Recommender System Methods and Feedback Mechanisms: A Survey

Megharani Misal, Uma Nagaraj

M.E. Student, Dept. of Computer Engineering., MIT Academy of Engineering., Pune University, India

Professor& Head, Dept. of Computer Engineering., MIT Academy of Engineering., Pune University, India

ABSTRACT: In recent years Recommender Systems(RS) has gained popularity to solve the problems of web information overload on web as an information recovering and classification techniques RS help users in finding the contents of their interest with minimum efforts by filtering the information available on the web. Various methods used to recommend the contents such as Collaborative filtering (CF), Content-based filtering, and Hybrid methods. We present in this paper all these methods and limitations of it in brief. RS uses two types of feedback mechanism based on explicit rating and implicit rating. Implicit feedback mechanism is more convenient for users who do not like to rate the contents explicitly. Therefore goal of this paper is to study RS types (methods) and implicit feedback mechanism which is based on user interaction to improve RSs performance.

KEYWORDS: Recommender Systems; Collaborative Filtering;

I. INTRODUCTION

RSs become very popular since the appearance of first paper on CF in the mid of 1990s [1].RSs are used in various application areas such as Electronic Commerce, Online Social Networks (Facebook, You-Tube, LinkedIn etc). RSs are mainly designed to handle information overload and provide personalized recommendation of contents and services to users across the internet. There are various techniques or methods used to filter a large amount of information on the web and recommend the information in which user is interested. Collaborative filtering which considers the similarity among the users   to recommend the contents. While Content-based methods consider the user as well as item profiles. To overcome the limitations of both methods hybrid methods are used .Over the time recommendation methods have improved to consider the different parameters to improve recommendation process such as contextual information [5], multicriteria ratings, user behaviour.

Although RSs are performing good still there is scope of improvement in the case of information feedback mechanism. All present systems predict contents based explicit rating provided by user .But in many applications of RS using explicit rating for the recommendation purpose is inconvenient because it may affect the user performance as user has to rate contents to get recommendations .To overcome this limitation implicit feedback approach is developed based on user interaction to the application.

Main objective of this paper are

1. To Study in brief recommendation methods (Collaborative filtering, Content-based filtering and Hybrid methods).
2. Studying Feedback mechanisms of RS.
3. Proposing a general architecture for RS application based on implicit feedback mechanism.

II. RELATED WORK

A. *Recommender Systems*

Since 1990s recommender systems were considered major research area after CF was introduced. RSs are widely used in many application areas for various purposes such as commercial or experimental or scientific. For example,

Fab: Content-based collaborative recommendations [6], Amazon.com recommendations: item-to-item collaborative filtering [7], Hybrid news recommendation system [9] that use contextual information to generate recommendations. RSs are classified into different types according to information used to recommend 1) Collaborative filtering: Which calculate the similarity between the users.

CF further classified into two types [2] such as Memory-based and Model-based approach. Memory-based approach uses similarity between user profiles and item profiles. While Model-based approach use the hidden users or items characteristics 2) Content-based: Generate recommendations based on the contents that to another user liked in the past.3) Hybrid Approach: It combine the characteristics of both approaches to overcome the limitations of both.

B. *Information Feedback Mechanisms*

To improve the performance of recommendation process requires efficient feedback mechanism. Now a day's almost all RSs using explicit feedback mechanism based on explicit rating which is collected through survey process or rating given by users explicitly. For example In Facebook 'like' rating used or star ratings used by online shopping portals like Amazon, Flipkart etc. However this may be inconvenient to users, as they typically do not like to rate contents. Implicit feedback which consists of evaluating the objects without intervention of users. Information is captured from the actions made by the users in the application. Such information is used to generate the ratings without intervention of user. These techniques use user behavior to understand user interests and preferences .The use of this feedback technique helps to improve the users experience and satisfaction .Implicit feedback techniques are attractive candidate to improve the information feedback mechanism as there is not required a further effort from the user .In [3] [4] presented how to collect the information form the user interaction and convert them into explicit rating.

III. RECOMMENDER SYSTEM AND ITS METHODS

A. *Recommender System*

Recommendation problems explicitly rely on rating systems.RS need to calculate rating for unrated items and then recommend this item(s) to user with highest estimated rating. Recommendation problem can be calculated as Let C is the set of all users and S is the set of items that can be recommended such as books, movies, mobiles, news etc. The space S of items can be very large in some applications as well as user space can also be very large. Let u be the utility function which measures the utility of particular item s'$\in$S To user c. i.e. u: C*S->R. where R is total ordered set .Then for each c$\in$C user we can choose item s'$\in$S which maximizes the user utility .

In RSs usually rating is used to represent the items utility. For example some user gave a book named "alchemist" rating of 4(out of 5).Each user of user space can be defined by using profile which consider various user characteristics   such as age,gender,marital status, occupation. Similarly each item in the item space can be represented using item profile. To generate ratings on unrated items we usually use heuristic which define the utility function and then estimating utility function to optimize the performance criterion. Once unknown ratings are generated then item with highest rating is recommended to the user. New ratings can be calculated by using different methods of approximation theory, machine learning and various heuristics.RSs are classified into following types based on the information used to make recommendations.

B. *Methods of Recommender System*

a) **Content-Based Methods**: In content-based methods the utility of item s for user c is calculated based on the utilities $u(c,s_i)$ assigned by the user c to items $s_i \in S$ that are similar to item s. For example in book recommender system to recommend the books to user c the Content-based RS try to find the similarity among the books user c has highest rated in the past. Then the book that has high degree of the similarity to user preferences would be recommended to user. Most Content-based RS use the techniques of information retrieval and information filtering. Due to the significant improvement of the information retrieval and filtering techniques Content-based

systems focus on the textual information such as documents, web sites for recommending items. Further RSs use the user profile which contains the information about the user. Such information can be collected explicitly or implicitly. Similarly item profile can also be maintained to determine the item for recommendation process. As content-based RSs are based on the text contents in the system can be represented as keywords. In the Fab System [6] which recommends the web pages to user. Such a system uses "100" important keywords to represent the web pages .Therefore Content-based RS uses (TF-IDF) measure to find important words in the document. Here we considered there are N documents that can be recommended to the user and keyword k_j appears in one of the n_i. TF can be calculated as follows:

$$TF_{i,j} = f_{i,j} / \max_z f_{z,j} , \tag{1}$$

Where,

$f_{i,j}$ is a frequency of the keyword k_i in the document d_j.

$TF_{i,j}$ is a term frequency of the keyword k in the document d_j.

$\max_z f_{z,j}$ is a maximum of frequencies $f_{z,j}$ of all keywords k_z in the document d_j.

To find the relevant and non-relevant documents more accurately inverse term frequency is used

$$ITF_i = \log N / n_i \tag{2}$$

TF-IDF weight for keyword k_i in document d_j is given by

$$w_{i,j} = TF_{i,j} * ITF_I \tag{3}$$

And content of document d_j is defined as follows:

Content $(d_j) = (w_{i,j} \ldots . w_{k,j})$.

Content-Based RS recommend items to users based on users past likings. So utility function can also be defined using ContentBasedProfile(c) of user as follows:

$$u(c,s) = score(ContentBasedProfile(c), Content(s)). \tag{4}$$

Content-Based RSs can use various techniques such as Bayesian classifiers and various machine learning techniques as clustering, Decision tree and Artificial neural networks. These techniques are different from traditional heuristic techniques based on information retrieval in terms of methods used. Unlike the heuristic methods this techniques are based on the modal learned by using statistical and machine learning methods.

Limitations

- **Limited Content Analysis:**
  Content-Based systems have limitations in the explicit features associated with object. By using manual or automatic extraction methods we can get sufficient amount of features. But these methods also have limitations in terms of extracting the multimedia data. Another problem related to limited content analysis is that it is indistinguishable if two items are represented by same set of resources.

- **Overspecialization**

  Problem of overspecialization is that content-based systems can not recommend the items that are different from the items that are seen by the users previously. It is also important that recommender system should not

recommend item that are very similar to previously seen items. Therefore good recommender system should constitute the diversity feature.

- **New User Problems**

To recommend efficiently, there should be the large set of ratings. As new users get very few ratings therefore it is difficult to get accurate recommendations.

b) **Collaborative Filtering Methods:** In CF based RS recommends the items for the users based on the previously rated items by the other users.More formally utility u(c,s) of item s for user c is calculated based on the utilities $u(c_j,s)$ assigned to the item s by users cj∈C who are similar to user for example in the book recommender system collaborative recommender system try to find out users that have similar interest(i.e., same rating given to similar books).And books that are liked by other similar user is recommended. Generally CF classified into two classes as follows:

- **Memory-Based CF**

The memory Based approach is also known as neighbourhood-based techniques, a subset of users are chosen based on their similarity between the user profiles, and a weighted combination of their ratings is used to make predictions for this user.

The value of the unknown rating$r_{c,s}$ for user c and item s is aggregate of N most similar   users for same item s.

$$r_{c,s}=aggr_{c'\in C}\, r_{c',s} \tag{5}$$

Such predictions are commonly calculated through a Pearson coefficient or a Variant thereof. We can also use item-specific memory-based approach which is calculated by considering the similarity between the item features.

- **Model-Based CF**

Model-based techniques provide recommendations by calculating parameters of statistical models for user ratings. Unlike neighbourhood based methods that generate recommendations based on statistical notions of similarity between users, or between items. Model based approach uses hidden features unlike memory based approach using the statistics between users or items. Model based approach using various techniques such as Bayesian networks, Clustering, Artificial neural networks, Linear regression, Probabilistic models are used to generate recommendations. Recently Matrix factorization [10] and Latent analysis techniques are also used in model based approach. On the basis of empirical study it is shown that model-based approach outperforms than the memory-based approach in terms of performance.

Limitations

- **Cold Start Problem**
  This problem comes because nobody has rated contents previously either explicitly or implicitly. Therefore there are very few ratings available to generate accurate recommendations.

- **Sparsity**
Sparsity problem is that there is insufficient data to identify the similar   users.

- **New Item Problem**
  This problem comes because new item is not yet rated therefore it is not considered while making the recommendations.

- **Popularity Bias**
  This problem states that system will often recommend the popular items or highly rated items. Users with unique taste will have problem to get items of their interest.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015

c) **Hybrid Techniques**: To overcome the limitations of both collaborative and content-based techniques several recommendation systems use hybrid approach. This approach combines the both collaborative and content-based methods to predict recommendations. There are various ways to combine these two approaches such as

- Implementing collaborative and content-based techniques separately and finally combining the results of both to get correct recommendations of contents.
- Incorporating some of the features or characteristics of collaborative filtering into content-based filtering.
- Incorporating some of the features or characteristics of content-based filtering into collaborative filtering.
- Building a one unified model by incorporating the characteristics of both approaches.

IV. RSs FEEDBACK MECHANISMS

The recommender systems collect user information by using the feedback techniques. User information is stored in the user profiles which show the users interest which is useful in making recommendations. There are two types of feedback techniques: Explicit and Implicit feedback techniques [8].

- **Explicit feedback**
  To evaluate the system user assigns some value or rating to some objects or a set of objects by using survey process. Explicit feedback technique is used to explicitly state the interest of user in particular system or object. There are various rating systems are used to rate the contents. For example ,Amazon online store, Film affinity use the star rating system, Online social networks such as Facebook,YouTube,LinkedIn use the like rating system.

- **Implicit feedback**
  This process consist of the evaluating the system without the user being aware. In this process information is captured through the users actions that are performed by user in the application. Then by using this information users behavior is analyzed to find the users interest. It is not required to rate the contents explicitly therefore this approach improves the user satisfaction and experience while searching contents.

  Recently, most of the implemented recommender systems are based on the explicit ratings. But it is inconvenient as users do not like to rate contents .Implicit approach do not require any effort from user side therefore it seem attractive approach. In this paper we study general architecture for the application using recommender system based on the implicit feedback.

V. PROPOSED SYSTEM

a) *Architecture of RS using implicit feedback*
  In [5], defined a set of implicit parameters on which comparative analysis performed to find the correlation between the user actions and explicit ratings given by user on the contents. The results of comparative analysis showed that it is possible to determine user interest through the analysis of its behaviour. By considering these results RS can be implemented with this architecture that contains algorithm to transform implicit behaviour into explicit ratings. Fig.1 shows the high level architecture of the RS application using implicit feedback. This architecture has three main levels.

- **Client Applications(User interface)**
  At this level graphical user interaction interface is designed where users interact with the application and perform the desired actions. The main objective of this module is to discover and share the contents within the internet community.

- **Feedback and Explicitation system**
   It is intermediate level composed of feedback and explicitation system.

  Feedback System: It obtains the information from users .This module presents an application that enables to register the actions performed by users. The main aim of doing this is that to analyze the behavior of user from user actions. To gather the quantity of information during users interactions, one web platform is developed. This will allows collecting

and storing users actions in an implicit way through web services. With this implementation, when a user performs an action from the web, e.g. highlight, remark, add comments, etc., a controller is invocated and it executes the necessary logic.

Explicitation System: Explicitation system performs analysis and conversion of implicit actions into explicit values. In order to analyze and evaluate the different users behaviour according to their actions performed with the platform, a User Interactions Converter Algorithm (UICA) is developed. This algorithm evaluates implicit actions previously performed on the platform using a set of procedures to convert these values in explicit ratings. These ratings are generated by considering some rating system that indicates users interest. To implement the explicitation system need to develop an application that obtains the implicit data and converts it into explicit ratings by using the implementation of a converter algorithm (UICA). As seen in [4] users use various rating systems such as star rating system, like rating system to show the users interest.
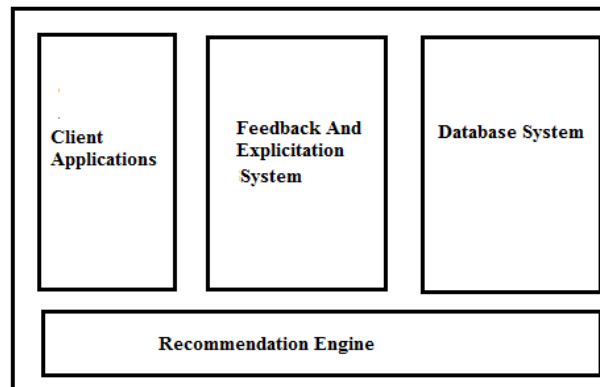

Fig. 1.High Level Architecture

- **Database systems**
    Third level of this architecture is the database systems used. Database   systems are formed by three databases.

    Implicit Feedback Database: Stores all the information resulting from users actions with the application.

    Explicit Feedback Database: Stores data obtained from data explicitation processes.

    Operative Database: Stores operative data from the web applications as well as stores the data generated by recommender engine.

    Recommender engine use the data processed and analyzed to generate recommendations based on user profiles.

VI. CONCLUSION

In this paper we have studied the recommender system and its methods such as collaborative, content-based and hybrid methods with their limitations. We also studied the feedback mechanisms which are key area of improvement in recommender process. RSs using implicit feedback approach are more convenient than those who are using explicit one. We proposed here the general high level architecture for RS using implicit feedback.

REFERENCES

1. Adomavicius, G., &Tuzhilin, A. "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions". IEEE Trans. Knowl. Data Eng., 17(6), 734–749,2005.
2. P. Melville and V. Sindhwani, "Recommender systems". Encyclopedia of machine learning, pp. 829–838, 2010.

3. González Crespo, R., SanjuánMartínez, O., CuevaLovelle, J. M., PelayoGarcía- Bustelo, B. C., Gayo, J. E. L., & de Pablo, P. O. "Recommendation System based on user interaction data applied to intelligent electronic books". Computers in Human Behavior, 27(4), 1445–1449,2011.

4. Nuñez-Valdéz, E. R., CuevaLovelle, J. M., SanjuánMartínez, O., García-Díaz, V., ordoñez de Pablos, P., & Montenegro Marín, C. E." Implicit feedback techniques on recommender systems applied to electronic books". Computers in Human Behavior,2012.

5. Verbert, K., Manouselis, N., Ochoa, X., Wolpers, M., Drachsler, H., Bosnic, I., et al. "Context-aware recommender systems for learning: A survey and future challenges". IEEE Transactions on Learning Technologies, 6(1), 2007,2012.

6. Balabanovic´ , M., &Shoham, Y.." Fab: Content-based, collaborative recommendation". Communications of the ACM, 40(3), 66–72,1997.

7. Linden, G., Smith, B., & York, J. "Amazon..com recommendations: Item-to item collaborative filtering". Internet Computing, IEEE, 7, 76–80,2003.

8. Adomavicius, G., Sankaranarayanan, R., Sen, S., &Tuzhilin, A." Incorporating contextual information in recommender systems using a multidimensional approach". ACM Transactions on Information Systems (TOIS), 23(1), 103–145,2005.

9. Montes-García, A., Álvarez-Rodríguez, J. M., Labra-Gayo, J. E., &Martínez-Merino, M. "Towards a journalist-based news recommendation system: The Wesomender approach". Expert Systems with Applications,40(17),67356741,2013.

10. Z. Cheng and N. Hurley, "Robust collaborative recommendation by least trimmed squares matrix factorization" .in Proc. 22nd IEEE Int. Conf. ICTAI, pp. 105–112,2010.

BIOGRAPHY

MegharaniVitthalMisalis a Graduate Student in the Computer Engineering Department, College of MIT Academy Of Engineering, Pune University. She is pursuing her Master (ME) degree from MITAOE, Alandi (D), Pune, State Maharashtra, India. Her research interest is in Data Mining.