Support Ranked Keyword on Remote Encrypted Data in Cloud

S. Usha[#1], Dr. A. Tamilarasi[#2], R. Vijayakumar[*3]

[#1]*Assistant Professor, Dept. of CSE, University College of Engg., Panruti Campus, Tamilnadu, India*
[#2]*Professor & HOD, Kongu Engineering College, Perundurai,Tamilnadu, India*

[*]*PG Scholar, Dept. of CSE, Anna University, Trichy, India.*

***Abstract*— Data owners outsource their complex data management systems from local sites to commercial public space for great flexibility and economic savings. However, the sensitive data should be kept extremely private from the users. Thus, every datum is needed to be encrypted before outsourcing the data. Also, the search and utilization of the outsourced data should be easier. An effective data retrieval need is met with the server, which performs result relevant ranking to give back the most relevant information. This is done by a principle named coordinate matching, which is used to capture the similarity between the search query and data documents. Existing system focuses on single keyword search or Boolean keyword search, which cannot serve the purpose and also no differentiation among the results, are done. In this work, every keyword of the user's query is taken into consideration and the ranked relevant information is provided, based on coordinate matching. User can download the data, only with the activation code that is sent through email. Thus, the privacy is also preserved.**

***Keywords*— Cloud, encrypted data, outsourced data, encrypted search**

I. INTRODUCTION

Cloud computing provides a simple way so as to reduce the memory by outsourcing the data to a cloud, managed by cloud provider. The data owners have to take certain efforts, to manage their data and also a considerable amount has to be spent for data maintenance. Also, the data grows with respect to time, so a huge memory space is needed and is not affordable by many, so they go for a good option called outsourcing.

With this, the data is outsourced to any commercial public space. Certain amount has to be paid for the public space being used. However, it is more economic to pay rent for the memory in use rather than own excessive memory. Thus, the economic savings motivate both the individuals and enterprises to outsource their data. In this scenario, many challenges like data security, searching and retrieval of data hit the scene.

Many data owners may go for data outsourcing; however every piece of data is very sensitive and has to be kept private. Thus, the data owners encrypt their data and index before outsourcing the data. Searching and utilization of data should be effective, since the data is in encrypted form. Downloading and decrypting locally is impractical. Also, outsourcing will not serve the purpose, unless the data can be easily searched and utilized.

Effective data retrieval has to perform ranking operation, such that the most relevant information appears first. This ranked search system, allows data users to find the most relevant information quickly. In the past, hackers generally targeted larger companies that stored government information or financial data. However, hackers today are becoming increasingly brazen and expanding their sights to information that can lead to more valuable data down the line.

According to a magazine named "Information Week", 73 percent of small and medium sized businesses report that they have been the victims of cyber attacks in the past year and 42 percent have lost confidential and proprietary information [1].

Effective security requires continuous monitoring and management as well as security experts who are able to identify and respond to network threats immediately. Outsourcing data to a loyal public space provides data owners, with a team of experienced security analysts with the ability, to assess the severity of a cyber attack, to formulate the proper response and implement the optimal defence.

Outsourcing allows the data owners to lower costs, save time and increase efficiency. To reduce space complexity, it is advised to outsource the data than to maintain by own. A strong assurance is provided for security.

Some of the benefits in outsourcing data are as follows. When the data is outsourced, economic savings and an effective search is experienced. Around 60 percent of the cost is saved, through outsourcing and also a faster and a better service is provided [2].

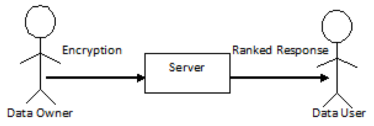


Fig 1: Data Outsourcing

However, data owners are needed to be cautious while choosing the commercial public space. Before outsourcing, the security policy of the public space should be checked with an eagle eye. Selection should be done, only on the basis of a strict and a sound security policy. Small businesses lose billions of dollars each year, due to data loss from hardware or system failure, human error and software corruption.

Natural disasters account for one percent of all data loss. Although the probability is small, the impact can be overwhelming. In fact, companies that are not able to resume operations, within ten days of a disaster are not likely to survive.

Catastrophes aside, being fast and nimble is one of the defining characteristics of small and medium sized businesses. In any business however, finding the right version of last month financial statements, can sometimes be difficult. If it takes two weeks or more to access and validate an operating unit's financials, that may have just lost the deal.

It is not only important to provide data protection, but quick and easy access to the right information. In today's increasingly wired business world, the best way to recover from data loss due to a natural disaster is to store backup data in a physically separate location.

There are companies that specialize in preserving documents on tapes and discs that are stored within other structures. Although, they too are susceptible to natural disasters, these companies have helped to save numerous businesses where the damage has only been internal. Another option that has been gaining attention recently is to store files remotely.

Once, services are available only to large companies because of certain cost considerations, but now, there are several firms specializing in offering offsite data storage to small businesses. Such systems offer their clients the opportunity to store, archive and even authenticate their data over a remote network for less than dollar 100 a month, so that important files can be accessible, even if the office is not.

The global storage rate increases by seventy percent every year [3]. Outsourcing in this manner often makes the most sense. The growing data demands of small and large corporations continue to outspace their current storage capacities. So, several storage solutions are provided to allow for the expansion and growth of business. High data availability and reliability are the major concerns here.

The global data increases around 70 percent every year and hence data storage is a million dollar question to small and medium scale business. Data backup is another thing to be worried of. Small and medium scale industries cannot afford a huge amount for storage.

A famous magazine named Information week reports that about 73 percent of small and medium scale businesses are the victims of cyber attacks and 42 percent among them lost their confidential and sensitive information.

Data security is the major concern here. Also, due to natural disasters, one percent of data loss hit the scene in small businesses. Apart from this, software corruption, hardware or system failure also paves way for the data loss.

The afore said problems can be addressed by a solution named 'Outsourcing', which takes care of the data of any entity based on some contract. However, searching and utilization of data must be made easy, such that the user can have an effective search experience.

Searching involves ranking in this work, hence the best relevant information has its room first. Multiple keywords can be given as search string, based on which the results are given back. The data owner submits his data to the public space in an encrypted format. Also, the data security is ensured by strict authorization.

II. Literature Survey

Small businesses lose billions of dollars each year, due to data loss from hardware or system failure, human error and software corruption. Natural disasters account for one percent of all data loss. Although the probability is small, the impact can be overwhelming.

In fact, companies that are not able to resume operations, within ten days of a disaster are not likely to survive. Catastrophes aside, being fast and nimble is one of the defining characteristics of small and medium sized businesses.

In any business however, finding the right version of last month financial statements, can sometimes be difficult. If it takes two weeks or more to access and validate an operating unit's financials, that may have just lost the deal. It is not only important to provide data protection, but quick and easy access to the right information.

In today's increasingly wired business world, the best way to recover from data loss due to a natural disaster is to store backup data in a physically separate location. There are companies that specialize in preserving documents on tapes and discs that are stored within other structures. Although, they too are susceptible to natural disasters, these companies have helped to save numerous businesses where the damage has only been internal.

Another option that has been gaining attention recently is to store files remotely. Once, services are available only to large companies because of certain cost considerations, but now, there are several firms specializing in offering offsite data storage to small businesses. Such systems offer their clients the opportunity to store, archive and even authenticate their data over a remote network for less than dollar 100 a month, so that important files can be accessible, even if the office is not.

High data availability and reliability are the major concerns here. Starting from the local level to massive Federal agencies, a wide range of cost-effective backup and restore products are available to guarantee the delivery of high performance and scalability.

The most common reasons for why companies decide to outsource their data are cost reduction and cost savings, the ability to focus on its core business and increased profits. Companies are able to focus their money and resources more towards improving the core aspects of its business when outsourced. The landscaping is performed by an expert outsourced organization and the insurance company can focus on doing what it specializes in. This allows the outsourcing company to build onto its core functions that keep the business running smoothly.

Revenue and profit plays a large role in the reason for a company outsourcing. Since, the costs are cheaper in different countries for a corporation to run it, as well as to train the employees; this saves the company a large sum of money. More profit comes in, when the vendors are able to purchase products at a less expensive rate and continue to sell them at a reasonable price for consumers.

The prices are reduced for services as well as products when purchased at a cheaper price. Companies are able to provide services and products to consumers at a cheaper price, while still having a large margin for profit.

This profit margin benefits both the company as well as the consumer. The cheaper prices lead to an increase a company's economy. Although losing jobs hurts the economy because more citizens become unemployed, the cheaper prices allows customers to purchase more products and services which helps to rebuild an economy.

III. Proposed Work

In this paper, we propose a new mechanism that assures confidentiality and reliability of data outsourced. We use AES algorithm to ensure security. The data is encrypted so as to assure security for the data owners. The secured data is outsourced to the cloud server. Thus, we need to provide a mechanism that makes data search and retrieval easier. To make this true, we provide multi-keyword search support and ranked response to the cloud users. We support ranked response by a ranking algorithm that considers location and frequency of the occurrence of the search keyword. Based on the number of occurrence, relevant response is returned.

IV. Algorithm & Architecture

Algorithm of the proposed system is presented in Fig 2 and its corresponding architecture is presented in Fig 3.

```
Apply AES algorithm over data;
Outsource encrypted data to cloud provider;
Support multiple keyword search by concatenation;
Support ranked search with location and frequency;
End process;
```

Fig 2: Algorithm

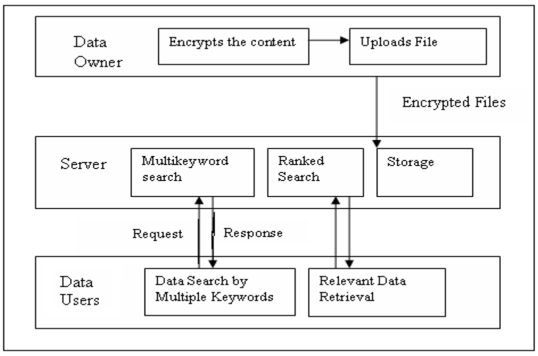


Fig 3: Architecture

Data owners outsource their data to a public space to save economy and memory space. However, they encrypt their data for security and are submitted to the server. To improve the utilization of data, multi-keyword and ranked search are enabled in this system, to enhance the searching experience of users.

The system entitled 'Privacy preservation of encrypted data with a multi-keyword rank based search, deals with the security, searching and utilization of outsourced data. Security is assured by strong encryption of the outsourced data.

This system employs multi keyword searching, in which multiple keywords can be given as search string, through which an effective search experience is felt by the clients.

Also, ranked results are given back to the clients, such that the best result that well suits the search string is provided. The data owner encrypts all the data before outsourcing. Also, the client can access the data only with the activation code that is sent through email. Thus, the privacy is also preserved.

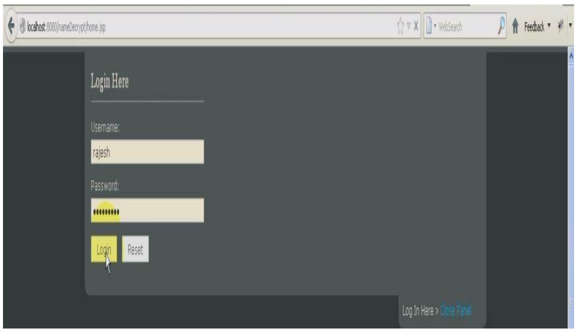
V. Discussion

Data is the major resource of every business. So, keen attention has to be paid towards data storage. However, storing the daily growing data is not that easy, since the small and medium scale business cannot afford much to data storage alone.

The best way to recover from data loss due to a natural disaster is to store backup data in a physically separate location. Thus, they go for another option named 'Outsourcing'. When the data is outsourced to a public space, then some amount has to be paid for the memory consumed by the data owner.

Remote storage of important information is made possible only through outsourcing. However, searching and utilization of the outsourced data is an important concern. This system employs multi keyword and ranked search through the most suitable results are listed out and thus the overall performance is enhanced.

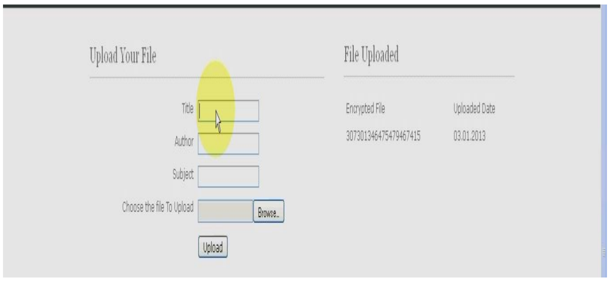
The following screen shows the implementation of our proposed work. We have implemented and tested with a system configuration on Intel Dual Core processor, Windows XP and Using Netbeans 7.0. The details of each module for this system are as follows:



Username: rajesh

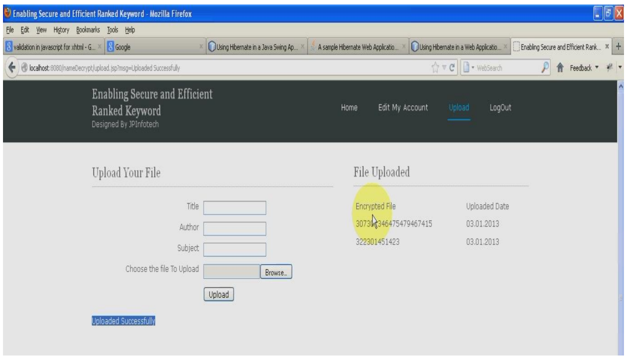
Password: ●●●●●●●●●

Fig 4: Cloud User Login



Encrypted File	Uploaded Date
307301346475479467415	03.01.2013

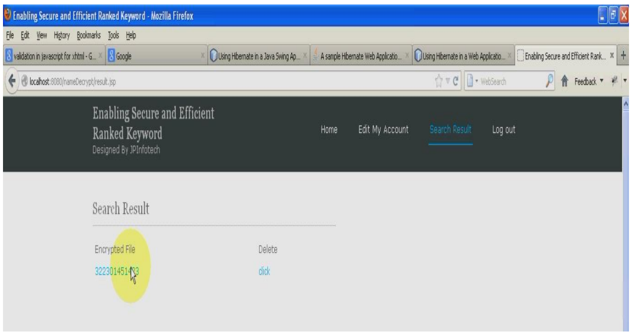
Fig 5: Cloud User Upload page


Enabling Secure and Efficient Ranked Keyword - Mozilla Firefox
File
Edit
View
History
Bookmarks
Tools
Help
validation in javascript for xhtml - G...
Google
Using Hibernate in a Java Swing Ap...
A sample Hibernate Web Applicatio...
Using Hibernate in a Web Applicatio...
Enabling Secure and Efficient Rank...
localhost:8080/nameDecrypt/upload.jsp?msg=Uploaded Successfully
WebSearch
Feedback
Enabling Secure and Efficient
Ranked Keyword
Designed By JPInfotech
Home
Edit My Account
Upload
LogOut
Upload Your File
Title
Author
Subject
Choose the file To Upload
Browse..
Upload
File Uploaded

Encrypted File	Uploaded Date
307301346475479467415	03.01.2013
322301451423	03.01.2013

Uploaded Successfully

Fig 6: After Upload page


Enabling Secure and Efficient Ranked Keyword - Mozilla Firefox
File
Edit
View
History
Bookmarks
Tools
Help
validation in javascript for xhtml - G...
Google
Using Hibernate in a Java Swing Ap...
A sample Hibernate Web Applicatio...
Using Hibernate in a Web Applicatio...
Enabling Secure and Efficient Rank...
localhost:8080/nameDecrypt/result.jsp
WebSearch
Feedback
Enabling Secure and Efficient
Ranked Keyword
Designed By JPInfotech
Home
Edit My Account
Search Result
Log out
Search Result

Encrypted File	Delete
322301451423	click

Fig 7: Encrypted file details

VI. Conclusions

This work exploits the multi-keyword and ranked search of outsourced data. It ensures security by means of encryption and privacy preservation, through email activation code that is, only intended user can gain access to the information and others cannot. Using this work, an effective search experience is gained by the client. Thus, this project makes searching and utilization of outsourced data simpler.

VII. References

[1] Kamara S. and Lauter K. (2010) ‘Cryptographic cloud storage’, in RLCPS, LNCS Heidelberg.

[2] Li J., Wang Q., Wang C., Cao N., Ren K. and Lou W., (2010) ‘Fuzzy keyword search over encrypted data in cloud computing’, in Proc. of IEEE INFOCOM’10 Mini Conference, San Diego, CA, USA.

[3] Wang C., Cao N., Li J., Ren K. and Lou W. (2010) ‘Secure ranked keyword search over encrypted cloud data’, in Proc. of ICDCS’10.

[4] Vaquero L.M., Rodero Merino L., Caceres J. and Lindner M. (2009) ‘A break in the clouds: towards a cloud definition’, ACM SIGCOMM Comput Commun. Rev., vol. 39, no. 1, pp. 50–55.

[5] Abdalla M., Bellare M., Catalano D., Kiltz E., Kohno T., Lange T, MaloneLee J., Neven G., Paillier P. and Shi H. (2008) ‘Searchable encryption revisited consistency properties, relation to anonymous ibe, and extensions’, J. Cryptol., vol. 21, no. 3, pp. 350–391.

[6] Bellare M., Boldyreva A. and ONeill A. (2007) ‘Deterministic and efficiently searchable encryption’, in Proc. of CRYPTO.

[7] Boneh D. and Waters B. (2007) ‘Conjunctive, subset, and range queries on encrypted data’, in Proc. of TCC, pp. 535–554.

[8] Brinkman R. (2007) ‘Searching in encrypted data’, in University of Twente, PhD thesis.

[9] Curtmola R, Garay J.A., Kamara S. and Ostrovsky R.(2006) ‘Searchable symmetric encryption: improved definitions and efficient constructions’, in Proc. of ACM CCS.

[10] Chang Y.C. and Mitzenmacher M. (2005) ‘Privacy preserving keyword searches on remote encrypted data’, in Proc. of ACNS.

[11] Boneh D., Crescenzo G.D., Ostrovsky R. and Persiano G. (2004) ‘Public key encryption with keyword search’, in Proc. of EUROCRYPT.

[12] Golle P., Staddon J. and Waters B. (2004) ‘Secure conjunctive keyword search over encrypted data’, in Proc. of ACNS, pp. 31–45.

[13] Goh E.J.(2003) ‘Secure indexes’, Cryptology ePrint Archive.

[14] Singhal A. (2001) ‘Modern information retrieval: A brief overview’, IEEE Data Engineering Bulletin, vol. 24, no. 4, pp. 35–43.

[15] Song D., Wagner D. and Perrig A. (2000) ‘Practical techniques for searches on encrypted data’, in Proc. of S&P.

[16] Witten I.H., Moffat A. and Bell T.C. (1999) ‘Managing gigabytes: Compressing and indexing documents and images’, Morgan Kaufmann Publishing, San Francisco.