DOI: https://dx.doi.org/10.26808/rs.ca.i7v5.01
International Journal of Computer Application (2250-1797)
Volume 7– No.5, September-October 2017

Novel Algorithm for File Sharing using Grid Password and Keyword Search

Ishita Mathur[1], Ram Lal Yadav[2]
1 M.Tech. Scholar, Department Of CSE, Kautilya Institute of Technology & Engineering,Jaipur ,Rajasthan.
2 M.Tech Coordinator, Department Of CSE, Kautilya Institute of Technology & Engineering,Jaipur ,Rajasthan.

ABSTRACT

Accidents especially to look at the railroad catastrophes happen dependably in India yet then can be diminished on the off chance that they examine the clarification behind misfortunes. In our investigation we are thinking about a couple of parts like Road which are associated with the Junctions where the accidents happened, environment in which the setback happened, day time or evening time when the episode happens and more factors. Additionally, consequently they attempt to discover the mixes utilizing the Modified Checkpoint based Apriori Algorithm the mixes of the factors which makes the most mishaps and attempt think of them as.

Key words: Apriori Algorithm, Data Mining, Accident Analysis..

Corresponding Author: Ishita Mathur ,M.Tech Scholar

INTRODUCTION

Big Data security and privacy includes Big Data management and investigation for cyber security. While Big Data has establishes in numerous technologies, database management is at its heart. Consequently in this area we will examine how data management has advanced and will then focus on the Big Data security and privacy issues.[1]
Database systems technology has propelled an extraordinary arrangement during the previous four decades from the legacy systems in light of network and various leveled models to relational and question database systems. Database systems can likewise now be gotten to by means of the web and data management administrations have been executed as web administrations. Because of the blast of electronic administrations, unstructured data management and web-based social networking and versatile registering, the measure of data to be taken care of has expanded from terabytes to petabytes and zetabytes in only two decades. Such immeasurable measures of complex data have come to be known as Big Data. Not exclusively does big data need to be overseen effectively, such data additionally must be analyzed to remove helpful chunks to upgrade businesses and enhance society. This has come to be known as Big Data Analytics [1].
Capacity, management and examination of huge amounts of data likewise result in security and privacy violations. Frequently data must be held for different reasons

including for administrative consistence. The data held may have touchy information and could disregard client privacy. Moreover, controlling such big data, for example, consolidating sets of various sorts of data could bring about security and privacy violations. For ex-abundant, while the crude data evacuates by and by identifiable information, the determined data may contain private and touchy information. For instance, the crude data about a man might be consolidated with the per-child's address which might be adequate to distinguish the individual.

Distinctive communities are taking a shot at the Big Data challenge. For instance, the systems community is developing technologies for enormous stockpiling of big data. The network community is developing solutions for overseeing huge networked data. The data community is developing solutions for effectively man-maturing and examining extensive arrangements of data. Big Data research and development is being completed both in the scholarly world, industry and government research labs. Nonetheless, little consideration has been given to security and privacy contemplations for Big Data. Security cuts over various zones including systems, data and net-works. We require the numerous communities to meet up to develop solutions for Big Data security and privacy.

1.2 Big data management and analytics

Big Data management and analytics [1] research is continuing in three headings. They are:

(i) Building foundation and superior registering strategies for the capacity of big data;

(ii) Data management procedures, for example, incorporating numerous data sources (both big and little) and ordering and questioning big data;

(iii) Data analytics procedures that control and investigate big data to concentrate pieces.

In rundown, Big Data management and analytics systems incorporate expanding current data management and mining strategies to deal with gigantic measures of data and in addition developing new methodologies including diagram data management and digging procedures for keeping up and breaking down vast networked data.

1.3 Security and privacy

The collection, storage, manipulation and retention of massive amounts of data have resulted in serious security and privacy considerations. Different directions are being proposed to deal with Big Data so that the privacy of the people is not disregarded. For instance, regardless of the possibility that by and by identifiable information is expelled from the data, when data is consolidated with other data, an individual can be distinguished. This is basically the deduction and collection issue [4] that data security researchers have been investigating for as far back as four decades. This issue is exacerbated with the management of Big Data as various wellsprings of data now exist that are identified with different people.

Now and again, directions may make privacy be abused. For instance, data that is gathered (e.g., email data) must be held for a specific timeframe (more often than not 5 years). For whatever length of time that one keeps such data, there is a potential for privacy violations. An excessive number of directions can likewise smother development. For ex-adequate, if there is a control that crude data must be kept as is and not controlled or models can't be worked out of the data, then enterprises can't dissect the data in imaginative approaches to improve their business. Along these lines development might be smothered.

Hence, one of the primary difficulties for guaranteeing security and privacy when managing big data is to concoct an adjusted approach towards directions and analytics. That is, by what means can an association complete helpful analytics and still guarantee

the privacy of people? Various methods for privacy-protecting data mining, privacy-safeguarding data joining and privacy-saving [5] information recovery have been developed. The test is to amplify these methods for taking care of huge measures of regularly networked data.

Another security challenge for Big Data management and analytics is to secure the frameworks. A large number of the technologies that have been developed including Hadoop, MapReduce, Hive, Cassandra, PigLatin, Mahout and Storm don't have satisfactory security assurances. The question is, in what capacity can these technologies be secured and in the meantime guarantee elite processing?

1.4 Keyword Query System Model

This area describes, in general terms, the structure of a keyword query system[7]. A keyword query system is an unpredictable system which is equipped for taking as input, an arrangement of words, called keywords and give a proper answer. Here, the structure and semantics of the appropriate response is particular to the query answer system. A system for keyword query comprises of the accompanying:

(i) Data Model: It describes the high-level representation of the data in the system, with the end goal that it mirrors the imperatives, affiliations, and association of the data. The genuine execution of the representation is not of worry, here.

(ii) Query Model: It specifies the structure of the input that can be given to the system. For keyword questions, the most widely recognized type of input is an arrangement of words or terms. This improves the assignment of querying, since, the client is required to know neither any query dialect nor the pattern of the database. An all the more effective type of querying is by utilizing chart or tree designs.

(iii) Answer Model: It specifies the structure of a response to a query and the necessities that it must fulfill as per the semantics of the system. The answers are normally spoken to as a chart or tree, or it might be only a tuple or a term.

(iv)Scoring Model: In general, there will be many answers to a similar keyword query and thus a large portion of the systems utilize a scoring model, which doles out a score to each of the answers, in light of their pertinence. The idea of importance is extremely ambiguous, since it relies on upon the client. Likewise, since keyword querying is not a capable technique as far as expressiveness, the clients can't well-spoken their prerequisites precisely. Henceforth, rather than a solitary answer, the system must return beat few records with the highest scores. The score is reliant on the semantics of the system. A straightforward strategy utilized, is to give higher score to an answer with more modest number of joins. Be that as it may, most systems utilize complex tenets to relegate scores, to enhance the nature of the top positioned answers.

LITERATURE SURVEY

Yue Lu, Chew Lim Tan [1] suggested that a huge amount of document images are accessible in the Internet and digital libraries. They find that, most of them are packed in PDF files and are compressed using CCITT Group 4 standards for saving storage space and speeding up transmission. There is thus significant meaning to develop the methods of directly searching keywords from these documents. In this paper, they present a compressed pattern matching method for searching keywords from the CCITT Group 4 compressed document images, without explicit decompression.

According to the CCITT Group 4 standards, each coded position indicates that the current pixel colour is different from its previous pixel, except for the next coded positions of the

pass mode. In their work, they extract these changing elements from the compressed images directly. The changing elements are utilized to segment and bound the word objects, and are used for measuring the similarity of two word images. The connected components are labelled based on the line- by-line strategy according to the relative positions between the changing elements of the current coding line and the changing elements of the reference line.

SanketS.PawarAbhijeetManepatilAniketKadamPrajaktaJagtap[2], This research work is devoted to keyword inquiry and gives two viewpoints of its application in IR and database system. Article show prototype of Machine An and B, where A presents Innovative IR system and B presents Discover approach relational database management system. Article concentrates more on stretching out keyword hunt to database management system as it less tended to subject and all the more difficult. Examination of Machine B demonstrate that execution assessment need to address with successful assessment like inquiry workload memory utilization for adaptable and versatile advanced machine improvement.

Instead of assessment parameters like time delay and so forth mixture versatile report retrieval system is construct and evaluated on memory utilization and inquiry space is decreased significantly with two layer algorithm. Assist extent of system is creating hybridization at machine level and working with pictures as information question.

Qiuxiang Dong, Zhi Guan, ZhongChen[3] In this paper, they grow new techniques that split the computation for the keyword encryption and trapdoor/token era into two stages: an arrangement stage that does by far most of the work to encrypt a keyword or make a token before it knows the keyword or the property list/access control strategy that will be utilized. A moment stage then quickly collects a middle figure content or trapdoor when the specifics get to be distinctly known. The readiness work can be performed while the cell phone is connected to a power source, then it can later quickly perform keyword encryption or token era operations moving without fundamentally depleting the battery. We name our plan Online/Offline ABKS. To the best of our insight, this is the primary work on building productive multi-client searchable encryption conspire for cell phones through moving most of the cost of keyword encryption and token era into a disconnected stage.

DrKehinde K. Agbele, Eniafe F. Ayetiran, Kehinde D. Aruleba and Daniel O. Ekong [4] proposed this article to create algorithms that improve the positioning of records recovered from IRS as per client seek setting. Specifically, the positioning assignment that drove the client to take part in information-chasing conduct amid inquiry errands. This article examines and portrays a Document Ranking Optimization (DROPT) algorithm for IR in an Internet-based or assigned databases environment. On the other hand, as the volume of information accessible on the web and in assigned databases is developing persistently, positioning algorithms can assume a noteworthy part with regards to list items. In this article, a DROPT technique for archives recovered from a corpus is produced as for report list keywords and the question vectors. This depends on figuring the weight ( ) of keywords in the report list vector, ascertained as a component of the frequency of a keyword over a record. The motivation behind the DROPT technique is to reflect how human clients can judge the setting changes in IR result rankings as per information significance. This article demonstrates that it is workable for the DROPT technique to beat a portion of the limitations of existing conventional (×) algorithms by means of adjustment. The observational assessment utilizing measurements measures on the DROPT technique helped out through human client cooperation indicates change over the conventional importance input technique to show enhancing IR viability.

XiaoliLian, Mona Rahimi,Remo Ferrari and Michael Smith [4] ,In this paper they first investigate the exertion expected to physically fabricate an abnormal state space demonstrate catching the utilitarian segments. They then present MaRK (Mining Requirements Knowledge), which recognizes and recovers the records containing depictions of practical parts in the area demonstrate. Area investigators can utilize this information to indicate prerequisites. They present and assess an algorithm which positions space archives as indicated by their significance to a part and after that highlights segments of content which are probably going to contain prerequisites related information. They portray prepare inside the setting of the Positive Train Control (PTC) area with a vault of 523 archives, speaking to 852MB of information. They experimentally assess the MaRK significance algorithm and its capacity to recover important prerequisites information for necessities identified with PTC's On-Board Unit.

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 permits the clients to input multiple query keywords to ask for reasonable records. Among these works, conjunctive keyword search plots just give back the records that contain the greater part of the query keywords. Disjunctive keyword search plans give back the greater part of the records that contain a subset of the query keywords.
Positioned search can empower fast search of the most significant information. Sending back just the top-k most significant archives can adequately diminish organize activity.
System demonstrate
The system demonstrates in this paper includes three unique elements: information proprietor, information client and cloud server.

Data owner: A data owner has a collection of records F = {f1, f2, ...,fn} that he wants to outsource to the cloud server in encrypted shape while as yet keeping the capability to search on them for powerful utilization. In our plan, the data owner firstly fabricates a secure searchable tree file I from record collection F, and then generates an encrypted archive collection C for F. Afterwards, the data owner outsources the encrypted collection C and the secure file I to the cloud server, and securely conveys the key information of trapdoor generation (counting keyword IDF values) and archive unscrambling to the authorized data clients. In addition, the data owner is in charge of the update operation of his reports put away in the cloud server. While updating, the data owner generates the update information locally and sends it to the server.

Data users:Data users are authorized ones to access the reports of data owner. With the query keywords, the authorized client can generate a trapdoor TD according to search control mechanisms to get k encrypted records from cloud server. Then, the data client can unscramble the records with the shared mystery key.

Cloud server: This server stores the encrypted report collection C and the encrypted searchable tree index I for data owner. After receiving the trapdoor TD from the data client, the cloud server executes search over the index tree I, and finally gives back the corresponding collection of topk ranked encrypted reports. Moreover, after receiving the

update information from the data owner, the server needs to update the index I and record collection C according to the got information. The cloud server in the proposed plan is considered as "genuine yet inquisitive", which is utilized by loads of takes a shot at secure cloud data search [25], [26], [27]. Specifically, the cloud server sincerely and accurately executes search control (trapdoors) access control (data decoding keys) Semi-trusted cloud server encrypted index tree search ask for encrypted records best k ranked outcome Fig. 1. The architecture of ranked search over encrypted cloud data −0.2 0 0.2 0.4 0.6 0.8 0 2 4 6 8 10 Term frequency # of reports(×10 2) (a) −0.2 0 0.2 0.4 0.6 0.8 0 2 4 6 8 10 12 Term frequency # of archives(×10 2) (b) Fig. 2. Appropriation of term frequency (TF) for (a) keyword "subnet", and (b) keyword "have". instructions in the designated convention. Meanwhile, it is interested to infer and analyze got data, which helps it acquire additional information. Depending on what information the cloud server knows, we adopt the two threat models proposed by Cao et al.

Gaps in the Base Paper:

Firstly, all the users usually keep the same secure key for trapdoor generation in a symmetric SE conspire. In this case, the revocation of the client is huge challenge. In the event that it is expected to repudiate a client in this plan, we have to modify the index and disperse the new secure keys to all the authorized users.

Also, symmetric SE conspires usually assume that all the data users are dependable. It is not practical and an exploitative data client will lead to many secure issues. For example, a deceptive data client may search the reports and convey the unscrambled archives to the unauthorized ones. Considerably more, an unscrupulous data client may disseminate his/her secure keys to the unauthorized ones. Later on works, we will attempt to enhance the SE plan to handle these challenge 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,

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

PROPOSED WORK AND IMPLEMENTATION

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

Algorithm 1: For Keyword Search
Step1: 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,

DOI: https://dx.doi.org/10.26808/rs.ca.i7v5.01
International Journal of Computer Application (2250-1797)
Volume 7– No.5, September-October 2017

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.

i.e. 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 order to just get the matched document will the keyword.

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.

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.

DOI: https://dx.doi.org/10.26808/rs.ca.i7v5.01
International Journal of Computer Application (2250-1797)
Volume 7– No.5, September-October 2017

Step4: Then mail the generated pin to the user and user then reenter the pin using the same process as mentioned in the step 1.

We have created the implementation on line on the website, using the PHP and MYSQL.

4.4 Keyword Search

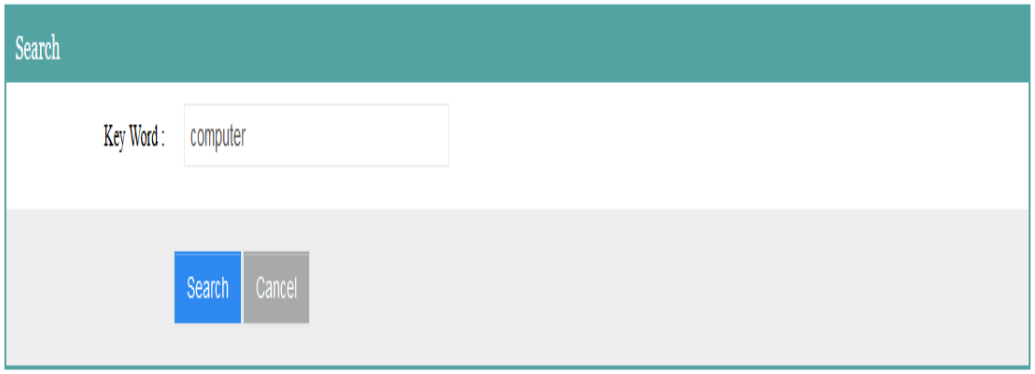

Fig 4.6 Keyword Search Form

This is the keyword search form and we will specify the keyword we want to search in the file. (This module is partially completed; we are working on refining the algorithm for the keyword search)
Then the files which are matched for the keyword are listed and we cannot directly download the file we have to request the access for downloading the file.

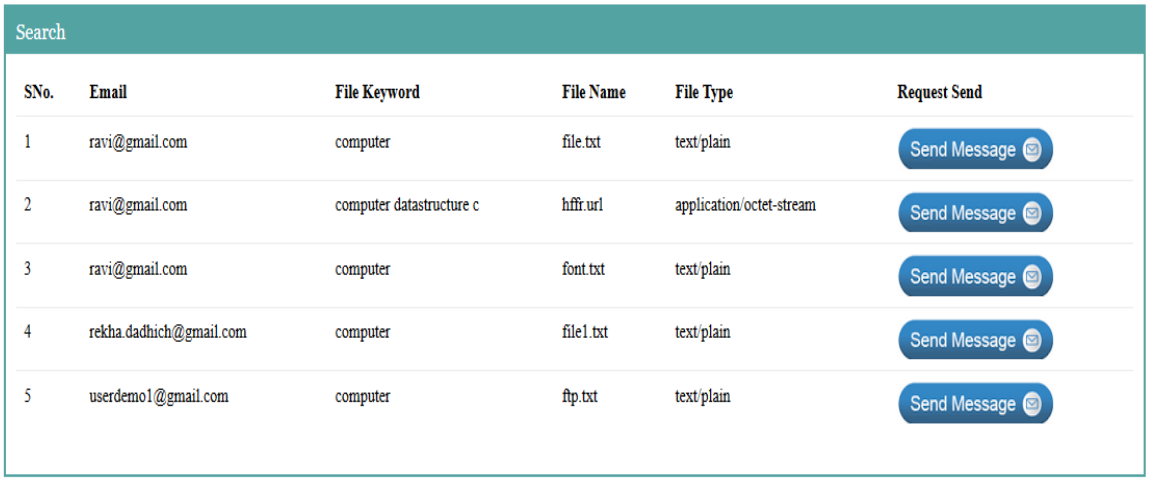

Fig 4.7 Result of Keyword Search

When we click on the send message then the request for accessing the file is generated and then the role of admin will come into play that will generate the secure graphical pin for accessing the file. (Here the userdemo2 has requested for the access)

DOI: https://dx.doi.org/10.26808/rs.ca.i7v5.01
International Journal of Computer Application (2250-1797)
Volume 7– No.5, September-October 2017

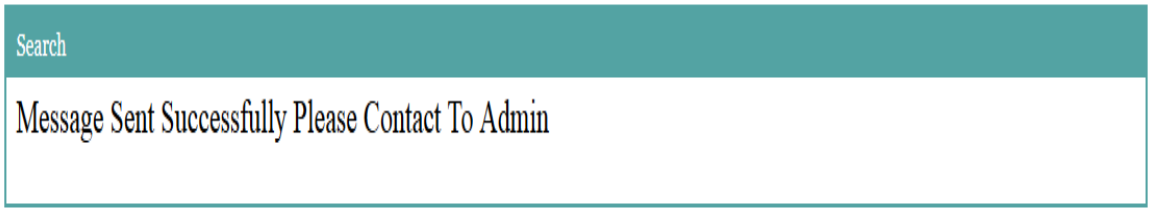

Fig 4.8 Message for Successful Request

4.1 View Requests

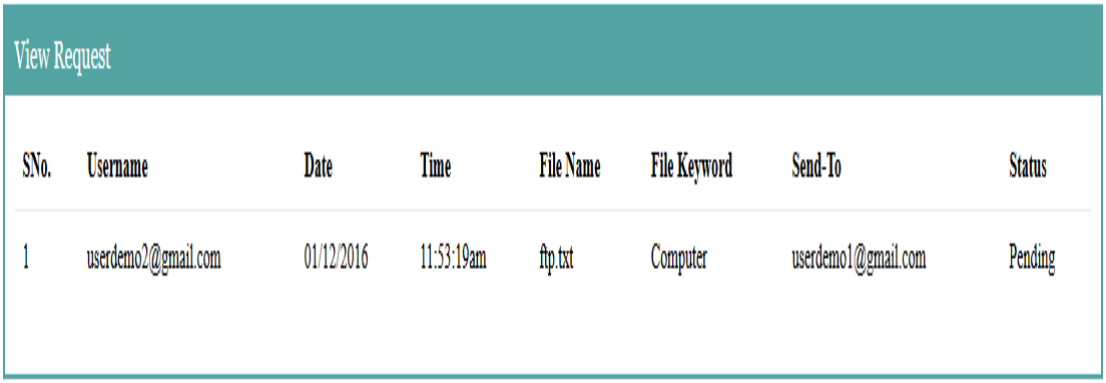

Fig 4.9 Request Status Form

In this section we will check and view the status of our requests which we or the user have requested for access. The status is shown pending until the administrator will grant the access to the use.

4.2 Admin Login page

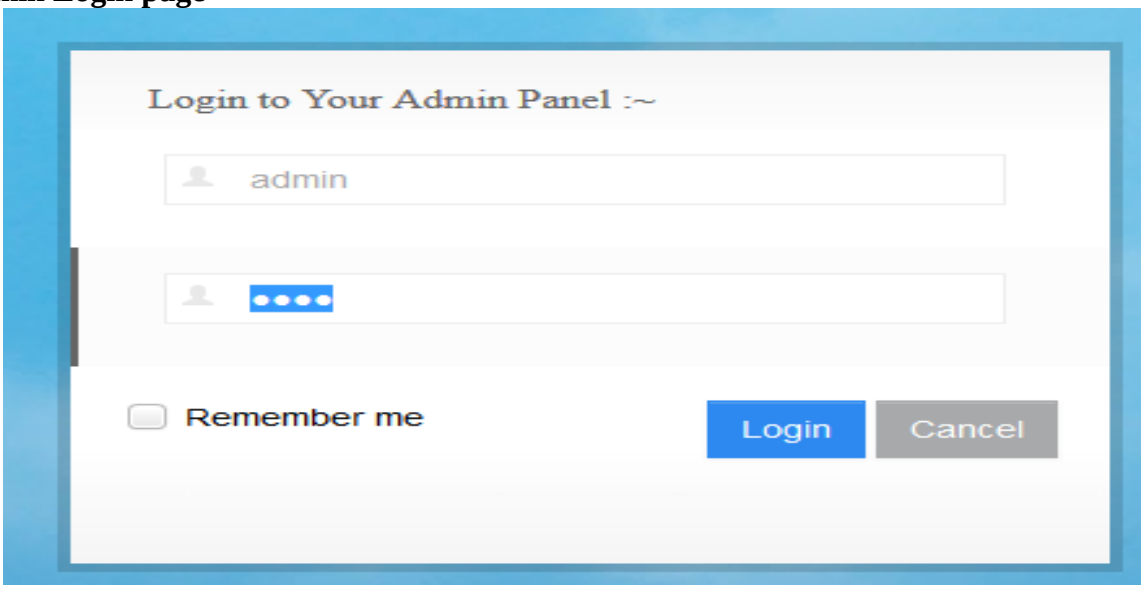

Fig. 4.10 Admin Login Form

This is admin login page in order to access the operations which an admin can perform. The administrator is come into role to control the access and granting the access to the requested file.

DOI: https://dx.doi.org/10.26808/rs.ca.i7v5.01
International Journal of Computer Application (2250-1797)
Volume 7– No.5, September-October 2017

4.3 Admin welcome page

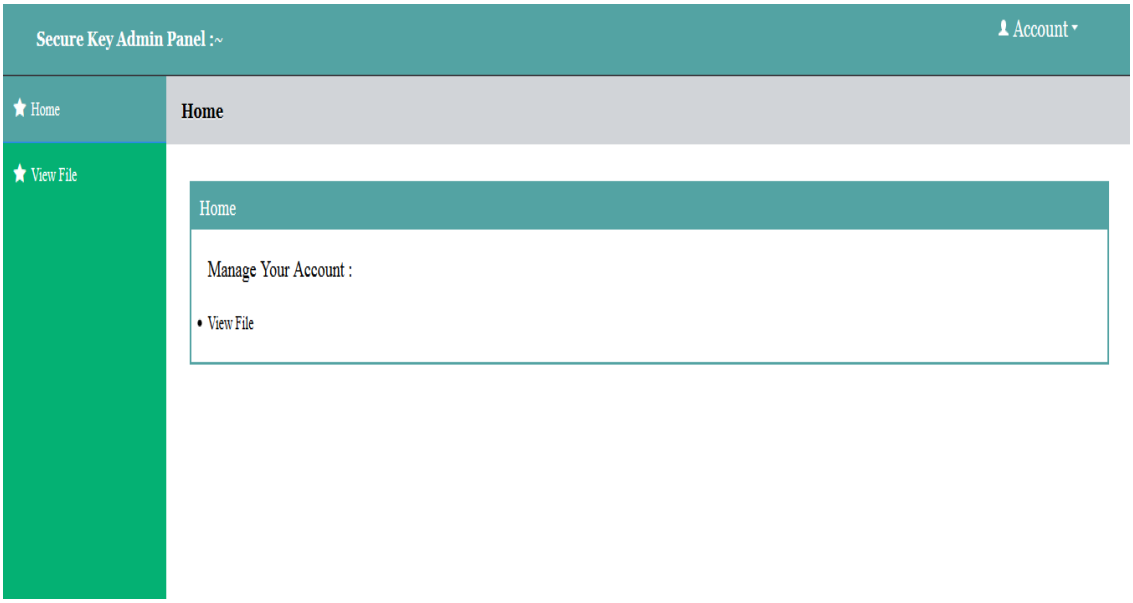

Fig 4.11 Administrator Welcome Page

This is the admin welcome page. This page will appear when the administrator will get the access after the successful login.

4.4 View Requests from Users

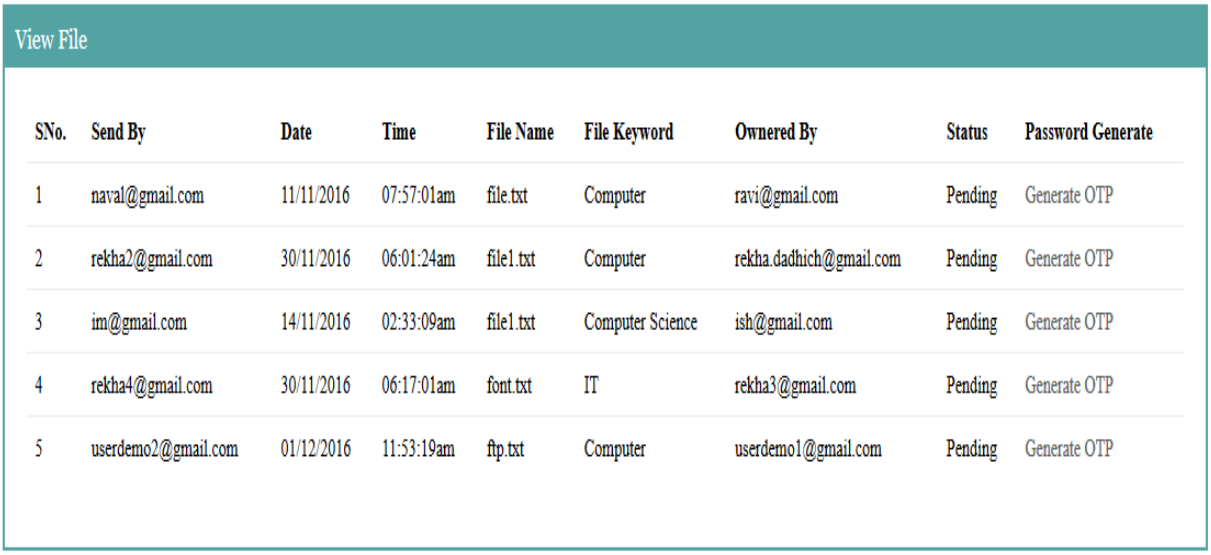

Fig 4.11 View Request Page for Administrator

This section is used for accessing requests for the files made by the client and admin will provide the access by generating the graphical pin.
For this purpose we have to click on the generate OTP

DOI: https://dx.doi.org/10.26808/rs.ca.i7v5.01
International Journal of Computer Application (2250-1797)
Volume 7– No.5, September-October 2017

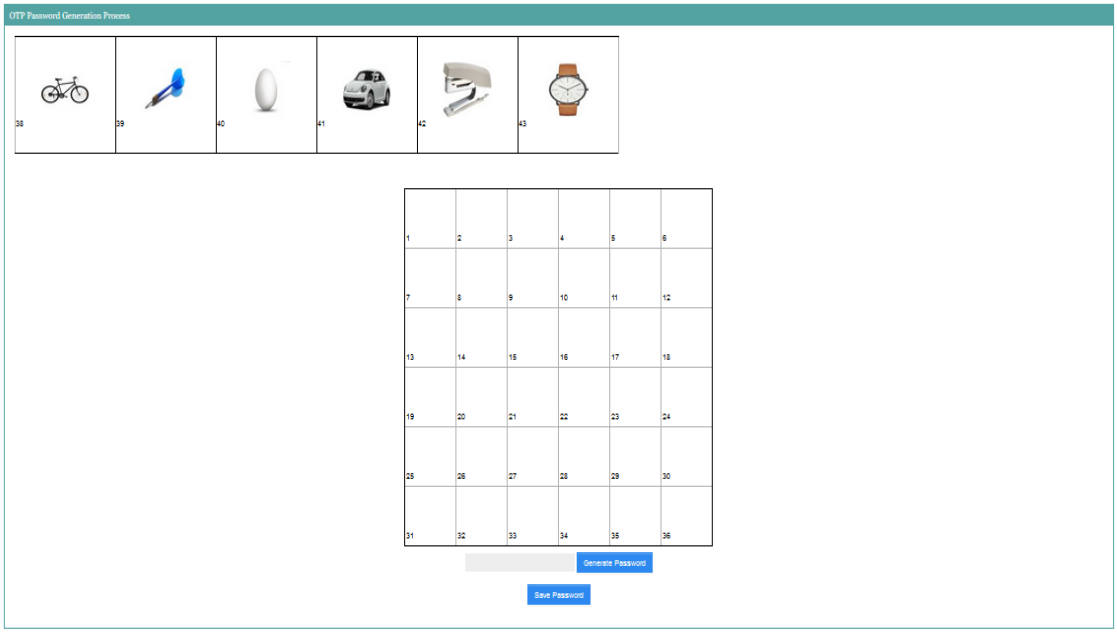

Fig 4.12 OTP Generation Form.

Now click on the image and then click on the position where you want to place the image in order to generate the OTP.

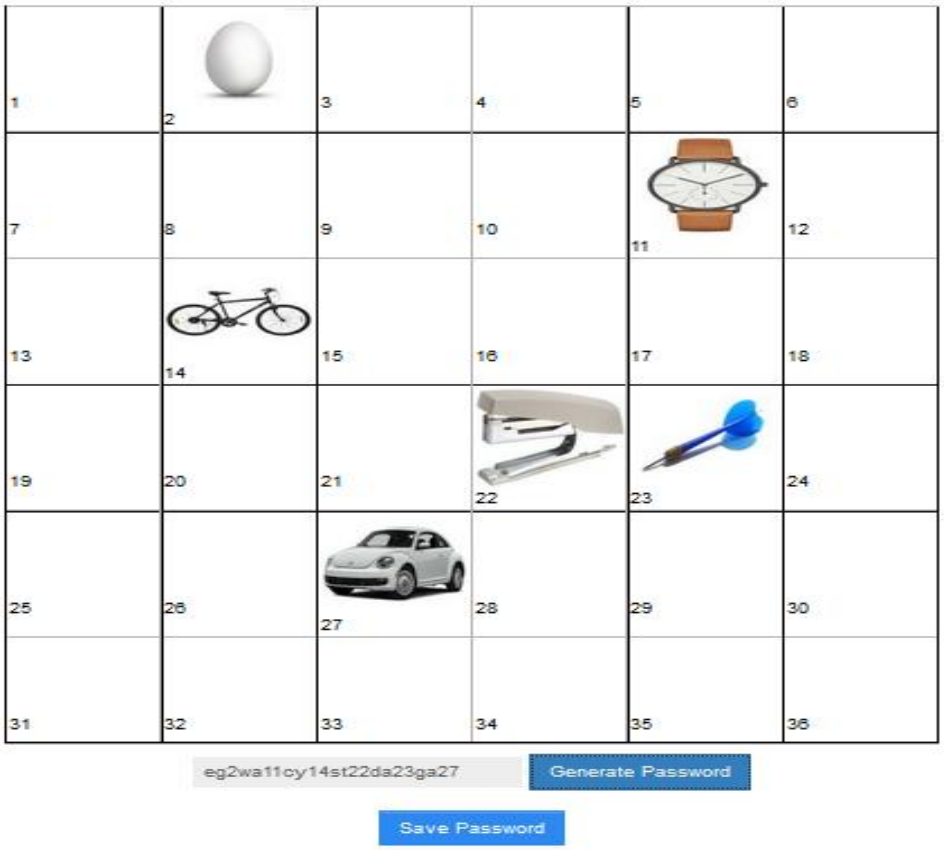

Fig 4.13 OTP Generation Form after Image Placement

Then click on save password, the OTP will be send by mail to the client or the user requesting the access.

DOI: https://dx.doi.org/10.26808/rs.ca.i7v5.01
International Journal of Computer Application (2250-1797)
Volume 7– No.5, September-October 2017

TESTING AND RESULTS

5.1 Test Results

We have tested the OTP generated by our proposed implementation using the various tools to check its strength. Below is presented the some of the test analysis presented on the OTP.

5.1.1 Password Meter

The website www.passwordmeter.com is an online site which tests the strength of the password. This application is designed to assess the strength of password strings. The instantaneous visual feedback provides the user a means to improve the strength of their passwords, with a hard focus on breaking the typical bad habits of faulty password formulation. Since no official weighting system exists, they created formulas to assess the overall strength of a given password.

Test Your Password		Minimum Requirements
Password:	•••••••••••••••••••	• Minimum 8 characters in length • Contains 3/4 of the following items: - Uppercase Letters - Lowercase Letters - Numbers - Symbols
Hide:	☑	
Score:	100%	
Complexity:	Very Strong	

Additions	Type	Rate	Count	Bonus
Number of Characters	Flat	+(n*4)	18	+ 72
Uppercase Letters	Cond/Incr	+((len-n)*2)	0	0
Lowercase Letters	Cond/Incr	+((len-n)*2)	12	+ 12
Numbers	Cond	+(n*4)	6	+ 24
Symbols	Flat	+(n*6)	0	0
Middle Numbers or Symbols	Flat	+(n*2)	5	+ 10
Requirements	Flat	+(n*2)	3	0
Deductions				
Letters Only	Flat	-n	0	0
Numbers Only	Flat	-n	0	0
Repeat Characters (Case Insensitive)	Comp	-	5	- 1

Fig 5.1 Test Results for website www.passwordmeter.com

The test password which is given is:
OTP: cy1da2eg3ga4st5wa6
Result: Very Strong

5.1.2 Password Checker

Password Checker Online helps you to evaluate the strength of your password. More accurately, Password Checker Online checks the password strength against two basic types of password cracking methods – the brute-force attack and the dictionary attack. It also analyzes the syntax of your password and informs you about its possible weaknesses. This tool can thus also help you create stronger password from a weak one.

Using Password Checker Online is safe in both the syntax analyzing mode and the dictionary attack mode. When you type your password to the Password field its syntax is analyzed on the client side, by JavaScript in your browser – i.e. the password is not transferred over the network to our server. You can also have your password checked against the dictionary attack. In this case the password is sent to our server in an encrypted form so you do not need to worry about attackers sniffing on your network. However, the implementation is not safe against man-in-the-middle attacks.

The score computation is mostly based on the time that a middle size botnet would need in order to crack your password if it employs the brute-force attack. An attacker typically tries several most common passwords first therefore if your password belongs to the list of 10000 most common passwords your password receives score 0 because these passwords are extremely weak.

There are important attributes of your password that an automatic tool such as Password Checker Online cannot evaluate. When you create your password you should keep in your mind that it should not contain any information that is related to you or the system in which the password is used. You should also avoid using same or similar passwords in different systems – e.g. you should use completely different passwords for your social network account and for your email account.

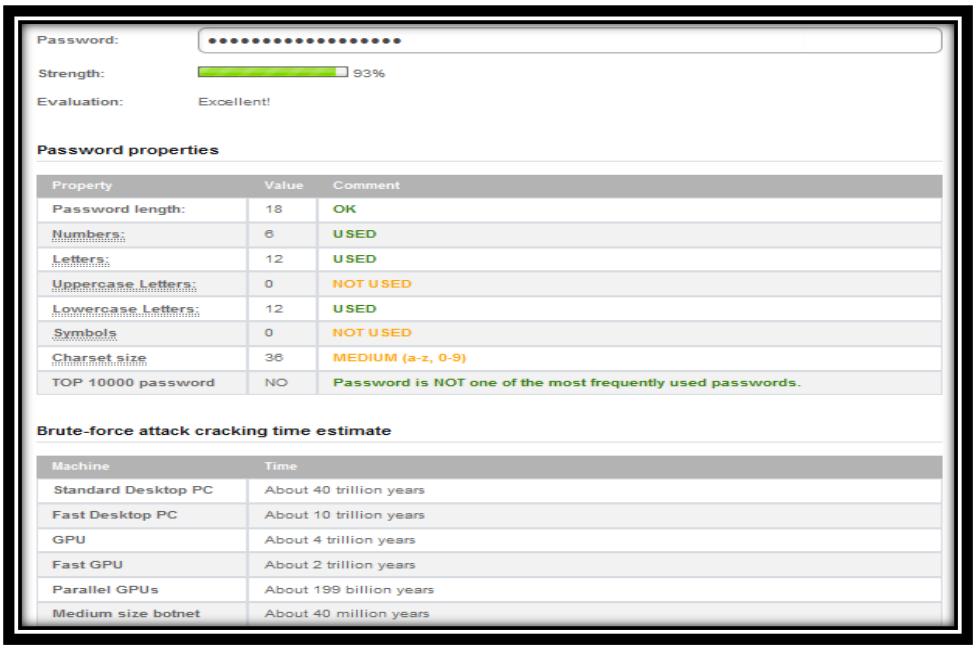

Fig 5.2 Test result for password on site http://password-checker.online-domain-tools.com/

5.1.3 CryptTool2

DOI: https://dx.doi.org/10.26808/rs.ca.i7v5.01
International Journal of Computer Application (2250-1797)
Volume 7– No.5, September-October 2017

CrypTool is a program for learning cryptographic algorithms. It provides a graphical user interface for visual programming. Thus, workflows can be visualized and controlled to enable intuitive manipulation and interaction of cryptographic functions. The vector-oriented GUI is based on the Windows Presentation Foundation (WPF) and gives users the ability to scale the current view at will.

Fig 5.3 Crypttool2

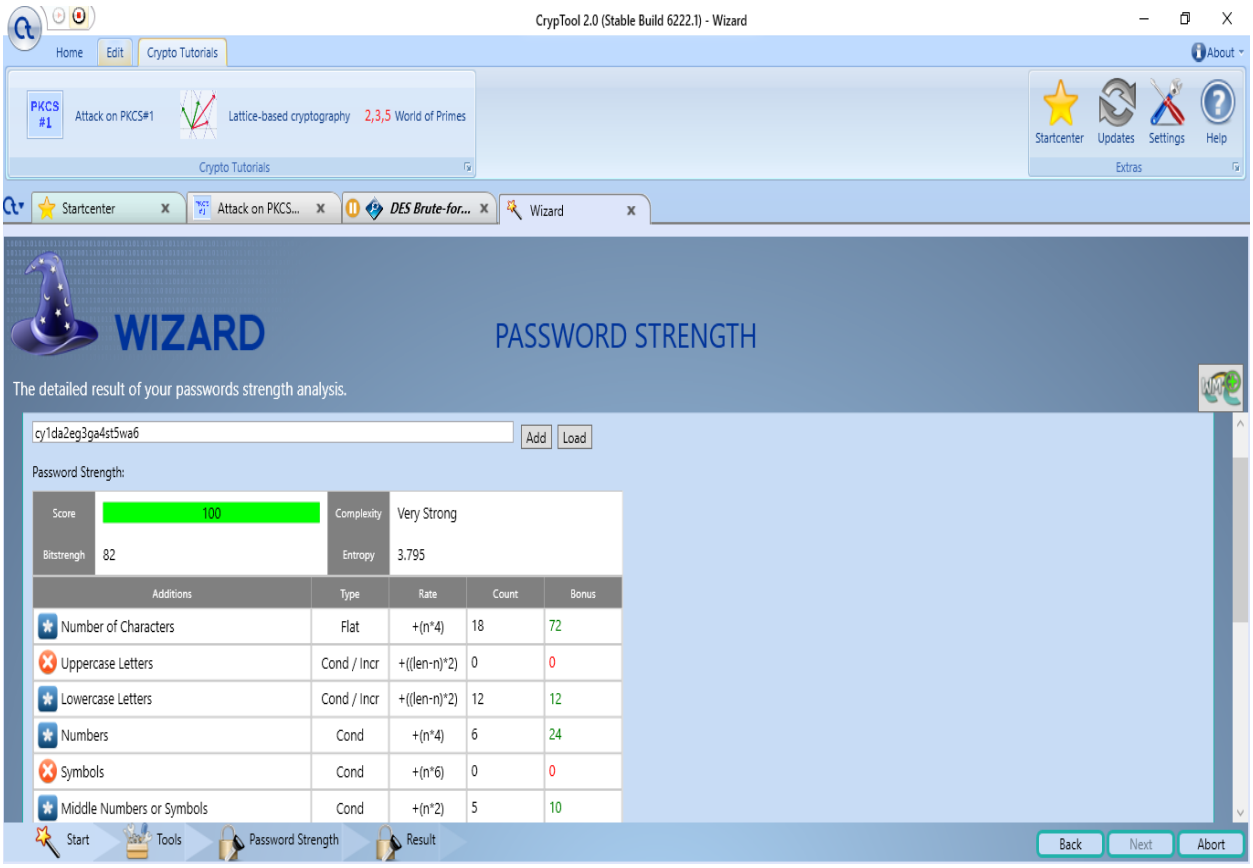

Fig 5.4 Test Result using Cryptool2

The above test can be summarized using the table 5.1.

Test OTP	Website/Tool	Result
cy1da2eg3ga4st5wa6	PasswordMeter	Very Strong
cy1da2eg3ga4st5wa6	Password Checker	Excellent Strength
cy1da2eg3ga4st5wa6	Cryptool2	Entropy 3.75 Strength 82 Very Strong

TABLE 5.1 TEST RESULT ANALYSIS TABLE

CANCLUSION AND FUTURE SCOPE

As of now information security is crucial to all organization to ensure their information and behaviors their business. Information security is characterized as the assurance of information and the framework, and equipment that utilization store and transmit that information. Information security performs four vital for an organization which is ensure the organization's capacity to work, empower the sheltered operation of utilizations actualized on the organization's IT systems, secure the information the organization gather and uses, and in conclusion is shields the technology resources being used at the organization. There are likewise difficulties and hazard includes in actualized information security in organization.

In an organization, information is vital business resources and basic for the business and along these lines require fitting ensured. This is particularly essential in a business domain progressively interconnected, in which information is presently presented to a developing number and a more extensive assortment of dangers and vulnerabilities. Cause harm, for example, noxious code, PC hacking, and disavowal of administration assaults have turned out to be more typical, more goal-oriented, and more complex. Along these lines, by executed the information security in an organization, it can ensure the technology resources being used at the organization.

In term of ensuring the functionality of an organization, both general management and IT management are in charge of executing information security that secures the organization capacity to work. Information is the most critical component in organization to work together. Other than that an organization is kept their clients information, so it is crucial for them to ensure the information. Without information, the business can't be run. By secure the information store; it can empower the organization to run business also. That is the reason the information security is critical in organizations.

Securing the password to be cracked is also an issue when sending the password online. In our proposed implementation we have created the algorithm for Visual Method of password generation and have tested the password strength of the OTP which is generated after the process.
Thus, we can say that our proposed implementation provides a better way to securely share the data online.

In the further studies, we will like to extend our research to use the real time password like live pictures, video and retina verification concepts for sharing the file to further enhance the security in the suggested framework.

REFERENCE

[1]. Gary Pan, SeowPoh Sun, Calvin Chan and Lim Chu Yeong,"Analytics and Cyber security: The shape of things to come",CPA ,2015

[2]. ErolGelenbe and Omer H. Abdelrahman,"Search in the Universe of Big Networks and Data",IEEE ,2014

[3]. 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

[4]. 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 ,2015

[5]. Zhihua Xia, Member, Xinhui Wang, Xingming Sun, and Qian Wang,"A Secure and Dynamic Multi-keyword Ranked Search Scheme over Encrypted Cloud Data",IEEE,2015

[6]. 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

[7]. N. L. Sarda and A. Jain. Mragyati: A system for keyword-based searching in databases. http://arxiv.org/abs/cs.DB/0110052.

[8]. Marcos D. Assuncao, Rodrigo N. Calheiros, Silvia Bianchi, Marco A.S. Netto, RajkumarBuyya "Big Data computing and clouds: Trends and future directions",J. Parallel Distrib. Comput. 79–80 (2015) 3–15.

[9]. Venkata NarasimhaInukollu, SailajaArsi and Srinivasa Rao Ravuri "SECURITY ISSUES ASSOCIATED WITH BIG DATA IN CLOUD COMPUTING",International Journal of Network Security & Its Applications (IJNSA), Vol.6, No.3, May 2014.

[10]. Puneet Singh Duggal, Sanchita Paul "Big Data Analysis: Challenges and Solutions" International Conference on Cloud, Big Data and Trust 2013, Nov 13-15, RGPV.

[11]. Zan Mo, Yanfei Li "Research of Big Data Based on the Views of Technology and Application" American Journal of Industrial and Business Management, 2015, 5, 192-197.

[12]. K.Arun, Dr.L.Jabasheela "Big Data: Review, Classification and Analysis Survey" International Journal of Innovative Research in Information Security (IJIRIS) Volume 1 (September 2014) ISSN: 2349-7017(O) ,ISSN: 2349-7009(P)

[13]. KalyaniShirudkar, DilipMotwani "Big-Data Security" International Journal of Advanced Research in Computer Science and Software Engineering Volume 5, Issue 3, March 2015.

[14]. Nishu Arora, Rajesh Kumar Bawa "A Review on Cloud to Handle and Process Big Data" International Journal of Innovations & Advancement in Computer Science IJ IACS ISSN 2347 – 8616 Volume 3, Issue 5 July 2014

[15]. A B M Moniruzzaman, Syed Akhter Hossain "NoSQL Database: New Era of Databases for Big data Analytics-Classification, Characteristics and Comparison" International Journal of Database Theory and Application Vol. 6, No. 4, August, 2013.Shilpa, Manjit Kaur "BIG Data and Methodology-A review" Volume 3, Issue 10, October 2013 ISSN: 2277 128X.