

A Review on Patient Centric Meta-Search Engine for Quality of Care Services

Sagar A. Raut, Sandeep
B. Gaware
UG Students
Computer Department
JSPM JSCOE, Pune

Deepak D. Shivarkar,
Gaurav S. Hambarikar
UG Students
Computer Department
JSPM JSCOE, Pune

Syed Ahmed
Assistant Professor
Computer Department
JSPM JSCOE,
Pune

ABSTRACT

Current scenario in India people don't know the actual service cost of any hospital. About 50 to 60 percent people in India are poor and middle class. They cannot afford the service charges provided by the hospital. The service charges of every hospital varies. People face difficulties in searching the hospital which provides quality services in minimum price. Most of the people pay unnecessary charges to those hospitals rather than other hospitals which provides same services in low price. Proposed system gives solution to all the above problems. The system compares multiple hospitals based on the area, services, cost, ratings or reviews. According to the people requirements it provides list of hospitals in filtered format. Based on the results people can decide which hospital provides the same services in minimum cost. Our proposed system works on the basis of Meta Search Engine, web mining algorithms like content based and usage based.

Keywords

Content based page ranking, usage based page ranking, meta-search engine, two phase page ranking phase page ranking.

1. INTRODUCTION

In India it is very difficult to find a hospital which gives good services and care in minimum cost. The hospital charges are varies from place to place but services are the same so it is very difficult task to find those hospitals as well as user don't have a time to analyze all hospitals according to cost wise and then its service quality wise. Because there is lot more data on the World Wide Web about each and every hospital. Proposed system solve all this difficulties, proposed system retrieve all the results of user's query. Then the main task of system is arranging the result as per user's query. Till now for web mining different page ranking algorithms are used. But in these system two page ranking algorithms are combined together to find better and efficient results [1]. We combine content based page ranking and usage based page ranking. This paper presents a well-organized web page ranking algorithm. It is based on content, and usage information of the pages to improve the sequence of the results. First, we fetch the relevant web pages from the search engine result page for the given user query. Extracted pages are stored in our database. After getting the relevant web pages from our database, our proposed algorithm is applied on the pages to order the result. Finally ordered results are displayed to the user [2].

2. LITERATURE SURVEY

Hospital search websites like practo gives information about hospitals, it books the online doctor's appointment, doctors visiting charges, reviews and timing. Other websites of each hospital give information about specific things like hospital location, hospital address, hospital phone number, number of

ambulance, number of general wards, etc. but none of the above application gives detail information of services provided by hospital and the cost of the services. Also it is very difficult to user search each website and compare them to each other and find best one among them according to user's budget. So above all difficulties are overcome by proposed system. In our system we used web mining techniques such as page ranking for data collection. Till now for web mining different techniques used like content based, usage based, hybrid, structure based and many others. In our system for data mining we combined two different algorithms together content based and usage based to get efficient and appropriate result [3]. User put his query in our system to find better hospital facility in minimum cost. Then this query is passed to the search engine. Now query is processed in search engine to fetch the results. Top n results are fetched from search engine result page. Our proposed algorithm computes the rank of a web page in two phases. In the first phase, score will be calculated based on the content relevancy and in the second phase rank will be given based on the user access time [4]. By adding these two scores the total rank of the web page can be obtained. At last, the normalized value of each result page is sorted in descending order to get the most relevant page on the top most place. Similarity rank determines the relevance of a page with respect to query terms by counting the number of occurrences of the query terms within the web document. It gives weight based on the locality of the keyword. Now this results are shown through the proposed system [5].

Table 1. Survey table of Meta Search engines

Sr No.	Meta Search Engine Name	Search engines	Main Features
1	Clusty	Ask.com, Gigablast, Live, NY Times, ODP, Shopzilla, Yahoo news, Yahoo stocks[6]	Term clusters
2	Dogpile	Google, Yahoo, Ask, Live	Search suggestions(related terms); recent searches
3	IxQuick	All the web, Exalead, Qkport, Ask, Gigablast, Wikipedia, Bebo, MSN,	"Star" system(the more stars=the more search engines agree on the listing rankings).

		Winzy, CNN,NBC, Yahoo, EntireWeb, OPD[7]	
4	IBoogie	All the web, MSN	Term clusters
5	Kartoo	N/A	Search results on a map
6	SurfWax	CNN, Yahoo news, Hotboot, ODP, MSN, All the web	Snapping=displaying the summary of the page containing the search query
7	Mamma	Ask.com, Abut.com, Entireweb, Business.com, Gigablast, Wisenut, ODP	Add/Exclude any of the search engine
8	Search	Google, Ask.com, MSN, ODP	-
9	Metacrawler	Google, Ask.com, MSN, Yahoo	-
10	Fuzz Find	Google, MSN, yahoo	Sort the result based on any of the sources

Table 2. Survey table of web mining

Paper No.	Mining technique	Algorithm
1	Web content mining, usage mining	Tag analyzer, Event explore
2	Web mining	Structure mining, content mining, usage mining
3	Web content mining, structure mining	Page rank, weighted page rank, HITS
4	Web content mining, usage mining, structure mining	Page rank, weighted page rank, HITS
5	Web structure mining	Page rank, weighted page rank, query dependent page rank, SQD page rank, distance rank
6	Web structure mining	Page rank, HITS, weighted page rank
7	Hybrid page ranking	Page rank, HITS, weighted page rank, distance rank
8	Hybrid web mining	Content based, usage based, matching and

		ranking algorithms
9	Search engine optimization	Meta search engine, page rank
10	Web content mining	Weight of outbound link page ranking algorithm[8]

3. PROPOSED SYSTEM

The HTML source code of a user query results or web pages are downloaded into our database and convert it into a DOM tree. DOM is an interface which allows scripts and programs to dynamically access and handles all the elements such as content, structure and style of web pages. We traverse through Dom tree to identify tags like title, Meta, heading and paragraphs. Title tag is used to identify title of the web document, meta tag gives information about web document, It is used to state page description, keywords of the page, author of the document, last modified date etc. the h1 tag generally used for title also. This tags gives information about the content in the document [9].

3.1 Tag analyzer algorithm steps

- Build a DOM tree of downloaded web page
- Traverse each node of a DOM tree
- Find title, heading, meta, paragraph tag

Calculate the page rank according to content of data which are in above tag

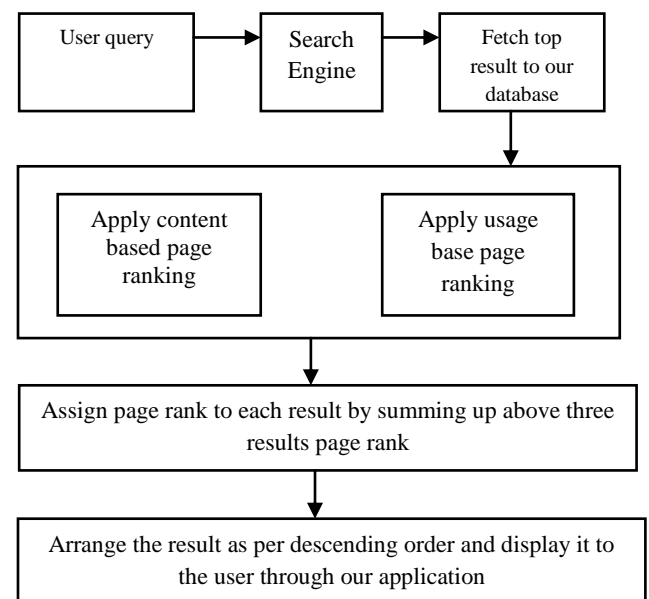


Fig 1. Flowchart of proposed system

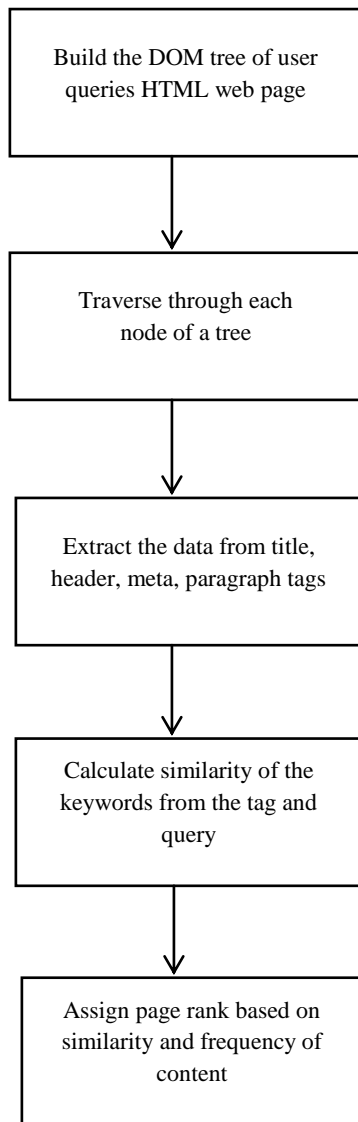


Fig 2. Flowchart of tag analyzer

3.2 Event Explore Algorithm

This technique is used to compute the interest of the user on certain web page. It find the how much time user spend on certain web page. If user spend much more time on specific web document then it consider an important document and its page rank should be more. When web page is open in user's browser, timer will be triggered. Every second timer will invoke these event explore function to check whether the user is idle or active. This verification is done by binding mouse events and keyboard events. If the user is continuously idle for 120 seconds i.e. user does not perform any activity on web page then timer is again reset. Otherwise user access time computed using timer value [10].

Steps-

- Web document open in browser
- Start timer
- Invoke event explore function every second to check whether user is idle or not
- If user is not idle for 2 minutes reset time
- Else user access time will calculate

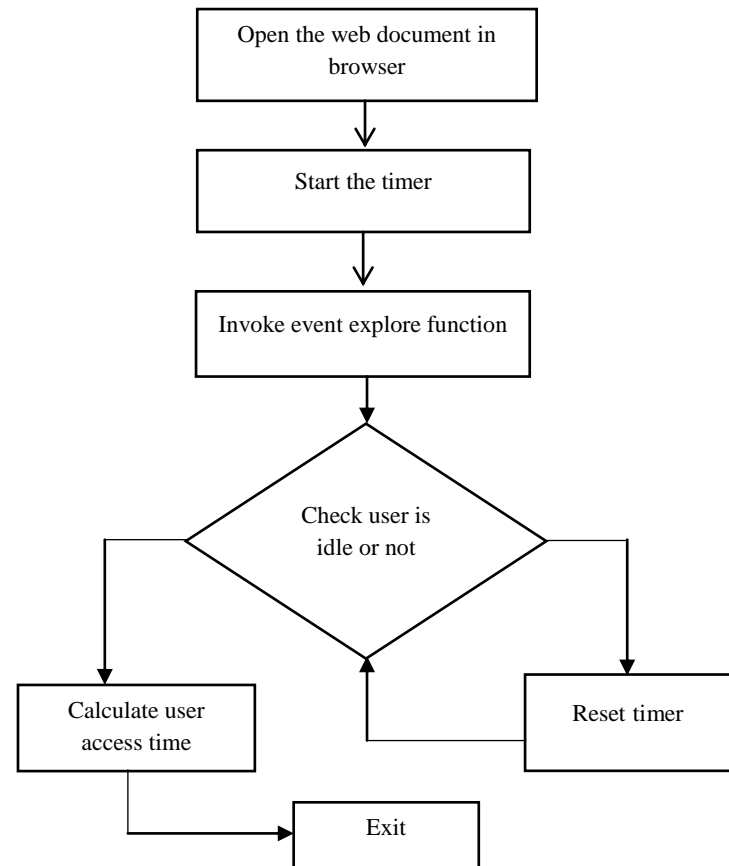


Fig 3. Flowchart of event explore algorithm

4. CONCLUSION AND FUTURE SCOPE

This web application compares different hospitals according to users query and gives best hospital services information which are in user's budget. This project reduces user's effort and time for comparing different hospital websites and find best from it which are given quality services in minimum cost. It also gives information about hospitals, its quality of services through user's reviews and gives information about available services in hospital like number of available ICU, general wards, ambulance, etc. For comparing different hospital application used web mining techniques like content based page ranking algorithm and usage based page ranking algorithm. And when the structure based algorithm will use in system the searching time of the hospital related information will be reduced.

5. ACKNOWLEDGMENTS

We are thankful to our department staff and head of the department for continuous motivation and support while making this article.

6. REFERENCES

- [1] Syed Ahmed, Dr.PVRD Prasad, 2018, A Framework For Decision Making And Quality Improvement By Data Aggregation Techniques On Private Hospitals Data, ARPN, 4337-4345.
- [2] Syed Ahmed,Dr.PVRD Prasad,2018, Analysis Of Single And Hybrid Data Mining Techniques For Prediction Of Heart Disease Using Real Time Dataset,IJET, 97-99.M.Usha, Dr. N.Nagadeepa , 2018 “Combined two phase page ranking algorithm for sequencing the web pages” .
- [3] Anuradha, G.Lavanya Devi and M.S Prasad Babu,2014, “Role of Web Mining Algorithms for Ranking Web Pages”.
- [4] Monica Sehgal, “Analysis of Link Algorithms for Web Mining”,2014,
- [5] Rekha Jain, Dr. G. N. Purohit, 2011,“Page Ranking Algorithms for Web Mining”.
- [6] Ms.M.Sangeetha, Dr.K.Suresh Joseph,2014, “Page Ranking Algorithms used in Web Mining”
- [7] N. V. Pardakhe, Prof. R. R. Keole,2013, “Analysis of Various Web Page Ranking Algorithms in Web Structure Mining”.
- [8] Sharmila Subudhi, 2013,“Designing a Hybrid Page Ranking Algorithm for Semantic Web Search Engine” ISSN: 2249 – 8958, Volume-3, Issue-1.
- [9] Nirali Arora, Sharvari Govilkar, 2016,“Efficient Hybrid Ranking Algorithm for Search Engine” 624-1629.
- [10] GangLuo,et al. (2008)IntelligentOutput Interfacefor Intelligent Medical SearchEngine, Association for the Advancement of Artificial Intelligence, Intelligence, pp 1201 1206
- [11] NIDHI SHALYA, SHASHWAT SHUKLA, et al.2015.“An Effective Content Based Web Page Ranking Approach”.
- [12] Syed Ahmed Yasin, Nikhil Kumar, Ravindra P Bachate,2016, Comparative Analysis of ID3 and Naïve Bayes for Heart Disease Prediction,IJCA.