Edge Analytics—Real-time Analysis of Social Media

Objective

First responders at the edge want to monitor open source social media feeds in real-time to enhance public safety.

Operational Challenges

Increased cognitive load

Volume and velocity of raw data increases the cognitive load on edge users.

Dynamic environment

Relevance is dynamic

Discovery

What needs to be monitored is unknown and changing

Local Context

Global trends may not be relevant to local events.

Hypothesis and Technical Challenges

Hypothesis:

Given a local context where patterns of interest are scattered without any dominant trend or pattern, events and emerging trends can be identified in sufficient time to inform operations.

Technical Challenges:

How can we summarize streaming data? How fast can we determine changes?



Given a local context, Significant Events are detectable as dominant trends or patterns

Boston dominated the news agenda Tweets per day April 2013, ET



Can the peak be detected in time to influence reaction?

Solution

Monitor social media streams (e.g., Twitter) in real time

- Sentiment analysis
- Topic modeling
- Social network analysis

	Topic Alerts	
Sentiment Summary	Between Dates 08/15/2013 12:35 and	
Trends	08/26/2013 12:35 Time Interval by hour	
Sentiment Over Time	Search	
Торіс	8/21/2013 20:11	
Sentiment	0 5 10 15 20 25 30 35 40 45 Williamsport	50
Pictures	gun 7 acs 3	
Location	Lamade Stadium 3 help 3	
Social Network	IED 2	
Relationships	I saw 2 JRL 2	
Groups	bomb 2 wet 1	
Alerts	shoot 1 fight 1	
Search	I just saw 📕 1	
Settings	Results	

Flexible infrastructure

- Performance: Process real-time streams
- Scalability: Quickly from a laptop to enterprise cloud
- Platform: Support multiple programing languages
- Development Time: Rapid and parallel development



Reduce Cognitive Load

- Flexible searching
- Custom alerting
- Transform data into patterns
- Visualize patterns

Continient Cuminary	Between Dates 06/25/2013 12:53 and
Trends	06/30/2013 12:53 Time Interval by hour \$
	Search
Sentiment Over Time	
Торіс	6/29/2013 18:57
Sentiment	0 2 4 6 8 10 12 14 christian
Pictures	kingsfest 6 huntingdon 3 hardrock 3
Location	irvine 2 creation2013 2
Social Network	socalfamilytun 2 guitarsax88 2
Relationships	christian music 2 love 2
	great christian 2 candle 1
Groups	cabal lounge 1
Groups Alerts	cabal lounge 1 mbar26 1 mike 1 texasleftist pay 1

- - Reuse: Existing & open source components
 - Fault tolerant: Ability to support failures
 - Modifiability: Quickly add new data/analysis

Types of Trend Analysis

Macro

- Collective trends
- Noise-tolerant
- Privacy protecting
- Large data volume



Micro

- Targeted and context specific
- Requires higher accuracy
- Privacy issues

Sleep patterns of the Boston bomber based on *Twitter Metadata*



Field Experiments and Results

Pilots

- 1. Huntingdon County EMA (Creation Festival 2013)
- 2. Weapons of Mass Destruction Civil Support Team (Little League World Series)

Observations & Feedback

HCEMA and WMD-CST showed the potential

- HCEMA "Harlem Shake": Event took 7 minutes to verify after Edge Analytics picked it up
- WMD-CST "Bomb threat": Detected and determined to be a hoax

Challenges and Future Work

- False positives from keywords
- Dynamic and adaptive discovery of relevant artifacts
- Non-textual inputs
- Voice
- Multi-lingual challenges and Leetspeak
- Images and Video
- Feedback to support preventative/influence operations

Contact: info@sei.cmu.edu) http://www.sei.cmu.edu/mobilecomputing/