Computer Engineering and Intelligent Systems
ISSN 2222-1719 (Paper) ISSN 2222-2863 (Online)
Vol.6, No.8, 2015

Enhancing Information Retrieval Relevance Using Touch Dynamics on Search Engine

Athman A. Masoud* Dr. Wilson Cheruiyot Kennedy Ogada
School of Computing and Information Technology, Jomo Kenyatta University of Agriculture and Technology, P.O. Box 62000-00200 Nairobi, Kenya

Abstract
Using Touch Dynamics on Search Engine is an attempt to establish the possibilities of using user touch behavior which is monitored and several unique features are extracted. The unique features are used for identifying users and their traits according to the touch dynamics. The results can be used for defining automatic user unique searching behavior. Touch dynamics has been discussed in several studies in the context of user authentication and biometric identification for security purposes. This study establishes the possibility of integrating touch dynamics results for identifying user searching preferences and interests. This study investigates a technique of combining personalized search with touch dynamics results information as an approach for determining user preferences, interest measurement and context.
Keywords: Personalized Search, Information Retrieval, Touch Dynamics, Search Engine

1. Introduction

The enormous volume of web data and information makes it difficult for search engines to determine the exact results expected by users. Many information retrieval (IR) systems suffer from a radical variance in performance when responding to users' queries. Even for systems that succeed very well on average to solve the query ambiguity issues, the quality of results returned for some of the queries is poor. Thus, it is desirable that IR systems will be able to identify relevant queries so they can be handled properly (Carmel & Yom-Tov, 2010). Understanding why some queries are inherently more irrelevant than others through user predefined preferences and interest identification is essential for IR, and a good answer to this important question will help search engine to reduce the variance in performance.

1.2 Definitions

1.2.1Information Retrieval

Information Retrieval (IR) is the science of searching for documents for information within documents and for metadata about documents as well as that of searching relational database and World Wide Web (Manning et al, 2008).

1.2.2 Search Engine

Search Engine is a tool that let users explore the databases containing the text from hundreds of millions of web pages. When the search engine software finds pages that match the user request (often referred to as hits) it presents them with brief descriptions and clickable links for user to access the relevant information (Glossbrenner & Glossbrenner, 2001).

1.2.3 Touch Dynamics

Touch dynamics are user touch behaviors on touch screens, which are monitored and several unique features are extracted including left versus right hand dominance, one handed versus bimanual operation, stroke size, stroke timing, symmetry, stroke speed and timing regularity. The unique features are used for identifying users and their traits according to the touch dynamics (F.Eika, 2012).

1.3Background

Information retrieval (IR) is devoted to finding relevant documents, not finding simple matches to patterns. Yet, often when information systems are evaluated they are found to miss numerous relevant documents. Moreover users become complacent in their expectation of accuracy of information retrieval systems (Grossman & Frieder, 2004).

The variability in performance is due to a number of factors. There are factors related to the query itself, such as *term ambiguity*. For example, consider the ambiguous query Golf. Without any context, search engines will be unable to identify the desired information need (the sport or the car).

In such cases, poor results are expected as answers related to different meanings of the query are interleaved. Other factors are related to the discrepancy between the query language and the content language (also known as the *vocabulary mismatch* problem) when inconsistency exists between the way users express their needs and the way the content is described. Other problematical cases are missing content queries, for which there is no relevant information in the corpus that can satisfy the information needs (Carmel & Yom-Tov, 2010).

1.4Problem Statement

Large population of the mobile device users rely on touch screens and search engines to retrieve relevant information. While a number of techniques for adaptation exist (Kumaran & Allan, 2006) the study is particularly interested in those that target personalized user searching preferences. Modifications to queries, either in the form of expansion or relaxation, have been widely studied or known to contribute to significant improvements in performance. Automatic query expansion (AQE) refers to the process of including related terms to the original query, while automatic query relaxation (AQR) refers to the dropping or down-weighting of terms from the original query. While the former is well-suited for short queries, the latter is known to work well for longer queries (Xu & Croft, 1996; Kumaran & Allan, 2007).

Systems should be able to anticipate and adapt to situations as well as invoke user interaction in a judicious manner. Forcing the user to interact during every query session irrespective of whether there is utility in doing so can degrade overall user experience, and lead to increased cognitive load (Bruza et al, 1998).

In spatial personalized search, the geographic information systems suffer from lack of data quality since the presented data is highly complex and diverse. Manipulated objects have a very rich semantic and the degrees of user's interests towards them vary depending on context and on personal tastes (Miriam et al, 2009)

The proposed framework of personalized information retrieval system for a wearer (a person who is equipped with a wearable computer) is only indispensable for a wearer because desired or undesired information will be flooded around the wearer. However a large percentage of non-wearers will still be unable to obtain personalized information (Hong et al,2005).

Most existing context-sensitive IR systems base their retrieval decision solely on queries, keywords, topics, and document collections. Others use full-fledged ontology driven approach for an enhanced representation of the semantic context of information objects and user actions, in order to better interrelate user-sought meanings with available meanings in the search space. As a consequence, either the wider perspective of overall user trends, or the ability of the system to focus on temporary user priorities, is often lost (Vallet et al, 2004)

Most studies concentrate on taming the database management system through various well researched algorithms and techniques but still to some extent users are still, most of the time not retrieving relevant information. Least has been done to tame the users or utilize user session during search for relevant information to assist the database management systems. Users on session when typing relevant keywords/query to the database, the system does not know user preferences, hobbies and favorite sites, hence systems fail to give relevant information and jeopardize system performance in information retrieval. Therefore this paper discusses the use of touch dynamics results on search engines to enhance information retrieval relevance with pre-defined preferences, interest measurement and context to decrease cognitive load and upgrade user experience.

1.5 Justification

This study, attempts to design user input interactive method that accurately captures specific keywords and preferred links and sites. These Preferences are then embedded to touch dynamics results which users themselves define during searching sessions in search engine domain. When the user is able to use the portal successfully, hobbies, favorite links, and sites are captured either explicitly during the signing up process or explicitly during searching sessions. The portal automatically captures the user logs containing the list of preferred links and keywords and site embedded with the touch dynamics analysis of that particular user. With the increase in number of touch screen device users; and since different users have unique touch dynamic results then it is possible for search engine to identify different users according to those unique features hence retrieve relevant information based on the previous search preferences, hobbies and favorite sites during search sessions. This research is unique in its methods such that it is personalized search that can identify users on their touch dynamics analysis results across different touch screen devices to solve both query ambiguity and preferences

information retrieval problems suffered by different search engines whereas other personalized search studies do not implicitly embed user identification information in their techniques to solve the same.

1.6 Significance of the study

Users will be able to retrieve relevant information; their preferences will be captured in the logs by the system embedded with touch dynamic results information, the blend of searching preferences together with the touch dynamics results can always be used when a visit is made to the site.

This paper therefore investigates Touch Dynamics method to enhance relevance in information retrieval and therefore raises the following questions:-

1.7 Research Questions

i) Are the existing methods effective enough for resolving query difficulty and ensuring relevancy in information retrieval on search engine?

ii) Can touch dynamics results be more user-friendly than database systems-oriented techniques?

1.8Broad objective

To Study and analyze different personalized search methods that have been used for resolving query difficulty and enhancing relevancy in information retrieval and identify the importance of touch dynamics results in personalized information retrieval on search engines.

1.9 Specific objectives

i) To study the effectiveness and relevancy of existing information retrieval techniques for personalized search.

ii) To investigate user friendliness of touch dynamics result on search engines against Query relaxation and expansion methods.

2. User Identification based on Touch Dynamics and Search Behavior Profiling

Touch interaction has quickly become the de-facto means of interacting with handheld devices due to its perceived attractiveness and low hardware cost. The studies which propose strategies for identifying users based on touch dynamics monitor users' touch behavior and several unique features that are extracted including left versus right hand dominance, one handed versus bimanual operation, stroke size, stroke timing, symmetry, stroke speed and timing regularity. The strategies successfully identified users and their traits according to the touch dynamics. The results can be used for automatic user interface customization. However, more research is needed before touch characteristics can be applied to increasing the security of handheld touch-based devices. (Eika, etal, 2012)

2.1 Contextual Search Engine Model

The purpose of Contextual Search Engine technique is to avoid irrelevant preferences at run time, it is explained that it is not always that all user preferences may be important during the time of search. User may switch to a different subject which has never been experienced by the search engine and still expects relevant results. The diagram (Figure 2.1) below illustrates the above mentioned approach.

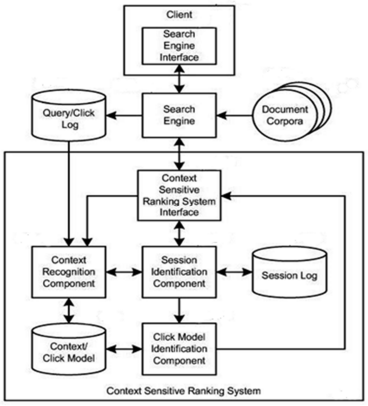

Figure 2.1, Contextual Search Engine Model, *adopted from Ashutosh Garg, 2009*

2.2Fully Connected Information Space

Fully Connected Information Space, can be created and maintained by integrating legacy database systems, and wired and wireless systems (PCS, Cellular system, and GSM), the figure2.3 below illustrate shows Integration of these components:-

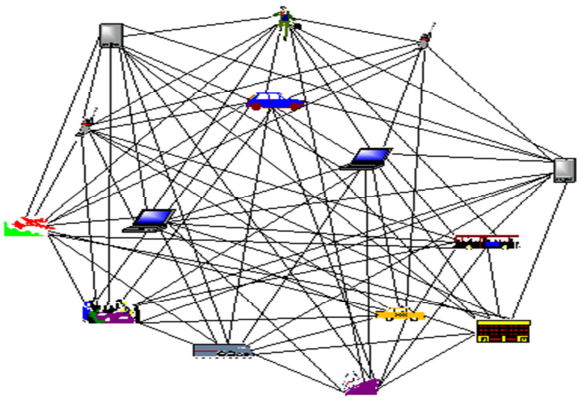

Figure 2.2 Interconnection of Web Searching Stations, *adopted from Kluwer Academic Publishers, 1998*

While users are searching for information, they have no idea of the original platform or source of the information they retrieve. Figure2.2above shows how a particular search engine retrieves the relevant information.

2.3 Related work

A number of existing studies have explored the topic of Personalized Information Retrieval aim to provide users with more personalized information provision. Personalization involves capturing the search interests of individuals and using these to train individual user interest models (Leung et al, 2008). There are two broad methods of capturing information for personalization:

1. Implicit feedback, where user interests are inferred from their behavior such as which documents they click on in the output of a search, their reading time for retrieved documents or their scrolling actions on a document (Mobasher & Anand, 2005).

2. Explicit feedback where users manually confirm document relevance or their topical interests (Mobasher & Anand, 2005). In therefore we look first at existing work in Personalized Information Retrieval and then review relevant studies.

2.3.1 Enhanced Information Retrieval Using Domain-Specific Recommender Models

The recommender systems attempt to recommend items that are likely to be of interest to users. Typically, a recommender system compares user profiles with some reference characteristics, and uses these to predict the rating that a user may give to a new item which he has not considered yet. These characteristics may be associated with the information contained in the item (the content-based approach) or the user's social environment (the collaborative filtering approach). In this paper we consider only the collaborative filtering approach to Recommender System collects user profile information in the same ways as IR systems. Since, as described earlier, the method is unable to collect explicit feedback in its environment, therefore considers collection of implicit feedback (Li et al, 2009).

This method for integrating domain-specific recommender models with information retrieval proceeds as follows:

1. The system records each query entered by previous users to search the available document archive, and implicit feedback from the users of the relevance rating of each retrieved document to viewed by each document, indicated by the time that the searcher spends on viewing the document.
2. The ratings of each viewed document for each topical query domain are then used to train a recommender model.
3. When a query is entered into the combined search method, a standard Information Retrieval technique is used to retrieve search results from the available document collection.
4. This query is also used to select an appropriate recommender model from the available domain-models generated from previous search data. The Recommender-Specific is then used to give predictions of potentially relevant documents based on the selected recommender model. (Li et al,2009)

2.3.2 Dynamic (Contextual) Personalized Information Retrieval

This personalized content retrieval aims at improving the retrieval process by taking into account the particular interests of individual users. However, not all user preferences are relevant in all situations. It is well known that human preferences are complex, multiple, heterogeneous, changing, even contradictory, and should be understood in context with the user goals and tasks at hand. This paper propose a method to build a dynamic representation of the semantic context of ongoing retrieval tasks, which is used to activate different subsets of user interests at runtime, in such a way that out of context preferences are discarded. The approach is based on an ontology-driven representation of the domain of discourse, providing enriched descriptions of the semantics involved in retrieval actions and preferences, and enabling the definition of effective means to relate preferences and context (Mylonas et al, 2008).

2.3.3 Spatial personalized information retrieval

This study presents a personalized information retrieval approach based on end user modeling. The proposed approach personalizes data retrieval using implicit user information and interests measurements. As the data manipulated is expressed by attributes and values, it defines several similarity measures. These measurements consider both semantic and spatial user contexts. The approach personalizes Web content and especially spatial information focusing on its spatial semantic aspects (Myriam et al, 2011).

2.3.4 Query Performance Prediction Method

Existing prediction approaches are roughly categorized to pre-retrieval methods and post-retrieval methods. Pre-retrieval approaches predict the quality of the search results before the search takes place,

thus only the raw query, and statistics of the query terms gathered at indexing time, can be exploited for prediction. In contrast, post-retrieval methods can additionally analyze the search results (Carmel & Yom-Tov, 2010).
Figure 2.3.4presents a general taxonomy of existing prediction approaches.

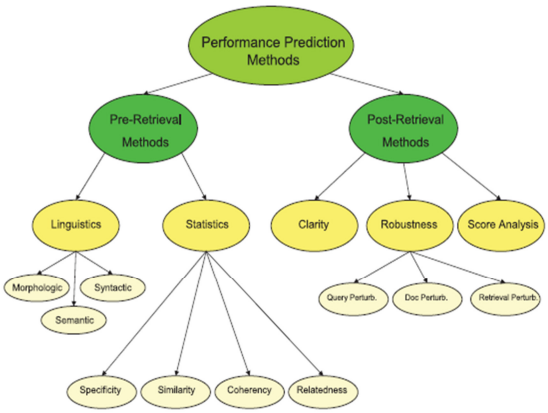

Figure 2.3.4, General taxonomy of query performance prediction methods. *Adopted from Estimating Query Difficulty for Information Retrieval, 2010.*

2.3.5 Personalized Concept-Based Clustering of Search Engine Queries

The underlying idea on figure 2.3.5is the work of (Leung et al., 2008), the technique is based on concepts and relations extracted from the submitted user queries, the web-snippets and the clickthrough data.

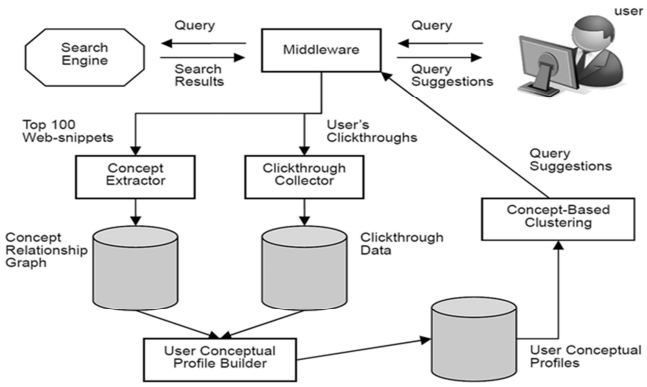

Figure 2.3.5, **Personalized Concept-Based Clustering of Search Engine Queries,** *Adopted from IEEE Transactions on Knowledge and Data Engineering, 2008.*

Clickthrough data was exploited in the personalized clustering process to identify user preferences: A user clicks on a search result mainly because the web snippet contains a relevant topic that the user is interested in. Moreover, clickthrough data can be collected easily without imposing extra burden on users, thus providing a low-cost means to capture user's interest. Our approach consists of the following four major steps. First, when a user submits a query, concepts (i.e., important terms or phrases in web-snippets) and their relations are mined online from web-snippets to build a concept relationship graph. Second, clickthroughs are collected to predict user's conceptual preferences. Third, the concept relationship graph together with the user's conceptual preferences is used as input to a concept-based clustering algorithm that finds conceptually close queries. Finally, the most similar queries are suggested to the user for search refinement. Figure 2.3.5shows the general process of their approach.

3. Critical Analysis of Existing Approaches/Techniques

Table 3.0, Critical Analysis of Existing Methods.

No.	Approach	Strengths	Limitations	Results
1	Query performance Prediction Method	System based method capable of searching all the relevant information available on the web (Carmel & Yom-Tov, 2010).	Does not consider user preferences, and previous interest	Lacks user search personalization
2	Dynamic Personalized Search	Searches according to user personal interest while omitting unnecessary preferences (Mylonas et al, 2008).	Users cannot explicitly define preferences	Retrieves information based on the context of the search session (out of context preferences are discarded)
3	Personalized Concept-Based Clustering of Search Engine Queries	Uses snippets to make users learn about new topics of the same context (Leung et al, 2008).	User can easily deviate from what initially intended to search	User learns new search keywords from web snippet for search refinement
4	Spatial Personalized Search	Captures user personalized search of information according to geographical locations. Hence discards information results from unselected regions (Myriam et al, 2011).	Lacks the universal layout of the information available on the world wide web.	Information retrieval results are based on the preferred region of user.
5	Personalized information retrieval based on context and ontological knowledge	The models and techniques address the automatic extraction of persistent, content-based user preferences, as well as live ad-hoc user interests, in such a way that the combination of both produce contextualized user models, which are then applied to gain accuracy in the personalization of retrieval results (Jiang and Tan, 2009).	Incapable of relating user search history to the live ad-hoc user interests. Mostly concentrates on the user predefined preferences	Lack a comprehensive personalized search results
6	Personalized Information Retrieval in Context	User interests that are out of focus for a given context are disregarded, and only those that are in the semantic scope of the ongoing user activity (a sort of intersection between user preferences and	There is no user search history or profiles considered at all	Contains only runtime user preferences

		runtime context) are considered for personalization (Vallet, 2004)		
7	Personalized information retrieval for ubiquitous computing environments	The proposed framework of personalized information retrieval system for a wearer (a person who is equipped with a wearable computer) is only indispensable for a wearer because desired or undesired information will be flooded around the wearer.(Hong et al, 2005)	A large percentage of non-wearers will still be unable to obtain personalized information	Depends on the information available around the wearable, which differs from one location to the other.
8	Interest-Based Personalized Search	Automatically maps a set of known user interests onto a group of categories in the Open Directory Project (ODP) and takes advantage of manually edited data available in ODP for training text classifiers that correspond to, and therefore categorize and personalize search results according to user interests (Ma et al, 2007).	Works on assumption that some information about the user interests is already known to the database and can be queried	Provides different results for different users or organize the same results differently for each user
9	Personalized Social Search Based on the User's Social Network	Uses a user profile that is derived from user feedback such as bookmarking, rating, commenting, and blogging which provide a very good indication of the user's interests (Carmel et al, 2009).	Lacks the universal layout of the information available on the world wide web.	Search results are solely based on user's social networks. Search results depend on the interests of other users in the network
10	Learning and inferencing in user ontology for personalized Semantic Web search	The support of a richer structure as well as more precise definitions of semantics. (Jiang and Tan , 2009)	Contextual or live, ad-hoc interests are not considered.	Search results are based on user previous search behavior through inferencing and learning

3.1 *Summary*

The different approaches mentioned in this paper as shown on table 3.0are very important for both personalized and non-personalized search, in fact they complement one another to fill some of the gaps users experience when retrieving information. However none of the existing personalized search technique has considered embedding user identification information to user search behavior profile.

3.2 Approach of the Study

In this study, the focus was on the combination of touch dynamic analysis result and personalization methods to improve the performance of personalized information retrieval. The key aspects in this proposed approach are:-

a) The explicit distinction between historic user preferences and query keyword technique.

b) The use of ontology-driven representations of the domain of discourse, as a common, enriched representational ground for content meaning, user interests, and contextual conditions, enabling the definition of effective means to relate the three of them, and
c) Introduction of the implicit combination of the historic user preferences and touch screen behavior identification to properly handle the uncertainty and imprecision involved in the automatic interpretation of meanings, user preferences, and user interest to enhance reliability and relevance in information retrieval.

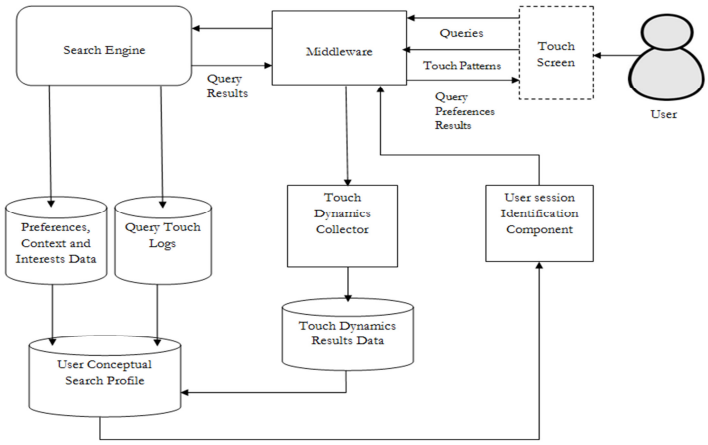

Figure 3.2, Combination of Touch Dynamics and Search Behavior, *Personalized Search Approach.*

4. Conclusion

Personalized search has been adopted by several search engines using the above mentioned methods to ensure that users get search results that satisfy their interests. The approach is well suited for a workplace setting or college environment where information about professional or student interests can be obtained automatically across touch devices of similar platforms no matter where the search result are needed.
This approach has several advantages, in that it (a) collect a user's browsing or search history, (b) ask a user to provide explicit or implicit feedback about the search behaviors together with touch dynamics,(c) does not depend on the search engine to give search results but gives suggestive queries to user for better results (d) using touch dynamics, user search behaviors and interests are automatically updated to any touch device.
Furthermore, the study works on assumption that users will always be using touch screen devices to relate search behaviors with their respective touch logs. Therefore speedy and accurate identification through touch dynamics of user data such as the search keywords, search results links, touched search results links, book marked pages and sites, previous keywords that are used to automatically suggest keyword choices for the user's search interests and all the visited sites, If this data is preserved across platforms and can be access from any device with the assistant of information synchronization narrows down further the gaps which exist in search results expected by the user. The system for enhancing information retrieval relevance using touch dynamics on search engine is a recommended application for this paper

REFERENCES

Carmel, D., & Yom-Tov, E. (2010).*Estimating the Query Difficulty for Information Retrieval*.Morgan & Claypool Publishers.
Grossman, D. A., & Frieder, O. (2004). *Information Retrieval: Algorithms and Heuristics*. Springer Science &Business Media.
Leung, K. W., Ng, W., & Lee, D. L. (2008).Personalized concept-based clustering of search engine queries.*IEEE Transactions on Knowledge and Data Engineering*, 1505–1518.
Manning, C. D., Raghavan, P., & Schütze, H. (2008).*Introduction to Information Retrieval*.Cambridge University Press.

Mylonas, P., Vallet, D., Cas [TEL] ls, P., Fernández, M., &Avrithis, Y. (2008). Personalized information retrieval based on context and ontological knowledge. *The Knowledge Engineering Review*, *23*(Special Issue 01), 73–100. http://doi.org/10.1017/S0269888907001282

Myriam, H., Hajer, B., Aude, A. M., & Henda, B. G. (2011).User Modeling-based Spatial Web Personalization.In *Proceedings of the 15th International Conference on Knowledge-based and Intelligent Information and Engineering Systems - Volume Part II* (pp. 41–50). Berlin, Heidelberg: Springer-Verlag. Retrieved from http://dl.acm.org/citation.cfm?id=2041341.2041347

F. Liu, C. Yu, and W. Meng, "Personalized Web Search for Improving Retrieval Effectiveness," *IEEE Trans. Knowledge and Data Eng.*, vol. 16, no. 1, pp. 28-40, Jan. 2004.

Q. Tan, X. Chai, W. Ng, and D.L. Lee, "Applying Co-Training to Clickthrough Data for Search Engine Adaptation," *Proc. Ninth Int'l Conf. Database Systems for Advanced Applications* (DASFAA), 2004.

Elad Yom-Tov, Shai Fine, David Carmel, and Adam Darlow. Learning to estimate query difficulty: including applications to missing content detection and distributed information retrieval. *In Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval* (SIGIR 2005), pages 512–519, Salvador, Brazil, 2005.

Web Search Engine," User Modeling and User-Adapted Interaction, vol. 14, no. 5, pp. 383-423, 2005.

M. Speretta and S. Gauch, "Personalized Search Based on User Search Histories," *Proc. IEEE/WIC/ACM Int'lbConf. Web Intelligence* (WI), 2005.

Vallet D., I. Cantador, and J. M. Jose.Personalizing Web Search with Folksonomy-based User and Document Profiles. *In Proceedings of ECIR* 2010, pages 420-431 (2010)

Song Y., N. Nauyen, L.He, S. Imig, and R. Rounthwaite.Searchable Web Sites Recommendations.*In Proceedings of the Fourth ACM international conference on Web Search and Data Mining* (WSDM 2011). (2011)

Liu J., and N. J. Belkin, Personalizing Information Retrieval for Multi-Session Tasks.*The Roles of Task Stage and Task Type.In Proceedings of ACM SIGIR* 2010. (2010)

Belkin. N. J. Some (what) Grand Challenges for Information Retrieval. *ACM SIGIR Forum* 42(1):47-54. (2008)

E. Pitoura and G. Samaras, "Data Management for Mobile Computing", *Kluwer Academic Publishers*, 1998.

GiridharKumaran and Vitor R. Carvalho. Reducing long queries using query quality predictors.*In Proceedings of the 32nd international ACM SIGIR conference on Research and developmentin information retrieval* (SIGIR 2009), pages 564–571, Boston, MA, USA, 2009. ACMDonna Harman and Chris Buckley.Overview of the reliable information access workshop. Information Retrieval, 12(6):615–641, 2009

Claudia Hauff. Predicting the Effectiveness of Queries and Retrieval Systems.PhD thesis, University of Twente, February 2010.xi, 19, 20, 22, 23, 24, 27, 29.

W. Magdy, P. Lopez, and G. J. F. Jones. Simple vs. sophisticated approaches for patent prior-art search. In ECIR, pages 725–728, 2011.

P. Lopez and L. Romary. Experiments with citation mining and key-term extraction for prior art search. In CLEF-2010

Ying Zhao, Falk Scholer, and YohannesTsegay.Effective pre-retrieval query performance rediction using similarity and variability evidence.*In Proceedings of the European Conference on Information Retrieval* (ECIR 2008), pages 52–64, 2008.

2012 9th International Conference on Ubiquitous Intelligence and Computing and 9th International Conference onAutonomic and Trusted Computing.

The IISTE is a pioneer in the Open-Access hosting service and academic event management. The aim of the firm is Accelerating Global Knowledge Sharing.

More information about the firm can be found on the homepage:
http://www.iiste.org

CALL FOR JOURNAL PAPERS

There are more than 30 peer-reviewed academic journals hosted under the hosting platform.

Prospective authors of journals can find the submission instruction on the following page: http://www.iiste.org/journals/ All the journals articles are available online to the readers all over the world without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. Paper version of the journals is also available upon request of readers and authors.

MORE RESOURCES

Book publication information: http://www.iiste.org/book/

Academic conference: http://www.iiste.org/conference/upcoming-conferences-call-for-paper/

IISTE Knowledge Sharing Partners

EBSCO, Index Copernicus, Ulrich's Periodicals Directory, JournalTOCS, PKP Open Archives Harvester, Bielefeld Academic Search Engine, Elektronische Zeitschriftenbibliothek EZB, Open J-Gate, OCLC WorldCat, Universe Digtial Library , NewJour, Google Scholar