International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2017 IJSRCSEIT | Volume 2 | Issue 2 | ISSN : 2456-3307

# Security Analysis on Cloud Data Search by using one to many Order Preserving Encryption

**Karishma Pathan, Krushna Phad, Nisha Waikar, Sonal Yenpure**
KJEI's Trinity Academy of Engineering, Pune, Maharashtra, India

## ABSTRACT

Cloud computing economically enables the paradigm of data service out-sourcing. However, to protect data privacy, sensitive cloud data have to be encrypted before outsourced to the commercial public cloud, which makes effective data utilization service a very challenging task .For ranked search in encrypted cloud data, order preserving encryption (OPE) is an efficient tool to encrypt relevance scores of the inverted index. When using deterministic OPE, the cipher texts will reveal the distribution of relevance scores. Therefore it is called one-to-many OPE, for applications of searchable encryption, which can flatten the distribution of the plain texts. We proposed a differential attack on one-to-many OPE by exploiting the differences of the ordered cipher texts. The experimental results show that the cloud server can get a good estimate of the distribution of relevance scores by a differential attack. Thus sensitive data have to be encrypted before being outsourced to a commercial public cloud. By using recurrence score we can retrieve the file from cloud.

**Keywords :** Access controls, Authentication, Cryptographic controls, Information flow controls, Invasive software (e.g. viruses, worms, Trojan horses), Security kernels, Verification.

## I. INTRODUCTION

Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model is composed of five essential characteristics, three service models, and four deployment models[1].

Cloud computing economically enables the paradigm of data service outsourcing. However, to protect data privacy, sensitive cloud data have to be encrypted before outsourced to the commercial public cloud, which makes effective data utilization service a very challenging task. For ranked search in encrypted cloud data, order preserving encryption (OPE) is an efficient tool to encrypt relevance scores of the inverted index. When using Deterministic OPE, the cipher texts will reveal the distribution of relevance scores. Therefore, proposed a probabilistic OPE, called one-to-many OPE, for applications of searchable encryption, which can flatten the distribution of the plaintexts. We proposed a differential attack on one-to-many OPE by exploiting the differences of the ordered cipher texts. The experimental results show that the cloud server can get a good estimate of the distribution of relevance scores by a differential attack. Furthermore, when having some background information on the outsourced documents, the cloud server can accurately infer the encrypted keywords using the estimated distributions. Thus sensitive data have to be encrypted before being outsourced to a commercial public cloud. By using recurrence score we can retrieve the file from cloud data to an untrusted third party. Thus, sensitive data have to be encrypted before being outsourced to a commercial public cloud.

Encryption on sensitive data presents obstacles to the processing of the data. Information retrieval becomes difficult in the encrypted domain because the amount of

outsourced files can be very large and traditional search patterns can not be deployed to cipher text retrieval directly. Users need to download all the data, decrypt it all, and then search keywords like plaintext retrieval. To overcome this, Searchable Encryption (SE) was proposed to make query in the encrypted domain possible while still preserving users' privacy. There are several problems in searchable encryption: fuzzy search, ranked search, multi-keyword search and so on. First proposed a search scheme only supporting single Boolean keyword search. After that plenty of searchable encryption methods arose to improve efficiency and reduce communication overhead.

Applying order preserving encryption (OPE) [4] is one practical way of supporting fast ranked search. This algorithm was first proposed in 2004 to solve encrypted query problems in database systems. OPE is a symmetric cryptosystem, therefore it is also called order-preserving symmetric encryption (OPSE). The order-preserving property means that if the plaintexts x1 < x2, then the corresponding cipher texts E(x1) and E(x2) satisfy E(x1) < E(x2).

However, the security definition and the constructions of OPE in and are based on the assumption that OPE is a deterministic encryption scheme which means that a given plaintext will always be encrypted as a fixed cipher text. However, deterministic encryption leaks the distribution of the plaintexts, so it cannot ensure data privacy in most applications. For instance, in privacy-preserving keywords search, OPE is used to encrypt relevance scores in the inverted index [5]. As noted by Wang et al. [5], when using a deterministic OPE, the resulting cipher text shares exactly the same distribution as the relevance score, by which the server can specify the keywords . Therefore, Wang et al. [5] improved the OPE in and proposed a "One-to-Many OPE" in their secure keyword search scheme, where they tried to construct a probabilistic encryption scheme and conceal the distribution of the plaintexts. we discover that the One-to-Many OPE [5] cannot ensure the expected security. In fact, although the Cipher texts of One-to-Many OPE conceals the distribution

**Table -1:** Example of Posting list of the Inverted Index

| Keyword | $hash(w)$ | | | |
|---|---|---|---|---|
| File ID | $F_1$ | $F_2$ | ... | $F_{f_w}$ |
| Relevance Score | $E'(8.6)$ | $E'(6.1)$ | ... | $E'(7.3)$ |

of the plaintexts, an adversary may estimate the distribution from the differences of the ciphertexts. So in this paper, we propose a differential attack on the One-to-Many OPE. Our experimental results show that, when applying this attack to the secure keyword search scheme of [5], the cloud server can get an estimation of the distribution of the the relevance scores, and furthermore accurately reveal the encrypted keywords.

## II. METHODS AND MATERIAL

### 1. Plaintext Searching Model

In practice, to realize effective data retrieval on large amount of documents, it is necessary to perform relevance ranking on the results. Ranked search can also significantly reduce network traffic by sending back only the most relevant data. In ranked search, the ranking function plays an important role in calculating the relevance between files and the given searching query. The most popular relevance score is defined based on the model of T F×I DF, where term frequency (TF) is the number of times a term (keyword) appears in a file and inverse document frequency (IDF) is the ratio of the total number of files to the number of files containing the term. There are many variations of T F × I DF-based ranking functions.

### 2. Ciphertext Searching Model

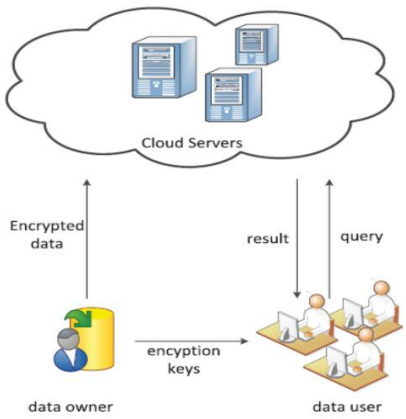


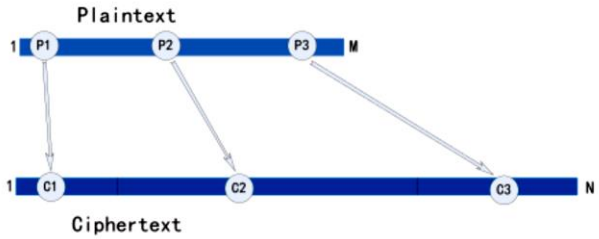

**Figure 1.**Framework of retrieval over encrypted cloud data.

Due to the special background of cloud computing, unlike traditional plaintext information retrieval, there are usually three entities in cloud data retrieval as shown in Fig-1: data owner, remote cloud server and users. A data owner can be an individual or a corporation, i.e., it is the entity that owns a collection

of documents Dc = {D1, D2 . . . DNd} that it wants to share with trusted users. The keyword set is marked as W = {w1,w2 . . .wNw}. For security and privacy concerns, documents have to be encrypted into ξ = {E(D1), E(D2) . . . E(DNd )} before being uploaded to the cloud server. Additionally, the plaintext index has to be encrypted into I to prevent information leakage.

### 3. Order Preserving Encryption

OPE is a symmetric cryptosystem, so it is also called order-preserving symmetric encryption (OPSE). The order preserving property means that if the plaintexts have such a relationship as x1 < x2, then the corresponding cipher texts E(x




1) and E(x2) satisfy E(x1) < E(x2).

**Figure 2.** Deterministic OPE

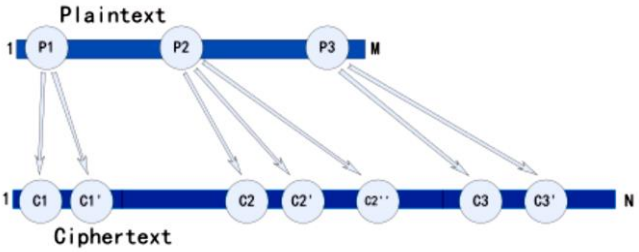



**Figure 3.** One-to-Many OPE

### 2. Literature Survey

[1] S. Subashini and V. Kavitha, "A survey on security issues in service delivery models of cloud computing," 2011, in this paper, a survey of the different security risks that pose a threat to the cloud is presented. This paper is a survey more specific to the different security issues that has emanated due to the nature of the service delivery models of a cloud computing system.

[2]A. Boldyreva, N. Chenette and A. ONeill, "Order-preserving encryp-tion revisited: improved security analysis and alternative solutions," 2011, this paper propose a simple and efficient transformation that can be applied to any OPE scheme. Our analysis shows that the transformation yields a scheme with improved security in that the scheme resists the one-wayness and window one-wayness attacks.

[3] L. Xiao, I.-L Yen, "Security analysis for order preserving encryption schemes,"2012, in this paper we analyze the security of the OP E encryption scheme SEm,n and give the upper bound on the probability for the adversary to recover the plain text encrypted by SEm,n under chosen plain text attacks.

[4] C. Wang, N. Cao and K. Ren, "Enabling secure and efficient ranked keyword search over outsourced cloud data" 2012, in this paper, he define and solve the problem of secure ranked keyword search over encrypted cloud data. Ranked search greatly enhances system usability by enabling search result relevance ranking instead of sending undifferentiated results, and further ensures the file retrieval accuracy.

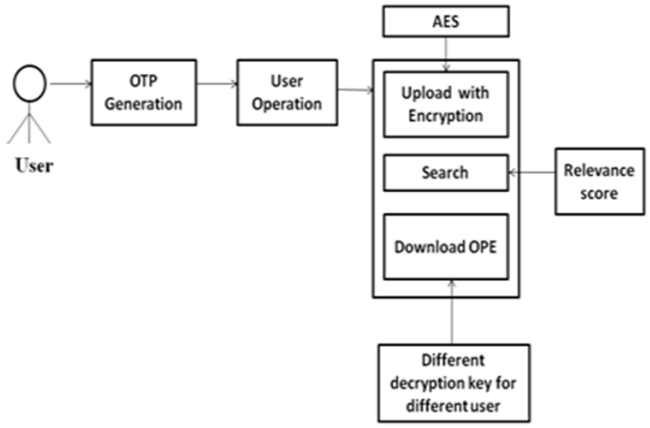
[5] S. Yu, C. Wang and K. Ren, "Achieving secure, scalable, and fine-grained data access control in cloud computing", 2010, this paper addresses this challenging open issue by, on one hand, defining and enforcing access policies based on data attributes, and, on the other hand, allowing the data owner to delegate most of the computation tasks involved in fine grained data access control to untrusted cloud servers without disclosing the underlying data contents. he achieve this goal by exploiting and uniquely combining techniques of attribute-based encryption.

### 3. Algorithm's to be Used

- ✓ Binary Search.
- ✓ RSA algorithm
- ✓ One to many order preserving encryption

## III. RESULTS

System Architecture




## IV. CONCLUSION

One-to-Many OPE is designed for encrypted data over the cloud and to preserve the order of relevance scores. cloud server can estimate the distribution of relevance scores by change point analysis on the differences of cipher texts of One-to-Many OPE. In this system we have described to improve One-to-Many OPE using this method. The system provides query privacy in search process under encrypted cloud data services. Search duration is reduced in the semantic relationship based encrypted keyword search process. Accuracy is improved with relevance score and semantic query model.

## V. FUTURE SCOPE

In future we will try to Provide Security using one-to-many Order Preserving Encryption for audio and video data on cloud.

## VI. REFERENCES

[1] S. Subashini and V. Kavitha, "A survey on security issues in service delivery models of cloud computing," J. Netw. Comput. Appl., vol. 34, no. 1, pp. 1–11, 2011.

[2] A. Boldyreva, N. Chenette, Y. Lee, and A. O'Neill, "Order-preserving symmetric encryption," in Advances in Cryptology. Berlin, Germany: Springer-Verlag, 2009, pp. 224–241.

[3] A. Boldyreva, N. Chenette, and A.O'Neill, "Order-preserving encryption revisited: Improved security analysis and alternative solutions,"in Advances in Cryptology. Berlin, Germany: Springer-Verlag, 2011, pp. 578 595.

[4] L. Xiao and I.-L. Yen, "Security analysis for order preserving encryption schemes," in Proc. 46th Annu. Conf. Inf. Sci. Syst., Mar. 2012, pp. 1–6

[5] C. Wang, N. Cao, K. Ren, and W. Lou, "Enabling secure and efficient ranked keyword search over outsourced cloud data," IEEE Trans. Parallel Distrib. Syst., vol. 23, no. 8, pp. 1467–1479, Aug. 2012.

[6] S. Yu, C. Wang, K. Ren, and W. Lou, "Achieving secure, scalable, and fine-grained data access control in cloud computing," in Proc. IEEE INFOCOM, Mar. 2010, pp. 1–9..