BANK SECURE FILE SEARCHING AND SHARING USING GRAPHICAL OTP AND SHA

Dharmendra Singh[1], Pooja[2], Dhawal Vyas[3]
[1]M.Tech Scholar, [2,3]Assistant Professor, Department Of Computer Science & Engineering,
Chandravati Education Charitable Trust Group of Institution Bharatpur.

Abstract: ***In the today's world of the information technology, the data is growing very rapidly. In any organization, whether any corporate or banks thousands of documents are required to be stored or exchanged. In such an environment, two main things which are of upmost requirement, fast search and secure access. In our dissertation, we focused on these two aspects. For the faster search we have processed the novel data structure , which is similar to the concept of the associative memory, to link the documents which are searched frequently with the associated keywords, when the user search the keyword the best matched documents are searched directly from the associate memory structure, leading to the faster access. And regarding the secure access, we have classified the documents are the secured and highly secured.***
For the secured, we have implemented the grid based organization of images which results in the password pattern used for sharing of the files and for the highly secured files apart from the grid based password, the SHA code for the file shared is also used as the transaction key to further increase the security.
In this way, the dissertation helps the organization like banks to share and search the documents more effectively and securely.
Keywords: Secure Search, Associate Memory Structure, Grid Password, SHA.

I. INTRODUCTION

Cryptography, a word with Greek starting implies "release making," cryptography is the preparation and examination of strategy for secure correspondence in the closeness of data correspondence with security so dark individual neither access nor change any data. In the sharing e of the documents , for example institutions like Banks we have to secure these documents by some sort of the encryption techniques , so that the intruders will not able to hack the important information related to the banking transactions.

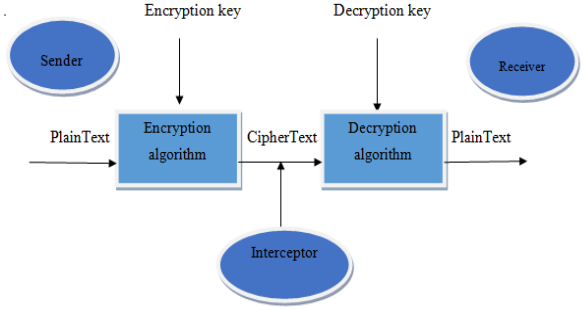

Fig 1. Cryptography

Encryption and decryption change over the principal message into proper game plan and sends the message over an unverifiable channel. All systems are being presented, interconnected to the overall system. The data is text, and sound picture are the main components of the message to be send. The modernized pictures are typically used are addressed in the 2-D bunch [1].
To secure our data during the season of transmission cryptography answers. The term cryptography got from a Greek word called "Kryptos" which signifies "Concealed Secrets."Cryptography can be characterized as the craft of defending archives and it ensures that exclusive the planned individuals can examine its substance. It is the Art of Science of changing over a plain clear data and again retransforming that message into its interesting shape. The five standard destinations behind using Cryptography join Confidentiality, Authentication, Integrity, Non-Repudiation, Service Reliability and Availability.[1]These goals ensure that the private data remains private, the data is not changed unlawfully and assurances against a social occasion denying a data or a correspondence that was begun by them.
There are on a very basic level two sorts of cryptography
(i) Symmetric or Secret Key Cryptography
(ii) Asymmetric or Public Key Cryptography
In Symmetric key cryptography, both the sender and recipient know a comparable riddle code called key. Messages are mixed by the sender using the key,and the recipient unscrambles it using a comparative key. E.g., Data encryption standard (DES), Triple DES, Advanced Encryption Standard (AES) and Blowfish Encryption Algorithm [1].

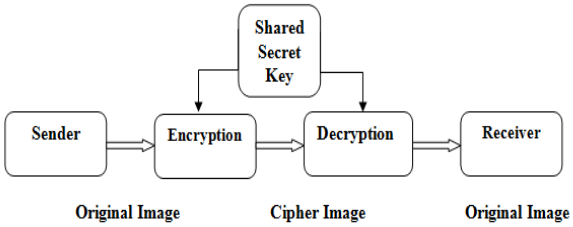

Fig 2. Symmetric Key Cryptography

In Asymmetric key cryptography, sender and recipient utilizes differing key for encryption and decryption. The sender scrambles the data using an open key and this key will be known by each one of the social affairs consolidated into the correspondence. The beneficiary unscrambles the data using a private key and it should be kept as a secret e.g.RSA.

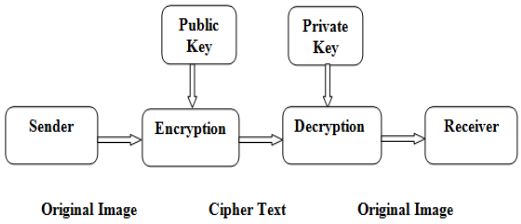

Fig 3.Asymmetric Key Cryptography

SHA-1

In cryptography, SHA-1 (Secure Hash Algorithm 1) is a cryptographic hash work arranged by the United States National Security Agency and is a U.S. Government Information Processing Standard distributed by the United States NIST. SHA-1 creates a 160-piece (20-byte) hash esteem known as a message process. SHA-1 hash esteem is typically rendered as a hexadecimal number, 40 digits long.

SHA-1 creates a message process in view of standards like those utilized by Ronald L. Rivest of MIT in the outline of the MD4 and MD5 message process algorithms, yet has a more traditionalist plan.

SHA-1 was created as a feature of the U.S. Government's Capstone project. The unique determination of the algorithm was distributed in 1993 under the title Secure Hash Standard, FIPS PUB 180, by U.S. government gauges organization NIST (National Institute of Standards and Technology).

This rendition is currently frequently named SHA-0. It was pulled back by the NSA soon after production and was superseded by the overhauled variant, distributed in 1995 in FIPS PUB 180-1 and regularly assigned SHA-1. SHA-1 contrasts from SHA-0 just by a solitary bitwise turn in the message calendar of its pressure work. As indicated by the NSA, this was done to redress a defect in the first algorithm which lessened its cryptographic security, yet they didn't give any further clarification. Openly accessible strategies did in reality trade-off SHA-0 preceding SHA-1[5].

II. RELATED STUDY

Yue Lu, Chew Lim Tan [1] recommended that an enormous measure of document images are available in the Internet and computerized libraries. They locate that, the greater part of them are pressed in PDF documents and are packed utilizing CCITT Group 4 principles for sparing storage room and accelerating transmission. There is therefore noteworthy importance to build up the strategies for straightforwardly looking catchphrases from these documents. In this paper, they show a compacted design coordinating strategy for seeking catchphrases from the CCITT Group 4 packed document images, without unequivocal decompression.

As per the CCITT Group 4 principles, each coded position shows that the present pixel shading is unique in relation to 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 compacted images straightforwardly. The changing components are used to section and bound the word questions, and are utilized for estimating the likeness of two word images. The associated segments are named 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 exploration work is given to catchphrase request and gives two perspectives of its application in IR and database framework. Article indicate model of Machine An and B, where A presents Innovative IR framework and B presents Discover approach social database administration framework. Article focuses more on extending catchphrase chase to database administration framework as it less tended to subject and all the more troublesome. Examination of Machine B exhibits that execution evaluation needs to address with effective appraisal like request workload memory use for versatile and adaptable propelled machine change.

Rather than appraisal parameters like time deferral et cetera blend flexible report recovery framework is build and assessed on memory usage and request space is diminished fundamentally with two layer calculation. Help degree of framework is making hybridization at machine level and working with pictures as data question.

Qiuxiang Dong, Zhi Guan, Zhong Chen[3] In this paper, they develop new systems that split the calculation for the catchphrase encryption and trapdoor/token time into two phases: a game plan organize that does by a wide margin the vast majority of the work to encode a watchword or make a token before it knows the catchphrase or the property list/get to control technique that will be used. A minute stage at that point rapidly gathers a center Fig. content or trapdoor when the specifics get the chance to be particularly known. The availability work can be performed while the wireless is associated with a power source, at that point it can later rapidly perform catchphrase encryption or token time activities moving without generally draining the battery. We name our arrangement Online/Offline ABKS. To the best of our understanding, this is the essential work on building beneficial multi-customer accessible encryption contrive for PDAs through moving the vast majority of the cost of catchphrase 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 calculations that enhance the situating of records recuperated from IRS according to customer look for setting. In particular, the situating task that drove the customer to partake in data pursuing behavior in the midst of request errands. This article looks at and depicts a Document Ranking Optimization (DROPT) calculation for IR in an Internet-based or appointed databases condition. Then again, as the volume of data open on the web and in allocated databases is growing diligently, situating calculations can expect a vital part with respect to list things. In this article, a DROPT system for files recuperated from a corpus is

delivered with respect to report list catchphrases and the inquiry vectors. This relies upon figuring the heaviness of watchwords in the report list vector, found out as a part of the recurrence of a catchphrase over a record. The inspiration driving the DROPT method is to reflect how human customers can judge the setting changes in IR result rankings according to data centrality. This article exhibits that it is workable for the DROPT strategy to beat a part of the restrictions of existing traditional calculations by methods for modification. The observational evaluation using estimations measures on the DROPT method assisted through human customer collaboration demonstrates change over the customary significance input strategy to indicate upgrading IR reasonability.

Sanjay Agrawal , Surajit Chaudhuri,Gautam Das[5],In this paper, they look at DBXplorer, a framework that engages watchword based interest in social databases. DBXplorer has been executed using a business social database and web server and grants customers to participate through a program front-end. They format the challenges and discuss the utilization of our framework including eventual outcomes of wide trial evaluation.

Anjaneyulu Karapakula, M. Puramchand and G. Mohammad Rafi[6], In this paper, strikingly, they portray and deal with the testing issue of assurance sparing multi-watchword situated look for over scrambled cloud data (MRSE), and set up a course of action of strict insurance requirements for such a safe cloud data usage framework to twist up unmistakably a reality. Among various multi-watchword semantics, they pick the capable manage of "encourage planning", i.e., an indistinguishable number of matches from conceivable, to get the likeness between request question and data reports, and further use "inner thing closeness" to quantitatively formalize such rule for similarity estimation. We initially propose a major MRSE contrive using secure inner thing calculation, and a short time later out and out improve it to meet particular assurance requirements in two levels of hazard models. Watchful examination investigating security and efficiency confirmations of proposed plans is given, and trials on this present reality dataset also demonstrate proposed schemes no ifs ands or buts introduce low overhead on calculation and correspondence.

III. PROBLEM DESCRIPTION

Right off the bat, every one of the clients as a rule keeps the same secure key for trapdoor age in a symmetric SE plot. For this situation, the renouncement of the customer is tremendous test. If it is relied upon to deny a customer in this arrangement, we need to alter the record and scatter the new secure keys to all the approved clients.

Additionally, symmetric encryption SE schemes more often than not expect that every one of the data clients are tried and true. It isn't down to earth and an exploitative data customer will prompt numerous safe issues. For instance, a complex data customer may search the reports and pass on the unscrambled documents to the unapproved ones. Significantly more, a deceitful data customer may scatter his/her safe keys to the unapproved ones. Later on works, we will endeavor to upgrade the SE intend to handle these test issues.

Additional Solution

1.Fast Multi-keyword search algorithm is proposed to actualize the search using the Associative Mapping so will take lesser time as compare to the normal search.

2.For the secure key we have contrived the novel approach of the password generation or the key generation of the access of the records shared,

3.Use the Matrix password as well as SHA Code for the secure file access.

We take a matrix of the 6x6 in which we will place the 6 pictures at any of the random locations.

IV. PROPOSED APPROACH

4.1 Algorithm 1: For Keyword Search

Step 1: Capture the Keyword String user entered for Searching

Step 2: Split the multi-keyword string into an array. Now each element of array is the keyword to be searched.

Step 3: In the keyword search, we will maintain the following data structures,

Structure 1 :
- Filename
- Uploaded By
- Keyword matched
- Line Number

By making this structure we will get access the lines of the file containing the keyword.

In further we will modify the concept of uploading the document on the category basis.

Structure for File Details
- Filename
- Uploaded By
- Date Time

Structure for Keywords
- CategoryId
- Category Name
- Keywords

When the user uploads the file then on the basis of the category a detailed record is stored in the following table structure
- FileName
- Keyword Matched
- Line Number

This structure can contain multiple entries for the same keyword as the same keyword can appear in the various lines.

In order to speed up the search we can use an associative memory structure
- Filename
- Keyword
- MatchTimes
- Uploaded By

In this it will return the documents which contain the matched keyword.

4.2 Algorithm 2: Secure Graphical OTP pin generation

Step 1: Place the Images in the Grid first by clicking on the

image and then on the position in the grid where we want to place the image.
Step 2 : After all the images are arranged in the grid the code will scan the grid starting from the first row and then processing to the last row and scanning each column in the row.
Step 3: If the column contains an image then it will participate in creating the pin and the concept involved First letter of the image following by row and the column number and this process is repeated for all the images in grid.
Step 4: Then mail the generated pin to the user and user then renter the pin using the same process as mentioned in the step 1.

4.3 Algorithm 3: Highly Secure Files

Step 1: Place the Images in the Grid first by clicking on the image and then on the position in the grid where we want to place the image.
Step2 : After all the images are arranged in the grid the code will scan the grid starting from the first row and then processing to the last row and scanning each column in the row.
Step 3: If the column contains an image then it will participate in creating the pin and the concept involved First letter of the image following by row and the column number and this process is repeated for all the images in grid.
Step 4: Files are Encrypted and they are not directly readable , the SHA code of the file will act as the key and then at the receiver end the file is required to be decrypted using the SHA code which is also required to be entered by the user. As it is the 40 characters hexadecimal code so the 8 characters each selected after 5th position will be used to simply the password.
Step 5: Both the SHA based code and the graphical code is shared.

V. TEST RESULTS

The implemented system on Visual Studio 2013 is run on various sample data and the result is noted on the basis of the graphs of time taken to perform the search. We have perform the test run on around 100 of sample data. And in the section we have shown some results of the test run.

Sample Data Set

In this test run we have provided the data string “data” as the search string and the comparison in the Associative Based Search and Normal Search is show in the Fig 4.

Fig. 4. Result for sample string “Data”

The graph of the comparison is show in Fig 4.

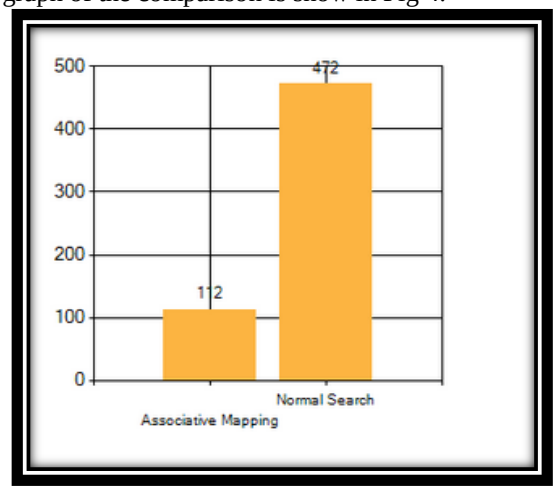

Fig. 5 Graph for sample string “Data”

This comparison is also show in the table 1.

Table 6.1 Table Showing Comparisons For Sample String “Data”

Associative Mapping	Normal Search
112 ms	472

The ms stand for Milliseconds.

VI. CONCLUSION

Security is the main concern in the transaction, the proposed dissertation has increased the security first by offering the concept of the gird based in which images are organised and the password is formed for the sharing purpose and for the high secure files the SHA algorithm is used to generate a unique code which will act as the transaction id. Thus the security and speed both have enhanced.
Thus, we can say that our proposed implementation provides a better way to share the data securely.
Together with the green concept in the dissertation, our work also focus on the security, by making the use of the visual password, we have enhance the security, thus overcoming the probability of the easily cracked passwords.
Future Work : In the future work we will try to extend our research towards the search of the images and other complex data and in segment of security we will try to extend our work with the DNA Passwords, ECG cryptography and similar concepts.

REFERENCES

[1] Lu, Yue & Tan, Chew Lim.,” Keyword searching in compressed document images”. DCC,2003..
[2] S. S. Pawar, A. Manepatil, A. Kadam and P. Jagtap, "Keyword search in information retrieval and relational database system: Two class view," International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), 2016.
[3] Q. Dong, Z. Guan and Z. Chen, "Attribute-Based

Keyword Search Efficiency Enhancement via an Online/Offline Approach," IEEE 21st International Conference on Parallel and Distributed Systems (ICPADS), 2015.

[4] Kehinde K. Agbele, Kehinde Daniel Aruleba, Eniafe F. Ayetiran,"Efficient schema based keyword search in relational databases." University of Computer Studies, Mandalay, Myanmar, International Journal of Computer Science, Engineering and Information Technology (IJCSEIT) 2.6 (2012).

[5] Sanjay Agrawal, Surajit Chaudhuri, Gautam Das,"DBXplorer: enabling keyword search over relational databases",SIGMOD,2002.

[6] A. Karapakula, M. Puramchand and G. M. Rafi, "Coordinate matching for effective capturing the similarity between query keywords and outsourced documents," IET Chennai 3rd International on Sustainable Energy and Intelligent Systems (SEISCON 2012), Tiruchengode, 2012.

[7] W. Tang, L. Yan, Z. Yang and Q. H. Wu, "Improved document ranking in ontology-based document search engine using evidential reasoning," in IET Software, vol. 8, no. 1, pp. 33-41, February 2014.

[8] Shengli Wu, Jieyu Li, "Merging Results from Overlapping Databases in Distributed Information Retrieval",PDP,2013.

[9] A. Lakhani, A. Gupta and K. Chandrasekaran, "IntelliSearch: A search engine based on Big Data analytics integrated with crowdsourcing and category-based search", International Conference on Circuits, Power and Computing Technologies , 2015.

[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 Network, 28.4(2014): 20-25.

[13] ShengliWu,ChunlanHuang,JieyuLi,"Combining Retrieval Results for Balanced Effectiveness and Efficiency in the Big Data Search Environment",Computer and Information Technology (CIT), 2014 IEEE International Conference on. (pp. 555-560) IEEE, 2014.

[14] Ajeet Lakhani, Ashish Gupta, K. Chandrasekaran, "IntelliSearch: A Search Engine based on Big Data Analytics integrated with Crowd sourcing and category-based search", International Conference on Circuit, Power and Computing Technologies (ICCPCT),2015. (pp. 1-6).

[15] Zhihua Xia, Xinhui Wang, Xingming Sun, and QianWang,"A Secure and Dynamic Multi-keyword Ranked Search Scheme over Encrypted Cloud Data." IEEE Transactions on Parallel and Distributed Systems 27.2 (2016): 340-352.

[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 Ankur Jain. "A system for keyword-based searching in databases."Report No. cs. DB/011052 on CORR (http://xxx.lanl.gov/archive/cs) (2001).

[18] Sarita Kumari,"A Research Paper on Cryptography Encryption and Compression Techniques," International Journal Of Engineering And Computer Science ISSN:2319-7242Volume 6 Issue 4 April 2017.

[19] Gopal D. Dalvi, Dr.D.G.Wakde, "Facial Images Authentication In Visual Cryptography Using Sterilization Algorithm," 2nd International Conference for Convergence in Technology (I2CT),2017.

[20] Ekta Agrawal,Dr. Parashu Ram Pal,A New and More Authentic Cryptographic Based Approach for Securing Short Message, International Journal of Advanced Research in Computer Science,2017

[21] Arpit Agrawal, Gunjan Patankar,"Design of Hybrid Cryptography Algorithm for Secure Communication," International Research Journal of Engineering and Technology (IRJET),2016

[22] Rahman MM, Akter T, Rahman A,"Development of Cryptography-Based Secure Messaging System," J Telecommun Syst Manage,2016.

[23] Disha Shah,"Digital Security Using Cryptographic Message-Digest Algorithm," International Journal of Advance Research in Computer Science and Management Studies,Volume 3, Issue 10, October 2015.

[24] Saikumar Manku And K. Vasanth,Blowfish Encryption Algorithm For Information Security,ArpnJournal of Engineering And Applied Sciences,2015.

[25] Snehal Javheri,Rahul Kulkarni,"Secure Data communication and Cryptography based on DNA based Message Encoding," International Journal of Computer Applications (0975 – 8887)Volume 98– No.16, July 2014.

[26] Md Asif Mushtaque, Harsh Dhiman, Shahnawaz Hussain,"A Hybrid Approach and Implementation of a New Encryption Algorithm for Data Security in Cloud Computing," International Journal of Electronic and Electrical Engineering,2014

[27] Maulik P. Chaudhari and Sanjay R. Patel,"International Journal of Advanced Research in Computer Science and Management Studies," International Journal of Advanced Research in Computer Science and Management Studies,2014.

[28] Pia Singh Prof. Karamjeet Singh,"Image encryption and decryption using Blowfish algorithm in Matlab," International Journal of Scientific & Engineering Research, 2013.

[29] Himanshu Gupta and Vinod Kumar Sharma,"Multiphase Encryption: A New Concept in Modern Cryptography",International Journal of Computer Theory and Engineering, Vol. 5, No. 4, August 2013

[30] Obaida Mohammad Awad Al-Hazaimeh,"A New Approach for Complex Encrypting and Decrypting Data",International Journal of Computer Networks & Communications (IJCNC) Vol.5, No.2, March 2013

