American Journal of Networks and Communications
2015; 4(3): 25-31
Published online April 13, 2015 (http://www.sciencepublishinggroup.com/j/ajnc)
doi: 10.11648/j.ajnc.20150403.11
ISSN: 2326-893X (Print); ISSN: 2326-8964 (Online)

Enhanced Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data

Pourush, Naresh Sharma, Manish Bhardwaj

Department of Computer Science and Engineering, SRM University, NCR Campus, Modinagar, India

Email address:
pulkittyagi1991@gmail.com (Pourush), nrssharma@gmail.com (N. Sharma), aapkaapna13@gmail.com (M. Bhardwaj)

To cite this article:
Pourush, Naresh Sharma, Manish Bhardwaj. Enhanced Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data. *American Journal of Networks and Communications.* Vol. 4, No. 3, 2015, pp. 25-31. doi: 10.11648/j.ajnc.20150403.11

Abstract: To protect the privacy, sensitive information has to be encrypted before outsourcing to the cloud. Thus the effective data uct keyword search. Related works on searchable encryption emphasis on single keyword based search or Boolean keyword based search, and hardly work on sorting the search results. Our work focuses on realizing secure semantic search through query keyword semantic extension. We mix-ups and used architecture of two clouds, explicitly private cloud and public cloud. The search process is distributed into two steps. The leading step develops the question keyword upon warehoused database in the private cloud. The subsequent step uses the drawn-out query keywords set to recover the index on public cloud. Finally the matched files are resumed in order. Complete security analysis shows that our explanation is privacy-preserving and secure. Trial evaluation determines the efficiency and effectiveness of the scheme.

Keywords: Motion Detection, Background Modeling (BM), Block Based, Human Detection, Wavelet Threshold Algorithm, Confusion Matrix

1. Introduction

We are existing in an exceedingly organized environment, where enormous measures of information are warehoused in confined, yet not basically trusted servers. There are various protection issues concerning to getting to information on such servers; two of them can just be perceived: affectability of

i. Keywords sent in questions and
ii. The information recuperated; both need to be concealed.

A related convention, Private Information Retrieval (PIR) empowers the client to get to public or private databases without uncovering which information he is extricating. Since protection is of an incredible concern, PIR conventions have been widely considered previously.

In today's information technology scene, clients that need high warehousing and processing power have a tendency to out- source their information and administrations to clouds. Clouds empower clients to remotely store and access their information by bringing down the expense of hardware possession while giving strong and quick administrations. The significance and need of protection saving pursuit procedures are significantly more claimed in the cloud applications. Because of the way that extensive organizations that work people in public clouds like Google or Amazon may get to the delicate information and hunt examples, concealing the question and the recovered information has extraordinary imperativeness in guaranteeing the protection and security of those utilizing cloud administrations.

This research paper concentrates on to the arrangement of multi-keyword ranked search encrypted (MRSE) over cloud information while protecting strict framework perceptive security in the cloud computing ideal model. The query to be tended to here is, given an arrangement of keywords, how would we use as ranking framework to secure cloud information storage and access?

To actuate the ranked search for successful use of outsourced cloud information under the previously stated model, the framework ought to be proposed by considering the security contemplations too. The framework is required to give the accompanying security and execution ensures as follows:

- Multi-keyword Ranked Search: To plan search plans which permit multi-keyword query and give result similarity ranking to authoritative information recovery, as opposed to returning undifferentiated results.
- Privacy-Preserving: To keep the cloud server from taking in extra data from the dataset and the record, and to meet the essential security necessities

- Efficiency: Ranked search should ensure privacy and computation overhead and also low communication.

2. Literature Review

The objective of this literature review is to summarize the data utilization and security issues of various searching techniques in the encrypted cloud data.

Qin Liu proposed this search that gives keyword security, semantic secure and information protection and public key encryption. Here, CSP is included in halfway decipherment by lessening the computational overhead and correspondence in deciphering for clients. The clients submit the keyword trapdoor encrypted by clients' private key to CS safely and recover the scrambled reports. Cong Wang proposed this search which understands transforming overhead, information and keyword protection, least correspondence and processing overhead. The executive construct file alongside the keyword recurrence based pertinence scores for records. Client demand "w" to CS with unrestricted "k" as Tw utilizing the private key. The CS seeks the record with scores and sends scrambled document focused around ranked grouping.

Wenhai Sun proposed this analysis that gives resemblance based query element ranking, keyword security, Index and Query privacy and Query Unlink ability. The encrypted record is assembled by vector space model supporting disjunctive and conjunctive document search. The searchable list is invented utilizing Multidimensional B tree. Holder makes encrypted inquiry vector $\bar{Q}$ for record keyword set. Client gets encrypted question vector of W from holder which is given to CS. Presently CS pursuits list by MD algorithm and compares at cosine measure of record and question vector and returns top k encoded records to client.

J. Baek proposed this strategy, in which CS makes its own particular public-private key pair. Sender encrypts all documents, keyword utilizing servers' and clients' public key before outsourcing. Client demands keyword trapdoor Tw to CS utilizing its private key. CS checks the Tw utilizing servers' private key and returns encrypted record. H. S. Rhee proposed this pursuit in which the outsourcing is carried out as SCF-PEKS. Client demands Tw to CS encoded with servers' open key and clients' private key. CS checks Tw utilizing servers' private key and returns encrypted record matching the keyword. Here the untouchable can't perform KGA without server's private key.

PengXu proposed this search, in which client makes fuzzy keyword trapdoor Tw and definite catchphrase trapdoor Kw for W. Client demands Tw to CS. At that point CS checks Tw with fluffy keyword file and sends superset of matching cipher messages by Fuzztest calculation that is executed by CS. The client process Exacttest calculation for checking ciphertexts with Kw and recover the encrypted records. Ning proposed this search for known cipher content model and foundation display over encoded information giving low calculation and correspondence overhead. The direction matching is picked for multi-keyword search for. They utilized internal item likeness to quantitatively assess comparability for ranking records. The disadvantage is that MRSE have little standard deviation which debilitates keyword protection.

Wenhai Sun proposed Verifiable that gives multi-keyword seeks Privacy-Preserving Multi-keyword Text Search by similitude pursuit based result ranking. Owner outsources encoded report $\check{D}$ utilizing vector space model and confirmed secure list tree manufactured utilizing Multidimensional B-tree encrypted utilizing RSA and SHA-1. Client submits W to holder and gets encoded question vector $\bar{Q}$ for W. The question $\bar{Q}$ alongside inquiry parameter k is given to CS. Presently CS looks $\bar{Q}$ utilizing MD algorithm and thinks about cosine measure of $\bar{Q}$ and $\check{D}$ and returns top k encrypted records to client. At that point client looks this base tree utilizing the same inquiry calculation as CS and checks the question results.

3. Problem Statement

The issue in recovering the important records is that clients may not need archives, which they demand, to be uncovered, since their substance may be sensitive and they are typically specifically identified with query terms in their inquiries. In our plan, the server can furnish a proportional payback records to the client.

In the proposed approach, we require the information manager or its delegate that does not stratagem with the server, to be dynamic. The use of a dynamic agent for the information manager is a typical approach that is cognizant with earlier works. As clarified in proposed architecture, the information holder encodes archives with a symmetric-key encryption technique utilizing an alternate secret key for each one record. The server ought not to have the capacity to decrypt those ciphers since this would suggest that the server takes in the substance of the archive the client demands for. Accordingly, the information manager encodes the symmetric-keys with an open key encryption strategy, which has blinding capacity, and stores the encoded symmetric-keys in the server. In cryptography, blinding is a procedure, whereby a specialists can register a cryptographic capacity (e.g. marking and unscrambling), without knowing either the genuine info or the genuine yield of the capacity. We pick the RSA as the general population key encryption, which assistance blinding.

Accept that the client asks for the archive R. He gets the RSA encryption of the symmetric-key (sk), in particular RSAe(sk), where e means the general public key of the information holder. The client does not know the private key of the RSA (i.e. d), in this way he needs the information holder to perform the unscrambling of sk without indicating y = RSAe(sk), which would uncover the report he recovers. The client utilizes the blinding strategy and interfaces with the information holder for unscrambling the RSA encryption without taking in the private key d. Firstly, y is blinded by an irregular blinding variable c picked by the client as $z = c^e y \bmod N$, where N is the RSA modulus. At that point, the client

sends the blinded result z to the information manager, who unscrambles it utilizing his private key and gives back where its due (ž = zd mod N) over to the client. At last the client un-blinds the result utilizing the blinding variable as sk = ž c-1 mod N. The information manager can't figure out which mystery key it is decoding following the ciphertext is blinded, thus arbitrary looking.

4. Existing Technology

The protection definition for analysis routines in the related text is that the server must study only the query items. We further tighten the protection over this general security definition and build a set of security necessities for privacy-preserving search protocols. A multi-keyword look system must give the accompanying client and information security properties (First instincts and afterward formal definitions are given):

1. Data Privacy: No one however the client can take in the genuine recovered information.
2. Index Privacy: The search file or the inquiry list does not release any data about the comparing keywords.
3. Trapdoor Privacy: Given one trapdoor for a set of keywords, the server can't produce an alternate legitimate trapdoor.
4. Non-Impersonation: No one can imitate an authentic client.

Definition 1- (Data Privacy) A multi-keyword search approach has information protection, if there is no polynomial time foe A that, given the recovered encrypted information and the comparing encoded key, realizes any data about the information.

Definition 2- (Index Privacy) A multi-keyword search approach has record protection, if for all polynomial time foes A that, given two distinctive keyword records L1 and L2 and a list $I_b$produced from the magic word list L_b where b ϵ_R {0, 1}, the playing point of An in discovering b is insignificant. The playing point of A is irrefutably the estimation of the distinction between its prosperity likelihood and 1/2.

Definition 3- (Trapdoor Privacy) A multi-keyword search approach has trapdoor security, if for all polynomial time enemies A that, given a legitimate trapdoor for a set of keywords, A can't create a substantial trapdoor for its subset.

Definition 4- (Non-Impersonation) A multi-keyword search approach has non-mimic property, if there is no foe A that can imitate a genuine client U with likelihood more noteworthy than Є where Є is the likelihood of breaking the signature scheme

5. Proposed Methodology

Under the above system, we propose two answers for searching on encrypted information, in particular APKS and APKS+. We make novel utilization of a late cryptographic primitive; various hierarchical predicate encryption (HPE), which offers assignment of search capacities. Both of our answers empower effective multi-dimensional questions with correspondence, subset and a class of basic extent inquiries. Since the PKC-based SE plans experiences a keyword reference assault that uncovers the basic magic words in an inquiry to the server, in APKS+ we upgrade the question security by keeping that sort of attack with the assistance of extra intermediary servers. To the best of our insight, the APKS+ plan is the first to accomplish productive multi-dimensional extent question, ability appointment and inquiry protection at the same time.

- Architecture:

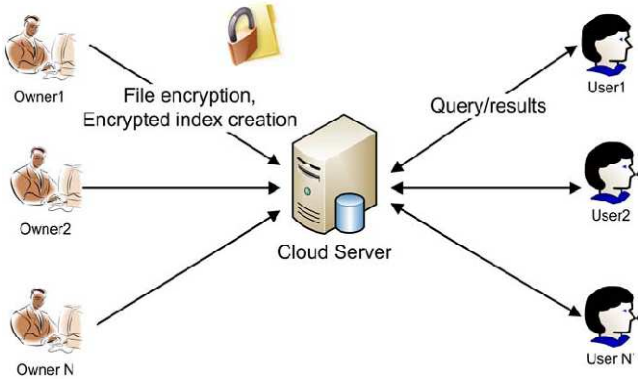

Fig. 1. *Architecture of proposed work.*

- Modules: The substances in the framework are: information managers/clients, trusted powers, and the cloud server.

1. Data Owner: In this paper, information manager depicts somebody who claims the data, e.g., a patient encrypts her information and needs them to be put away in the cloud server while protecting her information. The framework ought to permit numerous owners to encrypt and help information, while empowering a substantial number of clients to search over different managers' information. In accomplishing this, the framework ought to have high versatility, i.e., low key administration overhead. Additionally, productivity ought to be adequate for every inquiry operation from a client's perspective.
2. Cloud Server: The cloud server stores the encrypted information helped by various owners in a database and performs look for the clients.
3. User: The "clients" for the most part allude to the individuals who can perform search for required data over the encrypted database. At the point when a client asks for an ability for inquiry $\hat{} \ Q$ from a LTA, the LTA checks whether a client either really has the quality worth set W basic the $\hat{}$, or is "qualified" for those qualities. One approach to accomplish this is to keep up a database of attribute qualities for all clients in the LTA's neighboring space. On the other hand, the LTA can issue to every client in its area a set of authorizations confirming the client's characteristic values, and confirms those qualifications upon a solicitation for capacity. With a specific end goal to

demonstrate its approval on ability, a TA/LTA can issue a signature that will be identity-based in light of every capacity it created/assigned. The server needs to confirm that a got ability has a legitimate signature from an enrolled LTA before performing quest for a client.

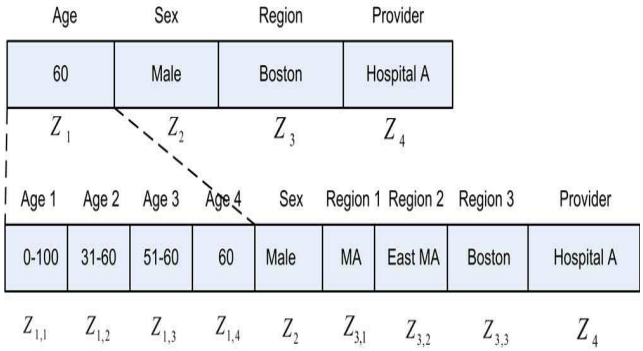

Fig. 2. *Index conversion- Age and region are hierarchical fields.*

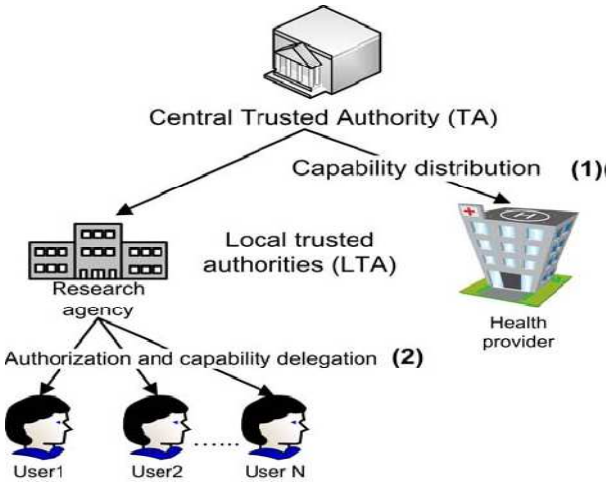

Fig. 3. *Query processing and search using ranked multi keywords.*

4. Encrypted Index Generation and query privacy-Index and Query Privacy: The essential security objective is to keep the cloud server from realizing any helpful data about the encrypted reports, records, and the clients' questions, with the exception of what can be acquired from the query items. List protection refers to confidentiality of the record, while inquiry security ensures clients' inquiries.
5. Search: Multi-dimensional Keyword Search: The framework ought to support multi-dimensional keyword search for usefulness, in particular, we need to help conjunctions among distinctive measurements where in each one measurement there can be different keywords (counting balance, subset and extent inquiries). The normal hunt handling time on single scrambled list under diverse n values It can be seen that the inquiry is much speedier than encryption and is straight to n, since it just takes n + 3 matching operations.

6. Results

"Direction matching", i.e., whatever number matches as would be prudent, is a proficient similitude measure between such multi-keyword semantics to refine the result importance, and has been broadly utilized as a part of the plaintext data recovery (IR) group.

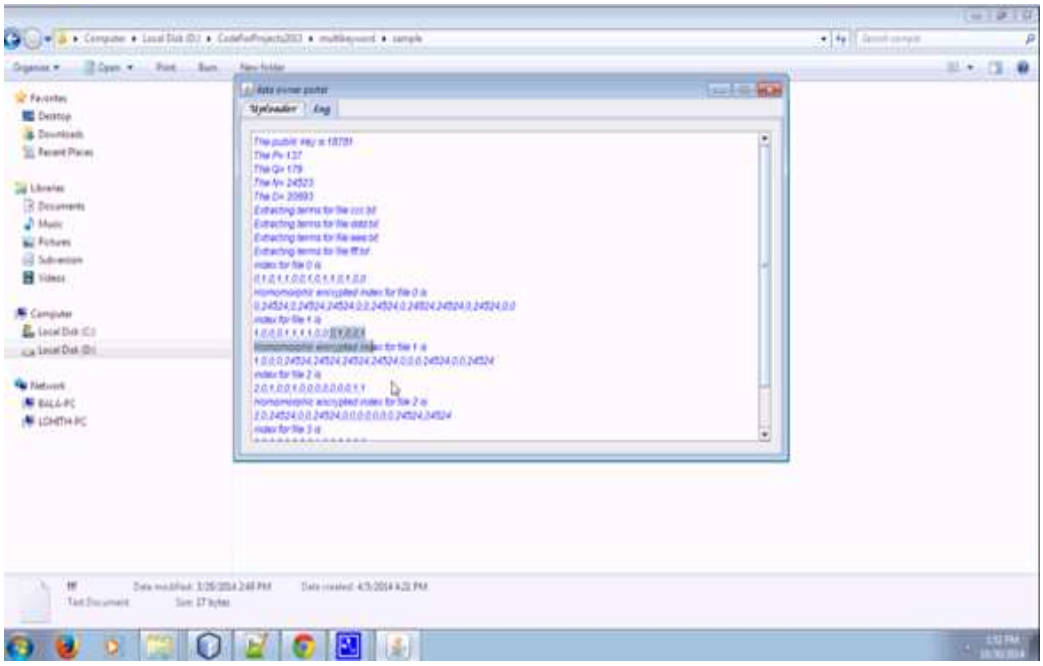

Fig. 4. *Profile of different data users at current time.*

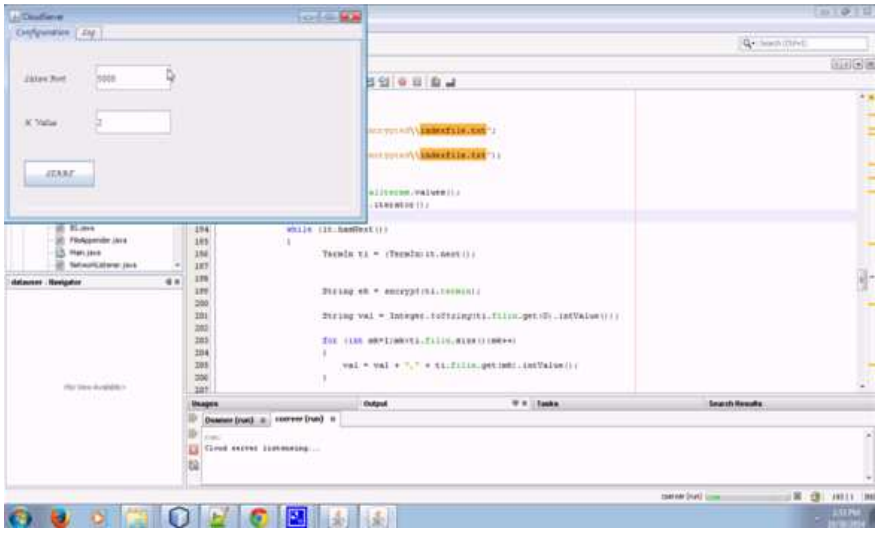

Fig. 5. *Log file calling means index generation.*

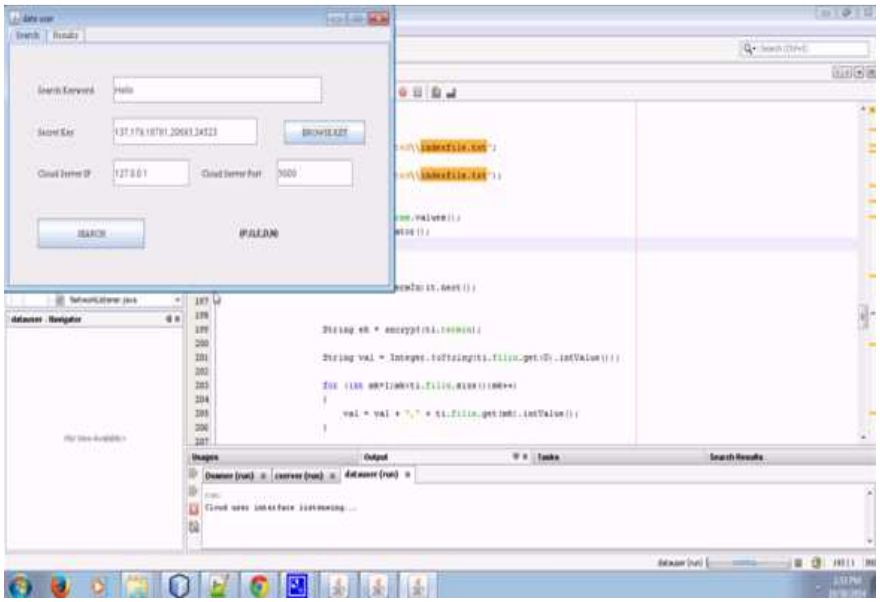

Fig. 6. *Key generation for a particular profile.*

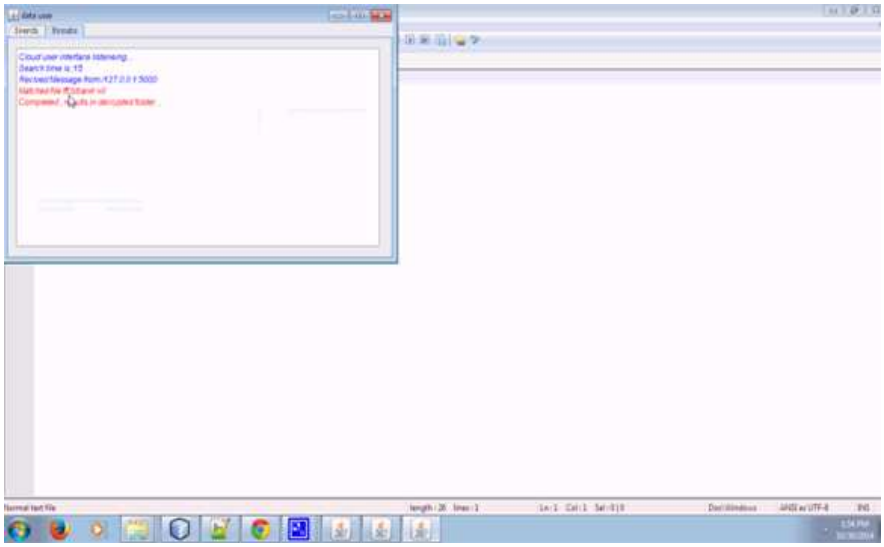

Fig. 7. *Encrypted profile search.*

Fig.4 shows the key pair of public and private are generated and then the using the keys the file data changed into cipher text shown by fig.5. The keyword is indexed for key protection in fig6. The results of user queries are search and ranked shown by fig 7 and fig 8.

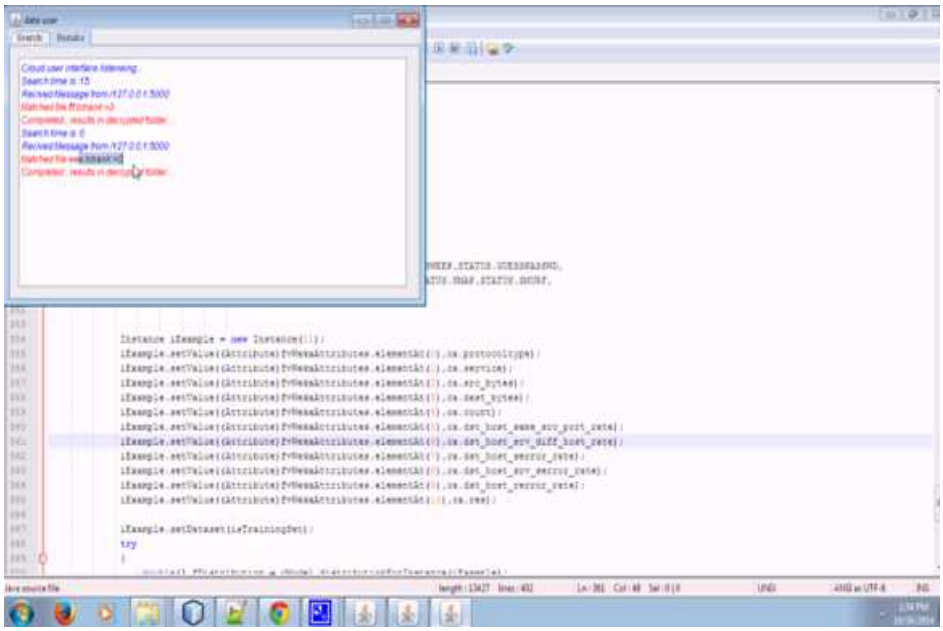

Fig. 8. *Decrypted profile search.*

In any case, how to apply it in the encrypted cloud information search framework remains an extremely difficult assignment as a result of immutable security and protection problems, including different strict necessities like the information protection, the file protection, the keyword protection, and numerous others.

7. Conclusion

Ranked Search can likewise richly eliminate unnecessary system traffic by sending back just the most important information, which is exceedingly attractive in the "pay-as-you- utilization" cloud standard. For protection assurance, such ranking operation, notwithstanding, ought not to release any keyword related data. To put it plainly, none of existing Boolean keyword searchable encryption plans help different essential words ranked inquiry over scrambled cloud information while saving security as we propose to investigate. Our proposed work recognizes the new approach and architecture algorithm to upgrade the execution of multi keyword ranked inquiry.

Firstly, the client side of proposed framework will be actualized on cell phones running Android and ios working frameworks since the potential application situation imagines that clients acquire to the information anyplace and whenever. Also besides, the proposed technique will be tried on a genuine dataset with a specific end goal to think about the execution of our ranking strategy with the ranking strategies utilized as a part of plain datasets that don't include any security or privacy-preserving techniques.

References

[1] Oxford dictionaries, the oec: (June 2011).“Facts about the language.”http://oxforddictionaries.com/page/oecfactslanguage/the-oec-facts-about-the-language.

[2] D. Boneh, E. Kushilevitz, R. Ostrovsky, and W. Skeith. (2007). “Public key encryption that allows pirqueries”; In Advances in Cryptology - CRYPTO, volume 4622 of Lecture Notes in Computer Science, pages 50{67. Springer Berlin / Heidelberg, 2007.

[3] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou. (2011). “Privacy-preserving multi-keyword ranked search over encrypted cloud data”. In IEEE INFOCOM.

[4] Y.-C. Chang and M. Mitzenmacher. (2005). “Privacy Preserving Keyword Searches on Remote Encrypted Data”. In Applied Cryptography and Network Security, pages 442{455. Springer.

[5] D. Chaum. (1982). “Blind signatures for untraceable payments”. In Advances in Cryptology: Proceedings of CRYPTO'82, pages 199{203.

[6] B. Chor, E. Kushilevitz, O. Goldreich, and M. Sudan. (November 1998). “Private information retrieval”. J. ACM, 45:965{981.

[7] L. E. Dickson. (2003). “Linear Groups with an Exposition of Galois Field Theory”. Dover Publications, New York.

[8] M. J. Freedman, Y. Ishai, B. Pinkas, and O. Reingold. (2005). “Keyword search and oblivious pseudorandom functions”. In Theory of Cryptography Conference -TCC 2005, pages 303{324.

[9] J. Groth, A. Kiayias, and H. Lipmaa. (2010).“Multi-query computationally-private information retrieval with constant communication rate”. In PKC, pages 107{123.

[10] W. Ogata and K. Kurosawa. (2004). “Oblivious keyword search”. In Journal of Complexity, Vol.20, pages 356{371.

[11] J. T. Trostle and A. Parrish. (2010). “Efficient computationally private information retrieval from anonymity or trapdoor groups”. In ISC'10, pages 114{128.

[12] L. M. Vaquero, L. Rodero-Merino, J. Caceres, and M. Lindner. (December 2008). “A break in the clouds: towards a cloud definition”. SIGCOMM Computer. Commun. Rev., 39:50{55.

[13] C. Wang, N. Cao, J. Li, K. Ren, and W. Lou. (2010). “Secure ranked keyword search over encrypted cloud data”. In ICDCS'10, pages 253{262.

[14] P. Wang, H. Wang, and J. Pieprzyk. (2009). “An efficient scheme of common secure indices for conjunctive keyword-based retrieval on encrypted data”. In Information Security Applications, Lecture Notes in Computer Science, pages 145{159. Springer.

[15] J. Zobel and A. Mo_at. (1998). “Exploring the similarity space”. SIGIR FORUM, 32:18{34.