Twitter Hashtag Recommendation using Matrix Factorization

Hamidreza Alvari
Department of EECS, University of Central Florida, Orlando, FL, USA
halvari@eecs.ucf.edu

ABSTRACT

Twitter, one of the biggest and most popular microblogging Websites, has evolved into a powerful communication platform which allows millions of active users to generate huge volume of microposts and queries on a daily basis. To accommodate effective categorization and easy search, users are allowed to make use of hashtags, keywords or phrases prefixed by hash character, to categorize and summarize their posts. However, valid hashtags are not restricted and thus are created in a free and heterogeneous style, increasing difficulty of the task of tweet categorization. In this paper, we propose a low-rank weighted matrix factorization based method to recommend hashtags to the users solely based on their hashtag usage history and independent from their tweets’ contents. We confirm using two-sample t-test that users are more likely to adopt new hashtags similar to the ones they have previously adopted. In particular, we formulate the problem of hashtag recommendation into an optimization problem and incorporate hashtag correlation weight matrix into it to account for the similarity between different hashtags. We finally leverage widely used matrix factorization from recommender systems to solve the optimization problem by capturing the latent factors of users and hashtags. Empirical experiments demonstrate that our method is capable to properly recommend hashtags.

Categories and Subject Descriptors
H.4 [Information Systems Applications]: Miscellaneous; D.2.8 [Software Engineering]: Metrics—complexity measures, performance measures

General Terms
Theory

Keywords
Hashtag recommendation, Hashtag Correlation, Matrix Factorization, Twitter

1. INTRODUCTION

Twitter is one of the prevalent and well-known microblogging Websites with millions of active users interacting with each other and posting tweets, a message up to 140 characters, per day on either computers or mobile devices. It is popular for massive spreading of tweets and the nature of freedom. Daily bursts of news, gossip, rumors, discussions and many others are all exchanged and shared by users all over the world, no matter where they come from, civilized or uneducated, or even what religion they hold. Consequently, users on Twitter are easily overwhelmed by the tremendous volume of data. The proliferation of such an unstructured user-generated data as opposed to the traditional structured data, has enabled researchers to study and analyze human behavior and develop complex systems such as hashtag recommendation systems that has recently drawn few researchers attention.

On Twitter, users are freely allowed to assign valid hashtags to their tweets, i.e. strings prepended with the hash ”#” character, to categorize their posts and represent a coarse-grained topic of the content. Hashtags are indeed a community-driven convention for adding additional context to tweets. This mechanism helps tweet search and quickly propagation of the topic among millions of users by allowing them to join the discussion. Although facilitating the task, the heterogeneity and arbitrariness of the hashtags, due to the fact that users do not face any restrictions while creating them, can immediately make mess and hence make subsequent searches for tweets difficult.

To tackle the problem, few approaches has been proposed in the literature [14, 18, 19, 26, 28], most of them rely on either tweet similarity, i.e. content similarity or focus on adopting Latent Dricllet Allocation (LDA) [9] and Latent Semantic Indexing (LSI) [12] to capture the abstracted topics of tweets. The problem with these approaches is that not all the time, the contents of tweets are available due to the fact that many users do not usually have public timelines. Even if all users’ timelines were public, these methods suffer from computation overloads and lack of strong natural language processing techniques. As opposed to the existing methods, in this paper, we rather take another simple yet effective approach and propose a new hashtag recommendation method which works only based on the users’ interests and their hashtag usage history without incorporating the network structure and using tweet/user similarity or LDA.

In more details, we treat the problem of hashtag recommendation as an optimization problem and solve it via

citation here
widely employed matrix factorization borrowed from recommendation systems. Moreover, we envision that users are more likely to adopt similar hashtags while posting on Twitter and to verify that, we use two-sample t-test. Our contributions are thus summarized as follows:

- We perform two-sample t-test to verify that users keep adopting related and similar hashtags and hence possess consistent hashtag usage history.

- We address the hashtag recommendation problem with an optimization problem and propose a weighted matrix factorization based method,hWMF, to recommend hashtags to the users. To ease the process of optimization, we use alternating least square scheme for updating the corresponding matrices as finding the optimal values for them is tedious.

- We integrate the concept of correlation between the related hashtags into the matrix factorization equation. We quantify the correlation between hashtags based on the times they have been used together and incorporate these values into the optimization problem to weigh the contribution of training samples over test samples.

- We collect and build a dataset of tweets which contains at least five hashtags very close to a predefined list of 25 trending hashtags (hereafter called seeds list) selected from different categories. We evaluate the model on this dataset and demonstrate its ability to recommend hashtags that best aligns with the users’ hashtag usage history.

The remainder of the paper is organized as follows. In Section 2, we first define the problem of interest and then detail our proposed matrix factorization method to solve the problem. We explain the experimental settings and discuss the results in Section 3. Then, we discuss the related work in Section 4 and conclude the paper in Section 5.

2. PROPOSED FRAMEWORK

In this section, prior to describing the proposed method, we first provide the formal definition of the problem and the notations used throughout the paper. Then we detail the proposed model hWMF for Twitter hashtag recommendation and explain the time complexity of the method. We finally describe two sample t-test to verify our initial assumption about users’ hashtag usage.

2.1 Problem Statement

Given a set of N users and a set of M hashtags adopted by them, we aim to recommend the most related hashtags to the users based on their interests and the hashtag usage history. Suppose we have a very sparse and low-rank user-hashtag matrix \( X = [x_{ij}] \in \mathbb{R}^{N \times M} \). We denote by \( x_{ij}, 1 \leq i \leq N, 1 \leq j \leq M \) the ith row and jth column of \( X \), which represents if user \( i \) has adopted hashtag \( j \), if \( x_{ij} = 1 \), but has nothing to say otherwise. The zeros in the matrix demonstrate unknown or missing values; a user might have adopted a hashtag in the past but we could not figure it out since our data collection method does not always return all data. The reason that \( X \) is sparse is because in Twitter, users do not usually adopt hashtags while tweeting and despite the availability of this feature, only 8% of the tweets contain hash "#" character [18]. Also, suppose we have a hashtag-hashtag matrix \( Y = [y_{ij}] \in \mathbb{R}^{M \times M} \) with \( y_{ij}, 1 \leq i \leq M, 1 \leq j \leq M \) equals to the number of times hashtags \( i \) and \( j \) are used together.

We treat the problem of hashtag recommendation as a collaborative filtering based one and formulate it into an optimization problem and employ low-rank matrix factorization to solve it as this method has been widely and successfully employed in various applications such as collective filtering [17] and document clustering [29]. In its basic form and in the context of recommander systems, matrix factorization, one of the realizations of latent factor models, captures both items and users by vectors of factors inferred from the ratings. Here, instead we characterize users and hashtags by inferring vectors of factors from user’s hashtag usage history.

2.2 Matrix factorization model

Based on matrix factorization scheme, we seek two low-rank and non-negative matrices \( U \in \mathbb{R}_+^{d \times N} \) and \( V \in \mathbb{R}_+^{d \times M} \) with dimensionality of the latent space \( d \ll N, M \) via solving the following optimization problem:

\[
\min_{U, V} ||W \odot (X - UV^T)||_F^2 + \gamma_1 ||U||_F^2 + \gamma_2 ||V||_F^2 \tag{1}
\]

where \( \odot \) is Hadamard product (element-wise product) where \((X \odot Y)_{ij} = X_{ij} \times Y_{ij}\) for any two matrices \( X \) and \( Y \) with the same size, \(||.||_F\) is the Frobenius norm of a matrix, \(||A||_F = \sqrt{\sum_{i} \sum_{j} A_{ij}^2}\) and \( W = [w_{ij}] \in \mathbb{R}_+^{N \times M}, 1 \leq i \leq N, 1 \leq j \leq M \) is an indicator matrix (i.e. weight matrix) to control the learning process. Also, \( \gamma_1 > 0 \) and \( \gamma_2 > 0 \) are non-negative regularization parameters and \(||U||_F^2\) and \(||V||_F^2\) are two smoothness regularization terms to avoid overfitting. The row vectors \( u_i, 1 \leq i \leq N \) and \( v_j, 1 \leq j \leq M \) denote the low-dimensional representations of users and hashtags respectively.

We integrate \( W \) into the optimization equation to avoid impacts of unknown elements of \( X \), i.e. increase the contribution of the elements with known values in the optimization process over the elements with the missing information. Therefore, for those hashtags that we have information for, i.e. whether a user has adopted them before, we use \( w_{ij} = 1 \) and for those with missing information we use the average of their correlation with other hashtags with information.

In other words, the indicator matrix is formally defined as:

\[
w_{ij} = \begin{cases} 1, & x_{ij} = 1 \\
\frac{\sum_{k \neq i} \text{corr}(h_i, h_k)}{|A|}, & \text{otherwise}
\end{cases}
\tag{2}
\]

where \( A = \{ j \mid x_{ij} = 1 \} \) and \( \text{corr}(h_i, h_k) \), i.e. correlation between hashtags \( h_i \) and \( h_k \), is calculated by:

\[
\text{corr}(h_i, h_k) = \frac{\sum_{j \neq i} y_{jk}}{\sum_{j \neq i} y_{jj}}
\tag{3}
\]

Optimization. The coupling between \( U \) and \( V \) in the optimization problem, makes it difficult to find the optimal solutions for both matrices. Therefore, in this work, we adopt the alternating least squares method [13] to solve the optimization problem, where the objective function is iteratively optimized with respect to one of the variables \( U \) and \( V \) while fixing the other one until convergence. Optimizing the equation \( \text{1} \) with respect to \( U \) and \( V \) corresponds to the
computation of their derivatives via the following equations. Given the following objective function:

$$L = ||W \odot (X - UV^T)||_F^2 + \gamma_1 ||U||_F^2 + \gamma_2 ||V||_F^2$$  \hspace{1cm} (4)

the update equations for $U$ and $V$ are computed according to the following equations:

$$U = U - \lambda \frac{\partial L}{\partial U}$$  \hspace{1cm} (5)

$$V = V - \lambda \frac{\partial L}{\partial V}$$  \hspace{1cm} (6)

where $\lambda > 0$ is the learning step and the partial derivatives of $L$ with respect to $U$ and $V$ are then obtained using:

$$\frac{\partial L}{\partial U} = -2(W \odot X)V + 2(W \odot (UV^T))V + \gamma_1 U$$  \hspace{1cm} (7)

$$\frac{\partial L}{\partial V} = -2(W \odot X)^T U + 2(W \odot (UV^T))^T U + \gamma_2 V$$  \hspace{1cm} (8)

Upon the convergence, we approximate $X$ by multiplying the low-rank matrices $U$ and $V$:

$$\tilde{X} = UV^T$$  \hspace{1cm} (9)

Algorithm. The detailed algorithm for the proposed matrix factorization framework is shown in Algorithm 1. After randomly initializing matrices $U$ and $V$ and constructing the indicator matrix $W$ in lines 3-5, in lines 7-10, we alternately update $U$ and $V$ based on the equations to 8 until we reach convergence. Practically, convergence is achieved whenever predefined maximum number of iterations has been reached or there is little change in the objective function value. Finally, $\tilde{X} = UV^T$ is the low-rank representation of user-hastag matrix $X$ and also is non-negative as $U$ and $V$ are both non-negative matrices.

Algorithm 1 The proposed framework hWMF

1: Input: User-hastag matrix $X$, hashtag-hastag matrix $Y$, $d$, $\gamma_1$, $\gamma_2$, $\lambda$
2: Output: Modeled matrix $\tilde{X}$
3: Initialize $U$ randomly
4: Initialize $V$ randomly
5: Construct the indicator matrix $W$ according to eq. 2
6: while Not convergent do
7: \hspace{1cm} $\frac{\partial L}{\partial U} = -2(W \odot X)V + 2(W \odot (UV^T))V + \gamma_1 U$
8: \hspace{1cm} Update $U \leftarrow U - \lambda \frac{\partial L}{\partial U}$
9: \hspace{1cm} $\frac{\partial L}{\partial V} = -2(W \odot X)^T U + 2(W \odot (UV^T))^T U + \gamma_2 V$
10: \hspace{1cm} Update $V \leftarrow V - \lambda \frac{\partial L}{\partial V}$
11: end while
12: Set $\tilde{X} = UV^T$

2.3 Time Complexity

We discuss the time complexity of the proposed method, hWMF, here. Obviously, the complexity burden of hWMF depends mainly on the computation of the derivatives in Equations 7 and 8. In each iteration in our algorithm, the time complexities of the computation of the derivatives in lines 7 and 9 are calculated as follows: first note that $W \odot X$, $UV^T$ and $(W \odot (UV^T))$ need to be calculated once for both equations. The time complexity of $W \odot X$ is $O(N_x)$ where $N_x$ is the number of non-zero elements of the sparse matrix $X$. Also, the time complexity of $UV^T$ is $O(NdM)$. For $(W \odot (UV^T))V$, we need $O(N_xd)$. For the second term in Eq. 8 i.e. $(W \odot (UV^T))V$, we need $O(NM + NMD)$. Thus in each iteration, the calculation of $\frac{\partial L}{\partial U}$ takes $O(N_xd + NMD)$. With the similar computations, the calculation of $\frac{\partial L}{\partial V}$ has the time complexity of $O(N_xd + NMD)$ in each iteration.

In the next section, we provide statistical evidence that users are indeed more willing to adopt hashtags that have correlations with each other and consequently maintain consistent hashtag usage history. This confirms the correctness of our intuition on both incorporating the indicator matrix $W$ into the optimization equation and solving the equation via matrix factorization model as collaborative-filtering based models captures well the relations between items (here, hashtags).

2.4 t-test

We perform two-sample t-test and verify the existence of hashtag usage consistency. In particular, we seek to answer the question: Do users in Twitter possess consistent hashtag usage history?

We construct two vectors $hc_u$ and $hc_v$ with the equal number of elements where each element in $hc_u$ is obtained by calculating the correlation between hashtags $h_i$ and $h_j$ used by user $u$ using Eq. 3 and similarly each element in $hc_v$ is the calculated correlation score between hashtags $h_i$ used by user $u$ and $h_j$ used by a random user $r$.

We perform a t-test on vectors $hc_u$ and $hc_v$. The null hypothesis here is that the correlation of hashtags used by a given user does not differ from those of different users, i.e., these two vectors are the same, $H_0 : hc_u = hc_v$, while the other hypothesis is that the hashtags used by the same user are more correlated than those used by different users, $H_1 : hc_u > hc_v$. Therefore we have the following two-sample t-test:

$$H_0 : hc_u = hc_v, H_1 : hc_u > hc_v$$  \hspace{1cm} (10)

The t-test result suggests a strong evidence with the significance level $\alpha = 0.001$ to reject the null hypothesis and as a result, confirms that users tend to use a consistent set of hashtags while posting tweets on Twitter. Therefore the answer to the above question is positive. This aligns well with our findings and equations in our matrix factorization model in the previous section.

3. EXPERIMENTS

We first collect and build our dataset by making use of the Twitter streaming API, which provides 1% random tweets from the total volume of tweets at a particular moment. We then conduct experiments to compare the performance of our proposed method with the baselines. In this section, we begin by introducing our dataset and the evaluation metric and then we design experiments and discuss the results.

3.1 Dataset

In general, the evaluation of hashtag recommendation approach is challenging due to difficulty of collecting the appropriate and standard dataset, while also human annotation is almost impossible and unreliable because of the tedious

2https://dev.twitter.com/streaming/overview
workload for evaluating data. Therefore we need a more systematic way of generating the dataset.

We rather collect the data in the following way: we first picked a set of 25 trending hashtags of different categories as our initial seeds list including: News, Obama, Iran, Yemen, Gaza, Islam, Terrorism, Shooting, BlackLivesMatters, Youtube, Apple, Google, Microsoft, iPad, Android, Internet, FIFA, JohnNash, LadyGaga, Movie, Weekend, MemorialDay, Love, Hate, Care. We then expanded this list to the 6,814 related hashtags by collecting a set of tweets containing at least five related hashtags to the hashtags in the seeds list (see Table 2 for some examples). This way, we obtain a coarse-grained dataset of very close related hashtags which ease the task of evaluation of the proposed method. The description of the resulting dataset with %0.99 spareness is shown in Table 1.

The power-law [11] like distributions of the paired hashtags used by users and seen together are depicted in Figures 1 and 2 respectively. For the sake of clarity, we remove from Fig. 2 pairs of hashtags which have never appeared together. As we observe from Fig. 1 very few users have used more than 20 paired hashtags while the peak in this plot shows that most of the users have only used less than 20 paired hashtags. This is somehow demonstrated in Fig. 2 as well; roughly 5,424 paired hashtags are used together only once while we observe the decrease in the number of adopted paired hashtags as we move further in the plot. These figures together show the severe sparsity in our dataset; not all the hashtags have paired together at all; most paired hashtags are adopted together rarely.

We further plot the distribution of hashtag correlation scores in Fig. 3 (based on the Eq. 2). Since the number of uncorrelated hashtags are much higher than those with correlation, once again, for clarity we remove uncorrelated hashtags. This figure demonstrates that many hashtag correlation scores fall in (0,0.1] which shows most of the hashtags do not have strong correlation with each other, suggesting a power-law like distribution of hashtag correlation scores in our dataset.

### 3.2 Evaluation Metric

Similar to the most recommendation systems, we are unable to use popular metrics such as precision as they require to know the values of all entries in the ground truth matrix to correctly evaluate the returned values against them. Remember we had zero elements in the matrix representation of our dataset which do not necessarily show users did not adopt the hashtags, i.e. they demonstrate unknown or missing values; a user might have adopted a hashtag in the past but we could not figure it out as our data collection method does not always return all data from Twitter. Instead, we use widely used metric for evaluating collaborative filtering results, root mean square error (RMSE), which is defined as:

$$RMSE = \sqrt{\frac{\sum_{i,j} (\hat{x}_{ij} - x_{ij})^2}{n}} \quad (11)$$

where $n$ is the number of test instances used for the evaluation and $\hat{x}_{ij}$ and $x_{ij}$ are corresponding test set elements of $\hat{X}$ and $X$ selected for the evaluation respectively.

### 3.3 Discussion

As discussed before, only few works have addressed the problem of hashtag recommendation and none of them appear to be comparable to our method since they all recommend hashtags based on the tweets contents while we do that based on the hashtag usage history. Consequently, we build two baselines and compare the results as follows.

**MF:** this is a variant of our matrix factorization model
recommendation approaches on the new representation of setting users and hashtags to predict. We remove these relations by have adopted. We choose percentage of test set is fixed to %30.

For brevity, we only demonstrate this in Table 3, when and observe the best performance is achieved when $d = 0.001$. Also, we apply different dimensions of latent space parameters. In particular, we set both regularization and report the best performance, while other methods do not

with the indicator matrix $W$ defined as follows:

$$w_{ij} = \begin{cases} 1, & x_{ij} = 1 \\ 0, & x_{ij} = 0 \end{cases} \quad (12)$$

which is basically the same as the user-hashtag matrix $X$.

$k$-Nearest Neighbors with correlation ($k$NN - Correlation): given the whole set of potential hashtags for each user, this method recommends $k$ most correlated hashtags with those each user has used before. In details, we rank hashtags according to the Eq. 3 in descending order and recommend top $k$ hashtags to the user.

Random: as the name suggests, this method recommends hashtags on a random basis to the users.

For our proposed method, we try various parameters and report the best performance, while other methods do not have parameters. In particular, we set both regularization parameters $\gamma_1$ and $\gamma_2$ to 0.2 and set the learning step $\lambda$ to 0.001. Also, we apply different dimensions of latent space $d$ and observe the best performance is achieved when $d = 10$. For brevity, we only demonstrate this in Table 3 when percentage of test set is fixed to %30.

With the parameters chosen as above, our experimental setting is as follows: Suppose we have $\ell = \{(u_i, h_j) \mid x_{ij} = 1\}$ is the set of pairs of users and hashtags that we know they have adopted. We choose $%x$ of $\ell$ as new relations $\Upsilon$ between users and hashtags to predict. We remove these relations by setting $x_{ij} = 0, \forall (u_i, h_j) \in \Upsilon$ and then apply the hashtag recommendation approaches on the new representation of $X$. We vary $x$ as $\{10, 20, 30, 40, 50\}$.

We make the following observations:

- As stated before, for brevity, we only report the best performance of our method when the dimension of latent space $d$ is set to 10. To demonstrate the effect of $d$ on the results in Table 3, we fix the percentage of test set to %30 and vary $d$ as $\{5, 10, 15, 20, 25\}$. The observation suggests that $d = 10$ is the best dimension among others.

- The performance comparison for different methods in terms of RMSE is shown in Table 3. We observe the best results for all percentages of test set are achieved by our proposed method. The next best method is MF which is a variant of the proposed framework with a different indicator matrix. One observation is that defining the indicator matrix based on the correlation between hashtags perform better than the one solely defined similar to the matrix $X$. Furthermore, the reason our method recommend hashtags more accurately than kNN - Correlation is because we incorporate the correlation between hashtags into widely used matrix factorization while kNN - Correlation recommends hashtags merely based on the correlation of the hashtags.

- Table 3 shows two potential practical examples of our method for two anonymized Twitter users. As we observe, both of them have been offered one hashtag (to be concise, we do not show other recommended hashtags) and they actually adopt them in their future tweets. Obviously, these predicted hashtags are just those from the ground truth set that we use as test instances, i.e. those that we know users have previously used. This means we have not applied our method in the real world yet.

### 4. RELATED WORK

As opposed to the large body of works that has been devoted to the study of various social networks based applications and in particular recommendation systems in social networks, very few works have addressed the problem of hashtag recommendation for easy categorization and retrieval of tweets. The previous works on hashtag recommendation mostly rely on the similarity between tweets and none of them have exploited the correlation between the topics of the hashtags nor the users’ history of hashtags usage. In particular, they focused on two main directions: majority of them recommend hashtags based on content similarity having storage problems, fewer adopt LDA, which fails in short texts and is
unsupervised and hence requires efforts to associate tweets with hashtags, and use abstracted topics of tweets. Another problem with these approaches is the contents of tweets are not always available since many users do not usually make their timelines publicly available due to the privacy concerns. Moreover, these methods suffer from computation overloads and lack of strong natural language processing techniques which are required to analyze the contents of tweets.

[19] employs WordNet similarity information and Euclidean distance to suggest hashtags from similar tweets. Given a tweet from a user, [25] first discover similar tweets in their dataset based on the TF-IDF representation of tweets. Then they recommend hashtags by three different approaches, relying on ranking the hashtags based on the overall popularity of tweets, the popularity within the most similar tweets, and third one, the most similar tweets, reported to perform the best among the others. As opposed to their method which is solely based on the tweets similarities, [23] considers terms in tweets and their relevance to candidate hashtags. In more details, they propose a ranking scheme, a variant of TF-IDF that takes into account the relevance of hashtags and data sparsity. Similar to their method, [14] proposes an approach focusing on detection of hidden topics for the tweets and recommending the use of those general topics as hashtags based on LDA model. In details, an unsupervised and content based hashtag recommendation based on LDA is proposed in this study. They first design and develop a binary language classifier for tweets based on the Naive Bayes and Expectation-Maximization and then apply LDA in the context of tweet hashtag recommendation in a fully unsupervised manner.

Different from the other studies, [26] proposes a supervised topic modeling based approach for recommendation of hashtags by treating hashtags as labels of topics and leveraging topic model to discover relationship between words, hashtags and topics of tweets. with the assumption that each tweet is about one local topic and there is a global background topic for the corpus. They finally recommend k most probable hashtags based on the probability that a new tweet would contain a hashtag. [22] uses probability distributions for hashtag recommendation. In particular, they uses Bayes rule to compute the maximum a posteriori probability of each hashtag class given the words of the tweet.

In different attempt, [18] proposes a method that recommends hashtags based on similar users and tweets by calculating the preference weight of a user towards a certain hashtag based on the TF-IDF and then selecting the top users with high cosine similarities with other users. Also, they use the same approach for calculating the top similar tweets. Since many tweets do not contain hashtags, the recommended hashtags may be from similar tweets instead of similar users. They claim that their method is able to recommend more personalized hashtags compared to the other methods and their method suits both user preferences and the tweet content.

In contrast, this study seeks to recommend hashtags by incorporating hashtag usage history of each user into the hashtag recommendation problem and address the problem with an optimization problem while using matrix factorization model to solve it. In particular, we do not rely on the network structure or similarity between tweets/users; instead we exploit the correlation between hashtags by focusing on the probability of their adoption together.

5. CONCLUSION AND FUTURE WORK

In this study, we presented an approach to recommend hashtags to the users solely based on their interests and hashtags usage history, which in contrast to the approaches in the literature, does not require the contents of tweets as they are not always available since many users do not usually make their timelines publicly available due to the privacy concerns. In addition to that, all other methods suffer from severe computation burden and lack of strong natural language processing techniques to analyze the contents of tweets. We formulated the problem into an optimization problem and integrated correlation between different hashtags into the equation and used widely used matrix factorization technique to solve the problem via alternating least squares scheme. Experiments demonstrate that our method outperforms other baselines in recommending more accurate hashtags to the users.

In future, we plan to replicate the study by investigating the correlation between hashtags based on first discovering their topics from the content of their corresponding tweets through using topic modeling techniques such as LDA, and then computing their topic similarity based on the existing popular similarity measures. That way, we can integrate

| User | Hashtags adopted in past | Suggested hashtags | Tweets posted with suggested hashtags |
|------|--------------------------|--------------------|--------------------------------------|
| 1    | Obama, News, Democrats   | HillaryClinton     | #HillaryClinton how explain ALL BUSH FAULT? How still #rap 1993-2001 #BillClinton 100% INNO-CENT? |
| 2    | Weekend, Party, Club    | MemorialDay        | Military Hero Joey Jones Is Honored with The Real Deal Award https://... #MemorialDay |

Table 5: Example hashtag usage history and suggested hashtags

Table 4: Performance Comparison for different approaches in terms of RMSE with d = 10 and different testbed sizes

| User | Hashtags adopted in past | Suggested hashtags | Tweets posted with suggested hashtags |
|------|--------------------------|--------------------|--------------------------------------|
| 1    | Obama, News, Democrats   | HillaryClinton     | #HillaryClinton how explain ALL BUSH FAULT? How still #rap 1993-2001 #BillClinton 100% INNO-CENT? |
| 2    | Weekend, Party, Club    | MemorialDay        | Military Hero Joey Jones Is Honored with The Real Deal Award https://... #MemorialDay |
We are investigating other interesting research directions such as time series analysis of the adoption of hashtags by different users and extend the model to work well in presence of time slots in order to design seasonal hashtag recommendation systems as well as incorporating demographics of users specially gender of the users to recommend personalized hashtags to them.

6. REFERENCES

[1] G. Adomavicius and A. Tuzhilin. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. Knowledge and Data Engineering, IEEE Transactions on, 17(6):734–749, June 2005.

[2] H. Alvari, A. Hajibagheri, G. Suthankar, and K. Lakkaraju. Identifying community structures in dynamic networks. Social Network Analysis and Mining (SNAM), 2016.

[3] H. Alvari, S. Hashemi, and A. Hamzeh. Detecting overlapping communities in social networks by game theory and structural equivalence concept. In International Conference on Artificial Intelligence and Computational Intelligence, pages 620–630. Springer Berlin Heidelberg, 2011.

[4] H. Alvari, S. Hashemi, and A. Hamzeh. Discovering overlapping communities in social networks: A novel game-theoretic approach. AI Communications, 26(2):161–177, 2013.

[5] G. Beigi, X. Hu, R. Maciejewski, and H. Liu. An overview of sentiment analysis in social media and its applications in disaster relief. In Sentiment Analysis and Ontology Engineering: An Environment of Computational Intelligence, volume 639, pages 313–340. Springer International Publishing, 2016.

[6] G. Beigi, M. Jalili, H. Alvari, and G. Suthankar. Leveraging community detection for accurate trust prediction. 2014.

[7] G. Beigi, J. Tang, and H. Liu. Signed link analysis in social media networks. In International AAAI Conference on Web and Social Media (ICWSM), 2016.

[8] G. Beigi, J. Tang, S. Wang, and H. Liu. Exploiting emotional information for trust/distrust prediction. In SIAM International Conference on Data Mining (SDM), 2016.

[9] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. J. Mach. Learn. Res., 3:993–1022, Mar. 2003.

[10] J. Bobadilla, F. Ortega, A. Hernando, and A. Gutíerrez. Recommender systems survey. Knowl.-Based Syst., 46:109–132, July 2013.

[11] A. Clauset, C. R. Shalizi, and M. E. J. Newman. Power-law distributions in empirical data. SIAM Rev., 51(4):661–703, Nov. 2009.

[12] S. Deerwester, S. T. Dumais, G. W. Furnas, T. K. Landauer, and R. Harshman. Indexing by latent semantic analysis. Journal of the American society for information science, 41(6):391–407, 1990.

[13] C. H. Q. Ding, T. Li, and M. I. Jordan. Nonnegative matrix factorization for combinatorial optimization: Spectral clustering, graph matching, and clique finding. In JCDM, pages 183–192. IEEE Computer Society, 2008.

[14] F. Godin, V. Slavkovikj, W. D. Neve, B. Schrauwen, and R. V. de Walle. Using topic models for twitter hashtag recommendation. In WWW (Companion Volume), pages 593–596. International World Wide Web Conferences Steering Committee / ACM, 2013.

[15] A. Hajibagheri, H. Alvari, A. Hamzeh, and S. Hashemi. Community detection in social networks using information diffusion. In Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2012), pages 702–703. IEEE Computer Society, 2012.

[16] A. Hajibagheri, H. Alvari, A. Hamzeh, and S. Hashemi. Social networks community detection using the shapley value. In Artificial Intelligence and Signal Processing (AISP), 2012 16th CSI International Symposium on, pages 222–227. IEEE, 2012.

[17] Y. Koren. Factorization meets the neighborhood: a multifaceted collaborative filtering model. In Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining, pages 426–434. ACM, 2008.

[18] T. Li, Y. Wu, and Y. Zhang. Twitter hash tag prediction algorithm. In ICOMP’11-The 2011 International Conference on Internet Computing, 2011.

[19] T. Li, Y. Wu, and Y. Zhang. Twitter hash tag prediction algorithm. In Proceedings of the 2007 ACM Conference on Recommender Systems, RecSys ‘07, pages 17–24, New York, NY, USA, 2007. ACM.

[20] P. Massa and P. Avesani. Trust-aware collaborative filtering for recommender systems. In R. Meersman and Z. Tari, editors, On the Move to Meaningful Internet Systems 2004: CoopIS, DOA, and ODBASE, volume 3290 of Lecture Notes in Computer Science, pages 492–508. Springer Berlin Heidelberg, 2004.

[21] P. Massa and P. Avesani. Trust-aware recommender systems. In Proceedings of the 2007 ACM Conference on Recommender Systems, RecSys ’07, pages 17–24, New York, NY, USA, 2007. ACM.

[22] A. Mazzia and J. Juett. Suggesting hashtags on twitter. In EECS 545 Project, Winter Term, 2011.

[23] E. Otsuka, S. A. Wallace, and D. Chiu. Design and evaluation of a twitter hashtag recommendation system. In Proceedings of the 2009:4:2–4:2, Jan. 2009.

[24] A. Rajaraman and J. D. Ullman. Mining of massive datasets. Cambridge University Press, Cambridge, 2012.

[25] G. Salton and M. J. McGill. Introduction to modern information retrieval. 1986.

[26] J. She and L. Chen. Tomoha: Topic model-based hashtag recommendation on twitter. In WWW (Companion Volume), pages 371–372. ACM, 2014.

[27] X. Su and T. M. Khoshgoftaar. A survey of collaborative filtering techniques. Adv. in Artif. Intell., 2009:4:2–4:2, Jan. 2009.

[28] E. Zangerle, W. Gassler, and G. Specht. Recommending #-tags in twitter. In Proceedings of the Workshop on Semantic Adaptive Social Web.
[29] S. Zhu, K. Yu, Y. Chi, and Y. Gong. Combining content and link for classification using matrix factorization. In SIGIR '07: Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval, pages 487–494, New York, NY, USA, 2007. ACM.