IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 1, Issue 5, Oct-Nov, 2013
ISSN: 2320 - 8791
www.ijreat.org

Concealed Data Aggregation Scheme for Multiple Application in Database as a Service Model

Subanivedhi N K[1] and Rekha M[2]

[1]PG Scholar, Department of Computer Science and Engineering, Vel Tech Multi Tech Engineering College, Avadi, Chennai

[2]Asst Professor, Department of Computer Science and Engineering, Vel Tech Multi Tech Engineering College, Avadi, Chennai.

Abstract

Recent evolution in networking and internet technologies has emerged "software as a service" model for enterprise computing."Database as a service" model allows user to create, store, modify, and retrieve data from anywhere in the world through internet. The major issue addressed is data privacy. Our approach is to process as much query as possible at the service providers' site, without decrypting the data. Decryption of the query processing is performed at the client site. In our proposed system we apply CDAMA (Concealed Data Aggregation Scheme for Multiplicative Application) to understand aggregation query in Database-As-a-Service (DAS) model. In DAS model, a client stores her database on an unauthorized service provider. The security to database is provided by the client through PH (Privacy Homomorphism) schemes because PH scheme keep compatible properties than standard ciphers. By PH schemes, the provider can aggregate queries without decryption. The most significant is that we need not consider the computation cost and the impact of compromising secret keys. These drawbacks is not an issue in CDAMA.

IndexTerms—*Data aggregation, Database as a Service, privacy homomorphism, query processing*

1.INTRODUCTION

The internet allows all computers to connect to one another around the world. The internet also has impact on the software industry. It has enabled an opportunity to provide software usage over the internet, and has led to a new class of business called "application service providers" or ASPs. ASPs offer privilege to use software over the Internet. ASPs provide their service to small and large organisation over the Internet. It has potential to reduce the service cost because the fixed cost is amortized among large number of users. The storage and file access are offered as a services. The question is the possibility of providing the next value-add layer in data management. The model allow organizations to control hardware and software solutions provided by the service providers, without having to develop them on their own DBMSs. Signifantly it provide a method for organization to share the knowledge of database professionals, thereby reducing the cost of managing a complex information infrastructure.

The data aggregation reduce transmission and it is vulnerable to some attacks. Thus compromising a service provider will allow to forge aggregated results. To solve this, various scheme such as delay aggregation, SIA, ESPDA, and SRDA have been proposed.

An alternative is to aggregate encrypted messages directly from node, thus avoiding the forgery of aggregated result. Since node are not capable of encrypting messages, compromising a node earns nothing in forging aggregated results. Based on this concept, Wu et al. [4] gave the proposal to allow node to classify encrypted data without decrypting them. Following this concealed data aggregation (CDA) sustaining operations on aggregation. Unlike Wu et al.'s work, CDA exploits the privacy homomorphism encryption (PH) to facilitate aggregation in encrypted data.

By the additive and multiplicative homomorphism properties, node are able to execute algebraic operations on encrypted numeric data. Further, Mykletun et al. [8] adopted several public- key-based PH encryptions to

construct their systems. In similar fashion, Girao et al. [9] extended the ElGamal PH encryption to construct theirs.

In this paper, we propose CDAMA, provides CDA between multiple groups. The CDAMA is a modification from Boneh et al.'s [7] PH scheme.The scenario is designed for multi-application. In practice, Client having different purposes in the same environment. If we apply conventional concealed data aggregation schemes the ciphertexts of different applications cannot be aggregated together; otherwise, the decrypted aggregated result will be incorrect. The only solution is to aggregate the ciphertexts of different applications separately. As a result, the transmission cost increases. By CDAMA, the ciphertexts from different applications can be encapsulated into "only" one ciphertext. Conversely the node can extract application-specific plaintexts via the corresponding secret keys.

2.RELATED WORKS

This section overviews pervious work on the concealing the data and performing aggregation on the data. Manyof such approaches have been proposed are as follows.Lingxuan Hu and David Evans [1] proposed a delayed aggregation and delayed authentication methodology that is resilient to intruder device and single device key compromises. BartoszPrzydatez, Dawn Song, and Adrian Perrig[2] proposed Aggregate-commit-prove method allows aggregator to commit and prove to the home server that it uses the data collected from the sensor.

Hasan Cam, SuatOzdemir, Prashant Nair and OzgurSanali [3] proposed sleep-active mode coordination that identify the node having overlapping sensing range and turn off the sensing unit of some of the node. This method consumes less energy. Dirk Westhoff, Joao Girao and MithunAcharya [5] proposed Privacy Homomorphism technique that allows to perform algebraic or statistical operation on encrypted data.

Yongdong Wu, Di Ma, Tieyan Li and Robert H.Deng [4] proposed that classifier has set of keywords and match these keyword with the received encrypted message and classify the message based on it. HakanHacigumus, BalaIyer, Chen Li and SharadMehrotra [6] proposed a technique that allows to perform SQL operation on encrypted data.

3.PROPOSED SYSTEM

The basic architecture comprised of three entities. A user offer query to the client. A server stores the encrypted database which is presented by service provider. The encrypted database is boosted with extra information. This threating data privacy. Client also maintains metadata for translating user queries to applicable presentation on the server and performs post query processing on server results. Based on the supplementaryinformation stored, we develop techniques to split an original query over unencrypted relations into a corresponding query over encrypted relations to run on the server and a client query for post processing results of the server query.We discover the feasibility and efficiency of our scheme by testing the performance of our approach over number of query. Our method achieve privacy with an sensible overhead.

4.METHODOLOGY

4.1 PRIVACY HOMOMORPHISM

Privacy homomorphism (PH) is an encryption scheme that has anhomomorphic property. Privacy homomorphisms allows to map set of operations on cleartext to set of operations on ciphertext. DK(EK(m1)*EK(m2))=m1+m2, where EK( ) is the encryption with key k, Dk( ) is the decryption with key k, and * and + denote the operations on plaintext and ciphertexts. The PH scheme can be classified as symmetric cryptosystem where similar key is used for encryption and decryption and asymmetric cryptosystem where different keys are used for encryption and decryption. This scheme is consider to be more secure when an ciphertext space is larger then an cleartext space.

4.2 CONCEALED DATA AGGREGATION

Conventional aggregation scheme allow any intermediate node to be forged thus they are considered to be insecure. PH scheme allows to perform aggregation operation on the ciphertext without decryption and they are more secure. The compromising any intermediate node gains no advantage.Girao et al extended PH scheme to construct an aggregation in which all share a same key thus compromising any intermediate node will be able to forge. To solve this problem, Castelluccia et al. proposed an method in which each node share unique key in beginning of the transmission. By this forging an single node have no effect on it.

4.3 CDAMA

CDAMA comprises of four processes namely Key generation, aggregation and decryption. For simplicity we take two user groups ie., (k=2). Consider three points P, Q, H and their orders are q1, q2, q3. The scalar of the two points carry aggregated message in a G_A and G_B and last point is used to carry random number for security. The aggregated ciphertext is multiplied by their order to get the aggregated message. The ciphertext from different applications are aggregated however they are not mixed. The message of each group can be obtained by simply decrypting the ciphertext with their private keys.

Pseudo code for key generation:

Step1: Based onsecurity parameter τ, compute the values of (q1,q2,q3,E). E denote set of the elliptic curve points that form cyclic group, ord(E) = n and n=q1q2q3, such that q1, q2, q3 are large prime numbers, length of q1, q2, q3 are of same
Step2: Collect three generators G1, G2, G3accidently and their orders are ord(G1) = ord(G2) = ord(G3) = n.
Step3: Calculate the value of H as q1q2*G3 and ord(H) = q3.
Step4: Choose T factor as the maximum plaintext boundary in such a way that Pollard's method is viable and figure the value of TA and TB.
Step5: Calculate the value of P as q2q3*G1 and ord(P) = q1 and G_A group public key as PK_A = (n, E, P, H, TA).
Step6: Calculate the value of Q as q1q3*G1 and ord(Q) = q2 and G_B group public key as PK_B = (n, E, P, H, TB).
Step7: Now output the value of G_A and G_B as SK_A and SK_B

Pseudo code for message encryption in G_A:

Step1: Ensure that message is from one of the user in group G_A
Step2: Select R such that it belongs to 0 to n-1.
Step3: Produce the ciphertext C as M*P+R*H.
Step4: Now output the value of C.

Pseudo code for message encryption in GB:

Step1: Ensure that message is from one of the user in group GB.
Step2: Select R such that it belongs to 0 to n-1.
Step3: Produce the ciphertext C as M*Q+R*H.
Step4: Now output the value of C.

Pseudo code for aggregating the two ciphertext:

Step1: Calculate the aggregated value of two ciphertext C^1as C1+C2 ie., $C^1=(\Sigma M_i)*P + (\Sigma M_j)*Q + (\Sigma Ri)*H$ where ΣMi is GA's aggregated result, ΣMj is GB's aggregated result, ΣRi is both groups aggregated randomness.
Step2: The output value of C^1 is returned.

Pseudo code for message decryption in GA:
Step1: Calculate the value of M as ΣM_i which in turn equals to $\log_{P'}$ (q2q3 * C)and P' = q2q3 * P.
Step2: Now output the value of M.

Pseudo code for message decryption in GB:
Step1: Calculate the value of M as ΣMj which in turn equals to $\log_{Q'}$ (q1q3 * C) and Q' = q1q3 * Q.
Step2: Now output the value of M.

4.4 CDAMA SIMPLIFICATION

CDAMA can be simplified where k>2. This uses distinctive generator to construct pair of key for various groups. The order of set of elliptic curve points should be larger for security issues. Thus when k become larger ,theciphertext length also increases. This may also result in considerable overhead. The users within the particular group are assigned same group public key. The improvement is that ciphertext from different applications are not mixed yet they are aggregated. The appropriate client can decrypt the aggregated result to extract the valuable data.

Pseudo code for key generation:

Step1: Based on security parameter τ, compute the values of (q1,q2,…,qk+1,E). E denote set of the elliptic curve points that form cyclic group, ord(E) = n and n is the product of q1,…,qk+1 , such that q1,…, qk+1 are large prime numbers, length of q1,…, qk+1 are of same
Step2: Collect k+1 generators G1,…, Gk+1 accidently and their orders are ord(Gi) = n ,i=1…k+1.
Step3: Calculate the value of H and ord(H) = qk+1.
Step4: Choose T factor as the maximum plaintext boundary in such a way that Pollard's method is viable and figure the value of TA and TB.
Step5: Calculate the value of Pi and ord(Pi) = qi.
Step6: The output of Gi's public key is PKi.
Step7: The output of Gi's private key is Ski.

Pseudo code for message encryption in Gi:

Step1: Ensure that message is from one of the user in group Gi
Step2: Select R such that it belongs to 0 to n-1.
Step3: Produce the ciphertext C as M*Pi + R*H

where Pi belongs to PKi
Step4: Now output the value of C.

Pseudo code for aggregating the two ciphertext:
Step1: Calculate the aggregated value of two ciphertext C1 as C1+C2 .
Step2: The output value of C1 is returned

Pseudo code for message decryption in Gi:
Step1: Calculate the value of M.
Step2: Now output the value of M.

4.5 KEY DISSEMINATION

The key disseminated among the client in the user group in two ways. They are Key predissemination in which we know the exact client who want to access the data so that we provide the appropriate key to that particular client and Key postdissemination in which client who are authorized to access are not known in the prior and provide it omly when they are in need.

5. DISCUSSION

5.1 SCALAR MULTIPLICATION EFFICIENCY

In CDAMA, the performance of scalar multiplication on elliptic curve determine the effectiveness of encryption and decryption. Decryption is not focused here because it is done at the client side. Calculate K*P where P is a random point on elliptic curve and K is an integer for scalar multiplication. The following technique has been used to enhance the scalar multiplication

1. By using endomorphism structures.
2.By using different integer value for scalar multiplication.
3. By using different coordinate values.
4. By using optimal extension fields.

5.2 SIZE OF CIPHERTEXTS

The one another metrics for performance and cost evaluation is size of the ciphertexts. In CDAMA, the ciphertext is stored as a couple of elliptic curve affine points. If the finite field of elliptic curve is FP, the size of ciphertext is p+1 bits because we only store the x-coordinates of curve points and the additional one bit for the sign of y-coordinates. CDAMA also require specific curves of the given order.

5.3 SECURITY ANALYSIS

Comparing CDAMA with other conventional method, this provide an single scheme of encryption from the point of transmitting from client side to the provider and storing it the server side. Whereas previous conventional method provide an separate method of security for transmission and for storage. Thus CDAMA is consider to be more secure than an other standard scheme

6.CONCLUSION

We propose the CDAMA scheme for storing the data confidently in server side. Client encrypts the data and stores it service provider in encrypted form. When user issues query, server perform operation on encrypted data and sends the result in encrypted form and client carry out the operation on encrypted result and sends the result to the user. Here most of the work is done in the client side. In
future, we can reduce the amount of operation cost on the client side

REFERENCES

[1]Lingxuan Hu and David Evans, “Secure Aggregation For Wireless Networks”, Proc. Symp. Applications and the Internet Workshops, pp. 384-391, 2003.

[2] B. Przydatek, D. Song, A. Perrig , “SIA: Secure Information Aggregation in Sensor Networks,” Proc. First Int’l Conf . Embedded Netwrked Sensor Systems, pp. 255-265, 2003.

[3] H. Cam, S.Ozdemir, P. Nair, D. Muuthuvinashiappan H.O. Sanli, “Energy-Efficient Secure Pattern Based Data Aggregation for Wireless Sensor Networks,” Computer Comm., vol. 29, no. 4, pp. 446-455, 2006.

[4] Y. Wu, D. Ma, T. Li, and R.H. Deng, “Classify Encrypted Data in Wireless Sensor Network,” Proc. IEEE 60th Vehicular Technology Conf., 2004.

[5] D. Westhoff, J. Girao, and M. Acharya, “Concealed Data Aggregation for Reverse Multicast Traffic in Sensor Networks: Encryption, Key Distribution, and Routing Adaptation,” IEEE Trans. Mobile Computing Vol. 5, no. 10, pp 1471-1431, Oct. 2006.

[6] B. Iyer, C. Li, and S. Mehrotra, “Executing Sql over Encrypted Data in the Database-Service-Provider Model,” Proc. ACM SIGMOD Int’l Conf. Mgmt of Data, pp. 216-227, 2002

[7] D. Boneh, E. Goh, and K. Nissim, "Evaluating 2-DNF Formulas on Ciphertexts," Proc.Second Int'l Conf. Theory of Cryptography (TCC), vol. 3378, pp. 325-341. 2005.

[8] E. Mykletum, J. Girao, and D. Westhoff , "Public Key Based Crytoschemes for Data Concealment in Wireless Sensor Networks," Proc. IEEE Int'l Conf . Comm., vol.5 , 2006

[9] J. Girao, D. Westhoff and T.Araki, "Tinypewords: Tiny Persistent Encryped Data Storage in Asynchronous Wireless Sensor Networks," Ad Hoc Networks, vol. 5, No. 7, pp. 1073-1089, 2007

[10] H. Hacigumus, "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

[11] J. Domingo-Ferrer, " A Provably Secure Addition and Multiplicative Privacy Homomorphism," Proc. Fifth Int'l Conf. Information Security, pp. 471-483, 2002.