



ISSN: 2350-0328

**International Journal of Advanced Research in Science,  
Engineering and Technology**

**Vol. 3, Issue 7 , July 2016**

# **XML Retrieval Improve the Result of Search in Personalization-Result**

**Jayati Kale , Prof. Fazeel Zama**

Department of CSE, Wainganga College of Engineering, Nagpur, Maharashtra, India.

Department of CSE, Wainganga College of Engineering, Nagpur, Maharashtra, India.

**ABSTRACT:** Searching information nowadays very important for us. Visual information increases day by day and so many times we want to search something we want but result find very different. Many times, information is close to what we want but not the same or different links served by search engine. For this, we want different type of search engine to search information as per our preference or our background history. They are helping us to find exact what we want. In our personalized search engine, we used XML retrieval; it is different from our information retrieval. It can help us to find particular part of document or only the thing we want instead of all the data, we want. In personalization of search engine we used users profile information to search information. In the basis of users account, they can search information in minimum time. In personalization, it takes too much time to search because of finding exact result to users profile matching. For this problem, we use query suggestion and re-ranking for result set to minimize the time.

**KEYWORDS:** Query suggestion, Re-ranking, User profile, Personalization, xml retrieval.

## **I. INTRODUCTION**

In our everyday need, we use to search the information on search engine. We use it as blindly to search. But so many times it is very difficult to search the proper information. When we search for exact information it will show us something different than we want and it takes more time to search. Here we find the solution to search as per our requirement and as per our need. Personalization in search is different and new concept to search information in search engine. For this, we develop personalized search engine to find information as per user's requirement and their preferences. Finding information as per users requirement is based on user's profile. Users first need to create their account on that search engine and fix their preferences and their background history of them. All the data stored in database their account profile. We use XML retrieval [1] instead of our typical information retrieval because XML retrieval used to show only required part of document instead of full length document and it is also known as fastest search retrieval. With the help of XML retrieval used every user's history to search information. This can be retrieve fastest from the database to search information. Also we use XML to stored profile wise search information of users when every time user searches for information and this will be stored in user profile in XML. In this system, we use three personalization techniques named as Query Suggestion, Re-ranking Technique, and Feedback-Based Technique. Query suggestion is used to find query from the set of query which is exact similar to the keyword type by the user at the time of search. Whenever user type some keyword in search box it will automatically suggest query list in search box. In re-ranking technique, system will automatically re-rank the result set from the result and arranged them rank-wise, this will happen in very less time. This technique also called as after search because it is used after the search result-set to give ranks every result. We develop third important technique to give feedback on products and company, and they can be maintain or improve the quality of products.

To find the result within the very few time, we need to create mathematical model for quick and exact result. Today there are different types of mathematical models are available to calculate result. Different mathematical model and language model are our need to find the exact results.

## **II. METHODOLOGY**

Here are we use different methods and models to re-rank the result-set. This are used to calculate time required or taken by the system to search the query set or result for users. This mathematical model calculate different times of re-ranking , finding the category of product preferred and product search count and rating finding of the product. This can

be very useful to find calculation and to overcome time problem. There are given models are shown to calculate the result-set is as follow-

- Mathematical Model for Search Re-ranked Result-Set

$$R(Sr)=Rpref(Rsc(products)). Rprofile(Rsc(products)).Rsc(products)$$

Where R(Sr)->Search Result re-ranking

Rpref-> Preferences wise Re-ranking

Rsc->Product ratings and search count wise re-ranking

Products->Search Keywords wise matching products

- Mathematical Model for User Preferred Product Category Finding  
Searched Category wise User Preferences (in %)=SC(x)

$$SC(x) \% = \left[ \left( \sum_{i=1}^n \text{Category} [C_i=x] \right) / \left( \sum_{i=1}^n S_i \right) \right] * 100$$

Where

x → Previously Searched Product Category

S → Search Attempts

C → Product Category

- Mathematical Model for Total Product Search Count and Rating Finding
  - Product Search Count

$$SC(x) = \sum_{i=1}^n [P_i=x]$$

Where

Sc → Search Count

x → Searched products

P → Product

- Product Ratings

$$AvgR(x) = \left[ \sum_{i=1}^n \text{RatingTotal} [P_i=x] \right] / \sum_{i=1}^n [P_i=x]$$

Where

AvgR(x) → Average ratings

x → searched Product

P → Product

### III.IMPLEMENTATION STRATEGY

After searching, system will fetch keyword specific data from database and data shown to the users. It is used to Automatic Ranking Finding, which can be very useful to calculate the time taken by the system to re-rank the result-set. System will find ranking of fetched result with their categories by using XML data, this is used XML retrieval to find the fastest result. This re-ranking is important part of system because it will used to re-rank all the available result set. The algorithm of re-ranking is used to find the result-set in very few time. This algorithm used all users data available in their account to find the appropriate result. At the base of users account they can match with users all priority and their personal information search all the data. The given algorithm shows the working of this personalization method is as follow-



ISSN: 2350-0328

# International Journal of Advanced Research in Science, Engineering and Technology

Vol. 3, Issue 7 , July 2016

## Re-ranking the Result Set

Proposed Re-ranking method will work for the result-set as:-

Step 1:- Input matching result set as per specified query

Step2:- Searched Category wise User Preferences (in %)=SC(x)

$$SC(x) \% = \left[ \left( \sum_{i=1}^n \text{Category } [C_i=x] \right) / \sum_{i=1}^n S_i \right] * 100$$

Step 3:- Where x → Previously Searched Product Category

Step 4:- S → Search Attempts

Step 5:- C → Product Category

Step 6:- Fetch users profile from database

Step 7:- Gender=user's gender

Step 8:- Branch=user's branch

Step 9:- Re-rank given result set as per the branch gender of the user

Step 10:- Again Re-rank result set using product rating

Step 11:-  $R(Sr)=Rpref(Rsc(products)). Rprofile(Rsc(products)).Rsc(products)$

Where R(Sr)→Search Result re-ranking

Rpref→ Preferences wise Re-ranking

Rsc→Product ratings and search count wise re-ranking

Products→Search Keywords wise matching products

## IV. EXPERIMENTAL RESULTS

### **NDCG of P-HRR(Result):**

In this diagram, admin can see companies' product or information search by users as per the profile wise. How many profile terms used by the user at the time of search. Here users search many data at a time. All the users search result-set used to normalize query expansion. This term is used to normalize query expansion profile wise and this query expansion term used to calculate normalized data gain which can be used to calculate re-ranking of the result-set at a time. In this page actual results show to the admin who can calculate the searching time for re-ranking for result-set.



ISSN: 2350-0328

# International Journal of Advanced Research in Science, Engineering and Technology

Vol. 3, Issue 7 , July 2016

localhost:50502/NDCGReport.aspx



## Search Engine Personalization

using XML techniques

Home Change Password Logout

Logged in as webadmin(admin)

### Companies Comparison Report

Select NDCG@ : <--Select-->

NDCG@	Profile Terms	Query Expansion	Normalized Query Expansion	NDCG of Query @	NDCG of P-HRR
0	0	12.8829	18.402991986299	171	0.2449
0	1	14.5384	10.3919727960802	158	0
0	2	14.5384	10.3919727960802	158	0
0	3	16.6793	17.5800682559021	177	0
3	1	10.536	34.8180646664591	200	1
3	2	10.536	34.8180646664591	200	1.6309
3	3	9.3257	12.2776501765714	149	2.1309
4	1	8.6081	34.0319960670141	194	1
4	2	8.6081	34.0319960670141	194	1.6309
4	3	8.6081	9.35459028065439	194	2.1309
6	0	6.029	8.79811560778717	157	1
6	1	6.029	8.79811560778717	157	1
6	2	6.029	8.79811560778717	157	1.6309

Fig1: Users search as per the profile wise re-ranking report to admin(Result)

### NDCG of P-HRR for Users:

In this given, shows that NDCG@ used to how many records used to search by the users at a time. At the time of search this records increased by time to time because users search added more records in it and they can used their profile terms for search every time profile terms matches with their search and find the re-ranked result-sets what they want to find for. When users select NDCG@, they can show the record used shown to them. It can be 3,4,6,7 and so on. This field add more when user used more for search.



## Search Engine Personalization

using XML techniques

Home Change Password Logout

Logged in as jayati(customer)

### Companies Comparison Report

Select NDCG@ : 6

<--Select-->  
 3  
 4  
 6  
 7  
 Advanced Profile

NDCG@	Profile Terms	Query Expansion	Norm	NDCG of Query @	NDCG of P-HRR
6	0	6.029	8.79811560778717	157	1
6	1	6.029	8.79811560778717	157	1
6	2	6.029	8.79811560778717	157	1.6309
6	3	6.029	8.79811560778717	157	2.1309

Fig 2:Users search re-ranking as per profile

**V. CONCLUSION**

This system is proposed for personalization of searching information using XML retrieval. This system is very useful searching data while xml is powerful used to retrieve fast information from database without wasting time because in personalization, we focused on build user profile and their preferences for search but every time system compare query with users profile and it will be lengthy and time wasting. The approach of using xml retrieval in search engine is very beneficial to search fast and quick data without wasting time. It is the process where evaluation methods that can be normalized all the result set with using PHRR. PHRR used for hard re-ranking in users profile wise search for NDCG which can be used to re-rank all the result sets. At the beginning we develop this search engine for products so that at the initial level it can give proper result and for the association we need internet connection to show associated products. This type of search engine is very useful for those who like personalization in their search. So many users wish to have search engine like they can search as per their profile but no one search engine work as personalized.

**REFERENCES**

- [1] L. M. de Campos, J. M. Fernández-Luna, J. F. Huete , And Eduerdo Vicente-Lopez, "Using Personalization to improve Xml Retrieval," in IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 26, NO. 5, MAY 2014.
- [2] G. Chernishev, "Personalization of XML text search via search histories," in *Proc. SYRCODIS 2008 Colloq. Databases Information Systems*.
- [3] L. M. de Campos, J. M. Fernández-Luna, J. F. Huete , and C. Martín-Dancausa, "A content-based approach to relevance feedback in XML-IR for content and structure queries," in *Proc.Int. Conf. Knowledge Discovery Information Retrieval*, Roskilde, Denmark, 2010, pp. 418–427.
- [4] B. W. Croft, S. Cronen-Townsend, and V. Lavrenko, "Relevance feedback and personalization: A language modeling perspective," in *Proc. 2nd DELOS Workshop Personalisation Recommender Systems Digital Libraries*, 2001.
- [5] N. Matthijs and F. Radlinski, "Personalizing web search using long term browsing history," in *Proc. 4th ACM Int. Conf. Web Search Data Mining*, Hong Kong, China, 2011, pp. 25–34.
- [6] L. Meister, O. Kurland, and I. G. Kalmanovich, "Two are better than one! Re-ranking search results using an additional retrieved list," Technion - Israel Instit. Technology, Haifa, Israel, Tech. Rep. IE/IS-2009-01, 2009.
- [7] L. Tamine-Lechani, M. Boughanem, and M. Daoud, "Evaluation of contextual information retrieval effectiveness: Overview of issues and research," *Knowl. Inform. Syst.*, vol. 24, no. 1, pp. 1–34, 2010.
- [8] R. Schenkel and M. Theobald, "Feedback-driven structural query expansion for ranked retrieval of XML data," in *Proc. 10<sup>th</sup> Int. Conf. Extending Database Technology*, Munich, Germany, 2006, pp. 331–348, LNCS 3896.
- [9] H. Pan, "Relevance feedback in XML retrieval," in *Proc. EDBT Workshops PhD, DataX, PIM, P2P&DB, ClustWeb*, Heraklion, Greece, 2004, pp. 187–196, LNCS 3268.
- [10] X. Shen, B. Tan, and C. Zhai, "Implicit user modeling for personalized search," in *Proc. 14th ACM Int. Conf. Information Knowledge Management*, Bremen, Germany, 2005, pp. 824–831.
- [11] A. Sieg, B. Mobasher, and R. Burke, "Web search personalization with ontological user profiles," in *Proc. 16th ACM Int. Conf. Information Knowledge Management*, 2007, pp. 525–534.
- [12] B. Steichen, H. Ashman, and V. Wade, "A comparative survey of personalised information retrieval and adaptive hypermedia techniques," *Inform. Process. Manag.*, vol. 48, no. 4, pp. 698–724, 2012.
- [13] L. Tamine-Lechani, M. Boughanem, and M. Daoud, "Evaluation of contextual information retrieval effectiveness: Overview of issues and research," *Knowl. Inform. Syst.*, vol. 24, no. 1, pp. 1–34, 2010.
- [14] L. Zighelnic and O. Kurland, "Query-drift prevention for robust query expansion," in *Proc. 31th Annu. Int. ACM SIGIR Conf.*, Singapore, 2008, pp. 825–826.
- [15] J. Teevan, S. T. Dumais, and E. Horvitz, "Personalizing search via automated analysis of interests and activities," in *Proc. 28th Annu. Int. ACM SIGIR Conf.*, Salvador, Brazil, 2005, pp. 449–456.
- [16] J. Teevan, S. T. Dumais, and E. Horvitz, "Potential for personalization," *ACM Trans. Comput. Hum. Interact.*, vol. 17, no. 1, Article 4, 2010.
- [17] A. Trotman and B. Sigurbjörnsson, "Narrowed extended XPath I (NEXI)," in *Proc. 3rd Int. Workshop Initiative Evaluation XML Retrieval*, Dagstuhl Castle, Germany, 2005, pp. 16–40, LNCS 3493.

**AUTHOR'S BIOGRAPHY**

**Miss Jayati Kale** did her MCA from Amravati University. She is also pursuing M.Tech in Computer science and Engineering from Wainganga College Of Management And Engineering, Dongargaon, Nagpur, India. Her area of interest is Data Mining.

**Prof. Fazeel I. Z. Kureshi** is HOD of Computer Science and Engineering, Wainganga College of Management And Engineering, Dongargaon, Nagpur, India. He had completed his M.Tech from Nagpur University, Nagpur.