Secure And Optimized Cloud Storage With Efficient Search And User Revocation Features

Rakeshreddy Gurrala[1], Srinivas Rao Adabala[2]
[1, 2] Dept of Information Technology
[1, 2] JNTUH/Keshav Memorial Institute of Technology

Abstract- *Cloud computing has been helping organizations to outsource their bulk of data to cloud. Storage is one of the widely used facilities used by cloud users. Though cloud service provider takes care of security dynamics of outsourced data, many security concerns remain in the minds of users. Secure search and retrieval of data are to be given high importance. Cloud data is stored in encrypted format. In such cases, it is not possible to perform search operations easily when traditional retrieval approaches are followed. Towards solving this problem, many researchers contributed in cloud computing arena. Recently Xia et al. proposed a multi-keyword search for flexible retrieval of encrypted data that has been outsourced to cloud. In this paper a suitable methodology is proposed in order to have optimized search with user revocation features. We built a model that supports dynamic queries. The application makes use of TF/IDF in order to have efficient search results. The application demonstrates the utility of the proposed methodology. The results showed that the application is useful to store and retrieve user data securely.*

Keywords- Cloud data storage, encrypted cloud data, outsourcing, searching on encrypted cloud data, user revocation

I. INTRODUCTION

Cloud computing has become a reality and organizations are outsourcing their data to cloud for having different services. When data is outsourced it may be subjected to theft or any attack. Therefore the owners of the data are supposed to encrypt data before outsourcing it. This approach can protect data from attacks. However, the search operations become difficult as the data is encrypted and stored. Therefore it is essential to have mechanism for having multi-keyword ranked search. The conceptual overview of the approach is shown in Figure 1

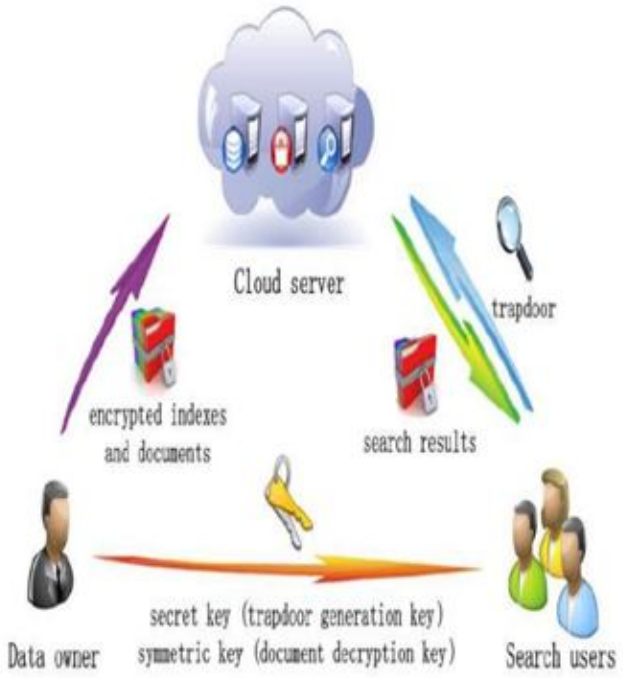

Figure 1: Overview of Scenario

The outsourced data is encrypted and stored in cloud servers. The data is indexed and the encrypted files are saved in cloud. The users can make search request and gain ranked results. In this paper we proposed a multi-keyword ranked search fine grained approach for efficient query processing. Many schemes came into existence as found in literature. When cloud data is outsourced, it is encrypted before sending to cloud for security reasons. Many researchers

Contributed towards performing search on encrypted outsourced data into cloud as explored in [1]-[25] where different techniques are found. The remainder of the paper is structured as follows. Section II provides review of literature. Section III presents the proposed system in detail. Section IV presents experimental results while section V concludes the paper and provides recommendations for future work.

II. RELATED WORKS

Encryption schemes with search capability are very important in the context of cloud computing where data is outsourced to cloud. When data is encrypted, it is converted into un-understandable format. This format is used to store in cloud storage. Searchable encryption schemes are explored in [1] and [2]. There are many techniques that make use of symmetric cryptography. They are explored in [3], [5], [6], and [8]. The first searchable encryption with symmetric keys is found in [3]. The search time is found to be linear based on the size of data collection. Goh [4] explored definition and research on the SSE which was based on the concept known as Bloom Filter. Their scheme achieved O(n) in terms of performance. The cardinality of data collection is denoted by n. There are two schemes found in [6]. They are known as SSE-1 and SSE-2 for providing optimal search time performance. The former is able to withstand CKA1 attacks known as chosen-keyword attack. The latter is better to withstand other kinds of chosen attacks known as CKA2.

All these researches focused on the single-keyword Boolean search. After this many researchers as explored in [7], [8], [9], and [10] focused on similarity search with single keyword. A Boolean search with multiple keywords with ranked search capabilities is found in [22], [23], [24], and [25]. This kind of keywords provides flexible search mechanisms to have resultant documents. There are other search mechanisms known as conjunctive keyword search schemes explored in [15], [12], and [13]. These schemes returned the documents that contained all keywords. In this paper we explored flexible and optimized multi-keyword search with user revocability.

III. WRITE DOWN YOUR STUDIES AND FINDINGS

Here is our methodology which provides mechanisms for encryption and decryption of docs for secure outsourcing. It also provides details of data user, data owner and their respective activities. Data owner and data user are the two roles involved in the system. The former is the owner of data who is responsible to encrypt and protect data from unauthorized access while the latter takes permissions or decryption keys from the data owner in order to gain access to legitimate data. Data owner provides decryption keys to data users. There might be multiple data users to whom permissions are granted in order to have top-k multi-keyword ranked query results.

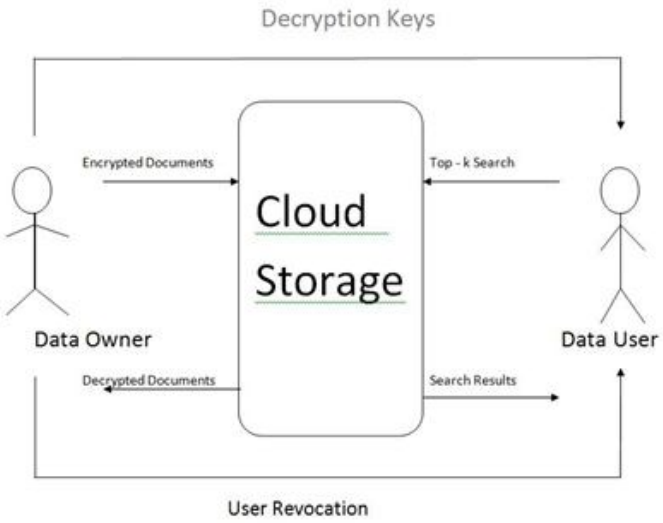

Figure 2: Proposed methodology for optimized search on encrypted cloud data

As shown in Figure 2, it is evident that the cloud server is used by users of two roles. They are data owner and data user. Data owner performs encryption on the files to be outsourced to cloud. Once files are encrypted, they are sent to cloud. The data owner can access the files and perform operations on them as well. The data owner needs to give access rights to data users. Then data users can gain access to required data using search operations. Generally top-k ranked search is performed by data users. Internally the proposed system makes use of TF/RDF for every file in order to have search operations faster. When any user is no longer associated with data owner for any reason, that user is revoked by taking away permissions given.

IV. IMPLEMENTATION

We built a prototype application using Microsoft .NET platform. The proposed system supports multi-keyword ranked search on encrypted cloud data. The application has functionalities that are associated with two roles namely cloud service provider and user. The activities of both the users are shown in Figure 3.

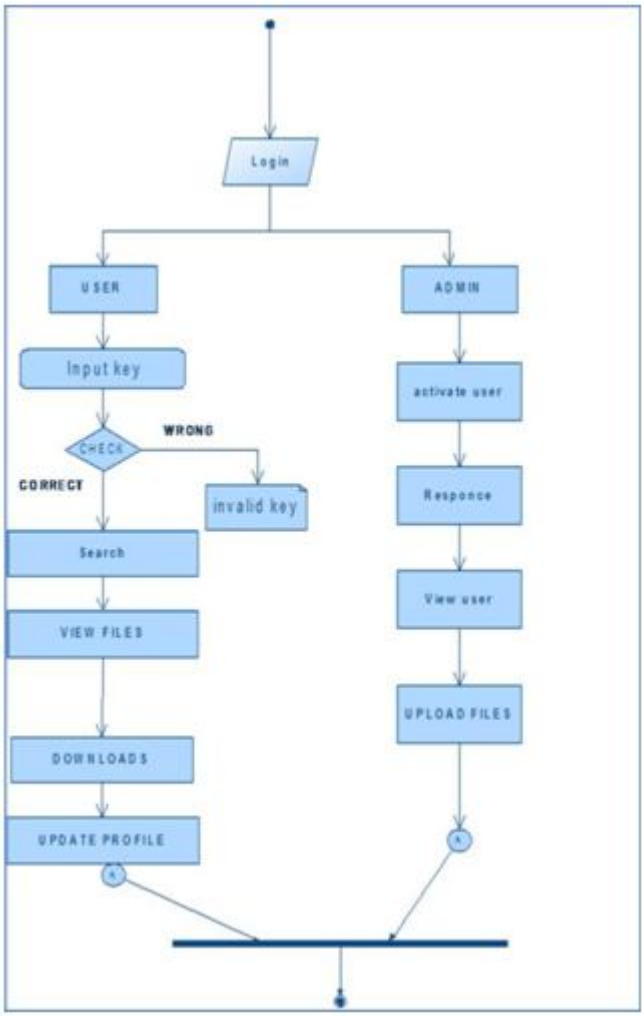

Figure 3: Shows Activities of Two Roles

As shown in Figure 3, the user can have registration, authentication, uploading data and making multi-keyword ranked search. The cloud service provider can have access to system functionalities such as key generation; viewing user details viewing file details, and performing encryption and decryption.

V. EXPERIMENTAL RESULTS

We built the prototype application using Microsoft .NET platform. The technologies used are ASP.NET for designing and implementing web based application, C# for coding functionality, ADO.NET for interacting with relational databases and SQL Server for storing data permanently and help in data manipulations. The functionalities of the system are divided into two roles for which users exist. The two roles include administrator and user.

EDMRS	BSMRS
0.3	0.2
0.4	0.3
0.5	0.4
0.6	0.5
0.8	0.7
1	0.9
1.3	1.1

Table 1: Performance comparison of schemes

As shown in Table 1, it is evident that the two schemes are tested for finding the time taken for trapdoor generation and presented with different number of keywords in the dictionary.

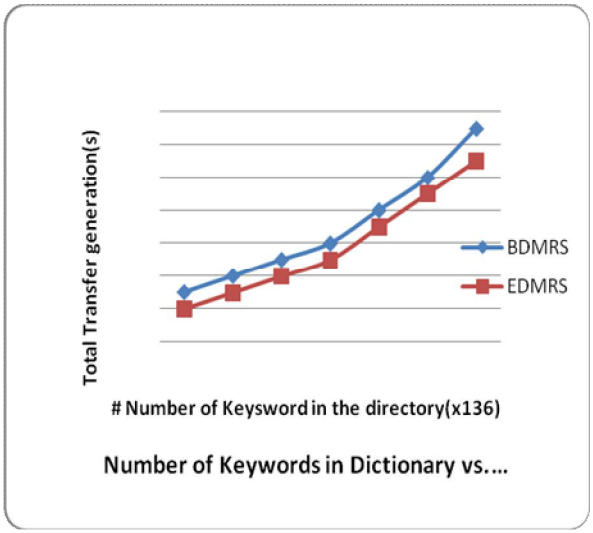

Figure 4: Number of Keywords in Dictionary vs. Time Taken for Trapdoor Generation

As shown in Figure 4, it is evident that the two schemes are tested for finding the time taken for trapdoor generation and presented with different number of keywords in the dictionary.

Table 2: Performance comparison based on number of threads
As shown in Table 2, the two schemes are tested with different number of threads and number of keywords in the dictionary.

Number of Documents in the Collection	EDMRS with 16 threads	EDMRS with 8 threads	EDMRS with 4 threads	Sun-1 st level	BD MR S	ED MR S
1	9	9	10	25	25	26
2	14	15	18	30	30	31
3	16	20	24	33	35	36
4	18	20	23	29	35	36
5	20	25	30	29	50	51
6	22	25	50	27	60	61
7	25	30	40	25	65	66
8	25	30	40	23	70	71

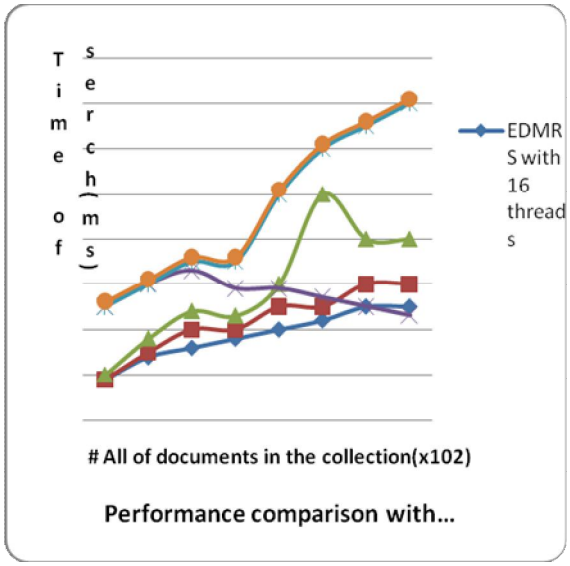

Figure 5: Performance comparison with number of documents in the collection.

As shown in Figure 5, the two schemes are tested with Different number of threads and number of keywords in the dictionary.

# of docs in collection	RDRMS	BDRMS
1	4.5	5
2	5	5.2
3	5.5	6.1
4	5.5	6.1
5	5.5	6.2
6	5.9	6.8
7	5.9	6.8
8	6	6.8

Table 3: Number of Keywords in the Dictionary vs. Performance of Schemes

As shown in Table 3, it is evident that the two schemes showed different performance when number of keywords in the dictionary is increased gradually for time taken to delete documents.

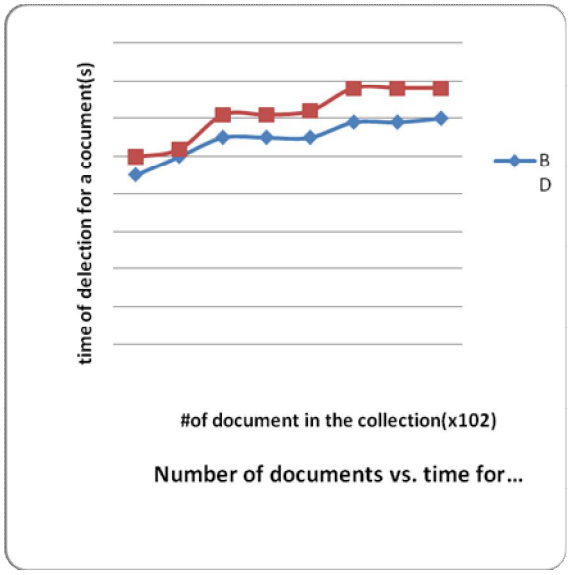

Figure 6: Number of documents vs. time of deletion.

As shown in Table 6, it is evident that the two schemes showed different performance when number of keywords in the dictionary is increased gradually.

Number of documents in the collection	BDMRS	EDMRS
1	4.5	5
2	5	5.2
3	5.5	6.1
4	5.5	6.1
5	5.5	6.2
6	5.9	6.8
7	5.9	6.8
8	6	6.8

Table 4: Number of Keywords in the Dictionary vs. the Performance of Both Schemes

As shown in Table 4, it is evident that the time of deletion of a document varies for the two schemes when the number of keywords in the dictionary is changed.

As shown in Table 7, it is evident that the time of deletion of a document varies for the two schemes when the number of documents in the collection is changed.

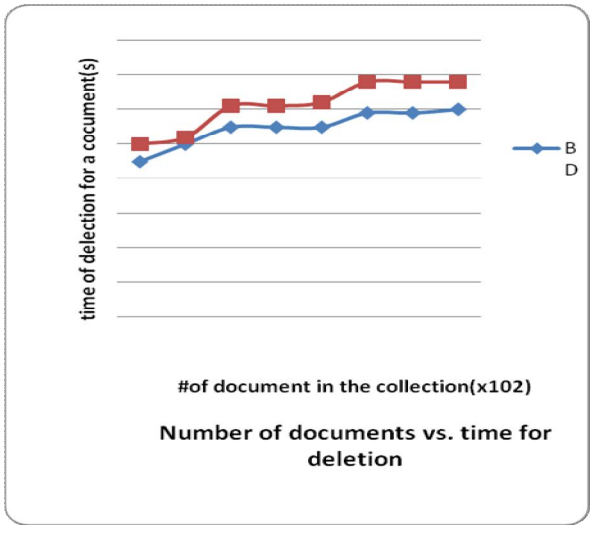

Figure 7: Number of documents vs. time for deletion

VI. CONCLUSONS AND FUTURE WORK

In this paper, we focused on the secure cloud storage and retrieval which is very important in the wake of widespread usage of cloud computing technologies across the globe. Cloud data users often store data in encrypted format. This is the main concern while retrieving data. Data dynamics and search operations become difficult when outsourced data needs to be retrieved. Moreover, the security of outsourced data is another important concern. To overcome these drawbacks many researchers proposed different techniques. The data owner performs encryption and sends data to cloud. Then the data owner and data user should have coordination and ability to perform search operations as per the permissions given to users. In this paper, we focused on creation of a methodology that allows efficient search on the encrypted data. We built a prototype to show this. The results showed significant improvement in the access control and also the efficiency in the retrieval of outsourced data securely. In future we provide more importance to validating our approach and underlying methodology for effective data storage and retrieval in cloud.

REFERENCES

[1] D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, “Public key encryption with keyword search,” in Advances in Cryptology- Eurocrypt 2004. Springer, 2004, pp. 506–522.

[2] D. Boneh, E. Kushilevitz, R. Ostrovsky, and W. E. Skeith III, “Public key encryption that allows pir queries,” in Advances in Cryptology-CRYPTO 2007. Springer, 2007, pp. 50–67.

[3] D. X. Song, D. Wagner, and A. Perrig, “Practical techniques for searches on encrypted data,” in Security and Privacy, 2000. S&P 2000. Proceedings. 2000 IEEE Symposium on. IEEE, 2000, pp. 44– 55.

[4] E.-J. Goh et al., “Secure indexes.” IACR Cryptology ePrint Archive, vol. 2003, p. 216, 2003.

[5] Y.-C. Chang and M. Mitzenmacher, “Privacy preserving keyword searches on remote encrypted data,” in Proceedings of the Third international conference on Applied Cryptography and Network Security. Springer-Verlag, 2005, pp. 442–455.

[6] R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, “Searchable symmetric encryption: improved definitions and efficient constructions,” in Proceedings of the 13th ACM conference on Computer and communications security. ACM, 2006, pp. 79–88.

[7] J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, “Fuzzy keyword search over encrypted data in cloud

computing," in INFOCOM, 2010 Proceedings IEEE. IEEE, 2010, pp. 15.

[8] M. Kuzu, M. S. Islam, and M. Kantarcioglu, "Efficient similarity search over encrypted data," in Data Engineering (ICDE), 2012 IEEE 28th International Conference on. IEEE, 2012, pp. 1156–1167

[9] C. Wang, K. Ren, S. Yu, and K. M. R. Urs, "Achieving usable and privacy-assured similarity search over outsourced cloud data," in INFOCOM, 2012 Proceedings IEEE. IEEE, 2012, pp. 451–459.

[10] B. Wang, S. Yu, W. Lou, and Y. T. Hou, "Privacy-preserving multikeyword fuzzy search over encrypted data in the cloud," in IEEE INFOCOM, 2014.

[11] P. Golle, J. Staddon, and B. Waters, "Secure conjunctive keyword search over encrypted data," in Applied Cryptography and Network Security. Springer, 2004, pp. 31–45.

[12] Y. H. Hwang and P. J. Lee, "Public key encryption with conjunctive keyword search and its extension to a multi - user system," in Proceedings of the First international conference on Pairing-Based Cryptography. Springer-Verlag, 2007, pp. 2–22.

[13] L. Ballard, S. Kamara, and F. Monrose, "Achieving efficient conjunctive keyword searches over encrypted data," in Proceedings of the 7th international conference on Information and Communications Security. Springer-Verlag, 2005, pp. 414–426.

[14] D. Boneh and B. Waters, "Conjunctive, subset, and range queries on encrypted data," in Proceedings of the 4th conference on Theory of cryptography. Springer-Verlag, 2007, pp. 535–554.

[15] B. Zhang and F. Zhang, "An efficient public key encryption with conjunctive-subset keywords search," Journal of Network and Computer Applications, vol. 34, no. 1, pp. 262–267, 2011.

[16] J. Katz, A. Sahai, and B. Waters, "Predicate encryption supporting disjunctions, polynomial equations, and inner product s," in Advances in Cryptology–EUROCRYPT 2008. Springer, 2008, pp.146–162.

[17] E. Shen, E. Shi, and B. Waters, "Predicate privacy in encryption systems," in Proceedings of the 6th Theory of Cryptography Conference on Theory of Cryptography. Springer-Verlag, 2009, pp. 457–473.

[18] A. Lewko, T. Okamoto, A. Sahai, K. Takashima, and B. Waters, "Fully secure functional encryption: attribute-based encryption and(hierarchical) inner product encryption," in Proceedings of the 29th Annual international conference on Theory and Applications of Cryptographic Techniques. Springer-Verlag, 2010, pp. 62–91.

[19] A. Swaminathan, Y. Mao, G.-M. Su, H. Gou, A. L. Varna, S. He, M.Wu, and D.W. Oard, "Confidentiality-preserving rank-ordered search," in Proceedings of the 2007 ACM workshop on Storage security and survivability. ACM, 2007, pp. 7–12.

[20] S. Zerr, D. Olmedilla, W. Nejdl, and W. Siberski, "Zerber+ r: Topk retrieval from a confidential index," in Proceedings of the 12th International Conference on Extending Database Technology: Advances in Database Technology. ACM, 2009, pp. 439–449.

[21] C. Wang, N. Cao, K. Ren, and W. Lou, "Enabling secure and efficient ranked keyword search over outsourced cloud data," Parallel and Distributed Systems, IEEE Transactions on, vol. 23, no. 8, pp. 1467–1479, 2012.

[22] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-preserving multi-keyword ranked search over encrypted cloud data," in IEEE INFOCOM, April 2011, pp. 829–837.

[23] W. Sun, B. Wang, N. Cao, M. Li, W. Lou, Y. T. Hou, and H. Li, "Privacy-preserving multi-keyword text search in the cloud supporting similarity-based ranking," in Proceedings of the 8th ACM SIGSAC symposium on Information, computer and communications security. ACM, 2013, pp. 71–82.

[24] C. Orencik, M. Kantarcioglu, and E. Savas, "A practical and secure multi -keyword search method over encrypted cloud data," in Cloud Computing (CLOUD), 2013 IEEE Sixth International Conference on. IEEE, 2013, pp. 390–397.

[25] W. Zhang, S. Xiao, Y. Lin, T. Zhou, and S. Zhou, "Secure ranked multi -keyword search for multiple data owners in cloud computing," in Dependable Systems and Networks (DSN), 2014 44th Annual IEEE/IFIP International Conference on. IEEE, 2014, pp. 276–286.

AUTHORS

Mr.Rakeshreddy Gurrala, working as Assistant Professor in the Department of Information Technology at Keshav Memorial Institute Of Technology (KMIT), Hyderabad. I did my M.Tech in Information *Technology* from JNTUH and B.Tech in Information Technology from JNTUH. My research interests network security and wireless sensor networks and mobile computing.

Mr.Srinivas Rao Adabala, working as Assistant Professor in the Department of Information Technology at Keshav Memorial Institute Of Technology (KMIT), Hyderabad. I did my M.Tech in Computer science and engineering *from* JNTUH and B.Tech in Information Technology from JNTUH. My research interests network security and wireless sensor networks and cloud computing.