Novel Approach for Fast Search and Random Grid Password for File Sharing

Deepshikha Bhati[1], Sonal Sharma[2]

[1]*Mtech. Scholar, Computer Science, Rajasthan college of engineering for women, Jaipur Rajasthan*

[2]*Assistant professor, Computer Science, Rajasthan college of engineering for women, Jaipur Rajasthan*

Abstract- **In dissertation , the new approach is proposed for the secure document search , in this an associative technique for keyword search is proposed with the new style of password applying using the random number blocks containing check mark to be considered for password formation. This type of concept will further increased up the security.**

***Index Terms*- Keyword Search , Grid Password , Random Password.**

I. INTRODUCTION

Data recuperation is the route toward get-together data by using catchphrases from the appropriate record and that report can be unstructured or organized data. It covers its versatile quality from customer by giving reasonable view. As customer don't have any data [1] about example and some other inquiry taking care of tongue, he can seek through unique interface by putting catchphrases. By using Keyword Search customer can submit catchphrase to look engines (Internet Search) or organized data and therefore it reestablishes an once-over of records to customer according to situating. Situating of reports are given in perspective of the watchwords match and occasion of catchphrase facilitate particularly record. Situating is given in plunging solicitation of occasion of watchword arrange and the file with most outrageous occasion get higher need[1]

Information based techniques are the most broadly utilized authentication techniques and incorporate both content based and picture-based passwords. The photo based techniques can be additionally separated into two classes: recognition - based and recall-based graphical techniques. Utilizing recognition-based techniques, a client is given an arrangement of pictures and the client passes the authentication by perceiving and distinguishing the pictures he or she chose amid the enrollment organize. Utilizing recall-based techniques, a client is requested to recreate something that he or she made or chose before amid the enrollment arrange.

As we probably am aware graphical pictures are all the more effectively recalled then content. In this segment, graphical password framework based on recognition and recall based are talked about as beneath:-

Recognition-Based Technique: In this kind of technique, clients will choose pictures, logos or any images from prestored picture. For authentication process client need to perceive the picture, which he pick as a password.

Recall-Based Technique: Again recall-based password authentication are arrange in two sections [2]:

(I) Pure Recall Based Technique (ii) Cued Recall Based Technique

Recognition based technique require the client to distinguish and perceive the mystery, or part of it, that the client chose previously. By and large amid password creation the clients are required to remember a progression of pictures, and afterward should perceive their pictures from among baits to sign in. Phishing assaults are to some degree more troublesome with recognition-based frameworks as a right arrangement of pictures must be exhibited to the client before password passage. Shoulder-surfing is by all accounts of specific worry in recognition-based frameworks when an aggressor is remaining behind the client and sees or watches the pictures chose by clients amid login [3][4]. Different recognition based password blueprint are clarified beneath:

(a) Passfaces: The recognition-based framework considered most widely to date is Passfaces. By and large amid setting a password the client chooses an arrangement of human appearances. A board of competitor faces is displayed amid his/her login.

Among the given arrangement of fakes the client must choose the faces he/she chose amid setting the password. Pass faces just works by having the client select a subgroup of x faces from a gathering of k faces. For authentication, the framework indicates p countenances and one of the faces has a place with the subgroup q. The client needs to do the determination ordinarily to finish the authentication procedure [5].

(b) Story: The Story plot, which requires the determination of pictures of items (individuals, autos, nourishments, planes, touring, and so on.) to frame a story line.

Cued-recall based password history is for the most part commanded by passpoints. In passpoints the client needs to tap on the five unique positions or territories of a similar picture. Thus it is clicked based graphical password. The snap is mouse based and client must recollect the right arrangement or arrangement of snap focuses on that foreordained picture for the following effective login. It is a tick based plan where clients select a single tick point on every one of 5 pictures in succession, each one in turn; this gives one-to - one signaling. Amid the following login the client must recollect that specific snap point on the offered picture to open the following right picture, if the snap isn't right the following opened picture will be a phony one and not from the picked arrangement of pictures. This will stop current client authentication [7].

The plan expectation of the randomized viewport positions is to smooth the appropriation of snap focuses over different clients, to diminish the impacts of hotspots[8].Two authentication techniques are based on content and hues proposed for PDA in this they create the session passwords and impervious to word reference assault. Once the session is ended, the session password is never again valuable. For each login procedure, clients input distinctive passwords. To evacuate the disadvantage of literary password expelled by graphical password plans which give a method for making more easy to understand passwords, while expanding the level of security, they are defenseless against bear surfing .Here content was consolidate with picture and shading to create the session password and each time client needs to enter new password as session closes..

2. CRYPTOGRAPHY

Cryptography, a word with Greek beginning infers "discharge making," cryptography is the planning and examination of procedure for secure correspondence in the closeness of data correspondence with security so dull individual neither access nor change any data [1].

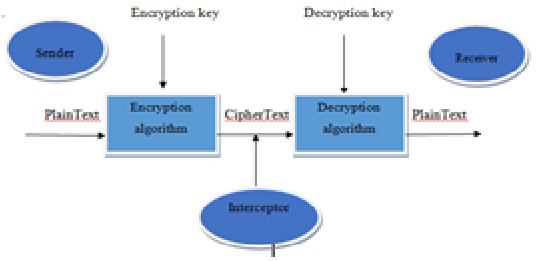

Fig 1. Cryptography

Encryption and decryption change over the principal message into proper game plan and sends the message over an unverifiable channel. All frameworks are being exhibited, interconnected to the overall framework. The data is content, and sound picture are the fundamental segments of the message to be send. The modernized pictures are commonly utilized are tended to in the 2-D bundle [1].

To secure our data amid the period of transmission cryptography answers. The term cryptography got from a Greek word called "Kryptos" which implies "Hid Secrets." Cryptography can be described as the specialty of protecting chronicles and it guarantees that selective the planned people can look at its substance. It is the Art of Science of changing over a plain clear data and again retransforming that message into its intriguing shape. The five standard goals behind utilizing Cryptography join Confidentiality, Authentication, Integrity, Non-Repudiation, Service Reliability and Availability.[1]These objectives guarantee that the private data stays private, the data isn't changed unlawfully and affirmations against a social event denying a data or a correspondence that was started by them.

3. LITERATURE SURVEY

Wenhu Tang, Long Yan, Zhen Yang, Qinghua Henry Wu[7] look at presents as a novel technique to oversee document arranging in a theory based record web list (ODSE) utilizing evidential thinking (ER).

At first, a space introspective philosophy appear, utilized for ask for extension, and a connection interface to an ODSE are made. A different quality basic activity (MADM) tree demonstrate is proposed to sort out widened question terms. The outcomes demonstrate that the proposed approach gives a reasonable reaction for report arranging and the exactness at a near overview levels for ODSE searches for have been updated essentially with ER presented, in examination with a standard catchphrase sorting out web crawler, an ODSE without ER and a non-haphazardness based weighting model.

Shengli Wu, Chunlan Huang, Jieyu Li [8], suggested that for data recovery structures, the social event of reports persuades the chance to be especially more prominent and more noteworthy. For some demand, a data recovery system needs to recover boundless as the outcome to the request. In actuality, all the time individuals on an exceptionally essential level think of some as best arranged records rather than the total huge summary of reports. In such a condition, how to build up a recovery structure with engaging capacity and appropriateness is an exploration issue. In this paper, they center around the data blend way to deal with oversee data recovery, in which each piece recovery structure contributes an outcome and every last one of the outcomes are converged by a blend framework. The objective of this examination is to locate a doable mix framework that can adjust sufficiency and proficiency. Utilizing 3 get-togethers of chronicled keeps running from TREC for the examination, they find that with the weights organized by weighted straight break faith, the quick blend system can accomplish wonderful outcomes in plentifulness and gainfulness.

Ajeet Lakhani, Ashish Gupta, K. Chandrasekaran [9], proposed that Big Data is the advancement that has changed the world. Its practical capacities to set up the data and make the basic data out of that has attempted the expression learning can be basically controlled by the general population who have souls. In spite of the way that the colossal data examination is to a great degree convincing in mining the inspiration out of the monstrous pool of data yet Big Data alone is lacking if there should rise an occasion of web-examination. The present web crawlers don't have the portrayal based demand highlight and a standard arranging system for site page. Crowdsourcing by including dynamic web based gathering on web is changing the state of web applications by refining the data and enhancing the suggestion for the things. The paper investigates about the web look devices, web-examination toolbars and swarm sourcing to enhance the web-examination. The paper in like way talks about how breaker of gathering sourcing with Big Data Analytics can accomplish 'IntelliSearch', a strong and solid web searcher.

Zhihua Xia, Member, IEEE, Xinhui Wang, Xingming Sun, and Qian Wang [15], demonstrate a safe multi-watchword arranged look think up over encoded cloud data, which in the meantime underpins dynamic overhaul exercises like cancelation and thought of records. In particular, the vector space appears and the overall utilized TF IDF demonstrate is joined in the record change and question time. They fabricate a momentous tree-based record structure and propose a "Greedy Depth-first Search" computation to give fruitful multi-watchword arranged look. The secured KNN figuring is used to encode the summary and demand vectors, and a short time later guarantee revise importance score computation between mixed archive and question vectors. With a specific extreme goal to negate quantifiable ambushes, ghost terms are added to the record vector for blinding summary things. Because of the utilization of their unprecedented tree-based report structure, the proposed plan can accomplish sub-arrange pursue time and manage the cancelation and thought of records adaptably. Far reaching examinations are composed to show the reasonability of the proposed plan.

4. PROPOSED WORK

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: Take a Range of Values.

Step 2: Take a grid of 8X8 grid elements.

Step 3: Generate random number and provide the checkbox in front of each block.

Step 4: User check the checkboxes and click on generate button to generate a new password.

The implementation is done in VS 2013 ,

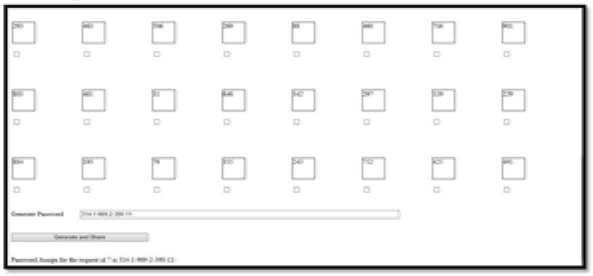

Fig 2. Implementation of Secure Password

5. TEST ANALYSIS

In this test run we have provided the data string "identifier device path" as the search string and the comparison in the Associative Based Search and Normal Search is show in the Fig 6.5.

Fig. 3 Result for sample string "identifier device path"

The graph of the comparison is show in Fig 6.6.

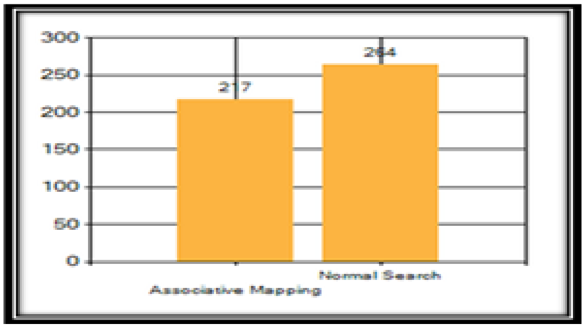

Fig. 4 Graph for sample string "identifier device path"

This comparison is also show in the table 6.3.

TABLE 1 TABLE SHOWING COMPARISON FOR SAMPLE STRING "IDENTIFIER DEVICE PATH"

Associative Mapping	Normal Search
217 ms	264

The ms stand for Milliseconds.

Table 2, which shows the result of the Key strength using the three tools which we have taken for the testing purpose.

TABLE 2 TEST RESULT ANALYSIS TABLE

Test Key	Website/Tool	Result
955-1-381-11-275-18-575-19-	Password Meter	Very Strong
955-1-381-11-275-18-575-19-	Password Checker	Excellent Strength
955-1-381-11-275-18-575-19-	Cryptool2	Entropy 3.24 Strength 102 Very Strong

6. CONCLUSION

Thus in the proposed work, the security is further increased and the searching time is also reduced to the considerable extent. Thus, this work moved a step in the security.

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,SurajitChaudhuri,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] ShengliWu,JieyuLi,"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] AjeetLakhani,AshishGupta,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(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.