A System for real-time twitter sentiment analysis of 2012 U.S. presidential election cycle

Wang, Hao, et al.
University of Southern California, Los Angeles, CA

ACL ‘12 Proceedings of the ACL 2012 System Demonstrations

29 October, 2014

Jaehwan Lee

Outline

- Introduction
- Related Work
- The System
  - Input/Data Source
  - Preprocessing
  - Sentiment Model
  - Aggregation
  - Display and Visualization
  - Annotation Interface
- System Evaluation
- Conclusion

Introduction

- Social media and politics
  - Social media platforms have become an important site for political conversations throughout the world
- Twitter
  - Twitter allows users to post tweets, messages of up to 140 characters, on its social network
  - Twitter usage is growing rapidly
  - 100 million active users worldwide, over 250 million tweets each day (Twitter, 2012)
  - It was actively used by 13% of on-line American adults
- "More than two thirds of U.S. congress members have created a Twitter account and many are actively using Twitter to reach their constituents."

Introduction

- Most work to date
  - They has focused on post-facto analysis of tweets
  - with results coming days or even months after the collect

- But, tweets are
  - short and easy to send
  - lend themselves to quick and dynamic expression of instant reactions to current events

- Therefore, we expect "automated real-time sentiment analysis"

Introduction

- Two additional issues
  - The vernacular used on Twitter
    - differs significantly from common language
    - we have trained our sentiment model on its idiosyncrasies
  - Tweets in general and political tweets in particular
    - tend to be quite sarcastic
    - presenting significant challenges for computer models

Related Work

- Growing interest in mining sentiment and opinions in text
  - due in part to the availability of documents and messages expressing personal opinions
- Sentiment in Twitter data
  - used for prediction or measurement in various domain
    - such as stock market, politics and social movements (Bollen et al., 2011; Tumasjan et al., 2010; Zeitzoff, 2011)
- Political sentiment on social networks
  - either post-hoc and/or carried out on small and static samples

The System

- For accuracy and speed,
  - we built our real-time data processing infrastructure on the IBM's InfoSphere Streams platform

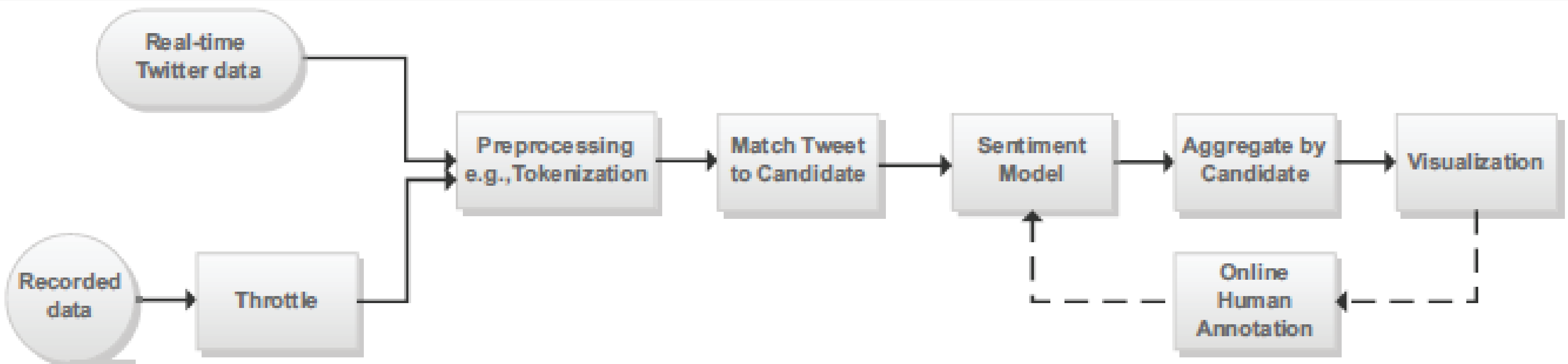

Figure 1. The system architecture for real-time processing Twitter data

The System - (1) Input/Data Source

- Twitter as our data source
  - because it is a major source of online political commentary and discussion in the U.S.
  - It played a significant role in political events worldwide
    - such as the Arab Spring Movement and the Moldovian protests

- How to collect the tweets
  - Twitter's public API provides only 1% or less of its entire traffic
  - Thus, we collect all relevant tweets in real-time from the entire Twitter traffic via Gnip Power Track-a commercial Twitter data provider

The System - (1) Input/Data Source

- Specify domain using rules
  - manually construct rules that are simple logical keyword combinations to retrieve relevant tweets
  - those about candidates and events including common typo
    - E.g. rules for Mitt Romney
    - {Romney, @MittRomney, @PlanetRomney, @MittNews,
    - @believeinromney, #romney, #mitt, #mittromney, #mitt2012}
  - 200 rules in total
    - for Barack Obama and nine Republican candidates

The System - (2) Preprocessing

- As in NLP practices, tweets is tokenize for later processing
  - we use certain rules to handle the special cases in tweets
  - found Christopher Potts' basic Twitter tokenizer best suited throughout several Twitter-specific tokenizers

Tweet	WAAAAAH!!! RT @politico: Romney: Santorum's 'dirty tricks' could steal Michigan: http://t.co/qEns1Pmi #MIprimary #tcot #teaparty #GOP
Tokens	WAAAAAH !!! RT @politico : Romney : Santorum's ' dirty tricks ' could steal Michigan : http://politi.co/wYUz7m #MIprimary #tcot #teaparty #GOP

Figure 2. The output tokens of a sample tweet from our tokenizer

The System - (3) Sentiment Model

- Design of the sentiment model
  - based on assumption
    - the opinions expressed would be highly subjective and contextualized
  - we used a crowd-sourcing approach to annotate

- To create a baseline sentiment model
  - we used Amazon Mechanical Turk
  - they participated anonymously
  - they were asked their age, gender and political orientation
  - they annotated the tweets' sentiment
    - positive, negative, neutral or unsure
    - whether it is sarcastic or humorous

The System - (3) Sentiment Model

- To create a baseline sentiment model (cont.)
  - training data consists of nearly 17,000 tweets
    - (16% positive, 56% negative, 18% neutral, 10% unsure)
  - nearly 2,000 tweets were multiply annotated
  - about 800 Turkers contributed
  - we use naïve Bayes model on unigram features

- Performance of model
  - performs at 59% accuracy on the four category classification
  - exceed the baseline of classifying all the data as negative
  - our model was not strictly motivated by global accuracy
  - but took into account class-wise performance so that the model performed well on each sentiment category

The System - (4) Aggregation

- For volume
  - the system outputs the number of tweets every minute for each candidate

- For sentiment
  - the system outputs the number of positive, negative, neutral and unsure tweets in a sliding five-minute windows

The System - (5) Display and Visualization

- an Ajax-base HTML dashboard
  - to display volume and sentiment by candidate as well as trending words and system statistics

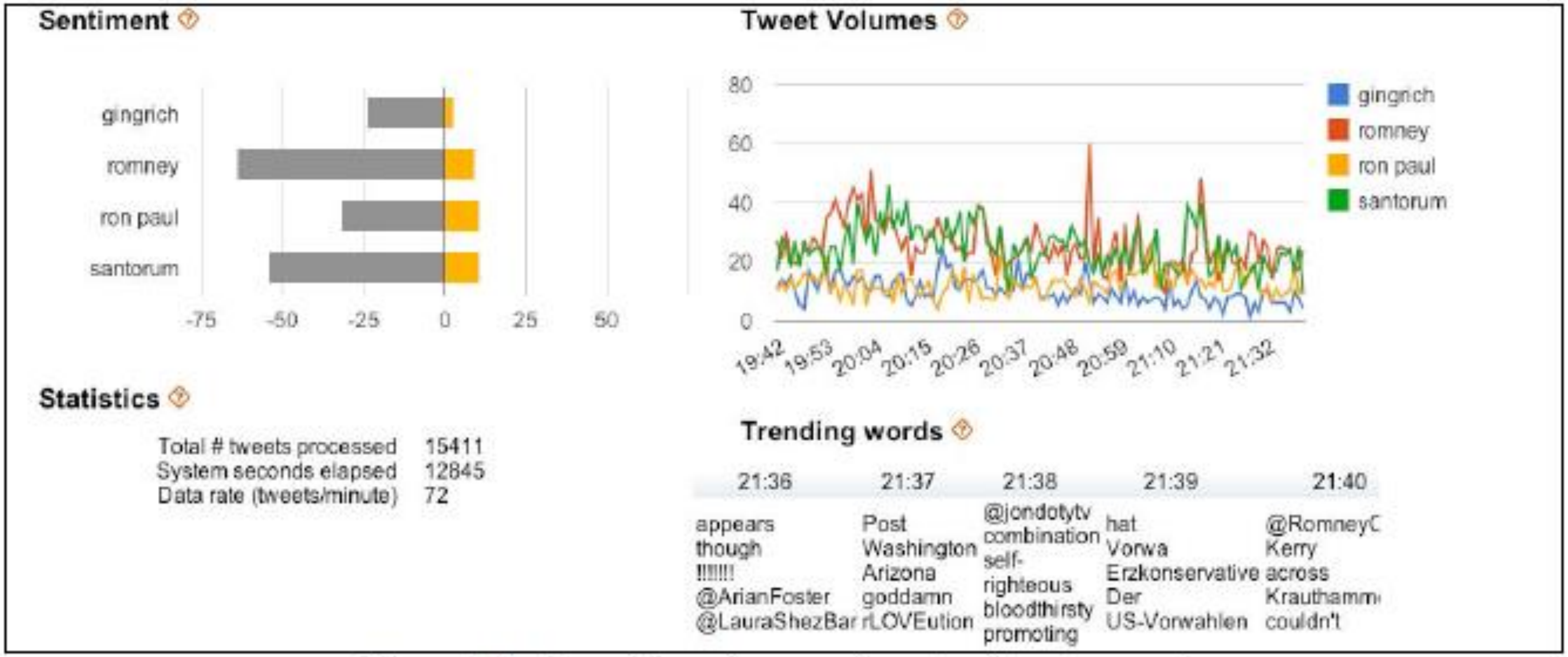

Figure 3. Dashboard for volume, sentiment and trending words

The System - (5) Display and Visualization

- refreshes display every 30 seconds

the # of positive and negative tweets about each candidate in the last five minutes

the # of tweets for each candidate every minute in the last two hours

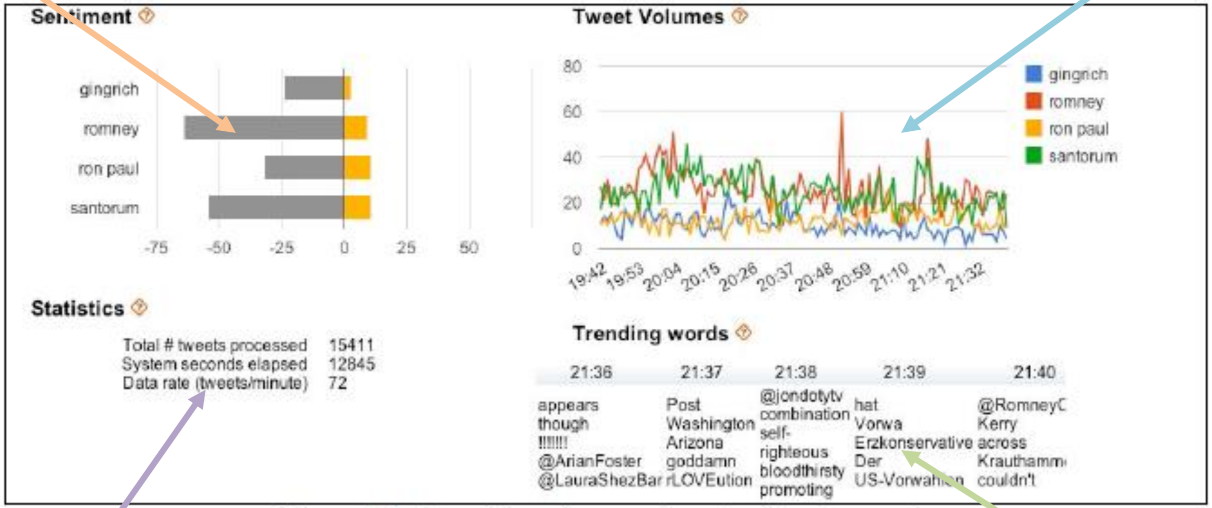

Figure 3. Dashboard for volume, sentiment and trending words

system statistics including the total # of tweets, # of seconds since system starts and the average data rate

trending words of the last five minutes, computed using TF–IDF

The System - (5) Display and Visualization

total volume over time

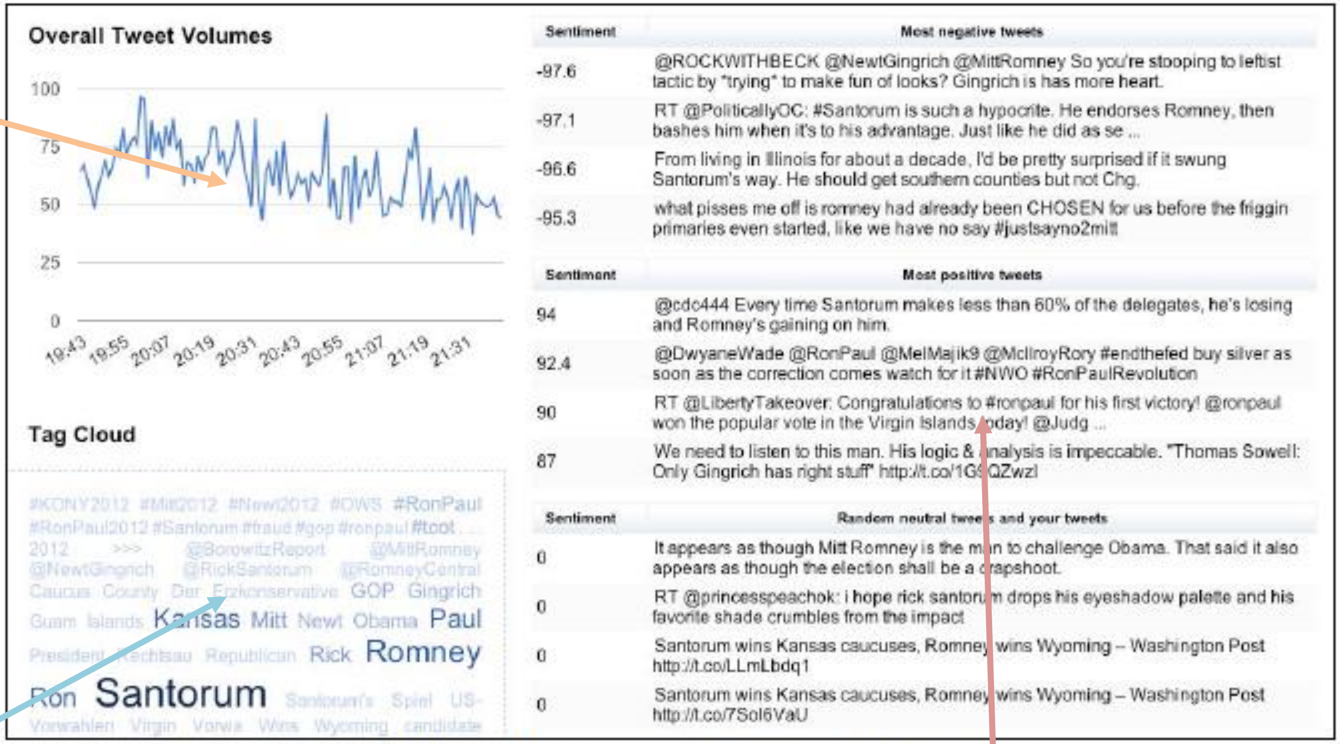

Figure 5. Dashboard for most positive, negative and frequent tweets

a tag cloud of the most frequent words in the last five minutes across all candidates

the most positive tweets
the most negative tweets
some random neutral tweets

The System - (6) Annotation Interface

- Annotation Correction
  - users can annotate tweets by clicking
  - feedback makes the model trained actively and iteratively

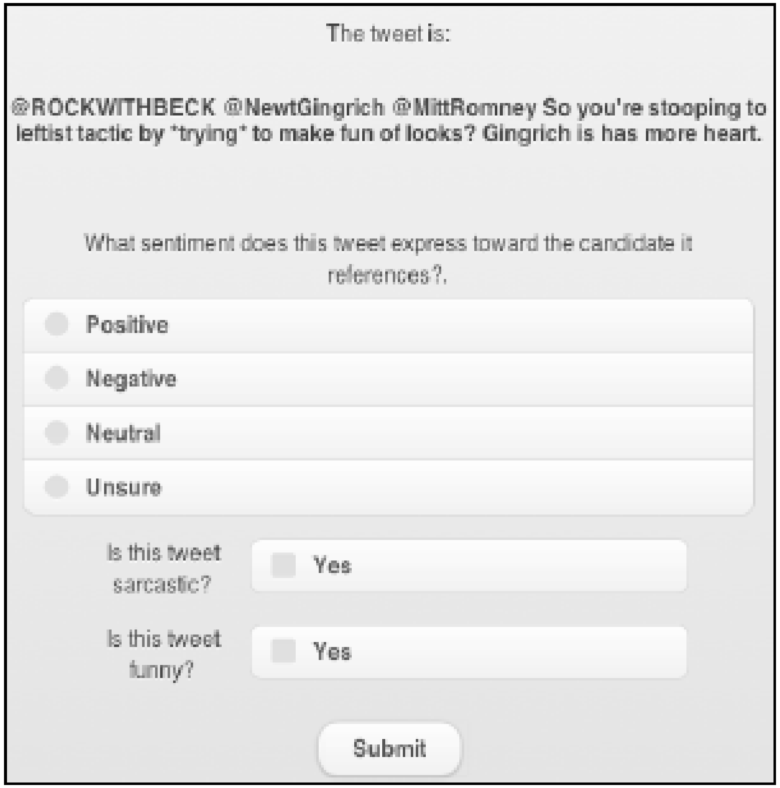

Figure 4. Online sentiment annotation interface

System Evaluation

- To evaluation the model
  - correlational analysis of sentiment with political events/news
  - this quantitative analysis is **ongoing work**!

- Some findings
  - tweet volume is largely driven by campaign events
    - of the 50 top hourly intervals except two correspond …
  - sentiment changes on emerging events
    - … within minutes, Newt Gingrich's negative sentiment increase rapidly – it became three times more negative in just two minutes

- "How tweet volume and sentiment are extremely responsive to emerging events in the real world (Vergeer et al., 2011)"

Conclusion

- Presented a system for real-time Twitter sentiment analysis
  - evaluates public sentiment changes in response to emerging political events and news

- Architecture and method are generic
  - can be easily adopted and extended