ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)

User-Based Collaborative Filtering System for Tourist Attraction Recommendations

[I]Alisab B Havaragi, [II]Prathibha B S
[I]PG Student, The National Institute of Engineering, Mysuru, India
[II]Assistant Professor, ISE Dept., The National Institute of Engineering, Mysuru, India

Abstract

User-based tourist attraction recommender system is developed in this paper. The recommender system is constructed as an online application which is capable of generating a personalized list of preference attractions for the tourist. Modern technologies of classical recommender system, such as collaborative filtering are considered to be effectively adopted in the tourism domain. On the basis of collaborative filtering principle, the recommendation process of tourist attractions divided into three steps, representation of user (tourist) information, generation of neighbor users (tourists) and the generation of attraction recommendations. In order to calculate the similarities between each user, the Cosine method is adopted during the process of the generation of neighbors. And then the recommendations of attractions are generated according to the visiting history of the user's neighbors. In order to demonstrate the calculation process of the system, a case is demonstrated in detail. In case if user does not have any visiting history or new member here the recommendation process takes place based on the User Area of Interest (AOI).

Keywords

Tourist Attraction Recommendation; User-Based Recommendation; Recommender 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.

Most of the tourism industries have a huge tourist's potential, unfortunately, too little valued and exploited. As a result, one of the strategic developments of the economy aimed the tourism industry. But strategies are based on different trends obtained from sophisticated analysis of data. Providing the managers in the tourism industry with information about and insight into the existing data is the key function of the data warehouse systems.

Data mining - techniques for exploration and analysis of large quantities of data in order to discover meaningful patterns and rules - helps businesses sift through layers of seemingly unrelated data for meaningful relationships, where they can anticipate, rather than simply react to, environment challenges.

A system which enables the use of data mining techniques on data stored in a data warehouse is ideal for high quality analyzes to support strategic decision. But designing and implementing a Data Warehouse is a complex and expensive process, so we can apply the data mining algorithms on large volumes of data from relational databases.

The aim of this is to present the opportunity to use data mining methods on data from tourism.

II. Literature Servey

Lots of technique is carried out some of discussed here.

A. A Study of Mixture Models for Collaborative Filtering [1]

The rapid growth of information over Internet demands intelligent information agents that can sift through all the available information and find out the most valuable to us. These intelligent systems can be categorized into two classes: Collaborative Filtering (CF) systems and Content-based Filtering (CBF) systems. The difference between them is that collaborative filtering systems utilize the given ratings of training users to make recommendation for test users while content-based filtering systems rely on contents of items for recommendation.

Most collaborative filtering methods fall into two categories: Memory-based algorithms and Model-based algorithms. In memory-based algorithms, rating examples of different users are simply stored in a training database, and the rating of a test user on a specific item is predicted based on the corresponding ratings of training users who share similar tastes as the test user. In contrast, in model-based algorithms, statistical models are learned from the given ratings of training users and ratings of test users are estimated using the learned model.

B. A collaborative filtering framework based on both local user similarity and global user similarity [2]

Collaborative filtering is a classical method of information retrieval widely used in helping people to deal with information overload. Generally, there are two major classes of collaborative filtering algorithms, memory-based algorithms and model-based algorithms. Because of their simplicity and robustness, memory-based algorithms are widely applied in practice, To estimate a prediction for a particular user (i.e., an active user), the memory-based algorithms first find users from the database that are most similar to this active user, and then combine those ratings together. The measurement techniques of the similarity between users include the Pearson Correlation Coefficient Vector Space Similarity (VSS) algorithm and the extended generalized vector space model these algorithms can be considered as user-based algorithms.

However in practice, systems based on collaborative filtering algorithms often face the problem of having at their disposal only an insufficient amount of preferences ratings of their individual users. Therefore, one of the biggest challenges of designing a collaborative filtering system is how to provide accurate recommendations with the sparse user profile data. To estimate an active user's rating of a particular item, traditional user-based methods first find the user's neighbors (the users who are similar to the active user). Then, the active user's rating is predicted by averaging the (weighted) known ratings on the item by his/her neighbors. This kind of methods is based on the assumption that similar users have similar rating patterns. Unfortunately, due to the data sparsity problem, firstly, often there does neither exist a sufficient amount of similar neighbors, nor a sufficient amount of ratings of the particular item.

The measurement of the similarity between users plays a fundamental role in user-based algorithms. Traditional methods of computing similarity, however, have two important shortcomings. Firstly, usually all items are treated the same when computing the similarity of users. This is addressed by), which assign different weights to items in order to allow for items to contribute in different strength to the user similarity calculation. The second problem is that the similarity of two users cannot be calculated if they have not rated any identical item. In other words, due to the data sparsity problem, the neighbors of active user cannot be found. To solve this problem, it seems promising to transitively examine whether the neighbors of the two users are similar. That means we should estimate similarities between any two users from a global perspective

C. Using location for personalized POI recommendations in mobile environments [3]

Internet-based recommender systems have traditionally employed collaborative filtering techniques to deliver relevant "digital" results to users. In the mobile Internet however, recommendations typically involve "physical" entities (e.g., restaurants), requiring additional user effort for fulfillment. Thus, in addition to the inherent requirements of high scalability and low latency, we must also take into account a "convenience" metric in making recommendations.

III. Proposed System

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. 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.

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.

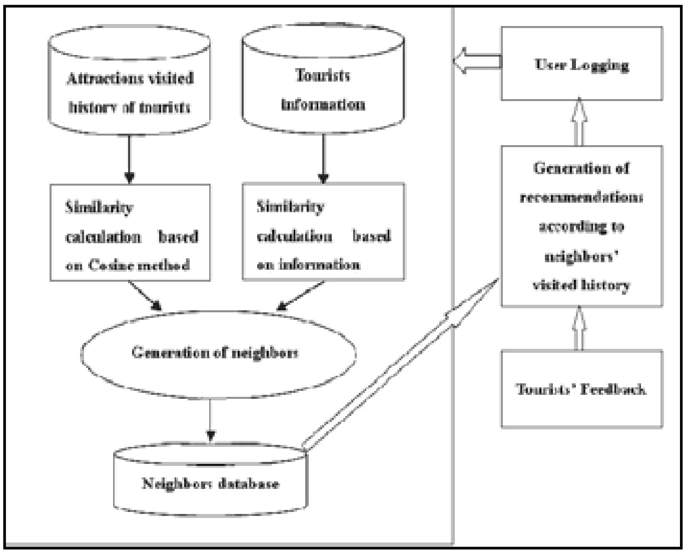

Fig 1: Generation process of attraction recommendations.

A. 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 preprocessing 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.

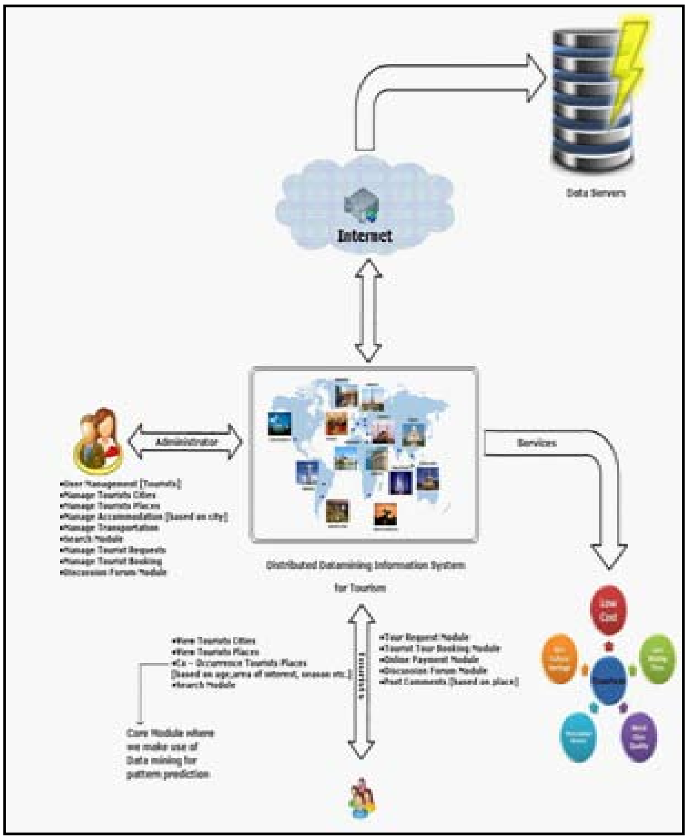

Fig 2 : Proposed System View

ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)

IV. Data Flow Diagram

A data flow diagram (DFD) is a graphical representation of the "flow" of data through an information system. DFDs can also be used for the visualization of data processing (structured design).

Data flow diagram of Admin:
Here the admin can login into the system by using His/her Id and Password. Here Admin can able to add Cities, Tourist Places, Accommodation, Transportation cost, Discussion Forum and FAQ's .Here admin have authority to Add or remove the cities or tourist places. Figure 4 shows DFD of Member.

Data flow diagram of Member:
Tourists can be login into the system by using their ID's and Passwords. Here member can able to see the tourist places by selecting a particular city, here this system can recommend the tourist places based on his/her visiting history. If the member is new then the recommendation is based on his/her AOI (Area of Interest) on the recommendation of tourist places member can book. Figure 5 shows DFD of Member.

Data flow diagram for Visitor:
Here visitor has limited access visitor only can able to view some basic functions which is located on the main page such as about us, contact us,FAQ's and Registration option to become member.
.

Figure 6 shows DFD of Visitor.

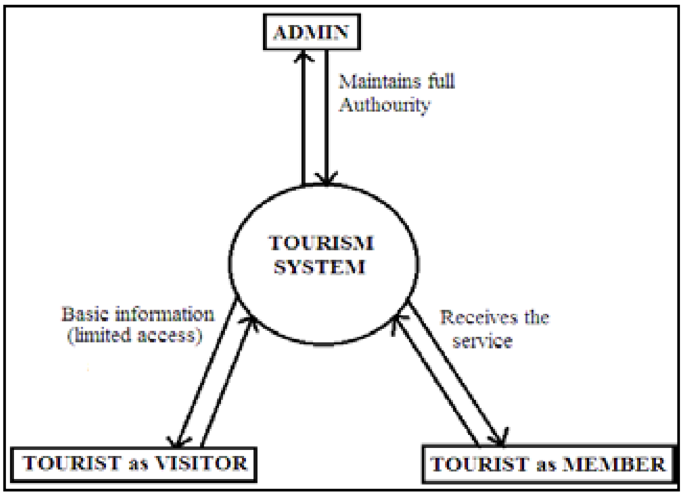

Fig 3 : Context data flow diagram

It is common practice to draw a context-level data flow diagram first, which shows the interaction between the system and external agents which act as data sources and data sinks. On the context diagram (also known as the 'Level 0 DFD') the system's interactions with the outside world are modeled purely in terms of data flows across the system boundary. The context diagram shows the entire system as a single process, and gives no clues as to its internal organization

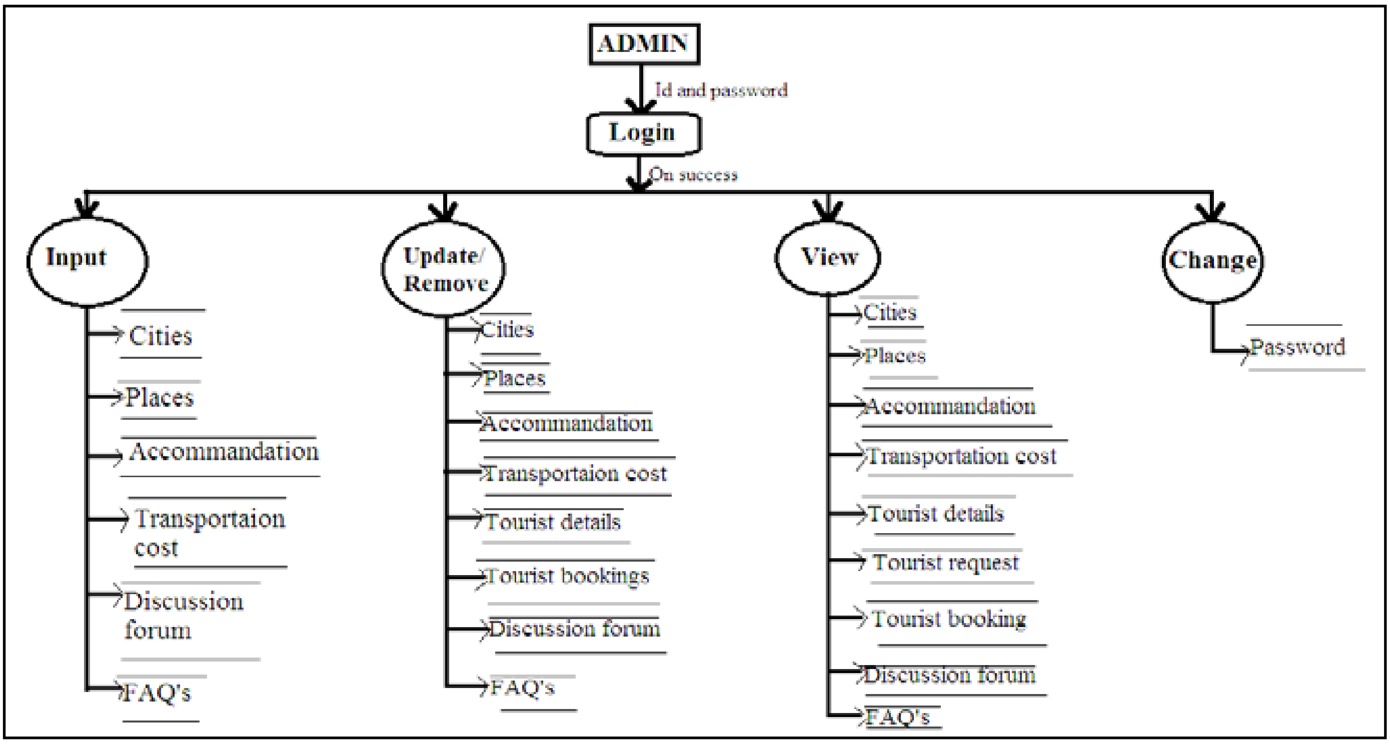

Fig 4: Data Flow Diagram (Level 1) – Admin

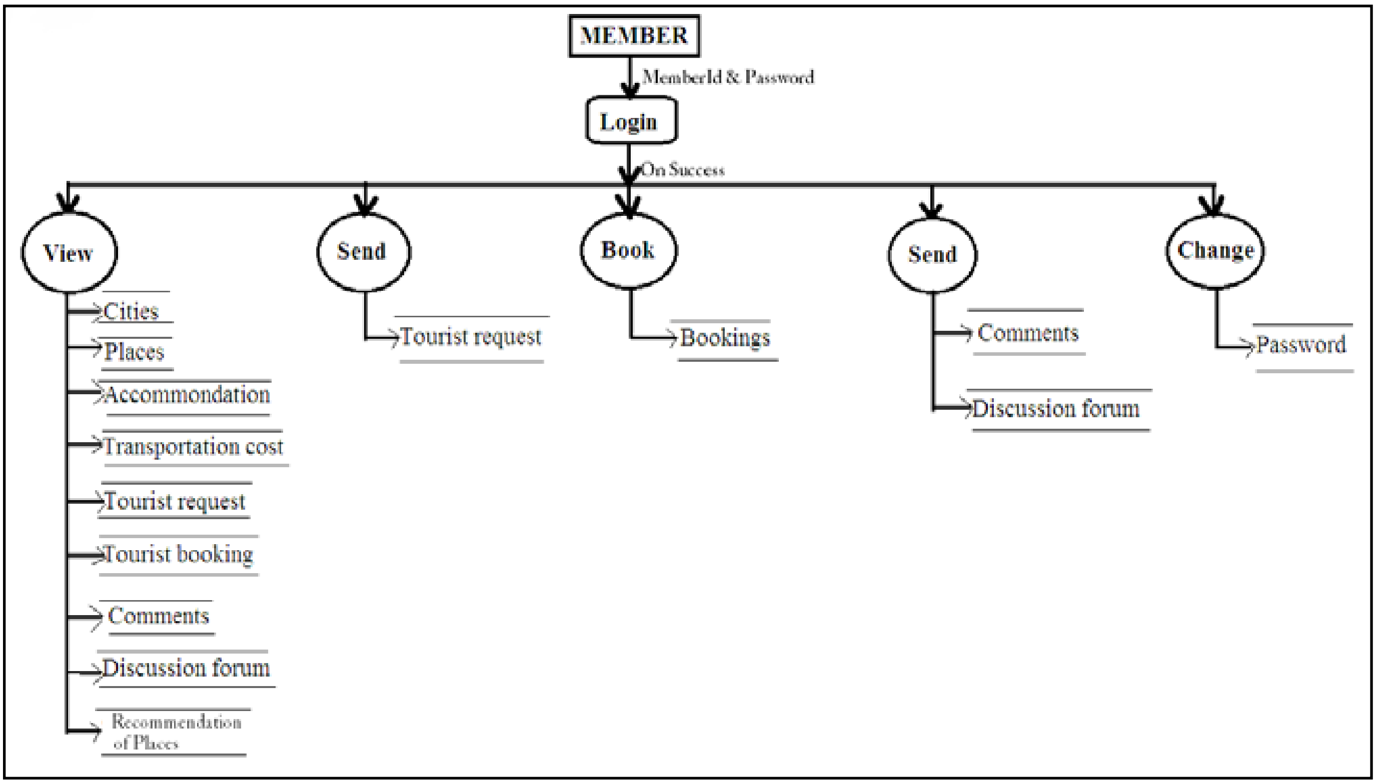

Fig 5: Data Flow Diagram (Level 1) – Member

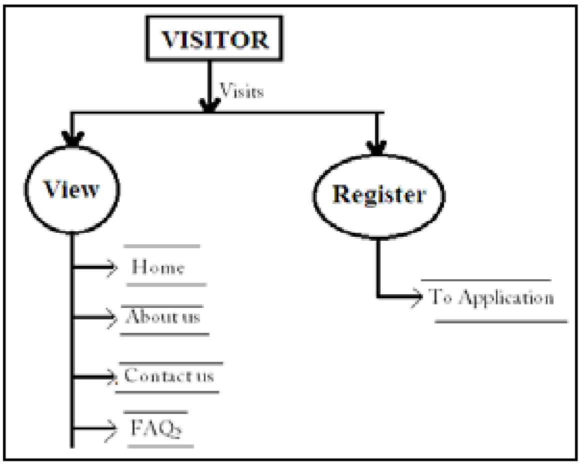

Fig 6: Data Flow Diagram (Level 1) – Visitor

V. Conclusion and Future Enhancements

In our proposed system "User-Based Collaborative Filtering System for Tourist Attraction Recommendations ". Tourist can easily book tourist places based on his/her interest or also book the tourist places based on recommendation which is generated based on his/her visiting history or on AOI (Area of Interest).

We can add up the Online Payment Module as a future enhancement to the application where tourists can pay online. We can add up the Feedback module as a future enhancement to the application where tourists can send the feedbacks regarding the website services to the administrator. We can add up the visitor query module as a future enhancement to the application where new visitors to the application can post the queries to the administrator. And we can add up user can get booking detail through the E-mail or Message to his\her cell.

References

[1] *Rong Jin, Luo Si, and Chengxiang Zhai, "A study of mixture models for collaborative filtering," Information Retrieval, vol. 9, pp. 357 382, June 2006*

[2] *Heng Luo, Changyong Niu, Ruimin Shen, Carsten Ullrich, "A collaborative filtering framework based on both local user similarity and global user similarity," Machine Learning, vol. 72, pp. 231-245, September 2008.*

[3] *Horozov, Tzvetan, Nitya Narasimhan, and Venu Vasudevan, "Using location for personalized POI recommendations in mobile environments," Proc. International Symposium on Applications and the Internet (SAINT 2006), IEEE Press, January 2006, pp. 129, doi: 10.1109/SAINT.2006.55*

[4] *Anhai Doan, Raghu Ramakrishna, and Alon Y, Halevy, "Crowdsourcing systems on the World-Wide Web," Communications of the ACM, vol. 54, pp. 86-96, April 2011*

[5] *Hsiu-Sen Chiang, and Tien-Chi Huang, "User-adapted travel planning system for personalized schedule recommendation," Information Fusion, vol. 21, pp. 3-17, January 2015.*

[6] *Lu, EH-C., Chih-Yuan Lin, and Vincent S. Tseng, "Trip-mine: An efficient trip planning approach with travel time constraints," Proc. 12th IEEE International Conference on Mobile Data Management (MDM 2011), IEEE Press, June 2011, pp. 152-161, doi:10.1109/MDM.2011.13.*

[7] *G. D. Abowd, C. G. Atkeson, J. Hong, S. Long, R. Kooper and M.Pinkerton, "Cyberguide: A Mobile Context Aware Tour Guide," Wireless Networks, vol. 3, pp. 421-433, October 1997*

[8] *Y. Zheng, X. Xie, "Learning Travel Recommendations from UserGenerated GPS Traces," ACM Transactions on Intelligent Systems and Technologies, vol. 2, article 2, January 2011.*

[9] *Haifeng Liu, Zheng Hu, Ahmad Mian, Hui Tian, and Xuzhen Zhu, "A new user similarity model to improve the accuracy of collaborative filtering," Knowledge-Based Systems, vol. 56, pp. 156-166, January 2014.*

ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)

[10] Michael D. Ekstrand, John T. Riedl, and Joseph A Konstan. "Collaborative Filtering Recommender Systems," Foundations and Trends in Human-Computer Interaction, vol. 4, pp. 81-173, February 2011.

[11] Arthur J. Lin, Chien-Lung Hsu, and Eldon Y. Li, "Improving the effectiveness of experiential decisions by recommendation systems," Expert Systems with Applications, vol. 41, pp. 4904-4914, August 2014.

[12] Cane Wing-ki Leung, Stephen Chi-fai Chan, and Fu-lai Chung, "An empirical study of a cross-level association rule mining approach to cold-start recommendations," Knowledge-Based Systems, 21(7), pp. 515-529, October 2008.

[13] Zhi-Dan Zhao and Ming-Sheng Shang, "User-Based CollaborativeFiltering Recommendation Algorithms on Hadoop," Proc. Third International Conference on Knowledge Discovery and Data Mining (WKDD 10), IEEE Press, January 2010, pp 478-481, doi: 10.1109/WKDD.2010.54.

[14] Alan Said, Brijnesh J. Jain, and Sahin Albayra, "Analyzing weighting schemes in collaborative filtering: cold start, post cold start and power users," Proc. 27th Annual ACM Symposium on Applied Computing (SAC 12), March 2012, pp. 2035-2040, doi: 10.1145/2245276.2232114.

[15] Danielle H. Lee and Peter Brusilovsky, "Reinforcing Recommendation Using Implicit Negative Feedback," Lecture Notes in Computer Science Volume, vol. 5535, pp. 422-427, 2009.