Time-aware Personalized Hashtag Recommendation on Social Media

Qi Zhang, Yeyun Gong, Xuyang Sun, Xuanjing Huang
Shanghai Key Laboratory of Intelligent Information Processing
School of Computer Science, Fudan University
825 Zhangheng Road, Shanghai, P.R.China
{qz, 12110240006, 13210240106, xjhuang}@fudan.edu.cn

Abstract

The task of recommending hashtags for microblogs has been received considerable attention in recent years, and many applications can reap enormous benefits from it. Various approaches have been proposed to study the problem from different aspects. However, the impacts of temporal and personal factors have rarely been considered in the existing methods. In this paper, we propose a novel method that extends the translation based model and incorporates the temporal and personal factors. To overcome the limitation of only being able to recommend hashtags that exist in the training data of the existing methods, the proposed method also incorporates extraction strategies into it. The results of experiments on the data collected from real world microblogging services by crawling demonstrate that the proposed method outperforms state-of-the-art methods that do not consider these aspects. The relative improvement of the proposed method over the method without considering these aspects is around 47.8% in F1-score.

1 Introduction

Over the past few years, social media services have become one of the most important communication channels for people. According to the statistic reported by the Pew Research Center's Internet & American Life Project in Aug 5, 2013, about 72% of adult internet users are also members of at least one social networking site. Hence, microblogs have also been widely used as data sources for public opinion analyses (Bermingham and Smeaton, 2010; Jiang et al., 2011), prediction (Asur and Huberman, 2010; Bollen et al., 2011), reputation management (Pang and Lee, 2008; Otsuka et al., 2012), and many other applications (Sakaki et al., 2010; Becker et al., 2010; Guy et al., 2010; Guy et al., 2013). In addition to the limited number of characters in the content, microblogs also contain a form of metadata tag (hashtag), which is a string of characters preceded by the symbol (#). Hashtags are used to mark the keywords or topics of a microblog. They can occur anywhere in a microblog, at the beginning, middle, or end. Hashtags have been proven to be useful for many applications, including microblog retrieval (Efron, 2010), query expansion (A.Bandyopadhyay et al., 2011), sentiment analysis (Davidov et al., 2010; Wang et al., 2011). However, only a few microblogs contain hashtags provided by their authors. Hence, the task of recommending hashtags for microblogs has become an important research topic and has received considerable attention in recent years.

Existing works have studied discriminative models (Ohkura et al., 2006; Heymann et al., 2008) and generative models (Krestel et al., 2009; Blei and Jordan, 2003; Ding et al., 2013) based on textual information from a single microblog. However, from a dataset containing 282.2 million microblogs crawled from Sina Weibo[1], we observe that different users may have different perspectives when picking hashtags, and the perspectives of users are impacted by their own interests or the global topic trend. Meanwhile,the global topic distribution is likely to change over time. To better understand how the topics vary over time, we aggregate the microblog posts published in a month as a document. Then, we use a Latent Dirichlet Allocation (LDA) to estimate their topics. Figure 1 illustrates an example, where ten active topics are selected. We can observe that the topics distribution varies greatly over time.

[1]http://www.weibo.com. It is one of the most popular microblog services in China.
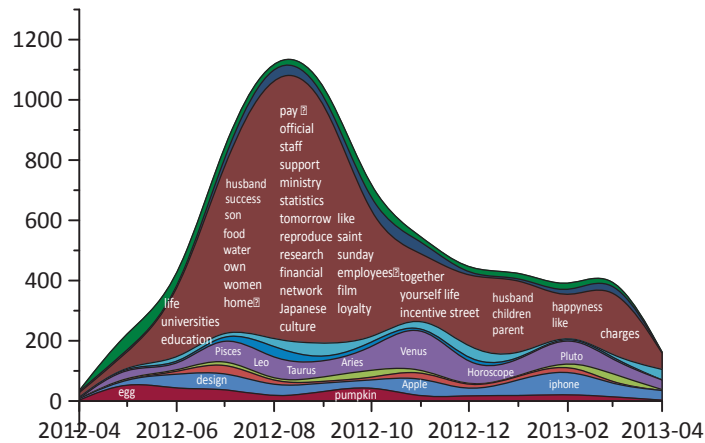
Figure 1: An example of the topics of retweets in each month. Each colored stripe represents a topic, whose height is the number of words assigned to the topic. For each topic, the top words of this topic in each month are placed on the stripe.

Motivated by the methods proposed to handle the vocabulary gap problem for keyphrase extraction (Liu et al., 2012) and hashtag suggestion (Ding et al., 2013), in this work, we also assume that the hashtags and textual content in a microblog are parallel descriptions of the same thing in different languages. To model the document themes, in this paper, we adopt the topical translation model to facilitate the translation process. Topic-specific word triggers are used to bridge the gap between the words and hashtags. Since existing topical translation models can only recommend hashtags learned from the training data, we also incorporate an extraction process into the model.

This work makes three main contributions. First, we incorporate temporal and personal factors into considerations. Most of the existing works on hashtag recommendation tasks have focused on textual information. Second, we adopt a topical translation model to combine extraction and translation methods. This makes it possible to suggest hashtags that are not included in the training data. Third, to evaluate the task, we construct a large collection of microblogs from a real microblogging service. All of the microblogs in the collection contain textual content and hashtags labeled by their authors. This can benefit other researchers investigating the same task or other topics using author-centered data.

The remaining part of this paper is structured as follows: We briefly review existing methods in related domains in Section 2. Section 3 gives an overview of the proposed generation model. Section 4 introduces the dataset construction, experimental results and analyses. In Section 5, we will conclude the paper.

2 Related Works

Due to the usefulness of tag recommendation, many methods have been proposed from different perspectives (Heymann et al., 2008; Krestel et al., 2009; Rendle et al., 2009; Liu et al., 2012; Ding et al., 2013). Heymann et al. (Heymann et al., 2008) investigated the tag recommendation problem using the data collected from social bookmarking system. They introduced an entropy-based metric to capture the generality of a particular tag. In (Song et al., 2008), a Poisson Mixture Model based method is introduced to achieve the tag recommendation task. Krestel et al. (Krestel et al., 2009) introduced a Latent Dirichlet Allocation to elicit a shared topical structure from the collaborative tagging effort of multiple users for recommending tags. Based on the the observation that similar webpages tend to have the same tags, Lu et al. proposed a method taking both tag information and page content into account to achieve the task (Lu et al., 2009). Ding et al. proposed to use translation process to model this task (Ding et al., 2013). They extended the translation based method and introduced a topic-specific translation model to process the various meanings of words in different topics. In (Tariq et al., 2013), discriminative-term-weights were used to establish topic-term relationships, of which users' perception were learned to suggest suitable hashtags for users. To handle the vocabulary problem in keyphrase extraction task, Liu et al. proposed a

204

topical word trigger model, which treated the keyphrase extraction problem as a translation process with latent topics (Liu et al., 2012).

Most of the works mentioned above are based on textual information. Besides these methods, personalized methods for different recommendation tasks have also been paid lots of attentions (Liang et al., 2007; Shepitsen et al., 2008; Garg and Weber, 2008; Li et al., 2010; Liang et al., 2010; Rendle and Schmidt-Thieme, 2010). Shepitsen et al. (2008) proposed to use hierarchical agglomerative clustering to take into account personalized navigation context in cluster selection. In (Garg and Weber, 2008), the problem of personalized, interactive tag recommendation was also studied based on the statics of the tags co-occurrence. Liang et al. (2010) proposed to the multiple relationships among users, items and tags to find the semantic meaning of each tag for each user individually and used this information for personalized item recommendation.

From the brief descriptions given above, we can observe that most of the previous works on hashtag suggestion focused on textual information. In this work, we propose to incorporate temporal and personal information into the generative methods. Further more, to over the limitation that translation based method can only recommend hashtags learned from the training data, we also propose to incorporate an extraction process into the model.

3 The Proposed Methods

In this section, we firstly introduce the notation and generation process of the proposed method. Then, we describe the method used for learning parameters. Finally, we present the methods of how do we apply the learned model to achieve the hashtag recommendation task.

3.1 The Generation Process

We use D to represent the number of microblogs in the given corpus, and the microblogs have been divided into T epoches. Let $t = 1, 2, ..., T$ be the index of an epoches, θ^t is the topic distribution of the epoch t. Each microblog 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. A microblog is a sequence of N_d words denoted by $w_d = \{w_{d1}, w_{d2}, ..., w_{dN_d}\}$. Each microblog contains a set of hashtags denoted by $h_d = \{h_{d1}, h_{d2}, ..., h_{dM_d}\}$. A word is defined as an item from a vocabulary with W distinct words indexed by $w = \{w_1, w_2, ..., w_W\}$. Each hashtag is from the vocabulary with V distinct hashtags indexed by $h = \{h_1, h_2, ..., h_V\}$. The notations in this paper are summarized in Table 1.

The original LDA assumes that a document is contains a mixture of topics, which is represented by a topic distribution, and each word has a hidden topic label. Although, it is sensible for long document, due to the limitations of the length of characters in a single microblog, it tends to be about a single topic. Hence, we associate a single hidden variable with each microblog to indicate its topic. Similar idea of assigning a single topic to a short sequence of words has also been used for modeling Twitters (Zhao et al., 2011)

The hashtag recommendation task is to discover a list of hashtags for each unlabeled microblog, In our method, we first learn a topical translation model, and then we estimate the latent variables for each microblog, finaly recommending hashtags accord to the learned model.

Fig. 2 shows the graphical representation of the generation process. The generative story for each microblog is as follows:

3.2 Learning

To learn the parameters of our model, we use collapsed Gibbs sampling (Griffiths and Steyvers, 2004) to sample the topics assignment $\mathbf{z}$, latent variables assignment $\mathbf{x}$ and $\mathbf{y}$.

Given the current state of all but the variable x_d and z_d for the dth microblog, we can jointly sample

1. Draw $\pi \sim Beta(\delta)$, $\eta \sim Beta(\lambda)$

2. Draw background word distribution $\phi^B \sim Dirichlet(\beta^w)$

3. Draw global trendy topic distribution $\theta^t \sim Dirichlet(\alpha)$ for each time epoch $t = 1, 2, ..., T$

4. Draw personal topic distribution $\psi^u \sim Dirichlet(\alpha)$ for each user $u = 1, 2, ..., U$

5. Draw word distribution $\phi^z \sim Dirichlet(\beta^w)$ for each topic $z = 1, 2, ..., K$

6. Draw hashtag distribution $\varphi_{z,w} \sim Dirichilet(\beta^h)$ for each topic $z = 1, 2, ..., K$ and each word $w = 1, 2, ..., W$

7. For each microblog $d = 1, 2, ..., D$

   a. Draw $x_d \sim Bernoulli(\eta)$

   b. If $x_d = 0$ then

   Draw a topic $z_d \sim Multinomial(\psi^u)$

   End if

   If $x_d = 1$ then

   Draw a topic $z_d \sim Multinomial(\theta^t)$

   End if

   c. For each word $n = 1, ..., N_d$

   i. Draw $y_{dn} \sim Bernoulli(\pi)$

   ii. If $y_{dn} = 0$ then

   Draw a word $w_{dn} \sim Multinomial(\phi^B)$

   End if

   If $y_{dn} = 1$ then

   Draw a word $w_{dn} \sim Multinomial(\phi^{z_d})$

   End if

   d. For each hashtag $m = 1, ..., M_d$

   i. Draw $h_{dm} \sim P(h_{dm}|w_d, z_d, \varphi_{z_d, w_d})$
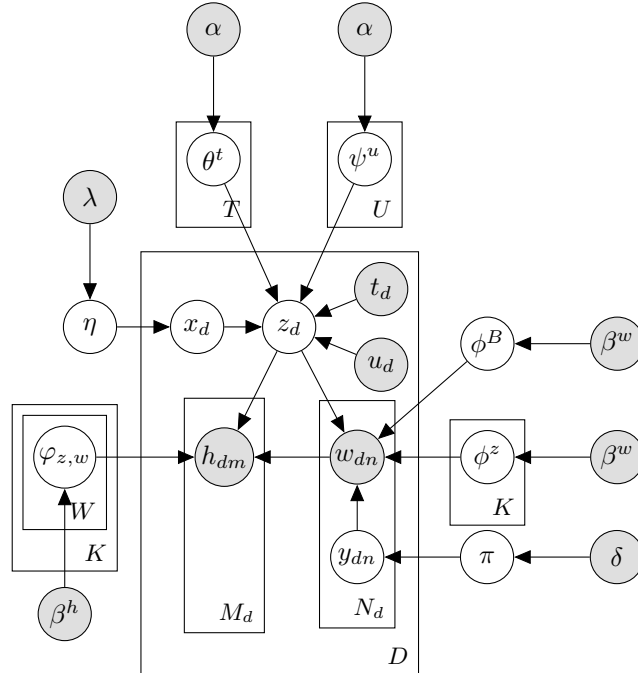

Figure 2: The graphical representation of the proposed model. Shaded circles are observations or constants. Unshaded ones are hidden variables.

Table 1: The notations used in this work.

D	The number of training data set
W	The number of unique word in the corpus
V	The number of unique hashtag in the corpus
K	The number of topics
T	The total number of time epoches
U	The total number of users
N_d	The number of words in the dth microblog
M_d	The number of hashtags in the dth microblog
z_d	The topic of the dth microblog
x_d	The latent variable decided the distribution category of z_d
y_{dn}	The latent variable decided the distribution category of w_{dn}
π	The distribution of latent variable y_{dn}
η	The distribution of latent variable x_d
ϕ^z	The distribution of topic words
ϕ^B	The distribution of background words
θ^t	The distribution of topics for time epoch t
ψ^u	The distribution of topics for user u
t_d	The time epoch for microblog d
u_d	The user of the microblog d
φ	The topic-specific word alignment table between word and hashtag or itself

x_d and z_d, the conditional probability of $x_d = p$,$z_d = k$ is calculated as follows:

$$Pr(x_d = p, z_d = k | \mathbf{z}_{\neg \mathbf{d}}, \mathbf{x}_{\neg \mathbf{d}}, \mathbf{y}, \mathbf{w}, \mathbf{h})$$

$$\propto \frac{N_p^\eta + \lambda}{N_{(.)}^\eta + 2\lambda} \cdot \frac{N_k^l + \alpha}{N_{(.)}^l + K\alpha} \cdot \prod_{n=1}^{N_d} \frac{N_{w_{dn}}^k + \beta^w}{N_{(.)}^k + W\beta^w} \cdot \prod_{m=1}^{M_d} \sum_{n=1}^{N_d} \frac{M_{\neg d,k}^{w_{dn},h_{dm}} + \beta^h}{M_{\neg d,k}^{w_{dn},(.)} + V\beta^h}, \tag{1}$$

where $l = u_d$ when $p = 0$ and $l = t_d$ when $p = 1$. N_0^η is the number of microblog generated by personal interests, while N_1^η is the number of microblog coming from global topical trends, $N_{(.)}^\eta = N_0^\eta + N_1^\eta$. $N_k^{u_d}$ is the number of microblogs generated by user u_d and under topic k. $N_{(.)}^{u_d}$ is the total number of microblogs generated by user u_d. $N_k^{t_d} = \sum_{t'=1}^{t_d} e^{\frac{-t'}{\rho}} N'_k{}^{t-t'}$, $N'_k{}^{t-t'}$ is the number of microblogs assigned to topic k at time epoch $t - t'$, $e^{\frac{-t'}{\rho}}$ is decay factory, and $N_{(.)}^{t_d} = \sum_{k=1}^{K} N_k^{t_d}$. $N_{w_{dn}}^k$ is the times of word w_{dn} assigned to topic k, $N_{(.)}^k$ is the times of all the word assigned to topic k, $M_{\neg d,k}^{w_{dn},h_{dm}}$ is the number of occurrences that word w_{dn} is translated to hashtag h_{dm} given topic k. All the counters mentioned above are calculated with the dth microblog excluded.

We sample y_{dn} for each word w_{dn} in the dth microblog using the following equation:

$$Pr(y_{dn} = q | \mathbf{z}, \mathbf{x}, \mathbf{y}_{\neg \mathbf{dn}}, \mathbf{w}, \mathbf{h}) \propto \frac{N_q^\pi + \delta}{N_{(.)}^\pi + 2\delta} \cdot \frac{N_{w_{dn}}^l + \beta^w}{N_{(.)}^l + W\beta^w}, \tag{2}$$

where $l = B$ when $q = 0$ and $l = z_d$ when $q = 1$. N_0^π is the number of words assigned to background words and N_1^π is the number of words under any topic respectively. $N_{(.)}^\pi = N_0^\pi + N_1^\pi$, $N_{w_{dn}}^B$ is a count of word w_{dn} occurs as a background word. $N_{w_{dn}}^{z_d}$ is the number of word w_{dn} is assigned to topic z_d, and $N_{(.)}^{z_d}$ is the total number of words assigned to topic z_d. All counters are calculated with taking no account of the current word w_{dn}.

In many cases, hashtag dose not appear in the training data, to solve this problem, we assume that each word in the microblog can translate to a hashtag in the training data or itself. We assume that each word

have aligned σ (we set $\sigma = 1$ in this paper after trying some number) times with itself under the specific topic. After all the hidden variables become stable, we can estimate the alignment probability as follows:

$$\varphi_{h,w,z} = \begin{cases} \frac{N_{z,w}^h + \beta^h}{N_{z,w}^{(.)} + \sigma + (V+1)\beta^h} & \text{if } h \text{ is a hashtag in the training data} \\ \frac{\sigma + \beta^h}{N_{z,w}^{(.)} + \sigma + (V+1)\beta^h} & \text{if } h \text{ is the word itself} \end{cases} \tag{3}$$

where $N_{z,w}^h$ is the number of the hashtag h co-occurs with the word w under topic z in the microblogs.

For the probability alignment φ between hashtag and word, the potential size is $W \cdot V \cdot K$. The data sparsity poses a more serious problem in estimating φ than the topic-free word alignment case. To remedy the problem, we use interpolation smoothing technique for φ. In this paper, we employ smoothing as follows:

$$\varphi_{h,w,z}^* = \gamma \varphi_{h,w,z} + (1-\gamma)P(h|w), \tag{4}$$

where $\varphi_{h,w,z}^*$ is the smoothed topical alignment probabilities, $\varphi_{h,w,z}$ is the original topical alignment probabilities. $P(h|w)$ is topic-free word alignment probability. Here we obtain $P(h|w)$ by exploring IBM model-1 (Brown et al., 1993). γ is trade-off of two probabilities ranging from 0.0 to 1.0. When $\gamma = 0.0$, $\varphi_{h,w,z}^*$ will be reduce to topic-free word alignment probability; and when $\gamma = 1.0$, there will be no smoothing in $\varphi_{h,w,z}^*$. For the word itself there are no smoothing, because it is a pseudo-count.

3.3 Hashtag Extraction

We perform hashtag extraction as follows. Suppose given an unlabeled dataset, we perform Gibbs Sampling to iteratively estimate the topic and determine topic/background words for each microblog. The process is the same as described in Section 3.2. After the hidden variables of topic/background words and the topic of each microblog become stable, we can estimate the distribution of topics for the dth microblog in unlabeled data by: $\chi_{dk}^* = \frac{p(k)p(w_{d1}|k)...p(w_{dN_d}|k)}{Z}$ where $p(w_{dn}|k) = \frac{N_1^\pi + \delta}{N_{(.)}^\pi + 2\delta} \cdot \frac{N_{w_{dn}}^k + \beta^w}{N_{(.)}^k + W\beta^w}$ and $N_{w_{dn}}^k$ is the number of words w_{dn} that are assigned to topic k in the corpus, and $p(k) = \frac{N_0^\eta + \lambda}{N_{(.)}^\eta + 2\lambda} \cdot \frac{N_k^u + \alpha}{N_{(.)}^u + K\alpha} + \frac{N_1^\eta + \lambda}{N_{(.)}^\eta + 2\lambda} \cdot \frac{N_k^t + \alpha}{N_{(.)}^t + K\alpha}$ is regarded as a prior for topic distribution, Z is the normalized factor. With topic distribution χ^* and topical alignment table φ^*, we can rank hashtags for the dth microblog in unlabeled data by computing the scores:

$$P(h_{dm}|w_d, \chi_d^*, \varphi^*) \propto \sum_{z_d=1}^{K} \sum_{n=1}^{N_d} P(h_{dm}|z_d, w_{dn}, \varphi^*) \cdot P(z_d|\chi_d^*) \cdot P(w_{dn}|w_d), \tag{5}$$

where h_{dm} can be a hashtag in the training data or a word in the dth microblog, $p(w_{dn}|w_d)$ is the weight of the word w_{dn} in the microblog, which can be estimated by the IDF score of the word. According to the ranking scores, we can suggest the top-ranked hashtags for each microblog to users.

4 Experiments

In this section, we introduce the experimental results and the data collection we constructed for training and evaluation. Firstly, we describe how do we construct the collection and statics of it. Then we introduce the experiment configurations and baseline methods. Finally, the evaluation results and analysis are given.

4.1 Data Collection

We use a dataset collected from Sina Weibo to evaluate the proposed approach and alternative methods. We random select 166,864 microblogs from Aug. 2012 to June 2013. The unique number of hashtags in the corpus is 17,516. We use the microblogs posted from Aug. 2012 to May 2013 as the training data. The other microblogs are used for evaluation. The hashtags marked in the original microblogs are considered as the golden standards.

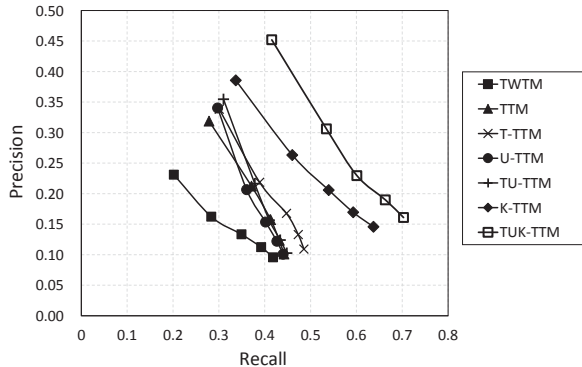Figure 3: Precision-recall curves of different methods on this task.

Table 2: Evaluation results of different methods on the evaluation collection.

Methods	Precision	Recall	F_1
TWTM	0.231	0.202	0.215
SVM	0.418	0.366	0.390
TTM	0.319	0.279	0.297
T-TTM	0.338	0.301	0.319
U-TTM	0.341	0.307	0.323
K-TTM	0.386	0.337	0.360
TU-TTM	0.355	0.310	0.331
TUK-TTM	**0.452**	**0.415**	**0.433**

4.2 Experiment Configurations

We use precision (P), recall (R), and F1-score (F_1) to evaluate the performance. Precision is calculated based on the percentage of "hashtags truly assigned" among "hashtags assigned by system". Recall is calculated based on the "hashtags truly assigned" among "hashtags manually assigned". F1-score is the harmonic mean of precision and recall. We do 500 iterations of Gibbs sampling to train the model. For optimize the hyperparmeters of the proposed method and alternative methods, we use 5-fold cross-validation in the training data to do it. The number of topics is set to 70. The other settings of hyperparameters are as follows: $\alpha = 50/K$, $\beta^w = 0.1$, $\beta^h = 0.1$, $\lambda = 0.01$, and $\delta = 0.01$. The smoothing factor γ in Eq.(3) is set to 0.6. For estimating the translation probability without topical information, we use GIZA++ 1.07 to do it (Och and Ney, 2003).

For baselines, we compare the proposed model with the following alternative models.

- **TWTM**: Topical word trigger model (TWTM) was proposed by Liu et al. for keyphrase extraction using only textual information (Liu et al., 2012). We implemented the model and used it to achieve the task.

- **TTM**: Ding et al. (2013) proposed the topical translation model (TTM) for hash tag extraction. We implemented and extended their method for evaluating it on the corpus constructed in this work.

4.3 Experimental Results

Table 2 shows the comparisons of the proposed method with the state-of-the-art methods on the constructed evaluation dataset. "TUK-TTM" denotes the method proposed in this paper. "T-TTM" and "U-TTM" represent the methods incorporating temporal and personal information respectively. "K-TTM" represents the method incorporating the extraction factor. From the results, we can observe that the proposed method is significantly better than other methods at 5% significance level (two-sided). Comparing to results of the TTM, we can observe that the temporal information, personal information and extraction strategy can all benefit the task. Among the three additional factors, the extraction strategy achieves the best result. The limitation of only being able to recommend hashtags that exist in the training data can be overcome in some degree by the proposed method. The relative improvement of proposed TUK-TTM over TTM is around 47.8% in F1-score.

Table 3 shows the comparisons of the proposed method with the method "K-TTM" in two corpus NE-Corpus and E-Corpus. NE-Corpus include microblogs whose hashtags are not contained in the training data. E-Corpus include the microblogs whose hashtags appear in the training data. We can observe that the proposed method significantly better than "K-TTM" in the E-Corpus. Another observation is that the method incorporating the extraction factor achieves better performances on the NE-Corpus than E-Corpus. We think that the reason is that the fewer times hashtag appear, the greater weight it has. Hence, we can extract this kind of hashtags more easier.

Figure 3 shows the precision-recall curves of TWTW, TTM, T-TTM, U-TTM, TU-TTM, K-TTM, and TUK-TTM on the evaluation dataset. Each point of a precision-recall curve represents extracting

Table 3: Evaluation results of two different corpus.

Corpus	Methods	P	R	F
NE-Corpus	K-TTM	0.631	0.553	0.589
	TUK-TTM	0.641	0.561	0.598
E-Corpus	K-TTM	0.172	0.162	0.167
	TUK-TTM	0.288	0.271	0.279

Table 4: The influence of the number of topics K of TUK-TTM.

K	Precision	Recall	F_1
10	0.410	0.382	0.396
30	0.435	0.380	0.406
50	0.448	0.413	0.430
70	**0.452**	**0.415**	**0.433**
100	0.439	0.404	0.421

Table 5: The influence of the smoothing parameter γ of TUK-TTM.

γ	Precision	Recall	F_1
0.0	0.379	0.354	0.366
0.2	0.405	0.372	0.388
0.4	0.433	0.398	0.415
0.6	**0.452**	**0.415**	**0.433**
0.8	0.426	0.386	0.405
1.0	0.423	0.381	0.401

different number of hashtags ranging from 1 to 5 respectively. In the figure, curves which are close to the upper right-hand corner of the graph indicate the better performance. From the results, we can observe that the performance of TUK-TTM is in the upper right-hand corner. It also demonstrates that the proposed method achieves better performances than other methods.

From the description of the proposed model, we can know that there are several hyperparameters in the proposed TUK-TTM. To evaluate the impacts of them, we evaluate two crucial ones, the number of topics K and the smoothing factor γ. Table 4 shows the influence of the number of topics. From the table, we can observe that the proposed model obtains the best performance when K is set to 70. And performance decreases with more number of topics. We think that data sparsity may be one of the main reasons. With much more topic number, the data sparsity problem will be more serious when estimating topic-specific translation probability. Table 5 shows the influence of the translation probability smoothing parameter γ. When γ is set to 0.0, it means that the topical information is omitted. Comparing the results of $\gamma = 0.0$ and other values, we can observe that the topical information can benefit this task. When γ is set to 1.0, it represents the method without smoothing. The results indicate that it is necessary to address the sparsity problem through smoothing.

5 Conclusions

In this paper, we propose a novel method which incorporates temporal and personal factors into the topical translation model for hashtag recommendation task. Since existing translation model based methods for this task can only recommend hashtags that exist in the training data of the topical translation model, we also incorporate extraction strategies into the model. To evaluate the proposed method, we also construct a dataset from real world microblogging services. The results of experiments on the dataset demonstrate that the proposed method outperforms state-of-the-art methods that do not consider these aspects.

6 Acknowledgement

References

A.Bandyopadhyay, M. Mitra, and P. Majumder. 2011. Query expansion for microblog retrieval. In *Proceedings of The Twentieth Text REtrieval Conference*, TREC 2011.

S. Asur and B.A. Huberman. 2010. Predicting the future with social media. In *WI-IAT'10*, volume 1, pages 492–499.

Hila Becker, Mor Naaman, and Luis Gravano. 2010. Learning similarity metrics for event identification in social media. In *Proceedings of WSDM '10*.

Adam Bermingham and Alan F. Smeaton. 2010. Classifying sentiment in microblogs: is brevity an advantage? In *Proceedings of CIKM'10*.

D.M. Blei and M.I. Jordan. 2003. Modeling annotated data. In *Proceedings of SIGIR*, pages 127–134.

Johan Bollen, Huina Mao, and Xiaojun Zeng. 2011. Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1):1 – 8.

Peter F Brown, Vincent J Della Pietra, Stephen A Della Pietra, and Robert L Mercer. 1993. The mathematics of statistical machine translation: Parameter estimation. *Computational linguistics*, 19(2):263–311.

Dmitry Davidov, Oren Tsur, and Ari Rappoport. 2010. Enhanced sentiment learning using twitter hashtags and smileys. In *Proceedings of COLING '10*.

Zhuoye Ding, Xipeng Qiu, Qi Zhang, and Xuanjing Huang. 2013. Learning topical translation model for microblog hashtag suggestion. In *Proceedings of IJCAI 2013*.

Miles Efron. 2010. Hashtag retrieval in a microblogging environment. In *Proceedings of SIGIR '10*.

Nikhil Garg and Ingmar Weber. 2008. Personalized, interactive tag recommendation for flickr. In *Proceedings of RecSys '08*.

T. L. Griffiths and M. Steyvers. 2004. Finding scientific topics. *Proceedings of the National Academy of Sciences*.

Ido Guy, Naama Zwerdling, Inbal Ronen, David Carmel, and Erel Uziel. 2010. Social media recommendation based on people and tags. In *Proceedings of SIGIR '10*.

Ido Guy, Uri Avraham, David Carmel, Sigalit Ur, Michal Jacovi, and Inbal Ronen. 2013. Mining expertise and interests from social media. In *Proceedings of WWW '13*.

Paul Heymann, Daniel Ramage, and Hector Garcia-Molina. 2008. Social tag prediction. In *Proceedings of SIGIR '08*.

Long Jiang, Mo Yu, Ming Zhou, Xiaohua Liu, and Tiejun Zhao. 2011. Target-dependent twitter sentiment classification. In *Proceedings of ACL 2011*, Portland, Oregon, USA.

Ralf Krestel, Peter Fankhauser, and Wolfgang Nejdl. 2009. Latent dirichlet allocation for tag recommendation. In *Proceedings of RecSys '09*.

Lihong Li, Wei Chu, John Langford, and Robert E Schapire. 2010. A contextual-bandit approach to personalized news article recommendation. In *Proceedings of the 19th international conference on World wide web*, pages 661–670. ACM.

Ting-Peng Liang, Hung-Jen Lai, and Yi-Cheng Ku. 2007. Personalized content recommendation and user satisfaction: Theoretical synthesis and empirical findings. *Journal of Management Information Systems*, 23(3):45–70.

Huizhi Liang, Yue Xu, Yuefeng Li, Richi Nayak, and Xiaohui Tao. 2010. Connecting users and items with weighted tags for personalized item recommendations. In *Proceedings of the 21st ACM conference on Hypertext and hypermedia*, pages 51–60. ACM.

Zhiyuan Liu, Chen Liang, and Maosong Sun. 2012. Topical word trigger model for keyphrase extraction. In *Proceedings of COLING*.

Yu-Ta Lu, Shoou-I Yu, Tsung-Chieh Chang, and Jane Yung-jen Hsu. 2009. A content-based method to enhance tag recommendation. In *Proceedings of IJCAI'09*.

Franz Josef Och and Hermann Ney. 2003. A systematic comparison of various statistical alignment models. *Computational Linguistics*, 29(1):19–51.

Tsutomu Ohkura, Yoji Kiyota, and Hiroshi Nakagawa. 2006. Browsing system for weblog articles based on automated folksonomy. *Workshop on the Weblogging Ecosystem Aggregation Analysis and Dynamics at WWW*.

Takanobu Otsuka, Takuya Yoshimura, and Takayuki Ito. 2012. Evaluation of the reputation network using realistic distance between facebook data. In *Proceedings of WI-IAT '12*.

Bo Pang and Lillian Lee. 2008. Opinion mining and sentiment analysis. *Found. Trends Inf. Retr.*, 2(1-2):1–135, January.

Steffen Rendle and Lars Schmidt-Thieme. 2010. Pairwise interaction tensor factorization for personalized tag recommendation. In *Proceedings of the third ACM international conference on Web search and data mining*, pages 81–90. ACM.

Steffen Rendle, Leandro Balby Marinho, Alexandros Nanopoulos, and Lars Schmidt-Thieme. 2009. Learning optimal ranking with tensor factorization for tag recommendation. In *Proceedings of KDD '09*.

Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo. 2010. Earthquake shakes twitter users: real-time event detection by social sensors. In *Proceedings of WWW '10*.

Andriy Shepitsen, Jonathan Gemmell, Bamshad Mobasher, and Robin Burke. 2008. Personalized recommendation in social tagging systems using hierarchical clustering. In *Proceedings of the 2008 ACM Conference on Recommender Systems*, RecSys '08, pages 259–266, New York, NY, USA. ACM.

Yang Song, Ziming Zhuang, Huajing Li, Qiankun Zhao, Jia Li, Wang-Chien Lee, and C. Lee Giles. 2008. Real-time automatic tag recommendation. In *Proceedings of SIGIR '08*.

Amara Tariq, Asim Karim, Fernando Gomez, and Hassan Foroosh. 2013. Exploiting topical perceptions over multi-lingual text for hashtag suggestion on twitter. In *FLAIRS Conference*.

Xiaolong Wang, Furu Wei, Xiaohua Liu, Ming Zhou, and Ming Zhang. 2011. Topic sentiment analysis in twitter: a graph-based hashtag sentiment classification approach. In *Proceedings of CIKM '11*.

Wayne Xin Zhao, Jing Jiang, Jing He, Yang Song, Palakorn Achananuparp, Ee-Peng Lim, and Xiaoming Li. 2011. Topical keyphrase extraction from twitter. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1*, pages 379–388. Association for Computational Linguistics.