I act, therefore I judge: Network sentiment dynamics based on user activity change

Kathy Macropol
Arcadia University
macropolk@arcadia.edu

Petko Bogdanov, Ambuj K. Singh, Linda Petzold, Xifeng Yan
University of California Santa Barbara
{petko,ambuj,petzold,xyan}@cs.ucsb.edu

***Abstract*—The study of influence, persuasion, and user sentiment dynamics within online communities has recently emerged as a highly active area of research. In this paper, we focus on analyzing and modeling user sentiment dynamics within a real-world social media such as Twitter. Beyond text and connectivity, we are interested in exploring the level of topical user posting activity and its effect on sentiment change. We perform topic-wise analysis of tweeting behavior that reveals a strong relationship between users' activity acceleration and topic sentiment change. Inspired by this empirical observation, we develop a new generative and predictive model that extends classical neighborhood-based influence propagation with the notion of user activation. We fit the parameters of our model to a large, real-world Twitter dataset and evaluate its utility to predict future sentiment change. Our model outperforms significantly (1 order of magnitude in accuracy) existing alternatives in identifying the individuals who are most likely to change sentiment based on past information. When predicting the next sentiment of users who actually change their opinion (a relatively rare event), our model is twice more accurate than alternatives, while its overall network accuracy is 94% on average. We also study the effect of inactive users on consensus efficiency in the opinion dynamics process both analytically and in simulation within the context of our model.**

I. Introduction

The proliferation of social media, forums, and communities has brought about networks that propagate news, opinions, and stances on a scale and speed that has never been seen before. Websites such as Facebook and Twitter produce novel information daily – not only in terms of friendship relations among users, but also abundant metadata (such as timestamps, tags, and links) as wells as textual and rich media content including users' updates and opinions. The textual information and tags within these datasets reveal users' sentiments on topics, as well as the influence of their network neighbors. Hence, studies of influence within networks are of natural importance with applications ranging from targeted advertising to the prediction of election results or the stock market.

Existing research has focused on the theoretical analysis and modeling of sentiment dynamics and influence within networks [1], [2], [3]. In this paper, we focus on sentiment dynamics in the context of a real-world network dataset, and specifically, incorporating behavioral traits such as timing and activity information available for users. Posting activity is a valuable source of information that reflects the temporal interest and attention of a user. By extending our analysis beyond the traditional text and user connectivity, we find a strong new connection between topic-based user activity levels, and changes in sentiment.

We hypothesize that a drastic *change* in the online activity level of a user is a precursor for a shifting sentiment/opinion. For example, in the period before elections, undecided voters may sway to one side on a given issue thanks to televised debates, political campaigns and their social neighborhood. Naturally, such voters who become increasingly polarized will express their opinion by posting more often than on average in public forums and social media. Similarly, dissatisfaction with a favored politician may decrease the number of posts of supporters from their normal levels.

We first set out to validate our hypothesized relation between change in activity and sentiment, based on empirical user behavior in Twitter. We discover a strong relationship between a user's *acceleration* in topic-based activity, and a subsequent change in the same user's topic sentiment (correlation values fifty times greater than expected at random). User and neighbors activity acceleration couples much stronger with future opinion shifts than the raw activity levels.

Building on our empirical observations in Twitter, we develop a new generative model for sentiment dynamics within social networks. Our model integrates both user activity acceleration, as well as neighborhood sentiment state in order to better fit and predict future changes of opinion. Similar connections between user sentiment and user message activity exist in the sociology literature: Henry et al [4] demonstrate that opinion and interest are more dynamic during activity-ramping phases of a topic, while Newman et al. [5] relate topic-based thresholds for user attention to topic activity. These connections, combined with our empirical evidence from Twitter, are the inspiration for our novel activity-based model for opinion dynamics.

When fitted to a real-world Twitter dataset, our model generates opinion dynamics which closely mirror the actual network behavior. It identifies users who are likely to change opinion an order of magnitude more accurately than a recently introduced alternative [1]. When predicting the next sentiment state of all users in a network our model achieves an average of 94% accuracy, while for users who change opinion it is twice more accurate than alternatives.

We make three main contributions in this work: (i) We demonstrate the relationship between opinion change and users' topic-based activity acceleration within a real-world network. (ii) We introduce a new model for activation-based sentiment dynamics and analyze the effect of inactive users on the efficiency of reaching sentiment consensus. (iii) We demonstrate the *applicability* and *accuracy* (2 orders of magnitude higher than alternatives) of our model for future sentiment change prediction in Twitter.

II. Related Work

Numerous prior studies focus on detecting positive/negative sentiment within user posts in social media [6], [7], [8]. These techniques are mainly concerned with static networks, and utilize both text processing, as well as graph and network-based approaches (for details see the Supplement [9]). Our approach focuses on the dynamics of sentiments as opposed to determining the sentiment state of a user/message. As such, our model is complementary to text- or network-based sentiment classification.

Exisiting opinion diffusion models employ Markov chain theory, neighbor opinion averaging, and others [2], [3], [10]. The majority of the previous work revolves around the connection between network structure and user sentiment, with individuals being influenced by their network neighbors' opinion, a phenomena explored in more detail for Twitter in [6]. Manipulation and control of user opinion and the effect of "stubborn" users have been the target of further research on the the established models [11], [12]. Unlike previous approaches that are limited to two sentiment/opinion states, the Kimura model [1] allows for the inclusion of more than 2 opinion labels by extending a basic voter model [13], where every node probabilistically copies its state from one of its neighbors. Different from all previous models, our goal is to incorporate the activity acceleration of network users for improved predictive power of future sentiment change.

The importance of user activity and its influence on multiple aspects of network behavior have been realized by Wilson et al [14], demonstrating that frequency of interaction may impact information flow within a network. Patterson et al [3] introduced a new opinion dynamics model, weighing network links according to interaction frequency. In [15], changes in individual user activity patterns were related to user interest. Sociologists have also discovered links between topic-based activity changes and user's opinion/interest in a topic. For example, a study in [4] discovered that users were more likely to focus attention on a topic when the level of topic-based activity was undergoing rapid *acceleration*, as opposed to simply correlating to high activity value levels. Despite the realization of the importance of user activity for opinion dynamics, the relation between *per-topic* activity changes and sentiment shifts has not been investigated within the context of large-scale real networks. Our model and analysis contributes to this important research area.

III. Data

Network and posts from Twitter. For our evaluation, we crawled Twitter (via its API) to obtain text, hashtag, and timestamps from posted tweets, as well as the follower structure among users. Our crawl began from a set of 10 randomly selected seed nodes whose recent posts had contained one or more of the topics listed in Table I. For each visited user, we obtained all available tweets, and proceeded to followees in a breadth first search manner. In this way, a complete collection of Twitter posts and information to which crawled users are exposed can be collected. Our crawled dataset includes $48M$ tweets sent across 6 years and $26M$ users. Approximately $12.4M$ of the crawled tweets contain hashtags (user-defined annotations of their post).

TABLE I: Statistics for Topic-Based Datasets

Topic	# of Tweets	# of Users	# of Hashtags	# of Graph Links
Obama	100,939	2,510	10,856	602,339
GOP	100,203	2,108	10,756	529,032
Liberal	51,474	2,072	5,632	532,589
Romney	8,321	1,045	1,879	172,509
Limbaugh	9,200	1,089	1,252	195,306

Detecting message and user sentiment. In order to study and model the change in topic sentiments across time within our data, we must first discover individual user sentiments within a fixed time point based on their posted content. To this end, we utilize two recently introduced sentiment mining techniques [16], [8] to label tweets with sentiment. As mentioned previously, numerous methods for discovering sentiment within text have been proposed, including averaging word polarity [17], topic-based analysis of words, emoticons, hashtags, and context information [7]. In this paper, we employ the techniques introduced in [16] because of their specialization to Twitter data, as well as the model introduced by Srivatsa and colleagues [8] because of its high classification quality. Extending the latter techniques, by incorporating emoticons and other textual/network features or better classification algorithms, may improve the quality and predictive power of our model. However, such directions are not within the scope and target of the current study.

We adopt a hashtag-based approach to identify broad "topics" within a Twitter dataset, similar to methods introduced in [16]. Hashtags, words prefixed by a hash symbol #, are widely used within tweets (26% of tweets within our collected dataset contained hashtags), and frequently identify concepts, subjects, and sentiment information contained within the text (for example #obama, or #awesome).

We annotate individual tweets following the approach in [8], [16]. Details of how we apply the approach to our dataset are available in the Supplement [9]. Given tweet sentiment, we determine the user-level sentiment within a time period (3 months) by aggregating the sentiments of all the user's tweets within the same period. Users who sent 25% more positive (than negative) tweets are classified as having a positive sentiment for that quarter, and similarly for negative. All other cases were classified as neutral. The result of this preliminary sentiment analysis step is a time series of sentiment states for each user within out dataset.

We apply the user sentiment classification to our Twitter dataset to obtain user sentiments across time for five separate hashtag topics within: #Obama, #GOP / #Republican (represented in figures hereafter as "#GOP"), #Romney, #Liberal / #Democrat (represented hereafter as "#Liberal"), and #Limbaugh. These five topics were chosen based upon their high opinion volatility and intensity. Table II presents the sizes and features for the obtained topic-based subnetworks. To help validate the sentiment classification results, we compared the individual tweet classifications with a set of 1000 randomly selected tweets, manually annotated to be either positive, negative, or neutral in opinion. From this comparison, we

Measure	Obama	GOP	Romney	Liberal	Limbaugh
PCC w/ User Activity, $PCC(\delta_i(t), A_i(t))$	0.15	0.08	0.12	0.02	0.18
PCC w/ User Activity Change $PCC(\delta_i(t), C_i(t))$	0.52	0.47	0.51	0.47	0.48
PCC w/ Neighbors' Activity Change $PCC(\delta_i(t), C_{N_i}(t)$	0.23	0.17	0.24	0.16	0.26
Critical Value (0.05 probability)	0.01	0.02	0.03	0.03	0.03

TABLE II: Pearson Correlation Coefficient between changes in topic-specific user activity, and changes in sentiment

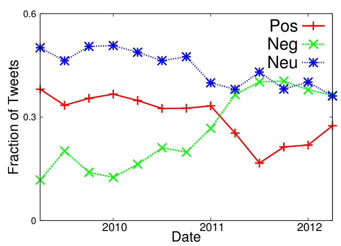


Fig. 1: Percentage of tweets classified as positive, negative, and neutral for topic "Obama".

found an overall accuracy of 85.6%, slightly higher than the accuracy found in [8], emphasizing the utility of combining both hashtag-level and tweet-level sentiment information for sentiment classification.

IV. Sentiments analysis over time

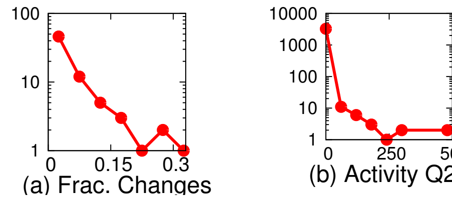
In this section we analyze user sentiments over time and their relationship to acceleration of topic-wise activity. Given our crawled and sentiment-annotated Twitter networks per topic, we analyze the sentiment evolution for the period from 2009 to the beginning of 2012 at a time scale of a quarter (3 months). Each included user is associated with a time series of sentiment states $x_i(t), x_i(t) \in \{Pos, Neu, Neg\}$ and another time series of topic-wise activity levels $A_i(t)$.

Figure 1 shows the fraction of users' tweets of each sentiment over time for the *obama* topic. The percentage of positive messages has a peak in the beginning of 2009, and the percentage of negative tweets has increased since, surpassing the positive in late 2011. These general trends mirror those seen in user polls and statistics[1], helping to further confirm the effectiveness of the sentiment classification.

Next, we turn our attention to user activity analysis. Let $A_i(t)$ denote the number of topic-based messages a user u_i posts in time period (quarter) t. The relative difference (acceleration) of user u_i's topic-based messages per quarter t is defined as $C_i(t) = |A_i(t-1) - A_i(t)|/max(A_i(t-1), A_i(t))$. User u_i's neighborhood relative acceleration $C_{N(u_i)}(T)$ is similarly measured as:

$$C_{N(u_i)}(t) = \frac{|\sum_{u_j \in N(u_i)} |A_j(t-1)| - \sum_{u_j \in N(u_i)} |A_j(t)||}{\max(\sum_{u_j \in N(u_i)} |A_j(t-1)|, \sum_{u_j \in N(u_i)} |A_j(t)|)}$$

[1]Gallup Polls: http://gallup.com/poll/113980/gallup-daily-obama-job-approval.aspx



(a) Frac. Changes (b) Activity Q20

Fig. 2: Distributions of fraction of users changing opinion in all time periods (a) and user activity (b) for quarter 20 across all topics. Merely 5% of the users change their opinion in Q20.

where $N(u_i)$ is the set of neighbors of user u_i. Let also $\delta_i(t)$ denote the binary time series taking a value 1 if $x_i(t) = x_i(t-1)$ (i.e. the user changes sentiment at time t) and 0 otherwise.

Table II shows the Pearson Correlation Coefficient (PCC) among $\delta(t), A(t), C(t)$ and $C_N(t)$ averaged across all users in our topics of interest. All three user activity series correlate significantly (as compared to the corresponding critical values shown in the last row). Two random variables (with no correlation) would have a 95% probability of PCC greater than a critical value or lower i.e. the critical represents the score at which the null hypothesis (random variables with no correlation) could be rejected with 95% probability. We observe a high correlation between the change in user activity and change in sentiment, with the PCC value on the topic of *obama* over 50 times higher than expected at random. On the other hand, the correlation of activity level ($A_i(t)$) and sentiment change $\delta_i(t)$ per topic is much weaker (close to random). This corroborates our hypothesis that the *change* in activity levels relates to a higher susceptibility to opinion change. The neighborhood acceleration is also significantly (well separated from random) correlated with sentiment change across all topics.

We also analyze the observed fraction of users who change sentiment at a given time. Figure 6(a) shows the distribution of the fraction of users changing sentiment per quarter (period of 3 months) in all topics. The distribution decreases exponentially with most quarters observing few changed opinions. The average fraction of users who change opinion within a quarter is on average less than 5% of all active users. User-specific activity levels $A_i(t)$ within a quarter also exhibit a skewed distribution, with a large fraction of "inert" non-active users in any given period. Figure 6(b) summarizes the distribution of activity levels across topics in the 20th quarter of our dataset (the quarters observe similar behavior).

V. Activity-based model

As we saw in the previous section, activity acceleration (change) is correlated with sentiment change. Classical network models for opinion formation rely on the assumption that all nodes are active at every step of the opinion dynamics process. As we demonstrated in the previous section (see Fig. 6), this is an unrealistic assumption for real-world online social media such as Twitter, where very few users are actively participating (change opinion) in the sentiment formation process at a given time. By incorporating the notion of inactive users in existing opinion dynamics models, we first demonstrate that converging to opinion consensus may be significantly slowed and, in general, consensus may not be guaranteed. Hence, we develop and evaluate a model that predicts realistic short-term opinion dynamics based on user activity acceleration.

A. Activation-based opinion dynamics

Since we have access to a real network and actual activity and sentiment dynamics, we can model the probability of users "not participating" in the opinion formation process. In established models such as DeGroot's [2], a user makes a decision for their sentiment based on their neighborhood and/or previous sentiment states. In our model, a user u_i is further associated with an activation probability $a_i(t)$ at time t that controls the likelihood of participation in the opinion dynamics process at time t. First, we analyze the effect of such activation probability on the speed of convergence to consensus of the opinion dynamics process, while in the second part of this section we tie the activation probability $a_i(t)$ to the acceleration of user activity $C_i(t)$.

In the following analysis, we augment DeGroot's [2] model with activation probability. The original model postulates that a node's opinion $x_i(t+1)$ at time step $t+1$ depends only on the previous opinion of its neighbors $N_i(t-1)$, where the neighborhood includes the node itself. All neighbors affect a node according to the edge strength encoded in the stochastic adjacency matrix W: $x(t+1) = W * x(t)$, where $x(t)$ is the real-valued opinion vector at time t and W is the row-stochastic adjacency matrix.

Patterson and colleagues [3] show that a necessary and sufficient condition for consensus is that W is primitive, i.e. 1 is a simple eigenvalue of W and all other eigenvalues have a smaller than 1 absolute value. Another way to state this condition from a graph-theoretic point of view is that the graph corresponding to W needs to be strongly connected. The rate at which the opinion vector approaches consensus depends on the second eigenvalue of W. Let the deviation from consensus sentiment $\bar{x}(t)$ be defined as the difference vector of every node's sentiment state from the average sentiment state at time t, i.e. $\bar{x}_i(t) = x_i(t) - \tilde{x}(t)$. Then an ϵ consensus is defined as reaching a time t such that $||\bar{x}(t)||/||\bar{x}(0)|| < \epsilon$. The number of rounds required to reach ϵ-consensus for all active nodes is $\frac{\log \epsilon}{\log \lambda_2(W)}$ [3].

We extend the basic model by incorporating an activation probability $a_i(t)$ for each node and time capturing the likelihood of this node participating in the opinion dynamics process:

$$x(t+1) = diag(a(t))Wx(t) + (I - diag(a(t)))x(t),$$

where $diag(a(t))$ is a diagonal matrix holding the activation vector and I is the identity matrix. At time t a user either participates in the consensus process with probability $a_i(t)$ or remains with the same opinion with probability $1 - a_i(t)$. Admitting the activation probability to be 1 for all time steps and nodes is equivalent to adopting the basic DeGroot model. We focus on the cases in which the $a(t) \neq \bar{1}$, and evaluate its effect on convergence and efficiency.

We can rewrite the activation model as

$$x(t+1) = [diag(a(t))W + I - diag(a(t))]x(t) = W'(t)x(t),$$

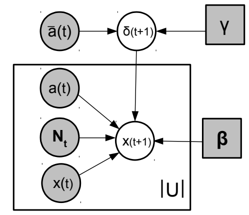


Fig. 3: Figure representing the relationships between the introduced model's variables and parameters using plate notation.

where $W'(t) = diag(a(t))W + I - diag(a(t))$ is the time-dependent adjacency matrix incorporating activation. In the case of constant activation probability for all nodes and time steps, i.e. $0 < a_i(t) = \alpha \leq 1$, W' is primitive (or equivalently corresponds to a strongly connected graph) if and only if W is primitive. As a result, the introduction of constant non-zero activity probability does not affect the existence of *eventual* consensus.

We also show that there is a significant overhead in the time to reach consensus due to "inactivity". We define the overhead of reaching an ϵ consensus o_α, as the ratio of number of steps to reach consensus in our activation based model with (i.e. $\alpha \leq 1$) and the number of steps required if all nodes are active $\alpha = 1$ (i.e. the original DeGroot model). We can show the following relationship between the overhead and the activation α:

Theorem 1. *[Activation Consensus Overhead] For a fixed activation $0 < a_i(t) = \alpha \leq 1$, the overhead to reach ϵ consensus is $o_\alpha = \frac{log(\lambda_2(W))}{log(\alpha\lambda_2(W)+1-\alpha)}$, where $\lambda_2(W)$ is the second eigenvalue of the original adjacency matrix W.*

Proof: For $a_i(t) = \alpha, 0 < \alpha \leq 1$, W' simplifies to $W'(t) = \alpha W + (1-\alpha)I$. Note, that the uniform activity assumption leads to W' being a scaled (by α) and translated (by $1-\alpha$) version of W. For such transformations, the eigenvalues of the two matrices are shifted and scaled in the same manner, i.e.

$$\lambda(\alpha W + (1-\alpha)I) = \alpha\lambda(W) + 1 - \alpha. \tag{1}$$

Then, using 1, we can derive the overhead as:

$$o_\alpha = \frac{\log \epsilon}{\log \lambda_2(W)} / \frac{\log \epsilon}{\log \lambda_2(W')} = \frac{log(\lambda_2(W))}{log(\alpha\lambda_2(W) + 1 - \alpha)}.$$

■

For fixed second eigenvalue $0 < |\lambda_2(W)| < 1$, the overhead o_α increases exponentially for decreasing activation probability α, meaning that the consensus rate slows down very fast with decreasing α. Intuitively, this is due to the fewer number of nodes behaving according to the original DeGroot model.

While we can show convergence and bound the overhead for a fixed activation probability, this is not possible for a general activation that changes per user and across time. In fact, one can design a trivial schedule of $a(t)$ such that convergence is never achieved, by keeping all nodes inactive, or deactivating the prevailing sentiment nodes' neighbors just before ϵ consensus is reached. In our real-world data, very few nodes tend to update their sentiment/opinion at every round. In addition, the activation is related to the topic-wise activity acceleration which may change arbitrary at every time step and, in a real-world network, may be affected by external channels (traditional media, personal information/opinion exchange, etc). As a result, we study the practical utility of our model for predicting future sentiment changes in short intervals, as opposed to stationary behavior upon convergence.

B. Model for prediction of short-term sentiment change

In what follows, we adapt our activity-based model for short-term prediction of sentiment change in Twitter. We consider a discrete state space (as opposed to continuous and without loss of generality) and tie the activation probability to the acceleration of topic-based posts. Since sentiment states (positive, neutral, negative) are not equally likely, we learn their effect in a node's neighborhood from past data. Following the observation that activity and sentiment changes are related, we now turn to instantiating our activation model such that it captures this observation in real data. As we discussed earlier, we make the Markovian assumption that the next sentiment state of a user depends only on the current state of the network and the activation probability, governed by the acceleration of topic based activity levels.

Our generative model updates the opinion of $\delta(t+1)$ users that are most likely to change their sentiments $x_i(t+1)$ at time $t+1$. An overall figure displaying the relationships between the parameters and variables within our model, for each network user $u_i \in U$ can be seen in Figure 3. We describe in detail each portion of this model in what follows.

Number of Changes. Within our generative model, the variable $\delta(t)$ represents the expected number of changed opinions likely to occur in the graph within a time step. For our real data, we train this value using linear regression and the average activation of users at the previous time step $\bar{a}(t)$. The linear dependence of $\delta(t)$ is controlled by the parameters γ (i.e. $\delta(t) = \gamma_1 \bar{a}(t-1) + \gamma_2$). Dependence on more signals from the previous time stamp can be considered, but for out Twitter data using the average activation from the previous period resulted in the best predictive accuracy (see experiments).

Updating the opinion of a single user. To produce each individual's next state within the evolving opinion network $x(t+1)$, a two-step process is employed. First, a user is randomly selected, according to their activation probability $a(t)$, to change their opinion. Second, the next state opinion of this user, $x_i(t+1)$ is selected using the current state of the node, $x_i(t)$, and a log-linear model parametrized by 1) the current states of neighbors in its neighborhood ($N(t)$) and 2) a vector of coefficients β.

Consider a network of users $U = u_i, |U| = n$ with m possible user states $X = x_i, |X| = m$ and a user u_i's current state at time t denoted $x_i(t)$. Recall that $A_i(t)$ is the number of topic-specific messages sent by user u_i at time t while her activity acceleration is denoted by $C_t(u_i)$. We define the activation probability as $a_i(t) = C_t(u_i)/\sum_{u_k \in U} C_t(u_k)$. If user u_i is selected to have her state change, we build a feature vector f_i, representing the current states of the users in its neighborhood $N(u_i)$ counting the frequency neighbor sentiments state:

$$f_i(j) = |\{u_k \in N(u_i) \wedge x_t(u_k) = x_j\}|, \forall j = 1..m$$

Similar frequency histograms have been used in previous graph models, such as majority voting models [13], which update the new state based on the most frequent previous state, or value-weighted voting models [1], which use a weighted combination of the histogram values. In our case, we choose to combine the histogram values (similar to the value-weighted voting models), adopting a log-linear model to estimate the probability of each next state:

$$Pr_t(x_t(u_i) = x_j) = \frac{1}{Z} e^{\beta_j \cdot f_i(j)}, j = 1..m,$$

where β is a coefficient vector parameter, and $Z = \sum_{j=1}^{m} e^{\beta_j \cdot f_i(j)}, j = 1..m$ is a normalization factor. This model is also known as logistic regression, and is commonly used to predict categorical variables (in our case, the user opinions) based on continuous predictor variables (the number of opinions among neighboring users).

VI. Experimental evaluation

To evaluate the effectiveness of our activity-based model, we implement its algorithm in Java, fitting it to our Twitter dataset. To obtain the β coefficients, we regress using a training set spanning August 2006 to December 2011. We then predict future opinion evolution for two separate quarters (Q1 2012: Jan–Mar 2012, and Q2 2012: Apr–June 2012). For both time periods, their previous quarter's network snapshot was obtained as a starting point, and the introduced model used to generate new sentiment dynamics upon it.

We compare our predictions with those obtained from the recently introduced opinion prediction model by Kimura et al. [1] which utilizes a modified voter model of opinion dynamics. To allow for fare comparison of accuracy between these two models and the actual network evolution, the precise number of user sentiment changes $\delta(t)$ is used for each quarter, and both models are separately run until this number of changes is produced. On average ~ 200 changes per quarter were observed.

Predicting users who would change sentiment. Before comparing prediction of future sentiment states, we compare the accuracy of predicting which nodes are most likely to change their sentiment. Figure 4 shows the ROC50 curves obtained by choosing the top-k most likely users to change their opinion. We compare the users predicted by either our Acceleration-based model (the users with the highest acceleration, $C_i(t)$), the Kimura model and a random baseline (selecting random users). The predicted users are compared to the actual set of users who change their opinions. The ROC50 curve (which continues until the first 50 false positives) is used rather than the full ROC curve, as it is generally viewed as a more useful measure in cases where the true negatives greatly outnumber the true positives [18]. As can be seen from Figure 4, activity acceleration is a much better predictor than the Kimura's model and random user selection.

Future sentiment prediction. Next, we evaluate the prediction of next sentiment states again in comparison to Kimura's approach and a random model. The random model selects

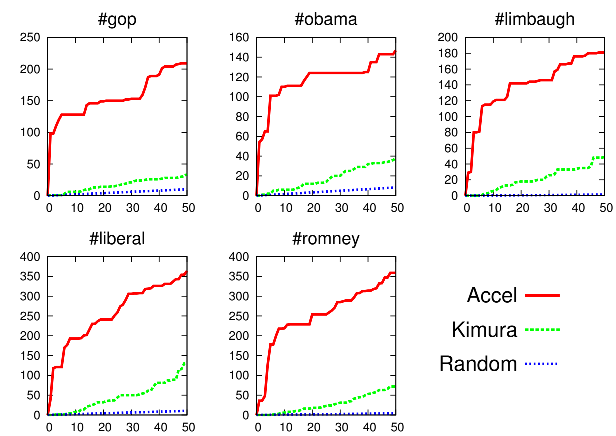


Fig. 4: ROC50 curves (TP as a function of FP) for prediction of users who change opinion according to our Acceleration model (Accel), the Kimura approach and Random user selection.

uniformly a node to change opinion and it's new opinion state. For the sake of fair comparison the number of changes in each quarter, $\delta(t)$, is assumed to be given (as the Kimura approach has no method for estimation of this value). We report results on learning $\delta(t)$ in the following subsection.

Figure 5 shows the results for running the competing models on the five different topics across two separate quarters. "Global Accuracy" at each timestep is defined as the percentage of correctly predicted users sentiments for the whole graph:

$$Acc_{global} = \frac{|\{x_{t+1}(u_i)|x_{m,t+1}(u_i) = x_{g,t+1}(u_i)\}|}{|\{x_{g,t+1}(u_i)\}|},$$

where $x_{m,t+1}(u_i)$ is the predicted sentiment state, at time $t+1$, for user u_i using model m, and $x_{g,t+1}(u_i)$ is the actual sentiment state, at time $t+1$, for user u_i from the original Twitter graph.

Since only a small fraction of the network changes opinion in each quarter on average (see Fig 6), we also analyze the accuracy of changes made by each model. These "Accuracy of Changes" are computed as the percentage of changes by model m that switched users to a correct new opinion:

$$Acc_{changed} = \frac{|\{y_{m,t+1}(u_i) = x_{g,t+1}(u_i)\}|}{\delta_{t+1}}$$

where $y_{m,t+1}(u_i)$ are the sentiments of nodes whose opinions actually changed from time t to $t+1$, using model m.

As can be seen from Figure 5, our introduced model ("Activity Model") has a high accuracy in predicting Global future sentiment evolution, consistently outperforming both Random and Kimura's model. The average of Acc_{global} across the topics and quarters is 94%, a significant increase when compared to an average Overall Accuracy of 86% for the Kimura, and 80% for the Random. In addition, the accuracy of changes made by our Activity model far outperform the alternatives, having an average accuracy of 50%, as opposed to 18% for the Kimura and 7.5% for Random. In fact, the accuracy of opinion changes in the Activity model are up to 87 times higher than that of Kimura (in topic limbaugh for Q2), and 110 times higher than that expected at random, (for topic romney, in Q2). This highly significant increase in the quality of prediction results helps to emphasize the effectiveness of user activity acceleration as an indicator for user opinion change.

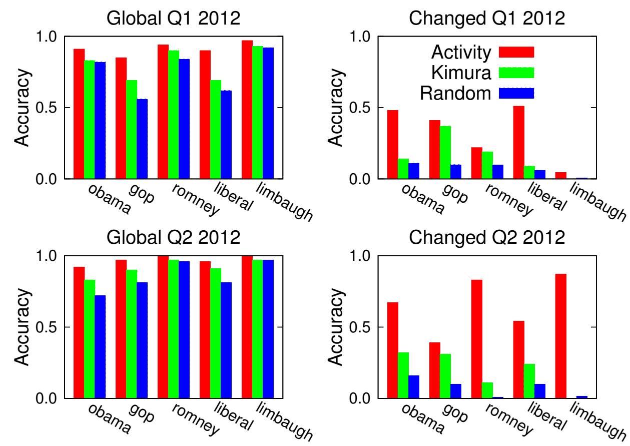


Fig. 5: Comparison of models across all nodes (Global Accuracy left), and only for nodes who change sentiment (Accuracy of Changes right) in Q1 and Q2 of 2012.

Figure 6 compares the overall per-quarter L1 error for negative and positive predicted fractions with actual values acriss topics. The activity model dominates Kimura, except for an overestimation of positives in gop at the expense of neutral.

Changed sentiment δ_t estimation. In the sentiment prediction experiments we assumed that the actual number of users who change opinion δ_t at time t is known. This assumption ensures a fair comparison to the Kimura model which does not have a mechanism of estimating this number. However, both when predicting future states of real networks or when generating sentiment dynamics according to our model, one needs to estimate δ_t. We propose to learn its dependence on the previous state of the network. In general, one can use different characteristics of the network's previous state including distribution of activity levels $A(t-1)$, distribution of acceleration $C(t-1)$, previous fraction of changed opinions, etc. For our data, we experimented with learning a linear regression model for δ_t based on past data on average activity $\bar{A}(t-1)$ and acceleration $\bar{C}(t-1)$. A regression model using acceleration $\bar{C}(t-1)$ results in 20% lower root-mean squared error than regressing on raw activity levels $\bar{A}(t-1)$ (0.0403 v.s. 0.032 respectively). A regression model based on both variables does not decrease the error further. Hence, we utilize a linear regression model (parametrized by γ) for δ_t as a function of the acceleration $\bar{C}(t-1)$. For other datasets, one can turn to more complex models to learn the fraction of changed opinions.

Simulating sentiment dynamics. Apart from predicting future sentiment dynamics for real networks of known user activity, our model can also be employed for generating realistic dynamics in synthetic networks. We study the model's behavior under different activation regimes assuming independent randomized node activation. Given a fixed network structure, we generate activation probabilities according to a truncated exponential distribution $a(t) \sim E_{tr}(\lambda), P[a_i(t) > 1] = 0$ with an average activation $\bar{a}(t) = 1/\lambda$. In order to control for the

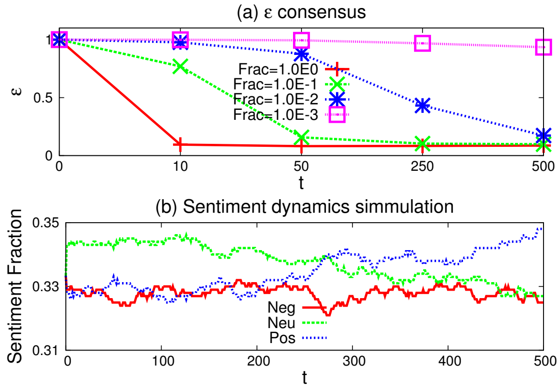


Fig. 7: Comparison of negative node graph evolution across one year for our Activity Model vs. the actual graph

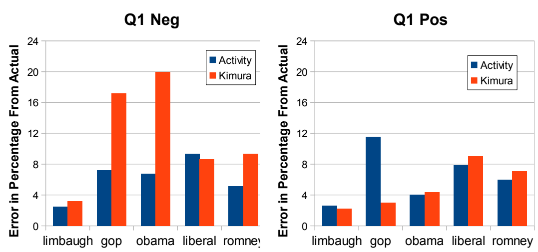


Fig. 6: L1 error from the actual fraction of negative (left) and positive (right) nodes in obama for Q1 2012.

fraction of activated users at any given time, we set $\gamma = [|U|, 0]$ and hence the activated fraction $Frac = \delta/|U| = 1/\lambda$ is exponentially distributed with the same average ($|U|$ is the number of nodes in the network). We use three states (Pos,Neu,Neg) with symmetric transition factors β and uniform representation of the states at time 0.

Figure 7(a) presents the average progress to sentiment consensus ϵ. Recall that progress is tracked as $||\bar{x}(t)||/||\bar{x}(t)||$ and lower values correspond to being closer to consensus. The different traces correspond to different expected fractions of activated users (Frac). We average the behavior of 50 sentiment dynamics evolution instantiations in 1000-node random preferential attachment networks. When all users are expected to be active (Frac=1.0E0), consensus is reached in as few as 10 time steps. For decreased fractions of activated users (between 1.0E0 and 1.0E-2) the overhead for reaching consensus increases, while when only a few nodes are expected to change opinion at a time (1.0E-3) consensus is not reached within 500 time steps. Figure 7(b) shows an example sentiment dynamics generation trace for the regime Frac=1.0E-3.

VII. Conclusion

In this paper, we proposed a novel user activity-based model for topic-specific sentiment evolution in social media networks. We analyzed a large Twitter dataset and discovered a significant connection between user activity acceleration and changes in user sentiment (correlations 50 times higher than expected at random). Following these observations, we introduced a novel activity-based generative and predictive model for opinion dynamics. We generalized existing models by incorporating the possibility of inactive users and showed both theoretically and in simulation that opinion consensus is reached slower due to inactivity. When employed to predict future user sentiment evolution, our model's prediction accuracy dominates (by 1 order of magnitude) that of recently introduced alternatives. It is capable of predicting the next sentiment state with 90% (averaged over all users) and 50% (averaged over hard-to-predict users who change opinion) accuracy. These results help confirm the utility of user activity acceleration as an important precursor to sentiment change, and the ability of our model to capture the sentiment dynamics within real-world social media networks.

Acknowledgements This work was supported by the Institute for Collaborative Biotechnologies through grant W911NF-09-0001 from the U.S. Army Research Office. The content of the information does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

References

[1] M. Kimura, K. Saito, K. Ohara, and H. Motoda, "Learning to predict opinion share in social networks," in *AAAI*, 2010.

[2] M. H. Degroot, "Reaching a Consensus," *Journal of the American Statistical Association*, vol. 69, no. 345, pp. 118–121, 1974.

[3] S. Patterson and B. Bamieh, "Interaction-driven opinion dynamics in online social networks," in *SOMA*, 2010, pp. 98–105.

[4] G. Henry and C. Gordon, "Tracking issue attention: Specifying the dynamics of the public agenda," *Public Opinion Quarterly*, 2001.

[5] R. Neuman, "The threshold of public attention," *Public Opinion Quarterly*, vol. 54, no. 2, pp. 159 – 176, 1990.

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

[7] L. Jiang, M. Yu, M. Zhou, X. Liu, and T. Zhao, "Target-dependent twitter sentiment classification," in *HLT*, 2011, pp. 151–160.

[8] S. L. Mudhakar Srivatsa and T. Abdelzaher, "Mining diverse opinions," in *MILCOM*, 2012, pp. 1–7.

[9] K. Macropol, P. Bogdanov, A. K. Singh, L. Petzold, and X. Yan, "I act, therefore i judge: Network sentiment dynamics based on user activity change (supplemental material), http://cs.ucsb.edu/%7edbl/papers/sentimentappendix.pdf," 2013.

[10] J. D. Farley and G. Markowsky, "Evolutionary dynamics of revolution: Opinion modeling in society based on graphs and rules," in *IDAACS (2)*, 2011, pp. 780–784.

[11] C. Qian, J. Cao, J. Lu, and J. Kurths, "Adaptive bridge control strategy for opinion evolution on social networks." *Chaos*, 2011.

[12] J. Xie, J. Emenheiser, M. Kirby, S. Sreenivasan, B. K. Szymanski, and G. Korniss, "Evolution of opinions on social networks in the presence of competing committed groups," *PLoS ONE*, 2012.

[13] E. Even-Dar and A. Shapira, "A note on maximizing the spread of influence in social networks," *Inf. Process. Lett.*, vol. 111, no. 4, pp. 184–187, 2011.

[14] C. Wilson, B. Boe, A. Sala, K. P. Puttaswamy, and B. Y. Zhao, "User interactions in social networks and their implications," in *EuroSys*, 2009.

[15] A. Ahmed, Y. Low, M. Aly, V. Josifovski, and A. J. Smola, "Scalable distributed inference of dynamic user interests for behavioral targeting," in *KDD*, 2011, pp. 114–122.

[16] X. Wang, F. Wei, X. Liu, M. Zhou, and M. Zhang, "Topic sentiment analysis in twitter: a graph-based hashtag sentiment classification approach," in *CIKM*, 2011, pp. 1031–1040.

[17] S.-M. Kim and E. Hovy, "Determining the sentiment of opinions," in *Proc. of the 20th Intl. Conf. on Computational Linguistics*, 2004.

[18] M. Gribskov and N. Robinson, "Use of receiver operating characteristic (roc) analysis to evaluate sequence matching," *Computational Chemistry*, vol. 20, pp. 25 – 33, 1996.

I act, therefore I judge: Network sentiment dynamics based on user activity change

Supplemental Material

Kathy Macropol
Arcadia University
macropolk@arcadia.edu

Petko Bogdanov, Ambuj K. Singh, Linda Petzold, Xifeng Yan
University of California Santa Barbara
{petko,ambuj,petzold,xyan}@cs.ucsb.edu

I. ANNOTATING INDIVIDUAL TWEETS

We annotate individual posts following the approach in [1], [2]. Tweets are grouped by topic based on included *topic hashtags*. For example, tweets relating to the topic of president *Barack Obama* contain the hashtag #obama within them. The topic-related tweets often contain other hashtags which we assign a preliminary sentiment probability (positive and negative) using the Multinomial Naive Bayes classifier method introduced in [1]. This classifier is trained on a document set consisting of positive / negative tweets discovered using manually collated sets of strongly positive and negative topic-based hashtags (for example "#voteForObama" or "#impeachObama"). The results from the classifier give the sentiment probabilities on a per-tweet basis (in other words $Pr(t = s)$, the probability that tweet t has sentiment s), but may be extended to the contained hashtags in order to obtain the probability of any hashtag h, having sentiment s using the following equation:

$$Pr(h = s) = \frac{\sum_{t \in T_h} Pr(t = s)}{|T_h|}, \quad (1)$$

where T_h is the set of all topic-based tweets which contain the hashtag h within them.

The hashtag-level sentiments are obtained using the iterative technique introduced in [2]. Initial sentiments for a hashtag are fixed using the previously mentioned manually collated set of positive and negative seed hashtags, as well as the probabilities calculated using Equation 1 for all other hashtags. Then, an iterative process based on Relaxation Labeling (RL) from [2] is utilized to update co-occurring hashtags' sentiment values to all hashtags (except the manually collated fixed hashtags) until convergence.

We compute the final sentiment annotation (positive, negative, or neutral) for each tweet as a linear combination of the polarities of its contained hashtags. We apply a sentiment threshold rendering tweets with sentiments exceeding $|0.75|$ as positive or negative respectively and the rest as neutral.

II. DISCUSSION AND FUTURE EXTENSIONS

For our analysis, we assume a set of user states, $X = \{P, N, R\}$, indicating whether a user has a positive, negative, or neutral (Pos,Neg,Neu) sentiment on a particular topic. However, this model can extend to work with larger sets of states. In addition, our model can be fit to a particular network by finding the unknown coefficients, β using multinomial logistic regression on historical data. This allows for the estimation of expected number of changes and next user opinion given the current network state. The parameter γ can be similarly learned using linear regression.

In this model, we ignored the influences of media or individuals outside the network on influencing users' opinions. Taking into account these external factors, however, may lead to more realistic and accurate models in the future. Additionally, previous research has shown that "interaction networks", composed of links between individuals who interact together, can be better indicators for user relationship and analysis than the following networks of users [3]. Further work could look at the connection between interaction, as opposed to following, relationships, and sentiment change, integrating this into an extended generative model.

III. MORE RELATED WORK

Here we discuss a longer list of message classification methods. They have either focused specifically on Twitter [4], [5], [6], [2], [7] or on other social media websites [8], [9], [10], [11], [1]. In particular, the work of Tan et al. [5] introduced multiple related Twitter sentiment classification approaches, each of which combined message-level processing with a hashtag-graph based approach. By utilizing the fact that hashtags contained within the same tweets tend to indicate similar sentiments, they were able to obtain methods with accuracies of almost 80% while predicting sentiment within tweets. In addition, it has been found that training classifiers in a topic-dependent manner can further increase their performance during sentiment classification, as well [7]. Our approach is complementary to all the above as it focuses on the dynamics of sentiments as opposed to determining the sentiment state of a specific user/message.

There have also been many recent models in the area of opinion dynamics and information spread. Influence models such as Independent Cascade and Linear Threshold have been applied to real-world data [12], [13], [14], [15].

REFERENCES

[1] S. L. Mudhakar Srivatsa and T. Abdelzaher, "Mining diverse opinions," in *MILCOM*, 2012, pp. 1–7.

[2] X. Wang, F. Wei, X. Liu, M. Zhou, and M. Zhang, “Topic sentiment analysis in twitter: a graph-based hashtag sentiment classification approach,” in *CIKM*, 2011, pp. 1031–1040.

[3] C. Wilson, B. Boe, A. Sala, K. P. Puttaswamy, and B. Y. Zhao, “User interactions in social networks and their implications,” in *EuroSys*, 2009.

[4] A. Go, R. Bhayani, and L. Huang, “Twitter sentiment classification using distant supervision,” *Processing*, vol. 150, no. 12, pp. 1–6, 2009.

[5] C. Tan, L. Lee, J. Tang, L. Jiang, M. Zhou, and P. Li, “User-level sentiment analysis incorporating social networks,” in *KDD*, 2011.

[6] R. Mehta, D. Mehta, D. Chheda, C. Shah, and P. Chawan, “Sentiment analysis and influence tracking using twitter,” *J. of Adv. Research in CS and EE*, vol. 1, no. 2, pp. 72 – 79, 2012.

[7] L. Jiang, M. Yu, M. Zhou, X. Liu, and T. Zhao, “Target-dependent twitter sentiment classification,” in *HLT*, 2011, pp. 151–160.

[8] B. Pang, L. Lee, and S. Vaithyanathan, “Thumbs up?: sentiment classification using machine learning techniques,” in *EMNLP*, 2002.

[9] H. Cui, V. Mittal, and M. Datar, “Comparative experiments on sentiment classification for online product reviews,” in *AAAI*, 2006.

[10] B. Pang and L. Lee, “Opinion mining and sentiment analysis,” *Found. Trends Inf. Retr.*, vol. 2, no. 1-2, pp. 1–135, Jan. 2008.

[11] X. Glorot, A. Bordes, and Y. Bengio, “Domain adaptation for large-scale sentiment classification: A deep learning approach,” in *ICML*, 2011.

[12] D. Kempe, J. Kleinberg, and E. Tardos, “Maximizing the spread of influence through a social network,” in *KDD*, 2003, pp. 137–146.

[13] M. Kimura, K. Saito, R. Nakano, and H. Motoda, “Extracting influential nodes on a social network for information diffusion,” *Data Min. Knowl. Discov.*, vol. 20, no. 1, pp. 70–97, Jan. 2010.

[14] W. Chen, Y. Wang, and S. Yang, “Efficient influence maximization in social networks,” in *KDD*, 2009, pp. 199–208.

[15] Z.-T. Zhu, J. Zhou, and X.-G. Chen, “Opinion evolution in different groups of a mega-project,” in *ICNC*, 2011, pp. 2188–2193.