2013 IEEE 13th International Conference on Data Mining

Spatio-Temporal Topic Modeling in Mobile Social Media for Location Recommendation

Bo Hu
School of Computing Science
Simon Fraser University
Burnaby, BC, Canada
boh@cs.sfu.ca

Mohsen Jamali
School of Computing Science
University of British Columbia
Vancouver, BC, Canada
jamalim@cs.ubc.ca

Martin Ester
School of Computing Science
Simon Fraser University
Burnaby, BC, Canada
ester@cs.sfu.ca

***Abstract*—Mobile networks enable users to post on social media services (e.g., Twitter) from anywhere and anytime. This new phenomenon led to the emergence of a new line of work of mining the behavior of mobile users taking into account the spatio-temporal aspects of their engagement with online social media.**

In this paper, we address the problem of recommending the right locations to users at the right time. We claim to propose the first comprehensive model, called STT (Spatio-Temporal Topic), to capture the spatio-temporal aspects of user check-ins in a single probabilistic model for location recommendation. Our proposed generative model does not only captures spatio-temporal aspects of check-ins, but also profiles users. We conduct experiments on real life data sets from Twitter, Gowalla, and Brightkite. We evaluate the effectiveness of STT by evaluating the accuracy of location recommendation. The experimental results show that STT achieves better performance than the state-of-the-art models in the areas of recommender systems as well as topic modeling.

***Keywords*-spatio-temporal; topic model; location recommendation;**

I. Introduction

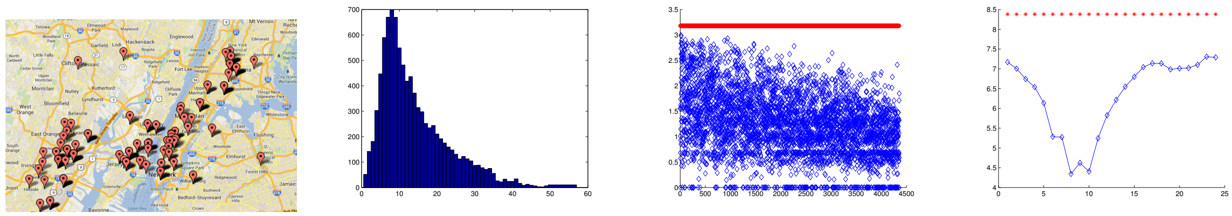
Location recommender systems can recommend a set of locations that users may be interested in, based on the history of user check-ins. A recent work [9] extends LDA (Latent Dirichlet Allocation) for location recommendation and addresses the spatial aspects of user check-ins by capturing the phenomenon of *geographical influence* [15]. Geographical influence suggests that locations that are closer to the user's visited locations are recommended with higher probabilities, and the existing methods [15], [9] assume that the location recommendations to a user should be geographically regularized by the set of all of the user's check-ins. However, this assumption does not hold when the check-ins of a user are spread over multiple regions. Let us take an example of a user, who commutes between two cities (regions). A good recommendation should be one of those locations in either one of the regions, but the existing models will recommend locations along the commute route, since they are on average closer to the user's check-ins in both regions.

Thus, our first observation is that users may have multiple visited regions. To further illustrate this point, let us take an example of a collection of check-ins from NYC (New York City) on Twitter. Figure 1(a) shows an example of a user with all his/her check-ins, which shows that this user has three frequently visited regions (Newark, Jersey City and Lower Manhattan). To further analyze the number of regions that users have visited, we cluster all check-ins into 100 regions by k-means based on their coordinates and plot the histogram in Figure 1(b), where the x-asis represents the number of regions and the y-asis represents the number of users. We observe that users have visited 14 regions on average, and most users have visited 10 to 20 regions.

Our second observation is that different locations have different "temporal activity patterns", i.e., probability distributions of check-ins over relative timestamps. For example, we use 1-24 to represent the time of the day in hours. Business, residence, and entertainment locations have different daily activity patterns, and their daily activity normally peaks in the morning, evening, and late night, respectively. In Figure 1(c), we plot the entropies of the daily activity patterns of all locations in NYC from the same Twitter data set. Note that higher entropies indicate more uncertainty (more Uniform distribution of check-ins over the day) than lower ones, and the top red curve represents the entropies of Uniform distributions (maximum entropies). The figure shows that a large number of locations have very small entropies, which implies that daily activity patterns do affect the check-ins at these locations. Figure 1(d) shows the entropies of the distribution of check-ins over locations for the 24 different hours, indicating that user movements are more predicable in the period from 5AM to 12AM than in the rest of the day.

Based on the above observations, we propose the **S**patio-**T**emporal **T**opic (**STT**) model of check-ins that takes the geographical influence and temporal activity patterns into account. Basically, a check-in is represented by a user, a location with a pair of coordinates, and a relative timestamp, which are all considered as observed random variables. Similar to LDA, a set of latent topics is defined. Each user is associated with a probability distribution over topics, which captures the user interests, and each topic has a probability distribution over locations, which captures the semantic relationship between locations. Topics are assumed to represent sets of locations that have similar functions such as parks, night clubs, or restaurants. Each check-in is assigned to a topic. To model multiple regions of users, STT assumes that there is a set of

1550-4786/13 $31.00 © 2013 IEEE
DOI 10.1109/ICDM.2013.139



(a) Check-ins of a sample user are from three regions (Newark, Jersey City and Lower Manhattan).
(b) The histogram of the number of users versus the number of regions that users have visited.
(c) The entropy of daily activity patterns at 3,518 different locations.
(d) The entropy of the distribution of check-ins over locations for the 24 different hours.

Fig. 1. An example of check-ins in New York City on Twitter.

latent regions, and each user is associated with a probability distribution over regions. Instead of regularizing locations to be close to all the user's check-ins, STT regularizes them to the center of the sampled region. Additionally, STT considers temporal activity patterns. It selects a topic of a check-in based on its user's and time's topic distributions, and it generates (recommends) a location based on the topic and time dependent location distributions.

We propose an EM (Expectation-Maximization) algorithm to learn the latent random variables and parameters of STT that maximizes the likelihood of observed random variables. We perform experiments on real life data sets from Twitter, Gowalla and Brightkite. We evaluate the effectiveness of STT and of state-of-the-art models in terms of the perplexity of the test data set, and the accuracy of location recommendation.

The major contributions of this paper are as follows:

- We propose the first spatio-temporal topic model for location recommendation, capturing the geographical influence between user regions and locations, and temporal activity patterns of different topics and locations.
- We employ the sparse coding technique which greatly speeds up the learning process.
- We demonstrate that the proposed STT model consistently improves the test perplexity and the average accuracy@1,5,10 for location recommendation, compared to existing state-of-the-art recommendation algorithms and geographical and temporal topic models.

II. Related Work

Location Recommendation. Many traditional recommendation algorithms, such as MF (Matrix Factorization) [8], [11], can be employed for location recommendation. LDA (Latent Dirichlet Allocation) [1] can also be used for location recommendation. However, none of them considers the spatio-temporal aspects for location recommendation.

Recent works [14], [15], [17], [2], [9] focus on capturing the geographical influence for location recommendation. [15], [14], [17] extend CF, [2] extends MF, and [9] proposes GLDA (Geo Latent Dirichlet Allocation) that extends LDA. Although [14], [17], [15], [2] capture the geographical influence, and the user and item factors (interests), they fuse these two parts in a linear regression framework with user input weights. As a result, the non-unified framework cannot exploit the benefit of the mutually reinforcement between both parts. Hence, the most recent related work [9] models both parts in a probabilistic framework. The intuition is that locations that are closer to the user's visited locations are recommended with higher probabilities. The idea works when users' locations are geographically cohesive to each other, but it fails when users visit multiple regions (as discussed in the introduction).

Note that all the above works ignore the temporal aspect of user check-ins which is an essential component in our problem definition.

Geographical and Temporal Topic Modeling. Recently, there are many works [5], [16], [6] in the area of geographical topic modeling, which detect geographical topics from documents that are associated with locations. [13], [7] propose temporal topic models, which assume the words of a document are drawn from the user specific topic distribution and the timestamps are drawn from that topic as well.

All the above works model either spatial or temporal aspects of mobile social media, while our proposed model takes both of them into account.

User Movement Analysis. Some works [3], [4], [10] have been studying user movements in location-based social networks. Since most user posts are not associated with coordinates, these works address the following problem: given a set of geo-tagged posts from many users, learn a model of region specific words, and apply this model to predict the user location of un-tagged posts based on their content. This is different from the problem in this paper, which is to recommend the locations to users.

III. Spatio-Temporal Topic Model

A. Problem Definition

We first introduce the notations needed in our problem and listed in Table I. We assume that all the online activities (also known as "check-ins") of mobile users are authored by a user from a fixed set U with size $|U|$. Note that we use capital letters to represent the sets and the $|.|$ sign to represent the size of the sets in this paper. We associate each user u with a set of check-ins D_u, and each check-in is represented by a user, a location with a pair of latitude and longitude coordinates, and a timestamp. Formally, a check-in d is defined by $d = \{u, i, l_i, t\}$, where u, i, t represents the (index of) user, location, and timestamp, respectively. l_i represents the coordinates of location i, and the values of t are discrete, e.g., 1-7 represent Sunday, Monday, ..., Saturday in a week and 1-24 represent hour 1, 2, ..., 24 in a day. A check-in collection D is defined as a set of check-ins from all users.

We assume that there is a set of latent topics Z and a set of latent regions R in the collection D. Each check-in d is assigned to one of the topics (z_d) and regions (r_d). To model user interests and movements, users are associated with topic distributions θ^{user} and region distributions η^{user}. A "semantically" coherent topic in the collection D is associated with a probability distribution over all locations ψ^{topic}. Additionally, we use θ^{time} to represent time dependent topic distributions, e.g., check-ins from a night club topic usually happen on weekends and at late hours of the day, and ψ^{time} to represent time dependent location distributions. A region has a geographical center μ, and it is associated with a set of check-ins, which are coherent in topics and close to the center geographically. Finally, θ^0, η^0, and ψ^0 represent the background distributions for topics, regions, and locations, respectively.

TABLE I
NOTATIONS OF PARAMETERS

Variable	Interpretation
$i_{u,d}$	location index of the d^{th} check-in by the u^{th} user
l_i	latitude and longitude coordinates of the i^{th} location
$t_{u,d}$	relative timestamp in the d^{th} check-in by the u^{th} user
$z_{u,d}$	topic assignment of the d^{th} check-in by the u^{th} user
$r_{u,d}$	region assignment of the d^{th} check-in by the u^{th} user
D_u	set of check-ins of the u^{th} user
θ_u^{user}	topic distribution of the u^{th} user
θ_t^{time}	topic distribution of the t^{th} timestamp
η_u^{user}	region distribution of the u^{th} user
ψ_z^{topic}	location distribution of the z^{th} topic
ψ_t^{time}	location distribution of the t^{th} timestamp
μ_r	region mean of the r^{th} region

Based on the above definitions, we formalize our research problem as follows:

Problem 1: **Spatio-Temporal Topic Modeling for Location Recommendation**. Given a check-in collection D, and numbers $|Z|$ of topics and $|R|$ of regions, the task is to model and learn the spatio-temporal parameters of users, topics, regions and locations for location recommendation.

B. Model

Figure 2 shows the graphical model of STT. Locations **i** and timestamps **t** are modeled as observed random variables, shown as shaded circles, while the latent random variables of topics **z** and regions **r** and all parameters listed in Table I are shown as unshaded circles.

Similar to LDA [1], the topic of check-ins is considered as a latent random variable. Topic distributions of users θ^{user} model the latent user interests, from which the topics of check-ins are sampled. Topics are associated with location distributions ϕ^{topic}, which model the latent location factors.

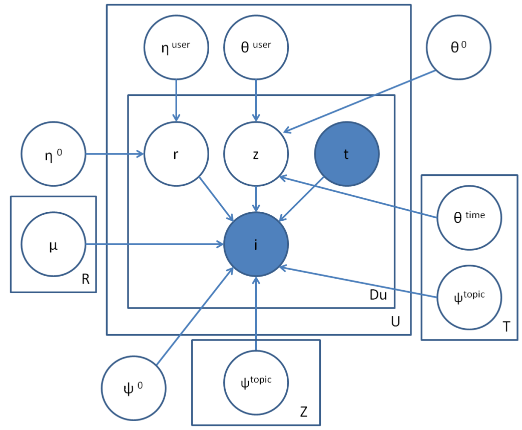


Fig. 2. The graphical model of STT

Given the sampled topic, locations are drawn from the location distribution of that topic. GLDA [9] extends LDA that the generated location index is regularized by the set of all check-ins by the given user.

As discussed in Section I, we observe that users have visited multiple regions, which has not been considered by existing models. Therefore, as shown in Figure 2, an important difference between STT and existing models is that users are associated with region distributions η^{user}, and the regions of check-ins are sampled from the mixture of users' region distributions η^{user} and a background region distribution η^0. Instead of regularizing locations to be close to all the user's check-ins, STT regularizes them to the sampled region center μ_r. This model design is more meaningful when observing that users have visited multiple regions.

Another major difference in STT is that it models the impact of temporal activity patterns on location recommendations (see the discussion in Section I). Given a timestamp t of check-in d, STT assumes that the topic $z_{u,d}$ depends not only on the user u's topic distribution θ_u^{user} but also on the time's topic distribution θ_t^{time}. Moreover, the probability of generating (recommending) the index of location $i_{u,d}$ depends on the given time's location distribution ϕ_t^{time}. STT also models the popularity of locations by using the background distribution of locations ψ^0.

Next, we describe the generative process of the STT model for a single check-in d of a given user u and time $t_{u,d}$.

- Draw a topic index $z_{u,d}$
  - $z_{u,d} \sim p(z_{u,d}|u, t_{u,d}, \theta^0, \theta^{user}, \theta^{time})$
- Draw a region index $r_{u,d}$
  - $r_{u,d} \sim p(r_{u,d}|u, \eta^0, \eta^{user})$
- Draw a location index $i_{u,d}$, given the region index $r_{u,d}$ and topic index $z_{u,d}$
  - $i_{u,d} \sim p(i_{u,d}|r_{u,d}, z_{u,d}, t_{u,d}, \psi^0, \psi^{topic}, \psi^{time}, \mu)$

For each check-in, the STT model first samples a topic from the set of topics. To generate a topic z, the model uses multinomial distributions of the background, user, and time's topic together as follows:

$$p(z|u,t,\theta^0,\theta^{user},\theta^{time}) = \frac{\exp(\theta^0_z + \theta^{user}_{u,z} + \theta^{time}_{t,z})}{\sum_{zz=1}^{|Z|}\exp(\theta^0_{zz} + \theta^{user}_{u,zz} + \theta^{time}_{t,zz})} \quad (1)$$

where θ^0 is the background topic distribution, and θ^{user} and θ^{time} are the topic distributions of user u and time t, respectively. To simplify the notations, we use $p(z|u,t,\theta^0,\theta^{user},\theta^{time}) = \alpha_{u,t,z}$. This approach employs the sparse coding technique introduced in the SAGE (Sparse Additive Generative) model [5]. The major advantage of SAGE is that it does not require additional latent "switching" variables when the model needs to take multiple factors into account.

Similarly, for generating the region index, we use a multinomial distribution as follows:

$$p(r|u,\eta^0,\eta^{user}) = \frac{\exp(\eta^0_r + \eta^{user}_{u,r})}{\sum_{rr=1}^{|R|}\exp(\eta^0_{rr} + \eta^{user}_{u,rr})} \quad (2)$$

where η^0 is the global distribution of regions and η^{user}_u is the region distribution of user u. To simplify the notations, we use $p(r|u,\eta^0,\eta^{user}) = \beta_{u,r}$.

Each location index i is drawn depending on the sampled topic z and sampled region r as follows:

$$\begin{aligned} p(i|r,z,t,\psi^0,\psi^{topic},\psi^{time},\mu) \\ = p(i|\psi^0,\psi^{topic}_z,\psi^{time}_t) \times p(l_i|\mu_r) \\ = \frac{\exp(\psi^0_i+\psi^{topic}_{z,i}+\psi^{time}_{t,i})\exp(-\frac{\rho}{2}||\mu_r-l_i||)}{\sum_{ii=1}^{|I|}\exp(\psi^0_{ii}+\psi^{topic}_{z,ii}+\psi^{time}_{t,ii})\exp(-\frac{\rho}{2}||\mu_r-l_{ii}||)} \end{aligned} \quad (3)$$

where $p(i|r,z,t,\psi^0,\psi^{topic},\psi^{time},\mu) = \delta_{z,r,t,i}$. The probability of a location index is the product of the probability of drawing the index of the location from the mixture of location distributions $\psi^0 + \psi^{topic}_z + \psi^{time}_t$, and the probability of drawing the coordinates l_i of the location i, which is inversely proportional to the distance between μ_r and $l_{i_{u,d}}$, i.e., the L^2-norm $||\mu_r - l_i||$. ρ controls the trade-off between the geographical factor and the topic and time factors. By increasing the value of ρ, the model gradually puts more weights on the geographical influence and recommends more locations nearby. When the value of ρ decreases, the model recommends more locations based on the topic and time factors.

C. Parameter Learning

Our goal is to learn parameters that maximize the marginal log-likelihood of the observed random variables $\mathbf{i},\mathbf{t}$, which is hard to be maximized directly. Therefore, we apply the MCEM (Monte Carlo Expectation Maximization) algorithm to maximize the complete data likelihood $p(\mathbf{z},\mathbf{r},\mathbf{i}|\mathbf{t},\mathbf{u},\Theta)$, where $\Theta = \{\theta^0, \theta^{user}, \theta^{time}, \eta^0, \eta^{user}, \psi^0, \psi^{topic}, \psi^{time}, \mu\}$.

According to the MCEM method, we sample the latent variables $\mathbf{r},\mathbf{z}$ in the E step and maximize the parameters Θ in the M step. To update the parameters, we use the gradient descent learning algorithm PSSG (Projected Scaled Sub-Gradient) [12], which is designed to solve optimization problems with L1 regularization on the parameters. More importantly, PSSG is scalable because it uses the quasi-Newton strategy with line search that is robust to common functions. The derivative functions for the parameters are omitted because of the page limit.

IV. Experiments

A. Data Sets

We used three publicly available data sets: a Twitter data set from [3][1], and Gowalla and Brightkite data sets from [4][2]. We generate subsets from a representative city NYC (New York City) in the US for Twitter, Gowalla, and Brightkite, where all check-ins contain a location label and geographical coordinates. Some statistics about the data sets are presented in Table II.

TABLE II
Statistics of data sets from Twitter, Gowalla, and Brightkite.

#	Twitter	Gowalla	Brightkite
Unique users	9,508	5,588	1,820
Locations	3,518	4,358	348
Check-ins	607,885	89,294	34,710
Avg. check-ins/user	64.93	15.97	19.07
Avg. check-ins/location	172.79	20.48	99.74

B. Experimental Setup

In our data sets, we randomly select 70% of observed data for each user as the training data, and the remaining 30% as the test data. We focus on the tasks of location recommendation for users. Location recommendation is by far the most commonly used performance measure for spatial models in the literature [14], [2], [9]. Given a check-in with a user, our task is to recommend top-k locations, that user will visit in the future. More precisely, given the user u and time t of a check-in d, the probability that user u visits location i at time t is computed by $p(i|t,u,\Theta) \propto \sum_r^R \sum_z^Z p(z|\theta^0,\theta^{user},\theta^{time}) \times p(r|\eta^0,\eta^{user}) \times p(i|z,r,t,\psi^0,\psi^{topic},\psi^{time},\mu)$. We rank the locations in descending order of $p(i|t,u,\Theta)$.

Evaluation Metrics. Perplexity is the standard for measuring how well a probabilistic model fits the data, and is monotonically decreasing in the likelihood of the test data set, so that a lower perplexity indicates better performance of the model. The top-k accuracy for a test check-in is one when the ground truth location is in the top-k recommendations, and zero otherwise. The accuracy@k is the average top-k accuracy over all test check-ins.

Comparison Partners. In our experiments, we evaluate the following comparison partners, which all model (and can predict) either the coordinates or the index of locations:

- MR (Multi-Region). This is a simplified version of the STT model. It assumes that users are associated with region distributions, and the coordinates of locations are drawn from 2D Gaussian distributions. As a result, it

[1] http://infolab.tamu.edu/data/
[2] http://snap.stanford.edu/data/

generates the coordinates of locations **l**, and the only latent variable is the region **r**. Intuitively, the MR model recommends closest locations to users based on their regions.

- PMF (Probabilistic Matrix Factorization). This is a well-known matrix factorization model proposed in [11].
- LDA (Latent Dirichlet Allocation). This is a modified LDA model, where the only observed variable is the index of locations **i** and the only latent variable is the topic **z**. Note that this model is equivalent to the original LDA model that generates index of locations instead of words.
- GT (Geographical Topic). This is one of the state-of-the-art geographical topic models proposed in [6].
- TOT (Topics Over Time). The TOT model is a temporal topic models proposed in [13], and it assumes that the continuous timestamp of a document is drawn from a topic-specific Beta distribution. Since we consider only relative timestamps in hours or days, the modified model assumes the timestamp is drawn from a Multinomial distribution.
- $GLDA$ (Geo LDA). This is a spatial extension of the LDA model, which is one of the state-of-the-art methods for location recommendation proposed in [9].
- STT (Spatio-Temporal Topic). This is the full spatio-temporal topic model proposed in this paper. The default number of relative timestamps is 24 (hours). Optionally, we use STT_{week} to denote the model considering 7 timestamps (7 days).

C. Experimental Results

1) Perplexity: Figure 3 shows the perplexity of the comparison partners for different numbers of topics. For all models, the number of regions is set to 50 for the Twitter and Gowalla data sets and to 30 for the Brightkite data set. We have similar results for different numbers of regions, which are omitted because of the page limit. To establish a fair comparison, we only compare the LDA, TOT, GLDA, and STT models, as they have the same observed random variables (the index of locations and timestamps).

We observe that the STT model consistently achieves the smaller (better) perplexity than the LDA, TOT, and GLDA models in all the three data sets for different numbers of topics and regions, which means that the STT model fits the data better than the other models. As expected, we observe that the perplexity of all models decreases as the number of topics and regions increases.

2) Location Recommendation: Figure 4(a), 4(b), and 4(c) show the accuracy@1,5,10 results for location recommendation in the Twitter, Gowalla and Brightkite data sets, respectively.

We observe that our STT and STT_{week} models consistently outperform all other models on all the three data sets. PMF and LDA yield similar results in all the data sets since they are conceptually analogous (as discussed in Section II). Both MR and GT perform worse than PMF and LDA, because they do not consider the "semantic" meaning of the locations and do not model the correlation between user movements. TOT is sometimes better or worse than PMF and LDA, because only the temporal aspect is considered, and it is dependent on only topics (not on both topics and locations as in STT). GLDA outperform PMF and LDA, because it considers the spatial aspect of user check-ins. Compared to the most competitive and state-of-the-art method in the area of recommender systems, GLDA, STT improves the accuracy@1 by 19% (Twitter), 30% (Gowalla), and 12% (Brightkite). This indicates that modeling the multiple regions of users and the impact of temporal activity patterns can help improve the accuracy of location recommendation. The results of STT are slightly better than those of STT_{week}, indicating that the timestamps in hours produce more distinctive temporal activity patterns than the ones in days.

We also observe that the accuracy difference between STT and the other models on Twitter and Gowalla is larger than on Brightkite. We argue that as the number of locations in the Brightkite data set is much smaller than in the other two data sets, the performance of baseline methods, such as MR, PMF, and LDA, is sufficiently good, so that the room for improvement is limited.

To analyze the impact of the input parameters, Figure 5 shows the accuracy@5 of the comparison partners for different numbers of topics. Again, the similar results for different numbers of regions are omitted because of the page limit. The results for accuracy@1,10 are similar to the results for accuracy@5.

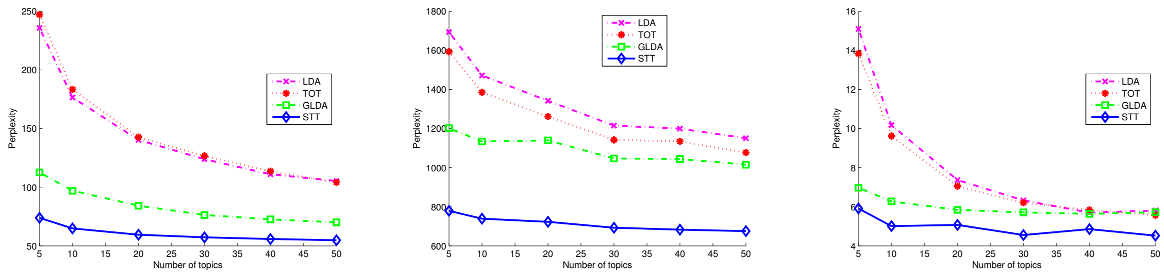
We observe that STT consistently outperforms the other comparison partners for all numbers of topics and regions. Furthermore, as the number of topics increases, the accuracy@5 of all the models increases and then plateaus when the number of topics reaches 20 or 30. Similarly, as the number of regions increases, the accuracy of STT increases. Overall, the results of STT are fairly robust to the choice of the input parameters.

V. Conclusion

In this paper, we propose the first spatio-temporal topic model to capture the spatial and temporal aspects of check-ins, as well as user profiles (topic and region distributions) in a single probabilistic model, called Spatio-Temporal Topic (STT) model. STT exploits the interdependencies between users' regions and their locations, and between temporal activity patterns and locations. We perform an experimental evaluation on Twitter, Gowalla, and Brightkite data sets from New York City. We compare STT against state-of-the-art methods in the areas of recommender systems, and geographical and temporal topic modeling. Our experiments demonstrate substantially improved performance in location recommendation.

References

[1] D. M. Blei, A. Ng, and M. Jordan. Latent dirichlet allocation. *JMLR*, 3:993–1022, 2003.
[2] C. Cheng, H. Yang, I. King, and M. R. Lyu. Fused matrix factorization with geographical and social influence in location-based social networks. AAAI, 2012.
[3] Z. Cheng, J. Caverlee, K. Lee, and D. Z. Sui. Exploring millions of footprints in location sharing services. ICWSM, 2011.

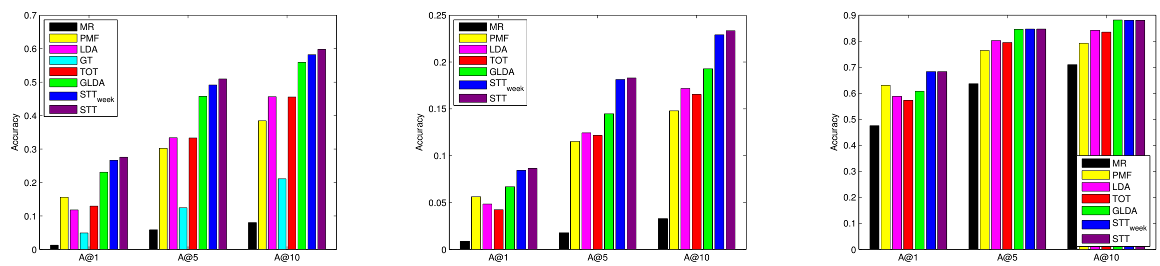


(a) Twitter data set.

(b) Gowalla data set.

(c) Brightkite data set.

Fig. 3. Perplexity of the comparison partners for 5,10,20,30,40,50 topics.

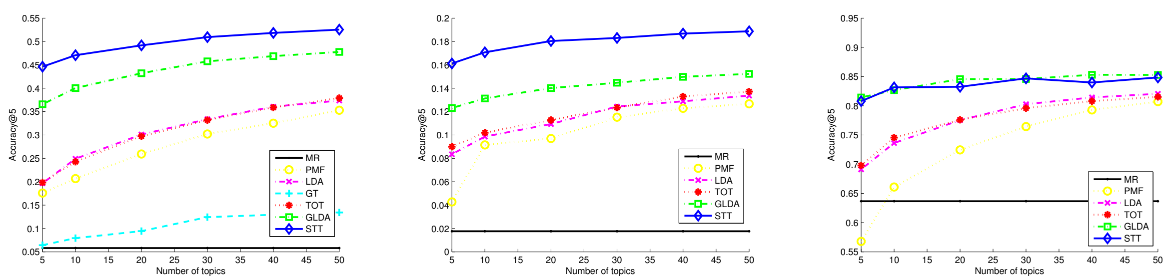


(a) Twitter data set.

(b) Gowalla data set.

(c) Brightkite data set.

Fig. 4. Accuracy@1,5,10 of location recommendation.



(a) Twitter data set.

(b) Gowalla data set.

(c) Brightkite data set.

Fig. 5. Accuracy@5 of location recommendation for different number of topics.

[4] E. Cho, S. A. Myers, and J. Leskovec. Friendship and mobility: user movement in location-based social networks. KDD, pages 1082–1090, 2011.

[5] J. Eisenstein, A. Ahmed, and E. P. Xing. Sparse additive generative models of text. ICML, pages 1041–1048, 2011.

[6] L. Hong, A. Ahmed, S. Gurumurthy, A. J. Smola, and K. Tsioutsiouliklis. Discovering geographical topics in the twitter stream. WWW, pages 769–778, 2012.

[7] N. Kawamae. Trend analysis model: trend consists of temporal words, topics, and timestamps. WSDM, pages 317–326, 2011.

[8] Y. Koren, R. Bell, and C. Volinsky. Matrix factorization techniques for recommender systems. *Computer*, 42:30–37, 2009.

[9] T. Kurashima, T. Iwata, T. Hoshide, N. Takaya, and K. Fujimura. Geo topic model: joint modeling of user's activity area and interests for location recommendation. WSDM, pages 375–384, 2013.

[10] A. Noulas, S. Scellato, N. Lathia, and C. Mascolo. Mining user mobility features for next place prediction in location-based services. In *ICDM*, pages 1038–1043, 2012.

[11] R. Salakhutdinov and M. Andriy. Probabilistic matrix factorization. NIPS, pages 1257–1264, 2008.

[12] M. Schmidt, A. Niculescu-Mizil, and K. Murphy. Learning graphical model structure using l1-regularization paths. AAAI, pages 1278–1283, 2007.

[13] X. Wang and A. McCallum. Topics over time: a non-markov continuous-time model of topical trends. KDD, pages 424–433, 2006.

[14] M. Ye, P. Yin, and W.-C. Lee. Location recommendation for location-based social networks. GIS, pages 458–461, 2010.

[15] M. Ye, P. Yin, W.-C. Lee, and D.-L. Lee. Exploiting geographical influence for collaborative point-of-interest recommendation. SIGIR, pages 325–334, 2011.

[16] Z. Yin, L. Cao, J. Han, C. Zhai, and T. Huang. Geographical topic discovery and comparison. WWW, pages 247–256, 2011.

[17] V. W. Zheng, Y. Zheng, X. Xie, and Q. Yang. Collaborative location and activity recommendations with gps history data. WWW, pages 1029–1038, 2010.