Android Application for Safe and Efficient Search over cloud

Anil Kumar[1], Roopalakshmi S[2] and S V Krishna Reddy[3]

[1-2] *Asststant Prof. Vemana Institute of Technology/CSE Department, Bengaluru-34, India*

[3] *Asst. Prof. Don Bosco Institute of Technology /ECE Department, Bengaluru-74, India*

Abstract—Cloud Technology has changed the world of Information Technology by efficiently using resources like computational capabilities, storage and network. Cloud adoption by businesses will take a leap in coming years. The market of cloud is expected to grow by $20.5 billion by the year 2018. Mobile phones are used by everyone and bringing together mobile phones and cloud infrastructure can help mobile to have access to immense amount of computation and storage which helps mobile industry to overcome storage issues. The application helps in retrieving sensitive information one stores on cloud with much more efficiency than other traditional schemes. The languages used are java and android, android for the graphical user interface. Cloud deployment of the application is done on open source cloud platform. The software development methodology used in the development in this paper is agile way of development. The algorithm used in implementing the security is the Advanced Encryption Standard (AES) which helps in storing the sensitive data securely over any network or platform. The user's multi keyword query is accepted from user interface and fed to the cloud platform, which by making use of ranked binary search tree technique displays results in a ranked order onto the mobile screen. To reduce the network traffic caused by multi-keyword query, compression technique is used. The performance analysis in this paper gives 1 to 1.5 millisecond search time as compared to the traditional approach's 5 to 6 millisecond which does not do any optimization like compression, usage of search tree. Because of data compression techniques, mapping table, encryption schemes and binary search, the overall performance of system is better than the conventional way of search techniques.

Index Terms—Android, Mobile cloud computing, Efficient search

I. INTRODUCTION

Cloud servers pose many security and privacy threats to the data moved from client location to service provider's location. Before moving any sensitive data to cloud, the owner of the data should make sure below few things are in place at cloud provider. The questions that should be asked to the service provider are:

1. Where your data will be placed and in what format?
2. Whether the data is encrypted or not?
3. How the data can be retrieved whenever we need it?

Make sure all your questions are answered by the cloud provider like data will be stored on one of the servers along with other customer data with proper demarcation, so that others cannot sneak into your data. While requesting the data from the cloud, you will be going through a series of authentication policies to avoid unauthorized data access. The data is stored in the encrypted format and search techniques allow you to search by giving the keywords which will also be encrypted and sent to the cloud.

Mobile cloud computing:

The first thing that comes to our mind when we hear the word mobile cloud computing is the computational and storage capacity of mobile which does not match cloud capabilities. Mobile faces many challenges like limited storage capacity compared to computers and little processing capabilities as compared to computers and also limited battery backup. Similarly many network issues such as wireless bandwidth, response latency, availability and heterogeneous nature. By moving all the data and applications to the cloud, we can give power of the computer to the mobile which lacks similar capabilities.

Android:

Android is an open source operating system which is suitable for smart phones and tablets with touch screen option. It is solely based on direct manipulation interface to make it easy to use by everyone. Direct manipulation involves touching, swapping and pinching those are in relation to the human nature. Applications used on android are written using android software development kit and the java programming language.

Security measures used in mobile cloud computing:

Mobile cloud computing faces many issues such as data breach, repudiation, etc. To avoid all these, a better framework should be in place which involves authenticating cloud service users and providing a mechanism for retrieving the data stored using single keyword search techniques wherein a single keyword entered is converted to the encrypted format and sent onto the network so that data breach is avoided. The data should be stored on the cloud in the encrypted format either using RSA or AES encryption techniques so that data will not be seen by other sharing the storage in cloud environment. Till now various options are used such as calculating a score for each document uploaded with the help of term frequency and inverted document frequency and provides the results in the ordered manner according to the relevance of the score. Till now only single keyword search is available for searching the cloud for the sensitive data.

II. LITERATURE SURVEY

The main idea in this paper is to make use of cloud computing, storage and various services that the cloud offers. Many people and their businesses are attracted towards cloud services and computation facilities, because of pay as per the usage facility (or Multi-tenant), easily scalable according to the business needs [1]. It helps businesses avoid the initial setup cost. With all the advantages that the cloud promises, it also brings many challenges to the table. This paper focused mainly on avoiding issues like security concerns that the users may have while using cloud services, privacy of user's data [2]. The second major area that this paper tries to explore is the mobile technology, the market of devices such as smart phones, tablets is booming. Every person wants to have one. The most popular devices in this era are Android devices. Our main area of focus is android mobile phones which make use of wireless networks such as 2G, 3G, 4G, and LTE [3]. These wireless networks also pose few challenges such as intermittent network connectivity, slow transfer speeds, delay in responding to a request as compared to Ethernet cables (or LAN) on desktop systems which have higher bandwidth when compared to wireless networks. The concern for the data security of data owner originates from the fact that the data service provider may maintain many data owner data on same storage devices with privacy protections, but still there is a risk that others may eavesdrops on owners data, this situation arises because of multi-tenant architectures of cloud platforms. To avoid this better privacy preserving techniques must be incorporated [4].Mobile devices also face the issue of low memory, slow processor speeds, smaller cache when compared to traditional desktop platforms. Because of these limitations, the encryption and decryption techniques cannot be implemented efficiently. The better idea is to outsource the computation and storage overhead required to encrypt and decrypt to the cloud platform. Cloud provides flexibility to scale as per business demands as such when number of users using our service through mobile increases overnight. In traditional systems, the network traffic is more because the mobile client needs to communicate with service provider for authentication, search

request (Keyword) and again search request (En-Keyword) to the cloud for preserving privacy of user data. Imagine the network traffic when more than one user is requesting [5].

To reduce the network, this paper provides an En-keyword Mapping Table (EMT) which will be stored on mobile devices upon requesting for the search service for the first time. This will avoid one round trip time, which will considerably reduce the network traffic. This paper also provides a different way of reducing the network traffic, by using an efficient compression technique for compressing the keywords used by the user whiles earching for documents from cloud [6]. Compression will help reduce the overall network traffic by some considerable amount.

Objective:

This Paper aims at improving search experience by identifying loopholes in conventional search techniques. Improve search time incurred after submitting the search keywords. It also aims at reducing network traffic incurred because of huge data that result from document search. It is implemented to aid the search done by users on mobile devices which meets the following objectives:

- Providing a secure and convenient way of accessing the user confidential data stored at cloud.
- Check if it performs better than the conventional methods.
- Ensure that the application meets the improvement in latencies incurred while searching documents.

III. SYSTEM ARCHITECTURE

System architecture gives detailed interactions between sub-modules of the whole system. It helps in clarifying the interactions between the various components which normally will be very complex to understand. By seeing the system architecture one can make out the services offered by each sub-module to contribute to the overall system. It involves identifying individual major modules and interactions between them. Fig 1shows the system architecture for the system, which clearly specifies the major modules, sub-modules inside the modules, services offered by each sub-module and interactions among major modules to constitute the whole system [34]. The modules are Android Application, Cloud Service Provider and Data Owner. Application interfaces the user with an interface which is easy to understand and use for any average android mobile user. Cloud Service Provider will help in reducing the total initial cost incurred while opting for cloud for any user. Data Owner gives services such as uploading documents and their indexes for faster search and associating the documents uploaded with an individual authentic user and storing them securely and also helps in retrieving them securely, because the indexes are sliced and stored which will help in expediting the search process.

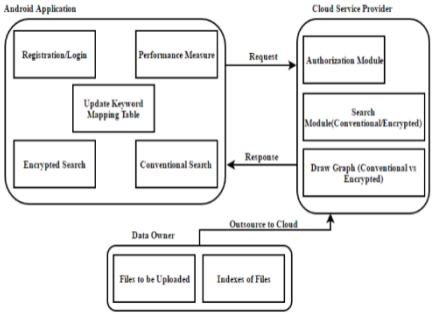

Fig 1 System Architecture for the Android Application over Cloud

Services offered to the end user are Authorization, Encrypted Search, and Conventional Search and manually updating the keyword mapping table stored on the mobile devices. Services offered by the cloud involve deploying the user sensitive data, provide access to valid user and give performance measure about the search. Data Owner gives service such as uploading user sensitive data onto cloud and indexes of those files also.

IV. DETAILED DESIGN

Detailed design gives all the information about various modules with the help of flowcharts and algorithms. This chapter gives more importance to the internal logic involved in the whole system. The internal logic will be explained based on individual modules. Functional description about the major modules or sub-modules is included in this section. Flow charts must specifically explain the various function calls involved in making the whole module work. Flow chart must follow all the rules like process should be drawn with the help of a rectangle, conditional decision with diamond, input/output with parallelogram and start, end with circle or oval shaped blocks.

The detailed design explains the following:

a) Structure Chart of the whole system

b) Functional Flow Chart of major modules

c) Detailed Description of major modules

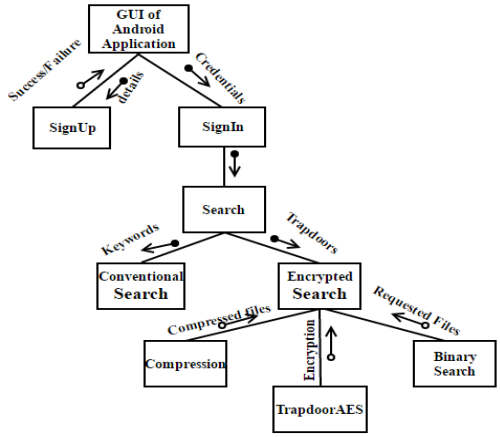

Fig 2 Structure Chart for Android Application

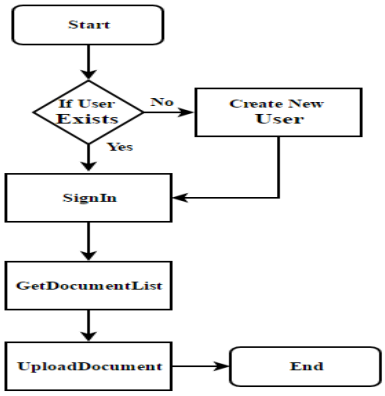

Fig 3 Flowchart of Authorization and Upload Document Module

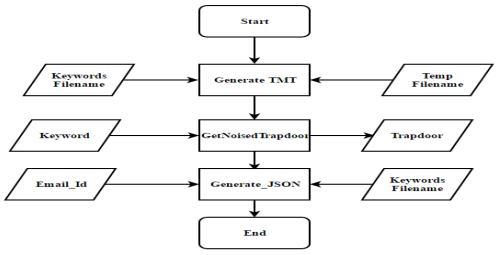

Fig 4 Flowchart of Encrypted Keyword Mapping Table

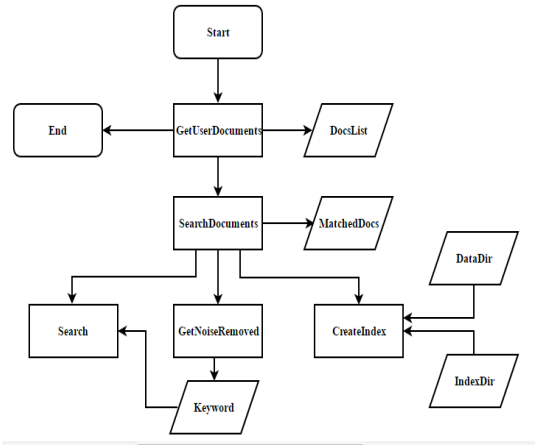

Fig 5 Search Module

V. CHALLENGES ENCOUNTERED AND STRATEGIES

The challenges faced while designing and developing the application are enlisted below to give a glimpse of the strategies used to tackle the problem faced.

Reducing the Network Traffic

Finding a way to reduce the network traffic was a big challenge. Network traffic is reduced by using a compression technique such as gzip feature provided by java programming language [8]. By compressing the files sent onto public network, the overall network traffic can be reduced. By avoiding the unnecessary requests sent onto public network, the network traffic can also be reduced. Instead of allowing the user to send request for trapdoors for keywords, a table can be accessed in the local memory to avoid the unnecessary network traffic.

Reducing the Search Time needed

There was a lot of confusion as to what search technique to be used. There are many search techniques each with their own merits and demerits. It was hard to choose among all, the one that best suits this papers need. Ranked binary search is used to search the files in O (log n) time. Ranking the files based on the term frequency in the document helps while sorting the files before displaying onto the user's screen.

Creating the Indexes of files

Java provides API's for creating indices named as apache lucene. By using the functions provided by these one can easily create indices of files for faster access. Indexing the filenames and storing only indices need limited amount of memory, thereby removing the constraint of memory.

VI. EXPERIMENTAL ANALYSIS AND RESULTS

To verify that the application developed works better than the existing applications, we needed few performance metrics that we can use to compare and make sure that it does. The main performance metric that was considered in this application

development is the search time required to access the data stored on cloud. Experimental analysis reveals the parameters which will help us in demonstrating the performance of the newly built system with that of existing one.

Evaluation Metric

Evaluation metrics help in getting insights about the performance of the application. Metrics are solely based on the factors considered in the application for benchmarking it against existing system [50]. This uses encryption time, search time, and ranks assigned to documents based on their relevance to the entered keywords. Most relevant will be listed first and so on. In Fig 6, x-axis has number of keywords used and y-axis has the time taken to encrypt the keywords. It is clearly shown in the figure that the time taken will vary as the number of keyword varies.

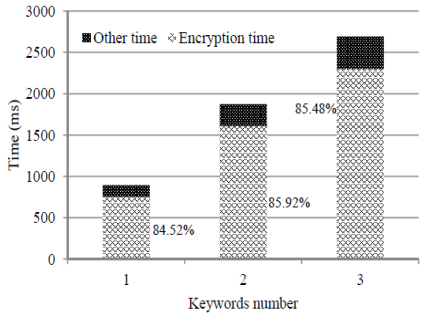

Fig 6: Time required to encrypt keywords varies with number of keywords

Performance Analysis

All the parameters considered for evaluating performance is being shed light on below. Parameters like search time for encrypted search compared to traditional search, cloud search time and encrypted keyword compression.

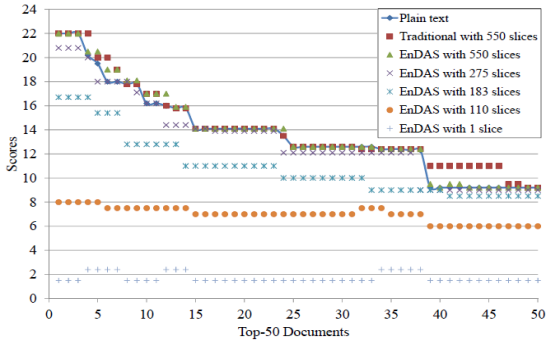

Fig 7 Cloud data search time in encrypted keyword search

In the above fig 7, the rating for each type of search types is represented on y-axis and the number of documents on x-axis. It clearly depicts the fact that, the search algorithm shows better accuracy against traditional schemes when used fine-grained slice number. The network traffic is evaluated based on the throughput given out by the system which in case of encrypted search is supported by keyword mapping table on mobile devices and compression of keywords before sending them out on network thereby reducing the network traffic. Network traffic is less as compared to the traditional schemes as depicted in below fig 8.

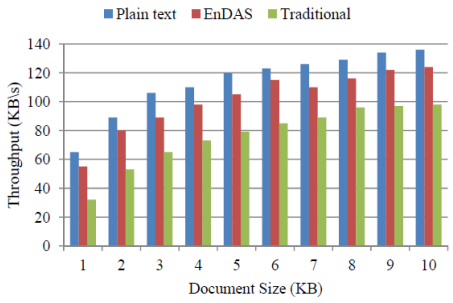

Fig 8 Throughput comparison between various schemes

As you can see in above fig 8, the comparison is done between plain text, Encrypted search and Traditional search schemes. Encrypted search tries to reduce the throughput to the network as compared to plain text search scheme. The effectiveness of keyword compression can be shown with the help of a graph which clearly shows the size differences between various schemes as in below fig 9.

After compression the size of keyword gets reduced to smaller sizes, the data selected for demonstrating the experiment is 5000 keywords selected randomly and fed to the compression scheme in encrypted search, traditional and plain text search. The noise is added to the pure keywords and later on they are compressed. The size of the keyword in case of encrypted search is reduced by 87% as compared to plaintext, and 65% when compared to the traditional search schemes. This is the result of using an optimized keyword mapping table and compression technique to ensure the reduction in network traffic inefficiencies.

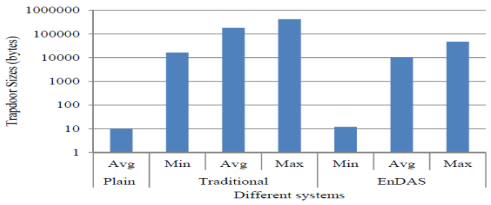

Fig 9 Keyword sizes compared to different systems

As you can see in the above fig 9, clearly the keyword size difference is seen when compared with plain text and traditional search schemes. Above figures all will tell us the same story saying that the encrypted keyword scheme is better than the traditional and plaintext search methods. The above comparisons are mainly focused on the search time and network traffic and so is the result [7]. The results are the reduction in network traffic because of keyword compression and keyword mapping table, search time efficiency because of binary search algorithm.

VII. CONCLUSION

Android application development poses many challenges like network in efficiencies, limited bandwidth of data networks available and limited memory available. This paper addresses all these issues. Network traffic is controlled by making use of keyword mapping table which will avoid unnecessary traffic on network. Keyword mapping table designed and implemented in such a way that the memory is utilized efficiently. Throughput from the application going onto network is also controlled by compressing the request given by user for searching their data on cloud. Rather than sending keywords directly onto network, they are compressed to reduce network traffic. Binary search tree mechanism helps in bringing down the overall time complexity of the system. The output of the requested document come in order of their relevance to the

search terms entered as part of request. This is achieved by considering the term frequency and inverted document frequency of all the documents uploaded by the end user.

VIII. FUTURE ENHANCEMENT

The android application can be enhanced further by including following features:

The user can be allowed to upload other types of data like personal and private images, multimedia and any other type of data with better compression techniques.

There is a scope for optimization in search technique used in this paper; a better search technique such as B tree, B+ tree can be used to reduce the search time.

REFERENCES:

[1]. D. Huang, "Mobile cloud computing", IEEE COMSOC Multimedia Commun. Tech.Committee (MMTC) E-Letter, vol. 6, no. 10, 2011, pp. 27-31.

[2]. N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, " Privacy-preserving multi-keywordranked search over encrypted cloud data ", in Proc. Int. Conf. Compute. Commun.,Apr. 2011, pp. 829–837.

[3]. C. Wang, N. Cao, K. Ren, and W. Lou, " Enabling secure and efficient rankedkeyword search over outsourced cloud data ", IEEE Trans. Parallel Distrib. Systems,vol. 23, no. 8, 2012, pp. 1467–1479.

[4]. C. Wang, N. Cao, J. Li, K. Ren, and W. Lou, " Secure ranked keyword search overencrypted cloud data ", in Proc. IEEE Int. Conf. Distrib. Comput. Syst. (ICDCS), Jun.2010, pp. 253–262.

[5]. C. Gentry and S. Halevi, " Implementing gentrys fullyhomomorphic encryptionScheme ", in Advances in Cryptology–EUROCRYPT 2011, 2011, pp. 129–148.

[6]. C. O¨ rencik and E. Savas,, " Efficient and secure ranked multikeyword search onencrypted cloud data ", in Proc. Joint EDBT/ICDT Workshops, Mar. 2012, pp. 186–195.

[7]. C. Wang, Q. Wang, K. Ren, and W. Lou, "Privacy-Preserving Public Auditing forData storage Security in Cloud Computing," Proc. IEEE INFOCOM, 2010.

[8]. Y. Ishai, E. Kushilevitz, R. Ostrovsky, and A. Sahai, "Cryptography from Anonymity," Proc. IEEE 47th Ann. Symp. Foundations of CS, pp. 239-248, 2006.

