Preserving Data in Cloud Computing

Nasrin Dalil[1] and Ahmed Kayed[2]

**[1] College of Computer Science and Information Technology, Sudan University of Science and Technology (SUST)
Khartoum, Sudan**

**[2] Faculty of Information Technology, Middle East University (MEU)
Amman, Jordan**

Abstract

Cloud computing is basically virtual pool of resources and it provides these resources to its users via internet as services who use them as when needed basis. On demands of user's data confidentiality and privacy the data is stored in cloud server as encrypted. Ideally, to maintain the security of user's data queries should be processing over encrypted data. There are different schemes recommended to support this issue. One of them is Ordered Preserving encryption (OPE) which allows comparison operations to be directly executed over encrypted data. This paper surveys various encryption functions provided to support querying over encrypted data, especially Order Preserving Encryption.

Keywords: *Cloud Server, Security, Order Preserving Encryption (OPE).*

1. Introduction

Cloud computing is the emerging field in the modern era. Cloud computing is defined as the set of resources or services offered through the internet to the users on their demand by cloud providers. It conveys everything as a service over the internet based on user demand, for instance operating system, network hardware, storage, resources, and software. As each and every organization is moving its data to the cloud, means it uses the storage service provided by the cloud provider. There is a need to protect that data against unauthorized access, modification or denial of services etc [1].

The security problems with the outsourced databases can be solved if the critical data are encrypted. By this way, the service provider or an attacker only can see the meaningless encrypted data. However, after encrypted, a database cannot be easily queried. The problem of how the service provider can process the queries is arises. It is not acceptable to decrypt the entire database before performing each query because the decryption might be very slow for a large database and the decrypted database is again at the risk of having its security and privacy breached. Ideally, a query should be executed directly over the encrypted database [2].

This paper concentrates on various encryption algorithms which used for processing queries over encrypted data to support better security. And it focuses on order preserving encryption algorithm, since the ordered encrypted data allows comparison operations and range queries to be perform over encrypted data. In fact [3], OPE not only allows efficient range queries, but allows indexing and query processing to be done exactly and as efficiently as for unencrypted data. The paper organized as follow, section 2 is cloud computing overview. Section 3 is about encryption of outsourcing data that used to protect data in the cloud. Finally, section 4 is about order preserving encryption.

2. Cloud Computing

Cloud computing refers to applications and services that run on a distributed network using virtualized resources and accessed by common Internet protocols and networking standards. It is distinguished by the notion that resources are virtual and limitless and that details of the physical systems on which software runs are abstracted from the user [4]. Cloud computing is basically broken down into three segments: "application" "storage" and "connectivity." Each segment serves a different purpose and offers different products for businesses around the world [5].

In cloud computing, resources are provided as a service over the Internet to customers who use them as when needed basis [6]. As many organizations outsourcing their data to the cloud, they need it to be secure and confidential. In addition, rather than outsourcing our data to the cloud server, we need to provide security to it. Therefore, security is a biggest concern of users when using cloud computing [7]. Security goals of data includes [7], Authentication, Authorization, Auditing, Confidentiality, Integrity, Availability and Non-repudiation.

3. Encryption of Outsourcing Data

Outsourcing data to cloud server leads to some security problems. The fact that data owners and cloud server are no longer in the same trusted domain may put the outsourced unencrypted data at risk. The cloud server may leak data information to unauthorized entities or even be hacked [8]. Therefore, prevention of unauthorized access to outsourced data is represented by encryption.

However, after encryption, the data isn't easy to query. Ideally, to maintain the security of data in cloud server a query should be executed directly over the encrypted data. To make this acceptable the data should be encrypted in some way to satisfy performing the queries over encrypted data like search and join queries. The following are some techniques to clarify this idea:

3.1 Deterministic (DET)

A deterministic encryption scheme (as opposed to a randomized probabilistic encryption scheme[9]) is a cryptosystem which always produces the same ciphertext for the same plaintext, even over separate executions of the encryption algorithm. Examples of deterministic encryption algorithms include the RSA cryptosystem (without encryption padding), and many block ciphers when used in ECB mode or with a constant initialization vector [10].

Deterministic encryption permits logarithmic time search on encrypted data, while randomized encryption only allows linear time search, meaning a search requires scanning the whole database. This difference is crucial for large outsourced databases which cannot afford to slow down search [11]. This encryption type allows the server to perform equality checks, which means it can perform selects with equality predicates, equality joins, GROUP BY, COUNT, DISTINCT, etc [9].

3.2 Word search (SEARCH)

As The plaintext keyword search techniques are not suitable for the cloud computing, the secure searchable encryption scenarios were developed. They are followed by indexing the each keyword in encrypted data file and by associating the indexed file with the keyword [12].

The system proposes in [12] return the matching data file when user searching input keywords exactly matching the predefine keywords or the closest matching files if there exist typos and/or format inconsistency in the searching input. They use distance editing and wildcard_based techniques to construct it. In [13], a system for Secured Multi-keyword search (SMS) over encrypted cloud data (ECD). This keyword search technique allows users to selectively retrieve encrypted files from cloud server. It able cloud server to sends back only top-k documents that are most relevant to the search query. [14] Shows a system for Ranked keyword search that provides security and efficiency. Data owner outsources data and index table in encrypted form. When the server receives a query it encrypts the query (by using the same encryption algorithm that used to encrypt data) and compares it with the encrypted keyword in the index table. SEARCH in [9] is used to perform searches on encrypted text to support operations such as MySQL's LIKE operator. The encryption does not reveal to the DBMS server whether a certain word repeats in multiple rows, but it leaks the number of keywords encrypted with SEARCH; an adversary may be able to estimate the number of distinct or duplicate words (e.g., by comparing the size of the SEARCH and RND ciphertexts for the same data) [9].

3.3 Homomorphic encryption (HOM)

Homomorphic encryption is the conversion of data into ciphertext that can be analyzed and worked with as if it were still in its original form. It allows complex mathematical operations to be performed on encrypted data without compromising the encryption. Homomorphic encryption is expected to play an important part in cloud computing, allowing companies to store encrypted data in a public cloud and take advantage of the cloud provider's analytic services [15].

There are two types of homomorphic encryption: fully homomorphic encryption (FHE) and somewhat homomorphic encryption (SHE) [16]. Each type differs in the number of operations that can be performed on encrypted data. FHE allows for an unlimited, arbitrary number of computations (both addition and multiplication) to be performed on encrypted data. SHE cryptosystems support a limited number of operations (i.e., any amount of addition, but only one multiplication) and are faster and more compact than FHE cryptosystems [16].

3.4 Order Preserving Encryption (OPE)

In this scheme, the i^{th} value in the plaintext domain is mapped to the i^{th} value in the ciphertext domain, such that the order between plaintexts is preserved between ciphertexts [2]. Order Preserving Encryption (OPE) allows performing order comparisons, so it ensures that ciphertexts retain the order established between plaintexts. So, if a field is encrypted in this way, SQL range queries can still be performed efficiently [17]. The server can also perform ORDER BY, MIN, MAX, SORT, etc [9].

The rest of the paper focuses on the Order Preserving Encryption (OPE). Since the ordered encrypted data

increases the speed of search operation and allows performing order comparisons over encrypted data while maintaining reasonable degree of security.

4. Order-preserving symmetric encryption (OPE)

The important class of methods to realize search is by using order preserving encryption (OPE) schemes. An OPE scheme is a deterministic symmetric-key encryption scheme that preserves the order of the plaintexts [18]. The concept of order-preserving symmetric encryption (OPE) was proposed in the database community by Agrawal et al. [19].

The reason for new interest in such schemes is that they allow efficient range queries on encrypted data. That is, a remote untrusted database server is able to index the (sensitive) data it receives, in encrypted form, in a data structure that permits efficient range queries [3]. It allows comparison operations to be directly applied on encrypted data without decrypting the operands. Thus, equality and range queries as well as the MAX, MIN, and COUNT queries can be directly processed over encrypted data. Similarly, GROUP BY and ORDER BY operations can also be applied. Only when applying SUM or AVG to a group do the values need to be decrypted [19].

OPE is a weaker encryption scheme than DET because it reveals order [9]. It is not a perfectly secure encryption scheme since ciphertexts inevitably leak the order information of the plaintexts [18]. An OPE scheme can therefore be a good choice when it is necessary to simultaneously maintain a reasonable performance for range query processing and to achieve a certain degree of security [18].

At the server side, there are different styles to preserve the order of outsourced encrypted data. A reasonable way to preserve the order of encrypted data is by arranging the ciphertexts in alphabetical or numerical order, as in [3]. To understand the idea behind this, consider the following encryption scheme explained in [19]: Generate |P| unique values from a user-specified target distribution and sort them into a table T. The encrypted value c_i of p_i is then given by $c_i = T[i]$. That is, the i^{th} plaintext value in the sorted list of |P| plaintext values is encrypted into the i^{th} value in the sorted list of |P| values obtained from the target distribution [19].

Sequential scan may not be efficient enough when the data size is large [20]. To deal with this problem another scheme is proposes. To preserve the order of encrypted data in cloud server another scheme is to have the encoded values organized at the server in a search tree [21]. Interestingly, one common form of tree construction is the binary tree (it has logarithmic worst-case cost for insert, delete, and lookup). The ciphertexts are arranged in the tree based on the order of the plaintext values, Called the OPE Tree. We can see that the plaintext values in the left subtree of each node are smaller than the node value, and the values in the right subtree are larger [21].

To understand the tree construction we illustrate the MIT's example in [21], which called mutable Order Preserving Encryption (mOPE). They use b_tree to construct ordered values. Figure1 is the MIT's example to explain their approach. The client will help the server find the location in the tree where to insert a value. To illustrate how the client does this, suppose the client wants to encode 55 using the tree state shown in Fig. 1. The client first requests the root node of the tree, and the server returns x93d12a, which the client decrypts to 32. Since 55 > 32, the client requests the right child of the root from the server, and the server responds with x27716c, which the client decrypts to 69. Finally, the client requests the left child of the last node requested, and the server responds that there is no such child. This means that the client can insert a new node in this position, containing the DET encryption of 55[21].

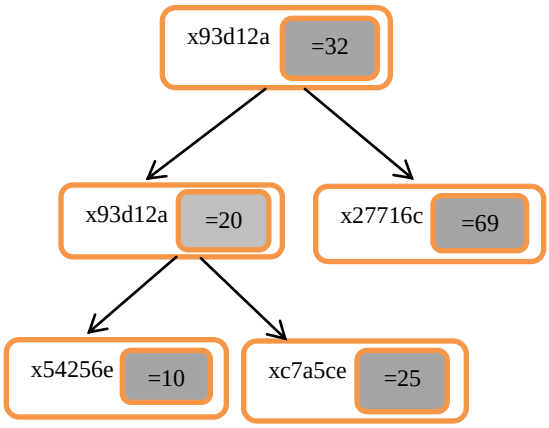

Figure 1. [21] Overview of mOPE's data structures. Each node in the OPE Tree contains a DET ciphertext (the hexadecimal value); for the reader's information, the gray block shows the corresponding plaintext value, but this is not stored in the tree.

Table [1] summarizes the advantages and disadvantages of preserving order of encrypted data in cloud server as it mention earlier.

Table 1 advantages and disadvantages of preserving order of encrypted data.

Advantages	*Disadvantages*
Allows indexing processing	Leak the order information of plaintext.
Allows order and comparison relation on encrypted database (equality, range, MIN, MAX and COUNT queries, ORDER BY and GROUP BY operations can be directly processed over encrypted data).	Doesn't perform SUM or AVG to a group of values unless it decrypted.
Provide security.	
Increase the performance of the search operations.	

5. Conclusion

To meet the security goals for the outsourcing data it should be outsourced as encrypted. Moreover, queries should be performed over encrypted data. There are different encryption functions which satisfy processing queries over encrypted data. Importantly, ordered encrypted data facilitate the comparisons and relation operation between data items.

References

[1] Rashmi Nigoti, Manoj Jhuria and Dr.Shailendra Singh, "A Survey of Cryptographic Algorithms for Cloud Computing", International Journal of Emerging Technologies in Computational and Applied Sciences (IJETCAS), 2013, ISSN (Print): 2279-0047.

[2] Dongxi Liu and Shenlu Wang, "Programmable Order-Preserving Secure Index for Encrypted Database Query", Proceeding in IEEE Fifth International Conference on Cloud Computing, 2012.

[3] Alexandra Boldyreva, Nathan Chenette, Younho Lee and Adam O'Neill, "Order Preserving Symmetric Encryption", In Proceedings of the 28th Annual International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT}, Cologne,Germany, April 2009 .

[4] Barrie Sosinsky, "Cloud Computing Bible", Wiley Publishing Inc, 2011.

[5] Ramandeep Kaur, Pushpendra Kumar Pateriya, "Encryption Techniques for Secure Database Outsourcing", Proceeding in Biskup, J., Lopez, J. (Eds.)ESORICS, Springer-Verlag, Berlin Heidelberg, 2007 .

[6] Safiriyu Eludiora, Olatunde Abiona, Clement Onime and Lawrence Kehinde, "A user Identity Management Protocol for Cloud Computing Paradigm", Int. J. Communications Network and System Sciences, Vol. 4, 2011, 152-163.

[7] MANDEEP KAUR and MANISH MAHAJAN, "Using Encryption Algorithms to Enhance the Data Security in Cloud Computing", International Journal of Communication and Computer Technologies, , Volume 01 – No.12, January 2013 , Issue: 03, ISSN NUMBER: 2278-9723.

[8] Treesa Maria Vincent and J.Sakunthala, "Data Storage in Cloud Environment Enhance Privacy", International Journal of Computer Trends and Technology, ISSN: 2231-2803, volume4, Issue3.

[9] Raluca Ada Popa, Catherine M. S. Redfield, Nickolai Zeldovich, and Hari Balakrishnan, "CryptDB: Protecting Confidentiality with Encrypted Query Processing ", In: Proceedings of the 23rd ACM Symposium on Operating Systems Principles (SOSP), 2011.

[10] http://www.Wikipedia.com/Deterministic encryption, May 2014.

[11] MIHIR BELLARE, MARC FISCHLIN, ADAM O'NEILL and THOMAS RISTENPART, "Deterministic encryption: Definitional Equivalences and Constructions without Random Oracles", CRYPTO 08, Lecture Notes in Computer Science, D. Wagner ed., Springer-Verlag, 2008 ,Vol. xx.

[12] V. Sravan Kumar Reddy and Professor C. Rajendra, "EKST: Efficient Keyword Searching Technique for Encrypted Data in CipherCloud", International Journal of Advanced Research in Computer Engineering Technology, Vol. 1, Issue 4, June 2012, ISSN: 2278 – 1323.

[13] C. R. Barde, Pooja Katkade, Deepali Shewale and Rohit Khatale, "Secured Multiple-keyword Search over Encrypted Cloud Data", International Journal of Emerging Technology and Advanced Engineering, Vol. 4, Issue 2, February 2014, ISSN 2250-2459.

[14] P.Rama Lakshmi, Nagabhushana Rao Chunduru and Amarendra Kothalanka,"An Efficient Encrypted Data Searching Over Out Sourced Data", International Journal of Engineering Trends and Technology (IJETT), Vol. 4, Issue 10, Oct 2013, ISSN: 2231-5381.

[15] Homomorphic encryption, http://searchsecurity.techtarget.com/definition, June 2014.

[16] Research Directorate staff, "Securing the cloud with Homomorphic encryption", The Next Wave, 2014, Vol. 20 No.3.

[17] Santi Mart´ınez∗, Josep M. Miret, Rosana Tom`as and Magda Valls, “Security Analysis of Order Preserving Symmetric Cryptography”, Applied Mathematics & Information Sciences, 2013, Vol. 7, No. 4.

[18] Liangliang Xiao and I-Ling Yen, “Security Analysis for Order Preserving Encryption Schemes”, IEEE, 2012.

[19] Rakesh Agrawal, Jerry Kiernan, Ramakrishnan Srikant and Yirong Xu, “Order Preserving Encryption for Numeric Data “, 2004.

[20] Dawn Xiaodong, Song David Wagner, Adrian Perrig, "Practical Techniques for Searches on Encrypted Data", Proceeding in IEEE Symp. Security and Privacy, 2000.

[21] Raluca Ada Popa, Frank H. Li, Nickolai Zeldovich, "An Ideal-Security Protocol for Order-Preserving Encoding", In: Proceedings of the 34th IEEE Symposium on Security and Privacy (SP), 2013.