A Muilt-Keyword Ranked Based Search and Privacy Preservation of Distributed Documents in the Network

Jyoti Muthreja[1], Arvind Bhagat Patil[2]

[1]*Student, Department of Computer Science and Engineering, Yeshwantrao Chavan College of Engineering, Nagpur, India*
[2]*Associate Professor, Department of Computer Science and Engineering, Yeshwantrao Chavan College of Engineering, Nagpur, India*

---***---

Abstract - *In Information Networks, proprietors can store their archives over conveyed various servers. It encouraging clients to store and get to their information in and from numerous servers by settling down anyplace and on any gadget. It is an extremely difficult assignment to give productive seek on disseminated records additionally give the privacy on owners archives. The current framework gives one conceivable arrangement that is privacy preserving indexing (PPI). In this framework, records are disseminated over various private servers which are all in all controlled by cloud/open server. At the point when client need a few reports, they inquiry to open cloud, which then returns the hopeful rundown that is private server rundown to clients. In the wake of getting rundown, client can seek the records on particular private server however in this framework; reports are put away in plain content shape on private server that is privacy is bargained. In any case, proposed framework improves this current framework to make it more secure and proficient. To start with records are put away in encoded frame on the private servers and after that utilization Key Distribution Center (KDC) for permitting decoding of information got from private server, at customer side. The proposed framework additionally executes TF-IDF, which gives the positioning of results to clients.*

***Key Words*: Information Network, Private Server, Public Cloud, Distributed Databases, Ranking Results**

1. INTRODUCTION

In the time of distributed computing, information clients, while appreciating a huge number of advantages from the public server (e.g. taken a toll viability and information accessibility), are all the while hesitant or even flexible to utilize the mists, as they lose information control. The current research and mechanical endeavours towards returning information control back to public server clients have brought forth an assortment of multi-space public server stages, most outstandingly developing data systems. In a data system, an information proprietor can hold the full control of her information by having the capacity to look over a variety of specialist organizations one that she can apparently trust or even have the capacity to dispatch an individual server administrated straightforwardly without anyone else. The data organize does not require shared trusts between servers, that is, a proprietor just needs to trust her own server and nothing more. Data systems develop in an assortment of utilization regions. For a case, in the undertaking intranet (e.g. IBM YouServ framework [1], [2]), representatives can store and deal with their own particular records on by and by administrated machines. While the representatives have their own privacy concerns and could set up get to control arrangements on the nearby records, they might be required by the corporate level administration group to share certain data for advancing potential joint efforts [2]. For another illustration, a few circulated informal communities e.g. Diaspora [3], Status [4] and Persona [5]) as of late rise and turn out to be progressively well known, which depend on the plan of decoupling the capacity of social data and long range informal communication usefulness. Not at all like the brought together solid long range informal communication (e.g. Facebook and LinkedIn), the appropriated interpersonal organizations permit a normal social client to dispatch an individual server for putting away her own particular social information and implementing self-characterized get to control rules for privacy-mindful data sharing [6]. Different cases of data systems incorporate electronic Healthcare over the general population Internet (e.g. the open-source NHIN Direct venture [7]), distributed document imparting to get to controls [8] and others. In every one of these systems, an information proprietor can have a select area for organization of physical assets (e.g., a virtual machine) and information administration of individual information under the full client control. Spaces situated inside numerous servers are disengaged and questioned between each other.1 Information sharing and trades over an area limit are attractive for different application needs.

For privacy-mindful inquiry and data partaking in the data organizes, an applicant arrangement is a privacy preserving file on get to controlled circulated records [9], [10], [11], or PPI for short. In Fig. 1, a PPI is a catalogue benefit facilitated in a third-gathering substance (e.g. an open cloud) that serves the worldwide information to various information customers or searchers. To discover reports of intrigue, a searcher would take part in a two-arrange look strategy: First she represents an inquiry of significant catchphrases against the PPI server, which gives back a rundown of applicant proprietors (e.g. p0 and p1) in the system.

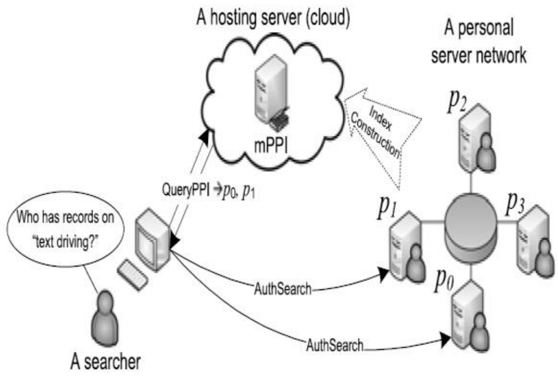

Fig -1: PPI system

At that point for every hopeful proprietor in the rundown, the searcher contacts its server and solicitations for client confirmation and approval before seeking locally there. Take note of that the confirmation and approval just happen inside the data arrange, yet not on the PPI server.

Contrasting with existing work on secure information serving in the cloud [12], [13], [14], the PPI plan is extraordinary as in 1) Data is put away in plain-content (i.e. without encryption) in the PPI server, which makes it feasible for proficient and adaptable information presenting with rich usefulness. Without utilization of encryption, PPI jam client privacy by adding clamours to cloud the touchy ground truth data. 2) Only coarse-grained data (e.g. the ownership of a sought expression by a proprietor) is put away in the PPI server, while the first substance which is private is still kept up and ensured in the individual servers, under the client determined get to control rules.

In the PPI framework, it is alluring to give separated privacy protection with respect to various search queries and proprietors. The information demonstrates utilized as a part of a PPI framework and a data system is that every server has different records, each comprising numerous terms. What is esteemed private and ought to be secured by a PPI is the ownership data as "whether a proprietor has no less than one record significant to a multi-term express." Under this model, the importance of separated privacy conservation is of two folds: 1) Different (single) terms are not conceived rise to as far as how touchy they are. For instance, in an eHealthcare organize, it is normal for a lady to think of her as therapeutic record of a "premature birth" operation to be substantially more delicate than that of a "hack" treatment. 2) A multi-term state, as a semantic unit, can be a great deal increasingly (or less) delicate than a solitary term contained in the expression. For example, "content" and "driving" are two terms that might be regarded non-delicate in their lone appearances; however a record of "content driving" can be viewed as more touchy.

The current PPI work [9], [10], [11], while intended to ensure privacy, is not ready to separate privacy conservation on various terms. Because of the quality-rationalist strategies utilized for developing these PPIs, they cannot convey a quantitative certification for privacy safeguarding for inquiry of a solitary term, not to mention that of a multi-watchword express.

In this paper, we propose ϵ-MPPI, another PPI deliberation which can quantitatively control the privacy spillage for multi-watchword record look. In the ϵ-MPPI framework, distinctive expressions, be it either a solitary term or a multi-term expression, can be designed with a proposed degree on privacy, meant by and can be of any an incentive from 0 to 1; Value 0 speaks to minimal worry on privacy conservation, while esteem 1 goes for the best privacy safeguarding (conceivably to the detriment of additional inquiry overheads). By this implies, an aggressor, seeking a multi-term state on ϵ-MPPI, can just have the certainty of mounting effective assaults limited by what the expression's privacy degree permits.

Building a ϵ-MPPI from a data system is trying from the points of both the calculation and framework outlines. Computationally, the ϵ-MPPI development requires watchful plan to legitimately include false positives (i.e. a proprietor who does not have a term or an expression erroneously claims to have it) so that a genuine positive proprietor can be covered up among the false positive ones, in this manner preserving privacy.

Regarding framework outlines, in a genuine data organize which needs shared trusts between self-rulingly worked servers; it is vital and attractive to develop ϵ-MPPI safely without a put stock in expert. The assignment of disseminated secure development would be extremely testing. On one hand, developing ϵ-MPPI to meet the stringent privacy imperatives under various multi-term looks while limiting additional hunt expenses can be basically displayed as an enhancement issue, tackling which requires complex calculations, for example, a non-straight programming or NLP.

Then again, while the basic intelligence for secure calculations (as required by the safe ϵ-MPPI development) is to utilize a multi-party calculation (MPC) system or MPC [15], [16], [17], [18] which ensures input information privacy, the current MPC strategies can work practically well just with a basic workload in a little system. For instance, FairplayMP [16], an agent useful MPC stage, "needs around 10 seconds to assess (extremely straightforward) capacities" [19] which should generally be possible inside milliseconds by the consistent non-secure calculation. Straightforwardly applying the MPC procedures to the ϵ-MPPI development issue which includes a mind boggling calculation and a substantial number of individual servers could prompt to a cost that is genuinely stupendous and for all intents and purposes unsatisfactory. To address the difficulties of proficient secure ϵ-MPPI development, our centre thought is to draw a line between the safe part and non-secure part in the calculation show. We limit the safe calculation part however much as could reasonably be expected by investigating different strategies (e.g. calculation reordering).

By along these lines, we have effectively isolated the perplexing NLP calculation from the MPC part to such an extent that the costly MPC in our ϵ-MPPI development convention just applies to an extremely straightforward computational errand, therefore advancing general framework execution.

The contribution of this paper can be abridged as taking after.

•We proposed ϵ-MPPI to address the necessities of separated privacy security of multi-term expresses in a PPI framework. To best of our insight, ϵ-MPPI is the principal chip away at the issue. ϵ-MPPI ensures the quantitative privacy insurance via precisely controlling the false encouraging points in a PPI and in this manner successfully constraining an aggressor's certainty.

•We proposed a suite of down to earth ϵ-MPPI development conventions material to the system of commonly untrusted individual servers. We particularly thought to be both the single-term and multi-term state cases, and improved the execution of the secure ϵ-MPPI development from both edges of calculation model and framework configuration by investigating the thoughts of rearranging the protected calculation undertakings however much as could be expected while without giving up the nature of privacy safeguarding.

•We executed a working model for ϵ-MPPI, in light of which a trial consider affirms the execution favourable position of our list development convention.

2. MODULES AND METHODOLOGY

Framework comprises of open cloud server, numerous private servers and different clients. The proprietors archives are store on private servers in disperse way. The records are put away in scrambled configuration. AES calculation is utilized for information encryption. Every private server made its file record of information. Observing framework gathers all records and combining them. This consolidated file is then put away at open cloud. Presently, if customer needs some record from server, it represents an inquiry to open cloud. In returns, open cloud gives the consolidated record got from observing framework. Presently from this last consolidation list, customer having the rundown of private server at which question related information is put away. At that point to get to the information at server, customer sends the confirmation asks for with client name and watchword.

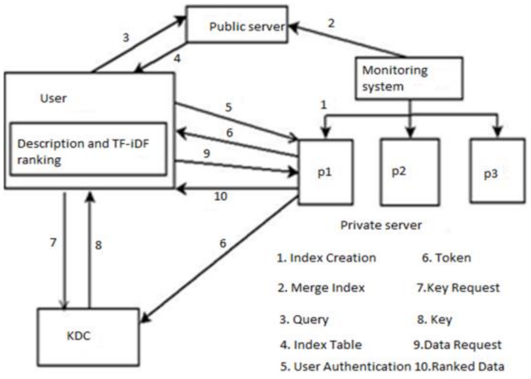

Fig -2: System Architecture

Private server confirms this subtle elements store in its database. After fruitful check, private server creates the token and sends it to customer and Key Distribution Center (KDC). In the wake of getting these token, clients demand to KDC for a key. KDC confirm this token with its token which is as of now getting from private server. After check, KDC gives encryption key to the customer. At that point customer send information demand to private server in returns server gives all coordinating scrambled documents. Utilizing key customer can unscramble the information. Lastly apply the TF-IDF positioning calculation, to get all outcomes in positioning configuration.

System consisting of following modules:

• **System Deployment**

Registration And Login with Database, Client and Server with attachment programming and information exchange AES Encryption and Decryption with Client side GUI.

• **MPPI Index creation algorithm**

MPPI calculation is utilized for making list of all private servers. List speaks to the detail portrayal of information store at private server.

• **Index combining and Upload on Public Server**

Checking framework is in charge of joining list of every private server and transfers this last consolidation file record on open cloud.

• **Input Query and Response from Public Server**

Client represents an inquiry to cloud server for receiving specific information from private server consequently open cloud gives consolidate file.

• **Client Authentication and token generation**

Subsequent to getting file, client needs to associate with private server to get the outcomes. Client login to the server and in the wake of finishing effective validation, private server create and disseminate the token to client and KDC.

• **Key Distribution and File Decryption**

After check of tokens, KDC give the way to client to decoding of results got from private server.

• **TF IDF Ranking Results**

After confirmation, client gets the outcomes from private server in scrambled organization. These scrambled outcomes are then unscrambled utilizing key acquired from KDC. At long last create the positioning of comes about by utilizing TF IDF.

3. MATHEMATICAL MODEL FOR PROPOSED WORK

Let S be a System.

S= {I, P, O}

Where,

• Input I: The input for the system is multi word query from the user.

• Output O: Ranking results.

• Process P:

(a) Single-Term Publication

Where, βj is number of probability values produces by source analytical computation for term.

(b) False Positive Rate

FP (0; 1) = F (0; 1)

Where, FP (0, 1) is the false positive values, β0; β1 are the probability at which a non-positive owner publishes data as a positive owner.

(c) Index Generation

I= {I1, I2... In}

Where I is the set of all index of all private servers

(d) Merge and Upload Index at Private Server

MI= {MI1, MI2... Min}

Where MI is the set of all merge indexes collected from monitoring system.

(E) User Query to Public Server

Q= {Q1, Q2... Qn}

Where, Q is the set of all queries poses to public cloud.

(F) User Authentication at Private Server

U= {U1, U2... Un}

Where U is the set of all authenticated users of private server.

(G) Token Generation and Distribution

T= {T1, T2... Tn}

Where T is the set of all tokens generated by private server for its authenticated users.

(H) Key Generation at KDC

G= {G1, G2... Gn}

Where G is the set of all keys stored at KDC, used for decryption of data at user side.

(I) Data Decryption and TF IDF Ranking

D= {D1, D2... Dn}

Where D is the set of all ranked results for particular input query

4. ALGORITHMS

4.1 Advanced Encryption Standard (AES) Algorithm

AES is a block cipher with a square length of 128 bits. AES licenses for three differing key lengths: 128, 192, or 256 bits. The encryption procedure utilizes an arrangement of especially inferred keys called round keys. AES is an iterative as opposed to Feistel figure. AES utilizes 10 rounds for 128-piece keys, 12 rounds for 192-piece keys and 14 rounds for 256-piece keys. The piece to be encoded is only an arrangement of 128 bits. Each round of handling contains one single-byte based substitution step, a line savvy stage step, a segment insightful blending step, and the expansion of the round key. The request in which these four stages are executed is diverse for encryption and decryption.
Encryption Steps:-
(a) Byte Substitution (SubBytes)
(b) Shift rows
(c) Mix Columns
(d) Add round key

Decryption Steps:-
(a) Add round key
(b) Mix columns
(c) Shift rows
(d) Byte substitution

4.2 TF-IDF

The term frequency inverse document frequency (TF IDF), is a numerical statistic that is proposed to reflect how significant a word is to a document in a corpus or collection. The TF-IDF value increases proportionally to the number of times a word appears in the document, but is equalizing by the frequency of the word in the corpus, which assist to regulate for the information that some words appear more frequently in general.
TF: Term Frequency, which measures how frequently a term occurs in a document. Since every document is different in length, it is possible that a term would appear much more times in long documents than shorter ones.
TF (t) = (Number of times term t appears in a document) / (Total number of terms in the document).
After calculating the TF values for the entire terms top 5 terms will be selected for generating the index. A table will be creating a table and the keyword obtained for index generation will be inserted. The generated table will contain the filename, keywords i.e., the word which will be used for index generation server Id and the size of the file. In further processing this table will be uploaded and sent to monitoring server for further processing.
IDF: Inverse Document Frequency, which measures how important a term is. While computing TF, all terms are considered equally important. However it is known that certain terms, such as "is", "of", and "that", may appear a lot of times but have little importance. Thus we need to weigh down the frequent terms while scale up the rare ones, by computing the following:
IDF (t) = loge (Total number of documents) / (Number of documents with term t in it).

4.3 Iterative-Publish (Owner Pi, set β0 (rk))

```
for all k ϵ [0; l -1] do β' (rk) is topologically sorted
    if match(cur-memvec, getStartingState(rk)) then Bcur⍰memvec is the current membership vector
        cur-memvec publish (cur-memvec, β' (rk))
    end if
end for
```

To publish data with multiple probabilities for overlapping phrases, we propose to use the IBeta approach. Algorithm illustrates how the index publication approach iteratively runs, phrase by phrase.

5. RESULT

By using non-grouping based approach of E-PPI the proposed system will going to provide better preservation of user's privacy in terms of data confidentiality through encryption and better quality of results i.e. relevant results to the queried using ranking techniques.

(a) Time Measurement Graph

Time measures for different process like uploading the file, query searching, encryption time, token generation, and ranking time.

(i) Time Graph for Different Servers

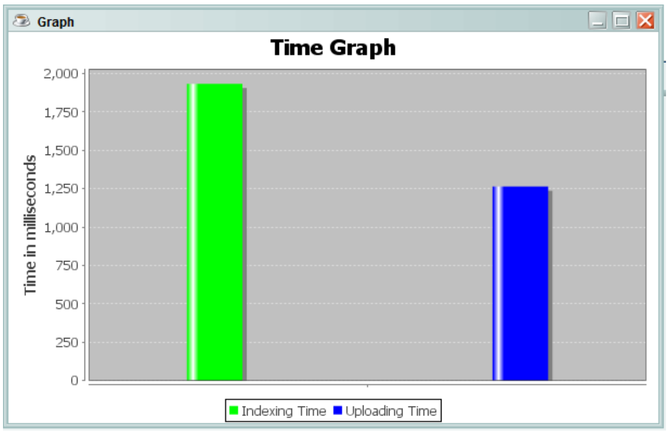

Chart -1: Time Graph for Private Server 1

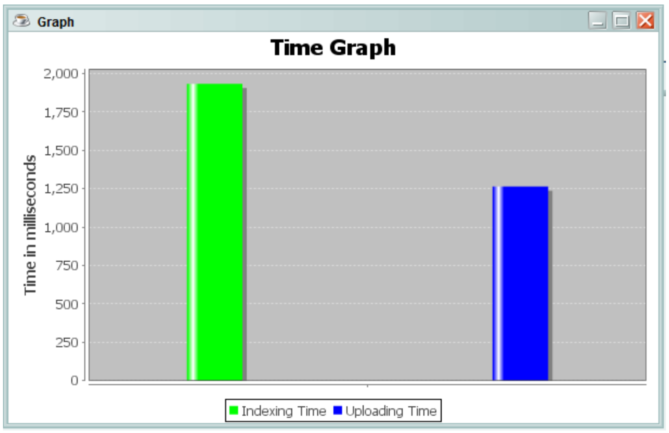

Chart -2: Time Graph for Private Server 2

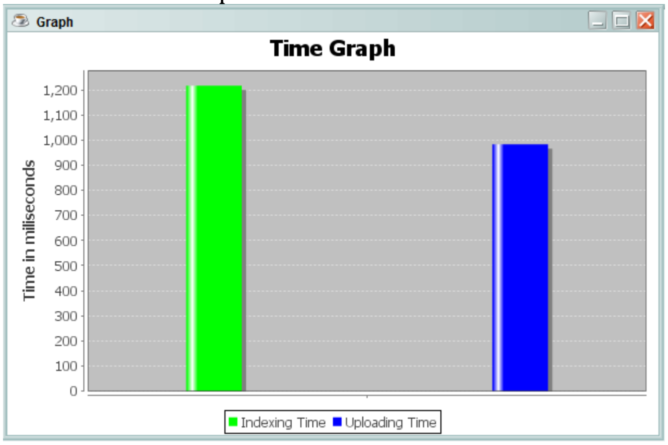

Chart -3: Time Graph for Private Server 3

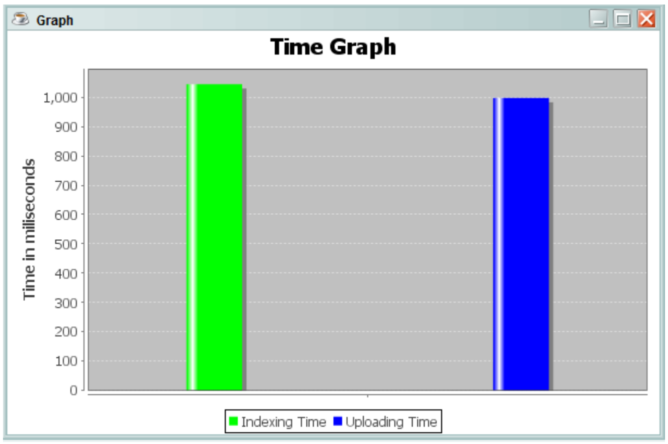

Chart -4: Time Graph for Private Server 4

(ii) Time Graph at User Side

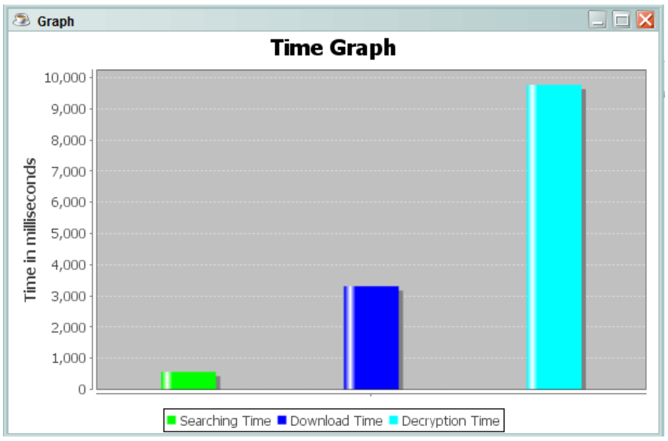

Chart -5: Time Graph at User Side

(b) Accuracy Graph for Keyword Count

Chart -6 represents the accuracy of the system with respect to the keywords searched. Its shows the search accuracy i.e., the count for each keyword available on all the servers.

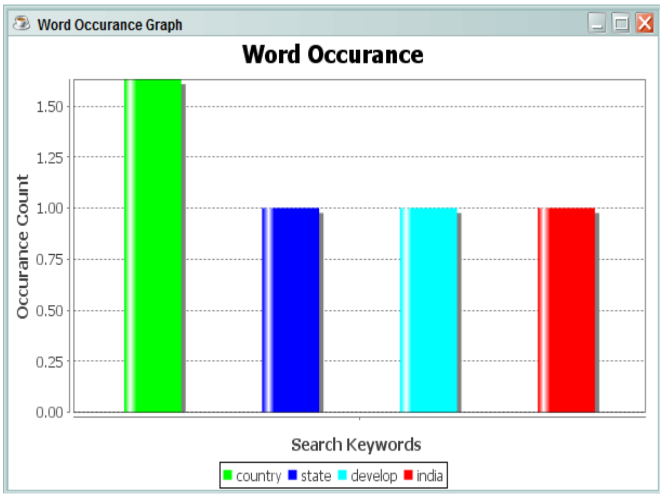

Chart -6: Accuracy Graph

6. CONCLUSIONS

The proposed framework is about connecting between neighborhood server and cloud server for information sharing among the clients. Some validation is required to get to particular information or data. This validation is dealt with through encryption framework. For sensible execution of secure calculations, it proposes Associate in Nursing MPC lessening system bolstered the conservative utilization of mystery sharing plans. Along these lines, through the proposed framework client can get an entrance to required information in positioned arrange utilizing PPI and encryption method.

ACKNOWLEDGEMENT

I take this opportunity to express my profound gratitude and deep regards to my guide Prof. Arvind Bhagat Patil for his exemplary guidance, monitoring and constant encouragement which helped me in completing this task through various stages. The Blessings, help and guidance given by his time to time shall carry me a long way in the journey of life which I am about to embark.

REFERENCES

[1] Yuzhe Tang and Ling Liu, “Privacy - Preserving Multi-Keyword Search in Information Networks,”IEEE Transactions On Knowledge And Data Engineering, Volume 27, Issue 9, 2015.
[2] Deepali D. Rane and Dr.V.R.Ghorpade “Multi-User Multi-Keyword Privacy Preserving Ranked Based Search Over Encrypted Cloud Data,”International Conference on Pervasive Computing (ICPC), Volume 2, Issue 4,2015.
[3] C. Wang, N. Cao, J. Li, K. Ren, and W. Lou, “Secure ranked keyword search over encrypted cloud data,”In Proceeding of ICDCS’10, Volume 4, Issue 1, 2010.
[4]Zhihua Xia, Member, IEEE, Xinhui Wang, Xingming Sun, Senior Member, IEEE, and Qian Wang, Member, IEEE “A Secure and Dynamic Multi-keyword Ranked Search Scheme over Encrypted Cloud Data,”IEEE Transactions On Parallel And Distributed Systems, Volume 1, Issue 3,2015.
[5] Ning Cao, Cong Wang, Ming Li, Kui Ren, and Wenjing Lou, “Privacy-Preserving Multi-keyword Ranked Search over Encrypted Cloud Data,”IEEE Infocom, Volume 3, Issue 8, 2014.
[6] Yuzhe Tang , Ling Liu , Arun Iyengar , Kisung Lee, Qi Zhang, “ E-PPI: Locator Service in Information Networks with Personalized Privacy Preservation,”IEEE Transactions On Knowledge And Data Engineering, Volume 7, Issue 6, 2015.
[7] Yuzhe Tang and Shuigeng Zhou, “LHT: A Low-Maintenance Indexing Scheme over DHTs,”28th International Conference on Distributed Computing Systems, Volume 10, Issue 3, 2008.
[8] Randy Baden, Adam Bender, Neil Spring, Bobby Bhattacharjee, “Persona: An Online Social Network with User-Defined Privacy,”ACM SIGCOMM, Volume 9, Issue 6, 2009.
[9] K.S.Sureh, Mrs. SaritaChowdary, T. Balachary. “ A Cloud Based System for Patient Health Records Using Symmetric Encryption,”International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 11, 2013.
[10] Yuzhe Tang, Ting Wang, Ling Liu, Shicong Meng, and Balaji Palanisamy. “Privacy-Preserving Indexing for eHealth Information Networks,” ACM CIKM, Volume 2, Issue 4, 2011.
[11] Mayank Bawa, Roberto J. Bayardo Jr., Rakesh Agrawal, “Privacy-Preserving Indexing of Documents on the Network,”29th VLDB Conference, Volume 4, Issue 6, 2003.
[12] Assaf Ben-David, Noam Nisan, Benny Pinkas, “Fairplay MP – A System for Secure Multi-Party Computation,” ACM CCS, Volume 7, Issue 5, 2008.
[13] Sergej Zerr, Elena Demidova, Daniel Olmedilla, Wolfgang Nejdl, Marianne Winslett2 and Soumyadeb Mitra, “Zerber: r-Confidential Indexing for Distributed Documents,”ACM EDBT, Volume 8, Issue 6, 2008.
[14] Ming Li, Shucheng Yu, Yao Zheng, Kui Ren, and Wenjing Lou, “Scalable and Secure Sharing of Personal Health Records in Cloud Computing using,”IEEE Transactions On Parallel And Distributed Systems Volume 24, Issue 1, 2013.
[15] Preethi Mathew, Dr. S. Sasidhar Babu , “Secure Fuzzy Multi-Keyword Ranked Search over Encrypted Cloud Data,” International Journal of Innovative Research in Computer and Communication Engineering, Volume 3, Issue 8, 2015.
[16] C. Gentry, “Fully Homomorphic encryption using ideal lattices,”41st Annual ACM Theory Computer, Volume 10, Issue 6, 2009.
[17] Ben-David, N. Nisan, and B. Pinkas, “Fairplaymp: A system forsecure multi-party computation,” ACM Conference Computer Commuication Security, Volume 6, Issue 9, 2008.
[18] W. Henecka, S. Kogl, A.-R. Sadeghi, T. Schneider, and I. Wehren- Eberg, “TASTY: Tool for automating secure two-partycomputations,”ACM Confernce Computer Communication Security, Volume 7, Isssue 3, 2010.
[19] Damgard, M. Geisler, M. Krøigaard, and J. B. Nielsen,“Asynchronous multiparty computation: Theory andimplementation,”International Conference Practice Theory Public Key Cryptography, Volume 5, Issue 4, 2009.
[20] M. Kantarcioglu and C. Clifton, “Privacy-preserving distributed mining of association rules on horizontally partitioned data,”IEEE Transaction Knowledge Data Engineering, Volume 16, Issue 9, 2015.
[21] Y. Tang, J. Xu, S. Zhou, and W. Lee, “m-LIGHT: Indexing multi-dimensional data over DHTs,” IEEE International Conference Distributed Computing System, Volume 7, Issue 2, 2009.
[22] Y. Tang, S. Zhou, and J. Xu, “LIGHT: A query-efficient yet low-maintenance indexing scheme over DHTs,”IEEE Transaction Knowledge Data Engineering, Volume 22, Issue 1, 2010.
[23] Machanavajjhala, J. Gehrke, D. Kifer, and M. Venkitasubrama-niam, “L-diversity: Privacy beyond k-

anonymity,"International Conference Data Engineering, Volume 2, Issue 5, 2006.

[24] Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibratingnoise to sensitivity in private data analysis,"Conference Theory Cryptography, Volume 2, Issue 6, 2006.

[25] M. K. Reiter and A. D. Rubin, "Crowds: Anonymity for web trans-actions,"ACM Transaction Information System Security, Volume 1, Issue 1, 1998.

[26] B. Bloom, "Space/time trade in hash coding with allowable errors," Communications of ACM, Volume 2, Issue 7, 1970.

[27] R. J. Bayardo Jr., A. Somani, D. Gruhl, and R. Agrawal. Youserv, "A web hosting and content sharing tool for the masses,"Conference on World Wide Web (WWW), Volume 5, Issue 7, 2002.

[28] M. K. Reiter and A. D. Rubin. Crowds:, "Anonymity for Web transactions,"ACM Transactions on Information and System Security, Volume 5, Issue 3, 1998.

[29] Kui Ren et al.,"Towards Secure And Effective Data utilization in Public Cloud,"IEEE Transactions on Network, Volume 26, Issue 6, 2012.

[30] Ning Cao et al.," Privacy-Preserving Multi- Keyword Ranked Search over Encrypted Cloud Data," IEEE Transactions on Parallel and Distributed Systems, Volume 25, Issue1, 2014.

[31] Ming Li et al., "Toward Privacy-Assured and Searchable Cloud Data Storage Services,"IEEE Transactions on Network, Volume 27, Issue 4, 2013.

[32] A. Singhal, "Modern Information Retrieval: A Brief Overview," IEEE Data Engineering Bull., Volume 24, Issue 4, 2001.

[33] Jianfeng Wang et al., "Efficient Verifiable Fuzzy Keyword Search over Encrypted Data in Cloud Computing,"Journal of Computer Science and Information System, Volume 10, Issue 2, 2013.

[34] Wei Zhou et al., "K-Gram Based Fuzzy Keyword Search over Encrypted Cloud Computing ,"Journal of Software Engineering and Applications, Scientific Research, Volume 29, Issue 6, 2013.

[35] Jiadi Yu, Peng Lu, Yanmin Zhu, Guangtao Xue, Member, IEEE Computer Society, and Minglu Li, "Toward Secure Multi keyword Top k Retrieval over Encrypted Cloud Data," IEEE Transactions on Dependable and Secure Computing, Volume 10, Issue 4, 2013.

[36] Peng Xu et al., Public-Key Encryption with Fuzzy Keyword Search: A Provably Secure Scheme under Keyword Guessing Attack,"IEEE Transactions on Computers, Volume 62, Issue 11, 2013.

[37] Cong Wang et al., "Enabling Secure and Efficient Ranked Keyword Search over Outsourced Cloud Data," IEEE Transactions on Parallel and Distributed Systems, Volume 23, Issue 8, 2012.

[38] Shih-Ting Hsu et al., "A Study of Public Key Encryption with Keyword Search," International Journal of Network Security, Volume 15, Issue 2, 2013.

[39] D. Boneh et al.,,"Public key encryption with keyword search," In Advances in Cryptology – EUROCRYPT, Volume 30, Issue 5, 2004.

[40] H. S. Rhee et al.,"Trapdoor security in a searchable public-key encryption scheme with a designated tester,"The Journal of Systems and Software, Volume 83, Issue 5, 2010.