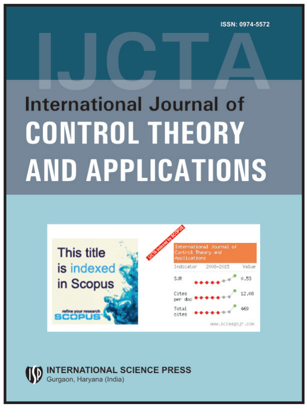

International Journal of Control Theory and Applications

ISSN : 0974-5572

© International Science Press

Volume 10 • Number 12 • 2017

A Personalized Web Search System using Frequency Based Ranking Method (FBRM) for Web Log

D. Kavitha[a] and B. Kalpana[b]

[a]*Assistant Professor, Department of Computer Science, KGiSL Institute of Information Management, Coimbatore - 641035*
[b]*Professor, Department of Computer Science, Avinashilingam Institute of Home Science, Coimbatore - 641043*

Abstract: Pattern mining extends constructive tools for exploration and analysis of data. Copious efficient algorithms are available to discover various types of patterns from large datasets. Challenge lies in identifying patterns that are interesting to users. Data mining expertise and effort are a must for existing approaches, thus making it difficult for typical domain experts. To address this, a generic Frequency Based Ranking Method (FBRM) is introduced for ranking user-specific pattern. FBRM aims to rank user preferred queries by calculating the weight based on the frequency of occurrence of the queries. It provides the personalized or preferred results by reweighting the relevant results in accordance with user's interest. When the user issues the query, the search engine retrieves the set of results. From which, top K frequent patterns are selected and it serves as the initial input to the FBRM. The proposed FBRM includes certain functionalities to extract the relevant result for personalized search. Experiment and analysis shows that proposed FBRM is more accurate in predicting user preferences and prove to be efficient when compared to existing ranking methods such as Personalized Reranking Algorithm (PRA) and Ranking Model Adaptation Framework (RMAF). The ranking result of the proposed FBRM provides higher values for F-measure, Precision, and Accuracy.

Keywords: Pattern mining, FBRM, Frequent patterns, Ranking.

1. INTRODUCTION

The increase of searchable information on the web has made ranking an essential component of information retrieval systems. Preference to the search tool is based on the quality of ranking order. The task of ranking is to collate the retrieval results as per the given criteria of providing most relevant documents to user needs. Information Foraging Theory states, people must assess the relevance or efficacy of information based on existing cues, like bibliographic citations, abstracts, keywords, titles, etc. Ranking predicts how the relevance of a set of documents to a given query is judged by users. The perception is that ranking is viewed as a subjective service which is strongly related with specific users. For example, user X_a is interested in locating the most classical paper in a specific domain in an attempt to understand the source and fundamental problem description. While the other user X_b is keen to identify the most recent publications to determine recent research trends and new findings. Though both users issue the same query, they have diverse preferences (X_a focus the number of citations feature while X_b focus the date feature). Subsequently, we need a customizable ranking mechanism to better serve users with diverse interests.

To discover user preferences through log mining, web server logs serves as a significant resource. Log mining is extensively used in Web personalization, recommender systems, and Web site design and evaluation [1-3]. Information like IP addresses, time stamps, and requested pages can be extracted from Web logs, which is in turn applied in Web application to deduce hidden user feedbacks such as motivations, goals and preferences. Considerable research to study hidden feedback to improvise ranking has been done to log mining and machine learning technologies [4-6]. Some ranking methods taken for discussion are as follows.

In search engines, documents are ranked by applying machine learning technology [4-5]. To train ranking models and globally optimize the search results, click stream data extracted from search engine query logs are used. They are useful for static ranking and to show an increase in relevance compared to Page Rank, Rank Net, an algorithm based on neural network back-propagation, is trained on the web page features [4].

Personalized ranking research has focused on "personalization vector" of the PageRank algorithm in recent times. Based on URL features, User profiles are used to find a geographically biased PageRank [6]. Papers focusing on the scalability and performance of personalized PageRank algorithms using graph-mining methods to generate personalized views of document importance do exists [7]. Algorithms taken for survey present a resounding perspective on personalization. Expanding this to a large number of categories is difficult, since these algorithms require a pre-computed set of ranking scores for each category. Preference learning [8], a research area encompassing several tasks related to learning preferences within the field of machine learning is widely used. An instance of the object ranking problem acquires ranking functions from sample orders [9]. To improve the top results of search engines, a number of general heuristics were developed in the recent works [9-10]. Methods that specifically target object ranking algorithms exploit probabilistic models of a document collection or relations between documents [11-12]. A theoretical analysis of query complexity of active object ranking has been presented recently [13].

Web systems utilize the User Relevance Feedback to interpret the user's information needs. Vector space model computes the similarity between the query and the document and is based on the terminological overlap between them [14]. Relevance Feedback requires the user to classify the documents in to relevant and irrelevant groups. Rocchio algorithm is used to expand the queries from the feedback thus obtained. Users are generally reluctant to provide information on whether they are interested with a particular document or not, so relevance feedback is not satisfying mechanism to fulfill the user needs. Historical query logs are learned and from which the results are optimized so that user intended pages are ranked higher. Queries from the logs are clustered using the similarity function and the sequential patterns from the selected web pages are captured and based on the patterns the results are re-ranked[15]. Similarly the frequent phrases from the past queries are obtained using frequency mining based algorithm and accordingly the appropriate results are re-ranked [16]. In [14][17], various web mining techniques are widely used for search result personalization. Weighted URL Ranking algorithm is used to rank the web search results based on the features extracted from hyperlinks, anchor terms and user interested domains. The retrieved results from the search engines are weighed according to the occurrence of tokens and are again weighed in accordance with the user interested domain and the same are retained for re-ordering the results according to the match with the query weight. However effective Weighted URL Ranking algorithm, to achieve fully visualized pattern in Web search practice, there are still several challenges in applying the principles. However, such work has not modeled the preferences for each individual user to provide personalized ranking.

In this work, we propose a personalized Frequency Based Ranking Method (FBRM) using web log mining results of a large-scale AOL's research team. The solution is based on the observation that a user's personalized frequency which can then be decomposed into a set of unit preferences, each of which is placed on specific data features. From click stream data harvested from Web logs, we periodically compute and update user preference vectors in a given feature space so that the ranker can always be used to describe a user's needs as exhibited in recent logs. This work proposes a personalized FBRM to improve the accuracy in predicting user preferences. The patterns of user click stream history that indicate users' subjective judgments of document relevance and a method of efficiently extracting and summarizing user preferences from large volumes of Web logs are identified. The results of extensive experiments on a real-world data of AOL's research team are presented to show the applicability and quality.

2. PROPOSED FREQUENCY BASED RANKING METHOD (FBRM) METHODOLOGY

The impulsive growth of information on the internet has attracted an enormous variety of users. Though the search engines present a well organized way to search the relevant information from the web, results acquired might not always be helpful to the users, as it fails to identify the user intention behind the query. A query means different things in varying context and the user alone can interpret the anticipated context. The web users are not satisfied with the search results in spite of recent development on web search technologies. Therefore, the requirement arises to have personalized web search system which could produce output as highly ranked pages suitable to the users. A personalized web search has various levels of efficiency for different users, queries and search contexts.

To solve the problem of search result personalizatione this problem, various web mining techniques are used [18]. Weighted URL Ranking algorithm ranks the web search patterns based on the features extracted from hyperlinks, anchor terms and user interested domains. The patterns results are weighed as per token occurrence, and in accordance with the user interested domain, the same are retained for re-ordering the results according to the match with the query weight. Though effective, Weighted URL Ranking algorithm still has several challenges in applying principles, to achieve fully visualized pattern in Web search practice. First, a user session may contain more than two queries, but the principles formulated focus only on two queries. Second, given a query and its patterns, there might be multiple principles applicable, there exists problem in jointly executing it. Third, user sessions contain rich information. Factors like the positions of the documents returned by the search engine, the terms shared by the current query and the previous, etc., may all be helpful in ranking documents. Incorporation of these factors into a ranking model remains a challenge. However, the problem of identifying user interested patterns is yet to be addressed. The Weighted URL Ranking algorithm ranking is done based on consideration of the features extracted from hyperlinks and not for patterns.

Problem formulation: A pattern consists of a description p and the subset of S defined by this description. In this paper work focus on the one particular pattern mining setting that is Subgroup Discovery. Subgroup Discovery is a supervised pattern mining task concerned with finding subsets of a dataset that have a substantial deviation in a property of interest as compared to the entire dataset, with a strong emphasis on obtaining comprehensible descriptions. Formally it is defined as follows. Let $A = \{A_1, ..., A_{1-1}; A_1\}$ denote a set of attributes, where each attribute A_j has a domain of possible values $Dom(A_j)$. Then a session $S = (s_1, ..., s_n) \subseteq Dom(A_1), ..., Dom(A_l)$ is a bag of tuples over A. The attribute A_1 is a binary target attribute, i.e. the property of interest, while the other attributes are description attributes. A subgroup description p is a conjunction of boolean atoms over description attributes, e.g. $A_1 = a \wedge A_2 > 0$. The Subgroup Discovery problem is defined as follows: given a session S, a quality measure Weight of pattern (W_{FP}), and an integer SOD(K), find the set of W_p highest-quality subgroups. Rely on the following assumption to define the problem: for a given session S and a Frequent pattern FP, there exists a ranking R of all Frequent patterns in FP according to their subjective interestingness for the current user, i.e. $p_i \succ p_j$ implies that the user considers the frequent pattern fp_i more interesting than fp_j. Also assume that R is consistent with the structure in session S and FP, i.e. that relative positions of frequent patterns are determined by their observable properties. Obviously, the user cannot generate the entire ranking R explicitly. According to their relative interestingness, small sets of patterns could be ranked. The problem is then defined as follows: given a session S and a set of sample pattern rankings F fully consistent with R, learn a ranking function R over frequent patterns in P so that the learned ranking is maximally consistent with R. The number of sample rankings that the user has to explicitly provide, i.e. user effort, has to be minimized.

Frequency Based Ranking Method (FBRM)

In the proposed FBRM weight values are computed for each patterns, discovered from Enhanced Fuzzy Apriori Algorithm (EFAA). Here weight values are computed by using both frequent patterns FP = $\{fp_1, ..., fp_j\}$ and each user session S = $\{s_1, ..., s_i\}$ .The FBRM ranks user preferred queries of user specific pattern ranking functions, where a user ranks small sets of patterns and a general weight based Frequency ranking function is inferred from this work by conventional ranking methods. Present a general framework for learning pattern ranking functions and propose a number of active learning heuristics that aims at minimizing the required user effort. The proposed Personalized FBRM ranks the query from users and provides the personalized or preferred results by reweighting the relevant results in accordance with user's interest. When the user issues the query the search engine retrieves the set of results. From the results retrieved top K frequent patterns are selected and it serves as the initial input to the FBRM. The proposed Frequency Based Ranking Method (FBRM) includes the following functionalities to extract the relevant result for personalized search.

1. Set of documents that matches the user query is fetched from the search engine.
2. The terms in the initial set of documents are preprocessed and, minded frequent patterns using EFAA.
3. The result set of frequent pattern is ranked based on computed weight of pattern value.
4. Pattern analysis is done based on the Weight of the pattern (W_{FP})

In the previous stage, we extract all the frequent patterns as preferred user patterns. However, there are too many preferred user patterns and it is difficult for users to browse all the candidates. Therefore, a ranking mechanism is needed for these user query patterns. For this purpose all these user query patterns are ranked by their frequencies. Frequency Based Ranking Method (FBRM) is developed for ranking user preferred queries by calculating the weight based on the frequency of occurrence of the queries.

Weight of pattern: The top K Frequent patterns from the web server are analysed for each term W_{FP} measure is computed and the same could be retained in the W_p store. Frequent Patterns are sorted based on the W_{FP} value measured and from this the top N Frequent patterns with higher weights are used for further processing and it is considered as important Frequent patterns, thus results as user preferred query. From the above Frequent patterns set, the identical Frequent patterns in all documents are collected and their weights are added up and from the outcome the higher weighted terms are again selected for building the personalized user preference.

$$\text{Weight of the pattern } (W_{FP}) = \text{No. of occurrence/Sum of distance} \tag{1}$$

$$\text{SOD(K)} = \text{Sum of distance} = \sum_{i=1}^{N-1} \left| p(i)\, p(i+1) \right| \tag{2}$$

here p is the position in a session S. The user preferred query frequent pattern score(PUQ_{FP}) is calculated by the summation of the edit distance of each frequent pattern in the dataset to the closest one in the top result. The score is normalized by the summation of the edit distance of each frequent pattern to the closest one in the dataset. The larger the weight value score is, the more representative the result list. Here merge the User sessions and ranking patterns we find a Prefer User Query Frequent patterns (PUQ_{FP}).

Algorithm description: In this section present an Frequency Based Ranking Method (FBRM) algorithm for learning a frequent pattern ranking function (Algorithm 1). It receives a collection of frequent patterns FP and user session log S as input. A standard Enhanced Fuzzy Apriori Algorithm (EFAA) can be used to mine the frequent patterns of user session S. It is initially ranked according to an objective measure (W_p). In practice, any measure can be used, e.g. coverage. This ranking is referred to as the source ranking and computed based on the FP results. At each iteration a subset of frequent pattern FP is shown to the user and ranked by her according to the subjective interestingness of the frequent patterns. Ranked sets of frequent patterns are used as training data for a ranking function. The computed ranking function can then be used to rank the input sessions as well as to score unobserved patterns.

Frequent Pattern representation (Line 2): In order to apply FBRM Algoitrhm , frequent patterns (FP) are represented as vectors of numeric features. Frequent Pattern features have to capture properties that make patterns interesting to a user.

FBRM learning (Line 4): Query selection methods select sets of patterns that will be shown to the user. Assuming that the query size is fixed, the goal is to minimize the number of queries required to attain a certain ranking accuracy. The methods take into account such factors as the current estimated quality of a pattern, the estimation uncertainty, the diversity of the query, or the structure of the data. Here in additon Ranking value is also computed using W_{FP}.

Prefered query frequent pattern format: A user provides her Prefer User Query Frequent patterns (PUQ_{FP}) in the form of rankings. For example, if the query is $\{q_1, q_2, q_3\}$ the PUQ_{FP} implies that q_3 based pattern the most interesting frequent pattern, and q_2 based pattern is the least interesting frequent pattern. This PUQ_{FP} format is computationally more expensive for a user than graded PUQ_{FP}, i.e. assigning scores from a predefined scale. However, argue that it has two advantages relevant to the iterative setting. First, it requires neither a deep understanding of the scale by a user, nor a thorough scoring function. Second, graded PUQ_{FP} can be converted to the ordered format, albeit at a cost of reduced granularity.

Learning rankings (Line 6) Ranked queries, i.e. sets of frequent patterns that are ranked by the user according to their Weight of the pattern (W_{FP}), are used as training data for an object ranking algorithm. It learns a ranking function R that returns a number for any possible queries as PUQ_{FP}. Any freqeunt pattern can be scored using R, therefore the ranking function is essentially a general important factor measure.

Stopping criteria (Line 7) Stopping criteria can consider marginal effects of additional queries on the learned ranking or limit the maximal user effort. In the simplest case, the user manually stops the algorithm, as soon as she considers her information need satisfied.

Algorithm 1: Frequency Based Ranking Method (FBRM)

Input: user Sessioned logs S, Discovered frequent Pattern (FP = fp_1, ..., fp_n)

Output: Ranking Function (R) for frequent Pattern (FP) over S

1. PUQ = ϕ, R = Source Ranking (W_{FP})
2. FPV = Convert to Vector (FP, S)
3. Repeat

4. q = select query (FPV, R)
5. PUQ = PUQ $\cup$ get (PUQ_{FP})
6. R = Learn Ranking Function (FPV, PUQ)
7. Until stopping criterion is met
8. Sort and return R
9. Results PUQ_{FP}

3. SIMULATION RESULTS

The evaluation of proposed FBRM is compared to existing ranking methods such as Personalized Reranking Algorithm (PRA) and Ranking Model Adaptation Framework (RMAF). FBRM is evaluated for frequent itemset mining and is compared to the existing methods such as PRA, RMAF methods. The study of personalized ranking helps to find the way to create the framework for ranking the web results based on user search behaviour in [19-20]. The concepts in [19-20] help to study in case of how to maintain the user search history and web directories in order to analyse it for user satisfactory results. The dwell time metric and other web based evaluation of user interests is obtained from [13] and to mine the user tasks based on several web based metrics.

The various concepts of ranking methods are studied from the literature discussed in introduction and that also tells a way to rank effectively to prioritize the web results. The idea of personalization of user search is studied in the previous work. Thus previous studies [19-20] clearly provides an idea to explore the personalize user search behaviour effectively and dynamically rank web results. So it used for proposed work comparison since the proposed FBRM is experimented and evaluated to dynamic web log session. AOL's research team has published a huge data collection of 20,000,000 search queries from 650,000 users sampled over three months for the public to see, dig around and analyse. This collection consists of ~20M web queries collected from ~650k users over three months. Where the data is sorted by anonymized user id: The data set includes (UserID, Query, QueryTime, ClickedRank, and Destination DomainUrl). The goal of this collection is to provide a real query log based on users. In initial stage, preprocessing is done to reduce the repetition of the queries for same users. The preprocessed dataset is categorized into four dataset D1, D2, D3 & D4 for testing purpose.

Characteristics of each test data set are shown Table 1.

Table 1
Testing data sets

Name	*Number of Samples*	*Number of sessions*	*Average session length*
D_1	3000	64	46.9
D_2	2500	55	46.5
D_3	4000	116	34.5
D_4	5000	151	33.1

In order to measure the performance accuracy of the ranking methods, F-measure, Precision, and accuracy are used.

F-measure: The mean average value of precision and recall results of test sequence is named as the F-Measure. The precision is defined as the percentage of the truly identified web session correctly to the entirety number of estimated web session. The recall is defined as the percentage of truly identified web session correctly, so the F-measure is defined as follows:

$$\text{F-measure} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \tag{3}$$

The following Figure 1, Figure 2 and Figure 3 illustrates the pattern analysis results of the ranking methods for, Precision, F-measure and Accuracy. By Merging the User sessions and frequent patterns, Preferred User Query patterns are ranked. The ranking result of the proposed Frequency Based Ranking Method (FBRM) provides higher values for, Precision, F-measure, and Accuracy. The proposed FBRM is compared to existing ranking methods such as PRA and RMAF.

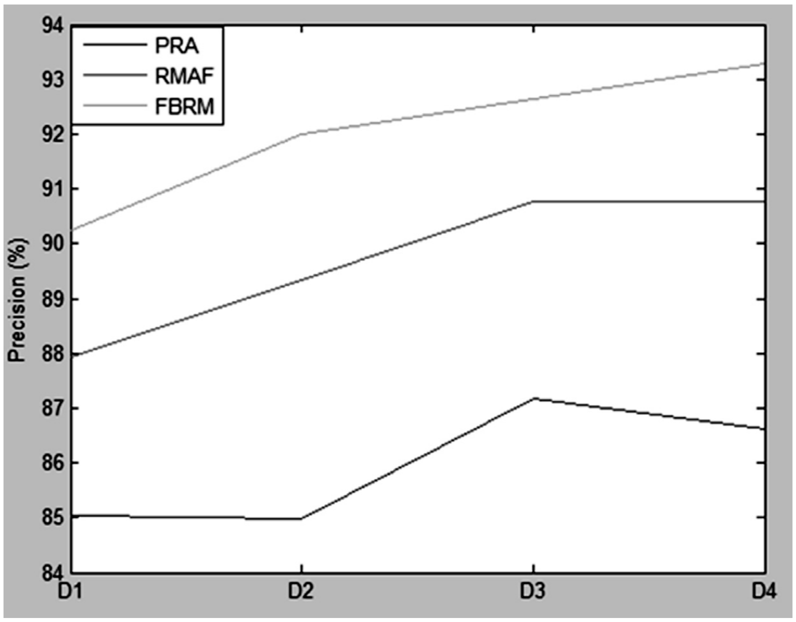

Figure 1: Pattern Evaluation Graph of ranking methods under precision

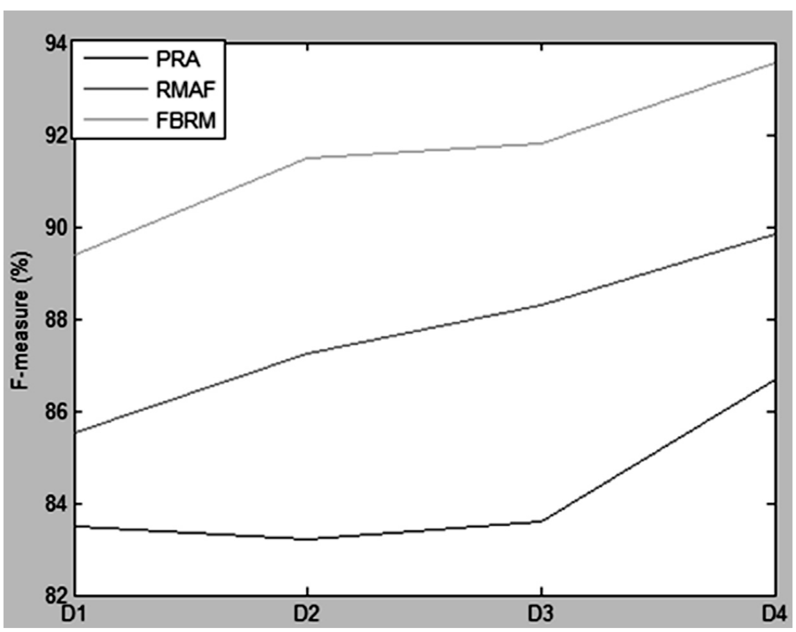

Figure 2: Pattern Evaluation Graph of ranking methods under F-measure

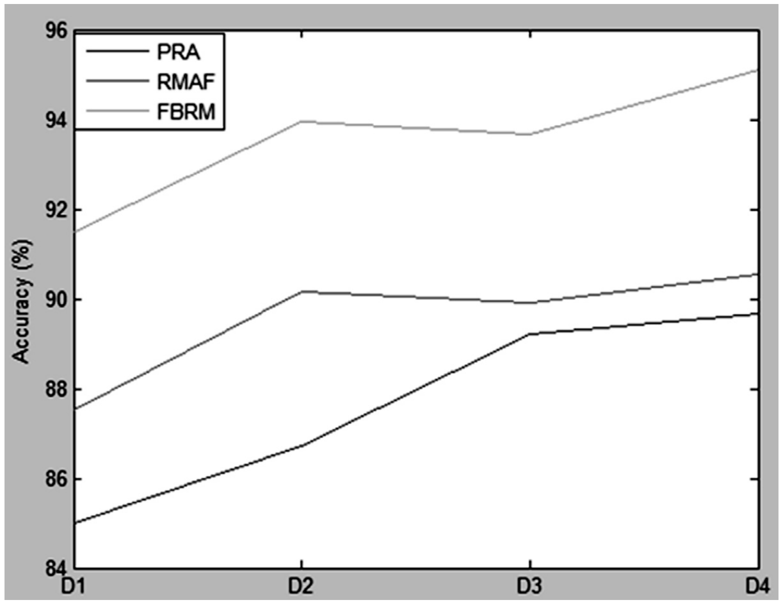

Figure 3: Pattern Evaluation Graph of ranking methods under accuracy

Table 2 shows the pattern analysis results of the ranking methods for F-measure, Precision, and Accuracy. The ranking result of the proposed Frequency Based Ranking Method (FBRM) provides higher values for F-measure, Precision, and Accuracy. The proposed FBRM is compared to existing ranking methods such as PRA and RMAF. Table 2 shows that the proposed FBRM achieves 91.56193% F-Measure, existing RMAF and PRA methods achieves 87.7358% and 84.24093% F-Measure respectively. Also, proposed FBRM achieves 92.04573% Precision, existing RMAF and PRA methods achieves 89.69983% and 85.95163% Precision respectively. The proposed FBRM achieves 93.54015% accuracy, existing RMAF and PRA methods achieves 89.54423% and 87.6538% accuracy values respectively. From the results it concludes that the proposed FBRM performs better for all parameters.

Table 2
Evaluating the effectiveness of ranking methods

Dataset	*Ranking methods*	*Parameters (%)*		
		F-Measure	*Precision*	*Accuracy*
D_1	PRA	83.4827	85.036	85.0035
	RMAF	85.5326	87.932	87.539
	FBRM	89.3821	90.255	91.475
D_2	PRA	83.21	84.989	86.722
	RMAF	87.2626	89.3311	90.152
	FBRM	91.499	92.0113	93.944
D_3	PRA	83.581	87.1715	89.2176
	RMAF	88.313	90.7625	89.9276
	FBRM	91.7906	92.6323	93.6516
D_4	PRA	86.69	86.61	89.6721
	RMAF	89.835	90.7737	90.5583
	FBRM	93.576	93.2843	95.09
All	PRA	84.24093	85.95163	87.6538
	RMAF	87.7358	89.69983	89.54423
	FBRM	91.56193	92.04573	93.54015

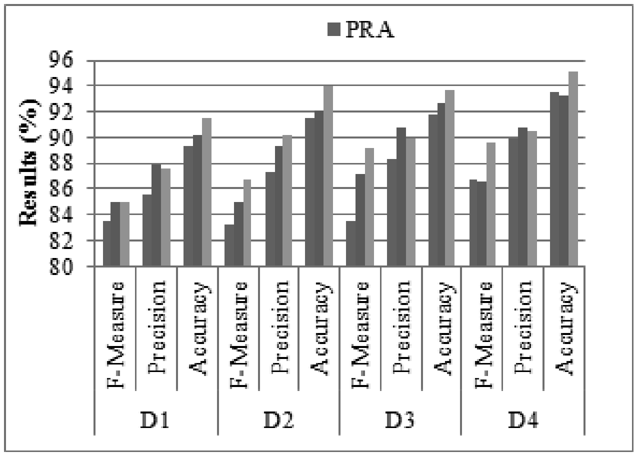

Figure 4: Pattern Evaluation Graph of ranking methods under study

Figure 4 illustrates the pattern analysis results of the ranking methods for F-measure, Precision, and Accuracy. By Merging the User sessions and frequent patterns, Preferred User Query patterns are ranked. The ranking result of the proposed Frequency Based Ranking Method (FBRM) provides higher values for F-measure, Precision, and Accuracy. The proposed FBRM is compared to existing ranking methods such as PRA and RMAF. Figure 5 shows that the proposed FBRM achieves 91.56% F-Measure which is higher by 3.83% and 7.32% when compared to RMAF and PRA methods respectively. Also, FBRM achieves 92.06% Precision which is higher by 2.34% and 6.09% when compared to RMAF and PRA methods respectively. The proposed FBRM achieves 93.54% accuracy, which is higher by 3.99% and 5.88 % when compared to RMAF and PRA methods respectively.

4. CONCLUSION AND FUTURE WORK

In FBRM, user sessions and frequent patterns are used for ranking users preferred query patterns. A general weight based frequency ranking function is inferred from this work by conventional ranking methods. The proposed personalized FBRM ranks the query from users and provides the personalized or preferred results by reweighting the relevant results in accordance with user's interest. For this purpose, here user preferred query frequent pattern score is calculated by the summation of the edit distance of each frequent pattern in the dataset to the closest one in the top result. The score is normalized by the summation of the edit distance of each frequent pattern to the closest one in the dataset. The larger the weight value score is, the more representative the result list. The ranking results of the proposed Frequency Based Ranking Method (FBRM) prove to be efficient when compared to existing ranking methods such as Personalized Reranking Algorithm (PRA) and Ranking Model Adaptation Framework (RMAF). Directions for future work include investigating the effect of coarse-grained or noisy feedback on learning performance, learning preferences over sets of patterns instead of individual patterns, and shifting from the pool-based active learning to query synthesis, i.e. directly mining patterns for queries. A user study is required to evaluate the practical applicability of the proposed framework.

REFERENCES

[1] Das S, Datar M, Garg A, and Rajaram S., "Google news personalization: scalable online collaborative filtering, " *In Proceedings of the 16th international WWW conference*, pp. 271-280, New York, NY, USA, 2007. ACM Press.

[2] Spiliopoulou M., "Web usage mining for web site evaluation", *Communication* , Vol. 43, No. 8, pp. 127-134, 2000.

[3] Srivastava J, Cooley R, Deshpande M, and Tan PN, "Web usage mining: Discovery and applications of usage patterns from web data", *ACM SIGKDD Explorations Newsletter*, Vol. 1, No. 2, pp. 12-23, 2000.

[4] Richardson M, Prakash A, and Brill E, "Beyond PageRank: machine learning for static ranking", *In Proceedings of the 15th international conference on World Wide Web*, pp. 707-715, 2006.

[5] Xu J,Li H, "Adarank: a boosting algorithm for information retrieval", *In Proceedings of the 30th ACM SIGIR, ACM Press*, pp. 391-398, 2007.

[6] Xin D, Shen X, Mei Q, Han J, "Discovering interesting patterns through user's interactive feedback," In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining* ,pp. 773-778, 2006.

[7] Rueping S, " Ranking interesting subgroups," In *Proceedings of the 26th Annual International Conference on Machine Learning*, pp. 913-920.

[8] "A Study on Existing Protocols and Energy-Balanced Routing Protocol for Data Gathering in Wireless Sensor Networks" published in International Journal of Computing and Technology on Nov 10, 2013. Impact Factor 1.213 (Refereed Journal). www.cirworld.com/index.php/ijct/article/view/2780/pdf_293

[9] "Challenges and Authentication in Wireless Sensor Networks by using promising Key Management Protocols" at International Conference in Kristu Jayanthi College, Bangalore on Feb 19th & 20th 2015 and Published in International Journal of Computer Applications. Impact Factor 0.814. www.ijcaonline.org/icctac2015/number1/icctac2005.pdf

[10] Xu Z, Akella R, Zhang Y, "Incorporating diversity and density in active learning for relevance feedback", *In European Conference on Information Retrieval , Springer Berlin Heidelberg*, pp. 246-257, 2007.

[11] Radlinski F, Joachims T, "Active exploration for learning rankings from clickthrough data," In *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 570-579, 2007.

[12] Xu Z, Kersting K, Joachims T, "Fast active exploration for link-based preference learning using gaussian processes. *In Joint European Conference on Machine Learning and Knowledge Discovery in Databases* , *Springer Berlin Heidelberg,* pp. 499-514, 2010.

[13] Hu J, Chan PK, "Personalized Web Search by Using Learned User Profiles in Reranking," *In: Internal Workshop on Web Mining and Web Usage Analysis*, pp. 78-83, 2008.

[14] Dali L, Fortuna B, Rupnik J, "Learning to Rank for Personalized News Article Retrieval", In *WAPA*, pp. 152-159, 2010.

[15] Sharma AK, Aggarwal N, Duhan N, Gupta R, " Web search result optimization by mining the search engine query logs," *2010 International Conference on Methods and Models in Computer Science (ICM2CS),* pp. 39-45, 2010.

[16] Agichtein E, Brill E, Dumais S, Ragno R, "Learning user interaction models for predicting web search result preferences," *In Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 3-10, 2006.

[17] Satokar KD, Gawali SZ, " Web Search Result Personalization using Web Mining," *International Journal of Computer Applications*, Vol. *2*, No. 5, pp. 29-32, 2010.

[18] Ratnakumar AJ, "An implementation of web personalization using web mining techniques", *Journal of Theoretical and applied information technology*, Vol. 18, No. 1, pp. 67-73, 2010.

[19] Xu, S., Jiang, H. and Lau, F.C.M., "Mining user dwell time for personalized web search re-ranking," *In Twenty-Second International Joint Conference on Artificial Intelligence (IJCAI)*, Vol. 11, pp. 2367-2372, 2012.

[20] Wang H, He X, Chang MW, Song Y, White RW, Chu W, "Personalized ranking model adaptation for web search", In *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval*, pp. 323-332, 2013.