Using time series and natural language processing to identify viral moments in the 2016 U.S. Presidential Debate

Josephine Lukito[1], Prathusha K Sarma[2], Jordan Foley[3] and Aman Abhishek[4]

[1], [3], [4] School of Journalism and Mass Communication, UW-Madison
[2] Electical and Computer Engineering, UW-Madison
{jlukito,kameswarasar,jfoley5,aabhishek}@wisc.edu

Abstract

This paper proposes a method for identifying and studying viral moments or highlights during a political debate. Using a combined strategy of time series analysis and domain adapted word embeddings, this study provides an in-depth analysis of several key moments during the 2016 U.S. Presidential election. First, a time series outlier analysis is used to identify key moments during the debate. These moments had to result in a long-term shift in attention towards either Hillary Clinton or Donald Trump (i.e., a transient change outlier or an intervention, resulting in a permanent change in the time series). To assess whether these moments also resulted in a discursive shift, two corpora are produced for each potential viral moment (a pre-viral corpus and post-viral corpus). A domain adaptation layer learns weights to combine a generic and domain specific (DS) word embedding into a domain adapted (DA) embedding. Words are then classified using a generic encoder+classifier framework that relies on these word embeddings as inputs. Results suggest that both Clinton and Trump were able to induced discourse-shifting viral moments, though the former is much better at producing a topically-specific discursive shift.

1 Introduction

Though research across disciplines tends to analyze language cross-sectionally, or synchronically, we know that language use is temporally dependent. In other words, discourse about a subject can ebb and flow dynamically over time, peaking at salient moments or dropping when attention to that subject is low. This feature is especially noticeable on social media platforms during media storms. Here, "media storm" is defined as "an explosive increase in news coverage of a specific item (event or issue) constituting a substantial share of the total news agenda during a certain time" (Boydstun et al., 2014). Media storms can be unplanned, such as the coverage of a scandal (Walgrave et al., 2017), or planned, like presidential debates (Dayan and Katz, 1994).

Three components are important to Boydstun's definition: news coverage must be explosive, all-consuming ("constituting a substantial share" of media attention), and long-lasting. Presidential debates fulfill all three conditions of a media storm because they are explosive (attention to the debate explodes when it begins, large (a debate consume most media attention until it is over) and can be long-lasting (post-debate spin ensures that coverage of the debate lasts for longer than 24 hours) (Fridkin et al., 2008).

Though news media are obviously important to media storms, the modern hybrid media ecology ensures that what appears in news media is likely to also appear on social media platforms. After all, if media storms concentrate news media attention towards one news events, social media attention likely becomes concentrated as well. This is particularly true on Twitter, which journalists rely on for their professional work (McGregor and Molyneux, 2016). As a result, Twitter has become an essential platform for the sharing of news information, and for public discussion of media storm events.

The purpose of this study is to explore the temporal and linguistic dynamics of viral moments during the first 2016 U.S. Presidential Debate, between Donald J. Trump and Hillary R. Clinton. In a media storm, viral moments constitute important peaks of attentionthe most discussed moments in an event that already garners significant media

scrutiny. Our study relies on an inductive, three-step approach to identifying and studying these viral moments during a debate.

2 Related Work on Debates and Viral Moment

Previous studies of political debates using computational methods have largely focused on candidates' rhetoric and topic shifts. For example a candidate who is able to shift topics during a debate is perceived to have greater relative power, which increases their ranks compared to other candidates (Prabhakaran et al., 2014). A handful of studies have also analyzed social media in tandem with debates, acknowledging the increasing role of second-screens. In these studies, social media is used to asses debate performance in real time (Diakopoulos and Shamma, 2010; Pond, 2016). We deviate somewhat from these analysis by focusing specifically on key viral moments, rather than overall sentiment or topic shifts.

"Going viral" constitutes a process of quickly becoming popular on one or multiple (digital) platforms (Hong et al., 2011). Many things can "go viral", including hashtags (Bastos et al., 2013) and people (Pancer and Poole, 2016). A "viral moment", therefore, is a moment in time where a person, place, or thing "goes viral". Because we are focused on debates, we are primarily interested in viral moments induced by candidates in the debate, and not (for example) by social media discourse occurring independently from the debate.

In a debate, candidates are likely to try inducing viral moments to garner and sustain attention during a highly publicized discursive spar. They may do so by making salient comments or gestures that received widespread attention for their deviance. These moments are important to candidates, as they can garner attention and "produce memorable and highly referenced moments" (Shah et al., 2016). Previous studies have found that these moments tend to be gaffs (misspoken statements) and zingers (insults) (Freelon and Karpf, 2015).

One unique feature of debates in the digital age is the popularity of "second-screening", whereby audiences watching something on one screen (the "first" screen) interact with a "second" screen, sometimes to enhance their overall viewing experience (Schirra et al., 2014). The most common example of this is live-tweeting when one watches television. Given the televised nature of political debates, many viewers enjoy live-tweeting and discussing the debate in real-time, often on a platform like Twitter. This creates a unique media consumption experience that did not exist in the 1960's or 1970's (Chadwick et al., 2017).

Within a media storm, particularly salient moments in time come to represent the media storm as a whole. These "highlights"or viral momentsare important to post-election spin. For example, citizens who did not watch the full debate may still seek out highlights to get the "main gist" of the event. This is not unlike news coverage of other planned media events, which tends to focus on that events key moments (Fridkin et al., 2007).

Furthermore, because of ability for the viewing audienceeveryday citizens, journalists, influencers, and celebrities alike, aka the "viewertariat" (Anstead and OLoughlin, 2011)to engage with discourse, the audience becomes especially meaningful to the production of viral moments. No longer are news media the gatekeepers of determining what is or is not an important debate moment. Rather, this can now be gaged through social media interaction and commentary.

Previous studies of these moments have largely been inductive (Shah et al., 2016; Freelon and Karpf, 2015). In other words, these moments are typically identified through an assessment made by the researchers, with varying levels of specificity regarding what constitutes or does not constitute a viral moment. There are a handful of exceptions; for example, one study looks at what content media will highlight from a debate (Tan et al., 2018). They find that many feature sets (including emotion, contrast, personal pronouns, and superlatives) increase the likelihood of a statement being highlighted. However, this analysis does not consider the role of social media.

This study contributes to ongoing information communication research by proposing a more quantitatively-driven, context-free strategy that can be applied to study highlights across many planned events. More specifically, we posit that viral moments during media storms (like this debate) are likely to have both temporal qualities and discursive properties that makes such a moment unique relative to the rest of a media storm.

3 Methodology

This study relies on a combination of time series models and Natural Language Processing (NLP)

strategies to explore a set of possible viral moments induced by the debates candidates (see Figure 1).

Two primary data are used: the first is a corpus of English-language tweets about Hillary Clinton or Donald Trump at the time of the debate. This corpus was purchased through Gnip, a social media API aggregation company that is owned by Twitter. Through Gnip, Twitter sells historic "firehose" data (a census of tweets using a keyword search within a given time frame); the cost of this data varies with the number of tweets in the search. We purchased all the tweets within the debates 90-minute window using the following keyword search: ((clinton OR hillary) - (trump OR donald)) OR ((trump OR donald) - (clinton OR hillary)).

For the time series analysis, counts of tweets referencing either Clinton or Trump were aggregated at the 30-second-level. This resulted in two time series: one with the number of tweets about Clinton every 30-seconds, and one with the number of tweets about Trump every 30-seconds. Because Twitter activity was high during this time, there were no gaps in the time seriesall equally-spaced time points had at least one tweet.

To perform the NLP analysis, we took all the tweets posted two minutes before each temporal outlier and constructed a corpus. We then took all the tweets posted two minutes after each temporal outlier to construct a second corpus. Each potential viral moment identified by the time series analysis, therefore, would have a corpora-pair (one corpus representing pre-viral tweets, and one representing post-viral tweets).

The second dataset is a C-SPAN video recording of the debate itself (this is analyzed in tandem with a transcript that has been manually time-stamped for every 10-second increment). C-SPAN is a public affairs programming network which televises and records U.S. political events, including U.S. Presidential debates. C-SPAN footage is made publicly available on their website. We analyzed the video in the "split-screen" format, wherein one camera is pointed to each candidate. The videos of both candidates are then shown side-by-side.

3.1 Time Series Analysis

This viral-moment identification process takes place in three steps. We begin with an analysis of temporal outliers in the two time series: one for

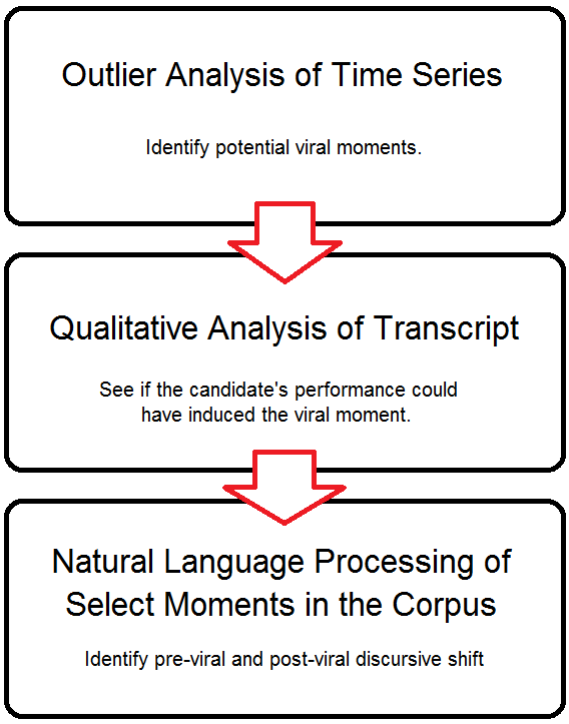

Figure 1: Three-step Method for identifying viral moments

mentions of Clinton and one for Trump. This is an inductive process. To identify outliers we estimate temporal outliers using an ARIMAX model. This is an extension of the popular univariate ARIMA model, which stands for a AutoRegressive, Integrated, Moving Average model (Brockwell et al., 2002). An ARIMA model attempts to identify and model the temporal data-generating process of a time series. In other words, to what degree (and how) is data at time "T" explained by its own prior values at time $t - 1$ or earlier $(t - n)$? It does so by looking at three possible dynamics, an autoregressive (AR) component, an integrated (I) component, and a moving average (MA) component. The ARIMAX model is an extension of the ARIMA that allows for control variables. Each of our models included one control: a dichotomous variable indicating the speaker. For the Clinton time series, the speaker was coded as "1" if Clinton was speaking, and "0" if she was not. For the Trump time series, the speaker was coded as "1" if Trump was speaking.

The R package $<$ tsoutliers $>$ estimates an ARIMAX model to identify three types of temporal outliers: a "pulse function", resulting in a quickly appearing or disappearing spike; a "transient change", where the series spikes quickly, but

the effect dissipates slowly; and an "intervention outlier", representing "a shock in the innovations of the model" (López-de Lacalle, 2016). More colloquially, a pulse results in a short-term change, a transient change reflects an immediate change that slowly disappears over time, and an intervention indicates a fundamental shift or change in the time series.

Positive outliers indicate a fast increase in attention towards a candidate. Negative outliers indicate a fast decrease in attention towards a candidate. Because this study posits that viral moments result in more attention (not less), we focus only on the positive outliers.

3.2 Analyzing Debate Discourse

The positive outliers moments identified through the time series analysis are then studied further. As the time series analysis relies entirely on count data, an outlier analysis cannot tell us why there would be a spike in attention. To analyze these moments further, we study the debate content around the time of the social media time series outlier, using the C-SPAN video and debate transcript, focusing on the speaker's rhetoric, both in terms of content and performance, as well as the opponent's non-verbal presentation). This is a necessary process to weed out temporal outliers triggered by things unrelated to the event (e.g., a celebrity's tweet going viral, unrelated to the debate in real-time).

In addition to this, we also explore the debate content around the time of the social media time series outlier.This is an important feature, as this study focuses on viral moments induced by candidate discourse during the debate. We use this qualitative analysis to identify key terms in the debate for which there is likely to be a discourse shift.

3.3 Natural Language Processing of Discursive Shift

To confirm that the debate-induced temporal outliers also induces a discursive shift, we apply a NLP strategy that identifies key words that have changed in their discursive use between two corpora. There exist several embedding algorithms that produce highly optimized and efficient embeddings for words in an n-dimensional vector space. Typically, such algorithms are trained on large-sized generic bodies of text (e.g., Wikipedia), as larger datasets are beneficial for capturing a wide range of the semantics of a word in its vector representation.

Recent work by (Sarma et al., 2018) demonstrates how one can perform 'Domain Adaptation' in word embeddings for small-sized data sets, by shifting the space of generic word embeddings. In their work (Sarma et al., 2018), two sets of word embeddings are obtained for a single vocabulary of words. One set of embeddings, called 'generic' embeddings are obtained from off-the-shelf solvers like word2vec (Mikolov et al., 2013), GloVe (Pennington et al., 2014) etc that are trained on a generic corpus such a Wikipedia. A second set of 'domain specific' (DS) word embeddings are obtained by either i)re-training algorithms like word2vec/GloVe on a target data domain or ii) use LSA (Deerwester et al., 1990) based embedding approach if the target domain is small in size. In the LSA approach, a documents by words ($d \times N$) matrix of word counts is constructed. Then, a SVD step is performed followed by projecting the left singular vectors on to the k largest singular values to obtain word embeddings for the N words. Once, the generic and DS embeddings are a obtained, a new adapted subspace is learned for the two sets of embeddings using Kernel Canonical Correlation Analysis (KCCA). The objective of KCCA (Hotelling, 1936) is to obtain a non-linear subspace such that the statistical correlations between two sets of variables is maximized. Domain Adapted (DA) embeddings are obtained by learning a non-linear subspace between the generic and DS embeddings. In their work (Sarma et al., 2018), the authors demonstrate that DA embeddings perform particularly well on sentiment analysis tasks applied to small sized target domains.

In our work, we obtain DA embeddings for words in tweets posted two minutes before and two minutes after an time series-identified viral moment. First, we tokenize texts from tweets before and after the viral moment and construct two sets of vocabularies corresponding to tweets before and after the viral moment. We obtain DA embeddings for the two vocabularies using KCCA. Then, we look for words that are common in both vocabularies and extract their corresponding DA embeddings. Once we have these DA embeddings, we measure the semantic shift in words that occur both before and after the viral moment by measuring the L2 distance between the pre- and post- vector representations of a given word.

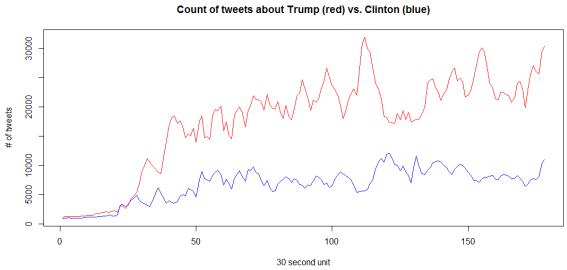

Figure 2: Time series of attention to Clinton (blue) and Trump (red)

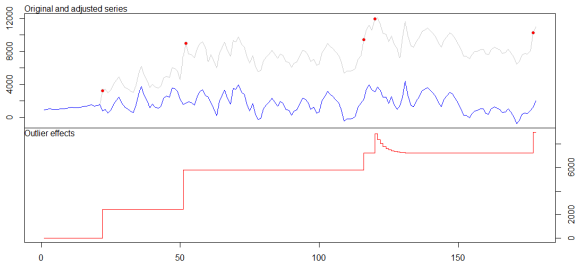

Figure 3: Outliers identified in the Clinton Time Series

4 Results

4.1 Time Series Analysis

The debate was 1 hours and 29 minutes long. When using a 30-second interval as the time-unit, this results in 178 points. Figure 2 displays two time seriesone showing the mentions of Clinton and one representing the mentions of Trump.

We begin our time series analysis with a test for unit roots. Unit roots are an indicator that a time series is non-stationary (i.e., that the time series' mean, variance and covariance vary over time). This is a problem for time series models that rely on stationary time series. For ARIMA models, a unit root also indicates that the series has at least one integrated component (the 'I' in 'ARIMA' will be 1 or greater). It is also possible that a time series could be fractionally integrated, which means its 'I' would be between 0 and 1.

Two tests are common for finding unit roots: a KPSS test and an ADF test (Culver and Papell, 1997). Both are available in the R package $<$ tseries $>$. These test confirm one another and show that both the Clinton and the Trump time series have one unit root. To ensure that these unit roots are indicative of full integration, and not of fractional integration, we calculated an estimation of the integrated component of each time series (Haslett and Raftery, 1989). In both instances, the time series were well over 0.7, suggesting that both series have full or near-full integration components.

To diagnose the data-generating properties of each candidates Twitter count, we build two univariate auto-regressive integrated moving-average (ARIMA) models. We use the R package $<$ forecast $>$ to test the fit of various ARIMA models on each time series, relying on the Bayesian Information Criterion to select the optimal model. This process yielded an optimal ARIMA model of (0,1,0) for Twitter attention to Trump (BIC = 3109.361) and an optimal ARIMA model of (0,1,1) for Twitter attention to Clinton (BIC = 2948.52).

Results of the outliers analysis identify several time series outliers. Because this study is only interested in positive spikes of attention, negative outliers are excluded from the subsequent analysis. For Clinton, there are four positive outliers (as a reminder: the time series is measured in 30-second intervals). The first occurs around 25:18 to 25:48 and is an intervention (coefficient = 2432.50, t = 5.37 , $p < 0.01$). The second occurs around 39:18 to 29:48 and is an intervention (coefficient = 3378.01, t = 9.09, $p < 0.01$). The third is between 1:12:18 to 1:12:48 and is another intervention (coefficient = 1048.00, t = 4.20, $p<0.01$). And finally, the fourth is between 1:14:48 to 1:15:18 and it is an intervention (coefficient = 1789.88, t = 3.11, $p < 0.01$).

For Trump, four outliers are also found. The first occurs between 42:18 and 42:48 and is a transient change outlier. The second happens between 46:18 and 46:48 and is a level shift. The third is between 1:09:18 to 1:09:48, and begins as a level-shift, but ends as a transient change. The fourth is a level shift that occurs between 1:22:48 to 1:23:18. Figure 2 displays Trumps positive outliers identified using this strategy. Figure 3 displays Clintons positive outliers identified using this strategy and Figure 4 displays Trump's positive outliers.

4.2 Analysis of Debate Discourse

To understand these outliers in more detail, we examine the candidates performative discourse at the seven aforementioned times. The first Clinton outlier occurs when she says, "Donald thinks that

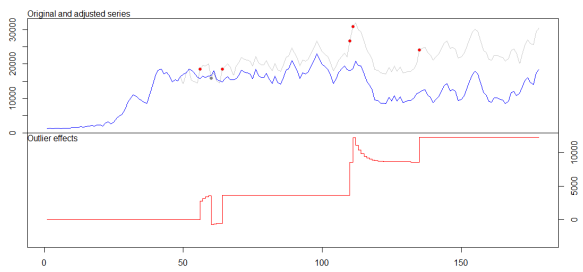

Figure 4: Outliers identified in the Trump Time Series

climate change is a hoax perpetrated by the Chinese" (0:25:37-0:25:49) in a response about climate change. The third outlier is a result of Clinton quoting First Lady Michelle Obama ("When they go low, we go high") during a question about Trump's Birther scandal (when Trump claimed that President Obama was not born in the United States). The fourth outlier happens in Clintons response to a question about cybersecurity. Curiously, the second outlier does not occur when Clinton is talking. Rather, mentions of Clinton increased when the moderator asked a question about Trumps tax returns:

> "Mr. Trump, we're talking about the burden that Americans have to pay, yet you have not released your tax returns. And the reason nominees have released their returns for decades is so that voters will know if their potential president owes money towho he owes it to and any business conflicts. Don't Americans have a right to know if there are any conflicts of interest?"

Because this event seemed to be induced by the moderators question, and not from either of the candidates, it was removed for subsequent analysis (this leaves us with 7 candidate-induced, temporally-identified viral moments). However, we posit that it remains a significant moment in the event worth noting.

The first outlier identified in the Trump attention time series occurred during the end of Trumps remarks regarding his taxes, and as Clinton begins her response. Following a back and forth that includes remarks about Clintons email scandal, the second outlier occurs when Trump uses the word "braggadocious" regarding his income: "I have a tremendous income. And the reason I say that is not in a braggadocios way" (0:46:28-0:46:33). He continues with a lengthy critique of Clinton, her foreign policy decisions, and the perceived economic consequences (this includes describing the United States as a third-world country). The third outlier occurs when Trump answers a question about the Birther scandal: "I figured you'd ask the question tonight, of course. But nobody was caring much about it. But I was the one that got him to produce the birth certificate. And I think I did a good job" (1:09:28-1:09:36). The fourth outlier occurs when Trump attributes the formation of Iraq to Clinton: "Well, President Obama and Secretary Clinton created a vacuum the way they got out of Iraq [. . .] once they got in, the way they got out was a disaster. And ISIS was formed" (1:22:08-1:22:22).

4.3 Natural Language Processing

Owing to space constraints, we only discuss three of the six potential viral moments. These are: Trump's use of the word "braggadocious" regarding his income and competency (Trump Viral Moment 2), Clinton's statement that Trump "thinks that climate change is a hoax" (Clinton Viral Moment 1), and Clinton quoting Michelle Obama. However, we present the NLP results of all seven viral moments identified through the time series analysis in our Appendix.

To look at the discursive shift prior to and after the temporally-identified viral moments, we subset our full corpus of tweets about Trump or Clinton during the first debate into three corpora-pairs. For viral each moment, there were two corpora: one from tweets in the pre-viral moment, and one for tweets in the post-viral moment; this produced fourteen corpora. Tweets from these were tokenized, and unique vocabularies were constructed using the two minute data from before (the pre-) and after (the post-) the viral moment. Final vocabularies were constructed by retaining words that appear at least two times across all the tweets from the pre- and post- viral moment. We then took the intersection of the two vocabularies and to identify words that occurred often among the tweets from before and after the viral moment. Previous studies have shown that the social media effects of a candidate's rhetoric tend to last no longer than two minutes (Bucy et al., 2019).

Words are ranked as 'most different' in use by measuring the l2 distance between the vector em-

bedding for a given word from the pre vocabulary and the corresponding embedding for the same word in the post vocabulary. Word embeddings for words in the pre and post vocabularies are obtained via the Kernel CCA projection method described in (Sarma et al., 2018). First domain specific word embeddings for both 'pre' and 'post' event vocabularies are constructed using LSA. Then, for words in common to both vocabularies a max-correlation subspace is constructed using KCCA. Projections of both sets of embeddings in this subspace are then compared to measure 'word-shift', i.e the l2 distance between the two projections of the same word in the KCCA derived vector subspace.

Among the three viral moments analyzed, many of the words that had the largest l2 distance in the pre-viral moment and the post-viral moment were words employed directly by a candidate during that time, or were relevant to the viral moment being discussed. This was especially true for Clinton's viral moments (regarding climate change and the Birther scandal). For example, in the first Clinton viral moment, words about climate change were among those with comparatively larger l2 distances, like green, climate, energy, and change.

Other discourse-specific words also had strong discursive shifts, such as hoax and China (words that originated directly from Clinton's statement). Similarly, the words with the largest l2-differences in Clinton's second viral moment were related to the Birther scandal, like Obama and Barack, or came directly from Clinton's statement, like response, high, go, and low. Clinton's statement also included a remark that Trump's accusation was "very hurtful" (the word "hurtful" also appeared to have a significant l2-difference in the pre-viral and post-viral corpora).

For Trump, the words with the largest l2-distance difference in the pre- and post-viral moment were related to the topic Trump was discussing. However neither the word braggadocious, nor the presumed root word brag, appeared on our list ("brag" appeared in this viral moment's pre-viral corpus, but "braggadocious" did not). Instead, the discourse shift on social media seemed to center around the foreign policy implications of his statement, which Trump pivoted to immediately following his statement about being a good businessman (this is what he was being 'braggadocious' about). Although Trump did not explicitly mention any countries in his statement, social media discourse focused on countries like China and Iraq (two countries that Trump mentions frequently elsewhere). However, the prevalence of more unrelated words suggests that this potential viral moment did not result in as strong of a discourse shift as Clinton's viral moments. More succinctly put: Trump's statement likely resulted in a spike of attention; however, this shock did not focus attention specifically on Trump's words the way shocks in attention to Clinton did.

n	Word	Δ L2 distance
1	**wrong**	102.62
2	**Iraq**	101.83
3	**should**	73.67
4	take	62.94
5	**China**	57.53
6	there	53.07
7	security	51.86
8	really	51.56
9	talking	45.79
10	**money**	45.33
11	wants	45.02
12	racial	44.28
13	only	41.38
14	plan	41.30
15	even	41.00
16	better	40.14
17	maybe	39.66
18	endorse	38.90
19	lost	36.91
20	**International**	36.15

Table 1: Words with the greatest l2 distance difference between the pre-viral and post-viral moment for Trump's Viral Moment 2

Figure 5, 6 and 7 provides words that changed the most between pre and post vocabularies and their corresponding differences in l2 distance for each viral moment studied.

5 Conclusion

Using a combination of time series techniques and natural language processing, this study finds several viral moments, or highlights, that have been induced by candidates during the first debate of the 2016 U.S. Presidential election. Though we find other spikes in attention towards either Trump or Clinton, they may be unrelated to the content of the debate itself (e.g., a celebrity watching the

n	Word	Δ L2 distance
1	blah	47.95
2	made	41.93
3	fuck	39.47
4	said	38.71
5	**green**	38.06
6	**climate**	37.57
7	**energy**	36.32
8	looks	36.28
9	again	35.19
10	**real**	33.80
11	because	33.71
12	sexist	33.68
13	**change**	33.54
14	**hoax**	33.38
15	important	32.93
16	please	32.21
17	bush	32.07
18	**china**	30.65
19	those	30.48
20	does	29.69

Table 2: Words with the greatest l2 distance difference between the pre-viral and post-viral moment for Clinton's Viral Moment 1

n	Word	Δ L2 distance
1	nothing	57.57
2	**response**	56.66
3	**high**	47.37
4	line	44.96
5	**go**	38.61
6	history	37.44
7	**they**	37.33
8	record	35.89
9	really	34.23
10	**hurtful**	33.45
11	vote	33.07
12	lester	31.75
13	**low**	31.67
14	went	31.64
15	**Obama**	31.26
16	**Barack**	31.12
17	**better**	30.77
18	there	30.75
19	watching	30.30
20	prepare	29.41

Table 3: Words with the greatest l2 distance difference between the pre-viral and post-viral moment for for Clinton's Viral Moment 2

debate that makes an unrelated comment that becomes popular). While those moments are important, we focus specifically on the stakeholders of the media storm, as they have the most to gain from viral moments.

Results of this analysis suggest that Trump and Clinton were both able to induce viral moments in the debate. Clinton's viral moments tended to produce a strong discursive shift that was directly related to her debate statement. This is indicated by the number of words that Clinton said which were also words that had the largest l2-difference in the pre-viral and post-viral corpus. By contrast, Trump's viral moment did not seem to have as prominent of a discursive shift. Nevertheless, Trump was seemingly able to focus attention on the topic of his interest. In the debate, Trump used his statement to pivot away from talk about his taxes to the present economic state of the country (in relation to other countries). On social media, attention also seemed to shift to his international critiques, reflecting Trump's ability to change public conversation at the time of the debate.

A more qualitative examination of the viral moments suggests that planned attacks or retorts, or those delivered in a more neutral tone, were not able to induce a viral event compared to unscripted words (e.g., braggadocious) and strong statements of condemnation (e.g., "Trump thinks climate change is a hoax perpetuated by the Chinese") were able to. We can note several instances where Trump or Clinton attempt to induce a viral moment, such as Clinton's use of Trumped-up, trickle-down economics and when Trump states: "Secretary Clinton doesn't want to use a few words like law and order." However, these statements did not induce temporally-evident viral moments, and likely did not result in a discursive shift.

Furthermore, even in instances where we may suspect the Twitter audience to focus on non-verbals or unique words (e.g., "braggadocious"), we find that the discourse shift occurs around words about policy issues, not words about the way a candidate behaves. This suggests that viral moments occur when the candidate makes a strong statement, often with critical audio or non-verbal cues but primarily if it relates to an already-salient political issue, such as a scandal (e.g., Birther scandal) or political decision (e.g., support for Iraq).

Combined, these results highlight the ability for debates to create politically salient viral moments, which carry symbolic meaning that lasts over the course of the debate, and beyond. As post-debate spin is important for audiences to understand how to interpret the debate ((Fridkin et al., 2007), (Shah et al., 2016)), we suspect that it is these viral moments that are subsequently identified as important highlights of the event. This is confirmed by news medias post-debate coverage of the top moment, which includes many of the viral moments identified here, though more of Trumps viral moments were listed at top moments by outlets like NBC, Fox News, and The New York Times. For example, both NBC and The New York Times highlighted Trumps remarks about Iraq, particularly when he attributed the creation of ISIS to Clinton and President Obama.

This debate may also be unique in its ability to induce viral moments. In particular, we found that the majority of the potential viral moments identified through the time series occurred during discussions of scandals, including the Birther Conspiracy, Trumps tax returns, and Clintons email and server scandal. Importantly, these scandals were not simple horserace stories. Rather, each candidate highlighted the other's scandals to emphasize their opponent's untrustworthiness or incompetence. The presence of so many scandals prior to, during, and after the election, likely fed into the ability for this debate to produce viral moments compared to other debates. Future work can explore this further by comparing insults that "go viral" in a debate to insults that do not.

5.1 Limitation

As with any study, there are several ways in which this work can be improved upon. In particular, the time series ends just as the debate ends. It is therefore difficult to interpret viral moments that occur early or late in the debate. Future studies on debates should therefore expand their time series into post-debate discourse so as to more accurately observe viral moments late in the debate.

While we highlight the importance of virality in spreading content, our study also does not empirically test the number of viral tweets that are produced as a result of viral moments. Though such analysis is beyond the scope of what this data can provide, future studies with more network information (i.e., retweets of a tweet over time) can also explore this phenomenon.

The addition of other control variables, such as interruptions, topic shifts and non-verbal features, would provide additional context that could help further explain why some insults or scandals induce viral events compared to others. For example, it is possible that insults induce virality when they are accompanied by aggressive gestures. Future studies can build upon this research by incorporating such data.

Nevertheless, we feel this study provides a substantive contribution to our understanding of debates as planned media storms that generate viral moments with potentially long-lasting implications in a political election. Rather than treating outliers as data to discard (for the purposes of better modeling), our research highlights the need to study why outliers appear the way they do, and to align these findings with our fields understanding of the media ecosystem.

Acknowledgments

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2016S1A3A2925033).

References

Nick Anstead and Ben OLoughlin. 2011. The emerging viewertariat and bbc question time: Television debate and real-time commenting online. *The international journal of press/politics* 16(4):440–462.

Marco Toledo Bastos, Rafael Luis Galdini Raimundo, and Rodrigo Travitzki. 2013. Gatekeeping twitter: message diffusion in political hashtags. *Media, Culture & Society* 35(2):260–270.

Amber E Boydstun, Anne Hardy, and Stefaan Walgrave. 2014. Two faces of media attention: Media storm versus non-storm coverage. *Political Communication* 31(4):509–531.

Peter J Brockwell, Richard A Davis, and Matthew V Calder. 2002. *Introduction to time series and forecasting*, volume 2. Springer.

Eric P Bucy, Jordan Foley, Josephine Lukito, Larisa Doroshenko, Dhavan V Shah, Jon CW Pevehouse, and Chris Wells. 2019. Performing populism: Trump's transgressive style and the dynamics of twitter response. In *2019 International Communication Association conference*. ICA, pages 1–28.

Andrew Chadwick, Ben OLoughlin, and Cristian Vaccari. 2017. Why people dual screen political de-

bates and why it matters for democratic engagement. *Journal of broadcasting & electronic media* 61(2):220–239.

Sarah E Culver and David H Papell. 1997. Is there a unit root in the inflation rate? evidence from sequential break and panel data models. *Journal of Applied Econometrics* 12(4):435–444.

Daniel Dayan and Elihu Katz. 1994. *Media events*. harvard University Press.

Scott Deerwester, Susan T Dumais, George W Furnas, Thomas K Landauer, and Richard Harshman. 1990. Indexing by latent semantic analysis. *Journal of the American society for information science* 41(6):391–407.

Nicholas A Diakopoulos and David A Shamma. 2010. Characterizing debate performance via aggregated twitter sentiment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, pages 1195–1198.

Deen Freelon and David Karpf. 2015. Of big birds and bayonets: Hybrid twitter interactivity in the 2012 presidential debates. *Information, Communication & Society* 18(4):390–406.

Kim L Fridkin, Patrick J Kenney, Sarah Allen Gershon, and Gina Serignese Woodall. 2008. Spinning debates: The impact of the news media's coverage of the final 2004 presidential debate. *The International Journal of Press/Politics* 13(1):29–51.

Kim L Fridkin, Patrick J Kenney, Sarah Allen Gershon, Karen Shafer, and Gina Serignese Woodall. 2007. Capturing the power of a campaign event: The 2004 presidential debate in tempe. *The Journal of Politics* 69(3):770–785.

John Haslett and Adrian E Raftery. 1989. Space-time modelling with long-memory dependence: Assessing ireland's wind power resource. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 38(1):1–21.

Liangjie Hong, Ovidiu Dan, and Brian D Davison. 2011. Predicting popular messages in twitter. In *Proceedings of the 20th international conference companion on World wide web*. ACM, pages 57–58.

Harold Hotelling. 1936. Relations between two sets of variates. *Biometrika* 28(3/4):321–377.

Javier López-de Lacalle. 2016. tsoutliers r package for detection of outliers in time series. *CRAN, R Package* .

Shannon C McGregor and Logan Molyneux. 2016. Twitters influence on news judgment: An experiment among journalists. *Journalism* page 1464884918802975.

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*. pages 3111–3119.

Ethan Pancer and Maxwell Poole. 2016. The popularity and virality of political social media: hashtags, mentions, and links predict likes and retweets of 2016 us presidential nominees tweets. *Social Influence* 11(4):259–270.

Jeffrey Pennington, Richard Socher, and Christopher Manning. 2014. Glove: Global vectors for word representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics, Doha, Qatar, pages 1532–1543. http://www.aclweb.org/anthology/D14-1162.

Philip Pond. 2016. Twitter time: A temporal analysis of tweet streams during televised political debate. *Television & New Media* 17(2):142–158.

Vinodkumar Prabhakaran, Ashima Arora, and Owen Rambow. 2014. Staying on topic: An indicator of power in political debates. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. pages 1481–1486.

Prathusha K Sarma, Yingyu Liang, and William A Sethares. 2018. Domain adapted word embeddings for improved sentiment classification. *arXiv preprint arXiv:1805.04576* .

Steven Schirra, Huan Sun, and Frank Bentley. 2014. Together alone: motivations for live-tweeting a television series. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. ACM, pages 2441–2450.

Dhavan V Shah, Alex Hanna, Erik P Bucy, David S Lassen, Jack Van Thomme, Kristen Bialik, JungHwan Yang, and Jon CW Pevehouse. 2016. Dual screening during presidential debates: Political nonverbals and the volume and valence of online expression. *American Behavioral Scientist* 60(14):1816–1843.

Chenhao Tan, Hao Peng, and Noah A Smith. 2018. You are no jack kennedy: On media selection of highlights from presidential debates. In *Proceedings of the 2018 World Wide Web Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, pages 945–954.

Stefaan Walgrave, Amber E Boydstun, Rens Vliegenthart, and Anne Hardy. 2017. The nonlinear effect of information on political attention: media storms and us congressional hearings. *Political Communication* 34(4):548–570.

6 Supplemental Material

n	Word	Δ L2 distance
1	nothing	61.44
2	high	41.52
3	well	38.51
4	back	37.39
5	election	33.89
6	time	32.60
7	they	32.59
8	senator	31.87
9	also	31.73
10	prepare	30.50
11	drop	28.67
12	watching	28.04
13	movement	27.98
14	birth	27.84
15	business	27.40
16	literal	26.99
17	them	26.87
18	hurtful	25.41
19	issue	25.00
20	there	24.94

Table 4: Words with the greatest l2 distance difference between the pre-viral and post-viral moment for Clinton's Viral Moment 3

n	Word	Δ L2 distance
1	paying	80.42
2	bubble	79.22
3	discurtir	75.57
4	smart	73.88
5	talk	71.58
6	Obama	69.45
7	federal	66.49
8	income	64.52
9	think	58.67
10	shit	57.66
11	rates	56.64
12	water	54.00
13	down	53.32
14	**ugly**	51.61
15	make	51.54
16	gold	51.42
17	need	51.41
18	interest	50.03
19	crook	48.75
20	tax	48.43

Table 5: Words with the greatest l2 distance difference between the pre-viral and post-viral moment for Trump's Viral Moment 1

n	Word	Δ L2 distance
1	healing	43.71
2	wasn't	36.56
3	ever	30.48
4	take	29.96
5	much	29.13
6	born	28.77
7	lying	28.09
8	even	27.75
9	here	26.63
10	profil	26.37
11	years	26.25
12	first	26.03
13	produced	25.80
14	very	24.95
15	chicago	24.31
16	politicians	24.23
17	white	23.78
18	must	23.57
19	communities	23.41
20	vote	23.38

Table 6: Words with the greatest l2 distance difference between the pre-viral and post-viral moment for Trump's Viral Moment 3

n	Word	Δ L2 distance
1	iraq	103.84
2	wrong	100.95
3	internet	94.33
4	hacker	86.05
5	take	70.37
6	china	59.39
7	really	49.59
8	america	47.04
9	they	45.31
10	does	45.03
11	security	43.57
12	year	43.08
13	racial	42.96
14	talking	42.82
15	wants	41.49
16	very	38.46
17	better	37.95
18	even	37.48
19	russia	35.39
20	jacking	34.81

Table 7: Words with the greatest l2 distance difference between the pre-viral and post-viral moment for Trump's Viral Moment 4