

A Survey on Personalized Service Recommendation Systems

Devdatta Godbole¹

Department of Information Technology
Walchand College of Engineering
Sangli, India

Manish Narnaware²

Department of Information Technology
Walchand College of Engineering
Sangli, India

Abstract—Service recommendation systems have been proved to be very efficient and valuable tools for giving appropriate recommendations to users in order to improve the user experience. recommendation systems have become extremely common in recent years. These systems are applied in a variety of applications such as music, books, movies, research articles, search queries, news, social tags and many more. In the last decade the amount of customers and services as well as online information has grown fast, resulting the big data analysis problem for the service recommendation systems on a big scale. recommendation systems normally generate a list of recommendations in any of these ways either collaborative or content-based filtering methods. Collaborative filtering approaches build a model from a user's past behavior as well as from similar decisions made by other similar users having similar interests. Such a model is then used to predict items that the user may be interested in. Content-based filtering approaches use a series of different attributes of an item in order to recommend more product items having similar set of properties associated with them. These approaches are often combined to form a hybrid recommendation system. Each type of system has its own pros and cons. This paper gives application survey and research scope to inspire researcher to solve service recommendation problems and help to choose a better algorithm strategy for recommendation depending upon the requirement and input sets of the system.

Keywords—Recommendation system, Personalization, Big data, Machine Learning, Hadoop

I. INTRODUCTION

Recommendation systems have become an important area of research nowadays. Much work has been done both in the industrial domain and academic domain for development of new approaches to recommendation systems over the last years [1],[2]. The number of people interested in this area is high because it is a problem rich and difficult yet challenging research area. Also because of the availability of tons of practical applications that help users to find the way out of information overload and also provide personalized service recommendations of services to them. Some examples of such applications include recommending gadgets, books, mobiles and other products at Amazon [1]. Many of the online shopping vendors also have integrated recommendation algorithms into their services [3]. Despite of all of these advances the recommendation systems of this generation still require many improvements to make recommendation techniques more efficient which can be applied to an even broader range of real life applications [2]. Such improvements include better and more efficient methods for representing user behavior [4]. Also they represent the information about

the items to be recommended precisely. Some advanced recommendation user modeling methods [5],[6], incorporate of various contextual information into the recommendation process. They provide better utilization of multiple criteria rating systems [5],[19], and develop less intrusive and more efficient and generic recommendation methods. These methods also rely on measures that more effectively determine performance of recommendation systems.

The aim of a Recommendation System is to generate a set of meaningful recommendations for a set of users regarding products or some items in which they have interests. The design of recommendation systems depends on the domain of the system and the particular characteristics of the data available [2]. For example, people watching movies on Netflix frequently provide ratings on a scale of 5 for example (3 / 5), etc. This dataset source keeps track of the quality of interactions between users and movies. Moreover, the system may have access to user-based [9] and item-based [7] profile attributes such as demographics and review or movie / product descriptions. recommendation systems are differentiated on the basis of the way they analyze these data sources to develop rules of relativity and closeness between users of services and product items which can be used to identify well matched pairs.

Content-based filtering [6] provides recommendations which considers the profile of the user and the item description as the features. Each item-content is represented as a set of keywords or terms. Normally these are the words that occur in a document or dataset. The user profile is represented with the same terms. The profile is then built up by making analysis of the item-content which has been seen by the user.

There any issues that are to be considered while implementing a content-based filtering system [5],[6]. First issue, terms can either be assigned in an automatic way or manual way. When terms are assigned automatically, a method or technique has to be chosen that can extract these terms from items. And the second issue, the terms have to be represented in such a way that both the user profiles and items can be compared in an efficient way [6]. Another issue is that a ML algorithm has to be chosen that is can analyze the user profile based on seen items or products and then can generate recommendations based on their profile.

The information source or data source that content based filtering systems mostly use are text documents. The standard approach for term parsing selects single words from the documents. Two methods, the vector space model [16] and latent semantic indexing use these terms to represent documents as vectors in a multiple dimensional space.

Genetic algorithms, Bayesian classifiers [17], Decision trees, and neural networks are some machine learning techniques for learning a user profile and extracting knowledge. A Vector Space Model can be used by these learning methods for representing documents [16]. Some learning methods also represent the user profile in one or more vectors belonging to same space which makes it very easy to compare documents and user profiles. Other learning methods such as the Bayesian classifier [17] and neural networks represent the user profile in their own way.

Collaborative filtering [7],[9],[18], performs filtering of the information by making use of recommendations of other similar users. It is based on the concept that the people who chose certain items in the past are likely to choose similar items again in the future. A person who wants to watch a movie for example, may ask for reviews and recommendations from his/her friends. The recommendations from friends having similar interests are more countable and trustworthy than those from others. Thus, information generated is used to make decision of which movie to watch.

In item-based CF recommendation systems, [7],[8] prediction of the possible rating is done on the basis of the ratings of other similar items by the same user. In user-based CF recommendation systems [9], the prediction of the rating of an item for a given user is done depending upon the ratings for the same item rated by different similar users.

II. RELATED WORK

There are three main categories in which recommendation systems are generally classified based upon their approach : Content based recommendation systems [6], Collaborative Filtering based recommendation systems [7],[9],[18], and hybrid approach [4],[18]. Also there are many more attributes based on which the recommendations can be generated. Paper [6] uses provides recommendations using user modeling. Paper [8],[9] implement collaborative filtering algorithms out of which [8] implements item-based collaborative filtering while [9] implements user-based collaborative filtering. Reference [10] provides suggestions based on community sharing activities. Paper [11] predicts services based upon user's consumption. Usage history is considered as an attribute to provide recommendations is used in paper [13]. Paper [14] provides recommendations based upon the location. Bayesian inference is considered in paper [17] which is as good as collaborative filtering. User reviews and hybrid approach is considered in paper [18] to give recommendations. Now we will check in detail these different approaches and under which conditions they are preferable over the others.

III. LITERATURE REVIEW

There are various techniques/algorithms to implement recommendation system and find the similarities between items. These techniques and algorithms are discussed in detail one by one next.

A. Content-based filtering

Content based filtering generates recommendations items based on a comparison between the item-content and user profile. Contents of each item are represented as a set of keywords or terms using words that occur in description or document. The user profile is represented with the same terms.

The recommendations are then provided based upon the items viewed by the user previously in the past.

M. Bjelica [6] analyzed and experimented recommendation system design to implement user modeling algorithm that efficiently learns the user interests. This learning strategy analyses the user behavior to find the patterns. By applying some efficient information retrieval algorithms, it is possible to build a simple yet robust system that provides helpful recommendations to its users.

B. Collaborative filtering

Collaborative filtering technique performs filtering of the information by using recommendations from other users who have similar interests. It is based on the idea that the people with similar interests, who bought certain items in the past are likely to buy similar items again in the future. Collaborative filtering has two types of filtering that is item-based and user-based.

Item based collaborative filtering can be used efficiently for systems where we want to recommend similar items for the same user. Jin et al. [8] presented a recommendation system which uses an item-based hybrid approach by combining collaborative filtering and content based methods with MapReduce. They present a large scale video recommendation system with MapReduce using item based collaborative filtering technique. This technique uses the user views of videos as dataset. The technique also performs analysis of the metadata associated with each video and extracts the feature information. These features are used to find the similarity between the videos.

User based collaborative filtering can be used to efficiently predict the ratings of an item to a user depending upon the ratings of the same item rated by similar users. Z. Zhao and M. Shang [9] presented in the paper a user based collaborative filtering recommendation system on a cloud computing platform. They have designed a user based CF algorithm using the MapReduce technique and have implemented the algorithm on the Hadoop platform.

Also different strategies can be implemented for providing efficient recommendations to users based on different approaches such as community [10], user history [13], consumption [11], location [14], review [18], user modeling [6], Bayesian Inference [17], etc.

C. Community Contribution based recommendation

Y. Chen, A. Cheng and W. Hsu [10] presented a system that is based on the community contribution done by the people. The system performs analysis of the community contributed that is socially shared photos of people and uses this information to derive attributes and travel group information. This information is later used to provide travel recommendations. This system provides personalized travel recommendation system by analyzing specific user profiles or attributes such as gender, occupation, nationality, age as well as travel group types such as family, colleagues, couple, friends. By using the photos and the comments and tags, the system interprets user features and travel group patterns. According to information retrieval theories, user profiles prove to be very informative and effective for travel recommendation specially while providing a promising personalization. This system considers that these attributes not

only provide a basic information which is beneficial for providing personalized recommendation but also explain the human behavior and human activities along with their associated contexts, e.g. location, season and time. The system considers landmarks of locations as examples to demonstrate the interesting findings from mining from community contributed and socially shared photos.

D. User Consumption based recommendation

F. Sánchez, M. Alduán, F. Álvarez, J. Menéndez and O. Báez [11] propose a recommendation system which is based on audio and visual consumption of users. This system provides sport video recommendations which considers the audio and visual consumption of users. When these sport videos are transmitted on the Internet the transmission is done as large scale events. The audio-visual consumption of the individual users is considered as a feature to provide recommendations. This technique is independent of the number of users. This feature collection is executed on client side. This avoids the computation issues and privacy issues. The study case considers large scenarios where huge number of users is associated for example Olympics. The base of this approach is based on the concept that each user has different approach and thus wants different things for many different reasons. The recommendation system takes into consideration multiple factors or attributes that can make a user interested in a particular video. The factors which are considered to characterize the interest of the users and extract the knowledge out of the system are : Athletes, their Nationality, Round and Sports. Each of these attributes has different behavior and nature so they cannot be modeled in a homogenized way. By making use of these attributes as feature values for each video, the system finds patterns and relevance between the user audio-visual consumption and their prediction accuracy. The final rating prediction per video for every user is generated based upon the combination of this relevance and the accuracy associated with it.

E. Usage History based recommendation

In the web application context, an application usage history of a user can be a good feature attribute in order to extract out patterns and derive knowledge regarding the user. These patterns and knowledge can be used to provide recommendations with good accuracy. G. Kang, J. Liu, M. Tang, X. Liu, B. Cao, Y. Xu [13] presented such a system which implements web service recommendation based on usage history of the given user. The system generates a list in which rankings for publicly available web services are generated. The top 'n' ranked web services in the list are recommended to a user. Analysis of usage history log and the associated preferences is done to generate knowledge about user interests and preferences. Then the system uses the user interests and preferences and then transforms them into the users possible request. Here the user interests are considered as users functional requirements and preferences are considered as users non-functional requirements respectively. The system generates a table for every user which incorporates usage history, preferences and Quality of Service weight for the same. The system performs normalization of the documents for the misspelled and abbreviated words. These words are replaced by these normalized forms. This difference is eliminated using stemming algorithms on these words. The trivial information such as tokens, frequently occurring trivial words and stop words are eliminated using

these stemming algorithms. And this generated knowledge and preferences are then used to find patterns and provide recommendations.

F. Location based recommendation

One of the most efficient aspect to consider while providing recommendations is location. Location can play a key role for extracting knowledge out of system and finding patterns and deriving relations for a given scenario. This location attribute can be a key concern in cases where geographical boundaries are one of priorities. Some services may be available for a particular geographic region only or users belonging to only a particular geographic area might be interested in a particular service. Z. Luo, Y. Li, J. Yin [14] propose a system that considers this location attribute for giving recommendation to users. The location of a service represents its place in the network graph which determines the transmission cost for using a particular service. The amount of requests attempts invoked is inversely proportional to transmission time required. Thus the system derives that location can be a basic feature in selection of services and providing recommendation.

G. Bayesian Inference

X. Yang, Y. Guo and Y. Liu [17] propose a recommendation system which is based on Bayesian Inference for online social networks. In this system users share their ratings with their friends and contacts. The similarity between ratings of a pair of users having similar interests is predicted based upon the probability distribution of similar ratings and the inference based upon the ratings provided in the past. A user sends a query along with his rating for a service to users having similar interests over the social network. Based upon the response to this query, a Bayesian network is constructed. This interprets the rating of the querying user. Authors have developed an algorithm that provides a closer to accurate recommendation predicts almost accurate recommendation. Authors have also designed a set of protocols for this system. This Bayesian inference based approach provides recommendation results comparable with collaborative filtering and is better than trust based recommendations. The system implements a Bayesian network to generate a closer to accurate rating starting at the root of the network. A Bayesian network is constructed from recommendation classification tree. Thus, recommendations are provided based on the Bayesian Inference network for online social networks.

H. User Review based recommendation

Most of existing service recommendation systems are based upon the same ratings and rankings of services for different users. This technique does not consider user's diverse preferences. Therefore sometimes they fail to meet user's personalized requirements. S. Meng, W. Dou, X. Zhang and J. Chen [18] propose a system which provides recommendations based on users reviews and uses hybrid approach [4]. The system presents a personalized service recommendation list and recommends the most appropriate services to the user. In this system keywords are used for representation of both users preferences and Quality of Service. Previous users preferences are extracted out from their reviews for a given service and a keyword set is generated. Specialized domain thesauruses are developed and maintained to run stemming algorithms for keyword extraction in each of different service domains. In

this system each review is processed and formalized into a corresponding keyword set. This operation is performed based upon the keyword candidate list and the domain thesaurus. If the review contains exact or similar word from the domain thesaurus then the corresponding keyword is extracted into the preference keyword set of the user. To calculate the similarity between two sets, Jaccard similarity coefficient is used. The domain thesaurus is developed using hadoop mapreduce framework.

Now, a brief comparison is provided below between papers reviewed and the attributes they propose on the basis of which their respective systems provide recommendations.

IV. DISCUSSIONS

Recommendation system can be defined as system that produces individualized recommendations as output or has the ability of providing the user a personalized way to the services of his/her interests in a large space of possible options. Recommendation systems are used nowadays in almost every domain to provide efficient recommendations predictions. Various algorithms are available to implement recommendation systems. There are many attributes based on which a particular recommendation system provides recommendations. The algorithm chosen for implementing recommendations plays an important role in prediction and providing recommendations for a given domain. Bigger the input dataset user, the more complex is recommendation processing. The attributes to provide recommendations may change from system to system based upon behavior of the system, the input set, the output to be generated and the domain in which the recommendation system is to be implemented. For every recommendation system, the aspects and the perspectives of the attributes based on which the recommendations are to be generated change. Thus, before implementing the system all these criteria are to be considered to decide the attributes and algorithms for implementation.

We have considered and compared some of the attributes based on which various systems provide recommendations. These attributes are as follows

- 1) User modeling
- 2) Item-based collaborative filtering
- 3) User-based collaborative filtering
- 4) Community based modeling
- 5) User consumption
- 6) Usage History
- 7) Location
- 8) Bayesian inference
- 9) User Review

TABLE I. COMPARISON BETWEEN VARIOUS RECOMMENDATION SYSTEMS

Study Paper	Advantages	Disadvantages
Towards TV Recommender System: Experiments with User Modeling [6]	Better recommendation based on user interests Better knowledge collection for every user using modeling	Cannot give recommendation based upon interests of similar users Based on only the same users interests
MySpace Video Recommendation with	Provides better recommendation for	Does not consider similar interests.

Study Paper	Advantages	Disadvantages
Map-Reduce on Qizmt [8]	same item using interests of similar users	Complex to implement
User-based Collaborative Filtering Recommendation [9]	Better for finding interests for similar items Personalized recommendation	Doesn't consider similar users interests
Recommendation by Mining People Attributes Group Types From Community-Contributed Photos [10]	Community contribution is a good attribute to consider for providing recommendations Lot of information can be extracted using community contributed photos	Privacy issues can arise when community contributed photos are processed. Very complex to execute
Recommender System for Sport Videos Based on User Audiovisual Consumption [11]	Better for video, audio recommendation systems Does not consider similar users	Not suitable for recommendations other than audio or video Very smaller sample space for implementation
Web Service Recommendation Based on Usage History [13]	Better recommendations using usage history of the user Accurate interest can be modeled	May fail when user changes interest in that particular domain Overhead of maintaining history
Location: a feature for service selection in the era of big data [14]	Better for location specific services. Reduces transmission cost overhead	Not suitable where location is not an attribute of concern Lack of similarity calculation thus not suitable for bigger datasets
Bayesian-Inference-Based Recommendation in Online Social Networks [17]	Improved recommendation using Bayesian inference. Provides efficient output comparable with collaborative filtering systems	Complex to implement. Bayesian network users complex tree data structure which can be difficult to maintain.
KASR: A Keyword-Aware Service Recommendation Method on MapReduce for Big Data Applications [18]	Better recommendations are provided due to hybrid approach Considers reviews thus deals with large sample space. More personalization	Stemming algorithm overhead Does not consider negative reviews Lacks in semantic processing of review text.

Also, in order to implement an effective recommendation system, the following aspects play an important role in choosing the right way for providing recommendations.

- The domain of the system.
- Type of services provided by the systems.
- The group of users that is to be benefited from the system.
- Set of inputs that is to be processed.
- Scope and dependencies in the recommendations.
- Attributes based on which the system is to be implemented.

Thus, for developing a recommendation system, all of the above and many other aspects should be considered along with their strengths and weaknesses to evaluate the most efficient way to implement the system for the given set of requirements.

V. CONCLUSIONS AND FUTURE WORK

In this paper, we have discussed various ways and aspects of recommendation system implementation based upon different attributes. We discussed different recommendation systems, the aspects they covered and the attribute on which they are based upon. We have compared between these different systems to derive the areas in which a particular strategy can be applied. We also have discussed the strengths and weaknesses of different systems. Thus, we can now select the recommendation system that fits best for our requirements and works efficiently with our domain using these comparisons.

In future, various other attributes and techniques can be developed and evaluated for efficient implementation of recommendation systems. Also by combining recommendation systems with machine learning (ML) and natural language processing (NLP), we can develop powerful and efficient recommendation systems which will consider various aspects. Using machine learning, we can train the system to provide best recommendations based upon its past experiences. This will result in a very efficient recommendation system with has its own intelligence to predict the best interest of user and hence provide recommendations with high accuracy.

ACKNOWLEDGMENT

We express our sincere thanks to all the authors, whose papers or algorithms are published in the area of Recommendation Systems, Big data and Machine Learning in various conference proceedings and journals. We also express our sincere thanks to all the researchers who are working in the area of Machine Learning and Recommendation Systems. And we express our most sincere thanks to the authors whose papers and algorithms have been studied for preparing the review.

REFERENCES

- [1] G. Linden, B. Smith, and J. York, "Amazon.com Recommendations: Item-to-Item Collaborative Filtering," *IEEE Internet Computing*, vol. 7, no. 1, pp. 76-80, Jan. 2003.
- [2] G. Adomavicius and A. Tuzhilin, "Toward the Next Generation of Recommender Systems: A Survey of the State of the Art and Possible Extensions," *IEEE Trans. Knowledge and Data Eng.*, vol. 17, no. 6, pp. 734-749, June 2005.
- [3] P. Resnick, N. Iakovou, M. Sushak, P. Bergstrom, and J. Riedl, "GroupLens: An Open Architecture for Collaborative Filtering of Netnews," *Proc. ACM Conf. Computer Supported Cooperative Work (CSCW '94)*, pp. 175-186, 1994.
- [4] R. Burke, "Hybrid Recommender Systems: Survey and Experiments," *User Modeling and User-Adapted Interaction*, vol. 12, no. 4, pp. 331-370, 2002.
- [5] K. Lakiotaki, N.F. Matsatsinis, and A. Tsoukis, "Multi-Criteria User Modeling in Recommender Systems," *IEEE Intelligent Systems*, vol. 26, no. 2, pp. 64-76, Mar./Apr. 2011.
- [6] M. Bjelica, "Towards TV Recommender System Experiments with User Modeling," *IEEE Trans. Consumer Electronics*, vol. 56, no. 3, pp. 1763-1769, Aug. 2010.
- [7] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-Based Collaborative Filtering Recommendation Algorithms," *Proc. 10th Int'l Conf. World Wide Web*, pp. 285-295, 2001.
- [8] Y. Jin, M. Hu, H. Singh, D. Rule, M. Berlyant, and Z. Xie, "MySpace Video Recommendation with Map-Reduce on Qizmt," *Proc. IEEE Fourth Int'l Conf. Semantic Computing*, pp. 126-133, 2010.
- [9] Z.D. Zhao and M.S. Shang, "User-Based Collaborative-Filtering Recommendation Algorithms on Hadoop," *Proc. Third Int'l Workshop Knowledge Discovery and Data Mining*, pp. 478-481, 2010.
- [10] Y. Chen, A. Cheng, and W. Hsu, "Travel Recommendation by Mining People Attributes and Travel Group Types from Community-Contributed Photos," *IEEE Trans. Multimedia*, vol. 25, no. 6, pp. 1283-1295, Oct. 2013.
- [11] M. Alduan, F. Alvarez, J. Menendez, and O. Baez, "Recommender System for Sport Videos Based on User Audiovisual Consumption," *IEEE Trans. Multimedia*, vol. 14, no. 6, pp. 1546-1557, Dec. 2012.
- [12] A. Chu, R. Kalaba, and K. Spingarn, "A Comparison of two Methods for Determining the Weights of Belonging to Fuzzy Sets," *J. Optimization Theory and Applications*, vol. 27, no. 4, pp. 531-538, 1979.
- [13] G. Kang, J. Liu, M. Tang, X. Liu, and B. Cao, "AWSR: Active Web Service Recommendation Based on Usage History," *Proc. IEEE 19th Int'l Conf. Web Services (ICWS)*, pp. 186-193, 2012.
- [14] Z. Luo, Y. Li, and J. Yin, "Location: A Feature for Service Selection in the Era of Big Data," *Proc. IEEE 20th Int'l Conf. Web Service*, pp. 515-522, 2013.
- [15] B. Issac and W.J. Jap, "Implementing Spam Detection Using Bayesian and Porter Stemmer Keyword Stripping Approaches," *Proc. IEEE Region 10 Conf. (TENCON '09)*, pp. 1-5, 2009.
- [16] P. Castells, M. Fernandez, and D. Vallet, "An Adaptation of the Vector-Space Model for Ontology-Based Information Retrieval," *IEEE Trans. Knowledge and Data Eng.*, vol. 19, no. 2, pp. 261-272, Feb. 2007.
- [17] X. Yang, Y. Guo, and Y. Liu, "Bayesian-Inference Based Recommendation in Online Social Networks," *IEEE Trans. Parallel and Distributed Systems*, vol. 24, no. 4, pp. 642-651, Apr. 2013.
- [18] S. Meng, W. Dou, X. Zhang, and J. Chen, "KASR: A Keyword-Aware Service Recommendation Method on MapReduce for Big Data Applications," *IEEE Trans. Parallel and Distributed Systems*, vol. 25, no. 12, pp. 3221-3231, Dec. 2014.
- [19] G. Adomavicius and Y. Kwon, "New Recommendation Techniques for Multicriteria Rating Systems," *IEEE Intelligent Systems*, vol. 22, no. 3, pp. 48-55, May/June 2007.