Personalized Hashtag Suggestion for Microblogs

Juan Xu, Qi Zhang, Xuanjing Huang

School of Computer Science,
Shanghai Key Laboratory of Intelligent Information Processing,
Fudan University
{13916430285, qz, xjhuang}@fudan.edu.cn

Abstract. In microblogging services, users can generate *hashtags* to categorize their tweets. However, a majority of microblogs do not contain hashtags, which has intrigued active research on the problem of automatic hashtag recommendation for microblogs. Previous work conducted on this problem mostly does not take the user's preference into consideration. In this paper, we propose a novel personalized hashtag recommendation method for microblogs based on a probabilistic generative model which exploits users' perspectives on microblog posts for hashtag generation. Our experiments on a real microblogs dataset show that the proposed method outperforms state-of-the-art methods. We also show some case studies that demonstrate the advantages of considering both the content and user's personal preferences for hashtag suggestion.

1 Introduction

Microblogging services overload us with information, bombarding us with thousands of tweets, blog posts, and status updates every day. For example, Twitter, one of the most popular microblogging tools, has grown rapidly, with an estimated 200 million users generating 400 million tweets per day recently[1]. To cope with the volume of information shared daily, hashtags ¡[a]keywords prefaced with "#" in microblogging services ¡[a]have been introduced to help users categorize and search for tweets. Past empirical research shows that hashtags can be useful in many applications, including sentiment analysis [1, 2], breaking event discoveries [3], query expansion [4], etc. In the spectrum of industry, Google started supporting Google+ hashtags in search queries on Sep 25, 2013[2]. Despite the availability and the usefulness of this feature, only 12.84% tweets are marked with hashtags. Inclusion of hashtags in tweets is completely voluntary and user dependent. Thus, how to automatically generate or recommend hashtags has become an important research topic and drawn increased attention recently.

The task of hashtag recommendation is to automatically generate a short list of relevant hashtags as suggestions for a given tweet. Since it was first introduced by Mazzia and Juett [5], several methods have been proposed to tackle this problem [6–9]. Indeed, these studies are successfully tackling many notable challenges (i.e., hashtag sparseness, content shortness because of the 140-character limit, the vocabulary gap between

[1] https://blog.twitter.com/2013/celebrating-twitter7
[2] http://techcrunch.com/2013/09/25/google-starts-supporting-google-hashtags-in-search-queries/

tweets and hashtags, and topic diversity because of the open access in social media [10]) in hashtag recommendation. Nevertheless, they do not take into account personal preferences when recommending hashtags.

However, we believe that hashtag recommendation should be personalized. Users often utilize very different hashtags for their tweets [11]. For example, regarding "*The quarterfinal match between Roger Federer and Jo-Wilfried Tsonga in the French Open (a major tennis tournament held in Paris) 2013*," tweets posted by users about this match may contain diverse tags for different purposes, such as "**#FrenchOpen13**," "**#RolandGarros1/4" or "#tenni**," which are used by those who just watched the game and posted a tweet about it, or ¡° **#Roger, go**," "**#Roger, Allez & Come on**" or ¡° **#Hero, Federer**" which are used by Roger's fans to represent their personal perspectives on this topic. Therefore, we would like to consider users' preferences in the choice of hashtag suggestions for tweets.

In this paper, we propose a personal-topic-translation model (PTTM) exploring users' perspectives on tweets to recommend personalized hashtags for microblogs. Our proposed method is a comprehensive generative model. We introduce a user perspective latent variable to take users¡preferences into consideration as well as exploiting topical perceptions about microblogs in hashtag suggestion. We evaluate our model on a real-world dataset with posts that have been assigned with hashtags. We find that compared with other models, hashtags recommended by our model are more accurate and less redundant within the top-ranked results. We also use some examples to explain the advantages of our model.

2 Related Works

Recently, increased efforts have been made to address the problem of hashtags recommendation for a certain microblog post in microblogging platforms. Mazzia and Juett [5] provide a preliminary suggestion system, using a Naive Bayes approach, with much focus on pre-processing steps. Further, some methods exploit to compute the similarity between tweets based different similarity metric, and then to recommend hashtags from similar tweets. Zangerle et al. [11] investigate three different approaches to recommend hashtags based on a TF-IDF representation of the tweet. They rank the hashtags based on the overall popularity of the tweet, the popularity within the most similar tweets, and the most similar tweets. Li et al. use Euclidean distance as the similarity metric to suggest hashtags from similar tweets [12]. Zangerle et al. [13] explore five text similarity functions as the similarity measure for the computation of recommendations.

However, the suggested tags are sparse. Therefore, a few works have been conducted to address this issue. They propose methods for general hashtag recommendation based on the underlying topics of the tweets. For example, Ding et al. [6] propose a topic-specific translation model, which regards hashtags and tweets as parallel description of a resource, and then combine topic model and word alignment model to recommend the hashtags. Our method is partly based on their study. But the striking difference is that we model the hashtag generation with user factor. So that the hashtags would suit both user's preferences and the theme of tweet content. Godin et al. [14]

present an approach relies on Latent Dirichlet Allocation (LDA) to model the underlying topic assignment of tweets for general hashtags recommendation. But this approach have an inherit problem, that is the keyphrases extracted are too general to capture the tweet themes well.

All previous approaches always return the same list of tags for the same item regardless of user's preference. This problem is noted by [15]. Therefore, they propose a hashtag recommendation method based on collaborative filtering, which combines hashtags of similar users and similar tweets. TF-IDF approach is used to construct a feature vector for each tweet. Cosine similarity is used to compare the feature vectors. Although their approach considers user's preference, it ignores many other issues, such as tag sparse problem, etc. Therefore, in this paper, we propose a method attempting to address those challenges in hashtag recommendation problem.

3 Method

3.1 Preliminaries

We first introduce the notation used in this paper and formally formulate our problem. Let D to present an annotated corpus microblog posts, denoted as $d_1, d_2, \ldots, d_D$. Each post d_i is generated by a user u_i, where u_i is an index between 1 and U, and U is the total number of users. Each d_i consists of a pair of content words and assigned hashtags (w_i, t_i), where w_i and t_i are an index between 1 and D respectively. Each w_i contains a bag of words, denoted as $\{w_{i1}, w_{i2}, \ldots, w_{iN_i}\}$, where w_{iN_i} is an index between 1 and W , and W is the word vocabulary size. N_i is the number of words in w_i. Each t_i contains a bag of hastags, denoted as $\{t_{i1}, t_{i2}, \ldots, t_{iM_i}\}$, where t_{iMi} is an index between 1 and T , and T is the hashtag vocabulary size. M_i is the number of hashtags in t_i. Given an unlabeled data set, the task of personalized hashtag recommendation is to discover a list of hashtags for each post with perceptive of both users' preferences and tweet themes.

3.2 Model Formulation

We first describe how we address the vocabulary gap between hashtags and microblogs, and the topic diversity issue. Topical word trigger model proposed by Liu et al. [7] and Ding et al. [6] have been shown to be effective for solving these two issues, in which they assume that hashtags and tweets as parallel description of a resource, and a document contains a mixture of topics, and each word has a hidden topic label. From this perspective, hashtag suggestion can be regarded as a translation process from a given post content to tags under a specific topic. While this assumption works well on long documents, for short microblog posts, posts are noisy and a single post tends to be about a single topic. Recently, there has been much progress in modeling topics for short texts [8], which assumes a single topic assignment for an entire tweet and also assumes a background word distribution ϕ^B that captures common words. Similar idea has also been used in the works of Zhao et al. [10] and Diao et al. [16]. Based on these works, we introduce a topic model which is pretty suitable for microblogs in our method.

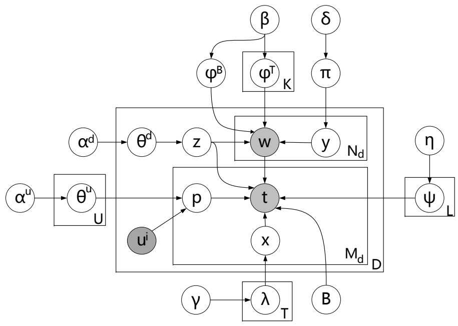

Fig. 1. The plate representation of the personal topical translation model.

As we discussed in Section 1, an important property of hashtag is that many hashtags are about user's personal perspectives on microblog posts rather than the themes of posts only. Thus our focus is to consider the impact of both microblog posts and user's preference for suggesting hashtags. To this end, an intuitive idea is that hashtags are either generated from posts or from user's perspectives. Therefore, we introduce a topic distribution θ^u for each user to capture her perspectives.

The proposed model is designed based on the following assumptions. When a user wants to write a tweet, she first generates the content, and then generates the hashtags. When she starts to write the content, she first chooses a topic based on the topic distribution θ^d. With the selected topic, words in the post are generated from the word distribution for that topic or from the background word distribution that captures white noise. During the generative process for hashtags, she first decides whether to tag about the post theme or her personal perspective. If she chooses the former, the hashtag is annotated according to the post topic. Otherwise, she selects a perspective according to her own perspective distribution θ^u to generate the hashtag. With the chosen generative source, hashtag t is either annotated according to the topic-dependent translation possibility $P(t_{dm}|w_d, z_d, \boldsymbol{B})$, where $P(t_{dm}|w_d, z_d, \boldsymbol{B}) = \sum_{n=1}^{N_d} p(t_{dm}|w_{dn}, z_d, \boldsymbol{B}) \cdot p(w_{dn}|w_d)$, and $\boldsymbol{B}$ presents the topic-specific word alignment table between a word and a hashtag which estimated by the combination of topic model and word alignment model, in which $B_{i,j,k} = P(t = t_j|w = w_i, z = k)$ is the word alignment probability between the word w_i and the hashtag t_j for topic k, or drawn from the tag distribution ψ^l of the perspective l. A variable x is introduced to decide the source of each hashtag, and we use λ to denote the probability of choosing to annotate according to the post theme rather than her personal perspective. Formally, the generation process of tweets is summarized in Figure 2. The plate representation of the proposed model is depicted in Figure 1.

3.3 Learning and Inference

We use collapsed Gibbs sampling [17] to obtain samples of the hidden variable assignment and estimate the model parameters from these samples. Due to the space limit, we leave out the derivation details and only show the derived Gibbs sampling formulas as follows. The major notations used in the following equations are explained in Table 1.

First, for the d-th tweet, we know its publisher u_d. The sampling probability of being a topic word or a background word for each word $v_i = w$ in dth tweet is sampled from:

$$p(y_{di} = s|v_i = w, z_{-i}, v_{-i}, y_{-di}, \delta, \beta) \propto \frac{M_{s,-i}+\delta}{M_{\cdot,-i}+2\delta} \cdot \frac{M_{h,-i}^{w_{di}}+\beta}{M_{h,-i}^{(\cdot)}+W\beta}, \tag{1}$$

where $h = B$ when $s = 0$ and $h = z_d$ when $s = 1$. $M_{0,-i}$ and $M_{1,-i}$ are counters to record the numbers of words assigned to the background model and any topic, respectively. $M_{B,-i}^{w_{di}}$ is the times of w_{di} that assigned to background words. $M_{z_d,-i}^{w_{di}}$ is the numbers of w_{di} that are assigned to topic z_d. $-i$ indicates taking no account of the current position i.

Then we sample the tweet topic variable z_d for d-th tweet using:

$$p(z_d = k|w_d, z_{-d}, y, \alpha^d, \beta) \propto \frac{M_{k,-d}+\alpha^d}{M_{\cdot,-d}+K\alpha^d} \cdot \prod_{i=0}^{N_d} \frac{M_{k,-d}^{w_{di}}+\beta}{M_k^{(\cdot)}+W\beta}, \tag{2}$$

where $M_{k,-d}$ is the numbers of tweets that are assigned with topic k in the corpus; $M_{k,-d}^{w_{di}}$ is the number of occurrences of topic word w_{di} that is assigned with topic k, here topic word refers to word whose latent variable y equal 1; $-d$ indicates taking no account of the current tweet w_d.

We still have two latent variables which are tag topic (or tag perspective) variable p and the generative source of each hashtag variable x. We can jointly sample them based on the values of all other hidden variables. As we described in our model when the tag source variable $X = 1$, topic for each tag is generated from it's tweet topic and that has been sampled in formulas (2). Therefore, we just need to sample the user's perspective variable for each tag $q_j = t$ when the tag source variable $X = 0$:

$$p(x_{dj} = 0, p_{dj} = l|q_j = t, t_{-j}, p_{-j}, \gamma, \alpha^u, \eta) \propto \frac{n_{t,-j}+\gamma}{n_{t,-j}+\tilde{n}_t+2\gamma} \cdot \frac{M_{u,-j}^l+\alpha^u}{M_{u,-j}^{(\cdot)}+L\alpha^u} \cdot \frac{M_{l,-j}^t+\eta}{M_{l,-j}^{(\cdot)}+T\eta}, \tag{3}$$

where $n_{t,-j}$ and $\tilde{n}_t$ are the number of times that tag t is generated from perspectives ($X_t = 0$) and topics ($X_t = 1$), respectively; $M_{u,-j}^l$ is the number of times that perspective l is adopted by user u; $M_{l,-j}^t$ is the number of times tag t is generated from perspective l; $-j$ indicates taking no account of the current position j.

After enough sampling iterations to burn in the Markov chain, we can estimate the eight parameters in our model:(1) the content-topic distribution θ^d, (2) the topic-word distribution ϕ, (3) the background-word distribution, (4) the binomial distribution π, (5) the topic-dependent word alignment table between a word and a hashtag B, (6) the user-perspective distribution θ^u, (7) the perspective-tag distribution ψ and (8) the binomial distribution λ for any single sample using the following equations:

$$
\begin{aligned}
&\theta^{(d)} = \frac{M_{k,-d}+\alpha^d}{M_{\cdot,-d}+K\alpha^d}, \quad \phi_{wk} = \frac{M_k^w+\beta}{M_k^{(\cdot)}+W\beta} \\
&\phi_{wb} = \frac{M_b^w+\beta}{M_b^{(\cdot)}+W\beta}, \quad \pi = \frac{M_{1,-i}+\delta}{M_{\cdot,-i}+2\delta} \\
&B_{z,w,t} = \frac{N_{t,w}^z}{\sum_{z'} N_{t,w}^{z'}}, \quad \theta^{(u)} = \frac{M_{u,-j}^l+\alpha^u}{M_{u,-j}^{(\cdot)}+L\alpha^u} \\
&\psi_{tl} = \frac{M_{l,-j}^t+\eta}{M_{l,-j}^{(\cdot)}+T\eta}, \quad \lambda_t = \frac{\widetilde{n}_{t,-j}+\gamma}{n_t+\widetilde{n}_{t,-j}+2\gamma},
\end{aligned}
\tag{4}
$$

where $N_{t,w}^z$ is the number of occurrences that w is translated to t given topic t.

1. Draw $\phi^B \sim$ Dirichlet (β), $\pi \sim$ Beta (δ), $\lambda^t \sim$ Beta (γ)
2. For each topic $k = 1, \ldots, K$
   (a) Draw $\phi^T \sim$ Dirichlet (β)
3. For each user $u = 1, \ldots, U$
   (a) Draw $\theta^u \sim$ Dirichlet (α^u)
4. For each perspective $l = 1, \ldots, L$
   (a) Draw $\psi^l \sim$ Dirichlet (η)
5. For each tweet $d = 1, \ldots, D$ created by $u = 1, \ldots, U$
   (a) Draw $\theta^d \sim$ Dirichlet (α^d)
   (b) Draw $z_d \sim$ Multinomial (θ^d)
   (c) for each word $n = 1, \ldots, N_d$
      i. Draw $y_{dn} \sim$ Bernoulli (π)
      ii. if ($y_{dn} = 1$) :
         Draw $w_{dn} \sim$ Multinomial (ϕ^{z_d})
      iii. if ($y_{dn} = 0$) :
         Draw $w_{dn} \sim$ Multinomial (ϕ^B)
   (d) for each hashtag $m = 1, \ldots, M_d$
      i. Draw flag $x_{dm} \sim$ Bernoulli ($\lambda^{t_{dm}}$)
      ii. if ($x_{dm} = 1$) :
         Draw $t_{dm} \sim P(t_{dm}|w_d, z_d, \mathbf{B})$
      iii. if ($x_{dm} = 0$) :
         Draw $p_{dm} \sim$ Multinomial (θ^u)
         Draw $t_{dm} \sim$ Multinomial (ψ^{dm})

Fig. 2. The generation process for all posts.

3.4 Personalized Hashtag Recommendation

In our model, we perform personalized hashtag recommendation as the followings. Given a tweet d with its content $w = \{w_n\}_{n=1}^N$ consisting of N words, paired with its user u. We first perform Gibbs Sampling to iteratively estimate the topic distribution of d (i.e., $\theta^{(d)}$) according to tweet content w. Afterwards, we can rank the hashtags for this post by computing the scores:

$$p(t_m \mid w, u) = p(\lambda_{t_m}) \sum_{c=1}^{K} \sum_{n=1}^{N} p(t_m \mid c_m, w, B) \cdot p(c_m \mid \theta_w^{(d)}) \cdot p(w_n \mid w) + [1 - p(\lambda_{t_m})] \sum_{l=1}^{L} p(t_m \mid p_l) \cdot p(p_l \mid u), \tag{5}$$

where $p(w_n \mid w)$ is the weight of the word w_n in post content w, which can be estimated by *TFIDF* or *IDF*, we apply *IDF* to compute this score. According to the ranking scores, we can suggest the top-ranked personalized hashtags for this post-user either.

4 Experiments and Results

4.1 Datasets

The dataset used for experiment is from a microblogging dataset we collected from Sina-Weibo[3], a popular Twitter-like microblogging system in China. The original dataset contains 20 million microblogs posted by $60,000$ users from Sep. 2009 to May. 2013. For prepossessing, we used ICTCLAS2009[4] to segment these microblogs and remove non-standard words(i.e. punctuation marks, urls, at-mentions, etc.) and stop words. Microblogs that do not contain hashtags or containing less than 3 words are removed from the dataset. Since we want to recommend personalized hashtags for microblogs, we filtered out users who posted microblog with hashtags less than 10 times. The remaining dataset contains $112,084$ microblogs posted by $6,661$ unique users and is used as our final dataset for training and evaluation. Some detailed statistics is shown in Table 1. We divided them into a training set of $101,644$ tweets posted by all users in our dataset and a test set of $10,440$ tweets. The hashtags actually annotated by users serve as the ground truth.

#Unique users	6,661
#Microblogs containing hashtags	124,707
#Vocabulary of words	10,7376
#Vocabulary of tags	33,777
#Average number of words in each microblog	27.23
#Average number of words in each hashtag	1.03

Table 1. Statistics of the dataset used in this paper

[3] http://weibo.com/
[4] http://ictclas.org/Down_OpenSrc.asp

4.2 Evaluation Criteria and Experimental Setup

We use Precision (P), Recall (R), and F-value (F) to evaluate the performance of hashtag recommendation methods. These metrics are computed as:

$$Precision = \frac{\#tags\ truly\ assigned}{tags\ assigned\ by\ system} \tag{6}$$

$$Recall = \frac{\#tags\ truly\ assigned}{tags\ manually\ assigned} \tag{7}$$

$$F = \frac{2P \cdot R}{P+R} \tag{8}$$

We ran our model with 500 iterations of Gibbs sampling. After trying a few different numbers of topics and user perspectives while one of them is fixed, we empirically set the number of topics to 10 and the number of perspectives to 80. We use $\alpha^d = 0.1, \alpha^u = 0.1, \delta = 0.1, \psi = 50/L$, and $\beta = 50/K$ as Griffiths and Steyvers [17] suggested. Parameter γ is set to 0.5. As two of our select baselines are variations of topic model, both of them are carefully tuned on the training data also.

4.3 Methods for Comparison

We compare our personal topic translation model (Figure 1) with 4 baselines and one variation of our model which are described as follows:

- **Naive Bayes(NB)**: This is a representative classification-based method. [18]. We applied this method to model the posterior probability of each hashtag given a tweet.
- **TSTM model**: This is a topical word alignment model proposed by Ding et al. [6], which assumes one tweet have multiple topics.
- **TTM model**: TTM model is a topical translation model described by Ding et al. [8]. We find it is kind of a variation model of TSTM model. To the best of our knowledge, it is the state of art method for hashtags recommendation for microblogs.
- **CF method**: This is a method based on collaborative filtering used by Kywe et al., [15], which combines hashtags of similar users and similar tweets to propose a more personalized set of tags.
- **PTTM_1**: PTTM_1 is a variation of our model, in which we consider one tweet have multiple topics and all the words in tweets are topic-related.

4.4 Experiment Results

Overall Performance In this subsection, We compared our personal topic translation model (Figure 1) with those baselines mentioned above. In Figure 3 we show the precision-recall curves of NB, TSTM, TTM, CF, PTTM_1 and PTTM on the dataset. Each point of a precision-recall curve represents suggesting different number of hashtags M^K, respectively ($M^k = \{1, 2, 3, 5, 10\}$).

Figure 3 clearly shows that PTTM outperforms all the other baseline methods. This indicates the effectiveness of our approach. On one hand, PTTM outperforms the state of art method TTM which implies that it is essential to take user's preference into

Method	Precision	Recall	F-measure
NB	0.236	0.231	0.233
TSTM	0.309	0.304	0.306
TTM	0.416	0.411	0.413
CF	0.276	0.267	0.271
PTTM_1	0.341	0.336	0.338
PTTM	**0.447**	**0.441**	**0.443**

Table 2. Comparison results of NB, TSTM, TTM, CF, PTTM_1 and PTTM, when suggesting top-1 hashtag.

consideration in the choice of suggesting hashtags for microblogs. On the other hand, PTTM, TTM and PTTM_1 outperform the CF method, it indicates that traditional collaborative filtering method is not suitable for personalized hashtag recommendation problem because of the the shortness of microblog and the diversity of microblogs topics as we concerned. There is an interesting phenomena, when M^k is getting smaller, the advantages of PTTM are more obvious compared to baselines. This implies that when a system is asked to suggest less hashtags for microblog, it is becoming important to take user's preference into account.

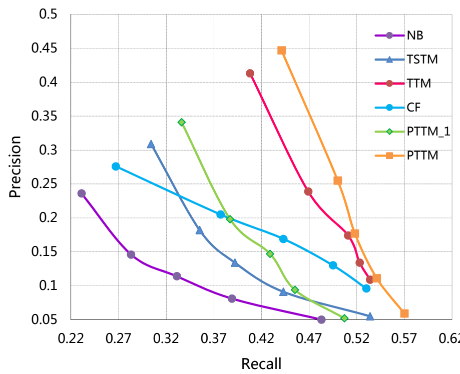

Fig. 3. Precision-recall curves for hashtags recommendation.

An additional observation is that TTM outperforms PTTM_1, which may imply that compared to consider user's preference, it is more important to assure that the hashtag is related to the tweet topic. It may illustrate that even though there exits a lot of hashtags are generated from user's perspective, a larger part of hashtags are generated from tweet themes. The last observation is that PTTM outperform PTTM_1 significantly just as TTM outperforms TSTM discriminately, it validates the observation that each microblog tends to cover only one topic.

To further demonstrate the performance of PTTM and other baseline methods, in Table 3, we show the Precision, Recall and F-measure of those models suggesting top-1 hashtag, because the number 1 is near the average number of hashtags in dataset. We find that the F-measure of PTTM comes to 0.443, outperforming the state of art method 7.26% relatively. Since the hashtags are very sparse, we owe the high performance of top 1 recommended hashtag to the incorporation of topic model, as it can suggest hashtags based on the underlying topics of the microblogs.

From Table 3 we observe that: (1) TTM can suggest hashtags that are closely related to the topic, such as “FrenchOpen13” and “Tennis”. However, due to not considering user’s preference, TTM recommends nearly the same set of hashtags for different users. (2) Taking advantage of considering user’s preference into account, PTTM can suggest representative and related hashtags and at the same time guarantee to suite user’s taste, such as ”Federer Allez”. We can see that PTTM can attain a good accuracy, and more, achieves the goal aiming at suggesting hashtags, which would suit both the content and user’s preference, for microblogs.

(User, Topic)	Top-5 hashtags
(2786, Frech Open)	**PTTM**: FrenchOpen13(*), Longines for FrenchOpen daily guess, Federer(+), 1/4FrenchOpen(+),Roger vs Tsonga(+)
	TTM: FrenchOpen13(*), Federer(+), Tennis(+), Longines for FrenchOpen daily guess, 1/4FrenchOpen(+)
(4556, Frech Open)	**PTTM**: Federer Allez(*), FrenchOpen13(+), Longines for FrenchOpen daily guess, Tsonga(+), Roger vs Tsonga(+)
	TTM: FrenchOpen13(+), Federer(+), Longines for FrenchOpen daily guess, Tennis(+), Tsonga(+)

Table 3. Examples of top-5 hashtags suggested by TTM, PTTM.

Parameter Influences There are two crucial parameters in PTTM, the number of topics T and the number of perspectives L. We first fix the number of perspectives to a certain number, and then test the performance of the trained model on the test data for different topic numbers. The smallest topic number which leads to the highest accuracy is selected. After the topic is chosen, the perspective number is selected similarly. We demonstrate the performance of PTTM for hashtag recommendation when parameters change in the table 4 and 5.

From Table 4, we can see that as the number of topics T varies from $T = 10$ to $T = 80$ when L is fixed to 80, the performance of hahstag suggestion roughly decreases. This shows that the granularity of topics will influence the recommendation performance. When $T = 10$, the performance achieves best, which properly due to the fact that this

T	Precision	Recall	F-measure
10	**0.447**	**0.441**	**0.443**
20	0.415	0.409	0.411
60	0.408	0.402	0.404
80	0.433	0.427	0.429

Table 4. The influence of topic number T of PTTM for hashtags suggestion when $M^k = 1$ and when $L = 80$.

topic number well covers the topics of the microblogs in the microblogging websites where our corpus crawled from. Hence we set $T = 10$ for our model.

As shown in Table 5, the topic number is fixed to 10, when the perspective number is set with $L = 20$, $L = 60$ or $L = 80$, PTTM achieves the relatively best performance. When $L = 10$, the performance is much poorer. This reveals that compared to the tweet topics, user's perspectives are more diversified. Therefore we set $L = 80$ for our model.

L	Precision	Recall	F-measure
10	0.419	0.413	0.415
20	0.443	0.437	0.439
60	0.445	0.438	0.440
80	**0.447**	**0.441**	**0.443**

Table 5. The influence of perspective number L of PTTM for hashtags suggestion when $M^k = 1$ and when topic number $K = 10$.

5 Conclusions

In this paper, we studied the problem of recommending hashtags for microblogs. Since most of existing work on this task does not take users' preference into consideration, we introduced a novel personal topic translation model which considers the impact of both content and users' preference on hashtag generation. We compared our model with a classification-based method, a topical translation method, the state-of-the-art models and a traditional collaborative filtering method as well as one variations of our model on a real microblogging dataset. Quantitative evaluations showed that our model could more accurately suggest hashtags for tweet. We also used some case studies to illustrate the effectiveness of the topic factor and the user factor of our model.

References

1. Davidov, D., Tsur, O., Rappoport, A.: Enhanced sentiment learning using twitter hashtags and smileys. In: Proceedings of the 23rd International Conference on Computational Linguistics: Posters, Association for Computational Linguistics (2010) 241–249
2. Wang, X., Wei, F., Liu, X., Zhou, M., Zhang, M.: Topic sentiment analysis in twitter: a graph-based hashtag sentiment classification approach. In: Proceedings of the 20th ACM international conference on Information and knowledge management, ACM (2011) 1031–1040
3. Cui, A., Zhang, M., Liu, Y., Ma, S., Zhang, K.: Discover breaking events with popular hashtags in twitter. In: Proceedings of the 21st ACM international conference on Information and knowledge management, ACM (2012) 1794–1798
4. Bandyopadhyay, A., Ghosh, K., Majumder, P., Mitra, M.: Query expansion for microblog retrieval. International Journal of Web Science **1**(4) (2012) 368–380
5. Mazzia, A., Juett, J.: Suggesting hashtags on twitter (2009)
6. Ding, Z., Zhang, Q., Huang, X.: Automatic hashtag recommendation for microblogs using topic-specific translation model. In: Proceedings of COLING 2012. (2012) 265–274
7. Liu, Z., Liang, C., Sun, M.: Topical word trigger model for keyphrase extraction. In: COLING. (2012) 1715–1730
8. Ding, Z., Qiu, X., Zhang, Q., Huang, X.: Learning topical translation model for microblog hashtag suggestion. In: Proceedings of the Twenty-Third international joint conference on Artificial Intelligence, AAAI Press (2013) 2078–2084
9. Tariq, A., Karim, A., Gomez, F., Foroosh, H.: Exploiting topical perceptions over multi-lingual text for hashtag suggestion on twitter. In: The Twenty-Sixth International FLAIRS Conference. (2013)
10. Zhao, W.X., Jiang, J., Weng, J., He, J., Lim, E.P., Yan, H., Li, X.: Comparing twitter and traditional media using topic models. In: Advances in Information Retrieval. Springer (2011) 338–349
11. Zangerle, E., Gassler, W., Specht, G.: Recommending#-tags in twitter. In: Proceedings of the Workshop on Semantic Adaptive Social Web (SASWeb 2011). CEUR Workshop Proceedings. Volume 730. (2011) 67–78
12. Li, T., Wu, Y., Zhang, Y.: Twitter hash tag prediction algorithm. ICOMP11 - The 2011 International Conference on Internet Computing (2011)
13. Zangerle, E., Gassler, W., Specht, G.: On the impact of text similarity functions on hashtag recommendations in microblogging environments. Social Network Analysis and Mining (2013) 1–10
14. Godin, F., Slavkovikj, V., De Neve, W., Schrauwen, B., Van de Walle, R.: Using topic models for twitter hashtag recommendation. In: Proceedings of the 22nd international conference on World Wide Web companion, International World Wide Web Conferences Steering Committee (2013) 593–596
15. Kywe, S.M., Hoang, T.A., Lim, E.P., Zhu, F.: On recommending hashtags in twitter networks. In: Social Informatics. Springer (2012) 337–350
16. Diao, Q., Jiang, J., Zhu, F., Lim, E.P.: Finding bursty topics from microblogs. In: Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers-Volume 1, Association for Computational Linguistics (2012) 536–544
17. Griffiths, T.L., Steyvers, M.: Finding scientific topics. Proceedings of the National academy of Sciences of the United States of America **101**(Suppl 1) (2004) 5228–5235
18. Garg, N., Weber, I.: Personalized, interactive tag recommendation for flickr. In: Proceedings of the 2008 ACM conference on Recommender systems, ACM (2008) 67–74