Comparison of Privacy Preserving Single-Keyword Search and Multi-Keyword Ranked Search Techniques over Encrypted Cloud Data

Madane S.A.
M.B.E.Society's College of Engineering.Ambajogai, Maharashtra.

B.M. Patil
M.B.E.Society's College of Engineering, Ambajogai, Maharashtra.

ABSTRACT
Cloud computing otherwise known as on demand computing. It provides the services over the internet. It has the provision of facilitating users to store and access their data in and from cloud server by sitting anywhere and on any device. Storing data in cloud server also opens up so many security threats as data is accessed over internet and client has no direct control over data once uploaded into cloud server. We first implement a basic idea for the Single Keyword Search Over Encrypted Data And then Multi-keyword Ranked. Search over Encrypted cloud data (MRSE) based on secure inner product computation and efficient similarity measure of coordinate matching, i.e., as many matches as possible, in order to capture the relevance of data documents to the search query, then we give two significantly improved MRSE schemes to achieve various stringent privacy requirements in two different threat models. Assignment of anonymous ID to the user to provide more security to the data on cloud server is done. To improve the search experience of the data search service, further extension of the two schemes to support more Search semantics is done.[5]

Keywords
Cloud computing, Encryption, Inner product similarity, Single Keyword Search, Multi-keyword search, ranking.

1. INTRODUCTION
Cloud computing is being intensively referred to as one of the most prominent innovations in information technology in recent epoch. By using resource virtualization cloud delivers us computing resources and services in a pay-as-you-go mode. Today world is moving on digitization and cloud computing is best concept to handle big datasets. Various cloud computing services are categorize into Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and last one is Software-as-a-Service (SaaS)[13].

Cloud computing is the elongated dreamed hallucination of computing as a efficacy, where cloud customers tenuously stock up their data into the cloud so as to take pleasure in the on-order far above the ground-eminence application and services from a public pool of configurable computing resources. Its great plasticity and financial savings are rousing both folks and enterprise to outsource subcontract their local multifaceted data management system into the cloud. To guard privacy of data and be in opposition to unwelcome accesses in the cloud and further than it, susceptible data, for illustration, e-mails, personal health records, photo albums, tax documents, and so on, may have to be encrypted by data owners before Outsourcing to the commercial public cloud; this, however, obsoletes the conventional data employment service based on plaintext keyword investigate. The irrelevant way out of downloading all the data and decrypting locally is obviously unreasonable, due to the big quantity of bandwidth cost in cloud scale systems. Images also be full of practical and vital information, so anticipated system also provides image cataloging in MRSE scheme [1]. Moreover, aside from eliminating the local storage management, storing data into the cloud doesn't hand out any point except they can be with no trouble searched plus utilized. Hence, explore privacy preserving and effective search service over encrypted cloud data is of immense consequence. Ranked search can also stylishly get rid of needless network traffic by sending back only the majority germane data, which is exceedingly enviable in the "pay-as-you-use" cloud paradigm. For privacy protection, such ranking process, however, should not leak any keyword related information. Besides, to improve search result accuracy as well as to enhance the user searching experience, it is also necessary for such ranking system to support multiple keyword search, as single keyword rummage around often yields far too coarse results. As a common practice indicate by today's web search engines (ex. Google search), data users may tend to provide a set of keywords instead of only one as the indicator of their search interest to retrieve the most applicable data. Along with the privacy of data and efficient penetrating schemes, real privacy is obtained only if the user's identity remains hidden from the Cloud Service Provider (CSP) as well as the third party user on the cloud server[6].

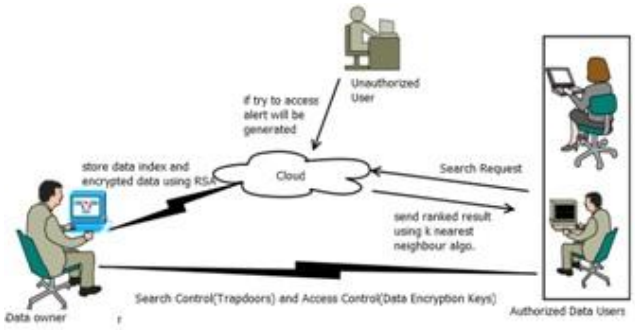

Fig1. Architecture of the search over encrypted cloud data.

The Three main donations of our planned work are described as follows

1. Multi keyword rank search
2. Single keyword search
3. Encryptions of data with AES
4. Cloud setup.
5. Comparision of Single keyword search and MRES Technic

2. LITERATURE SURVEY

2.1 Secured Multi-keyword Ranked Search over Encrypted:

In cloud compute data possessor are aggravated to farm out their multifaceted data organization systems from local sites to the marketable public cloud for superior give and economic savings. To make sure safety of stored data, it is have to to encrypt the data before storing. the cloud. In these existing systems the algorithms used are cryptographic [1].

2.2 Privacy Preserving Keyword Searches on Remote:

Encrypted Data: Consider the problem: a user *U* wants to stock up his files in an encrypted form on a far-flung file server *S*. afterward the user *U* wants to professionally get back some of the encrypted files contain exact keywords, keeping the keywords themselves clandestine and not to cause danger to the security of the tenuously store files. For example, a user may want to store old e-mail post encrypted on a server manage by Yahoo or one more large vendor, and later regain certain messages while travelling with a mobile device. In [2], solutions for this problem under well-defined safety requirements are obtainable.

2.3 Efficient and Secure Multi-Keyword Search on Encrypted Cloud Data:

On one pass, users who do not unavoidably have priorknowledge of the encrypted cloud data, have to post process every retrieved file in arrange to find ones most matching their interest; On the other hand, invariably retrieving all files containing the query keyword further incur unnecessary network traffic, which is absolutely undesirable in today's pay-as-you-use cloud archetype. This paper has definite and solved the problem of effectual yet safe and sound rank keyword search over Encrypted cloud data [2]. Ranked search greatly enhances system usability by returning the matching files in a ranked order regarding to certain relevance criteria (e.g., keyword frequency) thus making one step closer towards sensible consumption of privacy-preserving data hosting services in Cloud Computing. For the first time, the paper has defined and solved the challenging problem of privacy-preserving multi-keyword ranked search over encrypted cloud data (MRSE)[2], and establish a set of strict privacy requirements for such a protected cloud data utilization system to become a reality. The proposed ranking method proves to be efficient to go back extremely relevant documents corresponding to submitted search terms. The idea of proposed ranking method is used in our future system in order to enhance the security of information on Cloud Service Provider.

2.4 Providing Privacy Preserving in Cloud Computing:

Privacy is an important issue for cloud computing, both in terms of legal compliance and user trust and needs to be considered at every phase of plan. The [5] paper tells the Importance of protecting individual's privacy in cloud computing and provides some seclusion preserving technologies used in cloud computing services. Paper tells that it is very important to take privacy into account while designing cloud services, if these engage the collection, processing or sharing of personal data. From this paper, main theme taken is of preserving privacy of data. This paper only describes privacy of data but doesn't let indexed search as well as doesn't hide user's identity. Thus, these two drawbacks are overcome in our wished-for system.

2.5 Privacy Preserving Data Sharing With Anonymous ID obligation:

In this paper, an algorithm for mysterious sharing of private data among N parties is developed. This technique is used iteratively to dispense these nodes ID numbers ranging from 1 to N. This assignment is anonymous in that the identities received are strange to the other members of the group. In [6], obtainable and new algorithms for assigning anonymous IDs are examined with respect to trade-offs between communication and Computational requirements. These new algorithms are built on top of a secure sum data mining process using Newton's identities and Sturm's theorem. The main idea taken from this paper is of assigning unidentified ID to the consumer on the cloud

2.6 Single Keyword Search Over Encrypted data on cloud:

Obtainable searchable encryption scheme consent to a user to firmly look for over encrypted data through keywords without first decrypting it, these techniques support only conventional Boolean keyword search, without capturing any relevance of the files in the search result. When directly applied in large joint data outsourcing cloud environment, they go through next shortcoming.

Drawbacks of Single Keyword Search system*:*

1. Single-keyword search without ranking
2. Boolean- keyword search without ranking
3. Single-keyword search with ranking
4. Do not get relevant data.

In Fig1.There is one another unit is shown i.e. unlawful User. If that Unauthorized user try to access any data from clod then attentive will be generate in the form of mail and message. The alert is given to the authorized person who is owner of that data.

3. IMPLEMENTATION DETAILS

3.1 MRES System

For our organism, we choose the attitude of harmonize matching, to identify the correspondence amid search inquiry and data credentials. Particularly, we use internal data correspondence, i.e., the figure of query keywords appearing in a document, to appraise the similarity of that document to the search query in coordinate matching principle. Each document is connected with a binary vector as a sub index where each bit represent whether analogous keyword is contained in the document [6] The search reservation is also describe as a dual vector where each bit means whether corresponding keyword appears in this search request, so the similarity could be exactly measured by inner product of query vector with information vector. However, directly outsourcing data vector or query vector will infringe index privacy or search privacy.

To improve document retrieval accuracy, search result should be ranked by cloud server according to some ranking criteria. Cloud server only sends back top-k documents that are most relevant to the search query.

In the wished-for organism the stream starts from the user. The user has to register in CSP to get the amenities. Once user data is stored in CSS it has no unswerving control above it. User has to hire any auditor called TPA who will very regularly check the user data in CSS[. The TPA should be granted by the user to check the integrity for a specific data and for a unambiguous time without accessing the exact data. Below an algorithm is provide which describe how the TPA does the audit.

AES algorithm is used to store the data in encrypted form in cloud server. So when TPA does the audit it only gets the false impression of original files. The values on which TPA calculates or check the integrity is actually the hash value of encrypted file calculated cooperatively. The TPA is only allowed to check the uprightness nothing else. It checks the integrity for the first time then checks whether uprightness potted and lastly check whether integrity remains or lost. User can any time grant or revoke the concession from TPA. User has the privilege to upload, download and edit data. User's edit request is also served for a fastidious part of the file instead of retrieve the whole file[6].

AES Algoritham:

Notation and Definitions

AES(K, W)	Encrypt W using the AES codebook with key K
AES-1(K, W)	Decrypt W using the AES codebook with key K
MSB(j, W)	Return the most significant j bits of W
LSB(j, W)	Return the least significant j bits of W
B1 | B2	Concatenate B1 and B2
K	The key-encryption key K
s	The number of steps in the wrapping process, = 6n
P[i]	The ith plaintext key data block
C[i]	The ith cipher text data block
A	The 64-bit integrity check register
R[i]	An array of 64-bit registers where i = 0, 1, 2, ..., n

Algorithm

1. The set of round keys from the cipher key.
2. Initialize state array and add the initial round key to the starting state array.
3. Perform round = 1 to 9 : Execute Usual Round.
4. Execute Final Round.

5. Corresponding cipher text chunk output of Final Round Step

ii.Usual Round Execute the following operations which are described above.
1.Sub Bytes
2. Shift Rows
3.MixColumns
4.Add Round Key , using K(round)

iii.Final Round:
Execute the following operations which are described above.
1.Sub Bytes
2.Shift Rows
3. Add Round Key, using K(10)

In the key wrap algorithm, the concatenation occupation will be used to concatenate 64-bit quantity to form the 128-bit input to the AES codebook. The pulling out functions will be used to split the 128-bit amount produced from the AES codebook into two 64-bit quantities.

The arrangement of the key enfold algorithm requires the use of the AES codebook [AES]. The next sections will describe the key envelop algorithm, the key unwrap algorithm, and the inherent data integrity check.

Algorithm Key Wrap:

1) Sub Bytes : The first transformation, Sub Bytes, is used at the encryption site. To substitute a byte,
we interpret the byte as two hexadecimal digits.
2) Shift Rows : In the encryption, the transformation is called Shift Rows.
3) Mix Columns : The Mix Columns transformation operates at the column level; it transforms each
column of the state to a new column.
4) Add Round Key : Add Round Key proceeds one column at a time.

The inputs to the key wrapping procedure are the KEY and the plaintext to be wrapped. The plaintext consists of n 64-bit blocks, contain the key data life form wrapped. The key covering process is described beneath.

An substitute explanation of the key wrap algorithm involves indexing quite than shifting. This approach allows one to calculate the wrap key in place, avoiding the rotation in the previous account. This produces the same results and is more easily implement in software.

4. IMPLEMENTATION RESULT

In this secation we present comparision result of Single Key word Search Ranked search and Multi Keyword Ranked Search Over A Encrypted Data On Cloud as shown in following figers .In this Result Exch Ranked search and Multi Keyword Ranked Search Over A Encrypted Data On Cloud.In this Result Existing System is Single Keyword Search System and proposed System is nothing but MRES System

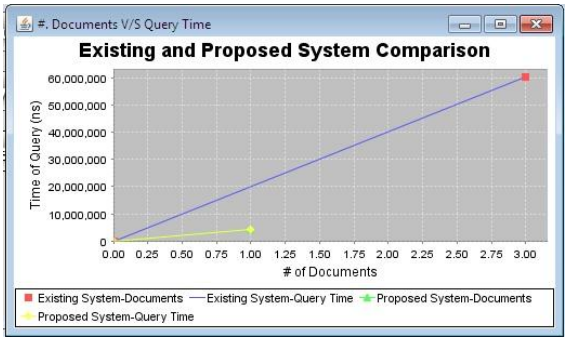

Fig 2: Comparison Graph- No. Of Documents V/S Query Time

Fig 2 is a Comparison graph of Existing System and Our System. The graph is plotted Number of Documents that the respective system's search result returned and Time required to return the documents; in respective System. As shown in the graph Our system requires less time which is less than around 5 ns with most specific result of number of document equal to one which is less than the documents returned by existing system which are three and time required is around 6ns

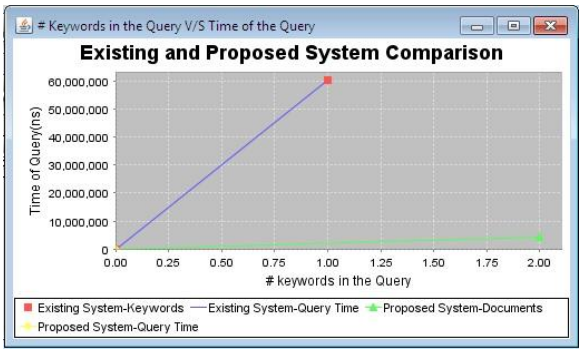

Fig 3: Comparison Graph- No. Of Keywords V/S Query Time

Fig 3 is a Comparison graph of Existing System and Our implemented System. The graph is plotted against Number of Keywords fired in the respective system's search and Time required in respective System. As shown in the graph Our system requires less time which is less than around 5 ns with multiple Keyword Query and existing system requires around 6ns even though a single Keyword query is fired. So Our System Works Better in each and every aspect then existing System.

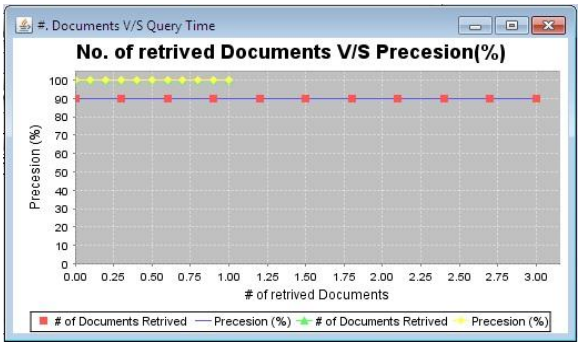

Fig 3: No. Of Retrieved Document V/S Precision.

Fig 4 shows The Comparison graph of Existing System and Our implemented System plotted against No of Documents returned by respective system V/S percentage Precision(Perfectness). Our system gives much better precession than existing system as shown in graph.

In this method the major merits are: (1) data security (2) privacy shield (3) Auditing details to the data owner (4) Audit aptitude aware data scheduling at this time we are going to evaluate the performance of our projected scheme in terms of the computation overhead introduce by each operation. Request and resources are taken as the computing parameter. When the number of requests increase at the same time, it is to check whether they are served within a particular time. The waiting time is measured for each request.

5. CONCLUSION AND FUTURE WORK:

In this document, for the primary occasion we term and crack the problem of multi-keyword ranked search in excess of encrypted cloud data, and institute a assortment of privacy necessities. in the midst of various multi-keyword semantics, we choose the efficient similarity measure of "coordinate matching," i.e., as many matches as possible, to effectively capture the relevance of outsourced documents to the query keywords, and use "internal product similarity" to quantitatively evaluate such similarity measure. For meeting the challenge of supporting multi-keyword semantic without privacy breaches, we propose a basic idea of MRSE using secure inner product computation. Then, we give two improved MRSE schemes to achieve various severe privacy requirements in two different threat models. We also investigate Some further enhancement of our ranked search mechanism, including supporting more search semantics, i.e., TF _ IDF, and dynamic data operations. Thorough analysis investigating privacy and efficiency guarantees of proposed schemes is given, and experiments on the real-world dataset show our proposed schemes introduce low overhead on both computation and communication.

In future we have to improve more on security issues of data storage on cloud storage service. On cloud computing this topic is not negotiable to improve. For implementing that process we increase the layers of authentications.In our future work, we will travel around checking theintegrity of the rank order in the search result assumingthe cloud server is untrusted

6. ACKNOWLEDGMENT

I would like to express my thanks to my guide Prof. Patil B.M. for his highly appreciable support and encouragement also for guidance is a force behind the completion of this paper. I am grateful for all the suggestions and hints provided by him. My acknowledgment of gratitude to all who supported to make it possible.

7. REFERENCES

[1] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data," Proc.IEEE INFOCOM, pp. 829-837, Apr, 2011.

[2] Secure Ranked Keyword Search over Encrypted Cloud Data , IEEE PAER, 2010.

[3] International Journal of Emerging Technology and Advanced Engineering Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 4, Issue 2, February 2014).

[4] Privacy preserving public auditing for Secure Cloud Storage", Cong Wang, Sherman S.-M. Chow, Qian Wang, Kui Ren.

[5] Shiba Sampat Kale et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (6) , 2014, 7093-7096

[6] Kuchi Ravi Kishore, et al International Journal of Computer and Electronics Research [Volume 4, Issue 2, April 2015]

[7] Li, S. Yu, N. Cao, and W. Lou. Authorized private keyword search over encrypted data in cloud computing. In Distributed Computing Systems (ICDCS), 2011 31st International Conference on, pages 383–392. IEEE, 2011

[8] Boldyreva, N. Chenette, Y. Lee, and A. O'Neill. Order-reserving symmetric encryption. In Proceedings of Eurocrypt'09, volume 5479 of LNCS. Springer, 2009.

[9] Wang, K. Ren, S. Yu, K. Mahendra, and R. Urs, "Achieving Usable and Privacy-Assured Similarity Search over Outsourced Cloud Data," Proc. IEEE INFOCOM, 2012.

[10] Boldyreva, N. Chenette, Y. Lee, and A. O'Neill. Order-preserving symmetric encryption. In Proceedings of Eurocrypt'09, volume 5479 of LNCS. Springer,

[11] International Journal of Advance Research, IJOAR.org Volume 3, Issue 2, February 2015, Online: ISSN 2320-9194