The 2016 US Presidential Election on Facebook:
An Exploratory Analysis of Sentiments

Saud Alashri
CIDSE
Arizona State University
Tempe, Arizona
salashri@asu.edu

Srinivasa Srivatsav Kandala
Decision Theater Network
Arizona State University
Tempe, Arizona
sskandal@asu.edu

Vikash Bajaj
Decision Theater Network
Arizona State University
Tempe, Arizona
vbajaj2@asu.edu

Emily Parriott
ISyE
Georgia Institute of Technology
Atlanta, GA
eparriott@gatech.edu

Yukika Awazu
IÉSEG School of Management
(LEM-CNRS 9221)
Paris, France
y.awazu@ieseg.fr

Kevin C. Desouza
School of Public Affairs
Arizona State University
Phoenix, Arizona
kev.desouza@gmail.com

Abstract

Social media platforms are valuable tools for political campaigns. In this study, we analyze a dataset representing over 22 thousand Facebook posts by candidates and over 48 million comments to understand the nature of online discourse. Specifically, we study the interaction between political candidates and the public during the 2016 presidential elections in the United States. We outline a novel method to classify commentators into four groups: strong supporters, supporters, dissenters, and strong dissenters. Comments by each group on policy-related topics are analyzed using sentiment analysis. Finally, we discuss avenues for future research to study the dynamics of social media platforms and political campaigns.

1. Introduction

Numerous studies show general citizens, voters and political representatives believe social media platforms to be legitimate spaces for active political participation [1]-[5]. Social media allows more people to participate in the political arena, involving more young people as well as those of different socio-economic statures [6]-[10]. Particularly, those users who are members of political parties use social media to gather information from fellow party members and have political discussions [11]. Users with party affiliation view social media sites as valid means of political engagement, expressing and discussing political views, and are positive about the effectiveness of participation through social media [11]. Politicians also use social media as a form of legitimate political engagement, particularly to market themselves and discuss issues with voters [12].

Facebook has already established itself as an important medium for political communication. For example, by 2011 in Norway, the use of Facebook was commonplace for political engagement [12]. Politicians use Facebook to broaden constituent accessibility and voter mobilization [12]. President Obama had higher activity levels and was portrayed more positively in Facebook groups than his counterpart, John McCain, during the 2008 presidential election [13]. Part of the success of President Barack Obama's campaign in the 2008 elections is thus credited to his campaign's ability to leverage Social Networking Sites (SNSs) [13], [4].

Although SNSs such as Facebook have become prominent means of communication for political campaigns and potentially play a crucial role in election results [14],[15], personal opinions and sentiments expressed by the general public on these platforms are often understudied [16]. In fact, data from the public can provide novel insights into online campaigning [17]. It can even be a summarizing indicator that could be used to predict the outcome of elections.

We studied the interaction between political candidates and the public during the 2016 presidential

HICSS

election in the United States. We analyzed content from the Facebook pages of the following candidates: Donald J. Trump, Hillary R. Clinton, Bernie Sanders, Ted Cruz, John R Kasich, Martin J. O'Malley, Marco A. Rubio, Ben S. Carson, Jeb E. Bush, Jim S. Gilmore, Chris J. Christie, Carly C. Fiorina, Rick J. Santorum, Rand H. Paul, and Mike D. Huckabee. We collected over 22 thousand Facebook posts, with over 48 million comments on those posts from the date the first candidate officially announced their intention to run for the U.S. presidency, (January 1st, 2015) until the time Donald J. Trump was assumed to be the winner of the 2016 election (end of day, November 8th, 2016). To understand the dynamics of interactions between the candidates and public, we divided data into four periods: 1) all candidates, 2) a selected group of Republicans and two Democrats, 3) one Republican and two Democrats, and 4) one candidate from each party. In this study, due to space limitations, we will highlight the results from time period II (a selected group of Republicans and two Democrats). The official Facebook posts of the candidates were analyzed using topic modeling, which allows us to identify core policy topics that each candidate discussed on their Facebook page. We put forth a novel method to classify commentators into four distinct groups: *Supporters*: expressed moderate positive opinions on a candidate's posts, *Strong Supporters*: expressed strong positive opinions, *Dissenters*: expressed moderate negative opinions, and *Strong Dissenters*: expressed strong negative opinions. We then study linguistic and psychological attributes of comments, to understand how different groups of commentators reacted to a given candidate's posts.

The rest of the paper is organized as follows. The next section provides a literature review. Section 3 describes our research methodology. The results and analysis of findings are presented in Section 4. Section 5 concludes the paper and outlines opportunities for future research.

2. Literature Review

Extant literature acknowledges the importance of SNSs for political engagements [1]-[5]. According to Enli and Skogerbø [12], SNSs enable the personalization of politics by broadening how candidates and the public are able to connect in new and different ways, which create different forms of political engagement.

Researchers are fascinated by the role that Facebook plays in political campaigns [4],[8],[18]. Williams and Gulati [19] suggests that Facebook is the leading platform for political campaigning. For example, in a study by Andersen and Medaglia [20],

over half of respondents (57%) in Denmark used Facebook to communicate with the candidate with whom they were Facebook friends, compared to 6% and 7% via mail and other chat mediums like Skype and MSN, respectively. Facebook is an attractive platform for digital natives who are developing their civic engagement skills and allows them to practice participating where they may not have been as inclined to participate before [21]. Studies show that students who show more political/civic participation on Facebook have higher rates of participation in the offline world [11], [22]. Those who are politically active in real life are those who are most likely to be politically active on Facebook [21].

Facebook is considered an attractive SNS for political campaigning because of its distinctive features. Candidates utilize it to campaign, interact with supporters, and to mobilize networks to advance their candidacy [3]. For example, Facebook offers personalized participation elements of sharing and net-working with "friends", which allows candidates to reach out to potential voters and connect with them [12], [18]. Politicians, in choosing "to be where their voters were," consider Facebook the most important medium for political campaigns compared to other platforms such as Twitter [12]. On the general citizens and voters side, Facebook is unique as it offers features such as the "newsfeed" and "wall" which allow users' thoughts and opinions to be displayed for their networks and thus increases more participations [23]. The ability to "comment" and "like" on Facebook directly impacts opinion-sharing and political engagement. "Likes" imply visibility and approval, agreement, or endorsement of the post in question and its content [24]. The larger the numbers of "likes", the more engagement Facebook users have with the post's content [24]. Comments, on the other hand, are vocalizations of user opinion and beliefs [24]. Debating and interacting through commenting and sharing on Facebook posts has become more appealing to younger users than traditional, time-consuming political engagement activities like canvassing and fund-raising [21],[24].

Lane and Dal Cin [25] found that sharing online on Facebook walls leads to engaging in offline helping behaviors (e.g. volunteering for an issue-related cause). These findings appear to deny that "slacktivism", or mere shallow gestures, is the result of political engagement online. Therefore, a person's commitment on Facebook is a reflection of his/her overall level of engagement, not a more flippant attitude than usual towards politics invoked by the medium. Let us take an example of the "Friends" feature available on Facebook. On the surface, this feature could provide candidates with a way "connect" with voters, to reach

out to them, and to hopefully mobilize them. According to a survey conducted by Andersen and Medaglia [20], when people who were listed as friends of the two main candidates for Prime Minister of Denmark in the 2007 election were surveyed, the majority of respondents (56%) said they friended the candidate in order to become better educated about the candidate's policy (45%) or be an "influence on their policy" (11%) versus the self-serving motives of "visibility on the Internet" (34%), and social prestige (19%). About one half of respondents were engaged online because they genuinely wanted to participate in the political process; therefore, one should take their behavior as indicative of their offline support or disapproval for a candidate. In this vein, Facebook appears to reproduce the traditional channels of supporting candidates, like party membership or connection to the campaign through employment or volunteering [20].

Although extant literature acknowledges the usefulness of Facebook's features for political engagements, little is known about how political engagements between candidates and the public are established. Sweetser and Lariscy [18] conducted content analysis of Facebook wall comments in the U.S. House and Senate races during the 2006 midterm election. They found that individuals who wrote on candidate walls consider themselves on friendly terms with the candidates, writing messages that are light hearted, supportive, and positive in tone. Candidates however, rarely responded to these messages. Gustafsson [11] studies three distinct user types, members of political parties or candidates, members of interest organizations, and those not politically active (non-members). He also finds that users do not appear to change their established political participation behaviors when exposed to political content and calls for political action on social media. Those users who were politically active before remain politically active while non-member users continued not to share their political views on social media or become politically involved. Because non- members did not change their engagement in politics, those users who are commenting on candidates' posts are invested in political engagement. Through their comments, they are showing support for the candidate they are reacting to, or are in support of another candidate and are attacking that candidate's rivals. These users are already involved and already feel strongly; almost all expressed a desire to vote.

The unique contribution of this paper can be found in 1) the study of interactions between political candidates and the public during the 2016 presidential election in the United States, 2) a novel method to classify commentators into four groups: *strong supporters, supporters, dissenters,* and *strong dissenters* and, 3) the analysis of linguistic and psychological attributes of these groups. Moreover, we hope to contribute to knowledge on computational political science both with our findings as well as our methodological approach (e.g. the use of topic modeling, commentators' classification, and the analysis of linguistic and psychological attributes).

3. Methodology

3.1. Dataset

Our dataset is comprised of 22,233 posts, and 48,991,502 comments spanning the entire period from when the first candidate announced their campaign (Jan 1st, 2015) until the time Donald J. Trump was assumed to be the winner of the 2016 election (end of day, November 8th, 2016). We created a python script that utilized Facebook's official Graph API[1] to collect posts and comments from every presidential candidate Facebook pages. We pre-processed our data by removing irrelevant posts and comments that were either not in English or did not have textual content. We also removed comments that were suspected to be from bots (i.e. VOTE FOR TRUMP, LET'S MAKE AMERICA GREAT AGAIN).

To understand the dynamics of topics and people interactions, we defined four periods as follows:

Table 1. Definition of Periods

Period	Dates	No. Posts	No. Comments	Candidates
I	01/01/2015 To 03/02/2016	16,696	14,732,578	(R) - Bush; Carson; Christie; Cruz; Fiorina; Gilmore; Huckabee; Kasich; Paul; Rubio; Santorum; Trump
II	03/03/2016 To 05/04/2016	1,901	5,233,383	(R) - Cruz; Kasich; Rubio; Trump
III	05/05/2016 To 07/25/2016	1,651	6,034,076	(R) – Trump (D) - Clinton; Sanders
IV	07/26/2016 To 11/08/2016	1,985	22,991,465	(R) - Trump (D) - Clinton

The first period includes all candidates who ran until the first vote (Iowa caucus)[2]. Subsequent periods were chosen to study how commentators changed their

---

[1] https://developers.facebook.com/docs/graph-api
[2] http://raviudeshi.com/16/02/2016-election-calendar

behavior as candidates dropped out and the field thinned.

3.2. Policy-related Topic Inference

Policy-related topics show political polarization among commentators. Thus, to discover the topics for each period, we utilized Latent Dirichlet Allocation (LDA) [26] on all candidates' posts. LDA fits our task as it is unsupervised probabilistic topic inference model that does not require a labeled dataset. This model assumes each document (post in this case) is a mixture of K latent topics, and each topic is a probability distribution over words. Therefore, a topic is the clustering of co-occurred words together. We manually examined the resultant topics, removed irrelevant topics, (such as those that call for attending events/debates) and then selected policy related topics. Next, the selected topics are labeled with their appropriate label (Healthcare, Taxes, etc.). Figure 1 shows an example of a post by Clinton on a policy-related topic (Climate Change) and below the post are examples of positive and negative comments identified by the algorithm, (pictures and names of commentators are covered). Next, for each candidate C and each period P, we select the top T topics that the candidate is actively posting, where the majority of the candidate's posts (more than 70%) are discussing these topics.

Figure 1. Example of a Post and its Comments

3.3. Identifying Four Categories of Commentators

Once we identified the topics and assigned candidates' posts to one of the topics, we examined the comments from commentators to identify: strong supporters, supporters, dissenters, and strong dissenters for a candidate. We measured both positive and negative sentiment scores for each comment using the Linguistic Inquiry and Word Count (LIWC) tool [27] where each score is measured from 0 to 100, the latter being the highest. For each candidate and each period, we measured the mean (μ) and standard deviation (σ) for positive μ_{pos}, σ_{pos} and negative μ_{neg}, σ_{neg} scores of the comments. Then we adjusted them based on the weight (W) of top topics the candidate discussed. For example, if Clinton mainly focuses on three topics during period 1 with 80% of her posts about these topics, then we multiply the mean (μ) and standard deviation (σ) by 0.8 to get a weighted mean and weighted standard deviation. Using the three-sigma rule [28], we defined the following categories of commentators who commented solely on one candidate:

- **Supporters**: a commentator (s) whose comments on a candidate (c) during a period (p) satisfies the following:

$$\frac{\sum_{s}(pos\ score)}{No.of\ Comments\ (s)} > \frac{\sum_{s}(neg\ score)}{No.of\ Comments\ (s)}$$

AND

$$\mu_{pos} \times W \leq \frac{\sum_{s}(pos\ score)}{No.of\ Comments\ (s)} \leq \sigma_{pos} \times W$$

- **Strong Supporters:** a commentator (ss) whose comments on a candidate (c) during a period (p) satisfies the following:

$$\frac{\sum_{ss}(pos\ score)}{No.of\ Comments\ (ss)} > \frac{\sum_{ss}(neg\ score)}{No.of\ Comments\ (ss)}$$

AND

$$\frac{\sum_{ss}(pos\ score)}{No.of\ Comments\ (ss)} > \sigma_{pos} \times W$$

- **Dissenters**: a commentator (*d*) whose comments on a candidate (*c*) during a period (*p*) satisfies the following:

$$\frac{\sum_{d}(neg\ score)}{No.of\ Comments\ (d)} > \frac{\sum_{d}(pos\ score)}{No.of\ Comments\ (d)}$$

AND

$$\mu_{neg} \times W \leq \frac{\sum_{d}(neg\ score)}{No.of\ Comments\ (d)} \leq \sigma_{neg} \times W$$

- **Strong Dissenters**: a commentator (*sd*) whose comments on a candidate (*c*) during a period (*p*) satisfies the following:

$$\frac{\sum_{sd}(neg\ score)}{No.of\ Comments\ (sd)} > \frac{\sum_{sd}(pos\ score)}{No.of\ Comments\ (sd)}$$

AND

$$\frac{\sum_{sd}(neg\ score)}{No.of\ Comments\ (sd)} > \sigma_{neg} \times W$$

In other words, if a commentator commented mostly positively on a candidate and his/her averaged positive score is between the mean and one standard deviation for all commentators' scores on that candidate, then he/she will be classified as supporter. If his/her score is above one standard deviation, then he/she is strong supporter. Similarly, the same rules applied on dissenters and strong dissenters but with negative scores. Figure 2 presents the flow chart of classifying a commentator.
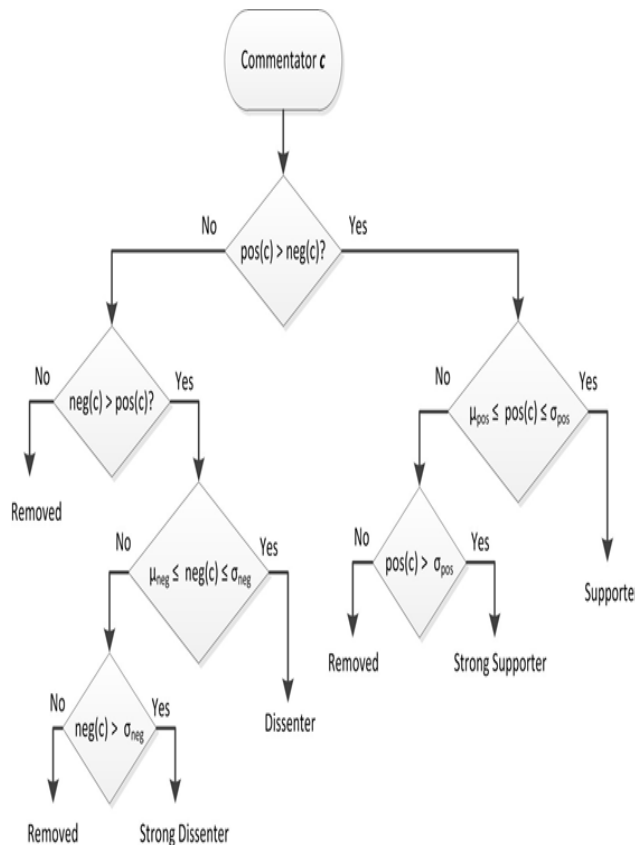

Figure 2. Flow Chart of Classifying a Commentator

3.4. Linguistic and Psychological Indices

We utilized the Linguistic Inquiry and Word Count (LIWC) tool [27] to infer the linguistic and psychological indices of commentators based on their comments to a candidate's posts. The algorithm processes each comment to search for and count words in psychology-relevant categories by comparing them with a dictionary file. Then, it assigns relevant words to one of the indices. Next, the indices are scored 0 to 100 based on the percentage of all words in the document.

The *Analytical Thinking* index examines how formal, logical, and hierarchical the writing is. This index is important in revealing how well-educated a person is [29]. The *Clout* index suggests that the commentator is writing from the perspective of high expertise and is confident [30]. The *Authenticity* index measures how authentic and honest the writing is; higher scores suggest honest writing and lower scores suggest deceptive writing [31]. *Anger, Anxiety,* and *Sadness* measure the tone expressed in writing [32]. The indices are scored based on their occurrences within the corpus.

4. Results and Findings

In this paper, due to space limitations, we will present the results of period II only. Results for other time periods can be obtained by contacting the authors.

4.1. Resultant Policy Related Topics

We utilized Gensim Latent Dirichlet Allocation (LDA) [33] to infer latent topics from the candidate posts. For each period, we ran the LDA algorithm iteratively on the posts to infer 10 topics. During each run of the algorithm we removed stop words (e.g. the, vote, candidates' names, states). Each topic consisted of 10 keywords. For example, the immigration topic was made up of the following keywords (illegal, immigration, border, build wall, immigrants) and Jobs/Taxes topic was made up the following keywords (Jobs, workers, working class, taxes, tax plan, millions). Next, we manually labeled topics with their appropriate label (e.g. taxes, healthcare, etc.) and removed irrelevant topics that are not related to public policies (e.g. campaigning, events). Table 2 (next page) shows the top topics for each candidate in period II.

Table 2. Top Topics of Candidates

Topic	Cruz	Kasich	Rubio	Trump	Clinton	Sanders
I	Jobs, Taxes	Wall Street, Climate Change	Immigration	Healthcare, Women Rights	Jobs, Taxes	Wall Street, Climate Change
II	Immigratio	Immigration	Terrorism/Security	Terrorism/Security	Immigration	Jobs, Taxes
III	-	Jobs, Taxes	-	Wall Street, Climate Change	Healthcare, Women Rights	Healthcare, Women Rights

Table 3. Distribution of Comments and Commentators by Candidate

Category	Unit	Cruz	Kasich	Rubio	Trum	Clinton	Sanders
Strong Supporters	Comments	2180	237	49	5290	168	360
	Commentators	2030	230	49	4481	154	345
Supporters	Comments	2002	553	114	9166	282	119
	Commentators	1808	520	114	7780	259	118
Dissenters	Comments	4676	174	79	8041	49	702
	Commentators	4112	169	79	7212	49	686
Strong Dissenters	Comments	4283	150	56	5700	130	449
	Commentators	3653	143	56	4754	115	446

Table 4. Ratio of Comments/Commentators by Candidate

Category	Unit	Cruz	Kasich	Rubio	Trum	Clinton	Sanders
Supporters/Strong Supporters	Comments	13141	1114	298	28197	629	1630
	Commentators	11603	1062	298	24227	577	1595
	Ratio	1.09	1.05	1.00	**1.18**	**1.09**	1.03
Dissenters/Strong Dissenters	Comments	8959	324	135	13741	179	1151
	Commentators	7765	312	135	11966	164	1132
	Ratio	**1.15**	1.04	1.00	1.14	**1.09**	1.02

4.2. Distribution of Comments and Commentators by Candidate

Table 3 shows distribution of commentators and their comments by candidate in period II. All candidates (except Cruz and Sanders) had more comments and commentators in the supporters and strong supporters categories compared to the ones in the dissenters and strong dissenters categories. For Cruz and Sanders, the opposite is true where they had larger number of dissenters and strong dissenters (with their comments) compared to the number of their supporters/strong supporters. Trump had the largest total number of comments and commentators while Rubio had the least. Clinton and Sanders had similar numbers of comments and commentators in supporters and strong supporters group. However, Clinton, compared to Sanders, had a much smaller number of comments and commentators in dissenters and strong dissenters categories. Both Trump and Cruz had a large number of comments and commentators across the different groups; however, unlike Trump, Cruz had a large number of comments and commentators in dissenters/strong dissenters.

Table 4 shows the ratio of comments/commentators for supporters/strong supporters and dissenters/strong dissenters. In this table, Trump is the leading Republican candidate with a score of 1.18 and Clinton is leading Democrat candidate with a score of 1.09. This ratio might be a plausible indicator of electoral success and can comes close to traditional polls. This is

in line with findings reported by Véronis [34] and Tumasjan et al. [35]. Véronis [34] studied the 2007 French elections on Twitter, and he observed that counting candidate's mentions can be a better predictor of electoral success than traditional polls. Similarly, Tumasjan et al. [35] analyzed tweets related to the 2009 German federal election and found that the mere number of tweets mentioning a political party reflects voter preferences. For Cruz, the ratio is higher in dissenters and strong dissenters than the ratio in supporters and strong supporters. All other candidates have either an equal ratio or a lower ratio.

4.3. Analysis of Indices

4.3.1. Analytical Thinking Index. (Table 5): Strong supporters always score the highest across all candidates. Strong supporters of Clinton scored higher than those of Trump. The score is relatively low for all candidates in dissenter and strong dissenter groups.

4.3.2. Authenticity Index. (Table 6): Authentic comments might be more understandable compared to deceptive comments, which are hinged on imagination [36]. They differ from each other based on the level of detail, where authentic comments are typically more detailed than deceptive ones [37]. The scores for dissenters and strong dissenters are high across all candidates. This indicates that honest writing is more likely to be shown in the comments written by dissenters or strong dissenters. Comments written by

Table 5. Analytical Thinking Scores

Category	Cruz	Kasich	Rubio	Trum	Clinton	Sanders
Strong Supporters	**72.32**	**66.91**	**85.27**	**89.26**	**92.84**	**89.18**
Supporters	53.03	53.70	54.02	57.98	56.03	44.03
Dissenters	64.82	59.66	45.11	54.30	52.20	51.95
Strong Dissenters	54.65	58.25	46.78	65.56	61.30	55.83

Table 6. Authenticity Scores

Category	Cruz	Kasich	Rubio	Trump	Clinton	Sanders
Strong Supporters	10.94	11.23	1.00	2.50	1.00	1.91
Supporters	21.45	31.03	11.13	11.09	8.96	14.44
Dissenters	20.34	24.84	**42.94**	**24.86**	16.37	**33.36**
Strong Dissenters	**26.80**	**38.04**	38.60	20.05	**17.23**	24.11

Table 7. Clout Scores

Category	Cruz	Kasich	Rubio	Trum	Clinton	Sanders
Strong Supporters	55.85	60.72	55.80	51.92	51.55	55.21
Supporters	67.87	**70.00**	**72.90**	**68.41**	64.24	**67.00**
Dissenters	**69.65**	69.37	67.30	64.33	77.09	57.23
Strong Dissenters	62.60	66.15	57.93	50.47	60.27	45.93

supporters and strong supporters may be deceptive or may not include details. A possible explanation being that the supporters and strong supporters already have a bias to supporting their candidate which comes across in their writing as deceptive' or at least not being authentic or credible due to lack of enough detail and evidence [18].

4.3.3. Clout Index. (Table 7): In contrast to Cruz and Clinton supporters, the supporters of the other candidates are confident. Confidence is greater among dissenters and strong dissenters for Clinton. Given the nature of the 2016 election campaign, and the rift between the supporters of the two democrat candidates, this result is interesting and points to the fact that Clinton's dissenters, and strong dissenters, were more confident in their remarks.

4.3.4. Indices of Top Topics. We further deepened our analysis to see the trends of indices over top topics per a candidate. Due to space limitations, we are showing only the tables for Trump and Clinton. For Trump, the highest scores for Analytical Thinking are typically in the Strong Supporters category (across all the top topics) (Table 8). For Authenticity and Clout, the Dissenters and Strong Dissenters score higher compared to his fans, where the highest scores are from Dissenters on Topic 1.

Table 9 presents the scores of indices of the four categories for Clinton's top topics. Her supporters score the highest on Analytical Thinking. The Dissenters and Strong Dissenters score higher scores on Authenticity and Clout, with Topic II receiving the highest scores.

Table 8. Linguistic Indices for Trump's Top Topics

Category	Index	Topic I	Topic II	Topic III
Strong Supporters	Analytical Thinking	**89.51**	89.24	88.54
	Authenticity	2.73	3.18	1.80
	Clout	52.86	51.82	52.15
Supporters	Analytical Thinking	60.32	60.30	56.42
	Authenticity	9.35	11.04	10.55
	Clout	62.06	69.04	67.64
Dissenters	Analytical Thinking	55.63	55.20	54.10
	Authenticity	**26.48**	23.55	25.33
	Clout	**64.99**	63.40	63.64
Strong Dissenters	Analytical Thinking	64.78	66.92	65.64
	Authenticity	21.61	19.57	20.08
	Clout	50.47	50.24	50.38

Next, we analyzed the topics that supporters supported the most and topics that brought most anger, anxiety, and sadness from dissenters. Table 10 (next page) shows positive sentiment of supporting groups and the negative emotions (Anger, Anxiety, Sadness) of dissenters on top topics for Trump (we omitted positive sentiment for the dissenters because these are typically zeros or negligible scores, similar applies to negative emotions scores for the supporting group). In this table, Trump's supporters supported him mostly on his Wall Street and climate change policies (Topic III). His opponents (strong dissenters) expressed the most anger on his healthcare and women rights policies, and anxious and sad on Wall Street and climate change policies.

Table 9. Linguistic Indices for Clinton's Top Topics

Category	Index	Topic I	Topic II	Topic
Strong Supporters	Analytical Thinking	**94.91**	90.89	92.74
	Authenticity	2.73	3.18	1.80
	Clout	52.86	51.82	52.15
Supporters	Analytical Thinking	56.67	48.49	58.54
	Authenticity	12.41	9.95	6.79
	Clout	55.99	62.61	69.37
Dissenters	Analytical Thinking	46.08	55.13	46.84
	Authenticity	4.23	**31.45**	13.81
	Clout	77.32	**86.69**	65.15
Strong Dissenters	Analytical Thinking	59.31	58.30	64.04
	Authenticity	19.71	20.97	14.32
	Clout	70.47	48.02	58.41

Table 10. Sentiment Indices for Trump's Top Topics

Category	Index	Topic I	Topic II	Topic III
Strong Supporters	Positive Sentiment	94.69	94.68	**95.33**
Supporters	Positive Sentiment	48.04	47.90	47.84
Dissenters	Anger	6.70	6.63	6.65
	Anxiety	1.22	1.35	1.82
	Clout	2.44	2.87	2.90
Strong Dissenters	Anger	**22.64**	20.21	19.05
	Anxiety	1.75	1.63	**2.08**
	Clout	6.18	8.66	**11.60**

Table 11 presents positive sentiment of supporting groups and the negative emotions (Anger, Anxiety, Sadness) of dissenters on top topics for Clinton (again we omitted negligible scores). Clinton's supporters supported her mostly on her jobs and taxes policy. Her dissenters expressed anger and anxious sentiment over her immigration policy. They expressed the most sadness on her healthcare and women rights policies.

Table 11. Sentiment Indices for Clinton's Top Topics

Category	Index	Topic I	Topic II	Topic III
Strong Supporters	Positive Sentiment	**99.69**	98.68	99.23
Supporters	Positive Sentiment	51.57	54.68	51.27
Dissenters	Anger	3.86	4.06	5.99
	Anxiety	0.46	0.5	1.67
	Clout	1.06	1.99	1.79
Strong Dissenters	Anger	8.23	**16.07**	12.03
	Anxiety	1.34	**2.47**	1.61
	Clout	2.78	1.98	**8.46**

5. Conclusion and Areas for Future Research

In this paper, we identified the characteristics of people who interacted with candidates during the 2016 U.S. elections on Facebook. First, we identified policy related topics that bring political polarization within candidates' posts using topic modeling. Second, we proposed a novel approach to classifying participants based on the positive and negative sentiments expressed in their comments. In addition, we also presented the analysis of different linguistic and psychological indices and how they differ across groups.

Our future research areas include: 1) understanding the contagion effect on the interactions between posts and comments, 2) building predictive models that link online and offline political activities, 3) connecting Facebook data with activity on other social media platforms, such as Twitter, and 4) studying the dynamics of networks of commentators in relation to policy topics and candidates they support.

Studying how a post by a candidate, especially the tone expressed in it (positive or negative) impacts the responses of the various groups of commentators will help us understand how sentiments spread around individuals (candidates) and the topics (policy viewpoints and/or priorities). Another area of future research is developing predictive models that link online and offline political engagements.

Past research shows mixed findings for predictability between online and offline behavior. There are studies that found that offline behavior and online behavior are consistent and each can be an indicator for political behavior of the other [25], [11], [38]; others found the opposite [39]. We are in the process of looking at critical incidents that took place offline during the election and linking them with online activities, i.e. particular posts or unusual activity with comments. The goal being to see if we can train machine learning models to alert us when online activities cross given thresholds which might indicate potential type of offline activity (e.g. rallies, change in key messages in an upcoming speech, etc.).

Fusing Facebook data with activities on other social media (e.g. Twitter) platforms opens up interesting opportunities for new research. Past research has also found strong use of Twitter in political campaigns and, in the U.S.; this effect was strengthened by President Trump's use. Véronis [34] studied the 2007 French elections on Twitter and he observed that counting candidate's mentions can be a

better predictor of electoral success than traditional polls. Similarly, Tumasjan et al. [35] analyzed tweets related to the 2009 German federal election and found that the mere number of tweets mentioning a political party reflects voter preferences. Standberg [40] analyzed both Twitter and Facebook to understand the use of social media in the 2011 Finland election cycle. The study concluded that differences in social media had much   to do with demographic characteristics such as age, income, gender, education, and Internet use. We plan to link our Facebook data with Twitter data around key events (e.g. during debates, campaign rallies, etc.), albeit in shorter timeframes (+/- 1 day). Linking data will allow us to see how online conversations on Twitter and posts by candidates reflect on Facebook immediately and drive future conversations in terms of comment.

Lastly, we have network level data linking commentators with candidates and the various policy topics. We intend to study the evolution of networks across the four time periods. Specifically, we are interested in looking at how network typologies change as candidates drop out and as the prominence of policy topics change over time.

6. References

[1] Rainie, Lee, et al. "Social media and political engagement." *Pew Internet & American Life Project* 19 (2012).

[2] Gil de Zúñiga, Homero, Nakwon Jung, and Sebastián Valenzuela. "Social media use for news and individuals' social capital, civic engagement and political participation." *Journal of Computer-Mediated Communication* 17.3 (2012): 319-336.

[3] Robertson, Scott P., Ravi K. Vatrapu, and Richard Medina. "Online video "friends" social networking: Overlapping online public spheres in the 2008 US presidential election." *Journal of Information Technology & Politics* 7.2-3 (2010): 182-201.

[4] Cogburn, Derrick L., and Fatima K. Espinoza-Vasquez. "From networked nominee to networked nation: Examining the impact of Web 2.0 and social media on political participation and civic engagement in the 2008 Obama campaign." *Journal of Political Marketing* 10.1-2 (2011): 189-213.

[5] Gil de Zúñiga, Homero, Logan Molyneux, and Pei Zheng. "Social media, political expression, and political participation: Panel analysis of lagged and concurrent relationships." *Journal of Communication* 64.4 (2014): 612-634.

[6] Pasek, Josh, Eian More, and Daniel Romer. "Realizing the social Internet? Online social networking meets offline civic engagement." *Journal of Information Technology & Politics* 6.3-4 (2009): 197-215.

[7] Valenzuela, Sebastián, Namsu Park, and Kerk F. Kee. "Is there social capital in a social network site?: Facebook use and college students' life satisfaction, trust, and participation." *Journal of Computer-Mediated Communication* 14.4 (2009): 875-901.

[8] Kushin, Matthew James, and Masahiro Yamamoto. "Did social media really matter? College students' use of online media and political decision making in the 2008 election." *Mass Communication and Society* 13.5 (2010): 608-630.

[9] Xenos, Michael, Ariadne Vromen, and Brian D. Loader. "The great equalizer? Patterns of social media use and youth political engagement in three advanced democracies." *Information, Communication & Society* 17.2 (2014): 151-167.

[10] Loader, Brian D., Ariadne Vromen, and Michael A. Xenos. "The networked young citizen: social media, political participation and civic engagement." (2014): 143-150.

[11] Gustafsson, Nils. "The subtle nature of Facebook politics: Swedish social network site users and political participation." *New Media & Society* 14.7 (2012): 1111-1127.

[12] Enli, Gunn Sara, and Eli Skogerbø. "Personalized campaigns in party-centred politics: Twitter and Facebook as arenas for political communication." *Information, Communication & Society* 16.5 (2013): 757-774.

[13] Woolley, Julia K., Anthony M. Limperos, and Mary Beth Oliver. "The 2008 presidential election, 2.0: A content analysis of user-generated political Facebook groups." *Mass Communication and Society* 13.5 (2010): 631-652.

[14] Fernandes, Juliana, et al. "The writing on the wall: A content analysis of college students' Facebook groups for the 2008 presidential election." *Mass Communication and Society* 13.5 (2010): 653-675.

[15] Gibson, Rachel K., and Ian McAllister. "Do online election campaigns win votes? The 2007 Australian "YouTube" election." *Political Communication* 28.2 (2011): 227-244.

[16] de Boer, Noortje, Hannah Sütfeld, and Jacob Groshek. "Social media and personal attacks: A comparative perspective on co-creation and political advertising in presidential campaigns on YouTube." *First Monday* 17.12 (2012).

[17] Groshek, Jacob, and Ahmed Al-Rawi. "Public sentiment and critical framing in social media content during the 2012 US presidential campaign." *Social Science Computer Review* 31.5 (2013): 563-576.

[18] Sweetser, Kaye D., and Ruthann Weaver Lariscy. "Candidates make good friends: An analysis of candidates' uses of Facebook." *International Journal of Strategic Communication* 2.3 (2008): 175-198.

[19] Williams, Christine B., and Girsh J. Gulati. "Facebook grows up: An empirical assessment of its role in the 2008 congressional elections." *Proceedings from Midwest Political Science Association, Chicago* 32 (2009): 53.

[20] Andersen, Kim Normann, and Rony Medaglia. "The use of Facebook in national election campaigns: politics as usual?." *International Conference on Electronic Participation*. Springer, Berlin, Heidelberg, 2009.

[21] Vitak, Jessica, et al. "It's complicated: Facebook users' political participation in the 2008 election." *CyberPsychology, behavior, and social networking* 14.3 (2011): 107-114.

[22] Vesnic-Alujevic, Lucia. "Political participation and web 2.0 in Europe: A case study of Facebook." *Public Relations Review* 38.3 (2012): 466-470.

[23] Carlisle, Juliet E., and Robert C. Patton. "Is social media changing how we understand political engagement? An analysis of Facebook and the 2008 presidential election." *Political Research Quarterly* 66.4 (2013): 883-895.

[24] Gerodimos, Roman, and Jákup Justinussen. "Obama's 2012 Facebook campaign: Political communication in the age of the like button." *Journal of Information Technology & Politics* 12.2 (2015): 113-132.

[25] Lane, Daniel S., and Sonya Dal Cin. "Sharing beyond Slacktivism: the effect of socially observable prosocial media sharing on subsequent offline helping behavior." *Information, Communication & Society* (2017): 1-18.

[26] Blei, David M., Andrew Y. Ng, and Michael I. Jordan. "Latent dirichlet allocation." *Journal of machine Learning research* 3.Jan (2003): 993-1022.

[27] Tausczik, Yla R., and James W. Pennebaker. "The psychological meaning of words: LIWC and computerized text analysis methods." *Journal of language and social psychology* 29.1 (2010): 24-54.

[28] Pukelsheim, Friedrich. "The three sigma rule." *The American Statistician* 48.2 (1994): 88-91.

[29] Pennebaker, James W., et al. "When small words foretell academic success: The case of college admissions essays." *PloS one* 9.12 (2014): e115844.

[30] Kacewicz, Ewa, et al. "Pronoun use reflects standings in social hierarchies." *Journal of Language and Social Psychology* 33.2 (2014): 125-143.

[31] Newman, Matthew L., et al. "Lying words: Predicting deception from linguistic styles." *Personality and social psychology bulletin* 29.5 (2003): 665-675.

[32] Cohn, Michael A., Matthias R. Mehl, and James W. Pennebaker. "Linguistic markers of psychological change surrounding September 11, 2001." *Psychological science* 15.10 (2004): 687-693.

[33] Rehurek, Radim, and Petr Sojka. "Software framework for topic modelling with large corpora." *In Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*. 2010.

[34] Véronis, Jean. "Citations dans la presse et résultats du premier tour de la présidentielle 2007." *Retrieved December* 15 (2007): 2009.

[35] Tumasjan, Andranik, et al. "Predicting elections with twitter: What 140 characters reveal about political sentiment." *Icwsm* 10.1 (2010): 178-185.

[36] Yoo, Kyung-Hyan, and Ulrike Gretzel. "Comparison of deceptive and truthful travel reviews." *Information and communication technologies in tourism 2009*(2009): 37-47.

[37] Hancock, Jeffrey T., et al. "On lying and being lied to: A linguistic analysis of deception in computer-mediated communication." *Discourse Processes* 45.1 (2007): 1-23.

[38] Tufekci, Zeynep, and Christopher Wilson. "Social media and the decision to participate in political protest: Observations from Tahrir Square." *Journal of Communication* 62.2 (2012): 363-379.

[39] Theocharis, Yannis, and Will Lowe. "Does Facebook increase political participation? Evidence from a field experiment." *Information, Communication & Society* 19.10 (2016): 1465-1486.

[40] Strandberg, Kim. "A social media revolution or just a case of history repeating itself? The use of social media in the 2011 Finnish parliamentary elections." *New Media & Society* 15.8 (2013): 1329-1347.