Heterogeneous Stream Processing for Disaster Detection and Alarming

Ryan Stern

stern@cs.colostate.edu

Big Data Lab

Department of Computer Science

Colorado State University

Outline

- Background

- System Goals

- System Architecture

Background

- Gather information from heterogeneous streams:
  - Mobile network utilization data
  - Twitter and other social media streams
  - Bus and train networks
  - Streams represent spatial time series

- Spatial time series: (attribute, event, time, location)
- Event: (event_i, object_n, time_n, location_n)
- Trajectory: (object_i, time_n, location_n)

- Automated fusion of heterogeneous stream is challenging:
  - Various data formats
  - Different spatiotemporal resolutions
  - Privacy concerns

System Goals

- Distributed

- Operate in real time

- Detect incidents by fusing information from multiple streams

- Combine related low-level incidents into high-level situations

- Scale with increasing number of data sources

- Handle uncertainty in the data

- Privacy preserving

System Architecture

- System consists of four layers:
  - Input layer
    - Capture input sensor streams
  - Intelligent Sensor Agents
  - Computation layer
    - Round Tables
   - Complex Event Processors
  - Output layer

Intelligent Sensor Agents (ISA)

- Responsible for cleaning data and detecting incident events

- May be deployed remotely or locally
  - Decentralization of data
  - To address privacy concerns and legal constraints
  - Shares abstract events with the rest of the system

- Specific to a each type of sensor
  - Able to understand the data format, acting as a translator
  - Individual storage strategies

- Produces events that the Round Tables are able to understand
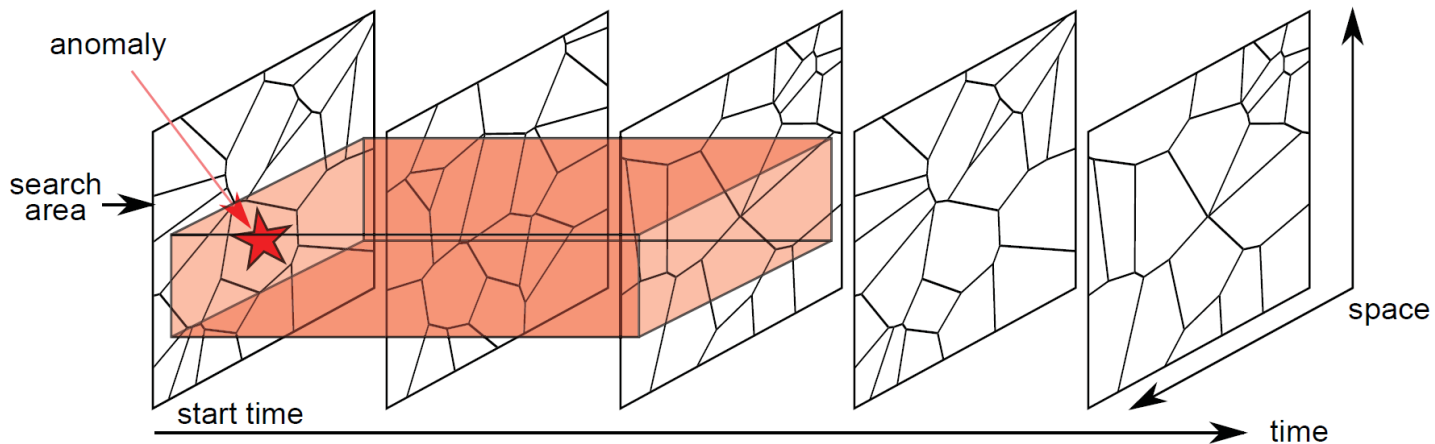
Round Tables (RT)

- Responsible for identifying, confirming, and labeling events

- Accomplished by fusing information from multiple ISAs

  – Iteratively queries information from participant ISAs to improve confidence in the incident being investigated

  – Where it gets its name: "A round table meeting of experts"
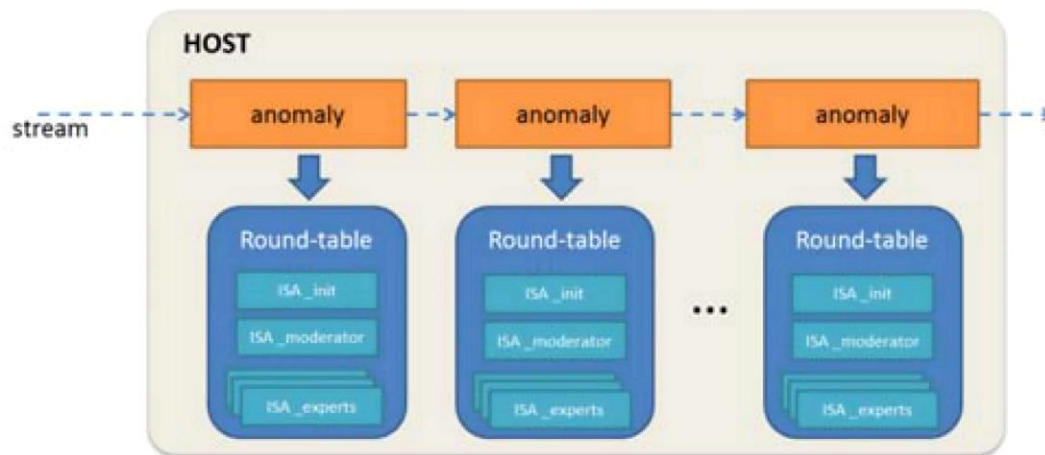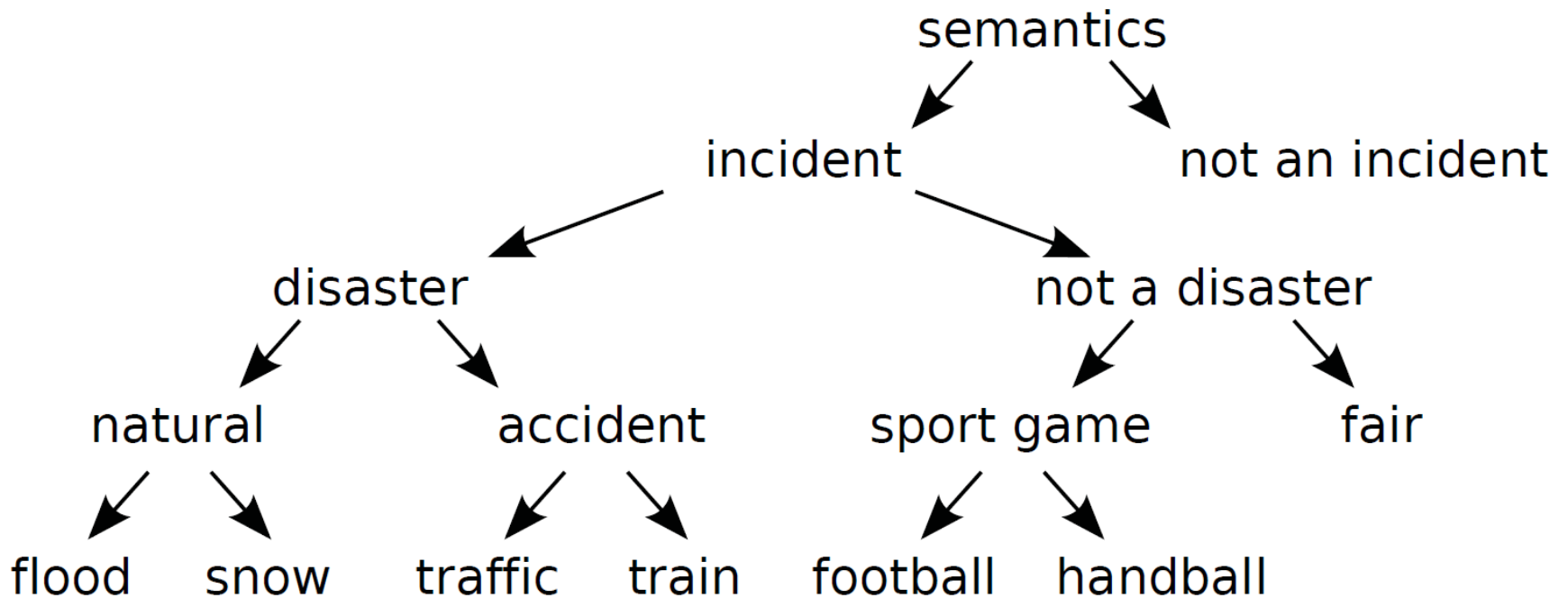
Round Tables (RT)

- Uses Maximum Likelihood to refine events
  - Iteratively compares candidate solution to find a local maximum

- The Round Tables may be unable to come to a decision
  - Able to request help from an expert or a crowd

Round Tables (RT)

Round Table Manager (RTM)

- Responsible for creating Round Tables
  - To prevent duplicate Round Tables from being created for the same incident

- Round Tables are created
  - When an anomaly is detected by an ISA
 - On the request of a human expert

- Assigns ISAs to Round Tables

Detection Example

1) An ISA detects an anomaly
2) The ISA requests the RTM to set up a RT session instance
3) The RTM selects multiple ISAs that may help evaluate the anomaly
4) The initiator ISA presents its finding to the RT
5) The RT iteratively requests information from the invited ISAs
6) The ISAs explore their sensor data history to verify an event
7) The RT receives and integrates information provided by each ISA
8) The most precise description of the anomaly is distilled
9) The RT sends decision information to the Complex Event Processor

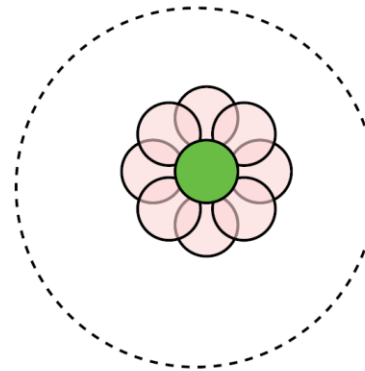# Detection Example
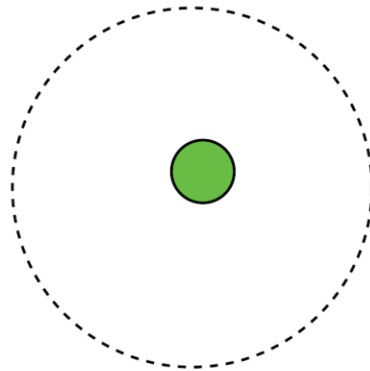
(a)   (b)   (c)

- *<(lat, long, radius), [starttime, endtime], {eventType, likelihood}>*

- *<(6.51, 9.87, 1500), [12:00, 13:20], {<incident, 1>,<not incident, 0.2>}>*

- *<(6, 9, 10000), [12:00, now], {<soccer game, 1>, <handball, 0.2>, <disaster, 0.1>, <not incident, 0.2>}>*

- *{<soccer game, 1>,<handball, 0.2>,<disaster, 0.1>,<not incident, 0.04>}*

Questions?

SLIDE SIDE

SLIDE OVER