Using Multi-Criteria Decision Making for Personalized Point-of-Interest Recommendations

Yan Lyu Chi-Yin Chow Ran Wang Victor C. S. Lee
Department of Computer Science, City University of Hong Kong, Hong Kong
{yanlv2-c@my., chiychow@, ranwang3-c@my., csvlee@}cityu.edu.hk

ABSTRACT

Location-based business review (LBBR) sites (e.g., Yelp) provide us a possibility to recommend new points of interest (POIs) for users. The geographical position and category of POIs have been considered as two major factors in modeling users' preferences. However, it is argued that the user's visiting behaviors are also affected by the attributes of POIs, which reflect the basic features of the POIs. Besides, a user may have different preference levels on the same POI with regard to different criteria. To this end, we propose a new personalized POI recommendation framework using Multi-Criteria Decision Making (MCDM). Firstly, preference models are built for the user's geographical, category, and attribute preferences. Then, an MCDM-based recommendation framework is designed to iteratively combine the user's preferences on the three criteria and select the top-N POIs as a recommendation list. Experimental results show that our framework not only outperforms the state-of-the-art POI recommendation techniques, but also provides a better trade-off mechanism for MCDM than the weighted sum approach.

Categories and Subject Descriptors

H.2.8 [**Database Management**]: Database Applications-Spatial databases and GIS; H.3.3 [**Information Search and Retrieval**]: Information Filtering

General Terms

Algorithms, Experimentation.

Keywords

Point-of-interest recommendations, location-based business reviews, multi-criteria decision making

1. INTRODUCTION

In Location-based business review (LBBR) sites (e.g., Yelp and Foursquare), recommending points-of-interest (POIs) for users based on their visiting preferences is one of the most popular services. Most existing POI recommendation techniques model user preferences based on only two criteria. (1) **The geographical positions of POIs**. The geographical criterion has a significant influence on the user's visiting behaviors, and has been considered as a major factor in POI recommendations [3, 5, 11]. (2) **The categories of POIs**. A POI usually belongs to a main category, such as restaurant, shopping, entertainment, etc. The category criterion also affects the user's visiting behaviors a lot. For instance, POIs are recommended to users by analyzing their category transition patterns [1], or by consulting local experts with similar category preferences [6].

Although both the geographical positions and categories of POIs have exhibited good performance on various recommenders, the user's preferences are also affected by **the attributes of POIs**. Consider a restaurant as an instance, people will consider a restaurant's price level before they dine in the restaurant. Furthermore, drinkers prefer places with a full bar, while Internet addicts prefer places with free Wi-Fi. Thus, we are motivated to incorporate the attributes of POIs into POI recommendations. To the best of our knowledge, our personalized POI recommendation framework is the first one to take all these three criteria into consideration.

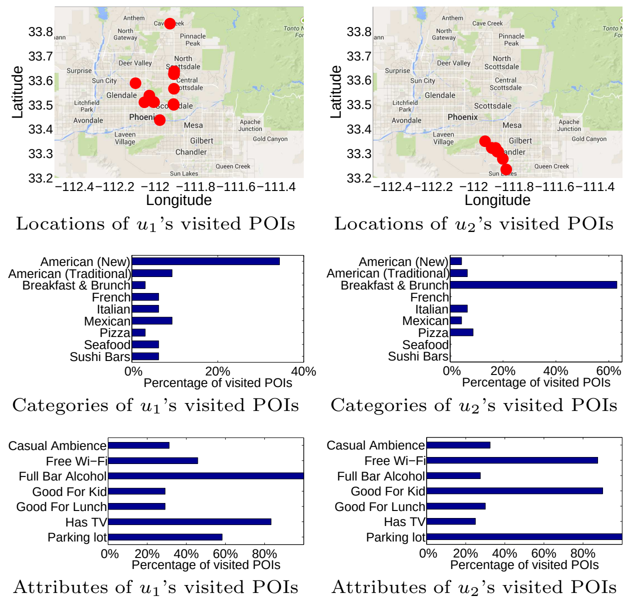
Figure 1 depicts the characteristics of the visited POIs of two different users, u_1 and u_2, which are collected from Yelp[1]. Obviously, the geographical positions, categories, and attributes of POIs can affect these users' preferences. For the geographical criterion, u_1 trends to visit more distant POIs compared to u_2. For the category criterion, u_1 likes restaurants serving new American food, while u_2 likes restaurants with breakfast and brunch, but shows no interest in French food, seafood, and sushi. Finally, for the attribute criterion, u_1 is interested in full bar alcohol and u_2 likes visiting restaurants that are good for kids and provide free Wi-Fi and parking lots.

A user may have different preference levels on the same POI with regard to different criteria. For example, given three candidate POIs, p_1, p_2, and p_3, for the geographical criterion, the user may prefer p_1. However, the user may prefer p_2 or p_3 based on its attributes or categories, respectively. As a result, it is undesirable to recommend POIs based on only one criterion; thus, we design our personalized POI recommendation framework by using multi-criteria decision making (MCDM) systems.

MCDM [4] is a well-known branch of decision making. It evaluates and ranks a set of alternatives based on multiple

[1]http://www.yelp.com/dataset_challenge/

SIGSPATIAL'14, November 04 - 07 2014, Dallas/Fort Worth, TX, USA
Copyright 2014 ACM 978-1-4503-3131-9/14/11 ...$15.00
http://dx.doi.org/10.1145/2666310.2666479.



(a) User u_1 (b) User u_2

Figure 1: Preferences on the geographical positions, categories and attributes of two users' POIs.

conflicting criteria, and selects the best one with a trade-off mechanism [9]. It is noteworthy that recommending POIs for a user based on multiple conflicting criteria is a decision making process, which selects the best POIs that match the user's preferences.

In this paper, we propose an MCDM-based personalized POI recommendation framework by aggregating all the three criteria. In particular, we propose the geographical preference model, category preference model, and attribute preference model to estimate the user's preference levels on a POI with regard to its geographical position, category, and attributes, respectively. Then, we define a preference preorder based on each criterion for all the POIs, and use the distance between preorders to quantify their dominant states. The POIs that have high dominating indices and low dominated indices in the preference preorders are recommended.

2. USER PREFERENCE MODELS

In this section, we describe how to model the three preference criteria, i.e., the geographical positions, categories, and attributes of POIs, and then present how to fuse these models with the user-based collaborative filtering.

2.1 Geographical Criterion

A kernel density estimation (KDE), which can be used for arbitrary distribution estimation, is employed to personalize the geographical preference for each user [11].

Geographical preference model. We use the most popular kernel function $K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. Given a user u, let $\mathcal{L}_u = \{l_i\}_{i=1}^m$ be the set of m POIs that have been visited by u, and D be the set of distances between every pair of the visited POIs. For a new POI candidate l_j, we define $d_{ij} = |l_i - l_j|$ as the Euclidean distance between l_i and l_j, where $l_i \in \mathcal{L}_u$. The kernel density of d_{ij} is defined as:

$$f(d_{ij}) = \frac{1}{|D|\, b} \sum\nolimits_{d' \in D} K\left(\frac{d_{ij} - d'}{b}\right), \tag{1}$$

where $b = \left(\frac{4\hat{\sigma}^5}{3m}\right)^{\frac{1}{5}}$ is an optimized bandwidth [7] and $\hat{\sigma}$ is the standard deviation of the samples in D.

Geographical preference estimation. The geographical preference of the user on POI l_j, denoted by $GeoRating(l_j)$, can be calculated by taking the average of all its probability densities, i.e.,

$$GeoRating(l_j) = \frac{1}{m} \sum\nolimits_{i=1}^{m} f(d_{ij}). \tag{2}$$

A higher value of $GeoRating(l_j)$ indicates that a user has a higher probability to visit l_j.

2.2 Category Criterion

The frequency of a POI category visited by a user can reflect his/her preference on this category. However, since a high visiting frequency of a category could also be caused by its popularity among all POIs, we adopt Term Frequency Inverse Document Frequency (TF-IDF) [10] to model the user's category preferences.

Category preference model. The TF-IDF value of a category c^* is calculated as:

$$tf \cdot idf(c^*) = \frac{n_{c^*}}{n} \cdot \log \frac{N}{N_{c^*}}, \tag{3}$$

where n_{c^*} is the user's visiting frequency of c^*, n is the number of the user's visit records, N is the number of POIs, and N_{c^*} is the number of POIs in c^*.

Category preference estimation. Given a new POI l_j with its categories in each level, $\{c_1^{(j)}, c_2^{(j)}, \ldots, c_H^{(j)}\}$, where H is the number of category levels of l_j, the category rating of the user on POI l_j, denoted by $CateRating(l_j)$, is calculated by a weighted sum of the user's preferences on the category of l_j at each level:

$$CateRating(l_j) = \sum\nolimits_{h \in \{1,2,\ldots,H\}} \beta \cdot tf \cdot idf(c_h^{(j)}), \tag{4}$$

where $\beta = \frac{1}{2^{H-h}}$, i.e., a lower weight is given to a category at a lower level.

2.3 Attribute Criterion

A user's preference on the specific value of an attribute can be deduced from his/her visiting frequency to POIs that have this value. Thus, we employ TF-IDF to estimate the user's preference on the values of an attribute. For different attributes, we employ an entropy weight to estimate the user's weighting on each attribute.

Attribute preference model. Given an attribute a and its possible values $\mathcal{V} = \{v_1, v_2, \ldots, v_{|\mathcal{V}|}\}$, the preference on value v_s is calculated as:

$$tf(a = v_s) = \frac{n(a = v_s)}{n} \text{ and } idf(a = v_s) = \log \frac{N}{N(a = v_s)},$$

where $n(a = v_s)$ and $N(a = v_s)$ are the number of visited POIs and the total number of POIs with v_s for attribute a, $v_s \in \mathcal{V}$, respectively.

Users may have different preference levels on different attributes. We learn the weight of a user's preference on an attribute from his/her visited POIs. If a user has a strong preference on the specific value of an attribute (i.e., this value has a much higher visiting frequency than other values), the values of this attribute for this user are less diverse. On the other hand, if a user has no specific preference on the values of an attribute, this attribute has a wide diversity of values. Therefore, we use entropy to describe the diversity of the values of an attribute a in a user's visited POIs as:

$$E(a) = -\frac{1}{\log |\mathcal{V}|} \sum\nolimits_{s=1}^{|\mathcal{V}|} tf(a = v_s) \cdot \log tf(a = v_s), \quad (5)$$

where $0 \leq E(a) \leq 1$. Obviously, a smaller entropy value indicates that the values of an attribute are less diverse, i.e., a user shows a stronger specific preference on the attribute. Thus, the weight of a is defined as:

$$EWeight(a) = 1 - E(a). \quad (6)$$

Attribute preference estimation. Given a new POI l_j with an attribute set $\mathcal{T} = \{a_1, a_2, \ldots, a_{|\mathcal{T}|}\}$ and their values are $\{v_{a_1}^{(j)}, v_{a_2}^{(j)}, \ldots, v_{a_{|\mathcal{T}|}}^{(j)}\}$, respectively, the attribute preference of a user on l_j is modeled as:

$$AttriRating(l_j) = \sum\nolimits_{t=1}^{|\mathcal{T}|} tf(a_t = v_{a_t}^{(j)}) \cdot idf(a_t = v_{a_t}^{(j)}) \cdot EWeight(a_t). \quad (7)$$

2.4 Fusing Preferences with User Opinions

We employ the standard user-based collaborative filtering (CF) model, in which, the rating of a user u on a new POI l_j, $\hat{r}(l_j)$, is calculated by considering opinion of other users who have similar check-in history [8]. Then, we fuse the three preference ratings with user-based CF through a product rule, and finally get the fused rating scores of a candidate POI l_j for user u with regard to the geographical, category and attribute criteria.

3. MCDM-BASED POI RECOMMENDATION FRAMEWORK

3.1 Preliminaries

Preference preorder. Given a set of POIs that have not been visited by a user, $\mathcal{L}^* = \{l_1, l_2, \ldots, l_{|\mathcal{L}^*|}\}$, and a set of criteria, $\mathcal{C} = \{C_1, C_2, \ldots, C_{|\mathcal{C}|}\}$, for each criterion $C_k \in \mathcal{C}$, three ordinal relations can be defined as:

- $l_i \succ l_j$: l_i is preferred to l_j, i.e., $C_k(l_i) - C_k(l_j) > T(C_k)$;
- $l_i \prec l_j$: l_j is preferred to l_i, i.e., $C_k(l_j) - C_k(l_i) > T(C_k)$;
- $l_i \approx l_j$: l_i is indifferent to l_j, i.e., $|C_k(l_i) - C_k(l_j)| \leq T(C_k)$,

where $C_k(l_j)$ represents the fused rating score of l_j with regard to a specific criterion C_k (e.g., the geographical, category or attribute criterion in our framework), and $T(C_k)$ is a threshold of indifference, $T(C_k) = \frac{\max_{l_j \in \mathcal{L}^*} C_k(l_j) - \min_{l_j \in \mathcal{L}^*} C_k(l_j)}{|\mathcal{L}^*|}$.

Based on these three relations, a preference preorder of the alternatives with regard to a criterion is given.

Definition 1. *(Preference preorder.) Given a set of alternatives $\mathcal{L}^* = \{l_1, l_2, \ldots, l_{|\mathcal{L}^*|}\}$ and a criterion C_k on them, the order $l_1^* \succ\approx l_2^* \succ\approx \ldots \succ\approx l_{|\mathcal{L}^*|}^*$ is called a preference preorder of the POIs in $\mathcal{L}^*$ with regard to C_k if and only if they satisfy $l_1^* \neq l_2^* \neq \ldots \neq l_{|\mathcal{L}^*|}^* \in \mathcal{L}^*$ and $C_k(l_1^*) \geq C_k(l_2^*) \geq \ldots \geq C_k(l_{|\mathcal{L}^*|}^*)$.*

We denote the preference preorder of $\mathcal{L}^*$ with regard to C_k as $Preorder(C_k)$. Obviously, $Preorder(C_k)$ indicates the priorities of the alternatives with respect to C_k. For example, l_1^* is the most preferred alternative and $l_{|\mathcal{L}^*|}^*$ is the least preferred one in this preorder.

Distance of ordinal relations. In a preference preorder, an alternative with more relations of $\succ$ and less relations of $\prec$ should have a higher priority. The distance between relations is employed to measure the priority of an alternative for a criterion [2], and the adopted values of these distances are given as following: (1) $dist(\prec,\prec) = dist(\succ,\succ) = dist(\approx,\approx) = 0$; (2) $dist(\prec,\approx) = dist(\approx,\prec) = 1$; (3) $dist(\succ,\approx) = dist(\approx,\succ) = 1$; and (4) $dist(\succ,\prec) = dist(\prec,\succ) = 2$.

Algorithm 1 MCDM-based POI Recommendation Framework

Input: A set of visited POIs for user u: $\mathcal{L} = \{l_i\}_{i=1}^{|\mathcal{L}|}$, a set of new POIs for u: $\mathcal{L}^* = \{l_j\}_{j=1}^{|\mathcal{L}^*|}$, and the number of POIs to be recommended: N.
Output: A set of top-N recommended POIs: $\mathcal{R}$.

for each $l_j \in \mathcal{L}^*$ **do**
  Compute $GeoRaing(l_j)$, $CateRating(l_j)$, and $AttriRating(l_j)$, fuse them with the user-based CF model in Section 2.4
end for
while $|\mathcal{R}| < N$ **do**
  Determine the threshold of each criterion
  Generate preference preorders for the three criteria: $Preorder(C_{geo})$, $Preorder(C_{cate})$, and $Preorder(C_{attri})$
  Compute the MCDM-based rating score of each l_j based on the three preference preorders by Eq. (12), i.e., $D(l_j) = \Phi^{\prec}(l_j) - \Phi^{\succ}(l_j)$
  Select the alternative l^* with the largest MCDM-based rating score, i.e., $l^* = \mathrm{argmax}_{l_j \in \mathcal{L}^*} D(l_j)$
  $\mathcal{R} \leftarrow \mathcal{R} \cup \{l^*\}$; $\mathcal{L}^* = \mathcal{L}^* - \{l^*\}$
end while

3.2 Recommendation Framework

The priorities of a given POI l_j in $Preorder(C_{geo})$, $Preorder(C_{cate})$, and $Preorder(C_{attri})$ may differ a lot. That is to say, the same POI may be ranked with a very high priority in one criterion but a very low priority in another one. In this case, it is important to have a tradeoff mechanism among different criteria, in order to prioritize the best alternatives as a recommendation result.

Assume the relation of l_i and l_j for a user with regard to criterion C_k is $R_{ij}^{(k)}$, where $R_{ij}^{(k)} \in \{\succ, \prec, \approx\}$. The k-th criterion dominated index of l_i is defined as:

$$\psi_k^{\succ}(l_i) = \sum\nolimits_{j \neq i} dist(\succ, R_{ij}^{(k)}), \quad (8)$$

and k-th criterion dominating index of l_i is defined as:

$$\psi_k^{\prec}(l_i) = \sum\nolimits_{j \neq i} dist(\prec, R_{ij}^{(k)}). \quad (9)$$

The k-th criterion dominated index and k-th criterion dominating index of l_i measure the degree of it being dominated and dominating others under the criterion C_k, respectively.

To further consider a set of criteria, i.e., $\mathcal{C} = \{C_1, C_2, \ldots, C_{|\mathcal{C}|}\}$, the dominated index of l_i is defined as:

$$\Phi^{\succ}(l_i) = \sum\nolimits_{k=1}^{|\mathcal{C}|} w_k \psi_k^{\succ}(l_i), \quad (10)$$

and the dominating index of l_i is defined as:

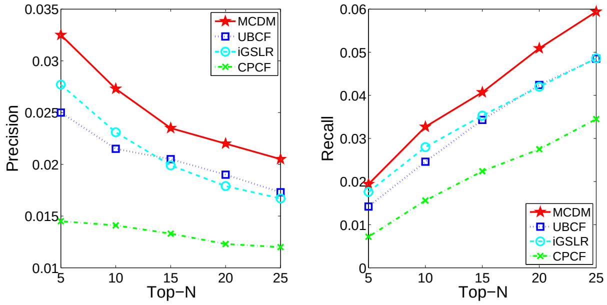
$$\Phi^{\prec}(l_i) = \sum\nolimits_{k=1}^{|\mathcal{C}|} w_k \psi_k^{\prec}(l_i), \quad (11)$$

where w_k is the weight of the criterion C_k. The weights of the geographical, category, and attribute preferences will be determined through empirical studies (Section 4).

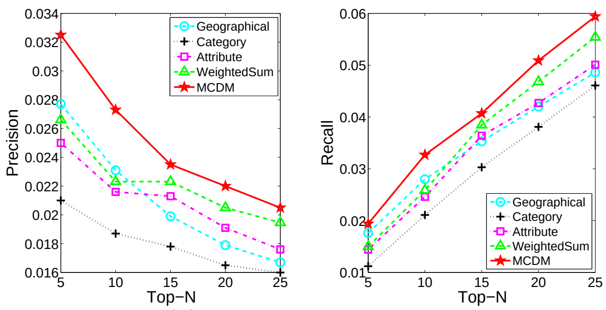
Since an alternative with a higher dominating index and a lower dominated index should have a higher priority, the MCDM-based rating score is developed as:

$$D(l_i) = \Phi^{\prec}(l_i) - \Phi^{\succ}(l_i). \quad (12)$$

An iterative framework is designed to recommend the top-N POIs with the highest MCDM-based rating score which



(a) Effect of the number of recommended POIs (Top-N)



(b) Effectiveness of MCDM

Figure 2: Experimental results.

is calculated by Equation (12), as depicted in Algorithm 1.

4. EXPERIMENTS

We conduct experiments on a real-world data set of POIs in the greater Phoenix region of USA provided by the "Yelp Dataset Challenge"[2]. The data set contains POIs with different categories, such as restaurant, shopping, etc. Due to space limitation, only the experimental results of restaurant data set are presented.

We compare our method with the following three existing POI recommendation methods: (1) User-based collaborative filtering (UBCF) [8], which recommends POIs by consulting the users with similar check-in histories; (2) iGSLR [11], which employs kernel density estimation to model the user's geographical preferences; and (3) Category preferred collaborative filtering (CPCF) [1], which learns the user's preferences based on the categories of visited POIs.

We conducted experiments to find the weights of the criterion of geographical w_g, category w_c, and attribute w_a. $w_g = 0.43$, $w_c = 0.14$ and $w_a = 0.43$ are set for the restaurant data set.

Effect of Top-N. Figure 2a depicts the average precision and recall of all the evaluated recommendation methods with respect to various values of N (i.e., the number of recommended POIs) from 5 to 25. Our proposed MCDM-based personalized POI recommendation framework (denoted as MCDM) outperforms the three baseline methods for all the values of N, in terms of precision and recall.

Effectiveness of MCDM. Figure 2b demonstrates the effectiveness of our MCDM by comparing its performance with the method that only considers the geographical, category, or attribute criterion (denoted as Geographical, Category, or Attribute, respectively) and the method that employs the weighted sum approach to combine the geographical, category and attribute preferences (denoted as WeightedSum). The results of this experiment verify our two important claims. (1) Our MCDM that considers all the three criteria provides better quality of POI recommendations than the method only considers one criterion. (2) Our MCDM has a better trade-off mechanism to combine the three conflicting criteria than the weighted sum approach.

5. CONCLUSION

We proposed a new personalized POI recommendation framework based on multi-criteria decision making (MCDM). We firstly designed the preference and estimation models for geographical, category and attribute criteria. Given a user and a set of candidate POIs, our method first arranges the candidate POIs in a preference preorder for each criterion, and then it iteratively combines the three preorders of the three criteria by computing dominated index and dominating index for each candidate, and recommends the top-N POIs for the user. Experimental results show that our framework significantly outperforms the state-of-the-art POI recommendation techniques and has a more effective trade-off mechanism for multiple criteria than the weighted sum approach.

6. ACKNOWLEDGMENTS

C.-Y. Chow and R. Wang were partially supported by a research grant (CityU Project No. 9231131). Y. Lyu and Victor C. S. Lee were partially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (CityU Project No. 115312).

7. REFERENCES

[1] J. Bao, Y. Zheng, and M. F. Mokbel. Location-based and preference-aware recommendation using sparse geo-social networking data. In *ACM SIGSPATIAL*, 2012.

[2] G. Campanella and R. A. Ribeiro. A framework for dynamic multiple-criteria decision making. *Decision Support Systems*, 52(1):52–60, 2011.

[3] C. Cheng, H. Yang, I. King, and M. R. Lyu. Fused matrix factorization with geographical and social influence in location-based social networks. In *AAAI*, 2012.

[4] J. Figueira, S. Greco, and M. Ehrgott. *Multiple criteria decision analysis: state of the art surveys*, volume 78. Springer, New York, 2005.

[5] B. Liu, Y. Fu, Z. Yao, and H. Xiong. Learning geographical preferences for point-of-interest recommendation. In *ACM SIGKDD*, 2013.

[6] X. Liu, Y. Liu, K. Aberer, and C. Miao. Personalized point-of-interest recommendation by mining users' preference transition. In *ACM CIKM*, 2013.

[7] B. W. Silverman. *Density estimation for statistics and data analysis*. Chapman and Hall, London, 1986.

[8] H. Wang, M. Terrovitis, and N. Mamoulis. Location recommendation in location-based social networks using user check-in data. In *ACM SIGSPATIAL*, 2013.

[9] R. Wang and S. Kwong. Active learning with multi-criteria decision making systems. *Pattern Recognition*, 47(9):3106–3119, 2014.

[10] H. C. Wu, R. W. P. Luk, K. F. Wong, and K. L. Kwok. Interpreting tf-idf term weights as making relevance decisions. *ACM TOIS*, 26(3):13, 2008.

[11] J.-D. Zhang and C.-Y. Chow. iGSLR: Personalized geo-social location recommendation: a kernel density estimation approach. In *ACM SIGSPATIAL*, 2013.

[2]http://www.yelp.com/dataset_challenge/