

Learning through Multimedia Answer Generation by Web Based Query System

Nayana V¹, Ajitha S²

¹Asst.Professor Dept. Of Computer Science and Engineering, MountZion Institute of Science and Technology, Chengannur, Kerala

²Asst.Professor Dept. Of Computer Science and Engineering, MountZion Institute of Science and Technology, Chengannur, Kerala

Abstract - In assessment to current strategies that provide the identical guidelines to all customers, our technique makes use of click through statistics to estimate consumer's conceptual possibilities after which affords personalised question guidelines for every individual consumer in keeping with his/her conceptual needs. The motivation of our studies is that queries submitted to a seek engine might also additionally have more than one meanings. For example, relying at the consumer, the question "apple" might also additionally discuss with a fruit, the organisation Apple Computer or the call of a person, and so forth. Thus, presenting personalised question suggestion (e.g., customers involved in "apple" as a fruit get guidelines approximately fruit, at the same time as customers involved in "apple" as a organisation get guidelines approximately the organisation's products) virtually enables customers formulate greater powerful queries according to their needs. Underlying concept of proposed method is primarily based totally on principles and their relations extracted from the submitted consumer queries, the web-snippets and the click on-via statistics. Click via statistics became exploited withinside the personalised clustering method to discover consumer possibilities: a consumer clicks on a seek end result particularly due to the fact the web-snippet carries a applicable subject matter which the consumer is involved in. Moreover, click on via statistics may be amassed effortlessly with out implementing extra burden on customers, and for this reason presenting a low-value manner to seize consumer's interest. In this concept- primarily based totally consumer profiling techniques which might be able to deriving each of the consumer's superb and terrible possibilities. All of the customers profiling techniques are question-oriented, which means that a profile is created for every of the consumer's queries. The consumer profiling techniques are evaluated and as compared with our formerly proposed personalised question clustering method.

Key Words: Multimedia, Profiling, Search Engine, Clustering

1. INTRODUCTION

Multimedia refers to content material that makes use of a mixture of different content material paperwork. By combining media and content material, the ones inquisitive about multimedia can tackle and paintings with numerous

media paperwork to get their content material across. This is an thrilling new discipline for the ones inquisitive about computers, technology, and innovative profession options. Multimedia consists of a mixture of textual content, audio, nonetheless images, animation, video, or interactivity content material paperwork. One instance of multimedia could be combining a internet site with video, audio, or textual content images. Multimedia can be widely divided into linear and non-linear categories. Linear lively content material progresses regularly with none navigational manage for the viewer together with a cinema presentation. Non-linear makes use of interactivity to govern development as with a video game or self-paced laptop primarily based totally training. Hypermedia is an instance of non-linear content material.

1.1 Classification of Multimedia search:

The text-primarily based totally seek methods use textual queries, a term-primarily based totally specification of the preferred media entities, to look for media information via way of means of matching them with the surrounding textual descriptions. To enhance the overall performance of text-primarily based totally seek, some gadget getting to know strategies that goal to robotically annotate media entities were proposed withinside the multimedia community. However, user-supplied text descriptions for media information are frequently biased in the direction of private views and contexts, and for that reason there's an opening among those tags. The content material of the media entities that not unusualplace customers are involved in. To address this issue, content material-primarily based totally media retrieval plays seek via way of means of studying the contents of media information in preference to the metadata.

1.2 CHALLENGES IN MULTIMEDIA SEARCH

Along with the proliferation and development of underlying verbal exchange technologies, network QA has emerged as an exceptionally famous opportunity to accumulate records online, proudly owning to the subsequent facts. First, records seekers are capable of submit their unique questions about any subject matter and achieve solutions supplied through different participants. By leveraging network efforts, they may be capable of get higher solutions than

virtually the usage of serps. Second, in assessment with computerized QA systems, cQA generally gets solutions with higher first-rate as they may be generated primarily based totally on human intelligence. Third, over times, a first-rate variety of QA pairs had been accrued of their repositories, and it helps the protection and seek of spoke back questions. For example, Wiki Answer, one of the maximum famous cQA systems, hosts greater than thirteen million spoke back questions allotted in 7,000 categories.

- Fully computerized QA nevertheless faces demanding situations that aren't smooth to tackle, consisting of the deep expertise of complicated questions and the state-of-the-art syntactic, semantic and contextual processing to generate solutions.

- A important hassle of modern Web seek is that seek queries are generally brief and ambiguous, and for that reason are inadequate for specifying the appropriate consumer needs.

- To alleviate this hassle, a few serps endorse phrases which might be semantically associated with the submitted queries in order that customers can pick from the pointers those that replicate their records needs.

2. LITERATURE REVIEW

Question-answering (QA) is a herbal route to deal with this records over-loading problem. The purpose of QA is to go back unique solutions to users' questions. Text-primarily based totally QA studies has been performed for the beyond 15 years with exact achievement specially for answering fact-primarily based totally questions. The purpose of this paper is to increase the text-primarily based totally QA studies to multimedia QA to address quite a number factoid, definition and "how-to" QA in a not unusualplace framework. The machine may be designed to locate multimedia solutions from Web-scale media assets which include Flickr and YouTube. This paper describes the structure and our current studies on numerous forms of multimedia QA for quite a number applications. Here additionally discusses guidelines for destiny studies.

Interactive query answering (QA), in which a speak interface allows follow-up and explanation questions, is a current despite the fact that long-endorsed subject of research. We record at the layout and implementation of YourQA, our open-domain, interactive QA system. YourQA is predicated on a Web seek engine to achieve solutions to each fact-primarily based totally and complicated questions, including descriptions and definitions. We describe the speak actions and control version making YourQA interactive, and talk the structure, implementation and assessment of its chat-primarily based totally speak interface. Our Wizard-of-Oz look at and very last assessment effects display how the designed structure can efficiently acquire open-domain, interactive QA.

With the speedy enlargement of video media reassets which include news, TV shows, and movies, there's an growing call for automated retrieval and surfing of video facts. Generally speaking, the only manner of handling video facts is to assist a video record retrieval primarily based totally at the enter key-word queries. In a cutting-edge video retrieval scenario, customers might question the device with brief or herbal language questions together along with his acquainted language, e.g., "Where is the start of the Chinese culture?" or "In China, what's the maximum extremely good pottery city?" They anticipate the device to go back brief solutions instead of the complete films which can be in exceptional language. To assist above goals, this means numerous studies fields which include, video content material extraction, data processing, and query answering. Extracting contents in films is a totally hard however complicated assignment and turning into an essential issue. Textual, visual, audio data is maximum regularly followed capabilities for this purpose.

we are looking for to apprehend YA's expertise sharing activity. Analyze the discussion board classes and cluster them in keeping with content material traits and styles of interplay some of the customers. While interactions in a few classes resemble information sharing boards, others include discussion, ordinary advice, and assist. With this type of range of classes wherein you possibly can take part, that a few customers cognizance narrowly on spetic topics, at the same time as others take part throughout classes. This now no longer most effective lets in us to map associated classes, however represent the entropy of the customers' interests. Decrease entropy correlates with receiving better solution ratings, however most effective for classes wherein authentic information is in general sought after. Here integrate each person attributes and solution traits to predict, inside a given category, whether or not a specific solution can be selected because the nice solution through the asker.

Existing network query-answering boards typically offer most effective textual solutions. However, for lots questions, natural texts can not offer intuitive data, at the same time as picture or video contents are greater appropriate. Here introduce a scheme this is capable of enhance textual content solutions with picture and video data. In this scheme investigates a wealthy set of strategies which includes query/solution classification, question generation, picture and video seek reranking, etc. Given a query and the network-contributed solution, our technique is capable of decide which kind of media data need to be added, after which mechanically collects facts from Internet to enhance the textual solution. Different from a few efforts that try to at once solution questions with picture and video facts, our technique is constructed primarily based totally at the network-contributed textual solutions and accordingly it's far greater possible and capable of address greater complicated questions. Here performed empirical observe on greater than 3,000 QA pairs and the effects display the effectiveness of our technique.

In this personalised idea-primarily based totally clustering endorse a brand new personalised idea-primarily based totally clustering approach that's capable of achieve personalised question recommendations for man or woman customers primarily based totally on their conceptual profiles. The approach uses clickthrough statistics and the idea courting graph mined from internet-snippets, each of which may be captured on the backend and as such do now no longer upload greater burden to customers. An tailored agglomerative clustering set of rules is hired for locating queries which can be conceptually near one another. Our experimental effects affirm that our method can effectively generate personalized question recommendations in step with man or woman consumer conceptual wishes. Moreover, it improves prediction accuracy and computational value evaluating to BB's set of rules, that's the today's approach of question clustering the usage of click on throughs for the same objective.

Current Web serps are constructed to serve all customers, impartial of the unique wishes of any man or woman consumer. Personalization of Web seek is to perform retrieval for every consumer incorporating his/her interests. We endorse a singular approach to analyze consumer profiles from customers' seek histories. The consumer profiles are then used to enhance retrieval effectiveness in Web seek. A consumer profile and a preferred profile are found out from the consumer's seek records and a class hierarchy, respectively. These profiles are blended to map a consumer question into a hard and fast of classes which constitute the consumer's seek purpose and function a context to disambiguate the phrases withinside the consumer's question. Web seek is performed primarily based totally on each the consumer question and the set of classes. Several profile gaining knowledge of and class mapping algorithms and a fusion set of rules are supplied and evaluated. Experimental effects imply that our approach to customize Web seek is each powerful and efficient.

This optimizing serps the usage of click on thru statistics affords a method to robotically optimizing the retrieval fine of serps the usage of click on thru statistics. Intuitively, an excellent statistics retrieval gadget have to gift applicable files excessive withinside the rating, with much less applicable files following below. While preceding methods to gaining knowledge of retrieval capabilities from examples exist, they commonly require schooling statistics generated from relevance judgments with the aid of using experts. This makes them hard and high priced to follow. The aim of this paper is to increase a way that makes use of click through statistics for schooling, particularly the question-log of the quest engine in reference to the log of hyperlinks the customers clicked on with inside the supplied rating. Such click through statistics is to be had in abundance and may be recorded at very low value.

Taking a Support Vector Machine (SVM) method, affords a way for gaining knowledge of retrieval capabilities.

In this internet seek rating explored the software of incorporating noisy implicit comments received in a actual internet seek placing to enhance internet seek rating. Finished a large-scale assessment over 3,000 queries and greater than 12 million consumer interactions with a major seek engine, organising the software of incorporating "noisy" implicit comments to enhance internet seek relevance. Here in comparison options of incorporating implicit comments into the quest process, particularly reranking with implicit comments and incorporating implicit comments functions at once into the educated rating function. Our experiments confirmed significant development over strategies that don't do not forget implicit comments. The profits are especially dramatic for the pinnacle $K=1$ bring about the very last rating, with precision enhancements as excessive as 31%, and the profits are good sized for all values of K . experiments confirmed that implicit consumer comments can similarly enhance seek performance, included at once with famous content material-ink-primarily based totally functions.

User profiles, descriptions of consumer interests, may be utilized by serps to offer personalised seek effects. Many methods to developing consumer profiles seize consumer statistics thru proxy servers (to seize surfing histories) or computing device bots (to seize all sports on a non-public computer). These each require participation of the consumer to put in the proxy server or the bot. In this observe, we discover using a much less-invasive manner of accumulating consumer statistics for personalised seek. In particular, we construct consumer profiles primarily based totally on pastime on the seek web website online itself and observe using those profiles to offer personalised seek effects. In our observe, we applied a wrapper for Google to look at one of a kind reasssets of statistics on which to base the consumer profiles: queries and snippets of tested seek effects. These consumer profiles had been created with the aid of using classifying the statistics into ideas from the Open Directory Project idea hierarchy after which used to re-rank the quest effects.

A seek engine question log introduces a method for mining a set of consumer transactions with an Internet seek engine to find out clusters of comparable queries and comparable URLs. The statistics we take advantage of is "clickthrough statistics": every report includes a consumer's question to a seek engine in conjunction with the URL which the consumer decided on from some of the applicants supplied with the aid of using the quest engine. By viewing this dataset as a bipartite graph, with the vertices on one facet similar to queries and on the alternative facet to URLs, you may follow an agglomerative clustering set of rules to the graph's vertices to discover associated queries and URLs. One noteworthy function of the proposed set of rules is that it is "content material-ignorant" "the set of rules makes little need of the real content material of the queries or URLs,

however handiest how they co-arise in the click on thru statistics.

3. METHODOLOGY

The evaluate of our seek personalization architecture. Most personalization strategies targeted at the advent of 1 unmarried profile for a consumer and implemented the equal profile to all the consumer's queries. We accept as true with that one-of-a-kind queries from a consumer need to be treated in a different way due to the fact a consumer's choices might also additionally range throughout queries. Web has delivered new demanding situations for constructing powerful engines like google. A main trouble of internet seek is that seek queries are generally quick and ambiguous, and as a consequence are inadequate for specifying the right consumer desires. When one of these question is issued, engines like google will go back a listing of files that blend one-of-a-kind subjects. It takes time for a consumer to pick out which statistics he/she wants. On every other question of "loose mp3 down load," even though maximum customers locate web sites to down load loose mp3s, their picks can diverge: one might also additionally pick out the internet site www.yourmp3.net, at the same time as every other might also additionally pick the internet site www.seekasong.com. Despite the life of bookmarking gear that might permit the consumer to attain this aim in an immediate manner, customers pretty regularly choose to re-look for statistics through re-filing a formerly issued question. Finding out approximately subjects of consumer hobby by thinking about the long time seek and skim records of a consumer, the principle subjects of consumer hobby emerge. Whenever a consumer question is ambiguous or large in nature, superposing the learnt consumer hobbies would possibly serve the consumer seek experience. However, as already located with inside the gain of such an technique would possibly range for re-curring compared to sparkling queries which motivates a differentiated utilization of long-time period consumer statistics. Serving an ad-hoc statistics want despite the fact that a consumer would possibly have sturdy recognition on numerous subjects of hobby, nevertheless would possibly transfer hobbies or increase a few quick-time period statistics desires out of doors the scope of her hobbies. As specifically the surfing sports past seek are out of doors the attain of a seek engine, client-aspect answers are favorable. Moreover, as all consumer facts is stored locally, consumer privateness isn't always violated. Consequently installation a client-aspect seek personalization with the usage of a proxy that is run in locally. It intercepts all HTTP traffic, extracts queries, question chains, i.e., finally posed queries, end result sets, clicked end result pages, in addition to the complete click on movement of finally visited internet pages, and shops this statistics to a neighborhood database document which we discuss with because the neighborhood index with inside the following. Accordingly, searches with Google (the equal technique may be without problems implemented to another

seek engine as well) are intercepted and seek consequences are re-ranked in keeping with private choices. offers a top level view of our seek personalization architecture.

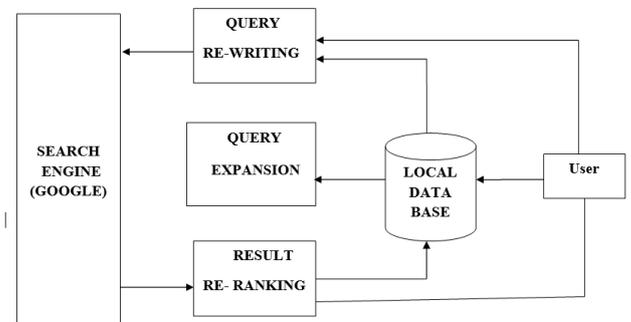


Fig 1: Methodology

A precise person profiling approach is an crucial and essential thing in seek engine personalization. We studied numerous person profiling techniques for seek engine personalization, and determined the subsequent troubles in present techniques. Most personalization strategies targeted at the introduction of 1 unmarried profile for a person and implemented the equal profile to all the person's queries. We accept as true with that distinct queries from a person need to be dealt with in a different way due to the fact a person's alternatives may also range throughout queries. There are approaches a person talks to an internet database thru a type-in container and with the aid of using choosing picks from a listing structure. This segment will take a look at the greater not unusualplace of the, the hunt engine, and could assist you to research the variations among seek engines (due to the fact, simply because the databases are distinct, so are the hunt engines) and a way to grow to be an powerful searcher. The 4 predominant steps in the use of a seek engine are: translation out of your seek request into seek phrases which you type, strategies of mixing the hunt phrases to provide the laptop greater data approximately your seek request, inspecting the effects presented, and re-looking if needed. This cycle may be repeated as essential till you've got got determined that that you are searching out or you make a decision which you have exhausted the choice from this seek engine. To deal with the above troubles with the aid of using featuring and analyzing idea-primarily based totally person profiling techniques which can be able to deriving each of the person's advantageous and terrible alternatives. The complete person profiling techniques is question-oriented, which means that a profile is created for every of the person's queries. The person profiling techniques are evaluated and as compared with our formerly proposed customized question clustering technique. Moreover, locate that terrible alternatives enhance the separation of comparable and distinctive queries, which enables an agglomerative clustering set of rules to determine if the foremost clusters had been obtained. To amplify the

question-oriented, idea-primarily based totally person profiling technique proposed in to recall each customers' advantageous and terrible alternatives in constructing customers profiles. In this proposed strategies use an RSVM to research from idea alternatives weighted idea vectors representing idea-primarily based totally person profiles. The weights of the vector elements, which may be advantageous or terrible, constitute the interestingness of the person at the concepts. In, the weights that constitute a person's pastimes are all advantageous, which means that the technique can most effective seize person's advantageous alternatives.

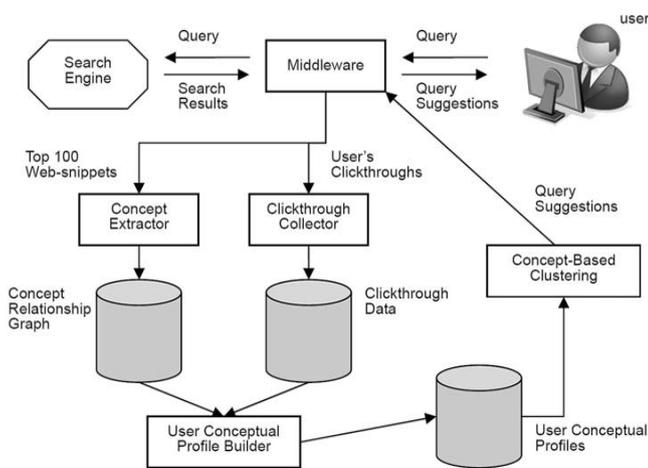


Fig 2: The general process of concept-based profiling

Although personalised seek has been beneathneath manner for decades and plenty of personalization algorithms were investigated, it's miles nevertheless doubtful whether or not personalization is continuously powerful on one-of-a-kind queries for one-of-a-kind customers and beneathneath one-of-a-kind seek contexts. We gift a large-scale assessment framework for personalised seek primarily based totally on question logs after which examine 5 personalised seek algorithms (consisting of click on-primarily based totally ones and 3 topical-interest-primarily based totally ones. It represents a widespread development over frequent Web look for a few queries, whilst it has little impact or even harms question overall performance beneathneath a few situations. Underlying concept of our proposed approach is primarily based totally on principles and their family members extracted from the submitted person queries, the web-snippets1 and the click on-via records. Click via records turned into exploited withinside the personalised clustering system to pick out person choices: a person clicks on a seek end result in particular due to the fact the web-snippet carries a applicable subject matter which the person is involved in. Moreover, click on via records may be gathered effortlessly with out enforcing greater burden on customers, and as a consequence offering a low-fee way to seize person's interest. presenting and analyzing seven idea-primarily based totally person profiling techniques which

are able to deriving each of the person's advantageous and terrible choices. All of the customers profiling techniques are question-oriented, that means that a profile is created for every of the person's queries. The person profiling techniques are evaluated and in comparison with our formerly proposed personalised question clustering approach. Enlarge the question-oriented, idea-primarily based totally person profiling approach proposed in to don't forget each customers' advantageous and terrible choices in constructing customers profiles. We proposed six person profiling strategies that take advantage of a person's advantageous and terrible choices to provide a profile for the person the usage of a Ranking SVM (RSVM). Our proposed strategies use an RSVM to examine from idea choices weighted idea vectors representing idea-primarily based totally person profiles. The weights of the vector elements, which might be advantageous or terrible, constitute the interestingness (or UN interestingness) of the person at the principles. In the weights that constitute a person's pursuits is all advantageous, that means that the approach can best seize person's advantageous choices. To remedy the problem, we divide clustering into steps. In the preliminary clustering step, an set of rules much like BB's set of rules is hired to cluster. All the queries, however it might now no longer merge equal queries from one-of-a-kind customers. After acquiring all of the clusters from the preliminary clustering step the network merging step is hired to merge question clusters containing equal queries from one-of-a-kind customers can see from will similarly in that the preliminary clustering step is capable of generate excessive precision charge as it preserves the choice of every person, whilst the network merging step is capable of enhance the don't forget charge due to the collaborative filtering impact.

To acquire applicable photo and video information from the web, we want to generate suitable queries from textual content QA pairs earlier than acting seek on multimedia serps. We accomplish the assignment with steps. The first step is question extraction. Textual questions and solutions are commonly complicated sentences. But often serps do now no longer paintings nicely for queries which are lengthy and verbose. Therefore, we want to extract a fixed of informative key phrases from questions and solutions for querying. The 2nd step is question selection. This is due to the fact we will generate distinctive queries: one from query, one from answer, and one from the mixture of query and answer. Which one is the maximum informative relies upon at the QA pairs. Since many questions incorporate a couple of sentences (really our records on Y!A display that at the least of the questions incorporate at the least sentences, and the variety is round For WikiAnswers) and a number of the sentences are uninformative, first rent the approach in to extract the middle sentence from every query. The category is carried out with steps. First, categorize questions primarily based totally on interrogatives (a few beginning phrases and finishing phrases), and on this manner can at once locate questions that must be spoke back with textual

content. Second, for the relaxation questions, carry out a category the use of a naive Bayes classifier.

Personalized seek is an critical studies vicinity that goals to clear up the anomaly of question terms. To growth the relevance of seek results, customized search engines like google like google and yahoo creat person profiles to seize the customers' non-public options and as such become aware of the real aim of the enter question. A precise person profiling method is an vital and essential issue in seek engine personalization. it then addresses each the issues of the person's nice and bad options. All of the customers profiling techniques are question-oriented, which means that a profile is created for every of the person's queries. It indicates that person profiles which seize each the person's nice and bad options carry out the pleasant amongst all the profiling techniques. A new method has been brought withinside the proposed machine is Personalized Ontology, which officially describes and specifies the person profile knowledge. Therefore, acquire 3 queries, and the following step is to pick out one from them. Introduce an powerful method that capturesthe person's conceptual options in an effort to offer customized question suggestions. We reap this aim with new techniques. First, we increase on line strategies that extract standards from the web-snippets of the hunt end result back from a question and use the standards to become aware of associated queries for that question. Second, we advise a brand new segment customized agglomerative clustering set of rules this is capable of generate customized question clusters. The approach proposed is primarily based totally on a question clustering manner wherein companies of semantically comparable queries are identified. The clustering manner makes use of the content material of historic options of customers registered with inside the question log of the hunt engine. The approach now no longer simplest discovers the associated queries, however additionally ranks them in keeping with a relevance criterion. Finally, we display with experiments over the question log of a seek engine the effectiveness of the approach .Query clustering is a manner used to find out often requested questions or maximum famous subjects on a seek engine. This manner is vital for search engines like google like google and yahoo primarily based totally on question-answering. Because of the quick lengths of queries, techniques primarily based totally on key phrases aren't appropriate for question clustering. This paper describes a brand new question clustering approach that uses person logs which permit us to become aware of the files the customers have decided on for a question. The similarity among queries can be deduced from the not unusualplace files the customers decided on for them. Our experiments display that a aggregate of each key phrases And person logs is higher than the use of both approach alone.

The intention of person interface layout is to make the person's interplay as easy and green as possible, in phrases of undertaking person goals—what's frequently referred to

as person-focused layout. Good person interface layout allows completing the undertaking handy with out drawing needless interest to it. Graphic layout can be applied to guide its usability. The layout technique ought to stability technical capability and visible elements (e.g., intellectual model) to create a device that isn't always simplest operational however additionally usable and adaptable to converting person needs. Interface layout is concerned in a huge variety of initiatives from pc systems, to cars, to business planes; all of those initiatives contain a whole lot of the identical primary human interactions but additionally require a few precise abilties and knowledge.

4. CONCLUSION

The strategies employ click on via statistics to extract from Web-snippets to construct concept-primarily based totally person profiles automatically. We carried out choice mining regulations to deduce now no longer handiest users' high-quality possibilities however additionally their terrible possibilities, and applied each sorts of possibilities in deriving person's profiles. The person profiling techniques have been evaluated and as compared with the customised question clustering approach that we proposed previously. Apart from enhancing the high-satisfactory of the ensuing clusters, the terrible possibilities with inside the proposed person profiles additionally assist to separate comparable and diverse queries into remote clusters, which assist to determine close to most effective terminating factors for our clustering algorithm.

REFERENCES

- [1] E. Agichtein, E. Brill, and S. Dumais, "Improving Web Search Ranking by Incorporating User Behavior Information," Proc. ACM SIGIR, 2006.
- [2] E. Agichtein, E. Brill, S. Dumais, and R. Ragno, "Learning User Interaction Models for Predicting Web Search Result Preferences," Proc. ACM SIGIR, 2006.
- [3] K.W. Church, W. Gale, P. Hanks, and D. Hindle, "Using Statistics in Lexical Analysis," Lexical Acquisition: Exploiting On-Line Resources to Build a Lexicon, Lawrence Erlbaum, 1991.
- [4] S. Gauch, J. Chaffee, and A. Pretschner, "Ontology-Based Personalized Search and Browsing," ACM Web Intelligence and Agent System, vol. 1, nos. 3/4, pp. 219-234, 2003.
- [5]] K.W.-T. Leung, W. Ng, and D.L. Lee, "Personalized Concept- Based Clustering of Search Engine Queries," IEEE Trans. Knowledge and Data Eng., vol. 20, no. 11, pp. 1505-1518, Nov. 2008..
- [6] F. Liu, C. Yu, and W. Meng, "Personalized Web Search by Mapping User Queries to Categories,"

Proc. Int'l Conf. Information and Knowledge Management (CIKM), 2002.

[7] W. Ng, L. Deng, and D.L. Lee, "Mining User Preference Using Spy Voting for Search Engine Personalization," ACM Trans. Internet Technology, vol. 7, no. 4, article 19, 2007.

[8] M. Speretta and S. Gauch, "Personalized Search Based on User Search Histories," Proc IEEE/WIC/ACM Int'l Conf. Web Intelligence, 2005.

[9] J.-R. Wen, J.-Y. Nie, and H.-J. Zhang, "Query Clustering Using User Logs," ACM Trans. Information Systems, vol. 20, no. 1, pp. 59- 81, 2002.

[10] Y. Xu, K. Wang, B. Zhang, and Z. Chen, "Privacy-Enhancing Personalized Web Search," Proc. World Wide Web (WWW) Conf., 2007.