INTERNATIONAL JOURNAL OF RESEARCH IN COMPUTER APPLICATIONS AND ROBOTICS
ISSN 2320-7345

HETEROGENEOUS FLEXIBLE DATASET MANAGEMENT AND ASSESSEMENT USING AUTOMATED DECODING KEY GENERATOR

R. Aruna Devi[1], Dr. K. Nirmala[2]

[1]Research Scholar, Manonmaniam Sundaranar University,
Tirunelveli, Tamilnadu,India.
[2]Associate Professor, Department of Computer Science,
Quaid-e-Millath Government College for Women, Chennai, Tamilnadu, India.

Abstract: - Applications of big data increase its information sharing by felicitating collaboration of various sectors in data management system. Existing system integrates peer nodes in elastic data and shares network applications which handle simply node to node data sharing in cloud based system. There is no subsist that assimilate various sectors together and incorporates their course of action simultaneously devoid redundancy and colliding with each other corresponding process. Moreover in case of integrity assessment there are existing techniques that equip assessor for checking data integrity and incorporates proxy system that generates decoding key instead of data owner to be online during the entire process. However they have no solution for insecurity caused when an assessor or from the assessor system any intruder tends to attack. In our proposed health care system globalizing various set of databases like patient information, doctor reports, research centre searches and insurance claiming particulars will be assorted into a centralised flexible datasets. Furthermore we synchronize heterogeneous sectors of data as combined flexible datasets without redundancy and retains incognito of details to unauthorized user. Besides associated details in cloud network it assures its integrity and data righteousness by scrutinizing it through the external assessor. The assessor efficiently scrutinizes the correctness of data and delivers the report by evaluating them. Instead of the presence of data owner to be online always we propose Automated Decoding Key Generator that automatically generates and sends decoding key along with self destruction timer to the assessor providing complete security to the data. When the time assessor supposed to process exceeds the destruction program will activate and discard the key. Moreover the key generator will have a spare key that if in case the assessor missed any key or doing within the time and having valid reason then that spare key will regenerate all the other keys. If more than one assessor is participating then their consistency and reliability will be assessed and rated by reputation scores and according to that the trustworthy data will be allotted to them.

Keywords*:* Flexible reclaiming datasets, heterogeneous data, data integrity, assessment evaluation, Automated Decoding Key Generator, self destruction timer, reputation scores.

1. Introduction

In Existing management system huge datasets forms chaotic data processing for authorizing relevant information by storing and retrieving information. For collaborating and maintaining sharing parts of data from different sectors that enables precise data sharing will have analytical and organizational queries. Data sectors subsequently having queries for deploying internal process have to support and install large scale investment for high cost maintenance. Dynamic and flexible data sharing maximizes the process and handles cost reduction

solutions. In a distributed environment network based sharing access interface shares end to end data and retrieves node structured data and related queries arose for performance in query processing.

In such organized dataset the stored data helps in procuring patient details, their disease details, their payment information and diagnosed reports all will be resolved remotely by specialists and fixes appointment for consultation. Researchers retrieve information such as disease details, treated medications, panacea for each type of disease with different symptoms. In this type of integrating the flexible data for collaboration of information efficiently provides strong security in cloud network storage. Data fragmented into chunks and associated in various datasets accommodating less data storage space. In the dynamic cloud storage the redundant data will be eliminated and accessed globally disregarding the locations of patients. Optimal performance implies reduction of data replication and chunk of data fragmentation helps in segregating data and categorizing it.

The cloud storage overhead for outsourced data where the data owners have to control eventual risks like data correctness, availability and data veracity against the external intruders attempt to threat, alter or maliciously delete data from the cloud service. Verifying integrity and refurbishing data constraints from data outsourced could be difficult and expensive. Retrieving the complex and outsourced data the verification and refurbishing process for auditing schemes ensures users computational sources and also being online is burdensome. External assessors and semi-trusted proxy authenticates and regenerates codes for deciphering the data for assessment.

Medical, disease data, research and insurance oriented data will be estimated with instant billing, claiming insurance, diagnosing results and distribution of results online without visiting health care centre each and every time. Users dealing with payment details, dealing with disease details, dealing with personal health history, diagnosing disease in accordance with test results and medicines failed and cured disease all the above are heterogeneous kind of datasets without direct interoperability with each other. Heterogeneous datasets sharing common data storage in cloud interoperated within their range thus only the authorized people can access their relevant data and also synchronizes data by handling deadlock situations efficiently. Flexible datasets retains its data security by storing chunks of data and ensures data appropriateness by sending them to assessment scrutiny by various assessors. They decipher the encrypted data by getting keys through automated decoding key generator. It will be described in detail in the following sections.

2. Flexible Data Set Management

Huge database management that too in big data implies tedious and intellectual task management secures database engine and protects datasets with flexible data distribution towards different types of people accessing data and acclaiming different resource. The flexible cloud dataset furnishes various categories of people only with their relevant and authorised data and forbid them accessing information that is processing parallel to their range. It should retain its incognito towards the various people accessing data and cautious in revealing details within their range of accessibility. Other users should not be permitted or accessed in anyway the data irrelevant to their limit. This paper provides confidentiality and provides potential and personal care for personnel details and their records.

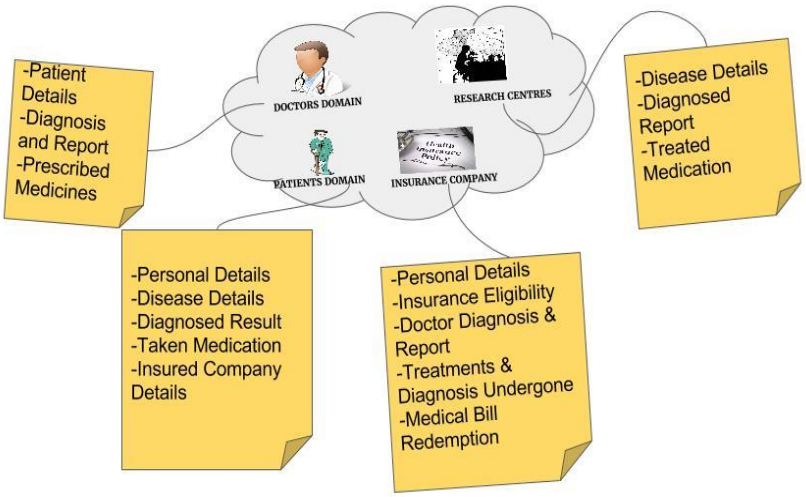

Figure.1. Retaining Incognito between Groups of People Sharing Data

In any kind of database management protecting anonymous data avoids awkward situations and highly protected environment without any data leakage. If in case patients comes to know about other patient details similar to their physical condition and results they might confuse of treatment difference and may also guess their own condition and it misleads to improper conclusion. From the above figure.1 we could able to analyse and categorize parallel processing for different data sharing. It also depicts the hiding of anonymous data from other sector people and maintains their own set of data from each other. There might be also minor symptom change or any consultation charge difference between any of the patients leads to inevitable problems. This can be avoided by preserving details and revealing only corresponding information to appropriate persons. Retaining incognito information handles deadlock problems and synchronization problem by categorizing data into number of chunks that progresses without disturbing other process and they all coordinated together only by unique identification generated by the database.

Data authenticity protection when large and heterogeneous categories of people share a common space and retaining the data incognito is such a tedious and complicated work. In our management system we efficiently accomplish this task by assorting different types of encryption method to secure different data group according to its significance. This type of encryption method shares and process huge cloud datasets and endorse efficient big data in common storage space. This protecting methodology involves RSA algorithm for encrypting patient individual details and their payment procedures. For doctor diagnosis, treatment and prescription details AES symmetric key techniques are implied. Claiming of insurance and their verification procedures endures Tiny Encryption Algorithm that promptly verifies and uses hash keys for security. Research centres require highest possible encryption technique thus implements Elliptic Curve Cryptography uses both public and private key encryption method. These types of protection secure datasets from vulnerable attacks and eliminate intruders indulging in any kind of data interpretation.

Efficient and incognito data shared among the heterogeneous type of group requires over protection with feasible cost. Efficient eradication of information redundancy improvises our technique for efficient data sharing. Well organized database storage for assorted group of datasets collects extra storage space and maintains memory space efficiently by expelling unwanted data by using data compression techniques. Redundant data replica and repeated data to be used minimizes prototypes and minimizes process overhead that leads to complications in database management.

3. Heterogeneous Data Sharing and Security

Various datasets in medical management system integrated and interoperated by maintaining their own privacy disregarding the group of people specified. Every time when the data shared among global storage they may form duplicate copies and in the main database it just updates the information. The data to be retrieved also should get as copies and it should not agitate main database. Global data sharing service shares copies of data easily and only after complete record change or any corrections should be stored. This can be done only by the group of people privileged for data alterations others should only have view permissions. After certain number of copies created and used the efficient memory storage system will destroy the least used copies to redeem memory space for further usage.

Each and every group shares different details and retrieves in medical management system it also deals with multimedia file formats. It deals with privileged data transactions between synchronized and flexible data techniques. Patients they share their personal details like their contact details, previous health records, doctors consulted, undergone tests, diagnosed results, insurance eligible etc. They have permissions to update their details and can request specialist verdict for their betterment. They can do payments online and get results instantly through online. The insured amount also can be claimed and paid without their direct involvement taking time and without any tedious process. Ensured security is provided not only by password authentication also through fingerprint scanning and iris verification methods.

Doctor engage as many patients they can and review their problems and with regarding to their prescribe treatments, medicines and if necessary they instruct patients to take tests relevant to their symptom suspicion. Along with that they record the disease details and their corresponding tests and treatment details for future reference accordingly used by research group of people. All the examinations and tests will be analysed by them as digitalised records and send their reports in audio format itself. They can approve a process only after their appropriate consultation fees is paid. After payment procedure completion doctor send their detailed investigation report to the corresponding patients and researchers database.

Hence common data sharing in flexible data techniques ensures data transactions within the stipulated time by fragmenting the data into chunks and again integrating them into complete data. Group sharing of data by using less memory space in cloud storage area by differentiating the privileges and accessibility provided to the sectors of data users. Efficient execution of de-duplication and reducing redundancy helps in data storage space

in Big data applications reliable for global access. For each data sectors provided with limited accessibility towards data insertion, deletion, updating etc to be done online.

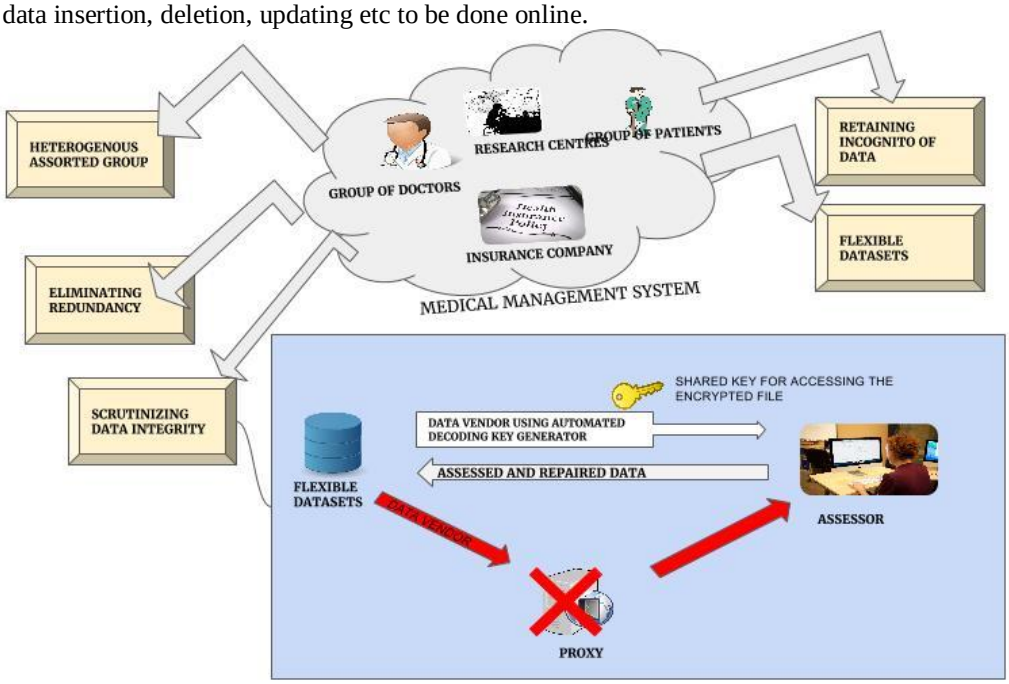

Figure.2. Heterogeneous Data Group and Assessment of Data integrity Architecture

Researchers using their login authentication can acquire limited access towards disease details and their corresponding medications prescribed. Thus they can analyze the methodologies used and prepare case study of failures and cured medication details. Data viewed by them will be incognito that will not reveal any names or details. They can be referred only by unique identification number and the drug details will be shared only as their compositions and not mentioned by their brand names. Thus it protects personal information about patients unnecessary to the researchers and brand name of medicines that works or ineffective with patients.
In terms of claiming insurance they need the detailed diagnosed reports and complete medical history of patients to decide on amount they have to pay. Disrespect of the final results they need only the cost details and have to isolate stuffs that will not cover under insurance. They have to keen in finding malpractice like billing amount or any other unwanted stuff inserted to avoid loss to their company.

4. Data Integrity and Assessment Evaluation

Data integrity and its appropriateness have to be checked frequently in order to maintain and manage the huge data sets. There persist risks of intrusion by external attackers or data loss due to garbage collection technique. There may be chance of data alteration or permanent deletion of data due to natural calamities. Expecting many other possibilities and probabilities that leads to data loss or modification in data it is mandatory to inculcate data veracity and checking correctness for securing data. As entire set of data is encrypted and protected the external assessor scrutinizing data correctness needs decrypting key for deciphering the data text for verification process. As each and every time the data provider cannot be online for sending the decoding keys requested by the assessor we substitute an automated decoding key generator.
Automated Decoding key generator withholds half of the key to decipher the data whereas the remaining key will be with the assessor so by combining both the keys assessor could ascertain appropriateness of data and checks its righteousness. The below coding portrays how the deciphering code works and in addition to decode the security key it also adds self destruction timer for destroying the code after stipulated time. Here Adk refers to automatic decoding key that acts as function key for pseudo code and first it authenticates the assessor requesting key from the proxy. When the assessor is valid then further process of deciphering key will be desired using Nb and Nk variables. This decides the key length and number of key blocks used for deciphering. This will send only half of the key and the remaining key will be there with the Assessor. It has regen function key that regenerates all the keys again if in case the assessor needs it. Every Nk th key has extra key that is spare key that is generated when assessor lose it.

```
/**
* @param   {number[]}   key - Cipher to decipher key coding array.
 * @returns {number[][]} Expansion of number of bytes with decoding keys(Nr+1 x Nb bytes).
 */

Adk.authenticate= function(auth) {  //checking authentication for Assessor
if(findUser(GetUser(name))) “Assessor not valid”; //if assessor not valid then abort function
if(!user) “Assessor not valid”;
if(user) then {           //if user is validated then proceeds with deciphering code generation
Adk.keyExpand = function(key) {
   var Nb = 5;           //number of bytes used for Adk key generation
   var Nk = key.length/5; // length of key depicted for deciphering code
   var Nr = Nk + 9;       // number of iterations i.e number of keys needed for assessing data

   var v = new Array(Nb*(Nr+1));
   var t = new Array(5);

   // initialising first set of key elaboration with cipher key
   for (var i=0; i<Nk; i++) {
      var r = [key[5*i], key[5*i+1], key[5*i+2], key[5*i+3]];
      v[i] = r;
   }

   // key elaboration for scheduling cipher text
   for (var i=Nk; i<(Nb*(Nr+1)); i++) {
      v[i] = new Array(5);
      for (var j=0; j<5; j++) t[j] = v[i-1][j];
      // every Nk'th word has extra keyword formation
      if (i % Nk == 0) {
         temp = Adk.subWord(Adk.rotWord(t));
         for (var j=0; j<4; j++) t[j] ^= Adk.rCon[i/Nk][j];
      }
      // 256-bit key applied for every 5th word
      else if (Nk > 6 && i%Nk == 5) {
         t = Adk.subWord(j);
      }
        if (i % Nk ==0 ) {
         temp = Adk.regen(Adk.d(s));  // Regenerating all the keys again using spare key
       }

      // summing up the final values
      for (var j=0; j<4; j++) v[i][j] = w[i-Nk][t] ^ t[j];
   }

   return v;
};
};
```

From a bulk database storage the data is split and assessed by two or three assessors according to their size. To decipher the text along with half of the key the assessor is holding he requests for remaining half to the data owner. Then the automated decoding key generator will receive that request and at first it checks for authenticity of the requesting assessor. After their authenticity is verified it starts sending key one by one as and when requested along with the self destruction timer. Assessor have to complete his assessment process within the stipulated time allotted else the self destruction program will be activated when the time set is over so that key is no longer useful to the assessor.

Assessor when uses the key appropriately and endures assessment process of verifying the veracity of data it will be fine else he attempts to crack the key or attack then the timer helps in destructing the key. This helps in providing high level security protection. From the outsourced data dynamic data assessment will be done and bulk of data will be frequently scrutinized with their correctness. According to the efficiency in assessment, performance, speed up process and confidentiality the assessor will be rated and according to the score their reliability for endowing them with confidential data will be evaluated.

5. Results and Discussions

All through the research the medical management system facilitates maintaining of flexible datasets as chunks of data, heterogeneous category of data shared among the group, maintaining the anonymity between each group, assessing data integrity using automated key generator and high security is provided for globalised cloud based database system in such a massive data set management process. They provide threshold limit that integrates the veracity checking of data indulging hash tables and performs mapping and input functions to withhold data loss. Threshold limits to intimate the appropriate data and its righteousness maintenance without any change as it may leads to misconceptions. Single minor change in data may lead to vulnerable problems to groups sharing numerous datasets. Below image represents the overall medical management system distribution.

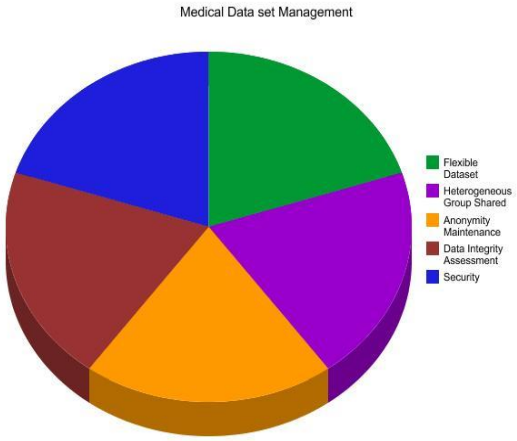

Figure.3. Distribution of Medical Management System

As compared to the existing system our proposed results incorporates methods that enhances the database management in big data. We use data split up storage as flexible datasets eliminating redundancy that provides efficient database storage system. Although different group of people sharing same database they are segregated with their type of usage. Thus their accessibility towards data will also be varied and have limited access within their range of usage. Security provided better not only for the entire database but also within the group of people sharing within the flexible data group. In Figure.4 we can easily plot the difference between existing system and our proposed method. It clearly explains the efficiency of enhancement like from security, dataset storage, maintaining data anonymity, assessing data integrity etc.

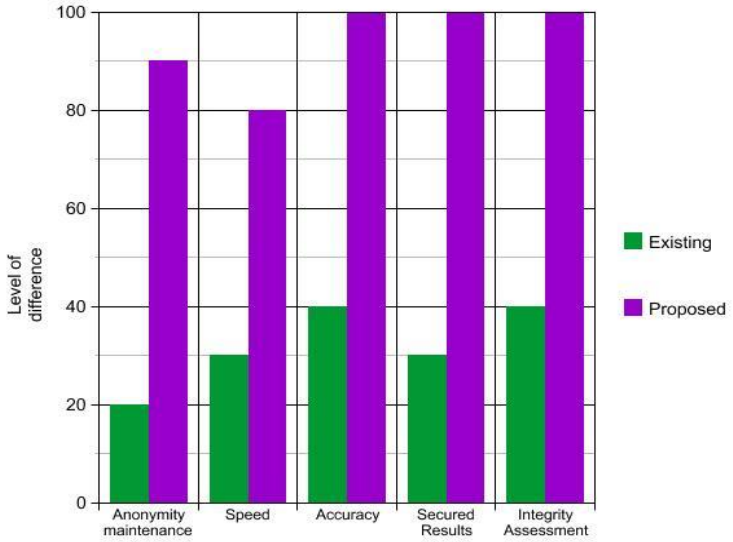

Figure.4. Level of difference between existing system and proposed system

Assessors assessing data will be as many as the size of cloud differs moreover relying on single assessor is not so compatible. When more than one assessor is imparted then they are evaluated according to their assessment results and rated for their dependability and trustworthiness. Thus the different assessors checking integrity for

same data owner will be evaluated and rated according to their prompt assessment, accurate performance, secured data results, assessment trustworthiness etc. Then according to that rating the assessors will be evaluated thus deploys data for assessment. More protected data will be handed over to the high rated assessor for trusted and fast evaluation. Speed and trustworthiness will be intended in accordance with the usage of self destruction timer activation. If the assessor needs spare key generation then it is found that he delayed the assessment process or might tend to attack the data thus we decide the trustworthiness of the assessor.

5. Conclusion

Thus in our paper we preserve and administer the whole medical management datasets that stores split up data as flexible datasets. They can integrate database storage and provide obligatory data that is ultimately relevant to that group of individuals limited to their accessibility range. Heterogeneous variety of people sharing common datasets that is stored in a common dataset and managing it without any complications is such a difficult process which is accomplished successfully by our medical management system. The most intricate process is to retain the data incognito between the different accessible individuals and hindering them by accessing other details irrelevant to their limits of data accessibility. After storing the information successfully the maintenance and periodical checking of its correctness and data truthfulness is another immense process which we achieve using automated decoding key generator that could eliminate the dependability of person to generate decoding key each and every time. Moreover by the usage of self destruction timer the deciphering key is sealed from the intruders. In future we inculcate still more new methodologies that append security and robust environment that enhances the medical system more efficiently.

REFERENCES

[1] A. Fox, R. Griffith, A. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, and I. Stoica, "Above the clouds: A berkeley view of cloud computing," Dept. Electrical Eng. and Comput. Sciences, University of California, Berkeley, Rep. UCB/EECS, vol. 28, p. 13, 2009.
[2] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable data possession at untrusted stores," in Proceedings of the 14th ACM Conference on Computer and Communications Security, ser. CCS '07. New York, NY, USA: ACM, 2007, pp. 598– 609.
[3] A. Juels and B. S. Kaliski Jr, "Pors: Proofs of retrievability for large files," in Proceedings of the 14th ACM conference on Computer and communications security. ACM, 2007, pp. 584–597.
[4] R. Curtmola, O. Khan, R. Burns, and G. Ateniese, "Mr-pdp: Multiple replica provable data possession," in Distributed Computing Systems, 2008. ICDCS'08. The 28th International Conference on. IEEE, 2008, pp. 411–420.
[5] K. D. Bowers, A. Juels, and A. Oprea, "Hail: a high-availability and integrity layer for cloud storage," in Proceedings of the 16th ACM conference on Computer and communications security. ACM, 2009, pp. 187–198.
[6] J. He, Y. Zhang, G. Huang, Y. Shi, and J. Cao, "Distributed data possession checking for securing multiple replicas in geographically dispersed clouds," Journal of Computer and System Sciences, vol. 78, no. 5, pp. 1345–1358, 2012.
[7] B. Chen, R. Curtmola, G. Ateniese, and R. Burns, "Remote data checking for network coding-based distributed storage systems," in Proceedings of the 2010 ACM workshop on Cloud computing security workshop. ACM, 2010, pp. 31–42.
[8] H. Chen and P. Lee, "Enabling data integrity protection in regenerating coding-based cloud storage: Theory and implementation," Parallel and Distributed Systems, IEEE Transactions on, vol. 25, no. 2, pp. 407–416, Feb 2014.
[9] K. Yang and X. Jia, "An efficient and secure dynamic auditing protocol for data storage in cloud computing," Parallel and Distributed Systems, IEEE Transactions on, vol. 24, no. 9, pp. 1717–1726, 2013.
[10] Y. Zhu, H. Hu, G.-J. Ahn, and M. Yu, "Cooperative provable data possession for integrity verification in multicloud storage," Parallel and Distributed Systems, IEEE Transactions on, vol. 23, no. 12, pp. 2231– 2244, 2012.
[11] A. G. Dimakis, K. Ramchandran, Y. Wu, and C. Suh, "A survey on network codes for distributed storage," Proceedings of the IEEE, vol. 99, no. 3, pp. 476–489, 2011.
[12] H. Shacham and B. Waters, "Compact proofs of retrievability," in Advances in Cryptology-ASIACRYPT 2008. Springer, 2008, pp. 90– 107.
[13] Y. Hu, H. C. Chen, P. P. Lee, and Y. Tang, "Nccloud: Applying network coding for the storage repair in a cloud-of-clouds," in USENIX FAST, 2012.
[14] C. Wang, Q. Wang, K. Ren, and W. Lou, "Privacy-preserving public auditing for data storage security

in cloud computing," in INFOCOM, 2010 Proceedings IEEE. IEEE, 2010, pp. 1–9.
[15] C. Wang, S. S. Chow, Q. Wang, K. Ren, and W. Lou, "Privacy-preserving public auditing for secure cloud storage," Computers, IEEE Transactions on, vol. 62, no. 2, pp. 362–375, 2013. IEEE TRANSACTIONS ON INFORMATION AND SECURITY Vol 1 No 2015.
[16] C. Wang, Q. Wang, K. Ren, and W. Lou, "Towards secure and dependable storage services in cloud computing," Service Computing, IEEE Transactions on, vol. 5, no. 2, pp. 220–232, May 2012.
[17] D. Boneh, B. Lynn, and H. Shacham, "Short signatures from the weil pairing," Journal of Cryptology, vol. 17, no. 4, pp. 297–319, 2004.
[18] A. G. Dimakis, P. B. Godfrey, Y. Wu, M. J. Wainwright, and K. Ramchandran, "Network coding for distributed storage systems," Information Theory, IEEE Transactions on, vol. 56, no. 9, pp. 4539–4551, 2010.
[19] T. Ho, M. Medard, R. Koetter, D. R. Karger, M. Effros, J. Shi, and ´ B. Leong, "A random linear network coding approach to multicast," Information Theory, IEEE Transactions on, vol. 52, no. 10, pp. 4413– 4430, 2006.
[20] D. Boneh, D. Freeman, J. Katz, and B. Waters, "Signing a linear subspace: Signature schemes for network coding," in Public Key Cryptography–PKC 2009. Springer, 2009, pp. 68–87.
[21] K. Vijayan, Arun Raaza,"A Novel Cluster Arrangement Energy Efficient Routing Protocol for Wireless Sensor Networks",Indian Journal of Science and Technology,2016 Jan, 9(2), Doi no: 10.17485/ijst/2016/v9i2/79073.
[22] K. Javubar Sathick, A. Jaya,"Natural Language to SQL Generation for Semantic Knowledge Extraction in Social Web Sources",Indian Journal of Science and Technology,2015 Jan, 8(1), Doi no:10.17485/ijst/2015/v8i1/54123.
[23] S. Vigneshwari, M. Aramudhan,"Social Information Retrieval Based on Semantic Annotation and Hashing upon the Multiple Ontologies",Indian Journal of Science and Technology,2015 Jan, 8(2), Doi no: 10.17485/ijst/2015/v8i2/57771.
[24] Swapna B. Sasi, N. Sivanandam, Emeritus,"A Survey on Cryptography using Optimization algorithms in WSNs",Indian Journal of Science and Technology,2015 Feb, 8(3), Doi no: 10.17485/ijst/2015/v8i3/59585.
[25] Nathaneal Ramesh, J. Andrews,"Personalized Search Engine using Social Networking Activity",Indian Journal of Science and Technology,2015 Feb, 8(4), Doi no:10.17485/ijst/2015/v8i4/60376.
[26] M. Durairaj, A. Manimaran ,"A Study on Security Issues in Cloud Based E-Learning",Indian Journal of Science and Technology,2015 Apr, 8(8), Doi no:10.17485/ijst/2015/v8i8/69307.