The Role of Social Media in the Discussion of Controversial Topics

Laura M. Smith, Linhong Zhu, Kristina Lerman and Zornitsa Kozareva
Information Sciences Institute
4676 Admiralty Way
Marina Del Rey, CA 90292
{lsmith,linhong,lerman,kozareva}@isi.edu

***Abstract*—In recent years, social media has revolutionized how people communicate and share information. Twitter and other blogging sites have seen an increase in political and social activism. Previous studies on the behaviors of users in politics have focused on electoral candidates and election results. Our paper investigates the role of social media in discussing and debating controversial topics. We apply sentiment analysis techniques to classify the position (for, against, neutral) expressed in a tweet about a controversial topic and use the results in our study of user behavior. Our findings suggest that Twitter is primarily used for spreading information to like-minded people rather than debating issues. Users are quicker to rebroadcast information than to address a communication by another user. Individuals typically take a position on an issue prior to posting about it and are not likely to change their tweeting opinion.**

***Index Terms*—Twitter, Political Opinions, Sentiment Analysis**

I. Introduction

The microblogging site Twitter has grown quickly in popularity due to its openness, ease of use, and flexibility. Twitter users share short text messages, known as tweets, with their followers and can also retweet messages they receive. In addition to disseminating information via retweets, Twitter users also engage in conversation using the "@username" mention convention. Twitter and similar social media sites are used heavily in the public sphere, for example, to disseminate information during emergencies and natural disasters and to mobilize support for social and political movements [1].

The use of social media, including blogosphere, for political activity has been intensely studied [2], [3], driven largely by hope to use these data to predict election results [4]. Despite the scrutiny, the role of new media in the public discourse about controversial issues is still largely unexplored [5]. Should taxes be raised to pay for schools? Should the death penalty be abolished? Should a third felony conviction result in a lifelong imprisonment? Should manufacturers be required to label foods containing genetically modified organisms? Complex questions such as these are routinely posed to the electorate of many states. How does the public use social media with respect to these issues? Do people leverage social media to gather information about the topic? Do they use this information to form their position or opinion? Do they engage others in discussion to help them frame their position? Do they use social media to promote their position, engage people holding the opposite opinion, and support those holding the same opinion? Evidence from the analysis of blogs [2] and Twitter [3] suggests that politically active users segregate according to partisan divisions, spreading ideologically acceptable messages within their ideological community and occasionally engaging users on the opposite side through mentions. It is not clear how much these findings carry over from the political domain. While some of the issues may have a clearly partisan association, e.g., taxation is often viewed as a liberal policy, others do not.

To investigate these questions, we have collected data from Twitter in the months leading up to the November elections. The November 2012 California ballot contained 11 initiatives, or *propositions*, on a variety of issues, including state taxation, corrections, and food labeling among others. This domain is a convenient choice for our study: discussions took place over a relatively short time period preceding the election, at the end of which voters had to decide on a position, either for or against the proposition. If the proposition won popular vote, it became law. Moreover, propositions had clearly identified advocates, both pro and con, with a presumed mission to champion their position on the issue.

Using these data, we study how positions on controversial issues are communicated on Twitter, how they evolve, and how they shape social interactions. Before we can address these questions, we have to automatically identify a user's position on the issue. In Section III we describe our method for classifying whether a given message expresses an opinion that is for or against a specific proposition. Using the classification results, we carry out an empirical study of user activity and interactions with other users (Sec. IV) and opinion dynamics (Sec. V). We found that users are quicker to rebroadcast messages sharing the same position on the topic than they are to engage others in discussions, which often cross ideological lines. Moreover, users who tweeted about propositions were unlikely to change their expressed opinions on topics, suggesting that Twitter played a minimal role in shaping public opinion about the California ballot propositions.

II. Description of Data

Our data consist of Twitter posts related to initiatives that appeared on the November 2012 California ballot. A ballot initiative, or *proposition*, is a political process that allows citizens of some states, including California, to place new legislation on the ballot. If the proposition wins the popular vote, it becomes law. A total of eleven propositions appeared on the 2012 California ballot, including two that would raise taxes to fund public schools (Propositions 30 and 38), one that would require manufacturers to label products that contain genetically modified organisms (GMOs) (Proposition 37), a proposition to abolish the death penalty (Proposition 34), and one to repeal the three strikes law (Proposition 36), among others. This domain is a convenient choice for our study: discussions about topics (propositions) occurred over a relatively limited time period, and propositions had clearly identified advocates, representing both pro and con positions, with a presumed intention to promote their positions on the issue and to get their followers to do the same.

We initiated data collection in August 2012 using the strategy described below. First, we created terms related to proposition names, including "proposition" or "prop," numeric value, and "yes" or "no" to indicate stance: e.g., "prop30", "proposition37", "yeson30", "noprop32," as well as terms related to propositions such as "ca2012" and "nonewtaxes." We later extended the set of terms to include those that occur frequently with these terms. This strategy allowed us to identify such relevant terms as "righttoknow", "LabelGMOs", "voteyeson37", and "stopspecialexemptions," for a total of 95 terms. Using the Search API, we monitored Twitter collecting tweets that contained these terms either as hashtags (e.g., "#prop30") or keywords.

A. Users

In addition to monitoring Twitter feed for mentions of the topics related to the propositions, we also identified one or more *advocates*, that is Twitter accounts that explicitly promoted a position on the issue, and added them to the list of monitored accounts. Some of the advocates were easy to identify, such as "@YesOnProp30" or "@NoProp37." We found others by reviewing similar account suggestions made by Twitter, e.g., "@CARightToKnow." We then manually reviewed these accounts to verify their position on the issue and added them to monitored accounts.

Advocates are users with an agenda, for or against a proposition. The goal of advocates is to influence others to support their cause in order to win the election. Examples of advocates include sponsors of the proposition measures, unions or special interest groups, companies with a stake in the outcome of the election, and prolific users with a strong opinion on the topic. Some of these users may be paid to campaign through Twitter. Most of the advocates listed a website in their user profile.

We collected tweets made by all monitored accounts, creating a complete record of their activity in the months preceding the November election. In addition, we periodically retrieved the names of followers of the advocate accounts and tracked their tweets.

Among the people tweeting about the propositions, there are *neutral users*. These users are typically journalists or work with news media and are frequently associated with a website. We tracked several of these accounts, all of their tweets, and the activity of their followers. Examples of these users include "@CADollarDollar," a website dedicated to listing all donor information to both sides of campaigns, and "@KCET," a public television station.

B. Tweets

We collected a total of 44M tweets from 600K users. The 81 advocates we monitored had 900K followers. In this set were 997,914 posts that contain at least one monitored term, and 314,193 posts that contain at least one term as a hashtag. We selected only the tweets in English (as some keywords have meanings in other languages), and excluded tweets where a keyword was included in a user's name with a different meaning, such as "gmo." The tweets were further classified by proposition, where posts can be included into multiple categories. The breakdown of tweets by proposition is presented in Table I.

TABLE I
DATA COLLECTED FROM TWITTER BY PROPOSITION

#	Proposition Title	Tweets	Users
30	Temporary Taxes to Fund Education	33,980	12,634
31	State Budget, State and Local Government	1,333	516
32	Political Contributions by Payroll Deduction	32,472	13,486
33	Auto Insurance Prices Based on Driver History	1,340	463
34	Death Penalty	6,499	3,023
35	Human Trafficking	6,528	2,940
36	Three Strikes Law	2,073	1,439
37	Genetically Engineered Foods, Labeling	627,299	178,236
38	Tax to Fund Education/Early Childhood Programs	5,351	1,893
39	Tax Treatment for Multistate Businesses	1,835	750
40	Redistricting, State Senate Districts	551	323

C. Hashtags

A hashtag is a Twitter convention to categorize tweets, making them more easily discovered through searches. For the California propositions, many tags were used to describe not only the proposition of interest, but also a user's stance on the issue.

Occasionally, hashtags indicating two opposite positions appear together, such as "#NoProp37" and "#Yeson37." This enables the tweet to be visible through searches by users on both sides. At times hashtags on different propositions appeared in the same tweet. The pair "#Yes30" and "#No32" were frequently posted together. Many groups advertised these positions as a unit through television, radio, and flyers. Hashtags on Propositions 30 and 38 were also often found in the same tweet. Both propositions raised taxes for education.

By adding multiple hashtags with different meanings to the same tweet, determining the user's stance on a proposition becomes very challenging. The position expressed in a tweet may differ from that indicated by the hashtags used. Therefore, we look to classify the position of each tweet through the methods of Section III.

III. CLASSIFYING POSITIONS ON TOPICS

We developed an approach, inspired by sentiment analysis, to classify the position on the issue that is expressed in a tweet. The position, or stance, could be for the issue, which results in a "yes" vote for the proposition, or it could be against the issue, resulting in a "no" vote. Following techniques developed for sentiment analysis, we leverage features, such as hashtags in tweets, as well as resources, such as a sentiment lexicon, for the classification task.

A. *Preprocessing*

Twitter users are likely to use colloquialisms, slang, and abbreviations in their posts due to the 140 character limit that is imposed on users. They are also likely to make grammatical and spelling errors. Therefore, before delving into details of the classifiers, we briefly discuss the data preprocessing, which consists of two steps: tokenization and normalization. In the normalization process, the presence of abbreviations within a tweet is noted, and then abbreviations are replaced with their actual meaning (e.g., tmr$\rightarrow$ tomorrow). In addition, we lowercase all letters (e.g., BREAKING $\rightarrow$ breaking), and replace repeated characters with a single character (e.g.,noooo $\rightarrow$ no). Finally, we remove all the URLs and user mentions in the tweets. Note that the presence of URLs, user mentions, and uppercase letters are all recorded before removing or converting.

B. *Features*

We use a variety of features for classification:

1) word features: TFIDF vectors,
2) lexicon features: the presence of words from "for/against" dictionary, and
3) micro-blog features: the presence of hashtags, URL links, retweets, replies, user mentions, punctuation (exclamation and questions marks), and emotion-icons.

Note that the dictionary here is different from the lexicons used in traditional sentiment analysis, such as MPQA [6] and sentiwordnet [7]. Traditional sentiment lexicon may not be appropriate for classifying stances on controversial topics. As an example, "against" should appear as a negative word in traditional sentiment lexicon, but "against monsanto" indicates support for labeling genetically modified foods (Proposition 37). To address this problem, we use the Bernoulli model to compute the likelihood of a bi-gram appearing in class $c \in \{\texttt{yes, no}\}$. More specifically, let $p(t \mid c) = \frac{tf_c(t)}{D(t)}$ where $tf_c(t)$ denotes the number of times term t appears in class c, and $D(t)$ denotes the number of tweets that contains term t. If $p(t \mid yes) > \epsilon \times p(t \mid no)$, we add the phrase t into the "yes" dictionary. Alternatively, if $p(t \mid no) > \epsilon \times p(t \mid yes)$, t is included in the "no" dictionary. In all of the following experiments, we set ϵ=3.

C. *Training/Testing Set Descriptions*

Each of the propositions 30, 32, 33, 35, and 37 had multiple advocates for and against the proposition, as well as neutral users. We selected tweets from the main advocates and their active followers with the same position, and placed them into the labeled sets "Advocates" and "Followers," respectively. Additionally, we selected posts with a hashtag indicating a particular stance and placed them into the set "Hashtags." The statistics of the labeled tweets are listed in Table II.

TABLE II
STATISTICS OF LABELED TWEETS FOR TRAINING AND TESTING

Prop	Advocates		Followers		Hashtags	
	Yes	No	Yes	No	Yes	No
30	1,940	3,804	1,024	397	5,813	813
32	1,834	1,435	86	249	157	5,269
33	282	419	48	34	105	453
35	451	442	213	283	51	286
37	6,612	1,534	1,043	42	27,134	1,011

For the two propositions with the most tweets, 30 and 37, we selected users that followed one of the main advocates ("@YesOnProp30", "@StopProp30", "@CARightToKnow", and "@NoProp37") for two months prior to the election, where the user had an easily identifiable position. The tweets of these labeled users were then used as a testing set. The number of labeled users is listed in Table III.

TABLE III
STATISTICS OF MANUALLY LABELED USERS

Prop	Yes Users	No Users	Neutral Users
30	107	32	—
37	233	61	8

D. *Performance Evaluation*

We trained our classifier with all of the above features using SVM-light [8] and compared its performance with a set of baselines with leave-one-out features: include everything but preprocessing (no-pre), include everything but lexicon features (no-lex), and include everything but micro-blog features (no-twi). Instead of using the sign of the decision function value δ [8] to determine the predicted class, we determine the classes according to the rule:

$$c = \begin{cases} yes & \text{if } \delta \geq 0.2 \\ no & \text{if } \delta \leq -0.2 \\ neutral & \text{otherwise.} \end{cases} \tag{1}$$

We report the classification accuracy for "yes/no" classes of the different classifiers in Figure 1. In this group of experiments, we use the labeled tweets from advocates for training and labeled tweets from followers for testing. Since the labels of testing data are extremely imbalanced (see Table II), we also report the classification accuracy if we classify all the testing data for a specific proposition according to the majority label. That is, if the majority of labeled tweets are "yes," then we label all tweets "yes" and refer to this classifier as (major). Based on Figure 1, we can conclude that using machine learning technology for classifying positions on issues achieves decent performance. The accuracy of all the classifiers is much higher than that of using majority voting based on prior labeling information. In addition, the results show that, similar to applications in other domains, preprocessing, sentiment lexicon, and micro-blog features are also quite helpful for position classification in Twitter. Excluding any feature significantly worsens the overall classification performance.

Finally, we evaluate the quality of classifying a user's position on an issue. To do this, we simply aggregate all

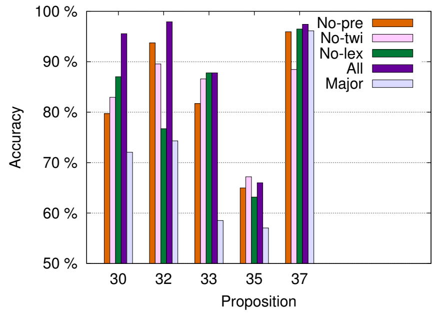


Fig. 1. Performance comparison over different types of classifiers in terms of accuracy. For each proposition, we used the following classifiers with the leave-one-out features: all but preprocessing (no-pre), all but lexicon features (no-lex), and all but micro-blog features (no-twi). We used a classifier with all features (all), as well as labeling according to the majority training set label (major).

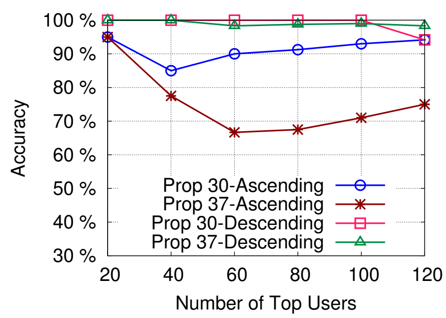


Fig. 2. User-level position classification performance for Propositions 30 and 37. We select the users discussed in Table III with known positions, and sort these labeled users in both descending order and ascending order of the user position score. We then select the top k users and calculate the percent of users that were correctly classified with k ranging from 20 to 120.

of a user's tweets and compute the average position score based on the positions predicted by the classifier. The user's position is determined using Equation (1). The performance results are presented in Figure 2. The classification of top-k "yes" users with highest position score (i.e., the top-k users in descending order), attains very good accuracy up to 95% for both Propositions 30 and 37. Unfortunately, for the top-k users who have the highest likelihood of a "no" position (i.e., the top-k in ascending order with lowest position score), the results are worse than expected. In the worst case, we can only achieve around 67% accuracy for the top-60 "no" users in Proposition 37. Overall, the performance of user-level position classification is worse than the tweet-level position classification. Since this work focuses more on the characteristics of social media for controversial topics, we leave the improvement of user-level position classification for future work.

IV. User Behavior Analysis

In this section, we examine the behavior of users tweeting about the propositions. The users presented in this data have various levels of activity and types of interactions. Users can interact by mentioning other users by name through an @mention. This is done by including "@username" in the tweet text. This action notifies the user that "@username" was mentioned in someone's post. The location of the @mention within the tweet is used for different things. If the tweet begins with the @mention, this typically indicates a directed tweet to the mentioned user as part of a dialogue. If the tweet begins "RT @username: ...," then the tweeter is reposting a tweet produced by the mentioned user. This type of action, called a retweet, spreads information to more people, giving credit to the originating user. Other locations in the tweet of the @mention are most often associated as a third-person mention, not explicitly directing the tweet to the mentioned user.

Table IV gives the general breakdown of tweet characteristics in our data set. These vary greatly from a random sample of tweets observed by the authors of [9]. This suggests that when discussing controversial topics, users are more likely to interact with other users through retweets and by mentioning other users. We further analyze these and other behaviors of advocates, their followers, and others discussing the propositions.

TABLE IV
Characteristics of Tweets by Proposition

Prop	Tweets	With URL	Retweets	With @mention
30	33,980	43.9 %	45.8 %	66.2 %
32	32,472	64.5 %	51.5 %	73.0 %
33	1,340	69.9 %	29.9 %	47.8 %
35	6,528	60.9 %	62.5 %	78.1 %
37	627,299	64.1 %	40.3 %	65.0 %

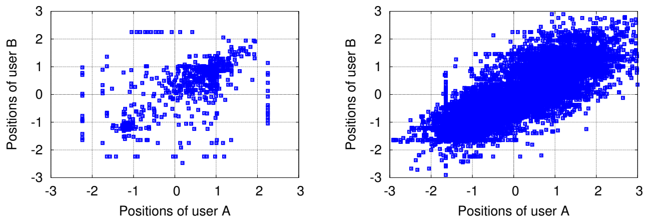
A. URLs

When discussing controversial topics on Twitter, users must express their opinions or present facts in only 140 characters. With a limited number of characters, it is challenging to make a case for one's argument. Thus, many people use Twitter to direct others to external websites and blogs, as well as in-depth newspaper and scientific articles using URLs embedded in tweets. The 43.9% – 69.9% of URL-containing tweets in our data set is substantially greater than the 22% of randomly sampled tweets found in [9].

Advocate users, such as "@YesOnProp30," direct others to webpages associated with their campaign or other articles endorsing their position. Neutral users, such as the public television station "@KCET," tweet links to in-depth analyses of arguments for both sides of the propositions. Many of the URLs posted in tweets on the topic "GMO food labeling" link to scientific articles as well as news articles and blogs. Frequently, users post URLs to their own associated websites, especially journalists and proposition advocates. For example, "@CADollarDollar" directed all 1785 URL postings to the website http://rhughes.com/dollardollarbill/. These URL posting behaviors are consistent with findings on political bloggers [10] and the political activity of Twitter users discussing the 2010 congressional midterm elections [3].

B. Dialogues

Macskassy [11] defined a dialogue as at least three tweets by at least two active users within K minutes through "@mention" actions. Here, we propose a more relaxed definition of dialogue, which is defined as a triple-tuple $D =< u, v, \mathcal{T} >$, where u and v are a pair of users who have mutually @mentioned each other, and $\mathcal{T}$ is a set of tweets that involve both u and v together, excluding retweets.



(a) User Position Correlation in Dialogues with Pearson Correlation Coefficient=0.565

(b) User Position Correlation in Retweets with Pearson Correlation Coefficient=0.851

Fig. 3. Emotion consistency: user position correlation for both (a) dialogues and (b) retweets. Emotion consistency examines whether a user A's position is close to another user B's position when A and B have communication. Each marker gives the average position of users A and B that (a) engaged in a dialogue with each other with mutual @mentions or (b) retweeted each other. The average position is the mean value of positions of all tweets posted by the individual user.

We examine the properties of dialogues via the aspects: emotion consistency theory, latency factor, and temporal factor. Initially, emotion consistency theory [12] suggests that two frequently co-occurring words should have similar sentiment polarity. Later, Tan et al. [13] extend this theory to user-level and conclude that the sentiment of a user is very close to the sentiments of @mentioning/@mentioned users. Now we study our dataset and validate whether the emotion consistency theory also holds for dialogue relations in controversial topics. While there is a positive correlation among user positions in mutual @mentioning dialogues (Pearson Correlation Coefficient of 0.565), in Figure 3(a) we observe that both agreements (the sign of A's position is the same with the sign of B's position) and arguments (the sign of A's position is opposite to the sign of B's position) occur. This is expected since our dataset includes conversations on controversial topics where communication between opposing positions happens.

We also study the relation between rounds of dialogues and the positions of dialogues, see Figure 4. We define one round of a dialogue as a user u @mention v and v @mention u in return. Generally, no matter whether the dialogue is

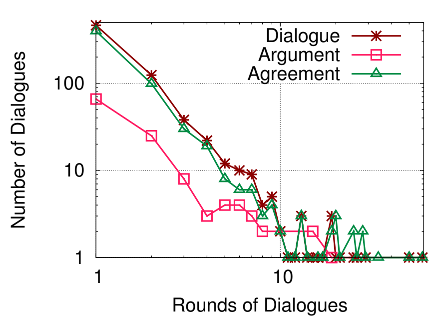


Fig. 4. Distribution of rounds of dialogues. The latency property of dialogues evaluates the relation between rounds of dialogues and the positions of dialogues. We define one round of a dialogue as a user u @mention v and v @mention u in return.

in argument or agreement, dialogues in controversial topics do not have as many rounds as stated in the work [11]. The distribution of rounds of dialogues has a long-tailed distribution, that is, most of the dialogues have only very few rounds. Interestingly, debating dialogues have much shorter rounds than normal dialogues: the former has a maximum of 19 rounds, while the latter has a maximum of 59 rounds. The results indicate that users do not typically utilize Twitter as a debating platform for controversial topics.

Finally, we look into the temporal characteristics of dialogues and its relation to the positions. Specifically, we want to know how quickly a user typically responds to a dialogue tweet. Through analysis, we find that the percent of dialogues that have an average response time of less than 1 hour are 36.5%, 51%, and 30% for all dialogues, argument dialogues, and agreement dialogues. This shows argument dialogues tend to have shorter response times than agreement dialogues.

C. Retweets

Twitter users often retweet to their followers the posts they receive from others. In the random sample of tweets observed by the authors of [9], only 3% of the tweets were retweets, whereas our data sets have a significantly higher proportion of retweets (see Table IV). Therefore, a large volume of tweets on these topics consists of communication and spreading information via retweeting.

By taking these posts and locating the original tweet, we calculated the time delay between the first user's tweet and the second user's retweet. We compare the average time delay retweet response (for user A to retweet user B's post) to the average time delay in a dialogue for the same users (user A to mention user B). The time delay difference between these types of responses are displayed in Figure 5. On average, it

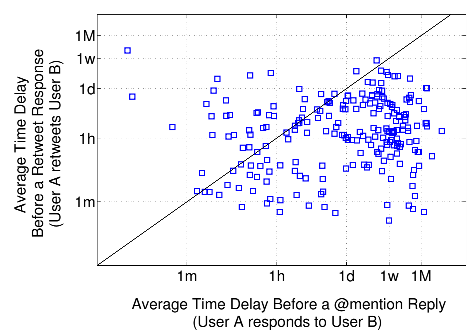


Fig. 5. This gives a time delay comparison for retweets and @mention dialogue responses. Each dot represents the time delay response for a user A to another user B. It compares the average time it takes for user A to retweet a post of user B to the average dialogue time between an @mention of B to an @mention of A (referred to as an @mention response). Points above the $y = x$ line indicate a longer time delay for a retweet. Points below the line indicate a longer @mention dialogue response time.

takes longer for a user to respond to an @mention in a dialogue than for a user to retweet a post by the same individual. This may be due to visibility limitations (users must access the "Connect" tab to view mentions) or the interest of the user in spreading information versus participating in dialogues.

Does emotion consistency theory hold for positions of users in retweets? Figure 3(b) shows the average position of user A inferred from classified positions of all her relevant tweets versus the average position of user B, for all pairs of users where A retweets B. The positions of the pairs are highly correlated (Pearson Correlation Coefficient of 0.851). Though there could be some bias since the position of a retweet and the original tweet are the same, emotion consistency theory appears to hold for user-user retweeting relations.

D. Behavior of Advocates

The goal of an advocate is to persuade others to have the same position on the issue and to convince their followers to do the same. The behavior of advocates, specifically their use of hashtags, URLs, @mentions, and retweets, varied greatly, as shown in Table V, though all had a high proportion of tweets containing URLs.

TABLE V
SELECTED ADVOCATES' BEHAVIORS

Prop	Advocate	Tweets	URL	RT	@mention
30	"@YesOnProp30"	371	58.0%	51.5%	82.7%
30	"@StopProp30"	20	90.0%	25.0%	25.0%
37	"@CARightToKnow"	1,070	68.7%	43.1%	63.8%
37	"@NoProp37"	469	86.8%	9.2%	30.1%

Figure 6 takes a closer look at the behaviors of advocates *for* ("@CARightToKnow") and *against* ("@NoProp37") Proposition 37, the most actively discussed proposition in our data set.

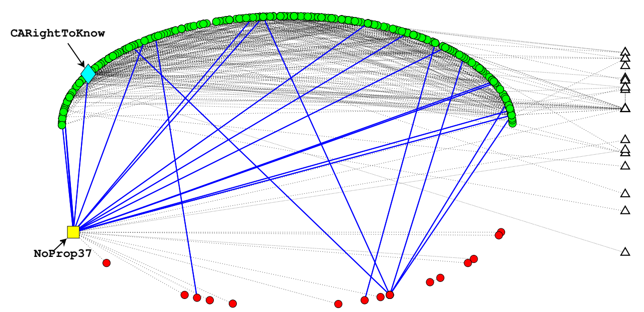


Fig. 6. The @mention network of user "@NoProp37." Each node represents a user, and its color and location indicate whether it was for prop. 37 (upper, green), against prop. 37 (lower, red), or neutral (right, white). Each edge indicates the two incident nodes mentioned each other in a tweet. Black edges are for communication between individuals of the same position or with neutral users, whereas blue edges show communication across the for and against communities. The main advocate for the proposition, "@CARightToKnow" (cyan diamond), and against the proposition, "@NoProp37" (yellow square), are labeled.

The network shows a set of pairwise communications between individual accounts through @mention in a tweet. Users who are against the proposition are positioned in the lower portion of the network, those that are for the proposition are in the upper portion, and the neutral users are positioned on the right. By separating the users in this manner, we can view the amount of communication between users of opposite positions, represented by the blue edges. The majority of this cross communication is associated with two users against Proposition 37, including the advocate "@NoProp37." Alternatively, "@CARightToKnow" fails to interact with opposing users other than "@NoProp37." These behaviors may be due to the large difference in the number of followers for each advocate. As of 9/10/2012, "@CARightToKnow" had over 64K followers, which doubled by 11/27/2012, while over the same period, "@NoProp37" had 191 and 418 followers, respectively. Despite these differences, it is interesting that the two advocates used different strategies. The user "@NoProp37" tweeted "#NoProp37" 240 times and either "#Yeson37", "#YesOnProp37", or "#YesProp37" 298 times. The benefit of posting a hashtag of the opposing opinion is that the tweet becomes more easily discovered through searches. The tendency of "@NoProp37" to address opposing users through @mentions, as well as by labeling tweets with a "#Yeson37" hashtag, reflects the strategy of actively engaging the opposite side in discussion, a strategy that may have ultimately helped to defeat the proposition.

The advocate "@YesOnProp30" was consistently active over the 2.5 months preceding the election. Looking at 110 users that followed this advocate since Sept. 1, 2012, we are able to measure how interested they were in the topic by looking at the proportion of tweets over time about Proposition 30. Figure 7 shows that although their overall activity level does not vary greatly over time, followers devote more of their tweets to the proposition as the election approaches.

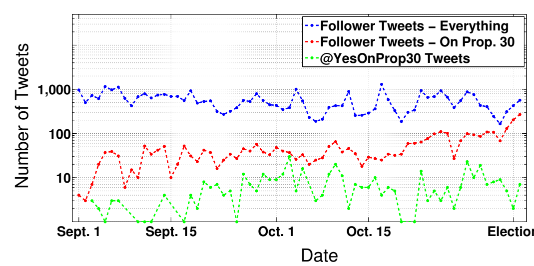
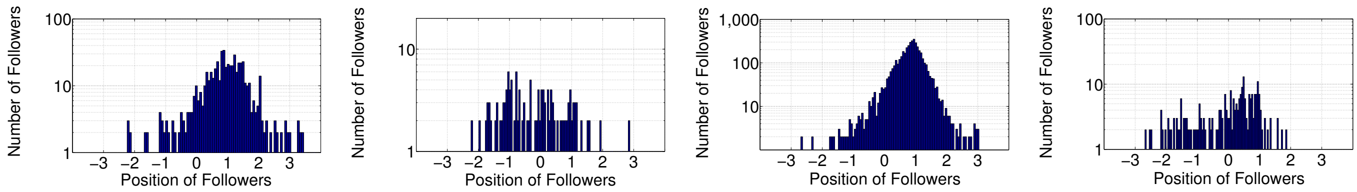


Fig. 7. Tweets by the original followers of @YesOnProp30 in the months preceding the election. The advocate's behavior does not vary much during this time period, but the followers devote a larger proportion of their tweets to Proposition 30 as the election approaches.

According to the findings of Tan et al. [13], we should expect that the sentiment of the advocate is close to the sentiments of the followers. Figure 8 gives the histograms of the followers' positions for the four main advocates of Propositions 30 and 37, namely "@YesOnProp30", "@StopProp30", "@CARightToKnow", and "@NoProp37." The majority of the followers of "@YesOnProp30" and "@CARightToKnow" have the same position in favor of the proposition. However, the followers of "@StopProp30" are split on both sides, and the majority of followers for "@NoProp37" have the opposite position as the advocate. Thus, the conclusions of [13] do not hold for all of these advocate accounts.

V. DYNAMICS OF USER POSITIONS

Do Twitter users change the position on the controversial issue expressed in their tweets over time? We address this question by looking at a user's tweets about a proposition and



(a) "@YesOnProp30" Followers (b) "@StopProp30" Followers (c) "@CARightToKnow" Followers (d) "@NoProp37" Followers

Fig. 8. Histogram of advocates' followers' average tweeting position. The plots give the user positions for followers of the advocates "@YesOnProp30", "@StopProp30", "@CARightToKnow", and "@NoProp37." The majority of followers have the same position as the advocate, except for the followers of "@NoProp37."

classifying the position expressed in these tweets. To quantify the degree to which positions remain the same, we analyze the correlation between positions of users in the past and those of users closer to the election date in Figure 9. It shows that users' earlier positions prior to Oct. 15, 2012 are highly correlated with their positions closer to election date (Pearson Correlation Coefficient of 0.779). It seems that users already took a position on the issue before tweeting about it. Perhaps this is because Twitter is a new social medium and people are just getting used to using it to express and discuss their opinions. It could be that only those people who feel strongly about their position are comfortable expressing it publicly.

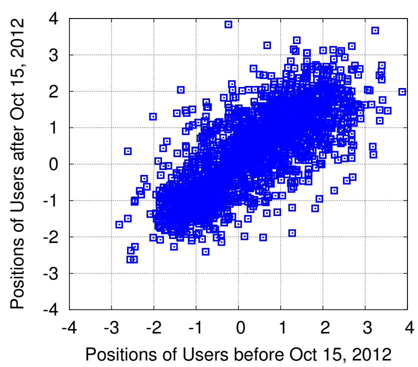


Fig. 9. Overall user position correlation with time for proposition 30, 32, 33, 35, and 37 with Pearson Correlation Coefficient=0.779. We split the tweets on these propositions into two parts: tweets posted before Oct. 15, 2012 and after Oct. 15, 2012. We compute the average position score of a user within each set and then report the average position scores of users before Oct. 15, 2012 versus their average position score after Oct. 15, 2012.

VI. Related Work

While there is a large amount of research in the area of sentiment classification for large pieces of text such as online reviews [14] and blogs [15], [16], how sentiments are expressed given the informal language and message-length constraints of microblogging has been much less studied. The state-of-the-art approaches for solving this problem, such as [17], [18], basically follow the traditional sentiment classification for reviews [14], who utilize machine learning techniques for sentiment classification of texts. Other approaches [19], [20] make use of sentiment WordNet and "emotion-icons"to conduct unsupervised sentiment analysis for Twitter data. Only a small branch of works look at the "hashtags" [21], [22], "mention" [13], and "follow-graph" [13] to perform better sentiment analysis of Twitter data.

In recent years, social media has become a platform for political discussion. People express their opinion in blog and microblog posts (i.e., tweets), and use them to spread information and engage in discussions. Researchers developed methods to identify political sentiment expressed in blog posts [15] and tweets [23], [24] and analyzed the patterns in and the structure of political interactions [2]. These studies revealed strong partisan asymmetries in online political interactions in the United States. Users within the same partisan group link to each other more [2], [22] and rebroadcast each other's posts more frequently [3] than members of opposing partisan groups. However, users are more likely to engage the opposing side in political dialogue [3], e.g., by using @mentions on Twitter. Such wealth of data about political opinions have inspired some researchers to develop methods to predict elections [25], though with mixed results [4].

While the focus of previous works was on analyzing online activity related to elections of candidates for political offices, the object of our study is analysis of opinions about controversial topics. We chose propositions that appeared on the California ballot as the controversial topics in our study. This domain is attractive because discussions about these issues took place over a relatively short time period immediately preceding the elections, at which time voters expressed their opinion by voting for or against each proposition. Though on the surface similar to the political domain, the propositions domain differs in several respects. Overall, this domain is non-partisan, although liberals may favor higher taxes to pay for education (yes on Prop30) and abolishing death penalty (yes on Prop34), while conservatives may favor tighter regulations of unions (yes on Prop32) and tougher sentencing (no on Prop36). Our domain is more general: political candidates may be controversial, but not all controversial issues are political. Despite these differences, our results are in general agreement with previous studies. Specifically, Twitter users tend to spread information from others holding a similar opinion and engage in some discussions with those holding the opposite opinion.

VII. Discussion

In this study, we examined tweets about controversial topics. The best classification performance was achieved with the three classes: "yes", "no", and "neutral." Although the tweet level classification gives good results, the user level classification is less accurate, particularly for the users against Proposition 37. Certain topics are more challenging for classification, like Proposition 37 where negative words are often associated with a "yes" vote.

We observed that users do not frequently utilize Twitter as a debating platform, even for controversial topics. The length of dialogues between users is short, with arguments being much shorter than conversations in agreement. Another indication that users are not debating is the correlation of user positions in dialogues (see Figure 3(a)). While this correlation is not as strong as the retweeting position correlation of Figure 3(b), it still suggests that users are more interested in conversing with others with the same opinion.

Since users are not focusing their efforts on debating, it is not surprising there is a strong correlation in opinions of users before and after Oct. 15, 2012. This suggests users do not change their observable opinion over time. Users typically take a position on an issue prior to publicly posting it.

While other studies have used Twitter to predict elections, our data suggest this is not possible for controversial topics. Proposition 37 is a prime example of differences in the voting population and the sample of users posting about the issue on Twitter. While 89.3% of the tweets were in support of the proposition, only 48.6% of voters favored it, ultimately causing the initiative to lose. Perhaps the general population of voters is unfamiliar with or uninterested in the technical aspects of genetically modified food, thus limiting the discussion on Proposition 37 to strongly opinionated users.

Here, users are primarily treating Twitter as a venue to spread information instead of having conversations on controversial issues. Users are apt to respond more quickly to retweeting a post than replying to a directed @mention in a dialogue. Individuals are more interested in spreading information and links to webpages than conversing about issues, demonstrated by the high proportion of tweets containing URLs and retweets.

A few users, advocates for the propositions, behave differently than the majority of ordinary users. They employ various strategies to interact with users of opposing opinions, attempting to persuade them to vote in their favor. The user "@NoProp37" actively engages in dialogues with the opposing side, frequently using a "#YesOn37" hashtag to capture the attention of users in favor of the initiative. Alternatively, the yes proponent, "@CARightToKnow," focused on interacting with users of the same position and accumulating followers.

Overall, social media is having an increasing role in political activity. The dynamics of user behavior on Twitter is different when discussing controversial topics than ordinary daily activities, with more information being spread through URLs and retweets. However, changing opinions and dialogues are less likely to be observed.

Acknowledgements

We are grateful to Suradej Intagorn for collecting data used in this study. This material is based upon work supported in part by the Air Force Office of Scientific Research under Contract Nos. FA9550-10-1-0569, by the Air Force Research Laboratories under contract FA8750-12-2-0186, and by DARPA under Contract No. W911NF-12-1-0034.

References

[1] G. Lotan, E. Graeff, M. Ananny, D. Gaffney, I. Pearce, and D. Boyd, "The revolutions were tweeted: Information flows during the 2011 tunisian and egyptian revolutions," *International Journal of Communications*, vol. 5, pp. 1375–1405, 2011.

[2] L. A. Adamic and N. Glance, "The political blogosphere and the 2004 us election: divided they blog," in *Proceedings of the 3rd international workshop on Link discovery*, 2005, pp. 36–43.

[3] M. D. Conover, B. Gonçalves, A. Flammini, and F. Menczer, "Partisan asymmetries in online political activity," *EPJ Data Science*, vol. 1, no. 1, pp. 1–19, 2012.

[4] D. Gayo-Avello, P. T. Metaxas, and E. Mustafaraj, "Limits of electoral predictions using twitter," in *ICWSM*, vol. 21, 2011, p. 2011.

[5] S. Aday, H. Farrel, M. Lynch, J. Sides, J. Kelly, and E. Zuckerman, "Blogs and bullets: New media in contentious politics," United States Institute of Peace, Tech. Rep., 2010.

[6] T. Wilson, J. Wiebe, and P. Hoffmann, "Recognizing contextual polarity: An exploration of features for phrase-level sentiment analysis," *Computational Linguistics*, vol. 35, no. 3, pp. 399–433, 2009.

[7] A. E. S. Baccianella and F. Sebastiani, "Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining," in *LREC*, 2010.

[8] T. Joachims, "Estimating the generalization performance of an svm efficiently," in *ICML*. Morgan Kaufman, 2000, pp. 431–438.

[9] D. Boyd, S. Golder, and G. Lotan, "Tweet, tweet, retweet: Conversational aspects of retweeting on twitter," in *System Sciences (HICSS)*. IEEE, 2010, pp. 1–10.

[10] L. McKenna and A. Pole, "What do bloggers do: an average day on an average political blog," *Public Choice*, vol. 134, no. 1-2, pp. 97–108, 2008.

[11] S. A. Macskassy, "On the study of social interactions in twitter," in *ICWSM*, 2012.

[12] R. P. Abelson, "Whatever Became of Consistency Theory?" *Personality and Social Psychology Bulletin*, vol. 9, no. 1, pp. 37–54, 1983.

[13] C. Tan, L. Lee, J. Tang, L. Jiang, M. Zhou, and P. Li, "User-level sentiment analysis incorporating social networks," in *KDD*, 2011, pp. 1397–1405.

[14] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up?: sentiment classification using machine learning techniques," in *ACL 02*, 2002, pp. 79–86.

[15] K. T. Durant and M. D. Smith, "Predicting the political sentiment of web log posts using supervised machine learning techniques coupled with feature selection," in *Advances in Web Mining and Web Usage Analysis*. Springer, 2007, pp. 187–206.

[16] P. Melville, W. Gryc, and R. D. Lawrence, "Sentiment analysis of blogs by combining lexical knowledge with text classification," in *KDD*, 2009, pp. 1275–1284.

[17] L. Barbosa and J. Feng, "Robust sentiment detection on twitter from biased and noisy data," in *COLING*, 2010, pp. 36–44.

[18] A. Go, R. Bhayani, and L. Huang, "Twitter sentiment classification using distant supervision," *Technical report*, pp. 1–6, 2009.

[19] A. Montejo-Ráez, E. Martínez-Cámara, M. T. Martín-Valdivia, and L. A. Ureña López, "Random walk weighting over sentiwordnet for sentiment polarity detection on twitter," in *WASSA*, 2012, pp. 3–10.

[20] X. Hu, J. Tang, H. Gao, and H. Liu, "Unsupervised sentiment analysis with emotional signals," in *WWW*, 2013.

[21] D. Davidov, O. Tsur, and A. Rappoport, "Enhanced sentiment learning using twitter hashtags and smileys," in *COLING*, 2010, pp. 241–249.

[22] M. D. Conover, B. Gonçalves, J. Ratkiewicz, A. Flammini, and F. Menczer, "Predicting the political alignment of twitter users," in *SocialCom*. IEEE, 2011, pp. 192–199.

[23] E. Mustafaraj, S. Finn, C. Whitlock, and P. T. Metaxas, "Vocal minority versus silent majority: Discovering the opionions of the long tail," in *SocialCom*. IEEE, 2011, pp. 103–110.

[24] F. M. F. Wong, C. W. Tan, S. Sen, and M. Chiang, "Quantifying political leaning from tweets and retweets," 2013.

[25] B. OConnor, R. Balasubramanyan, B. R. Routledge, and N. A. Smith, "From tweets to polls: Linking text sentiment to public opinion time series," in *ICWSM*, 2010, pp. 122–129.