

Aggregated Temporal Tensor Factorization Model for Point-of-interest Recommendation

Shenglin Zhao^{1,2(✉)}, Michael R. Lyu^{1,2}, and Irwin King^{1,2}

¹ Shenzhen Key Laboratory of Rich Media Big Data Analytics and Application, Shenzhen Research Institute, The Chinese University of Hong Kong, Shenzhen, China

² Department of Computer Science and Engineering, The Chinese University of Hong Kong, Hong Kong, China
{slzhao,lyu,king}@cse.cuhk.edu.hk

Abstract. Temporal influence plays an important role in a point-of-interest (POI) recommendation system that suggests POIs for users in location-based social networks (LBSNs). Previous studies observe that the user mobility in LBSNs exhibits distinct temporal features, summarized as periodicity, consecutiveness, and non-uniformness. By capturing the observed temporal features, a variety of systems are proposed to enhance POI recommendation. However, previous work does not model the three features together. More importantly, we observe that the temporal influence exists at different time scales, yet this observation cannot be modeled in prior work. In this paper, we propose an Aggregated Temporal Tensor Factorization (ATTF) model for POI recommendation to capture the three temporal features together, as well as at different time scales. Specifically, we employ temporal tensor factorization to model the check-in activity, subsuming the three temporal features together. Furthermore, we exploit a linear combination operator to aggregate temporal latent features' contributions at different time scales. Experiments on two real life datasets show that the ATTF model achieves better performance than models capturing temporal influence at single scale. In addition, our proposed ATTF model outperforms the state-of-the-art methods.

Keywords: Point-of-interest recommendation · Tensor factorization · Temporal influence · Location-based social network

1 Introduction

Temporal influence plays an important role in a point-of-interest (POI) recommendation system that recommends POIs for users in location-based social networks (LBSNs). Previous studies show that the user mobility in LBSNs exhibits distinct temporal features [3, 4, 17]. For example, users always stay in the office in the Monday afternoon, and enjoy entertainments in bars at night. The temporal features in users' check-in data could be summarized as follows.

- **Periodicity.** Users share the same periodic pattern, visiting the same or similar POIs at the same time slot [3, 17]. For instance, a user always visits restaurants at noon, so do other users. Hence, the periodicity inspires the time-aware collaborative filtering method to recommend POIs.
- **Consecutiveness.** A user’s current check-in is largely correlated with the recent check-in [2, 4]. Geo et al. [4] model this property by assuming that user preferences are similar in two consecutive hours. Cheng et al. [2] assume that two checked-in POIs in a short term are highly correlated.
- **Non-uniformness.** A user’s check-in preference changes at different hours of a day, or at different months of a year, or different days of a week [4]. For example, at noon a user may visit restaurants while at night he/she may have fun in bars.

By capturing the observed temporal features, a variety of systems are proposed to enhance POI recommendation performance [2, 4, 17]. They do gain better performance than general collaborative filtering (CF) methods [15]. Nevertheless, previous work [2, 4, 17] cannot model the three features together. Moreover, an important fact is ignored in prior work that the temporal influence exists at different time scales. For example, in day level, you may check-in POIs around your home in the evening morning, visit places around your office in the day time, and have fun at nightclubs in the evening. In week level, you may stay in the city for work in weekdays and go out for vacation in weekends. Hence, to better model the temporal influence, capturing the temporal features at different time scales is necessary.

In this paper, we propose an Aggregated Temporal Tensor Factorization (ATTF) model for POI recommendation to capture the three temporal features together, as well as at different time scales. We construct a user-time-POI tensor to represent the check-ins, and then employ the interaction tensor factorization [14] to model the temporal effect. Different from prior work that represents the temporal influence at single scale, we index the temporal information for latent representation at different scales, i.e., hour, week day, and month. Furthermore, we employ a linear combination operator to aggregate different temporal latent features’ contributions, which capture the temporal influence at different scales. Specifically, our ATTF model learns the three temporal properties as follows: (1) periodicity is learned from the temporal CF mechanism; (2) consecutiveness is manifested in two aspects—time in a slot brings the same effect through sharing the same time factor, and relation between two consecutive time slots could be learned from the tensor model; (3) non-uniformness is depicted by different time factors for different time slots at different time scales. Moreover, an aggregate operator is introduced to combine the temporal influence at different scales, i.e., hour, week day, and month, and represent the temporal effect in a whole.

The contributions of this paper are summarized as follows:

- To the best of our knowledge, this is the first temporal tensor factorization method for POI recommendation, subsuming all the three temporal properties: periodicity, consecutiveness, and non-uniformness.

- We propose a novel model to capture temporal effect in POI recommendation at different time scales. Experimental results show that our model outperforms prior temporal model more than 20 %.
- The ATTF model is a general framework to capture the temporal features at different scales, which outperforms single temporal factor model and gains 10 % improvement in the top-5 POI recommendation task on Gowalla data.

2 Related Work

In this section, we first review the literature of POI recommendation. Then we summarize the progress of modeling temporal effect for POI recommendation.

POI Recommendation. Most of POI recommendation systems base on the collaborative filtering (CF) techniques, which can be reported in two aspects, memory-based and model-based. On the one hand, Ye et al. [15] propose the POI recommendation problem in LBSNs solved by user-based CF method, and further improve the system by linearly combining the geographical influence, social influence, and preference similarity. In order to enhance the performance, more advanced techniques are then applied, e.g., incorporating temporal influence [17], and utilizing a personalized geographical model via kernel density estimator [18, 19]. On the other hand, model-based CF is proposed to tackle the POI recommendation problem that benefits from its scalability. Cheng et al. [1] propose a multi-center Gaussian model to capture user geographical influence and combine it with social matrix factorization (MF) model [12] to recommend POIs. Gao et al. [4] propose an MF-based model, Location Recommendation framework with Temporal effects (LRT), utilizing similarity between time-adjacent check-ins to improve performance. Lian et al. [9] and Liu et al. [11] enhance the POI recommendation by incorporating geographical information in a weighted regularized matrix factorization model [7]. In addition, some researchers subsume users' comments to improve the recommendation performance [5, 8, 16]. Other researchers model the consecutive check-ins' correlations to enhance the system [2, 10, 21].

Temporal Effect Modeling. In 2011, Cho et al. [3] propose the periodicity of check-in data in LBSNs. For instance, people always visit restaurants at noon. Hence, it is reasonable to recommend users restaurants he/she did not visit at noon. In 2013, Yuan et al. [17] improve the similarity metric in CF model by combining the temporal similarity and non-temporal similarity, which enhances the recommendation performance. At the same year, Gao et al. [4] observe the non-uniformness property (a user's check-in preference changes at different hours of a day), and consecutiveness (a user's preference at time t is similar with time $t - 1$). Further, Gao et al. propose LRT model to capture the non-uniformness and consecutiveness. Meantime, Cheng et al. [2] propose FPMC-LR model to capture the consecutiveness, supposing strong correlation between two consecutive checked-in POIs.

3 Aggregated Temporal Tensor Factorization (ATTF) Model

3.1 Formulation

The Aggregated Temporal Tensor Factorization (ATTF) model estimates the preference of a user u at a POI l given a specific time label t through a score function $f(u, t, l)$. In our case, the time label t is represented in three scales, month (from Jan. to Dec.), week day (from Monday to Sunday), and hour (from 0 to 23). Denote that \mathcal{U} is the set of users and \mathcal{L} is the set of POIs. In addition, \mathcal{T}_1 , \mathcal{T}_2 , and \mathcal{T}_3 are the set of months, days of week, and hours respectively. Further we define \mathcal{T} as the set of time label tuples, where a tuple $t := (t_1, t_2, t_3)$. For instance, a time stamp, “2011-04-05 10:12:35”, could be represented as $(3, 2, 10)$ since we index the temporal id from zero. Therefore, the goal of ATTF model is to learn the preference score function $f(u, t, l)$, where u denotes a user id, t denotes a time label tuple, and l denotes a POI id. Under this circumstance, the score function subjects to that a high score corresponds to high preference of a user u for the POI l at time t .

Furthermore, we formulate the Aggregated Temporal Tensor Factorization (ATTF) model through interaction tensor factorization. Thus the score of a POI l given user u and time t is factorized into three interactions: user-time, user-POI, and time-POI, where each interaction is modeled through the latent vector inner product. As different candidate POIs share the same user-time interaction, the score function for comparing preference over POIs could be simplified as the combination of interactions of user-POI and time-POI. Formally, the score function for a given time label t , user u , and a target POI l could be formulated as,

$$f(u, t, l) = \langle U_u^{(L)}, L_l^{(U)} \rangle + \langle A(T_{1,t_1}^{(L)}, T_{2,t_2}^{(L)}, T_{3,t_3}^{(L)}), L_l^{(T)} \rangle, \quad (1)$$

where $\langle \cdot \rangle$ denotes the vector inner product, $A(\cdot)$ is the aggregate operator. Suppose that d is the latent vector dimension, $U_u^{(L)} \in R^d$ is user u 's latent vector for POI interaction, $L_l^{(U)}, L_l^{(T)} \in R^d$ are POI l 's latent vectors for user interaction and time interaction, $T_{1,t_1}^{(L)}, T_{2,t_2}^{(L)}, T_{3,t_3}^{(L)} \in R^d$ are time t 's latent vector representations in three aspects: month, week day, and hour.

Aggregate operator combines several temporal characteristics together. In this paper, we propose a linear convex combination operator. It is formulated as follows,

$$A(\cdot) = \alpha_1 \cdot T_{1,t_1}^{(L)} + \alpha_2 \cdot T_{2,t_2}^{(L)} + \alpha_3 \cdot T_{3,t_3}^{(L)}, \quad (2)$$

where α_1 , α_2 , and α_3 denote the weights of each temporal factor, satisfying $\alpha_1 + \alpha_2 + \alpha_3 = 1$, and $\alpha_1, \alpha_2, \alpha_3 \geq 0$.

3.2 Inference

We infer the model via BPR criteria [13] that is a general framework to train a recommendation system from implicit feedback. We treat the checked-in POIs

as positive and the unchecked as negative. Suppose that the score of $f(u, t, l)$ at positive observations is higher than the negative given u and t , we formulate the relation that user u prefers a positive POI l_i than a negative one l_j at time t as $l_i >_{u,t} l_j$. Further, we extract the set of pairwise preference relations from the training examples,

$$D_S := \{(u, t, l_i, l_j) | l_i >_{u,t} l_j, u \in \mathcal{U}, t \in \mathcal{T}, l_i, l_j \in \mathcal{L}\}. \quad (3)$$

Suppose the quadruples in D_S are independent of each other, then to learn the ATTF model is to maximize the likelihood of all the pair orders,

$$\arg \max_{\Theta} \prod_{(u,t,l_i,l_j) \in D_S} p(l_i >_{u,t} l_j), \quad (4)$$

where $p(l_i >_{u,t} l_j) = \sigma(f(u, t, l_i) - f(u, t, l_j))$, σ is the sigmoid function, and Θ is the parameters to learn, namely $U^{(L)}$, $L^{(U)}$, $L^{(T)}$, $T_1^{(L)}$, $T_2^{(L)}$, and $T_3^{(L)}$. The objective function is equivalent to minimizing the negative log likelihood. To avoid the risk of overfitting, we add a Frobenius norm term to regularize the parameters. Then the objective function is

$$\arg \min_{\Theta} \sum_{(u,t,l_i,l_j) \in D_S} -\ln(\sigma(y_{u,t,l_i} - y_{u,t,l_j})) + \lambda_{\Theta} \|\Theta\|_F^2, \quad (5)$$

where λ_{Θ} is the regularization parameter.

Algorithm 1. ATTF model learning algorithm

Input: Training tuples $\{(u_i, t_i, l_i)\}_{i=1,\dots,N}$

Output: $U^{(L)}$, $T_1^{(L)}$, $T_2^{(L)}$, $T_3^{(L)}$, $L^{(U)}$, $L^{(T)}$

1: Initialize $U^{(L)}$, $T_1^{(L)}$, $T_2^{(L)}$, $T_3^{(L)}$, $L^{(U)}$, $L^{(T)}$

2: **repeat**

3: Draw (u, t, l_p) uniformly from training tuples

4: For $s = 1, 2, \dots, k$, where k is the number of sampled negative POIs

5: Draw (u, t, l_p, l_n) uniformly

6: $\delta \leftarrow 1 - \sigma(f(u, t, l_p) - f(u, t, l_n))$

7: Update parameters according to Eq. (6)

8: **until** convergence

9: **return** $U^{(L)}$, $T_1^{(L)}$, $T_2^{(L)}$, $T_3^{(L)}$, $L^{(U)}$, $L^{(T)}$

3.3 Learning

We leverage the Stochastic Gradient Decent (SGD) algorithm to learn the objective function for efficacy. Due to convenience, we define $\delta = 1 - \sigma(f(u, t, l_p) - f(u, t, l_n))$. As $T_{1,t_1}^{(L)}$, $T_{2,t_2}^{(L)}$, and $T_{3,t_3}^{(L)}$ are symmetric, they have the same gradient form. For simplicity, we use $T_t^{(L)} \in \{T_{1,t_1}^{(L)}, T_{2,t_2}^{(L)}, T_{3,t_3}^{(L)}\}$ to represent any

of them, $\alpha \in \{\alpha_1, \alpha_2, \alpha_3\}$ to denote corresponding weight, and $A(\cdot)$ to denote $A(T_{1,t_1}^{(L)}, T_{2,t_2}^{(L)}, T_{3,t_3}^{(L)})$. Then the parameter updating rules are as follows,

$$\begin{aligned}
 U_u^{(L)} &\leftarrow U_u^{(L)} + \gamma \cdot \left(\delta \cdot \left(L_{i_p}^{(U)} - L_{i_n}^{(U)} \right) - \lambda \cdot U_u^{(L)} \right), \\
 L_{i_p}^{(U)} &\leftarrow L_{i_p}^{(U)} + \gamma \cdot \left(\delta \cdot U_u^{(L)} - \lambda \cdot L_{i_p}^{(U)} \right), \\
 L_{i_p}^{(T)} &\leftarrow L_{i_p}^{(T)} + \gamma \cdot \left(\delta \cdot A(\cdot) - \lambda \cdot L_{i_p}^{(T)} \right), \\
 L_{i_n}^{(U)} &\leftarrow L_{i_n}^{(U)} - \gamma \cdot \left(\delta \cdot U_u^{(L)} + \lambda \cdot L_{i_n}^{(U)} \right), \\
 L_{i_n}^{(T)} &\leftarrow L_{i_n}^{(T)} - \gamma \cdot \left(\delta \cdot A(\cdot) + \lambda \cdot L_{i_n}^{(T)} \right), \\
 T_t^{(L)} &\leftarrow T_t^{(L)} + \gamma \cdot \left(\delta \cdot \alpha \cdot \left(L_{i_p}^{(T)} - L_{i_n}^{(T)} \right) - \lambda \cdot T_t^{(L)} \right),
 \end{aligned} \tag{6}$$

where γ is the learning rate, λ is the regularization parameter. To train the model, we use the bootstrap sampling skill to draw the quadruple from D_S , following [13]. Algorithm 1 gives the detailed procedure to learn the ATTF model.

Complexity. The runtime for predicting a triple (u, t, l) is in $O(d)$, where d is the number of latent vector dimension. The updating procedure is also in $O(d)$. Hence training a quadruple is in $O(d)$, then training an example (u, t, l) is in $O(k \cdot d)$, where k is the number of sampled negative POIs. Therefore, to train the model costs $O(N \cdot k \cdot d)$, where N is the number of training examples.

4 Experiments

4.1 Data Description and Experimental Setting

Two real-world datasets are used in the experiment: one is Foursquare data from January 1, 2011 to July 31, 2011 provided in [6] and the other is Gowalla data from January 1, 2011 to September 31, 2011 in [20]. We filter the POIs checked-in by less than 5 users and then choose users who check-in more than 10 times as our samples. Table 1 shows the data statistics. We randomly choose 80 % of each user’s check-ins as training data, and the remaining 20 % for test data. Moreover, we use each check-in (u, t, l) in training data to learn the latent features of user, time, and POI. Then given the (u, t) , we estimate the score value of different candidate POIs, select the top N candidates, and compare them with check-in tuples in test data. Finally, we exploit the **precision** and **recall** to evaluate the model performance, following [8, 21].

Table 1. Data statistics

Source	#users	#POIs	#check-ins	Density
Foursquare	10,180	16,561	867,107	0.0015
Gowalla	3,318	33,665	635,600	0.03

4.2 Model Comparison

To demonstrate the advantage of ATTF in aggregating several temporal latent factors, we propose three single temporal latent factor models as competitors: **TTFM**, **TTFW**, and **TTFH**. They correspondingly model the month, week day, and hour latent factor. We attain their results by setting the corresponding weight as 1, and others as 0 in ATTF.

We compare our ATTF model with state-of-the-art collaborative filtering (CF) methods: **WRMF** [7] and **BPRMF** [13], which are two popular MF models for processing large scale implicit feedback data. In addition, we also compare with state-of-the-art POI recommendation models capturing temporal influence: **LRT** [4] and **FPMC-LR** [2].

4.3 Experimental Results

In the following, we demonstrate the performance comparison on precision and recall for a top- N (N ranges from one to ten) POI recommendation task. We set the latent factor dimension as 60 for all compared models. We leverage grid search method to find the best weights in ATTF model. As a result, ATTF model on Foursquare data gets the best performance when $\alpha_1 = 0.7$, $\alpha_2 = 0.1$, and $\alpha_3 = 0.2$, while ATTF model on Gowalla data gets the best performance when $\alpha_1 = 0.2$, $\alpha_2 = 0.1$, and $\alpha_3 = 0.7$.

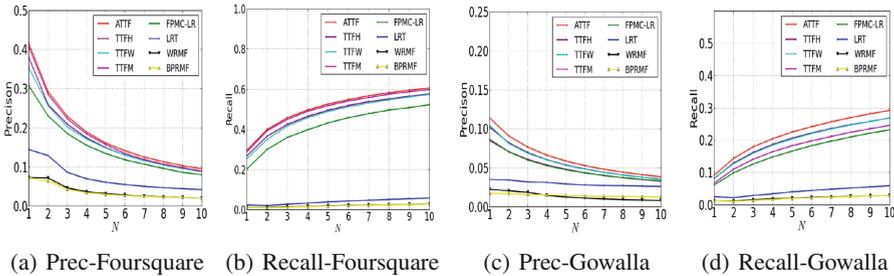


Fig. 1. Model comparison

Figure 1 demonstrates the experimental results on Foursquare and Gowalla data. We sihe following observations. **(1)** Our proposed models, ATTF, TTFH, TTFW, and TTFM, outperform state-of-the-art CF methods and POI recommendation models. Compared with the state-of-the-art competitor, FPMC-LR, ATTF model gains more than 20% enhancement on Foursquare data, and more than 36% enhancement on Gowalla data at Precision@5 and Recall@5. We observe that models perform better on Foursquare dataset than Gowalla dataset, even though it is sparser. The reason lies in Gowalla data contain much more POIs. **(2)** The ATTF model outperforms single temporal factor models. Compared with best single temporal factor model, ATTF gains about 3% enhancement on Foursquare data, and about 10% improvement on Gowalla data in a

top-5 POI recommender system. **(3)** Our proposed models and FPMC-LR perform much better than other competitors, especially at recall measure. They try to recommend POIs at more specific situations, which is the key to improve performance. Our models recommend a user POIs at some specific time, and FPMC-LR recommends POIs given a user’s recent checked-in POIs; while, the other three models give general recommendations.

In addition, different weight assignments on both data give us two interesting insights: (1) When data are sparse, the temporal characteristic on month dominates the POI recommendation performance. Foursquare dataset has high weight on month temporal factor. However, when data are denser, check-ins on hour are not so sparse. Consequently the characteristic on hour of day becomes prominent. (2) We usually pay much attention on characteristics on hour of day and day of week. Our experimental result indicates that the characteristic on month is important, especially for sparse data. The temporal characteristic on day of week is the least significant factor, which always has smallest weight in ATTF model.

5 Conclusion and Future Work

We propose a novel Aggregated Temporal Tensor Factorization (ATTF) model for POI recommendation. The ATTF model introduces time factor to model the temporal effect in POI recommendation, subsuming all the three temporal properties: periodicity, consecutiveness, and non-uniformness. Moreover, the ATTF model captures the temporal influence at different time scales through aggregating several time factor contributions. Experimental results on two real life datasets show that the ATTF model outperforms state-of-the-art models. Our model is a general framework to aggregate several temporal characteristics at different scales. In this work, we only consider three temporal characteristics: hour, week day, and month. In the future, we may add more in this model, e.g., splitting one day into several slots rather than in hour. In addition, we may try nonlinear aggregate operators, e.g., max, in our future work.

Acknowledgments. The work described in this paper was partially supported by the Research Grants Council of the Hong Kong Special Administrative Region, China (No. CUHK 14203314 and No. CUHK 14205214 of the General Research Fund), and 2015 Microsoft Research Asia Collaborative Research Program (Project No. FY16-RES-THEME-005).

References

1. Cheng, C., Yang, H., King, I., Lyu, M.R.: Fused matrix factorization with geographical and social influence in location-based social networks. In: AAAI (2012)
2. Cheng, C., Yang, H., Lyu, M.R., King, I.: Where you like to go next: successive point-of-interest recommendation. In: IJCAI (2013)
3. Cho, E., Myers, S.A., Leskovec, J.: Friendship and mobility: user movement in location-based social networks. In: SIGKDD (2011)

4. Gao, H., Tang, J., Hu, X., Liu, H.: Exploring temporal effects for location recommendation on location-based social networks. In: *RecSys* (2013)
5. Gao, H., Tang, J., Hu, X., Liu, H.: Content-aware point of interest recommendation on location-based social networks. In: *AAAI* (2015)
6. Gao, H., Tang, J., Liu, H.: gSCorr: modeling geo-social correlations for new check-ins on location-based social networks. In: *CIKM* (2012)
7. Hu, Y., Koren, Y., Volinsky, C.: Collaborative filtering for implicit feedback datasets. In: *ICDM* (2008)
8. Lian, D., Ge, Y., Zhang, F., Yuan, N.J., Xie, X., Zhou, T., Rui, Y.: Content-aware collaborative filtering for location recommendation based on human mobility data. In: *ICDM* (2015)
9. Lian, D., Zhao, C., Xie, X., Sun, G., Chen, E., Rui, Y.: GeoMF: joint geographical modeling and matrix factorization for point-of-interest recommendation. In: *SIGKDD* (2014)
10. Liu, X., Liu, Y., Aberer, K., Miao, C.: Personalized point-of-interest recommendation by mining users' preference transition. In: *CIKM* (2013)
11. Liu, Y., Wei, W., Sun, A., Miao, C.: Exploiting geographical neighborhood characteristics for location recommendation. In: *CIKM* (2014)
12. Ma, H., Zhou, D., Liu, C., Lyu, M.R., King, I.: Recommender systems with social regularization. In: *WSDM* (2011)
13. Rendle, S., Freudenthaler, C., Gantner, Z., Schmidt-Thieme, L.: BPR: bayesian personalized ranking from implicit feedback. In: *UAI* (2009)
14. Rendle, S., Schmidt-Thieme, L.: Pairwise interaction tensor factorization for personalized tag recommendation. In: *WSDM* (2010)
15. Ye, M., Yin, P., Lee, W.C., Lee, D.L.: Exploiting geographical influence for collaborative point-of-interest recommendation. In: *SIGIR* (2011)
16. Yin, H., Sun, Y., Cui, B., Hu, Z., Chen, L.: Lcars: a location-content-aware recommender system. In: *SIGKDD* (2013)
17. Yuan, Q., Cong, G., Ma, Z., Sun, A., Thalmann, N.M.: Time-aware point-of-interest recommendation. In: *SIGIR* (2013)
18. Zhang, J.D., Chow, C.Y.: iGSLR: personalized geo-social location recommendation: a kernel density estimation approach. In: *SIGSPATIAL* (2013)
19. Zhang, J.D., Chow, C.Y.: GeoSoCa: exploiting geographical, social and categorical correlations for point-of-interest recommendations. In: *SIGIR* (2015)
20. Zhao, S., King, I., Lyu, M.R.: Capturing geographical influence in POI recommendations. In: Lee, M., Hirose, A., Hou, Z.-G., Kil, R.M. (eds.) *ICONIP 2013, Part II*. LNCS, vol. 8227, pp. 530–537. Springer, Heidelberg (2013)
21. Zhao, S., Zhao, T., Yang, H., Lyu, M.R., King, I.: Stellar: spatial-temporal latent ranking for successive point-of-interest recommendation. In: *AAAI* (2016)