A Novel Tourist Attraction Recommendation System Based on Improved Visual Bayesian Personalized Ranking

Yi Liang, Nan Chen

Shijiazhuang University of Applied Technology, Shijiazhuang 050081, China

Corresponding Author Email: mzcnn@126.com

https://doi.org/10.18280/isi.250413

Received: 18 March 2020
Accepted: 30 June 2020

Keywords: recommendation system, Bayesian personalized ranking (BPR), stratified sampling, tourist attractions

ABSTRACT

Statistics show that most tourists log into the main tourism websites to view user reviews or scores before selecting their destinations. However, the existing tourist destination recommendation models neither consider the implicit user preferences nor mine the potential semantics of tourist attractions. To solve the problems, this paper predicts user scores of tourist attractions through stratified sampling, and optimizes the predicted scores with Bayesian personalized ranking (BPR) and improved visual BPR (VBPR). Then, the recommendation system was optimized by the improved VBPR, which decomposes the prediction score matrix and considers visual features. Experimental results fully demonstrate the excellence of the proposed tourist attraction recommendation system. The research findings provide a good reference for online travel agencies to recommend tourist attractions.

1. INTRODUCTION

With the development of Internet technology, the amount of information on the Internet is exploding, making the problem of information overload increasingly prominent. It is difficult for a user without a clear demand to obtain the information of interest from a large amount of information. Meanwhile, lots of valuable information are submerged in the sea of data, becoming invisible to potential users. Faced with the massive online data, the traditional search algorithm, capable of filtering information for users, cannot provide personalized services that meet the interests and preferences of each user. To solve the problem, recommendation system [1-2] has emerged as the bridge between users and Internet information. On the one hand, the system enables users to find interesting information from massive data; on the other hand, it could transfer valuable information to potential users. Relevant surveys have shown that more than 75% of tourists log into the main tourism websites to view user reviews or scores before selecting their destinations and travel routes. However, many of them have difficulty in finding valuable content out of the huge amount of information on these websites. Therefore, the tourism websites need to further improve their recommendation systems, in order to satisfy the growing demand for personalized tourism.

In this paper, the latent semantic space model [3] is combined with visual Bayesian personalized ranking (VBPR) [4] into a novel VBPR recommendation model, which mitigates the problem of data sparsity and improves the interactive experience of users.

2. LITERATURE REVIEW

The research into recommendation system started in the early 1990s. The earliest recommendation systems could only suggest the products that interest users, using a recommendation algorithm. With the proliferation of such systems, the recommendation problem has gradually evolved into the score prediction of the recommended object. Many scholars at home and abroad have studied recommendation systems. For instance, Goldberg et al. [5] innovatively introduced collaborative filtering to Tapestry system. Pang et al. [6] proposed a score-based collaborative filtering recommendation model, which derives user preferences from their scores, and analyzes user similarity through clustering to complete the recommendation. Combining user score with visual information, Huang et al. [7] solved the recommendation problem with matrix factorization (MF) model. Based on convolutional neural network (CNN), Tsai et al. [8] built a recommendation system coupling user score with video. Drawing on deep structured semantic models (DSSMs), Yoon et al. [9] constructed a location-aware personalized news recommendation system. Zhao et al. [10] recommended commodities with a high accuracy, using a self-designed Bayesian personalized ranking (BPR) model. Pan et al. [11] applied the BPR model to hotel recommendation. He et al. [12] introduced visual information to commodity recommendation, and extended the BPR model into the VBPR model. Based on attention mechanism, Han et al. [13] proposed a CNN model for the recommendation of Weibo information.

In recent years, the tourist attractions recommendation system has become a research hotspot. For example, Yuan et al. [14] designed a case-based tourist attractions recommendation system, and developed a web-based intelligent recommendation framework for travel agencies, which integrates reasoning with multi-criteria decision-making technology; the recommendation quality of their system was verified through experiments. Based on collaborative filtering, Kirn et al. [15] set up a decision support system for tourist attractions, which predicts user preference for tourist attractions by Bayesian model, and demonstrated...
the prediction accuracy by the receiver operating characteristic (ROC) curve. Hsu et al. [16] designed a tourist attractions recommendation system based on multi-criteria collaborative filtering; the system relies on multi-objective collaborative filtering to process more information on user preferences, thereby meeting user demand more effectively.

Despite the above breakthroughs, the recommendation systems for tourist attractions still face multiple challenges: (1) the data on user scores are extremely sparse, causing a prominent problem of data sparsity; (2) the tourist attractions are recommended solely based on the historical data of users, without considering the latent information of user preferences; (3) the potential semantic information of tourist attraction images or users is not mined or analyzed from the perspective of multimodality.

3. A NOVEL TOURIST ATTRACTIONS RECOMMENDATION SYSTEM BASED ON BPR MODEL

The literature review indicates that the previous recommendation systems have not considered the role of tourist attraction images across heterogeneous media. To make up for the gap, this section mines the latent semantic space of users and tourist attractions with MF model, and optimizes the space in BPR or VBPR model to generate the prediction score matrix.

3.1 Tourist attraction recommendation system based on stratified sampling and BPR model

To design the tourist attraction recommendation system, the user preferences were captured through stratified sampling, and the latent semantics of users and tourist attractions were mined by the BPR model. To begin with, the data on tourist preferences were collected through a questionnaire survey, and subjected to stratified sampling. Under the preset collection rules, the data on user scores of tourist attractions were acquired automatically from Ctrip.com, and preprocessed. Next, a user score matrix \( R \) was generated for tourist attractions. Based on the BPR model, MF model, and matrix \( R \), a prediction score matrix was established by predicting user scores. Then, the recommendation \( R_A \) from the BPR model was combined with that \( R_U \) obtained through stratified sampling into the mixed recommendation \( R_A + R_U \).

3.1.1 Stratified sampling model

In the stratified sampling model, the overall unit was divided proportionally into \( U \) independent layers \( (H_1, H_2, \ldots, H_U) \). Then, sampling was performed layer by layer. The results of all layers were added up to obtain the overall distribution of samples. The specific workflow is as follows:

Step 1. Describe the differences in user preferences with random target variables, namely, travel time, interest category, and travel mode.

Step 2. Based on the influencing factors, stratify the overall unit into \( U \) layers, each of which has \( H_i \) individuals: layer \( i \) \((i = 1, 2, \ldots, U)\). Hence, the overall distribution of samples \( H \) can be calculated by:

\[
H = \sum_{i=1}^{U} H_i
\]  

Step 3. Determine the sampling number of each layer. Let \( N \) be the total number of samples. Then, the sampling number of layer \( i \) can be calculated by:

\[
i = N \times \frac{H_i}{H}
\]

To reduce intra-layer difference and magnify inter-layer difference, the samples were classified through stratified sampling, in the light of the feature distribution of the overall unit. After that, a certain number of samples were extracted from each layer to describe the distribution of that layer, forming the sample population.

Because of the reasonable stratification of samples, the stratified sampling model could capture user preferences for tourist destinations in an accurate manner, laying a solid basis for producing a suitable recommendation list.

The stratified sampling results were weighted through analytic hierarchy process (AHP), a subjective weighting method, aiming to adjust the weight of each user attribute. Specifically, the relative importance between attributes on the same layer was compared to a new discriminant matrix, which was used to determine the weight of each attribute.

In this paper, six user attributes are selected, including gender \( G_1 \), region \( G_2 \), age \( G_3 \), educational background \( G_4 \), job type \( G_5 \), and monthly income \( G_6 \). The relative importance between two attributes was measured against a seven-point scale: strongly important \( =6 \), moderately important \( =4 \), slightly important \( =2 \), equally important \( =1 \), slightly unimportant \( =1/2 \), moderately unimportant \( =1/4 \), and strongly unimportant \( =1/6 \). On this basis, the discriminant matrix \( G \) can be constructed as:

\[
G = \begin{bmatrix}
G_1 & G_2 & G_3 & G_4 & G_5 & G_6 \\
1 & 1 & 1 & 1 & 1 & 1 \\
2 & 4 & 4 & 2 & 6 & 6 \\
2 & 1 & 2 & 2 & 4 & 2 \\
4 & 2 & 1 & 1 & 1 & 2 \\
4 & 2 & 1 & 1 & 2 & 4 \\
4 & 4 & 2 & 2 & 4 & 2 \\
4 & 4 & 2 & 2 & 1 & 1 \\
\end{bmatrix}
\]

Based on matrix \( G \), the weight \((i = 1, \ldots, 6)\) of the \( i \)-th attribute was calculated to generate user preference:

\[
W_i = \prod_{j=1}^{6} G_{ij} / \sum_{i=1}^{6} \prod_{j=1}^{6} G_{ij}
\]

3.1.2 Stratified sampling results

The survey data on user preferences for tourist attractions were analyzed through stratified sampling, from the perspective of various user attributes. Figures 1 and 2 present the stratified sampling histograms of the preferences of 1,000 users.

Three conclusions could be drawn from Figures 1 and 2: Most users prefer to travel in spring and autumn, when the temperature is comfortable and the scenery is pleasant to the eyes; The users under 20, most of whom are students, prefer to travel in summer, as they have plenty of free time during summer vacation; Males prefer to travel in spring, while females prefer to travel in autumn.
The travel time difference among users with different genders, regions, and ages

Figure 1.

The travel time difference among users with different educational background, job types, and monthly incomes

Figure 2.

3.1.3 BPR model

The traditional MF model to predict user scores of tourist destinations can be defined as:

\[ \hat{s}_{u,i} = \beta + \delta_u + \delta_i + \theta_u^T \sigma \]

where, \( \beta \) is the global offset; \( \delta_u \) and \( \delta_i \) are the offsets of user \( u \) and tourist attractions \( i \) respectively; \( \theta_u \) and \( \sigma \) are \( k \)-dimensional vectors of the latent semantic spaces of users and tourist attractions, respectively. The adaptability of user \( u \) and tourist attractions \( i \) can be described by the inner product \( s_u^T \sigma \). The user preference of a tourist attraction is negatively correlated with the adaptability.

Then, the MF model was optimized in the BPR model, which is an optimization framework for pairwise sorting through stochastic gradient descent (SGD). Pairwise sorting has a much better optimization effect than single sample.

Next, a training set \( D_S \) of triplets \( (u, i, j) \) was established, where \( u \) is the user, \( i \) is a tourist attraction viewed positively by the user, and \( j \) is a tourist attraction viewed negatively by the user:

\[ D_S = \{ (u, i, j) | u \in U \land i \in I_u^+ \land j \in I \setminus I_u^+ \} \]

Let \( P_m \) be the parameter of the BPR model, and \( \hat{s}_{u,i,j}(P_m) \) be the relationship between triplets \((u, i, j) \). Then, the optimization process of the BPR model can be depicted as:

\[ \sum_{(u,i,j) \in D_S} \ln \sigma(\hat{s}_{u,i,j}) - \epsilon \| P_m \|^2 \]

where, \( \sigma \) is a logical function (e.g. sigmoid function); \( \epsilon \) is a regularized hyperparameter. For MF-based prediction, \( \hat{s}_{u,i,j} \) can be defined as:

\[ \hat{s}_{u,i,j} = \hat{s}_{u,i} - \hat{s}_{u,j} \]

After randomly sampling \((u,i,j) \) from \( D_S \), the relevant parameters can be obtained by the BPR model, coupled with the SGD:

\[ P_m \leftarrow P_m + \eta \cdot (\sigma(-\hat{s}_{u,i,j}) \frac{\partial \hat{s}_{u,i,j}}{\partial P_m} - \epsilon \xi(P_m)) \]

where, \( \eta \) is the learning rate.

3.2 Experimental verification

To verify the tourist attractions recommendation system based on BPR model, the user preferences for tourist destinations were obtained through stratified sampling, and weighted subjectively. The obtained data indicate that the user preferences vary greatly with user attributes.

By formula (2), the results of stratified sampling were weighted to establish a stratified sampling model. By the weight, the user attributes could be ranked in descending order as monthly income =0.3702, age =0.1948, educational background =0.1372, region =0.1360, job type =0.1081, and gender =0.0520. Obviously, monthly income and age have the greatest impact on traveling, highlighting the importance of economic factors. On the contrary, gender and job type have a relatively small impact on traveling. This is consistent with our objective cognition.

Then, the recommendation \( R_H \) from stratified sampling model and that \( R_A \) from the BPR model were synthesized into a mixed recommendation \( R_{Mix}(R_A + R_H) \). Then, the precision of the mixed recommendation (MR) was compared with that of stratified sampling (SS) model and BPR model, as well as that of traditional models, including item-based collaborative filtering (IBCF), user-based collaborative filtering (UBCF), location-based collaborative filtering (LBCF), horizontal collaborative filtering (HC), and fuzzy c-means clustering (FCM) (Table 1).

As shown in Table 1, the FCM had slightly superior precision than most traditional models, because the clustering-based prediction is better guided by the prior knowledge in the eight types of tourist attractions.

Besides the FCM, the SS model achieved a high precision. This means the user preferences obtained through questionnaire survey are quite accurate, eliminating the need for any machine learning (ML). Thus, it is very important to explore user preferences for tourist attractions.

The BPR model boasted much better precision than the contrastive models, thanks to the mining of the latent semantic spaces of users and tourist attractions. Based on existing scores, these spaces accurately illustrate user preferences and popularities of tourist attractions.

The MR was even more accurate than the result of the BPR model. The excellence is achieved through pairwise sorting of the BPR result.
Tables 1. The comparison of precisions among different recommendation models

<table>
<thead>
<tr>
<th>No.</th>
<th>ICF</th>
<th>UBCF</th>
<th>LBCF</th>
<th>HC</th>
<th>FCM</th>
<th>BPR</th>
<th>SS</th>
<th>MR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.893</td>
<td>9.985</td>
<td>4.221</td>
<td>4.372</td>
<td>4.516</td>
<td>37.721</td>
<td>5.721</td>
<td>38.281</td>
</tr>
<tr>
<td>3</td>
<td>2.588</td>
<td>7.603</td>
<td>6.355</td>
<td>7.086</td>
<td>7.238</td>
<td>39.841</td>
<td>7.816</td>
<td>40.182</td>
</tr>
<tr>
<td>5</td>
<td>2.672</td>
<td>6.336</td>
<td>6.482</td>
<td>7.003</td>
<td>7.672</td>
<td>40.051</td>
<td>8.831</td>
<td>40.982</td>
</tr>
<tr>
<td>6</td>
<td>2.816</td>
<td>6.606</td>
<td>6.892</td>
<td>7.277</td>
<td>7.309</td>
<td>40.093</td>
<td>8.775</td>
<td>42.183</td>
</tr>
<tr>
<td>7</td>
<td>2.933</td>
<td>5.762</td>
<td>7.119</td>
<td>7.537</td>
<td>7.832</td>
<td>40.580</td>
<td>9.021</td>
<td>42.985</td>
</tr>
<tr>
<td>8</td>
<td>2.971</td>
<td>5.492</td>
<td>7.325</td>
<td>7.688</td>
<td>8.062</td>
<td>41.287</td>
<td>9.027</td>
<td>43.187</td>
</tr>
<tr>
<td>9</td>
<td>2.763</td>
<td>5.504</td>
<td>7.332</td>
<td>7.642</td>
<td>7.882</td>
<td>42.487</td>
<td>8.578</td>
<td>42.167</td>
</tr>
<tr>
<td>10</td>
<td>2.705</td>
<td>4.998</td>
<td>7.233</td>
<td>7.683</td>
<td>8.125</td>
<td>41.872</td>
<td>9.082</td>
<td>43.982</td>
</tr>
</tbody>
</table>

![Figure 3. The precision advantage of BPR model](image)

![Figure 4. The precision advantage of the MR](image)

4. A NOVEL TOURIST ATTRACTIONS RECOMMENDATION SYSTEM BASED ON IMPROVED VBPR MODEL

The BPR model can outperform most traditional recommendation models. However, the image information of tourist attractions is not considered in that model, resulting in serious data sparsity. To overcome this problem, this section improves the VBRP model by innovatively introducing visual features to tourist attraction recommendation.

Under the preset collection rules, the data on user scores of tourist attractions were acquired automatically from Ctrip.com, and preprocessed. Next, a user score matrix $R$ was generated for tourist attractions. In the meantime, the images on the tourist attractions were downloaded from Ctrip.com automatically. Then, the visual features (e.g. color and texture) were extracted from these images, and used to improve the VBP model. Based on the MF model, matrix $R$ and visual features, the recommendation $R_B$ from the improved VBP model was combined with the $R_H$ from the SS model into a mixed recommendation $R_B + R_H$. Owing to the improvement, the $R_B + R_H$ reflects the merits of multimodal analysis, supplementing the result of $R_A + R_H$.

4.1 Improved VBPR model

In theory, the traditional MF model can find the potential features (latent semantics) of users or tourist attractions in the relevant dimensions. Nevertheless, the score data acquired by the MF model are too sparse to generate a robust recommendation. This problem can be mitigated by adding the auxiliary information of visual features. Hence, the MF of the VBPR model was improved as:

$$
\hat{s}_{u,i} = \beta + \delta_u + \delta_i + \theta_u^T I_i + I_u^T I_i + \epsilon^1 I_i
$$

where, $I_u$ and $I_i$ are newly introduced $D$-dimensional visual features; $I_u^T I_i$ is the visual interaction between user $u$ and tourist attraction $i$.

Let $\epsilon^1 I_i$ be a user’s overall evaluation of the visual features of tourist attraction $i$. Introducing a visual offset term $\epsilon$, the MF model can be finalized as:

$$
\hat{s}_{u,i} = \beta + \delta_u + \delta_i + \theta_u^T I_i + I_u^T I_i + \epsilon^1 I_i + \epsilon^1 I_i
$$

The VBP framework was adopted to optimize the final MF model. The SGD method was adopted by the VBP model to update the relevant parameters:

$$
I_u \leftarrow I_u + \eta \cdot (\sigma(-\hat{s}_{u,i}) - \hat{s}_{u,i})
$$

$$
\epsilon \leftarrow \epsilon + \eta \cdot \sigma(-\hat{s}_{u,i}) - \hat{s}_{u,i}
$$

To sum up, the VBPR model was optimized from multiple visual angles, using VGG features and related visual features like shape, color, and texture.
4.2 Experimental verification

To verify its effectiveness, the improved VBPR model was evaluated on the Wisdom Tourism dataset [17] by metrics like root mean square error (RMSE) [18], mean absolute error (MAE) [19], precision, and recall. Different visual features were selected as the bases of recommendation, including scale-invariant feature transform (SIFT), GIST descriptor, hue saturation value (HSV), red green blue (RGB), local binary pattern (LBP), and VGG. The RMSE and MAE of our model based on different visual features are presented in Figure 5.

As shown in Figure 5, the RMSE and MAE of improved VBPR model varied with visual features. The highest RMSE and MAE were observed on RGB features, because RGB features cannot accurately reflect our objective cognition of colors (mainly from the angles of hue, saturation, and brightness). By contrast, the lowest RMSE and MAE were achieved on HSV features, indicating that the inclusion of HSV features can improve the recommendation effect to a certain extent.

Furthermore, the improved VBPR model was compared with IBCF, UBCF, LBCF, HC, FCM, nonnegative matrix factorization (NMF), k-th nearest neighbor (KNN), and BPR in terms of RMSE and MAE (Figure 6).

As shown in Figure 6, the contrastive models like FCM, NMF, and KNN achieved good performance, but the improved VBPR with HSV features was more outstanding. The improved VBPR outperformed the BPR model, for the recommended objects were well depicted by the visual features of tourist attraction images. Overall, the improved VBPR model had better RMSE than any other model.

In summary, the recommendation effect could be promoted by the improved VBPR model with HSV features, which helps to predict user scores of recommended tourist attractions more in line with user preferences.

In addition, the recommendation $R_{B}$ from the SS model was synthesized with the recommendation $R_{P}$ from the improved VBPR model into a mixed recommendation $R_{B} + R_{P}$ based on visual features (VMR). The precision of the VMR was compared with that of many other models (Table 2).

Table 2. The comparison of precisions among different recommendation models

<table>
<thead>
<tr>
<th>N</th>
<th>IBCF</th>
<th>UBCF</th>
<th>LBCF</th>
<th>HC</th>
<th>FCM</th>
<th>BPR</th>
<th>SS</th>
<th>VMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.893</td>
<td>9.985</td>
<td>4.221</td>
<td>4.372</td>
<td>4.516</td>
<td>37.721</td>
<td>5.721</td>
<td>50.622</td>
</tr>
<tr>
<td>2</td>
<td>1.821</td>
<td>8.082</td>
<td>6.381</td>
<td>6.733</td>
<td>6.971</td>
<td>37.852</td>
<td>8.082</td>
<td>52.391</td>
</tr>
<tr>
<td>3</td>
<td>2.588</td>
<td>7.603</td>
<td>6.355</td>
<td>7.086</td>
<td>7.238</td>
<td>39.841</td>
<td>7.816</td>
<td>51.588</td>
</tr>
<tr>
<td>5</td>
<td>2.672</td>
<td>6.336</td>
<td>6.482</td>
<td>7.003</td>
<td>7.672</td>
<td>40.051</td>
<td>8.831</td>
<td>53.729</td>
</tr>
<tr>
<td>7</td>
<td>2.933</td>
<td>5.762</td>
<td>7.119</td>
<td>7.537</td>
<td>7.832</td>
<td>40.580</td>
<td>9.021</td>
<td>55.287</td>
</tr>
<tr>
<td>8</td>
<td>2.971</td>
<td>5.492</td>
<td>7.325</td>
<td>7.688</td>
<td>8.062</td>
<td>41.287</td>
<td>9.027</td>
<td>55.209</td>
</tr>
<tr>
<td>9</td>
<td>2.763</td>
<td>5.504</td>
<td>7.332</td>
<td>7.642</td>
<td>7.882</td>
<td>42.487</td>
<td>8.578</td>
<td>56.672</td>
</tr>
<tr>
<td>10</td>
<td>2.705</td>
<td>4.998</td>
<td>7.233</td>
<td>7.683</td>
<td>8.125</td>
<td>41.872</td>
<td>9.082</td>
<td>57.689</td>
</tr>
</tbody>
</table>

As shown in Table 2, the improved VBPR greatly outshined the traditional recommendation models in precision. The good performance is attributable to the following factors: First, the latent semantic spaces of users and tourist attractions are mined deeply by the BPR; these spaces, grounded on the existing scores, accurately describe user preferences and the popularity of tourist attractions. Second, a set of visual features (e.g. HSV, LBP, and VGG) are introduced to describe the visual contents of tourist attractions, which alleviates the data sparsity.

It can also be seen from Table 2 that the recommendation precision of the improved VBPR model increased steadily with the rising number of tourist attractions. Compared with NMF and KNN, the improved VBPR model only exhibited a small volatility. Hence, the improved VBPR model is a stable method for recommending multiple tourist attractions to users.

Figures 7 and 8 present the precision advantages of improved VBPR model and VMR, respectively. As shown in Figure 7, the precision advantage of VBPR model increased steadily with the growing number of tourist attractions.

As shown in Figure 8, the precision advantage of VMR also increased steadily with the growing number of tourist attractions. The VMR precision is mainly resulted from the improved VBPR model through matrix decomposition to the inclusion of visual features. Meanwhile, the user preferences obtained by the SS model play an auxiliary role.
ted scores were optimized in BPR and

5. CONCLUSIONS

This paper designs a tourist attraction recommendation system based on SS, BPR and improved VBPR models. Firstly, user scores of tourist attractions were predicted by MF model. Then, the predicted scores were optimized in BPR and improved VBPR frameworks, resulting in a high-quality recommendation of tourist attractions. Experimental results show that the tourist attraction recommendation system based on SS and improved VBPR achieved a high accuracy, satisfied user demand, and greatly alleviated data sparsity. The future research will further improve the recommendation performance of the proposed system based on the multimodal semantic correlation between different visual features.

REFERENCES


