Homeland Security: What Can Data Science Do?

Fred Roberts
Director of CCICADA

CCICADA

Command, Control and Interoperability Center for Advanced Data Analysis

A Department of Homeland Security University Center of Excellence Based at Rutgers University

CVADA

- CCICADA is one of two coordinated "halves" of the DHS Center for Visual and Data Analytics (CVADA), founded as the CCI Center by DHS in 2009.
- CCICADA is based at Rutgers University
- CCICADA emphasizes data analysis.
- The other half of CVADA is based at Purdue and emphasizes visual analytics.

Why CCICADA?

- *Virtually all of the activities in the homeland security enterprise require the ability to reach conclusions from large flows of data and require data-driven decision support*

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

Why CCICADA?

- We apply methods of mathematics, computer science, statistics and operations research to problems of homeland security.
- We partner with behavioral scientists, economists, biologists, epidemiologists, physicians, sociologists, industrial engineers, etc.
- Problems we face are fundamentally multidisciplinary.

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

CCICADA Partners

- Alcatel-Lucent Bell Labs
- AT&T Labs - Research
- City College of NY
- Howard University
- Princeton University
- Rensselaer Polytechnic Inst.
- Texas Southern University
- University of Massachusetts, Lowell
- University of Medicine & Dentistry of NJ
- Applied Communications Sciences
- Carnegie-Mellon Univ.
- Geosemble Technologies
- Morgan State University
- Regal Decision Systems
- Rutgers University (Lead)
- Tuskegee University
- University of Illinois, Urbana Champaign
- University of So.California

CCICADA Works with Many Partners

- Federal Emergency Management Agency (FEMA)
- Coast Guard
- Customs and Border Protection (CBP)
- Transportation Security Administration (TSA)
- US Citizenship & Immigration Service
- Immigration & Customs Enforcement (ICE)
- FBI
- CDC
- NJ Office of Homeland Security & Preparedness
- NJ Dept. of Health and Senior Services
- NJ State Police
- LAPD, NYPD
- All major sport Leagues


U.S. Department of Homeland Security
Federal Emergency Management Agency
United States Coast Guard
CDC
CREATE
Homeland Security Center
Domestic Nuclear Detection Office
Office of the Attorney General
State of New Jersey
Department of Justice
Federal Bureau of Investigation
U.S. Customs and Border Protection
U.S. Department of Homeland Security
POLICE
University of Illinois
Urbana-Champaign
1911 • 100 Years • 2011
LAPD
Los Angeles Police Department
Homeland Security
Blue Campaign
UMASS LOWELL
at&t
TUSKEGEE UNIVERSITY
UMDNJ
USC
University of Southern California
1867
HOWARD UNIVERSITY
ILLINOIS
Princeton
DEI SUB NUMINE VIGET
TSU
Texas Southern University
"Excellence in Achievement"
Applied Communication Sciences
MORGAN
STATE UNIVERSITY
Rensselaer
THE STATE UNIVERSITY OF NEW JERSEY
RUTGERS
CCICADA
Command, Control, and Interoperability
Center for Advanced Data Analysis
A Department of Homeland Security Center of Excellence
Carnegie-Mellon University
Pittsburgh Pennsylvania 1900
Geosemble
the City College of New York
REGAL DECISION SYSTEMS,
Alcatel·Lucent
THE PORT AUTHORITY OF NY & NJ
MetLife Stadium
MALL OF AMERICA
OEM
New York City Office of Emergency Management
intuidex
SUPER BOWL XLVI
INDIANAPOLIS 2012
Office of Homeland Security & Preparedness
The Great Seal of the State of New Jersey

CCICADA Application Areas

- Data Science (DS) methods are applicable to a wide variety of homeland security applications
- CCICADA researchers already heavily involved in these with homeland security practitioner partners, e.g.:
  - **Intelligence analysis of text (IC)**
  - **Disease event detection (CDC)**
  - **Port of entry inspection (CBP)**

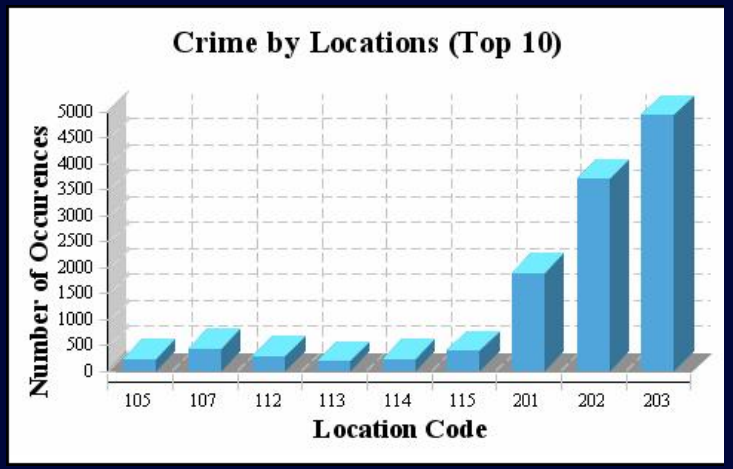

The New York Times.

ALLIED ARMIES LAND IN FRANCE IN THE HAVRE-CHERBOURG AREA; GREAT INVASION IS UNDER WAY

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

CCICADA Application Areas

- DS methods are applicable to a wide variety of homeland security applications
- CCICADA researchers already heavily involved in these with homeland security practitioner partners, e.g.:
  - **Author identification (IC)**
  - **Rapid summarization of crime data (PANYNJ)**
  - **Bioterrorism sensor location (DTRA)**

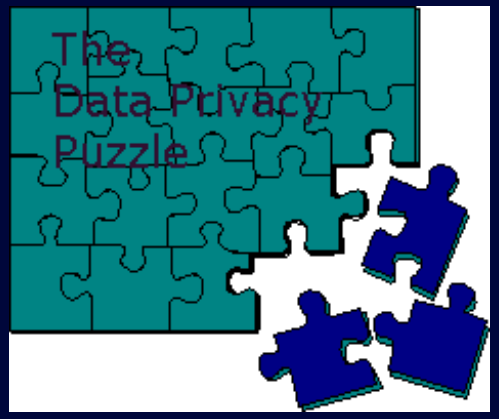


B-T sensor, Salt Lake City Olympics

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

CCICADA Application Areas

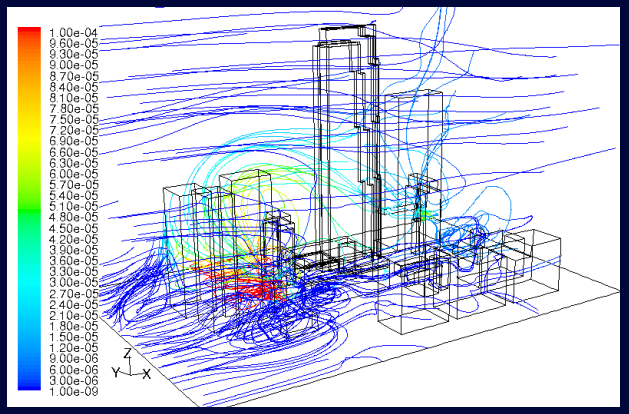
- DS methods are applicable to a wide variety of homeland security applications
- CCICADA researchers already heavily involved in these with homeland security practitioner partners, e.g.:
  - **Nuclear detection with moving detectors (DNDO)**
  - **Risk scoring methods for infrastructure protection (USCG, FEMA)**
  - **Privacy-preserving data sharing**

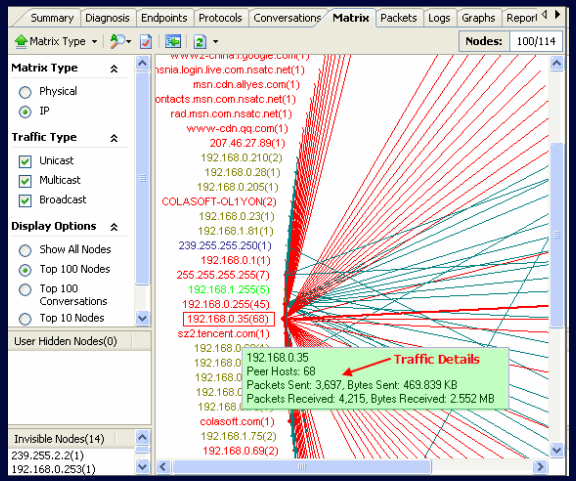


CCICADA
Command, Control, and Interoperability
Center for Advanced Data Analysis
A Department of Homeland Security Center of Excellence

CCICADA Application Areas

- DS methods are applicable to a wide variety of homeland security applications
- CCICADA researchers already heavily involved in these with homeland security practitioner partners, e.g.:
  - **Predicting paths of plumes**
  - **Using social media for alerts & warnings**
  - **Thwarting attacks on cyberinfrastructure by detecting abnormalities in network traffic**





CCICADA Application Areas

- DS methods are applicable to a wide variety of homeland security applications
- CCICADA researchers already heavily involved in these with homeland security practitioner partners, e.g.:
  - **Assessing risks in waterways (NJ DOT)**
  - **Planning for evacuations during heat events from climate change (CDC)**
  - **Understanding economic impact of terrorist events and natural disasters (PANYNJ)**

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

CCICADA Application Areas

- DS methods are applicable to a wide variety of homeland security applications
- CCICADA researchers already heavily involved in these with homeland security practitioner partners, e.g.:
  - **Countering human trafficking (FBI, LAPD, NJ Attorney General)**
  - **Mall security (NJ OHSP)**

- **Developing tools for health emergency situational awareness by first responders (NJ DHSS)**

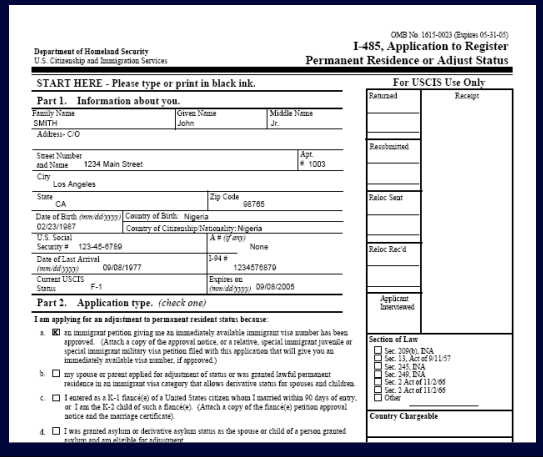
CCICADA
Command, Control, and Interoperability Center for Advanced Data Analysis
A Department of Homeland Security Center of Excellence

CCICADA Application Areas

- DS methods are applicable to a wide variety of homeland security applications
- CCICADA researchers already heavily involved in these with homeland security practitioner partners, e.g.:
  - **Flood mitigation (FEMA)**
  - **Fisheries law enforcement (USCG)**
  - **Tools for stadium security (DHS OSAI, major sports league security)**

CCICADA Application Areas

- DS methods are applicable to a wide variety of homeland security applications
- CCICADA researchers already heavily involved in these with homeland security practitioner partners, e.g.:
  - **Immigration form volumes projection (USCIS)**
  - **Virtual reality games to train first responders**
  - **Natural language processing for early detection of plans for cyber attack**



Department of Homeland Security
U.S. Citizenship and Immigration Services

OMB No. 1615-0023 (Expires 05-31-05)

I-485, Application to Register Permanent Residence or Adjust Status

START HERE - Please type or print in black ink.

For USCIS Use Only

Part 1. Information about you.

Family Name	Given Name	Middle Name
SMITH	John	Jr.

Address- C/O

Street Number and Name	Apt.
1234 Main Street	# 1203

City: Los Angeles

State	Zip Code
CA	98765

Date of Birth (mm/dd/yyyy)	Country of Birth: Nigeria
02/23/1967	Country of Citizenship/Nationality: Nigeria

U.S. Social Security #	A # (if any)
123-45-6789	None

Date of Last Arrival (mm/dd/yyyy)	I-94 #
09/08/1977	1234576879

Current USCIS Status	Expires on (mm/dd/yyyy)
F-1	09/08/2005

Part 2. Application type. (check one)

I am applying for an adjustment to permanent resident status because:

a. ☒ an immigrant petition giving me an immediately available immigrant visa number has been approved. (Attach a copy of the approval notice, or a relative, special immigrant juvenile or special immigrant military visa petition filed with this application that will give you an immediately available visa number, if approved.)

b. ☐ my spouse or parent applied for adjustment of status or was granted lawful permanent residence in an immigrant visa category that allows derivative status for spouses and children.

c. ☐ I entered as a K-1 fiancé(e) of a United States citizen whom I married within 90 days of entry, or I am the K-2 child of such a fiancé(e). (Attach a copy of the fiancé(e) petition approval notice and the marriage certificate.)

d. ☐ I was granted asylum or derivative asylum status as the spouse or child of a person granted asylum and am eligible for adjustment

Returned | Receipt

Resubmitted

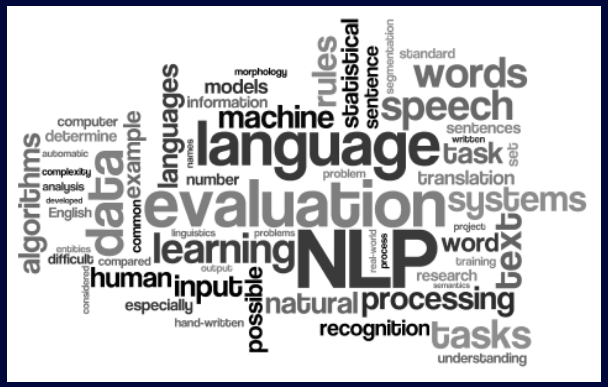
Reloc Sent

Reloc Rec'd

Applicant Interviewed

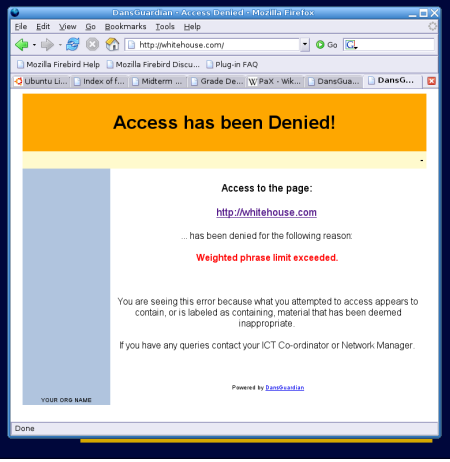
Section of Law
☐ Sec. 209(b), INA
☐ Sec. 13, Act of 9/11/57
☐ Sec. 245, INA
☐ Sec. 249, INA
☐ Sec. 2 Act of 11/2/66
☐ Sec. 2 Act of 11/2/66
☐ Other

Country Chargeable



CCICADA Application Areas

- DS methods are applicable to a wide variety of homeland security applications
- CCICADA researchers already heavily involved in these with homeland security practitioner partners, e.g.:
  - **Coast Guard resource allocation: boat allocation and aircraft allocation (USCG)**
  - **Choosing algorithms for nuclear detection (DNDO)**
  - **Detecting Chinese censorware**



CCICADA Application Areas

- DS methods are applicable to a wide variety of homeland security applications
- CCICADA researchers already heavily involved in these with homeland security practitioner partners, e.g.:
  - **Risk scoring for containers (DNDO)**
  - **Botnet detection using biometrics**
  - **Placing unaccompanied alien children caught at the border (Border Patrol, ICE, HHS)**

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

CCICADA Application Areas

- DS methods are applicable to a wide variety of homeland security applications
- CCICADA researchers already heavily involved in these with homeland security practitioner partners, e.g.:
  - **Communication security for energy delivery systems**
  - **Bio-inspired cyber defense**
  - **Resource allocation in Arctic supply chains (Coast Guard)**

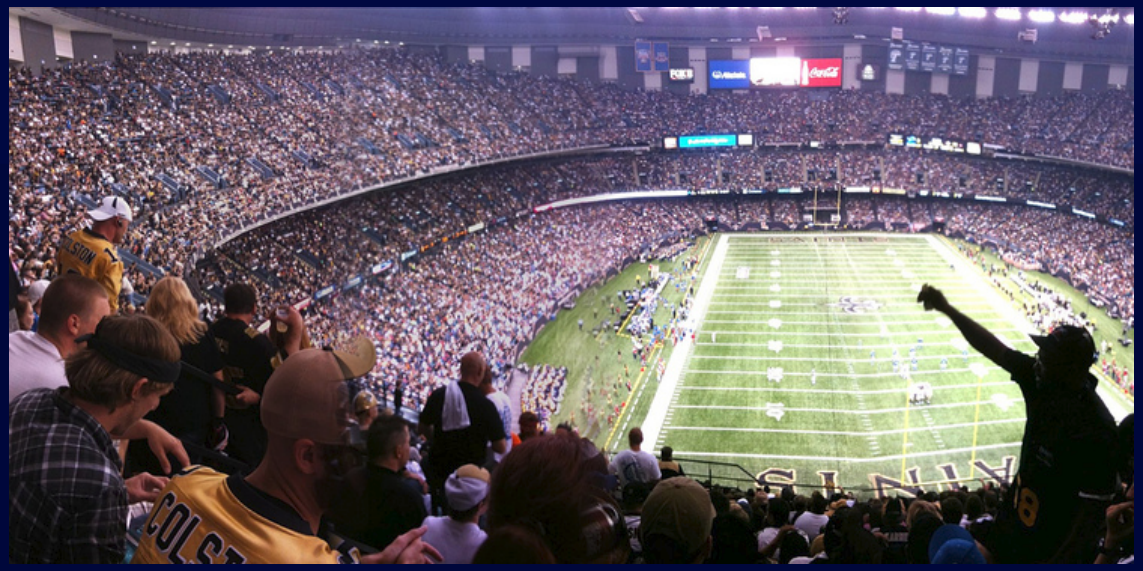
CCICADA
Command, Control, and Interoperability Center for Advanced Data Analysis
A Department of Homeland Security Center of Excellence

Outline

- Sports Stadium Security
- Collaborations with Coast Guard Law Enforcement: Scoring Rules for Fisheries Violations
- Biosurveillance: Early Warning through Entropy
- Inspection at Ports and Borders: Risk Scoring and Anomaly Detection
- Climate and Health
- Maritime Cyber Security

Example 1: Inspections at Sports Stadiums & Large Gathering Places

- Millions of Americans spend time as spectators at sports events.
- How safe are they?



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

Credit kitv.com

Boston, April 15, 2013

Could data science have helped?

Jury selection beginning in trial of Dzhokhar Tsarnaev

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

Could Data Science Have Helped?

- We think so.
- CCICADA has a project on stadium security.
- We work with all major sports leagues (NFL, NBA, NHL, MLB, MLS, USLTA, NASCAR) + college football & basketball + minor league baseball & hockey, etc.

Boston, April 15, 2013

Credit
consortiumnews.com

The day after the Boston attack, we met with security at MetLife Stadium in NJ
Then, we met with NFL Security in NYC

C2ICADA
Command, Control, and Interoperability
Center for Advanced Data Analysis
A Department of Homeland Security Center of Excellence

Example I: Inspections at Sports Stadiums & Large Gathering Places

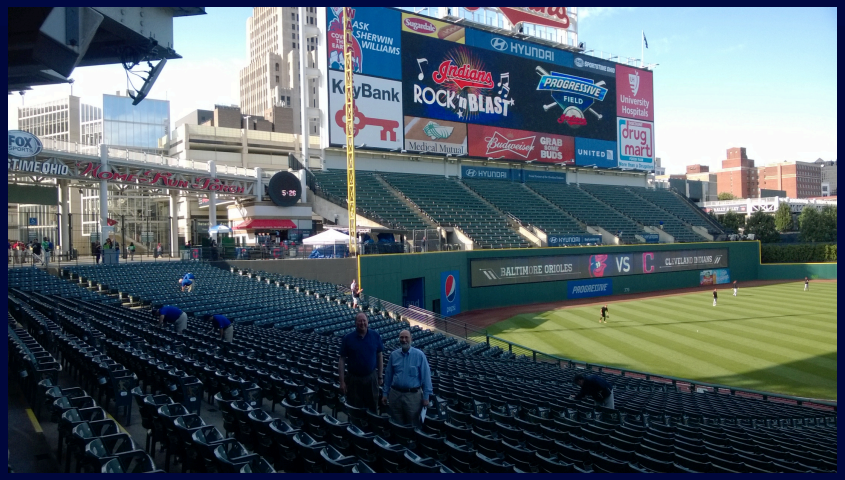
- Earlier work: modeling and simulation of sports stadium evacuation (in collaboration with Regal Decision Systems) led us to close collaborations with National Football League security and stadium operators.
  - Worked with 6 NFL stadiums and several SuperBowls

  - Work applied during lightning storm at MetLife Stadium in NJ

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

Inspections at Sports Stadiums & Large Gathering Places

- Our work has taken us to Progressive Field (Cleveland) for baseball, MetLife Stadium (NJ) for football and world cup soccer friendlies, etc.



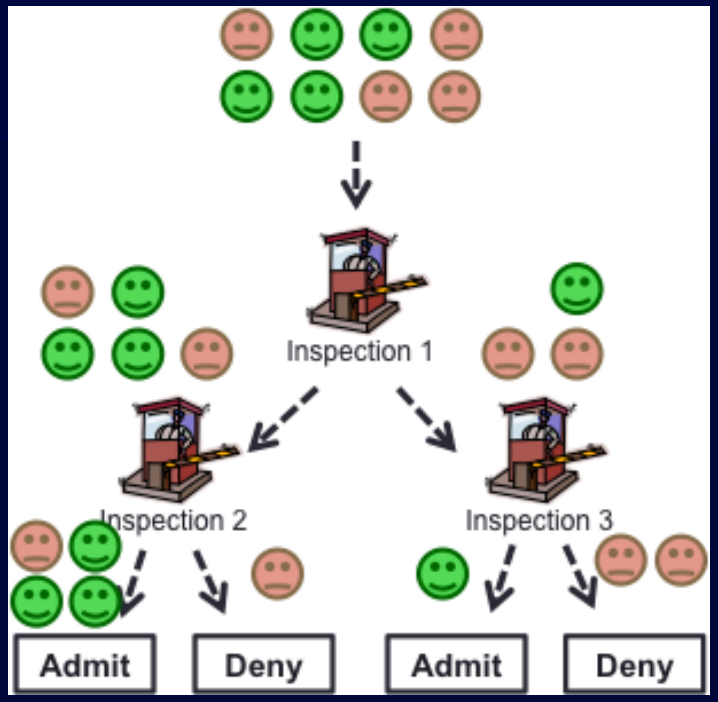
Inspections at Sports Stadiums & Large Gathering Places

- We trained first responders at the 2014 USA Special Olympics in NJ

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

Security at NFL Stadiums

- Working with NFL stadiums
- Looking at variety of inspection problems
- Gathering data about how they do layered defense and building simulation models



Security Project Goals

Improve: Effectiveness, Efficiency & Satisfaction

- Maintain and improve the ***effectiveness*** of patron inspection procedures and processes: identify *contraband* items
- Improve **efficiency**: reduce resource costs (financial, time, staffing, etc.) associated with the procedures/processes; and
- Maintain and improve patron ***satisfaction*** as enhanced procedures are applied to individuals attending stadium events.

Security at NFL Stadiums

- In practice: Started by looking at three types of inspection:
  - ***Wanding***
  - ***Pat-down***
  - ***Bag inspection***

- Observed stadium inspections and gathered data about each type of inspection, in particular length of time it takes.
- Data shows major differences depending on inspector, time before kickoff, etc.

Stadium Inspection

- NFL asked all stadium security operators to perform 100% wanding of patrons.
- This didn't always work. Close to kickoff time, lines got too long.
- Met with NFL Security
- Began analysis of security procedures at one stadium

Approach

- Data Collection, Examination, and Analysis of:
  - Efficiency (inspection times) and Effectiveness (detection of contraband).
  - Comparison of pat-down, wanding, and bag check
  - Anonymous comparison of different inspectors
  - Comparison of different gates
  - Physical design of pods
  - Ticket scanning process and related data
  - Arrival patterns of patrons over time

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

Data Analysis - SUMMARY

- We evaluated the ***effect of several important factors on the inspection times***:
  - **Inspection method** (pat-down, wanding, or bag check)
  - **Location** (gate, pod, lane ~ inspector)
  - **Time before event** (early wave vs. late wave)
    - Early wave = from time of gate opening until waiting line is cleared
    - Late wave = from time of crowd accumulation until event start
  - **Type of event/crowd demographics** (soccer match, monster truck)

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

Data Analysis

- Since there is a lot of (random) variation, we analyzed the results using statistical methods.

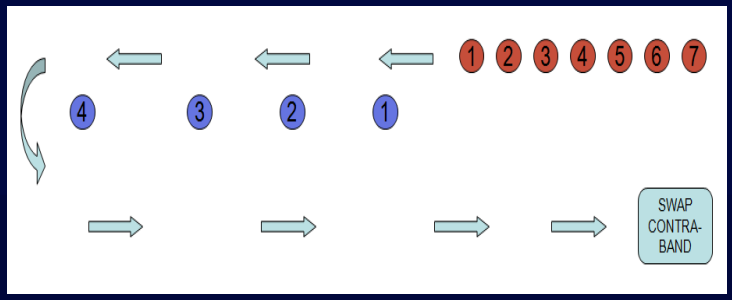
CONCLUSIONS

- Inspection time distributions differ significantly according to
  - Inspection methods
  - Times
  - Inspectors
  - Gates
  - Events
- *Statistical analysis shows that the differences are much greater than can be explained by random chance.*

C4ICADA
Command, Control, and Interoperability
Center for Advanced Data Analysis
A Department of Homeland Security Center of Excellence

Data Analysis: Training

- We designed protocol for evaluating effectiveness of training wanders at a stadium
- We observed training of wanders
- Findings reported to NFL Security



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

Data Analysis

- Our data analysis and analysis of training procedures helped lead NFL to explore alternatives to the wanding strategy.
- Now planning use of walk-through metal detectors.

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

Data Analysis

- We are in the process of analyzing the strategy of going to 100% use of walkthrough metal detectors (WTMDs)
- Issues:
  - ***Project*** number of WTMDs needed to deal with largest expected throughput challenges
  - Observe time required for throughput
  - Model physical location
  - Consider effect of weather on performance
  - Observe how WTMDs work in less than ideal conditions

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

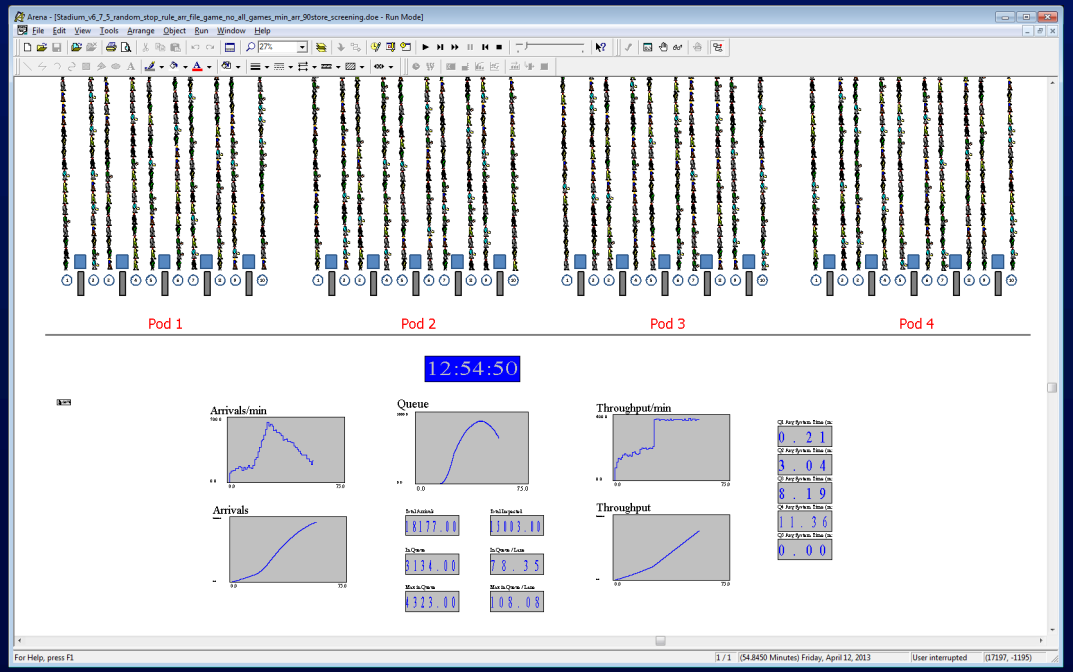
Simulation as a Planning Tool

- **Simulation modeling – strategic planning:**
  - Based on the information obtained from the data collected during in-person observation and video analysis, we have developed a ***simulation*** of entrance queues.
  - Using the data from actual distributions, we have used the simulation to evaluate the speed and cost of inspection for various alternative policies.

The Simulation Model

Most of the parameters can be obtained by ***choosing a representative game***

- Parameters
  - Arrival rates
  - Number of lanes
  - Wanding times
  - Pat-down times
  - Magnetometer times
- Screening Strategy
  - Switching inspection type (Y/N)
    - Number of patrons in queue to switch the process, or
    - Time of switch
  - Does phase 2 include randomization? (Y/N)
    - Ratio of patrons in each type of inspection in the randomization



The model output file includes:

- Total Arrivals
- Total Arrivals @ kickoff
- Maximum number in Queue
- **In Queue @ kickoff**
- **Queue clearance time**
- Screening switch time
- Number of patrons inspected by different procedures
- **Max Waiting Time per patron**

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

Magnetometer Scenarios (Queue Clearance)

No	Game Time	Queue Clearance Times as function of Number of Lanes					
		Base Case (*Wanding & switch to Patdown*)	Magnetometer Scenarios (Number of Lanes)				
		40	20	25	30	35	40
1	9/16/12 1:00 PM	64.65	97.76	83.57	72.18	63.19	56.57
2	10/7/12 1:00 PM	72.79	113.38	95.87	81.07	72.39	64.66
3	10/21/12 1:00 PM	68.67	108.49	92.53	82.13	71.48	65.03
4	11/4/12 4:25 PM	66.80	114.18	94.48	79.75	71.21	61.03
5	11/25/12 8:20 PM	72.40	111.95	94.56	82.52	74.22	65.96
6	12/9/12 4:25 PM	75.40	118.88	99.42	85.81	76.06	67.32
7	12/30/12 1:00 PM	82.67	128.82	108.36	95.27	85.81	76.99
8	9/9/12 1:00 PM	65.46	108.92	89.23	77.64	67.33	58.04
9	9/30/12 1:00 PM	71.33	111.08	94.26	83.39	74.11	65.91
10	10/8/12 8:30 PM	60.80	94.76	76.65	58.19	55.00	55.00
11	10/14/12 1:00 PM	66.50	109.20	91.91	79.01	65.45	55.00
12	10/28/12 1:00 PM	70.82	112.12	93.47	81.09	69.53	61.86
13	11/22/12 8:20 PM	65.94	93.41	79.52	55.12	55.00	55.00
14	12/2/12 1:00 PM	64.45	105.51	91.92	77.06	55.00	55.00

Worse than the Base and does not meet the goal

Similar to Base or better, but does not meet the goal

Meets the goal

Goal: Queue clears by 65 minutes

Future Directions/Next Steps

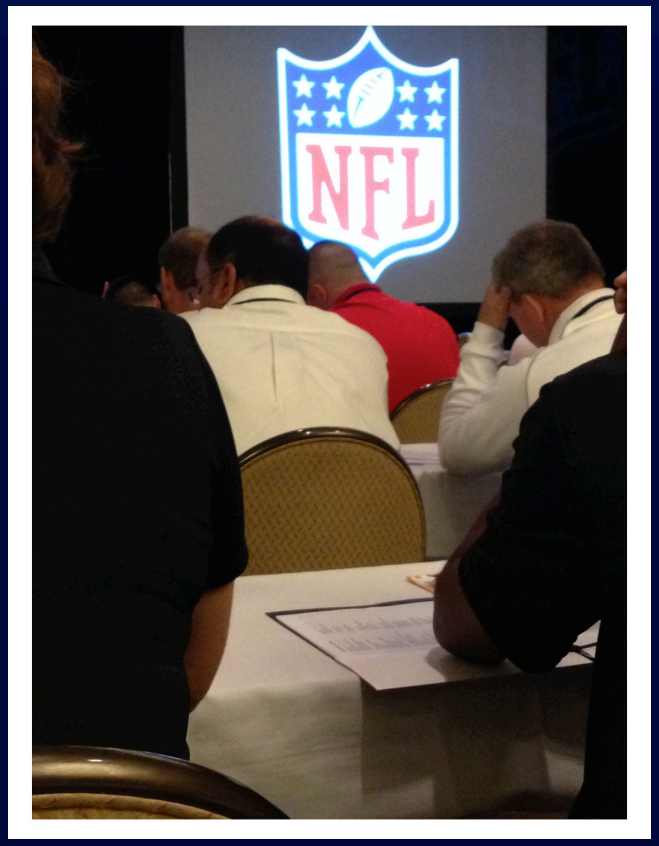
- ***Simple randomizations and how to implement and test in practice:***
  - When 100% inspection is not feasible, is there a randomized inspection scheme that ensures equal or greater security protection and deterrence benefit?
  - Use our simulation model to help with percentages that can be inspected at each stage before kickoff?
  - Is there a way to implement such a scheme that is practical and not subject to being interpreted as profiling?
    - Random beeper
    - Deck of cards
    - Credit card number to present later in the queue

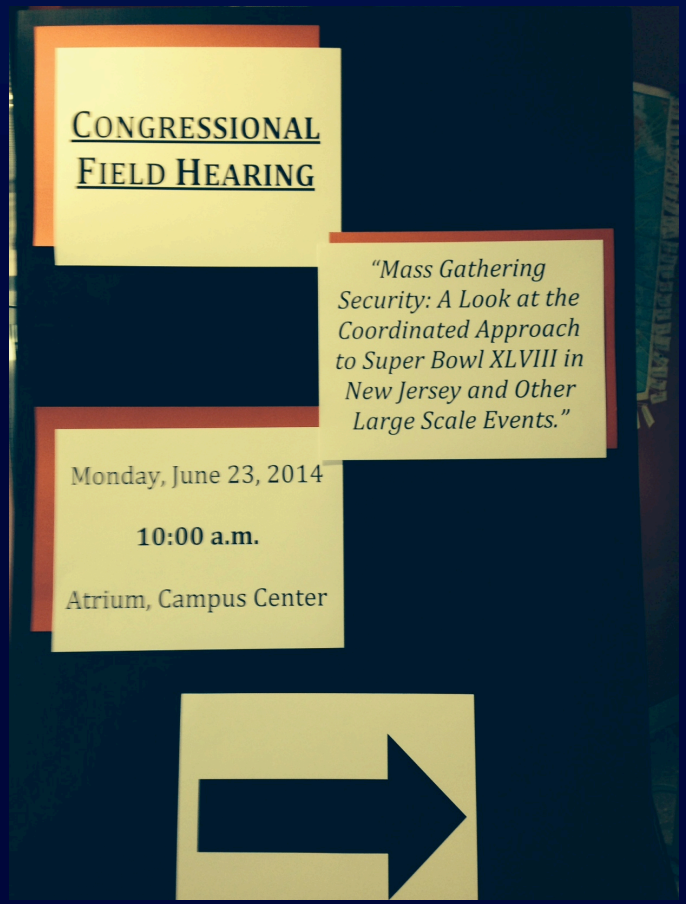
Stadium Inspection

- **Other Key Security Issues:**
  - Identify key components of successful stadium protection plan
    - Access control
    - Credentialing
    - Communications
    - Transportation access
    - Inspection methods
    - Perimeter control
    - Training and exercising
    - Evacuation planning

Stadium Inspection

- Fred Roberts was invited to address the NFL Security Summit in 2014 on the work by CCICADA and to testify on stadium security before a Congressional hearing





Boston April 15, 2013

Credit people.com

So would data science have helped?
We hope it will to prevent the next one.

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

Example 2: Collaborations with Coast Guard Law Enforcement

- We have worked with the Coast Guard on a variety of projects involving information-based modeling and simulation and other advanced data analysis tools

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

Rutgers group touring Port of Philadelphia with Coast Guard Sector Delaware Bay

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

Work with the Coast Guard

- We have served on Sector Delaware Bay Area Maritime Security Committee.
- CCICADA asked to chair the committee.

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

A Selection of Relevant Prior or Current Work at CCICADA

- Accident Risk in Waterways
- Port Operation Simulation
- Container Inspection
- Machine Learning Tools for Manifest Data and Risk Scoring
- Fisheries Law Enforcement
- Port Reopening after Storms/Events
- Leading Indicators of Maritime Safety
- Automatic Identification Systems
- Economic Impact of Coast Guard Decisions
- Boat and Aircraft Allocation

Fisheries Law Enforcement in Coast Guard District 1

- Project sponsor:
  - First District Response Enforcement
- Multidisciplinary team: statisticians, computer scientists, mathematicians, economists, ecologists/ fisheries experts
- Close collaboration: Weekly calls with USCG sponsors
- Similar tools/methods may be relevant to variety of prediction/risk assessment problems

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

Fisheries Law Enforcement

- Coast Guard developed a scoring system (OPTIDE) to determine which commercial fishing vessels to board to look for violations.
- Was having about 20% success rate in finding violations
- Can this be improved by use of sophisticated methods of data analysis?

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

Fisheries Law Enforcement

Some of the Scoring Rules (OPTIDE)

- Points for current Derogatory Report
- Points for past Derogatory Reports (depending upon how many)
- Points depending upon date last boarded
- Points for missing currently valid CG safety sticker
- Points depending upon history of finding fisheries violations in past boardings
- ***Board if total score (number of points) exceeds a threshold***

Fisheries Law Enforcement

- USCG provided us with data from 10,000 boardings and sightings: date, information about the vessel, types of violations found (fisheries, safety)
- We looked at features used by USCG in their scoring methods and also introduced others.

- Scores were not calculated prior to recently, so have to impute the scores.

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

Many Goals of Fisheries Law Enforcement

- While our project was concerned with increasing success rates from boarding, there are many other goals of fisheries law enforcement and we took those into consideration as well:
  - Balanced deterrent
  - Balanced policing
  - Balanced maintenance of safe operations

53

A Variety of Relevant Features

- ***In addition to working with Coast Guard features, we explored introducing other features***, such as:
  - Weather
  - Seasonality
  - Fish migration
  - Key fish species
  - Home port
  - Detailed vessel history

- Economic data (e.g., fish prices)
- Socioeconomic factors (such as type of family boat vs. large commercial fishing boat, or attitudes toward law enforcement)

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

54

The Punchline

- OPTIDE is very clever, works well, and is a significant improvement over random choice of boardings.
- Our new approach called RIPTIDE, developed from sophisticated methods of "machine learning" designed to automatically learn decision rules, could improve on OPTIDE's boarding efficiency by as much as 87%
- Our alternative new approach called DE-OPTIDE, developed using methods of statistical science, uses same "features" as OPTIDE, and could also raise boarding efficiency considerably

55

Machine Learning Approaches

- Looked at machine learning methods to see if other features, or combination of present features and new ones, can lead to decision rules that obtain higher success rate from boardings.
- Represent boarding activities by a set of features
- Aim to learn a classifier that will output “board” or “don’t board” based on the features
- Choosing the features: Combination of data analysis, intuition, and a lot of trial and error

56

Machine Learning Approaches: RIPTIDE

- RIPTIDE is our classifier obtained from USCG data
- RIPTIDE = Rule Induction OPTIDE
- Our best model for RIPTIDE uses some new features, such as type of vessel (General, Trawler, Pot/Trap) and prior violations per boarding
- Much experimentation.
- ***Best model for RIPTIDE found so far outperforms OPTIDE up to 87% in an experiment***

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

Machine Learning Approaches: RIPTIDE

- ***Best model for RIPTIDE found so far outperforms OPTIDE up to 87% in an experiment***
- Experiment:
  - Choose random set of *k* vessels, rank elements according to the model (OPTIDE, RIPTIDE), test whether top-ranked vessel has a violation.
  - If *k* = 30, RIPTIDE does 87% better than OPTIDE
  - If *k* = 10, RIPTIDE does 38.5% better than OPTIDE
- RIPTIDE is best at larger k, maybe larger than what Coast Guard would use; but consider batching vessels and choose best to board
- Riptide software delivered to USCG D1; USCG R&D Center looking to expand to national use

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

- ***Delivering RIPTIDE to Admiral Daniel Abel of Coast Guard District 1, Boston***

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

Logistic Regression Modeling

- Logistic regression allows one to work with the same features as OPTIDE does, but to estimate different weights for those features
- Then we develop a new decision rule for boarding, called DE-OPTIDE (Data Enhanced OPTIDE), with the new weights and different threshold for boarding.

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

Beyond the Predictive (OPTIDE) Model: Alternative Approaches

The optimal strategy is not necessarily to board the boat with the highest predicted probability of being in violation. Here are three reasons why:

1. A desire to check on every boat in the fishery at least once a year

2. Consideration of the time it will take to find and board the vessel

3. Search/wait and improve?

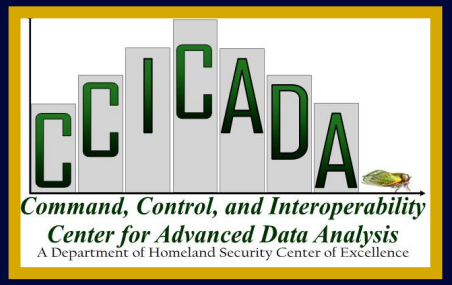
CCICADA
Command, Control, and Interoperability Center for Advanced Data Analysis
A Department of Homeland Security Center of Excellence

Beyond the Predictive (OPTIDE) Model: Alternative Approaches

1. A desire to check on every boat in the fishery at least once a year
   - *Developed model including both probability of finding violation and # days since last boarded*

2. Consideration of the time it will take to find and board the vessel
   - *Explored notion of violations found per hour*

3. Search/wait and improve?
   - *Will a "better" boat be out there or should we inspect now?*



Many Interactions with USCG Continue

Exchanging a ceremonial coin with the Commandant of the US Coast Guard, Admiral Paul Zukunft

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

Example 3: Biosurveillance: Early Warning through Entropy

- Early detection of disease outbreaks critical for public health.
- ***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 ***early*** into the onset of an outbreak

smallpox

Our Ultimate Goal: Effective Biosurveillance

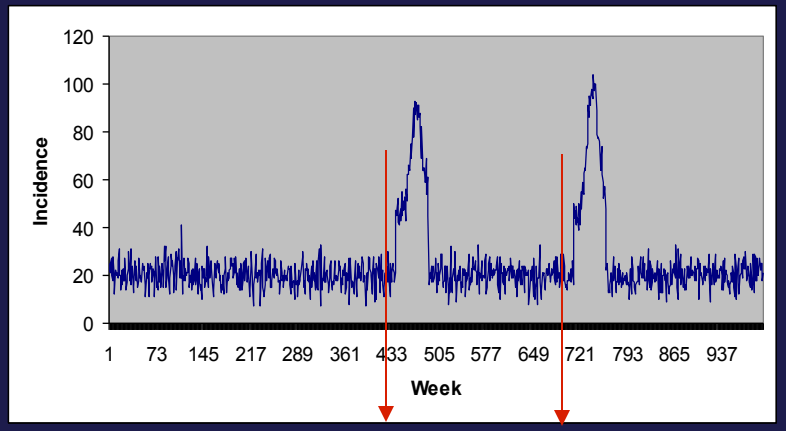
Shannon's Entropy Formula

$H(X) = E(I(X)) = \Sigma_i\, p(x_i) log_2(x_i)$

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

$p(x_i)$ = Prob $(X = x_i)$ 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

65

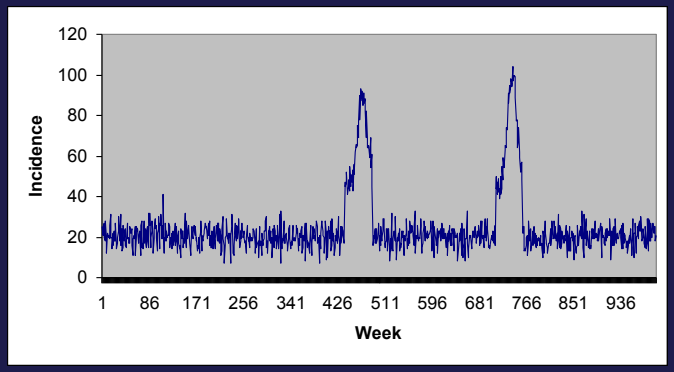
Biosurveillance Methods

- 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

66

Biosurveillance Using Entropy

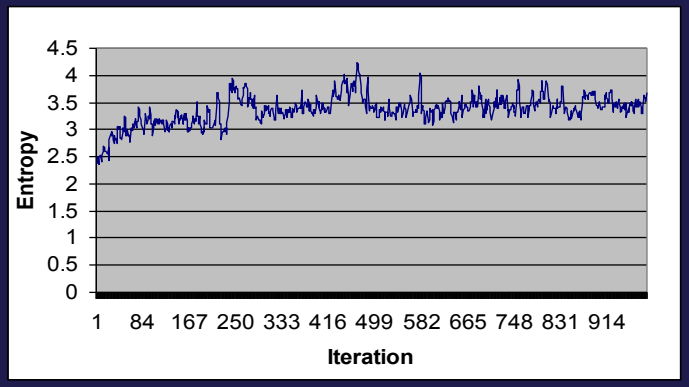
Reported Incidence Data



We apply 3 preprocessing steps

We stream the processed data through our entropy calculation

Entropy Outcome



67

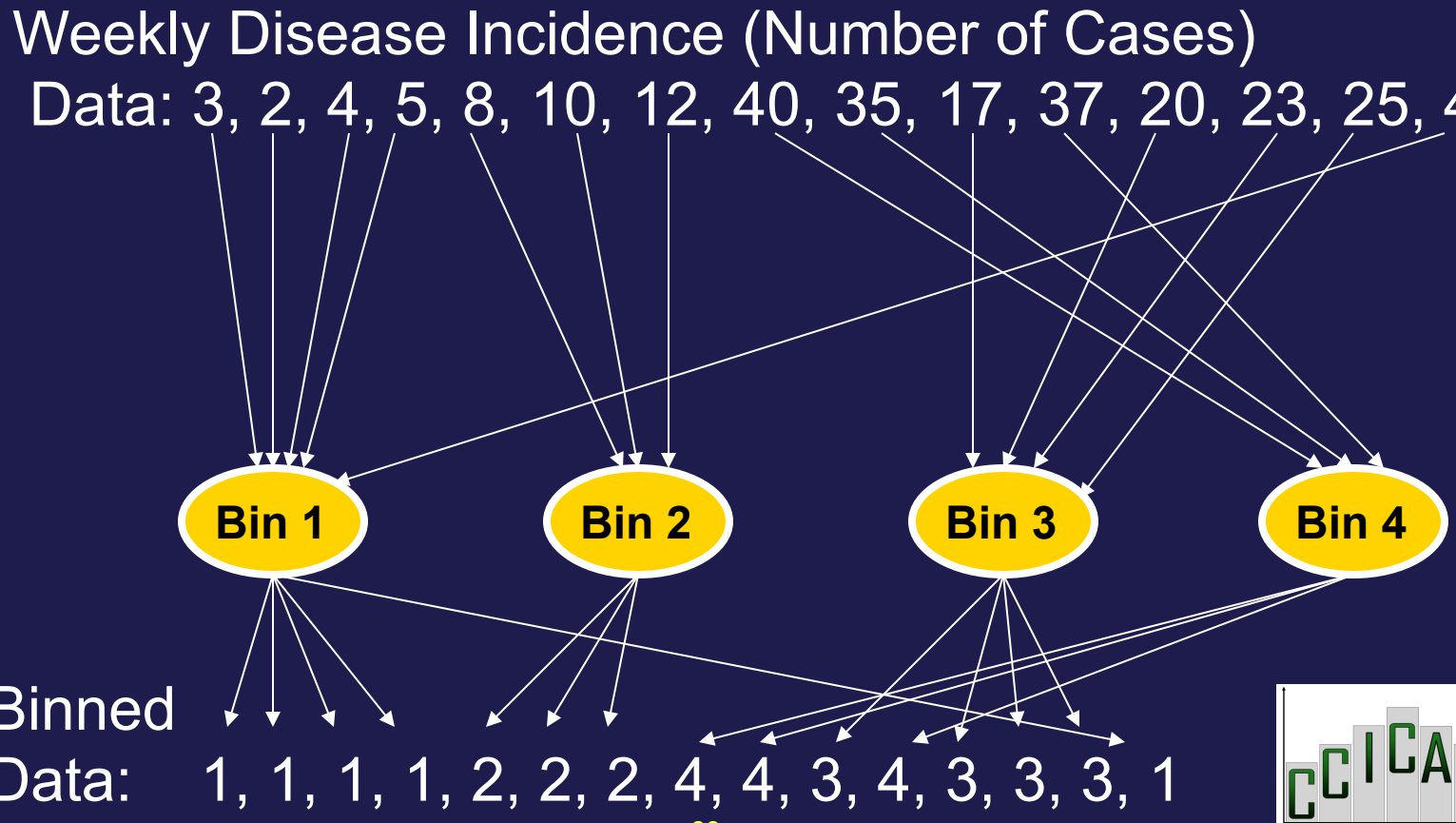
Biosurveillance Using Entropy: The 3 Preprocessing Steps

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**

68

Binning

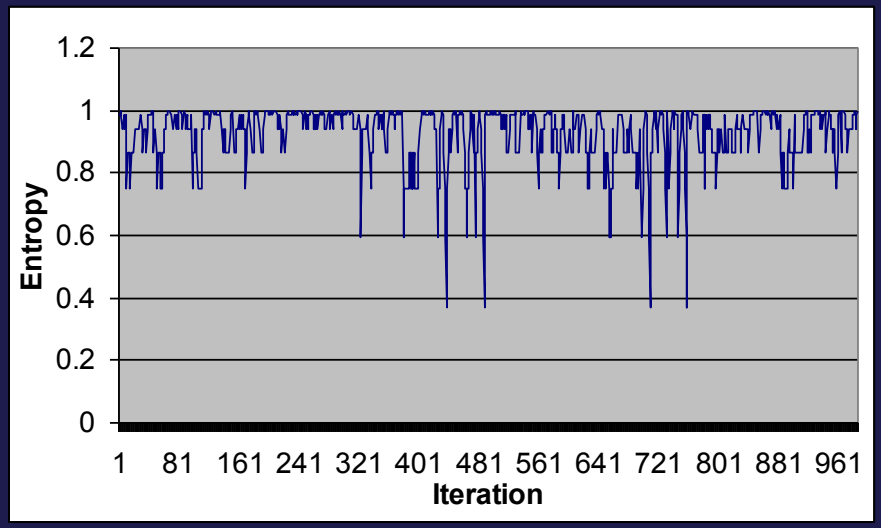
- Assign each "count" to a bin or category.
- Binning lets us try to focus on ***biologically meaningful differences.***



69

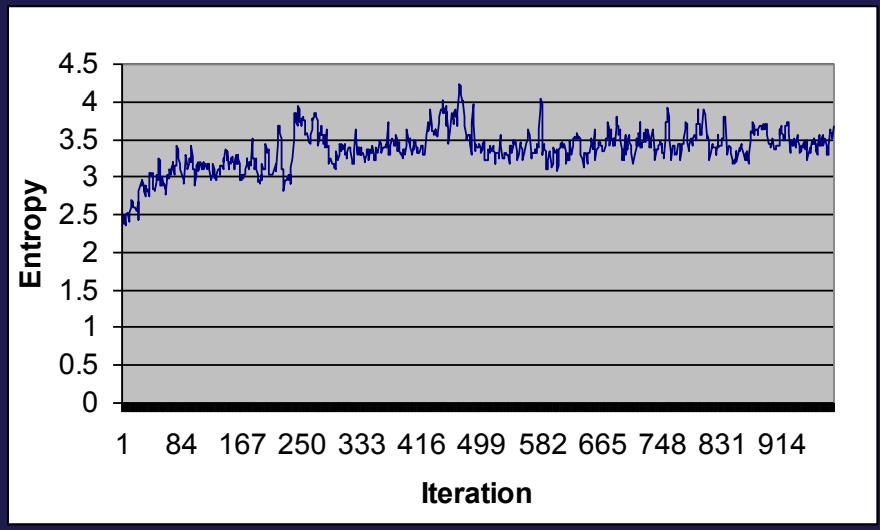
Method of Binning can Really Change the Outcome

Method 1



Number of bins = 2

Method 2



Number of bins = 14

70

Command, Control, and Interoperability Center for Advanced Data Analysis

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

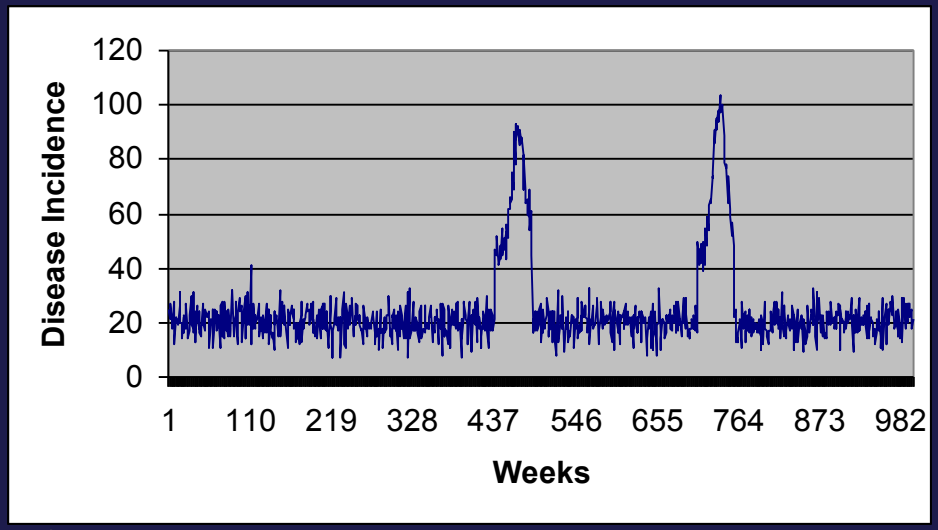
Step Size = 1

Incidence Data:

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

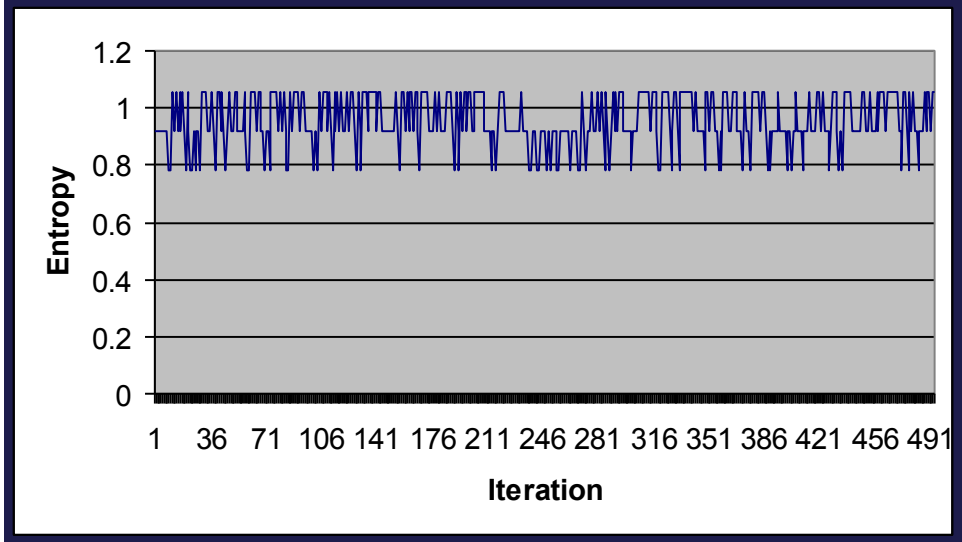
Calculate Entropy

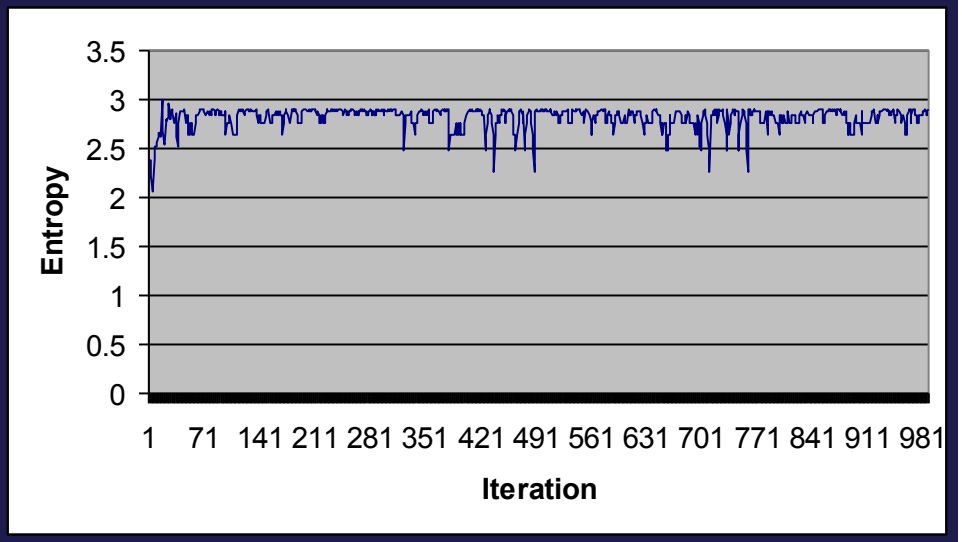
E(1)



71

Window Size can Also Have a Huge Impact





Method 1

Window size = 3

Method 2

Window size = 14

Command, Control, and Interoperability Center for Advanced Data Analysis

72

Step Size

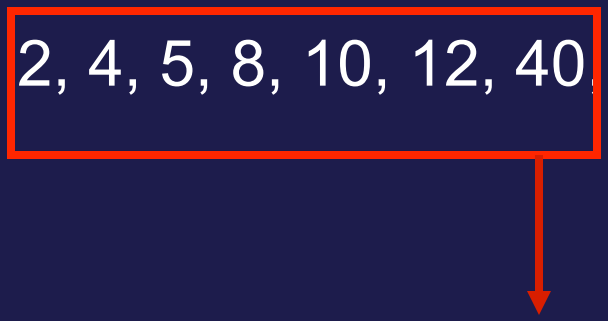
We allow windows to overlap. The window might need to ‘walk along’ the data, not just expand to always include more and more history. Step size tells us how continuous the process is (e.g. how much overlap with the last window)

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)

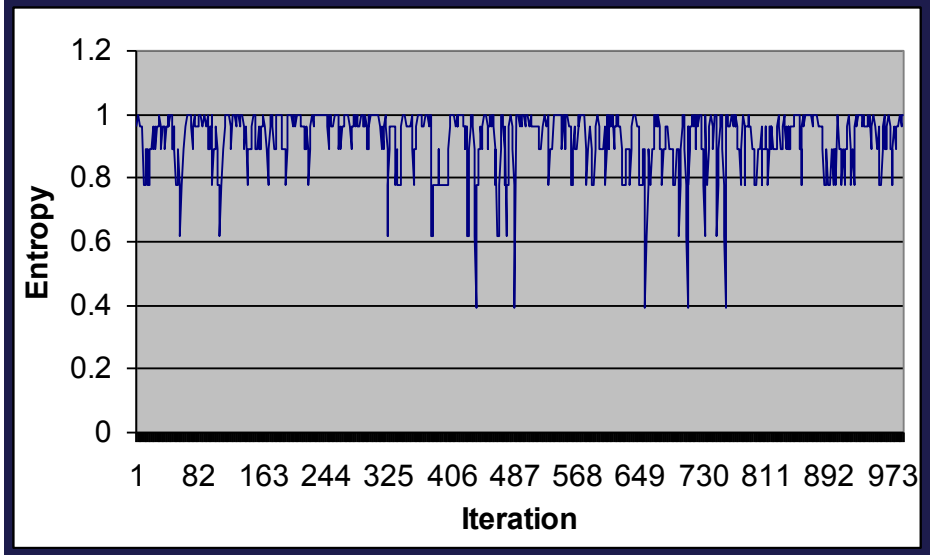
etc. for all the data

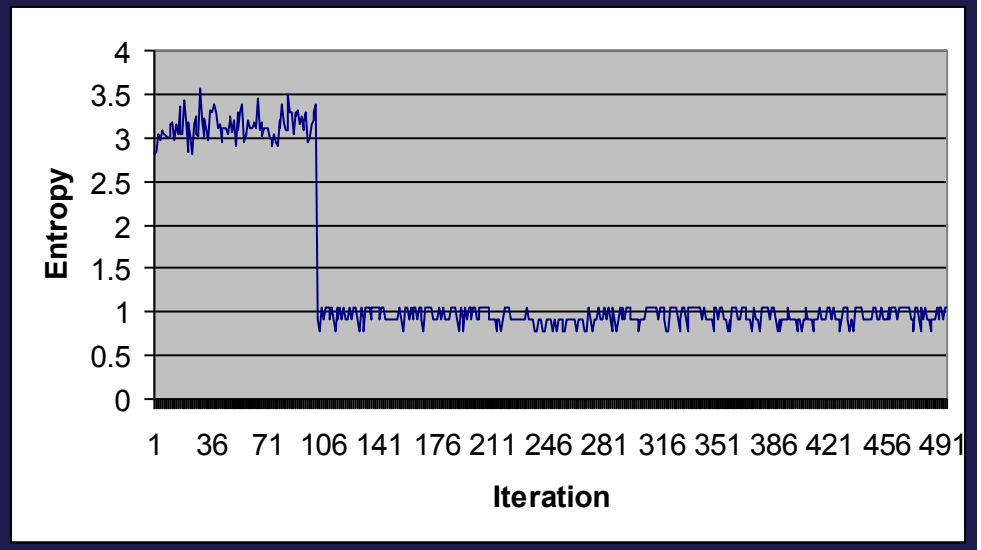
73

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

Step Size

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





Step size = 1

Step size = 5

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

74

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

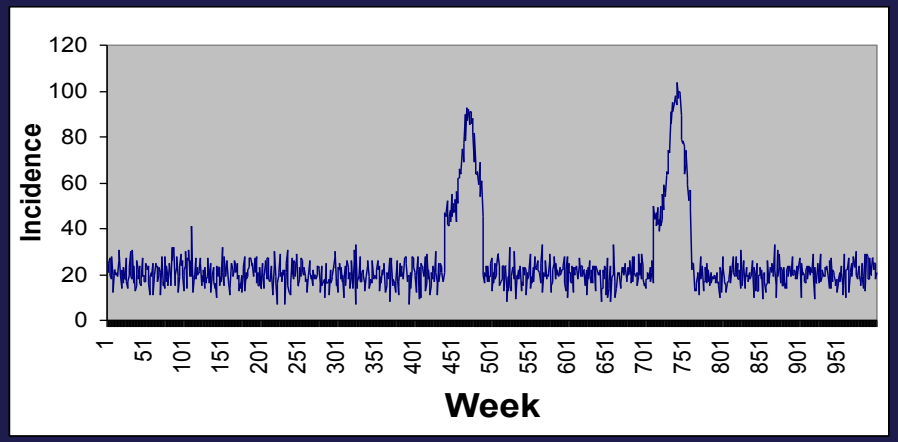
Calculate Entropy

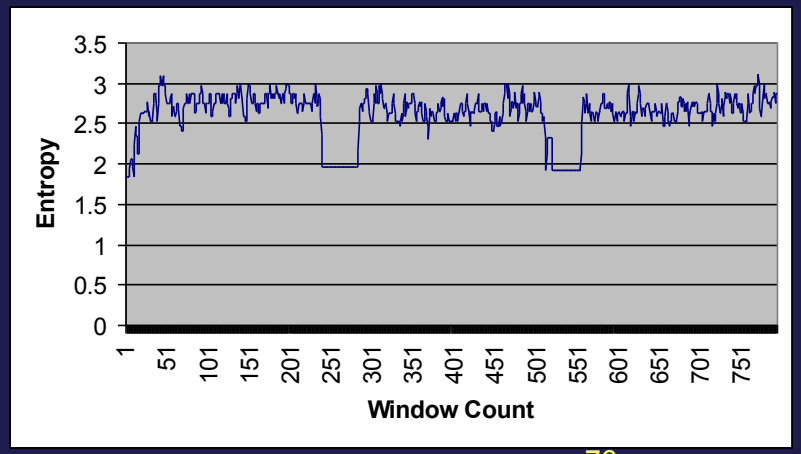
E(1), E(2), E(3), E(4), E(5), E(6), E(7), E(8), E(9),

75

Biosurveillance using Entropy

- Our preliminary 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

76

Command, Control, and Interoperability Center for Advanced Data Analysis

Biosurveillance using Entropy

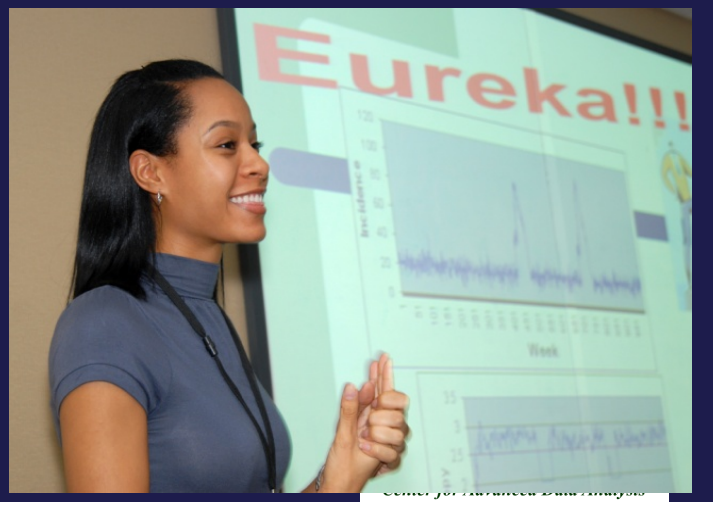
Next step: Make selection of preprocessing parameters automatic

- Right now, all parameters (number of bins, how to bin, how large the window size, how long the step size) is all determined manually by trial and error
- To make this useful for actual surveillance, we are working to design algorithms to select optimal parameters for these three preprocessing steps based on small samples of training data and known outbreak definitions

77

Biosurveillance using Entropy

- Student/faculty teams very involved in this project.
- This is a team combining students/faculty from our partners at Howard University and Morgan State University
- Ashley Crump from Howard presented the project to the Deputy Undersecretary of DHS



78

Example 4: Inspection at Ports and Borders

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

Examples of Work on Inspection at Ports and Borders

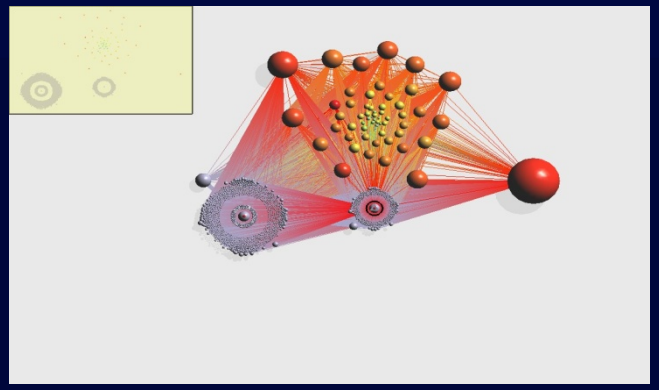
A. Algorithms for more efficient container inspection (Office of Naval Research)
B. Modeling port operations (CBP)
C. Analysis of new off-site container inspection protocols (CBP)
D. Risk scoring of containers (DNDO)
E. Modeling of inspection of arriving international passengers at the airport (CBP)

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

Risk Scoring of Containers

- Developing Tools for Risk Assessment and Anomaly Detection
  a. Machine Learning Tools
  b. Visualization of Data

Project supported by DNDO

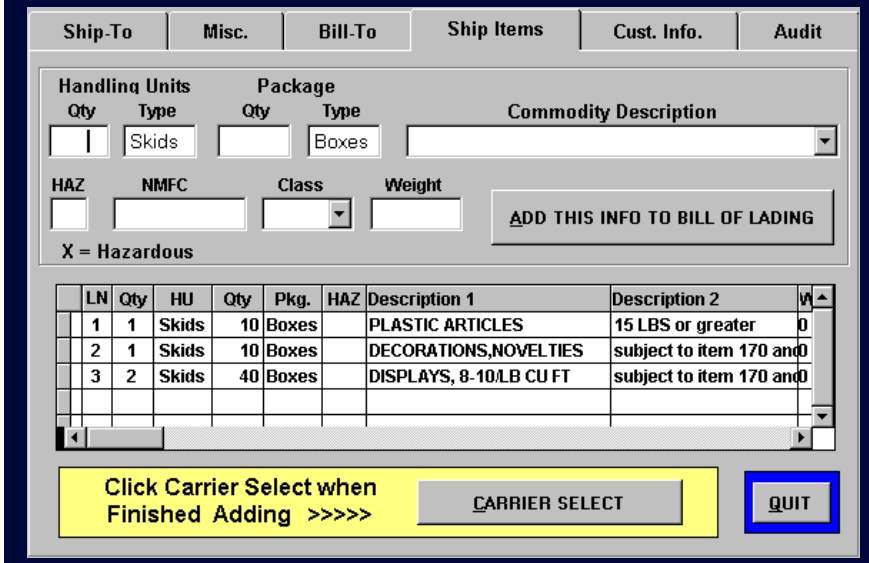


Visualization of Port to Port Shipments

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

Manifest Data

- We obtained from CBP one month's data consisting of manifests for all cargo shipments to all US ports
- Jan 30, 2009 – Feb 28, 2009
- Later obtained more data so could compare effect of Japanese tsunami



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

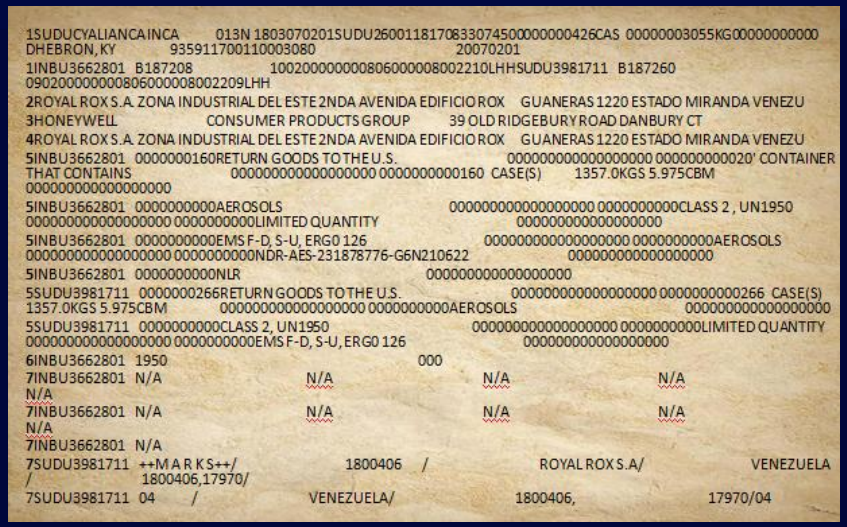
Mining of Manifest Data

- ***Goal: Predict risk score for each container***
  - Quantify the likelihood of need for inspection
  - ***Based on covariates/characteristics of a container's manifest data.***
- Methods:
  - ***We have developed machine learning algorithms to detect anomalies in manifest data.***
  - Text mining on verbiage fields leads to useful characteristics.
  - Then regression based on the useful characteristics or "covariates"
  - "Penalized regression" using LASSO and Bayesian Binary Regression software developed by our group gives us a risk score for a container.

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

Manifest Data

- We obtained from CBP one year's data consisting of manifests for all cargo shipments to all US ports from container ships – every Wed.
- Goal: Identify mislabeled or anomalous shipments through scrutiny of manifest data
- Goal: compare effect of Japanese tsunami



```
1SUDUCYALIANCA INCA        013N 1803070201SUDU2600118170833074500000000426CAS 00000003055KG00000000000
DHEBRON, KY          935911700110003080                    20070201
1INBU3662801 B187208                1002000000008060000008002210LHHSUDU3981711 B187260
0902000000008060000008002209LHH
2ROYAL ROX S.A. ZONA INDUSTRIAL DEL ESTE 2NDA AVENIDA EDIFICIO ROX  GUANERAS 1220 ESTADO MIRANDA VENEZU
3HONEYWELL                  CONSUMER PRODUCTS GROUP        39 OLD RIDGEBURY ROAD DANBURY CT
4ROYAL ROX S.A. ZONA INDUSTRIAL DEL ESTE 2NDA AVENIDA EDIFICIO ROX  GUANERAS 1220 ESTADO MIRANDA VENEZU
5INBU3662801 0000000160RETURN GOODS TO THE U.S.                  0000000000000000000 000000000020' CONTAINER
THAT CONTAINS                0000000000000000000 0000000000160 CASE(S)       1357.0KGS 5.975CBM
000000000000000000
5INBU3662801 0000000000AEROSOLS                                  0000000000000000000 0000000000CLASS 2 , UN1950
0000000000000000000 0000000000LIMITED QUANTITY                   000000000000000000
5INBU3662801 0000000000EMS F-D, S-U, ERG0 126                    0000000000000000000 0000000000AEROSOLS
0000000000000000000 0000000000NDR-AES-231878776-G6N210622        000000000000000000
5INBU3662801 0000000000NLR                                       000000000000000000
5SUDU3981711 0000000266RETURN GOODS TO THE U.S.                  0000000000000000000 0000000000266 CASE(S)
1357.0KGS 5.975CBM           0000000000000000000 0000000000AEROSOLS                  000000000000000000
5SUDU3981711 0000000000CLASS 2, UN1950                           0000000000000000000 0000000000LIMITED QUANTITY
0000000000000000000 0000000000EMS F-D, S-U, ERG0 126             000000000000000000
6INBU3662801 1950                                                000
7INBU3662801 N/A                  N/A                 N/A                 N/A
N/A
7INBU3662801 N/A                  N/A                 N/A                 N/A
N/A
7INBU3662801 N/A
7SUDU3981711 ++M A R K S++/              1800406   /          ROYAL ROX S.A/          VENEZUELA
/            1800406,17970/
7SUDU3981711 04      /          VENEZUELA/          1800406,          17970/04
```


84

Manifest Data

- Test of our risk scoring methods: looked at manifest data from before and after the Japanese tsunami. Expected to find differences.

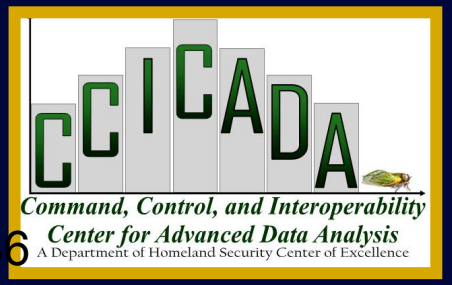
Credit: National Geographic News

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

85

Manifest Data

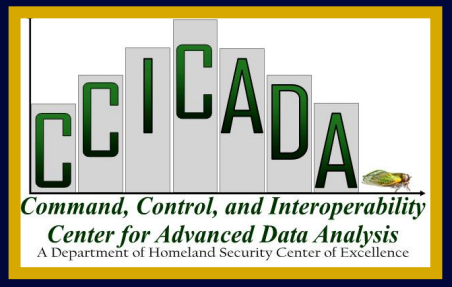
- We used statistical analysis tools (Poisson regression) to detect patterns or time trends of important variables.
- Found that pattern of frequency data based on "domestic port of unlading" is statistically different before and after the tsunami.
- But the pattern based on distribution of carrier is not
- ***Conclusion: Don't depend on just one variable to uncover anomalies.***



86

Machine Learning & Manifest Data: Visualization Tools

- Visualizing data can give us insight into interconnections, patterns, and what is “normal” or “abnormal”
- Our visual analysis methods are based on tools originally developed at AT&T for detection of anomalies in telephone calling patterns – e.g., quick detection that someone has stolen your AT&T calling card.
- The visualizations are interactive so you can “zoom” in on areas of interest, get different ways to present the data, etc.

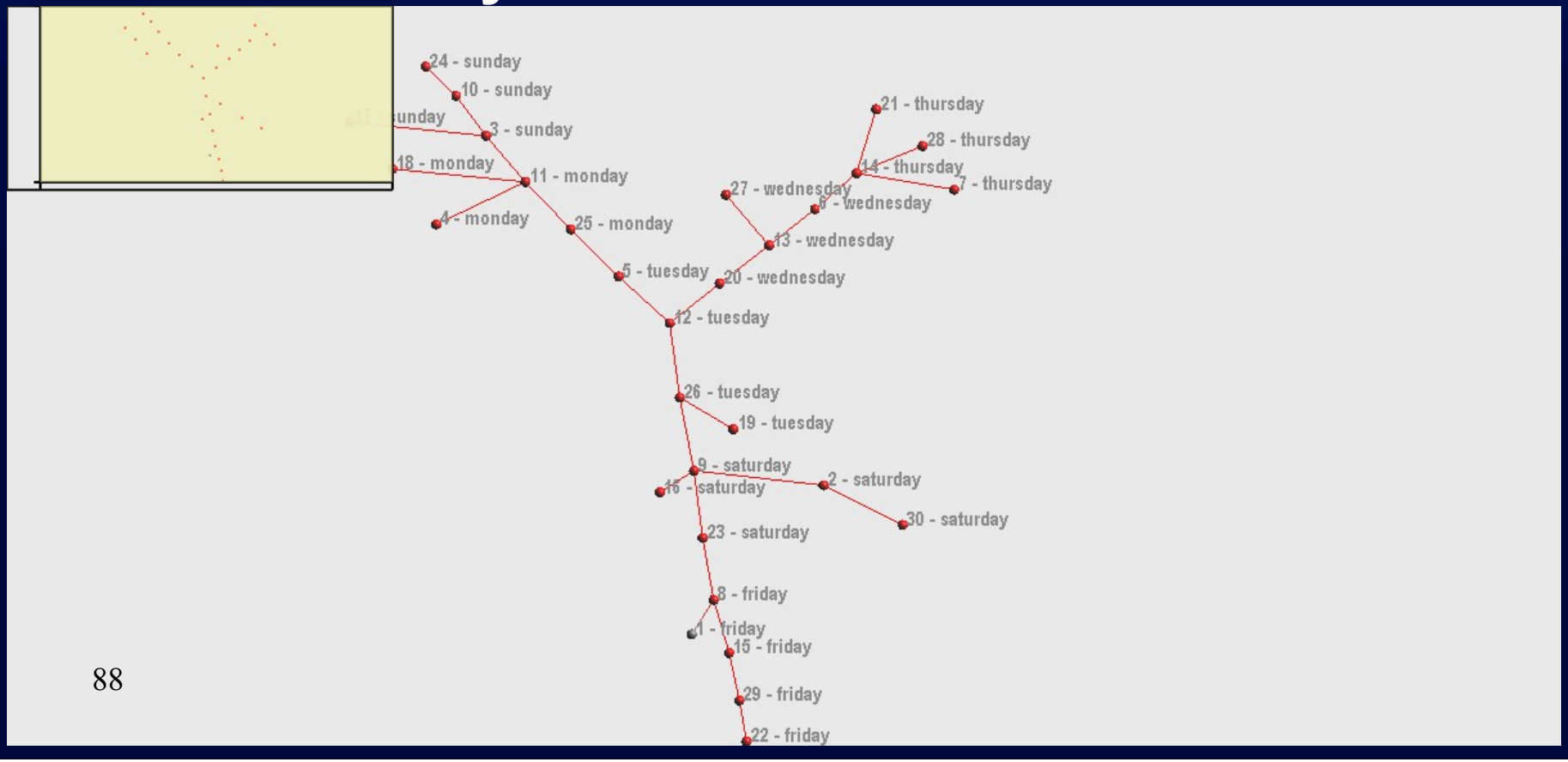


Temporal Evolution of Manifest Data

Fix a commodity.

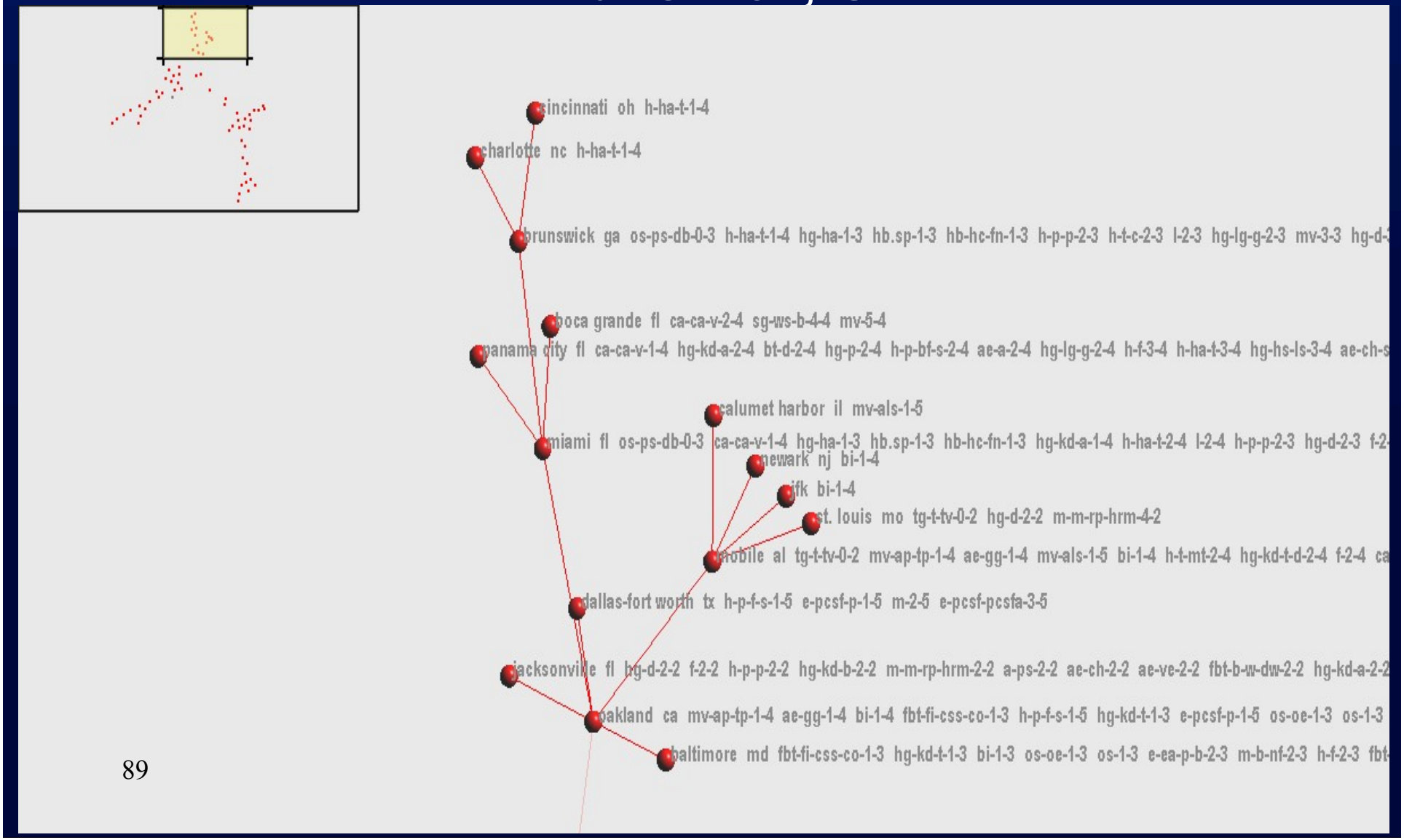
Each node represents all shipments from foreign to US ports on a given day.

Cluster by similarity. Notice how all Tuesdays and Wednesdays are well clustered



Can also Cluster by Ports

Note similarity, e.g., Cincinnati, OH and Brunswick, GA



Port of Entry Inspection Algorithms

- Aim: Develop decision support algorithms that will help us to "optimally" intercept illicit materials and weapons subject to limits on delays, manpower, and equipment
- ***Find inspection schemes that minimize total "cost" including "cost" of false alarms ("false positives") and failed alarms ("false negatives")***

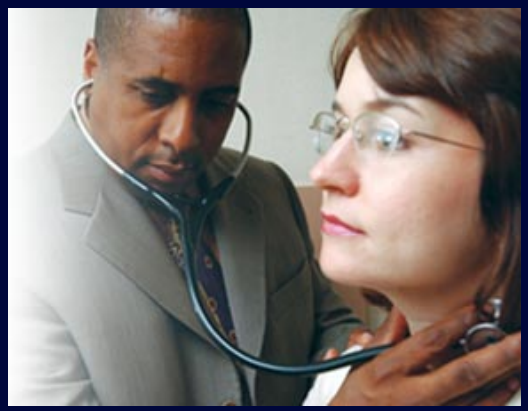
CCICADA
Command, Control, and Interoperability Center for Advanced Data Analysis
A Department of Homeland Security Center of Excellence

Inspection Algorithms: Sequential Decision Making Problem

- Stream of containers arrives at a port
- **The Decision Maker's Problem**:
  - Which to inspect?
  - Which inspections next based on previous results?
- **Approach**:
  - "***decision logics***"
  - ***combinatorial optimization methods***
  - Builds on ideas of Stroud and Saeger at LANL

Inspection Algorithms: Sequential Diagnosis Problem

- Such ***sequential diagnosis problems*** arise in many areas:
  - Communication networks (testing connectivity, paging cellular customers, sequencing tasks, …)
  - Manufacturing (testing machines, fault diagnosis, routing customer service calls, …)
  - Medicine (diagnosing patients, sequencing treatments, …)



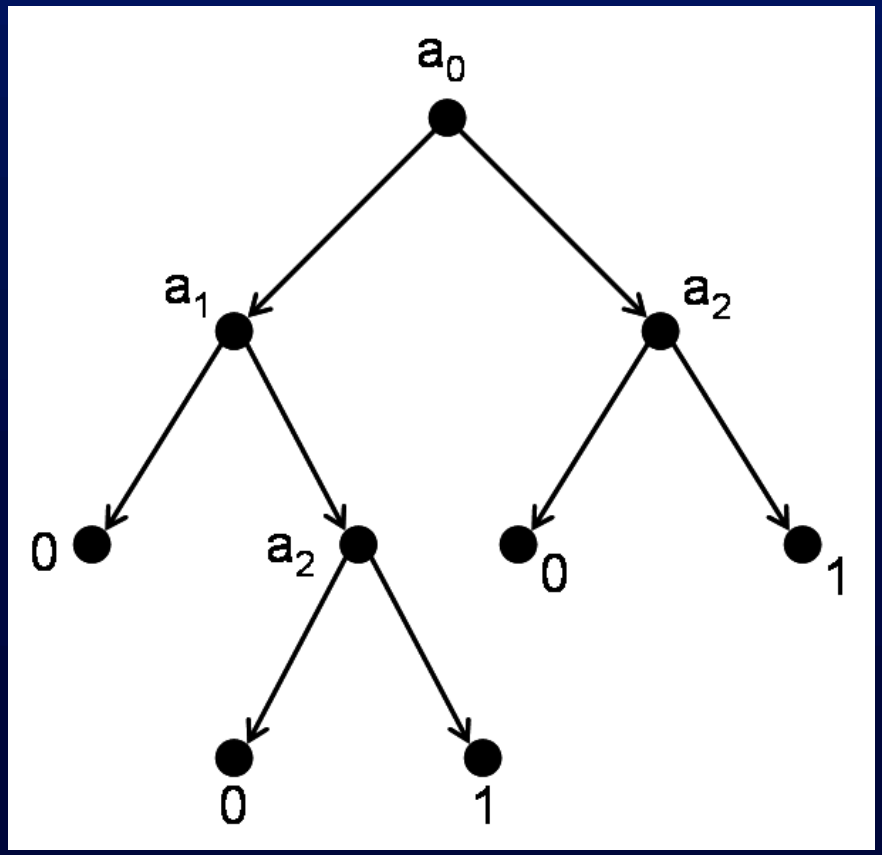
CCICADA
Command, Control, and Interoperability Center for Advanced Data Analysis
A Department of Homeland Security Center of Excellence

Binary Decision Tree Approach

- a_i = i^{th} type of test
- Category 1 = suspicious, 0 =ok
- Go left if pass test a_i, right otherwise
- Reach category 1 from the root by:

a_0 L to a_1 R a_2 R 1 or
a_0 R a_2 R1

- ***Container classified in category 1 iff it has a_1 and a_2 and not a_0 or a_0 and a_2 and possibly a_1.***
- Corresponding decision function:
- $F(111) = F(101) = F(011) = 1$, $F(abc) = 0$ otherwise.



Binary Decision Tree Approach

- Finding the "least cost" binary decision tree is computationally intractable once the number n of types of tests gets too large.
- Stroud and Saeger (LANL) limit the possible decision functions to ones corresponding to "complete monotone Boolean functions" and then find a method for finding the least cost corresponding binary decision trees.
- Their method practical for n up to 4, not n = 5
- n = 4 at Port of Long Beach-Los Angeles.

Binary Decision Tree Approach

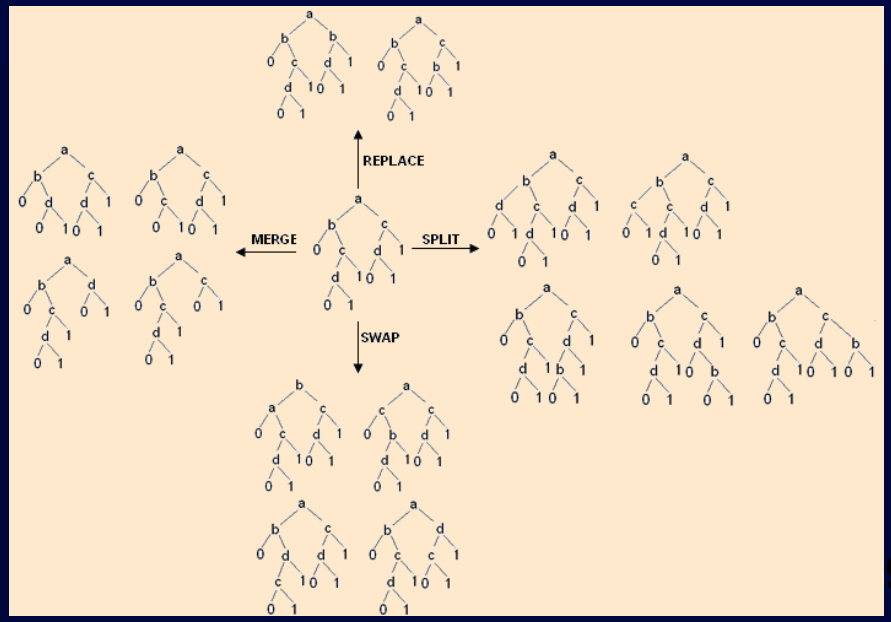
- n = 5
  - *263,515,920* binary decision trees for complete, monotone Boolean functions.
- Combinatorial explosion!
- Need alternative approaches; enumeration not feasible!
- Even worse: compare *5 x 10^{18}* BDTs corresponding to all Boolean decision functions

95

Binary Decision Tree Approach

Our results:

- New search algorithms work for n = 5.
- Genetics-algorithms search allows us to go to n = 10.
- Our SNSRtree software allows us to go to n = 20 on laptop



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

CBP Project on Modeling VACIS Inspection Processes at APM Terminal Port of Elizabeth

- Project Goal: study the VACIS operation at the APM terminal in Port Elizabeth, NJ using ***simulation modeling and analysis to improve VACIS operational efficiency and throughput.***
- A simulation model was built to capture
  - vessel arrivals
  - container storage at the yard
  - presentation of containers to CBP officers
  - and the actual inspection processes.

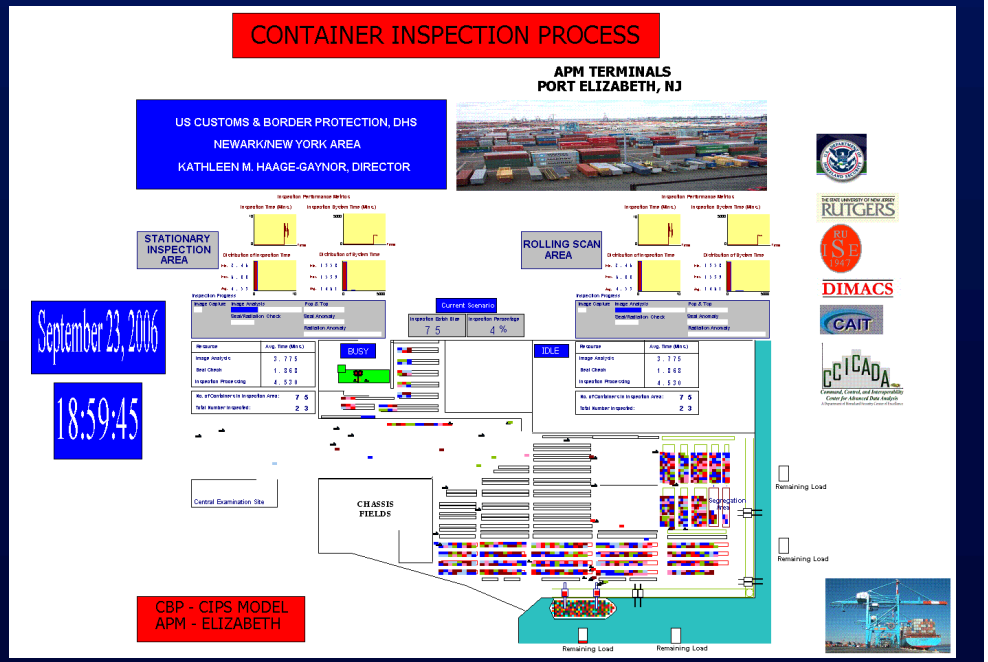
- A number of scenarios were analyzed to understand the capabilities of the inspection process under various surge conditions

The Simulation Model

Use Discrete Event Simulation with ARENA software

The animation displays:

- The incoming workload with ship arrivals and departures
- Loading and unloading of containers by cranes
- Shuttling of containers to storage areas
- Transfer of CBP-specified containers to the inspection area
- Container inspection processes
  - Stationary Scan
  - Moving Scan



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

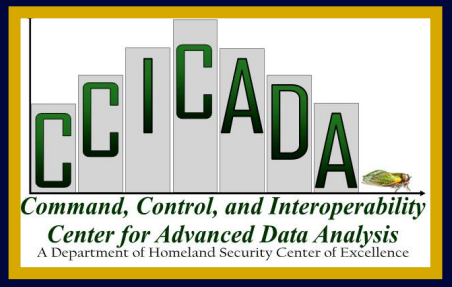
Performance Metrics

- ***Impact**: A revision was proposed in the way the hourly throughput is calculated in CBP's inspection operations to better reflect CBP operational metrics.*
  - ***Overall throughput per hour***: Hourly throughput based on the total time spent at the terminal.
  - ***Effective throughput per hour***: Hourly throughput based on the actual hours worked.

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

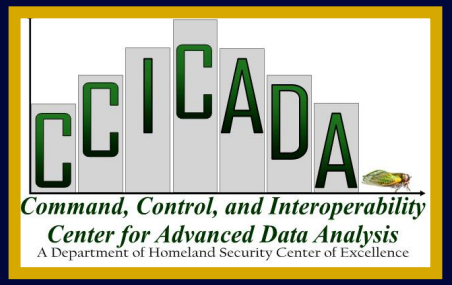
CBP: Off-site Container Inspection at Ports

- CBP of New York/Newark approached CCICADA to help with new initiative.
- CBP was looking for new ways to do container inspection – offsite in warehouse.
- CBP was experimenting with the new approaches
- Questions: Does this make inspection more efficient (faster throughput)? Does it make them less costly?
- CCICADA project: modeling and analysis of new approaches
- Frequent communication with CBP, monthly reports on data analysis from CCICADA



Container Inspection Off-site

- Project aimed to help evaluate and identify benefits for CBP in moving forward with this new initiative
- ***Initiative has major national implications for way we inspect containers at ports***



Example 5: Climate and Health

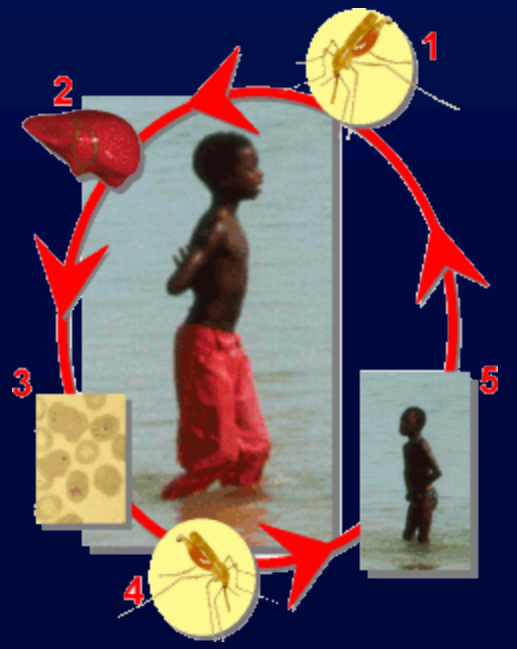
CCICADA
Command, Control, and Interoperability
Center for Advanced Data Analysis
A Department of Homeland Security Center of Excellence

Climate and Health

- Climate change is a homeland security issue:
  - More storms and natural disasters
  - Effects on health – more diseases in places where they weren't present before
  - Future conflicts over natural resources (water, land that is not under water, etc.)

Climate and Health

- Some early warning signs:
  - Malaria in the African Highlands
  - Dengue epidemics
  - Floods, hurricanes



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

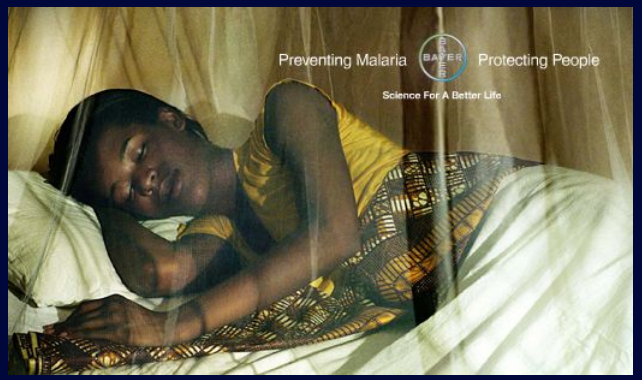
Malaria

- ***For each disease, effect of climate raises its own complex modeling challenges***
- Malaria a case in point (Mercedes Pascual).
- Climate change impacts:
  - Transmission rate changes with climate conditions
  - Affects dynamics of both host and mosquito
  - Affects time lag for parasite development as a function of temperature
  - Affects time lag for development of symptoms
  - Affects length of time patient remains immune
  - Rainfall affects the carrying capacity for larvae

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

Malaria

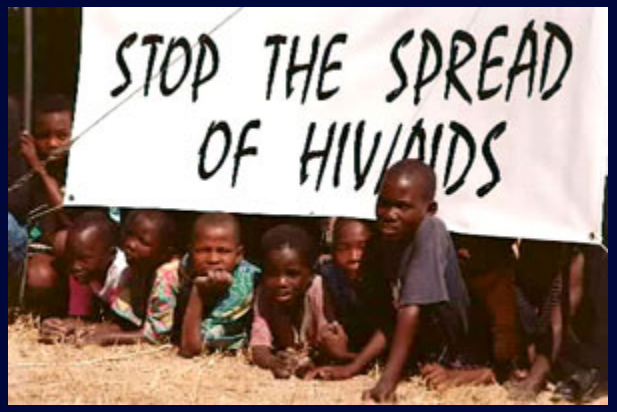
- But, there are complexities (David Rogers)
  - Models that depend on average change in temperature or rainfall miss out on spatial variation
  - There are changes in the time of the malaria cycle during which rainfall changes result in changes in impact
  - With global warming, there are areas where malaria would appear and others where malaria would disappear



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

Malaria

- Other complexities:
  - Confounding factors: climate change not the only variable
    - Changing land use
    - Migration of people (perhaps related to changing climate)
      - Emergence of other diseases (HIV)
    - Drug resistance (TB)



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

Malaria

- Other complexities:
  - WHO's early warning system for malaria
    - Monitors weather and climate
    - Weather is not the same as climate
    - Temporal scale issues:
      - But how far in the future can we reliably predict?
      - ***Short range***: one year?
      - ***Long range***: one decade???
    - Spatial scale issues
      - What is an appropriate resolution for predictions/ models?
      - By the kilometer? The region? The continent?

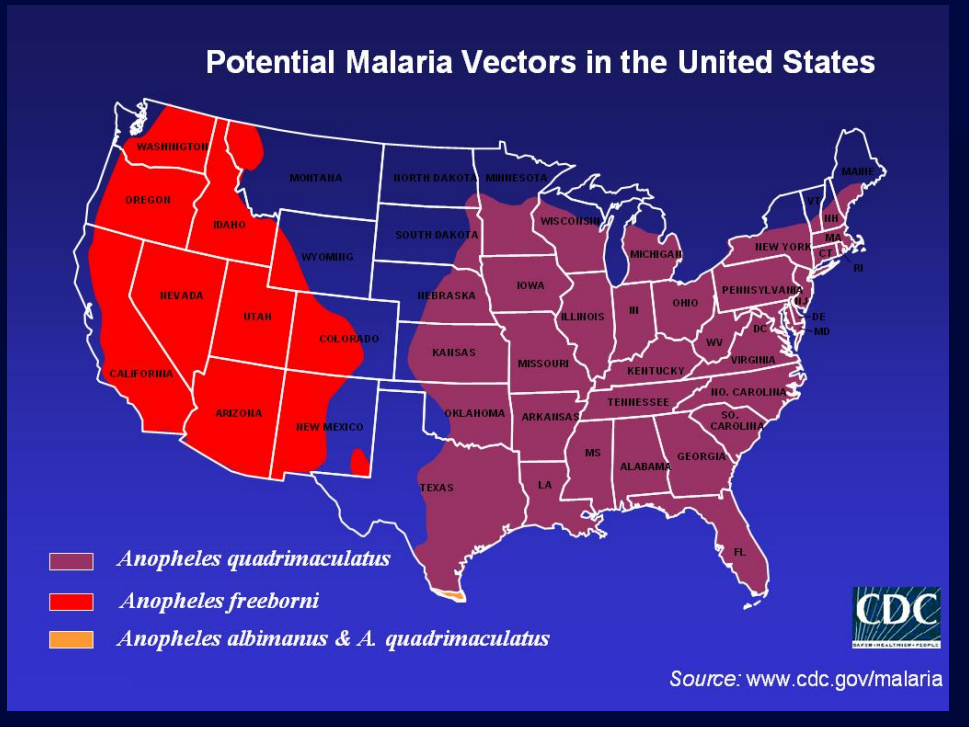
CCICADA
Command, Control, and Interoperability Center for Advanced Data Analysis
A Department of Homeland Security Center of Excellence

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 key role for modelers:

Aid in early warning:

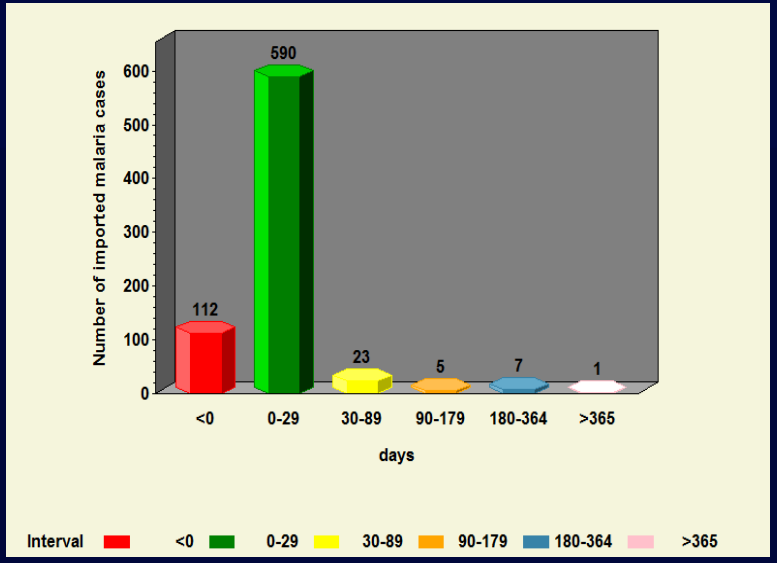
Surveillance.



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

Malaria

- Current CCICADA project at Howard University and Morgan State University
- Using parameterized math models to assess impact of climate change on spread of insect/tick-borne diseases in US and assess malaria prevention/ control
- Fitting models to CDC data



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

Cholera

- Complex interaction between climate and disease
- Regional climate variations have significant effect on cholera forecast
- Prevalence of the bacteria affected by conditions of sea water.
- Real correlation between January sea surface temperature and incidence of cholera in Bangladesh (Rita Colwell, Ben Cash)
- But: how consistent is the “signal”?
- Other relevant factors (related to climate)
  - Water level
  - Salinity
  - Nutrients in the water

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

Meningoccocal (Epidemic) Meningitis

- Occurs when bacteria penetrate the mucus membrane and enter the bloodstream
- Transmission of disease spikes in dry season, decreases when humidity hits.
- Modeling based on land cover and seasonal humidity profile (Madeleine Thomsen)
- But, also relevant are seasonal dust profile, soil type, population density
- Disease incidence clearly

climate related

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

Meningoccocal (Epidemic) Meningitis

- In Mali, clear relationship between onset of dry winds and seasonal meningitis epidemic (Madeleine Thomsen)
- But, this is a statistical relationship.
- We don't understand the mechanism.
- Is loss of integrity of mucus membrane due to physical damage from humidity? Dust?

Dust storm
Bamako, Mali

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

Dengue Fever

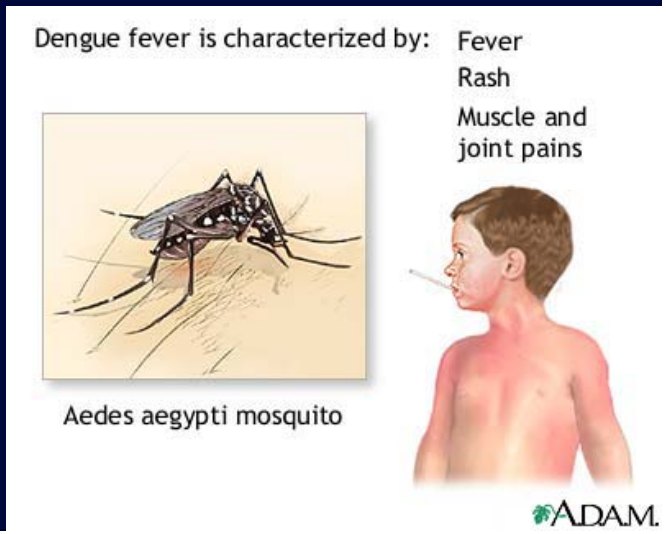
- Mosquito-borne disease.
- Appearing in places it hasn't appeared before.
- Large outbreaks in Brazil.
- Seems climate-related.
- Excessive rainfall leads to excessive cases (Timothy Desole)
- However, counterintuitive:
- Drought also leads to excessive cases.
- Why? Challenge for modellers.

Female *Aedes aegypti*
WHO/TDR/Stammers

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

Dengue Fever

- One explanation of drought case:
- In drought, people store water in containers.
- This provides breeding grounds for mosquitoes.
- Complex interaction among weather, mosquito populations, people's responses to weather conditions, etc.
- Another conundrum: In 2008, 50,000 cases along the Rio Grande between Mexico and Texas.
- Almost all in Mexico.
- Why?



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

Lyme Disease in Canada

- Increase in Lyme Disease in Canada.
- Is it due to climate change or normal range expansion for the ticks that carry the disease?
- Migratory birds carry the disease into Canada.

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

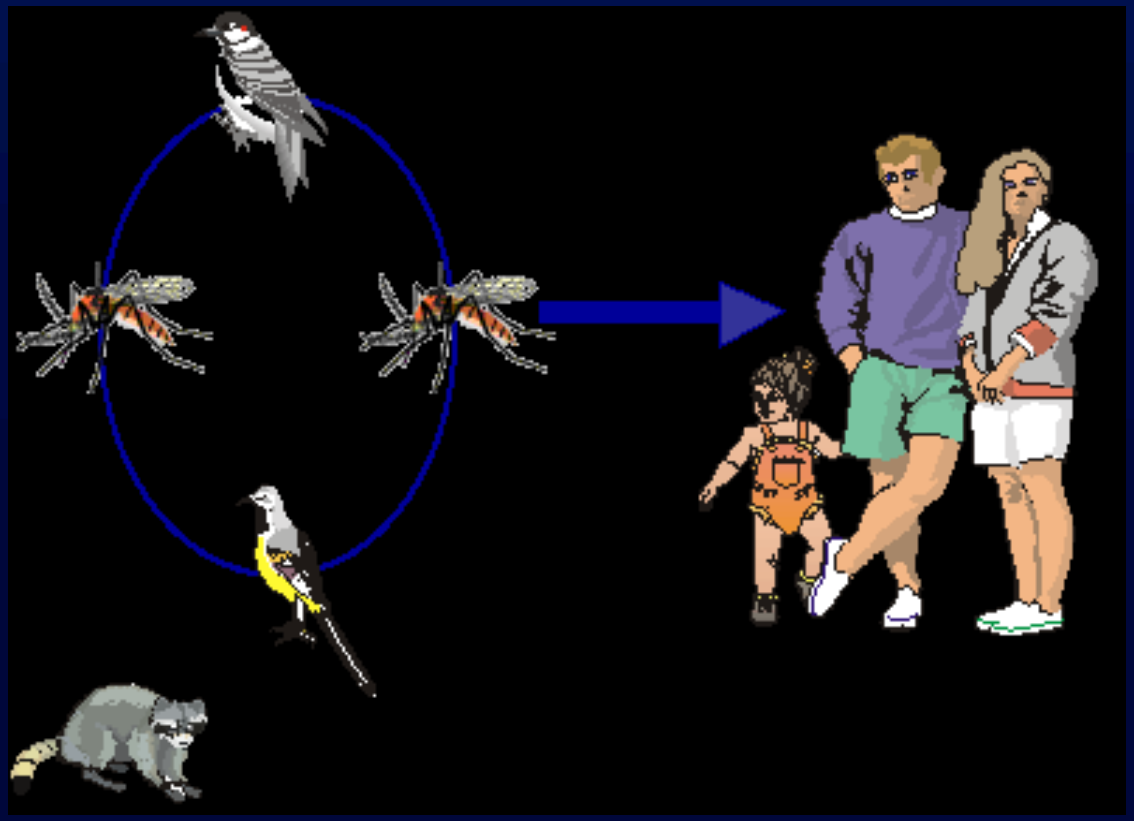
Lyme Disease in Canada

- Complex interactions among tick life cycles, bird migrations, climate, etc. (Michael Ogden)
- Temperature changes affect patterns of bird migrations.
- Do temperature changes also affect speed of bird migration?
- Temperature changes affect life cycle of ticks, birds
- Different climate change scenarios yield different tick life cycles

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

St. Louis Encephalitis

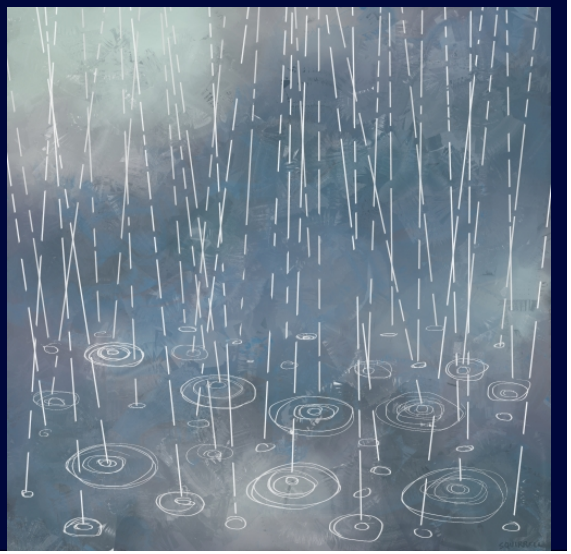
- Cycling of a virus between mosquitoes and birds is necessary to affect humans.



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

St. Louis Encephalitis

- Rainfall has heavy influence on mosquito population.
- SLE outbreak in Florida in 1990. Chickens bitten.
- Complex interactions among mosquito life cycle, rainfall cycles (Jeffrey Haman).
- Precipitation not only determinant of soil moisture.
- Model movement of water between layers of ground, model runoff, connect to nesting preferences for birds and reproductive cycle of mosquitoes.



C2ICADA
Command, Control, and Interoperability Center for Advanced Data Analysis
A Department of Homeland Security Center of Excellence

Ebola

- Connection between Ebola & climate change is subtle.
- Related to rainfall, not temperature
- 2002 study in the journal of Photogrammetric Engineering and Remote Sensing: sudden shifts from dry to wet conditions were associated with Ebola outbreaks from 1994 to 1996 in tropical Africa.
- With global warming come more sudden, more intense rain events.
- Interplay between climate change and deforestation positions humans closer to infected animals. People are often left to hunt "survivor species," like bats, which are one of the most common natural reservoirs of the Ebola virus.

Source: Washington Post 8/5/14

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

Ebola

- CCICADA project on Ebola is concerned with comparing the risk of different types of interventions by US on African soil:
  - Sending in military to build hospitals
  - Sending in a hospital ship
  - Using mothballed merchant marine vessels to bring in hospital materials

Credit: CDC

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

Some Animal Diseases

- Much data is anecdotal.
- Anthrax: related to rainfall.
- African horse sickness: related to combination of heavy rainfall and drought.
- West Nile Virus: Mild winter followed by severe Spring
- Rabies: Affected by rainfall – related to crowding around the waterhole in dry season

Blue Tongue

- Source: David Rogers
- Until the 1960s, restricted to Africa.
- Then, an outbreak in Europe.
- In the 1990s, permanently established In Southern Europe.
- The African vector (C. Imocale) moved there.
- Complex factors involved:
  - Climate
  - Weather
  - Vector/host ratio
  - Transmission coefficients
  - Biting rate
  - Incubation rate

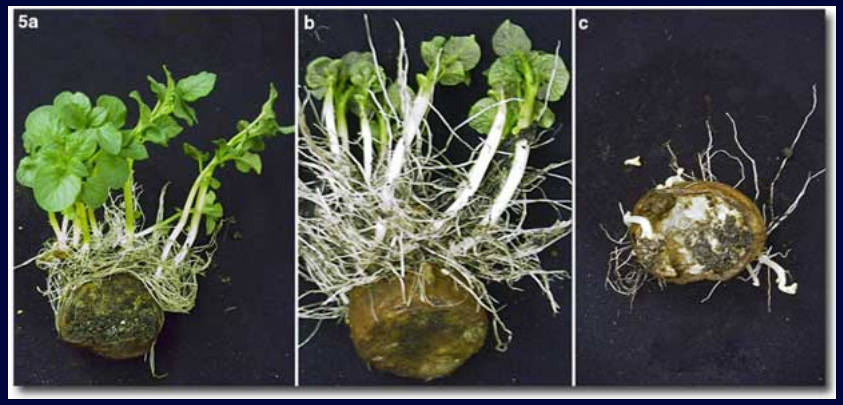
CCICADA
Command, Control, and Interoperability
Center for Advanced Data Analysis
A Department of Homeland Security Center of Excellence

Diseases of Plants

- Clear relation between drought/rainfall and health of plants.
- Effect of fungicides, pesticides, etc. complicates matters for the modeler.
- Possibility of genetic modification of crops to help them deal with global warming – presents complex modeling challenges

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

Potato disease



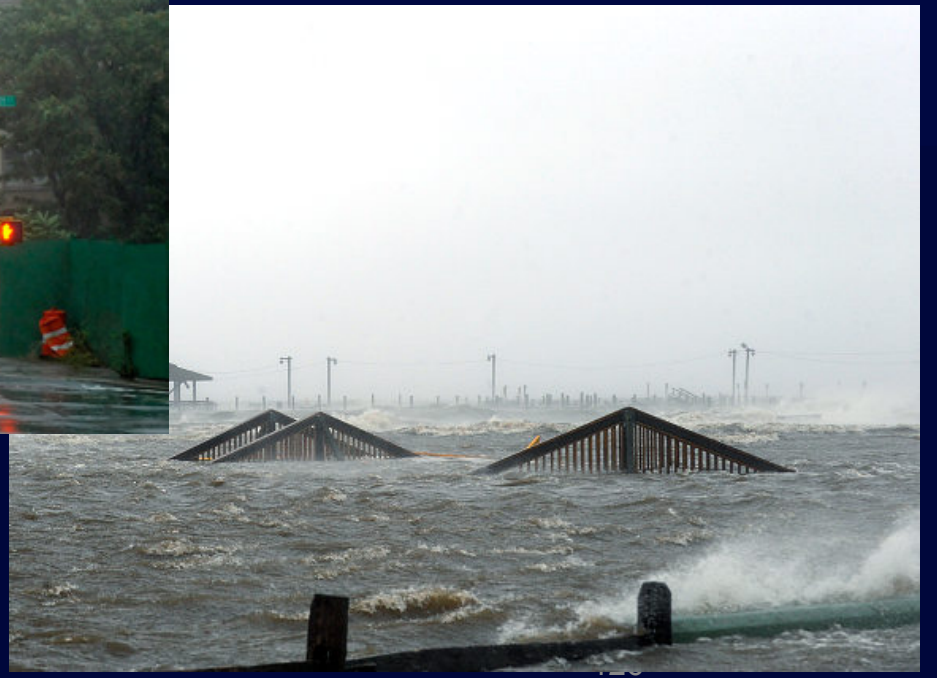
Extreme Events due to Global Warming

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

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

Extreme Events due to Global Warming: More Hurricanes

Irene hits NYC – August 2011



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

Extreme Events due to Global Warming: More Hurricanes

Irene hits NYC – August 2011

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

Extreme Events due to Global Warming: More Hurricanes

Irene hits NYC – August 2011

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

Extreme Events due to Global Warming: More Hurricanes

Sandy Hits NJ Oct. 29, 2013

My backyard

My block

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

Extreme Events due to Global Warming: More Hurricanes

Sandy Hits NJ Oct. 29, 2013

My neighborhood

My block

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

Extreme Events due to Global Warming: More Hurricanes

Sandy Hits NJ Oct. 29, 2013

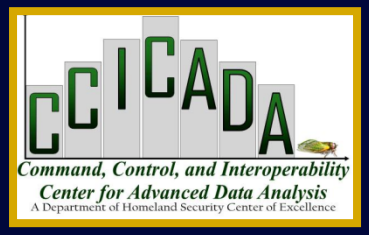
NJ Shore – from Jon Miller

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

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?

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

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?

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

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

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.
  - Relevance of mathematical tools of operations research – location theory

– CCICADA project on evacuation during heat events in connection with City of Newark

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

Extreme Heat Events

- A side effect of such events: Extremes in energy use lead to need for rolling blackouts.
- Modeling challenges:
  - Understanding health impacts of blackouts and bringing them into models
    - Lack of air conditioning
    - Elevators no work: vulnerable people
    - over-exert
    - Food spoilage
  - Minimizing impact on most vulnerable populations
  - Some optimization problems here

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

Example 6: Maritime Cyber Security

- The maritime transportation system is critical to the US economy.
- 95% of goods in international trade are still transported by sea.
- Disruption of global supply chain for commodities such as oil could cause dramatic problems for the world-wide economy.
- Disruption of the maritime transportation system could cause billions of dollars in damage to the economy.

Thanks to Capt David Moskoff, US Merchant Marine Academy, for many of the following examples

Hacking into a Ship

- A recent demonstration by a UT Austin team showed how a potential adversary could remotely take control of a vessel by manipulating its GPS.
- The yacht "White Rose of Drachs" was successfully spoofed while sailing on the Mediterranean.
- The team's counterfeit signals slowly overpowered the authentic GPS signals until they ultimately obtained control of the ship's navigation system.
- "The ship actually turned and we could all feel it, but the chart display and the crew saw only a straight line."

Source: UT Austin "Know"

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

Modern Ship Cyber Physical Systems

- For modern ships: dependence on a proliferation of sophisticated technology – that is subject to cyber attack
  - ECDIS (Electronic Chart Display and Information System)
  - AIS (Automatic Identification System)
  - Radar/ARPA (Radio Direction and Ranging) (Automatic Radar Plotting Aid)
  - Compass (Gyro, Fluxgate, GPS and others)
  - Steering (Computerized Automatic Steering System)
  - VDR (Voyage Data Recorder –"Black Box")
  - GMDSS (Global Maritime Distress and Safety System)
  - Numerous other advanced units and systems

Electronic Chart Display & Info System (ECDIS)

- Computer-based navigation system
- Can be used as an alternative to paper navigation charts
- Integrates a variety of real-time information
- Automated decision aid - continuously determining ship's position in relation to land, charted objects, navigation aids and unseen hazards
- Includes electronic navigational charts and integrates position information from the Global Positioning System (GPS) and other navigational sensors, such as radar, fathometer and automatic identification systems (AIS).
- May also display additional navigation-related information, such as sailing directions.

Electronic Chart Display & Info System

- Electronic Chart Display and Information System enables solo watchstanding

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

Electronic Chart Display & Info System

- World's largest container ship: Triple E Maersk under construction
  - 18,000 containers
  - 400 meters long!
  - Crew size: Can operate with 13 crew members!!
    - Thanks to ECDIS & other such systems.

Credit: http://www.worldslargestship.com/

Electronic Chart Display & Info System

- ECDIS flaws might allow an attacker to access and modify files and charts on board or on shore; could cause serious environmental and financial damage, even loss of life.
- In Jan. 2014, the NCC Group tried to penetrate an ECDIS product from a major manufacturer.
- Several security weaknesses were found: ability to read, download, replace or delete any file stored on the machine hosting ECDIS, etc.
- Once such unauthorized access is obtained, attackers could be able to interact with the shipboard network and everything to which it is connected.

Sources: templarexecs.com 2014, CyberKeel 2014

Automatic Identification System (AIS)

- AIS transceivers on over 400,000 ships (2013 estimate).
- Estimated that the number will soon reach a million.
- Installation is mandatory for all passenger ships and commercial (non-fishing) ships over 300 metric tons.
- Tracks ships automatically by electronically exchanging data with other ships, AIS base stations, and satellites.

Source: Help Net Security

Credit: wikipedia.org

Automatic Identification System

- An attacker with a $100 VHF radio could exploit weaknesses in Automatic Identification System which transmits data (e.g. vessels' identity, type, position, heading and speed to shore stations).
- The attacker could also tamper with the data, impersonate port authorities, communicate with the ship or effectively shut down communications between ships and with ports.

Source: templarexecs.com 2014

Automatic Identification System

- In Oct. 2013 Trend Micro demonstrated how easy it is to penetrate a ship's AIS.
- Plausible scenarios (CyberKeel 2014):
  - Modification of all ship details, position, course, cargo, speed, name
  - Creation of "ghost" vessels at any global location, which would be recognized by receivers as genuine vessels
  - Trigger a false collision warning alert, resulting in a course adjustment

Dr. Marco Balduzzi of Trend Micro discussing potential scenario
Credit: Help Net Security

Automatic Identification System

- In Oct. 2013 Trend Micro demonstrated how easy it is to penetrate a ship's AIS.
- Plausible scenarios continued (CyberKeel 2014):
  - Send false weather information to a vessel to have them divert around a non-existent storm
  - The ability to impersonate marine authorities to trick the vessel crew into disabling their AIS transmitter, rendering them invisible to anyone but the attackers themselves
  - Cause vessels to increase the frequency with which they transmit AIS data, resulting in all vessels and marine authorities being flooded by data. Essentially a denial-of-service attack

Automatic Identification System

- Somali pirates help choose their targets by viewing navigational data online, prompting ships to either turn off their navigational devices, or fake the data so it looks like they're somewhere else. (Reuters 4/23/14)

Credit: wikipedia.org

Automatic Identification System

- In Oct. 2013 Trend Micro demonstrated how easy it is to penetrate a ship's AIS.
- Why? (CyberKeel 2014):
  - The key problem with AIS is that it has no built-in security. All information is automatically assumed as being genuine and hence treated as correct piece of information.
  - Additionally, AIS messages are not encrypted and therefore very easy for outsiders to tap into and manipulate.

Automatic Identification System

- Potential Countermeasures to AIS Vulnerability:
  - Addition of authentication in order to ensure that the transmitter is the owner of the vessel
  - Creating a way to check AIS messages for tampering
  - Making it impossible to enact replay attacks by adding time checking
  - Adding a validity check for the data contained in the messages (e.g. geographical information)

Source: Help Net Security

GPS Jamming

- GPS Jamming can wreak havoc with modern ships.
- The UK & Irish General Lighthouse Authority directed GPS jamming equipment at a specific patch of ocean.
- On a vessel entering the jamming zone, a range of services failed: the AIS transponder, the dynamic positioning system, the ship's gyro calibration system and the digital selective calling system.
- The crew was able to cope with multiple alarms as they had been expecting this.
- However, on a modern vessel the bridge might in some cases be single-manned at night, causing significant problems should such a situation occur.

Source: CyberKeel 2014

GPS Jamming

- GPS jamming is possible with low cost jammers available over the Internet (though illegal).
- Many devices are battery-operated or can be plugged into a cigarette lighter and cost as little as $20.

Credit: CAPT David Moskoff

Oil Rigs

- Not just ships – ***vulnerabilities extend to the entire maritime transportation system.***
- Hackers recently shut down a floating oil rig by tilting it. (Reuters 4/23/14)
- Another rig was so riddled with computer malware that it took 19 days to make it seaworthy again. (Reuters 4/23/14)

Credit: www.peakoil.net

Cargo

- Cargo is also affected.
- 2011-2013: Hackers infiltrated computers connected to the Port of Antwerp, located specific containers, made off with their smuggled drugs and deleted the records.
- Attackers obtained remote access to the terminal systems; released containers to their own truckers without knowledge of the port or the shipping line.
- Access to port systems was used to delete information as to the existence of the container after the fact.

Source: Reuters 4/23/14, CyberKeel

Credit: wikipedia.org

Cargo

- In 2012 it was revealed that crime syndicates had penetrated the cargo systems operated by the Australian Customs and Border protection.
- The penetration of the systems allowed the criminals to check whether their shipping containers were regarded as suspicious by the police or customs authorities.
- The consequence was that containers with contraband were abandoned whenever such attention was identified by the criminals.

Source: CyberKeel

Credit: commons.wikipedia.org

Cargo

- The Iranian shipping line IRISL suffered from a successful cyber attack in 2011.
- The attacks damaged all the data related to rates, loading, cargo number, date and place.
- This meant that no one knew where containers were, whether they had been loaded or not, which boxes were onboard the ships or onshore.
- Even though the data was eventually recovered, it led to significant disruptions in operations and resulted in sending cargo to wrong destinations causing severe financial losses.
- Additionally, a considerable amount of cargo was lost.

Source: CyberKeel

Ports

- Today, ports rely as much on computer networks as on human stevedores.
- Complex networked logistics management systems track maritime cargo from overseas until reaching a U.S. retailer.
- Networked control systems are also often involved in the loading and unloading of these goods.
- Modern gantry cranes and other systems use optical recognition and other technologies to locate, scan, and manage all facets of port terminal operations.
- Automated container terminal systems use GPS to facilitate the automatic placement and movement of containers.

Source: CDR Joe Kramek, Brookings Report 2013

Credit: wikipedia.org

Ports

- The entire port is vulnerable – from cargo handling to truck and crane movement.
- Easily available jammers could close down a port at cost of more than $1B per day.

Maritime Cyber Security

- CCICADA will hold the first-ever tutorial and symposium on Maritime Cyber Security
- March 2-3, 2015 at Rutgers University, Piscataway, NJ
- Keynote by Admiral Chuck Michel, US Coast Guard
- Registration is limited
- Register at ccicada.org

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

Questions?