IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727,
PP 53-58
www.iosrjournals.org

Automotive Tools for Making Effective Recommendations for E-commerce Websites: An In-Depth Comparative Study

Jyoti[1], Pran Dev[2], Dr. Sanjeev Dhawan[3], Dr. Kulvinder Singh[4]
[1, 2, 3, 4]*(Department of Computer Science & Engineering, University Institute of Engineering & Technology, Kurukshetra University, Kurukshetra, Haryana, India)*

ABSTRACT : *Deployment of recommendation engines in social websites and e-commerce Websites has greatly facilitated the growth of number of users and the customer satisfaction. Recommenders are automotive tools that make a significant contribution to a better understanding of customers' behavior and their interest, so as to provide them useful suggestion of items from a plethora of available information. Item is a generic term which may refer to video, song, book, friend list, location etc. There exist various recommender algorithms that have been proposed in recent years for generating different kind of user recommendations. This paper surveys the existing recommender algorithms that effectively and extensively produce suggestions to users. A comparison is made between content based and collaborative filtering recommendation engines that helps in alleviating the severe issues related to them and their effectiveness in making recommendations to users. In nutshell, an attempt has been made to provide an overview of recommenders, their evolution, and the exposures in areas of future implementation.*

Keywords - *Automotive Tools, Collaborative, Content-based, Hybrid, Items, Recommendations*

I. INTRODUCTION

Online The Internet has breached into all areas of human activities, making the rapid expansion of information and growing number of users in the modern era. Thus, because of overwhelming avalanche of information all around, there is a heavy need to dissect relevant information from irrelevant, so as to reduce the tiresome surfing in finding what is needed at a particular time and that too in a requisite manner, which best meets users' needs. The filtering of appropriate information has been done by manual searching or through automatic search engines which can condense information exploration practicable even within chaotically muffled and anarchistic environments such as the Web. Now-a-days, Recommendation systems have become another efficient means of providing relevant information with reduced efforts and low complexity. Recommender systems are automotive software tools which deal with information overload, and provide people with relevant suggestions based upon some parameters like users purchase history, similarity of content, ratings and reviews of other users etc. Many commercial sites and social websites now-a-days embed recommendation engine into their websites, to achieve some business objectives by knowing the taste of its users, and giving them optimum useful suggestions by framing the website according to their preferences. Some of the applications of the recommenders include recommending videos on YouTube, friend recommendations on Facebook, recommending items of interest to a user in online shopping sites, game recommendations, book recommendations etc. Recommendation Systems are mainly categorized into content-based, collaboration-based, and hybrid recommenders. Other recommendation techniques include personalized systems, explicit and implicit raters, demographic approaches etc.

The remaining paper is structured as follows: section II introduces the goals and motives for forming a recommendation engine, section III describes background analysis of related work, section IV presents the comparison of content based recommender and collaborative filtering based on advantages and pitfalls and finally section V concludes the paper.

II. GOALS AND MOTIVES FOR MAKING A RECOMMENDATION ENGINE

There are several reasons why recommendation systems have come to be an imperative part of commercial sites and social networking.

2.1 To make users find an item of interest in a time saving manner

There is a vast volume of available information on the web, which is unstructured and indirect. Finding unerring and appropriate information in the gigantic space is a burdensome job for a user. Recommendation Engines aid in the fast retrieval of interesting data with improved quality and efficiency, using push strategy. Users can even find less popular items with more accuracy.

2.2 To increase sale and earn high profits

Recommender systems tend to recommend items according to users' taste and its past purchase history. Better the quality of recommendations more will be sale of items and hence business persons will earn more profit.

2.3 User-centric approach

Recommenders are automotive user centric tools to tackle with the problem of recursive web surfing. They tend to transform the content on a website into a user friendly framework in which whole information is presented according to users' taste and access habits. The consumers with not as much of product knowledge and less shopping experience can also go for online shopping with ease of use and alleviated gratification.

III. BACKGROUND ANALYSIS OF RELATED WORK: STATE OF ART

Having introduced the basic requirements of recommendation systems, it can be said that the knowledge of custom-made search systems and recommendation engines has been extensively acknowledged amid users who necessitate support in probing, arrangement, categorizing, sieving and sharing the massive amount of information. A wide-ranging and orderly scrutiny of previous work is conceded out in this section.
The different approaches:

3.1 Collaborative Filtering Technique

It is the most renowned approach for making recommendations. The recommendations are generated according to the taste of like-minded users. It uses the rating and review given by users. The most common techniques are nearest neighbor method, k-means etc. In March 1997, Resnick and Varian [1] conferred the collaborative filtering model with its social consequences and commercial models casted-off to produce profits to shelter the maintenance costs. The main idea for recommendation was grounded on individuals providing and consuming them. The exercise of blind and double blind umpiring was given as an explanation to secrecy complications so as not reveal the behaviors of persons. To produce profits, they carried out three steps. First, pay-per- use model was well-thought-out. Second, they introduced the publicist's sustenance to produce comprehensive marketing information about the consumer. Third, they brought the concept of charging a fee to the owners, of items being estimated. In addition, Cai-Nicolas Ziegler *et.al* [2], in 2005 focused on properties of recommendation lists rather than focusing on the correctness of peculiar suggestions. They showed the "intra-list similarity metric". The sphere in which they made contribution are Topic diversification, Intra-list similarity metric and accuracy versus satisfaction. It even showed that user-based CF is less prone to topic diversification than item-based CF. The improvement of neighborhood based approach for recommender systems based on collaborative filtering for increasing accuracy and running time was given by Bell and Koren [3], 2007. In this approach synchronized interpolation leads to interaction between neighbors. This increases accuracy and improvises optimization problem. Here explicit profile creation is not required. Before start of KNN method the data is normalized, different ratings are brought closer and interfering variability is removed. This offers improvement in the quality of estimation with increase in running time. Lemire and Maclachlan [4] in 2008, proposed three slope one instances of having type, $F(x) = x+a$. The first one was The Slope-One scheme which depends on user's average rating and items on which user has rated. The drawback of this technique that the number of ratings visualized is not taken into account. This is overcome by "The Weighted Slope One scheme" [4] where numbers of ratings are considered. In third The Bi-Polar Slope One scheme one prediction is formulated from things user liked and predicting the other, using items that user hated. These approaches proved to be better in implementation than more expensive memory-based scheme.

Furthermore, WSRec(Web Service Recommender system), which has its implementation using Java language was presented by Zibin Zheng *et al* [5], *in 2009.* It had a method with user-involvement for collecting "Web service Qos information" [5] and an algorithm for "Web service Qos value prediction". This algorithm was hybrid collaborative filtering algorithm. This approach helped to overcome the data scarcity problem. It overcomes the drawback of service invocation and enforce cost for service users, consuming the service provider's resource. They worked to give a systematic approach to collect Qos information of web services, gave a hybrid collaborative filtering method to enhance the quality of suggestions and verified the algorithm experimentally. Then in 2010, RegionKKN, an innovative scalable hybrid algorithm for collaborative filtering based recommender was given by Xi Chen *et al.* [6], which grouped users on basis of regions into physical locations and on basis of Qos similarities. Services are identified by regions and predict Qos of candidate web services from information in past. It proposed models for the nearest-neighbor grouping and identifying region-sensitive web services correspondingly. It has scalability advantages as well. Again, In 2011, XIE Hongtao *et al.* [7], presented a collaborative filtering algorithm that made use of extremely valued ratings (EVRs). These EVRs calculated similarity between users in the user-item matrix. These EVRs used valued ratings by 1 of 5, or 5 of 5 other ratings are not considered. This was an approach to improve the prediction accuracy along with cutback in computing quantity. Three recognized similarity algorithms namely “cosine-based similarity”, “adjusted cosine based similarity”, and “Pearson correlation similarity” are used for similarity calculation [7]. For prediction output weighted sum and linear regression methods are considered. Combination of Bayesian non-parametric and max-margin learning was exhibited by Minjie Xu *et.al* [8], in 2012. These are two different paradigms but when applied together give balancing returns. Their paper gave infinite probabilistic max-margin matrix factorization. This was a nonparametric Bayesian style model. It determined unknown number of latent factors automatically. The maximum entropy discrimination principle is the key step towards this model. They also developed block wise coordinate descent algorithms pro variation inference. Threshold based Similarity Transitivity (TST) was proposed by Feng Xie *et.al* [9], in 2013. The setting of an intersection threshold removes the dissimilarities and those are then replaced by transitivity similarity. This method was meant to be in accordance with MapReduce Framework which is based on cloud computing platform. It balances the quality and quantity of similarity by using an adequate brink. Further, in 2013, Tianzi and Minchang [10] proposed the improvisation of “Follow the Leader” model, where the authorities got allocated to various fields according to the categories of rated items. The collaborative filtering algorithm they proposed was supported with prioritized domain expert trust in the process of recommendation. It has its application on actual set of data which is accessible publicly. It uses the idea of expert trust. The amount of computation is also condensed to a great extent.

3.2 Content-Based Approach

Content based methods are simple in functionality and fast in retrieval. They make recommendations according to the degree of similarity in the content of items. If user searches for an item or it has liked some item in the past, and then based on the keywords, the items similar to that will be recommended to him. The most common approaches are clustering, probability model, and “TF-IDF (term frequency – inverse document frequency) methods. Huan-Ming Chuang *et al.* [11], in 2008 conducted a comparison of content and preference based approach for recommendation. The importance of this comparison was to section the customers according to CLV (Customer Life Time Value), scrutinize effectual modified recommendations to improve the response of customers and maintaining the customers by correlation marketing. The result that it is advantageous to sector customers than to make available modified service to different customer sectors was obtained from comparison of ARM (Association Rule Mining) and RFM-CF. In 2009, Zenebe and Norcio [12], presented improvised state of fuzzy modelling technique. Improvements were on user behaviour and information about items as well as advancement in models for recommendation. Using movie recommendation as domain they further developed and assessed “Fuzzy set theoretic method (FTM)” for content-based recommendation engines [12]. This enhanced precision without loss in recalls and gives practical description for how the fuzzy set theory has to be applied in a new domain. The algorithms given are - item representation using fuzzy set, user feedback representation using fuzzy set, inference engine and algorithm.

3.3 Hybrid Methods

Hybrid methods combine the best of two, retaining their advantages and by-passing their pitfalls, hence giving sound quality recommendations. Balabanovic and Shoham [13], 1997 explained how a hybrid system, Fab included the advantages of both collaboration based and content-based techniques excluding the disadvantages of both. It involves accumulating the user databases and stalking their varying interests. The design comprises of three chief components- collection representatives, selection representatives and a central router. This also accomplishes personalization by exploiting group opinion. The Fab system improves its pursuance with time. Further, in 1998, Chumki Basu *et al.* [14], used hybrid features that united social and content based information so as to attain more accurate outcomes. They deliberated the domain of movie recommendation to walk around the complications of pure social-filtering. They functioned to cultivate a system which could use both ratings and content information. They formalized the problem as:

$$f((user,movie)) \rightarrow \{liked,disliked\}$$

For new movies, which were not rated, recommendations could be foretold as liked or disliked and output was not an ordered list of movies, in its place, it was the predicted one’s which could be liked by the user. This approach was supple and delivered better-quality enactment over collaborative method. Filterbolt, a different hybrid approach to address the rating sparsity and early rater problems was investigated in 1998 by Badrul M. Sarwar *et al* [15]. This approach was different from Fab as it incorporated semi-intelligent filtering agents called filterbolts into ratings- based collaborative filtering system. They proposed this method to solve problems of partitioning, implicit ratings, dimensionality reduction and sparsity problem. Filterbolts are automated rating robots that evaluate new published documents. It is not necessary for collaborative filtering agents to know whether the rater is filterbolt or human. It improved both coverage and accuracy. Results by filterbolt show SpellCheckerBot, IncludedMsgBot and LengthBot. This was the first swot with combination of collaborative filtering techniques and syntactic filtering technique. Further, ClustKNN algorithm which overcomes the problem of expense for large set of data was given by Al Mamunur Rashid *et al.* [16], in 2006. It first compresses data to a large extent making a clustering model and then applying Nearest Neighbour-Based Approach. It provides good accuracy, is well instinctive and scalable. The pros of both "memory-based and model-based algorithms" [16] are exploited in it. It can easily be tuned i.e. clusters could be adapted for making precision for time and space requirements it has construction of the model as off-line stage and generation of forecast suggestions as on-line. Then, In 2008 Xuan Nhat Lam *et al.* [17], developed a hybrid model combining collaborative filtering with user's information that gives forecast predictions to persons who have no priority on any item. They addressed the user side problems by using strength of vector aspect model along with user's personal details such as age, gender and occupation. This model resolves user-side cold-start problems. This model is effective only for small data set. Based on MovieLens dataset, this increased the excellence of user recommendations with "NMAE" approximately equal to 0.44 [17]. Furthermore, with the use of software architecture concepts and multi-agent system; Elammari and Elfrjany [18], 2012, made a trial to lessen the complexity of recommender systems. The architecture uses switching hybrid technology in order to switch between three recommendation algorithms which is "collaborative filtering", "content based" and "knowledge based" [18]. This enhances these three algorithms. Together they offer an influential recommendation. The recommendation process becomes faster, based on preferences of user but with least involvement from user. Even the developers benefit from this architecture by inheriting the advantage of the CF, CBF and KB approach, reducing sparisty problem, reducing the recommendation computation time by offline item retrieval.

3.4 Other Approaches

Some other methods include demographic methods, knowledge based methods, location based techniques, implicit and explicit rating based recommenders, personalized recommendation systems etc. In order to bridge the gap in the middle of user's priority and auto generated recommendations, Ja-Hwung Su *et al.* [19], in 2010 proposed FRSA, a novel recommender. It grouped a number of content based and collaborative based info in order to determine the user's inclination combining the rough set and average category rating [19]. Attempts have been made to overcome problems such as cold-start, first-rater, sparsity and extendibility. For removal of these problems FRSA was considered along with rule based imputation. It used switched based prediction method and grouping the users into clusters reducing the prediction cost significantly. Then, Camille Salinesi *et al.* [20], in 2012 proposed an interactive product line configuration approach which combined configuration and recommendation. These are two complementary form of guidance that informs the customer in real time if he/she could or could not get what all they desire.

This gives suggestions so as to make choices with reasoning along with known configurations. Every time series of decision are given to customer, choice is made by customer, user configuration is tested, recommendation is made, configuration and constraint propogation are dealt with and then final decision is taken. It recommends the configuration which he/she had specified initially. Another concept was brought into consideration for tourism. With increase in tourism industry, to provide the tourists with requested information Huang Yu *et al.* [21], 2013, presented a recommender system which provided services based on user's requirement, personality and travel habits. After designing a broad frame for personalized recommender system, comparison and analyzing of existing recommendation strategies, they used Apriori algorithm for the completion of intelligent recommendation module. Again, Shyi-Ming Chen *et al.* [22], 2014, offered a technique for group decision making by making the use of group recommendations using interval fuzzy preference relations along with consistency matrices. With the use of collective consistency matrix and preference relations group consensus degree of all experts is planned. The preference values are tailored according to marked consensus values so that its degree is larger than or equal to the already defined threshold value.

IV. COMPARISON OF RECOMMENDERS: ADVANTAGES AND PITFALLS

Table 1: Comparison of Recommenders

S.No.	Content based Recommenders	Collaborative Recommenders
1.	They rely only upon the unique ratings specified by current user to create their personal profile, which provides it an advantage of user-independence.	They depend upon ratings of other neighborhood users which have similar tastes to generate recommendations of interesting items.
2.	Content based recommendations are transparent to the users as they are usually what a user expects according to the degree of match of the content specified by him.	Collaboration based recommendations are based upon the taste of other unknown users, with similar tastes. So they generally act as black-box to an active user.
3.	The new items can be recommended by content based recommender, which are yet not rated by anyone.	This method heavily suffers from the problem of first rater. The unrated novel items may never get recommended.
4.	Content based methods have a problem of limited content analysis for discriminating items which do not have enough domain knowledge and parametric information.	They use neighborhood methods which are simple to implement and tune.
5.	They are not much reliable for a new user when enough ratings are not available. There is a problem of scalability because of high memory requirements.	They are more scalable and efficient than model based methods which require more memory and repetitive training for making recommendations.
6.	Enough ratings are required for a content based recommender to accurately understand users' preferences, which makes it suffer from cold start problem.	Collaborative filtering method also needs users past preferences to recommend items of interest. Hence, it also has a problem of cold start for new users, and data scarcity problem too.
7.	It uses model based approaches and clustering methods, so it can recommend items to a user with unique taste.	It uses nearest neighbor approach, which makes it unable to generate recommendations for someone with unique taste.

V. CONCLUSION

Recommendation systems are applications which aid users' with useful items' suggestion in the gigantic space of possible alternatives. Recommenders are programmed in such a way that they work in an automotive manner so as to reduce users' efforts in searching interesting and valuably appropriate information from the enormous pool of chaotic information. In this paper, an effort is made to make an in-depth study of the existing recommenders. Various recommendation techniques have been discussed and a comparison is made to describe the advantages and pitfalls of the different methods. Content based methods uses model based approaches and they suffer from the disadvantages of limited content analysis, scalability, and cold start problem. While collaborative filtering methods have first rater problem, data scarcity problem and they are not much transparent, as discussed in table 1. In nutshell, it can be said that combining both approaches to make a hybrid recommender is a more powerful and efficient tool in generating recommendations. In future, this approach can further be combined with other methods like demographic based, personalized methods and recommendations based upon frequency of visit of an item to generate more viable recommendation results.

REFERENCES

[1] P. Resnick and H. R. Varian, Recommender Systems, *Communications of the ACM*, 40(3), 1997.
[2] C-N Ziegler, S. M. McNee, J. A. Konstan and G. Lousen, Improving Recommendation Lists Through Topic Diversification, *GroupLens Research in Minneapolis, International World Wide Web Conference Committee (IW3C2)*, 2005.
[3] R. M. Bell and Y. Koren, Improved Neighbourhood- based Collaborative Filtering, *AT &T Labs Research,* 2007.
[4] D. Lemire and A. Maclachlan, Slope One Predictors for Online Rating-Based Collaborative Filtering, *In SDM,* Vol. 5, 2008, 1-5.

[5] Z. Zheng, H. Ma, M. R. Lyu and I. King, WSRec: A Collaborative Filtering Based Web Service Recommender System, *IEEE International Conference on Web Services*, 2009.

[6] X. Chen, X. Liu, Z. Luang and H. Sun, RegionKKN: A Scalable Hybrid Collaborative Filtering Algorithm For Personalized Web Service Recommendation, *IEEE International Conference on Web Services*, 2010.

[7] X. Hongtao, Z. Qingsheng, Q. Hongchun and Y. Ruilong, User-Based Collborative Recommendation Filtering Algorithm Using Extremely Valued Ratings, *International Journal of Digital Content Technology and its Applications(JDCTA)*, 5(9), 2011.

[8] M. Xu, J. Zhu and B. Zhang, Nonparametric Max-Margin Matrix Factorization for Collaborative Prediction, In *Advances in Neural Information Processing Systems*, 2012, 64-72.

[1] F. Xie, Z. Chen, H. Xu, X. Feng and Q. Hou, TST: Threshold based Similarity Transitivity Method in Collaborative Filtering with Cloud Computing, *Tsinghua Science and Technology ISSN*, 18(3), 2013.

[2] Z. Tianzi and J. Minchang, A CF Recommendation Method based on Domain Expert Trust, *International Journal of Knowledge and Language Processing*, 4(2), 2013.

[3] H-M Chuang, L-C Wang and C-C Pan, A Study on the Comparison between Content-Based and Preference-Based Recommendation Systems, *Fourth International Conference on Semantics, Knowledge and Grid, IEEE*, 2008.

[4] A. Zenebe and A. F. Norcio, Representation, Similarity Measures and Aggregation Methods using Fuzzy Sets for Content-Based Recommender Systems, *ScienceDirect, Elsevier, Fuzzy Sets and Systems*, 76(94), 2009.

[5] M. Balabanovic and Y. Shoham, Fab: Content based, Collaborative Recommendation, *Communications of the ACM*, 40(3), 1997.

[6] C. Basu, H. Hirsh and W. Cohen, Recommendation as Classification: Using Social and Content- Based Information in Recommendation, *In AAAI/IAAI, 1998, 714-720.*

[7] B. M. Sarwar, J. A. Konstan, A. Borchers, J. Herlocker, B. Miller and J. Riedl, Using Filtering Agents to Improve Prediction Quality in the GroupLens Research Collaborative Filtering System, *GroupLens Research team, Net Perceptions Inc., National Science Foundation under grant IRI-9613960*, 1998.

[8] A. M. Rashid, S. K. Lam, G. Karypis and J. Riedl, ClustKNN: A Highly Scalable Hybrid Model-& Memory-Based CF Algorithm, *Proceeding of WebKDD*, 2006.

[9] X. N. Lam, T. Vu, T. D. Le and A. D. Duong, Adressing Cold Start Problems in Recommendation Systems, In *Proceedings of the 2nd international conference on Ubiquitous information management and communication*, ACM, 2008, 208-211.

[10] M. Elammari and R. Elfrjany, Reducing the Complexity of Recommender System Development, *International Journal of Computer and Communication Engineering*, 1(4), 2012.

[11] J-H Su, B-W Wang, C-Y Hsiao and V. S. Tseng, Personalized Rough- Set-based Recommendation by Integrating Multiple Contents and Collaborative Information, *Information Sciences*, 180(1), 2010, 113-131.

[12] C. Salinesi, R. Triki and R. M. Cri, Combining Configuration and Recommendation to Define an Interactive Product Line Configuration Approach, *arXiv preprint arXiv: 1206.2520*, 2012.

[13] H. Yu, Y. Dan, L. Jing and Z. Mu, Research on Personalized Recommender System For Tourism Information Service, *Computer Engineering and Intelligent Systems, Science and Technology Planning Project of Guangdong province, China* , 4(5), 2013.

[14] S-M Chen, S-H Cheng and T-E Lin, Group Decision Making Systems Using Group Recommendations Based on Interval Fuzzy Preference Relations and Consistency Matrices, *Elsevier, Information Sciences*, 2014.