2015 IEEE International Conference on Data Mining

Learning User Preferences across Multiple Aspects for Merchant Recommendation

Xin Li [1], Guandong Xu [2], Enhong Chen *[1], Lin Li [3]
[1] *University of Science and Technology of China,* [2] *University of Technology, Sydney,* [3] *Wuhan University of Technology*
[1] *{leexin,cheneh}@ustc.edu.cn,* [2] *Guandong.Xu@uts.edu.au,* [3] *cathylilin@whut.edu.cn*

***Abstract*—With the pervasive use of mobile devices, *Location Based Social Networks* (LBSNs) have emerged in past years. These LBSNs, allowing their users to share personal experiences and opinions on visited merchants, have very rich and useful information which enables a new breed of location-based services, namely, *Merchant Recommendation*. Existing techniques for merchant recommendation simply treat each merchant as an item and apply conventional recommendation algorithms, e.g., *Collaborative Filtering*, to recommend merchants to a target user. However, they do not differentiate the user's real preferences on various aspects, and thus can only achieve limited success. In this paper, we aim to address this problem by utilizing and analyzing user reviews to discover user preferences in different aspects. Following the intuition that a user rating represents a personalized rational choice, we propose a novel utility-based approach by combining collaborative and individual views to estimate user preference (i.e., rating). An optimization algorithm based on a Gaussian model is developed to train our merchant recommendation approach. Lastly we evaluate the proposed approach in terms of effectiveness, efficiency and cold-start using two real-world datasets. The experimental results show that our approach outperforms the state-of-the-art methods. Meanwhile, a real mobile application is implemented to demonstrate the practicability of our method.**

I. Introduction

Location-Based Social Networks (LBSNs) allow users to explore various merchants, generally known as Points-of-Interest (POIs), such as restaurants, stores, and theatres, by sharing their experiences, opinions and ratings on those merchants. As a result, a new breed of location-based services, namely, *merchant recommendation*, has emerged aiming to recommend possibly preferred merchants to targeted users. To support merchant recommendations, conventional neighborhood-based or model-based approaches, along with their variant forms, can be directly applied by simply treating merchants as ordinary items in traditional recommender systems. To alleviate the data sparsity and cold start problems, which are commonly encountered in real applications, some studies further incorporate contextual information, such as geographical location, social networks and temporal information, to enhance recommendation quality [2], [4], [12], [20], [22].

Despite the successes and improvements, there are still several open issues. Firstly, existing recommendations mainly rely on one overall rating without differentiating user preferences across various aspects, although distinct preferences on different aspects of a merchant are observed[1]. We argue that the overall rating is merely a preference fusion where the rich

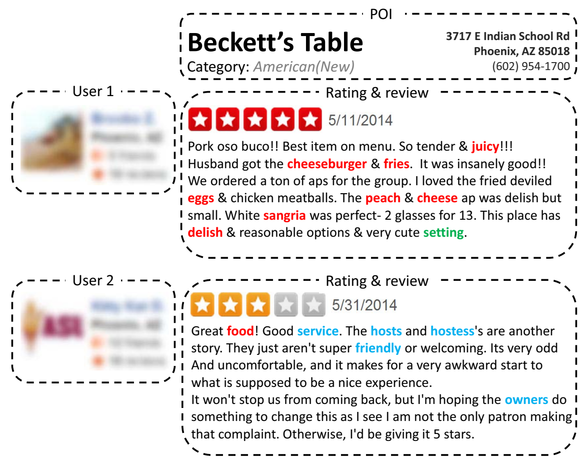


Figure 1. Example of Yelp Rating and Review Clip

information of individual preferences in multiple aspects is lost. Consider the example shown in Figure 1 where different ratings are given by two users on the same restaurant. We can observe from their reviews that these two users exhibit distinctive preferences over aspects of food and service - *user 1* pays more attention to the taste of food (marked red) while *user 2* cares more about the services (marked blue). As indicated in the example, the restaurant may offer good dishes but disappointing services, i.e., the quality in different aspects of a merchant may vary. Secondly, in most LBSNs, apart from ratings that offer aggregated views of users, user reviews reveal the accurate and valuable spectrum information about individual user preferences. Both numeric ratings and detailed reviews provide users' overall and individual views on merchants. Thus, they should be jointly and simultaneously considered for merchant recommendations. Thirdly, although a spectrum of user preference aspects may be deduced from reviews through text processing techniques, how to use them for modelling user preferences is also a problem. Inspired by the above observations, we propose to capture multiple aspects of user preferences to make precise merchant recommendations.

In this paper, we particularly explore the notion of *utility*, a concept widely adopted in economics and psychophysics, to describe, explain and quantitatively measure how a particular item satisfies a user's needs. Generally, review writing can be deemed as a procedure of evaluating all aspects of a merchant in which users' satisfaction across multiple aspects is directly reflected. We argue that utility could be a measure of satisfaction of one user with a merchant over a spectrum

[1]The overall rating is usually an aggregated average of the preferences over various ratings on the merchant.
* denotes the corresponding author.

1550-4786/15 $31.00 © 2015 IEEE
DOI 10.1109/ICDM.2015.10

IEEE computer society

of preferences. Therefore, we specify such aspect-based satisfaction as *individual utility*. We assume that a larger utility on a specific merchant indicates a more preferable choice to the user. On the other hand, we argue that employing matrix factorization on the rating matrix of users to merchants may derive a different kind of utility, termed *collaborative utility*, which reflects a collaborative view. Finally, we fuse the aspect-based individual utility (deduced from reviews) and the latent factor-based collaborative utility (from ratings) to form an overall utility. By optimizing the utility-based objective function via a Gaussian model, we can accurately estimate the user preference spectrum, and in turn predict the utility of the target user on a specific merchant (i.e., rating). In summary, we make the following contributions:

- We propose to capture user preference across multiple aspects from review text in order to enhance merchant recommendation. Different users place different weights on those aspects, which are learned through our parameter estimation algorithm.
- We define a new user utility function by combining the individual utility derived from reviews and the collaborative utility derived from the rating matrix to better model user rating on a merchant.
- We devise a learning algorithm based on Gaussian distribution with two various priors to obtain the parameters in our utility function. The learned utility is served as a rating for merchant recommendation.
- We conduct extensive experiments to evaluate our approach in terms of effectiveness, efficiency, and cold-start using two real data sets.
- We implement a real mobile application to show the practicability of our model, which has an intuitive exhibition of user preference across multiple aspects.

The remainder of this paper is organized as follows. Section II provides an overview of the related work. Section III gives a formulation of the problem. Section IV details our model together with the learning and inference. Section V presents the experimental results and the demonstration. Finally, Section VI concludes and discusses the future work.

II. Related Work

In this section, we review a number of studies on recommendations, especially in POI recommendation and review-based recommender systems.

1) Traditional Recommender Systems: Traditional research on recommender systems mainly employ memory-based [10] and model-based collaborative filtering (CF) [8], [5] techniques on user-item ratings. In the past few years, contextual information like social or trust networks have been incorporated into various models to further improve recommendations [14], [13]. Our research is different from these traditional recommendation techniques since traditional ones rely greatly on the overall ratings while ignoring different aspects of the rating process.

2) POI Recommendation: Recently, owing to the prevalence of portable devices, techniques for supporting POI recommendation in LBSNs have attracted much research interest. Studies on POI recommendation have many applications, such as user behavior study [24], online retail store placement [7] and both event and activity recommendations [23], [25]. Conventional CF methods can be applied by treating a POI as an item [19]. In addition, a wide range of information available in LBSN has been extensively explored, such as geographical location [1], [11] , social connections [2], [12], temporal information [4], [22] and POI descriptions [21]. Our study has two distinctions : (1) we focus on user reviews which are not adequately studied in the previous work on LBSNs; (2) we address the rating prediction problem from an individual and collaborative view of utility.

3) Review-based Recommender Systems: Review is an important resource for understanding user opinions. Several studies explored this information to make predictions for ratings [18], [16]. Fan *etc.,* predict ratings for merchants from reviews through a regression method [3]. Huang in [6] extract subtopics from Yelp reviews to expose breakdown ratings for topics through topic models. Empirical experiments are conducted on sentiment rating prediction from reviews by incorporating author preference through the matrix approach [16]. The study most closely related to our work is [18] which analyses latent aspects from reviews via regression to predict overall ratings. Different to these methods, our approach focuses on aspect-based learning and the matching of user preference and merchant quality through review analysis.

III. Problem Formulation

In an LBSN, we only utilize the rating and review information, and leave other information out of consideration. All the notations are shown in Table I.

Formally, let $C = \{u_1, ..., u_N\}$ and $P = \{v_1, ..., v_M\}$ be the set of users and merchants, respectively. User u_i can give a rating, denoted as R_{ij} in a range of 1-5, to merchant v_j. In addition to rating a merchant, a user may write a review about his experience with the merchant, denoted as review $d_j^{(i)}$, which implicitly infers user preference and merchant quality. Specially, we can aggregate reviews written for merchant v_j, i.e., $\boldsymbol{d}_j = \{d_j^{(1)}, \cdots, d_j^{(i)}, \cdots, d_j^{(N)}\}$, to analyze merchant v_j's quality as commented by all users, which will be detailed in the following section on the model.

To better understand our model, we introduce the following definitions: 1) **Multiple Aspects**: Aspect is a high level concept derived from user reviews to explicitly show which factors may influence a user's rating, $A = \{A^1, ..., A^k, ..., A^{|A|}\}$, such as decor, service or food taste in the case of a restaurant. 2) **Aspect Representation**: Each aspect A^k can be represented as a collection of words, i.e., $A^k = \{w_1^k, w_2^k, \cdots, w_n^k\}$. Aspect words are selected from the review corpus. Words in the same aspect may be synonyms. For example, we have $decor = \{atmosphere, ambiance, feel, decor, ...\}$. 3) **User Preference**: For a given user u_i, user preference is defined as a vector of weights over all aspects, denoted as: $\boldsymbol{\beta}_i = [\beta_i^1, \cdots, \beta_i^k, \cdots, \beta_i^{|A|}]$. Likewise, every aspect of each merchant has a score to indicate how well the merchant performs in such aspect, which we termed merchant quality. 4) **Merchant Quality**: Merchant Quality z_j is a vector of scores, within which each score quantitatively measures the merchant's quality in such aspect. That is $\boldsymbol{z}_j = [z_j^1, ..., z_j^k, ..., z_j^{|A|}]$. More details on gaining $\boldsymbol{z}_j$ by utilizing aspect representation will be discussed in Section IV-A.

Symbol	Size	Description
C	N	all the users
P	M	all the merchants
R	$N \times M$	rating matrix
U	$N \times K$	user latent factor matrix
V	$M \times K$	merchant latent factor matrix
D	$N \times M$	review matrix
$d_j^{(i)}$	-	reviews on merchant j written by user i
$\mathcal{U}^i, \mathcal{U}^c$	$\mathbb{R}$	individual and collaborative utility
A	(.)	multiple aspects from user reviews
W_j	$n \times \lvert A \rvert$	aspect word matrix for merchant j
z_j	$1 \times \lvert A \rvert$	aspect score vector for merchant j
γ	$n \times \lvert A \rvert$	word sentiment polarity matrix
β_i	$1 \times \lvert A \rvert$	preference vector of user i
$\sigma, \sigma.$	$\mathbb{R}$	variance of the priors

Table I. NOTATIONS USED IN THIS PAPER

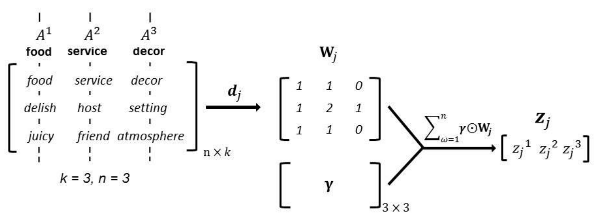


Figure 2. Aspect Representation and Merchant Quality Calculation

IV. MODEL SPECIFICATION

In this section, we introduce our learning model based on utility, followed by a complexity analysis of our algorithm.

A. Utility-based Model

In economics, discrete choice models are derived from a random utility model (RUM), where user behavior is under the assumption that maximizes the utility. As discussed previously, utility is a representation of satisfactions over a set of considerations. In [17], utility $\mathcal{U}$ is modeled as $\beta z^T + \varepsilon$, where z is a feature vector comprised by the user consideration, β is a parameter vector to weigh the utility of each considerations and ε is the error term without observations. In the context of merchant recommendations, we can easily treat rating as an equivalence of satisfaction (i.e., utility) - z is the aspect quality of merchants affecting user decision making and β is the user preferences over aspects. Thus, the overall rating given by a user to a merchant is decided by aggregating the aspect-wise inner products of β and z.

As aforementioned, we have individual and collaborative utility and thus obtain the overall utility as follows:

$$\mathcal{U}_{ij} = \mathcal{U}_{ij}^c + \mathcal{U}_{ij}^i + \varepsilon_{ij} \tag{1}$$

where $\mathcal{U}_{ij}^c$ is the collaborative utility (equivalent to $U_i V_j^T$) while $\mathcal{U}_{ij}^i$ is the individual utility (equivalent to $\boldsymbol{\beta}_i \boldsymbol{z}_j^T$). A tuning parameter should be added to control the contributions of the two utilities. However, due to the presence of the individual utility parameter $\boldsymbol{\beta}$ in the calculation, this tuning parameter can be integrated into $\boldsymbol{\beta}$, thus there is no need to involve a new parameter.

Defining meaningful aspects is the first step in our work, here we follow the boot-strapping aspect segmentation algorithm proposed in [18] to obtain A^k. Firstly, for each dataset we generally define a certain number of aspects[2], which are factors considered important by customers in that category. Aspects defined for our both datasets are different due to the reason that merchants categories are distinct between two datasets we have, where most merchants in Yelp are restaurants while TripAdvisor merchants are hotels. Though restaurant and hotel share some of the same aspects, we define six aspect for each of the datasets in order to better reveal users' consideration upon restaurant and hotel. Secondly, we manually select several seed words with high frequencies for each aspect. After segmenting the sentences by words, we count the frequency for each word in the entire review corpus and then sort these words according to the frequency descendingly. By removing the stop words, we manually select those words that satisfy two criteria: 1) The word has a high frequency, 2) The word unambiguously belongs to one of the aspects aforementioned. Afterward, in order to expand the words in each aspect, we adopt the aforementioned algorithm in [18] to find similar words and partition them into corresponding aspects, to form an aspect representation A^k. Thirdly, from $\boldsymbol{d}_j$ which is the review corpus in regarding to merchant j, we calculate $\mathbf{W}_j \in \mathbb{R}^{n \times k}$, a word frequency matrix with each column corresponding to an aspect, in terms of the aspect representation. Take Figure 2 as an example, let us suppose that $\boldsymbol{d}_j$ contains two reviews shown in Figure 1 and we only show three aspects and present corresponding representative words due to space limitation. Every element in $\mathbf{W}_j$ represents the word frequency in corpus $\boldsymbol{d}_j$ in terms of the aspect representation words shown in the figure. Last, we multiply the $\mathbf{W}_j$ with a parameter matrix $\boldsymbol{\gamma}$, which characterizes the sentiment polarities of all words and will be learned in the objective function optimization, to derive the merchant's quality, denoted as $\boldsymbol{z}_j = \sum_{w=1}^{n} \boldsymbol{\gamma} \odot \mathbf{W}_j$, where n is the length of the word vector in the aspect and $\odot$ is the Hadamard product.

Eventually, we thus reformulate Eqn. (1) as:

$$\mathcal{U}_{ij} = U_i V_j^T + \sum_{k=1}^{|A|} \boldsymbol{\beta}_i^k (\gamma^{:,k^T} W_j^{:,k})^T + \varepsilon_{ij} \tag{2}$$

Under the assumption that the error term follows Gaussian distribution, we present the likelihood for our utility-based model.

$$p(R|U, V, \boldsymbol{\beta}, \gamma, \sigma^2) = \prod_{i=1}^{N} \prod_{j=1}^{M} [\mathcal{N}(R_{ij}|U_i V_j^T + \boldsymbol{\beta}_i \boldsymbol{z}_j^T, \sigma^2)]^{I_{ij}} \tag{3}$$

B. Priors on Parameters

We could apply Maximum Likelihood Estimation (MLE) to estimate the parameters in each task. Nevertheless, parameters such as $U, V, \boldsymbol{\beta}, \gamma$ lack regularization, which may lead to poor prediction performance due to overfitting. Hence, we employ a Bayesian model by placing priors $p(\Omega|\Theta)$ on parameters $\Omega = \{U, V, \boldsymbol{\beta}\}$, where Θ denotes hyper-parameters. In order to simplify the discussion, we place a uniform distribution on parameter γ. Here, Gaussian prior and laplace prior are placed for regularization.

[2]Aspects and its representation words can be downloaded from url: https://www.dropbox.com/s/05y5mq25gv7rutb/Aspects.txt?dl=0

C. Optimization Problem

The posterior can be obtained according to Bayes theorem as follows:

$$p(\Omega|R, D, \Theta) \propto p(R, D|\Omega)p(\Omega|\Theta) \tag{4}$$

Maximizing the Eqn. (4) is to learn the parameter Ω, which is equivalent to minimizing its negative log form, i.e., minimizing the log-likelihood objective function.

$$log\mathcal{L} = \sum_{i,j} I_{ij}(R_{ij} - U_i V_j^T - \boldsymbol{\beta}_i \boldsymbol{z}_j^T)^2 + Reg(\Omega) \tag{5}$$

For Gaussian prior, the regularizer term corresponds to the L2-norm regularization ($Reg_N(\Omega) = \frac{\lambda}{2}\|\Omega\|_F^2$) while it is L1-norm regularization for Laplace prior ($Reg_L(\Omega) = \lambda\|\Omega\|_1$), where λ is the parameter to be tuned by cross-validation through experiments.

Compared with L2-norm, L1-norm regularizer induces sparse values, i.e., most values of Ω are zero. This is closer to the realistic choice process since only a few aspects matter for a given user. Hence, the prediction with sparser Ω may better describe the fact.

A local minimum of the objective function given by Eqn. (5) can be achieved by performing gradient descent.

$$\frac{\partial log\mathcal{L}}{\partial \gamma^k} = \sum_{i,j}[(U_i V_j^T + \boldsymbol{\beta}_i \boldsymbol{z}_j^T) - R_{ij}]\boldsymbol{\beta}_i^k W_j^{:,k} \tag{6}$$

$$\frac{\partial log\mathcal{L}}{\partial U_i} = \sum_{j}[(U_i V_j^T + \boldsymbol{\beta}_i \boldsymbol{z}_j^T) - R_{ij}]V_j + Reg(U_i) \tag{7}$$

$$\frac{\partial log\mathcal{L}}{\partial V_j} = \sum_{i}[(U_i V_j^T + \boldsymbol{\beta}_i \boldsymbol{z}_j^T) - R_{ij}]U_i + Reg(V_j) \tag{8}$$

$$\frac{\partial log\mathcal{L}}{\partial \boldsymbol{\beta}_i} = \sum_{j}[(U_i V_j^T + \boldsymbol{\beta}_i \boldsymbol{z}_j^T) - R_{ij}]\boldsymbol{z}_j + Reg(\boldsymbol{\beta}_i) \tag{9}$$

where γ^k is a word sentiment orientation vector for k^{th} aspect.

D. Complexity Analysis

The overall complexity of our models is relevant to the computation of the objective function and its gradient descent.

According to Eqn. (3), the indicator matrix I_{ij} shows that only the observed values are involved in the model. With each observed value being decomposed to F dimensions of latent factors, the integrated computation for the first part is $\mathcal{O}(|I_{ij}| \cdot F)$. On the other hand, the review preprocessing only costs a fixed time and the word vector for each aspect is a constant, and the computation for individual utility is $\mathcal{O}(|I_{ij}| \cdot |A|)$. The overall computation of the objective function is $\mathcal{O}(|I_{ij}| \cdot (F + |A|))$. In summary, the algorithm is computationally efficient and linear in the size of the dataset.

E. Recommendation

Merchant recommendations can be performed by selecting merchants according to their gained utilities. The algorithm of proposed utility-based merchant recommendation is detailed in Algorithm 1. Particularly L1-6 calculate utility through iteratively optimizing Eqn.(5), e.g., L3 corresponds to Eqn.(6)-(9). L7-9 select merchants with higher utilities for recommendation.

Algorithm 1: merchant recommendations based on utility

input : Reviews D containing W, Rating Matrix R
output: $\Omega = \{U, V, \boldsymbol{\beta}, \gamma, \theta\}$
1 Random Ω;
2 **for** *step = 1 to MAX STEP* **do**
3     $\Omega_l \leftarrow \Omega_l - \eta\nabla_{\Omega_l}$;
4   **if** *converge* **then**
5   return $utility = UV^T + \sum \boldsymbol{\beta}\gamma W$;
6     $\eta \leftarrow \frac{\eta}{1+step/MAXSTEP}$;
7 **for** *each user* **do**
8   sort $utility$ of all merchants;
9   recommend top merchants with highest utilities;

V. Experiment Results

In this section, we compare our approach with up-to-date methods in terms of effectiveness and efficiency.

A. Dataset & Setting

The experiments are based on the Yelp Dataset[3] and TripAdvisor Dataset[4]. Yelp is a well-known website for business ratings and reviews and TripAdvisor is a travel website providing reviews of travel-related content, thus making them ideal sources for our experiments. We adopt a subset of the whole published data. For Yelp, we choose the restaurant category because it covers more than 2/3 of the total reviews. The final dataset consists of 4074 business merchants along with 5881 users making 218484 ratings and reviews in total (after filtering out users and merchants with less than 20 records). For TripAdvisor, only the hotel category is collected. Similar to Yelp, we filter out users and hotels with less than 10 records and finally have 2510 hotels and 7106 users who have made 81046 reviews. Our experiments are conducted on a personal computer equipped with an Intel Core2 Duo CPU(2.67GHz) and 2G memory.

In our experiments, we particularly choose six aspects[5] along with their representation words for each dataset. The iteration step is set to 0.0001 and regularization parameter λ is 0.05 through cross validation. We investigate the impact of dimension selection the dimension is fixed to 10.

B. Metrics

Two widely used metrics, i.e., Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), are utilized to measure the performance. The definitions are as follows:

$$MAE = \frac{1}{|T|}\sum_{i,j}|R_{i,j} - \widehat{R}_{i,j}|$$

$$RMSE = \sqrt{\frac{1}{|T|}\sum_{i,j}(R_{i,j} - \widehat{R}_{i,j})^2}$$

where $R_{i,j}$ and $\hat{R}_{i,j}$ denote the observed and predicted rating, and T is the testing set. Apparently the smaller both metrics are, the better the approaches perform.

[3] http://www.yelp.com.au/dataset_challenge

[4] http://sifaka.cs.uiuc.edu/wang296/Data/index.html

[5] Aspects and its representation can be downloaded from url: https://www.dropbox.com/s/05y5mq25gv7rutb/Aspects.txt?dl=0

Data Set	Training Data	Metrics	ES	IAP	LRR	Dimensionality = 10 PMF	GM-L1	GM-L2
Yelp	80%	MAE improved	0.82758 10.87%	0.78920 6.54%	0.76032 2.99%	0.75729 2.60%	**0.73758**	0.74002
		RMSE improved	1.03427 9.65%	1.00212 6.76%	0.99983 6.54%	0.95078 1.72%	**0.93442**	0.94039
	60%	MAE improved	0.82831 10.04%	0.79019 5.70%	0.77790 4.22%	0.76081 2.06%	0.74865	**0.74511**
		RMSE improved	1.03391 8.74%	1.00406 6.03%	0.99545 5.22%	0.97598 3.33%	**0.94350**	0.94897
TripAdvisor	80%	MAE improved	0.79595 9.07%	0.77562 6.69%	0.75413 4.03%	0.73759 1.88%	0.72438	**0.72376**
		RMSE improved	1.04198 11.74%	1.00701 8.68%	0.95992 4.20%	0.94081 2.25%	0.92746	**0.91961**
	60%	MAE improved	0.79859 6.54%	0.78777 5.26%	0.77472 3.66%	0.76969 3.03%	0.75567	**0.74637**
		RMSE improved	1.04949 8.08%	1.00131 3.66%	0.99964 3.50%	0.98255 1.82%	0.97710	**0.96465**

Table II. PREDICTION ACCURACY COMPARISON

C. Baselines

We compare the following methods with our approach:

1) **PMF**: This method is proposed in [15] by factorizing the user-merchant matrix, which is the only input used for the algorithm.
2) **LRR**: **L**atent aspect **R**ating **R**egression is proposed by [18] to make predictions to ratings via analyzing reviews.
3) **IAP**: **I**ncorporating **A**uthor **P**reference for rating prediction is presented in [16].
4) **ES**: This method won the 2014 Yelp challenge [6]. We averagely aggregate **E**xtracted **S**ubtopic ratings to predict the overall rating.
5) **GM-L1 and GM-L2**: These are our utility-based Gaussian models with L1 and L2 regularization.

We use different amounts of training data to test all algorithms and the results are tabulated in Table II. The selection process is conducted three times and the average results are reported to avoid bias.

From Table II, we can observe that of the compared methods, our approach generally achieves the best performance, with improvement by up to 11% and 10% in MAE and RMSE. This is mainly because some baselines only utilize the rating resource for model training, resulting in vulnerability to sparsity and cold-start. Those which use both resources estimate the rating by merely using a simple regression method without jointly considering individual and collaborative behaviour. In constrast our approach combines both and the improvements on the two metrics verify our claims.

Different proportion of the training data affect the result, which is evident to all and is also reported in the previous research papers. Higher percentage of training data may deduce more precise models thus resulting in accurate predictions, i.e., small errors.

D. Cold-Start Problem

In order to compare the robustness of our approach in the cold-start problem with baseline methods, we first partition the users into groups based on the sparsity ratio of observed ratings in the training data set, i.e., 1-10, 10-20, 20-40, 40-80, 80-160, >160, and then calculate the MAE and RMSE of the testing data set with different groups respectively. The entire cold-start experiment indicates that our approach outperforms other methods consistently, especially for the sparser groups. The result of TripAdvisor is consistent with that of Yelp, thus we report the results using 70% and 90% of Yelp data for training, which is shown in Figure 3.

Figure 3(a) and 3(d) describes the distribution of the observed ratings for groups in testing dataset in regarding to 70% and 90% training dataset in use respectively. Figure 3(b)-3(c) and 3(e)-3(f) report the MAE and RMSE results for different groups.

In comparison with PMF, especially when user ratings are very sparse, the reviews in our approach become a crucial auxiliary resource to help to capture the user's preferences thus providing good results, which coincides with our aforementioned assumption.

E. Mobile Application Demonstration

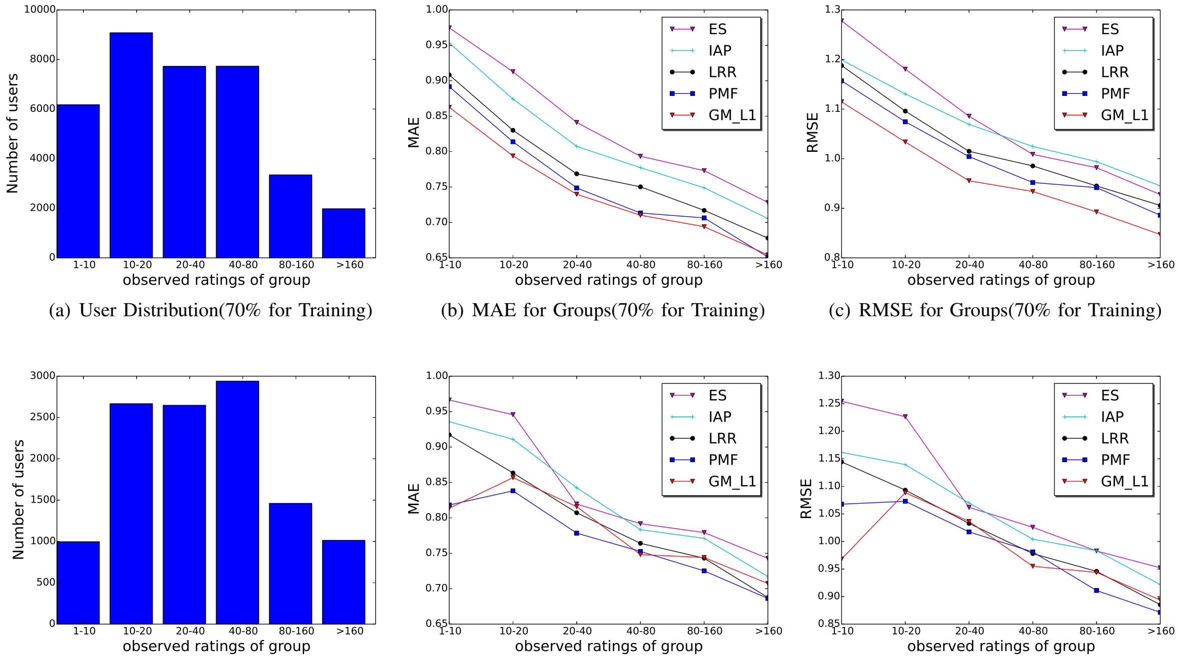
We build a real mobile recommender system based on our utility-based model, which can be detailed in [9]. It can not only present visualization for users and merchants via analyzing reviews but also do personalized recommendations according to users' preferences.

VI. CONCLUSION

In this paper, we address merchant recommendation by jointly exploring the ratings and reviews on LBSNs. Different from conventional methods, we aim to model the user rating decision process on merchants by leveraging individual and collaborative utility. A novel and efficient utility-based model is proposed. Experiments on two real world datasets have demonstrated that our approach outperforms other methods, especially in the cold-start problem. Meanwhile, a real mobile application is implemented to show the practicability of our model. In future we will examine the impact of geographic information and further improve the prediction accuracy.

VII. ACKNOWLEDGEMENT

This research was partially supported by grants from the National Science Foundation for Distinguished Young Scholars of China (Grant No. 61325010), the National High Technology Research and Development Program of China (Grant No. 2014AA015203), and the Fundamental Research Funds for the Central Universities of China (Grant No. WK2350000001).



(a) User Distribution(70% for Training) (b) MAE for Groups(70% for Training) (c) RMSE for Groups(70% for Training)

(d) User Distribution(90% for Training) (e) MAE for Groups(90% for Training) (f) RMSE for Groups(90% for Training)

Figure 3. Cold Start Problem Result

REFERENCES

[1] 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, pp. 17–23.

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

[3] M. Fan and M. Khademi, "Predicting a business star in yelp from its reviews text alone," *CoRR*, vol. abs/1401.0864, 2014.

[4] H. Gao, J. Tang, X. Hu, and H. Liu, "Exploring temporal effects for location recommendation on location-based social networks," in *RecSys*, 2013, pp. 93–100.

[5] Y. Hu, Y. Koren, and C. Volinsky, "Collaborative filtering for implicit feedback datasets," in *ICDM*, 2008, pp. 263–272.

[6] S. R. James Huang and E. Joo, "Improving restaurants by extracting subtopics from yelp reviews," *iConference, Berlin*, 2014.

[7] D. Karamshuk, A. Noulas, S. Scellato, V. Nicosia, and C. Mascolo, "Geo-spotting: mining online location-based services for optimal retail store placement," in *KDD*, 2013, pp. 793–801.

[8] Y. Koren, R. M. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *IEEE Computer*, vol. 42, no. 8, pp. 30–37, 2009.

[9] X. Li, G. Xu, E. Chen, and L. Li, "Mars: A multi-aspect recommender system for point-of-interest," in *ICDE*, 2015.

[10] G. Linden, B. Smith, and J. York, "Industry report: Amazon.com recommendations: Item-to-item collaborative filtering," *IEEE Distributed Systems Online*, vol. 4, no. 1, pp. 76–80, 2003.

[11] B. Liu, Y. Fu, Z. Yao, and H. Xiong, "Learning geographical preferences for point-of-interest recommendation," in *KDD*, 2013, pp. 1043–1051.

[12] B. Liu and H. Xiong, "Point-of-interest recommendation in location based social networks with topic and location awareness," in *SDM*, 2013, pp. 396–404.

[13] H. Ma, I. King, and M. R. Lyu, "Learning to recommend with social trust ensemble," in *SIGIR*, 2009, pp. 203–210.

[14] H. Ma, T. C. Zhou, M. R. Lyu, and I. King, "Improving recommender systems by incorporating social contextual information," *ACM Trans. Inf. Syst.*, vol. 29, no. 2, pp. 9:1–23, 2011.

[15] A. Mnih and R. Salakhutdinov, "Probabilistic matrix factorization," in *Advances in neural information processing systems*, 2007, pp. 1257–1264.

[16] S. Mukherjee, G. Basu, and S. Joshi, "Incorporating author preference in sentiment rating prediction of reviews," in *WWW*, 2013, pp. 47–48.

[17] J. Neumann and O. Morgenstern, *Theory of games and economic behavior*, commemorative edition ed. Princeton Univ. Press, 2007.

[18] H. Wang, Y. Lu, and C. Zhai, "Latent aspect rating analysis on review text data: A rating regression approach," in *KDD*, 2010, pp. 783–792.

[19] M. Ye, P. Yin, and W.-C. Lee, "Location recommendation for location-based social networks," in *GIS*, 2010, pp. 458–461.

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

[21] H. Yin, Y. Sun, B. Cui, Z. Hu, and L. Chen, "Lcars: a location-content-aware recommender system," in *KDD*, 2013, pp. 221–229.

[22] Q. Yuan, G. Cong, Z. Ma, A. Sun, and N. Magnenat-Thalmann, "Time-aware point-of-interest recommendation," in *SIGIR*, 2013, pp. 363–372.

[23] W. Zhang, J. Wang, and W. Feng, "Combining latent factor model with location features for event-based group recommendation," in *KDD*, 2013, pp. 910–918.

[24] V. W. Zheng, B. Cao, Y. Zheng, X. Xie, and Q. Yang, "Collaborative filtering meets mobile recommendation: A user-centered approach," in *AAAI*, 2010, pp. 236–241.

[25] V. W. Zheng, Y. Zheng, X. Xie, and Q. Yang, "Collaborative location and activity recommendations with gps history data," in *WWW*, 2010, pp. 1029–1038.