International Journal of
Emerging Research in Management &Technology
ISSN: 2278-9359 (Volume-4, Issue-7)

Research Article

July 2015

Advance Multi-Factor Authentication and Multi-keyword Ranked Search for Encrypted Data

Snehal Rahul Patil
Department of Computer Engineering, DPCOE,
Wagholi, India

Abstract—

This paper proposes a novel biometrics-based Single Sign On user authentication scheme in Telematics system, in order to enable to protect user account privacy. The focus during this work is to develop a secure and versatile multi-factor authentication to telematics environments with key management. Considering the massiveform ofdata users and documents inside the cloud, it's a necessityto allow multiple keywords inside the search request and are available documents inside the order of their connectedness to those keywords. Connected works on searchable encodingspecialize in single keyword search or man of science keyword search, and barely sort the search results. Among varied multikeyword linguistics, during this paper, for the first time we'vean inclinationto stipulate and solve the matter of multi-keyword stratified search over encrypted cloud knowledge, and establish a variety of privacy wants. Among varied multi-keyword linguistics, we tend to elect the economical similarity live of "coordinate matching", i.e., as many matches as accomplishable, to effectively capture the connectedness of outsourced documents to the question keywords, and use "inner product similarity" to quantitatively decide such similarity live. For meeting the challenge of supporting multi-keyword linguistics while not privacy breaches, we'vean inclination to propose a basic arrange of MRSE mistreatment secure complex quantity computation. Then we'vean inclinationto supplya pair of improved MRSE schemes to achieve varied tight privacy requirements in a pair ofcompletelytotally different threat models. Thorough analysis work privacy and efficiency guarantees of projected schemes is given, and experiments on the real-world dataset show our projected schemes introduce low overhead on every computation and communication

Keywords— MRSE, LNCS, PEKS

I. INTRODUCTION

To protect knowledge privacy and combat uninvited accesses within the cloud and on the far side, sensitive knowledge, strong authentication methods to key management involves cryptography devices or biometrics e.g., emails, personal health records, picture albums, tax documents, financial transactions, etc., mightneed to be encrypted by knowledge owners before outsourcing to the business public cloud [2]; this, however, obsoletes the standardknowledge utilization service based on plaintext keyword search. The trivial answer of downloading all the info and decrypting regionally is clearly impractical, as a result ofthe largequantity of information measurevalue in cloud scale systems. Moreover, other than eliminating the local storage management, storing knowledge into the cloud serves no purpose unless they will be simply searched and utilized. Most of the literature on the subject works with some sort of multifactor authentication and brings concepts of biometrics and cryptographic devices to increase the security. Thus, exploring privacy-preserving and effective search service over encrypted cloud knowledge is of predominant importance

Several matches as attainable, to capture the relevancy of knowledge documents to the search query. Specifically, we have a tendency to use "inner product similarity" [4], i.e., the number of question keywords showing in a very document, to quantitatively valuate such similarity live of that document to the search question. Throughout the index construction, each document is related to a binary vector as a sub-index wherever every bit represents whether or not corresponding keyword is contained within the document. The search question is also delineate as a binary vector wherever every bit suggests that whether corresponding keyword seems during this search request, so the similarity may well be precisely measured by the inner product of the question vector with the information vector. However, directly outsourcing the information vector or the question vector can violate the index privacy or the search privacy. To fulfill the challenge of supporting such multi-keyword linguistics while not privacy breaches, we have a tendency to propose a basic plan for the MRSE using secure real computation that is tailored from a secure k-nearest neighbor (kNN) technique [23]

MRSE Framework:

For easy presentation, operations on the data documents are not shown in the framework since the data owner could easily employ the traditional symmetric key cryptography to encrypt and then outsource data. With focus on the index and query, the MRSE system consists of four algorithms as follows.

- Setup(1ℓ) Taking a security parameter ℓ as input, the data owner outputs a symmetric key as SK.
- BuildIndex(F , SK) Based on the dataset F , the data owner builds a searchable index I which is encrypted by the symmetric key SK and then outsourced to the cloud server. After the index construction, the document collection can be independently encrypted and outsourced.

- Trapdoor($\widetilde{w}$) With t keywords of interest in$\widetilde{w}$ as input, this algorithm generates a corresponding trapdoor T$\widetilde{w}$.
- Query(T $\widetilde{w}$, k, I) When the cloud server receives a query request as (T$\widetilde{w}$, k), it performs the ranked search on the index I with the help of trapdoor T$\widetilde{w}$, and finally returns F$\widetilde{w}$, the ranked id list of top-k documents sorted by their similarity with$\widetilde{w}$.

Neither the search control nor the access control is within the scope of this project. While the former is to regulate how authorized users acquire trapdoors, the later is to manage users' access to outsourced documents.

KNN Algorithm :
1.For each training example <xM(x)>, add the
Example to the list of training examples.
Given a query instance xq to be classified,

2. Let x1, x2....xk denote the k instances from training examples that are nearest to xq.

3. Return the class that represents the maximum of the k
Instances.

DES Algorithm :
/// add DES algo which i have sent you in pdf ...

II. LITERATURE SURVEY

L. M. Vaquero, L. Rodero-Merino, J. Caceres, and M. Lindner, "A break in the clouds: towards a cloud definition," ACM SIGCOMM Compute. Commun. Rev., vol. 39, no. 1, pp. 50–55, 2009[1].Next generation computing stared with the advent of Cloud computing. In cloud computing data possessor are goaded to farm out their complex data management systems from local sites to the commercial public cloud for great flexibility and economic savings. To ensure the safety of stored data, it becomes must to encrypt the data before storing. In cloud the data search arises only with the plain data. But it is essential to invoke search with the encrypted data also. The specialty of cloud data story age should allow copious keywords in a solitary query and results the data documents in the relevance order. This paper focuses on multi keyword search based on ranking over an encrypted cloud data (MRSE). The search uses the feature of similarity and inner r product similarity matching. The experimental results show that the overhead in computation and communication are considerably low.

S. Kamara and K. Lauter, "Cryptographic cloud storage," in RLCPS, January 2010, LNCS. Springer, Heidelberg [2].Cloud computing is becoming more interesting day by day. As the use of cloud services increases it's now important to do something for improving efficiency and security of cloud computing. Cloud storage contains huge amount of data, in such case to search that data efficiently becomes a challenging task. Also security vulnerability of such online storage systems is always non trustable. The recent researches are trying to so love this problem by the method of keyword search. But these methods solves this problem to some extend but some methods increases the computational burden on the cloud server or makes the retrieval of files the costly by means of bandwidth efficiency by se ending all similar files to the requesting user. This paper discuss this problem and later gives the solution to solve this problem. To solve this problem the method of keyword search has been used. This paper tries to solve the problem of searching files through the huge amount of files securely and efficiently. The previous methods make the search non efficient by means of time and computational cost, but the method discussed in this paper makes the searching very efficient and secure [2].

E.-J. Goh, "Secure indexes," Cryptology ePrint Archive, 2003, http://eprint .iacr.org /2003 /216[6].A secure index is a data structure that allows a queried with a "trapdoor" for a word x to test In O (1) time only if the index contains x;the index reveals no information about its contentswithout valid trapdoors, and trapdoors can only be generated with a secret key. Secure indexes are a natural extension of the problem of constructing data structures with privacy guarantees such as those provided by oblivious and history independent data structures. In this paper, we formally define a secure index and formulate a security model for indexes known as semantic security against adaptive chosen keyword attack (Ind-cka). We also develop an efficient Ind-cka secure index construction called z-idx using pseudo-random functions and Bloom filters, and show how to use z-idx to implement searches on encrypted data. This search scheme is the most efficient encrypted data search scheme currently known; it provides O (1) search time per document, and handles compressed data, variable length words, and Boolean and certain regular expression queries. The techniques developed in this paper can also be used to build encrypted searchable audit logs, private database query schemes, accumulated hashing schemes, and secure set membership tests [6].

Y.-C. Chang and M. Mitzenmacher, "Privacy preserving keyword searches on remote encrypted data," in *Proc. of ACNS*, 2005[7].

Due to its low cost, robustness, flexibility and ubiquitous Naturea, cloud computing is chang- in the way entities manage their data. However, various privacy concerns arise whenever potentially sensitive data is outsourced to the cloud. This paper presents a novel approach for coping with such privacy concerns. The proposed scheme prevents the cloud server from learning any possibly sensitive plaintext in the outsourced databases. It also allows the database owner to delegate users to conducting content-level fine-grained private search and decryption. Moreover, our scheme supports

private querying whereby neither the database owner nor the cloud server learns query details. Additional require mint that user's input be authorized by CA can also be supported [7].

D. Boneh, G. D. Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in *Proc. of EUROCRYPT*, 2004[8].The public key encryption with keyword search (PEKS) scheme recently proposed by Boneh, Di Crescenzo, Ostrovsky, and Persiano enables one to search encrypted keywords without compromising the security of the original data. In this paper, we address three important issues of a PEKS scheme, \refreshing keywords", \removing secure channel" andprocessing multiple keywords", which have not been considered in Boneh ET. al.'s paper. We argue that care must be taken when keywords are used frequently in the PEKS scheme as this situation might contradict the security ofPEKS. We then point out the ine±ciency of the original PEKS scheme due to the use of the secure channel. We resolve this problem by constructing an E_cient PEKS scheme thatremoves secure channel. Finally, we proposea PEKS scheme that encrypts multiple keywords anciently [9].

Roy et al., 2008] Roy, A., Datta, A., and Mitchell, J. (2008). Formal proofs of cryptographic security of diffie-Hellman-based protocols. *Trustworthy Global Computing*, pages 312–329.The EAP-GPSK protocol may be alight-weight, versatile authentication pro- toolhoping oncentrosymmetric key cryptography. It'sa part ofANin progress IETF process to develop authentication ways for the EAP framework. We analyze the protocol and realize3 weaknesses: a serviceable Denial-of-Service attack, an Anomaly with the keyderivation operatewont toproduce a short mas- term session key, and a cipher suitedowngrading attack. We have a tendency to propose fixes to those anomalies, and use a finite-state verification tool to look for remaining prob- elmswhencreating these repairs. We have a tendency to then prove the fastened version correct employing a protocol verification logic. We have a tendency tomention the attacks and our prompt fixes with the authors of the specification document that has afterwards been changed to include our projectedchanges [10].

Record on, D. and Fitzpatrick, B. (2011). Opened authentication 1.1, may 2006. *URLhttp: //openid. Net/specs/openid-authentication-1 1. Html*. Open IDAuthentication provides a way to prove that an End User owns an Identity URL. It does this without passing around their password, email address, or anything they don't want it to. Open ID is completely decentralized meaning that anyone can choose to be a Consumer or Identity Provider without having to register or be approved by any central authority. End Users can pick which Identity Provider they wish to use and preserve their Identity as they move between Providers. While nothing in the protocol requires JavaScript or modern browsers, the authentication scheme plays nicely with "AJAX"-style setups, so an End User can prove their Identity to a Consumer without having to leave the page they are on.

The Open ID Authentication specification does not provide any mechanism to exchange profile information, though Consumers of an Identity can learn more about an End User from any public, semantically interesting documents linked thereunder (FOAF, RSS, Atom, card, etc.). Extensions are being built on top of the foundation created by Open ID Authentication to provide mechanisms to exchange profile information.

III. PROPOSED APPROACH FRAMEWORK AND DESIGN

3.1 Problem Definition

The problem of multiple keyword ranking searches over encrypted data on cloud, and establish a set of strict privacy requirements for such a secure cloud data utilization system. Here we propose two MRSE schemes based on the similarity measure of "coordinate matching" while meeting different privacy requirements in two different threat models. Thorough analysis investigating privacy and efficiency guarantees of the proposed schemes is given, and experiments conducted on the real-world dataset furthermore shows that the proposed schemes indeed introduce low overhead on computation and communication.

3.2 Proposed System

The proposed technique averts the cloud server from learning any possibly sensitive plaintext in the outsourced databases. This system also allows the database owner to envoy users to conducting content-level fine-grained private search and decryption.

3.3 Mathematical Model

Input:

Data files

Output=

{Multikeyword Search}

Query:

$$T_W, K, I$$

Trapdoor:

$$T_W = \{M_1^{-1}\overrightarrow{Q}, M_1^{-1}\overrightarrow{Q}''\}$$
$$= (D_i, \in_i, 1).(rQ, r, t)$$
$$= r(D_i \ . \ Q + \in_i) + t$$

Data Vector:

$$I_i T_W = \{M_1^T D_i' M_2^T D_i''\} . \{M_1^{-1} \vec{Q}', M_2^{-1} \vec{Q}''\}$$
$$= D_i' . \vec{Q}' + \vec{D}'' . \vec{Q}''$$
$$= \vec{D}_i . \vec{Q}$$
$$= (D_i, \epsilon_i, 1) . (rQ, r, t)$$
$$= r(D_i . Q + \epsilon_i) + t$$

Scale Analysis Attack:

$$\begin{cases} y_1 - y_2 = r(1 + \epsilon_1 - \epsilon_2); \\ y'1 - y'2 = r'(1 + \epsilon_1 - \epsilon_2); \\ y_2 - y_3 = r(1 + \epsilon_2 - \epsilon_3); \\ y'2 - y'3 = r'(1 + \epsilon_2 - \epsilon_3); \end{cases}$$

Similarity:

$$\frac{y_1 - y_2}{y'1 - y'2} = \frac{y_2 - y_3}{y'2 - y'3} = \frac{y_3 - y_1}{y'3 - y'1}$$
$$\frac{x_1 - x}{x'1 - x'2} = \frac{x_2 - x_3}{x'2 - x'3} = \frac{x_3 - x_1}{x'3 - x'1}$$

NP-hard and NP-Complete Analysis

NP-hard

What does NP-hard mean? A lot of times you can solve a problem by reducing it to a different problem. I can reduce Problem B to Problem an if, given a solution to Problem A, I can easily construct a solution to Problem B. (In this case, "easily" means "in polynomial time.")

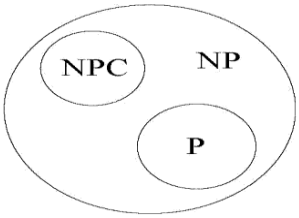


Fig 3.3 NP hard and complete

NP-complete

A problem is **NP-complete** if the problem is both

- NP-hard, and
- In NP.

A technical point: O (n) actually means the algorithm runs in asymptotically linear time, which means the time complexity approaches a line as n gets very large. Also, O (n) is technically an upper bound, so if the algorithm ran in sub linear time you could still say it's O (n), even if that's not the best description of it

IV. WORK DONE

In this section we are discussing the practical environment, scenarios, performance metrics used etc.

4.1 Input

In this user signature is the input for our practical experiment.

4.2 Hardware and Software Configuration

Hardware Requirements:

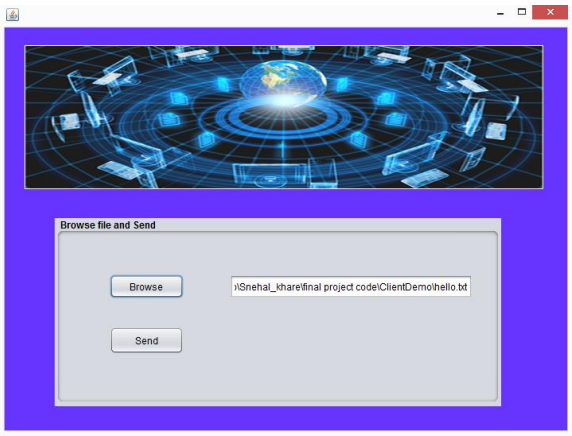
Processor : Pentium IV 2.6 GHz
Ram : 512 MB DD RAM
Monitor : 15" COLOR
Hard Disk : 20 GB

Software Requirements:

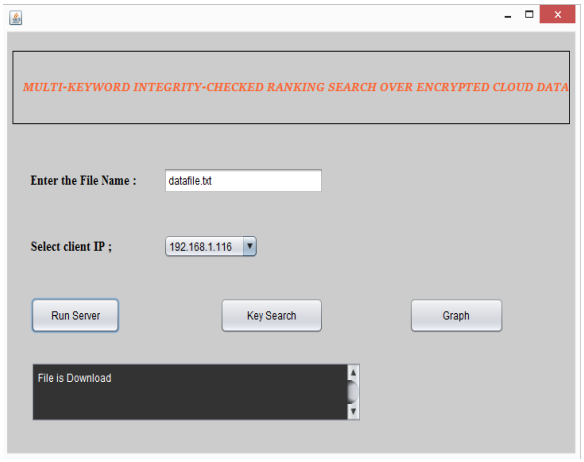
Front End : Java
Tools Used : NetBeans
Operating System : Windows 7/8
Database : MySQL
Cloud server : Amazon.com, somee.com

V. RESULTS

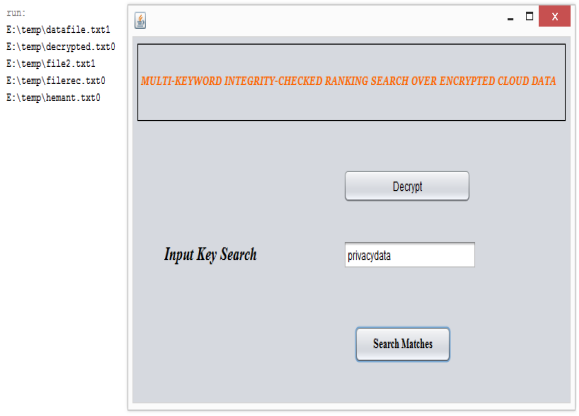
1. Client Side:



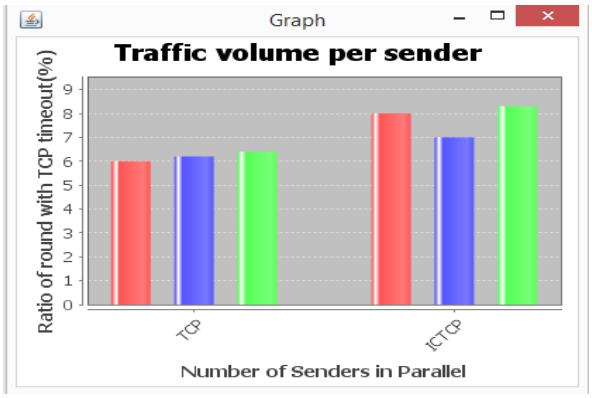
2. Server Side:



Multikeyword Search Result:



TCP Ration:



VI. CONCLUSION

Among various multi-keyword linguistics, we elect the economical similarity measure of "coordinate matching", i.e., as severalmatches asdoable, to effectively capture the connectedness of outsourced documents to the question keywords, and use "inner product similarity" to quantitatively valuate such similarity live. For meeting the challenge of supporting multi-keyword linguistics without privacy breaches, we tend to propose a basic plan of MRSEusing secure dot product computation. Then we tend toimproved MRSE schemes to realize varieddemanding privacy requirements in completely different threat models. Thorough analysis investigating privacy and potency guarantees of projected schemes is given, and experiments on the real-world dataset show our projected schemes introduce low overhead on each computation and communication.

REFERENCES

[1] L. M. Vaquero, L. Rodero-Merino, J. Caceres, and M. Lindner, "A break in the clouds: towards a cloud definition," ACM SIGCOMM Compute. Commun. Rev., vol. 39, no. 1, pp. 50–55, 2009.

[2] S. Kamara and K. Lauter, "Cryptographic cloud storage," in RLCPS, January 2010, LNCS. Springer, Heidelberg. [3] A. Singhal, "Modern information retrieval: A brief overview," IEEE Data Engineering Bulletin, vol. 24, no. 4, pp. 35–43, 2001.

[4] I. H. Witten, A. Moffat, and T. C. Bell, "Managing gigabytes: Compressing and indexing documents and images," Morgan Kaufmann Publishing, San Francisco, May 1999.

[5] D. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in Proc. of S&P, 2000.

[6] E.-J. Goh, "Secure indexes," Cryptology ePrint Archive, 2003, http:// eprint.iacr.org/2003 /216.

[7] Y.-C. Chang and M. Mitzenmacher, "Privacy preserving keyword searches on remote encrypted data," in Proc. of ACNS, 2005.

[8] R. Curtmola, J. A. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: improved definitions and efficient constructions," in Proc. of ACM CCS, 2006.

[9] D. Boneh, G. D. Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in Proc. of EUROCRYPT, 2004.

[10] Roy.A., Datta A. And Mitchell,J(2008) Formal certification of cryptographic security of Deffen Hellman Based protocol Pages 312-329

[11] Record on D.and Fitzpatrick, B (2011) Opened authentication 1.1, May 2006 URL http: //opened. Net/specs/openid-authentication-1. Html.

[12] J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, "Fuzzy keyword search over encrypted data in cloud computing," in Proc. of IEEE INFOCOM'10 Mini-Conference, San Diego, CA, USA, March 2010.

[13] D. Boneh, E. Kushilevitz, R. Ostrovsky, and W. E. S. III, "Public key encryption that allows pir queries," in Proc. of CRYPTO, 2007.

[14] P. Golle, J. Staddon, and B. Waters, "Secure conjunctive keyword search over encrypted data," in Proc. of ACNS, 2004, pp. 31–45.

[15] L. Ballard, S. Kamara, and F. Monrose, "Achieving efficient conjunctive keyword searches over encrypted data," in Proc. of ICICS, 2005.

[16] D. Boneh and B. Waters, "Conjunctive, subset, and range queries on encrypted data," in Proc. of TCC, 2007, pp. 535–554.

[17] R. Brinkman, "Searching in encrypted data," in University of Twente PhD thesis, 2007.

[18] Y. Hwang and P. Lee, "Public key encryption with conjunctive keyword search and its extension to a multi-user system," in Pairing, 2007.

[19] J. Katz, A. Sahai, and B. Waters, "Predicate encryption supporting disjunctions, polynomial equations, and inner products," in Proc. Of EUROCRYPT, 2008.

[20] A. Lewko, T. Okamoto, A. Sahai, K. Takashima, and B. Waters, "Fully secure functional encryption: Attribute-based encryption and (hierarchical) inner product encryption," in Proc. of EUROCRYPT, 2010.

[21] E. Shen, E. Shi, and B. Waters, "Predicate privacy in encryption systems," in Proc. of TCC, 2009.

[22] C. Wang, N. Cao, J. Li, K. Ren, and W. Lou, "Secure ranked keyword search over encrypted cloud data," in Proc. of ICDCS'10, 2010.

[23] W. K. Wong, D. W. Cheung, B. Kao, and N. Mamoulis, "Secure knn computation on encrypted databases," in Proceedings of the 35th SIGMOD international conference on Management of data, 2009, pp. 139–152.

