



Aggregation Queries Processing In Database-as-A-Service (DAS) Model

DR.R.GEETHA, KESAVAN G R

Associate Professor, Department of Computer Applications (MCA), K.S.R. College of Engineering (Autonomous), Tiruchengode, India

Department of Computer Applications (MCA), K.S.R. College of Engineering (Autonomous), Tiruchengode, India

ABSTRACT: Today, efficient data processing is a fundamental and vital issue for almost every scientific, academic, or business organization. The main challenge is preserving the DAS from unauthorized access and to protect the data from being accessed. Here we focus on second challenge. The ciphertext of each user input data are aggregated to form a single ciphertext. Privacy homomorphism technique is used to execute the query over encrypted data. User need to process as much of query as possible without need to decrypt the data. Our strategy is to extract the individual data from the aggregated result without decrypting the data. This technique rapidly increases the strength of existing security schemes. To prove the proposed scheme's efficiency, we also conducted comparisons in the end.

KEYWORDS: Privacy Homomorphism, Cipher block chaining, Database-as-a-service, Aggregation.

I. INTRODUCTION

Concealed data aggregation is mainly used in the field of wireless sensor networks. WSNs can be classified on the basis of their mode of operation or functionality, and the type of target applications. To enhance the lifetime of the network we reduce the unnecessary traffic and energy consumption of sensor nodes. Data aggregation is a technique which tries to alleviate the localized congestion problem. It attempts to collect various information from the sensors surrounding the event. It then transmits only the useful information to the end point thereby reducing congestion and its associated problems. Cluster based WSN have been proposed for energy utilization. Instead of being sent directly to the sink, every sensor sends the data to the cluster head. For security each sensor encrypts the data and transforms the encrypted data to the Cluster head. Cluster head aggregates the data without decryption. Generally, aggregative operations are algebraic, such as the addition or multiplication of received data, or statistical operation, such as a median, a minimum, or a maximum of a data set. By CDAMA[11], the ciphertexts from different applications can be encapsulated into "only" one ciphertext. But the major problem in CDAMA whereas the ciphertext of CDAMA with a greater k has to consider the extra cost for retransmission of lost fragments. When this scenario is used for aggregating the ciphertext in DAS model it has to support even with the ciphertext greater than k and need to extract the individual data from the aggregated ciphertext.. The user's query is executed over the database without decrypting the data even if the ciphertext is greater than k .

II. RELATED WORK

Different indexing methods have been proposed to execute query over encrypted data each one suitable for of a particular kind of query. These techniques are needed to enable external servers to execute queries on encrypted data. The storage of encrypted data do not work for all criteria. For confidentiality it is stated that that data decryption must be possible only at the client. The author proposed query execution techniques to select the data to be return in responds to a query without the need of decrypting the data themselves based on additional indexing information [1,2,3,6, 7]. The Other Query Execution Techniques in [1, 3] is a hash-based method suitable for selection queries. In [3] author proposed order preserving encryption schema (OPES) to support equality and range query. The major drawback is it only operates on integer values. In [4, 5] proposed execution of aggregation queries is done over encrypted data. Limitation of the above technique is the data is protected only at the server side. Access control mechanism enforce access restriction to different users, sets of users, or applications. It consists of grouping users with the same access privileges and in encrypting each group of tuples with the key associated with the set of users that can access it. This mechanism limited to the static groups. In this case outsourced database has to be re-encrypted each time group membership changes. Server hosted by the service provider stores encrypted database. The encrypted database is augmented with additional information this allows certain amount of query processing to occur at the server where client maintains metadata.



Strategy: Split the original query as the following condition

- A corresponding query over encrypted relations to run on the server
- A client query for post processing the results of the server query

Various overheads are Metadata at client is amount of filtering ,Bandwidth consumed and Storage wasted. Our proposed work is to encrypt the data using PH cryptosystem and to aggregate the ciphertext of each data to a single ciphertext using CBC and individual data can be retrieved using RCBC.

III. ENCRYPTION SCHEME

For confidentiality demands the user encrypts the data. The key used for encryption is asymmetric key that is sender and receiver uses different key for encryption and decryption. The encryption data is sent to the collector where the ciphertext is generated for each user data. The generation of ciphertext process continued when user continuously sends the multiple data. The ciphertext of each data are aggregated to form a single ciphertext. CBC is a mode of operation used for aggregating sequence of bits into an single unit of bits.

3.1 GENERATION OF SINGLE CIPHERTEXT (CIPHER BLOCK CHAINING)

Cipher block chaining (CBC) is a mode of operation for a block cipher (one in which a sequence of bits are encrypted as a single unit or block with a cipher key applied to the entire block). Cipher block chaining uses what is known as an initialization vector (IV) of a certain length. One of its key characteristics is that it uses a chaining mechanism that causes the decryption of a block of ciphertext to depend on all the preceding cipher text blocks. As a result, the entire validity of all preceding blocks is contained in the immediately previous ciphertext block.

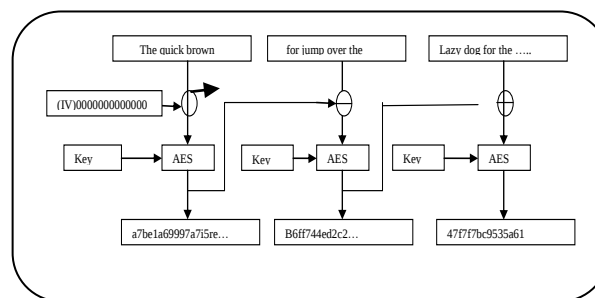


Fig.1 CBC operation

Step 1: The first step with CBC is to convert the data and keys to binary.
 Plaintext: 10011 00111 00100 00001 10100 00010 01010
 10010 10011 01110 01111 10010 00111 00100 10001 00100

Key: 11000 00100 00000 00111

Step 2: Break the plaintext up into some larger (12 bit) blocks and

remove whitespace from key :

Plaintext: 100110011100 100000011010 000010010101 001010011011 1001111100
 001110010010 00100100

Key: 11000001000000000111

Four bits are added as padding so the final block will be '000000000100'. This additional block is used to assist in deciphering the message. Without it, there is no way to determine how many extra bits were added.

Padded string

Plaintext: 100110011100 100000011010 000010010101 001010011011 100111110010 001110010010 001001001010 000000000100

Step 3: Each block is encrypted, one by one. Our basic encryption method for this example will be to reverse the block, and XOR each bit with the corresponding bit of the key. Here is how the first block is encrypted:

Plaintext : 100110011100

Reversed : 001110011001

First 12 bits of key: 110000010000

Ciphertext: 001110011001 ([XOR])



Step 4: The standard method of cipher block chaining uses the ciphertext of one block to assist in encrypting the next block. This is done by XORing the ciphertext of the previous block with the plaintext of the next block, before the normal encryption technique is executed on the plaintext. Here is how we would XOR the plaintext of block 2 with the ciphertext of block 1:

Ciphertext(Block 1): 111110001001

Plaintext(Block 2): 100000011010

Exclusive or (XOR): 011110010011

Step 5: The result of the exclusive or is now treated as the plaintext, and encrypted normally as above

Result of XOR(New Plaintext) 011110010011 Reversed : 110010011110

First 12 bits of key : 110000010000

Ciphertext(by XOR): 000010001110

Step 6: This process repeats until the last block is encrypted. Here is the result of encrypting all 8 blocks using this method:

Ciphertext : 111110001001 000010001110 000110010000
000100011100 101101100001 000011100001 000101000100 110000111000.

Since each of these blocks (except the first one) was encrypted using the ciphertext of the previous block and the key, it becomes very difficult to decrypt.

3.2 PH CRYPTOSYSTEM

After generation of single ciphertext user passes the query for accessing the data. The query get executed over encrypted data. Hence Privacy homomorphism technique is used. Privacy homomorphic encryption (PH) is a encryption scheme with homomorphic property. It allows to execute the query over encrypted data. The homomorphic property implies that algebraic operations on plaintexts can be executed by manipulating the corresponding ciphertexts; for instance, $DK(EK(m1) \odot EK(m2)) = m1 \odot m2$, where $EK(.)$ is the encryption with key K , $DK(.)$ is the decryption with key K , and $\odot$ denote operations on ciphertexts and plaintexts, respectively.

EXAMPLE:

Consider A is the domain of unencrypted values, ϵ_k is an encryption function using key k and D_k is the corresponding decryption function.

Encryption - $\epsilon_k(a) = (a \bmod p, a \bmod q)$ where $a \in \mathbb{Z}_n$

Decryption - $D_k(a) = d1q^{-1} + d2p^{-1} \pmod{n}$

($d1 = a \bmod p$ & $d2 = a \bmod q$)

Consider

$p = 5, q = 7$ ($n = 35$)

$a1 = 5$ & $a2 = 6$

$\epsilon(a1) = (0, 5)$ $\epsilon(a2) = (1, 6)$ are stored on the server

Compute $(a1 + a2)$

Server computes $\epsilon(a1) + \epsilon(a2)$ component wise

$(0+1, 5+6) = (1, 11)$

Client decrypts $(1, 11)$ as $d1q^{-1} + d2p^{-1} \pmod{n}$

$(1.7.3 + 11.5.3) \pmod{35} = 186 \bmod 35 = 11$

3.3 REVERSE CIPHER BLOCK

The single ciphertext get disaggregated and it is checked whether the user is authorized to access the data. If the user is authorized then the user query get executed over encrypted data to retrieve the user requested data. The requested data is returned to the respective user where the user can decrypt the data at client side.

Step 1: This will give us the original first block:

Ciphertext(1st block) 111110001001

First 12 bits of key: 110000010000

Exclusive or (XOR): 001110011001

Reversed: 100110011100 (block 1).

Step 2: The second block is XOR'd with the key and reversed, and XOR'd with the ciphertext of block 1:

Ciphertext(2nd block): 000010001110

First 12 bits of key: 110000010000

Exclusive or (XOR): 110010011110



Reversed : 011110010011
 Ciphertext(1stblock): 111110001001
 Exclusive or (XOR): 100000011010-block2
 Ciphertext(2ndblock): 000010001110
 First 12 bits of key: 110000010000
 Exclusive or (XOR) : 110010011110
 Reversed : 011110010011
 Ciphertext(1stblock): 111110001001
 Exclusive or (XOR): 100000011010

Step 3: Each block can now be deciphered using this method. Once the complete plaintext is found, the padding can easily be removed.

IV. RESULTS

Here the proposed method is tested with the ratio of the number of packets received by the destination and the number of packets transmitted by the source. And the fig.2 states delay in the proposed is smaller than the existing system.

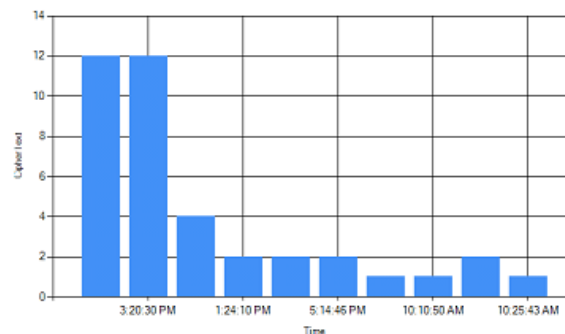


Fig.2 Aggregation delay

V. CONCLUSION

As a result, privacy and security of data at a service provider site is a paramount. In this paper, we addressed a specific data-privacy challenge. Our solution is to aggregate the each ciphertext and to provide a single ciphertext. We have developed a technique of executing the queries can be done by the provider without decrypting the aggregated result. The individual data can be retrieved from the aggregated result. This increases the strength of the existing security schemes.

REFERENCES

- [1] . Ernesto Damiani, S. De Capitani di Vimercati, Sara Foresti, Sushil Jajodia, Stefano Paraboschi, Pierangela Samarati, Key Management for Multi-User Encrypted Databases, Proceedings of the 2005 ACM workshop on Storage security and survivability, November 2005.
- [2] . H. Hacigumus, B. Iyer, S. Mehrotra, and C. Li. Executing SQL over encrypted data in the database-service-provider model. In Proc. of the ACM SIGMOD'2002, Madison, WI, USA, June 2002.]
- [3]. E. Damiani, S. De Capitanidi Vimercati, S.Jajodia, and S.Paraboschi, and P. Samarati. Balancing confidentiality and efficiency in untrusted relational DBMSs. In Proc. of the 10th ACM Conference on Computer and Communications Security, Washington, DC, USA, October 27-31 2003.
- [4] . R. Agrawal, J. Kierman, R. Srikant, and Y. Xu. Order preserving encryption for numeric data. In Proc. of ACM SIGMOD 2004, Paris, France, June 2004.
- [5] . S. Akl and P. Taylor. Cryptographic solution to a problem of access control in a hierarchy. ACM Transactions on Computer System, 1(3):239–248, August 1983.
- [5]. D. Boneh, G. D. Crescenzo, R. Ostrovsky, and G. Persiano. Public-key encryption with keyword search. In Proc. of Eurocrypt 2004, Interlaken, Switzerland, May 2004.



- [6] . H. Hacigümüş, B. Iyer, and S. Mehrotra. Providing database as a service. In Proc. of 18th International Conference on Data Engineering, San Jose, CA, USA, February 2002.
- [7]. H. Hacigümüş, B. Iyer, and S. Mehrotra. Ensuring the integrity of encrypted databases in the database-as-a-service model. In DBSec, pages 61–74, 2003.
- [8]. B. Iyer, C. Li, and S. Mehrotra, “Executing Sql over Encrypted Data in the Database-Service-Provider Model,” Proc. ACM SIGMOD Int’l Conf. Management of Data, pp. 216-227, 2002.
- [9]. H. Hacigümüş, “Efficient Execution of Aggregation Queries over Encrypted Relational Databases,” Proc. Ninth Int’l Conf. Database Systems for Advanced Applications (DASFAA ’04), vol. 9, p. 125, 2004.
- [10] . E. Damiani, S. De Capitani di Vimercati, M. Finetti, S. Paraboschi, P. Samarati, and S. Jajodia. Implementation of a storage mechanism for untrusted DBMSs. In Proc. of the Second International IEEE Security in Storage Workshop, Washington DC, USA, May 2003.
- [11] CDAMA: Concealed Data Aggregation Scheme for Multiple Applications in Wireless Sensor Networks Yue-Hsun Lin, shih-ying chang, and hung-min sun, iee transactions on knowledge and data engineering, vol. 25, no. 7, july 2013.