An Effective Fuzzy Association Rule Mining Algorithm for Collaborative Web Recommendation System

Dr A. Kumar
Associate Professor
Department of Computer Science and Engineering,
Perunthalaivar Kamarajar Institute of Engineering and Technology,

Karaikal.
E-Mail: kumar24462@outlook.com

Abstract--- *As the increasing curiosity of web recommendation systems those are applied to convey modified data for their users. In general the recommendations system are separated into two main types are collaborative recommendation system and content based recommendation system. In case of collaborative recommendation systems, these try to seek out others who share same tastes that of are given user as well as recommend the websites according to the liking are given user. Main drawback of the same is that the web pages are given equal significance. Here the importance of pages will change according to the frequency of calling the web page as well as the measure of time user spends on that page. To overcome this drawback, the research works have used the web usage log within the adaptive association rule based web mining wherever the association rules were applied to personalization. Though, this approach also suffers from some unavoidable problems. This work introduces an approach, "effective Fuzzy Association Rule Mining Algorithm (CWRS-FARMA)", a new hybrid algorithm for web recommendation system was proposed based on Association Rule Mining. This is improved algorithm which has semantic knowledge to the results for more effectiveness and thus gives better performances and quality are as compared to preceding methods.*

Keywords-Web Mining; Association Rules; Web Recommender System; CWRS-FAARMA; CWRS-FARMA.

I. INTRODUCTION

The pool of data available on the Internet is booming rapidly with the extensive growth of the World Wide Web (WWW). Since, the users are provided with enormous data and service options, it has become tougher for them to identify the "right" or "interesting" data. In fact, this susceptible difficulty is widely addressed as information overload. Recommender systems are optional, user-centric techniques, capable of dealing with the issue of information overload by adapting the content and structure of websites to the requirements of the users by taking the benefit of the information acquired from the analysis of the users' access behaviors. The Recommender Systems are broadly employed in several fields such as information filtering, web-browsing, movie recommender or net-news and e-commerce. The essential factor of all recommender systems is the user-model that comprises information bank of the individual preferences which determines his or her performance in a multipart environment of web-based schemes.

Even though a wide range of techniques and approaches are available in the recommender systems, still it is largely separated into two categories namely, Content-based Filtering and Collaborative Filtering Recommender System. As the Content based recommendation focuses on the web sites that the user has shown interest, the Collaborative Recommendation examines the similarity of interests among some users besides suggesting some other web sites both have used. As a matter of fact, both the Content-based recommendation and Collaborative recommendation have their own merits and demerits. For better results some recommender systems combines different techniques of collaborative approaches and content based approaches. Using the effective techniques like CWRS-FARMA is used to finding the problems through frequent itemsets using fuzzy sets in order to treat the quantitative attributes in relational and transactional databases. It also describes some of the baseline algorithms that are used in developing the web recommendation systems.

A rest of this research is organized as follows: Section II presents the background concepts about the techniques used in this research work. Section III represents about the proposed work in detail and Section IV presents about simulation result. Finally in Section V illustrates the conclusion are given.

II. LITERATURE SURVEY

[1] presents an adaptive web site may offer automated recommendations generated through any number of well-studied techniques including collaborative, content-based and knowledge-based recommendation. The surveys the space of two-part hybrid recommender systems, comparing four different recommendation techniques and seven different hybridization strategies. The study finds that cascade and augmented hybrids work well, especially when combining two components of differing strengths. [2] they proposed a novel Web recommendation system in which collaborative features such as navigation or rating data as well as the content features accessed by the users are seamlessly integrated under the maximum entropy principle. Both the discovered user patterns and the semantic relationships among Web objects are represented as sets of constraints that are integrated to fit the model. In the case of content features, they use a new approach based on Latent Dirichlet Allocation (LDA) to discover the hidden semantic relationships among items and derive constraints used in the model.

[3] presents a web usage mining has become the subject of exhaustive research, as its potential for Web-based personalized services, prediction of user near future intentions, adaptive Web sites, and customer profiling are recognized. Recently, a variety of recommendation systems to predict user future movements through Web usage mining have been proposed. Indeed, the basic process performed by a content-based recommender consists in matching up the attributes of a user profile in which preferences and interests are stored, with the attributes of a content object (item), in order to recommend to the user new interesting items. They provides an overview of content-based recommender systems, with the aim of imposing a degree of order on the diversity of the different aspects involved in their design and implementation was represented by [4].

[5]presents with increasing presence and adoption of Web services on the World Wide Web, Quality-of-Service (QoS) is becoming important for describing nonfunctional characteristics of Web services. They presented a collaborative filtering approach for predicting QoS values of Web services and making Web service recommendation by taking advantages of past usage experiences of service users. They represent a powerful method for enabling users to filter through large information and product spaces. Nearly two decades of research on collaborative filtering have led to a varied set of algorithms and a rich collection of tools for evaluating their performance. Research in the field is moving in the direction of a richer understanding of how recommender technology may be embedded in specific domains was presented by [6].

[7] presents with the increasing popularity of location-based services, such as tour guide and location-based social network, they now have accumulated many location data on the Web. They show that, by using the location data based on GPS and users' comments at various locations, they can discover interesting locations and possible activities that can be performed there for recommendations. They present an effective personalized collaborative filtering method for Web service recommendation. A key component of Web service recommendation techniques is computation of similarity measurement of Web services. Different from the Pearson Correlation Coefficient (PCC) similarity measurement, they take into account the personalized influence of services when computing similarity measurement between users and personalized influence of services was suggested by [8].

[9] presents in the rapidly changing financial market, investors always have difficulty in deciding the right time to trade. In order to enhance investment profitability, investors desire a decision support system. The proposed artificial intelligence methodology provides investors with the ability to learn the association among different parameters. After the associations are extracted, investors can apply the rules in their decision support systems.

[10] presents after that rough association rule mining (Rough ARM) came which seemed to be a good alternative of fuzzy association rule mining in terms of performance. But day by day their mining task is becoming huge. So, performing mining task efficiently and accurately over a large dataset is still a big challenge to us. Data and knowledge on the Web may, however, consist of imprecise, incomplete, and uncertain data. Because fuzzy-set theory is often used to handle such data, several fuzzy Web-mining techniques have been proposed to reveal fuzzy and linguistic knowledge. This paper reviews these techniques according to the three Web-mining categories above—fuzzy Web usage mining, fuzzy Web content mining, and fuzzy Web structure mining was presented by[11].

[12] suggested that the importance of pages changes according to the frequency of visiting the web page as well as amount of time user spends on that page. Also recommendation of newly added web pages or the pages those are not yet visited by users are not included in the recommendation set. To over-come this problem, they have used the web usage log in the adaptive association rule based web mining where the association rules were applied to personalization. Further, the present work explores the indirect (transitive) association between users as well as between items for providing more accurate recommendations even with highly sparse history of transactions was developed by [13].

[14] presented a novel web classification algorithm using the principles of fuzzy association rule mining to classify the web pages into different web categories, depending on the manner in which they appear in user sessions. This paper describes a collaborative educational data mining tool based on association rule mining for the ongoing improvement of e-learning courses and allowing teachers with similar course profiles to share and score the discovered information was presented by [15].

III. PROPOSED METHODOLOGY

This research focuses on the usage of collaborative filtering for web recommendation system based on Association Rule Mining which of course is studied through the subsequent applications or avenues. The proposals of the techniques are as follows:

A. A Collaborative Web Recommendation Systems Based on Fast Adaptive Association Rule Mining Algorithm

To provide web recommendation, a new fast algorithm for mining association rule has been proposed. This current approach is able to creating web recommendation more successfully and effectively with accurate result against the conventional method. By joining similarity among rules and currently present user and confidence of the weighted rules, the recommendation engine will select only the most relevant pages. Therefore, it will increase the effectiveness of the recommendation engine. The proposed algorithm also guides the society by offering the most common web sites that is needed. Where adaptive associative rule mining and fast adaptive association rule mining are the techniques used for the web recommendation processes. These techniques are very much useful for web recommendation. The modifications using other algorithms to get the association rules that can be assumed in the preceding web recommendation system to make them practically more efficient. To improve the effectiveness of the web recommendation system using fuzzy association rule mining technique.

B. Effective Fuzzy Association Rule Mining Algorithm for Web Recommendation Systems

Mining fuzzy association rules is the detection of association rules using fuzzy set models such that the quantitative attributes can be treated. In this research work views each attribute as a linguistic variable, and the variables divided into various linguistic terms. An efficient algorithm was required for the FARM methodology because numerous preprocessing (conversions, filtration, normalization) and mining steps are regarded in the web recommendation system. During the fuzzy ARM process, a number of fuzzy partitions are defined on the domain of each quantitative attribute, as result of which the original dataset is transformed into an extended one with attribute values in the interval [0, 1].

An association rule 'r' denotes an expression of the form $X \Rightarrow Y$ (σr, σr), where X and Y are item sets, σr is the support of $X \cup Y$, and σr is the confidence for the rule 'r' given by $\sigma (X \cup Y)/ \sigma (X)$. Additionally, association rules that do not fulfill a minimum lift threshold are pruned, where elevation is determined as $\sigma r/\sigma (Y)$. If there's no enough support for a specific item that may never start up in any frequent item set. The importance is that such associate item can never be suggested. The difficulty of coverage could be an exchange. Lowering the support threshold can make sure that a lot of things are often suggested, however at the chance of recommending associate item while not ample proof of a pattern.

FARM algorithm consists of four major steps:

Step 1: Filtration and transformation of common web logs database into a database with average (Web System Attributes) WSA logins.

Step 2: Accurate and proper transformation of WSA logs into a database having fuzzy extensions. Normalization is presented in this database.

Step 3: Nominee generation and look for the entire fuzzy frequent itemsets within contestant that have fuzzy which maintain greater than the user specified minimum support.

Step 4: Use of frequent itemsets to evaluate the preferred most significant websites through evaluating the correlation values and fuzzy confidence.

The Fuzzy ARM algorithmic rule belongs to the breadth initial traversal relations of ARM algorithms, expanded by tree data structures and it functions in an exceedingly fashion quite the same as the Apriori algorithm. Also, implementation of this proposed method is different from avoiding an additional database scan to search out correlation values, so increasing potency.

IV.   EXPERIMENTAL RESULTS

In this chapter, experimental evaluation is presented for the proposed approaches CWRS-FAARMA and CWRS-FARMA and performed using MATLAB environment. It is used for the computation of the numerical analysis and is considered as a fourth-generation programming language. Matlab is a commercial Matrix Laboratory package which performs as an interactive programming environment. In general, these experimentations have verified the effectual of the proposed methods in web page recommendation.

$$Precision = \frac{correctly \ recommended \ items}{total \ recommended \ items}$$

$$Recall = \frac{correctly \ recommended \ items}{total \ useful \ recommendations}$$

$$Coverage = \frac{items \ with \ recommendations}{total \ number \ of \ items}$$

TABLE 1
PRECISION, RECALL AND CONVERGE RATE FOR WEB RECOMMENDATION TECHNIQUES

Methods	Precision (%)	Recall (%)	Converge rate (Seconds)
CWRS-FAARMA	82	69	39
CWRS-FARMA	98	47	25

In table 1 shows the precision, recall and converge rate for collaborative web recommendation based on effective fuzzy ARM algorithm.
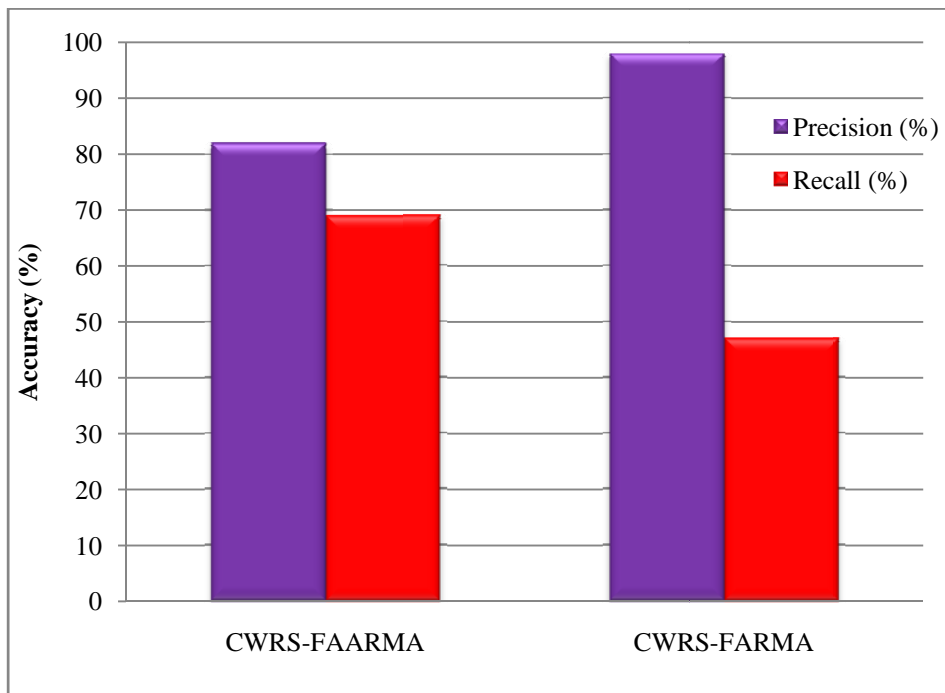

Figure 1 Precision and Recall for CWRS methods

Figure 1 and Figure 2 shows the Precision and Recall for CWRS-FAARMA and CWRS-FARMA. The proposed method of CWRS-FARMA have high precision and less recall rate. Figure 3 shows the convergence rate for web recommendation techniques.
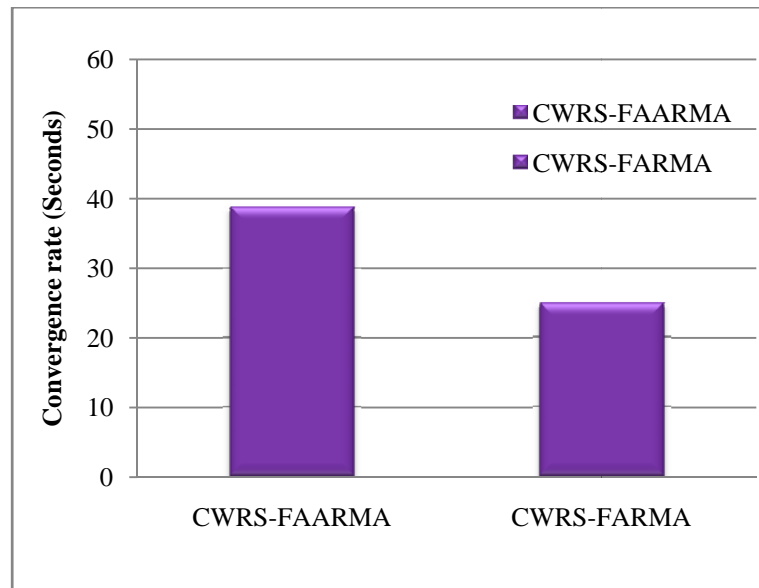

Figure 2 Converge rate for CWRS methods

V.   CONCLUSION

In this research conclude that the effective algorithm for mining association rule to present web recommendation has been suggested. It can be fulfilled that the proposed method is accomplished of making web recommendation in effectively and more accurately against the preceding approaches. The experimental results prove that the presentation of proposed CWRS-FARMA outperforms the existing collaborative web recommendation algorithm through convergence rate, precision and recall. The proposed CWRS-FARMA proceeds accurate value in their performance and renders best outcomes for web recommendation systems.

REFERENCES

[1]   Burke, Robin (2007), "Hybrid web recommender systems", in The adaptive web, pp. 377-408.
[2]   Jin, Xin, Yanzan Zhou, and Bamshad Mobasher (2005), "A maximum entropy web recommendation system: combining collaborative and content features", In Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining, pp. 612-617.
[3]   Jalali, Mehrdad, Norwati Mustapha, Md Nasser Sulaiman, and Ali Mamat (2010), "WebPUM: A Web-based recommendation system to predict user future movements", Expert Systems with Applications 37, no. 9: 6201-6212.
[4]   Lops, Pasquale, Marco De Gemmis, and Giovanni Semeraro (2011), "Content-based recommender systems: State of the art and trends." In Recommender systems handbook, pp. 73-105.
[5]   Zheng, Zibin, Hao Ma, Michael R. Lyu, and Irwin King (2011), "Qos-aware web service recommendation by collaborative filtering", Services Computing, IEEE Transactions on 4, no. 2: 140-152.
[6]   Ekstrand, Michael D., John T. Riedl, and Joseph A. Konstan (2011), "Collaborative filtering recommender systems", Foundations and Trends in Human-Computer Interaction 4, no. 2: 81-173.
[7]   Zheng, Vincent W., Yu Zheng, Xing Xie, and Qiang Yang(2010), "Collaborative location and activity recommendations with gps history data", In Proceedings of the 19th international conference on World wide web, pp. 1029-1038.
[8]   Jiang, Yechun, Jianxun Liu, Mingdong Tang, and Xiaoqing Liu (2011), "An effective web service recommendation method based on personalized collaborative filtering", In Web Services (ICWS), 2011 IEEE International Conference on, pp. 211-218.
[9]   Ho, George TS, W. H. Ip, C. H. Wu, and Y. K. Tse (2012), "Using a fuzzy association rule mining approach to identify the financial data association", Expert Systems with Applications 39, no. 10: 9054-9063.
[10] Roy, Aritra, and Rajdeep Chatterjee (2014), "Introducing New Hybrid Rough Fuzzy Association Rule Mining Algorithm", In Int. Conf. on Recent Trends in Information, Telecommunication and Computing, ITC.
[11] Lin, Chun Wei, and Tzung Pei Hong (2013), "A survey of fuzzy web mining", Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 3, no. 3: 190-199.
[12] Wanaskar, Ujwala, Sheetal Vij, and Debajyoti Mukhopadhyay (2013), "A Hybrid Web Recommendation System Based on the Improved Association Rule Mining Algorithm", arXiv preprint arXiv: 1311.7204.
[13] Tyagi, Shweta, and Kamal K. Bharadwaj (2013), "Enhancing collaborative filtering recommendations by utilizing multi-objective particle swarm optimization embedded association rule mining", Swarm and Evolutionary Computation 13: 1-12.
[14] Thomas, Binu, and G. Raju (2013), "A Novel Web Classification Algorithm Using Fuzzy Weighted Association Rules", International Scholarly Research Notices2013.
[15] García, Enrique, Cristóbal Romero, Sebastián Ventura, and Carlos de Castro (2011), "A collaborative educational association rule mining tool", The Internet and Higher Education 14, no. 2: 77-88.