ICIC Express Letters
Volume **10**, Number **7**, July **2016**

ICIC International ©2016 ISSN 1881-803X
pp. 1747–1754

TIME-AWARE COLLABORATIVE LOCATION RECOMMENDATION IN LOCATION-BASED SOCIAL NETWORKS

Miao Zhang, Shiqun Yin*, Sheng Gao and Zhanhao Han

Faculty of Computer and Information Science
Southwest University
No. 2, Tiansheng Road, Beibei District, Chongqing 400715, P. R. China
miaozhangswu@163.com; *Corresponding author: qqqq-qiong@163.com

Received December 2015; accepted March 2016

Abstract. *The popularity of smart phones and maturity of positioning technology spawn location-based social network (LBSN). LBSNs provide a platform for users to share their location information with each other. Location recommendation aims to help users find interesting locations by utilizing users' visiting histories, users' profiles and so on. Several techniques have been recently proposed for location recommendation. However, few works have considered the new city problem, thus how to recommend locations in a new city where people have never visited before. In this paper, a method to solve the new city problem is given on the basis of user-based collaborative filtering recommendation. It uses the guide mechanism, selecting a series of guides for each category in each city, then calculates the similarity between users and candidate guides and infers the score of candidate locations based on the opinions of these local guides. Besides, we believe that time plays an important role in location recommendation, so temporal information is fully considered during the similarity calculation. Finally, we conduct a performance evaluation over a large dataset from Gowalla. Experimental results show that our method can offer more accurate recommendation in the new city situation than other traditional location recommendation methods.*

Keywords: Location-based social network, Location recommendation, New city problem, Similarity calculation

1. **Introduction.** The popularity of smart phones and the development of positioning technology spawn the location-based social network (LBSN), for example, Foursquare, Gowalla, and Brightkit [1]. Users can share their locations or point of interests (POI) such as restaurants, attractions, cinemas in LBSNs. In February 2013, Foursquare already has 300 million users, and the attendance is up to three billion [2].

Collaborative filtering has been widely used in point of interest recommendation [3]. B. Liu et al. [4] proposed a new framework based on collaborative filtering recommendation with a full integration of user preferences, social impact and geographical factors. Q. Yuan et al. [5] further pushed collaborative recommendation by considering the time factor. H. Gao et al. [6] made full use of various content information in LBSN, such as user labels and user comments using semantic analysis to further capture user preferences.

While with the improvement of people's living standard and transportation, more and more people tend to travel in the holidays. How to recommend suitable locations for a user in a city where he has never visited before is becoming an urgent demand. However, most previous studies were based on the premise that people just visit local locations, and they merely consider how to solve the local recommendation problem. In this new city case, methods in [4-6] are not feasible.

This paper proposes a location recommendation method to solve the new city problem by adopting a guide mechanism. Guides always have the best understanding of a certain type of location in each city, so users can have access to the city's most valuable locations by referring to the advice of these guides. Besides, it models user's behavior and mines his temporal patterns across different location categories. Thus the similarity computation incorporates temporal factors.

The remainder of this paper is organized as follows. In Section 2, we introduce notations and present background information of location recommendation in LBSN. Section 3 details our time-aware collaborative location recommendation method. The results of an empirical experiment are presented in Section 4, followed by conclusions in Section 5.

2. Basic Collaborative Filtering for Location Recommendation.

As one of the most successful recommendation technologies, collaborative filtering (CF) is also widely used in location recommendation. It comprises user-based CF and item-based CF. As user-based CF performs better in location recommendation than item-based CF [7], the following will explicitly describe how to apply it in location recommendation.

The principle of user-based collaborative filtering is to find similar users to the target user and then according to these similar users' check-in history to recommend. This method consists of three steps: similar users discovery, candidate locations rating and Top-N result acquisition.

Given a recommendation system consists of M users and N locations, it can be denoted by an M N matrix, which is called the user-location matrix as Table 1 shows. $C_{i,j}$ represents whether user i has visited location j or not. If user i has visited location j, then $C_{i,j} = 1$. If he has not, $C_{i,j} = 0$. Each user is represented by $U_i = [C_{i,1}, C_{i,2}, \ldots, C_{i,n}]$.

Table 1. User-location matrix

	L_1	...	L_j	...	L_n
U_1	1	1	0	1	$C_{1,n}$
...	...	...	...	...	...
U_i	1	0	$C_{i,j}$	1	0
...	...	...	...	...	...
U_m	$C_{m,1}$	1	1	0	1

Step1 Similarity users discovery. We first should compute the similarity between target user and other users, and then choose the most similar users. In this paper, Cosine Similarity is used like Equation (1).

$$w_{i,k} = \frac{\sum_{l_j \in L} c_{i,j} c_{k,j}}{\sqrt{\sum_{j \in L} c_{i,j}^2} \sqrt{\sum_{j \in L} c_{k,j}^2}} \tag{1}$$

$w_{i,k}$ denotes the similarity between users i and k.

Step2 Candidate locations rating. Suppose choosing P similar users and these users have visited Q locations in total. Candidate locations are those which were among the Q meanwhile not visited by target user. Equation (2) is used to infer the probability for target user i to visit candidate location j.

$$p_{i,j} = \sum_{k \in Candidate} w_{i,k} \times r_{k,j} \tag{2}$$

$p_{i,j}$ means the probability for user i to visit location j, $w_{i,k}$ denotes the similarity between users i and k, and $r_{k,j}$ means the interest of user k to location j.

Step3 Top-N result acquisition. After rating all the candidate locations, select the top-k locations as recommendation results and return them to target user.

3. **Time-Aware Collaborative Location Recommendation.** Through above illustration, we can infer that in the new city situation, a user-based CF method is not feasible any more. In this case, it is likely that all the similar users live in the same city with target user, most of them almost only visit the local locations so few of the candidate locations are in the new city.

We define the new city problem as follows: given a querying user u with a querying new city l, find k interesting locations within city l which perfectly match the preference of u. In order to solve this problem, we proposed a time-aware collaborative location recommendation shown as Figure 1. It comprises 4 parts, i.e., guide expertise discovery, candidate guide selection, similarity computation and candidate location rating, which will be given explict illustrations in this section.

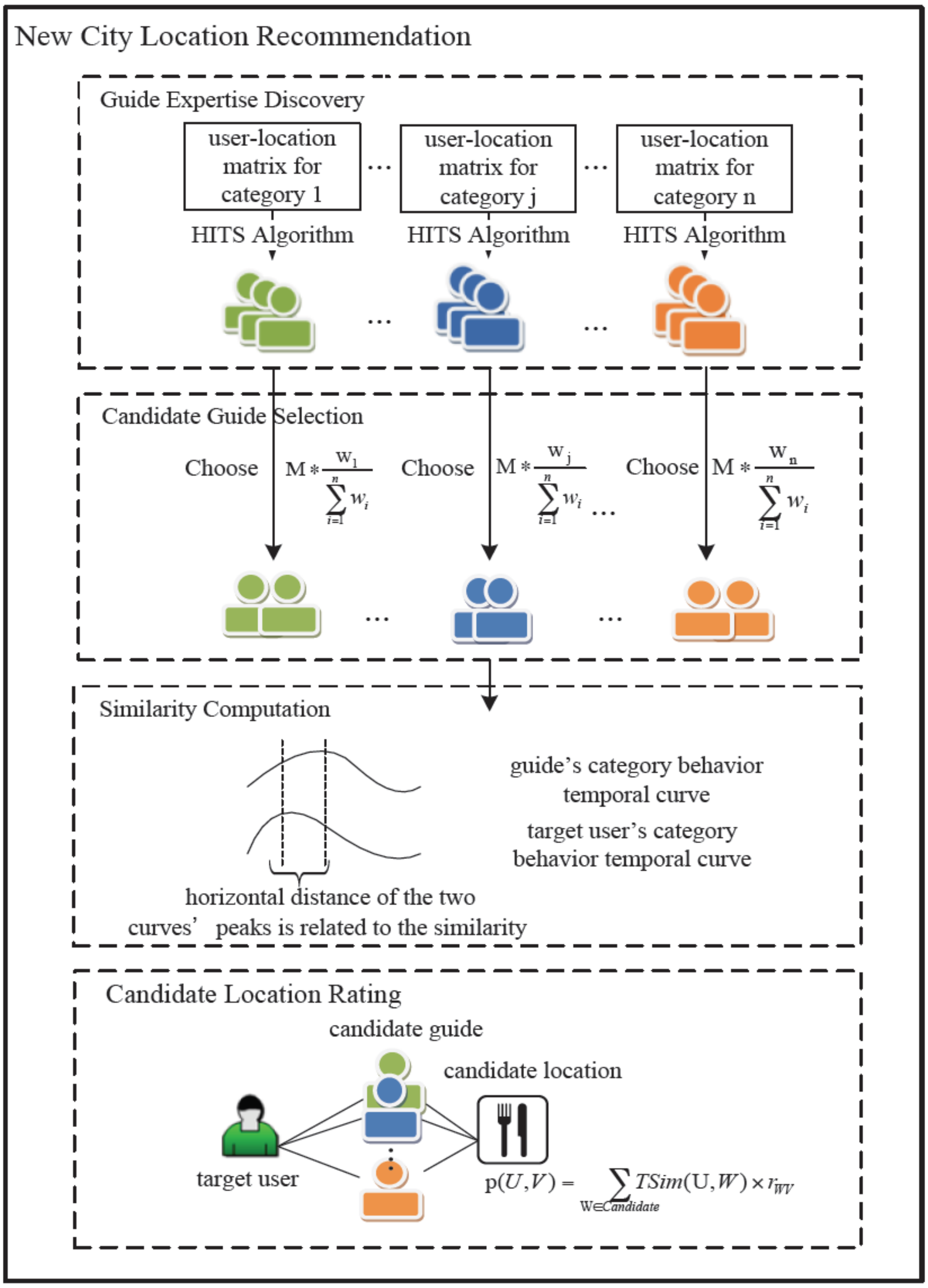

FIGURE 1. Time-aware collaborative location recommendation model

3.1. **Guide expertise discovery.** When we travel to a new city to play, we always tend to find our acquaintances in that city in advance, regarding them as the local guides to recommend interesting places. This phenomenon is called guide mechanism (local guide) in this paper. In order to determine whether a user is a guide to a specific category, we need to calculate the user's expertise in this category. This process is called guide expertise discovery.

A guide has two characteristics. (1) A guide is respect to a specific category in a city. If someone is familiar with all the restaurants in this city, we may regard him as a guide for the restaurant category. (2) Compared to the average residents, locations visited by a guide are of higher quality, and therefore their references are more valuable.

Assuming the new city has n location categories. First, divide the city's user-location matrix into n sub user-location matrix by category. For each sub-matrix, utilize Hypertext-Induced Topic Search (HITS) algorithm [8] to calculate the user's expertise for this category shown as Equations (3) and (4).

$$a(v, C) = \sum_{u \in U} h(u, C) \tag{3}$$

$$h(u, C) = \sum_{v \in V} a(v, C) \tag{4}$$

v denotes a location in this new city, C is the category of location v and u means user u. $a(v, C)$ represents the authority of place v in category C, and it is the sum of all the users' expertise who have visited location v for category C; $h(u, C)$ indicates the user u's expertise in terms of category C, and it is the sum of all locations' authority which were visited by user u for category C.

3.2. **Candidate guide selection.** Assuming we need to recommend K locations to user u, his history check-in records are $\{l_1, l_2, \ldots, l_i, \ldots\}$, Hub_j is a set containing all users sorted by their expertise for category j from big to small. We select M candidate guides as follows.

Step1 User category preference learning. Use TF-IDF [9] method to calculate the user category preference and get the preference vector $P(w_1, w_2, \ldots, w_n)$, where w_j denotes users' interest in category j.

Step2 Candidate guide selection by proportion. The distribution of M candidate guides should be in accordance with user's category preferences as much as possible, so choose the candidate guides from each category by proportion, thus the first $M * \frac{w_j}{\sum_{i=1}^{n} w_i}$ users from Hub_j.

Suppose a user's preference for categories $c1$, $c2$, $c3$ are respectively $(0.4, 0.1, 0.5)$, and need $M = 10$ candidate guides. Then there are $10 * 0.4 = 4$ candidate guides form $c1$, $10 * 0.1 = 1$ from $c2$ and $10 * 0.5 = 5$ from $c3$. As can be seen, the more the target user prefers a category, the larger the proportion of candidate guides in this category accounts for.

3.3. **Similarity computation and candidate location rating.** Users have different behavior habits for different categories of locations [10]. For example, user A likes morning exercise while user B prefers night exercise. So for category j, user's behavior can be represented as a vector $U^j = (u_1^j, u_2^j, \ldots, u_m^j, \ldots, u_{24}^j)$, and u_m^j represents the probability of the user u to visit category j at the m-th hour of a day. We can see the probability values of a user consist of a curve which is called category temporal curve in this paper, shown as Figure 2. This curve can describe a user's behavior for a specific category. So the similarity computation between the two users is converted to the similarity computation between two curves.

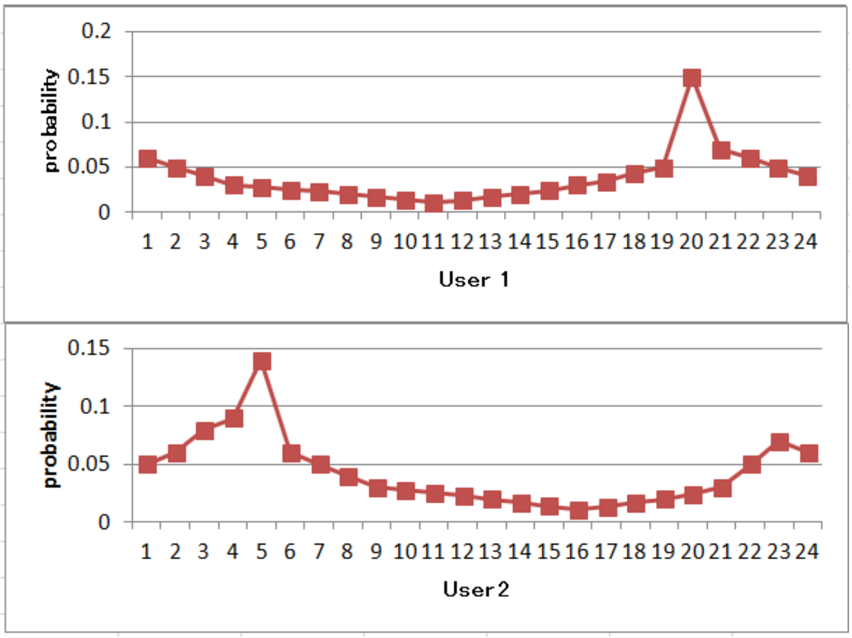

FIGURE 2. User category temporal curve

For a specific category, users' check-in activities generally concentrate in one or several time periods; thus there will be a peak in the category temporal curve, which represents the time zone of the maximum visit. The peak is the most distinct point in a curve, so we use it for curve similarity calculation.

For category j, the category temporal curves of users u and w are represented as $U^j = (u_1^j, u_2^j, \ldots, u_m^j, \ldots, u_{24}^j)$, $W^j = (w_1^j, w_2^j, \ldots, w_m^j, \ldots, w_{24}^j)$, u reaches a peak at time $T1$, w reaches a peak at time $T2$, and the similarity between the two curves is signed by $CSim(U^j, W^j)$.

$$CSim(U^j, W^j) = \begin{cases} \frac{1}{|T1-T2|}, & T1 \neq T2 \\ 1, & T1 = T2 \end{cases} \tag{5}$$

As a user may have visited different categories of location, the similarity between users u and w similarity $TSim(U, W)$ is defined as follows:

$$TSim(U, W) = \frac{\sum_{j_i \in C} CSim(U^{j_i}, W^{j_i})}{|C|} \tag{6}$$

C is a category set $\{j_1, j_2, j_3, \ldots\}$; U^{j_i}, W^{j_i} are the temporal curves of users u and w for category j_i. We can see $TSim(U, W)$ accounts of the behavior of users u and w in all categories.

After getting the similarity between target user and his candidate guides, the following formula measures the probability for user u to visit candidate location v:

$$p(U, V) = \sum_{W \in Candidate} TSim(U, W) \times r_{WV} \tag{7}$$

$TSim(U, W)$ is the temporal similarity between two users, and r_{WV} denotes the interest of candidate guide w to place v.

4. The Experimental Results and Analysis.

4.1. Experiment settings.

Dataset Description. We use Gowalla dataset provided by [5], including 1,617,811 check-in activities, 41326 users, 630,000 locations and 355 categories. New York and Austin are America's most active cities, with check-in records up to 765,642 times and

570,306 times. So these two cities are selected as sample cities in this paper. 1000 users are chosen respectively from New York and Austin to test.

Comparative Approaches. Our recommendation method presented will compare with the following three recommendation methods: (1) Most Popular (MP): it recommends the most popular locations to target user; (2) Geographic Distance (GD): it recommends the nearest locations to target user [11,12]; (3) User Interest, Social and Geographical Influences (USG): it is proposed in [5] using collaborative filtering techniques based on linear fusion of user preferences, and geographic factors affecting friends.

Performance Evaluation. The dataset is split into a training set and a test set. User's local check-in records are regarded as the training set and the check-in records in target city treated as the test set. We adopt the evaluation method and the metrics $Recall@K$ applied in references [13-15].

For each test case (u, v) in S_{test}, the recommendation result has two states: hit or miss. $Recall@K$ is defined as Equation (8)

$$Recall@K = \frac{\#hit}{num(S_{test})} \tag{8}$$

$\#hit$ represents the number of hit test cases in S_{test}, and $num(S_{test})$ denotes the number of all test cases.

4.2. **Experimental results.**

Influence of Parameters. The recommendation method in this paper has a parameter M mentioned in Section 3.2. M represents the total number of candidate guides. By adjusting the value of M, we can get the best recommendation result. The result is also affected by K, the number of recommendation items.

Combined with real life experience, this paper confined K in $[1, 20]$ and explored how the different M values affect the accuracy of recommendation in a number of K ($K = 5, 10, 15, 20$). The results are displayed in Figure 3.

Regardless of the value of K, the increase of M leads to the rise of $Recall@K$ and when M increases to a certain extent, $Recall@K$ changes little. For further exploration, we calculate the optimal value of M from $K = 1$ to 20, and find when $M = K$ we get the best result.

Comparison. Figure 4 shows the performance of these four methods in new city location recommendation scene. It is apparent that they have significant performance, our method performs best and USG is the least effective.

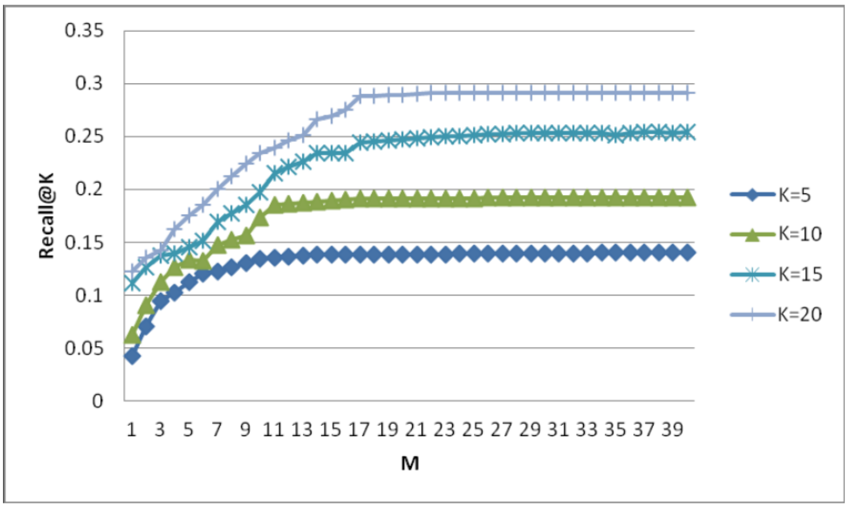

FIGURE 3. Tuning parameter M

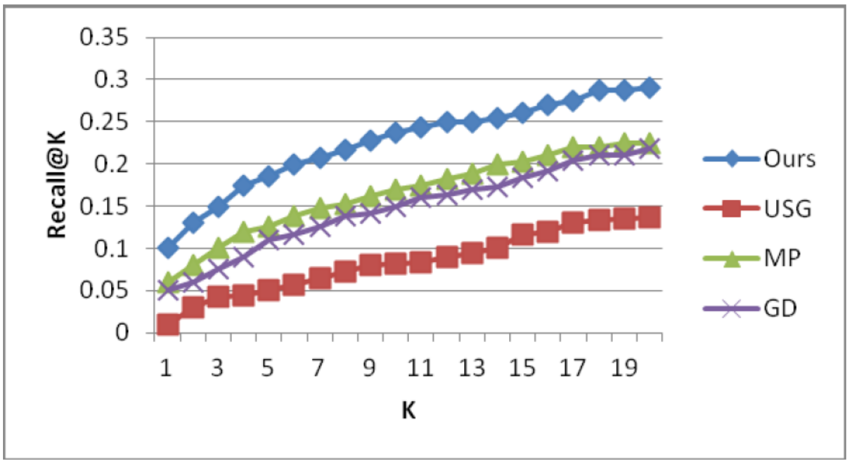

Figure 4. Performance comparison in new city scene

Through the foregoing analysis, the most important things in new city recommendation are user's preference and local preference of the new city. Our method performs well because, on the one hand, it learns user's preference by category information which can be transferred to a new city, on the other hand it uses the guide mechanism which captures the preference in new city. The worst performance belongs to USG which is also based on user-based CF. We know the most similar neighbors for target user should all live in the same city with home, thus these neighbor users rarely go to places in the new city, so USG can hardly recommend locations in new city to the target user.

5. **Conclusions.** This paper mainly discusses how to solve the location recommendation in a new city scene. We put forward a time-aware collaborative location recommendation on the basis of user-based collaborative filtering. It uses the guide mechanism, selecting a series of guides for each category in each city, and then calculates the similarity between users and candidate guides. During the similarity calculation, it fully considers the temporal information. By experimental results, it is shown that our method generates a more accurate recommendation in the new city scene. For future work, we plan to conduct semantic analysis on location tags and user comments to capture user preferences and make recommendation.

Acknowledgment. This work is supported by the Fundamental Research Funds for Central Universities (XDJK2009C027) and Science & Technology project (2013001287).

REFERENCES

[1] M. Li, X. Wang and J. Zhang, Study on check-in and related behaviors of location-based social networks, *Computer Science*, vol.40, no.10, pp.72-76, 2013.
[2] *https://foursquare.com/about/*.
[3] J. Bao and Y. Zheng, A survey on recommendations in location-based social networks, *ACM Transaction on Intelligent Systems and Technology*, 2013.
[4] B. Liu, Y. Fu, Z. Yao and H. Xiong, Learning geographical preferences for point-of-interest recommendation, *Proc. of KDD 2013*, pp.1043-1051, 2013.
[5] Q. Yuan, G. Cong, Z. Ma, A. Sun and N. M. Thalmann, Time-aware point-of-interest recommendation, *Proc. of SIGIR 2013*, pp.363-372, 2013.
[6] H. Gao, J. Tang, X. Hu and H. Liu, Content-aware point of interest recommendation on location-based social networks, *The 29th AAAI Conference on Artificial Intelligence*, Austin, TX, 2015.
[7] M. Ye, P. Yin, W. Lee and D.-L. Lee, Exploiting geographical influence for collaborative point-of-interest recommendation, *Proc. of SIGIR 2011*, pp.325-334, 2011.
[8] *http://blog.csdn.net/hguisu/article/details/8013489*.
[9] *http://zh.wikipedia.org/wiki/TF-IDF*.

[10] D. Preoţiuc-Pietro and T. Cohn, Mining user behaviours: A study of check-in patterns in location based social networks, *Proc. of the 5th Annual ACM Web Science Conference*, pp.306-315, 2013.
[11] J. J. Levandoski, M. Sarwat, A. Eldawy and M. F. Mokbel, LARS: A location-aware recommender system, *Proc. of ICDE 2012*, pp.450-461, 2012.
[12] M. Sarwat, J. J. Levandoski, A. Eldawy and M. F. Mokbel, LARS*: An efficient and scalable location-aware recommender system, *IEEE Transactions on Knowledge and Data Engineering*, pp.99-106, 2013.
[13] P. Cremonesi, Y. Koren and R. Turrin, Performance of recommender algorithms on top-N recommendation tasks, *ACM Conference on Recommender Systems*, pp.39-46, 2010.
[14] W. Y. Chen, J. C. Chu, J. Luan, H. Bai, Y. Wang and E. Y. Chang, Collaborative filtering for orkut communities: Discovery of user latent behavior, *International Conference on World Wide Web*, pp.681-690, 2009.
[15] H. Yin, B. Cui, J. Li, J. Yao and C. Chen, Challenging the long tail recommendation, *Proc. of VLDB Endow.*, vol.5, no.9, pp.896-907, 2013.