PERSONALIZED TRAVEL RECOMMENDATION SYSTEM BY USING USER-BASED COLLABORATIVE FILTERING METHOD

Tin Nilar Win[1], Dr. Wint Pa Pa Kyaw[2]

Abstract

The huge amount of data in mobile business processes and physical limitations have increased the importance of personalization process. Recommendation system can aid users in discovering information or items in a personalized manner. Personalization constitutes tailoring a service or a product to accommodate specific individuals. The users are flooded with so much of choices that it is hard for them to find appropriate and suitable information in tourism field. Personalized travel recommendation system can help customers in travel planning because it may be so complicated and confusing to process a lot of information on the travel sites. Collaborative filtering method compares the user's past ratings with those of other users (neighbors) to find users with similar preferences. Highly rated items by these neighbors will be recommended. In this research, the system suggests personalized travel locations to users based on their rating profiles and interests by using user-based collaborative filtering method.

Keywords: Recommendation System, Personalization, User-based Collaborative Filtering

1. Introduction

Increasing amounts of information on traveling are available on the web. As is the case for many other domains, the web is becoming the most important information source for planning a holiday. Specialized web sites, such as WikiVoyage or Frommers, are providing information and travel advice on different destinations. Others, such as Expedia or SkyScanner, exist for finding the best deals, flight tickets or travel packages. On TripAdvisor websites, reviews and evaluations of hotels, restaurants, and attractions can be read. Although these services are all valuable information sources, they typically give no personal advice which holiday destination to choose.

Now-a-days recommendation system is becoming very popular and people are getting attracted to it, as it is assisting them in discovering interesting items over huge amount of information. Recommendation systems are used in digital libraries, electronic stores, travel tours, restaurants, hospitals and in general can be useful in any decision-making process to provide predictions of appropriate items to specific users. During a commercial interaction, recommendation systems have advantages for both customers and merchants.

---

1. Assistant Lecturer, Department of ICT, Co-operative University (Thanlyin)
2. Associate Professor, Department of Computer Studies, University of Yangon

In a business interaction through the online shopping, recommendation systems can help customers to find their favorite items among an overwhelming number of items in an electronic department store. Therefore, recommendation systems can facilitate and accelerate shopping for users. Merchants proffer their products and hereby they can increase their sales and customers satisfaction by offering the new and preferable items. Similarly, in a digital library, recommendation systems can manage information overload by helping users to choose appropriate information items from a large set of alternatives.

In the tourism field, travel recommendation systems purpose to match the characteristics of tourism and leisure resources or attractions with the user needs. The travel companies have to aware of these preferences from different tourists and serve more attractive packages to get more business and profit. Therefore, the demand for intelligent tour services, from both travellers and tour companies, is expected to increase dramatically. Since recommendation systems have been successfully applied to enhance the quality of service for customers in a number of fields, it is natural direction to develop recommendation systems for personalized travel package recommendation.

1. Anatomy of Recommendation System

Recommendation systems are made to help users in their search for a fitting product from an overwhelming array of options. Recommendation systems can nowadays be found in a broad range of applications and are very common in e-business solutions. Every recommendation system needs to at least consist of two base elements: the user profile and the information filtering technique.

The user profile is needed for the system to represent the user's information and preferences. Without a user profile, it becomes impossible to generate personalized recommendations. Based on this user profile, the recommendation will need a certain matching (filtering) approach to match users with items. Figure 1 shows the process of recommendation system.

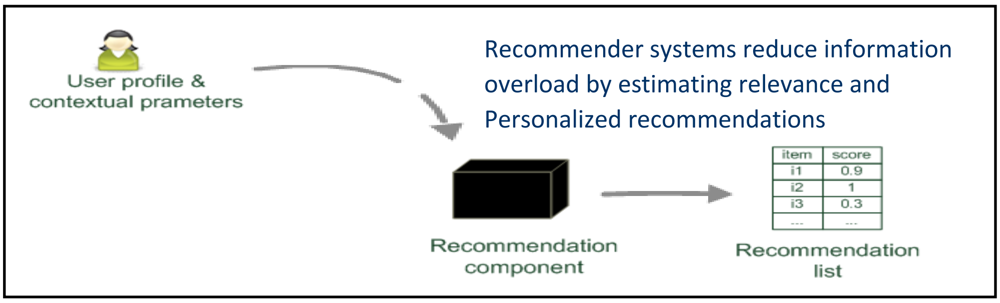

Figure 1: Paradigm of Recommendation System

2.1 User profiles

A user profile includes all personal information needed of a user to help with making recommendations. For constructing and maintaining user profiles, there are many different ways to represent the user's preferences. Two of the most successfully used techniques are to save a user-item matrix with ratings a user made in the past combined with the use of a feature vector, representing the affinity of a user to predefined features. For example, in a travel recommendation, feature can represent hobby. The user profile would then represent what are user hobbies. The user can also be asked for explicit input. In such case the users can state their preferences by answering questions or indicating their interests from a list.

Recommendation systems always try to improve their user's profiles and adapt to changing user's preferences over time. So, another important aspect of any system is user feedback. Explicit feedback can be attained by asking the user to rate items or ask his opinion (like/dislike) on a recommendation. Explicit feedback is the most accurate information but ask the user for to make an effort for the system. The more user-friendly approach is to collect implicit feedback from a user's behaviour and (natural) interactions with the system. Processing this information can also give insight to the user's preferences.

2.2. Filtering approaches

With the user profile and a database of items available, the final step to make a recommendation is to match users with suitable items. The filtering method determines how these are found. This work categorizes recommendation systems by their filtering approach and distinguishes between four different ones.

(1) In Content-Based Filtering, where the system makes use of the user's profile to recommend items that exhibit similar characteristics to what he has liked in the past.

(2) In Collaborative Filtering, the recommendation compares the user's past ratings with those of other users to find users with similar taste. Highly rated items by these neighbors will be recommended.

(3) Knowledge-Based recommendations make use of domain specific information to match user interests with items.

(4) Hybrid systems represent any system that combines two or more of the above approaches to a more complex whole.

2. Rating Estimations

An important element in recommendation systems is the user-item ratings. Ratings in recommendation systems represent how pleasing or useful a certain item is to a user. User can give the product an explicit rating when he has experienced it. But for most products, such rating is not known. Most recommendation approaches reduce the problem of making a recommendation to estimating ratings for items a user hasn't rated yet. Given these estimations, the system can then recommend the highest scoring items to the user.

3.1 Explicit ratings

Asking for explicit item ratings is probably the most precise one among the existing alternatives for gathering users' opinions. In most cases, five-point or seven-point Likert response scales ranging from "Strongly dislike" to "Strongly like" are used; they are then internally transformed to numeric values so the previously mentioned similarity measures can be applied. Some aspects of the usage of different rating scales, such as how the users' rating behavior changes when different scales must be used and how the quality of recommendation changes when the granularity is increased.

Explicit ratings require additional efforts from the users of the recommendation system and users might not be willing to provide such ratings as long as the value cannot be easily seen. So, the number of ratings could be very small, recommendation results might be poor. Figure 2 shows the five-point interval ratings scale.

Figure 2: Example of 5-point interval ratings

3.2 Implicit ratings

Implicit ratings are typically gathered by the web shop or application in which the recommendation system is embedded. When a customer buys a product, many recommendation systems interpret this behavior as a positive rating. The system could also track the user's browsing manner. If the user retrieves a page with detailed item information and remains at this page for a longer period of time, for example, a recommendation could interpret this behavior as a positive orientation toward the item.

Although implicit ratings can be gathered continuously and do not need more efforts from the side of the user, one cannot be sure whether the user manner is correctly interpreted. Still, if a sufficient number of ratings is available, these particular cases will be factored out by the high number of cases in which the interpretation of the behavior was right. In some domains (such as personalized online radio stations) collecting the implicit feedback can even result in more accurate user models than can be done with explicit ratings.

3. Collaborative Filtering

The major purpose of collaborative filtering approaches is to exploit information about the past behavior or the opinions of an existing user community for predicting which items the current user of the system will most probably like or be interested in. These kinds of systems are in widespread fields use today, in particular as a tool in online retail sites to customize the content to the needs of a particular customer and to thereby encourage more items and promote sales.

From a research viewpoint, these types of systems have been explored for many years, and their advantages, their performance, and their limitations are nowadays well understood. Years ago, many types of algorithms and techniques have been proposed and successfully evaluated on real-world and artificial test data.

A matrix of given user–item ratings is taken as the only input and typically produced the following types of output in pure collaborative approaches. These are (a) a numerical prediction indicating to what degree the current user will like or dislike a certain item and (b) a list of *n* recommended items. Such a *top-N* list should, of course, not include items that the current user has already bought.

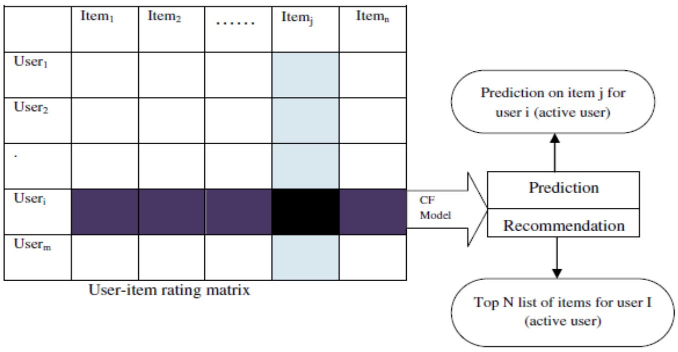

Figure 3: Collaborative Filtering Process

4.1 User-based collaborative filtering

User–based collaborative filtering is a straightforward algorithmic interpretation of the core premise of collaborative filtering: find other users whose past rating behavior is similar to that of the current user and use their ratings on other products to forecast what the active user will like. In a travel recommendation system, to predict Mary's preference for an item she has not rated, user–based collaborative filtering looks for other users who have high agreement with Mary on the items they have both rated. These users' ratings for the product are then weighted by their level of agreement with Mary's ratings to predict Mary's preference.

Besides the rating matrix R, a user–based collaborative filtering system needs a similarity functions: U×U → R calculating the similarity between two users and a method for using similarities and ratings to generate predictions. The main idea of user-based collaborative filtering is that given a ratings database and the ID of the current (active) user as an input, identify other users referred to as peer users or nearest neighbors that had similar likes to those of the current user in the past. Then, in travel recommendation, for every tour package *p* that the active user has not yet seen, a prediction is computed based on the ratings for *p* made by the peer users. The underlying assumptions of such methods are that (a) if users had similar preferences in the past, they will have similar preferences in the future and (b) user tastes remain stable and consistent over time.

With respect to the determination of the set of similar users, one common measure used in recommendation systems is Pearson's correlation coefficient. The similarity *sim* (*a*, *b*) of users *a* and *b*, given the rating matrix *R*, is defined in the following formula. The symbol $\overline{r_a}$ corresponds to the average rating of user *a*.

$$sim(a,b) = \frac{\sum_{p \in P}(r_{a,p} - \overline{r_a})(r_{b,p} - \overline{r_b})}{\sqrt{\sum_{p \in P}(r_{a,p} - \overline{r_a})^2}\sqrt{\sum_{p \in P}(r_{b,p} - \overline{r_b})^2}}$$

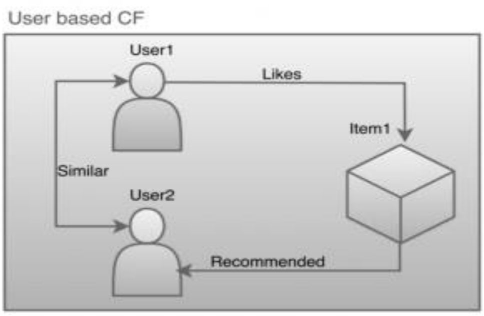

Figure 4: User-based Collaborative Filtering Process

4. Results and Discussion

User-based collaborative filtering recommendation based on nearest-neighbors enjoys a large amount of popularity, due to its simplicity, efficiency, and ability to give accurate and personalized recommendations. Table 1 shows a database of ratings of the current user, Mary, and some other users. Mary has, for instance, rated “Package1” with a “5” on a 1-to-5 scale, which means that she strongly liked this item. The task of a recommendation system in this simple example is to determine whether Mary will like or dislike “Package5”, which Mary has not yet rated or seen.

In this sample, $U = \{u_1, . . . , u_n\}$ to denote the set of users, $P = \{p_1, . . . , p_m\}$ for the set of tour packages (items), and R as an $n \times m$ matrix of ratings $r_{i,j}$, with $i \in 1 . . . n, j \in 1 . . . m$. A numerical scale from 1 (strongly dislike) to 5 (strongly like) can be defined as the possible rating values. If an item j has not been rated by a certain user, the corresponding matrix entry $r_{i,j}$ remains empty.

Table 1: Sample Ratings Database for Collaborative Recommendation

	Package1	Package2	Package3	Package4	Package5
Mary	5	3	4	4	?
User1	3	1	2	3	3
User2	4	3	4	3	5
User3	3	3	1	5	4
User4	1	5	5	2	1

By substituting the rating value from Table 1 in Pearson’s correlation coefficient formula, the similarity of Mary to User1 is thus as follows: ($\overline{r_{Mary}}$ = $\overline{r_a}$=4, ($\overline{r_{User1}}$ = $\overline{r_b}$=2.4:

$$\frac{(5-\overline{r_a})*(3-\overline{r_b})+(3-\overline{r_a})*(1-\overline{r_b})+\cdots+(4-\overline{r_a})*(3-\overline{r_b}))}{\sqrt{(5-\overline{r_a})^2+(3-\overline{r_a})^2+\cdots}\sqrt{(3-\overline{r_b})^2+(1-\overline{r_b})^2+\cdots}}=0.85$$

The values are taken from +1 (strong positive correlation) to −1 (strong negative correlation) in Pearson correlation coefficient. The results 0.70, 0.00, and −0.79 are the similarities to the other users, *User2* to *User4* respectively.

Based on these calculations, *User1* and *User2* were somehow similar to *Mary* in their rating behavior in the past. The Pearson measure regards the fact that users are different with respect to how they interpret the rating scale. Some users tend to give only high ratings, whereas others will never give a 5 to any package. The Pearson coefficient factors these averages out in the calculation to make users comparable – that is, although the absolute values of the ratings of *Mary* and *User1* are completely different, a rather clear linear correlation of the ratings and thus similarity of the users is detected.

This fact can also be seen in the visual representation in Figure 5, which both illustrates the similarity between *Mary* and *User1* and the differences in the ratings of *Mary* and *User4*. To make a prediction for *Package5*, which of the neighbors' ratings shall be taken into account and how strongly shall be valued their opinions. In this example, an obvious choice would be to take *User1* and *User2* as peer users to predict Mary's rating.

A formula for calculating a prediction for the rating of user *a* for tour package *p* that also factors the relative *proximity* of the nearest neighbors *N* and *a*'s average rating r_a is the following:

$$pred(a, p) = \overline{r_a} + \frac{\sum_{b \in N} sim(a, b) * (r_{b,p} - \overline{r_b})}{\sum_{b \in N} sim(a, b)}$$

In the sample, the prediction for Mary's rating for *Package5* based on the ratings of near neighbors *User1* and *User2* will be:

$4 + 1/(0.85 + 0.7) * (0.85 * (3 - 2.4) + 0.70 * (5 - 3.8)) = 4.87$

Given these calculation schemes, rating predictions for Mary can be computed for all items she has not yet seen and include the ones with the highest prediction values in the recommendation list. It will most probably be a good choice to include *Package5* in such a list.

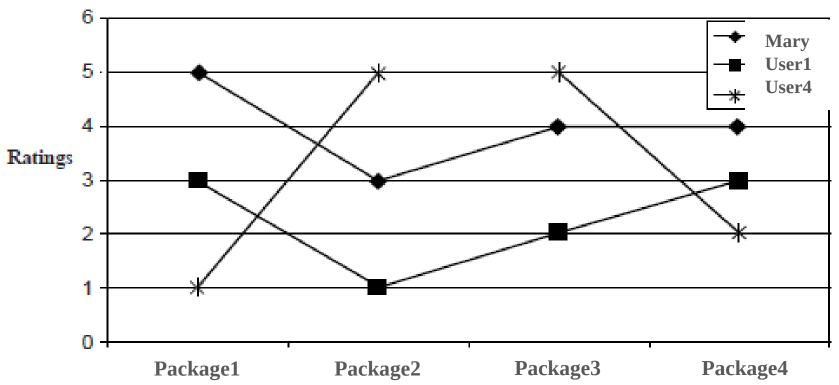

Figure 5: Comparing Mary with two other users

5. Conclusion

Recommendation systems open new opportunities of retrieving personalized information on the Internet. Recommendation techniques have coped with the information overload problem and have proven their usefulness as a tool in many classical domains such as movies, books, and music. A variety of approaches have been used to perform recommendations in these domains, including content-based, collaborative, and knowledge-based.

This paper proposes a recommendation system that offers personalized recommendations for travel destinations to individuals. It can help overcome information overload problem by exposing users to interesting, novel, surprising and relevant items based on preferences users have expressed either explicitly or implicitly. It can introduce users to new items that have not been known or have not been retrieved. So recommendations can help users in meeting their information needs.

On the whole, it is primarily an intelligent application, created to provide users by personalized recommendations in search process and their decision-making while interacting with large information spaces. These recommendations are based on the users' rating profile, personal interests, and specific demands for their travel destination by using user-based collaborative filtering approach. Recommendation system automates some of these strategies with the goal of supplying affordable, personal, and high-quality recommendations.

Acknowledgement

Firstly, I would like to express my profound gratitude to Dr. Yi Yi Win, Rector, Co-operative University (Thanlyin) for her kind permission to carry out this research. Secondly, I also would like to thank U Oo Tin Thein, Pro-rector, Co-operative University (Thanlyin) and Daw Myint Myint Sein, Pro-rector, Co-operative University (Thanlyin) for encouraging to complete my research. Thirdly, I want to thank Dr. Soe Mya Mya Aye, Professor and Head of the Department of Computer Studies, Yangon University, for performing as a chairperson. And then, I want to thank Dr. Ohnmar San, Professor and Head of the Department of ICT, Co-operative University (Thanlyin) for her kind guidance. Finally, I would like to thank my supervisor Dr. Wint Pa Pa Kyaw, Associate Professor, Department of Computer Studies, Yangon University, for her close supervision and invaluable suggestions. I wish to thank all who help and give me advice on this research.

References

[1] Aggarwal C.C. (2016), *Recommendation System: The Textbook*, ©Springer International Publishing Switzerland 2016, ISBN 978-3-319-29657-9

[2] Christos K. Georgiadis and Athanasios Manitsaris (2008), Personalized Recommendations in Mobile Commerce, https://www.researchgate.net/publication/249675053_personalized_recommendations_in_mobile_commerce

[3] Dietmar Jannach, Markus Zanker, Alexander Felfernig, and Gerhard Friedrich (2011), *Recommendation Systems: An Introduction*, Cambridge University Press, ISBN 978-0-521-49336-9

[4] F.O. Isinkaye, Y.O. Folajimi, B.A. Ojokoh (2015), Recommendation systems: Principles, methods and evaluation, *Egyptian Informatics Journal*, Vol.16, Page 261~273

[5] Francesco Ricci, Lior Rokach, Bracha Shapira, Paul B. Kantor (2011), *Recommendation Systems Handbook*, Springer New York Dordrecht Heidelberg London, ISBN 978-0-387-85819-7

[6] Jeroen Dhondt (2015), A hybrid group recommender system for travel destinations, https://lib.ugent.be/fulltxt/RUG01/002/224/826/RUG01-002224826_2015_0001_AC

[7] Michael D. Ekstrand, John T. Riedl and Joseph A. Konstan (2010), Collaborative Filtering Recommender Systems, Foundations and Trends in Human–Computer Interaction, Vol. 4, No. 2, Page 81~173

[8] Zohreh Dehghani Champiri, Siti Salwah Binti Salim, and Seyed Reza Shahamiri (2015), The Role of Context for Recommendations in Digital Libraries, *International Journal of Social Science and Humanity*, Vol.5, No.11, Page 948~953