

A Fast and New Collaborative Web Recommendation System Using Fast Adaptive Association Rule Mining

Dr A. Kumar

Associate Professor

Department of Computer Science and Engineering,
Perunthalaivar Kamarajar Institute of Engineering and Technology,
Karaikal, India.

Abstract--- World Wide Web has become one of the most extensive information resources in a recent span of time. It mostly covers all the information needed for any user. But, finding data on a large web site is a not an easy task. The users of the web sites mostly suffer from the problem of finding the required data in time. In fact, locating the required dataset on the web has become one of the difficult and time consuming tasks today. Massive development of internet in recent years necessitates the improvement of recommender systems that is to be user friendly in web applications. This work introduce an approach, “Collaborative Web Recommendation Systems Based on Fast Adaptive Association Rule Mining Algorithm (CWRS-FAARMA)”, a new Fast algorithm for web recommendation system was proposed based on Adaptive Association Rule Mining. Experimental result shows this algorithm used to speeds up the web recommendation process with high accuracy.

Keywords--- Recommendation system, CWRS-ARMA, CWRS-AARMA, CWRS-FAARMA, Association Rule Mining Algorithm.

I. INTRODUCTION

There are two approaches that can be used when generating recommendation systems. One is content based web recommender system and the other is collaborative web recommender system. Existing recommendation systems exhibit the problems of less coverage, lack of accuracy, extended and high run time. To overcome this problem this work uses three methods. The main objective of this research is to develop an efficient web recommendation system with better recommendation results when compared to previously recommendation systems.

The approach “Collaborative Web Recommendation Systems Based on Association Rule Mining Algorithm (CWRS-ARMA)” employs the association rule mining technique. The process of personalization based on web usage data, the association rules were applied. This method generates association rules by utilizing apriori algorithm. In general, association rule mining is a method common in data mining, which tries to discover patterns of products that are purchased together.

Another approach “Collaborative Web Recommendation Systems Based on Adaptive Association Rule Mining Algorithm (CWRS-AARMA)” uses the adaptive

association rule mining techniques for web recommendations. In order to select desirable and appropriate minimum confidence and support for each web page before the mining process a new rule to supplement association rule mining for recommender systems namely, adaptive association rule mining for the web recommendation system is adapted. Another proposed method is called CWRS-FAARMA which is used for fast performance.

II. LITERATURE SURVEY

[1] suggested association rule mining based on concept lattice in bioinformatics research. Concept lattice characterizes information with the associations between the intent and the extent of concepts, and the associations among the overview and the specialty of concepts, then information can be revealed on the Hasse diagram with hierarchical arrangement, thus it is correctly applied to the description of mining association rules in databases.

Web recommendation system is an essential research area of web mining. In this approach [2] proposed a novel web recommendation system model derived from cluster mode to understand the real-time online recommendation. [3] focuses on offering suggestions to learners in addition to web masters to develop overall efficiency of web based teaching and learning. This approach deals with examination of web log data and improvement of recommendation framework using web usage mining.

[4], [5] done researches on Personalized Recommendation System based on Web Mining with the intention of overcoming the information overload, personalized recommendation systems were planned to recommend products and to offer consumers with information to assist them in make a decision which products to purchase. [6] presented a personalized Web page recommendation system namely personalized web paGerEcommendatiON (PIGEON) through collaborative filtering and a topic-aware Markov system.

[8] proposed a collaborative filtering technique for predicting QoS values of Web services and making Web service recommendation by the benefits of previous experiences of service users. [7] presented an induction-based approach for generating a compressed representation of sets of association rules, mainly proposed for publishing the compact representation in the Web. [6] proposed a novel technique for mining such rules derived from hidden

Markov model (HMM). In this technique, the contents of the pages and Web server's log are required to be preprocessed initially and subsequently some definitions of users' access interest in a particular Web site are provided.

[9] proposes a new approach to mine context based positive and negative spatial association rules as they might be applied to hydrocarbon prospection. Many researchers are currently using an Apriori algorithm on spatial databases but this algorithm does not utilize the strengths of positive and negative association rules and of time series analysis, hence it misses the discovery of very interesting and useful associations present in the data.[10] propose a novel collaborative filtering algorithm designed for large-scale web service recommendation.

[11] investigate the sick and healthy factors which contribute to heart disease for males and females. Association rule mining, a computational intelligence approach, is used to identify these factors and the UCI Cleveland dataset, a biological database, is considered along with the three rule generation algorithms – Apriori, Predictive Apriori and Tertius.

[12] develops a better conceptual base for improving the application of association rule mining methods to extract knowledge on operations and information management. The emphasis of the paper is on the improvement of the operations processes. Association rule discovery is an ever increasing area of interest in data mining. Finding rules for attributes with numerical values is still a challenging point in the process of association rule discovery are shown by [13].

[14] propose a hybrid personal trust model which adaptively combines the rating-based trust model and explicit trust metric to resolve the drawback caused by insufficient past rating records. Moreover, users with similar preferences usually form a group to share items (knowledge) with each other; thus, users' preferences may be affected by group members. Recommendation systems support users and developers of various computer and software systems to overcome information overload, perform information discovery tasks, and approximate computation, among others are introduced by [15].

III. RESEARCH METHODOLOGY

The approaches determine the preferences of the user from their implicit feedbacks, especially the web pages of their browsing history. Clustering and collaborative filtering techniques are prepared to integrate both binary and non-binary weights of pages, even if binary weights are generally used for computing efficiency. The following approaches that can be used for generating recommendation systems.

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

The application and utility of the internet is gaining control over wider audience in recent span of time. Consequently, it is necessary that a user need to expend much time to choose essential information from large web pages created every day. Addressing this issue, various web page recommender systems are developed which

automatically choose and for user's support the recommends web pages is suitable. There are two approaches which can be used when generating recommendations such as Content based techniques and collaborative recommendation techniques.

Collaborative recommendation technique tries to find some users who share similar tastes with the given user. The proposed approach involves algorithm of healthy association rule mining for Association Rule Mining. To obtain the web sites which are most relevant to user this approach is used.

1) Association Rule Mining Algorithm (ARMA) for Web Recommendation

The association rule mining is a technique which is used in data mining that attempts to determine patterns of products that are purchased together. This approach adapts the Apriori algorithm to collaborative filtering is an effort to determine patterns of items that have ordinary ratings. Association rules capture associations between things related on patterns of co-occurrence across transactions. The association rules were used to personalization based on web usage data. That is, every user profile as a transaction, it is possible to use the Apriori algorithm and generate association rules for groups of commonly liked items.

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

The Apriori Algorithm has its individual merits and demerits. Though it competently sustains support to us in mine rules, the time complexity and the space complexity of the algorithm increases proportional to the size of the database. The Apriori algorithm uses join operations and stores the results in a table for each pass resulting in a large space complexity as well as time complexity.

1) A New Adaptive Algorithm Association Rule Mining Algorithm (AARMA)

This approach relies on information about relationships among different users' preferences in order to suggest the web site of potential interest to the target user. This algorithm adjusts the limited support of the rules during mining in order to obtain an appropriate number of significant rules for the target predicate.

The new AARMA-CRS consists of two parts: AARMA-CRS-1 and AARMA-CRS-2.

AARMA-CRS-1 With the intention of mining only a specified number of most capable rules for each target web site, AARMA-CRS-1 is used to control the minimum support count and discover the rules with the highest supports. The minimum support count is the smallest amount of transactions that convince a rule with the aim of making that rule frequent, specifically; it is the multiplication of the minimum support and the whole number of transactions. AARMA-CRS-2 is an alternative of CBA-RG and as a result of the Apriori algorithm also. AARMA-CRS-2 is an alternative of CBA-RG in the sense that rather than mining rules for all target classes, it only mines rules for one target item. It varies from CBA-RG in that it will simply mine a number of rules within a particular range.

C. 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.

1) Proposed Association Rule Mining Methodology

In this section, two new algorithms, Apriori and AprioriTid are given, that differ fundamentally from ordinary algorithms. Later discussion is on how the best features of Apriori and AprioriTid can be joined into a hybrid algorithm, called AprioriHybrid.

The difficulty of discovering all association rules can be decayed into two sub problems:

1. Find all sets of web sites (itemsets) that have transaction support over minimum support. The support for an itemset is defined as the number of transactions that include the itemset. Itemsets that has minimum support are considered as large itemsets, and remaining are small itemsets. In this section, novel approaches Apriori and AprioriTid are presented for solving this difficulty.

2. Use the large itemsets to produce the preferred rules. The algorithms for this difficulty are specified in this section. The universal idea is that if, say, ABCD and AB are large itemsets, then can decide if the rule $AB \Rightarrow CD$ holds by calculating the ratio $conf = \frac{support(ABCD)}{support(AB)}$. If $conf \geq minconf$, then the rule holds.

IV. EXPERIMENTAL RESULTS

In this chapter, experimental evaluation is presented for the proposed approaches CWRS-ARMA, CWRS-AARMA and CWRS-FAARMA and is compared with the standard algorithm called First-order Markov Model. MATLAB is used for the computation of the numerical analysis and is considered as a fourth-generation programming language.

MATLAB allows matrix operations, plotting of functions and data, creation of user interfaces, execution of algorithms and interfacing with programs written in other languages, including C, C++, and FORTRAN. The performance of the proposed approaches are evaluated using the following parameters like Precision, Recall and Coverage.

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

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

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

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

| Methods | Precision (%) | Recall (%) | Converge rate (Sec) |
|-------------|---------------|------------|---------------------|
| CWRS-ARMA | 74 | 84 | 39 |
| CWRS-AARMA | 86 | 79 | 31 |
| CWRS-FAARMA | 92 | 64 | 25 |

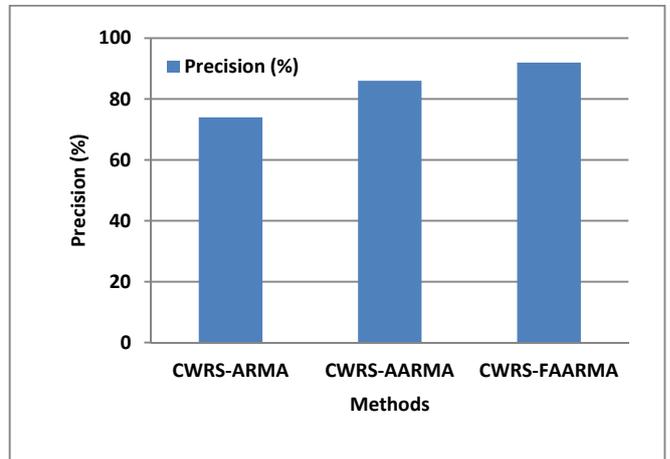


Figure 1: Precision for CWRS approach

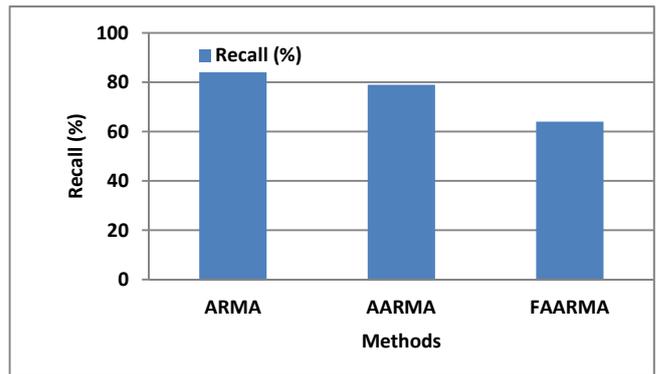


Figure 2: Recall for CWRS approach

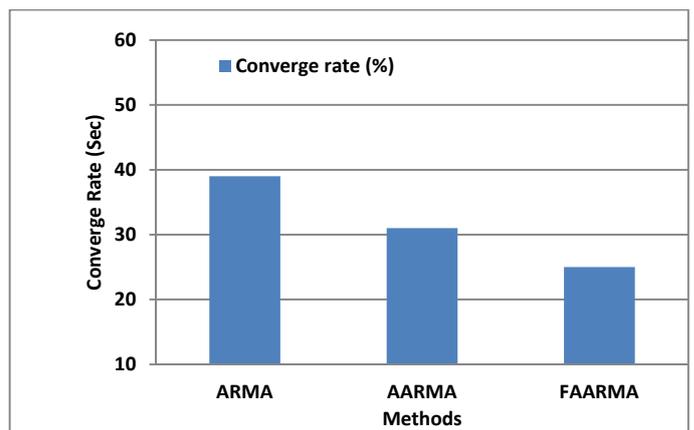


Figure 3: Converge rate for CWRS approach

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

V. CONCLUSION

The aspiration and focus of this research is mainly formulating a new perspective approach towards developing an efficient web recommendation system with an application of association rule mining. Through the implementation of this system, it is illustrated, how an efficient web recommendation system can be effectively integrated with an associative rule mining model. This proposed approach for web recommendation is capable of yielding good results and can be considered as an alternative to traditional web recommendation approaches.

REFERENCES

1. Dexing Wang, Lixia Cui, Yanhua Wang, Hongchun Yuan, Jian Zhang(2010), "Association Rule Mining Based on Concept Lattice in Bioinformatics Research", 2010 International Conference on Biomedical Engineering and Computer Science (ICBECS), Page(s): 1 – 4.
2. Chishe Wang, Qi Shen, LinjunZou(2010), "Research and Application of Web Recommendation System Based on Cluster Mode", International Conference on E-Business and E-Government (ICEE), Pp. 1445 – 1447.
3. Anitha, A., Krishnan, N.(2010), "A web usage mining based recommendation model for learning management systems", IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Pp. 1 – 4.
4. Xia Min-jie, Zhang Jin-ge(2010), "Research on personalized recommendation system for e-commerce based on web log mining and user browsing behaviors", International Conference on Computer Application and System Modeling (ICCASM), Vol. 12, Pp. 408-411.
5. Zhang Junyan, Shao Peiji(2010), "Personalized recommendation based on WAR", International Conference on Computer Application and System Modeling (ICCASM), Vol. 2, Pp. 108 - 111.
6. Qingyan Yang, Ju Fan, Jianyong Wang, Lizhu Zhou,(2010) "Personalizing Web Page Recommendation via Collaborative Filtering and Topic-Aware Markov Model", IEEE 10th International Conference on Data Mining (ICDM), Pp. 1145 – 1150.
7. MinutoEspil, Mauricio, Ale, Juan M.(2010), "An induction-based compaction of sets of association rules among Web concepts", Fourth International Conference on Research Challenges in Information Science (RCIS), Pp. 311 – 318
8. Zheng, Zibin, Ma, Hao, Lyu, Michael R., King, Irwin(2011), "QoS-Aware Web Service Recommendation by Collaborative Filtering", IEEE Transactions on Services Computing, Vol. 4, No. 2, Pp. 140 – 152.
9. Shaheen, Muhammad, Muhammad Shahbaz, and Aziz Guergachi. "Context based positive and negative spatio-temporal association rule mining." *Knowledge-based systems* 37 (2013): 261-273.
10. Chen, Xi, ZibinZheng, Xudong Liu, Zicheng Huang, and Hailong Sun. "Personalized qos-aware web service recommendation and visualization." *Services Computing, IEEE Transactions on* 6, no. 1 (2013): 35-47.
11. Nahar, Jesmin, Tasadduq Imam, Kevin S. Tickle, and Yi-Ping Phoebe Chen. "Association rule mining to detect factors which contribute to heart disease in males and females." *Expert Systems with Applications* 40, no. 4 (2013): 1086-1093.
12. Kamsu-Foguem, Bernard, Fabien Rigal, and Félix Mauget. "Mining association rules for the quality improvement of the production process." *Expert Systems with Applications* 40, no. 4 (2013): 1034-1045.
13. Minaei-Bidgoli, Behrouz, R. Barmaki, and Mahdi Nasiri. "Mining numerical association rules via multi-objective genetic algorithms." *Information Sciences* 233 (2013): 15-24.
14. Lai, Chin-Hui, Duen-Ren Liu, and Cai-Sin Lin. "Novel personal and group-based trust models in collaborative filtering for document recommendation." *Information Sciences* 239 (2013): 31-49.
15. Avazpour, Iman, TeeratPitakrat, Lars Grunske, and John Grundy. "Dimensions and metrics for evaluating recommendation systems." In *Recommendation Systems in Software Engineering*, pp. 245-273. Springer Berlin Heidelberg, 2014.