Temporally Aligning Clusters of Social Media Reaction to Speech Events

Brian Amanatullah, Greg Barish, Matthew Michelson and Steve Minton
InferLink Corporation
2361 Rosecrans Avenue, Suite 348
El Segundo, California 90245

Abstract— *This paper describes our approach to analyzing social media response to speech events, such as the President's State of the Union address. Our approach clusters Tweets by topic and in time, and this paper specifically focuses on analyzing the temporal dynamics of the topically clustered Tweets. We classify topical clusters as either being temporally clustered around a specific period of time or not (e.g., referring to the whole speech). For instance, a topical cluster related to a specific line in the speech is likely clustered in time as well. In contrast, a cluster about the speakers suit could be composed of Tweets written anytime during the speech. We demonstrate the approach on a number of real world data sets, and show that an analysis of these temporal dynamics can lead to a structuring of the social media responses that can support deeper analysis than just topical clustering alone.*

1. Introduction

Recently, the analysis of social media and politics has become an interest in the data mining community. However, much of this previous work focuses on the sentiment of the social media reaction (e.g., [1], [2], [3]). In this work we take a different approach, focusing instead on finding the points of the speech that generate the most reaction and the topics that correspond to these spikes in activity. Specifically, this paper presents an approach for analyzing social media responses to live speech events, such as the US President's State of the Union address. In particular, we focus on Twitter users' responses to the speech by analyzing their Tweets during the duration of the speech. We designed the approach to allow analysts, social scientists, and policy researchers to measure public reaction to various talking points (and visuals) in a speech. One can measure which points generate the most reaction, including those that may be surprising to the speaker and his/her staff. Further, marketers, journalists, political junkies and the general public can also use the system (and its related, public facing website) to better understand the speech's effect on different groups of people, their opinions, etc.

Our method clusters Tweets both topically and temporally. By topically, we mean that Tweets are clustered because they refer to the same subject. For instance, one cluster of Tweets discusses the "President's tie," while another cluster focuses on "universal pre-school education." If a cluster refers specifically to a line (or lines) in the speech, we call this cluster "referent." This is in contrast to clusters whose topic is not directly related to the speech itself, such as reaction to what the viewers are seeing on the screen at that moment (e.g., the President's tie). We call these clusters "non-referent." By temporally clustered, we mean clusters whose Tweets exhibit "burstiness" within a short interval of time. For instance, if many of the Tweets in a cluster happen to fall within a short time-window of one another, we call them "temporal." This usually happens when there is reaction to something specific in the speech's time-line (such as a particular line in the speech or non-referent event). In contrast are "non-temporal" clusters, which are clusters that refer to the speech generally. For instance, a cluster where people are simply noting that they are watching a speech seem to occur at various points throughout the speech, and do not exhibit the similar "bursty" behavior. By definition, we note that non-temporal clusters are also non-referent.

By aligning Tweets both topically and temporally, we can then overlay the topic-clusters onto the time-line of the speech itself to get a two-dimensional representation of responses to the speech. Then, for a given time in the speech, we can see the major topics of discussion at that particular time. For instance, if the cluster is Temporal/Referent, then we know it refers to that part of the speech at that time, and therefore that part sparked social reaction. If a cluster is Temporal/Non-Referent, then something in the broadcast outside of the speech itself, such as what is on-screen at that time, prompted reaction. Finally, we can exclude Non-Temporal clusters from the timeline analysis, since they would provide broad color (possibly), but not provide much deeper temporal analysis. The key then is to classify the clusters.

This paper focuses on the particular classification task of determining whether a cluster of Tweets is temporal or non-temporal.[1] Intuitively, our algorithm works by assuming that temporal clusters will exhibit bursty behavior, while non-temporal clusters will not. For instance, if during a speech the speaker says something controversial or resonating, then we assume lots of Tweets on that specific topic will be generated, and therefore should be grouped in time (e.g, classified as temporal). The key to the classifier, therefore, is to define "burstiness" and then measure it. We define such a

[1] While there exists a body of past research on determining topical-clustering of text, for example, k-means clustering [4], Latent Dirichlet Allocation (LDA) [5], Hierarchical Agglomorative Clustering (HAC) , (e.g., [6]) among others, that is not the main focus of this paper.

measurement and show that it can form a classifier for defining temporal and non-temporal clusters, and that indeed it can even find non-temporal clusters that are also non-referent.

Table 1 makes the clustering clear with a few examples taken from the 2012 State of the Union speech. The first cluster in the table is non-temporal. It reflects a number of Tweets from users watching the State of the Union speech, and Tweeting that they are doing so. The Tweets occur at various times throughout the speech, and do not, as a whole group, refer to a specific time period in the speech. The second cluster reflects Tweets about a specific topic that occurs at a specific time period in the speech. That is, the cluster is both temporal and referent (it refers to a line about the auto industry). The final example cluster in the table is temporal but non-referent. Instead, the time period reflected by the cluster is a point during the speech when the camera showed President Obama giving Gabrielle Giffords a hug.

Our overall approach for analyzing the social media response to a speech is given in Figure 1. While we focus on the temporal classification aspect in this paper, briefly, the full system works as follows. During a speech, the system sources, collects and then cleans a set of Tweets. The Tweets are then clustered by topic, and also broken down by cohort, where each sub-cohort represents a group of users responding (for instance each cluster is further sub-divided into cluster members provided by men and those provided by women). Finally, the Tweet clusters are aligned temporally. This temporal alignment is the temporal classification we will focus on in this paper. Once the data is processed, we display the results on a webpage[2] where users can explore and analyze the results.

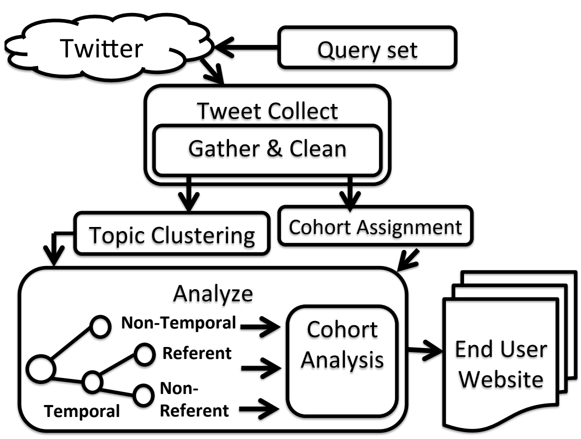

Fig. 1
THE 2SR ARCHITECTURE

As stated above, we implemented the approach in a public-facing website, where users can select various speeches and analyze the output themselves. One of the important aspects of our user interface is the alignment of the clusters to specific times and parts of the speech. This enables users to dig into the social media reaction at specific times of the speech. The Website is shown in Figure 2. The top of the figure shows a timeline with vertical bars. The bars represent clusters linked temporally to that part of the speech. The size of the bar reflects the relative volume of Tweets such that a taller bar reflects more Tweets (larger cluster) than a smaller bar. At a glance, this allows users to zoom into the sentences of the speech that generated the most reaction. In the figure, a user clicked on one of the taller bars (highlighted in grey) and the site automatically scrolled to the line in a transcript of the speech reflected by the time of the cluster. On the bottom right, example Tweets from the cluster scroll by the user to examine.

The rest of this paper is organized as follows. Section 2 describes our approach in detail, Section 3 presents our results and discussion, and Section 4 contains our conclusions and future directions for this research.

2. Temporal Clustering of Social Media Reaction

In this section, we detail our approach to temporal clustering. As we stated above, we assume that Tweets are already clustered by topic, and that forms the input to our process. Before we give the algorithm a more formal treatment, we discuss the intuition behind the approach. Intuitively, a set of responses to a particular item in the speech will cluster around (though after) that item in the speech. That is, assume we have a given cluster, and assume that we choose to define bursty behavior as having most of the tweets in a cluster fall within a 4-minute window. Then, we can analyze a time-line of all of the Tweets that belong to the cluster, and if some defined proportion (such as simply the majority) of them falls within the rolling window, we would say the cluster exhibits bursty behavior. This situation is shown in Figure 3.

In the figure, we see the Tweets that define the cluster, $\{T_0, \ldots, T_{10}\}$. Each Tweet is then aligned in time, and the timeline is shown along the bottom of the figure. Since our given rolling window size is four minutes in the example, the figure also shows example rolling windows as horizontal bars across the top, along with the range they represent. The first cluster, 8:00 to 8:04, contains 8 of the 10 cluster members and we therefore classify this cluster as temporal. If none of the rolling windows for a given cluster satisfy the constant that the proportion of Tweets contained in the window is above the proportion threshold, then we classify the cluster as non-temporal.

Therefore, we define define bursty clusters as those that contain Tweets that flare up in a specific time range, and more or less remain local to that time range. In some sense, one can think of members of a bursty set as having a tight temporal relationship. We define a tight temporal relationship as one where a majority (or a proportion above a threshold) of the

[2]http://www.socialreactiongroup.com

Table 1

DIFFERENT CLUSTER CLASSIFICATIONS

Non-temporal cluster
RT @tjholmes: Unless ur watching CSPAN, u might not know President of the United States is delivering State of the Union address n 50 mi...
What did you think of President Obama's State of the Union address?
#NowWatching President @BarackObama's "State Of The Union Address"...& you should be too!
Referent & Temporal cluster: reflects specific part(s) of the speech
Reacting to the line: "And tonight, the American auto industry is back"
#SOTU - Cites General Motors return to #1."Tonight, the American auto industry is back" -applause.
The American auto industry is back. #manufacturing #SOTU
The American auto industry is back! (to making cars that last 30k miles) #SOTU #honda4life
RT @DPerkDT: Republicans initially opposed the bailout of the American auto industry, now applaud jobs created. #SOTU
"The American auto industry is back. What is happening in Detroit can happen in other industries." -POTUS #SOTU
Pres. Obama says General Motors is "back on top as the world's number one automaker....the American auto industry is back." #SOTU #Detroit
Non-Referent & Temporal: Does not reflect specific part(s) of the speech
#whchat..just saw the President embrace Gabby. I'm grabbing my tissue?snif...So good to see her.
RT @jbarro: Obama hugging Gabrielle Giffords very sweetly.
#SOTU it was nice seeing President Obama hug Gabrielle Gifford.
Obama hugs Rep. Giffords. #SOTU #WHTweetup
President gave a huge hug to Gabby Giffords. How sweet. #sotu
Obama and Gabby Giffords hugging...#tearjerker

Fig. 2

THE SOCIALREACTIONGROUP WEBSITE

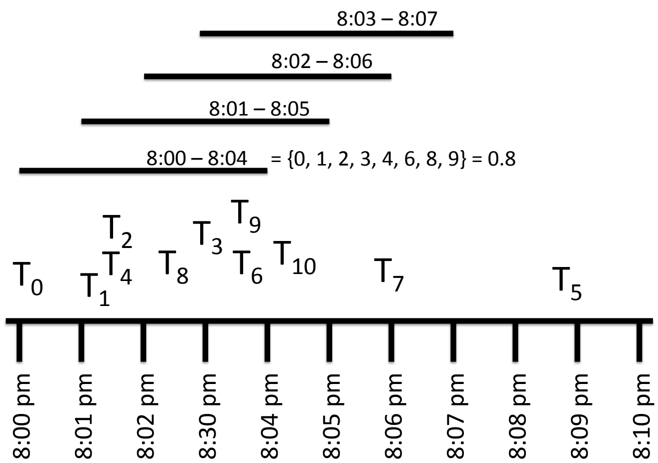

Fig. 3
A CLUSTER OF TWEETS ON A TIME LINE

cluster members (Tweets) occur within a short time period. This implies two parameters: the window size of the short time period, and the proportion of members in the cluster that must occur within that window. This definition allows us to apply a discrete measurement of tightness. Namely, the proportion of cluster members that fall within this window. This forms the basis of our classifier, if the proportion is above a threshold, we define the cluster as temporal, and otherwise not.

More formally, given a rolling window size, W, and a set of topical clusters, $\{C_0, \ldots, C_n\}$, each of which, C_i, is defined by the Tweets $\{T_0, \ldots, T_n\}$ contained within it (noting that each T_j also has a time-stamp), we define the cluster's temporal tightness as:

$$\textsc{TemporalTightness}(C_i) = \frac{\|T_j^{time} \in W\|}{\|C_i\|}$$

Where T_j^{time} is the time differential from the earliest time-stamped Tweet in the cluster, C_i to this Tweet, T_j. Then, given a proportion threshold P_{thresh}, we define a cluster C_i as being temporally clustered if $\textsc{TemporalTightness}(C_i) \geq P_{thresh}$.

The whole algorithm is given in Table 2.

Table 2
CLASSIFYING TEMPORAL CLUSTERS

$\forall C_i \in \{C_0, \ldots, C_n\}$
If $\textsc{TemporalTightness}(C_i) \geq P_{thresh}$
$C_i \leftarrow$ Temporal Cluster
Else
$C_i \leftarrow$ Non-Temporal Cluster

The key parameter, is therefore the window size (the proportion can simply be a majority), and this has certain implications. For instance, if a window size is set to a ridiculously large value, then general comments about the whole speech will be assigned as temporally relevant, but the point-in-time they refer to would be the whole speech. If the window size is way too small (e.g., 20 seconds) this would imply that the social response is almost instantaneous. Understanding the properties of this parameter informs a reasonable choice of a few minutes. That is short enough to capture the dynamics of temporal tightness, but long enough to give social media users time to respond to the same point, even if they response time differs. As we show below in our experiments, five minutes is an adequate choice.

One advantage of our approach is that it can discover tightly temporal clusters that refer to something about the speech, but not in the text itself (e.g., Non-referent clusters). For example, in Table 1 the algorithm discovered that the cluster describing President Obama's hug was temporally clustered as well, reflecting that this action had a specific time-period, namely when it was on screen during the televised speech. That is, by examining the temporal dynamics separately from the topical clustering, the approach does not need to take the content of the cluster into account.

3. Experimental Results

Above we described our approach to classifying topical clusters as having a temporal relationship or not. Here we apply our approach to a number of real world speeches and analyze the resulting clusters. We find that indeed, we are able to discern topics that have both a topical relationship and a temporal relationship to the speech.

We analyzed four specific speeches: President Obama's State of the Union (SOTU) address in 2012, Benjamin Ne-

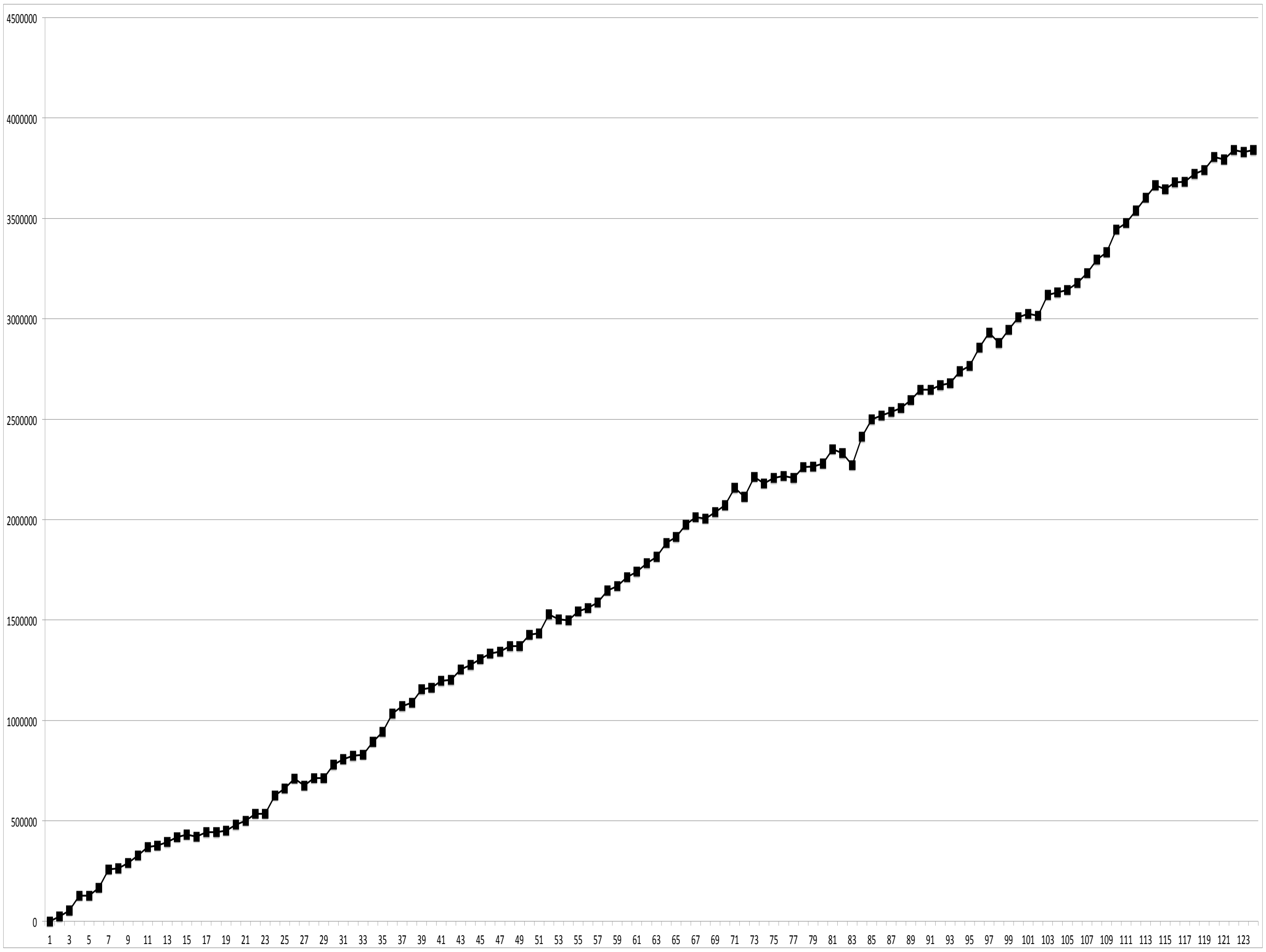

Fig. 4
SOTU 2012 TWEETS VS SPEECH TIME

Table 3
SPEECH DATA FOR ANALYSIS

Speech	Start Time (UTC)	Stop Time (UTC)	Tweets	# Clusters
SOTU 2013	2/12/2013 2:05:00	2/12/2013 3:20:00	145,913	182
Inaguration 2013	1/21/13 16:55	1/21/13 17:12	38,937	55
Netanyahu UN 2012	9/27/12 17:40	9/27/12 18:18	11,800	43
SOTU 2012	1/25/12 2:05	1/25/12 3:25	164,547	124

tanyahu's address to the UN General Assembly in 2012, President Obama's Inauguration address in 2013, and President Obama's State of the Union (SOTU) address in 2013. Table 3 describes the data, showing the start and end time of the speech (in UTC), along with the number of Tweets analyzed during that time period, for the given speech.

As shown in the table, we were able to discover a number of temporal clusters for each speech. We also found a number of clusters that were temporal but non-referent. As we stated above, these may correspond with events shown on the screen during the televised speech. For instance, for the Netanyahu speech, the temporal/non-referent cluster occurred after he showed a picture of a cartoon bomb during the speech. For each speech we found 34, 4, 1, and 15 temporal/non-referent clusters, respectively.

Figure 4 validates our premise that our window size and threshold are reasonable. In the figure, we see a plot of the time of the speech-line of the 2012 SOTU speech and the time associated with the topical cluster assigned to that speech line. The y-axis of the figure shows the number of milliseconds from the start of the speech, increasing up the axis (origin is the start of the speech). The x-axis is the minutes from the start of the speech of the cluster associated with that line of the speech (e.g., the cluster that represents that line in the speech), again, ordered from the start of the speech. For instance, when Obama says, "And tonight, the American auto industry is back." this is associated with a specific cluster. The figure shows the time of the first Tweet in this temporal cluster (x-axis), plotted against the timeline of the speech (y-axis). Therefore, if the topical clusters were well aligned in time with the speech, then the line should be relatively straight and sloping up and to the right (we assign time as monotonically increasing in seconds from the start of the speech). Indeed, we see this is the case in the figure, and there are very few dips downward, which would signify a mis-aligned cluster in time with the speech.

4. Conclusion

This paper describes an approach for clustering Social Media responses to speech events not just by topic, but also temporally. We present an approach for taking a topical cluster and deciding whether the members of that cluster refer to a specific point in time, or do not. The approach relies on dynamically setting a window of time, within which we consider the set of social media responses to be "temporally tight," and therefore clustered in time as well. Clusters whose members fall outside of this range are non-temporal and may refer a more general time-frame, such as the whole speech itself.

We validated our approach both empirically, and also through a public facing website, where users can analyze social media reaction to speeches themselves, in both the topical and temporal clustering dimensions. In the future we plan to investigate this topic further to try and more deeply understand the temporal dynamics of social media reaction.

References

[1] N. A. Diakopoulos and D. A. Shamma, "Characterizing debate performance via aggregated twitter sentiment," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '10. New York, NY, USA: ACM, 2010, pp. 1195–1198. [Online]. Available: http://doi.acm.org/10.1145/1753326.1753504

[2] A. Livne, M. P. Simmons, E. Adar, and L. A. Adamic, "The party is over here: Structure and content in the 2010 election," *Proc. of 5th ICWSM*, 2011.

[3] P. H. Calais Guerra, A. Veloso, W. Meira, Jr., and V. Almeida, "From bias to opinion: a transfer-learning approach to real-time sentiment analysis," in *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, ser. KDD '11. New York, NY, USA: ACM, 2011, pp. 150–158. [Online]. Available: http://doi.acm.org/10.1145/2020408.2020438

[4] J. MacQueen *et al.*, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, vol. 1, no. 281-297. California, USA, 1967, p. 14.

[5] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *the Journal of machine Learning research*, vol. 3, pp. 993–1022, 2003.

[6] F. Murtagh, "A survey of recent advances in hierarchical clustering algorithms," *The Computer Journal*, vol. 26, no. 4, pp. 354–359, 1983.