Mathematics of Planet Earth: Natural Disasters

Fred Roberts
Rutgers University

Mathematics of Planet Earth 2013

- A joint effort initiated by North American Math Institutes: **MPE2013**
- More than 100 partner institutes, societies, and organizations in UK, France, South Africa, Japan, and all over the world
- www.mpe2013.org

Mathematics of Planet Earth 2013

- Activities world-wide throughout 2013
- Sponsorship by UNESCO
- Support from Simons Foundation
- Workshops, tutorials, competitions, distinguished lectures, educational programs

Mathematics of Planet Earth Beyond 2013

- Problems of the planet will not go away in one year.
- We are organizing a series of events to continue beyond 2013.
- New initiative world-wide now called **MPE**
- In the US, we call it **MPE2013+**
- NSF support

Mathematics of Planet Earth Beyond 2013

Goals of MPE2013+

- Involve mathematical scientists in addressing the problems of the planet
- Enhance collaborations between mathematical scientists and other scientists
- Involve students and junior researchers in the effort
- Encourage life-long commitment to working between disciplines to solve the problems of society

Mathematics of Planet Earth 2013+

- Opening Introduction to Problems of the Planet and involve students and junior faculty: Arizona State University, Jan. 7-10, 2014
- Five ***Research Clusters***, beginning with workshops:
  - ***Sustainable Human Environments*** (Rutgers U.), April 23-25, 2014
  - ***Global Change*** (UC Berkeley), May 19-21, 2014
  - ***Data-aware Energy Use*** (UC San Diego), Sept. 29 – Oct. 1, 2014
  - ***Natural Disasters*** (GA Tech), May 13-15, 2015
  - ***Management of Natural Resources*** (Howard University), June 4-6, 2015

Mathematics of Planet Earth 2013+

Follow-up cluster activities:

- Sustainable Human Environments cluster:
  - Pre-workshop: Urban Planning for Climate Events Sept. 2014; Post-workshop: Tentatively in Paris, Fall 2015
  - Cluster activities of various kinds
- Natural Disasters cluster: working with several potential partners in Mexico and Colombia.
- Global Change cluster: considering a follow-up event at the National Center for Atmospheric Research (NCAR) and one at Old Dominion U. on communication of global change challenges

Mathematics of Planet Earth 2013+

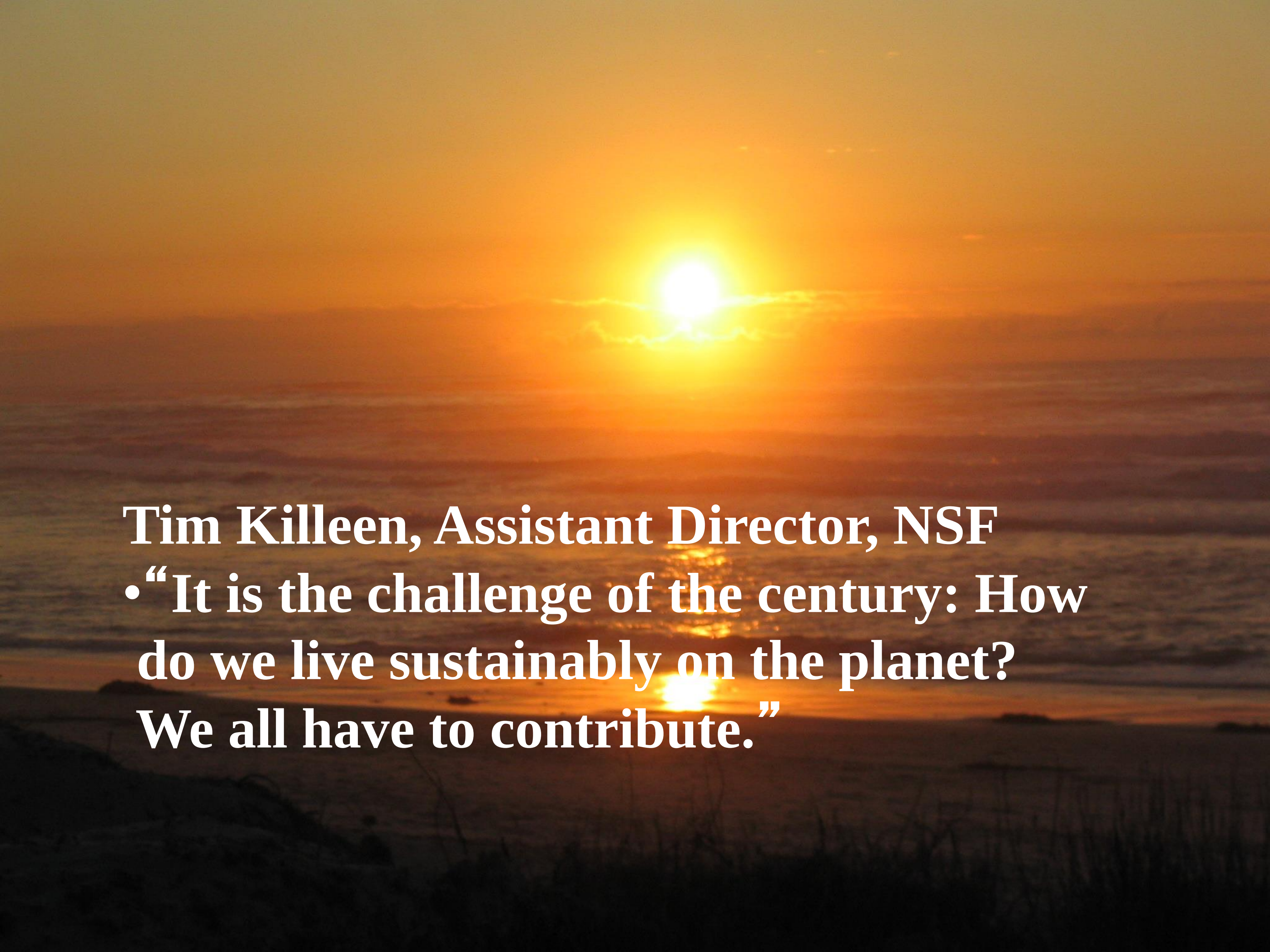
Follow-up cluster activities:

- Management of Human Resources cluster and Global Change cluster:
  - Looking into the possibility of follow up in Africa
- All clusters:
  - Looking into possibility of research groups ("squares") at American Inst. of Mathematics (AIM)

Mathematics of Planet Earth 2013+

- Education is a crucial piece of this and of the sustainability effort
  - Workforce development
  - Public literacy

- Need education at all levels, starting with K-12.
- Education issues in each workshop
- Special Education cluster: ***Education for the Planet Earth of Tomorrow***
- Cluster workshop: U. of Tennessee, Sept. 30 – Oct. 2, 2015.



Tim Killeen, Assistant Director, NSF

- “It is the challenge of the century: How do we live sustainably on the planet? We all have to contribute.”

Natural Disasters

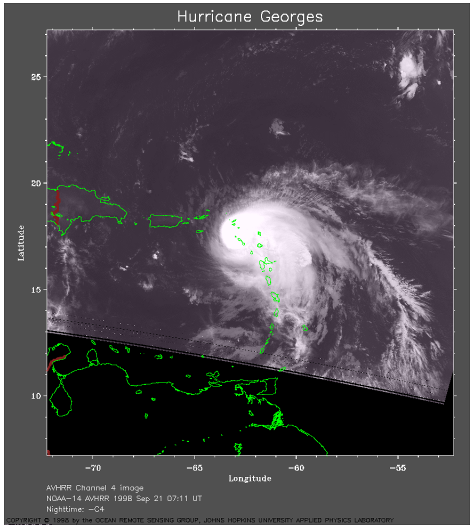
- ***No part of the world is impervious to natural disasters***
  - Epidemics
  - Earthquakes
  - Floods
  - Hurricanes
  - Tornadoes
  - Wildfires
  - Tsunamis
  - Extreme temperatures
  - Drought
  - Oil spills

Nepal 2015: www.circleofblue.org

- ***Mathematical sciences can help in predicting, monitoring, and responding to such events, and mitigating their effects.***

Natural Disasters

- Will discuss sample math challenges from some of these disasters
- Will describe some examples of projects I've been involved with
- All by way of introducing key themes for the workshop



Haitian earthquake

Japan earthquake & tsunami

Global Change

From MPE2013+ Global Change Workshop:

- The planet is constantly changing.
- But the pace of change has accelerated as a result of human activity:
  - Construction and deforestation change habitats
  - Over-fishing reduces wild populations
  - Fossil fuel combustion leads to atmospheric greenhouse gas buildup
  - Commerce and transport introduce non-native species.

Planning for a Changing Environment

- ***Planning for a Changing Environment: Math Science Challenges:***
  - Changing climate could lead to more extreme weather events
  - Changing climatic and environmental conditions could affect where and when diseases break out
  - Changing conditions could lead to more extended droughts
  - How do we plan for these events/changes and mitigate their effects?

Urban Planning for Climate Events

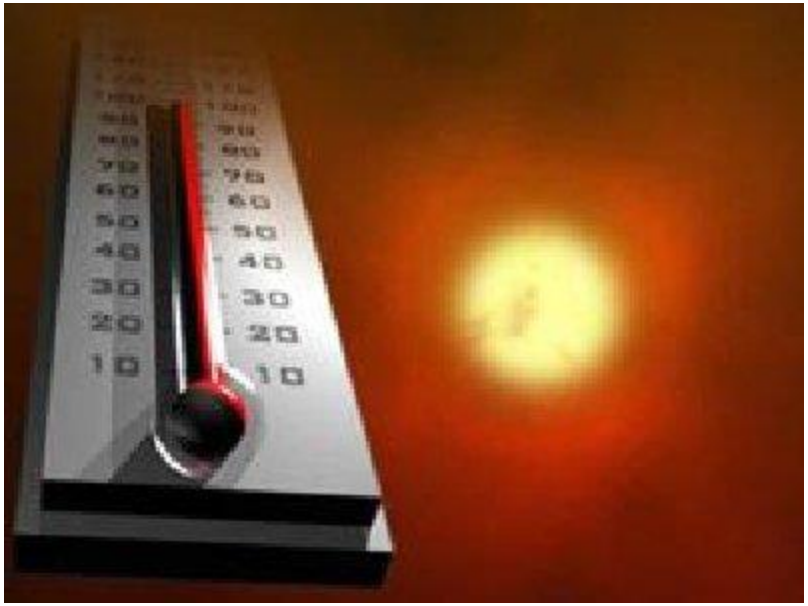
- ***Example: Climate events***: Super Storms, heat, drought, floods – all could be increasing in number and severity.
- What can urban areas do to prepare for them?

Urban Planning for Climate Events

- MPE 2013+ Sustainable Human Environments cluster: Pre-workshop: Urban Planning for Climate Change, Sept. 2013, at DIMACS-Rutgers University
- *What can urban areas do to prepare for/mitigate changes due to climate and in particular the effect of future climate events?*

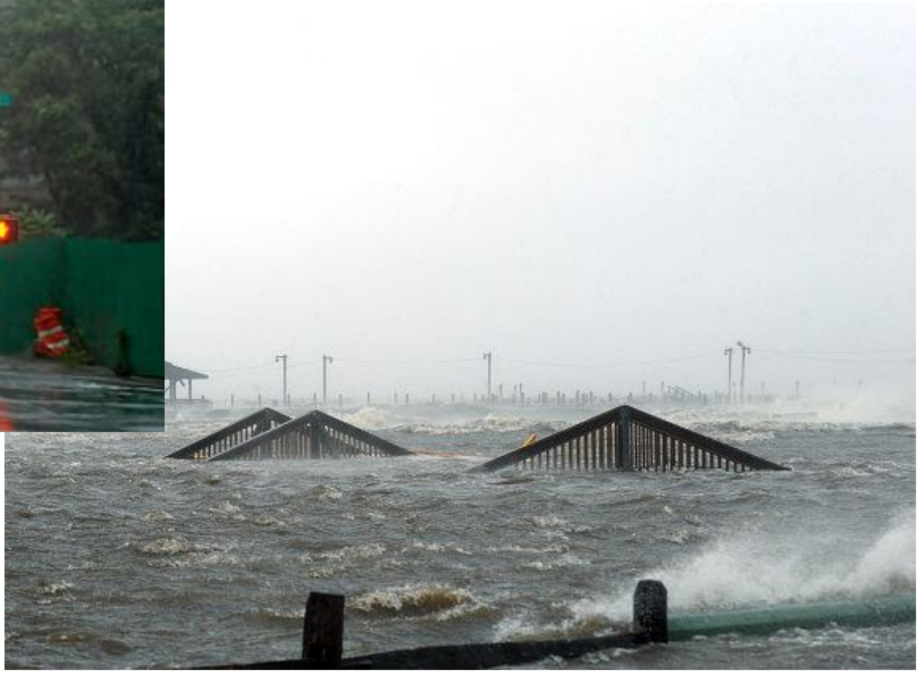
Extreme Events due to Global Warming

- We anticipate an increase in number and severity of extreme events due to global warming.
- More heat waves.
- More floods, hurricanes.



Extreme Events due to Global Warming: More Hurricanes

Irene hits NYC – August 2011



Extreme Events due to Global Warming: More Hurricanes

Irene hits NYC – August 2011

Extreme Events due to Global Warming: More Hurricanes

Irene hits NYC – August 2011

Extreme Events due to Global Warming: More Hurricanes

Sandy Hits NJ Oct. 29, 2013

My backyard

My block

Extreme Events due to Global Warming: More Hurricanes

Sandy Hits NJ Oct. 29, 2013

My neighborhood

My block

Extreme Events due to Global Warming: More Hurricanes

Sandy Hits NJ Oct. 29, 2013

NJ Shore – from Jon Miller

Extreme Events due to Global Warming: More Hurricanes

Future Storms

- To plan for the future, what do we need to do?
- How can we use mathematical modeling, simulation, and algorithmic tools of risk assessment to plan for the future?
- To plan for more extreme events
- To plan for rising sea levels

Extreme Events due to Global Warming: More Hurricanes

- Using mathematical modeling, simulation, and algorithmic methods of risk assessment to plan for the future:
  - What subways will be flooded?
  - How can we protect against such flooding?

Extreme Events due to Global Warming: More Hurricanes

•Using mathematical modeling, simulation, and algorithmic methods of risk assessment to plan for the future:

– What power plants or other facilities on shore areas will be flooded?

- Do we have to move them?

Extreme Events due to Global Warming: More Hurricanes

- Using mathematical modeling, simulation, and algorithmic methods of risk assessment to plan for the future:
  - How can we get early warning to citizens that they need to evacuate?
  - How can we plan such evacuations effectively?

Extreme Events due to Global Warming: More Hurricanes

- Using mathematical modeling, simulation, and algorithmic methods of risk assessment to plan for the future:
  - –How can we plan placement of utility lines to minimize down time?

Extreme Events due to Global Warming: More Hurricanes

- Using mathematical modeling, simulation, and algorithmic methods of risk assessment to plan for the future:
  - How can we plan for getting people back on line after a storm?

Bringing in help from out of state

Extreme Events due to Global Warming: More Hurricanes

- Using mathematical modeling, simulation, and algorithmic methods of risk assessment to plan for the future:
  - How can we set priorities for cleanup?

Sea Level Rise

- Sectors affected by sea level rise include:
  - Transportation
  - Communications
  - Energy
  - Construction
  - Water supply
  - Waste

- How do we prioritize among them?
- How are they interrelated? E.g., better communications might help us reroute transportation more efficiently.

Photo Source: Updating Maryland’s Sea-Level Rise Projections, Maryland Climate Change Commission, June 26, 2013

Example: Extreme Heat Events

- Subject of a DIMACS project.
- Nina Fefferman, et al.
- Key initial emphasis by CDC's modeling unit
- Extreme heat events result in increased incidence of heat stroke, dehydration, cardiac stress, respiratory distress
- Hyperthermia in elderly patients can lead to cardiac arrest.
- Effects not independent: Individuals under stress due to climate may be more susceptible to infectious diseases

Extreme Heat Events

- One response to such events: evacuation of most vulnerable individuals to climate controlled environments.
- Modeling challenges:
  - Where to locate the evacuation centers?
  - Whom to send where?
  - Goals include minimizing travel time, keeping facilities to their maximum capacity, etc.
  - All involve tools of Operations Research: location theory, assignment problem, etc.
  - Long-term goal in “smart cities”: Utilize real-time information to update plans

Extreme Heat Events

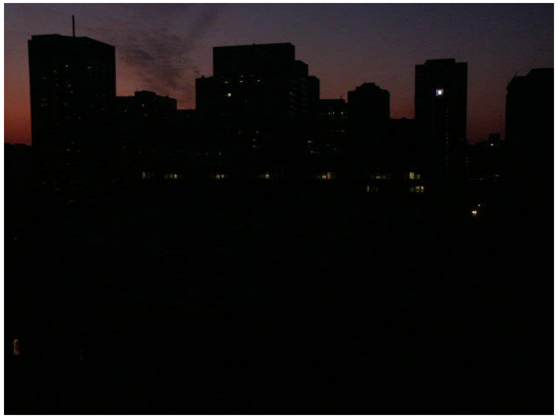
- The DIMACS project on heat event evacuation was based in Newark, NJ – collaboration with Newark city agencies.
- Data includes locations of potential shelters, travel distance from each city block to potential shelters, and population size and demographic distribution on each city block.
- Determined "at risk" age groups and their likely levels of healthcare needed to avoid serious problems

Extreme Heat Events

- DIMACS project was also concerned with computing optimal routing plans for at-risk population to minimize adverse health outcomes and travel time
- Question: Can routing plans be implemented quickly; can we get information to people quickly?
- Can we develop evacuation plans that are quickly modifiable given data from evacuation centers, traffic management, etc.?
- Far from what happens in evacuations today.

Extreme Heat Events

- A side effect of such events: Extremes in energy use lead to need for rolling blackouts.
- Modeling challenges (another DIMACS research project)
  - Understanding health impacts of blackouts and bringing them into models
  - How do we design efficient rolling blackouts while minimizing impact on health?
    - Lack of air conditioning
    - Elevators don't work: vulnerable people suffer from over-exertion
    - Food spoilage
  - Minimizing impact on the most vulnerable populations



- Math science challenge: How to utilize "smart grid" to update plans

Floods

• How do we plan emergency rescue vehicle routing to avoid rising flood waters while still minimizing delay in provision of medical attention and still getting afflicted people to available hospital facilities?

•In graph theory terms, how do we find a minimal spanning tree – minimal number of edges so every node can reach every other node.

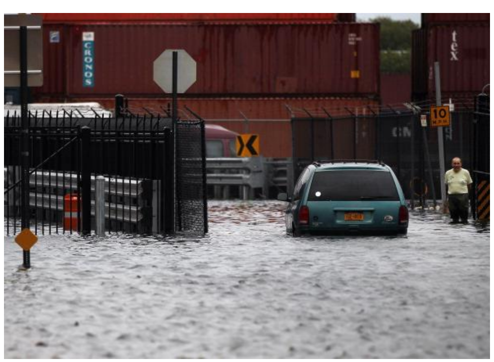
Floods

- Which flood mitigation projects to invest in?
  - Buyouts
  - Better flood warning systems
  - “Green infrastructure” (cisterns & rain barrels)
  - Pervious concrete
  - Etc.

Raritan River flood
Bound Brook, NJ
August 2011

Floods

- This requires data-driven/Model-driven Decision Support
- Data-driven. Assemble data about:
  - Precipitation (duration, amount)
  - Antecedent conditions (soil moisture content, ground cover, seasonality)
  - River gage levels
  - Flood maps
  - Property damage data – FEMA payouts



Floods: Example Project

- Data-driven, Model-driven Decision Support
- Model-driven: A CCICADA project with FEMA
  - CCICADA is our homeland security center based at DIMACS.
  - Hydrological modeling to predict effect on river gage levels of different flood mitigation projects
  - Econometric modeling to predict insurance costs at different flood gage levels
  - Combine the two

CCICADA
Command, Control, and Interoperability Center for Advanced Data Analysis
A Department of Homeland Security Center of Excellence

Floods: Example Project

- Hydrological Models
  - Goal: determine effect of a given mitigation strategy on water levels
  - Hydrological models combine information on:
    - ➢rainfall
    - ➢soil moisture
    - ➢seasonality
    - ➢river levels prior to a rain event
    - ➢elevation
    - ➢watershed properties
    - ➢land cover (natural and built environment)
  - Hydrological models produce flood inundation maps

Command, Control, and Interoperability Center for Advanced Data Analysis
A Department of Homeland Security Center of Excellence

Floods: Example Project

- Mathematical sciences can enhance hydrological models by:
  – Providing better methods of calibration
  – Developing improved models for land cover and soil types
  – Analysis of the effect of channel geometry and flow speed
  – More precise handling of uncertainty

CCICADA
Command, Control, and Interoperability Center for Advanced Data Analysis
A Department of Homeland Security Center of Excellence

Floods: Example Project

- Econometric models predict insurance costs at different flood gage levels
- By combining hydrological models that predict flood gage levels and econometric models, we can estimate effect on costs of a flood under different assumptions about mitigating investments.

CCICADA
Command, Control, and Interoperability
Center for Advanced Data Analysis
A Department of Homeland Security Center of Excellence

Disease Events

- Newly emerging diseases can threaten the health of millions of people.
- The 1918 influenza epidemic killed 50 million people around the world; WW I killed 16 million
- Great concern about the potential for a new influenza outbreak of similar proportions

Source: Dartmouth Medicine 2006

Disease Events

- Modern transportation systems make it much easier for diseases to spread around the world – e.g., Ebola
- Deliberate introduction of diseases by bioterrorists is a serious concern
- Climate change leads to diseases appearing in places the have not appeared in before – e.g., malaria in the highlands of Kenya & potential for malaria in the US

Disease Events

•Mathematical models of the spread of disease go back to Bernoulli's mathematical analysis of smallpox in 1760.

Disease Events

- Thousands of models since have:
  - Highlighted concepts such as core population in STDs
  - Made explicit concepts such as herd immunity for vaccination policies
  - Led to insights about drug resistance, epidemic trends, effect of different kinds of treatments

Disease Events

- Thousands of models since have:
  - Helped develop strategies for vaccination or quarantine in time of emergency

Waiting on line to get smallpox vaccine during New York City smallpox epidemic 1947

Disease Events

- Understanding infectious diseases requires being able to reason about highly complex biological systems, with hundreds of demographic and epidemiological variables
- Intuition is not sufficient.
- Experimental or field trials are impossible or unethical
- Need mathematical modeling

Disease Events

- Mathematical models have become important tools in analyzing the spread and control of infectious diseases, especially when combined with powerful, modern computer methods for analyzing and/or simulating the models

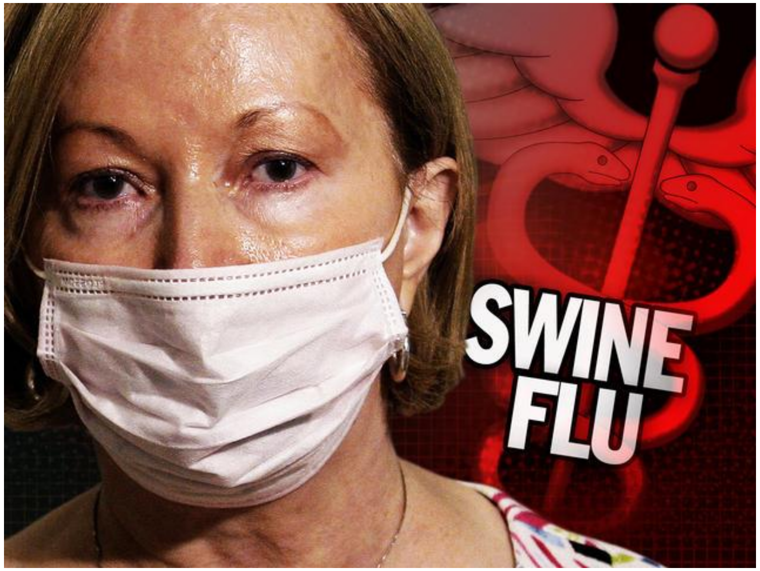
Disease Events

- Great concern about the deliberate introduction of diseases by bioterrorists has led to new challenges for mathematical modelers

anthrax

Disease Events

Great concern about possibly devastating new diseases like H1N1 influenza has also led to new challenges for mathematical modelers



Example: Climate and Health

- Some early warning signs:
  - Malaria in the African Highlands
  - Dengue epidemics along the Rio Grande & in Brazil

Dengue Fever

Malaria mosquito

Example: Climate and Health

- Some early warning signs:
  - Cholera affected by sea surface temperature
  - Increase in Lyme disease in Canada
  - St. Louis Encephalitis (Florida outbreak)
  - Animal and plant diseases too

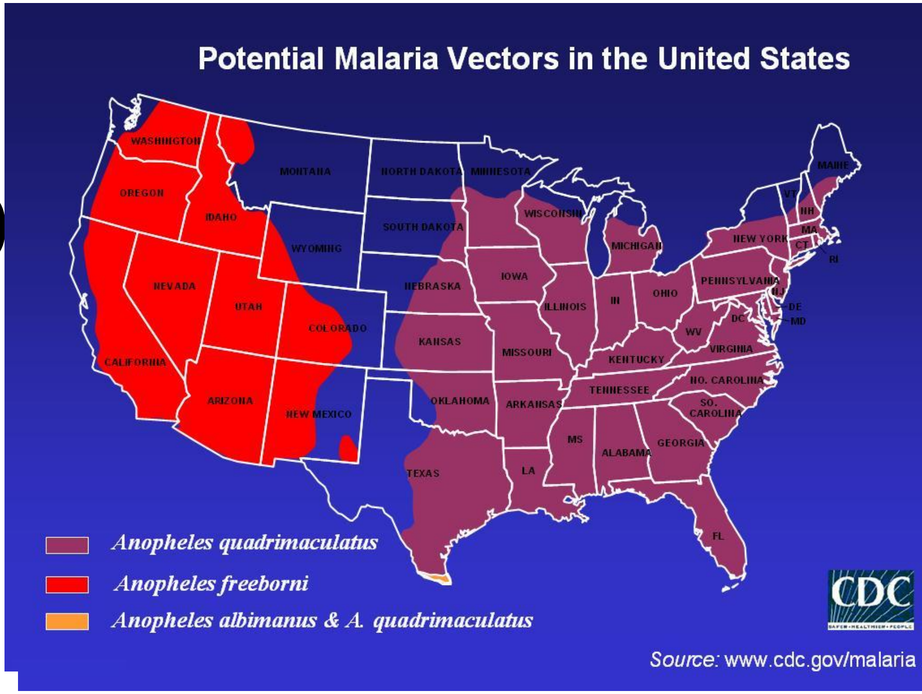
–Complex interaction among climate, life cycle of hosts and vectors, migration patterns, etc.

Tick carrying Lyme Disease

cholera

Example: Malaria

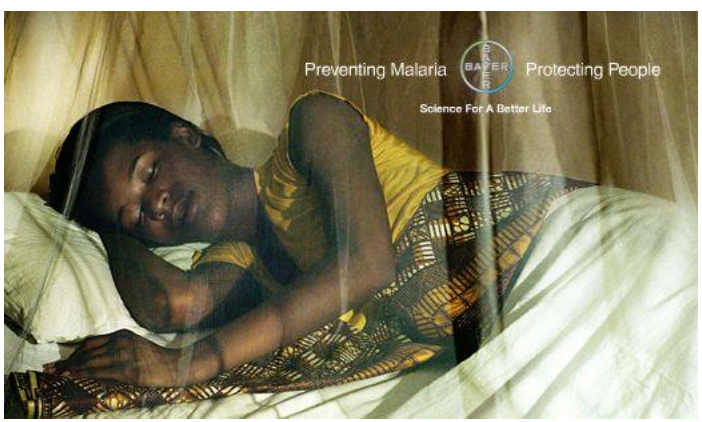
- The challenge of climate change: Malaria springs up in areas it wasn't in before.
- Highlands of Kenya
- Potential for Malaria in the US – Texas, Florida, Washington, …
- A CCICADA project
 (Nkwanta, Yakubu, et al.)
- A key role for modelers:
Aid in early warning:
Surveillance.



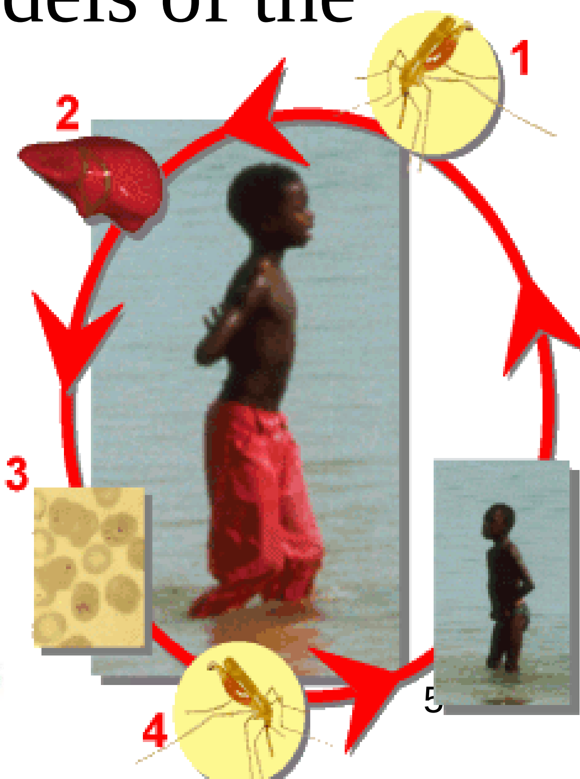
Command, Control, and Interoperability
Center for Advanced Data Analysis
A Department of Homeland Security Center of Excellence

Example: Malaria

- CCICADA: Modeling impact of interventions in Africa: bed nets, mosquito eradication, education.
- Modeling impact on malaria in the US of different interventions in Africa
- Complex interactions involving models of the malaria cycle and the impact of modern transportation systems



CCICADA
Command, Control, and Interoperability Center for Advanced Data Analysis
A Department of Homeland Security Center of Excellence



Oil Spills

- The Deepwater Horizon oil spill (2010):

Oil Spills

- The Exxon-Valdez Oil Spill –Alaska (1989)

Source: The Encyclopedia Of Earth eoearth.org

Source: Alaska Dept. Environmental Conservation

Example: Oil Spills

- With climate change, more vessel traffic in the Arctic and possibility of offshore oil drilling
- Increased risk of spills
- Arctic challenges: resource allocation in advance in case of oil spill
  - Necessary because of long transit times, lack of infrastructure, remote locations, lack of roads, distant airlift
- CCICADA Research project with US Coast Guard: Dynamic Modeling for Arctic Resource Allocation (DMARA): Graboski, Wallace, Sharkey, et al.

CCICADA
Command, Control, and Interoperability Center for Advanced Data Analysis
A Department of Homeland Security Center of Excellence

Example: Oil Spills

Arctic resource allocation a persistent and complex challenge in many Coast Guard missions

"The lack of infrastructure and oil spill response equipment in the U.S. Arctic is a significant liability in the event of a large oil spill" (National Academies, 2014).

United States Coast Guard
U.S. Department of Homeland Security

CCICADA
Command, Control, and Interoperability Center for Advanced Data Analysis
A Department of Homeland Security Center of Excellence

Example: Oil Spills

- DMARA: formulated oil spill resource allocation problem as large optimization problem.
- Model has two million variables and almost as many constraints.
- Need:
  - Modifications to the mathematical formulation
  - Algorithmic research to improve solvability
  - Heuristic programming techniques as possible solution procedures.

CCICADA
Command, Control, and Interoperability Center for Advanced Data Analysis
A Department of Homeland Security Center of Excellence

Oil Rigs

- Cyber security a real challenge to oil rigs.
- Hackers recently shut down a floating oil rig off the coast of Africa by tilting it. It took a week to identify and fix the problem. (Reuters 4/23/14)
- In 2010: drilling rig being moved at sea from South Korea to South America was infected by malicious software so its critical control systems couldn't operate. Took 19 days to fix matters. (Reuters 4/23/14)

Credit: www.peakoil.net

Oil Rigs

- In the Korean example: the computers controlling the blowout preventer were infected.
- If this had happened while the rig was engaged in drilling operations, there could have been a well blowout with possible explosion and oil spill.
- The blowout preventer failed during the Deepwater Horizon oil spill in the Gulf of Mexico in 2010.
- The malware involved might not have caused a problem for a smartphone, but that has much better security than an oil rig.

Credit: wikipedia.org, Shauk 2013

Oil Rigs

- The system that keeps an oil rig in position also has vulnerabilities.
- Dynamic positioning (DP) is a computer-controlled system to automatically maintain the position (and heading) of a vessel, in particular an oil rig.
- In DP, knowledge of the oil rig's position and angle, sensor information, wind direction, and speed feed into a computer program that contributes to the oil rig's stability.
- Disabling the DP of an oil rig by jamming its GPS could conceivably have a serious effect on the rig.
- In addition to safety and environmental impacts, large cost: Oil rigs are contracted for at close to $1M a day.

Shauk 2013

Some Approaches/Tools to Aid with Natural Disasters

- Stockpiling
- Quarantine/Vaccination
- Surveillance/Monitoring
- Social media for situational awareness

Logistics: Supply Chains/Stockpiles

- What supplies are needed during an emergency?
  - Water, Food, Fuel, Generators, Chainsaws?
- How and where can we stockpile them?
- What are good methods for getting these to those who need them in an efficient way?
- How can we better utilize the private sector?


Logistics: Supply Chains/Stockpiles

- How can we tell who needs what kinds of goods during an emergency?
- How can we locate stockpiles so as to be “agile” in allocating the resources when needed?
- E.g.: CDC’s strategic national stockpile of medicines for emergencies: how do we decide what medicines to include, how many doses, where to keep them?

Source: cdc.gov

Quarantine/Vaccination

- We can vaccinate in advance of an epidemic or during an outbreak.
- Vaccination strategies are highly math-based.
- For example: Do we just vaccinate those who have been exposed to a disease? Or do more widespread preventive vaccination?

Quarantine/Vaccination

- With quarantine, we tell some people to stay home.
- Or, we try to isolate from others those who might be likely to come down with the disease.
- There are many complex quarantine questions calling for mathematical analysis
- During the SARS epidemic in Singapore, people who were told to stay home didn't do so – until the government came up with a bribe" (something like $75 a day). How does one compute what will work?

Quarantine/Vaccination

- During the 2001 UK foot and mouth disease, the mathematical modelers came up with a “ring” strategy – culling livestock within a certain number of meters of an affected farm.
- That is effectively quarantining the farm.

Surveillance/Monitoring

- Early detection of disease outbreaks is critical for public health
- There are many warning systems in place.
- DIMACS faculty were among the pioneers in "syndromic surveillance" systems of various kinds
- We can monitor emergency room admissions, prescription drug purchases, hits on health-advice websites, physician reports, etc.

Example: Biosurveillance: For Early Detection

- CCICADA project (Fefferman, Nkwanta, Yakubu, et al.): Use entropy for biosurveillance
- ***Entropy quantifies the amount of information communicated within a signal***
- ***Signal strength may change when an outbreak starts***
- We are hoping to detect changes in signal strength **<u>early</u>** into the onset of an outbreak

smallpox

Command, Control, and Interoperability Center for Advanced Data Analysis

Example: Biosurveillance: For Early Detection

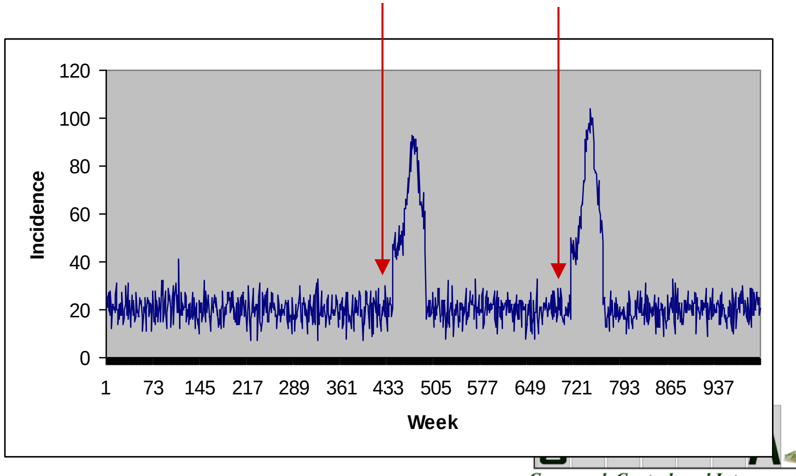
Shannon's Entropy Formula

$$H(X) = E(I(X)) = -\sum_{i=1}^{n} p(x_i)\log_2 p(x_i)$$

– $I(X)$ is the information content of X

– $p(x_i) = \Pr(X = x_i)$ is the probability mass function of X

We want to be able to take incoming disease data and, as early as possible, notice when an outbreak is starting



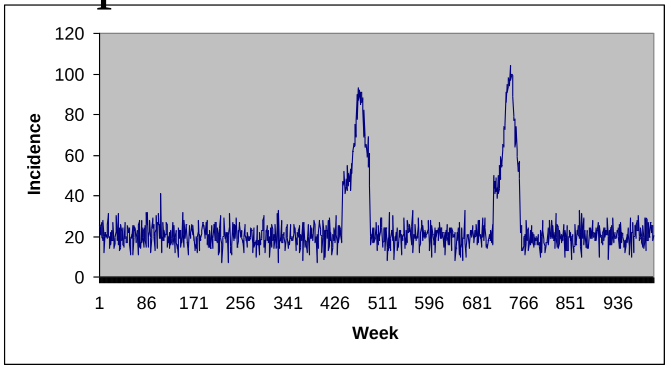
Command, Control, and Interoperability Center for Advanced Data Analysis

Example: Biosurveillance: For Early Detection

- Current methods of outbreak detection are often hit or miss.
- A frequently used method: ***CuSum***
  - Compares current cumulative summed incidence to average
  - Needs a lot of historical "non-outbreak" data (bad for newly emerging threats)
  - Has to be manually "reset"
- Other methods have similar problems

Example: Biosurveillance: For Early Detection

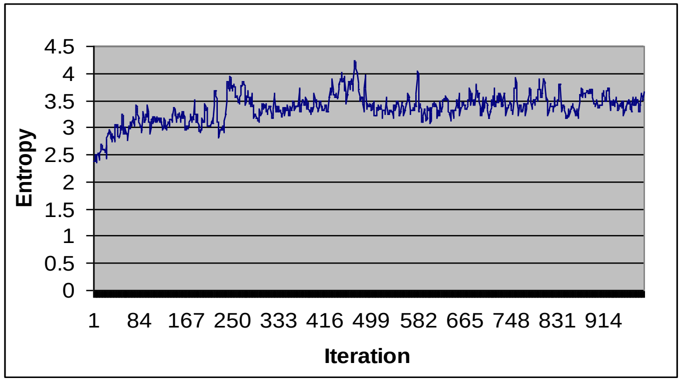
Reported Incidence Data



We apply 3 preprocessing steps

We stream the processed data through our entropy calculation

Entropy Outcome



Command, Control, and Interoperability Center for Advanced Data Analysis

Biosurveillance Using Entropy: The 3 Preprocessing Steps

1. **Binning** the Incidence Data: Number of categories

2. Analyzing within a Temporal **Window:** Number of time points lumped into one observation

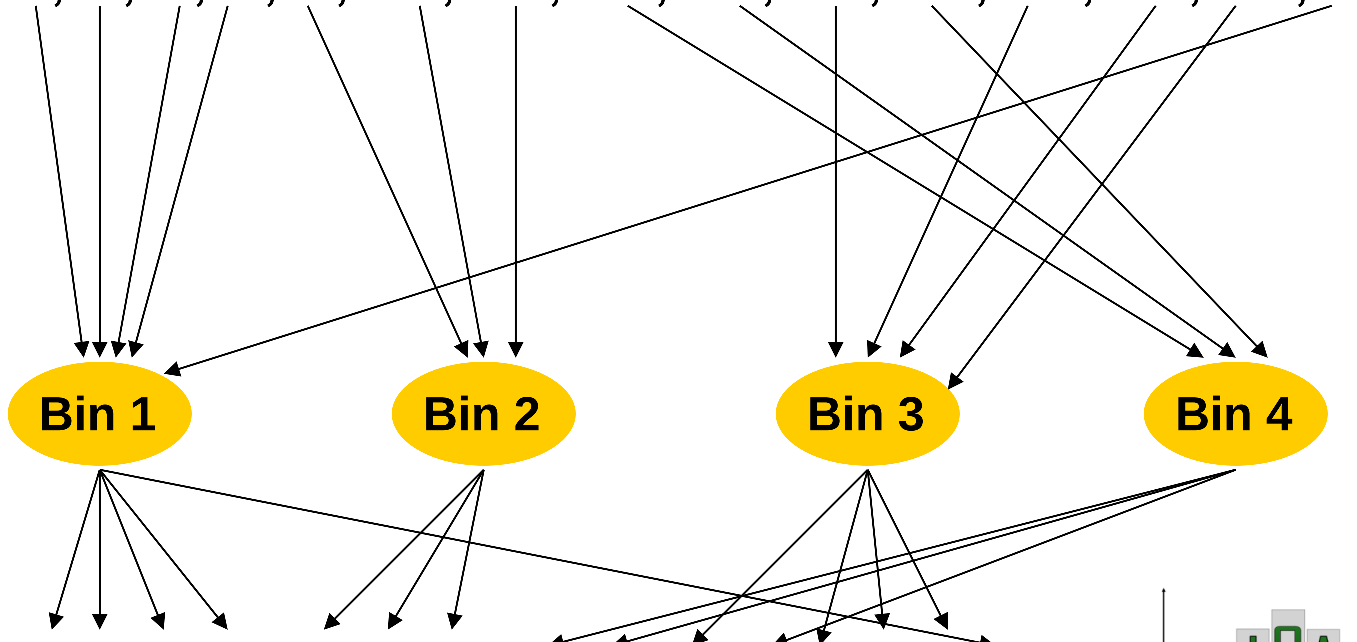
3. Moving the temporal window according to different **Step Sizes**

Binning

- Assign each "count" to a bin or category.
- Binning lets us try to focus on ***biologically meaningful differences***.

Weekly Disease Incidence (Number of Cases)

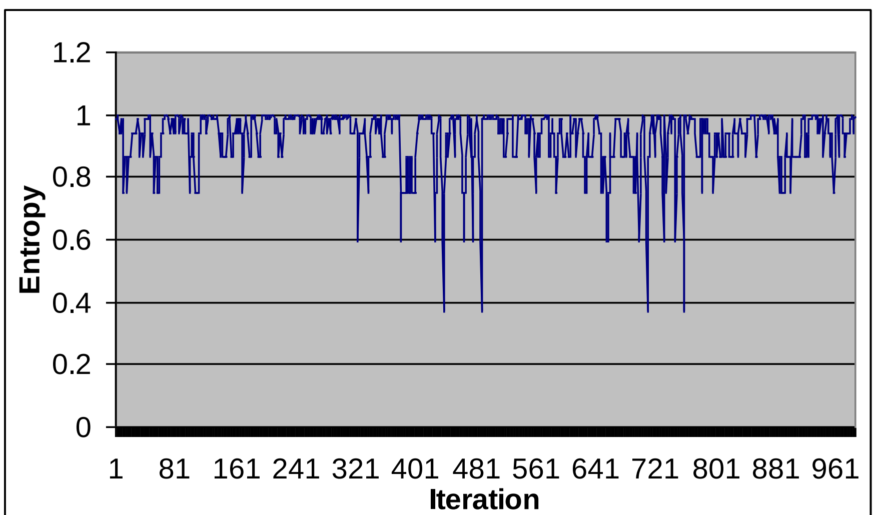
Data: 3, 2, 4, 5, 8, 10, 12, 40, 35, 17, 37, 20, 23, 25, 4,…



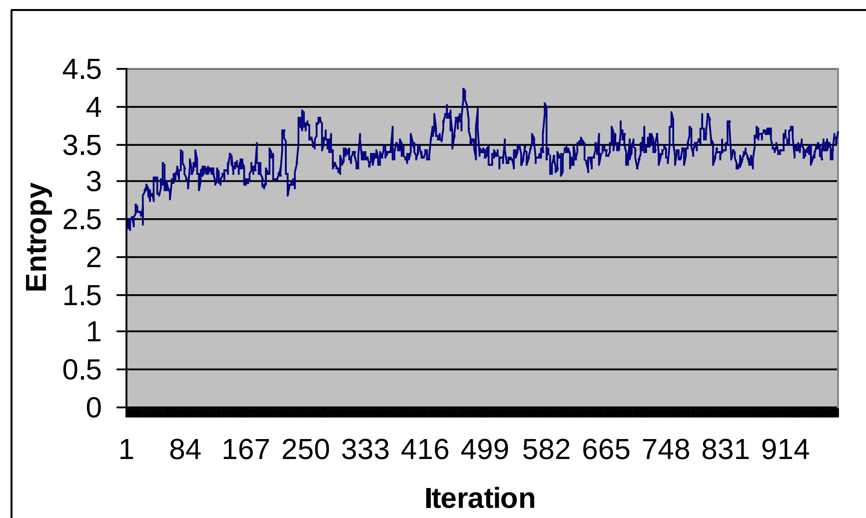
Binned Data: 1, 1, 1, 1, 2, 2, 2, 4, 4, 3, 4, 3, 3, 3, 1

Method of Binning can Really Change the Outcome

Method 1



Method 2



Window Size

The window for number of data points to look at at one time should be large enough to detect when a change has happened (some data from "before" and some from "after" the outbreak starts), but small enough that it can't entirely contain rapid peak.

Window Size = 7

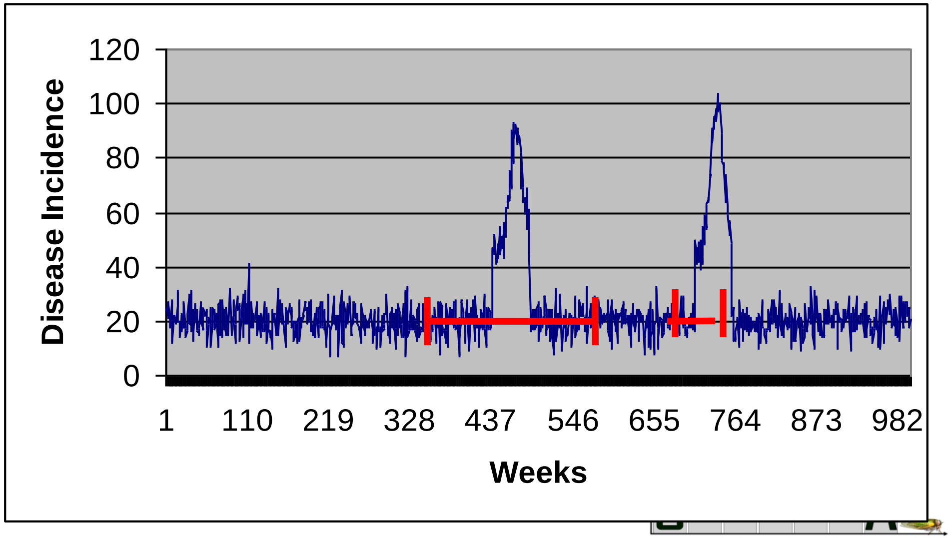
Step Size = 1

Incidence Data:

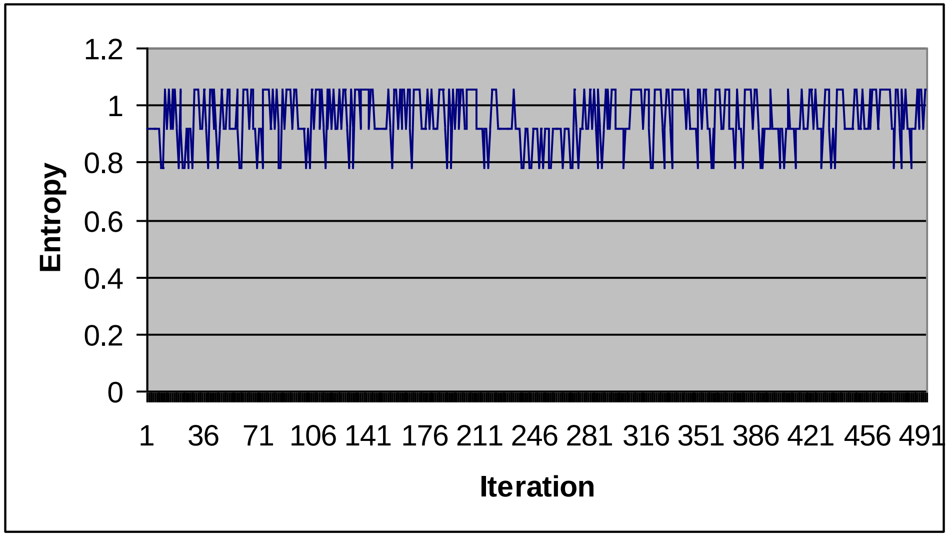
3, 2, 4, 5, 8, 10, 12, 40, 35, 17, 37, 20, 23, 25, 4,…

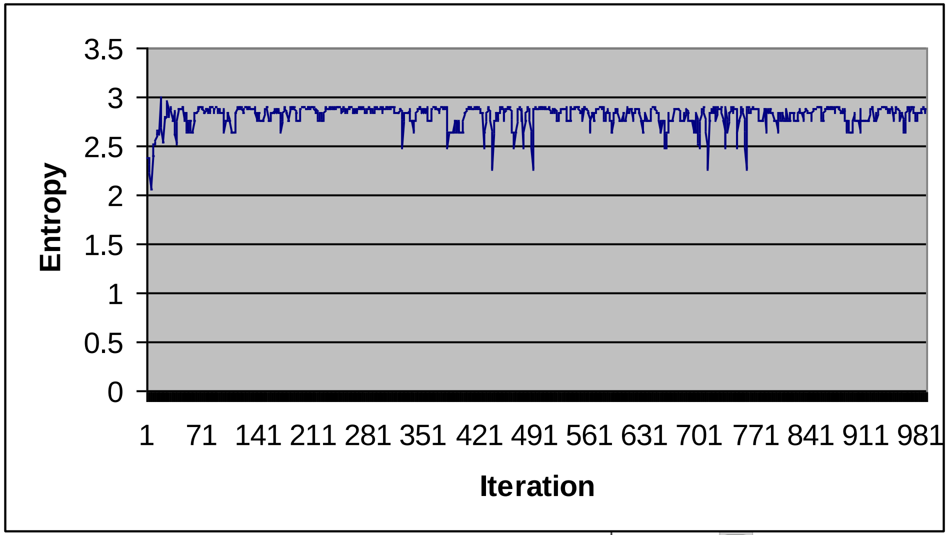
Calculate Entropy

E(1)



Window Size can Also Have a Huge Impact





Method 1

Method 2

14

Command, Control, and Interoperability Center for Advanced Data Analysis

Step Size

We allow windows to overlap. The window might need to ‘walk along’ the data, not just expand to always include more and more history. Step size tells us how continuous the process is (e.g. how much overlap with the last window)

Window Size = 7

Step Size = 1

Incidence Data:

3, 2, 4, 5, 8, 10, 12, 40, 35, 17, 37, 20, 23, 25, 4,…

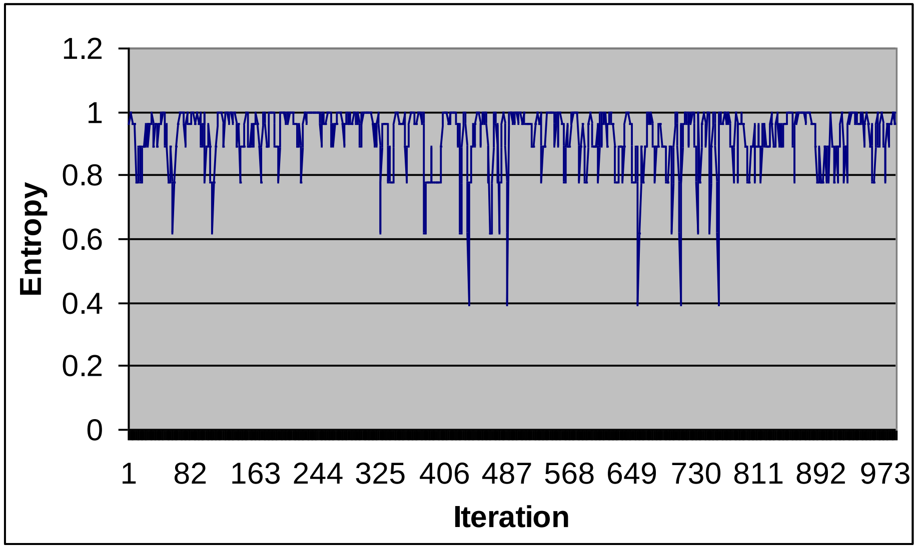
Calculate Entropy

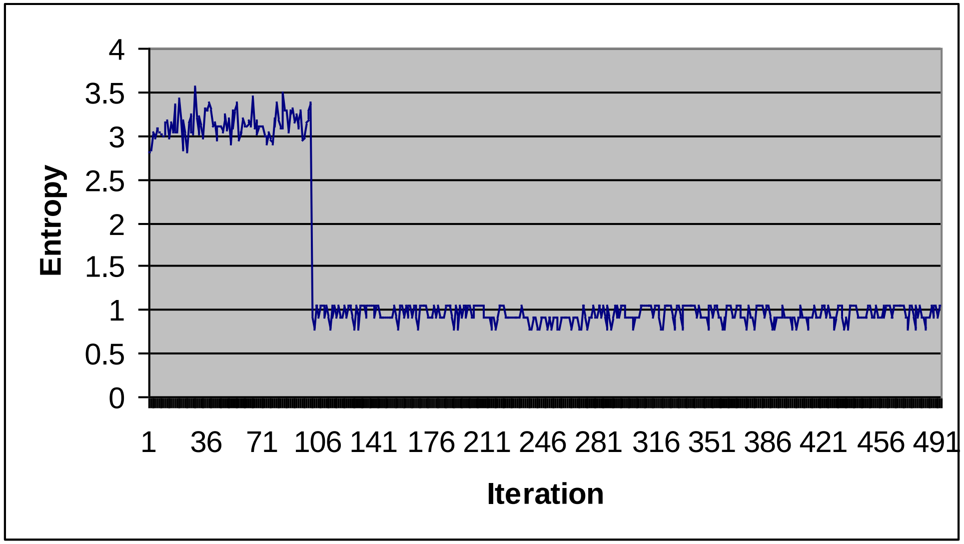
E(1), E(2)

etc. for all the data

Step Size

Adjusting step size can eliminate glitches like weekends and holidays in daily datasets.





Step size = 1

Step size = 5

Command, Control, and Interoperability Center for Advanced Data Analysis

Computing an Entropy Output

We produce a new data stream by doing this over again, walking the window along the binned data, using our step size

Window Size = 6

Step Size = 1

Incidence Data:

3, 2, 4, 5, 8, 10, 12, 40, 35, 17, 37, 20, 23, 25, 4,…

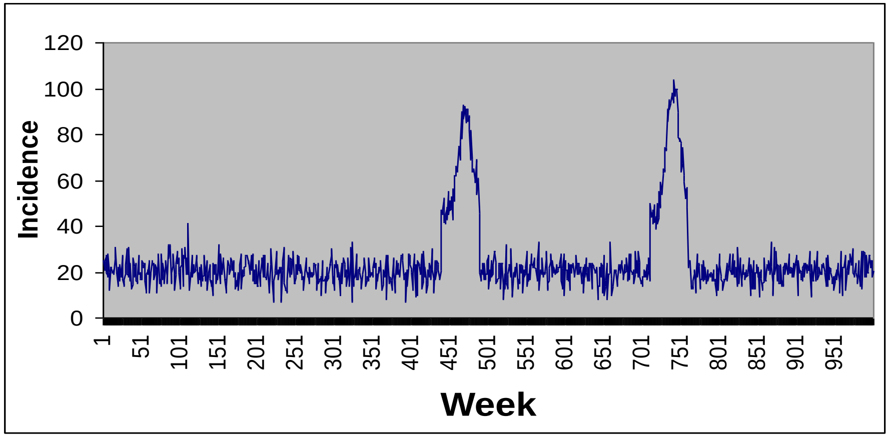
This is the 9th window since step size is 1

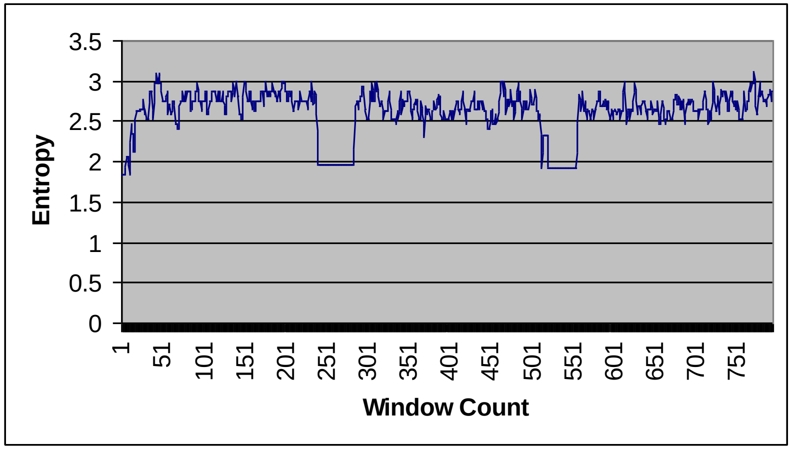
Calculate Entropy

E(1), E(2), E(3), E(4), E(5), E(6), E(7), E(8), E(9), …

Biosurveillance using Entropy

- Our results show this method can work.
- Favorable when compared to CuSum and other methods.





We need more work to test it to make sure it's sensitive and specific enough

Command, Control, and Interoperability Center for Advanced Data Analysis

Social Media for Situational Awareness

•Social media can help to provide situational awareness in case of developing emergencies

•CCICADA projects: Ji and Wallace, Hovy and Metzer - analyzed data from Twitter from Haitian Earthquake of 2010 & Japanese Earthquake & Tsunami of 2011 – over a billion tweets

CCICADA
Command, Control, and Interoperability Center for Advanced Data Analysis

Social Media for Situational Awareness

- Work in these projects has found:
  - Great diversity of communication
  - Interesting characteristics of network spread
  - People coordinate in different ways
  - People follow typical sequences when communicating in emergency situations
- Understanding typical sequence allows crisis responders and others to identify "relapses"

Command, Control, and Interoperability Center for Advanced Data Analysis

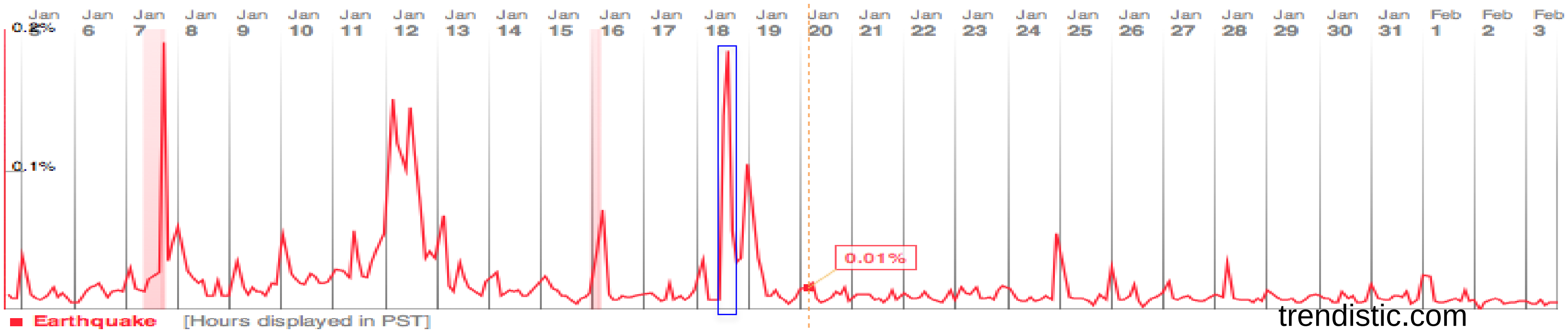
Social Media for Situational Awareness

- Developed tools to learn "topic signatures" that indicate when an event of a given type occurs
  - Discover 'burst' of topic-related words in timespan
  - Identify relevant tweets
  - Extract main ones to build a summary
  - Monitor as event unfolds
  - Pick out anomalies, etc.

CCICADA
Command, Control, and Interoperability Center for Advanced Data Analysis

Social Media Challenge 1. Detect Events

Metzler et al. 2011



Query: q

earthquake

Step 1: q' = EXPAND (q)

Event signature: *earthquake, earthquake., earthquake.., magnitude, epicenter, earthquake.., foreshocks, usgs, tsunami, indonesia, …*

Step 2: TS = RANK_TIMESPAN(q')

"Burstiness" of term *w* in timespan *T*: $\beta_T(w) = \frac{P(w|T)}{P(w)}$

Step 3: SUMMARIZE(TS)

Timespan 1: Jan 18, 2010 at 22 UTC, for 9 hours

Summary tweets:

i.#IWO Latest on Pakistan earthquake tonight: 7.3 mag quake in SW Pakistan http://bit.ly/guYOry

ii."@cnnbrk: USGS: An earthquake with preliminary magnitude of 7.4 strikes southwestern Pakistan http://on.cnn.com/gQcnRa

iii.Major quake hits Pakistan: An earthquake with a preliminary magnitude of 7.4 struck Wednesday morning in southwe… http://bit.ly/iildWP

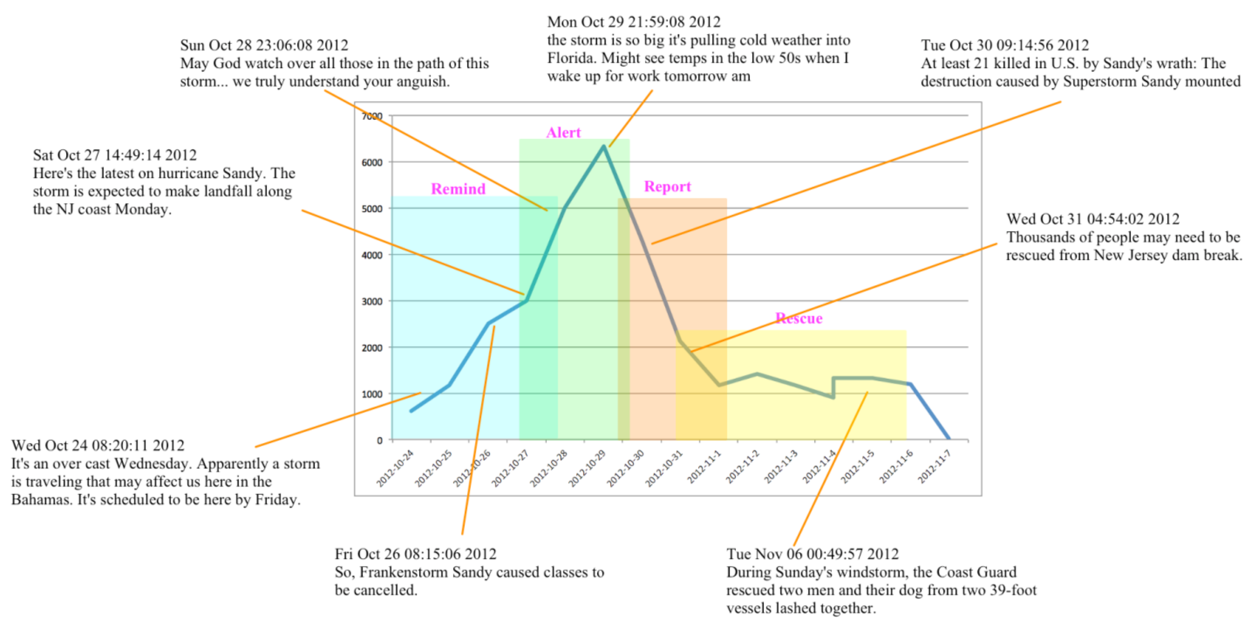
Command, Control, and Interoperability Center for Advanced Data Analysis

Social Media Challenge 2. Track Event Stages

- Model 1: simple linear timeline
  - Earthquake:
    - 3 mins of fear
    - 30 mins of family & friends outreach
    - 3 hours of planning immediate actions
    - 3–5 days of discussion about repair activities
- Model 2: multiple timelines, depending on relationship of tweeters to event
  - It's different if you are there than if you just heard about it

10 topics and 10 top words for each topic (generated by LDA)

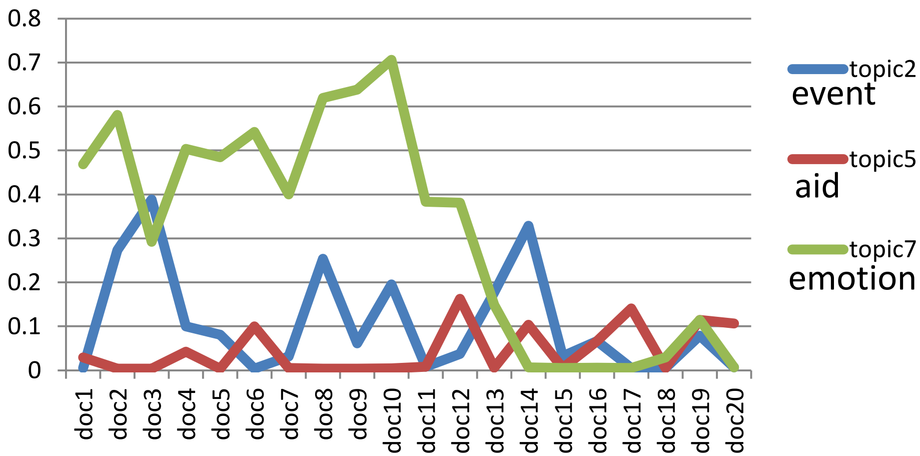
topic0	topic1	topic2	topic3	topic4	topic5	topic6	topic7	topic8	topic9
stupid	hate	magnitude	condolences	tsunami	donate	depth	feel	large	upgraded
supermoon	destroy	damage	bodies	affected	text	epicenter	felt	crack	triggering
search	teen	weird	islands	people	relief	coast	shit	photo	large
total	violence	make	found	hit	redcross	offshore	thought	showing	effects
utter	gangs	started	worst	prayers	victims	miles	shaking	death	unbelievable
offensive	protect	rocked	city	thoughts	red	strikes	tonight	toll	issued
post	depth	hurricane	sweeping	news	cross	survey	big	raise	flash
bollocks	thing	news	living	massive	support	bst	time	police	feed
site	stopped	back	slam	coast	record	region	scary	earth	suffered
response		fault	snow	pray	strongest	revised	back	unleashed	larger
		event		sympathy	aid	location	emotion		



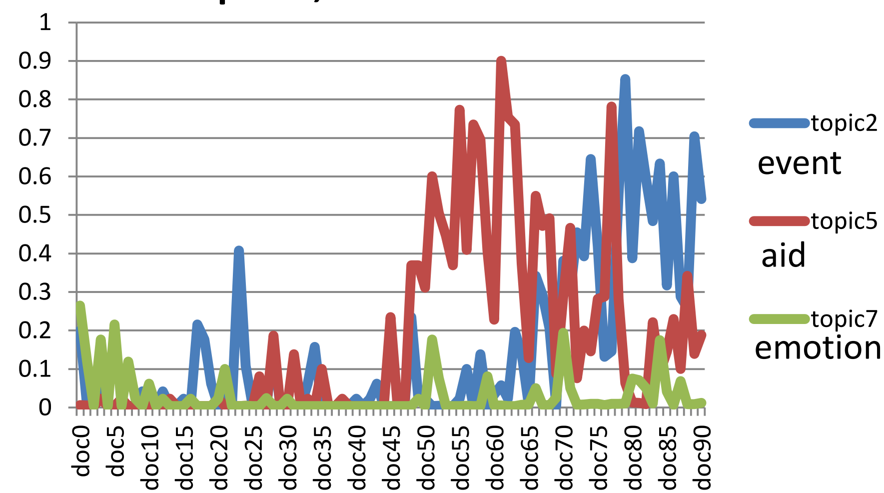
CCICADA
Command, Control, and Interoperability Center for Advanced Data Analysis

Social Media: Two Earthquakes on Twitter

S. California, Jun 2010



Japan, Mar 2011



Each ‘document’ is a bucket of 100 tweets, sorted in time order. Initial discussion about emotions, then focus shifts to aid

Less about emotion: English-language tweeters were not so present in Tokyo.
More about event: implications for Dai-ichi Nuclear Plant

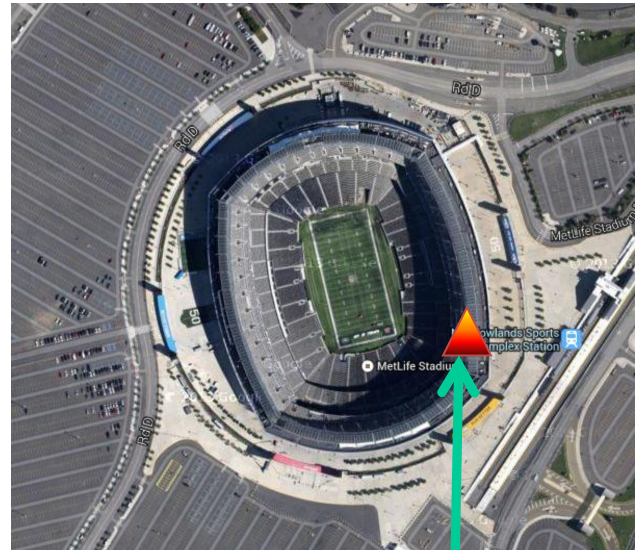
Need to determine details of location and participants of events

Real Time Optimization of Emergency Response Using Social Media

- Another CCICADA Project: Nelson and Pottenger
- Created a framework that:
  - Grouped messages by location (clustering)
  - Determined top requests by location using machine learning (Higher Order Naïve Bayes – HONB – or Higher Order Latent Dirichlet Assn. – HO-LDA)
  - Allocated aid based on integrated social media geolocations requests received
- Applied ideas to social media data from 2010 Haitian Earthquake

Command, Control, and Interoperability Center for Advanced Data Analysis
A Department of Homeland Security Center of Excellence

Real Time Optimization of Emergency Response Using Social Media



911 (cluster)

"ran out of h2o"

After several messages are received, first they are **CLUSTERED**

Supervised method → HONB learned on the cluster → Class: Water; *Predict resource need* → Mixed integer programming resource allocation model

OR

Unsupervised method → HO-LDA learned on the cluster → Topic 1 *(water, liquid, fluid, thirsty)* *Words matched to category*; *with xx% probability associated* → Mixed integer programming resource allocation model

Location *a* needs Water

Messages received; Messages CLUSTERED; Model learned with either HO-LDA or HONB; then predicted request put into RESOURCE ALLOCATION Model

Social Media: Trust

- Motivation:
  - Social media have great potential for early warning of emergencies and situational awareness in them.
  - Key is: tools for analyzing trustworthiness of data from such sources
- Challenges:
  - Data in disaster situations comes from multiple sources: sensors, tweets, speech signals, …
  - Data come from multiple locations, is dynamic and noisy and incomplete as well as redundant
  - Data is inconsistent and even conflicting
  - Want to develop a coherent view of what is happening and where and when

CCICADA
Command, Control, and Interoperability Center for Advanced Data Analysis

Social Media: Trust

- Specific challenge: Trustworthiness of sources (authors, sensors, etc.)
  - Fact-finding algorithms attempt to identify “the truth” among competing claims, but fail to take advantage of a user’s prior knowledge
- New Tools (Roth, et al. – through CCICADA):
  - Computational framework for addressing trustworthiness
  - Based on precise definition of factors contributing to trustworthiness: accuracy, completeness, non-bias.
  - Integrated background knowledge with information extraction
  - Introduced framework for incorporating prior knowledge into any fact-finding algorithm – using first order logic and linear programming
  - Scales well to even large problems

CCICADA
Command, Control, and Interoperability Center for Advanced Data Analysis

Thanks

- National Science Foundation
- DIMACS Staff
  - Bidisha Nag
  - Christine Spassione
- Organizing Committee
  - Lora Billings
  - Carlos Castillo-Chavez
  - Midge Cozzens
  - John Mitchell
  - Al Wallace
- Local Organization: Lisa Redding
- Especially: Eva Lee

MPE: Natural Disasters
How can Math Sciences help?
How can we each help?