Detecting and Analyzing Bursty Events on Twitter

by

Pau Perng-Hwa Kung

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning
in partial fulfillment of the requirements for the degree of

Master of Science in Media Arts and Sciences

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2016

Author . . . .
August 15, 2016

Certified by . . . .
Deb Roy
Associate Professor of Media Arts and Sciences
Thesis Supervisor

Accepted by . . . .
Pattie Maes
Academic Head, Program in Media Arts and Sciences

Detecting and Analyzing Bursty Events on Twitter

by

Pau Perng-Hwa Kung

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning
on August 15, 2016, in partial fulfillment of the
requirements for the degree of
Master of Science in Media Arts and Sciences

Abstract

This thesis presents *BurstMapper*, a system for detecting and characterizing bursts of tweets generated by multiple sources in order to understand interactions between Twitter users and the role of exogenous events (not directly observable on Twitter) in driving tweets. The first stage of the system finds temporal clusters, or *bursts* of tweets. The second stage characterizes bursts along two dimensions, *semantic coherence* and *causal influence*. Semantic coherence measures the semantic relatedness of the tweets in a burst to each other based on a deep neural network derived embedding of tweet contents. Causal influence measures the potential causal interaction between Twitter users using the Hawkes process model.

We introduce an annotated corpus of 7,220 tweets produced by five leading candidates in the 2016 U.S. presidential election. Evaluating the system on the annotated corpus shows that with a precision of 75%, tweets caused clearly by specific exogenous events (or *responsive tweets* hereafter) are detected by the burst detector components of our system. Furthermore, experiments show that the linear combination of semantic coherence and causal influence are predictive of the presence of responsive tweets in a burst, with the F1-score of 0.76. Examining bursts along the two dimensions reveals that (i) the measures are positively correlated with each other (corr=0.33, $p<0.001$), (ii) the measures allow us to understand how candidates tend to respond differently to exogenous events, e.g., by attacking opponents or making plan announcements, and (iii) the measures can be used to describe the influence dynamics between candidates over time. Plotting the bursts from a corpus of 1,470 Twitter accounts (the five leading candidates and the users followed by them) shows visual evidence that some user groups (e.g., campaign staffs, journalists, etc.) have a higher levels of semantic coherence and causal interactions. These experiments suggest that the bursts detected by our system provide a useful level of abstraction that summarizes tweet content, providing a solution for coping with massive amount of data on Twitter.

Thesis Supervisor: Deb Roy
Title: Associate Professor, Program in Media Arts and Sciences

Acknowledgments

A huge thanks to the following people for making this thesis possible:

I would like to thank my advisor, Professor Deb Roy, for providing the research support for making this thesis happening in the first place. I am truly grateful for Deb's patience and the willingness to give guidance where needed. The process has not just been producing a good research, but also learning the key factors to be innovative and inspiring.

I would also like to thank Doctor Allen Gorin, who has been actively involved in the research process and kindly shared many of his valuable research experience. I am also grateful of my thesis readers, Professor César Hidalgo, and Professor Ryan P. Adams for providing valuable comments and feedbacks that bring substance into this thesis. I would like to thank everyone in MASCOM: Chris and Sep for research and presentation feedbacks; Linda and Keira for patience and kindness to make sure the thesis progress goes accordingly to the institute rules.

It was fortunate to be around the truly great and inspiring LSM folks. I am truly thankful. In particular, I would like to thank Soroush who gave really valuable research feedbacks and pointers, Prashanth for kindly providing data and the awesome Tweet2Vec model, Mina and Neo for kind support and consultation, and Amber and Heather for providing every resource necessary to expedite the research process.

I would like to thank my dear friends for being a great source of emotional support, especially Benjamin and Ken, whom I can truly trust and rely to share visions, and know there are still amazing folks back in my home country who inspire me to also be a more intelligent, kind, empathetic person.

Also, many thanks to my family, Mom and Dad, for their support in pursuing a long-time goal/dream of coming to the MIT.

Lastly, and most importantly, I would like to thank my girlfriend, Ying-Ju (Ruthy), for continual support and love. If there is anyone to dedicate this thesis to, it would be her. I am in gratitude beyond words.

Detecting and Analyzing Bursty Events on Twitter

by Pau Perng-Hwa Kung

This masters thesis has been examined by the following:

Professor Ryan P. Adams . . . .

Thesis Reader

Associate Professor and Head of Advanced Technologies Group

Harvard University and Twitter

Detecting and Analyzing Bursty Events on Twitter

by Pau Perng-Hwa Kung

This masters thesis has been examined by the following:

Doctor Allen Gorin . . . .

Thesis Reader

Research Affiliate and Visiting Scholar

Johns Hopkins University and Massachusetts Institute of Technology

Detecting and Analyzing Bursty Events on Twitter

by Pau Perng-Hwa Kung

This masters thesis has been examined by the following:

Professor César A. Hidalgo . . . .

Thesis Reader

Associate Professor of Media Arts and Sciences

Massachusetts Institute of Technology

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

In recent years, the micro-blogging site Twitter[1] has become a major social media platform for hundreds of millions of users. These users tweet about their lives, share opinions, and discuss issues across a variety of different topics. The open-access and mutually visible nature [11] of the platform makes it an ideal place for interested people to explore how different individuals and organizations interact and respond to social signals.

Developing an understanding for how tweets are generated is non-trivial. Take the tweets posted by the 2016 U.S. leading presidential candidates as an example. A naïve way to explain the tweeting events is by fitting a Poisson Process, in which a candidate tweets without being influenced by other candidates or exogenous events. Figure 1-1 shows the daily tweet volumes for the candidates, with the fitted constant-rate Poisson Process. Chi-square test reveals that the fitted Poisson Process variance significantly underestimates the sample variance of the tweet series. This is also true when we consider tweets filtered by election topics[2] (see Figure 1-2). As an another challenge, explicit interactions between voices do not occur frequently. In our example, candidate mentions account only 12% of all topic-filtered tweets. Therefore,

[1]http://www.twitter.com

[2]Here we apply the trained election topic classification model as proposed in [51]

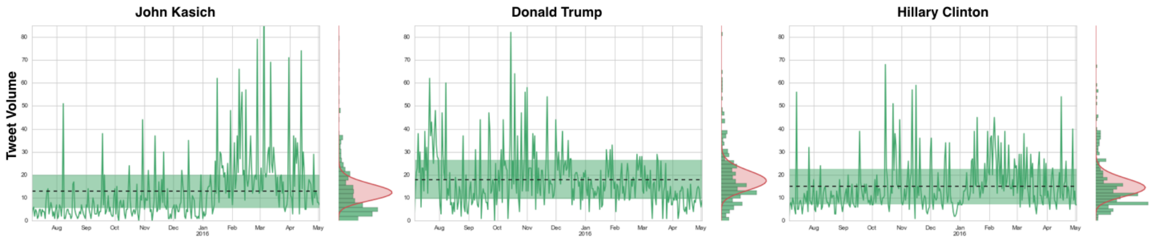


Figure 1-1: Mapping of three presidential candidates' daily tweeting volume from July 2015 to April 2016. X-axis is discretized with day as the time unit. The fitted Poisson Process mean is plotted as the dashed line, with green band being the $\pm 2\sigma$ range. Right panel shows the data histogram and fitted Poisson Process' distribution.

Figure 1-2: Tweet series with the same candidates and time span as in Figure 1-1, filtered by election topics terrorism and guns. The panels with colored vertical lines show tweets with explicit mentions of other candidates.

there may be influences not directly observable on Twitter that drive a candidate to decide *when* to tweet, and *what* to tweet about.

What if there was a tool to help us understand the interactions between Twitter users and the roles of exogenous events (not directly observable on Twitter) in driving tweets? This was the motivation for developing *BurstMapper*, a tool for detecting and characterizing tweet bursts to provide a useful level of abstraction for summarizing tweet events.

1.1.1 Bursts in tweets

Before describing our approach, we first explain what a tweet burst is. A burst of tweets is a temporal cluster with high density of tweeting activities. In the study of human dynamics, burst is a commonly observed phenomenon in everyday activities (e.g., E-mailing, web-browsing) [3][48][26]. For a bursty series of events, the distribution for inter-event arrival time often follows a heavy-tail distribution. We confirm this characteristic with our candidates tweeting example (α=1.13, reflective of the observed $\alpha \simeq 1$ for human activities [48])[3].

The occurrence of burst is often understood as a result due to the presence of some exogenous events. Therefore, we have reasons to believe bursts on Twitter capture significant information on user activities.

1.1.2 Properties of tweet bursts

Understanding the characteristics of a tweet burst can help us describe the roles user interactions and exogenous events play in driving tweets. To develop intuitions for characterizing bursts, let us zoom-in Figure 1-2 to examine a week-long series of tweets, as shown in Figure 1-3. For the tweets on topic *terrorism*, there is a burst of tweets right after the Paris shooting incident [17] in November 2015, in which the candidates expressed similar semantic contents by emphasizing their anti-ISIS stance. It is noteworthy that Hillary Clinton announced her anti-ISIS policy plan a few days after the shooting, in contrast to her usual disengagement on the subject. For the tweets on *gun control*, whenever Hillary Clinton tweets, notice how Ted Cruz quickly responds by supporting Second Amendment rights. By observing the coordination in relative tweet timing, one can speculate potential influence from Clinton to Cruz.

[3] α is the parameter for fitting $P(\tau) = \tau^{-\alpha}$ on inter tweet arrival v.s. probability log-log plot, with log-binning.

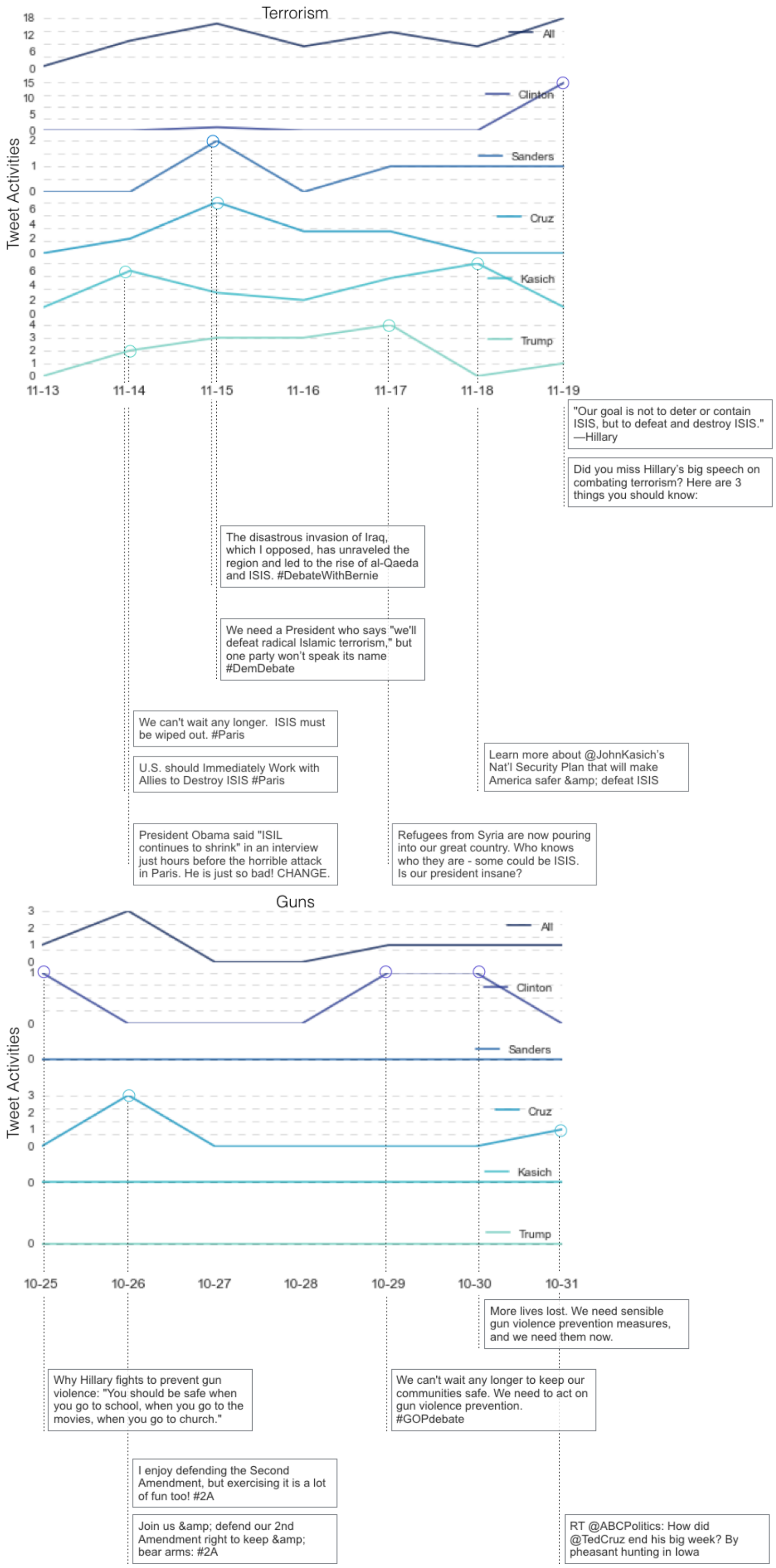


Figure 1-3: Two particular instances: (Top) Terrorism tweets after the Paris shooting in Nov, 2015. (Bottom) Gun control tweets debate in Oct., 2015.

1.2 Approach and Contributions

The main objective in this thesis is to develop a system that detects tweet bursts and analyzes them by their characteristics. We are particularly interested in bursts in which tweets due to exogenous events occur (*hereafter responsive bursts*)[4]. A successful system should be able to identify many responsive bursts, and the analysis of burst characteristics should reveal interesting ways a user responds to exogenous events.

In this thesis, we propose **BurstMapper**, a system to detect and analyze the potential drivers to a tweet time series by multiple user sources in terms of tweet bursts. *BurstMapper* first applies a topic classifier to generate topic-filtered tweet series. Next, a burst detection algorithm is used to extract temporal clusters with high density of tweets. We then characterize bursts along two dimensions of analysis: *semantic coherence* that measures the semantic relatedness of tweets in bursts using a deep neural network derived embedding of tweet contents, *Tweet2Vec* [54]; *causal influence* that measures the potential causal interactions between users by applying Hawkes Process [25], which infers an influence network based on relative event timing.

BurstMapper provides a new tool to examine how exogenous events play a role in how a user responds on Twitter. This approach differs from previous works that use only partial semantic information (e.g., hashtags, URLs) to analyze information influence on Twitter. Additionally, *BurstMapper* aims to cope with information overload by providing a useful abstraction that summarizes tweet content and potential user interactions. The framework therefore offers the capability for efficient data exploration. Lastly, our framework is *modular*, so we can substitute with better models of Twitter semantics and temporal causality if needed.

In a nutshell, *BurstMapper* makes the following contributions:

- We design a burst detection algorithm to extract temporal clusters of high density tweets. Also, to characterize bursts, two measures are designed along

[4] Tweets due to exogenous events do not occur very often. Based on our annotated corpus, only 8% of the tweets are influenced by other voices, with another 30% of tweets triggered by news events. It is therefore important to distinguish responsive bursts in order to examine exogenous event-caused tweets in more detail.

the dimensions of semantics and potential causal interactions. The two measures can be linearly combined as features for classifying responsive bursts. To our knowledge, this is the first work that applies state-of-the-art semantic and temporal-causal models to characterize tweet bursts.

- We create a manually annotated corpus of 7,220 election related tweets by five leading U.S. presidential candidates from 2015 to 2016, where each tweet is labeled as whether it is caused by exogenous events or not.

- Each component in our system is evaluated on the annotated corpus. Experimental results show improvements over benchmark methods in evaluating burst quality and responsive burst classification.

- Exploratory data analysis is performed on the annotated corpus, and a larger dataset that includes the candidates and the users followed by them on Twitter. Analysis reveals how the two measures can help us observe different ways a candidate responds to exogenous events. Testing on the larger dataset reveals how certain groups of users may have a higher level of potentially causal interactions and semantic coherence. Lastly, a small case study discusses how *BurstMapper* provides a more information-rich level of abstraction for summarizing tweet content, in comparison to alternatives using only partial semantic information like hashtags.

1.3 Thesis Overview

Chapter 2 discusses the related research works and explains how this thesis differs from them. Chapter 3 introduces the *BurstMapper* system. Chapter 4 describes data collection, and evaluation results of the system. Chapter 5 demonstrates exploratory data analyses on two datasets using *BurstMapper*. Lastly, we conclude by discussing the implications of our findings, limitations, and future directions.

Chapter 2

Related Works

This thesis shares connections with research works concerning communication influence and behavior modeling on Twitter. This chapter briefly explains these works and discusses their relations to the thesis.

2.1 Communication Influence

Influence is an important notion in communication theory that explains how people interact. Specifically, several theories of influence include: (i) cognitive dissonance [18]- one may act to correct the perception of surroundings that become inconsistent with her prior beliefs, (ii) social judgement theory [27]- a change in attitude is controlled by one's judgement of social context and the effects the change brings, (iii) elaboration likelihood model [43]- one's change of attitude is determined by how she thinks about the persuasive content communicated by others, and (iv) the theory of reasoned action [1]- one's reaction is the dual process of her own attitude and learning from attitudes of others. Influence is also a major concept in many other social science domains. In political theory for example, rhetoric, lexical expression, and interactions can dictate the progression of political discourse [7] or election [29].

In computational linguistics, there has been active interest in underdstanding what affects one's word choice in a conversation. Recent works have applied probabilistic topic modeling to map thematic change in language use; scholarly impact can then be measured by observing coordinating shifts in topic distributions between authors [19]. Another approach analyzes variations in word usage due to the power difference between individuals [9]. Additionally, a recent work [24] proposes probabilistic language model that directly considers person-to-person influence.

Notice that the mentioned concepts share a mental model in which influence is due to both *implicit bias* and *observing others*. In our thesis, *BurstMapper* is inspired by this mental model. Therefore, we hope the proposed framework can serve as an alternative tool for data exploration in these fields. Also, this work differentiates from works in computational linguistics in that this thesis focuses on how *content* could stimulate communication events to *occur*.

2.2 Detecting Bursts on Twitter

Detecting a temporally coherent, or high density region of tweets has close connections to event detection in data streams, where bursts of tweets correspondingly imply changes in data's global latent state [30]. In the context of Twitter platform, event detection is a well-researched topic. One popular way to detect events is to track use patterns of popular and informative words [36][8]. Another way is to segment time where text-based measure for *newsworthiness* is used to rank potential event segments [34]. Other works have proposed generative models of text to directly model varying semantic distribution given time [13].

In this thesis, burst detection is used to extract time regions with *unusual* intensities of tweeting activity. In particular, *BurstMapper* deploys a frequency filter [56], rather than incorporating semantic information. This design choice allows us to explore a greater variability in the semantic characterization of bursts.

2.3 Modeling Semantic Representation on Twitter

There has been a wealth of research on developing computational methods for modeling tweet contents. Many of these works have been task-driven, where specific methods are designed for sentiment classification [40][12][5], syntactic annotation [39], speech acts classification [53], stance detection [50], among many other use cases. Typically, these works construct machine learning models [28][55], or novel features [10][31][44] tailor-made for the given research tasks. These works provide a sample view of the vibrant, and fast changing research landscape in modeling tweet contents.

Very recently, a line of research emerges for modeling Euclidean embedding for tweet contents. Inspired by the widely popular *Word2Vec* model proposed by Mikolov et al. [37], researchers have started to apply different shallow to deep neural network architectures to learn the tweet semantic embedding [54][50]. This thesis focuses on *applying* the established models that provide an expressive representation of tweet semantics.

2.4 Modeling Influence and Propagation on Twitter

Since Twitter is an online *social* platform, an important research question is to understand influence. Many research works focus on the network structure of the following relations and interactions among users [2][58]. Other works have taken the idea further to summarize social context [6], rank users [57], detect rumors [52], etc. Our work differs by using both content and timing-based influence to provide an understanding of potential drivers for tweets, which provides a different perspective to influence on Twitter.

A particular case of influence that involves relative timing of events is information propagation. Propagation of information consists of a user exposing to information, and adopting the information based on her social neighbors. The seminal work of Gomez-Rogriguez et al. [22] models the observed cascades (e.g., retweets, photo reshares) with an underlying influence network. Subsequent works have extended [22] by providing more scalable parameter optimizations [15][20][21], using out-of-network information [38], as well as exploring other factors of influence like burst in product item popularity [47][60].

Amidst the research progress, Hawkes Process [25] is a recently popularized alternative model. This thesis focuses on Hawkes Process because it models multivariate discrete time series (and not just series of *cascades*), which is more general. Researchers have applied the model to explain dependencies among a set of social interaction [4][42][61], topic cascades [16], and tweet streams [14][59]. This thesis focuses on a purely application perspective, where in particular we adopt the discrete-time Hawkes Process proposed by Linderman et al. [35].

2.5 Modeling Behaviors on Twitter

The news-like [32] and openly accessible nature of Twitter has inspired works to understand Twitter users on a population level, such as predicting election [46], study politics [45], inferring health outcomes [41], etc. These works often concerns with profiling users in terms of their demographic attributes and use of words. This thesis focuses on tweet bursts and mapping them along the dimensions of semantic and temporal-causal properties. *BurstMapper* aims to serve as a general tool for providing useful levels of abstraction for summarizing tweet content, where the other works aim to solve case-specific data science problems.

Chapter 3

Methodology

3.1 Definitions

To start, we first describe the data setting used in the thesis. Assuming the data is given in the form of a discrete series of tweet events $\mathcal{S} = \{e_1, ..., e_{|\mathcal{S}|}\}$. Each tweet event is generated by a *source* user account $u \in \mathcal{U} = \{1, ..., |\mathcal{U}|\}$. A tweet e_n is denoted as a $e_n = (e_n^t, e_n^c, e_n^z, e_n^s, e_n^C, e_n^p)$ tuple, with tweet raw data: time e_n^t, user identity e_n^c, when the user previously tweeted e_n^p, and content C_n; as well as content-derived meta-data: topic classification e_n^z, and categorization of stimulus e_n^s.

Additionally, note we can subdivide or filter the original series $\mathcal{S}$ of tweet events. First, the tweet time series generated by user u is denoted by the superscript $\mathcal{S}^u$. Second, we use the underscript z to denote events filtered by topic z, so that $\mathcal{S}_z^u$ is the tweet series by user u regarding topic z. Figure 3-1 illustrates how a raw tweet time series can be subdivided into multiple tweet series *per candidate* and *by topic*.

We also define the discrete time series representation of $\mathcal{S}$ as τ_S. More formally, $\tau_S(i)$ describes the number of event counts occurring between $\Delta_t * i < t \leq \Delta_t * (i+1)$ for some i, where Δ_t is the size of discrete time step. All time series operations use discrete time series representation, with $\Delta_t = 1$ day unless otherwise noted.

A burst $\mathcal{B}_{t_x t_y}^S = \{\forall e_i \in S \; s.t. \; e_i^t \in [t_x, t_y]\}$ is defined as a temporal cluster of tweets in time interval $[t_x, t_y]$. Specifically, a burst is *responsive* if there exists some tweets in the burst that are caused by exogenous events, e.g., breaking news or tweets by

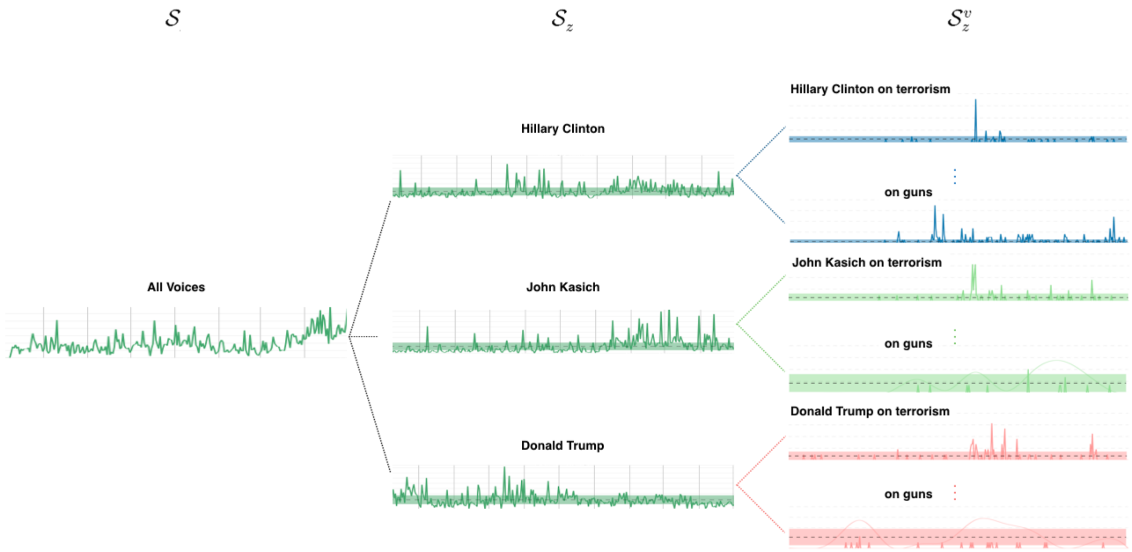


Figure 3-1: An example showing how tweet series $\mathcal{S}$ is divided into sub-user and sub-topic streams $\mathcal{S}^u$; $\mathcal{S}_z^u$. Δ_t=1 day.

other users.

3.2 Overall Objective

The main objective in this thesis is to develop a system that detects tweet bursts and analyze the bursts by their characteristics. Therefore, the first measure of success in the system is the ability to reliably produce bursts of tweets in which many tweets are caused by exogenous events. The second measure of success is how well analysis using the characterizations of tweet bursts can help reveal anything interesting about how a user responds in the face of exogenous events.

3.3 Overview of *BurstMapper*

With the objectives in mind, we introduce *BurstMapper*, a system to identify and characterize bursts in tweet time series. Figure 3-2 displays the overall structure of the system. There are several stages to *BurstMapper*

- *Topic Filtering*: Given tweet series by multiple sources, we filter the series into per-topic subbands using a trained topic classifier [51] to provide topic annotation e_n^z.

- *Burst Detection*: Given the topic-filtered series of tweets, a burst detection algorithm extracts high-density regions of tweet events.

- *Attribute Mapping of Tweets*: The content for each tweet is embedded in a high dimensional Euclidean space, based upon term-occurrence. This is then reduced onto a latent semantic feature space by applying the state-of-the-art Tweet2Vec model [54].

  For the topic-filtered series over multiple users u, Hawkes Process [25] is used to determine influence from user u' to u at time t by inferring an underlying influence network based on relative event timing.

- *Burst Characterization and Data Exploration*: Now, each burst can be mapped along the two dimensions of characterization, *semantic coherence* and *causal influence*, computed based on the semantic and temporal-causal feature mapping of tweets. Using the two measure of the tweet bursts, one can design a classifier that uses the linear combination of the features for responsive burst classification. Further, examining bursts along the two dimensions can reveal interesting ways users respond to exogenous events.

For the remainder of this chapter, each component of *BurstMapper* is explained in greater detail.

3.4 Topic Categorization

We want to annotate topic category for each tweet in the original tweet series $\mathcal{S}$. For this purpose, a multi-class topic classifier is used to label topics for the tweets.

BurstMapper adopts a convolutional neural network based word embedding model for training tweet topic classifier as suggested in Vijayaraghavan et al. [51]. Figure 3-3 illustrates the model architecture. It begins by mapping the tweets to n-gram word

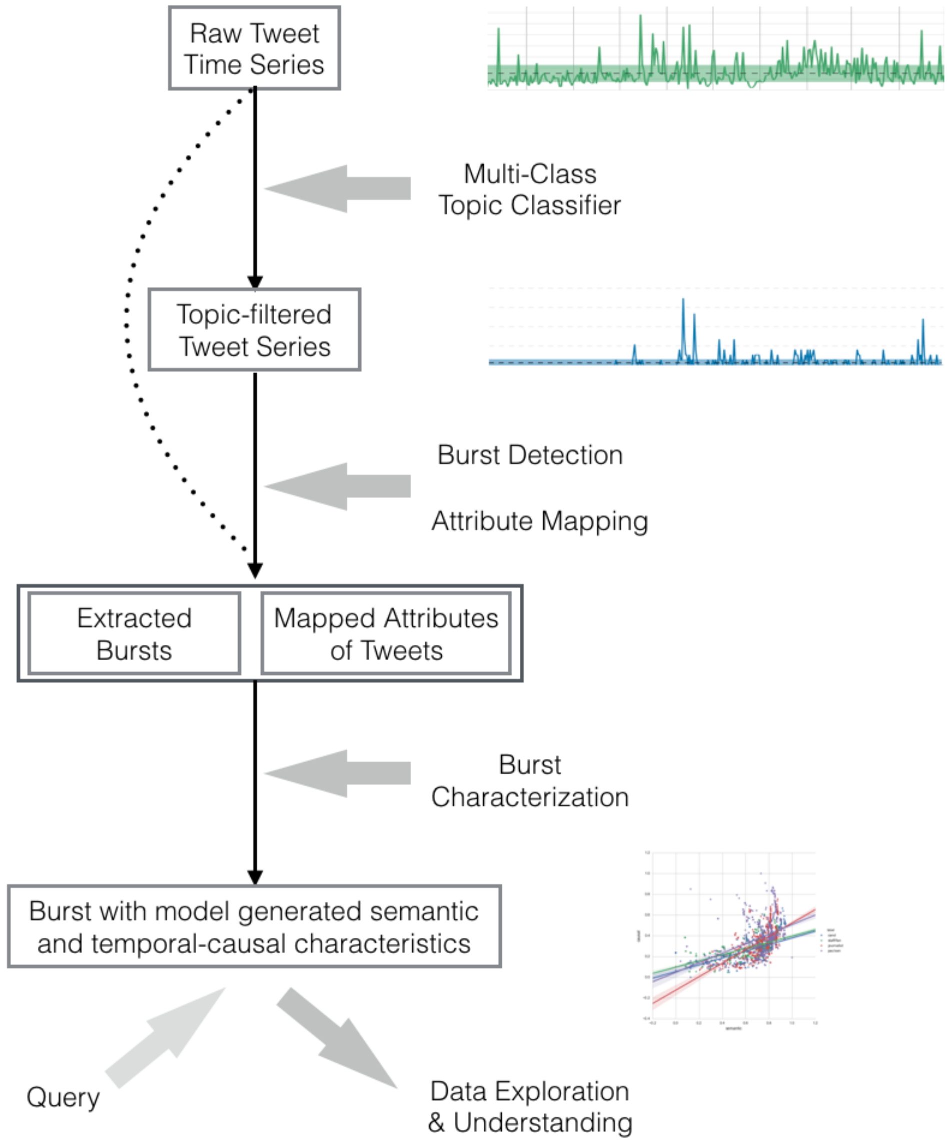


Figure 3-2: The overall pipeline for the proposed framework *BurstMapper*

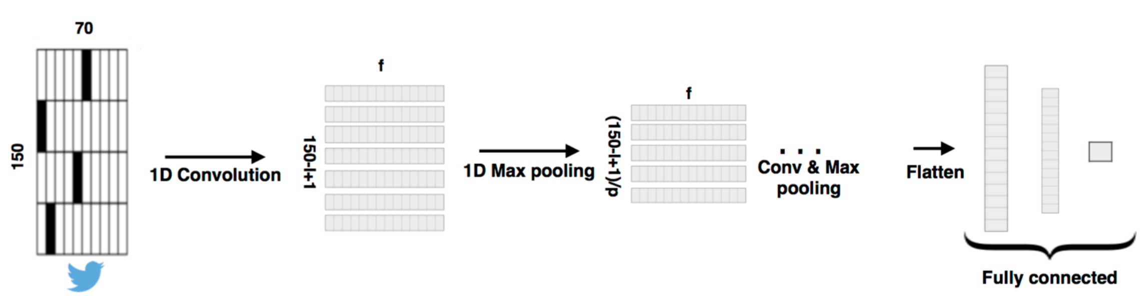


Figure 3-3: The schematic of the computational architecture used in Vijayaraghavan et al. [51] for developing a tweet topic classifier

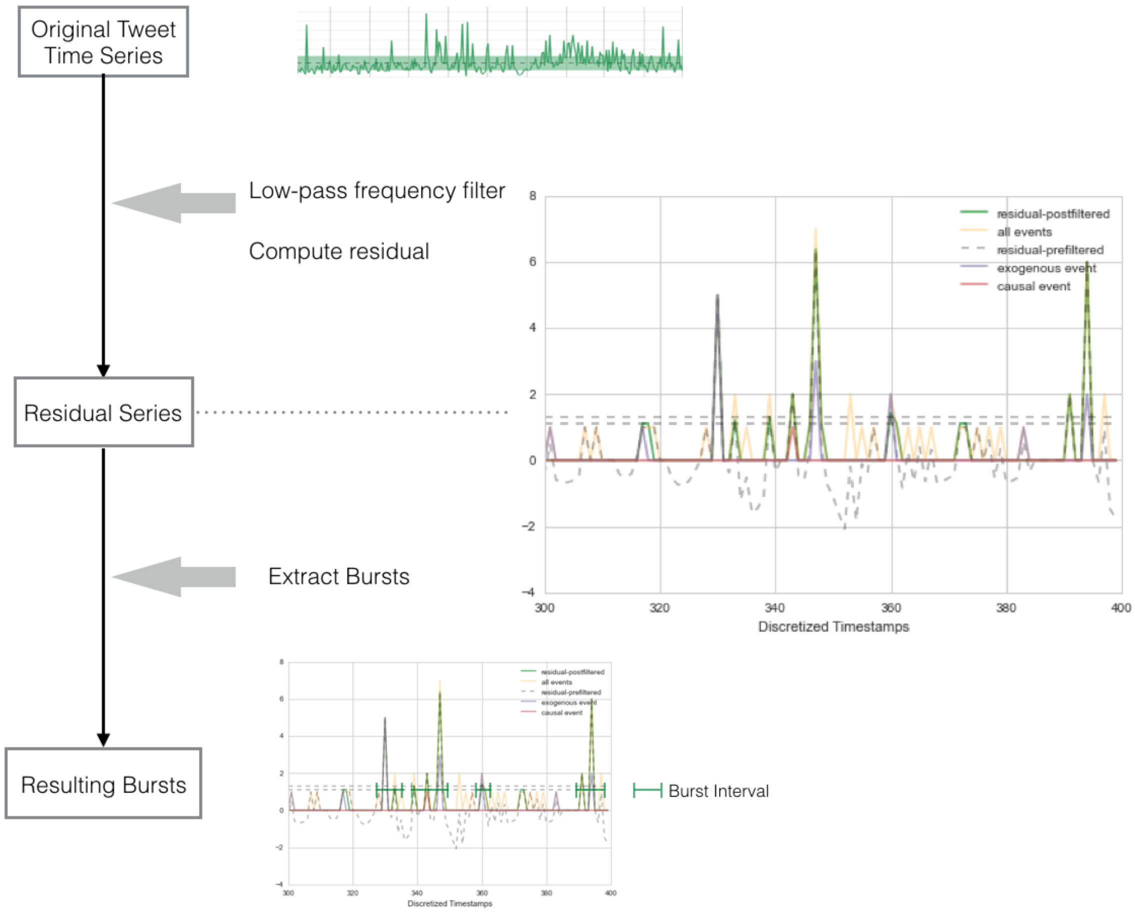


Figure 3-4: The schematic burst detection algorithm. The enlarged figure shows how using the residual signal can filter many tweets that are not due to exogenous events.

embeddings. Afterwards, the n-gram word embedding is passed to a convolutional neural network for supervised classification, with topic annotation determined by the highest score in the softmax layer. In this thesis, we use the pre-trained election topic classifier in [51], with reported F1 score for the topics to be 0.90.

3.5 Burst Detection

The aim of burst detection is to extract temporal clusters of high density tweets. *BurstMapper* applies a two-step procedure for detecting bursts:

(1) Computing residual signal- the series $S_z^{\cdot}$ filtered by topic is passed through a high-pass frequency filter. The filter is constructed by first transforming τ_S into

$\hat{\tau_S}$ using Butterworth filter (this removes sporadically occurring events). Next, the residual signal is computed as $r_S = \tau_S - \hat{\tau_S}$. The values of r_S are set to 0 if at any given time t, $r_S(t)$ is less than some threshold.

(2) Extracting bursts- now, given r_S, we apply a simple greedy algorithm to output bursts. We slide a time window through r_S, and concatenate timestamps t into time interval $[t_x, t_y]$ with $r_S(t) > 0$ into $\mathcal{B}^S_{t_x t_y}$.

Figure 3-4 illustrates the process. Note that the right diagram shows how the tweets caused by exogenous events can be captured by the residual series. The overall burst detection algorithm is summarized in Algorithm A-1.

3.6 Attribute Mapping of Tweets

3.6.1 Semantic Mapping

BurstMapper adopts the *Tweet2Vec* model [54] for mapping the tweet contents onto a 600-dimensional latent semantic feature space. Tweet2Vec is a recently proposed model for embedding tweet that achieves state-of-the-art results on measuring semantic similarity [54]. The model learns semantic embedding based on a character-level CNN-LSTM encoder-decoder architecture. After the model is trained, for any given tweet e_n with content e_n^C, we use the encoder portion of Tweet2Vec to produce the 600-dimensional latent semantic embedding features e_n^{sem}. Figure 3-5 displays the architecture.

3.6.2 Temporal-Causal Mapping

For temporal-causal mapping of tweets, we adopt the *Hawkes Process* model [25]. Hawkes Process is a point process in which the frequency of event occurrence depends on the event history. It is dictated by the *intensity* function of time, denoted by $\lambda(t)$ as describing the expected event occurrence frequency at time t. Hawkes Process corresponds to Poisson Process with varying-rate in discrete time domain.

For a tweet series S_z^u by user u on topic z, the corresponding intensity function of

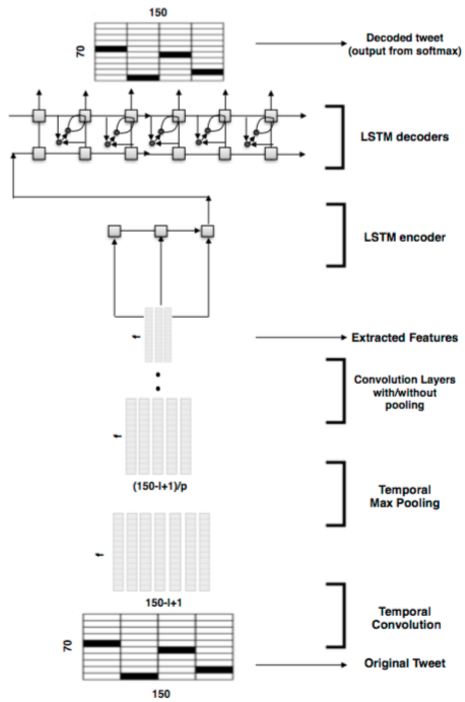


Figure 3-5: The schematic of the computational architecture used in Vosoughi et al. [54] for developing Tweet2Vec, vector embedding of tweet content.

time, $\lambda_u(t)$, that describes user u's frequency of tweets, is:

$$\lambda_u(t) = \overbrace{\mu_u}^{intrinsic\ influence} + \overbrace{\sum_{t_i<t, e_i^c=u'} \alpha_{u'u}\phi(t|t_i)}^{extrinsic\ influence} \tag{3.1}$$

μ_u is user u's tweeting frequency, independent of other users. $\sum \alpha_{u'u}\phi(t|t_i)$ accounts the historical influence of other users u' on u's tweeting frequency, where $\alpha_{u'u}$ describes user-to-user influence from u' to u. $\phi(t|t_i)$ is a time-decaying weight function to discount influence over time. For convenience, define $\lambda_{u'u}(t) = \sum_{t_i<t, e_i^c=u'} \alpha_{u'u}\phi(t|t_i)$. Note that Hawkes Process is a form of Generalized Linear Model, and the parameters $\alpha_{u'u}$ correspond to regression-based Granger Causality [23]. See [25][35] for details. In this thesis, *BurstMapper* applies the discrete-time Hawkes Process [35].

To fit a Hawkes Process model for S_z, jointly consider $\lambda_u(t)$ for all users u that corresponds to tweeting frequency for series S_z^u. One can learn the model parameters by MLE learning on data log-likelihood.

To account the dynamic nature of tweet time series, Hawkes Process is fitted multiple times on a sliding time window W (length W_l) through time. For each tweet e_n, we have its corresponding causal influence mapping e_n^{cau}:

$$e_n^{cau} = \frac{1}{W_l} \sum_{t:t \in W} \sum_{u'} \lambda_{u' e_n^c}(t) \tag{3.2}$$

essentially accounting user influence averaged over all time windows where $t \in W$.

3.7 Burst Characterization and Data Exploration

3.7.1 Burst Characterization

Given the extracted bursts, and the semantic and temporal-causal mapping of the tweets, we now compute the two measures of bursts characteristics.

Semantic Coherence

The first measure is *semantic coherence*. The intuition is that if users participate in similar sub-topic discussion, the pairwise semantic similarity among tweets should be high. Additionally, if many people participate in the same sub-topic, then it is likely to have wider variability in content expression. Specifically, for burst $\mathcal{B}_{t_x t_y}^{S_z}$, we have:

$$\begin{aligned} SC_{\mathcal{B}_{t_x t_y}^{S_z}} &= \frac{1}{Z} * Avg(\boldsymbol{m}) * Ent(\boldsymbol{m}) \\ where\ \boldsymbol{m} &= [e^{-|e_j^{sem} - e_i^{sem}|_F},\ j \in kNN(i)\ \forall e_i \in S\ s.t.\ e_i^t \in [t_x, t_y]] \end{aligned} \tag{3.3}$$

where Z is the normalizing constant. Basically, for each tweet e_i in the burst, we first compute the average of the semantic distances between e_i and its top-k nearest neighbors (which are tweets by other users). This results in an array, $\boldsymbol{m}$. To compute semantic coherence, the average (Avg) and entropy (Ent) of $\boldsymbol{m}$ are computed.

Causal Influence

Causal influence aims to measure temporal-causal activity level within a burst. Using the temporal-causal mapping of the tweets, we have:

$$
\begin{aligned}
CI_{\mathcal{B}^{S_z}_{t_x t_y}} &= \frac{1}{Z} * Avg(\boldsymbol{g}) \\
where\ \boldsymbol{g} &= [e_i^{cau} * \phi(e_i^t - e_i^p)\ \forall e_i \in S\ s.t.\ e_i^t \in [t_x, t_y]]
\end{aligned}
\tag{3.4}
$$

The basic intuition is that e_i^{cau} computes the level of potential causal influence of users u' on the tweet e_i. $\phi(e_i^t - e_i^p)$ computes the wait time between tweet e_i and when the user last tweeted.

Applying the Two Measures for Characterizing Bursts

In the later chapters, we will show that by linearly combining the two measures for a burst, one can develop a classifier for identifying responsive tweet bursts. Also, we will show that mapping the bursts along the two dimensions can reveal interesting ways different users generally respond to exogenous events.

3.7.2 Querying from the Bursts

As the system churns raw tweet series and outputs the detected bursts that are mapped along the two measures, we also want to be able to search the bursts according to some queries of interest. Naïve query pattern matching may fail to retrieve semantically and temporally similar tweets that do not contain the query terms. We apply a simple two-step procedure. First, we extract a set of *seed tweets* based on those with matching query terms. Second, we apply an exemplar-based clustering algorithm [33] using the seed tweets as examples. The similarity between pairs of events for clustering algorithm is computed as

$$
sim(e_1, e_2) = e^{-(\beta * |e_1^{sem} - e_2^{sem}|_F + (1-\beta) * |e_1^{cau} - e_2^{cau}|_F)}
\tag{3.5}
$$

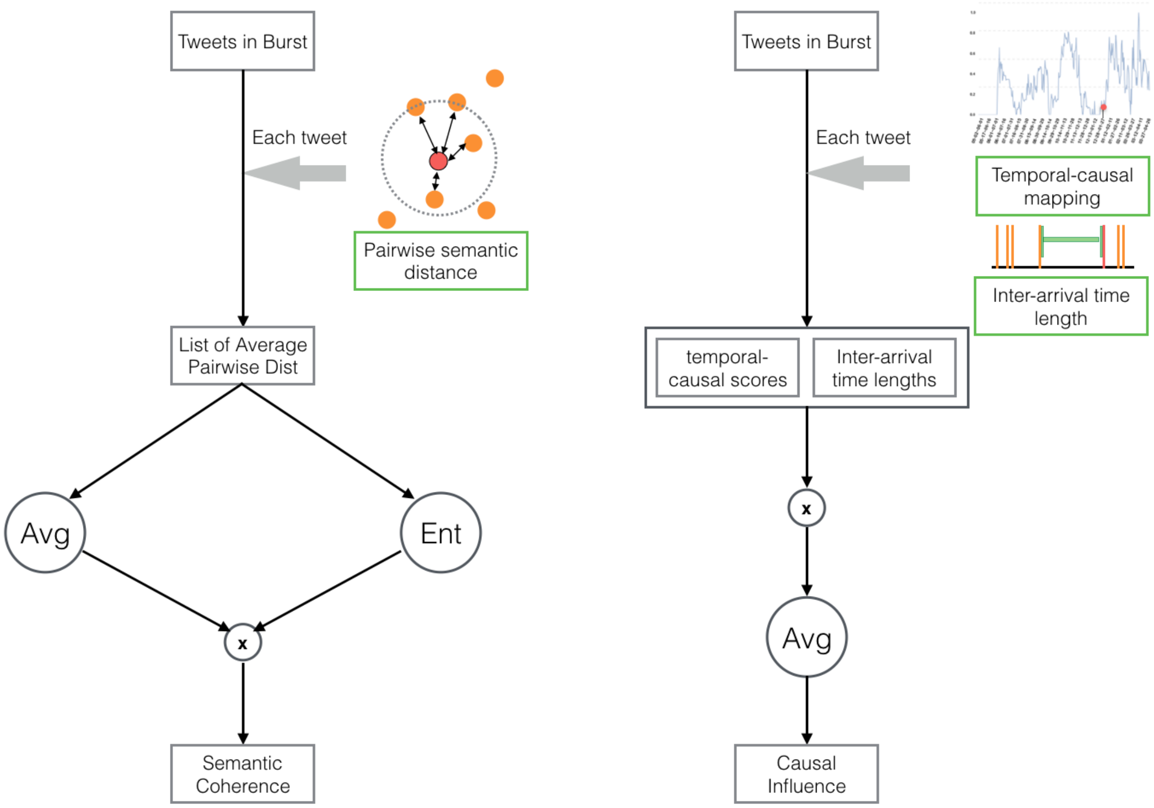


Figure 3-6: A graphical explanation of how the two proposed burst characters measures are computed given the tweets and their corresponding attributes in a burst. (Left) Computation of *semantic coherence*; (Right) Computation of *causal influence*.

. After the cluster IDs are assigned, instances that are far away from the assigned cluster centroids will be dropped. In the end, the process returns bursts that contain tweets closest to the center of the assigned cluster id. This process is described in Algorithm A-2. Figure 3-7 displays the procedure visually, as well as sample results based on two queries ("Sandy Hook" for the Sanders remark controversy; "Belgium" for the terror attack). Notice how the returned contents are semantically relevant to the query term.

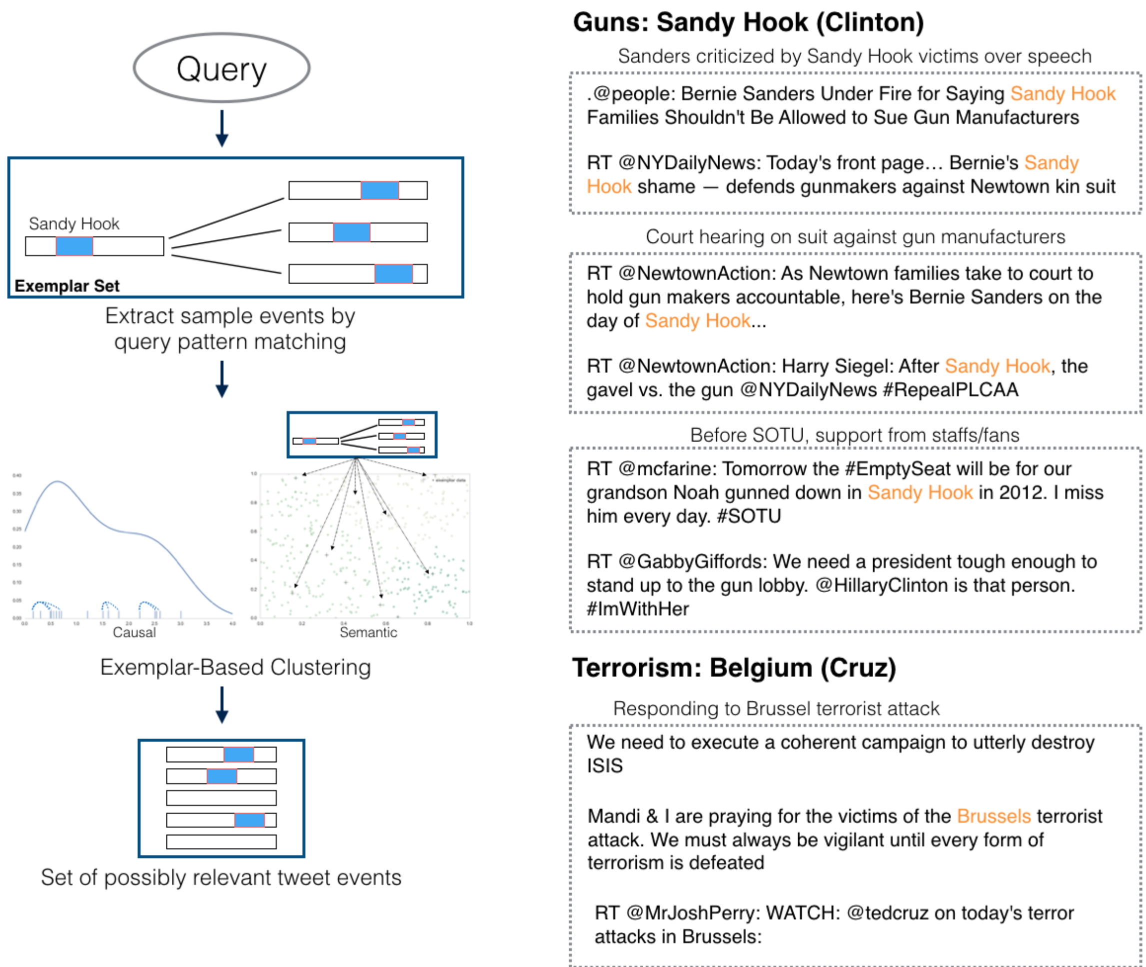


Figure 3-7: (Left) The schematic of the process to query tweet events based on the two measures; (Right) Two examples of returned tweets using the event-based queries.

Chapter 4

Evaluations

4.1 Datasets Description

To evaluate *BurstMapper*, we focus on tweets relevant to the 2016 U.S. presidential election. We map each tweet to one of 22 election topics (excluding the 'Others' category, see Table 4-1, Appendix B explains the definition for each topic in more details), based on the pre-trained model used in [51].

Racial Issues	Justice	Income Inequality
Budget/Taxation	Campaign Finance	Abortion
LGBT Issues	Terrorism	Veterans
Economy	Drugs	Guns
Jobs/Employment	Ethics	Foreign Policy/National Security
Immigration	Surveillance/Privacy	Health Care
Environment/Energy	Financial Regulation	Trade
Education		

Table 4.1: List of Topics

Two datasets are used for evaluation experiments. The first corpus, named *FiveCandidates*, provides annotated data based on historical tweets by five U.S. leading presidential candidates[1] for the 2016 election. The second dataset, called *FiveCandidatesExpanded*, is obtained by considering a larger set of users of both the candidates and the accounts followed by them on Twitter.

[1]Candidates: Hillary Clinton, Bernie Sanders, Ted Cruz, John Kasich, and Donald Trump. The list is selected based on the candidates that were still running in April, 2016.

4.1.1 Detailed Description of Datasets

FiveCandidates

In order to collect the tweets, we use GNIP's[2] Historical Power Track to obtain the five candidates' tweet history. We are able to obtain tweets from May 2nd, 2015 to May 2nd, 2016, totaling 29,016 tweets. After passing through topic categorization funnel that filters tweets in the 'Others' category[3], we are left with 7,220 tweets, comprising of 25% of all tweets. Figure 4-1 (Left) shows the tweet count distribution.

FiveCandidatesExpanded

To collect the tweets for the second dataset, we use the Twitter Public API to obtain users that the five candidates follow. We apply a profile-based examination to filter accounts that may potentially be celebrities (those with number of followers more than 500,000). The historical tweets by the new set of users are extracted from Jan 1st, 2016 to May 2nd, 2016. We further filter users with less than 3 tweets related to any of the 22 topics. The preprocessing steps result in a total of 1,470 users with 117,648 tweets after applying topic categorization filter.

4.1.2 Label Annotation

The *FiveCandidates* dataset additionally undergoes a manual labeling process to classify whether a tweet is caused by some exogenous events. We resort to Amazon Mechanical Turk to have online workers to do the labeling. Each tweet can be categorized as: (i) *Response* (tweet influenced by *tweets* of other voices), (ii) *External* (tweet influenced by breaking news events), (iii) *Internal* (tweet not due to exogenous events), and (iv) *Error* (tweet where topic categorization errs). Figure 4-1 (Right) shows the resulting proportion of tweet categories, with the worker agreement of 0.70

[2]GNIP(http://gnip.com) is a service provider (acquired by Twitter) for social media data ingestion. It has a "Historical Power Track" that enables search-back-in-time capabilities to Twitter's historical tweets.

[3]There are 21796 'Others' tweets, in which the candidates elaborate on personality/credibility, polls, campaign events, etc. An example: 'RT @FoxNews: .@tedcruz: "#SuperTuesday revealed pretty powerfully that the only candidate who has a path to beating @realDonaldTrump is me.'

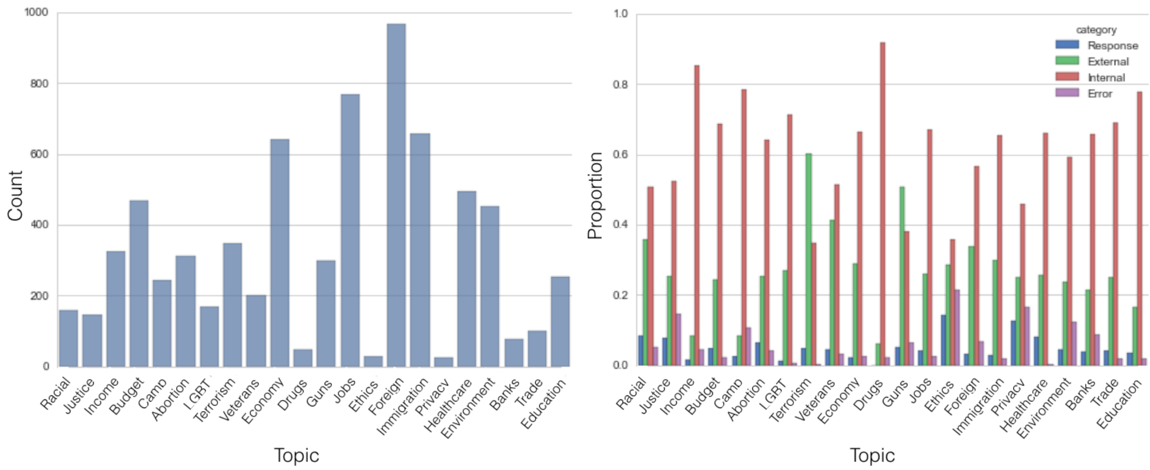


Figure 4-1: Left: Count of events for each of 22 topics. Right: the labeled proportion determined by manually labeling.

(3+1 AMT/dedicated workers per tweet)[4].

For the purpose of the experiments, the Response and External categories both account for the type of tweets caused by exogenous events. Note the relatively low proportion of External and Response tweets. It is therefore understandable that correctly identifying all tweets caused by exogenous events would be difficult.

4.2 Evaluating Burst Detector

The first evaluation tests how well the burst detector (**BurstDetector**) is able to produce bursts with reasonable quality. We want the time regions extracted by the bursts to include a high proportion of tweets caused by exogenous events (a.k.a. *responsive tweets*). Since the time region due to each burst is a collection of discrete timestamps, the evaluation metrics is computed as: (i) Precision- Prob(time stamps in the bursts of responsive tweets | time stamps of all bursts), and (ii) Recall- Prob(time stamps in the bursts of responsive tweets | all time stamps of responsive tweets).

To compare against the proposed method, two competing baselines are used. The

[4] To control for the subjective nature of online workers (their decision might be biased by their view of the election), we invited a dedicated worker who is knowledgeable of the election background, and gave him a few hours training with detailed instructions. We provided him an annotated corpus of 30 tweets (labeled by the author), in which he was able to answer with over 85% accuracy. We weighted his answers with 2 votes.

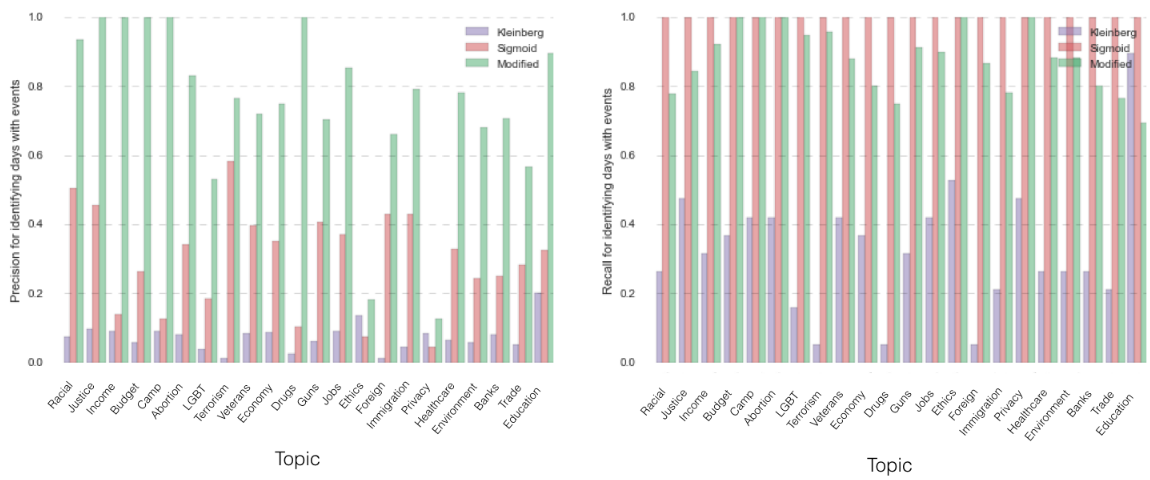


Figure 4-2: Left: Precision of correctly identify the timestamp containing externally stimulated tweets. Right: Recall plot.

first baseline (**Sigmoid**) applies a thresholded sigmoid activation to the tweet volume time series. The transformed series is passed to a greedy grouping procedure for producing bursts. The second baseline (**Kleinberg**)[30] is a commonly used method that captures burst as state transitions describing the tweet time series.

Figure 4-2 shows the results. The result shows that while naïve baselines such as **Sigmoid** has high recall, in exchange, the precision drops significantly. **Kleinberg** is not comparable in performance due to the model assumption that the events are homogeneous. On Twitter, tweets are due to a complex set of influence. The assumption behind Kleinberg's detection method wouldn't suit this dataset well.

4.3 Evaluating Measures for Burst Characterization

The next phase of evaluation applies the measures to classify whether a burst is responsive (for such bursts, let the default minimum number of responsive tweets within a burst be 1). The goal of this evaluation is to test whether the proposed measures for characterizing burst are in fact informative.

To get some inspiration, we run a logistic regression test to see how *semantic coherence* and *causal influence* measures, as well as feature interactions, are significant in characterizing bursts. For the regression test, the response variable holds the value

Temporal-Causal	Semantic	Interactions
$\overline{e_i^{cau}}$	$Avg(\boldsymbol{m})$	$\overline{e_i^{cau}}$ * $Avg(\boldsymbol{m})$
$\overline{\phi(e_i^t - e_j^t)}$	$Ent(\boldsymbol{m})$	$\overline{e_i^{cau}}$ * $Ent(\boldsymbol{m})$
CI	SC	$\overline{e_i^{cau}}$ * SC
		$\overline{\phi}(e_i^t - e_j^t)$ * $Avg(\boldsymbol{m})$
		$\overline{\phi}(e_i^t - e_j^t)$ * $Ent(\boldsymbol{m})$
		$\overline{\phi}(e_i^t - e_j^t)$ * SC
		CI * $Avg(\boldsymbol{m})$
		CI * $Ent(\boldsymbol{m})$
		CI * SC

Table 4.2: A list of features used in regression analysis.

of 1 if the datum is a *responsive burst*, and 0 otherwise. A test is performed on each topic-filtered series S_z (for a list of variables, see Table 4-2). Result shows that each feature, as well as the feature interaction across categories, show certain amount of significance for fitting the logit model (see Appendix D for more details).

Method	Precision	Recall	F1-Score
Random	0.574	1.0	0.678
Causal Interaction Only	0.664	0.755	0.702
Semantic Only	0.777	0.698	0.734
Joint Features	0.774	0.754	0.761

Table 4.3: Overall performance burst classification over 22 election topics

With the developed intuition, the next test aims to construct a *responsive burst* classifier to see how can the measures be practically applied. The classifier model being used is the standard Gradient Boosted Tree Classification[5] trained on the entire corpus. For the features, the components for each measure, shown in Table 4.2, are linearly combined as part of the feature vector. The feature vector also includes a binary vector of topic category ID.

There are 1,030 detected bursts, of which 574 are labeled responsive. A model is trained on each election topic, and evaluated with a 5-fold cross validation. The results are summarized in Table 4-3, where we also display the naïve result by randomly guessing 1. When employing features associated with both *semantic coherence* and *causal influence*, the performance provides non-trivial improvement over random

[5]Hyperparameters are tuned to be: estimator number-50, sub-sample rate-0.5, max depth-3

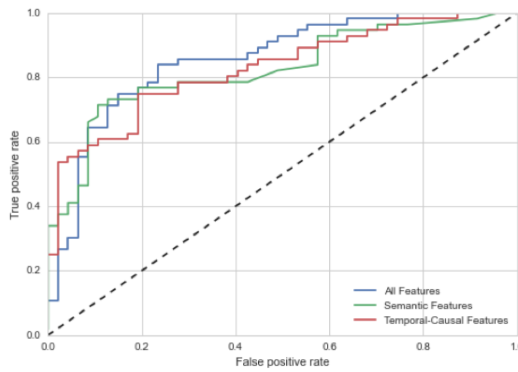


Figure 4-3: The receiver-operating characteristics (ROC) curve of the classifier on responsive burst classification task.

baseline, as well as using features associated with only a single measure (8%/3% improvement in F1-score).

To better understand the performance of the binary classifier, Figure 4-3 also shows the overall receiver-operating characteristics (ROC) curve that shows the performance of a classifier as the discrimination threshold is varied. This plot can help us to understand the trade-offs between *false-positive*(False Positives/(False Positives+True Negatives)) and *true-positive* (True Positives/(True Positives+True Negatives)) rates. Depending on the application and how sensitive one is to getting false positive errors, the classification threshold can be changed accordingly. For example, at a tolerance of 24% false positives, the system can achieve almost 90% precision.

Remark The burst detector/classifier extracts 1000 bursts from the 29,000 tweets in the *FiveCandidates* corpus. This is a data reduction of 1000/29000, or 30-fold. Furthermore, the system selects about ∼66% of those 1000 bursts, e.g. 660 bursts. This provides a data reduction of 660/29000 , or 44-fold. To summarize, this method of "coping with information overload" involves: first, an aggregation from tweets to voice/topic-conditioned bursts; second, a filter for whether a burst is responsive.

Chapter 5

Data Explorations

In this chapter, we want to examine the datasets in more detail to see how using the two measures to analyze burst can help reveal interesting ways users respond to exogenous events. In particular, three case studies are presented:

- We examine *FiveCandidates* corpus in greater detail. The detected bursts are mapped along the *semantic coherence* and *causal influence* measures to explore how exogenous events drive each presidential candidate to tweet differently.

- Data exploration is further tested on the *FiveCandidatesExpanded* dataset. Here we examine if the measures can reveal how certain *groups* of users may have a higher level of semantic coherence and potential causal interactions.

- The last case study discusses how *BurstMapper* delivers useful findings not provided by analyzing tweets with only partial semantic information like hashtags, URLs, etc.

5.1 Deeper Dive in FiveCandidates

Given the extracted bursts, a natural next step is to see how the bursts are characterized along the two dimensions, semantic coherence and causal influence. We focus particularly on candidate-level analysis to understand how a candidate becomes responsive to exogenous events.

For each burst, we compute the *semantic coherence* and *causal influence* measures. Specifically, the measures are computed with respect to the tweets by the particular candidate within a burst. We plot all bursts on a *semantic coherence* by *causal influence* plane. The motivation is that some bursts are not necessarily both highly semantically coherent and temporally causal, and it is thus interesting to explore under what situations might bursts not have high values for both measures[1]. The results are plotted in Figure 5-1.

We first examine the region with high semantic coherence and high causal influence measures for a burst. Comparing the bursts between Hillary Clinton and Bernie Sanders, observe how Sanders' bursts are mostly about addressing and attacking his fellow candidates. On the other hand, Clinton focuses more on addressing supporters and announcing policy plan when major breaking news events occur.

Upon closer inspections on Sanders' tweets, he has been relatively vocal on topics where other candidates tweet less frequently about (e.g., financial regulation, worker inequality, etc). Also, many of his tweets are not burst-related, but rather part of frequent promotions of his own ideology. However, it turns out Sanders becomes temporally responsive and semantically coherent on battleground topics to attack fellow candidates. Therefore, in terms of responsiveness to exogenous events, Sanders is actually more responsive than Clinton, where in contrast Clinton's tweets are considerably more composed and planned.

Now take a closer view at Clinton's tweets. When Clinton is temporally responsive and semantically coherent, she is usually responding to some important news events or controversial issues. One can observe that Clinton takes advantage of the opportunities to make speeches or plan announcements. Since these announcements take time to prepare, they often make Clinton slightly lag behind her competitors. To summarize, analyzing along the *semantic coherence* and *causal influence* measures can reveal different ways the candidates respond to exogenous events. We find that mapping the bursts along the two dimensions also tell us how other candidates re-

[1]Temporal-causality measures if relative tweet timing between voices follows a consistent pattern. It is possible to have a burst not having strong temporal-causality between voices, but still be 'responsive' (e.g., a terror attack, which happens only a few times in a year, is such an example).

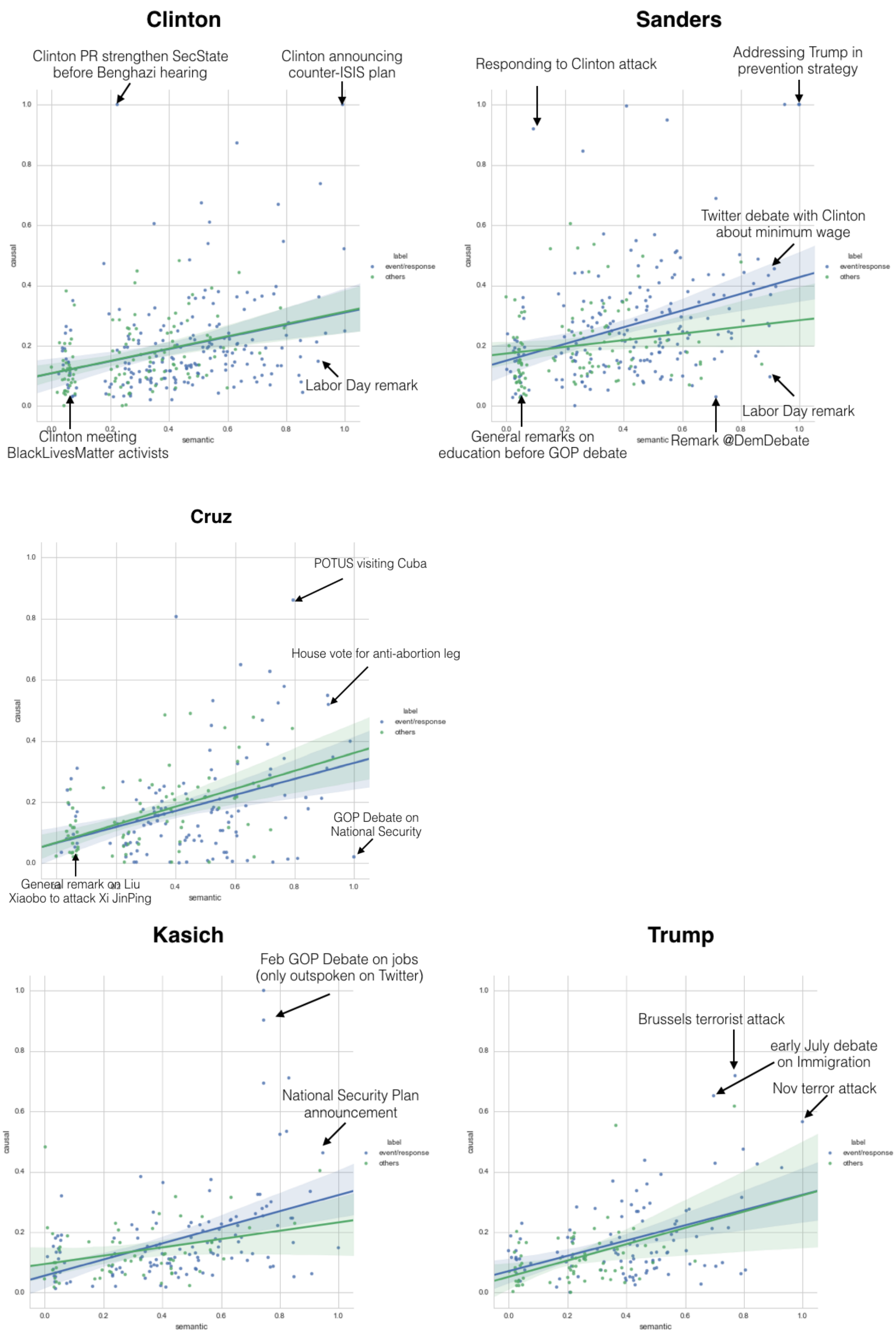


Figure 5-1: Mapping the bursts onto semantic coherence-causal influence plane with respect to each presidential candidate. Non-responsive bursts are also plotted as comparison.

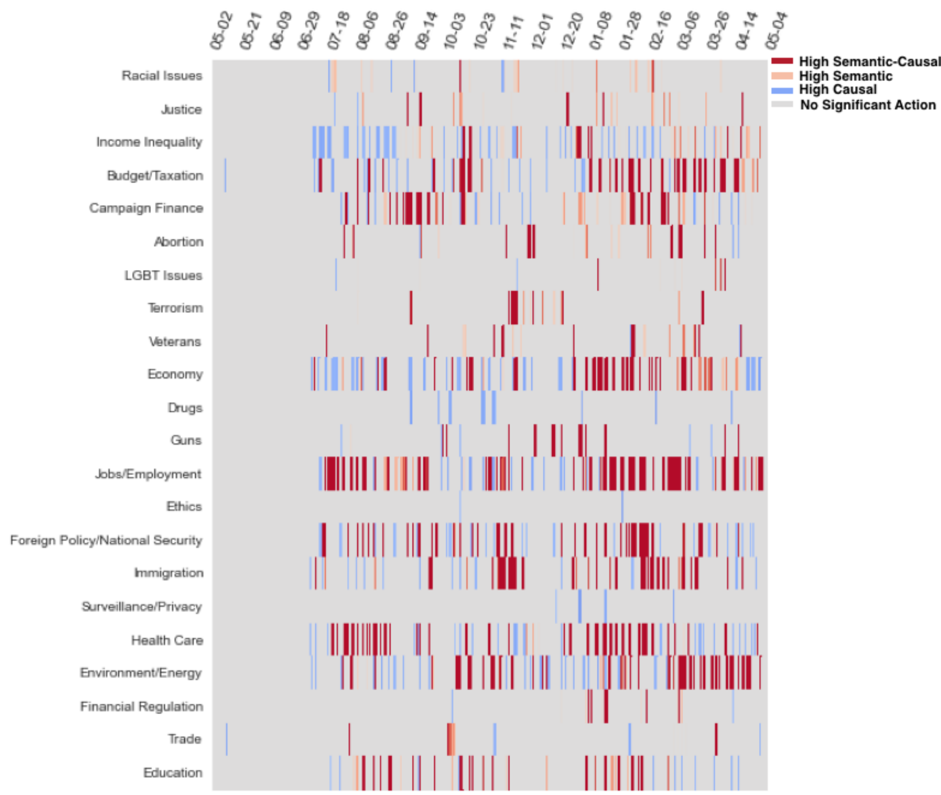


Figure 5-2: Heatmap of Sanders' tweeting activity in terms of the burst measures.

spond to exogenous events as well. Lastly, it is interesting to point out that the two dimensions are positively correlated to each other (with correlation coefficient of 0.33, $p < 0.001$). This suggests that as the candidates try to be more semantically coherent, it is likely that such semantic similarity is due to potential causal interactions between candidates.

Now we explore another way to examine bursts in terms of the two measures by plotting the measures over time with a heatmap. Figure 5-2 displays such heatmap for Sanders over the 22 election topics. In the heatmap, the color bands (red, orange, blue) imply (high-high, high-low, low-high) *semantic coherence-causal influence* measure pairs in terms of the measure value. Note Sanders typically becomes more semantically coherent and temporally responsive later in the election campaign. Also, in the topics Sanders engages less on, he tends to become responsive when important news events occur (e.g., Paris and Belgium terror attacks). These observations reflect a progression in social media strategy: as the election process got more heated,

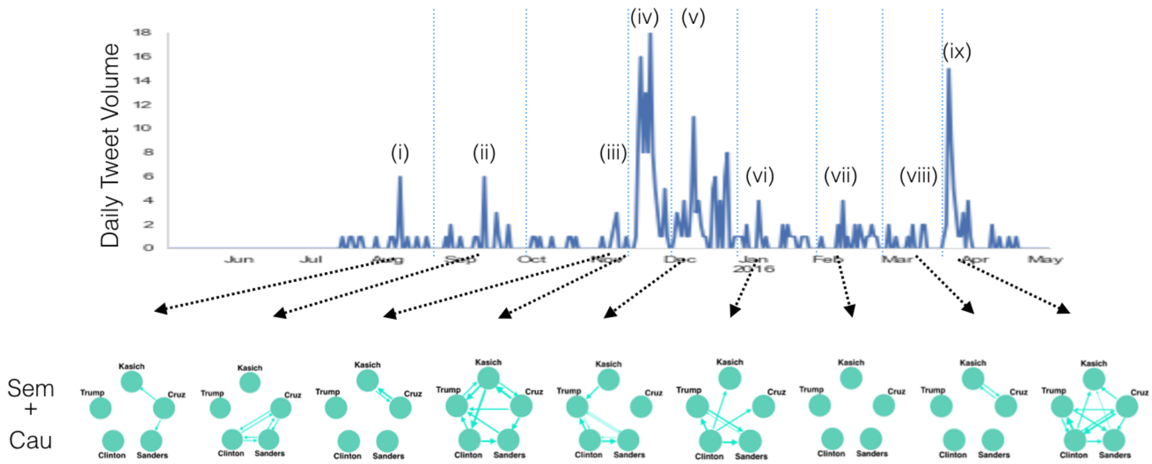


Figure 5-3: Mapping information influence networks over time for election topic terrorism. Top panel shows the tweet time series of daily tweet volume, divided into time regions. Bottom panel shows the inferred information influence network in each region. See Table 5.1 for descriptions of news events in each region.

Sanders could not focus solely on his own narrative without engaging on topics that other candidates enjoy strong presence in.

Lastly, we explore how the semantic and temporal-causal information can be used to derive a network of information influence between candidates. Figure 5-3 displays the tweet time series filtered by the topic *terrorism*. The time series is divided into nine segments describing major events relevant to the topic (Table 5.1 provides a detailed description of the events occurring in each segment). The user networks shown in the bottom panel describe the inferred information influence between user u' and u, defined by

$$I_{u',u} = \frac{1}{|\{e_i \in \mathcal{B} | e_i^c = u\}|} \sum_{e_i \in \mathcal{B}, e_i^c = u} \sum_{e_j \in \mathcal{B}, e_j^t < e_i^t, e_j^c = u'} \gamma \lambda_{e_j^c e_i^c}(e_i^t) + (1-\gamma) e^{-|e_i^{sem} - e_j^{sem}|_F}$$

, where $\mathcal{B}$ spans a set of bursts within a time region (γ is determined empirically as 0.75). In this experiment, information influence is determined by temporal-causality due to relative tweet timing, as well as how one conforms semantically to previously posted contents by other users. Take regions (iv) and (v) for example. The major news event in region (iv) is the Paris terror attack. Not only do we observe a surge

Region	Description
(i)	The first GOP Debate, Cruz is particularly active on Iran Deal and anti-ISIS stance.
(ii)	9-11 Anniversary
(iii)	Kasich and Cruz responding to the Russian Airline crash incident.
(iv)	Paris shooting: each candidate responds within a day or two; Cruz is the first vocal candidate; Kasich makes several speeches, interviews, and press releases; Trump becomes responsive by attacking Obama; Clinton finally responds with a policy plan announcement.
(v)	Clinton continues to be engaged in the topic; Trump attacks more on Democrats.
(vi)	DNC Debate with discussions on anti-terrorism, as well as responding to Trump's controversial remarks; other GOP candidates also respond to the Democratic Debate
(vii)	February GOP Debate, not as heated debate.
(viii)	GOP Debate, Trump not engaging.
(ix)	Brussels attack: candidates elevate engagements on the topic.

Table 5.1: Detailed description of events in each region specified in Figure 5-3.

of information flow between users, notice the thick edges from Kasich to Clinton and Trump. It turns out that Kasich made an important press statement on anti-ISIS plan. Clinton and Trump both actively responded, with Clinton releasing a policy plan as well. In region (v), notice how Trump is trying to rob the narrative from Kasich and Clinton by elevating his anti-ISIS stance. In addition, he starts to attack Muslim, President Obama and Clinton in an attempt to get media attention.

5.2 Exploring the FiveCandidatesExpanded Dataset

After demonstrating how *BurstMapper* can be used to explore tweet bursts for a small set of users, one follow-up question is: is it possible to apply this framework to a larger set of users? Since the dataset now becomes potentially too large to create a dataset with labeled ground-truth, it would be ideal if our system can still provide a useful level of abstraction that summarize tweet content. To address the questions, we test on the *FiveCandidatesExpanded* dataset. In particular, for each of the five candidates, we focus on the subset of the 1,470 users who are followed by the candidate on Twitter. Additionally, we categorize the users as either the candidate

Figure 5-4: For each tweet series filtered per candidate group, map the bursts with the semantic/causal measures with respect to each type of voice group.

herself, campaign staffs or supporters, or journalists. Hawkes Process model training is done on the subsets of users.

Similar to what was done in the *FiveCandidates* dataset, we map the bursts along the two characteristic dimensions, with respect to each user category. The result is shown in Figure 5-4. For Clinton, observe that campaign staffs and supporters are in general highly semantically coherent and temporally responsive in comparison to the journalists. This is evident in the actual tweets, where the campaign staffs are actively stating supports for Clinton over various election issues. In contrast, for a candidate like Kasich, the relative temporal-responsiveness for journalists is higher than his supporters. To explain this difference, a quick look into the set of voices reveals that Kasich tends to follow many journalists who are also well connected to each other (with an average degree of 25.2). For Sanders, who also follows many journalists, has a similar trend for journalists, while being slightly more semantically coherent. Interestingly enough, the journalists are also well connected (with an average degree of 40.0). By looking purely at contents and tweet timing, we can already discover some non-trivial aspects in the data.

We can also compare how the levels for potential causal interactions and semantic coherence differ for groups of users *within tweet bursts*. Figure 5-5 shows the average semantic distance and temporal-causal mapping e_n^{cau} for each tweet e_n relative to start of the burst for candidates Clinton, Sanders, and Kasich. For the journalists followed

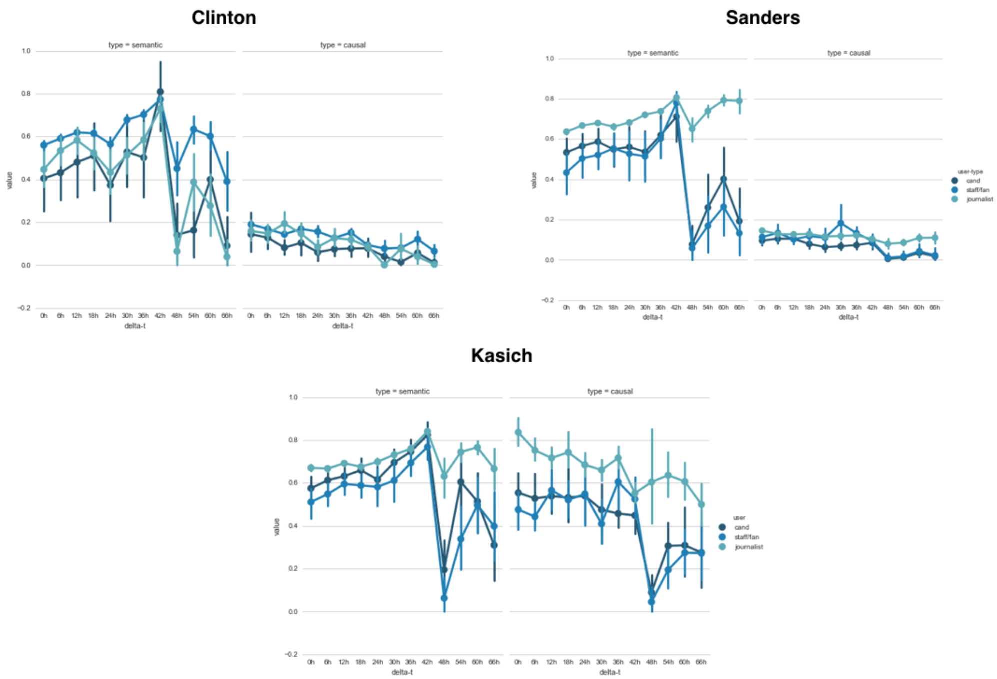


Figure 5-5: Mapping the within-burst variability using the burst measures with respect to series filtered per candidate group. Left/Right panel: semantic/causal measure; X-axis: time past since the burst first initiated.

by Kasich, the tweets are generally semantically coherent and temporally-responsive throughout a burst. As with Clinton, the journalists are semantically coherent in the early phase of a burst, but notice there is a small bump at the end of a burst, suggesting a "lagged" behavior for the journalists to respond. In short, taking the perspective of a within-burst view of the tweets can also reveal some descriptive response patterns of the users.

5.3 BurstMapper v.s. Traditional Approaches

We further elaborates how our system contrasts from using partial semantic information for analysis as in many previous works.

Referring again to the FiveCandidatesExpanded dataset. This case study focuses on Hillary Clinton and the accounts she follows on Twitter. Consider analyzing tweets using only popular hashtags. We find that popular hashtag occurrence do

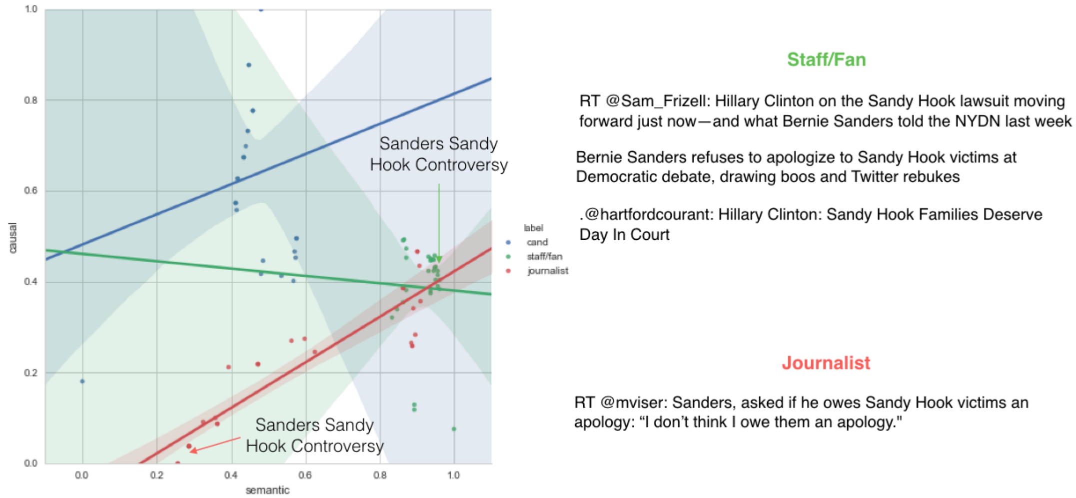


Figure 5-6: Mapping of bursts relevant to topic *Guns*. Those related to the incident of Sandy Hook controversy are pointed out. Right shows the tweets by journalists and campaign staffs.

not necessarily correlate with semantic coherence[2]. Take the election topic *Guns* for example. Many of the popular hashtags (e.g., #IamWithHer, #StopGunViolence) are not very informative of the possible exogenous events that may influence the users to tweet.

#DemDebate	#IamWithHer	#DemTownHall
#StopGunViolence	#gunviolence	#HillaryinNH

Table 5.2: List of Popular Hashtags in the topic Guns for voices Clinton follow.

Suppose we want to understand how different users respond to gun-related news events. Let's map the bursts along the *semantic coherence* and *causal influence* dimensions as in the previous sections. One can observe bursts associated with campaign staffs are more responsive in terms of semantically coherence and causal influence. Suppose one wants to examine the recent controversy in Sanders' remark regarding Sandy Hook court hearing [49]. Mapping the bursts shows that the journalists are not very actively engaged in the sub-topic. More evidence can be found in the actual tweets: the campaign staffs and supporters are openly skeptical on Bernie Sanders, whereas the journalists are not as interested in this news event. Therefore,

[2]In fact, hashtags are often sparse in these topic-filtered series, and several topics show zero to negative Pearson correlation.

by using a more complete information of the semantics as well as potential causal interactions, *BurstMapper* delivers a more expressive tool to perform data exploration on tweets.

Chapter 6

Conclusion

This thesis proposes the *BurstMapper* system to identify and characterize tweet bursts to analyze the roles exogenous events play in driving tweets. The system is evaluated on an annotated corpus, with each component showing reasonably good results. *BurstMapper* is further demonstrated to be able to characterize bursts in ways helping interested individuals to understand how different users respond differently in the face of exogenous events. The system is shown to provide informative results even in larger-scale datasets. We believe the proposed methodology can bring forth a new way for analyzing tweet communications.

More General Use of *BurstMapper*

While the evaluations are done specifically on election-related tweets, expanding *BurstMapper* to different tweet topics should be a straightforward change of topic classification model. In fact, the model architecture in [51] can be directly applied to data in the new topic domain with appropriate training data.

Limitation and Future Work

The proposed system is not without limitations. While the system seems promising to apply to different topic domains, it is not clear if the residual-based burst detection would work very well when facing an even larger-scale tweet series. It seems that concepts such as Dirichlet Hawkes Process [14] or models that consider latent state spaces with content generation should be able to tackle data with higher complexity data. A natural next step is to try to consider using a point process generative model to improve the currently proposed burst detection algorithm.

It is also not yet clear to what extent can Tweet2Vec model be generalized. A more detailed experiment is warranted to see how do finer-grained tweet expressions map to the latent semantic embedding space. Having a more detailed understanding of the model can be helpful to enhance the current framework to deal with semantic similarity with respect to forms of expressions (e.g., language formality, allusion, sarcasm).

Lastly, while Hawkes Process can acts as a proxy to measure temporal-causality between pairs of voices, it does not strictly follow the Granger causality by blocking out potential confounding factors. Therefore, certain degree of noise is inevitable. Also, in terms of computation, naïve Hawkes Process doesn't scale with many voices, as the learning/simulation worst-case complexities are $O(|V|^2 T_{train})/O(|V|^2 T_{sim})$. It might not scale to millions of voices without special treatment. Knowing that social network connection is sparse, it is natural to explore how to approximate the parameter learning by exploiting the sparse network structure.

Appendix A

Algorithms

Algorithm 1 Burst Detection Algorithm

procedure FINDBURSTS($\tau_{S^{\cdot}_z}$,f_c,f_c, w_{len}, b_{max})
    Given per-topic time series $\tau_{S^{\cdot}_z}$, frequency threshold f_c, residual cutoff threshold f_c, window length w_{len}, max burst length b_{max}
    Let $S' = S^{\cdot}_z$
    Set Butterworth filter $G(\omega|f_c)$
    Compute the transformed signal $\hat{\tau_{S'}}$ using frequency response $G(\omega|f_c)$
    Compute residual $r_{S'} = \tau_{S'} - \hat{\tau_{S'}}$
    Set $r_{S'}[r_{S'} < r_c] = 0$
    return GreedyGroup($r_{S'}$)
end procedure
function GREEDYGROUP($r_{S'}$, w_{len}, b_{max}) bursts = [], start position pos = 0, tmp = []
    while pos+w_{len} < $|r_{S'}|$ **do** ▷ Slide window through $r_{S'}$
        if $\sum_{pos:pos+w_{len}} r_{S'} > 0$ and $|tmp| < b_{max}$ **then**
            $j = max(\{k|\forall k \in [0, |r_{S'}|]\ s.t.\ r_{S'}(t) > 0\})$
            tmp = tmp + [pos, pos+j+1]
            pos += j+1
        else
            bursts = bursts + tmp
            tmp = []
        end if
        pos = pos + 1
    end while
    return bursts
end function

Algorithm 2 Query from Burst

```
procedure QueryBurst(q, S, β, c)    ▷ Query q, series S, weight β, threshold c
    Seeds = {e_k ∈ B, q ∈ e_k^C} for all bursts B
    Find cluster assignments C by ConvexExemplarClustering [33] with Exemplars=Seeds, metric M(e_1, e_2): eq (3.5)
    for all events e_i in detected bursts do
        if M(e_i, argmin_s M(e_i, s)) < c, s ∈ Exemplars then
            output += e_i
        end if
    end for
    return output
end procedure
```

Appendix B

Election Topic Descriptions

According to [51], the election topic definitions are determined by discussing with experts in the U.S. presidential election. We adopt the list of topics, and provide a more descriptive explanations based on the annotated corpus.

Topic	Description
Racial Issues	Contents about racial and gender equality.
Justice	Discussions concerning criminal justice, court ruling, or Supreme Court related.
Income Inequality	Discussion on income inequality, poverty issues, and low-income initiatives.
Budget / Taxation	Government budget, federal tax, corporate tax, social security, IRS, tax break, etc.
Campaign Finance	Source for candidate's campaign finance- self-funded, super PACs, special interest groups, big corporation funding.
Abortion	Planned Parenthood and general abortion discussion.
LGBT Issues	Specific discussions on LGBT rights.
Terrorism	Events/Policies regarding terrorism; issues that the user is suggesting to be related to terrorism (e.g., radical Islamic).
Veterans	Veteran benefits, events about veterans.
Economy	Boost economy, small business, Wall Street reform, middle class.
Drugs	Discussion on drug use and drug policy.

Table B.1: Explanation for each of the 22 election topics.

Topic	Description
Guns	Second Amendment debate, gun violence, gun rights, gun safety.
Jobs and Employment	Job creation, unemployment, worker rights, workplace equality, minimum wage.
Ethics	Discussion about integrity of the candidate (e.g., Benghazi hearing).
Foreign Policy/Security	Foreign policy, military intervention, national security, commander-in-chief.
Immigration	Immigration policy, refugees, Mexican immigrants, Mexican border/wall.
Surveillance/ Privacy	Patriot's Act, privacy control by government.
Health Care	Health care related discussions.
Environment/Energy	EPA, Pollution, climate change, alternative energy, petroleum/oil/gas industry.
Financial Regulation	Wall Street reform, banks reform.
Trade	Discussion on trade agreements, protectionism, trading with China, etc.
Education	Discussion on affordable university (student loans), education system reform etc.

Table B.2: Explanation for each of the 22 election topics (cont'd).

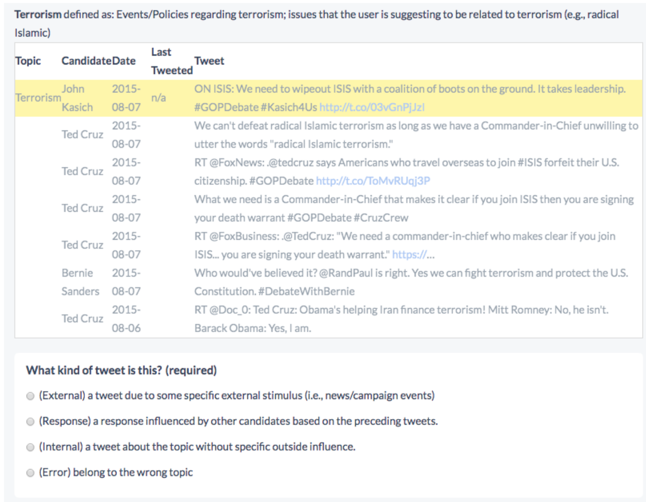
Appendix C

Data Annotation Details

We described in more detail the data labeling process. Recall that we task the workers to label each tweet into any of the following categories: (**External**) a tweet due to a recent newsworthy event (e.g., breaking news, major campaign events); (**Response**) a tweet content-wise influenced by, or triggered to respond to the preceding tweets by others; (**Internal**) a tweet not due to specific outside influence (news events or tweets by others); (**Error**) content not relevant to the given topic label.

Each labeling instance is given the particular tweet, with content, author, topic category, tweet time, and when author previously tweeted. Lastly, tweets that precede in time are also presented.

One difficulty is that it is hard to reliably know if one tweet is causally related to another tweet. Therefore, we specify a set of rules for the workers to follow. Additional suggestions are given to reduce false positives. See instruction template to see for more details the actual cases for each label category. Figure B-1 shows an actual example, notice some styling was done to highlight the importance of the tweet he/she is tasked to label. B-2 to B-3 shows the question template.



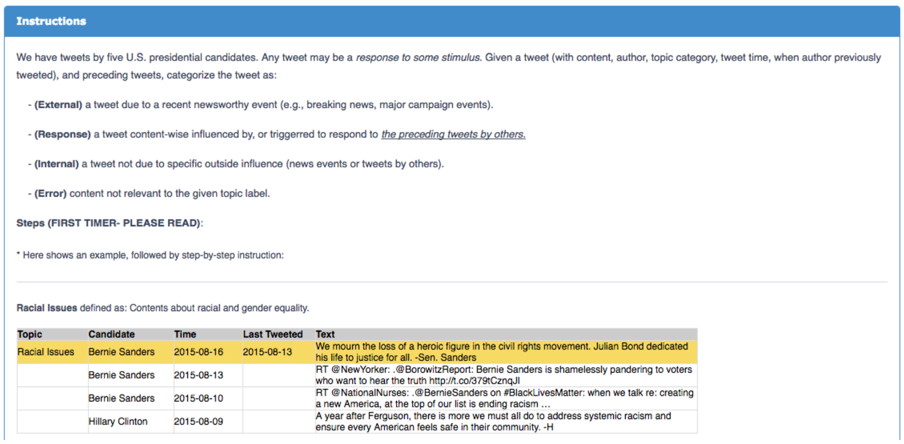
Terrorism defined as: Events/Policies regarding terrorism; issues that the user is suggesting to be related to terrorism (e.g., radical Islamic)

Topic	Candidate	Date	Last Tweeted	Tweet
Terrorism	John Kasich	2015-08-07	n/a	ON ISIS: We need to wipeout ISIS with a coalition of boots on the ground. It takes leadership. #GOPDebate #Kasich4Us http://t.co/03vGnPjJzI
	Ted Cruz	2015-08-07		We can't defeat radical Islamic terrorism as long as we have a Commander-in-Chief unwilling to utter the words "radical Islamic terrorism."
	Ted Cruz	2015-08-07		RT @FoxNews: .@tedcruz says Americans who travel overseas to join #ISIS forfeit their U.S. citizenship. #GOPDebate http://t.co/ToMvRUqj3P
	Ted Cruz	2015-08-07		What we need is a Commander-in-Chief that makes it clear if you join ISIS then you are signing your death warrant #GOPDebate #CruzCrew
	Ted Cruz	2015-08-07		RT @FoxBusiness: .@TedCruz: "We need a commander-in-chief who makes clear if you join ISIS... you are signing your death warrant." https://...
	Bernie Sanders	2015-08-07		Who would've believed it? @RandPaul is right. Yes we can fight terrorism and protect the U.S. Constitution. #DebateWithBernie
	Ted Cruz	2015-08-06		RT @Doc_0: Ted Cruz: Obama's helping Iran finance terrorism! Mitt Romney: No, he isn't. Barack Obama: Yes, I am.

What kind of tweet is this? (required)

- (External) a tweet due to some specific external stimulus (i.e., news/campaign events)
- (Response) a response influenced by other candidates based on the preceding tweets.
- (Internal) a tweet about the topic without specific outside influence.
- (Error) belong to the wrong topic

Figure C-1: A sample question that each worker is asked to label.



Instructions

We have tweets by five U.S. presidential candidates. Any tweet may be a *response to some stimulus.* Given a tweet (with content, author, topic category, tweet time, when author previously tweeted), and preceding tweets, categorize the tweet as:

- **(External)** a tweet due to a recent newsworthy event (e.g., breaking news, major campaign events).
- **(Response)** a tweet content-wise influenced by, or triggerred to respond to *the preceding tweets by others.*
- **(Internal)** a tweet not due to specific outside influence (news events or tweets by others).
- **(Error)** content not relevant to the given topic label.

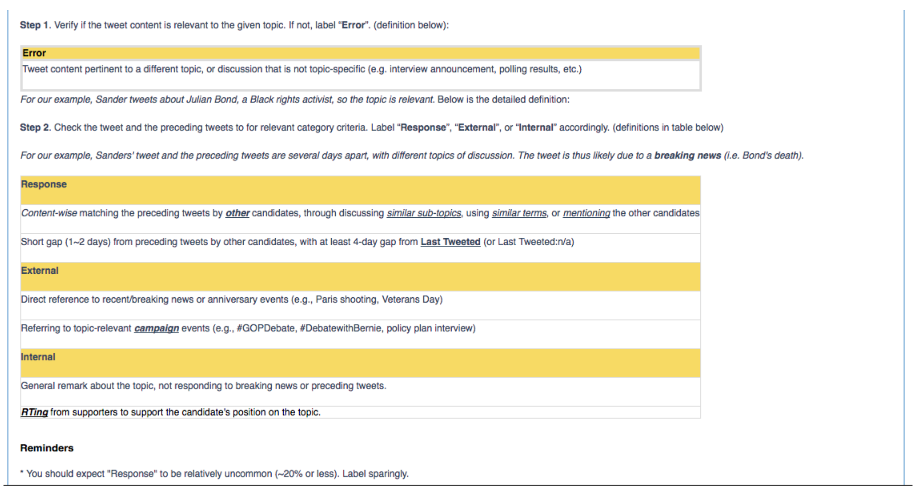
Steps (FIRST TIMER- PLEASE READ):

* Here shows an example, followed by step-by-step instruction:

Racial Issues defined as: Contents about racial and gender equality.

Topic	Candidate	Time	Last Tweeted	Text
Racial Issues	Bernie Sanders	2015-08-16	2015-08-13	We mourn the loss of a heroic figure in the civil rights movement. Julian Bond dedicated his life to justice for all. -Sen. Sanders
	Bernie Sanders	2015-08-13		RT @NewYorker: .@BorowitzReport: Bernie Sanders is shamelessly pandering to voters who want to hear the truth http://t.co/379tCznqJl
	Bernie Sanders	2015-08-10		RT @NationalNurses: .@BernieSanders on #BlackLivesMatter: when we talk re: creating a new America, at the top of our list is ending racism …
	Hillary Clinton	2015-08-09		A year after Ferguson, there is more we must all do to address systemic racism and ensure every American feels safe in their community. -H

Figure C-2: Question template.



Step 1. Verify if the tweet content is relevant to the given topic. If not, label "**Error**". (definition below):

Error
Tweet content pertinent to a different topic, or discussion that is not topic-specific (e.g. interview announcement, polling results, etc.)

For our example, Sander tweets about Julian Bond, a Black rights activist, so the topic is relevant. Below is the detailed definition:

Step 2. Check the tweet and the preceding tweets to for relevant category criteria. Label "**Response**", "**External**", or "**Internal**" accordingly. (definitions in table below)

*For our example, Sanders' tweet and the preceding tweets are several days apart, with different topics of discussion. The tweet is thus likely due to a **breaking news** (i.e. Bond's death).*

Response
Content-wise matching the preceding tweets by ***other*** candidates, through discussing *similar sub-topics*, using *similar terms*, or *mentioning* the other candidates
Short gap (1~2 days) from preceding tweets by other candidates, with at least 4-day gap from **Last Tweeted** (or Last Tweeted:n/a)
External
Direct reference to recent/breaking news or anniversary events (e.g., Paris shooting, Veterans Day)
Referring to topic-relevant ***campaign*** events (e.g., #GOPDebate, #DebatewithBernie, policy plan interview)
Internal
General remark about the topic, not responding to breaking news or preceding tweets.
RTing from supporters to support the candidate's position on the topic.

Reminders

* You should expect "Response" to be relatively uncommon (~20% or less). Label sparingly.

Figure C-3: Question template.

Appendix D

Regression Analysis Result

This chapter shows the regression analysis result discussed in Chapter 4. For the experiment, a logistic regression is performed on a set of features computed for each burst. The endogenous variable is a binary-value variable, saying whether the burst is responsive or not. For feature coding, see Table C-1.

Feature	Code	Feature	Code
$\overline{e_i^{cau}}$	C1	$Avg(\boldsymbol{m})$	S1
$\overline{\phi}(e_i^t - e_j^t)$	C2	$Ent(\boldsymbol{m})$	S2
CI	C3	SC	S3
$\overline{\phi}(e_i^t - e_j^t) * Avg(\boldsymbol{m})$	C2S1	$\overline{\phi}(e_i^t - e_j^t) * Ent(\boldsymbol{m})$	C2S2
$CI * Avg(\boldsymbol{m})$	C3S1	$CI * Ent(\boldsymbol{m})$	C3S2
Feature	Code	Feature	Code
$\overline{e_i^{cau}} * Avg(\boldsymbol{m})$	C1S1	$\overline{\phi}(e_i^t - e_j^t) * SC$	C2S3
$\overline{e_i^{cau}} * Ent(\boldsymbol{m})$	C1S2	$\overline{e_i^{cau}} * SC$	C1S3
$CI * SC$	C3S3		

Table D.1: Feature coding for the features used in regression analysis.

The test is performed on each topic-filtered tweet time series. Table C-2 shows the significance test for the coefficients for each feature. From the table, one can see that the interaction of features turn out to be significant in most topics, with individual semantic and temporal-causal features being significant at times.

Topic	C1	C2	C3	S1	S2	S3	S1C1	S1C2
Racial Issues				*	*	**	**	*
Justice		*		*	**	**	**	**
Income Inequality				*	*	*		**
Budget/Taxation						*		**
Campaign Finance					**	**		**
Abortion					**			**
LGBT Issues								
Terrorism			***		**	*		**
Veterans	*	*	*		**	**	**	*
Economy					*			
Drugs								
Guns					*		*	
Jobs/Employment						**	***	
Ethics								
Foreign Policy/Nat'l Security		*						**
Immigration								
Surveillance								
Health Care	*			*	*	*		
Environment					*	*	***	
Financial Regulation					***	***	***	***
Trade				***	***		***	***
Education								*

Table D.2: Feature significance over 22 election topics.

Topic	S1C3	S2C1	S2C2	S2C3	S3C1	S3C2	S3C3
Racial Issues	*	***	***	***	***	***	***
Justice	**	***	**	**	***	**	**
Income Inequality	***	*	**	***	**	**	***
Budget/Taxation		*	*	*	*	**	**
Campaign Finance	*	***	**	**	***	**	**
Abortion			**	*	**	*	*
LGBT Issues							
Terrorism	***	**	*	***	**	**	***
Veterans	*	***	**	***	***	**	***
Economy	*	**	*	*	*	*	**
Drugs							
Guns			**				*
Jobs/Employment		*	*	*	*	***	***
Ethics							
Foreign Policy/Security	*		**	*		**	*
Immigration	*	**	**	**	*	*	*
Surveillance							
Health Care		**	**	**	**	**	**
Environment		**	**	**	**	*	**
Financial Regulation	***			***		***	
Trade	***		***	***		***	***
Education		*	*	**	*	**	**

Table D.3: Feature significance over 22 election topics (cont').

Bibliography

[1] Icek Ajzen and Martin Fishbein. *Understanding attitudes and predicting social behaviour*. Prentice-Hall, 1980.

[2] Eytan Bakshy, Jake M. Hofman, Winter A. Mason, and Duncan J. Watts. Everyone's an influencer: Quantifying influence on twitter. In *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*, WSDM '11, pages 65–74, New York, NY, USA, 2011. ACM.

[3] Albert-Laszlo Barabasi. The origin of bursts and heavy tails in human dynamics. *Nature*, 435(7039):207–211, 2005.

[4] Charles Blundell, Katherine Heller, and Jeffery Beck. Modelling reciprocating relationships with hawkes processes. In *Advances in Neural Information Processing Systems*, 2012.

[5] Johan Bollen, Huina Mao, and Alberto Pepe. Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena, 2011.

[6] Yi Chang, Xuanhui Wang, Qiaozhu Mei, and Yan Liu. Towards twitter context summarization with user influence models. In *Proceedings of the Sixth ACM International Conference on Web Search and Data Mining*, WSDM '13, pages 527–536, New York, NY, USA, 2013. ACM.

[7] Paul Chilton. *Analyzing Political Discourse Theory and Practice*. Routledge, Taylor and Francis Group, London and New York, 2004.

[8] Anqi Cui, Min Zhang, Yiqun Liu, Shaoping Ma, and Kuo Zhang. Discover breaking events with popular hashtags in twitter. In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*, CIKM '12, pages 1794–1798, New York, NY, USA, 2012. ACM.

[9] Cristian Danescu-Niculescu-Mizil, Lillian Lee, Bo Pang, and Jon Kleinberg. Echoes of power: Language effects and power differences in social interaction. In *Proceedings of WWW*, pages 699–708, 2012.

[10] Dmitry Davidov, Oren Tsur, and Ari Rappoport. Enhanced sentiment learning using twitter hashtags and smileys. In *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*, COLING '10, pages 241–249, Stroudsburg, PA, USA, 2010. Association for Computational Linguistics.

[11] Daniel Dennett and Deb Roy. How digital transparency became a force of nature. *Scientific American*, 312, 2015.

[12] Nicholas A. Diakopoulos and David A. Shamma. Characterizing debate performance via aggregated twitter sentiment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 1195–1198, New York, NY, USA, 2010. ACM.

[13] Qiming Diao, Jing Jiang, Feida Zhu, and Ee-Peng Lim. Finding bursty topics from microblogs. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers - Volume 1*, ACL '12, pages 536–544, Stroudsburg, PA, USA, 2012. Association for Computational Linguistics.

[14] Nan Du, Mehrdad Farajtabar, Amr Ahmed, Alexander J. Smola, and Le Song. Dirichlet-hawkes processes with applications to clustering continuous-time document streams. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '15, pages 219–228, New York, NY, USA, 2015. ACM.

[15] Nan Du, Le Song, Manuel Gomez-Rodriguez, and Hongyuan Zha. Scalable influence estimation in continuous-time diffusion networks. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 26*, pages 3147–3155. Curran Associates, Inc., 2013.

[16] Nan Du, Le Song, Hyenkyun Woo, and Zha Hongyuan. Uncover topic-sensitive information diffusion networks. In *Proceedings of the 16th International Conference on Artificial Intelligence and Statistics*, 2013.

[17] Cyclonebiskit et al. November 2015 paris attacks. [Online; accessed 29-July-2016].

[18] L. Festinger. *A Theory of Cognitive Dissonance.* California: Stanford University Press., 1957.

[19] S.M. Gerrish and Blei D.M. A language-based approach to measuring scholarly impact. In *Proceedings of the 26th International Conference on Machine Learning*, 2010.

[20] Manuel Gomez-Rodriguez, David Balduzzi, and Bernhard Scholkopf. Uncovering the temporal dynamics of diffusion networks. In *Proc. of the 28th International Conference on Machine Learning*, JMLR:WCP, Bellevue, 2011. ACM.

[21] Manuel Gomez-Rodriguez, Jure Leskovec, David Balduzzi, and Bernhard Scholkopf. Uncovering the structure and temporal dynamics of information propagation. *Network Science*, 2014.

[22] Manuel Gomez Rodriguez, Jure Leskovec, and Andreas Krause. Inferring networks of diffusion and influence. In *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '10, pages 1019–1028, New York, NY, USA, 2010. ACM.

[23] C. W. J. Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37:424–438, 1969.

[24] Fangjian Guo, Charles Blundell, Hanna Wallch, and Katherine Heller. The bayesian echo chamber: Modeling social influence via linguistic accommodation. In *Proceedings of the 18th International Conference on Artificial Intelligence and Statistics*, 2015.

[25] A.G. Hawkes. Point spectra of some self-exciting and mutually exciting point processes. *Journal of the Royal Statistical Society: Series B (Methodology)*, 58:83–90, 1971.

[26] César A Hidalgo R. Conditions for the emergence of scaling in the inter-event time of uncorrelated and seasonal systems. *Physica A: Statistical Mechanics and its Applications*, 369(2):877–883, 2006.

[27] Carl Hovland and Muzafer Sherif. *Social judgment: Assimilation and contrast effects in communication and attitude change*. Westport: Greenwood, 1980.

[28] Long Jiang, Mo Yu, Ming Zhou, Xiaohua Liu, and Tiejun Zhao. Target-dependent twitter sentiment classification. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1*, HLT '11, pages 151–160, Stroudsburg, PA, USA, 2011. Association for Computational Linguistics.

[29] Marion Just, Ann Crigler, Dean Alger, and Timothy Cook. *Crosstalk: Citizens, Candidates, and the Media in a Presidential Campaign*. University of Chicago Press, 1996.

[30] Jon Kleinberg. Bursty and hierarchical structure in streams. In *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '02, pages 91–101, New York, NY, USA, 2002. ACM.

[31] Efthymios Kouloumpis, Theresa Wilson, and Johanna Moore. Twitter sentiment analysis: The good the bad and the omg!, 2011.

[32] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon. What is twitter, a social network or a news media? In *Proceedings of the 19th International Conference on World Wide Web*, WWW '10, pages 591–600, New York, NY, USA, 2010. ACM.

[33] Danial Lashkari and Polina Golland. Convex clustering with exemplar-based models. *Advances in neural information processing systems*, 20, 2007.

[34] Chenliang Li, Aixin Sun, and Anwitaman Datta. Twevent: Segment-based event detection from tweets. In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*, CIKM '12, pages 155–164, New York, NY, USA, 2012. ACM.

[35] Scott Linderman and Ryan Adams. Discovering latent network structure in point process data. In *Proceedings of the 31st International Conference on Machine Learning*, JMLR:WCP, volume 32, Beijing, 2014.

[36] Michael Mathioudakis and Nick Koudas. Twittermonitor: Trend detection over the twitter stream. In *Proceedings of the 2010 ACM SIGMOD International Conference on Management of Data*, SIGMOD '10, pages 1155–1158, New York, NY, USA, 2010. ACM.

[37] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. Distributed representations of words and phrases and their compositionality. *CoRR*, abs/1310.4546, 2013.

[38] Seth A. Myers, Chenguang Zhu, and Jure Leskovec. Information diffusion and external influence in networks. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '12, pages 33–41, New York, NY, USA, 2012. ACM.

[39] Olutobi Owoputi, Chris Dyer, Kevin Gimpel, Nathan Schneider, and Noah A. Smith. Improved part-of-speech tagging for online conversational text with word clusters. In *In Proceedings of NAACL*, 2013.

[40] Alexander Pak and Patrick Paroubek. Twitter as a corpus for sentiment analysis and opinion mining. In *LREc*, volume 10, pages 1320–1326, 2010.

[41] Michael Paul and Mark Dredze. You are what you tweet: Analyzing twitter for public health, 2011.

[42] P.O. Perry and P.J. Wolfe. Point process modeling for directed interaction networks. *Journal of the Royal Statistical Society: Series B (Methodology)*, 75:821–849, 2013.

[43] Richard E Petty and John T Cacioppo. Source factors and the elaboration likelihood model of persuasion. *Advances in Consumer Research*, 11:668.

[44] Hassan Saif, Yulan He, and Harith Alani. Alleviating data sparsity for twitter sentiment analysis. CEUR Workshop Proceedings (CEUR-WS. org), 2012.

[45] Stefan Stieglitz and Linh Dang-Xuan. Emotions and information diffusion in social media?sentiment of microblogs and sharing behavior. *Journal of Management Information Systems*, 29(4):217–248, 2013.

[46] Andranik Tumasjan, Timm Sprenger, Philipp Sandner, and Isabell Welpe. Predicting elections with twitter: What 140 characters reveal about political sentiment. In *Proc. of the 4th International AAAI Conference on Web and Social Media*, 2010.

[47] Isabel Valera and Manuel Gomez-Rodriguez. Modeling adoption and usage of competing products. In *Proceedings of the IEEE International Conference on Data Mining*, 2015.

[48] Alexei Vázquez, Joao Gama Oliveira, Zoltán Dezsö, Kwang-Il Goh, Imre Kondor, and Albert-László Barabási. Modeling bursts and heavy tails in human dynamics. *Physical Review E*, 73(3):036127, 2006.

[49] Elise Viebeck. ?shame on you': Bernie sanders draws sharp rebuke for sandy hook lawsuit comments, 2015. [Online; accessed 29-July-2016].

[50] Prashanth Vijayaraghavan, Ivan Sysoev, Soroush Vosoughi, and Roy Roy. Deepstance at semeval-2016 task 6: Detecting stance in tweets using character and word-level cnns. In *Proc. of the 10th International Workshop on Semantic Evaluation, NAACL*, 2016.

[51] Prashanth Vijayaraghavan, Soroush Vosoughi, and Roy Roy. Automatic detection and categorization of election-related tweets. In *Proc. of the 10th International AAAI Conference on Web and Social Media*, 2016.

[52] Soroush Vosoughi. *Automatic Detection and Verification of Rumors on Twitter*. PhD thesis, Massachusetts Institute of Technology, 2015.

[53] Soroush Vosoughi and Deb Roy. Tweet acts: A speech act classifier for twitter. *CoRR*, abs/1605.05156, 2016.

[54] Soroush Vosoughi, Prashanth Vijayaraghavan, and Deb Roy. Tweet2vec: Learning tweet embeddings using character-level cnn-lstm encoder-decoder. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '16, pages 1041–1044, New York, NY, USA, 2016. ACM.

[55] Xiaolong Wang, Furu Wei, Xiaohua Liu, Ming Zhou, and Ming Zhang. Topic sentiment analysis in twitter: A graph-based hashtag sentiment classification approach. In *Proceedings of the 20th ACM International Conference on Information and Knowledge Management*, CIKM '11, pages 1031–1040, New York, NY, USA, 2011. ACM.

[56] Jianshu Weng. Event detection in twitter. 2011.

[57] Jianshu Weng, Ee-Peng Lim, Jing Jiang, and Qi He. Twitterrank: Finding topic-sensitive influential twitterers. In *Proceedings of the Third ACM International Conference on Web Search and Data Mining*, WSDM '10, pages 261–270, New York, NY, USA, 2010. ACM.

[58] Shaomei Wu, Jake M. Hofman, Winter A. Mason, and Duncan J. Watts. Who says what to whom on twitter. In *Proceedings of the 20th International Conference on World Wide Web*, WWW '11, pages 705–714, New York, NY, USA, 2011. ACM.

[59] Shuang-hong Yang and Hongyuan Zha. Mixture of mutually exciting processes for viral diffusion. In Sanjoy Dasgupta and David Mcallester, editors, *Proceedings of the 30th International Conference on Machine Learning (ICML-13)*, volume 28, pages 1–9. JMLR Workshop and Conference Proceedings, May 2013.

[60] Ali Zarezade, Ali Khodadadi, Mehrdad Farajtabar, Hamid R. Rabiee, Le Song, and Hongyuan Zha. Correlated cascades: Compete or cooperate. In *WSDM*, 2016.

[61] Ke Zhou, Hongyuan Zha, and Le Song. Learning social infectivity in sparse low-rank networks using multi-dimensional hawkes processes. In *Proceedings of the International Conference on Artificial Intelligence and Statistics*, volume 16, 2013.