ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 5, Issue 5, May 2016

Text And Image Searching Method Based on Genetic and Page Rank Revised Algorithms

Khandve Dhanashri,Sudrik Snehal,Wagh Sonali,Yewale Savita,Prof.Borse.P.T.
Dr.D.Y.Patil School Of Engineering, Lohegaon,Pune
Computer Enginering Department

***Abstract*—Now days people usually access the information over the internet with the help of search engines. To get relevant search while searching over the internet is big challenge because of data patterns and different large terms. In this paper we are going to introduce a system with a search engine developed using Genetic algorithm, Page Rank revised algorithm. Search Engine is a program which gives relevant result according to the users query. Into days world there are number of search engines available like google, yahoo, Bing, ask etc. This paper presents a new rank Based system by using page Ranker Reviser algorithm with data mining techniques and also studied result of K-Nearest Neighbor and page rank revised algorithm.This technique work together to rank the searching data and provide a better result according to users query.**

Keywords-Search Engine, Genetic Algorithm, K-Nearest Neighbor Algorithm.

I. INTRODUCTION

Now days internet is the best way for searching. For searching over internet people uses search engine. Search engines are the programs which find the specific pages for users according to their query.The main purpose of search engine is to provide a result according to the users query.The amount of information is increasing day by day speedily that creates the challenge for information retrieval.There are so many tool for execute effective searching.

In proposed system, genetic algorithm and page rank reviser algorithm together provide a relevant search result according to users query. In the proposed system, user can login using username and password, when user want to search data over internet or on particular PC, whatever user search over that search engine , searching result will be store in his account. Next time when he wants to search something related to his history, this system will help to provide a relevant result according to his query.Ranking and Indexing plays important role to provide the relevant result. Ranking system gives more suitable result for user query.Ranking is the method which provide an order of substitutions which presents their relative preference. In search engine, indexing collects parse and store data to enable fast and accurate information retrieval. [1].

II. COMPARISON BETWEEN EXISTING SYSTEM AND PROPOSED SYSTEM

Existing system:

According to existing system objectives create the profile as per users interest. Existing system also provide a relevant search according to rank base but it take times. Existing system is not secured for searching. It do not provide any security to searching data. Anyone can easily access this system. Every time whenever user enter same query the system will provide the same result. It do not save any history of user searching.

Proposed system :

In proposed system, it provide two option either user want to search data over internet or particular Pc. A approach to access control in searching methods the user can login using user name and password or his/her admin login. More than one user can create account or profile on one system for recognition purpose and also security purpose. This system is secured. No one can easily access the system. It reduce the time period in searching process. This system provide the appropriate relevant data on user requirement.

III. PROPOSED SYSTEM

In proposed system , it provide login option. User can login using username and password. For security purpose and for appropriate result, more than one user can create account or profile on one system. This system provide relevant result according to users query. This system also provide option of searching for user interest. User can search on internet and also on Pc. The searching result will be saved and update for particular user, this result will be save on users profile. This system also update the user profile history according to the searching result. Because of that it provide more accurate and appropriate result according to user query.

Indexing is also done while searching process this will help to get quick result and it reduces time period of searching process.Using page rank reviser algorithms, user get relevant page on their rank. Multiple search can be done at the page ranking algorithm. Here we have indexing the file. when the file is indexing the path and address of this folder is indexed.

The feature vector from the client are then used in page rank reviser algorithm used to find a weight vector of the content and location contents.

Following figure 1 Shows the working of proposed system

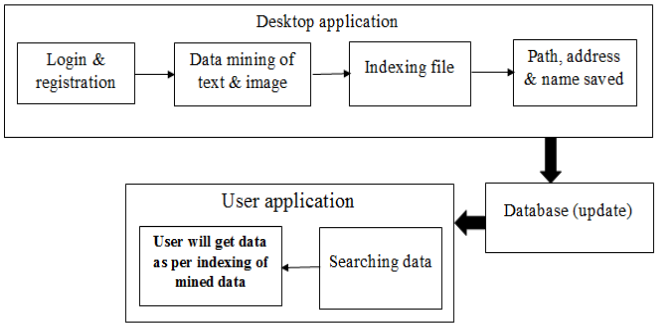

Figure1. Working of proposed system

IV. ARCHITECTURE

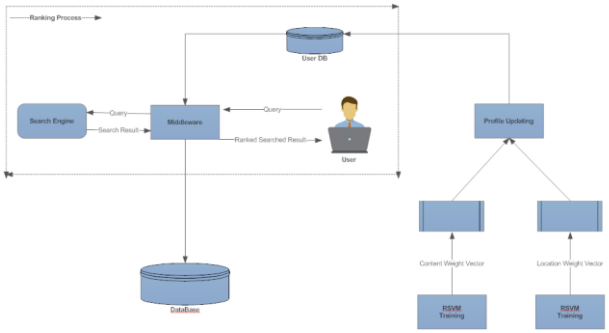

Fig:-System Architecture

Profiling:-
In this module,the System can generate the profile of user as per the users interest or as per requirement.Profiling is usefull for maintain the history of the searching data and working process.profiling provides the security to user data.[1]

Indexing:-
In this module, indexing collects, parses and stores data to facilitate fast and accurate information retrieval. The purpose of storing an index is to optimize speed and performs in finding the relevant documents for a search keyword.Without index process the search engine would check each document in the system so this process take a more time and computing power. [1]

Ranking :-
In this module,the ranking process is performed.ranking process help for find the relevant result.once a keyword is entered into a sarch box,search engines will check pages or another data with in there index that are a closest match;a score will be assign to this pages ,show the ranking result.[2][4]

V. ALGORITHMS

Genetic Algorithms

Genetic Algorithms (GAs) are adaptive heuristic search algorithm based on the evolutionary ideas of natural selection and genetics. An intelligent exploitation of a random search which is represented by Genetic algorithm is used to solve optimization problems.The basic techniques of the GAs are designed to simulate processes in natural systems necessary for evolution.[3].
Genetic Algorithm is better than conventional AI in that it is more robust. Unlike older AI systems, they do not break easily even if the inputs changed slightly, or in the presence of reasonable noise. Also, in searching a large state-space, multi-modal state-space, or n-dimensional surface, a genetic algorithm may offer significant benefits over more typical search of optimization techniques like linear programming, heuristic, depth-first, breath-first, and praxis.[5] [9].

Implementation Details:-

Based on Natural Selection:-

After an initial population is randomly generated, the genetic algorithm evolves the through three operators:
1)selection - which equates to survival of the fittest;
2)crossover- which represents mating between individuals;
3)mutation - which introduces random modifications.

1. Selection Operator

key idea: give prefrence to better individuals, allowing them to pass on their genes to the next generation.
The goodness of each individual depends on its fitness.
Fitness may be determined by an objective function or by a subjective judgement.

2. Crossover Operator

Prime distinguished factor of GA from other optimization techniques
Two individuals are chosen from the population using the selection operator
A crossover site along the bit strings is randomly chosen
The values of the two strings are exchanged up to this point
If S1=000000 and s2=111111 and the crossover point is 2 then S1'=110000 and s2'=001111
The two new offspring created from this mating are put into the next generation of the population
By recombining portions of good individuals, this process is likely to create even better individuals

3. Mutation Operator

With some low probability, a portion of the new individuals will have some of their bits flipped.
Its purpose is to maintain diversity within the population and inhibit premature convergence.
Mutation alone induces a random walk through the search space
Mutation and selection (without crossover) create a parallel, noise-tolerant, hill-climbing algorithms

Effects of Genetic Operators

Using selection alone will tend to fill the population with copies of the best individual from the population
Using selection and crossover operators will tend to cause the algorithms to converge on a good but sub-optimal solution
Using mutation alone induces a random walk through the search space.
Using selection and mutation creates a parrallel, noise-tolerant, hill climbing algorithm

Algorithm Steps:-

Step 1: randomly initialize population(t)

Step 2: determine fitness of population(t)
repeat

Step 3: select parents from population(t)

Step 4: perform crossover on parents creating population(t+1)

Step 5: perform mutation of population(t+1)

Step 6: determine fitness of population(t+1)
until best individual is good enough

CONCLUSION

In this developed system,we have proposed a improved searching method by using genetic algorithm and page rank reviser algorithm.
This system provide relevant searching result according to user query. System can be provide relevant result and reduce the time period for searching process.

ACKNOWLEDMENT

We would like to take this opportunity to express our profound gratitude and deep regard to our project guide "Prof. P.T.Borse" for his guidance, valuable feedback and for constant encouragement for the project. Working under him was extremely knowledgeable experience for us.

REFERENCES

[1] Michael E. Houle and Michael Nett, "Rank-Based Similarity Search: Reducing the Dimensional Dependence" , IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 37, No. 1, January 2015.
[2] Yuefeng Li, AbdulmohsenAlgarni, Mubarak Albathan, Yan Shen, and MochArifBijaksana,"Relevance Feature Discovery For Text Mining", IEEE Transactions On Knowledge And Data Engineering, Vol. 27, No. 6, June 2015.
[3] FabrizioLamberti, Member, IEEE, Andrea Sanna, and Claudio Demartini, Member, IEEE, "A Relation-Based Page Rank Algorithm for Semantic Web Search Engines" IEEE Transactions On Knowledge And Data Engineering, Vol. 21, No. 1, January 2009
[4] Narayan L. Bhamidipati and Sankar K. Pal, Fellow, IEEE, "Comparing Scores Intended for Ranking" IEEE Transactions On Knowledge And Data Engineering, Vol. 21, No. 1, January 2009
[5] NishaBansal And Dr. Paramjeet Singh," Improved Web Page Ranking Algorithm Using Semantic Similarity And HITS Algorithm", *International Journal Of Emerging Trends & Technology In Computer Science (IJETTCS)*,Volume 3, Issue 4 July-August 2014
[6] *Seema Rani***1**, *UpasanaGarg,"*A Ranking Of Web Documents Using Semantic Similarity And Artificial Intelligence Based Search Engine*",* International Journal Of Science, Engineering And Technology Research (Ijsetr), Volume 3, Issue 12, December 2014
[7] Asuncion and D. J. Newman. (2007). UCI machine learningRepository. [Online].Available: http:/www.ics.uci.edu/Mlearn/MLRepository.html
[8] M. Ankerst, M. M. Breunig, H.-P. Kriegel, and J Sander, "OPTICS:Ordering points to identify the clustering structure," in Proc. ACMSIGMOD Int. Conf. Manag. Data, 1999, pp. 49–60.
[9] Improved Searching Method Based on Genetic and Page Rank Revised Algorithms Khandve Dhanashri,Sudrik Snehal,Wagh Sonali,Yewale Savita,Borse P.T.IJRCET-2015