

A Case Study on Various Recommendation Systems

Athulya R. Krishnan

P G Scholar

Department of Computer Science & Engineering
College of Engineering, Perumon(CUSAT), Kerala, India

Remya R.

Assistant Professor in CSE

Department of Information Technology
College of Engineering, Perumon(CUSAT), Kerala, India

ABSTRACT

The goal of a recommender system is to generate relevant recommendations for users. It is an information filtering technique that assists users by filtering the redundant and unwanted data from a data chunk and delivers relevant information to the users. An information system is known as recommendation engine when the delivered information comes in the form of suggestions. Since different users have different interests, the information filtering system must be personalized to accommodate the individual users interests. This requires gathering of feedbacks from the user in order to make a user profile of his preferences. Recommender systems have become extremely common in the recent years, and are applied in a variety of fields. Usually recommender systems are based on the keyword search which allows the efficient scanning of very large document collections. Recommender systems typically produce a list of recommendations through collaborative or content-based filtering techniques. Different recommendation systems are available for different scenarios. In this paper, a detailed review of various recommendation systems is presented..

General Terms

Recommendation, Text mining

Keywords

Collaboration, Content based, Information filtering

1. INTRODUCTION

Information is the most powerful weapon in the modern society. Every day we are overloaded with a huge amount of data in the forms of electronic newspaper articles, emails, web pages and search results etc. But we often obtain incomplete data. Usually we need to do further search activities to have the correct acquisition of information. The importance of recommender systems comes to light in this scenario. The use of dimensionality reduction is to improve the performance for a new class of data analysis software called recommender systems. It deals with the detection and delivery of information that the user is likely to find interesting or useful. It assists users by filtering the data source and deliver relevant information to the users. Recommender systems have evolved from the extremely interactive environment of the Web. One typical application of recommendation systems is to help customers find which products they would like to purchase at E-Commerce sites. In general, every recommendation system follows a specific

process to produce product recommendations. It is depicted in Fig .1.

The authenticity of the information depends upon the interest of users. The recommendation system must be personalized to accommodate an individual users interest. These systems have achieved widespread success in the E- commerce field nowadays. Suggestions for books on Amazon, or movies on Netix, are real world examples of the operation of industry-strength recommender systems. For instance, a recommender system on Amazon.com suggests books to its customers based on other books the customers have told Amazon they are interested in. Another example for recommender system is CDnow, which helps the customers choose CDs to purchase, based on other CDs the recipient has liked in the past.

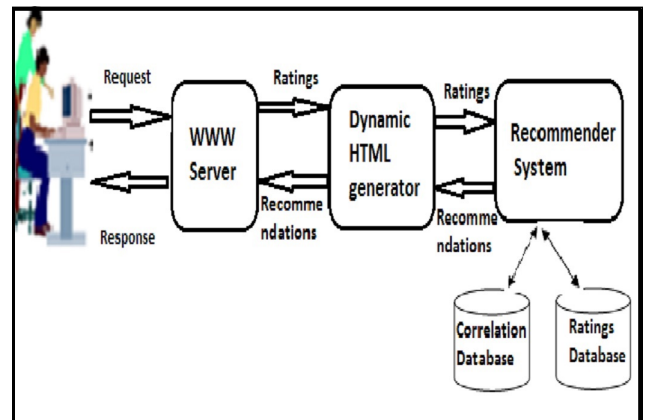


Fig. 1. Architecture of Recommender systems

There are mainly two approaches for the information filtering in recommendation systems: Content-based filtering and collaborative filtering. The Content-based filtering recommends items based on a comparison between the content of the items and a user profile. The content of each item is represented as a set of descriptors or terms, typically the words that occur in a document. The user profile is represented with the same terms and built by analyzing the content of items which have been seen by the user. They are independent on the presence of other users in the system. It also referred as cognitive filtering. Collaborative filtering or social filtering is a prominent approach used by some recommender systems

such as large, commercial ecommerce sites. Collaborative-based filtering systems can produce recommendations by computing the similarity between a users preferences and the preferences of other people. Such recommendation systems do not attempt to analyze or understand the content of the items being recommended.

2. LITERATURE SURVEY: EXISTING RECOMMENDATION SYSTEMS

There are several types of recommender systems available which are helpful in various scenarios. This section reviews some of the existing recommendation systems in detail.

2.1 Amazon.com Recommendation System

The amazon.com recommendation system [1] uses item to item collaborative technique that can be used for Ecommerce websites. They use recommendation algorithms to personalize the online store for each customer. The suggestions must depend upon the customers behavior. Here the suggestions are built by the conversation rating and the click-through process. The traditional collaborative filtering techniques compare customers based on their previous purchases and provide recommendations to similar customers. Few customers who are most similar to the user are found out. Suppose that there are two customers A and B. The similarity between these customers can be calculated by using cosine similarity (eq 1).

$$similarity(\vec{A}, \vec{B}) = \cos(\vec{A}, \vec{B}) = \frac{\vec{A} \cdot \vec{B}}{\|\vec{A}\| * \|\vec{B}\|} \quad (1)$$

In this system they use item to item collaborative technique to personalize its Web site to each customers interests. The most similar match for a given item is determined, and the proposed algorithm builds a similar table by finding items that customers tend to purchase together. They built a product to product matrix by iterating all item pairs and compute the similarity metric for each pair. Based on the calculated matrix, most appropriate suggestions are provided to the customers.



Fig. 2. The Your Recommendations feature on the Amazon.com website

The home page of amazon.com have a feature Your recommendations. By clicking on the Your Recommendations link, the customer will be taken to an area where he can see the product recommendations and rating of previous purchases as shown in Fig. 2. The system is scalable over very large customer bases and product catalogs. The main issue with this method is that it cannot provide suggestions to new items.

2.2 Remembrance Agent System

The Remembrance Agent [2],[3] is a software which augments human memory by displaying a list of documents which might be relevant to the users current context. It runs without user intervention. It continuously monitors the user activities and identifies the information needs. For continuous monitoring, it generates the explicit queries from words that are written or spoken by the user. Based on the explicit query generated, relevant information is extracted as suggestions. The suggestions are presented in the form of a one line summary to the users. The system has both front and back ends. The front end is a UNIX text editor which displays the one line suggestion along with ratings to the users for their local activity. When the user forwards a request, suggestions provide full text of suggested documents. The back-end is a program which produces suggestions of similar documents from a pool of documents when it is provided with a query.

2.3 Watson just-in-time-retrieval-system

Jay Budzik et al., [4] propose a Watson just-in-time-retrieval-system is a recommender system that assists the users by finding the relevant documents while browsing web or writing. The Watson just in system is efficient than the remembrance agent system. Because it taking the advantage of structure of written text in addition to word frequency. The web browsers provide rich contextual information that can be leveraged to task relevant information. It emphasizing the words with large fonts and words that specified in headings. The Watson system uses Information Management Assistant System (IMA) observes users interaction with everyday applications and anticipate their information needs using a task model of hand. IMAs then automatically fulfill these needs using the text of the document the user is manipulating.

The IMA is implemented by using several application adapters, which gain access to an applications internal representation of a document. The adapters produce a document representation. The documents are represented in different styles like normal, emphasized, de-emphasized or list item. Watson transforms the original document representation into a query and selects appropriate sources. This query takes the form of an internal query representation, which is then used to select information adapters. The information adapter translates the query into source specific language and performs the search. In parallel to this task, Watson ventures to detect conceptually atomic, lexically regular structures in the document. The generated query is ranked by using a term weighting algorithm. Based on the formulated query, the relevant documents are retrieved and the search engine results are clustered using an incremental algorithm to avoid the redundancy of pages.

2.4 Collaborative tourist information retrieval system

The tourist information retrieval system uses collaborative filtering technique [5][6]. Collaborative filters predict someone's personal preferences for information and/or products by keeping track of their likes and dislikes, and then connecting that information with a database of other peoples' preferences to check for matches, and to make predictions. The study reports that the CIR in tourism is still in its early stages. It is evident that the Information needs of tourists are characterized by highly dynamic and individual factors, such as preferences for sites, weather conditions, prices, transportation, and economical changes. The Internet has been steadily increasing as an important information source for visitors in recent years. In particular, the participants responses to the questions such as whether

they need and/or want to collaborate when searching on the Web are important for tourism information. Web searching is a process of querying and reformulating queries to satisfy certain information needs.

In Query reformulation, web searchers frequently modify their queries to obtain better results. Collaborative Tourism Information Search supports the searching for travel-related information in both standalone (a single user) mode and collaborative mode (multiple users). They proposed a system that uses remotely-located collaboration technique and the participants in the team could communicate with each other by sending instant text messages. The collaborative querying (CQ) formulation architecture (Fig.3) was drawn to illustrate how a CQ is initiated, reformulated and executed. In this scenario a query is initiated by any of the users. The query is visible to all other users and it can be modified by other users. At the end of modification, the query is executed to retrieve relevant information. If the results are not satisfactory, then the user can redefine the query. The query reformulation continues until a satisfactory result is obtained for the tourist.

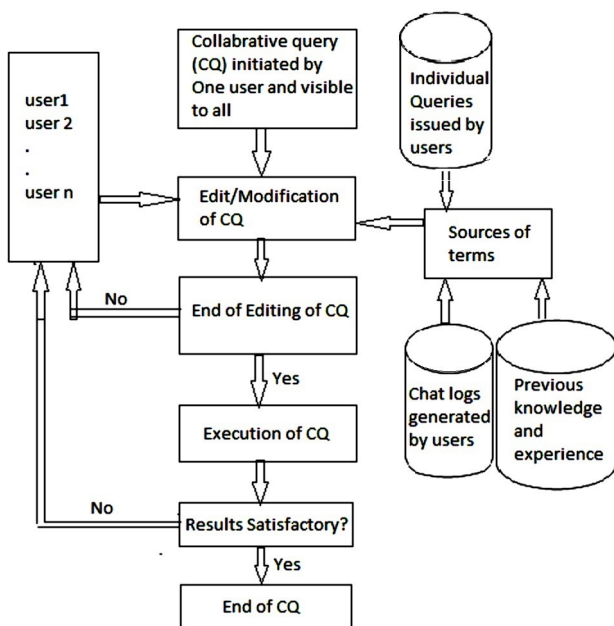


Fig. 3. Flow of collaborative query reformulation.

2.5 Ada and Grace: Direct Interaction with Museum Visitors

David Traum et al.,[7][8] presents Ada and Grace, a twin virtual guide that are used in the Museum of Science, Boston, to interact directly with museum visitors as shown in the Fig.4. The quality of a museum visitors experience depends upon a well informed guide or interpreter. Virtual guides designed to engage visitors in an interactive and increase their knowledge and promote excitement about museum content are used. In order to provide efficient recommendations to the visitors, large amounts of data was aggregated so as to promote the efficient identification and recognition of the questions that are asked frequently by the visitors.



Fig. 4. Ada&Grace:Guides at the Museum of Science, Boston

The virtual guide interacts with visitors using natural language input and produce output rather than the traditional menu driven approach because it makes the interface more user friendly. To interact with the system, an operator presses a push-to-talk button and speaks into a microphone. An audio acquisition client sends these captured audio into automatic speech recognition (ASR) module. The ASR module convert the audio into text and is sent to the Language Understanding (LU) module to understand the language. Dialogues are then analyzed by the DM module which later processes the responses to be forwarded to the user. The main inconvenience of this system is that it cannot provide recommendations for a new question, which is not in the domain specific library

2.6 Implicit Query System

Susan Dumais et al.,[9][10] suggest Implicit Query (IQ) prototype is a system which automatically generates context-sensitive searches based on a users current computing activities. This method demonstrates an IQ system running during reading or composing email. The system analyzes the email message and delivers the results to users. The important words are extracted using TF-IDF weights. TF-IDF stands for term frequency-inverse document frequency, is a numerical statistic that is intended to reflect how important a word is to a document in a collection or corpus. A sidebar contained in the current email reflects the search results as suggestions and users can adjust their task focus and use constructively peripherally presented suggestions.

3. CHALLENGES AND ISSUES

In real world the existing recommendation systems facing several issues. One of the main issue faced by the system is that cold start which is a problem related to recommendations for novel users or new items. In case of new users, the system does not have information about their preferences in order to make recommendations. So it becomes difficult for the system to recommend any suggestions. Another problem is that scalability. Because the number of users and items are increasing day by day so much that more resources are required for processing those information to get a perfect and efficient recommendation system. Sometimes the reviews and rates given by user who rarely uses his profile is irrelevant comparing to the profile having a great history. Privacy is the prime factor for all the systems, to give the user the most accurate recommendation, the system has to use the history of the user.

4. CONCLUSION

As more and more information became available electronically, people seek suggestions or recommendation from friends and experts who have knowledge about their needs. Due to the overload of information on the World Wide Web, the necessity of recommender systems to generate efficient solutions have evolved. In such circumstances the need for effective information retrieval and implementation of filtering tools have become essential for easy access of relevant information. Recommender Systems (RS) are software tools and techniques that providing suggestions for items to be of use to a user. In this paper, we discussed different approaches of recommender systems. Almost all the key techniques for recommender systems have been surveyed. Retrieval of information from huge volumes of data is a tedious process so in order to make the recommendation process easier, agent mediated recommender systems have evolved nowadays.

5. REFERENCES

- [1] Greg Linden, Brent Smith, and Jeremy York, "Amazon.com Recommendations :Item-to-Item Collaborative Filtering IEEE Computer Society, 2008.
- [2] B. Rhodes and T. Starner, "Remembrance Agent: A continuously running automated information retrieval system, in Proc. 1st Int. Conf. Pract. Applicat. Intell. Agents Multi Agent Technol., London, U.K., 1996, pp. 487495.
- [3] B. J. Rhodes and P. Maes, Just-in-time information retrieval agents, IBM Syst. J., vol. 39, no. 3.4, pp. 685704, 2000.
- [4] J. Budzik and K. J. Hammond, User interactions with everyday applications as context for just-in-time information access, in Proc. 5th Int. Conf. Intell. User Interfaces (IUI00), 2000, pp. 4451.
- [5] A. S. M. Arif, J. T. Du, and I. Lee, Towards a model of collaborative information retrieval in tourism, in Proc. 4th Int. Interact. Context Symp., 2012, pp. 258261.
- [6] A. S. M. Arif, J. T. Du, and I. Lee, Examining collaborative query reformulation: A case of travel information searching, in Proc. 37th Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval, 2014, pp. 875878.
- [7] D. Traum, P. Aggarwal, R. Artstein, S. Foutz, J. Gerten, A. Katsamanis, A. Leuski, D. Noren, and W. Swartout, Ada and Grace: Direct interaction with museum visitors, in Proc. 12th Int. Conf. Intell. Virtual Agents, 2012, pp. 245251.
- [8] David Traum, William Swartout, Ada and Grace: Toward Realistic and Engaging Virtual Museum Guides, Springer-Verlag Berlin Heidelberg 2010.
- [9] S. Dumais, E. Cutrell, R. Sarin, and E. Horvitz, Implicit queries (IQ) for contextualized search, in Proc. 27th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval, 2004, pp. 594594.
- [10] M. Czerwinski, S. Dumais, G. Robertson, S. Dziadosz, S. Tiernan, and M. Van Dantzich, Visualizing implicit queries for information management and retrieval, in Proc. SIGCHI Conf. Human Factors Comput. Syst. (CHI), 1999, pp. 560567.