



# *Tourist Attraction recommendations using collaborative filtering*

**Ms. Navaneetha H P**

PG Scholar,

Department of computer science and Engineering, Maharaja Institute of Technology (MIT), Mysore, Karnataka, India

**Dr. Sharath Kumar Y H**

Associate Professor & Head

Department of Information Science and Engineering, Maharaja Institute of Technology (MIT), Mysore, Karnataka, India

**Abstract**—Tourism has become the world's largest economy industry. More and more people share their travelogues on travel websites. Recommender system is an effective tool to provide travel services for tourists. In this paper, we present a personalized neighbourhood-based recommendation using Cosine method to recommend cities. This is a fundamental problem whose solution supports other tourism recommendations. To address these challenges, we develop a package generation in a simplified way. We present a system for personalized city recommendation that takes into account the user preferences for cities and the attractions. The proposed solution is based on Collaborative Filtering that relies only on past user behaviour (e.g., the cities each user has visited and liked) and does not assume explicit profiles. We propose a Cosine Method to generate new travel packages to improve the overall recommendation effectiveness. First we generate recommendation process, then generation of neighbours, finally generation of recommendation based on the area of interest and the visiting history of tourist neighbour.

**Keywords**—Recommendation system, collaborative filtering.

## I. INTRODUCTION

Tourism is a great income generator due to an increased demand for its services. Their components range from quality and wide range of transportation to infrastructure, accommodation, food and beverage, support services and travel distribution services. Tour recommendation is different from other recommendation as the tourist interest in package is directly affected by its cost. Cost aware recommendation of package is need of the

recommender system. The travel logs are collected from different agents of company then analyzed for time and financial cost connected to every travel package. The tourist has different level of affordability for aspect of cost. The recommendation system focuses on such factors to make it more effective like data sparsity, cold start problem, scalability and accuracy issues and accuracy in prediction.

The Recommendation systems are software tools providing suggestions for items for a user. The suggestions provided are aimed at supporting their users in various decision making processes, such as

where to plan a tour, what items to buy, for which season, what music to listen, or what news to read. Recommender systems have proven to be valuable means for online users to cope with the information overload and have become one of the most popular and powerful tools in electronic commerce field. Various techniques for recommendation system have been proposed like content-based, user-based and item based collaborative filtering and hybrid recommendation system. Many of them have also been deployed successfully in commercial environments. Added to it there are many evolutionary methods that could be incorporated to achieve better results in terms of handling various challenges of recommendation system. Collaborative filtering systems make recommendations based on groups of users with similar preferences. The similarity between users is computed by comparing the ratings that they give to some of the items. When the system finds out who are the people that share similar interests with the active user, then the items that those people liked are recommended to this user. In this approach, some feedback about the provided recommendations is necessary, so as to know which items the user has liked or disliked (e.g. which places he has enjoyed visiting).

The idea behind recommender systems is to automatically recommend items for a user, aiming to predict the user's interest level about the tourist places. These systems help users to deal with information overload, providing personalized recommendations of places. Recommender systems are typically classified in three categories: *content-based*, which recommends items that are similar to items that the user preferred in the past, *collaborative*, which recommends items that users

with similar profiles have preferred in the past history, and *hybrid* approaches, which combine those two to make recommendations. Collaborative Filtering (CF) is one of the most successful recommendation strategies to date, and is used in many domains, such as social streams and movies. This approach represents the state-of-the-art among recommendation techniques, and it is used as the basis of our proposed method.

Recommender systems propose items from different alternatives for user by analyzing travel history or behaviour. The user's behaviour has affect from unseen interests of user. To invest on getting information about the interest of tourist is not favourable to make good recommendations. The present recommender systems based on collaborative filtering focuses on user's interaction with the system. The information about inactive user is discarded. The topic model cooperated so that to find out the personalized ranking. The goal to create the item based collaborative filtering model.

## II. RELATED WORK

In this section we briefly present some of the research literature related to collaborative filtering, recommender systems, data mining and personalization. Collaborative Filtering (CF) has been proven successful in both research and practice. A few examples selected from different domains are introduced in this section. Amazon.com is one of the most successful and well-known online retailers. It makes use of purchase histories of customers to produce recommendations using the item-based technique. GroupLens is a pioneering and well known

project in automated CF. It had used KNN to recommend Usenet articles. Filtering agents were later integrated into the system to improve prediction quality. When the number of registered users increased, however, it was found that the processing time of the system increased linearly until it crashed. The constant time algorithm Eigentaste was therefore developed and adopted.

Some related applications in the tourism domain were also studied. They are mainly based on knowledge based technologies. CF is only used to rank the recommended products according to user preferences. Dietorecs and NutKing are two similar applications developed by the eCommerce and Tourism Research Laboratory (eCTRL). They recommend travel products and activities based on the Case-Based Reasoning (CBR) technology, which is a problem solving technique that solves the current problem that is manual process by adapting solutions for previously solved, similar problems. The CBR cycle consists of 4 stages, namely Retrieve, Reuse, Revise and Retain.

In recent years, it began to gain research attention in the tourism domain for its ability to produce personalized recommendations. However, existing content base techniques fall short when recommending tourism products which are heterogeneous and complex. With the objective to bridge the gap between content base and tourism products recommendation, a study was carried out to identify the problems in recommending tourism products, and to review the state-of-the-art in both areas. This paper details the results of the study, and suggests future research directions to work towards collaborative travel recommender systems.

The CF technique can be divided into three main components: (a) similarity computing, (b) neighbourhood selection, and (c) top- $N$  recommendation. The literature is rich in various efforts to develop and test algorithms that implement each of these components. In this paper we describe and discuss various techniques to compute similarity between different users. The neighbourhood selection step was addressed primarily and more recently by some online application, where we applied it in a collaborative system using cosine method. Solutions to the top- $N$  recommendation step are specially developed in this paper, where several techniques are empirically analyzed. Moreover, we compare evaluation metrics to the top- $N$  recommendation task, which consists in recommending only the best  $N$  items that would be suited to the user. CF has been a promising solution to the problem of information overload in the past decade.

### III. EASE OF USE

Collaborative Filtering algorithms are commonly used in early systems. It works in three major steps, namely similarity of users (which is the Area of Interest (AOI) of each user), neighbour selection and prediction computation. At the similarity stage, *each user* in the database is assigned a similarity using some measures which we considered as the area of interest. Such similarity is reflected in the prediction they have given on the time of registration. For Recommending places users to be comparable, only AOI that both users have same are counted. At the neighbour selection stage, a number of  $k$  nearest neighbours, who are the users having the visited similar places, of the active user are

selected as predictors for places. Based on the interests of the selected neighbours and some partial information of the active user, prediction of the places is computed.

The core of collaborative filtering algorithms is the calculation of users' similarity. The similarity between the tourist  $T_i$  and  $T_j$  is measured as follows: First the system needs to retrieval all the ratings of

tourist  $T_i$  and  $T_j$  over the attractions, and then calculate the similarity between the tourist  $T_i$  and  $T_j$  through different similarity models, denoted as  $\text{sim}(T_i, T_j)$ . There are mainly three ways to measure the similarity between tourists, including Cosine method, Correlation similarity method and Adjusted Cosine method.

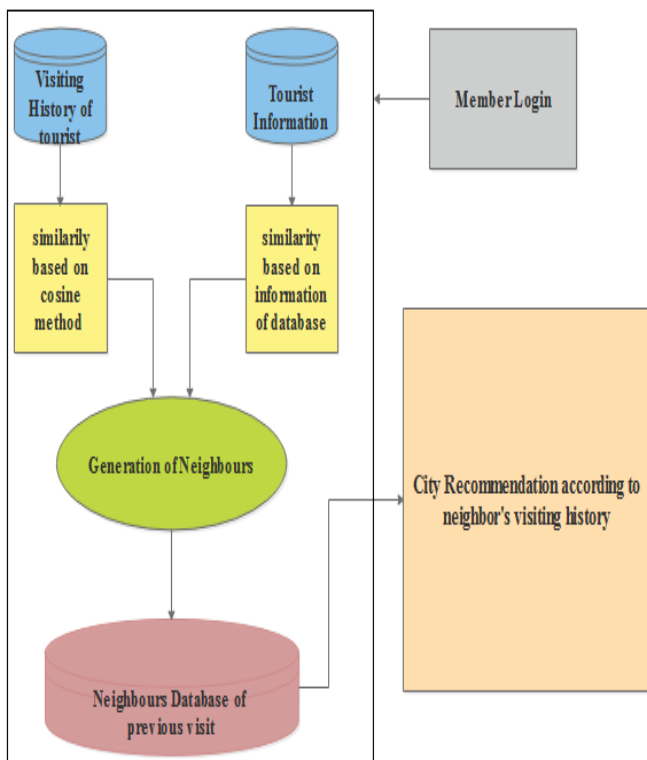


Fig: Overall Process

### III. Development of Recommender System

#### 1. Recommendation Process

On the basis of collaborative filtering principle, the recommendation process of tourist attractions can be divided into three steps;

1) The representation of user (tourist) information. The visiting history of attractions by tourist need to be analyzed and modeled.

2) The generation of neighbor users (tourists). The similarity of tourists can be computed according to the visiting history data and the collaborative filtering algorithm presented by us. A neighbor tourist list can be calculated on the basis of known similarities.

3) The generation of attraction recommendations. Top-N attractions will recommended to the tourist according to the visiting history of his neighbors.

According to above steps, user's basic information and past travel history can be used to calculate the user list of neighbors which are recorded in the corresponding record in the user database. When users log into the system, the recommendations of tourist attractions can be presented based on the travel experiences of his neighbors.

The process of recommendation is shown in Fig.1.

In order to increase the efficiency of the system, the system calculated users' neighbors' offline for each user, and calculated at regular intervals.

#### 2. Generation of Neighbors

Neighbor users generated mainly based on the similarity between each user. However, a larger number of users have become null and void, so the system needs to pre-process the user data before generating neighbors. The pre-processing process consists of data cleaning, data integration, data conversion and data reduction. The system then can get a more streamlined and representative list of users which are usually often logged on and recently recorded users, which are recognized as active users. For users already have travel records, the system can use the visiting history of tourist attractions to calculate the similarities between users. For the user there is no travel records, the system can use the user's basic information (such as logging times, sex, professional, vehicle, etc.) as the basis for the calculation of similarity.

Suppose that the set of all tourists  $T=\{T_1, T_2..T_n\}$ , for each tourist  $T_i$  ( $i=1, 2.. n$ ), the system can calculate the neighbors list including the top N tourists which similarity is higher than the given threshold\_.

The core of collaborative filtering algorithms [13] is the calculation of users' similarity. The similarity between the tourist  $T_i$  and  $T_j$  is measured as follows: First the system needs to retrieval all the ratings of

tourist  $T_i$  and  $T_j$  over the attractions, and then calculate the similarity between the tourist  $T_i$  and  $T_j$  through different similarity models, denoted as  $sim(T_i, T_j)$ . There are mainly three ways to measure the similarity between tourists, including Cosine method, Correlation similarity method and Adjusted Cosine method [14]. We have adopted the Cosine method in this paper. Tourist ratings are viewed as a vector of n-dimensional term space. If the tourist has ever visited the attraction, the rating is set to 1, otherwise set to 0. And the cosine similarity between tourists is computed by the angle between vectors. The similarity between tourist  $T_i$  and  $T_j$  can be calculated by:

$$sim(T_i, T_j) = \frac{R_{T_i} \bullet R_{T_j}}{\|R_{T_i}\| \times \|R_{T_j}\|} \quad (1)$$

1. Neighbor users generated mainly based on the similarity between each user.
2. Suppose that the set of all tourists  $T=\{T_1, T_2...T_n\}$ , for each tourist  $T_i$  ( $i=1, 2... n$ ), and  $C=\{C_1, C_2, C_3, C_4, C_5\}$  are the cities, Then the system can calculate the neighbors list including the top N tourists which similarity is higher than the given threshold.
3. There are mainly three ways to measure the similarity between tourists, including Cosine method, Correlation similarity method and Adjusted Cosine method.

| Tourists | Attractions |    |    |    |    |
|----------|-------------|----|----|----|----|
|          | C1          | C2 | C3 | C4 | C5 |
| T1       | 1           | 0  | 1  | 1  | 1  |
| T2       | 0           | 1  | 1  | 1  | 1  |
| T3       | 1           | 0  | 1  | 1  | 0  |
| T4       | 0           | 1  | 1  | 0  | 1  |
| T5       | 0           | 1  | 0  | 1  | 0  |

Table I: Attraction Visiting History of All Tourists.

Suppose that  $S_i$  is the set that composed of all the attractions which the tourist  $T_i$  has ever visited. So, we can give

out the calculation formula of similarity between  $T_i$  and  $T_j$ .

$$sim(T_i, T_j) = \frac{|S_i \cup S_j|}{|S_i \cap S_j|} \quad (2)$$

Based on (2) and Table 1, we can calculate the similarity between  $T_1$  and  $T_2$ ,  $T_1$  and  $T_3$ ,  $T_1$  and  $T_4$ ,  $T_1$  and  $T_5$  as follows

$$sim(T_1, T_2) = \frac{|S_1 \cup S_2|}{|S_1 \cap S_2|} \quad (3)$$

$$= \frac{3}{5} = 0.6$$

$$sim(T_1, T_3) = \frac{|S_1 \cup S_3|}{|S_1 \cap S_3|} \quad (4)$$

$$= \frac{3}{5} = 0.6$$

$$sim(T_1, T_4) = \frac{|S_1 \cup S_4|}{|S_1 \cap S_4|} \quad (5)$$

$$= \frac{2}{5} = 0.5$$

$$sim(T_1, T_5) = \frac{|S_1 \cup S_5|}{|S_1 \cap S_5|} \quad (6)$$

$$= \frac{1}{5} = 0.2$$

If the value of threshold  $\_$  is set to be 0.5, then the neighbors of  $T_1$  are  $T_2$  and  $T_3$ .

When computing recommendations for a particular tourist, the very basic approach is to select the objects favored by other tourists that are similar to the target tourist. When new tourists enter the system, there is usually insufficient information to produce recommendation for them, because there is no visiting history of the new tourists. The usual solution of the cold start problem is similarity calculation between each user by profile information, such as age, sex, professional, vehicle, income, etc.

### 3. Generation of Recommendations



Recommendations of attractions are computed by the visiting times of neighbors. According to the calculation above, we know that the neighbors of tourist T1 are T2 and T3, so we can list all the visiting history of all the attractions so as to summary the most popular ones. As listed in Table 2, we can find that the maximal visiting times of neighbors are attraction C3 and attraction C4.

| Tourists | Attractions |    |    |    |    |
|----------|-------------|----|----|----|----|
|          | C1          | C2 | C3 | C4 | C5 |
| T1       | 1           | 0  | 1  | 1  | 1  |
| T2       | 0           | 1  | 1  | 1  | 1  |
| T3       | 1           | 0  | 1  | 1  | 0  |
| Total    | 2           | 1  | 3  | 3  | 2  |

TABLE II: Recommended cities based on Visiting History

## Conclusion

Recommender systems have grown as an area of both research and practice. A personal tourist attractions recommender system is considered as the effective way for tourist to tourist attractions search. The recommender system compares the collected data to similar and dissimilar data collected from others and calculates a list of recommended attractions for the tourist. Collaborative filtering is considered to be memory-based and model based collaborative filtering. So as well known example of memory-based approaches is user-based algorithm that we adopt in this paper. Research work of this paper is to explore the recommendation of tourist attractions, and a lot of further work needs to be optimized on the basis of this paper. As a future enhancement we can extent this application for the online payment and visitor query.

## ACKNOWLEDGEMENTS

I would like to extend my gratitude to a person who helped me to bring this paper fruition. I would like to thank Dr Sharath Kumar Y H. I am so deeply grateful for his help, professionalism, and valuable guidance throughout this paper. This accomplishment would not have been possible without my guide. Thank you.

## References

1. F.Anishya<sup>1</sup>, M.Suresh kumar "Discovering Real Time Travel Package Recommendation Using Collaborative Tagging". *International Journal of scientific research and management (IJSRM)* || Volume || 3 || Issue || 1 || Pages || 2023-2028 || 2015 ||
2. VincentW. Zheng, Bin Cao<sup>1</sup>, Yu Zheng, Xing Xie and Qiang Yang "Collaborative Filtering Meets Mobile Recommendation: A User-centered Approach", *Association for the Advancement of Artificial Intelligence (www.aaai.org)* 2010.
3. Parnika Patil, Vaishali Kolhe. "An Approach for Travel Package Recommendation using Clustering and Typicality based Approach." *International Journal of Innovative Research in Computer and Communication Engineering*.
4. Monali Gandhi, Khushali Mistry, Mukesh Patel. "A Modified Approach towards Tourism Recommendation System with Collaborative Filtering and Association Rule Mining." *International Journal of Computer Applications (0975 – 8887)* Volume 91 – No.6, April 2014.
5. Masoumeh Mohammadnezhad and Mehregan Mahdavi "An Effective Model For Improving The Quality Of Recommender Systems In Mobile E –Tourism." *International Journal of Computer Science & Information Technology (IJSIT)* Vol 4, No 1, Feb 2012.
6. Gediminas Adomavicius And Alexander Tuzhilin "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions." *IEEE Transactions On Knowledge And Data Engineering*, VOL. 17, NO. 6, JUNE 2005.
7. Mr. Hrishikesh Teke, Mr. Parag Borle, Mr. Mayuresh Tiwari, Mr. Samadhan Kangale, Mr. Avinash Dandgule. "An Intelligent Recommendation System To Enhance The Travel Services" *IJARIIIE-ISSN(O)-2395-4396 Vol-1 Issue-4* 2015.
8. Xinhuan Chen, Yong Zhang, Pengfei Ma, Chao Li, Chunxiao Xing "A Package Generation and Recommendation Framework based on Travelogues." 2015 *IEEE 39th Annual International Computers, Software & Applications Conference*.
9. Alexander Smirnov, Alexey Kashevnik, Andrew Ponomarev, Maksim Shchekotov, Kirill Kulakov. "Application for e-Tourism: Intelligent Mobile Tourist Guide." 2015 *IIAI*



*4th International Congress on Advanced Applied Informatics.*

10. Liu Hai-ling, Li lun-huai, Peng lun" A Novel Recommendation System for the Personalized Smart Tourism" *2015 IEEE 14th Int'l Conl. on Cognitive Inlormatics & Cognitive Computing.*
11. Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl "Item-based Collaborative Filtering Recommendation Algorithms".
12. Alexander Smirnov, Alexey Kashevnik, Andrew Ponomarev, Nikolay Shilov, Maksim Schekotov, Nikolay Teslya "Recommendation System for Tourist Attraction Information Service" *proceeding of the 14th conference of fruct association*