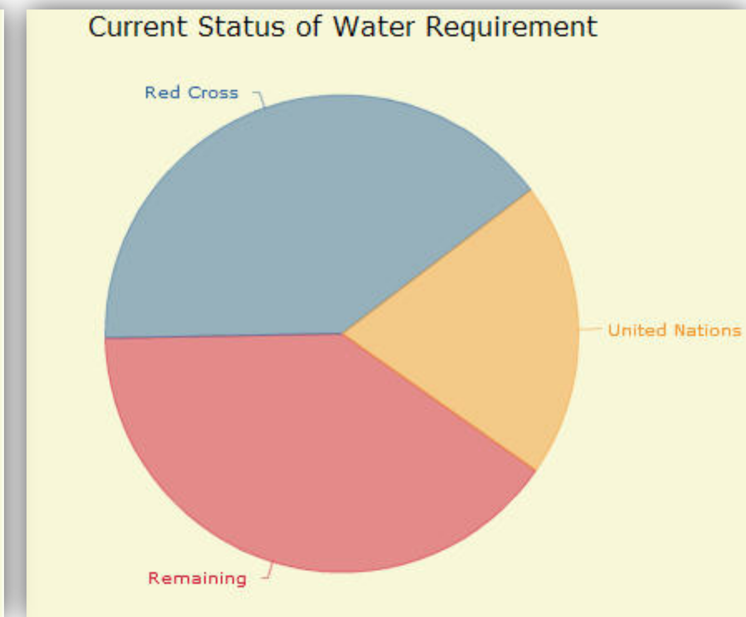




ACT Crisis Map

ACT Modules - Response Statistics





References

1. [Beigi SDM'16] Ghazaleh Beigi, Jiliang Tang, Suhang Wang, and Huan Liu. "Exploiting Emotional Information for Trust/Distrust Prediction". SIAM International Conference on Data Mining (SDM16), May 5-7, 2016. Miami, Florida.
2. [Dani ICDM'15] Harsh Dani, Fred Morstatter, Xia Hu, Zhen Yang, and Huan Liu. "Social Answer: A System for Finding Appropriate Sites for Questions in Social Media". Demo. In Proceedings of IEEE International Conference on Data Mining (ICDM2015), November 14 - 17, 2015. Atlantic City, NJ.
3. [Gao DTIC 11] Huiji Gao, Geoffrey Barbier, and Rebecca Goolsby. *Harnessing the crowdsourcing power of social media for disaster relief*. Arizona State Univ Tempe, 2011.
4. [Kumar ASONAM'15] Shamanth Kumar, Huan Liu, Sameep Mehta, and L. Venkata Subramaniam. "Exploring a Scalable Solution to Identifying Events in Noisy Twitter Streams", short paper, IEEE/ACM International Conference on Advances in Social Network Analysis and Mining (ASONAM2015), August 25-28, 2015, Paris, France.
5. [Kumar HT'14] Shamanth Kumar, Xia Hu, and Huan Liu. "A Behavior Analytics Approach to Identifying Tweets from Crisis Regions", short paper, 25th ACM Conference on Hypertext and Social Media (Hypertext2014), 1–4 Sep 2014, Santiago, Chile.
6. [Kumar ICWSM'11] Shamanth Kumar, Geoffrey Barbier, Mohammad Ali Abbasi, and Huan Liu. "TweetTracker: An Analysis Tool for Humanitarian and Disaster Relief, " Demo, 5th International AAAI Conference on Weblogs and Social Media (ICWSM-11), July 17-21, 2011. Barcelona, Spain.
7. [Morstatter ASONAM'16] Fred Morstatter, Liang Wu, Tahora H. Nazer, Kathleen M. Carley, and Huan Liu. "A New Approach to Bot Detection: Striking the Balance Between Precision and Recall", IEEE/ACM International Conference on Advances in Social Network Analysis and Mining (ASONAM2016), August 18-21, San Francisco, CA.
8. [Morstatter WWW'14] Fred Morstatter, Jürgen Pfeffer, Huan Liu. When is it Biased? Assessing the Representativeness of Twitter's Streaming API", WWW Web Science 2014.
9. [Morstatter ICWSM'13] Fred Morstatter, Jürgen Pfeffer, Huan Liu, Kathleen M Carley. Is the Sample Good Enough? Comparing Data from Twitter's Streaming API with Twitter's Firehose", ICWSM 2013.
10. [Ranganath WSDM'16] Suhas Ranganath, Xia Hu, Jiliang Tang, and Huan Liu. ``Understanding and Identifying Advocates of Political Campaigns on Social Media". ACM International Conference on Web Search and Data Mining (WSDM2016), February 22-25, 2016. San Francisco, CA.
11. [Ranganath ICDM'15] Suhas Ranganth, Suhang Wang, Xia Hu, Jiliang Tang, and Huan Liu. "Finding Time-Critical Replies for Information Seeking in Social Media". In Proceedings of IEEE International Conference on Data Mining (ICDM2015), November 14 - 17, 2015. Atlantic City, NJ.
12. [Sampson CIKM'16] Justin Sampson, Fred Morstatter, Liang Wu and Huan Liu. "Leveraging the Implicit Structure within Social Media for Emergent Rumor Detection", short paper, ACM International Conference of Information and Knowledge Management (CIKM2016), October 24-28, 2016. Indianapolis, Indiana.
13. [Sampson ICDM'15] Justin Sampson, Fred Morstatter, Reza Zafarani, and Huan Liu. "Real-Time Crisis Mapping Using Language Distribution". Demo. In Proceedings of IEEE International Conference on Data Mining (ICDM2015), November 14 - 17, 2015. Atlantic City, NJ.

Thank You and Acknowledgments

- Research efforts are, in part, supported by grants from ONR, ARO, NSF
- Former and current members in the DMML lab and and many collaborators

- Some further challenges in Social Media Mining
  - Evaluation without ground truth
  - Big data paradox
  - Noise removal fallacy
  - Inferring the implicit

Repositories and Recent Books

- scikit-feature – an open source feature selection repository in Python
- Social Computing Repository