J. Myanmar Acad. Arts Sci. 2019 **Vol. XVII. No.3**

A RECOMMENDER SYSTEM FOR INTERESTING PLACES IN MYANMAR BY USING COLLABORATIVE FILTERING METHOD

Yee Yee Aung*

Abstract

Tourists need information about places to visit, where to eat, activities, and so on. This information is typically found in guidebooks and on the Internet, requiring the tourist to actively search for relevant information. Instead, this information can be available on an electronic marketplace where it is personalized to suit each tourist. A recommender system is one of the types of personalization information filtering system that seeks to predict the "rating" or "preference" which a user would give to an item. Personalized recommender systems are becoming more interesting especially when not limited to just searching for information but they are also to recommend the items that would be more appropriate for the user’s needs or preferences. The aim of this research, this recommender system is then used to predict interesting places in Myanmar (or ratings for places of items) that the users may have an interest in.

Keyword: Personalization, Recommender system, Collaborative Filtering System

Introduction

Tourism information on the World Wide Web is dynamic and constantly changing. It is not easy to get the relevant and updated information to each individual users’ need. For the tourism domain, the internet has become a new media to deliver first-hand information about services to the customers from around the world. When surfing the Internet Web sites, users are demanding more powerful tools that are capable of integrating and interpreting the huge amount of mixed information available on the Web.

Personalized recommender systems are becoming more interesting, especially when not limited to just searching for information but that also to recommend the items that would be more appropriate for the user’s needs or preferences. There are mainly two types of Recommender Systems(RSs): Content-Based Filtering (CBF) and Collaborative Filtering (CF). CF is one of the most commonly used methods in personalized recommendation systems

*. Dr, Associated Professor, Department of Computer Studied, Dagon University

and is used for dealing with this problem. The overwhelming amount of data necessitates mechanisms for efficient information filtering. Collaborative filtering aims at predicting the users' interest for a given item based on a collection of user profiles.

Memory-based approaches to collaborative filtering that can be divided into user-based and item-based that approaches alike. In fact, there are various reasons why service providers may want to develop this technology: they want to increase the number of items which are sold, sell more diverse items, increase the user satisfaction, and better understand what the user wants.

A travel plan may be composed of various attractions, destinations, and accommodation services that are located in delimited areas. From the point of view of the user these various alternatives can be considered and selected as a single travel destination. A travel RS is typically suggested by a travel agent or a destination management organization to increase its income, i.e., sell more hotel rooms, or to increase the number of tourists to the destination. Whereas, the user's primary motivation for accessing this system is to find a suitable hotel and interesting events/attractions when visiting a destination. Myanmar has a variety of travel attractions that are the huge cultural and geographical diversity, but has retained much of its own historic and unique character. This research is to provide information and services to people for News information about interesting places in Myanmar and to get the recommended place of items easily about user's satisfaction.

Data And Knowledge Sources From Recommender System

Recommender Systems (RSs) are information processing systems that actively gather various kinds of data in order to build their recommendations. Data is primarily about the items to suggest and the users will receive these recommendations. But, since the data and knowledge sources available for recommender systems can be very diverse, ultimately, whether they can be exploited or not depend on the recommendation technique. Recommendation techniques can either be of poor knowledge or knowledge dependent. While poor knowledge makes use of simple and basic data such as user ratings/evaluations for items, and knowledge dependent uses ontological descriptions of the users or the items, or constraints, or social relations and activities of the users. In this system, the use of knowledge is poor technique. Thus, as a general classification, three kinds of elements namely items, users and transactions construct the data used by recommender systems.

(i) Items T = {t_1, t_2,…,t_n}: Items are the objects or products in the recommender system for making suggestion to the user. Items may be characterized by their complexity and their value or utility. Travels are the most complex items that have been considered. All domain relevant items are stored in set **T**.

(ii) Users U = {u_1,u_2,…,u_n}: Elements of U comprises of all the users that have browsed items or contributed to the item ratings in the sites. In this system, site browsing patterns in a web-based recommender system or travel search patterns in a travel recommender system can be used to describe the users.

(iii) Transactions: Transaction, a recorded interaction between a user and the recommender system, are log- like data that stored important information generated during human-computer interaction used for the recommendation generation algorithm by the system. A reference to an item selected by the user and a description of the context (e.g., the user goal/query) for that particular recommendation is an instance of a transaction log and an explicit feedback, such as the rating for the selected item, and during a transaction is provided by the user. The rating is in fact the most popular form of transaction data collected by a recommender system which may be explicitly or implicitly. In the explicit rating, the user is asked to provide about an item on a rating scale.

(i) Explicit Profiling: The most apparent way to collect a customer's preferences is simply to ask the user to provide the information. This explicit profiling is often achieved by demanding the user to complete a preliminary set of questions in detail any appropriate preferences.

(ii) Implicit Profiling: Implicit profiling techniques build individual information by inferring users rating from so-called interest indicators depending on customer's interactions with the system.

User-item matrix: User-item matrix is a matrix of customers against products that have components as the explicit ratings of customers to products (user to item). Some of the user-matrix cells are not loaded, as there are products that are not rated by any user. For m items and k users, the user profiles are represented in a k x m user-item matrix X. Each element $x_{k,m}$ = r indicates that user k rated item m by r, where $r \in \{1, ..., |r|\}$. If the item has been rated $x_{k,m} = 0$; means that the rating is unknown. The user-item matrix can be decomposed into row vectors:

$$X=[u_1, \ldots, u_k]^T, \quad u_k = [x_{k,1}, \ldots, x_{k,m}]^T, \; k = 1, \ldots, k \tag{1}$$

Where, T denotes transpose. Each row vector $\mathbf{u}^{\mathbf{T}}_{\mathbf{k}}$ corresponds to a user profile and represents a particular user's item ratings. As discussed below, this decomposition leads to user based CF. Alternatively, the matrix can also be represented by its column vectors:

$$X=[i_1, \ldots, i_m], \quad i_m = [x_{1,m}, \ldots, i_{k,m}]^T, \; m = 1, \ldots, m \tag{2}$$

where, each column vector i_m corresponds to a specific item's ratings by all k users.

Recommendation Techniques

A recommender system is a subclass of personalization information filtering system. There are mainly two types of recommender systems: Content-Based Filtering (CBF) and Collaborative Filtering (CF). CF is one of the most commonly used methods in personalized recommendation systems and used for dealing with this problem. This system use a memory-based collaborative filtering approach build a model from a user's past behavior (selected and/or numerical ratings given to those visited in Myanmar) as well as similar decisions made by other users.

Collaborative Filtering Methods

Collaborative filtering (CF) can be classified into two main methods as user-based collaborative filtering (memory-based) and item-based collaborative filtering (model-based). The memory-based methods are most popular prediction techniques in collaborative filtering applications. This approach uses user rating data to compute the similarity between users or items. This is used for making recommendations. They utilize the entire user-item database to generate predictions. The user-based collaborative filtering algorithm is a memory-based algorithm. The model-based collaborative filtering methods use the user's preferences to learn a model, which is then used for predictions. Model-based methods are not suitable for environments in which user preference models must be updated rapidly or frequently.

Collaborative Filtering System Architecture

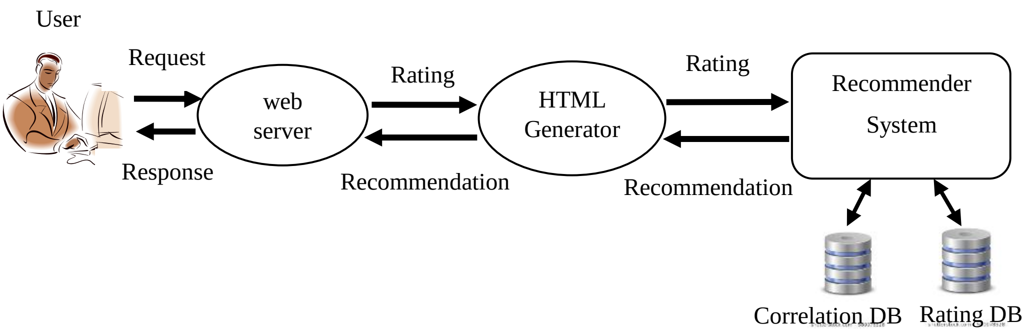

Figure 1 : Collaborative Filtering System Architecture

Collaborative filtering (CF) is a technique used by recommender systems, is shown in figure 1.

The implementation of this method recommends to the active user the items that other users with similar tastes liked in the past. The similarity in taste of two users is calculated based on the similarity in the rating history of the users. This system use a collaborative filtering method construct a model from a user's past behavior (selected and/or numerical ratings given to those visited in Myanmar) as well as similar decisions made by other users. In this system, implemented model is then used to ratings for items that the user may have an interest places in Myanmar.

Pearson Correlation Coefficient

It is used for converting similarity between two users or items by measuring obliquity of two series of preferences to act together in a comparative and linear manner. It considers preferences of conflicting users and items. It tries to find each users or items derivations from their average rates while recognizing linear adjustment between two items or users.

$$P,C_{(W,U)}= \frac{\sum_i (r_{w,i} - \bar{r}_w)(r_{u,i} - \bar{r}_u)}{\sqrt{\sum_i (r_{w,i} - \bar{r}_w)^2 \sum_i (r_{u,i} - \bar{r}_u)^2}} \tag{3}$$

w and **u** shows the two users or items for which the coefficient is calculated **i** is an item, $\boldsymbol{r_{w,i}}$ and $\boldsymbol{r_{u,i}}$ are individual rating from **w** and **u** for **i**, and average rating of $\overline{\boldsymbol{r_w}}$ and $\overline{\boldsymbol{r_u}}$ are, for user (or item) **w** and **u**.

Rule-based Personalization with Collaborative Filtering (RPCF) Algorithm

This RPCF algorithm aims to identify users that have relevant interest by calculating similarity between user profiles. In this system, Rule-based Personalization with Collaborative Filtering (RPCF) Algorithm is used and described as follows:

Step 1: Generate rules by using Rule-based Personalization from user query input.

This step generates the content rules based on the user query input. The content rules are in the form IF condition THEN action. Collaborative Filtering works by collecting user feedback in the form of ratings for items in a given domain and exploit similarities and differences among profiles of several users in determining how to recommend an item or how to give the prediction for the active user's interest. A subset of users is chosen based on their similarity to the active user, and a weighted combination of their ratings is used to produce predictions for the active user. Similarity is a powerful way to retrieve interesting information from large repository. The threshold of Correlation Coefficient ranges from -1.00 to +1.00. The value of -1.00 represents a perfect negative correlation while a value of +1.00 represents a perfect positive correlation. A value of 0.00 represents a lack of correlation.

Step 2: Compute the similarity measure between users by means of the Person Correlation Coefficient.

The following Pearson Correlation Coefficient function is used to compute the similarity measure between the user's preference functions.

$$\mathrm{S(a,b)} = \frac{\sum_{i=1}^{N}(x_{a,i} - \bar{x}_a)(x_{b,i} - \bar{x}_b)}{\sqrt{\sum_{i=1}^{N}(x_{a,i} - \bar{x}_a)^2 * \sum_{i=1}^{N}(x_{b,i} - \bar{x}_b)^2}} \tag{4}$$

Where, ***S (a, b)*** is the similarity of user ***a*** and ***b*, N** is the number of items. $x_{a,i}$ and $\boldsymbol{x_{b,i}}$ are the ratings given to the item ***i*** by user ***a*** and ***b***. $\bar{x}_a$ *and* $\bar{x}_b$ are the average ratings (mean) of user ***a*** and ***b.***

A higher collection value indicates more accurate recommendations. The Pearson's Correlation Coefficient only measures the overlapping items between users.

Step 3: Weight the similarity by the number of item ratings and select the neighboring user that have the highest similarity rating with the active user.

The Significance-Weighting method as shown in equation (5) is used to devalue the correlation based on few co-rated items,

$$W(a,b) = S(a,b) * \frac{n}{50} \tag{5}$$

If two users have less than 50 commonly rated items, we apply a significance weight of **n/50**, where, **n** is the number of co-rated items.

If there are more than 50 co-rated items, then a significance weight of 1 is applied which means we leave the correlation unchanged.

Step 4: Compute a prediction from the ratings of neighbour.

Predictions are computed from the weighted combination of the neighbour, which is defined in as:

$$P_{a,i} = \bar{x}_a + \frac{\sum_{b=1}^{M}(x_{b,i} - \bar{x}_b)}{\sum_{b=1}^{M} W(a,b)} * W(a,b) \tag{6}$$

where, ***M*** is the number of users in the neighbour hood. $\boldsymbol{P_{a,i}}$ is the prediction of the active user *a* on the target item ***i.*** [9]

Case Study of a Recommender System for Interesting Places in Myanmar by Using Collaborative Filtering Method

This system is implemented to recommend items from natural tourism/ecotorisum places and interesting places list to the user. In this system, user based filtering processes are as follow:

Step 1: Active user observe the information of the interested places in Myanmar, especially in natural places, historical places and other famous pagoda, beaches etc. In this system describe the interesting places, historical places and then user can do activities in Myanmar such as climbing, trekking, and dolphins and Ayeyarwady, …etc.)

Step 2: Active user give the rating (rating scale-1-5) for his/her preferences on the item places. Table 1 is shown for sample calculated rating based on seven interested places for number of twenty users.

Table 1: Sample Rating Scores for Interesting Places

User id	Chin	Kachin	Rhaing	Shan	Mandalay	Mgwae	Sagaing
1	1	2	3	4	5	5	4
2	3	2	2	1	5	4	3
3	2	2	3	3	4	5	1
4	2	4	5	2	3	1	5
5	5	5	4	3	2	2	3

User id	Chin	Kachin	Rhaing	Shan	Mandalay	Mgwae	Sagaing
6	3	1	4	2	3	5	4
7	1	4	5	2	3	5	4
8	4	5	2	3	4	2	1
9	1	5	3	2	1	4	2
10	2	4	5	3	2	1	3
11	1	4	5	3	1	5	1
12	3	3	1	4	5	3	2
13	4	5	1	2	3	4	5
14	2	1	3	4	5	3	1
15	4	5	3	1	2	4	1
16	3	5	2	1	4	3	5
17	2	3	4	5	2	2	3
18	5	2	3	1	4	2	4
19	3	4	5	2	1	4	3
20	1	2	3	4	5	2	3

Step 3: In this step, to estimate a prediction for an active user, the memory based algorithms first find the user's neighbours (the users who are similar to the active user). Then, the active user's rating is predicted by averaging the (weighted) known rating on the places item by his/her neighbors. It is based on the assumption that similar users have similar rating patterns.

Step 4: Find the measurement of the similarity between users by using Pearson's Correlation Coefficient. According to results of similarities measurement between user 20 (active user) and other users (user 2, user 3, …, user 20), the highest similarities is found between user 4 and user 20. According to results of similarities measurement between user 20 (active user) and other users (user 2, user 3, …, user 20), the highest similarities is found between user 4 and user 20.

Table 2: Show Similarity and User Pairs

User Pair	Similarity
user1, user 20	0.84
user2, user 20	0.70
user3, user 20	0.26
user4, user 20	**0.97**
user5, user 20	0.41
user6, user 20	0.61
user7, user 20	0.58
user8, user 20	0.60
user9, user 20	0.45
user10, user 20	0.68
user11, user 20	0.38
user12, user 20	0.82
user13, user 20	0.38
user14, user 20	0.89
user15, user 20	0.28
user16, user 20	0.58
user17, user 20	0.79
user18, user 20	0.56
user19, user 20	0.41

Similarity is greater than 0, this pair of users may be chosen and then user4 rating of each items and user20 rating of each items are the same, these items count are increased. These resulting counts are sorted and shown to the user in ascending order as shown in table 2. According to above the table result, Shan state is suitable place for visit in Myanmar for active user, is shown in figure 2.

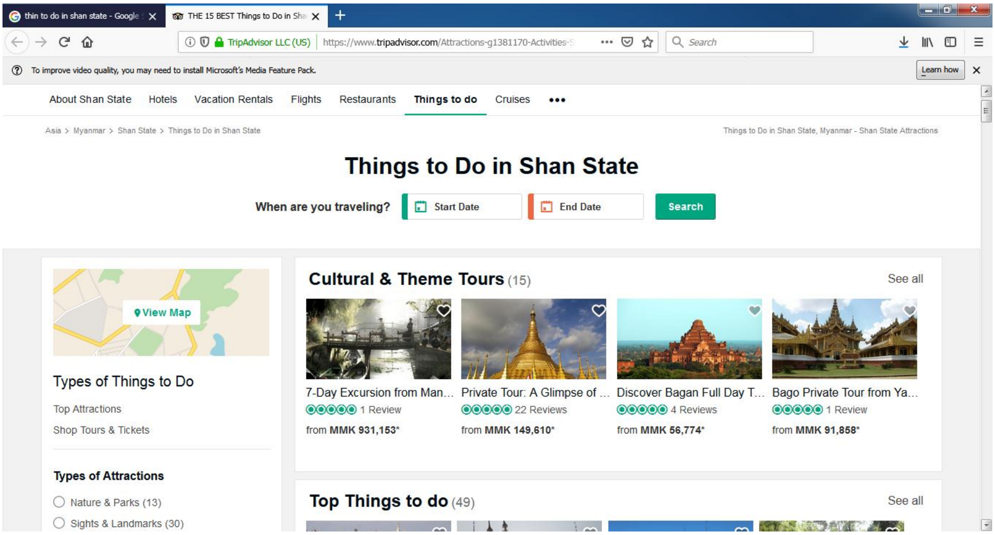

Figure 2: Recommended Place of Shan State Web Page

Discussion and Conclusion

A non-personalized recommender system is simpler to generate the popular places and top ten selections of places. This system is implemented by using memory-based technique, based prediction mechanisms and user based collaborative filtering for recommendations. This system will assist in achieving relevant recommendations combined with an active users' preferences and the same with past users' preferences and behaviors.

This recommender system helps to suggest and inform visitors about these natural tourism/ecotourism opportunities and in doing so, it hopes to promote the conservation of Myanmar's most beautiful areas and most extraordinary species. These suggestions relate to various decision-making processes, such as what to do and what places to visit. A tourist using this system can be helped to consider and select as a single travel destination. In future research, the community-based recommender system will implement for ecotourism and historical places in Myanmar. This type of RSs model acquires information about the social relations of the users and preferences of the user's friends.

Acknowledgements

First of all, I am grateful to Dr. Win Naing, Rector of Dagon University, and Dr.Yi Mon Win Associate Professor, Head of Department of Computer Studies, Dagon University, for their kind permission to carry out this work. I am greatly indebted to Professor Dr. Nwe Nwe Win, Head of Department of Computer Studies (Retired), University of Yangon, for her valuable guidance, overview and perpetual encouragement. Finally, I am very thankful to Dr. Soe Mya Mya Aye, Professor, Head of Department of Computer Studies, and University of Yangon, for her advices and comments.

References

Arsan, T., et al., (2016), "Comparison of Collaborative Filtering Algorithms with Various Similarity Measures for Movies Recommendation", International Journal of Computer Science, Engineering and Applications (IJCSEA) Vol.6, No.3

Neuhol d, E. J., 2003, Personalization and User profiling & Recommender Systems, Department of Computer Science and Business Informatics, University of Vienna

Nilashi M, et al., 2013, Collaborative Filtering Recommender System, Research journal of Applied Science, Engineering and Technology", 5(16), 4168-4182

Ricci, F. et al., 2011, Recommender Systems Handbook, Springer Science + Business Media

Sarwar, B., Karypis, G., Konstan, J., and Riedl, J., 2001, Item-Based Collaborative Filtering Recommendation Algorithms, Group Lens Research Group/Army HPC Research Center Department of Computer Science and Engineering University of Minnesota, Minneapolis

Yunkyoung. L., 2015, Recommendation System Using Collaborative Filtering, Master Thesis and Graduate Research, San Jose State University

http://www.csu.edu.au/division/studserv/maths/ index.htm, Charles Sturt University, December, 2008