# A Personalized System for Tourist Attraction Recommendation

Thara Angskun<sup>+</sup>, Thawatphong Phithak and Jitimon Angskun

School of Information Technology, Suranaree University of Technology, Nakhon Ratchasima, Thailand

**Abstract.** Currently, travelers can easily access travel information from the internet by themselves. Unfortunately, most of these tourism websites recommend the same content to every traveler. Hence, travelers may receive overwhelming number of options or may receive information that does not comply with their own interests. This article proposes a personalized system for tourist attraction recommendation using a clustering technique and an analytic hierarchy process (AHP). The clustering and AHP are combined to construct a ranking model of tourist attractions. The model is used to recommend the tourist attractions based on individual traveler's preferences and constraints. The ranking model evaluation uses 400 test cases consisting of 50 tourist attractions ranked by 400 travelers. The evaluation results reveal that the tourist attraction ranks obtained from the ranking model is similar to those ranks rated by experienced travelers. The Spearman correlation coefficient is equal to 0.907.

**Keywords:** clustering, AHP, personalization, attraction recommendation.

## 1. Introduction

In recent years, Travelers can easily access travel information from the internet by themselves. The tourism websites provide several travel-related information including the attractions. Unfortunately, most of these websites recommend the attractions by letting travelers search or browsing attractions through their electronic catalogues. Hence, travelers may receive overwhelming number of options [1]. These problems can be overcome by researching on personalized attraction recommendation.

The personalized attraction recommendation (PAR) [2] is a one-to-one recommendation system offered a ranked list of attractions. The PAR tries to understand each traveler and predict the most suitable attractions based on the traveler's preferences and constraints. Generally, the recommendation approaches can be classified into six categories [3] as shown in Table 1. One of the most successful recommendation approaches [4], [5] is "Collaborative filtering (*CF*)".

Approach	Description
Content-based	It recommends attractions that are similar to the one that the traveler liked in the past.
Collaborative filtering	It recommends attractions that other travelers with similar tastes liked in the past.
Demographic	It recommends attractions based on a demographic profile of the traveler.
Knowledge-based	It recommends attractions based on specific knowledge of how attraction features meet travelers' preferences.
Community-based	It recommends attractions based on the preferences of the travelers' friends (not anonymous with similar tastes).
Hybrid	It recommends attractions based on the combination of the above mentioned approaches

The goal of *CF* approach is to suggest new attractions to a target traveler based on the opinions of other like-minded travelers [5]. Typical *CF* scenario, there is a list of *m* travelers *T* such that  $T = \{T_1, T_2, ..., T_m\}$ 

<sup>&</sup>lt;sup>+</sup> Corresponding author. Tel.: +66 44 224 210. *E-mail address*: angskun@sut.ac.th.

and a list of *n* attractions *A* such that  $A = \{A_1, A_2, ..., A_n\}$ . Each traveler  $T_i$  (where  $T_i \in T$ ) has a list of  $A_{t_i}$  (such that  $A_{t_i} \subseteq T$ ), which each traveler has expressed his/her opinions. The opinions can be explicitly given by the travelers as a rating score. The task of *CF* is to find a list of *p* new attractions  $A_p$  (a.k.a. Top-*p* recommendation) that a target traveler  $(T_j)$  (where  $T_j \in T$ ) will like the most. These recommended attractions may not be visited by the target traveler, i.e.,  $A_p \cap A_{T_i} = \emptyset$ .

This article presents a personalized system for tourist attraction recommendation based on the collaborative filtering (*CF*) approach. The proposed *CF* approach is based on a combination of data clustering and analytic hierarchy process (*AHP*) for the model construction. The results of model are Top-p ranked lists of tourist attractions for each target traveler. Moreover, this article also presents a ranking evaluation method based on Spearman Correlation Coefficient.

### 2. Personalized Tourist Attraction Recommendation

The design and development of a personalized system for tourist attraction recommendation using a combination of data clustering [6] and analytic hierarchy process (AHP) [7] have a system framework as shown in Fig. 1. The framework consists of 3 principal components, which are users of the system, a data gathering process, and a model construction process.

Users of this recommendation system can be classified into 4 categories which are travelers, a system administrator, target travelers and a travel expert. Travelers are those who have experiences in traveling and can provide traveler profiles and opinions about criteria of tourist attractions. The system administrator is one who creates a tourist attraction recommendation model. The model can be recreated when new information is obtained from travelers. Target travelers are those who need the tourist attraction recommendation. The last but not the least users' category is the travel expert. The travel expert is one who can assign the scores for each tourist attraction. The scores from the expert along with information from travelers are used in the recommendation process.



Fig. 1: A framework of personalized tourist attraction recommendation.

Data gathering process collects travel-related data from travelers and experts. Those gathered data is stored in a database and then applied for recommendation model construction.

Model construction process of tourist attraction recommendation is performed by the system administrator as mentioned above. This process applies a clustering technique to find neighbors of target travelers. Neighbor information and an analytic hierarchy process are used to rank the tourist attractions because of its capability of embedding the experienced traveler's information in the ranking process. The experienced traveler's information plays an important role in the ranking of tourist attractions because each individual has different travel behavior, which can be learned from the experience of others. Finally, the ranked tourist attractions are recommended to those target travelers.

## 3. Experiments

#### **3.1.** Experimental Evaluation

The experimental evaluation focuses on evaluating the tourist attraction ranking model. The model recommends the most appropriate tourist attractions for each target traveler based on their preferences and constraints. The model is evaluated by testing data of ranking a list of 50 tourist attractions ranked by 400 travelers. The evaluation compares the difference between tourist attractions ranked by 400 travelers and tourist attractions recommended by the ranking model. The model employed clustering technique to predict cluster of target traveler based on traveler profiles. After that the model will recommend those pre-ranked tourist attractions for the cluster to each tourist. The evaluation process of ranking model is shown in Fig. 2.



Fig. 2: An evaluation process of the ranking model.

This research applies a statistical method called Spearman's correlation coefficient (or Spearman's rank or Spearman's Rho) [8-9] to measure a difference between traveler-based and model-based ranks of tourist attractions as shown in Equation 1.

$$\rho = 1 - \frac{6T}{n(n^2 - 1)} \tag{1}$$

where  $\rho$  is the Spearman's Rho.

 $T = \sum_{i=1}^{n} (r_i - s_i)^2$  is the summation square of difference between each score pair.

 $r_i$  is a tourist attraction priority that obtained from the ranking model.

 $s_i$  is a tourist attraction priority that obtained from each traveler.

n is the total number of travelers.

The Spearman's Rho has possible values between -1 and 1. If the result has a value closed to 1, it means that both tourist attraction priorities are very similar. In other words, the priority obtained from the model is correct (it matches with the priority obtained from each traveler). On the other hand, if the result has a value closed to -1, it means that both tourist attraction priorities are very different, i.e., the priority obtained from the model is incorrect.

#### **3.2.** Experimental Results and Discussion

An evaluation result of tourist attraction ranking model is shown in Fig. 3. The tourist attractions ranked by the model are compared with tourist attractions ranked by 400 experienced travelers. The difference numbers of tourist attractions (between 10 to 50 tourist attractions) is used to evaluate the ranking model. As shown in Fig. 3, the Spearman's Rho value of 50 tourist attractions is equal to 0.907 which is very close to 1.

The value indicates that the model-based attraction ranks are very similar with traveler-based attraction ranks. In fact, the Spearman's Rho values of every number of attractions are very close to 1 that means the number of attractions does not effect to the accuracy of the ranking model.



Fig. 3: An evaluation result of tourist attraction ranking model.

## 4. Conclusions

This article proposes a design and development of a personalized system for tourist attraction recommendation using a clustering technique and an analytic hierarchy process (AHP). The clustering and AHP are combined to construct a ranking model of tourist attractions. The model is applied to recommend the Top-*p* tourist attractions based on individual traveler's preferences and constraints. This article also proposes an evaluation method for ranking based on a Spearman correlation coefficient. The ranking model evaluation uses 400 test cases consisting of 50 tourist attractions ranked by 400 experienced travelers. The evaluation results reveal that the tourist attraction ranks obtained from the ranking model are similar to those ranks rated by experienced travelers. The Spearman correlation coefficient is equal to 0.907.

## 5. References

- Q. N. Nguyen, D. Cavadaand, and F. Ricci. On-tour Interactive Travel Recommendations, In Proceedings of the 11<sup>th</sup> International Conference on Information and Communication Technologies in Travel and Tourism (ENTER04). 2004, pp. 259-270.
- [2] F. Ricci, L. Rokach, B. Shapira, and P.B. Kantor. *Recommender Systems Handbook*, Berlin, Heidelberg: Springer-Verlag, 2011.
- [3] R. Burke. Hybrid Web Recommender Systems. In Proceedings of *the Adaptive Web*. Springer Berlin Heidelberg. 2007, pp. 377-408.
- [4] J.L. Herlocker, J.A. Konstan, L.G. Terveen, and J.T. Riedl. Evaluating collaborative filtering recommender systems, *ACM Transactions on Information Systems*. 2004, 22(1): 5-53.
- [5] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl. Item-based collaborative filtering recommendation algorithms, In Proceedings of *the 10<sup>th</sup> international conference on World Wide Web*: ACM. 2010, pp. 285-295.
- [6] J.A. Hartigan. Clustering Algorithms, New York: John Wiley & Sons, 1975.
- [7] T. Saaty. *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*, New York: McGraw-Hill, 1980.
- [8] E.C. Fieller, H.O. Hartley, and E.S. Pearson. Tests for rank correlation coefficients, *I. Biometrika*. 1957, 44: 470-481.
- [9] R.J.G.B. Campello and E.R. Hruschka. On comparing two sequences of numbers and its applications to clustering analysis, *Information Science*. 2009, *179*: 1025-1039.