An Efficient Privacy-Preserving Ranked Keyword Search Method

Pallavi[1] Dr. Basavaraj Mathpathi[2]
[1]PG Student [2]Head of Dept. & Professor
[1,2]Department of Computer Science & Engineering
[1,2]Appa Institute of Engineering and Technology Kalaburagi, Karanataka India

Abstract— The main concern in cloud computing is the searching on the encrypted data. The query is not known to the clouds and searching concept based on the Keyword should show all the files in the cloud. The main aim is to provide the privacy protection and safety to the data that is saved in cloud system. To upload and download a file securely from cloud and to encrypt and decrypt the files from the cloud .To provide keys to users for downloading file from cloud securely. To keep the identity of each user and data owners protected and to store all the information on cloud in protected manner. We proposed a Rank search concept is introduced which takes less time by choosing the desired files along with keyword categories. Data owner provides the file to browser, encrypt and uploads files into the Drivehq in encrypted form. Cloud server will provide following operations such as store encrypted data, view attachers, view owner files, give privileges to user to create index on searched data to view all the android users. User can get the files from the cloud by using their credentials and secret key. User can searches for files based on contents keyword request for secret key, requests for downloading files and retrieve and to store the data.

Key words: Cloud Computing, Encrypted Data and Rank Search Concept

I. INTRODUCTION

Cloud Computing is an emerging technology that is receiving a lot of attention from the generation of the academic and Industrial worlds. In cloud computing, people or the user can provide their resources such as computation and storage to servers via the internet. Sharing the user resources is provided by Clouds. By applying this, makes free to users from the maintaining the resources. Cloud computing provides many forms of services such as applications (e.g., Google products etc.), Infrastructures (e.g., window's services, Amazon's EC2 etc.), and provides platform for software engineers to write applications (e.g., windows azure, Amazon's S3 etc.).

Information stored in the cloud server is purely precise for example medical documents, and social networking sites.

Parameters such as safety and privacy are important parts in the cloud computing. First, user itself must be authenticated before any transaction initiated. Lastly, It should be verified that the cloud do not make mess with information that is outsourced by the users. Providing user privacy is also a main concern so that the protection is provided for the user's identity from cloud system as well as from other users.

The main concern in cloud computing is the searching on the encrypted data. The query is not known to the clouds and searching concept based on the Keyword should show all the files in the cloud. The cloud receives the encrypted keywords and results will be returned by the cloud without the knowing the correct keyword to search. The information records should have the keywords that are associated with records that enables for the search. While searching in clouds the actual keywords are returned when matched with exact keywords in the cloud storage.

Challenging task in cloud computing is the accountability and includes technical issues. In cloud computing the operations performed or requested cannot be denied by either clouds or by the users. It is important to maintain the log file for every transactions performed on the clouds and it important to maintain that log file so to decide how much information kept in that log file.

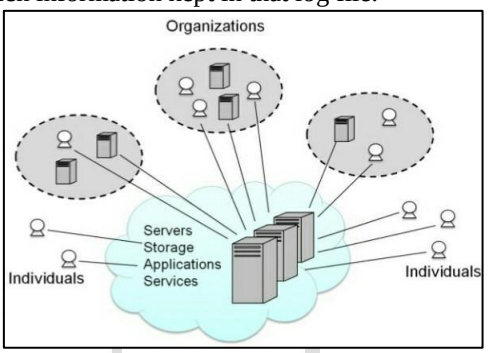

Fig. 1: Basic computing model

II. RELATED WORK

The authors S.Grzonkowski, et. al. in [1] explained a number of well-known authentication protocols are taken into consideration within the context of subsequent-generation cell and CE community services. The capacity weaknesses of present day protocols may be triumph over the usage of Zero Knowledge proof (ZKP) strategies to guard user passwords so an opportunity ZKP protocol, SeDiCi 2.zero, is described. This gives mutual and also two-component authentication that is taken into consideration extra at ease in opposition to numerous phishing tries than present relied on third party protocols.

The suitability of the sort of ZKP protocol for diverse CE-primarily based cloud computing packages is tested.

The authors D.Boneh, et. al. in [2] explained the issues of searching the files is encrypted on using a general key system. Consider any user can sends the email to any user encrypted under their general key. Email opening needs to test whether this email contains keyword could route the email accordingly. On the other hand any user does not like to give opening authority to decrypt users messages and so for discussed with general key by encrypting files. This mechanism helps us to use in other instance like seeking to provide general key with among different buildings.

The author Y. C. Chang et. al. in [3] explained in as followed surveys: Any registered user can store or save their files over the cloud in an encrypted form and as well as retrieve the documents which you need to download it from cloud by decryption general key efficiently retrieve few of the

encrypted files containing (or indexed by) particular keywords, keeping the keywords themselves as a secret and not jeopardizing the safety of the remotely saved file. They offer some solutions for this problem under some safety requirements. Indeed, our methodology is independent for the encryption method chosen for the remote files. They are also incremental user can submit proper files are secured with previous search.

The authors I. H. Witten, et.al. in [4] explained the emerging cloud technology, due to their diverse specific and appealing residences, are unexpectedly being adopted all through the IT industry. Authors identify protection challenges that get up in incorporation of cloud-based services, and gift a set of solutions to address them. To assure the person control over the get entry to their personal facts, it's far a promising technique to encrypt the records earlier than outsourcing on cloud. Fundamental problems which include privateness, scalability in key management, flexible get right of entry to and green consumer revocation which are the most crucial concerns for gaining high-quality grained, cryptographically used facts get right of entry to control.

III. EXISTING AND PROPOSED SYSTEM

A. *Existing System*

- Cryptographic Mark and Sun et al. utilize Merkle hash tree to make an demonstrable MDB-tree. In any case, their work can't be specifically utilized as a part of our design which is situated for protection saving various watchword look. Along these lines, an appropriate system that can be utilized to check the query items inside huge information situation is basic to both the CSPs and end clients.
- In the current years, analysts have projected many figure content inquiry plots by consolidating the cryptography procedures. What's more, the connection between archives is disguised in the above techniques. The connection between records speaks to the properties of the reports and subsequently keeping up the relationship is key to completely express an archive. For instance, the relationship can be utilized to express its classification. On the off chance that a record is free of whatever other archives with the exception of those reports that are identified with games, at that point it is simple for us to declare this archive has a place with the classification of the games.
- While doing encryption blindly, this vital property has been covered in the conventional techniques. Along these lines, proposing a strategy which can keep up and use this relationship to speed the hunt stage is attractive.

1) *Disadvantages of Existing System*

- The current techniques are with high security provider and takes lot of time to recover complex nature. In this way, previous techniques are not appropriate for the enormous information situation where information volume is huge and applications require online information preparing.
- The searching technique so for used is a time consuming in order to scan the whole documents collections in the cloud word by word.
- To avoid the deficiency of rank appliance, clients need to set aside a long opportunity to choose what they need when enormous records contain the inquiry watchword. Accordingly, the request saving methods is used to understand the rank Concept.

B. *Proposed System*

- Proposed system contains, a vector space model is utilized and each file is signified to by a vector, which implies each report can be viewed as a point in a high dimensional space. Because of the connection between various files, every one of the files subdivided into groups using multi keyword concept.
- By using ordinary course of action look method, a backtracking procedure is conveyed to look for goal records. At first server will discover diverse Files and get suitable sub records. By then cloud server will pick focused on k records from base needed sub-characterization. An estimation of k is previously picked by customer and sent to the cloud server. In occasion that present sub-characterization can't satisfy k reports, cloud server will take after back to its parent and select the pined for files from its kin classes.
- Check for reliability and result finding, obvious structure is developed based on the hash capacity is built.

1) *Advantages of Proposed System*

- Rank search concept introduced in our proposed system which takes less time by choosing the desired files along with keyword categories.
- Comparing with every one of the files with recorded documents, all number of reports stored in cloud in such case end user searchable is less. As technologies concerns the way we have done over classification of files can be additionally subdivided into few sub records.
- Information and classifications can be represented by virtual root. In order to check the query item related to file search from the cloud where confirming each root comparably better search than virtual root.

IV. IMPLEMENTATION

Implementation part includes the modules of the project which are divided so that it makes easy to understand and develop the project.

1) *Modules*

There are three modules that provide the detailed description about the project. They are listed below.

- Data Owner
- Cloud Server
- END User

a) Data Owner

Data Provider browser and uploads their encrypted files in the cloud that is drivehq and also has the capability to manipulate the encrypted file and performs the following operation that is to browse, encrypt and upload files, and also grant permission to the cloud consumer or end user.

Here first register the registration form by registering end user or data owner. After completion of registration user or data owner will get secret key with their registered mail id. Then it shows data owner login page with their credential along with secret key. The open source cloud that is Drive HQ to check the encrypted uploaded file will be stored on cloud using cloud credentials, and then it shows the list of encrypted file on cloud. The user will login with secret

key to check the credentials then will get the user home page, where we can search the files which were uploaded by the data owner using multi keyword along with user secret key. Then search files results by providing keyword, in order to download particular searched file click on file and request will be sent to the particular data owner and owner has to respond the request which we have sent. Again the data owner login with secret key and credentials, then will get data owner home page, click on requested file sent by user and respond to the same. The requested files details and also shows the number of file request details with response action and status. The user will login to download the files, then click on file download to view the download files, the files details of downloaded files and it shows the number of downloaded files. The key verification system in order to download particular file data owner has sent a decryption key to the user registered mail id.

b) Cloud Server

Cloud server manages data storage service for all owners those who registered. Registered owners can encrypt their files and saved those files into the drivehq by providing access to the user one who registered with cloud that is drivehq. Customer or user can download encrypted file by logged into the cloud with their credentials and decryption provided by the owner of the particular file which you want to download, the decryption key will be sent to the user mail id at the time of registration. User requests for file authorization to access and performs the following operations such as View all User Files, Give privileges to user View Search Transaction, View all attackers, View all End Users, View all Data Owners, Create Index on searched data and provide all related data related to corresponding keyword, View all android users.

Cloud server maintained the details of data owner and user. In cloud server home page, click on data owner details, data owner details cloud contains the details of the data owner and can be multiple data owners. And then click on user details cloud contains the details of the user and it can be multi user where, user can access any data owners' uploaded files from the cloud.

- Proxy Server: The role of proxy server is just to re-encrypt the owner upload file into the cipher text and outsource it to the user.
- Drive HQ: The CSP used in this work is DriveHQ which is an IT cloud service provider and have some unique features like Drive HQ file manager, Drive HQ online backup and cloud file sharing. Generator API's are used for the generation of a random six-digit secret key for both the user and the consumer and will be mailed to the registered email-ID.

c) End User

Users can get files from cloud and download using their secret key and credentials given by cloud at the time registration. User can searches for files based on contents keyword, request for secret key, request for downloading files and retrieve and store the data. The data which matches for a specified keyword will be indexed in the cloud server and then respond to the user.

Here first register the registration form by registering end user or data owner. After completion of registration user or data owner will get secret key with their registered mail id. Then it shows the user login page where user has to log in with his credentials along with secret key which we got whole registration, then will get user home page with password and logout link. To change the Password request, here we can change the password for security purpose.

V. CONCLUSION

The concept of the0privacy0preserving0keyword searching concept for the cloud0storage supports automated authorization cancellation. The conducted experiment and security Investigation demonstrates that this work holds a significant part higher security over the existing results with a traditional outlay for cloud requisitions. The competency inquiry shows that this work achieves high processing and storage ability together with its higher security, as compared with other classical searchable encryption schemes.

REFERENCES

[1] S.Grzonkowski, P. M. Corcoran, and T. Coughlin, "Security analysis of authentication protocols for next-generation mobile and CE cloud services," in Proc. IEEE Int. Conf. Consumer Electron., 2011, Berlin, Germany, 2011, pp. 83–87.

[2] D. X. D. Song, D. Wagner, and A. Perrig,"Practical techniques for searches on encrypted data," in Proc. IEEE Symp. Security Priv., BERKELEY, CA, 2000, pp. 44–55.

[3] D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in Proc. EUROCRYPT, Interlaken, SWITZERLAND, 2004, pp. 506–522.

[4] Y. C. Chang and M. Mitzenmacher, "Privacy preserving keyword searches on remote encrypted data," in Proc. 3rd Int.Conf. Applied Cryptography Netw. Security, New York, NY, 2005, pp. 442–455.

[5] E.-J. Goh, Secure Indexes, IACR Cryptology ePrint Archive, vol. 2003, pp. 216. 2003.

[6] A. Swaminathan, Y. Mao, G. M. Su, H. Gou, A. Varna, S. He, M. Wu, and D. Oard, "Confidentiality-preserving rank-ordered search," in Proc. ACM ACM Workshop Storage Security Survivability, Alexandria, VA, 2007, pp. 7–12.

[7] S. Zerr, D. Olmedilla, W. Nejdl, and W. Siberski, "Zerber+R: Topk retrieval from a confidential index," in Proc. 12th Int. Conf. Extending Database Technol.: Adv. Database Technol., Saint Petersburg, Russia, 2009, pp. 439–449.

[8] C. Wang, N. Cao, J. Li, K. Ren, and W. J. Lou, "Secure ranked keyword search over encrypted cloud data," in Proc. IEEE 30th Int. Conf. Distrib. Comput. Syst., Genova, ITALY, 2010, pp. 253–262.

[9] 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 Proc. 8th ACM SIGSAC Symp. Inform., Comput. Commun. Security, Hangzhou, China, 2013, pp. 71–82.

[10] I. H. Witten, A. Moffat, and T. C. Bell, Managing Gigabytes: Compressing and Indexing Documents and Images, 2nd ed. San Francisco, CA, USA: Morgan Kaufmann, 1999.

