International Journals of Advanced Research in Computer Science and Software Engineering
ISSN: 2277-128X (Volume-7, Issue-8)

Research Article | **August 2017**

Study of Dynamic Multikeyword Text Search Techniques over Encrypted Data in Cloud

Ankita Puri, Naveen Kumari
Punjabi University Regional Centre of Information and Technology Management, Mohali, Punjab, India

DOI: 10.23956/ijarcsse/V7I8/0144

Abstract — Day by Day ,with the advancement of modern technology over cloud computing motivating the data owners to outsource their data to the cloud server like Amazon, Microsoft, Azure etc .With the help of data outsourcing ,the organization can provide reliable data services to their user without any management of the overhead concern. Suppose, a large number of users that are on cloud and large number of documents on cloud, Its important for the service provider to allow multi-keyword query and provided the result that meet efficient data retrieval needs. In this paper, for the first time, we define and solve the challenging problem of privacy preserving multi-keyword ranked search over encrypted cloud data (MRSE), and establish a set of strict privacy requirements for such a secure cloud data utilization system to become a reality. Among various multi-keyword semantics, we choose the efficient principle of "coordinate matching", i.e., as many matches as possible, to capture the similarity between search query and data documents, and further use "inner product similarity" to quantitatively formalize such principle for similarity measurement.

Keywords— Cloud Computing, Multi Keyword Search, Ranked Search, Encryption method.

I. INTRODUCTION

The term "Cloud Computing" is the computing services in Information Technology like infrastructure, platforms, or applications could be arranged and used through the internet. Infrastructure upon which cloud is built upon is a large scaled distributed infrastructure in which shared pool of resources are generally virtualized, and services which are offered are distributed to clients in terms of virtual machines, deployment environment, or software. Hence it can be easily concluded that according to the requirements and current workloads, the services of cloud could be scaled dynamically. As many resources are used, they are measured and then the payment is made on the basis of consumption of those resources.

Cloud computing is set of resources that are being allocated on demand. Cloud computing proposes new ways to provide services. These new innovative, technical and pricing opportunities bring changes in the way business operated. Cloud computing is the matchless computing technology. Cloud computing is a new label to an old idea. Cloud computing is a collection of resources and serviced provided by cloud service provider through internet. Cloud services are distributed from data canters sited all over the world. Cloud computing makes possible for its users to use the virtual resources via internet as per requirements. Cloud computing grabbed the spotlight in few years.

General example of cloud services are Google Engine, Oracle Cloud, Office 365. As the cloud computing is growing rapidly this also leads to severe security concerns. Lack of security is the only barrier in wide adoption of cloud computing. The rapid growth of cloud computing has brought many security challenges for users and providers.

II. CLOUD DEPLOYMENT MODEL

Public Cloud: Public cloud describes the conventional meaning of cloud computing that is accessible, effective ways and means, which are accessible on internet from a minor party, which detached assets and charges its clients on the basis of utility. Cloud organization is possessed and accomplish by a supplier who suggest its retune to public domain.

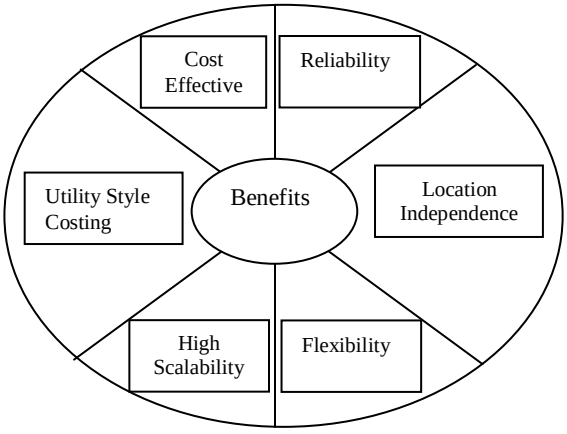


Fig 1: Benefits of Public Cloud

Private Cloud: Private cloud is a term used to donate a proprietary computing architecture provisioned services on corporate networks. Big enterprises usually used this type of cloud computing to permit their private network and information Centre administrators to effectively become in-house 'service providers' catering to customers within the corporation. Cloud organization is establishing for a particular aggregation and managed by a third party under a service level agreement. Only single organization preferred to operate via corporate cloud. There are advantages (benefits) of internal cloud model. The diagram given below depicts a few of these advantages (benefits):

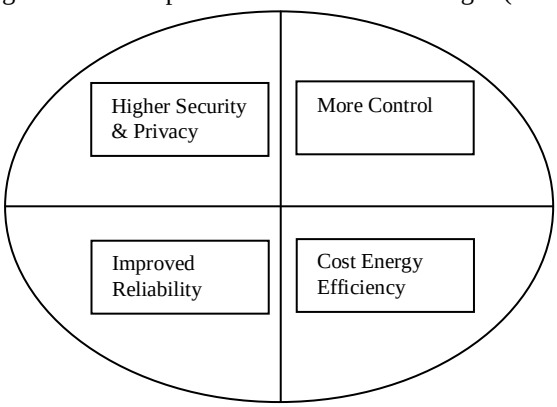


Fig 2: Benefits of Private Cloud

Hybrid Cloud: A hybrid cloud comprises assets from both corporate and public providers will definitely become the demanded choice for enterprises. The hybrid cloud is a combination of both corporate cloud and public cloud. For example, for general computing enterprise could selects to make usage of external services, and its own data Centre's comprises it own data Centre's. Hybrid cloud model has number of advantages (benefits).The diagram given below reveals some of those advantages (benefits):

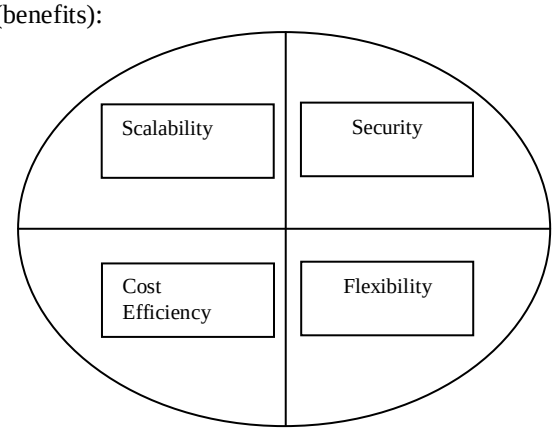


Fig 3: Benefits of Hybrid Cloud

III. PRIVACY IN CLOUD COMPUTING

As the usage of internet is increasing, the users prefers to store/upload data on cloud so that they can access the data from anywhere in the world. But keeping the privacy in mind, traditional data storage techniques for authentication are not that much reliable. For the protection of the data over cloud, the data needs to be encrypted before uploading them to cloud to avoid the escalation which may open up with the confidentiality of data.[1] Cloud storage is the very important and widely used cloud computing model where data is stored on remote servers and managed and accessed over internet. It is managed and operated by the CSP on a server which support data storage and is built on virtual machines. It works through the data centre virtualization which provides applications and data users a virtual architectural environment that is scalable according to its requirements. All the search schemes which support the multikeyword functionality retrieve search output.

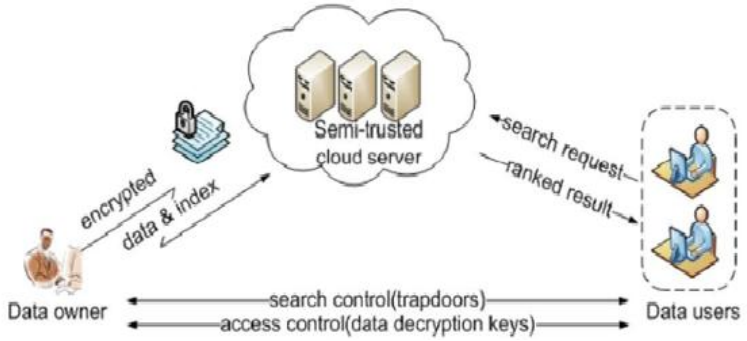


Fig 4: Privacy preserved Multi keyword Search[2]

IV. MULTI KEYWORD RANKED SEARCH IN CLOUD

Keyword search have been implemented in various databases to get the better result. One of the major requirement over the web is about the selection of best service and service provider over the web[1]. When we talk about cloud service the work is more specific and the parametric. The tag cloud is also used to search the relevant information on the basis of tag assignment to different kind of keywords and on the basis of these tags a query refinement is been performed. Finally a flexible search over the database is performed to derive the final outcome. The result analysis is based on the basis of effectiveness and efficiency of the cloud services[2]. Every keyword search follow some common step which can be explained below:

Keyword Extraction: A word used by search engine in its search for relevant web pages. In this architecture at first the query is performed by the user and on this query the query analysis is performed. The analysis includes the keyword extraction by removal of stop words. Once the keyword extracted the next work is about to perform the keyword summarization based on frequency of keywords. Once we get the summarized keywords it will be used as the content based analysis.

Indexing: Another program called indexer ,reads the documents and creates the index based on the words contained in each document. In information, indexing structure is used to store a list of mappings from keywords to the corresponding set of files that contain this keyword allowing full text search.

Ranking formula: For ranked search purposes ,the job of determining which files are most relevant is determined by assigning a numerical score ,which can be determined to each file based on some ranking formula. Ranking function is used to calculate relevance scores of matching files to a given search request. The most widely used formula for evaluating relevance score in the information retrieval is TF*IDF , where TF (term frequency) is simply the number of times a given term (keyword) present within sa file to represent the importance of term within a file.IDF (inverse document frequency) is given by dividing the number of files in the whole collection by number of files containing the keyword to measure the overall importance of the keyword in the whole collection. Now we will explain different methods which uses these above discussed steps in information retrieval. We have to search the files as well as maintain the privacy of the documents retrieved and stored at the server. Several protocols have also used to maintain the security in ranked search in cloud computing. One of them is private information retrieval(PIR) ,provides useful cryptographic tools to hide the queried terms and the the data retrieved from the database while returning most relevant documents to the user. With the growth of music collection ,music information retrieval (MIR) has been given in recent years. There are several ways to retrieve pieces of desired music. For example query by meta-information and query by tag. In content based MIR system ,user input a query of multiple tags with multiple level of preference by colorizing desired tags in a web based tag cloud interface to search music[4].Keyword search of PubCloud is also used in PubMed (database of biomedical literature). PubMed which is part of National Center for Biotechnology Information (NCBI), is a centralized database that indexes millions of biomedical publications. Responses to queries are presented by ranked list that are similar to responses of most web search engines.

V. K-MEANS CLUSTERING

Clustering is an important chore in data analysis and data mining applications. Data divides into similar object groups based on their features by clustering process. Each data group with similar objects are clusters. It means clusters are the ordered set of data which have the familiar characteristics. Clustering is a process of unsupervised learning. Highly superior clusters have high intra-class similarity and low inter-class similarity. Clustering algorithms have many categories like hierarchical- based algorithms, partition-based algorithms, density-based algorithms and grid based algorithms.

K-means clustering technique is a technique of clustering which iswidely used. This algorithm is the most popular clustering tool that is used in scientific and industrial applications. It is a method of cluster analysis which aims to partition observations into k clusters in which each observation belongs to the cluster with the nearest mean. The basics algorithm is very simple:

1. Select K points as initial centroids.
2. Repeat
3. Form K cluster by assigning each point to its closest centroid.
4. Recomputed the centroid of each cluster until centroid does not change.

Properties of k-means algorithm

1. Large data set are efficiently processed.
2. It often terminates at a local optimum.
3. It works only on numeric values.
4. The shape of clusters are convex.

VI. RELATED STUDY

Chi Chen et al. [1] developed the searchable encryption for multi-keyword ranked search over the storage data. Specifically, by taking the huge number of outsourced documents (data) on the cloud, algorithm utilize the k-nearest neighbor and relevance score techniques to design an efficient multi-keyword search scheme that can give back the ranked search results based on the accuracy. This system improve the search efficiency by supporting efficient index, and hide access pattern of the search user by adopt the blind storage system.

Keerthana G et al. [2]fabricated an application for enhancing cloud security utilizing partition and encryption technique which will enhance the cloud security. First of all record from client were taken and partition it into number of parts. After partition we encrypt the all record parts. At that point we send record parts to various cloud servers. At the point when client need that information back we take that information from cloud servers and decrypt that information. After decrypting, merging of that information is done and offer it to client. Our objective is that the application ought to have straight forward client interface for clients adaptability. The proficient method for giving security is the utilization of hybrid cryptography for more secured sending and receiving of information.

K.Ramadevi and L.Sunitha Rani [3]proposed schemes to deal with Privacy preserving Ranked Multi-keyword Search in a Multi-owner model (PRMSM). To enable cloud servers to perform secure search without knowing the actual data of both keywords and trapdoors and systematically construct a novel secure search protocol. To rank the search results and preserve the privacy of relevance scores between keywords and files, we propose a novel Additive Order and Privacy Preserving Function family. To prevent the attackers from eaves dropping secret keys and pretending to be legal data users submitting searches, we propose a novel dynamic secret key generation protocol and a new data user authentication protocol. Furthermore, PRMSM supports efficient data user revocation.

Khnd Sri Sandhya and K. Venkat Rao [4]scheme is based on multi-keyword ranked search which supports dynamic update operations. The data owner generates an exceptional tree-based catalog composition together with "Greedy Depth-first Search" criteria to make successful multi-keyword search. Achieving parallelism is the limitation of the existing system. The described technique extended the existing scheme with secure Dynamic Key generation along with the vector space model and it also includes TF_IDF model for index development as well for query generation. The dynamic key generation favors parallel search process by allowing multiple users to retrieve the same encrypted cloud data

Veerraju Gampala and Sreelatha Malempati [5]employed probabilistic public key encryption algorithm for encrypting the data and invoke ranked keyword search over the encrypted data to retrieve the files from the cloud. This technique aimed to achieve an efficient system for data encryption without sacrificing the privacy of data. Further, this ranked keyword search greatly improves the system usability by enabling ranking based on relevance score for search result, sends top most relevant files instead of sending all files back, and ensures the file retrieval accuracy. Thorough security and performance analysis, we prove that our approach is semantically secure and efficient.

Shreejit Pillai et al. [6]proposed schemes to deal with privacy preserving ranked multi keyword search in a multi owner environment to enable cloud servers to perform secure search without knowing the actual data of both keywords systematically construct a novel secure search protocol. To rank the search results and preserve the privacy of relevance course between keywords and files, and proposed a novel additive order and privacy preserving function family. To prevent the attackers from eavesdropping secret keys and pretending to be legal data owners submitting searches proposed a dynamic secret key generation key protocol and a new data user authentication protocol.

Shrilakshmi Prasad and B. S. Mamatha [7] defined and solve the problem of association attack by encrypting the index file using Paillier cryptographic algorithm. So cloud will have the challenge of searching the index file with the search query where both will be in an encrypted format. Hence privacy of the document will be preserved. Cosine similarity search is used to retrieve the top matching documents based on their relevance score and the beauty of the proposed system is the user can give multiple keywords in their search query. The enterprises are interested in storing their data in the public cloud. Before uploading the data on to the cloud, it needs to be encrypted to preserve privacy. In order to ease searching, the index file should be built for each document. The index file contains the keyword and its count in the particular document. The unencrypted index file leads to association attack since with the keywords and their count, the content of the document can be known.

Seema Ranga and Ajay Jangra [8]aimed of Intrusion detection System is to defend the security of the Computer system by a layer over the defense system. IDS systems sense the misuse, breach in the security system and also the malicious or unauthorized access to the system. Although Firewalls works for the same reason but the major difference between firewalls and the IDS is IDS suspect the source of the attack and signals the alarm to the system but a firewall directly stops the communication without informing the system. These attacks requires true concerns as they harm the data stored in system and also effect the network traffic, data packet etc.

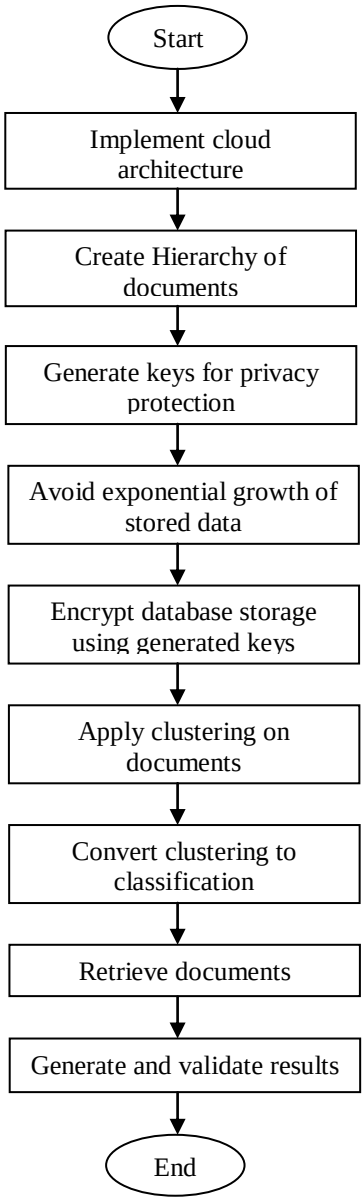
Wei Zhang et al. [9]proposed scheme to deal with privacy preserving ranked multi-keyword search in a multi-owner model (PRMSM). To enable cloud servers to perform secure search without knowing the actual data of both keywords and trapdoors, algo systematically construct a novel secure search protocol. To rank the search results and preserve the privacy of relevance scores between keywords and files and proposed a novel additive order and privacy preserving function family. To prevent the attackers from eavesdropping secret keys and pretending to be legal data users submitting searches, this approach developed a novel dynamic secret key generation protocol and a new data user authentication protocol.

VII. PROPOSED WORK

GA based K-Means Algorithm

Start
Input: Read unstructured data
Output: Clustered Data
Step 1: Read unstructured data
Step 2: Generate Initial Value for centroids

Step 3: Choose centroid randomly
Step 4: Choose Similarity matrix using Euclidian distance
Step 5: if cluster formed
Generate results
else
Apply GA algorithm
Step 6: Generate Result
Step 7: Compare results



The performance Analysis is done on the basis of two categories which are given below:
Comparison of K-means, GA and improved K-Means for various output parameters based on different data sets.
Performance evaluation of new designed Improved K-Means algorithm based on various output parameters using different data sets.

Comparison of K-means, GA and improved K-Means for various output parameters based on different data sets

In this type of comparison the various performance parameters like recall, purity, intra cluster distance, computation time and computational complexity are computed for k-means, GA and improved K-Means algorithm for one data set at a time and comparison is done with these three algorithms. The comparisons between these algorithms are shown with the help of bar charts. All these type of comparisons are described below.

Comparison through Recall for Wine Data set

The results for recall for wine data set using K-means, GA and improved K-Means are shown in Table 4.1. These results for recall for K-means, GA and improved K-Means are computed for different number of clusters. In this study, results are computed up to 10 numbers of clusters. The recall value is calculated in terms of percentage.

Table 7.1 Recall Using K-means, GA AND improved K-Means algorithm

Number of clusters	**Recall (%) K-means**	**Recall (%) GA**	**Recall (%)** improved K-Means
2	28.86	31.45	37
3	29.64	32.3	38
4	28.86	31.45	37
5	16.38	17.84	21
6	17.16	18.7	22
7	29.64	32.3	38
8	28.86	31.45	37
9	18.7	17.16	22
10	28.86	31.45	37

Table 7.1 shows the recall value for wine data set up to 10 numbers of clusters by using K-means algorithm, GA and improved K-Means algorithm.

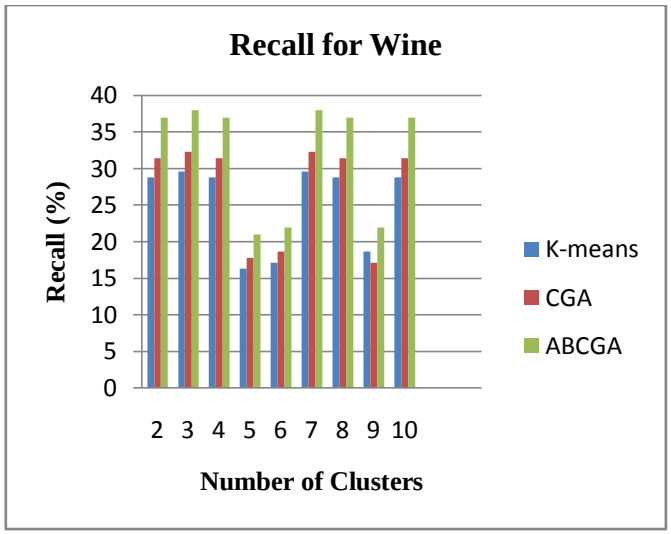


Figure 7.1 Comparison of Recall for K-means, GA and improved K-Means algorithm

Figure 7.1 shows that the Recall for IMPROVED K-MEANS algorithm is more than K-means and GA algorithms and recall is highest for IMPROVED K-MEANS when number of cluster is 7 and lowest when number of cluster is 5.

Comparison through Purity for Wine Data set

The results for recall for wine data set using K-means, GA and IMPROVED K-MEANS are shown in Table 7.2. These results for purity are computed for different number of clusters. In this study, results are computed up to 10 numbers of clusters.

Table 7.2 Purity using K-means, GA and IMPROVED K-MEANS Algorithm

Number of clusters	**Purity (%) K-means**	**Purity (%) GA**	**Purity (%) IMPROVED K-MEANS**
2	30.52	34.04	40.7
3	31.35	34.96	41.8
4	30.52	34.04	40.7
5	17.32	19.32	23.1
6	18.15	20.24	24.2
7	31.35	34.96	41.8
8	30.52	34.04	40.7
9	20.24	18.15	24.2
10	30.52	34.04	40.7

Table 7.2 shows the Purity value for wine data set up to 10 numbers of clusters by using K-means algorithm, GA and IMPROVED K-MEANS algorithm.

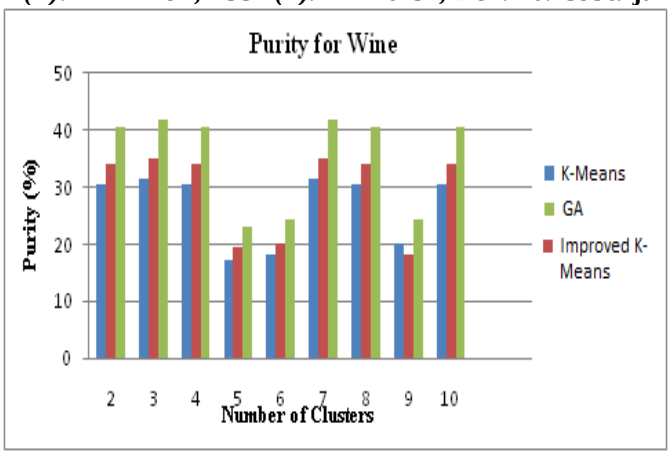


Figure 7.2 Comparison of Purity for K-means, GA and IMPROVED K-MEANS algorithm

Figure 7.2 shows that the purity for IMPROVED K-MEANS algorithm is more than K-means and GA algorithms and purity is highest for IMPROVED K-MEANS when number of cluster is 3, 7 and lowest when number of cluster is 5.

Comparison through Intra Cluster Distance for Wine Data set

The results for Intra Cluster Distance for wine data set using K-means, GA and AGA are shown in Table 4.3 respectively. These results for Intra Cluster Distance are computed for different number of clusters. In this study, results are computed up to 10 numbers of clusters.

Table 7.3 Intra Cluster Distance using K-means, GA and IMPROVED K-MEANS algorithm

Number of clusters	**Intra Cluster Distance K-means**	**Intra Cluster Distance GA**	**Intra Cluster Distance IMPROVED K-MEANS**
2	20.8	17.6	16
3	19.5	16.5	15
4	20.8	17.6	16
5	20.8	17.6	16
6	40.3	34.1	31
7	19.5	16.5	15
8	20.8	17.6	16
9	40.3	34.1	31
10	20.8	17.6	16

Table 7.3 shows the intra cluster distance value for wine data set up to 10 numbers of clusters by using K-means algorithm, GA and IMPROVED K-MEANS algorithm.

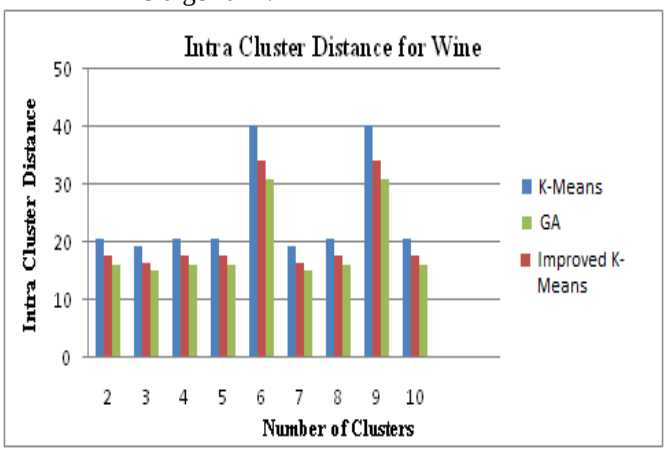


Figure 7.3 Comparison of Intra Cluster distance for K-means, GA and IMPROVED K-MEANS algorithm

Figure 7.3 shows that the intra cluster distance value for IMPROVED K-MEANS algorithm is less than K-means and GA algorithms and intra cluster distance value is highest for IMPROVED K-MEANS algorithm when number of cluster is 6, 9 and lowest when number of cluster is 3.

Comparison through Computation Time for Wine Data set

The results for Computation Time for wine data set using K-means, GA and IMPROVED K-MEANS are shown in Table 7.4 respectively. These results for Computation Time are computed for different number of clusters. In this study, results are computed up to 10 numbers of clusters.

Table 7.4 Computation Time Using K-means, GA and IMPROVED K-MEANS algorithm

Number of clusters	**Computation Time (m sec) K-means**	**Computation Time (m sec) GA**	**Computation Time (m sec) IMPROVED K-MEANS**
2	31	47	16
3	16	47	15
4	32	47	16
5	32	47	16
6	31	63	31
7	47	47	15
8	31	47	16
9	31	47	31
10	16	47	16

Table 7.4 shows the Computation time value for wine data set up to 10 numbers of clusters by using K-means algorithm, GA and IMPROVED K-MEANS algorithm.

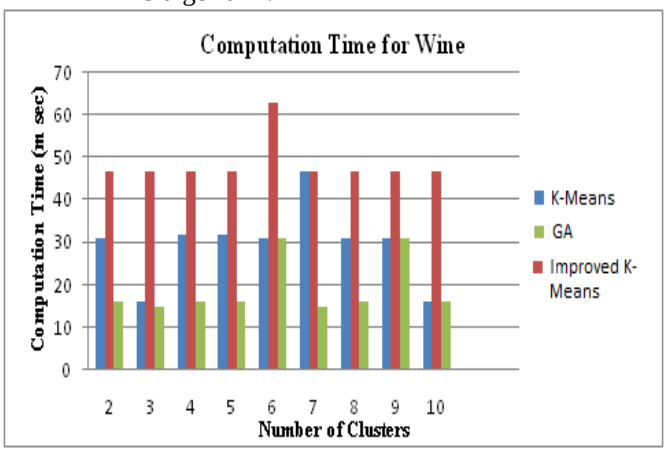


Figure 7.4 Comparison of Computation Time for K-means, GA and IMPROVED K-MEANS algorithm

Figure 7.4 shows that the Computation time for IMPROVED K-MEANS algorithm is less than K-means and GA algorithms and Computation time is highest for IMPROVED K-MEANS when number of cluster is 6, 9 and lowest when number of cluster is 3.

VIII. CONCLUSION

In this paper a review on various multikeyword set text search technique over encrypted cloud is studied. We assemble a special keyword balanced binary tree as the index, and intend a "Greedy Depth-first Search" algorithm to acquire preferable effectiveness over linear search. Likewise, the parallel search procedure can be completed to further lessen the time cost. Among the above defined technique Sequential and binary searach are least effective where as the Greedy DFS and MHR trees are efficient methods for Multi Keyword Ranked Searching. IN the multi-keyword ranked search some of the algorithm use K-Means clustering technique which has many drawbacks like efficient centroid selection, outliers. In the proposed work the drawback of K-Mean clustering may be removed for an efficient search algorithm.

REFERENCES

[1] C. Chen, X. Zhu, P. Shen, J. Hu, S. Guo, Z. Tari, A. Y. Zomaya, "An Efficient Privacy-Preserving Ranked Keyword Search Method", IEEE Transactions On Parallel And Distributed Systems, Vol. 27, No. 4, 2016

[2] G. Keerthana, S. Prabu, P. Swarnalatha, "An Efficient Data Security in Cloud Computing using Cryptography", International Journal of Advanced Research in Computer Science and Software Engineering, ISSN: 2277 128X, Volume 6, Issue 5, 2016, pp: 654-660

[3] K. Ramadevi, L. S. Rani, "Ranked keyword search over Cloud Storage by several owners using dynamic hidden keys", International Journal of Computer Science, ISSN: 2348-6600, Volume 4, Issue 2, No 7, 2016, pp: 894-900

[4] K. S. Sandhya, K. V. Rao, "Privacy-Preserving and Dynamic Multikey Generation over Encrypted Cloud Data", International Journal of Advanced Research in Computer Science and Software Engineering, ISSN: 2277 128X, Volume 6, Issue 10, 2016, pp: 78-82

[5] S. K. Pasupuleti, S. Ramalingam, R. Buyya, "An efficient and secure privacy-preserving approach for outsourced data of resource constrained mobile devices in cloud computing", Journal of Network and Computer Applications, ISSN: 1084-8045, Vol: 64, 2016, pp: 12-22

[6] S. Pillai, G. Ransing, N. Ransing, S. Markad, N. Sable, "Survey on Privacy Preserving Multi Keyword Search in Cloud Computing", International Journal of Advanced Research in Computer and Communication Engineering, ISSN (Online) 2278-1021 ISSN (Print) 2319 5940, Vol. 5, Issue 11, 2016

[7] S. Prasad, B. S. Mamatha, "Retrieving documents from encrypted cloud data in a secured way using cosine similarity search with multiple keyword search support", International Journal of Advance Research in Computer Science and Management Studies, ISSN: 2321-7782 (Online), Volume 4, Issue 5, 2016, pp: 108-115

[8] S. Ranga, A. Jangra, "A Study of IDS Technique Using Data Mining", International Journal of Technical Research & Science, ISSN No.: 2454- 2024 (online), Volume 1 Issue 6, 2016, pp: 152-158

[9] W. Zhang, Y. Lin, S. Xiao, J. Wu, S. Zhou, "Privacy Preserving Ranked Multi-Keyword Search for Multiple Data Owners in Cloud Computing", IEEE transaction on Computers, ISSN: 0018-9340, Vol: 65, Issue: 5, 2016, pp: 1566-1577

[10] I. Alsmadi, I. Alhami, "Clustering and classification of email contents", Journal of King Saud University – Computer and Information Sciences,ISSN: 1319-1578, Vol: 27, 2015, pp: 46-57

[11] C. Liu, Z. Peng, L. Wu, "Role of Time-Domain Based Access Control Model", Journal of Software Engineering and Applications, Volume: 9, 2016, pp: 57-62

[12] A. Shaheen, A. Slei, "Comparing between different approaches to solve the 0/1 Knapsack problem", International Journal of Computer Science and Network Security, Volume: 16, No: 7, July 2016, pp: 1-10

[13] Jaikishan Tindwani, Aruna Gupta, "A Survey on Multi-Keyword Ranked Query Search over Encrypted Cloud Storage", International Journal of Science and Research, ISSN (Online): 2319-7064, Volume 4 Issue 11, November 2015, pp: 2366-2370

[14] SonamDarda, Manasi. K. Kulkarni, "Study of Multi-keyword Ranked Searching and Encryption Technique over Cloud",) International Journal of Computer Science and Information Technologies, ISSN: 0975-9646, vol: 6, issue: 6, 2015, pp: 5417-5420

[15] Kalyani Sonawane, Rahul Dagade, "A Survey on Multi-Keyword Ranked Search over Encrypted Cloud Data with Multile Data Owners", International Journal of Computer Applications, ISSN: 0975–8887, vol: 162, no: 11, 2017, pp: 9-12

[16] Zhihua Xia, Xinhui Wang, Sun Xingming, Qian Wang, "A Secure and Dynamic Multi-Keyword Ranked Search Scheme over Encrypted Cloud Data", IEEE Transactions on Parallel and Distributed Systems, vol: 27, issue: 2, 2015

[17] Hongwei Li, Yi Yang, Tom H. Luan, Xiaohui Liang , Liang Zhou, Xuemin Sherman Shen, "Enabling Fine-Grained Multi-Keyword Search Supporting Classified Sub-Dictionaries over Encrypted Cloud Data", IEEE Transactions on Dependable and Secure Computing, Vol: 13, Issue: 3, 2016, pp: 312 – 325

[18] Zhangjie Fu, Kui Ren, Jiangang Shu, Xingming Sun, Fengxiao Huang, "Enabling Personalized Search over Encrypted Outsourced Data with Efficiency Improvement", IEEE Transactions on Parallel and Distributed Systems, Vol: 27, Issue: 9, 2016, pp: 2546-2559

[19] Shulan Wang, Junwei Zhou, Joseph K. Liu, Jianping Yu, Jianyong Chen, Weixin Xie, "An Efficient File Hierarchy Attribute-Based Encryption Scheme in Cloud Computing", IEEE Transactions on Information Forensics and Security, Vol: 11, Issue: 6, 2016, pp: 1265-1277

[20] Ning Cao, Cong Wang, Ming Li, Kui Ren, Wenjing Lou, "Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data", IEEE Transactions on Parallel and Distributed Systems, Vol: 25, Issue: 1, 2014, pp: 222-233

