International Journal of Research Available at https://edupediapublications.org/journals

e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue 23
December 2018

An ideal-security protocol for Numeric-Related SQL Queries over encrypted data

[1] P Swapna Madhuri Ray, [2] P. Gangadhara

[1]*M. Tech, Dept of CSE, ShriShirdiSai Institute of Science and Engineering, Affiliated to JNTUA, Ananthapuramu, AP, India.* swapnamadhuripdwh@gmail.com

[2]*AssistantProfessor, Dept of CSE, ShriShirdiSai Institute of Science and Engineering, Affiliated to JNTUA, Ananthapuramu, AP, India.* gangadhara115208@gmail.com.

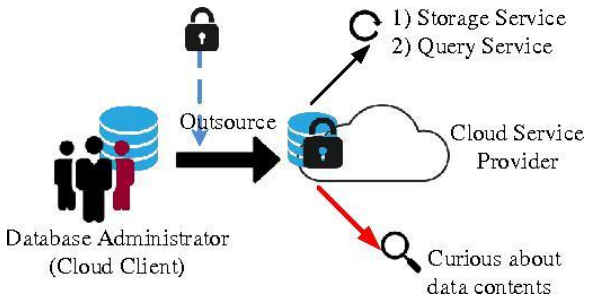
***ABSTRACT*: Industries and individuals outsource database to realize convenient and low-cost applications and services. In order to provide sufficient functionality for SQL queries, many secure database schemes have been proposed. However, such schemes are vulnerable to privacy leakage to cloud server. The main reason is that database is hosted and processed in cloud server, which is beyond the control of data owners. For the numerical range query (">", "<", etc.), those schemes cannot provide sufficient privacy protection against practical challenges,e.g., privacy leakage of statistical properties, access pattern. Furthermore, increased number of queries will inevitably leak more information to the cloud server. Here, we propose a two-cloud architecture for secure database, with a series of intersection protocols that provide privacy preservation to various numeric-related range queries. Security analysis shows that privacy of numerical information is strongly protected against cloud providers in our proposed scheme.**

Index Term: database, range query, privacy preserving, cloud computing

I.INTRODUCTION

THE growing industry of cloud has provide a service paradigm of storage/computation outsourcing helps to reduce users' burden of IT infrastructure maintenance, and reduce the cost for both the enterprises and individual users [1], [2], [3].

One straightforward approach to mitigate the security risk of privacy leakage is to encrypt the private data and hide the query/access patterns. Unfortunately, as far as we know, few academia researches satisfy both properties so far. CryptDB is the first attempt to provide a secure remote database application, which guarantees the basic confidentiality and privacy requirement, and provides diverse SQL queries over encrypted data as well. CryptDB uses a series of cryptographic tools to achieve this security functionality. Especially, order preserving encryption [15] is utilized to realize numeric related range query processes. From the perspective of query functionality, CryptDB supports most kinds of numerical SQL queries with such cryptology. However, such privacy leakage hasn't been well addressed thoroughly, since OPE is relatively weak to provide sufficient privacy assurance



e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue 23
December 2018

Some specific purpose cryptology like order preserving encryption(OPE) will expose some private information to the cloud service provider naturally: As it is designed to preserve the order on ciphertexts so that it can be used to conduct range queries, the order information of the data, the statistical properties derived therefrom, such as the data distribution, and the access pattern will be leaked.while the knowledge of query pattern is well partitioned into two parts, and knowing only one cannot reveal any private information; 2) We then present a series of intersection protocols to provide numeric-related SQL range query with privacy preservation, and especially, such protocols will not expose order-related information to any of the two non-colluding clouds.

II.RELATED WORK

Fuzzy query over encrypted data is becoming a popular topic, since in practical scenarios, some query requests usually want to retrieve data with similar, rather than exactly same indexes. Fuzzy searchable encryption has been introduced for cloud computing in many literatures. These schemes deal with the issue that search keywords allows small-scaled distinction in character/numeric level. Specifically for numerical keywords, the query predicate can get numerical records within a range. Some schemes targeted at spatial query, especially knn ,which focus on the distance between the query vector and the data. They usually inquire about certain spatial objects (or several numerical attributes) related to the others within a certain distance. Range query has been proposed for that purpose. . However, such existing range query schemes are not suitable for practical secure database due to high storage overhead to maintain the corresponding ciphertext.

III. SYSTEM ARCHITECTURE

SECURITY ASSUMPTION AND SECURITY REQUIREMENTS A. System Architecture Our proposed secure database system includes a database administrator, and two non-colluding clouds. In this model, the database administrator can be implemented on a client's side from the perspective of cloud service. The two clouds (refer to Cloud A and Cloud B), as the server's side, provide the storage and the computation service.

The two clouds work together to respond each query request from the client/authorized users (availability). For privacy concerns, these two clouds are assumed to be non-colluding with each other, and they will follow the intersection protocols to preserve privacy of data and queries (privacy). In our scheme, the knowledge of stored database and queries is partitioned into two parts, respectively stored in one cloud. The mechanism guarantees that knowing either of these two parts cannot obtain any useful privacy information. As shown in Fig. 2(a), to conduct a secure database, data are encrypted and outsourced to be stored in one cloud (Cloud A), and the private keys are stored in the other one (Cloud B). For each query, the corresponding knowledge includes the data contents and the relative processing logic. We utilize a prototype of knowledge partition, dividing application logic into two parts, which is firstly proposed by Bohli et al. in The application logic, as a secret knowledge, is partitioned into two parts, each of which is only known to one cloud. Intuitively, this two-cloud architecture increases some complexity to some extent, and we will analyze and point out that this overhead is acceptable in Section

B. **Security Assumption** Following the general assumption of many related works in public cloud, we assume the clouds to be honest-but-curious: On one hand, both of the two clouds will respond with correct information in the interactions of our proposed scheme (honest); on the other hand, the clouds try their best to obtain private information

e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue 23
December 2018

from the data that they process (curious). From the perspective of privacy assurance, here the data not only include permanently stored information (i.e., database), but also each temporary query request (i.e., queries). Additionally and importantly, as the assumption in some existing works, we assume that the two clouds A and B are non-colluding: Cloud A follows the protocol to add required obfuscation to protect privacy against cloud B, so that cloud B cannot obtain additional private information in the interactions with Cloud A. No private information is delivered beyond the scopes of protocols.

C. Potential Threats and Privacy Requirements

This section describes the potential threats and the privacy requirements when the database is outsourced to public cloud. The stored data contents and the query processes. Although there are many data encryption schemes, some fail to provide sufficient privacy preservation after statistical analysis: Repeated and large-amount query processes not only leak the access patterns but also disclose the stored encrypted data progressively.

Data contents The privacy of data contents includes the definition and description of each column (column name) in the table of the stored database, and the values of each record in the table. Some related works have mainly focused on this issue, in which the column names are blinded (such as CryptDB) and meanwhile the values are encrypted with some other encryption techniques (such as Order Preserving Encryption) and some deterministic encryption schemes, so that the adversaries cannot easily and directly guess the meaning of the column, or the values of the data. However, in an outsourced database, utilizing encryption alone, without other mechanisms, is far from being enough to preserve the privacy of the data contents. With the development of data analysis, by extracting features from data and queries, classification technique can help understand the definition of columns, and then breach of confidentiality of data contents.

Statistical properties. Besides the static properties candisclose the private information of data contents, such properties themselves are already sensitive and private for the client. Order Preserving Encryption(OPE), which is widely used in constructing the secure database, with support of range queries, directly exposes the statistical information in the encryption field. Furthermore, the leakage of statistic properties is part of the nature of outsourced cloud databaseservice: the cloud can learn the statistical properties (like order) by repeated query requests. As an example, Fig. 3 describes such an attack: After two simple queries over one same column, the order relationship of some data in certain column can be determined. There are also some other direct and indirect scenarios to leak statistical properties. In this way, even though the order property is not exposed to the semi-trusted cloud at the beginning, the cloud can gradually find out the order information after many query requests.

Query pattern. The query pattern also contains privacyinformation, as they can reveal the client's purpose of the query. Even worse, such pattern can leak some statistical properties, as discussed above.Based on the above discussion, we assert that an outsourced secure database providing numeric-related queries should prevent the following private information from being obtained by the honest-but-curious clouds.

Data contents. The data contents includes item values andcolumn names, which are the raw data that should be protected against any potential adversaries

Statistical properties. It includes the order of data and their probability distributions.

e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue 23
December 2018

Query pattern. Each query should be kept private against the honest-but-curious clouds and any unauthorized parties. The secrecy of such pattern should be well preserved even after many query processes.

IV. OUR PROPOSED TWO-CLOUD SCHEME

In thissection, we firstly give an overview of our proposed two-cloud scheme, and then present the detailed interaction protocols to realize range query with privacy preservation on outsourced encrypted database. A. Overview In our scheme, two clouds (refer to Cloud A and Cloud B, respectively) have been assigned distinct tasks in the database system: Cloud A provides the main storage service and stores the encrypted database. Meanwhile, Cloud B executes the main computation task, to figure out whether each numerical record satisfies the client's query request with its own security key. With the assumption of no collusion between two clouds, the knowledge of application logic can be partitioned into two parts in our proposed scheme, where each one part is only known to one cloud. As we will analyze Here, one single part of knowledge cannot reveal privacy of the data and the query. Based on the two-cloud architecture, our scheme provides an approach to query numeric-related data with privacy preservation. The client can retrieve the desired data from the cloud.The proposed mechanism can preserve the privacy of data and query requests against each of the two clouds. Specifically, Cloud A only knows the query request type and the final indexes, but due to dummy items appending, Cloud A cannot accurately understand the finally satisfied index set for each single request. Meanwhile, in order to prevent Cloud A from launching multiple specific-purpose query requests to deliberately to seek more knowledge about the data, we introduce a token based scheme, which can restrict the number of items and the range of columns that Cloud A can only process. For Cloud B, it knows the satisfied indexes of each single request, but after the proposed operations, it does not know the relationship of the corresponding items. Moreover, Cloud B can hardly distinguish whether two received columns are generated from one or more columns in the original database.

V.CONCLUSION

Here, we presented a two-cloud architecture with aseries of interaction protocols for outsourced databaseservice,which ensures the privacypreservation of datacontents, statistical properties and query pattern. At the sametime, with the support of range queries, it not only protectsthe confidentiality of static data, but also addresses potential privacy leakage in statistical properties or after large number of query processes. Security analysis shows that our scheme can meet the privacy-preservation requirements.
Further more,performance evaluation result shows that our proposed scheme is efficient. In our future work, we will consider to further enhance the security while ensuring practicality, and we will extend our Proposed scheme to support more operations, such as "SUM/AVG".

REFERENCES

[1]KAIPINGXUE, SENIOR MEMBER, IEEE, SHAOHUA LI, JIANAN HONG, YINGJIE XUE, NENGHAI YU, AND PEILIN HONG"TWO-CLOUD SECURE DATABASE FOR NUMERIC-RELATED SQL RANGEQUERIES WITH PRIVACY PRESERVING".IEEETRANSACTIONSONINFORMATI ONFORENSICSAND SECURITY,YEAR: 2017, VOLUME: 12, ISSUE: 7

[2]M.ARMBRUST, A. FOX, R. GRIFFITH, A. D. JOSEPH ET AL., "A VIEW OF CLOUD COMPUTING," COMMUNICATIONS OF THE ACM, VOL. 53, NO. 4, PP. 50–58, 2010.

[3]C.WANG, Q.WANG, K.REN, N.CAO, AND W. LOU,"TOWARD SECURE AND DEPENDABLE STORAGE SERVICES IN CLOUD COMPUTING," IEEE TRANSACTIONS ON SERVICES COMPUTING, VOL.5, NO. 2, PP. 220–232, 2012.

[4]K.XUE AND P. HONG, "A DYNAMIC SECURE GROUP SHARING FRAMEWORK IN PUBLIC CLOUD COMPUTING," IEEE TRANSACTIONS ON CLOUD COMPUTING, VOL. 2, NO. 4, PP. 459–470, 2014.

[5]J. W. RITTINGHOUSE AND J. F. RANSOME, CLOUD COMPUTING: IMPLEMENTATION, MANAGEMENT, AND SECURITY. CRC PRESS, 2016.

[6]D. ZISSIS AND D. LEKKAS, "ADDRESSING CLOUD COMPUTING SECURITY ISSUES," FUTURE GENERATION COMPUTER SYSTEMS, VOL.28, NO. 3, PP. 583–592, 2012.

[7] H. T. DINH, C. LEE, D. NIYATO, AND P. WANG, "A SURVEY OF MOBILE CLOUD COMPUTING: ARCHITECTURE, APPLICATIONS, AND APPROACHES," WIRELESS COMMUNICATIONS AND MOBILECOMPUTING, VOL. 13, NO. 18, PP. 1587–1611, 2014.

[8] R. A. POPA, C. REDFIELD, N. ZELDOVICH, AND H. BALAKRISHNAN, "CRYPTDB:PROTECTING CONFIDENTIALITY WITH ENCRYPTED QUERY PROCESSING," IN PROCEEDINGSOF THE 23RD ACM SYMPOSIUM ON OPERATING SYSTEMS PRINCIPLES.ACM, 2011, PP. 85–100.

[9] C. CURINO, E. P. JONES, R. A. POPA, N. MALVIYA ET AL., "RELATIONAL CLOUD:A DATABASE-AS-A-SERVICE FOR THE CLOUD," 2011, HTTP://HDL.HANDLE.NET/1721.1/62241.

[10] D. BONEH, D. GUPTA, I. MIRONOV, AND A. SAHAI, "HOSTING SERVICES ON AN UNTRUSTED CLOUD," IN ADVANCES IN CRYPTOLOGY-EUROCRYPT 2015. SPRINGER, 2015, PP. 404–436.

[11] X. CHEN, J. LI, J. WENG, J. MA, AND W. LOU, "VERIFIABLE COMPUTATION OVER LARGE DATABASE WITH INCREMENTAL UPDATES," IEEE TRANSACTIONS ON COMPUTERS, VOL. 65, NO. 10, PP. 3184–3195, 2016.

[12] X. CHEN, J. LI, X. HUANG, J. MA, AND W. LOU, "NEW PUBLICLY VERIFIABLE DATABASES WITH EFFICIENT UPDATES," IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING, VOL. 12, NO. 5, PP. 546–556, 2015.

[13] S. BENABBAS, R. GENNARO, AND Y. VAHLIS, "VERIFIABLE DELEGATION OF COMPUTATION OVER LARGE DATASETS," IN ANNUAL CRYPTOLOGY CONFERENCE. SPRINGER, 2011, PP. 111–131.

[14] W. LI, K. XUE, Y. XUE, AND J. HONG, "TMACS: A ROBUST AND VERIFIABLE THRESHOLD MULTI-AUTHORITY ACCESS CONTROL SYSTEM IN PUBLIC CLOUD STORAGE," IEEE TRANSACTIONS ON PARALLEL & DISTRIBUTED SYSTEMS, VOL. 27, NO. 5, PP. 1484–1496, 2016.

[15] K. XUE, Y. XUE, J. HONG,W. LI, H. YUE, D. S.WEI, AND P. HONG, "RAAC: ROBUST AND AUDITABLE ACCESS CONTROL WITH MULTIPLE ATTRIBUTE AUTHORITIES FOR PUBLIC CLOUD STORAGE," IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 12, NO. 4, PP. 953–967, 2017.

[16] R. A. POPA, F. H. LI, AND N. ZELDOVICH, "AN IDEAL-SECURITY PROTOCOL FOR ORDER- ON SECURITY AND PRIVACY (SP'13). IEEE, 2013, PP. 463–477.

[17] J.-M. BOHLI, N. GRUSCHKA, M. JENSEN, L. L. IACONO, AND N. MARNAU, "SECURITY AND PRIVACY-ENHANCING MULTICLOUD ARCHITECTURES," IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING, VOL. 10, NO. 4, PP. 212– 224, 2013.

[18] Y. ELMEHDWI, B. K. SAMANTHULA, AND W. JIANG, "SECURE K-NEAREST NEIGHBOR QUERY OVER ENCRYPTED DATA IN OUTSOURCED ENVIRONMENTS," IN 2014 IEEE 30TH INTERNATIONAL CONFERENCE ON DATA ENGINEERING. IEEE, 2014, PP. 664–675.

[19] F. HAO, J. DAUGMAN, AND P. ZIELINSKI, "A FAST SEARCH ALGORITHM FOR A LARGE FUZZY DATABASE," IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 3, NO. 2, PP. 203–212, 2008.

e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue 23
December 2018

[20] A. CASTELLTORT AND A. LAURENT, "FUZZY QUERIES OVER NOSQL GRAPH DATABASES: PERSPECTIVES FOR EXTENDING THE CYPHER LANGUAGE," IN INTERNATIONAL CONFERENCE ON INFORMATION PROCESSING AND MANAGEMENT OF UNCERTAINTY IN KNOWLEDGE-BASED SYSTEMS. SPRINGER, 2014, PP. 384– 395.

[21] J. LI, Q. WANG, C. WANG, N. CAO, K. REN, AND W. LOU, "FUZZY KEYWORD SEARCH OVER ENCRYPTED DATA IN CLOUD COMPUTING," IN PROCEEDINGS OF THE 29TH IEEE INTERNATIONAL CONFERENCE ON COMPUTER COMMUNICATIONS(INFOCOM2010). IEEE, 2010, PP. 1–5.

[22] C. WANG, N. CAO, J. LI, K. REN, AND W. LOU, "SECURE RANKED KEYWORDSEARCH OVER ENCRYPTED CLOUD DATA," IN PROCEEDINGS OF THE 30TH IEEEINTERNATIONAL CONFERENCE ON DISTRIBUTED COMPUTING SYSTEMS (ICDCS2010). IEEE, 2010, PP. 253–262.

[23] N. CAO, C. WANG, M. LI, K. REN, AND W. LOU, "PRIVACY-PRESERVING MULTIKEYWORD RANKED SEARCH OVER ENCRYPTED CLOUD DATA," IEEE TRANSACTIONSON PARALLEL AND DISTRIBUTED SYSTEMS, VOL. 25, NO. 1, PP. 222–233, 2014.

[24] B. WANG, S. YU, W. LOU, AND Y. T. HOU, "PRIVACY-PRESERVING MULTIKEYWORDFUZZY SEARCH OVER ENCRYPTED DATA IN THE CLOUD," IN PROCEEDINGSOF THE 33RD ANNUAL IEEE INTERNATIONAL CONFERENCE ON COMPUTERMMUNICATIONS(INFOCOM2014). IEEE, 2014, PP. 2112–2120.