

Search Engine++: A Novel framework for Future Information Sharing

Narayanan Srinivasn

CSE Department, Velammal Engineering College, Chennai -66, Tamil Nadu, India.
prahladhan@gmail.com

Abstract

Search engine plays a vital role in Information Retrieval from repository of web pages. State of the art search engines are working fine for almost all the time except few instances. Results obtained from the search engine may be a text document, multimedia file, web page. Tendency of the search engine user community is to search result for an answer known query to get different aspect of it. Few search instances may not give result page sets at all. Because it is not guaranteed that all bag-of-words in the query used at-least once in one page of the World Wide Web. Even if they are used, the explanation for those may not be available elsewhere. Few examples are bag-of-words related to a specific territorial usage, bag-of-words related to a specific religious practice in a closed circle, bag-of-words related to territorial specific food item. Under these circumstances search engine used to give only a very few result pages (probably 2 to 3 pages) with good or poor relevance or the search not all produces any result page. This paper addresses this issue; provide an innovative frame work to solve this issue. The software architectural frame work proposed in this paper is implementable and will be very effective for future information sharing.

Keywords: Author Guide, Article, Camera-Ready Format, Paper Specifications, Paper Submission.

1. Introduction

Currently many search engines provides not only simply searching facility but also other services like e-mail. Exploiting this facility a new frame work Search Engine++ is provided in this paper. Though the following discussions hold sense to all the search engines, “Google” is chosen for convenience. In Google, an user can feed a query either by text or by voice. In the case of text query, language of the query can be in any of the supported language. Different types of users are also using search engine for different purposes. Based on their interaction ability to retrieve and locate the relevant page, search engine users can be categorized as novice, intermediate, brilliant.

Novice and Intermediate users require assistance or training in getting the required result from the set of resultant pages irrespective of the language used for searching. Outcome of a search for a particular query falls in one of these categories:

1. The expected result available in the first few resultant pages and the user feels lucky.

2. The expected result requires sequences of navigation through the resultant page sets and the user feels lucky only after tiresome effort.

3. Not all but some of the bag-of-words in the search query appears in the resultant pages but reading through any one of the resultant page gives the expected content and the user feels lucky.

4. Not all but some of the bag-of-words in the search query appears in the resultant pages but even after reading through all of them the user could not get any expected results.

5. Not even a single page is returned as result.

For the category numbered 4 other way is available to continue the search until getting the expected result. From the resultant pages the user can guess the appropriate forum and post his query to the forum members and possibly get the response.

For the category numbered 5, novice and intermediate users might feel stuck. They may simply abandon the search or carrying it unconsciously until they get the result. Either case is not desirable and strongly indicates the weakness in the current Technology. For this same category brilliant users may approach Social Media Network which is not strict as that of a forum of a particular interest.

2. Social Media Network’s limitations to substitute Search Engine

Though Wikipedia like knowledge repositories are evolving over a period of time by its users community contributions and illumine plenty of searchers day in and day out, it is the observed tendency of many search users to access Wikipedia through search engine’s results. It also implies that such knowledge repositories are also to be expanded in the future.

Plenty of knowledge is shared with the help of social media nowadays. But category numbered 4 and 5 indicates that such share is limited and not suit to the contextual need of the searcher. Deliberately, if the user is of category Brilliant, then he would post a query in the social network like Facebook, Twitter. The possibility of getting expected result is very low.

3. Proposed framework

The following situations should prevail for the Search Engine++ to be implemented:

1. Search Engine shall provide web based e-mail facility of its own or through other web based e-mail services.
2. A reporting mechanism to be implemented in the search engine to inform the Search engine server whether or not the search is success or failure. Category 4 and 5 are deemed to be a failure in search.
3. Having decided the search as success or failure, a provision is implemented to submit the Search query to the Search Engine server for the category numbered 4 and 5. This “Submit” operation can also be implemented as user controlled. That is if the searcher do not want to report failure, he may turn off the auto submitting of the search query to the server.
4. Two data base is maintained to collect that query which was decided failure. One data base with hierarchy and another data base without hierarchy.
5. Every search engine user should have e-mail id and should have the e-mail id registered with the server. The registration process collects user’s spatial details in the following hierarchy: Country -> State -> District -> Taluk -> Residence, depending upon the Country’s administrative hierarchy.

4. Interaction sequence for the searcher (Information Recipient)

- Step 1 : Searcher register his e-mail id with the server.
- Step 2: Search Engine++ mechanism activated by clicking a link in response mail for that registration.
- Step 3: Searcher signed out from e-mail service.
- Step 4: Searcher searches for a query. If the result found no need of Search Engine++, continue here else go to step 5.
- Step 5: Sign in to the Search Engine++ interface available with the search engine.
- Step 6: Announce the search as failure and submit the search query to the server.
- Step 7: Sign out from Search Engine++ and continue other searching.

5. Interaction sequence for the Information donor

- Step 1 : User subscribe to be an Information donor with the search engine through a registration process, which collects all the relevant information required about the user.

Step 2: Once registered, a mechanism let the user to access the query data base at interested hierarchy. If user knows answer for any query, can reply for the query.

Step 3: User exit from the query data base.

6. Architectural diagram for the proposed solution

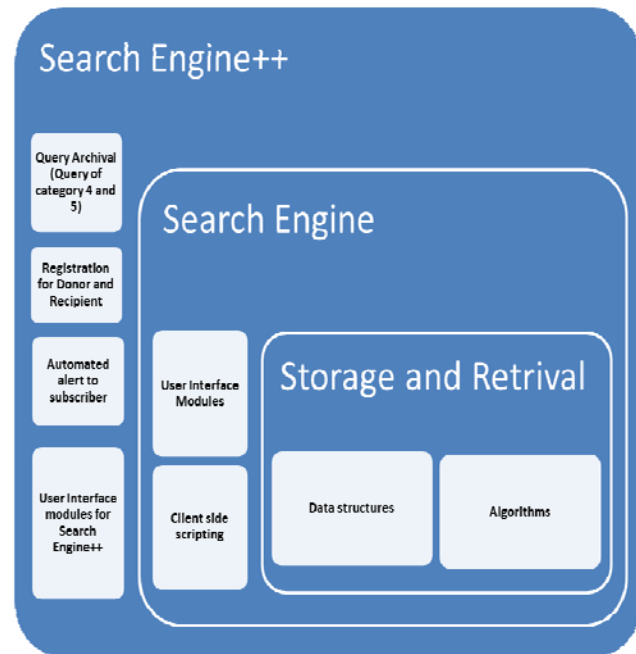


Fig. 1 Architecture of Search Engine++

4. Activities performed by the Search Engine++ server and future work

Search Engine maintains data base to collect queries and response from the user. Structure of the data base suitable for the current solution is a future work. Similarly privilege level and the interaction items (menu, text box etc.) are also issue for the future discussion. A user will be alerted through an e-mail when he gets an answer for his query.

5. Conclusion

Search Engine++ frame work is the need of the hour and will be very effective for Information Sharing. Microscopic details and implementation of it along with testing results are reserved as a future work.

References

[1] Kristen Purcell, Joanna Brenner, Lee Rainie, “ Search Engine Use 2012” , A project of the Pew Research Center, pp:3, March 9,2012.

Narayanan.S Completed Masters in Software Engineering from Anna University. Currently working as an Assistant Professor in the department of Computer Science and Engineering, in Velammal Engineering College. Successfully guided many projects in career.