Predicting the Users' Navigation Patterns in Web, using Weighted Association Rules and Users' Navigation Information

Zeynab Liraki
Department of Computer Engineering, Science and Research khozestan Branch, Islamic Azad University, Ahwaz, Iran

Ali Harounabadi
Department of Computer Engineering, Tehran Center Branch, Islamic Azad University, Tehran, Iran

Javad Mirabedini
Department of Computer Engineering, Tehran Center Branch, Islamic Azad University, Tehran, Iran

ABSTRACT
World Wide Web is developing in a chaotic and unfocused process, and this process has resulted in production of documents which are linked with each other, and which are not logically organized. Therefore, the aim of recommender systems is guiding users to find their favorite resources and meet their needs, by using the information obtained from the previous users' interactions. In this paper, to predict the users' navigation pattern with high precision, a hybrid algorithm of FCM fuzzy clustering techniques, weighted association rules, and fuzzy systems are presented. This algorithm is implemented in two phases, namely offline and online phases. In offline phase, using the recorded data in log file of the web server, the users' navigation patterns are extracted. In online phase, the recommender system suggests, as the initial proposed set, a list of the current user's favorite webpages which he/she has not visited yet. Then it expands this set using HITS algorithm so that the new webpages which have recently been added to the website have the chance to be present in the list of the proposed webpages. The results of the simulation in real-world data indicate the higher efficiency of the proposed algorithm in terms of precision and coverage comparing to other algorithms.

Keywords
Recommender System, FCM Clustering, Fuzzy Inference System, Weighted Association Rules.

1. INTRODUCTION
World Wide Web contains so much information and performs like a large data base. In various studies regarding web development, it has been estimated that there are more than a million webpages added to the World Wide Web daily, and more than 600 GB of the pages undergo changes during a month [1, 2]. This phenomena is called Information Overloading. Therefore, methods and techniques are needed to facilitate effective access to extract the information from these pages. The main challenge users are facing is to effectively find relevant information with the minimum effort and time invested. In this huge information store, users have difficulties finding their required information easily and at the right time, since on the one hand, they have to evaluate the extent of relevance of the content of the webpage to their needs, and on the other, they have to evaluate webpages in terms of their reliability. To resolve this problem, personalizing webpages to customize them has become very popular. Any action done to adapt the information or services provided by a webpage to a special group of users' needs through making use of the information obtained from the user's circulating behavior and their particular interests is called Web Personalization [3]. The aim of a web personalization system is to provide the desired or required information of the users without their explicit request.

Personalizing webpages is a functional area of web-mining, through which one can adapt the contents of the pages to the users' interests in order to offer services and provide the user's required information. There are many personalized webpages created based on web-mining, all of which include two main steps [4-7]. In the first step, which is offline, the users' access patterns and behavioral models are mined based on educational data retrieved from the users' behavior in webpages. In the second step, which is online, the models obtained from the first step are used to interpret and compare with the current user's navigation pattern, and advices are given based on such comparisons. The aim of personalizing webs is to recommend a series of items to the current user. Such recommendations include links, advertisements, texts, products, etc., in lines with the user's interests and preferences. It is done based on matching the user's current session (and probably along with his/her saved profile) with the functional patterns discovered from web-mining. This process is performed by the search engine which is an online component of personalization system.

In web-usage mining, the main data-mining techniques are used. These techniques include mining of association rules, extraction of consecutive patterns, and clustering in order to extract conductive patterns and offer recommendations based on them. Association rules have successfully been used in web recommender system [8]. In this paper, FCM is combined with usual association rules which is developed through assigning weight to items included in transactions in order to determine the significance of each item in transaction, and defined "Weighted Association Rules". In the proposed model, at first the web log files are preprocessed, and therefore the users' sessions are extracted. To produce the users' movement patterns, an algorithm is presented which is based on fuzzy clustering, weighted association rules and fuzzy system. After making movement patterns, the recommender system offers a list of the user's favorite webpages which he/she has not visited yet. The proposed algorithm has been simulated on real-world data, and the results indicate that this algorithm has significantly enhanced the quality of the recommendations. In the second section, the background material about clustering and association rules will be presented. In section 3, several works done in this field will be studied. In section 4, the proposed algorithm will be presented, in section 5, the proposed algorithm will be evaluated. And eventually in section 6, the results and conclusions are discussed.

2. BACKGROUND MATERIAL

2.1 FCM Clustering Algorithm

Like the classic C-Means algorithm, in this algorithm, the number of clusters (c) are determined in advance. The objective function defined for this algorithm is as presented in formula 1:

$$J = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{m} d_{ik}^{2} = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{m} \parallel x_k - v_i \parallel^2 \qquad (1)$$

In the above formula, m is a real number greater than 1, and in most cases number 2 is selected for m. X_k is the k-th sample, and V_i is the representative or the center of the i-th cluster. U_{ik} shows the extent the i-th sample belongs to the k-th cluster. ||*|| is the extent of the similarity (distance) of the sample to (from) the center of the cluster, for which we can use any function which indicates the similarity of the sample and center of the cluster. A U matrix can be defined from U_{ik} which has c rows and n columns, and its components can have any values between 0 and 1. If all the components of matrix u are 0 or 1, the algorithm will be similar to C-Means algorithm. Although the components of matrix U can have any values between 0 and 1, the total components of every column must be equal to 1, hence the formula 2:

$$\sum_{i=1}^{c} u_{ik} = 1, \forall\, k = 1, \ldots, n \qquad (2)$$

This condition means that the extent every sample belongs to c cluster must be equal to 1. According to the mentioned condition, and minimizing the objective functi ‹formula (3) and (4) will be presented.

$$u_{ik} = \frac{1}{\sum_{j=1}^{c}\left(\frac{d_{ik}}{d_{jk}}\right)^{2/(m-1)}} \qquad (3)$$

$$v_i = \frac{\sum_{k=1}^{n} u_{ik}^{m} x_k}{\sum_{k=1}^{n} u_{ik}^{m}} \qquad (4)$$

FCM algorithm steps:

1. Initializing c, m, and U_0. The initial clusters are guessed.
2. The center of clusters are calculated (calculation of v_i).
3. Calculation of matrix-attachment from the clusters calculated in step 2.
4. If $\|U_{l+1} - U_l\| \leq \varepsilon$, the algorithm ends, otherwise go to step 2.

2.2 Mining association rules

Frequent Items Mining was presented by Agarwal for the first time in 1993 as mining association rules in set of items. Association rules show the interactions among items based on their co-occurrence in transactions (regardless of their orders.)

The basic hypothesis are as follows [9]: I = I_1 ,I_2 ,I_3 ,… as the set of items are considered, and set D as the set of data, that is transaction of database in a way that every transaction includes a set of items, that is every T transaction is a subset of I (T⊆ I). Every transaction has an identifier called TID. If A is a set of items, then transaction T includes A if and only if A ⊆ T, in which A and B are two sets of items. A is called antecedent and B called consequent. An association rule is a proposition which is indicated as formula 5 [10]:

A→B [Support, Confidence] | $A \subseteq I$, $B \subseteq I$, A∩B = ∅ (5)

To review the value and acceptance criteria of association rules, two important parameters namely Support and Confidence are introduced, which are calculated as in formulae 6 and 7:

$$confidence\ (A \rightarrow B) = \frac{number\ \ of\ trasaction\ \ (A \cup B)}{number\ \ of\ trasaction\ \ (A)} \qquad (6)$$

$$Support(A \rightarrow B) = \frac{number\ \ of\ trasaction\ \ (A \cup B)}{number\ \ of\ trasaction} \qquad (7)$$

The rules which meet the min-sup and min-conf are called strong association rules, and the aim of all algorithms is to find such rules. About web transactions, association rules show the relationships between visiting webpages based on the users' circular patterns.

3. REVIEW OF THE PREVIOUS WORKS

Olson and Lee in [11] presented a data-mining algorithm to discover fuzzy association rules for quantitative and weighted data. In this algorithm, for every transaction and all features of the transaction, each feature is divided into fuzzy areas. To produce set of single-member candid items, for every fuzzy area of features, the fuzzy values are calculated based on determined membership functions, and the support values of every fuzzy area are calculated to produce items. Then it will be examined that if the support of that fuzzy area is greater than or equal to the minimum determined support threshold, then the area is placed within the large single-item set. In the next step, if the sum of large single-item sets is not empty, the algorithm will continue, otherwise the algorithm ends. To produce candid items of C_{r+1}, the important items of l_r , if having r-1 common items, are merged. Then, to produce large items having r+1 items, the membership value and fuzzy support of each area in C_{r+1} are calculated, and if greater than or equal to the determined fuzzy min-sup, it will be placed in l_{r+1}. These steps are repeated for the sum of all produced large items, and finally the produced association rules, if their confidence coefficient is greater than or equal to the user-defined min-conf, will be saved as interesting association rules.

Makker and Rathy In [12], using users' behavior analysis, an approach, which makes use of preprocessing of log-file, is presented in order to pre-fetch, predict, and improve the performance of the web server. Clustering, Marco Model, and association rules are the three techniques used in recommending webpages. Therefore, after preprocessing and identifying the sessions, they are clustered using K-means clustering algorithm and measuring similarities. Every data-set is grouped in a different cluster. Then, Marco Model predicts the results. In case of ambiguity, association rules are applied to present accurate results. Extracting website structure, it complements the website routes. The knowledge base in this system is a reservoir of features which are mined using data-mining techniques. These features include the number of users, the visited webpages, and the time to access the pages.

Tyagi et al. in [13], based on quantitative association rules (QARS) and basic users' information such as age, gender, occupation, etc., presented QARF and QARF/CF to recommend appropriate suggestions to the new users. It includes two phases of online and offline. Before offline phase, it categorizes the users' quantitative features. The items are also categorized in 4 groups based on the average of the ranks. Then during transformation process, the users' information are converted into binary format. In offline phase of association rules, some of the users' information are extracted using ASARM algorithm [14].

In online phase, to recommend new suggestions, if QARF method is used, using the basic rule derived from the previous step, the rules which their antecedent section is similar to the antecedent section of the basic rule selected in offline phase, and also which meet the minimum of quality, are selected and the rank of items are calculated. If the recommendations are based on QARF/CF method, the recommendations are a combination of the sum of the weight of the produced recommendations through QARF, and CF Pearson standard technique.

Regarding the fact that in web usage mining, association rules which are time-oriented in nature can produce useful knowledge about how the associations take place, and also to resolve the problem of loss of rules in border-crossings of fuzzy sets which were less supported in former fuzzy association rules mining algorithms, Mathews et al. in [15] proposed an algorithm which was based on genetic algorithm and made use of linguistic representation to discover rules in fuzzy sets borders. Due to its usage of representation graph, and also improving fitness function, this algorithm has improved in comparison with previous models. It has also made it possible to extract the rules which were lost in former approaches. This algorithm can be recommended as a complementary to the available algorithms in order to enable us to discover extra rules.

To mine association rules, Rodriguez et al. in [16], used different similarity functions which are different from the similarity functions which make use of the equality of direction of comparing two objects. Therefore, the extracted repeated patterns are also called repeated similarity patterns. This process has two steps. The first step is mining repeated similar patterns, and the second is producing interesting association rules out of repeated similar patterns. There are two algorithms presented for the first step: STreeDC-Miner algorithm to mine repetitive similar patterns when the similarity function holds the features of f-package down. And STreeNDC-Miner to mine repetitive similar patterns when the similarity function does not hold the features of f-package down. In the second step, Gen Rule Algorithm is adapted from the repetitive similar patterns in order to produce association rules.

4. THE PROPOSED ALGORITHM

The proposed method of this paper is implemented in online and offline phases. In the offline phase, at first server registries are preprocessed, and then the users' sessions are extracted. To produce the users' movement patterns, an algorithm is presented which is based on FCM fuzzy clustering, weighted association rules, and fuzzy inference system. After making movement patterns, the recommender system in online phase expands a list of the user's favorite webpages which he/she has not visited so far using HITS algorithm, and recommends it to the user. It is worth mentioning that the pages of the site are grouped in terms of their contents; the groups are computer, economic, spatial, historical, entertainment, and etc. The proposed algorithm includes the following steps:

1. Preprocessing of the server registries in order to extract the user's sessions. Preprocessing of the server registries include data cleaning, user identification, and session identification.
2. Dividing the obtained sessions into two categories of educational and experimental sessions.
3. Vectorizing the test sessions, and weighting the pages observed in sessions.
4. Applying FCM fuzzy clustering in order to cluster the similar sessions.
5. Extracting weighted association rules out of each cluster.
6. Making membership functions.
7. Using fuzzy inference system in order to predict the relevant cluster or clusters. If the current session is placed in the relevant cluster or clusters, the algorithm continues, otherwise it goes back to level 4.
8. Extracting weighted association rules with the highest similarities with the user's current session.
9. Producing a list of the user's favorite webpages which he/she has not visited so far and intends to visit, and expanding these pages using HITS algorithm.
10. Recommending the current user a list of suggested pages with the highest recommendation scores.

4.1 Clustering the Web Sessions

In order to determine the number of clusters in this algorithm, the (FCM) c-means fuzzy clustering is used to cluster the obtained sessions and the method which is suggested in [17] is made to be used. The users' sessions are clustered based on the similarities among the visited webpages. After applying the FCM method to the obtained sessions, several cluster centers are resulted as C= $\{c_1, c_2, \ldots, c_m\}$ in which every c_i $(1 \le i \le k)$ indicates a subset of the users' sessions, and k indicates the number of clusters. Then the degree of every user's attachment to every cluster is presented. Since the pages of the site are grouped in 6 categories, which include computer, economic, spatial, historical, entertainment, etc., and each cluster is shown through the average of the users' behavior of that cluster towards the defined groups.

4.2 Fuzzy Rules Inference

After obtaining the center of clusters and the degree of attachment of every profile to every cluster, the membership functions using the trapezoidal distribution is obtained. Then using fuzzy inference system based on Mamdany method, fuzzy rules are extracted. Since Mamdany is comprehensible, it is an appropriate method in the proposed system Therefore, by using fuzzy rules a user can be assigned to one or more clusters according to his/her interests. In our system, the relationships between the users and the cluster which is belongs is described by fuzzy rule. In these rules, the antecedent section contains the users' interests in each defined group of the pages, and the consequent section of the rules is the cluster or clusters of which the users are members.

4.3 Extracting Weighted Association Rules of Each Cluster

In this section, using the new presented fuzzy weighted benchmark a set of weighted rules from every cluster which include the navigation information of the users having similar navigation behavior and interests is extracted. Weighted association rules mining enables us to assign different weights to the items of the transactions, which is an appropriate method to improve the personalization process based on association rules. In this model, the larger weights indicate more important items, which provides extraction of more important items with fewer number of repetitions. Based on the steps in preprocessing of the registry files of the users' access, eventually a set of m pages $P = \{p_1, p_2, \dots, p_m\}$ and a set of user's n transactions $t = \{t_1, t_2, \dots, t_m\}$ are produced, in which every $t_i \in T$ is a subset of P. The user's every session is a vector with m dimensions of page and weight, which is represented as 8:

$$S = \{(p_1, w_1), (p_2, w_2), \dots, (p_m, w_m)\} \quad (8)$$

The priori algorithm was extracted to weighted items and its relevant definitions, and using them, weighted association rules were expanded from every cluster, and in the final section, the use of the extracted rules to recommend pages is made to the users. In this model, the weighted association rules as indicated in 9 are defined.

$$r = \langle (p_1, p_2, \dots, p_k), (q_{k+1}, q_{k+2}, \dots, q_{k+m}), (w_1, w_2, \dots, w_{k+m}), \delta, \alpha \rangle \quad \epsilon \quad R(9)$$

In this relation, $(p_1, p_2, \dots, p_k)$ is the antecedent and $(q_{k+1}, q_{k+2}, \dots, q_{k+m})$ the consequent of the weighted association rules, and $(w_1, w_2, \dots, w_{k+m})$ is the corresponding weight of every page. δ Indicates weighted support coefficient and α is the weighted confidence coefficient.

4.4 Webpage Recommendation

The To recommend webpages to the users based on weighted association rules, after improving the current user's session and weighting the pages visited by the current user according to the weighting processes done in the users' sessions, an appropriate length of the sequence of the pages viewed in the current session are selected which is used to participate in recommending pages. The former methods used a sliding window with a constant length on the user's current session, and offered their suggestions based on that. The sliding window mechanism is that it only considers the last w pages viewed during the session, and allows just the last w pages to have a role in recommendations. But using this method alone is not enough. Since it does not consider the different importance of different pages for the user, and all recent w pages viewed by the user have the same values, even if these pages are not of use to the user. The appropriate method is defined as follows: the pages which were not of use to the users and have not met the user's required information are omitted from the current session, and the process of recommendation takes place based on the pages which are important to the user and meet his/her needs. Therefore, the "place-page" and "weight-page" parameters is used in the user's current session simultaneously for selecting pages of the user's current session which must influence webpage recommendation. Then after assigning the user to the relevant cluster or clusters, the highest degree of the user's attachment to the identified clusters are obtained and select it as the "target cluster" and follow the procedures of discovering patterns using the weighted association rules extracted from the target cluster. To recommend pages with the highest similarity to the user's current sessions, instead of binary correspondence between the user's current session and the antecedent section of the weighted association rules extracted from the cluster, the degree of similarity of the user's current session (s) are calculated and the antecedent section of the weighted association rules (r_l) extracted from the target cluster. To do this, the formulae 10 and 11 [18] are used.

$$\text{Dissimilarity}(s, r_l) = \sum_{i:r_{li}} \left(\frac{2*(w(s_i) - w(r_{li}))}{(w(s_i) + (w(r_{li})} \right)^2 \quad (10)$$

$$Matchscore(s, r_l) = 1 - \frac{1}{4} \sqrt{\frac{\text{Dissimilarity} \quad (s, r_l)}{\sum_{i:r_l > 0} I}} \quad (11)$$

The aim of personalization system is to calculate a proposed set for the user's current session, which has the highest correspondence with the user's interests. To this end, scores are calculated for all the pages the user has not visited so far. While determining the score of the suggestion, the rate of the correspondence between the user's current session with the mentioned association rules which were examined in relation (11), and the weighted confidence coefficient are calculated based on relation (12):

$$Rec(s, x \to y) = Matchscore(s, x) \quad * \quad wconf(x \to p) \quad (12)$$

After calculating the score of the suggestion, m pages with the highest scores are ranked, and produced as the initial set of recommendations.

4.5 Expansion of the Proposed Pages via HITS Algorithm

As the number of the proposed pages increases, the precision decreases. The site structure is used to resolve this problem and the problem resulting from the recently added webpages which are nonetheless not present in the log file. The initial proposed webpages which were produced in previous section are selected as the root of the neighborhood graph in HITS algorithm [19]. Later, this root is complemented by its neighbors. Neighbors are a series of pages with which the root is linked, or the neighbors are linked with the root. Then, using HITS algorithm, we calculate degree of reliability and centrality for each node in neighborhood graph. The new pages and pages with low frequency are less observed in log file because there are not many links to them. Therefore they do not have a high credibility. So the nodes are arranged based on centrality score, and only the webpages with high centrality points are recommended to the user. The factor of the proposed threshold limit are used to control the number of proposed pages.

5. APPLYING AND EVALUATING THE PROPOSED ALGORITHM

Matlab software are used to apply proposed method. The standard data of Nasa website are used to evaluate the presented algorithm. The users' sessions in Nasa website are examined for two weeks in 1995. Then the users' sessions, vectorization, and weighting sessions are done, and the total data are divided into experimental and educational categories. The proposed system are Taught using the educational data. After educating the system, The experimental data are used which were not responsible in making the movement pattern, and simulated the active user. The aim of personalization is to calculate a proposed set (rs) for the user's current session which has the highest correspondence with the user's interests. This part is the only online component of the system and must be highly efficient and precise. If (rs) indicative the pages viewed by the user during the real session. Precision and coverage are effective parameters on the efficiency of the system, which are obtained using formulae (13) and (14). Precision shows the ability of the recommender system to suggest accurate suggestions, and coverage shows the ratio of the appropriate suggestions to the total webpages viewed by the user.

$$Precision(rs, rp) = \frac{|rs \cap rp|}{|rs|} \quad (13)$$

$$Coverage(rs, rp) = \frac{|rs \cap rp|}{|rp|} \quad (14)$$

6. COMPARING THE PROPOSED ALGORITHM WHIT OTHER METHODS

In this section, as shown in figures 1 and 2, the precision and coverage of the proposed algorithm are compared respectively, with the performance of the recommender system based on Diffusion Limited Aggregation [18] and Association Rules [20], and the number of pages with the window length of 4. As observed, as the number of the proposed pages increases, precision decreases and coverage increases. As it can be seen in Figure (1), except in the number of proposed pages of 3 and 6 pages, in which WAC method has higher accuracy than the proposed algorithm, in other parts the proposed algorithm is more accurate compared to the other two methods.

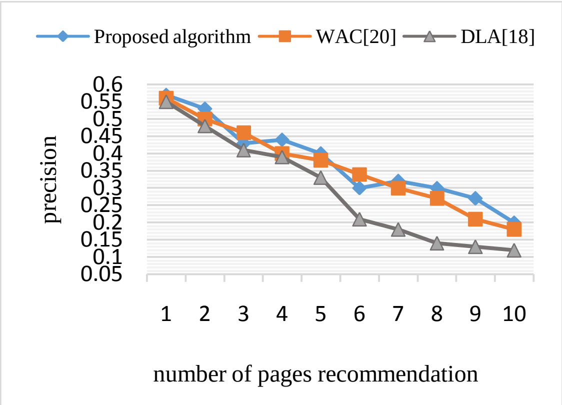

Figure 1. Comparison of the Precision Algorithms

Figure (2) indicates the coverage of the algorithms being compared with one another. By increasing the number of proposed pages in all areas, the proposed algorithm, as the figure (2) shows, has a greater coverage in comparison with the DLA method. Such improvement is also true compared with the WAC method except in the range of 8 to 10 pages.

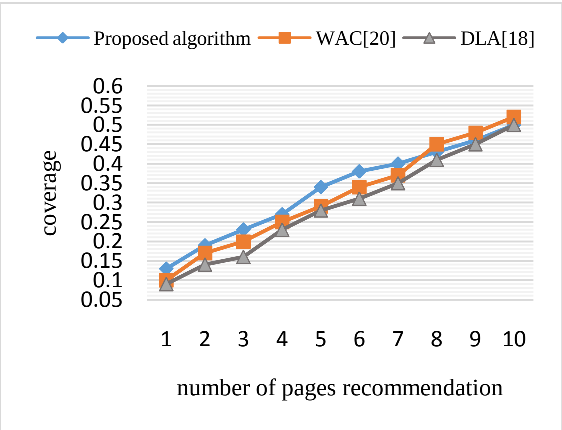

Figure 2. Comparison of the Coverage Algorithms

Figure (3) indicates the performances of the systems under study which have been compared with each other with regard to changes in the harmonic mean of precision and coverage in the standard format F. As it can be seen, the proposed method outperformed the other methods. The proposed system has created an average of %3 improvement regarding the standard F compared to the other two methods.

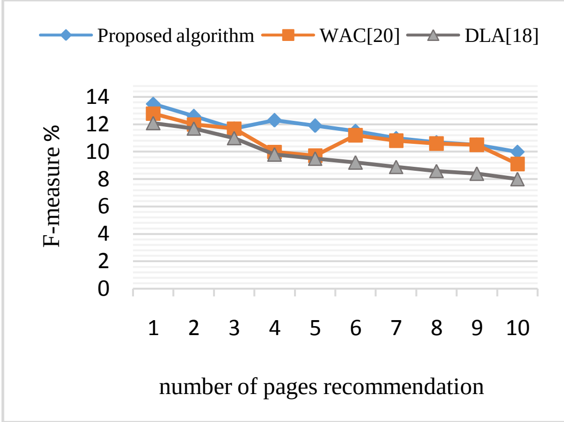

Figure 3. Comparison of the F-measure Algorithms

7. RESULTS AND SUGGESTIONS

In this paper, a new hybrid algorithm is presented, which is a combination of fuzzy clustering, weighted association rules, and fuzzy inference system. In this algorithm, the sessions of the users who have similar navigation behavior are clustered using FCM fuzzy clustering. Then, it extracts weighted association rules from every cluster using the developed a priori algorithm whose items we expanded to weighted items. In recommending pages to the current user which is done online, it assigns the current user's session to a cluster using fuzzy inference system. Then it extracts the weighted association rules which have the highest similarity with the user's current session, and produces a list of pages with the highest recommendation scores and expands these pages using HITS algorithm, and then recommends a list of pages with the highest recommendation score to the current user. The results show that the proposed algorithm has higher coverage and precision comparing with other algorithms. Also the proposed system has created an average of %3 improvement regarding the standard F compared to the other two methods. It also provides the possibility of the presence of the pages which have been recently added or have not been recorded in log file. It is recommended for further research to use improved algorithms in order to produce association rules. It is also recommended to make use of evolutionary algorithms to cluster the sessions, and use neural networks in order to learn the users' navigation patterns.

8. REFERENCES

[1] Achananuparp, P., Han, H., Nasraoui O. and Johnson, R., 2007, Semantically Enhanced User Modeling, Proceedings of the 2007 ACM Symposium on Applied Computing. SAC '07. ACM, New York, NY, pp. 1335-1339.

[2] Nasraoui, O., Soliman, M., Saka, E., Badia, A. and Germain, R., 2008, A Web Usage Mining Framework for Mining Evolving User Profiles in Dynamic Web Sites, IEEE Transactions On Knowledge And Data Engineering, Vol. 20, pp. 1041-4347.

[3] Eirinaki, M., and Vazirgiannis, M., 2003, web mining for web personalization, ACM Transactions on internet technologies(ACM TOIT), NY, USA, Vol. 3, issue 1, pp. 1-27.

[4] sirvastava, J., cooley, R., Deshpande, M. and tan, P., 2000, web usage mining: discovery and applications of usage patterns from web data, SIGKDD explorations, Vol. 1, issue 2, pp. 12-23.

[5] ishikawa, H., nakajima, T., mizuhara, T., yokoyama, S., nakayama, J., ohta, M. and katayama, K., 2002, an intelligent web recommendation system: a web usage mining approach," in: ISMIS, Vol. 2366, pp. 342-350.

[6] burke, R., 2002, hybrid recommender systems: survey and experiments, user modeling and user-adapted interaction, Vol. 12, issue 4, pp. 331-370.

[7] NetSizer: Main Page, 1998, http://www.netsizer.com.

[8] Demiriz, A., September 2004, Enhancing Product Recommender Systems on Sparse Binary Data, published in the Journal of Data mining and Knowledge Discovery, Vol. 9, Issue 2, pp. 147-170.

[9] Han, J., 2000, Data Mining: Principles and Concepts, Morgan Kauffman Publisher.

[10] Agrawal, R., and Sriknat, R., 1994, Fast Algorithms for Mining Assocaition Rules, IBM Research Report RJ9839, IBM almaden Research Center, pp. 1-32.

[11] Olson, L.D., Li, Y., 2007, Mining Fuzzy Weighted Association Rules, Proceedings of the 40th Hawaii International Conference on System Sciences, ISSN 1530-1605, 53 p.

[12] Makker, S., and Rathy, R.K., June 2011, Web Server Performance optimization using prediction prefetching Engine, International Journal of Computer Applications, Vol. 23, No.9, pp. 19-24.

[13] Tyagi, S., Bharadwaj, K.K., 2012, Enhanced New User Recommendations based on Quantitative Association Rule Mining, The 3rd International Conference on Ambient Systems, Networks and Technologies (ANT), Vol 10, pp. 102-109.

[14] Li, J. & Za¨ıane, O.R., 2004, Combining usage, content, and structure data to improve web site recommendation, 5th International Conference on Electronic Commerce and Web Technologies, Springer Berlin, Heidelberg, Vol. 3182, pp. 305-315.

[15] Matthews, S.G., Gongora, M.A., Hopgood, A.A. and Ahmadi, S., 2013, Web usage mining with evolutionary extraction of temporal fuzzy association rules, knowledge-based systems, Vol. 54, pp. 66-72.

[16] Rodríguez-González, A.Y., Martínez-Trinidad, J.F., Carrasco-Ochoa, J.A. and Ruiz-hulcloper, J., 2013, Mining frequent patterns and ssociation rules using similarities, International Journal of Expert Systems With Applications, Vol. 40, pp. 6823-6836.

[17] Fukuyama, Y. and Sugeno, M., 1989, A new method of choosing the number of clusters for the fuzzy c-means method, In: proceedings of fifth fuzzy system symposium, pp. 247–50.

[18] Forsati, R. and Meybodi, M.R., 2010, Effective page recommendation algorithms based on distributed learningautomata and weighted association rules, Expert Systems with Applications, Vol. 37, pp. 1316-1330.

[19] Kleinberg, J.M., 1999, "uthoritative sources in a hyper-linked environment, Proc. ACM-SIAM Symposium on Discrete Algorithms, Also appears as IBM Research Report RJ 10076(91892), Vol. 46 Issue 5, pp. 604-632.

[20] Singhal, V. and Pandey, G., May 2013, A Web Based Recommendation Using Association Rule and Clustering, International Journal of Computer & Communication Engineering Research (IJCCER), Vol. 1 , Issue 1, pp.1-5.