IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 3 Issue 11, November 2016
ISSN (Online) 2348 – 7968 | Impact Factor (2015) - 4.332
www.ijiset.com

Secure Cloud Data Store with Keyword Search

Desineni Bhuvan Kumar[1], S. Surekha[2]

[1] M.Tech Student, Dept.of CSE, Vaishnavi Institute of Technology, Andhra Pradesh, India,
[2] HOD, Vaishnavi Institute of Technology, Andhra Pradesh, India,

Abstract

Searchable Public-Key Ciphertexts with Hidden Structures (SPCHS) for keyword search as speedy as viable without sacrificing semantic protection of the encrypted keywords. In SPCHS, all keyword-searchable ciphertexts are structured by means of hidden relations, and with the search trapdoor corresponding to a keyword, the minimum statistics of the members of the family is disclosed to a search algorithm as the coaching to locate all matching ciphertexts efficiently. We construct a SPCHS scheme from scratch in which the ciphertexts have a hidden star-like structure. We prove our scheme to be semantically invulnerable in the Random Oracle (RO) model. The search complexity of our scheme is dependent on the authentic range of the ciphertexts containing the queried keyword, as an alternative than the variety of all ciphertexts. Finally, we existing a prevalent SPCHS construction from anonymous identity-based encryption and collision-free full-identity malleable Identity-Based Key Encapsulation Mechanism (IBKEM) with anonymity. We illustrate two collision-free full-identity malleable IBKEM instances, which are semantically impenetrable and anonymous, respectively, in the RO and preferred models. The latter occasion permits us to construct an SPCHS scheme with semantic safety in the well-known mannequin

Keywords: *Public-Key, k-NN Classifier, SPCHS, SMC, Cloud Storage.*

1. Introduction

PUBLIC-KEY encryption with keyword searches (PEKS), added by means of Boneh, and has the advantage that everybody who is aware of the receiver's public key can upload keyword-searchable ciphertexts to a server. The receiver can delegate the key-word search to the server. More specifically, every sender separately encrypts a file and its extracted key phrases and sends the resulting ciphertexts to a server; when the receiver wishes to retrieve the documents containing a specific keyword, he delegates a key-word search trapdoor to the server; the server finds the encrypted files containing the queried key-word without understanding the authentic archives or the keyword itself, and returns the corresponding encrypted files to the receiver; finally, the receiver decrypts these encrypted files1. The authors of PEKS additionally introduced semantic safety towards chosen keyword attacks (SS-CKA) in the sense that the server can't distinguish the ciphertexts of the key phrases of its desire before observing the corresponding keyword search trapdoors. It seems an appropriate protection notion, specifically if the keyword area has no excessive min-entropy. Existing semantically tightly closed PEKS schemes take search time linear with the whole number of all ciphertexts. This makes retrieval from large-scale databases prohibitive. Therefore, more efficient search performance is quintessential for practically deploying PEKS schemes.

In existing work, Existing work on Privacy-Preserving Data Mining (either perturbation or invulnerable multi-party computation primarily based approach) can't resolve the DMED problem. Perturbed data do not possess semantic security, so records perturbation methods cannot be used to encrypt exceptionally sensitive data. Also the perturbed records do now not produce very accurate statistics mining results. Secure multi-party computation (SMC) based totally approach assumes statistics are distributed and not encrypted at every taking part party.
The main drawbacks are,

- No classification trouble has been proposed underneath exclusive security models.
- Existing approach assumes facts are distributed and not encrypted.
- Data do no longer possess semantic safety

2. Proposed System

We center of attention on fixing the classification hassle over encrypted data. In particular, we advocate a secure k-NN classifier over encrypted facts in the cloud. The proposed protocol protects the confidentiality of data, privateness of user's input query, and hides the information get right of entry to patterns. To the excellent of our knowledge, our work is the first to strengthen an impervious k-NN classifier over encrypted data below the semi-honest model. Also, we empirically analyze the efficiency of our proposed protocol the use of a real-world dataset below one of kind parameter settings.
We proposed novel techniques to effectively remedy the DMED problem assuming that the encrypted records are outsourced to a cloud. Specifically, we focal point on the classification hassle on the grounds that it is one of the most common facts mining tasks. Because each

classification approach has their own advantage, to be concrete, this paper concentrates on executing the okay nearest neighbor classification technique over encrypted information in the cloud computing environment. The main advantages are,

- ✓ Solving the classification hassle over encrypted data
- ✓ Protects the confidentiality of data
- ✓ Secure k-NN classifier over encrypted statistics underneath the semi-honest mannequin

3. System Architecture

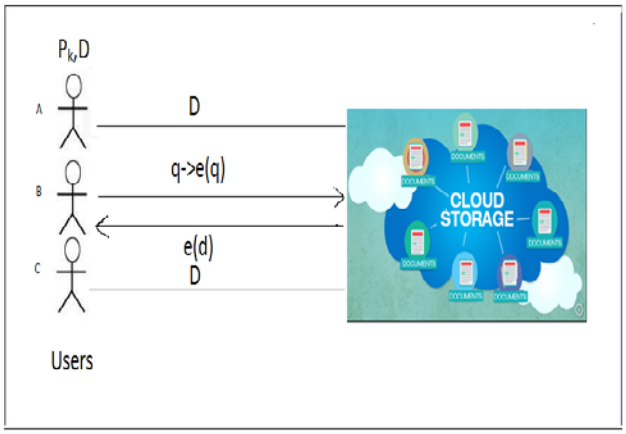

Fig.1. Cloud Storage

The above fig shows cloud storage architecture. This architecture contains many modules. They are, cloud server and AES Cryptosystem.

A cloud server is a logical server that is built, hosted and delivered via a cloud computing platform over the Internet. Cloud servers possess and show off comparable abilities and performance to a standard server however are accessed remotely from a cloud provider. A cloud server may additionally also be known as a virtual server or virtual private sever.

Stage 1 - Secure Retrieval of k-Nearest Neighbors (SRkNN): In this stage, Bob at the beginning sends his question q (in encrypted form) to C1. After this, C1 and C2 involve in a set of sub-protocols to securely retrieve (in encrypted form) the classification labels corresponding to the k-nearest neighbors of the enter query q. At the give up of this step, encrypted category labels of k-nearest neighbors are recognized only to C1.

Stage 2 - Secure Computation of Majority Class (SCMCk): Following from Stage 1, C1 and C2 mutually compute the category label with a majority vote casting amongst the k-nearest neighbors of q. At the end of this step, solely Bob is aware of the classification label corresponding to his enter query file q.

4. Motivation

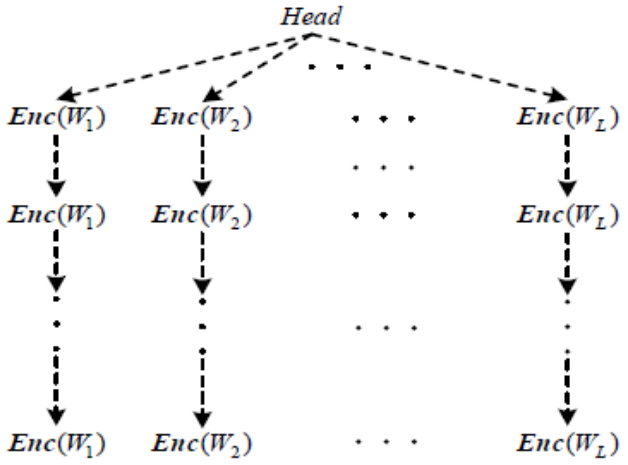

Fig.2. Hidden star-like structure shaped with the aid of keyword searchable ciphertexts. (The dashed arrows denote the hidden relations. Enc (Wi) denotes the searchable ciphertext of keyword Wi.)

We are involved in presenting quite environment friendly search overall performance without sacrificing semantic protection in PEKS.Observe that a key-word area is commonly of no high minentropy in many scenarios. Semantic security is crucial to warranty key-word privateness in such applications. Thus the linear search complexity of existing schemes is the predominant obstacle to their adoption. Unfortunately, the linear complexity appears to be inevitable due to the fact the server has to scan and test each ciphertext, due to the reality that this ciphertexts (corresponding to the same keyword or not) are indistinguishable to the server.

A nearer appear shows that there is still area to improve search performance in PEKS barring sacrificing semantic safety if one can arrange the ciphertexts with elegantly designed but hidden relations. Intuitively, if the keyword searchable ciphertexts have a hidden star-like structure, as shown in Figure 1, then search over ciphertexts containing precise keywords may additionally be accelerated. Specifically, suppose all ciphertexts of the same key-word structure a chain via the correlated hidden relations, and also a hidden relation exists from a public Head to the first ciphertext of each chain. With a keyword search trapdoor and the Head, the server seeks out the first matching ciphertext through the corresponding relation

from the Head. Then another relation can be disclosed with the aid of the found ciphertext and courses the searcher to be seeking out the subsequent matching ciphertext. By carrying on in this way, all matching ciphertexts can be found. Clearly, the search time depends on the actual number of the ciphertexts containing the queried keyword, as a substitute than on the complete quantity of all cipher texts.

4. Literature Survey

4.1 Study about Public Key Encryption with keyword Search

We learn about the trouble of looking out on facts that is encrypted using a public key system. Consider consumer Bob who sends e mail to user Alice encrypted below Alice's public key. An email gateway wants to take a look at whether the e-mail includes the keyword \urgent" so that it could route the e mail accordingly. Alice, on the other hand does no longer want to provide the gateway the potential to decrypt all her messages. We outline and assemble a mechanism that permits Alice to supply a key to the gateway that allows the gateway to test whether or not the phrase \urgent" is a keyword in the electronic mail besides learning something else about the email. We refer to this mechanism as Public Key Encryption with keyword Search. As every other example, reflect on consideration on a mail server that shops quite number messages publicly encrypted for Alice by others. Using our mechanism Alice can send the mail server a key that will allow the server to become aware of all messages containing some precise keyword, however learn nothing else. We define the thinking of public key encryption with key-word search and supply quite a few constructions.

4.2 Study about Security Proofs for Identity-Based Identification and Signature Schemes

We current as-strong-as-possible definitions of privacy, and constructions reaching them, for public-key encryption schemes the place the encryption algorithm is deterministic. We attain as a consequence database encryption techniques that allow speedy (i.e. sub-linear, and in reality logarithmic, time) search whilst provably providing privacy that is as strong as viable concern to this fast search constraint. One of our constructs, referred to as RSA-DOAEP, has the introduced characteristic of being size preserving, so that it is the first example of a public-key cipher. We generalize this to reap a thought of efficiently-searchable encryption schemes which permit extra bendy privacy to search-time trade-offs by a method referred to as bucketization. Our results answer much-asked questions in the database neighborhood and furnish foundations for work completed there.

5. Simulated Result

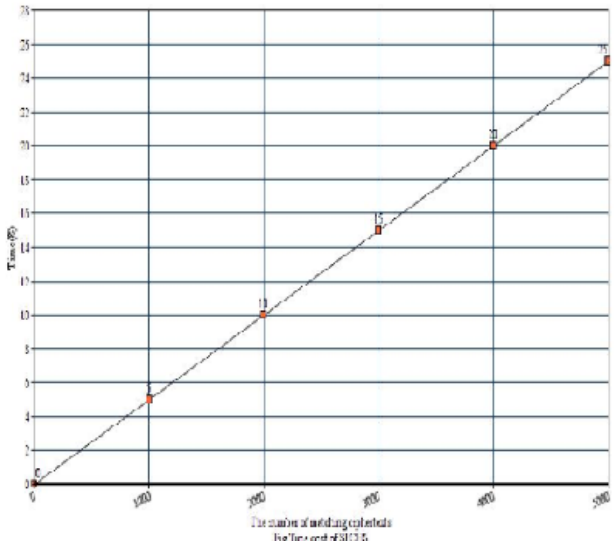

Fig.3. Time cost of SPCHS

Hardware	*Intel CPU E5300 @ 260GHz*
OS and Compiler	*Win XP and Microsoft VC++ 6.0*
Program library	*MIRACL version 5.4.1*
Parameters of bilinear map	
Elliptic curve	$Y^2=x^3+A.x+B.x$
Pentanomial basis	$T^m+t^a+t^b+t^c+1$
Base field: 2^m	$M=379$
A	*1*
B	*1*
Group Order :q	$2^m+2^{(m+1)/2}+1$
a	*315*
b	*301*
c	*287*
The default unit is decimal	

Table I: System parameters

We coded our SPCHS scheme, and examined the time fee of algorithm Structured Search to execute its cryptographic operations for specific numbers of matching ciphertexts. We also coded the PEKS scheme. Table I suggests the device parameters including hardware, software and the chosen elliptic curve. Assume there are in total 104 searchable ciphertexts. PEKS takes about 53.8 seconds search time per keyword, on the grounds that it must test all ciphertexts for every search. **Fig.3** suggests the experimental effects of SPCHS. It is clear that the time price of SPCHS is linear with the variety of matching ciphertexts, whereas for PEKS it is linear with the variety of complete ciphertexts. Hence, SPCHS is a whole lot more environment friendly than PEKS.

6. Conclusion and Future Work

This paper investigated as-fast-as-possible search in PEKS with semantic security. We proposed the concept of SPCHS as a variant of PEKS. The new idea allows Keyword-searchable ciphertexts to be generated with a hidden structure. Given a keyword search trapdoor, the search algorithm of SPCHS can disclose phase of this hidden shape for guidance on finding out the ciphertexts of the queried keyword. Semantic safety of SPCHS captures the privateness of the key phrases and the invisibility of the hidden structures. We proposed an SPCHS scheme from scratch with semantic safety in the RO model. The scheme generates keyword-searchable ciphertexts with a hidden star-like structure. It has search complexity basically linear with the specific variety of the ciphertexts containing the queried keyword. It outperforms existing PEKS schemes with semantic security, whose search complexity is linear with the variety of all ciphertexts. We recognized various interesting properties, i.e., collision-freeness and full-identity malleability in some IBKEM instances, and formalized these homes to build a frequent SPCHS construction. We illustrated two collision-free full-identity malleable IBKEM instances, which are respectively invulnerable in the RO and general models. SPCHS seems a promising tool to solve some difficult troubles in public-key searchable encryption. One utility might also be to gain retrieval completeness verification which, to the fine of our knowledge, has no longer been performed in current PEKS schemes. Specifically, by forming a hidden ring-like structure, i.e., letting the ultimate hidden pointer always point to the head, one can gain PEKS allowing to test the completeness of the retrieved ciphertexts through checking whether or not the pointers of the back ciphertexts shape a ring. Another application may also be to understand public key encryption with content search, a similar performance realized by symmetric searchable encryption. Such type of content searchable encryption is useful in practice, e.g., to filter the encrypted spam's. Specially, through forming a hidden treelike shape between the sequentially encrypted phrases in one file, one can gain public-key searchable encryption allowing content material search (e.g., to find whether or not there are precise contents in an encrypted file). The search complexity is linear with the size of the queried content.

References

[1] Boneh D., Crescenzo G. D., Ostrovsky R., Persiano G.: Public Key Encryption with Keyword Search. In: Cachin C., Camenisch J. (eds.) EUROCRYPT 2004. LNCS, vol. 3027, pp. 506-522. Springer,
Heidelberg (2004)
[2] Bellare M., Boldyreva A., O'Neill A.: Deterministic and Efficiently Searchable Encryption. In: Menezes A. (ed.) CRYPTO 2007. LNCS, vol. 4622, pp. 535-552. Springer, Heidelberg (2007)
[3] Boneh D., Boyen X.: Efficient Selective-ID Secure Identity-Based Encryption Without Random Oracles. In: Cachin C., Camenisch J. (eds.) EUROCRYPT 2004. LNCS, vol. 3027, pp. 223-238. Springer, Heidelberg (2004)
[4] Boyen X., Waters B. R.: Anonymous Hierarchical Identity-Based Encryption (Without Random Oracles). In: Dwork C. (ed.) CRYPTO 2006. LNCS, vol. 4117, pp. 290-307. Springer, Heidelberg (2006)
[5] Gentry C.: Practical Identity-Based Encyrption Without Random Oracles. In: Vaudenay S. (ed.) EUROCRYPT 2006. LNCS, vol. 4004, pp.445-464. Springer, Heidelberg (2006)
[6] Ateniese G., Gasti P.: Universally Anonymous IBE Based on the Quadratic Residuosity Assumption. In: Fischlin M. (ed.) CT-RSA 2009. LNCS, vol. 5473, pp. 32-47. Springer, Heidelberg (2009)
[7] Ducas L.: Anonymity from Asymmetry: New Constructions for Anonymous HIBE. In: Pieprzyk J. (ed.) CT-RSA 2010. LNCS, vol. 5985, pp. 148-164. Springer, Heidelberg (2010)
[8] Abdalla M., Catalano D., Fiore D.: Verifiable Random Functions: Relations to Identity-Based Key Encapsulation and New Constructions. Journal of Cryptology, 27(3), pp. 544-593 (2013)
[9] Freire E.S.V., Hofheinz D., Paterson K.G., Striecks C.: Programmable Hash Functions in the Multilinear Setting. In: Canetti R., Garay J.A. (eds.) Advances in Cryptology - CRYPTO 2013. LNCS, vol. 8042, pp. 513-530. Springer, Heidelberg (2013)
[10] Garg S., Gentry C., Halevi S.: Candidate Multilinear Maps from Ideal Lattices. In: Johansson T., Nguyen P. (eds.) Advances in Cryptology - EUROCRYPT 2013. LNCS, vol. 7881, pp. 1-17. Springer,
Heidelberg (2013)

ISSN (Online) 2348 – 7968 | Impact Factor (2015) - 4.332

www.ijiset.com

Desineni Bhuvan Kumar received the B.Tech Degree in Computer Science and Engineering from Vaishnavi Institute of Technology, JNTUA in 2014. He is currently working towards the Master's Degree in Computer Science and Engineering, in Vaishnavi Institute of Technology. He interest lies in the areas of Web Development Platforms, SQL, and Cloud Computing Technology.

S. Surekha received M.Tech degree in Chadalavawada Ramannamma Engineering College in the Year 2013 from JNTUA, A.P., and India. Currently she is HOD in the Department of Computer Science and Engineering at Vaishnavi Institute of Technology -Tirupati.