

# NOVEL KEYWORD SEARCH ALGORITHM WITH GRAPHICAL PASSWORD

Prerna Chopra<sup>1</sup>, Vidisha Agrawal<sup>2</sup>, Nidhi Sharma<sup>3</sup>, Yusera Farooq Khan<sup>4</sup>  
<sup>1,4</sup>M. Tech scholar, <sup>2,3</sup>Associate Professor

<sup>1,2,3,4</sup>Computer Science and Engineering Department, NIMS University Rajasthan, India.

**Abstract:** With the advance in the field of the information technology there are number of successful strategies and proficient techniques of keyword search which are right now being used. Keyword search technique is to a great extent utilized for searching unstructured information. With time it has brought about improvement of number of techniques of rating and positioning of question results and to evaluate the adequacy of those techniques. In database group principle concentrate is on gigantic gathering of the organized information which brought about advancement of number of counterfeit techniques and strategies for preparing or executing the organized questions on the database. In today's time, the blend of database techniques and the information recovery techniques is extremely fundamental. With the gigantic development of internet and expanding clients of internet requested necessity of keyword search techniques and to augment idea of keyword search over social information.

**Keywords:** Green Computing , Keyword Search , Visual Password.

## I. INTRODUCTION

Part of distributed computing is extremely uncommon and is one of the creating technology in the PC field, it is utilized as a part of many fields, for example, Education, Medical, Networking, Entertainment et cetera the rundown proceeds. In all the previously mentioned list control effectiveness assume a conspicuous part, as there is a need of energy in each field. Control utilization ought to be kept up as overabundance use of it will prompt carbon emanation discharge to condition which as its gigantic negative effect on the environment.[1] To conquer the issue one of the conceivable arrangement may be 'Cloud Green Computing', the characterized paper gives one vital answer for the issue by determining the algorithm in light of group technique which will lessen the Power utilization by for some degree, therefore of that carbon outflow into the earth will descend. Performing converging to streamline vitality utilization while giving required execution raises a few concerns. Right off the bat, combination strategies should deliberately choosing which workloads ought to be consolidated on a given regular physical server. Workload asset utilization, execution, and vitality uses are not vital. Attempting to comprehend the way of the creation is in this way basic to choose which workloads can be together stuffed. Furthermore, besides, there exists an ideal execution and vitality point. This happens in light of the fact that solidification prompts corruption in execution which causes the execution time to build, eating into the vitality investment funds from

decreased static sit out of gear vitality. Further, the perfect point changes with satisfactory debasement in execution and application blend. Finding the best point and following it as workloads change, in this way ends up noticeably essential for vitality productive consolidation.[2]. Keyword search techniques are exceptionally helpful for breaking down both the organized and also the unstructured information which contains the expansive measure of the textual information. In our research paper we will investigate different keyword search techniques and we will likewise attempt to break down the regions on which we can work to enhance execution of keyword search algorithms.

**Organized and Unstructured Data:**

Organized Data is one in which information is sorted out regarding structures i.e. relations or tables and that structure will take after a strict database blueprint Like in SQL. These tables further compose the information as far as lines and sections, where lines allude to tuples or records and segments alludes to qualities and These all tables are limited together with some cardinality or relationships e.g. one to many, many to many and so on as show in figure 1.

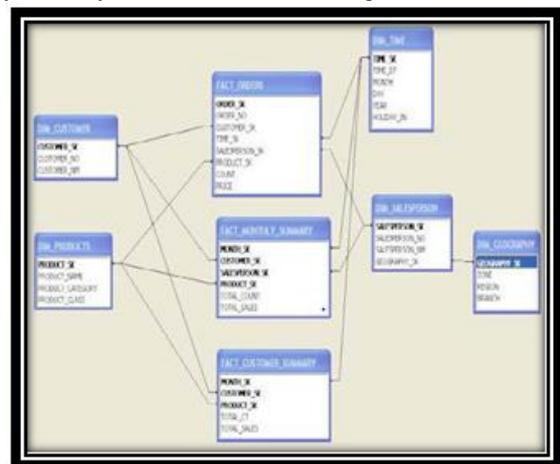


Figure 1: Relationship among Product, Customer, Sales person and Monthly Record via Structured Data.

Example: As appeared above in Figure 1: Structured information Contain information about Customer, Sales Person, their Orders, Monthly Records et cetera by Specific Relationship(Using Primary and Foreign Key) among these records.

Unstructured information is absolutely Opposite to organized information. It contains information that don't composed in any predefined Schema. It can be in any shape like Audio, Videos, JPEG Files, Pdfs , Text Files et cetera and it is normally alludes to information that doesn't dwell in a customary line segment database.



Figure 2: Unstructured Data

## II. LITERATURE SURVEY

Yue Lu, Chew Lim Tan [1] proposed that a tremendous measure of archive pictures are open in the Internet and computerized libraries. They locate that, the greater part of them are stuffed in PDF documents and are compacted utilizing CCITT Group 4 principles for sparing storage room and accelerating transmission. There is accordingly noteworthy intending to build up the strategies for straightforwardly searching keywords from these archives. In this paper, they display a compacted design coordinating technique for searching keywords from the CCITT Group 4 packed archive pictures, without express decompression. As per the CCITT Group 4 guidelines, each coded position shows that the present pixel shading is not the same as its past pixel, with the exception of the following coded places of the pass mode. In their work, they separate these changing components from the packed pictures specifically. The changing components are used to portion and bound the word questions, and are utilized for measuring the similitude of two word pictures. The associated parts are marked in view of the line-by-line procedure as indicated by the relative positions between the changing components of the present coding line and the changing components of the reference line. Sanket S. Pawar Abhijeet Manepatil Aniket Kadam Prajakta Jagtap[2], This research work is committed to keyword request and gives two perspectives of its application in IR and database framework. Article indicate model of Machine A and B, where A presents Innovative IR framework and B presents Discover approach social database administration framework. Article focuses more on extending keyword chase to database administration framework as it less tended to subject and all the more troublesome. Examination of Machine B show that execution appraisal need to address with fruitful evaluation like request workload memory use for versatile and adaptable propelled machine change. Rather than appraisal parameters like time deferral et cetera blend adaptable report recovery framework is build and assessed on memory use and request space is diminished altogether with two layer algorithm. Help degree of framework is making hybridization at machine level and working with pictures as information question. Qiuxiang Dong, Zhi Guan, ZhongChen[3] In this paper, they develop

new techniques that split the calculation for the keyword encryption and trapdoor/token period into two phases: a course of action stage that does by a wide margin a large portion of the work to scramble a keyword or make a token before it knows the keyword or the property list/get to control system that will be used. A minute stage then rapidly gathers a center figure content or trapdoor when the specifics get the chance to be unmistakably known. The preparation work can be performed while the mobile phone is associated with a power source, then it can later rapidly perform keyword encryption or token period operations moving without on a very basic level draining the battery. We name our arrangement Online/Offline ABKS. To the best of our understanding, this is the essential work on building profitable multi-customer searchable encryption contrive for mobile phones through moving a large portion of the cost of keyword encryption and token time into a disengaged organize. Dr Kehinde K. Agbele, Eniafe F. Ayetiran, Kehinde D. Aruleba and Daniel O. Ekong [4] proposed this article to make algorithms that enhance the situating of records recouped from IRS according to customer look for setting. In particular, the situating task that drove the customer to partake in information-pursuing behavior in the midst of request errands. This article inspects and depicts a Document Ranking Optimization (DROPT) algorithm for IR in an Internet-based or relegated databases condition. Then again, as the volume of information available on the web and in allocated databases is growing tenaciously, situating algorithms can expect an imperative part as to rundown things. In this article, a DROPT technique for chronicles recouped from a corpus is delivered with respect to report list keywords and the question vectors. This relies on upon figuring the weight ( ) of keywords in the report list vector, discovered as a segment of the recurrence of a keyword over a record. The inspiration driving the DROPT technique is to reflect how human customers can judge the setting changes in IR result rankings according to information importance. This article exhibits that it is workable for the DROPT technique to beat a segment of the restrictions of existing ordinary (  $\times$  ) algorithms by methods for alteration. The observational appraisal using estimations measures on the DROPT technique assisted through human customer collaboration demonstrates change over the ordinary significance input technique to indicate upgrading IR practicality. ShengliWu, Chunlan Huang, Jieyu Li [3] , recommended that for information recovery frameworks, the gathering of reports gets the opportunity to be unmistakably greater and greater. For some request, an information recovery framework needs to recoup innumerable as the result to the question. In fact, all the time people on a very basic level consider some top-situated files rather than the aggregate extensive rundown of reports. In such a condition, how to develop a recovery framework with appealing proficiency and reasonability is a research issue. In this paper, they focus on the information mix approach to manage information recovery, in which each fragment recovery framework contributes a result and each one of the results are merged by a mix procedure. The target of this research is to find an achievable blend technique that can

alter sufficiency and efficiency. Using 3 gatherings of chronicled continues running from TREC for the examination, they find that with the weights arranged by weighted straight backslide, the immediate mix methodology can finish awesome results in sufficiency and efficiency.

### III. PROBLEM DESCRIPTION

Base Paper Approach (Zhihua Xia, Member, IEEE, Xinhui Wang, Xingming Sun, and Qian Wang ,A Secure and Dynamic Multi-keyword Ranked Search Scheme over Encrypted Cloud Data) Multi-keyword Boolean search allows the customers to enter multiple inquiry keywords to request sensible records. Among these works, conjunctive keyword search plots simply give back the records that contain most of the question keywords. Disjunctive keyword search arranges give back most of the records that contain a subset of the question keywords. Situated search can enable quick search of the most noteworthy information. Sending back only the top-k most critical documents can satisfactorily lessen arrange movement.

Framework illustrate

The framework exhibits in this paper incorporates three one of a kind components: information proprietor, information customer and cloud server.

Information proprietor has an accumulation of records  $F = \{f_1, f_2, \dots, f_n\}$  that he needs to outsource to the cloud server fit as a fiddle while so far keeping the capacity to search on them for intense usage. In our arrangement, the information proprietor right off the bat manufactures a safe searchable tree document  $I$  from record gathering  $F$ , and after that creates a scrambled chronicle accumulation  $C$  for  $F$ . Subsequently, the information proprietor outsources the scrambled gathering  $C$  and the protected document  $I$  to the cloud server, and safely passes on the key information of trapdoor era (checking keyword IDF esteems) and chronicle unscrambling to the approved information customers. Moreover, the information proprietor is accountable for the refresh operation of his reports set away in the cloud server. While refreshing, the information proprietor creates the refresh information locally and sends it to the server. Information clients are approved ones to get to the reports of information proprietor. With  $t$  inquiry keywords, the approved customer can create a trapdoor  $TD$  as per search control instruments to get  $k$  scrambled records from cloud server. Then, the information customer can unscramble the records with the common puzzle key.

### IV. PROPOSED WORK

In our proposed approach we have make utilization of two algorithms,

Algorithm 1: For Keyword Search

Step1: Capture the Keyword String client entered for Searching

Step 2: Split the multi-keyword string into an exhibit. Presently every component of exhibit is the keyword to be searched.

Step 3: In the keyword search, we will keep up the accompanying information structures,

Structure 1 :

Filename

Transferred By

Keyword coordinated

Line Number

By making this structure we will get to the lines of the record containing the keyword.

In further we will adjust the idea of transferring the archive on the classification premise.

i.e. Structure for File Details

Filename

Transferred By

Date Time

Structure for Keywords

Classification Id

Classification Name

Keywords

At the point when the client transfers the document then on the premise of the classification an itemized record is put away in the accompanying table structure

Document Name

Keyword Matched

Line Number

This structure can contain multiple sections for an indistinguishable keyword from a similar keyword can show up in the different lines.

To accelerate the search we can utilize a cooperative memory structure

Filename

Keyword

Coordinate Times

Transferred By

To simply get the coordinated record will the keyword.

Algorithm 2: Secure Graphical OTP stick era

Step 1: Place the Images in the Grid first by tapping on the picture and after that on the position in the framework where we need to put the picture.

Step2 : After every one of the pictures are masterminded in the framework the code will examine the network beginning from the primary line and afterward handling to the last line and filtering every section in the column.

Step 3: If the section contains a picture then it will participate in making the stick and the idea included First letter of the picture taking after by line and the segment number and this procedure is rehashed for every one of the pictures in network.

Step4 : Then mail the produced stick to the client and client then reemerge the stick utilizing an indistinguishable procedure from specified in the progression.

### V. IMPLEMENTATION

We have reproduced the proposed Multi-Keyword Modified search algorithm and Visual Password Security algorithm utilizing Visual Studio 2013 and Microsoft SQL Server Express 2008. Microsoft Visual Studio is an incorporated advancement condition (IDE) from Microsoft. It is utilized to create PC programs for Microsoft Windows, and additionally sites, web applications and web administrations. Visual Studio utilizes Microsoft software improvement stages, for example, Windows API, Windows Forms, Windows

Presentation Foundation, Windows Store and Microsoft Silverlight. It can create both local code and oversee code.

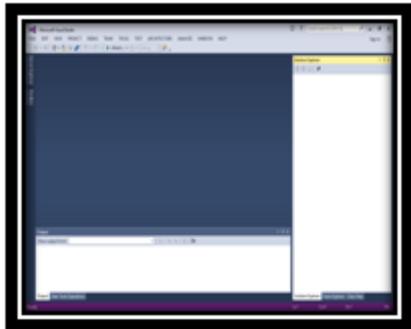


Fig 3 Visual Studio 2013 IDE



Fig 4 User Registration Form

Fig 4 demonstrates the Registration Form, in my proposed usage I have made a little web application for File sharing, for executing the Multi-Keyword Search and Visual Password. I have given access to just the enlisted clients to this framework. So this frame, will mimic the enlistment procedure for getting to the executed framework.

VI. RESULT

The executed framework on Visual Studio 2010 is keep running on different specimen information and the outcome is noted on the premise of the diagrams of time taken to play out the search. We have play out the trial on around 100 of test information. Furthermore, in the segment we have demonstrated a few aftereffects of the trial.

6.1 Sample Data Set 1:

In this trial we have given the information string "information" as the search string and the correlation in the Associative Based Search and Normal Search is show in the Fig 6



Fig 5 Result for sample string "Data"

The graph of the comparison is show in Fig 6.

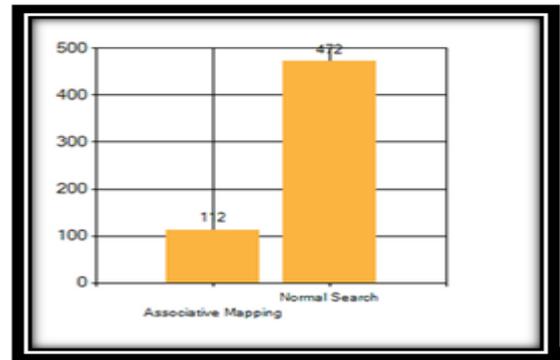


Fig 6 Graph for sample string "Data"

This comparison is also show in the table 1.

Associative Mapping	Normal Search
112 ms	472

Table 1 Table showing comparison for sample string "Data" In this test run we have provided the data string "string web" as the search string and the comparison in the Associative Based Search and Normal Search is show in the Fig 7



Fig 7 Result for sample string "Data"

The graph of the comparison is show in Fig 8

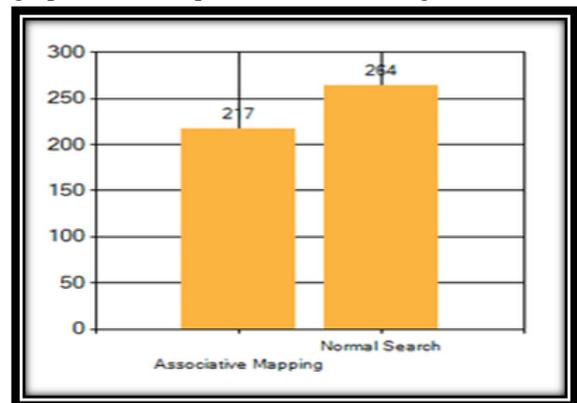


Fig 8 Graph for sample string "string web"

This comparison is also show in the table2.

Associative Mapping	Normal Search
217ms	264

## VII. CONCLUSION

In last we can conclude that the refined and the novel approach proposed by us, apply the concept of the green computing by reducing the CPU search by performing the fast and quick search by making use of the concept of the Associative Mapping.

## REFERENCES

- [1] A.Baid, I. Rae, J. Li, A. Doan, and J. Naughton "Toward Scalable Keyword Search over Relational Data," Proceedings of the VLDB En-dowment, vol. 3, no. 1, pp. 140–149, 2010.
- [2] Y. Chen, W. Wang, Z. Liu, and X. Lin, "Keyword Search on Structured and Semi-Structured Data," in Proceedings of the 35th SIGMOD International Conference on Management of Data, ser. SIGMOD '09, June 2009, pp. 1005–1010.
- [3] Vagelis Hristidis, Yannis Papakonstantinou "DISCOVER: Keyword Search In Relational Databases." In Proceedings of the 28th VLDB Conference, Hong Kong, China, 2002.
- [4] MyintMyintThein 1 and Mie Mie Su Thwin "Efficient Schema Based Keyword Search In Relational Database" in International Journal Of Computer Science, Engineering And Information Technology(IJCSEIT),Vol.2, No.6, December 2012.
- [5] Pradeep M. Ghige, Prof.Ruhi R. Kabra "Relational Keyword Search System" in International Journal of Engineering Research and General Science Volume 2, Issue 6, October-November, 2014.
- [6] Qi Zhou, Chong Wang, Miao Xiong, Haofen Wang and Yong Yu "SPARK: Adapting Keyword Query to Semantic Search" This work is funded by IBM China Research Lab.
- [7] Yi LuoXuemin Lin Wei Wang, Xiaofang Zhou "SPARK: Top-k Keyword Query in Relational Databases" in ACM June 11–14, 2007, Beijing, China.
- [8] Arvind Hulgeri, Gaurav Bhalotia Charuta NakheY, Soumen Chakrabarti "Keyword Search In Databases" in 2001 IEEE Computer Society Technical Committee on Data Engineering, Pune, India.
- [9] S. Dar, G. Entin, S. Geva, and E. Palmon. "DTL's DataSpot: Database exploration using plain language" In Proc. of the Int'l Conf. on VLDB, pages 645–649, 1998.
- [10] Roy Goldman, Narayanan ShivaKumar, Suresh Venkatasubramanian, Hector Gercia Molina "Proximity Search In Database" In Proceedings of the 24th VLDB Conference, New York, USA, 1998.
- [11] Gary Pan, SeowPoh Sun, Calvin Chan and Lim Chu Yeong,"Analytics and Cybersecurity: The shape of things to come",CPA ,2015
- [12] ErolGelenbe and Omer H. Abdelrahman,"Search in the Universe of Big Networks and Data",IEEE ,2014
- [13] ShengliWu,ChunlanHuang,JieyuLi,"Combining Retrieval Results for Balanced Effectiveness and Efficiency in the Big Data Search Environment",IEEE International Conference on Computer and Information Technology,2014
- [14] Ajeet Lakhani, Ashish Gupta, K. Chandrasekaran, "IntelliSearch: A Search Engine based on Big Data Analytics integrated with Crowdsourcing and category-based search", International Conference on Circuit, Power and Computing Technologies ,2015
- [15] Zhihua Xia, Member, Xinhui Wang, Xingming Sun, and QianWang,"A Secure and Dynamic Multi-keyword Ranked Search Scheme over Encrypted Cloud Data",IEEE,2015
- [16] Bing Wang, Wei Song, Wenjing Lou Y. ,Thomas Hou," Inverted Index Based Multi-Keyword Public-key Searchable Encryption with Strong Privacy Guarantee", IEEE Conference on Computer Communications (INFOCOM),2015
- [17] N. L. Sarda and A. Jain. Mragyati: A system for keyword-based searching in databases. <http://arxiv.org/abs/cs.DB/0110052>.