IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Vol. 7, Issue 4, April 2020

Challenges to Achieve Privacy and Secure Searchable Outsource Cloud Data Storage Services

Prof. Vaishali R Patil[1], Prof. Renuka R Gavli[2]

Lecturer, Computer Engineering, Bhivarabai Sawant Polytechnic, Pune, India[1,2]

Abstract: Cloud data have to be encrypted to protect data privacy, before outsourced to the commercial public cloud. The encryption process makes effective data utilization service a very challenging task. Traditional searchable encryption techniques allow users to securely search over encrypted data through keywords. They support only Boolean search and are not yet sufficient to meet the effective data utilization need that is inherently demanded by large number of users and huge amount of data files in cloud. The system facilitates server side ranking without keyword privacy. Search result authentication is provided in the system. The similarity analysis scheme is used to identify the query results under the cloud data storage.

Keywords: Encryption, Searchable encryption technique, Boolean search, Ranked Search

I. INTRODUCTION

Cloud computing is the long dreamed vision of computing as a utility, where cloud customers can remotely store their data into the cloud so as to enjoy the high quality networks, servers, applications and services from a shared pool of configurable computing resources . The advantages of cloud computing include on-demand self-service, ubiquitous network access, location independent resource pooling, rapid resource elasticity, usage-based pricing, transference of risk. The many advantages of cloud computing are increasingly attracting individuals and organizations to outsource their data from local to remote cloud servers. In addition to cloud infrastructure and platform providers, such as Amazon, Google and Microsoft, more and more cloud application providers are emerging which are dedicated to entering more accessible and user friendly data storage services to cloud customers. It is a clear trend that cloud data outsourcing is becoming a pervasive service. Along with the widespread enthusiasm on cloud computing, however, concerns on data security with cloud data storage are arising in terms of reliability and privacy which rise as the primary obstacles to the adoption of the cloud. To address these challenging issues, this dissertation explores the problem of secure and reliable data outsourcing in cloud computing. We aim at deploying the most fundamental data services including data management and data utilization, with built-in reliability and privacy assurance as well as high level service performance, usability, and scalability. Firstly, in addition to major cloud infrastructure providers, such as Amazon, Google, and Microsoft, more and more third-party cloud data service providers are emerging which are dedicated to offering more accessible and user friendly storage services to cloud customers. Secondly, to protect data privacy and combat unsolicited accesses in the cloud and beyond, sensitive data, e.g., emails, personal health records, tax documents, financial transactions, etc may have to be encrypted by data owners before outsourcing to the commercial public cloud

II. RELATED WORK

Cloud Computing economically enables a fundamental paradigm of data service outsourcing, where enterprises and organizations can benefit from lower up-front capital costs and less hands-on management. Despite the tremendous benefits, outsourcing data services to the commercial public cloud is also depriving customers' direct control over the systems that manage their data, raising security and privacy as the primary obstacles to the adoption of the cloud. To address these challenges and thus motivate the wide adoption of the cloud, how to safeguard the deployment of the most fundamental data services including data utilization, data storage, data sharing, and data computation outsourcing on the commercial public cloud, with built-in security and privacy assurance as well as high level data service performance, usability, and scalability.

Specifically, the following challenging questions are being investigated:

1. Privacy-assured and Effective Cloud Data Utilization: how to enable an encrypted cloud data search service with strong privacy-assurance, while enjoying high service-level performance inherently demanded by the large number of data users and huge amount data files in cloud. Two challenging research tasks are fuzzy and ranked keyword search over encrypted cloud data.

DOI 10.17148/IARJSET.2020.7416

2. Secure Cloud Storage Auditing: how to design efficient data integrity verification mechanisms for strong correctness assurance of cloud data storage service, given the challenge that data is no longer locally possessed by data owners. The design should be publicly auditable, privacy-preserving, and support data dynamics.
3. Scalable and Owner-controlled Cloud Data Sharing: how to enable data owners to reliably and efficiently enforce the dissemination of sensitive cloud data among large number of users in a fine-grained and scalable way, when the data reside in the open untrusted cloud environment.
4. Secure Data Computation Outsourcing: How can a computationally weak end-user securely outsource expensive data computation workloads to cloud, such that the mechanism protects both confidentiality of the sensitive workload information and integrity of the computation result while simultaneously ensuring end-users' substantial computational savings.

Design Goals

To provide secure and reliable cloud data storage services, our design should simultaneously achieve performance guarantees during data retrieval and repair.

- **Availability and Reliability**

By accessing any k-combination of n storage servers, the data user could successfully retrieve encoded data and recover all the original data. The data retrieval service remains functional when up to n − k storage servers are corrupted in one round, and corrupted servers can be repaired from other healthy servers.

- **Security**

The designed storage service protects the data confidentiality and periodically checks the integrity of data in cloud servers to prevent data dropout or corruption.

- **Online Data Owner**

Data owners can go onine immediately after data outsourcing, which means they are not required to be involved in tasks such as data integrity check and repair at a later stage.

- **Efficiency**

Above goals should be achieved with low storage, computation and communication cost for the data owner, data users and cloud servers.

III. PRIVACY-PRESERVING PUBLIC AUDITING FOR SECURE CLOUD STORAGE

Using Cloud Storage, users can remotely store their data and enjoy the on-demand high quality applications and services from a shared pool of configurable computing resources, without the burden of local data storage and maintenance. However, the fact that users no longer have physical possession of the outsourced data makes the data integrity protection in Cloud Computing a formidable task, especially for users with constrained computing resources. Moreover, users should be able to just use the cloud storage as if it is local, without worrying about the need to verify its integrity. Thus, enabling public audit ability for cloud storage is of critical importance so that users can resort to a third party auditor (TPA) to check the integrity of outsourced data and be worry-free. To securely introduce an effective TPA, the auditing process should bring in no new vulnerabilities towards user data privacy, and introduce no additional online burden to user. In this paper, we propose a secure cloud storage system supporting privacy-preserving public auditing. We further extend our result to enable the TPA to perform audits for multiple users simultaneously and efficiently. Extensive security and performance analysis show the proposed schemes are provably secure and highly efficient.

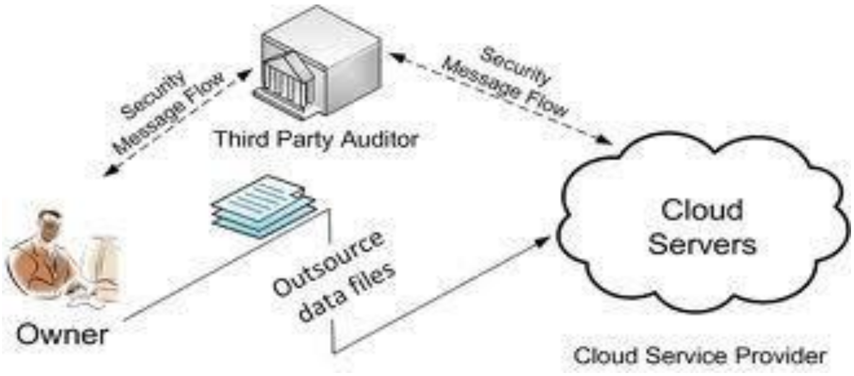

Fig. 1 Third Party auditing system

DOI 10.17148/IARJSET.2020.7416

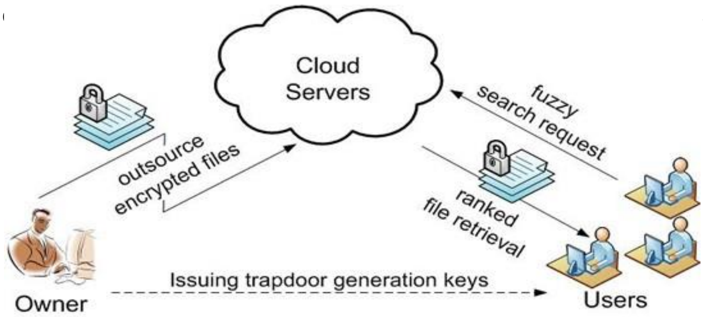

Fig. 2 Cloud data search outsourcing Model

As the data produced by enterprises and individuals that need to be stored and utilized is rapidly increasing, data owners are motivated to outsource their local complex data management systems into the cloud for its great flexibility and economic savings. To protect data privacy and combat unsolicited accesses in cloud and beyond, sensitive data has to be encrypted before outsourcing .To explore such a privacy-assured and effective cloud data utilization service with high service-level performance and usability, by investigating the two challenging research tasks: fuzzy keyword search and ranked keyword search over encrypted cloud data.

Fuzzy keyword search, opposing to exact keyword match, tolerates minor typos and format inconsistencies in user search request, and greatly enhances system usability and user searching experience. Its challenge lies in the fact that two words similar to each other would no longer be so after one-way cryptographic transformation (for encrypted keyword search).
Ranked keyword search further ensures the file retrieval accuracy and allows the user to find the most/least relevant information efficiently. We explore the statistical measure approach (i.e. relevance score) from information retrieval (IR), and properly hide the scores in an order-preserved manner. The resulting design is expected to facilitate efficient server-side ranking without losing keyword privacy.

Secure and Proof-carry Computation Outsourcing

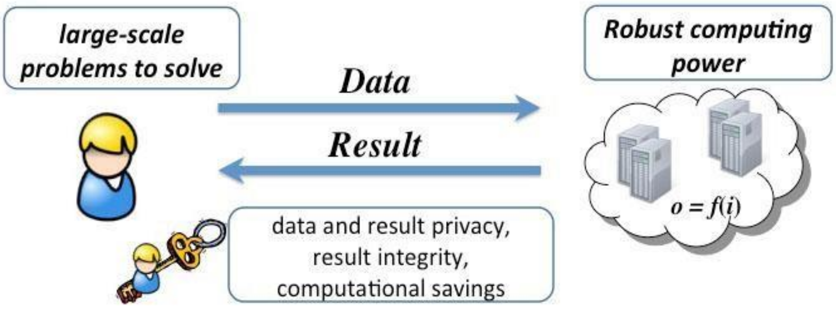

Fig. 3 Secure data computational model

A fundamental concern to move computational workloads from private resources to the cloud is the protection of the confidential data that the computation consumes and produces. Thus, secure computation outsourcing services are in great need to not only protect sensitive workload information but validate the integrity of the computation result. This is, however, a very difficult task due to a number of challenges that have to be met simultaneously. Firstly, such a service has to be practically feasible (immediate practicality) in terms of computational complexity. Secondly, it has to provide sound security guarantee without restricted system assumptions. Thirdly, it also has to enable substantial computational savings at the end-user's side as compared to the amount of the efforts that otherwise has to be committed to solve the problem locally. These challenges practically exclude the applicability of the existing techniques developed in the context of secure multi-party computation and fully homomorphic encryption.
With above challenges in mind, our proposed approach is to understand the nature of an application and its security requirements and develop application-specific solutions that are highly customized and achieve desirable trade-offs among privacy protection, performance, and other factors. We start from widely applicable engineering computing and optimization problems, which are essential for modern engineering designs, and usually require a substantial amount of computational power and involve confidential data. Aiming for a practical solution, our proposed methodology is to explicitly decompose computations into public programs running on the cloud and private data owned by the users. By organizing computation problems at various abstraction levels into a hierarchy, it is possible to leverage the structures of specific computations for achieving desirable trade-offs among security, efficiency, and practicality in a systematic manner. Critical applications that we are currently investigating include secure outsourcing large scale systems of linear equations, linear programming, and convex optimization in the cloud

Privacy-preserving Search for Cloud Data

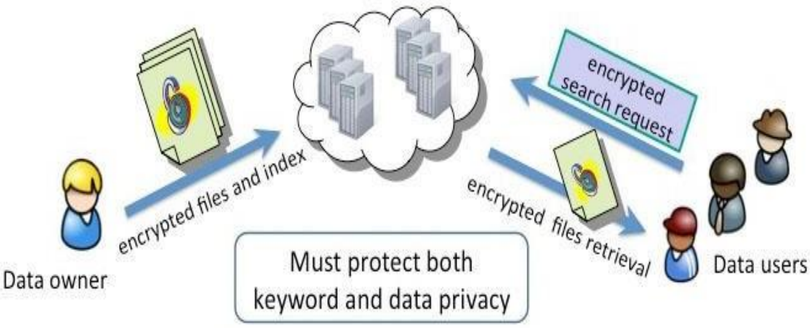
encrypted search request
encrypted files and index
encrypted files retrieval
Data owner
Data users
Must protect both keyword and data privacy

Fig. 4 Keyword and data privacy

When data services are increasingly outsourced to cloud for its greater flexibility and cost efficiency, sensitive data has to be encrypted before outsourcing to combat unsolicited accesses in cloud and beyond. However, the encryption makes deployment of traditional data utilization service, such as plaintext keyword search over textual data or query over database, a difficult task. Downloading all the data and decrypting locally is clearly impractical. Besides, aside from eliminating the local storage management, storing data into the cloud serves no purpose unless they can be easily searched and utilized. This necessitates the need for developing effective searching techniques over encrypted cloud data of massive scale. Such techniques should enable critical search functionalities that have long been enjoyed in modern search engine over unencrypted data, like Google, Bing, etc. The adequacy of such techniques is essential to the long- term success of the cloud services and the ultimate privacy protection of both individuals and organizations.
Our proposed research starts from enabling versatile keyword search over encrypted data that is highly usable, including the functionalities like fuzzy tolerance, result ranking, multi-keywords search, similarity search, etc. Beyond textual data, we also propose to enable privacy-preserving search over all kind of non-textual data, including search over graph-structured data, image, and/or multimedia, which are ubiquitous in modern life and are driving many new applications. Efficient techniques for searching such high dimensional encrypted data have to be developed. Our ultimate goal is to enable rich search semantics in a privacy preserving manner and efficiently support for large-scale and distributed nature of cloud data.

IV. CONCLUSION

In this chapter, for the first time, we define and solve the problem of query over encrypted graph-structured cloud data, and establish a variety of privacy requirements. For the efficiency consideration, we adopt the principle of “filtering-and verification” to prune as many negative data graphs as possible before verification, where a feature-based index is pre-built to provide feature-related information for every encrypted data graph. Then, we choose the inner product as the pruning tool to carry out the filtering procedure efficiently. To meet the challenge of supporting graph semantics, we propose a secure inner product computation technique, and then improve it to achieve various privacy requirements under the known-background threat model. Thorough analysis investigating privacy and efficiency of our schemes given, and the evaluation further shows our scheme introduces low overhead on computation and communication.

REFERENCES

[1]. S. Zerr, D. Olmedilla, W. Nejdl, and W. Siberski, =“Zerber+r: Top-k Retrieval from a Confidential Index,” Proc. Int’l Conf. Extending Database Technology: Advances in Database Technology (EDBT ’09), 2009.
[2]. J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, “Fuzzy Keyword Search over Encrypted Data in Cloud Computing,” Proc. IEEE Infocom 10, 2010.
[3]. N. Cao, C. Wang, K. Ren, and W. Lou, “PrivacyPreserving Multi-Keyword Ranked Search over Encrypted Cloud Data,” Proc. IEEE Infocom 11, 2011.
[4]. C. Wang, K. Ren, S. Yu, K. Mahendra, and R. Urs, “Achieving Usable and Privacy-Assured Similarity Search over Outsourced Cloud Data,” Proc. IEEE INFOCOM, 2012.
[5]. Cong Wang, Ning Cao, Kui Ren and Wenjing Lou, “Enabling Secure and Efficient Ranked Keyword Search over Outsourced Cloud Data” IEEE Transactions On Parallel And Distributed Systems, Vol. 23, No. 8, August2012.
[6]. Cong Wang, Qian Wang, Kui Ren, and Wenjing Lou, "Towards Secure and Dependable Storage Services in Cloud Computing," To appear, IEEE Transactions on Service Computing (TSC).

DOI 10.17148/IARJSET.2020.7416