ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Secured Confidential Query Services Mechanism in the Cloud Computing

Raid Abd Alreda Shekan Alsamaraii
Master of Science (Information System),
Nizam College (Autonomous), O.U,
Basheer Bagh, Hyderabad.

T. Ramdas Naik
Assistant Professor Dept, Computer Science (PG),
Nizam College (Autonomous), O.U,
Basheer Bagh, Hyderabad.

Abstract:

With the wide deployment of public cloud computing infrastructures, using clouds to host data query services has become an appealing solution for the advantages on scalability and cost-saving. However, some data might be sensitive that the data owner does not want to move to the cloud unless the data confidentiality and query privacy are guaranteed. On the other hand, a secured query service should still provide efficient query processing and significantly reduce the in-house workload to fully realize the benefits of cloud computing. We propose the random space perturbation (RASP) data perturbation method to provide secure and efficient range query and KNN query services for protected data in the cloud.

The RASP data perturbation method combines order preserving encryption, dimensionality expansion, random noise injection, and random projection, to provide strong resilience to attacks on the perturbed data and queries. It also preserves multidimensional ranges, which allows existing indexing techniques to be applied to speedup range query processing. The KNN-R algorithm is designed to work with the RASP range query algorithm to process the KNN queries. We have carefully analyzed the attacks on data and queries under a precisely defined threat model and realistic security assumptions. Extensive experiments have been conducted to show the advantages of this approach on efficiency and security. Characteristics Of Cloud Computing: The salient characteristics of cloud computing based on the definitions provided by the National Institute of Standards and Terminology (NIST) are outlined below:

On-demand self-service: A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service's provider. Broad network access: Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, laptops, and PDAs). Resource pooling:

The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. There is a sense of location-independence in that the customer generally has no control or knowledge over the exact location of the provided resources but may be able to specify location at a higher level of abstraction (e.g., country, state, or data center). Examples of resources include storage, processing, memory, network bandwidth, and virtual machines.

Rapid elasticity: Capabilities can be rapidly and elastically provisioned, in some cases automatically, to quickly scale out and rapidly released to quickly scale in. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be purchased in any quantity at any time. Measured service: Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be managed, controlled, and reported providing transparency for both the provider and consumer of the utilized service [1].

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Advantages:

1. Price: Pay for only the resources used.

2. Security: Cloud instances are isolated in the network from other instances for improved security.

3. Performance: Instances can be added instantly for improved performance. Clients have access to the total resources of the Cloud's core hardware.

4. Scalability: Auto-deploy cloud instances when needed.

5. Uptime: Uses multiple servers for maximum redundancies. In case of server failure, instances can be automatically created on another server.

6. Control: Able to login from any location. Server snapshot and a software library lets you deploy custom instances.

7. Traffic: Deals with spike in traffic with quick deployment of additional instances to handle the load

LITERATURE SURVEY:

"Security Modeling And Analysis AUTHORS: J. Bau and J.C. Mitchell [9]"

Security modeling centers on identifying system behavior, including any security defenses; the system adversary's power; and the properties that constitute system security. Once a security model is clearly defined, security analysis evaluates whether the adversary, interacting with the system, can defeat the desired security properties. Although the authors illustrate security analysis using model checking, analysts can use various methods and tools to evaluate system security, including manual and automated theorem-proving tools that provide assurance about the absence of attacks in a specified threat model. This article describes a uniform approach for evaluating system security and illustrates the approach by summarizing three case studies. Security modeling and analysis also provides a basis for comparative evaluation and some forms of security metrics. "Privacy-Preserving Multi-Keyword Ranked Search Over Encrypted Cloud Data AUTHORS: N. Cao, C. Wang, M. Li, K. Ren, and W. Lo [10] "

With the advent of cloud computing, data owners are motivated to outsource their complex data management systems from local sites to the commercial public cloud for great flexibility and economic savings. But for protecting data privacy, sensitive data have to be encrypted before outsourcing, which obsoletes traditional data utilization based on plaintext keyword search. Thus, enabling an encrypted cloud data search service is of paramount importance. Considering the large number of data users and documents in the cloud, it is necessary to allow multiple keywords in the search request and return documents in the order of their relevance to these keywords.

Related works on searchable encryption focus on single keyword search or Boolean keyword search, and rarely sort the search results. In this paper, for the first time, we define and solve the challenging problem of privacy-preserving multi-keyword ranked search over encrypted data in cloud computing (MRSE). We establish a set of strict privacy requirements for such a secure cloud data utilization system. Among various multi-keyword semantics, we choose the efficient similarity measure of "coordinate matching," i.e., as many matches as possible, to capture the relevance of data documents to the search query. We further use "inner product similarity" to quantitatively evaluate such similarity measure.

We first propose a basic idea for the MRSE based on secure inner product computation, and then give two significantly improved MRSE schemes to achieve various stringent privacy requirements in two different threat models. To improve search experience of the data search service, we further extend these two schemes to support more search semantics. Thorough analysis investigating privacy and efficiency guarantees of proposed schemes is given. Experiments on the real-world data set further show proposed schemes indeed introduce low overhead on computation and communication.

"Geometric Data Perturbation For Outsourced Data Mining AUTHORS: K. Chen and L. Liu [11] "

Data perturbation is a popular technique in privacy-preserving data mining.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

A major challenge in data perturbation is to balance privacy protection and data utility, which are normally considered as a pair of conflicting factors. We argue that selectively preserving the task/model specific information in perturbation will help achieve better privacy guarantee and better data utility. One type of such information is the multidimensional geometric information, which is implicitly utilized by many data-mining models. To preserve this information in data perturbation, we propose the Geometric Data Perturbation (GDP) method. In this paper, we describe several aspects of the GDP method. First, we show that several types of well-known data-mining models will deliver a comparable level of model quality over the geometrically perturbed data set as over the original data set. Second, we discuss the intuition behind the GDP method and compare it with other multidimensional perturbation methods such as random projection perturbation. Third, we propose a multi-column privacy evaluation framework for evaluating the effectiveness of geometric data perturbation with respect to different level of attacks. Finally, we use this evaluation framework to study a few attacks to geometrically perturbed data sets. Our experimental study also shows that geometric data perturbation can not only provide satisfactory privacy guarantee but also preserve modeling accuracy well.

"Towards Attack-Resilient Geometric Data Perturbation AUTHORS: K. Chen, L. Liu, and G. Sun [12]"

Data perturbation is a popular technique for privacy preserving data mining. The major challenge of data perturbation is balancing privacy protection and data quality, which are normally considered as a pair of contradictive factors. We propose that selectively preserving only the task/model specific information in perturbation would improve the balance. Geometric data perturbation, consisting of random rotation perturbation, random translation perturbation, and noise addition, aims at preserving the important geometric properties of a multidimensional dataset, while providing better privacy guarantee for data classification modeling. The preliminary study has shown that random geometric perturbation can well preserve model accuracy for several popular classification models, including kernel methods, linear classifiers, and SVM classifiers, while it also revealed some security concerns to random geometric perturbation.

In this paper, we address some potential attacks to random geometric perturbation and design several methods to reduce the threat of these attacks. Experimental study shows that the enhanced geometric perturbation can provide satisfactory privacy guarantee while still well preserving model accuracy for the discussed data classification models.

"RASP: Efficient Multidimensional Range Query On Attack-Resilient Encrypted Databases AUTHORS: K. Chen, R. Kavuluru, and S. Guo [13]"

In this paper we propose the random space encryption approach to efficient range queries over encrypted data and analyze the unique attacks to this approach. Our approach uses a random space transformation to generate indexable auxiliary data. The auxiliary data is exported to the service provider, indexed and used for processing range queries. We present an efficient server-side two-stage query processing strategy. Experimental results show that this processing strategy is highly efficient. In addition, we analyzed the attacks on encrypted data and queries. Experiments are performed to show the resilience of the encryption to estimation attacks. Note that this attack analysis is just the first step to rigorous analysis of security.

" Private Information Retrieval AUTHORS: B. Chor, E. Kushilevitz, O. Goldreich, and M. Sudan [14] "

Publicly accessible databases are an indispensable resource for retrieving up-to-date information. But they also pose a significant risk to the privacy of the user, since a curious database operator can follow the user's queries and infer what the user is after. Indeed, in cases where the users' intentions are to be kept secret, users are often cautious about accessing the database. It can be shown that when accessing a single database, to completely guarantee the privacy of the user, the whole database should be down-loaded; namely n bits should be communicated (where n is the number of bits in the database).In this work, we investigate whether by replicating the database, more efficient solutions to the private retrieval problem can be obtained. We describe schemes that enable a user to access k replicated copies of a database ($k \geq 2$) and privately retrieve information stored in the database.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

This means that each individual server (holding a replicated copy of the database) gets no information on the identity of the item retrieved by the user.

PROBLEM STATEMENT:

Requirements for constructing a practical query service in the cloud as the CPEL criteria: data confidentiality, query privacy, efficient query processing, and low in-house processing cost. Satisfying these requirements will dramatically increase the complexity of constructing query services in the cloud. Some related approaches have been developed to address some aspects of the problem. The crypto index and order preserving encryption (OPE) are vulnerable to the attacks. The enhanced crypto index approach puts heavy burden on the in-house infrastructure to improve the security and privacy.

Disadvantages Of Existing System :

» Do not satisfactorily addressing all aspects of Cloud.
» Increase the complexity of constructing query services in the cloud.
» Provide slow query services as a result of security and privacy assurance.

PROBLEM DEFINITION:

• We propose the random space perturbation (RASP) data perturbation method to provide secure and efficient range query and kNN query services for protected data in the cloud.
• The RASP data perturbation method combines order preserving encryption, dimensionality expansion, random noise injection, and random projection, to provide strong resilience to attacks on the perturbed data and queries.

Advantages Of Proposed System :

» The RASP perturbation is a unique combination of OPE, dimensionality expansion, random noise injection, and random projection, which provides strong confidentiality guarantee.
» The RASP approach preserves the topology of multidimensional range in secure transformation, which allows indexing and efficiently query processing.

The proposed service constructions are able to minimize the in-house processing workload because of the low perturbation cost and high precision query results. This is an important feature enabling practical cloud-based solutions.

IMPLEMENTATION:

User Module: In this module, Users are having authentication and security to access the detail which is presented in the ontology system. Before accessing or searching the details user should have the account in that otherwise they should register first. Multidimensional Index Tree Most multidimensional indexing algorithms are derived from R-tree like algorithms, where the axis-aligned minimum bounding region (MBR) is the construction block for indexing the multidimensional data. For 2D data, an MBR is a rectangle. For higher dimensions, the shape of MBR is extended to hyper-cube. The MBRs in the R-tree for a 2D dataset, where each node is bounded by a node MBR. The R-tree range query algorithm compares the MBR and the queried range to find the answers.

Performance Of KNN-R Query Processing In this set of experiments, we investigate several aspects of KNN query processing. (1) We will study the cost of (k, δ)-Range algorithm, which mainly contributes to the server-side cost. (2) We will show the overall cost distribution over the cloud side and the proxy server. (3) We will show the advantages of KNN-R over another popular approach: the Casper approach for privacy-preserving KNN search.

Preserving Query Privacy Private information retrieval (PIR) tries to fully preserve the privacy of access pattern, while the data may not be encrypted. PIR schemes are normally very costly. Focusing on the efficiency side of PIR, Williams et al. use a pyramid hash index to implement efficient privacy preserving data-block operations based on the idea of Oblivious RAM. It is different from our setting of high throughput range query processing. Hu et al. addresses the query privacy problem and requires the authorized query users, the data owner, and the cloud to collaboratively process KNN queries. However, most computing tasks are done in the user's local system with heavy interactions with the cloud server. The cloud server only aids query processing, which does not meet the principle of moving computing to the cloud.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

CONCLUSION:

We propose the RASP perturbation approach to hosting query services in the cloud, which satisfies the CPEL criteria: data confidentiality, query privacy, efficient query processing, and low in-house workload. The requirement on low in-house workload is a critical feature to fully realize the benefits of cloud computing, and efficient query processing is a key measure of the quality of query services. RASP perturbation is a unique composition of OPE, dimensionality expansion, random noise injection, and random projection, which provides unique security features. It aims to preserve the topology of the queried range in the perturbed space, and allows to use indices for efficient range query processing.

With the topology-preserving features, we are able to develop efficient range query services to achieve sublinear time complexity of processing queries. We then develop the kNN query service based on the range query service. The security of both the perturbed data and the protected queries is carefully analyzed under a precisely defined threat model. We also conduct several sets of experiments to show the efficiency of query processing and the low cost of in-house processing. We will continue our studies on two aspects: 1) further improve the performance of query processing for both range queries and kNN queries; and 2) formally analyze the leaked query and access patterns and the possible effect on both data and query confidentiality.

REFERENCES:

[1] P. Mell and T. Grance, ,,,,The NIST Definition of Cloud Computing, National Institute of Standards and Technology, Gaithersburg, MD, USA, Tech. Rep., 2009.
[2] Http://Explainingcomputers.Com/Cloud.Html.
[3] Sunita Rani And Ambrish Gangal, " Cloud Security With Encryption Using Hybrid Algorithm And Secured Endpoints ", (Ijcsit) ,Vol. 3 (3) , 2012 ,4302 – 4304.
[4] Introduction To Cloud Computing Architecture White Paper 1st Edition, June 2009
[5] " Advance Computer Technology " A Book By Dr. Deven Shah. Edition- 2011.
[6] Ramgovind S, Eloff Mm, Smith E ," The Management Of Security In Cloud Computing ", School Of Computing, University Of South Africa, Pretoria, South Africa ©2010 .
[7]Http://Searchcloudsecurity.Techtarget.Com/Tip/Clou D-Computing-Security-Choosing-A-Vpn-Type-To-Connect-To-The-Cloud.
[8] J. Kubiatowicz, D. Bindel, Y. Chen, P. Eaton, D. Geels, R.Gummadi, S. Rhea, H. Weatherspoon, W. Weimer, C. Wells, Andb. Zhao, " Oceanstore: An Architecture For Global-Scale Persistent Storage, " Proc. Ninth Intl Conf. Architectural Support For Programming Languages And Operating Systems (Asplos), Pp. 190- 201, 2000 . Secured Confidential Query Services Mechanism REFERENCES In The Cloud Computing 118 .
[9] J. Bau and J.C. Mitchell, "Security Modeling and Analysis," IEEE Security and Privacy, vol. 9, no. 3, pp. 18-25, May/June 2011.
[10] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data," Proc. IEEE INFOCOMM, 2011.
[11] K. Chen and L. Liu, "Geometric Data Perturbation for Outsourced Data Mining," Knowledge and Information Systems, vol. 29, pp. 657- 695, 2011.
[12] K. Chen, L. Liu, and G. Sun, "Towards Attack-Resilient Geometric Data Perturbation," Proc. SIAM Intl Conf. Data Mining, 2007.
[13] K. Chen, R. Kavuluru, and S. Guo, "RASP: Efficient Multidimensional Range Query on Attack-Resilient Encrypted Databases," Proc. ACM Conf. Data and Application Security and Privacy, pp. 249-260, 2011.
[14] B. Chor, E. Kushilevitz, O. Goldreich, and M. Sudan, "Private Information Retrieval," ACM Computer Survey, vol. 45, no. 6, pp. 965-981, 1998.
[15] R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, "Searchable Symmetric Encryption: Improved Definitions and Efficient Constructions," Proc. 13th ACM Conf. Computer and Comm. Security, pp. 79-88, 2006.
[16] H. Hacigumus, B. Iyer, C. Li, and S. Mehrotra, "Executing SQLover Encrypted Data in the Database-Service-Provider Model," Proc. ACM SIGMOD Intl Conf. Management of Data (SIGMOD), 2002.
[17] Z. Huang, W. Du, and B. Chen, "Deriving Private Information from Randomized Data," Proc. ACM SIGMOD Intl Conf. Management of Data (SIGMOD), 2005.

AUTHORS BIOGRAPHY:

Raid Abd Alreda Shekan Alsamaraii,-+pursuing his Master of Science in Information System, from Nizam College (Autonomous),O.U, Basheer Bagh, Hyderabad, India.