Predicting Online Protest Participation of Social Media Users

Suhas Ranganath
srangan8@asu.edu
Arizona State University

Fred Morstatter
Fred.Morstatter@asu.edu
Arizona State University

Xia Hu
hu@cse.tamu.edu
Texas A&M University

Jiliang Tang
jlt@yahoo-inc.com
Yahoo Labs

Suhang Wang
swang187@asu.edu
Arizona State University

Huan Liu
Huan.Liu@asu.edu
Arizona State University

Abstract

Social media has emerged to be a popular platform for people to express their viewpoints on political protests like the Arab Spring. Millions of people use social media to communicate and mobilize their viewpoints on protests. Hence, it is a valuable tool for organizing social movements. However, the mechanisms by which protest affects the population is not known, making it difficult to estimate the number of protestors. In this paper, we are inspired by sociological theories of protest participation and propose a framework to predict from the user's past status messages and interactions whether the next post of the user will be a declaration of protest. Drawing concepts from these theories, we model the interplay between the user's status messages and messages interacting with him over time and predict whether the next post of the user will be a declaration of protest. We evaluate the framework using data from the social media platform Twitter on protests during the recent Nigerian elections and demonstrate that it can effectively predict whether the next post of a user is a declaration of protest.

Social media has emerged as a popular information and communication channel for protest-related issues (Muthiah et al. 2015; Contractor et al. 2015; Tufekci and Wilson 2012). It provides an open and accessible platform for people to put forth views on issues affecting them. Millions of people, therefore, use social media to declare protest, mobilize opinion and participate in discussions on these issues. For example, electoral malpractice was suspected during the recent elections in Nigeria (Mark 2014), and users employed social media to express and mobilize viewpoints. Owing to this, social media can be used by protest organizers to recruit potential members.

We concentrate on posting behavior of users and define online protest participation as the act of declaring protest through status messages. However, it is a nontrivial problem to directly predict online protest participation of social media users. Millions of users post in social media during popular social movements, and protest organizers have to go through them to identify potential participants. The mechanisms by which the protest affects the population cannot be fully observed (Lin et al. 2013), and protest organizers cannot easily estimate the number of protestors. Designing algorithms to predict whether the next post of a user will be a declaration of protest will enable protest organizers to anticipate the behavior of the protestors and estimate the total number of participants.

This task faces several challenges. First, the ways in which protest-related events affect a person are not observable, resulting in a lack of knowledge of factors operating at that time causing his next post to be a declaration of protest. Second, a user is subject to various types of influence in his past, and many of them are in conflict with each other. This may lead to ambiguities on whether his posts will contain declarations of protest in the future. Finally, each user can post a large amount of content and interact with many people, leading to issues of scalability.

Sociological studies have theorized factors from an individual's history causing his next post to be a declaration of protest. A user will be more likely to protest if his social ties have reached out to him with protest-related messages in his past (Schussman and Soule 2005). The chances of him protesting are bolstered if these messages are sent by people interested in protest related issues (Lim 2008). The likelihood of protest is reduced if people uninterested in protest-related issues have reached out to him in the past with messages unrelated to protests. (Snow, Zurcher Jr, and Ekland-Olson 1980).

Inspired by these sociological theories, we utilize the user's previous status messages and messages interacting with him to predict whether his next post will be a declaration of protest. To model the effect of the interactions on the user's status messages, we build upon concepts from the Brownian motion theory of fluid particle motion (Zhou 2003). This theory models the path of fluid particles as other fluid particles come in contact with them over time. We draw analogies from these concepts to model the probability of the user declaring protest as other users reach out to him over time.

The primary contributions of this work are:

- Formally defining the problem of predicting whether the next post of a user will be a declaration of protest based on past status messages and messages interacting with him,

- Demonstrating the applicability of sociological theories of protest participation in online social media data,

- Proposing a framework that predicts online protest participation of a user, and

- Evaluating the framework using a real world dataset from Twitter on the Nigerian election protest in March 2015.
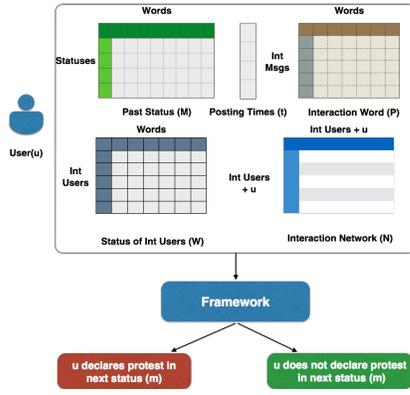
Figure 1: The proposed framework to predict if the next post of user u is a declaration of protest

Problem Statement

In this section, we present the notations and formally state the problem. Boldface uppercase letters (e.g., $\mathbf{X}$) denote matrices and lowercase letters (e.g., $\mathbf{x}$) denote vectors. $\mathbf{X}_{ij}$ signifies the element in the ith row and j^{th} column of matrix $\mathbf{X}$. $\mathbf{x}_i$ denotes the ith element of vector $\mathbf{x}$. The ith column of a matrix $\mathbf{X}$ is denoted by $\mathbf{X}_i$ and its ith row by $\mathbf{X}_{i'}$. The Frebonius norm of a matrix $\mathbf{X}$ is denoted as $||\mathbf{X}||_F = \sqrt{\sum_{i,j} \mathbf{X}_{ij}^2}$.

The terms related to proposed framework are illustrated in Fig 1. Let the set of candidate users be denoted as $\mathcal{U}$. For each user $u \in \mathcal{U}$, let $\mathbf{t}$ contain the posting times of a set of status messages denoted by $\mathcal{M}$. We construct a dictionary of words in the status messages of candidate users, messages that reach out to them and status messages of users posting them. Previous status messages of candidate users are contained in the matrix $\mathbf{M} \in \mathbb{R}^{l \times m}$. Here l is the number of status messages, and m is the total number of words in them. Messages reaching out to u are contained in the matrix $\mathbf{P} \in \mathbb{R}^{o \times r}$. Here o is the total number of such posts and r is the total number of words in them. Similarly, the status messages of users posting such messages are contained in $\mathbf{W} \in \mathbb{R}^{o \times q}$. Here q is the total number of words in the status messages. Let the user u and users who post messages interacting with them constitute the set $\mathcal{G}$. $\mathbf{N} \in \mathbb{R}^{d \times d}$ denotes the interaction network between these users, where d is the total number of users, and

$$\mathbf{N}_{ij} = \begin{cases} t & g_i \text{ interacts with } g_j \ t \text{ times}; g_i, g_j \in \mathcal{G}, \\ 0 & \text{otherwise} \end{cases}$$

The problem can be formally defined as *"Given a user u, his posting time vector $\mathbf{t}$, his status message matrix $\mathbf{M}$, interaction word matrix $\mathbf{P}$, the word matrix of users posting these interactions $\mathbf{W}$ and the interaction network $\mathbf{N}$, predict if the next post of user u will be a declaration of protest."*

Data Analysis

In this section, we describe the data collection methods and present statistics related to the collected dataset. We then propose postulates from sociological theories of protest participation that form the conceptual basis of the framework and utilize the datasets to verify them.

The data is from the Nigerian general election that took place on March 2015. This event took place amid insurgencies of Boko Haram (Puig 2015) and alleged discrepancies in the voting process (Mark 2014). We collected messages from the Twitter Streaming API geotagged within Nigeria using (Kumar et al. 2011). Assisted by experts, we prepared a set of keywords and hashtags that denote declarations of protests to filter out tweets that did not contain any of these words. The earliest tweet in our dataset was posted on Feb. 25, 2015, and the last tweet is posted on Apr. 27, 2015. We collected a total of 2,686 posts potentially containing declarations of protest by users.

Posts containing protest-related hashtags and keywords do not necessarily mean the user is declaring protest. For example the post "That same guy you can't ever resist." contains the protest-related keyword "resist", but is not a declaration of protest. To address this discrepancy, we used Amazon Mechanical Turk to categorize posts. The posts were given to the workers along with instructions to determine if the user is declaring protest in the given post or not. Three distinct workers evaluated each post, and we employed majority voting to choose the category of the post. In total, the workers categorized 626 posts where the user expressed protest, which we refer to as positive labels. The workers categorized that 2,060 posts did not contain a declaration of protest, which we refer to as negative labels. We use these posts as ground truth for future experiments and users who post them form the set $\mathcal{U}$.

The Twitter Streaming API shows a bias in keyword and geo-tagged distributions (Morstatter et al. 2013), so the data might not represent the actual spread of protests. To overcome this, we treat each user as a separate entity and use his historical information to predict whether his next post will be a declaration of protest. For each user $u \in \mathcal{U}$, we collect their status messages, a maximum of 200 posts, and extract the set of users he mentioned to construct the set $\mathcal{G}$. As we are interested in predicting whether the post of u at time t is a declaration of protest, we do not explicitly denote if his previous posts are a declaration of protest. We then collect the status messages of the users in $\mathcal{G}$ and retain only those users who have mentioned u back. We then construct the matrices $\mathbf{M}$, $\mathbf{W}$, $\mathbf{P}$ and $\mathbf{N}$. In total, the dataset contains 63,983 users, around 11 million interactions and 105 million posts by users who mentioned the candidates.

Our first postulate is regarding the ability of appropriate social ties that can enhance protest participation (Lim 2008). It can be stated that "if the user u is previously mentioned in a protest related message by people interested in it, the next post of the user is likely to be a declaration of protest". We first denote the set of words $\mathcal{L}$ in the positive labels by the crowdsourced workers and denote them as protest-related words. We then select columns corresponding to these words from $\mathbf{P}$ and $\mathbf{W}$ and denote them as $\mathbf{A}$ and $\mathbf{B}$ respectively. To measure the extent that user is mentioned in protest-related messages by people interested in them, we compute $\sum_{i=1}^{P} \mathbf{A}_{i'} \mathbf{B}_{i'}^T$, where P is the total number of messages. The higher the number of protest-related messages by interested users mentioning u in his past, the higher will be this quantity. We repeat this for all the users with positive labels and a similar number of users with negative labels. A paired t-test shows that positively labeled users have a significantly higher score with $p < 0.01$ than negatively labeled users,

Parameter	Value
Candidate Posts	2,686
Protest Posts	626
Non Protest Posts	2,060
Candidate Users	2,686
Previous Posts of Candidate Users	362,485
Interacted Users	63,983
Interactions	11,976,235
Posts of Interacted Users	105,513,184

Table 1: Dataset of protests during Nigerian elections.

thus validating our first postulate.

Our next postulate is regarding the ties to alternative networks that can hinder protest participation (Snow, Zurcher Jr, and Ekland-Olson 1980). It can be stated as "the more the user is mentioned in a message unrelated to the protest by people not interested in it in his past, the lower the chance of his next post being a declaration of protest". We select the columns from $\mathbf{P}$ and $\mathbf{W}$ corresponding to words not present in $\mathcal{L}$ and denote the matrices as $\mathbf{C}$ and $\mathbf{D}$. To measure the extent that a user is mentioned in matters unrelated to protests by people not interested in protests, we compute $\sum_{i=1}^{P} \mathbf{C}_{i'}\mathbf{D}_{i'}^{T}$. The greater the number of unrelated messages u is mentioned in by uninterested users in his past, the higher will be this sum. We repeat this for all the users given positive labels and a similar number of users given negative labels. A paired t-test shows that negatively labeled users have a significantly higher score than positively labeled users with $p < 0.05$, thus validating our second postulate.

The Proposed Framework

In this section, we build upon the presented postulates to design a computational framework that predicts whether the next post of a user will be a declaration of protest. It accomplishes this by modeling the interplay between status messages and messages interacting with him in his past.

The matrices in Fig 1 have a large number of elements, making scalability an important issue. To improve scalability, we represent them in a shared latent dimensional space. Latent representations also give a semantic meaning to the interactions and the status messages, as each dimension can be interpreted as a cluster. From the first postulate, the probability of u posting in dimension i increases if a user interested in dimension i mentioned him in the past in a post related to i. From the second postulate, the probability decreases if he is mentioned in a post unrelated to dimension i in his past by a user with low interests in i. Posts related to other dimensions might affect his probability of posting in i. For instance, a post in the dimension related to politics might bolster the likelihood of the next post declaring protest, and a post associated with fashion might diminish it. We model these concepts to develop a framework that predicts whether the next post of the user u will be a declaration of protest.

We first construct a latent dimensional representation for the matrices $\mathbf{P}$ and $\mathbf{W}$ from Fig 1. The latent dimension representation of message k mentioning the user is given by $\mathbf{P}_{k'}\mathbf{U} = \mathbf{L}_{k'}$ and the content of the users posting those messages by $\mathbf{W}_{k'}\mathbf{V} = \mathbf{X}_{k'}$. Here $\mathbf{U} \in \mathbb{R}^{r \times I}$ and $\mathbf{V} \in \mathbb{R}^{q \times I}$ are latent dimension representations of the content of messages and users posting them respectively, with I dimensions.

Let $\mathbf{s}_i^M$ and $\mathbf{s}_i^{M-1}$ give the predicted latent dimension representation of the Mth and $M-1^{th}$ previous status messages of u in dimension i, ordered chronologically from the earliest post. Let k be a message mentioning the user u posted between his M^{th} and $M-1^{th}$ previous status message. The change in $\mathbf{s}_i^M$ due to k can be modeled as $\mathbf{L}_{ki}\mathbf{T}_{i'}\mathbf{X}_{k'}^{T}$, where $\mathbf{T}$ is the dimension correlation matrix. This quantity has a high value when a message related to dimension i is posted by a person with high interests in i mentions the user, modeling the first postulate. It has a low value when a message unrelated to dimension i is posted by a user with low interests in i, thus modeling the second postulate. It also increases or decreases when a message related to dimension j is posted by a person interested in j and dimension j has a high positive or negative correlation with i respectively. We represent change in $\mathbf{s}_i^M$ due to all posts mentioning u between the Mth and $M-1^{th}$ previous status message in the form of a differential equation.

$$\frac{d\mathbf{s}_i^M}{\mathbf{s}_i^M} = \mu_i(1 + \sum_{k \in \mathcal{P}_M} \mathbf{L}_{ki}\mathbf{T}_{i'}\mathbf{X}_{k'}^{T})dt + \sigma_i dW. \quad (1)$$

Here μ is the default change of $\mathbf{s}_i^M$ when there are no messages interacting with him. $\mathcal{P}_M$ is the total set of messages aggregated for model simplicity. W is the Wiener process to account stochastic variations. (Øksendal 2003). Let $\mathbf{n}_i = 1 + \sum_{k \in \mathcal{P}_M} \mathbf{L}_{ki}\mathbf{T}_{i'}\mathbf{X}_{k'}^{T}$.

The intuition for Eq. 1 can be drawn from Geometric Brownian Motion (GBM) (Karatzas and Shreve 2012) which models the path of fluid particles as they interact with other particles over time. It models two quantities; drift, to account for the deterministic effects and volatility, for the unpredictable effects of the interactions. In Eq. 1, $\mu_i\mathbf{n}_i$ is equivalent to the drift, and σ_i to volatility. According to Itō's theorem (Øksendal 2003), the differential equation has the following solution:

$$\mathbf{s}_i^M = \mathbf{s}_i^{M-1}exp\Big(\mu_i(\mathbf{n}_i - \frac{\sigma_i^2}{2})t + \sigma_i W\Big), \quad (2)$$

where t is the time elapsed from when the $M-1^{th}$ status message was posted. $\mathbf{s}_i^M$ is related to $\mathbf{s}_i^{M-1}$ by an exponential function, and each message mentioning u changes the coefficient value according to the intuitions in the postulates.

We next present a method to optimally learn the parameters $\mathcal{K} = \{\mu, \sigma, \mathbf{U}, \mathbf{V}, \mathbf{T}\}$. From the properties of GBM (Karatzas and Shreve 2012), $ln(\mathbf{s}_i^M) \sim \mathcal{N}(\mathbf{m}_i, \mathbf{v}_i)$. Here $\mathbf{m}_i = ln(\mathbf{s})_i^{M-1} + \big(\mu_i\mathbf{n}_i - \frac{\sigma_i^2}{2}\big)t$ is the mean and $\mathbf{v}_i = \sigma_i\sqrt{t}$ is the variance of the Gaussian Distribution. We denote the predicted value of the Mth status message as $ln(\mathbf{s}_i^M)$ for all i and equate it to the latent representation of the actual message. Let the word vector of the Mth status message be $\mathbf{g}$ and its latent representation be $\mathbf{gV} = \mathbf{a}$. We set $ln(\mathbf{s}_i^M) \approx \mathbf{a}_i$ for all i. The likelihood function then becomes

$$\mathcal{L}(\theta) = \frac{1}{\sigma_i\sqrt{2\pi t}}exp\Big(-\frac{(\mathbf{a}_i - \mathbf{m}_i)^2}{2(\mathbf{v}_i)^2}\Big) \quad (3)$$

Drawing from the principles of network homogeneity (Marsden 1988), the latent dimensions of users can be similar to the users they interact with. We model this by intro-

Algorithm 1: Predicting Online Protest Participation

Data: $\mathcal{U}, \mathcal{M}, \mathcal{K}, w_{reg}$
Result: Labels of Protest Declarations
for $u \in \mathcal{U}$ **do**
    Randomly initialize $\mathcal{K}$;
    for $M \in \mathcal{M}$ **do**
        for *all latent dimensions i* **do**
            Compute $\mathcal{L}(\theta)$ from Eq. 3 and f_r from Eq. 4;
            Compute $f_t = -ln(\mathcal{L}(\theta)) + w_{reg}f_r$;
            do
                Update values of $\mathcal{K}$ using Eq. to 6.;
                Compute $\mathbf{m}_i, \mathbf{v}_i$ from updated values.;
                Update values of $\mathcal{L}(\theta)$, f_r and f_t.;
            while f_t *does not converge*;
            Draw $ln(\mathbf{s}_i^{M+1}) \sim \mathcal{N}(\mathbf{m}_i, \mathbf{v}_i)$;
        end
    end
    Predict feature vector $\mathbf{u}$ as $ln(\mathbf{s}_i^{size(\mathcal{M})+1}) \forall i$;
end
Construct $\mathbf{F}$ from $\mathbf{u}$ for all $u \in \mathcal{U}$;
Classify the feature set $\mathbf{F}$ to obtain labels;

---

ducing the regularization function

$$f_r = \sum_{\mathbf{N}_{ij}^M = 1} (sig(\mathbf{X}_{i'}\mathbf{X}_{j'}{}^T) - 1)^2, \quad (4)$$

where $sig(x) = \frac{1}{1+e^{-x}}$ denotes the sigma function, and $\mathbf{N}^M$ is the interaction network of user u at the time of the Mth status message. The overall optimization function can be written as

$$\min_{\mu, \sigma, \mathbf{U}, \mathbf{V}, \mathbf{T}} -ln(\mathcal{L}(\theta)) + w_{reg}f_r, \quad (5)$$

where w_{reg} affects the contribution of the regularizer. Letting $\mathbf{a}_i - \mathbf{m}_i = \mathbf{c}_i$, the update equations for each dimension i are

$$\mu_i \leftarrow \mu_i + \eta \frac{1}{\sigma_i^2}\mathbf{n}_i\mathbf{c}_i, \sigma_i \leftarrow \sigma_i + \eta \frac{1}{t\sigma_i^3}\big(\mathbf{c}_i(-\mathbf{c}_i + t\sigma_i^2) + t\sigma_i^2\big)$$

$$\mathbf{U}_i \leftarrow \mathbf{U}_i + \eta \frac{\mu_i}{\sigma_i^2}\mathbf{c}_i \sum_{k \in \mathcal{P}_M} \mathbf{P}_{k'}{}^T\mathbf{X}_{k'}\mathbf{T}_{i'}{}^T,$$

$$\mathbf{T}_{i'} \leftarrow \mathbf{T}_{i'} + \eta \frac{\mu_i}{\sigma_i^2}\mathbf{c}_i \sum_{k \in \mathcal{P}_M} \mathbf{L}_{ki}^T\mathbf{X}_{k'}$$

$$\mathbf{V} \leftarrow \mathbf{V} + \eta \left(\frac{2}{\sigma_i^2 t}\mathbf{c}_i(-\mathbf{m}_i + \mu_i t \sum_{k \in \mathcal{P}_M} \mathbf{W}_{k'}{}^T\mathbf{L}_{ki}\mathbf{T}_{i'}) + w_{reg}\mathbf{Y} \right).$$

Here $\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad (6)$

$$\mathbf{Y} = \sum_{\mathbf{N}_{ij}^M = 1} 2sig(g_{ij})(sig(g_{ij}) - 1)^2\mathbf{X}_{i'}{}^T\mathbf{X}_{j'}\mathbf{V}, \quad (7)$$

where $\mathbf{X}_i^T\mathbf{X}_j = g_{ij}$. We repeat this all for the status messages of u, updating values of $\mathbf{m}$ and $\mathbf{v}$ at every instance.

After learning the parameters, we predict the latent membership of the next post of u as $\mathbf{u}_i \sim \mathcal{N}(\mathbf{m}_i, \mathbf{v}_i)$ for all i. We repeat this for all the users in the dataset to build a feature set

$\mathbf{F}$, whose each row corresponds to the latent dimension $\mathbf{u}$ of a single user. We then classify $\mathbf{F}$ to decide whether the next post of the user is a declaration of protest. The algorithm is summarized in **Algorithm 1**.

Complexity Analysis
The computational complexity mainly arises from the objective function and the update equations from Eqs. 5 and 6. Let us first consider the objective function in Eq. 5. The dominant computation here comes from $\mathbf{n}_i$, which needs $O(IP)$ operations, where P is the total number of messages mentioning the user. The complexity is less due to the low value of I, the number of dimensions in latent dimensional space.

We next consider the computational complexities in the update equations. The update of μ and σ have been $O(I)$, which is very small. The update of $\mathbf{U}$ takes $O(WIP)$ operations. Here W is the total number of words in $\mathbf{P}_k$ and $\mathbf{X}_k$. The update of $\mathbf{T}$ takes $O(PI^2)$ operations, which is low due to the low value of I. The update equation for $\mathbf{V}$ contains two parts, the terms arising from differentiation of likelihood function and social regularization. The term from differentiating the likelihood function takes $O(I^2 + WI)$, which is low due to the low value of I. The social regularization term takes $O((W_X^2 + W_X I) * C)$ operations. Here C is the total number of edges in the interaction network and W_X is the total number of words in $\mathbf{X}_i$ and $\mathbf{X}_j$. This is computationally tractable due to the sparsity of $\mathbf{X}$ and $\mathbf{N}$.

From the analysis presented, we can say that the algorithm is scalable with tractable computational complexity. Hence, the algorithm can be used in large data sets usually encountered in social media.

Experiments
We evaluate the framework using a real world dataset presented in Table 1 by answering the following questions: How does the framework perform in determining whether the next post of the user is a declaration of protest? What is the effect of the varying proportions of the interaction network information on the performance of the framework? How does the framework perform for different proportions of training data? We first present the designed experimental settings and then answer each of these questions in detail.

Experimental Settings
We now present the metrics and baselines used for evaluating the algorithm. In our dataset, there is an imbalance of positive and negative examples with a ratio of about 1:6. Therefore, we choose metrics that are capable of dealing with data imbalance. We use three metrics: accuracy, AUC, and the F1 measure computed for the positive class. The following baselines are employed as performance benchmarks.

- **Random Label Assignment**: Whether the user is protesting or not protesting is randomly assigned. This is repeated for 100 trials, and the mean value is presented.

- **Info Prop** (Jin et al. 2014): The authors model declarations of protest as it spreads through the information propagation network with external influence.

- **Topic** (Godin et al. 2013): This work predicts the topics of the next post of the user from the topics of his previous posts. We present the predicted topics to classify whether the next post of the user is a declaration of protest.

- **Sim Net** (Kywe et al. 2012): The authors use collaborative filtering methods to predict features of the next post of the user. They incorporate the content of similar users and posts while computing the features.

- **Time Topic** (Ma et al. 2014): The paper predicts the topics of the next post of the user by integrating the past content of the user, content at the time of the candidate message and interaction network between candidate users.

- **Our Method -Int**: We implement **Algorithm 1** with only the past status messages of the user i.e. $\mathbf{n}_i = \mu_i$. This is to assess the value of user interactions in predicting whether the next post of the user is a declaration of protest.

Performance Evaluation

We now evaluate the performance of the algorithm using Accuracy, AUC, and F1 measure, and compare it with the baselines. In this experiment, we set the value of information network parameter at $w_{reg} = 0.1$ and the number of latent dimensions $I = 50$. We set the earliest 50% of the candidate posts for training and the rest for testing. Later in the section, we will experiment with different values of w_{reg} and training data size. We use a linear discriminant classifier on the feature set $\mathbf{F}$ and illustrate the results in Table 2.

From Table 2, we see that random assignment performs poorly in all the three metrics showing the difficulty of the problem. (Jin et al. 2014) performs slightly better than random assignment, showing that diffusion-based algorithms might not work in cases where the entire propagation network is not available. (Godin et al. 2013) gives better performance by building topic models from the user's previous posts. This shows the utility of harnessing the user's past content to determine whether his next post will be a declaration of protest.

The baseline (Kywe et al. 2012) utilizes collaborative filtering and gives a similar performance as (Godin et al. 2013), even though it uses word features rather than topic models. The similar performance showcases that using relational attributes of users assists in enhancing the performance. This is further illustrated by the improvement over previous baselines shown by (Ma et al. 2014) which integrates content, network and temporal information to predict characteristics of the user's next post.

The improvement by **Our Method -Int** demonstrates that the framework effectively models the user's past status messages using concepts from GBM for the task. The significant improvement in **Our Method** demonstrates that sociological theories of protest participation to model the interplay between interactions and user content are useful in determining whether the next post of a user will be a declaration of protest. It also demonstrates the effectiveness of the framework for modeling these theories. A paired t-test showed that the method improves significantly over the baselines.

In summary, the proposed framework significantly outperforms the baselines and is useful for predicting whether the next post of a user will be a declaration of protest. In the next section, we will examine the variation of the performance with different values of the interaction network parameter.

Effect of Varying Interaction Network Parameter

In **Algorithm 1**, w_{reg} represents the contributions of interaction network information to the framework. To understand

Method	Accuracy	AUC	F1
Random	0.491	0.116	0.607
Info (Jin et al. 2014)	0.531	0.120	0.629
Topic (Godin et al. 2013)	0.587	0.133	0.692
Sim (Kywe et al. 2012)	0.570	0.147	0.693
Time (Ma et al. 2014)	0.599	0.150	0.715
Our Method -Int	0.636	0.177	0.765
Our Method	**0.702**	**0.204**	**0.817**

Table 2: Performance of the proposed framework.

the effect of this on the performance of the framework, we vary the value of the regularization parameter as $w_{reg} =$[0, 0.1, 0.2, 0.5, 1, 2, 5, 10]. We use the earliest 50% of the data for training and the rest for testing. We study the variations of the performance of the framework for these values and illustrate the results in Table 3.

From Table 3, we can see that interaction network information is useful in improving the performance of the framework in all the three metrics. For very high values of the information network parameter ($w_{reg} \geq 2$) the performance slightly decreases, as this might suppress the contributions of the other terms. However, we can see that it still significantly outperforms the baselines for all values of the information network parameter. A paired t-test to compare the results with the baselines showed that the improvement is significant. Therefore, the framework is a robust to the variation of the information network parameter.

In summary, we can say that the information network helps in improving the overall performance of the framework, emphasizing the contribution of the regularization term in determining whether the next post of a user will be a declaration of protest. The framework performs consistently across all values of the parameter and hence is robust to its variation. We next study the effect of varying training data proportions on the performance of the framework.

Variation of Training Data Size

We now evaluate the variation of performance of the framework with different proportions of training data. This experiment enables us to assess the performance of the framework when less amount of training information is available. A possible scenario could be in the early stages of protest when only a few users are available for training. This experiment will also help assess the robustness of the framework for varying proportions of training data. We change the percentage of the training data set from 10% to 90% in steps of 10%, starting from the earliest candidate post and measure the framework performance. We illustrate the results of the experiment in Table 4 and make the following observations.

From Table 4, we can say that more training data is beneficial for increasing the performance of the framework. The framework outperforms the nearest baseline (Ma et al. 2014), for tiny proportions for training data (10%), demonstrating that it performs well for low training data sizes. We also can see a slight dip in performance for higher proportions of training data ($> 80\%$), and this may be due to insufficient testing data. The performance shows consistent trends in all three metrics, showing that the framework efficiently utilizes the training data points to determine whether the next post of a user will be a declaration of protest.

In summary, the results demonstrate that the framework

		w_{reg}							
		0	**0.1**	**0.2**	**0.5**	**1**	**2**	**5**	**10**
Metrics	**Accuracy**	0.6813	0.7029	0.6977	**0.7163**	0.6761	0.6910	0.6768	0.6895
	AUC	0.1922	0.2044	0.2013	**0.2134**	0.1894	0.1913	0.1895	0.1959
	F1 Measure	0.7989	0.8177	0.8134	**0.8285**	0.7947	0.8108	0.7947	0.8074

Table 3: Performance of the framework with varying values of information network parameter

		Training Data Percentage								
		10%	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**	**90%**
Metrics	**Accuracy**	0.6749	0.6729	0.6839	**0.7044**	0.6975	0.6989	0.6690	0.6737	0.6561
	AUC	0.1962	0.1964	0.2017	**0.2242**	0.2217	0.2193	0.2143	0.2129	0.1723
	F1 Measure	0.7920	0.7931	0.8039	0.8173	0.8124	**0.8180**	0.7970	0.8013	0.7896

Table 4: Performance of the framework with varying values of training size

can learn from a small amount of training data, and it effectively utilizes training data to determine whether the next post of a user will be a declaration of protest. The framework consistently performs across all proportions of training data and hence is robust to its variations.

Related Work

The use of social media for political campaigns has received considerable attention in the literature (Segerberg and Bennett 2011; González-Bailón et al. 2011; Ranganath et al. 2016). (Lee et al. 2013) study the problem of extracting campaigns from social media using textual co-similarity, and the authors here do not concentrate on the behavior of individual users. The use of social media for political mobilization has been studied in (Bond et al. 2012), and the authors analyze the effect of political messages on political self-expression and interactions. They, however, do not predict future declarations of protest of an individual user. Behavior adoption by social media users has been modeled as an effect of information propagation in (Myers and Leskovec 2012; Lin et al. 2013). The effect of external influence on information propagation in networks have been studied in (Myers, Zhu, and Leskovec 2012; Iwata, Shah, and Ghahramani 2013). Propagation networks might not be fully available in many scenarios, and we utilize the user history to predict whether his next post will be a declaration of protest.

Considerable attention has been given to predicting characteristics of a user's future status messages from his history (Yang et al. 2012; Ranganath et al. 2015; Liang et al. 2012; Kumar et al. 2013). The authors in (Godin et al. 2013) predict the characteristics of the user's future messages by modeling the topics from his previous posts. This utilizes only the content of the individual user. The authors in (Kywe et al. 2012) use collaborative filtering methods by incorporating the content of similar users and posts to compute characteristics of future messages. The authors of (Ma et al. 2014) predict the characteristics of future messages by integrating the past content of the user, temporal information with the effect of interactions between candidate users. We build upon these works to model how messages interact with the user over time to affect the probability of his protest declaration in his next post.

The study of Brownian motion on networks has been investigated in (Zhou 2003). The theoretical development for adapting Brownian motion in the network environment has been considered in (Zhou 2003), with applications in community discovery. It has also been used for model movements of stock price, due to its ability to model sharp changes (Tankov 2003). More recently, Brownian motion has been used to model information propagation processes in the presence of external influence in social networks (Jin et al. 2014). In most scenarios, the complete propagation network between the candidate users is not available. To address this, we take each user as a separate entity and model his past interactions and status messages to predict his protest participation.

Conclusions and Future Work

Social media provides a platform for people to declare protest on various socio-political issues. In scenarios where the spread of protest is not available, predicting if the next post of a given user will be a declaration of protest will help in estimating the number of protest participants, facilitating better protest organization. In this paper, we draw from theories of protest participation and propose a framework that models the user's history to predict whether the next post of a given user will be a declaration of protest. We incorporate these principles to model the interplay between his content and interactions over time to predict latent dimension representations of the user's future post. We feed the latent dimensions to a classifier and determine whether the next post of a user will be a declaration of protest. We evaluate the framework on data posted in Twitter during protests on the Nigerian Elections. Our experiments demonstrate its effectiveness in predicting whether the next post of a user will be a declaration of protest.

This work paves the way to interesting future directions of research. Collecting labels of a user's posts over time can assist us in understanding the dynamical performance characteristics of the proposed model. Studying the behavior of protest groups over time will give insights into their formation and dynamics. The impact of the protest groups on spreading protest related messages is an interesting area of future research. Studying information seeking patterns among protesters can give interesting insights into their recruitment process. Finally, linguistic analysis of the messages posted during protests can provide insight into the persuasive content adopted by the protesters.

Acknowledgements

References

Bond, R. M.; Fariss, C. J.; Jones, J. J.; Kramer, A. D.; Marlow, C.; Settle, J. E.; and Fowler, J. H. 2012. A 61-million-person experiment in social influence and political mobilization. *Nature* 489(7415):295–298.

Contractor, D.; Chawda, B.; Mehta, S.; Subramaniam, L. V.; and Faruquie, T. A. 2015. Tracking political elections on social media: Applications and experience. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015, Buenos Aires, Argentina, July 25-31, 2015*, 2320–2326.

Godin, F.; Slavkovikj, V.; De Neve, W.; Schrauwen, B.; and Van de Walle, R. 2013. Using topic models for twitter hashtag recommendation. In *Proceedings of the 22nd international conference on World Wide Web companion*, 593–596. International World Wide Web Conferences Steering Committee.

González-Bailón, S.; Borge-Holthoefer, J.; Rivero, A.; and Moreno, Y. 2011. The dynamics of protest recruitment through an online network. *Scientific reports* 1.

Iwata, T.; Shah, A.; and Ghahramani, Z. 2013. Discovering latent influence in online social activities via shared cascade poisson processes. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, 266–274. ACM.

Jin, F.; Khandpur, R. P.; Self, N.; Dougherty, E.; Guo, S.; Chen, F.; Prakash, B. A.; and Ramakrishnan, N. 2014. Modeling mass protest adoption in social network communities using geometric brownian motion. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, 1660–1669. ACM.

Karatzas, I., and Shreve, S. 2012. *Brownian motion and stochastic calculus*, volume 113. Springer Science & Business Media.

Kumar, S.; Barbier, G.; Abbasi, M. A.; and Liu, H. 2011. Tweettracker: An analysis tool for humanitarian and disaster relief. In *ICWSM*.

Kumar, S.; Morstatter, F.; Zafarani, R.; and Liu, H. 2013. Whom should i follow?: identifying relevant users during crises. In *Proceedings of the 24th ACM conference on Hypertext and social media*, 139–147. ACM.

Kywe, S. M.; Hoang, T.-A.; Lim, E.-P.; and Zhu, F. 2012. On recommending hashtags in twitter networks. In *Social Informatics*. Springer. 337–350.

Lee, K.; Caverlee, J.; Cheng, Z.; and Sui, D. Z. 2013. Campaign extraction from social media. *ACM Transactions on Intelligent Systems and Technology (TIST)* 5(1):9.

Liang, H.; Xu, Y.; Tjondronegoro, D.; and Christen, P. 2012. Time-aware topic recommendation based on micro-blogs. In *Proceedings of the 21st ACM international conference on Information and knowledge management*, 1657–1661. ACM.

Lim, C. 2008. Social networks and political participation: How do networks matter? *Social Forces* 87(2):961–982.

Lin, Y.-R.; Margolin, D.; Keegan, B.; Baronchelli, A.; and Lazer, D. 2013. # bigbirds never die: Understanding social dynamics of emergent hashtags. In *Seventh International AAAI Conference on Weblogs and Social Media*.

Ma, Z.; Sun, A.; Yuan, Q.; and Cong, G. 2014. Tagging your tweets: A probabilistic modeling of hashtag annotation in twitter. In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*, 999–1008. ACM.

Mark, M. 2014. Nigeria elections: protests erupt following knife-edge vote. http://bit.ly/1JoPETY. [Online; accessed 8-August-2015].

Marsden, P. V. 1988. Homogeneity in confiding relations. *Social networks* 10(1):57–76.

Morstatter, F.; Pfeffer, J.; Liu, H.; and Carley, K. 2013. Is the sample good enough? comparing data from twitter's streaming api with twitter's firehose. In *International AAAI Conference on Weblogs and Social Media*.

Muthiah, S.; Huang, B.; Arredondo, J.; Mares, D.; Getoor, L.; Katz, G.; and Ramakrishnan, N. 2015. Planned protest modeling in news and social media.

Myers, S. A., and Leskovec, J. 2012. Clash of the contagions: Cooperation and competition in information diffusion. In *Data Mining (ICDM), 2012 IEEE 12th International Conference on*, 539–548. IEEE.

Myers, S. A.; Zhu, C.; and Leskovec, J. 2012. Information diffusion and external influence in networks. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, 33–41. ACM.

Øksendal, B. 2003. *Stochastic differential equations*. Springer.

Puig, C. 2015. Boko Haram extremists disrupt Nigerian election. http://usat.ly/1Jg91wz. [Online; accessed 8-August-2015].

Ranganath, S.; Wang, S.; Hu, X.; Tang, J.; and Liu, H. 2015. Finding time-critical responses for information seeking in social media. In *ICDM*.

Ranganath, S.; Hu, X.; Tang, J.; and Liu, H. 2016. Understanding and identifying advocates for political campaigns on social media. In *WSDM*.

Schussman, A., and Soule, S. A. 2005. Process and protest: Accounting for individual protest participation. *Social forces* 84(2):1083–1108.

Segerberg, A., and Bennett, W. L. 2011. Social media and the organization of collective action: Using twitter to explore the ecologies of two climate change protests. *The Communication Review* 14(3):197–215.

Snow, D. A.; Zurcher Jr, L. A.; and Ekland-Olson, S. 1980. Social networks and social movements: A microstructural approach to differential recruitment. *American sociological review* 787–801.

Tankov, P. 2003. *Financial modelling with jump processes*, volume 2. CRC press.

Tufekci, Z., and Wilson, C. 2012. Social media and the decision to participate in political protest: Observations from tahrir square. *Journal of Communication* 62(2):363–379.

Yang, L.; Sun, T.; Zhang, M.; and Mei, Q. 2012. We know what@ you# tag: does the dual role affect hashtag adoption? In *Proceedings of the 21st international conference on World Wide Web*, 261–270. ACM.

Zhou, H. 2003. Network landscape from a brownian particles perspective. *Physical Review E* 67(4):041908.