VOICE ENABLE PERSONALIZED WEB SEARCH

FAUZIA PARVEEN M.F.RAHMAN
(Department of CSE, GHRAET/RTM Nagpur University, INDIA fauziaparveen16@gmail.com)

VIKRANT CHOLE
(Department of CSE,GHRAET/RTM Nagpur University, INDIA, vikrant.chole@raisoni.net)

ABSTRACT:

The technology of voice browsing is rapidly evolving these days. It is because the use of cell phones is increasing at a very high rate, as compared to connected PCs. Speech interface integrated browser is a web browser that helps users by using an interactive voice user interface ,useful to those who have difficulties in seeing and reading a web content. Listening and speaking are the natural modes of communication and information gathering. As a result we are now heading towards a more voice based approach of browsing rather than operating on textual mode. A voice browser or speech browser will take and presents the information in the form of voice as well as text, using text to speech and speech to text conversion to render information. People want to get accurate and appropriate data at the top of search results in a user friendly manner. People also want to get a personal space over the internet when they are browsing on web, from this arises a need of personalization of the search history. Thus there is a need of a highly efficient and effective ranking algorithm that provides search results according to user preferences. This paper concentrates on this new technique, voice browsing, which unites speech recognition and speech synthesis with better personalized search that can be very fruitful in the coming years. In this paper we provide personalization by creating individual search history for each user on the browser and also focused on the search results to get customized according to the user demand.

KEYWORDS: personalized web search, Voice Recognition, Speech Synthesizer, Page Rank, browsing history.

I. INTRODUCTION

It is evident that the information on World Wide Web is still inaccessible to many people. The web is primarily a visual medium that requires a keyboard and mouse to navigate. People, who lack motor skills to use a keyboard and mouse, find navigation troublesome. Visually impaired people have problems in accessing the web, and also those who temporarily cannot use a traditional web browser, as their eyes or hands are occupied or because they are not closer to their computer are at a minimum inconvenienced.

A speech browser presents an interactive voice user interface to the user, useful to those who face difficulties in primary senses like reading and seeing. A voice browser presents information aurally, either using prerecorded audio file playback or using text to speech conversion to render textual information as audio. Voice browser designed for users who wish to access the internet in a non-visual or combined auditory and visual way. Immediate beneficiary users will include blind or partially sighted users, people with dyslexia or learning difficulties, and users who are learning new languages. When a user submits a query to the search engine to find specific information, the search engine must be able to retrieve documents according to user specific need. However search engine shows the list of ranked documents in accordance with the words present in user's query. In several occasions the inappropriate search results are presented, this might either be due to many valid terms are absent from the query or query words are ambiguous. Moreover, different users use same word to find different information. For instance, two users searching for the term "Apple" may have different perspective. It is not clear whether the query refers to Apple Gadget Company or a fruit? Personalization is solution to this problem. Web personalization means customizing the web environment according to the user interest which can be inferred from user's action, browsed page history, user preferences etc. For example, someone searching for "Apple" will get results for the fruit if they tend to search for information on fruits earlier, on the other hand will get results for the electronic gadget company if past searches centered on devices. It takes the advantage of information that obtained by the analysis of the user's behavior moving around the web sites in correlation with other data collected within the web context.

In this paper we emphasize in making voice browser more accessible for partially blind person, hence, we will also include the feature of individual profile generation. In that, users have to create their profile in which they have to provide their information like their area of

interest, occupation, whether they are partially sighted or normal.

In order to make the browser more relevant to the blind users we are using different techniques such that personalization, page ranking algorithm, alternate query generation algorithms. In this , first user's query in the form of voice will be converted into text format. This textual query will be regenerated by using user's profile, senses and synonyms along with some pre- processing techniques. Search results will be re ranked using user's profile and search history. After re ranking, resultant query will be again converted into the voice.

II. LITERATURE REVIEW:

Kraft ., et.al [1] state that the context in its general form refers to any additional information associated with the query in the web search field and also present three different algorithms to implement the contextual search instead of modelling user profile. Generally speaking if the context information is provided by individual user in any form whether automatically or manually, explicitly or implicitly, search engine can use the context to custom tailor search results. The process is named as personalization search.

Ivan Marcialis ., et.al [2] proposed, User's behavior is used to personalize the human computer interaction during information seeking tasks. It describes a personal agent, embedded in the Web browser, which is able to sense user's interest and help him during query formulation phase and, finally filter and sort the results according to the user's need. Searchy uses the DART P2P network to store user's profiles in order to enable collaborative recommendations.

R.K Agrawal.,et.al[3]Due to the increasing use of audio visual communication tools invites people to use new modernized tools in education. For the success of students is to adapt the modernized educational equipment by voice. Multi tools are always beneficial then single tool for educational purposes in web application. This is a modernized period where the teaching based upon web technology and media is better than verbal teaching. It increases the level of learning, teaching and also provides solid information. Furthermore it improves students speaking and listening comprehension skills by the help this application. In this they have presents an investigation on the possibility to integrate different types of features such as MFCC, PLP and gravity centroids to improve the performance of ASR in the context of Hindi language. Our experimental results show a significant improvement in case of such few combinations when applied to medium size lexicons in typical field conditions.

Fang Liu et al., [4] recommended personalized Web search for improving retrieval effectiveness. Modern Web search engines are generated to serve all users, independent of the particular requirements of any individual user. Personalization of Web search is to perform retrieval for every user according to his/her interests. In this paper, the authors propose a novel approach to learn user profiles from users' search histories. The user profiles are then utilized to enhance retrieval efficiency in Web search. A user profile and a common profile are studied from the search history of the user's and a category hierarchy, respectively. These two profiles are integrated to map a user query into a group of categories which corresponds to the user's search intention and provide a context to disambiguate the words in the user's query. Web search is performed according to both the user query and the group of categories. A number of profile learning and category mapping approaches and a fusion algorithm are presented and evaluated. Experimental outcomes show that this approach to personalize Web search is very effective.

Patel Jay., et.al [5] a novel approach is proposed that personalize web search result throughquery reformulation and user profiling. First, a framework is proposed that identify relevant search term for particular user from previous search history by analysing web log file maintained in the server. These terms are appended to user's ambiguous query. Second, the proposed approach proceeds the user's search result and re-rank the retrieved result by identifying interest value of user on retrievedlinks.

Anjali Choudhary. , et.al[6] proposed the approach of custom personalized searching in this they used page ranking algorithm, social signals to rank the pages, they also provide personalization by creating individual search history for each user on the browser , they have also focused on the search result to get customized according to the user demand and parameters like page authority and domain authority.

ShilpaSethi., et.al[7] proposed design of personalized search system in which query generator is used to capture all the senses of the main query and assist the user with alternate queries. In this paper they proposed personalized search system which uses two tools, one is personalization and another is alternate queries. In this personalization based upon user profile, click history and last action performed by the user is used to improve the ranking of search results.

Venkatesh. A., et.al [8] proposed the architecture of a speech interface for the partially impaired person, in this they have used the text to speech conversion to render the textual information.

III. PROBLEM DEFINITION:

The internet includes all types of information and knowledge based on the various elements: text, graphic, numeric, and to a lesser extent, audio, video and animation. However, the visually impaired learners are deprived of this very important tool of learning. The design of information on the web is visually oriented. The reliance on visual presentation places high cognitive demands on a user to operate such a system. The interaction may sometimes require the full attention of a user.

The World Wide Web (WWW) is growing exponentially per year thereby search engine provides the low quality of search results. Thus, the users get difficulty in getting the relevant information from the obtained search results. The quality of web search results depends on the information needs of the user and the searching techniques employed in the web search systems. So, the personalization is a general need in web search now-a-days.There are some problems in current browser if we start to thinking about the disabled people: 1) current browser not support speech facility. 2) Many browser that provide personalization do not provide security to user's data. 3) Existing browsers do not provide the functionality of text to speech, speech to text and personalization together.

So, the main objective of the system is to help visually impaired users in searching over the internet through the voice enable browser.

IV. METHODOLOGY:

STT CONVERSION AND PROFILE GENERATION MODULE:

In this module, User's profile will be created by registering user into the system. After this, the system will provide the search engine interface to the registered user. The user's speech will be converted into the text format. While searching, system will gather required or meaningful information about a user i.e. user id, domain of interest no. of times a web page clicked, highest time spent on a web page, action performed (save, print, bookmark, etc.) on web page etc. To deal with changing behaviour of user's interest and ranking the searched pages accordingly, this personalized information is used.

ALTERNATE QUERY GENERATION MODULE:

In this module, an alternate query will be generated to retrieve more relevant pages for the searched query. Query entered by the user is being normalized by using stopping and stemming techniques. Then, this refined query is rearranged by using senses and synonyms from Word Net.

QUERY PROCESSING MODULE:

In this module, internally, generated query will be again executed on the search engine and related dynamic information, i.e. web pages, will be retrieved by using Google API.

CRAWLING DATA:

It will provide web content from the World Wide Web.

RE-RANKING MODULE:

In this module, extracted pages will be re-ranked by using user's personalized information that is maintained in profile generation module. Web pages will be re-ranked by using page ranking algorithms and other actions performed by user.
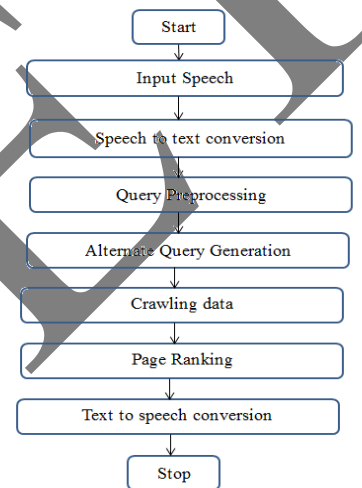

Figure: work flow of proposed architecture

TTS AND RESULT ANALYSIS:

Resultant re-ranked pages will be converted into speech format. So that user will be able to listen the search data. Finally, accuracy of the proposed approach will be calculated using performance measures.

V. CONCLUSION:

In order to make technology more familiar to the user its access should be made easier. As we know that visual internet access experiences various limitations, such as people who are physically handicapped (especially partially blind users) cannot use keypads or touch screens for giving instructions. Even after having crossed such impediments, it would be recommendable to additionally deal with the demands of present generation where ease of accessing internet independent of PC's and also hands free. For such task voice browsing appears to be more promising alternative. Temporal truncation in retrieving the appropriate data from internet according to user necessity is another major task in present circumstances.

In order to achieve this, the recommended approaches involved various techniques like personalization of query generation and web search. To increase relevancy, extracted search results can be re-ranked by using user's profile and search history. Hence, the users shall retrieve the desired information without any delay in time. Thus, in conclusion we can suggest that voice browsing accentuates on convenience of accessing webs.

REFERENCES:
1) Fauzia M. F. Rahman and Vikrant Chole "*A Voice Sensing Personalized Web Search,*" in International Engineering Journal For Research & Development Vol. 3, special issue 2, January 2017.
2) R. Kraft, C.C. Chang, F. Maghoul, and R. Kumar, "*Searching with context.* In proceeding of the 15th international conference on World Wide Web (WWW'06) pages 477-486, Edinburgh, Scotland, UK, 2006.
3) Ivan Marcialis and Emanuela De Vita, "*SEARCHY: An agent to Personalize Search results,*" Third International Conference on Internet and Web Applications and Services.
4) R. K. Aggarwal, and M. Dave, *"Performance evaluation of sequentially combined heterogeneous feature streams for Hindi speech recognition system",* Telecommunication Systems, Vol. 52, Issue 3, pp. 1457-1466 2013.
5) Fang Liu, C. Yu and WeiyiMeng, "*Personalized Web search for improving retrieval effectiveness,*" IEEE Transactions on Knowledge and Data Engineering,Vol. 16, No. 1, Pp. 28 –40, 2004.
6) Patel Jay, Pinal Shah, KamleshMakvana, Parth Shah, *"Review on Web Search Personalization Through Semantic Data"* 2015 IEEE 978-1-4799-608S-9.
7) Anjali Choudhary, Ankita Sadhwani, NidhiSaxena, "*A New Approach to Ranking Algorithm Custom Personalized Searching*", 2015 IEEE 978-9-3805-4416-8.
8) ShilpaSethi, Ashutosh Dixit," *Design of Personalized Search System Based on User Interest and Query Structuring"* 2015 IEEE 978-9-3805-4416-8. 4
9) Shanmugapriyan. K, S.K.V. Jayakumar, Venkatesh. A,"A Framework For Visually Impaired Web Users With Keyboard as User Interface", March 2015.
10) Dilip Kumar, Abhishek Sachan," *Bridging the Gap between Disabled People and New Technology in Interactive Web Application with the Help of Voice",* IEEE International Conference on Advances in Engineering &Technology Research (ICAETR - 2014), August 01-02, 2014.
11) Manny Rayner, Beth Ann Hockey, Nikos C Vhatzichrisafis, Kim Farrell, Jean Michel Renders, *"A Voice Enabled Procedure Browser for the International Space Station",* Association for Computational Linguistics, June 2005.
12) Dharani Perera, *"Voice Recognition Technology for Visual Artist with Disabilities in their upper limbs",* Conference on Computer Human Interaction, OZCHI'05, November 2005.
13) Ramin Halavati, Saeed Bagheri Shouraki and Saman Harati Zadeh, "*Recognition of human speech phonemes using a novel fuzzy approach",* Journal applied to soft computing, Volume 7, Issue 3, June 2007.
14) ZhengXu, Hai-Yan Chen, and JieYu, *"Generating Personalized Web Search Using Semantic Context",* the scientific journal, volume 2015.
15) Rathod Sweta D, Aakash Shah, Nilesh D. Solanki, *"personalized page ranking using genetic algorithm and dynamic profile",* IJARIIE volume 3, 2017.
16) Sumitro Ghatak, Avishikta Lodh, Enakshi Saha, Aakash Goyal, Aritra Das, SouravDutta, *"Development of a Keyboardless Social Networking website for visually impaired Social Web",* IEEE September 27.2014.