ISSN:2348-2079

Volume-6 Issue-2

International Journal of Intellectual Advancements and Research in Engineering Computations

Enable Central Keyword Base Semantic Search using UserLevel Encryption Extension Model

[1]Ms. N. ZahiraJahan, M.C.A., M.Phil., Associate Professor/MCA
[2]Mr.M.Loganathan, III MCA
Department of MCA,Nandha Engineering College (Autonomous), Erode-52
E-Mail ID :zahirajahan1977@gmail.com , loganathan.mmurugan@gmail.com

Abstract: **A cloud system is difficult to synchronize login and authentication data between external clouds and internal systems without exposing internal security data. The cloud technologies are rapidly being adopted throughout the Information Technology (IT) due to their various attractive properties. In spite of their spread, they have raised a range of significant security and privacy concerns which interrupt their adoption in sensitive environments. In existing scheme makes a good tradeoff between the functionality and the efficiency. To better express the relevance between the query and files, we introduce the TF-IDF rule into our design.**

***Index Terms*— cloud computing, Keyword search, privacy Protection, semantic search,**

I. INTRODUCTION

Cloud computing is internet-based computing in which large groups of remote servers are networked to allow sharing of data-processing tasks, centralized data storage, and online access to computer services or resources. Clouds can be classified as public, private or hybrid. Cloud computing is a type of computing that relies on sharing computing resources rather than having local servers or personal devices to handle applications. The main enabling technology for cloud computing is virtualization. Virtualization software allows a physical computing device to be electronically separated into one or more "virtual" devices, each of which can be easily used and managed to perform computing tasks. Cloud computing adopts concepts from Service oriented Architecture (SOA) that can help the user break these problems into services that can be integrated to provide a solution. Cloud computing provides all of its resources as services, and makes use of the well-established standards and best practices gained in the domain of SOA to allow global and easy access to cloud services in a standardized way.

An individual or an organization may not require purchasing the needed storage devices. Instead they can store their data backups to the cloud and archive their data to avoid any information loss in case of hardware / software failures. Even cloud storage is more flexible, how the security and Privacyis available for the outsourced data becomes a serious concern. Preserving authorized restrictions on information access and disclosure. The main there at accomplished when storing the data with the cloud.

Guarding against improper information modification or destruction and ensuring timely and reliable access to and use of information.To achieve secure data transaction in cloud, suitable cryptography method is used. The data owner must encrypt the file and then store the file to the cloud. If a third person downloads the file, he/she may view the record if he/she had the key which is used to decrypt the encrypted file. Sometimes this may be failure due to the technology development and the hackers. To overcome the problem there are lot of techniques introduced to make secure transaction and secure storage.

Anonymous authentication is the process of validating the user without the details or attributes of the user. So the cloud server doesn’t know the details or identity of the user, which provides privacy to the users to hide their details from other users of that cloud. Sp that proposes a secure cloud storage using decentralized access control with anonymous authentication. The files are associated with file access policies, that used to access the files placed on the cloud. Uploading and downloading of a file to a cloud with standard Encryption/Decryption is more secure. Revocation is the important scheme that should remove the files of revoked policies. So no one can access the revoked file in future. The policy renewal is made as easy as possible.

OBJECTIVE

- To best knowledge, this is the first work to take the relationship between query keywords into consideration in searchable encryption designs.
- a new central keyword semantic extension ranked scheme (CKSER scheme) based on the keyword weight and multi-keyword ranked search.
- Two secure searchable encryption schemes to meet different privacy requirements under fuzzy authentication different threat models
- Multi-user key distribution scheme is proposed to provide data to the same group of users.
- Encryption cost and thereby data transmission cost is less.
- Different kind of encryption is maintained for various data saved in the cloud nodes based on the security level requirement.

II. LITERATURE SURVEY

Mohamed Al Morsy et al describes the Central Keyword-based Semantic Extension Search introduces a detailed analysis of the cloud security problem. To investigate the problem from the cloud architecture perspective, the cloud offered characteristics perspective, the cloud stakeholders' perspective, and the cloud service delivery models perspective. Based on this analysis they derive a detailed specification of the cloud security problem and key features that should be covered by any proposed security solution.

Cloud computing provides the next generation of internet based, highly scalable distributed computing systems in which computational resources are offered 'as a service'. The most widely used definition of the cloud computing model is introduced by NIST as "a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction.". Multi-tenancy and elasticity are two key characteristics of the cloud model. Multi-Tenancy enables sharing the same service instance among different tenants.

Jon Brodkin et al stated Cloud computing transforms the way information technology (IT) is consumed and managed, promising improved cost efficiencies, accelerated innovation, faster time-to-market, and the ability to scale applications on demand (Leighton, 2009).

According to Gartner, while the hype grew exponentially during 2008 and continued since, it is clear that there is a major shift towards the cloud computing model and that the benefits may be substantial (Gartner Hype-Cycle, 2012).

However, as the shape of the cloud computing is emerging and developing rapidly both conceptually and in reality, the legal/contractual, economic, service quality, interoperability, security and privacy issues still pose significant challenges. In this chapter describe various service and deployment models of cloud computing and identify major challenges.

In particular, they discuss three critical challenges: regulatory, security and privacy issues in cloud computing. Some solutions to mitigate these challenges are also proposed along with a brief presentation on the future trends in cloud computing deployment.

Ramgovind S stated that the purpose of the Central Keyword-based Semantic Extension Search is to provide an overall security perspective of Cloud computing with the aim to highlight the security concerns that should be properly addressed and managed to realize the full potential of Cloud computing. Gartner's list on cloud security issues, as well the findings from the International Data Corporation enterprise panel survey based on cloud threats, will be discussed in this paper.

The success of modern day technologies highly depends on its effectiveness of the world's norms, its ease of use by end users and most importantly its degree of information security and control. Cloud computing is a new and emerging information technology that changes the way IT architectural solutions are put forward by means of moving towards the theme of virtualization: of data storage, of local networks (infrastructure) as well as software.

Pooja et al focus on cloud data storage security, which has constantly been an important feature of quality of service Data owner's stores encrypted data in the cloud to ensure security for his data in the cloud computing environment and issues decryption key to only authorized user to access the data from cloud.

As user is revoked, data owner has to re-encrypt the data so that revoked user cannot access the data again .To perform this operation data owner will issue re-encryption command to cloud so that data in cloud gets re-encrypted. Once encryption is done here is a need for generation of new decryption keys to legal user, so that they can go on to access the data.

Melissa Chase et al considered the problem of encrypting structured data (e.g., a web graph or a social network) in such a way that it can be efficiently and privately queried. For this purpose, introduce the notion of structured encryption which generalizes previous work on symmetric searchable encryption (SSE) to the setting of arbitrarily-structured data. Present the model for structured encryption, the formal security definition and several efficient constructions.

The present schemes for performing queries are two simple types of structured data, specifically lookup queries on matrix-structured data, and search queries on labeled data. To show how these can be used to construct efficient schemes for encrypting graph data while allowing for efficient neighbor and adjacency queries.

III.SYSTEM METHODOLOGY

PROBLEM DEFINITION

The existing system has provided comprehensive information regarding the cloud security problems. It has been estimated the security problem from cloud architecture point of view, the cloud stakeholders' point of view and at the end from cloud services delivery models point of view. From stakeholder prospective, the security configurations need to be organized and each service should be maintained a level at runtime. From service delivery model prospective, the cloud management security issues and cloud access method security issues are also highlighted.

The existing system has presented details about the security issues which cloud service providers are facing when they dig deep for cloud engineering. There are some serious issues and challenges which cloud computing are facing in the domain of cyber security. The paper also covers security management models for the cloud service providers in order to meet security compliance.

The existing system has identified the serious threats and risks related to privacy and security for the mass and corporate users when they will integrate their mobile hand held devices with the cloud infrastructure.

A.SYSTEM MODEL

Input central keywords with certain adjunct words as the query keywords when searching documents. The importance of each query keyword depends on the search intension of a user. So far many works have demonstrated the importance of keywords. The super-increasing sequence is to show the preference factors of keywords to indicate the importance of keywords in a query keyword set.However, users need to sort keywords according to their importance, which increases the users' input cost. Due to the lack of the super-increasing sequence, the last keyword the user inputs are more important than all the other keywords. The existing model built a user interest model for individual user by analyzing his search history.

However, when inputting unusual keywords, it needs to re-build a new interest model. In this project, our use the grammatical relations as standards to show the weight of each keyword, and this enables users to retrieve relevant documents from the cloud based on their own interests.In addition with all the existing system mechanism, a correlated Authentication aspect with combination of the cloud storage provider, application service provider and end user is also considered. In addition, time limit is provided to end user to access the Application Service Providers (ASPs). So at different time intervals, different kinds of tariffs can be applied to end users to access the service. Likewise, the security aspects provided by the cloud storage provider is also taken by ASPs to increase the security more. In addition, trusted third party authentication mechanism is included.

B.EXPERIMENTAL SETUP

The proposed system is designed and implemented with the following processes:

- Proxy Server (PS) Module
- Data Owner
- Term Frequency – Inverse Document Frequency
- End User Search

C.SETUP

In this process, the admin user can able to add the cloud service provider details, application provider details and data owner details, the details which are stored into the corresponding tables in the data baseCloud Service Provider details includes the Cloud service provider id, name of the cloud provider, website and password details will be stored into the CSProviders table. Application Service Provider details includes the Application service provider id, name of the application service provider, password are stored into the ASProviders table. Data Owner details includes the data owner id, name of the data owner and password details are stored into the DataOwner table.Also the admin user assigns the Cloud Service Provider to Application Service Provider and Assign Cloud Service Provider to Data Owner. And the admin user can able to view the Cloud Service Providers details, Application Service Providers details, Data Owners Details; View Users details and view downloads details.

D.PROXY SERVER (PS)

A solution is to apply a distributed self-proxy re-encryption technique, propose a Proxy Server (PS). It coordinates and chooses keys by Key Manager (KM) whenever group membership changes. The distributed SPS provides not only encryption and decryption keys but also immediate re encryption keys for shared data. After communicating with KM, it automatically receives necessary keys from KM by self-created algorithm. A distributed SPS scheme is one solution where multiple proxies are automatically deployed in several clouds.

In this Module, the cloud service provider can able to login with their provided credentials and can able to View Application Service Provider Details, View Data Owner Details. In this module the Payment from Data Owner will be performed. The details includes of the payments are Cloud storage provider id, data owner id, date of payment, file details and the payment amount.

E.DATA OWNER

The Data Owner (DO) has data to be stored in the cloud and rely on the cloud for data computation, consists of both individual consumers and organizations. The data owner of

MCP shares data to many other cloud users. The data is encrypted with a key from KM and then stored in the cloud along with access control list indicating the user group. Upon access request from a user, the cloud communicates with SPS, based on access control list, and Self Proxy Server (SPS) requests for the key.

According to the key request to the SPS, uses re-encryption to transfer the encrypted format that can be decrypted by the user's private key. The user can download the encrypted data from the cloud and use the decryption key. In this data owner module, the data owner can View the Application service provider Details, View CSP Details after logged into the system. The data owner can also Upload Content to the cloud storage with the description of file description, category of the file and cloud storage provider details along with application service provider details. The data owner can also View the Download Request from the users and Provide Keys to the download request files by the end users.

F.TERMFREQUENCY–INVERSEDOCUMENTFREQUENCY (TF-IDF)

The TF measures how frequently a particular term occurs in a document. It is calculated by the number of times a word appears in a document divided by the total number of words in that document. It is computed as TF (the) = (Number of times term the 'the' appears in a document) / (Total number of terms in the document). The IDF measures the importance of a term. It is calculated by the number of documents in the text database divided by the number of documents where a specific term appears. While computing TF, all the terms are considered equally important. That means, TF counts the term frequency for normal words like "is", "a", "what", etc. Thus we need to know the frequent terms while scaling up the rare ones, by computing the following: IDF (the) = log_e(Total number of documents / Number of documents with term 'the' in it).

For example, consider a document containing 1000 words, wherein the word give appears 50 times. The TF for give is then (50 / 1000) = 0.05. Now, assume that, 10 million documents and the word give appears in 1000 of these. Then, the IDF is calculated as log(10,000,000 / 1,000) = 4. The TF-IDF weight is the product of these quantities − 0.05 × 4 = 0.20.

G.END USER SEARCH

In this process, the end user search the data can able to view the Cloud Storage Provider Details, Application Service Provider Details and Data Owner Details. The user can also Search for Content from the cloud storage and they can download the file by means of send request to the data owner to obtain the key to download the contents. The user can download the encrypted data from the cloud and use the decryption key.

IV.CONCLUSION

In this project describe the relationship among the query keywords into consideration and designed a keyword weighting algorithm based on the relations. We also designed a central keyword semantic extension scheme according to the keyword weights. By choosing the central keyword instead of not all the keywords to extend, our scheme achieves a tradeoff between functionality and efficiency. To express the relevance between the query and the files better, we introduced the TF-IDF rule when building the trapdoor and index. By storing the IDF value in the dictionary, our scheme can support updates for adding new files In this experimental study, the existing system is describing the problem of secure authentication for storage in cloud. In this project, proposed FA which carries out a flexible file-sharing scheme between an owner who stores the data in one cloud party and applications which are registered within another cloud party. The security analysis shows that our N-FA (Novel Fuzzy Authorized) scheme provides a thorough security of outsourced data, including confidentiality, integrity and secure access control.

.

V.FUTURE ENHANCEMENTS

At present, the project presented a client-side privacy protection framework called UPS for web search. For future work, the project will try to resist adversaries with broader background knowledge, such as richer relationship among topics (e.g., exclusiveness, sequentiality, and so on), or capability to capture a series of queries (relaxing the second constraint of the adversary) from the victim. It will also seek more sophisticated method to build the user profile, and better metrics to predict the performance (especially the utility) of UPS.

The following enhancements are should be in future.

- The application if developed as web services, then many applications can make use of the records.
- The data integrity in cloud environment is not considered. The error situation can be recovered if there is any mismatch.
- The web site and database can be hosted in real cloud place during the implementation.

REFERENCES

[1] Alliance for Telecommunications Industry Solutions. Homepage URL: http://www.atis.org.

[2] Amazon S3 Availability Event: (2008). URL: http://status.aws.amazon.com/s3-20080720.html (Accessed on November 29, 2012).

[3] AOL Apologizes for Release of User Search Data (2006). URL: ews.cnet.com/2010-1030_3- 6102793.html. August 7, 2006.

[4] Brodkin J, "Gartner: Seven Cloud-Computing Security Risks", Mar. 2009.

[5] Frank Gens, Robert P Mahowald and Richard L Villars. (2009, IDC Cloud Computing 2010.

[6] Khana. A.N Kiaha M. L.M., Khanb S.U. and. Madanic S. A, "Towards Secure Mobile Cloud Computing: A Survey", Future Generation Computer Systems, vol.29, Issues 5, July 2013.

[7] Patterson D.A., Rabkin A., StoicaI, Zaharia M, Above the clouds: a Berkeley view of cloud computing, Technical Report UCB/EECS-2009-28, EECS Department, University of California, Berkeley, Feb. 2009.