iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 7 **Issue:** XI **Month of publication:** November 2019

DOI: http://doi.org/10.22214/ijraset.2019.11009

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue XI, Nov 2019- Available at www.ijraset.com

Analysis of Health Care Services using Meta Search Engine and Web Mining

Neha Badhe[1], Madhura Bhokare[2], Snehal Patil[3], Pooja Raut[4], Prof. Syed Ahmed[5]
[1, 2, 3, 4]*Computer Department, JSPM, JSCOE Pune Savitribai Phule Pune University*
[5]*Research Scholar and Assistant Professor, JSPM JSCOE Pune*

Abstract: ***Many existing page ranking algorithms are used in web mining to display the resultin search engine result page. But this existing algorithms are either best on the inlinks and outlinks of the page or content of the page. Current scenario in India is thatpeople don't know the actual service cost of any hospital. About 50 to 60 percentpeople in India are poor and middle class. They cannot afford the service chargesprovided by the hospital. The service charges of every hospital varies. People facedifficulties in searching the hospital which provides quality services in minimumprice. Most of the people pay unnecessary charges to those hospitals rather thanother hospitals which provides same services in lower price. Proposed system givessolution to all the above problems. The system compares multiple hospitals based onthe area, services, cost, ratings or reviews. According to the people's requirements it provides list of hospitals in filtered format. Based on the results people can decidewhich hospital provides the same services in minimum cost. Our proposed systemworks on the basis of Meta Search Engine, web mining algorithms like contentbased , usage based and structured based algorithm as well so as to find results in minimal optimal time .***
Keywords: Web mining , content based page ranking, usage based page ranking, meta-search engine,two phase page ranking phase page ranking.

I. INTRODUCTION

Now-a-days searching for hospitals that provide good services a affordable price is difficult. There are many page ranking algorithms that use web mining to displaythe result in search engine result page. But these existing algorithms are either basedon inlinks and outlinks of the page or content of the page which eventually becomesdifficult for the customers to decide. This leads to a need of web page rankingalgorithm concerning content and usage of the pages.TPPR (Two Phase Page Ranking) technique computes the score in two phases based on the output of TPPR algorithm, the URLs are sequenced and displayed to theuser. Event Explore techniques detects whetherthe user is idle or active on the page. The proposed algorithm produced better performance and displays the mostrelevant web pages in the top of the result.

II. 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 ofambulance, 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. 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. 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 thrown the proposed system .

A. Tag Analyzer Algorithm Steps

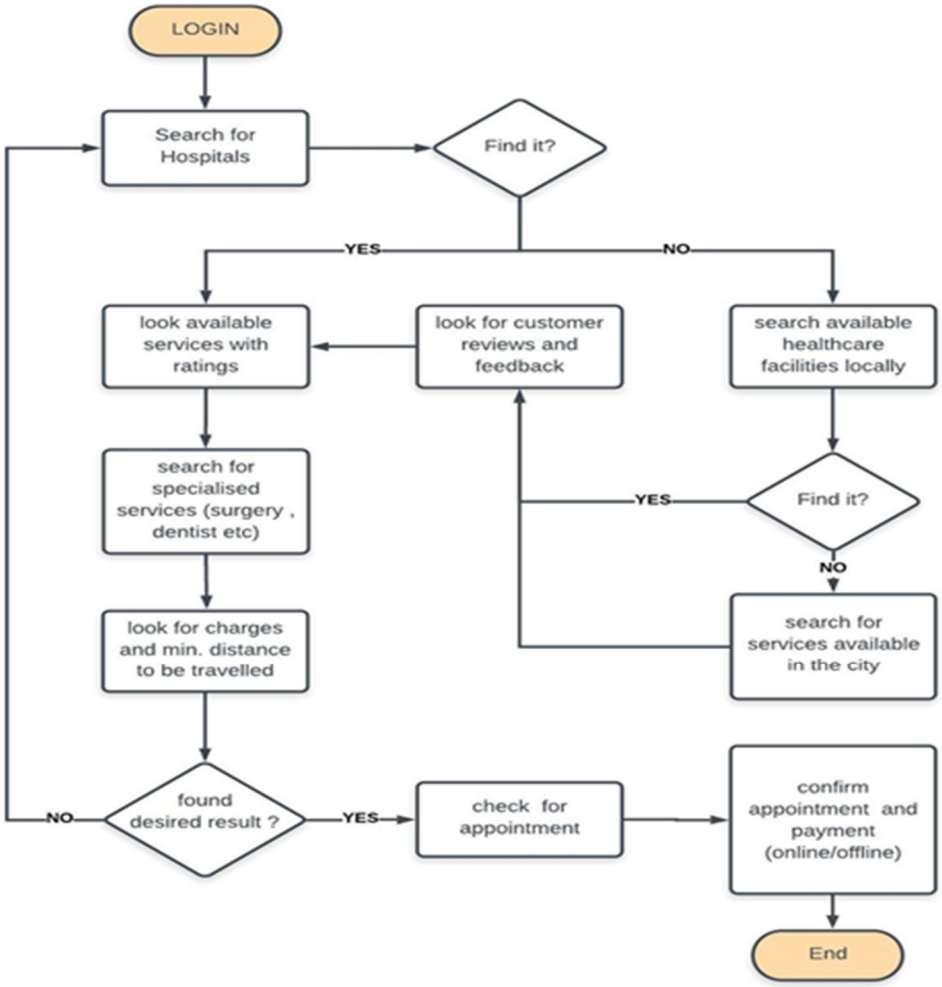
1) Build a DOM tree of downloaded web page .
2) Traverse each node of a DOM tree .
3) Find title, heading, meta, paragraph tag.
4) Calculate the page rank according to content of data which arein above tag

B. Event Execution

This technique is used to compute the interest of the user oncertain web page. It finds how much time a user spends oncertain web page. If user spends much more time on specific web document then it consider to be an important document and its page rank is updated. When web page is opened in user's browser, timer will be triggered. Every second timer willinvoke these event explore function to check whether the useris idle or active. This verification is done by binding mouse events and keyboard events. If the user is continuously idlefor 60 seconds i.e. user does not perform any activity on webpage then timer is again reset. Otherwise user access timeis computed using timer value .

Steps

1) Open the site in browser.
2) Start timer.
3) Invoke functionsafter every 60 seconds tocheck whether the user is idle or not.
4) If user is idle for more than 60 seconds then reset the time.
5) Else user access time will be calculated accordingly.



ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue XI, Nov 2019- Available at www.ijraset.com

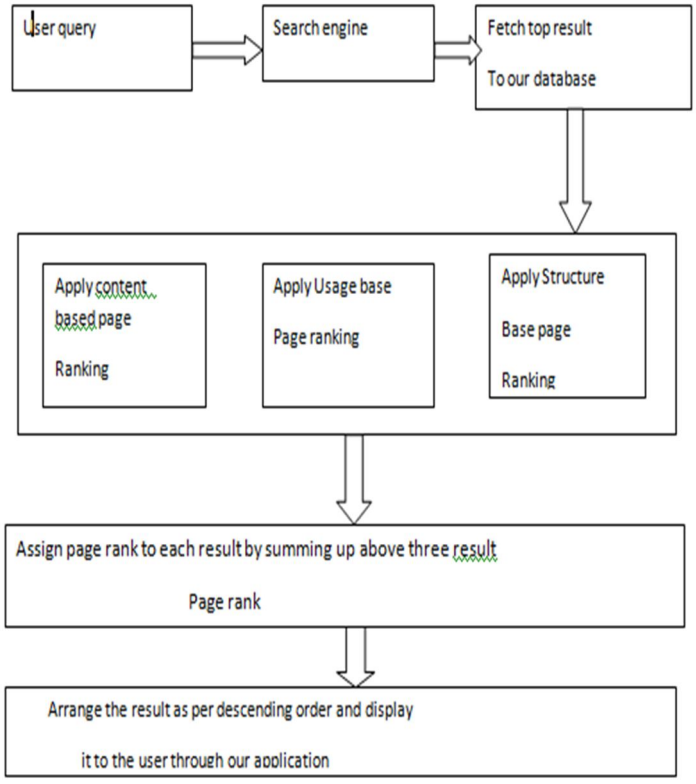
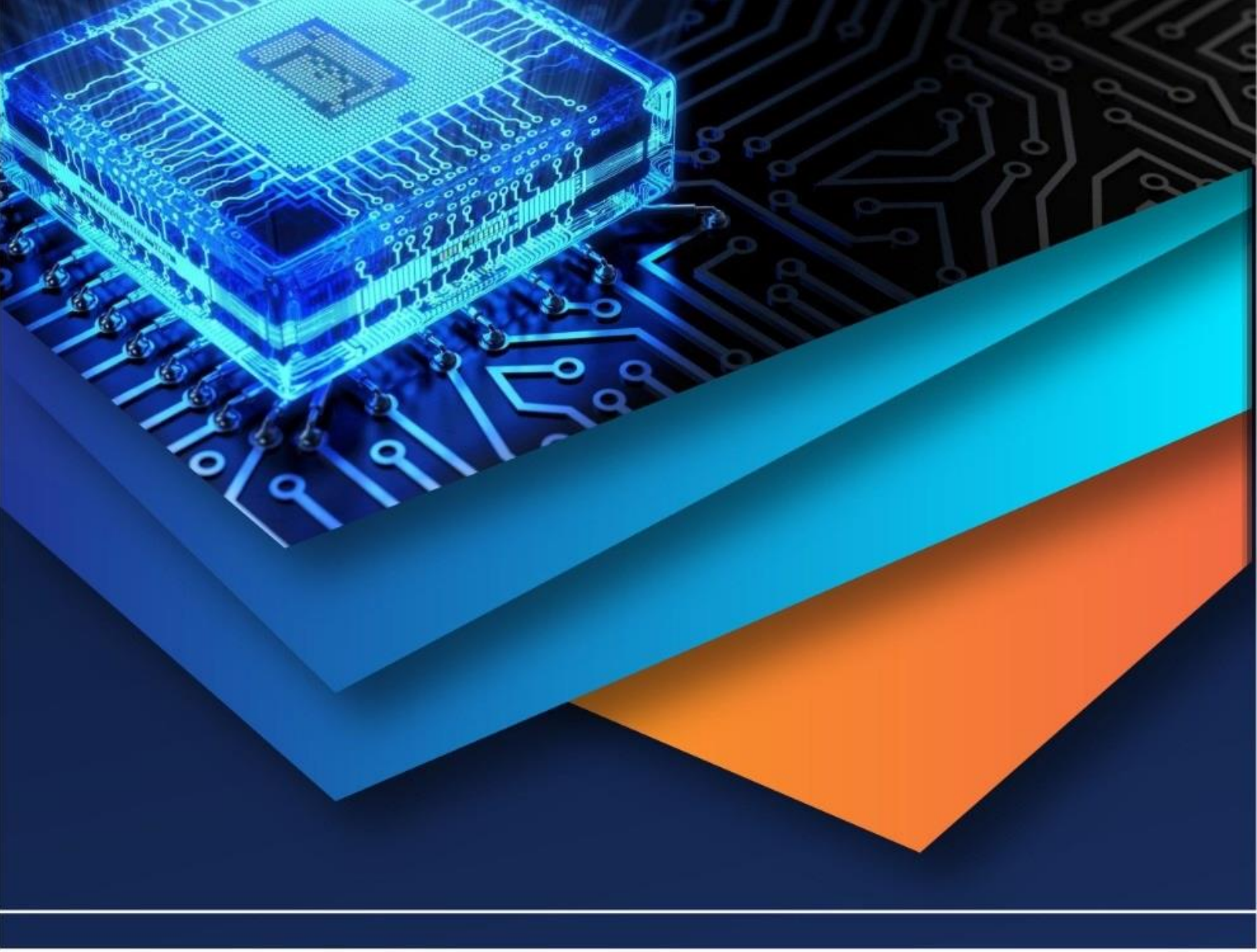


Fig.flow chart of proposed system

III. CONCLUSION AND FUTURE SCOPE

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



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY