A Secured Public Auditing Protocol with Dynamic Structure for Cloud Data

K.Vidhya Lakshmi, S. Thanga Ramya

Abstract*: *At present Cloud computing is a very successful paradigm for data computing and storage. It Increases the concerns about data security and privacy in the cloud. Paper covers cloud security and privacy research, while focusing on the works that protect data confidentiality and privacy for sensitive data being stored and queried in the cloud. As Survey enlist all the research carried out related to data security and users privacy preserving techniques in detail. Data sharing can be achieved with sensitive information hiding with remote data integrity auditing, propose a new concept called identity based shared data integrity auditing with sensitive information hiding for secure cloud storage. Initially every data would be outsourced to the cloud only after authorized or activated by the proxy. The key would be generated to the file randomly by the key generation Centre. The transaction details such as key mismatch, file upload and download, hacking details would be shown to the proxy and cloud server. If the match occurs, automatically file would be recovered by the user even if hacker access or tamper the file. The main motive is to ensure that when the cloud properly stores the user's sanitized data, the proof it generates can pass the verification of the third party auditor. And the paper provides various research work done in the field

Keywords*: *Auditing, Privacy, Cloudsecurity.

I. INTRODUCTION

Cloud computing has emerged as a successful paradigm that considerably simplifies the deployment of computing and storage infrastructures of both large and small enterprises. Major enabling features of the cloud computing infrastructure include pay per use and hence no up-front cost for deployment, perception of infinite scalability, and elasticity of the resources. The Software as a Service (SaaS) paradigm, such as web-based emails and online financial management, has been widely adopted for almost a decade. The launch of Amazon Web Services (AWS) in the second half of 2006, followed by a plethora of similar offerings such as Google AppEngine, Microsoft Azure, etc., have popularized the model of "utility computing" for other levels of the computing substrates such as Infrastructure as a Service (IaaS) and Platform as a Service (PaaS) models. The widespread popularity of these models is evident from the tens of cloud based solution providers [1] and hundreds of corporations hosting their critical business infrastructures and hence business data in the cloud [2].

Revised Manuscript Received on November 27, 2019.
*** Correspondence Author**
Ms. K.vidhya lakshmi, IT,Veltech Mutitech, Dr.Rangarajan and Dr.Sakunthala Engineering College, Chennai, India. Email: vidhyamecse@gmail.com
Dr. S. Thanga Ramya, IT, RMD Engineering College, Chennai, India. Email: str.it@rmd.ac.in

Concerns about data security and privacy in the cloud, however, are increasing, as vulnerabilities were found in cloud service provider sites [2], [3], and user data leakage incidents were reported for a number of cloud based application services [4], [5]. Users cannot control and audit their sensitive data stored in the cloud by themselves.

A. Security and Privacy Threats

First by presenting a general overview of various security and privacy threats that could arise in the context of data oriented services, and specifically data oriented services deployed in the cloud. consider cloud service providers and any unauthorized parties that can monitor and control data and activities in the cloud as adversaries. assume the adversaries are honest but curious. identify the desirable features for ensuring a secure and privacy-preserving database service in the cloud.

B. Data Confidentiality

Schemes and techniques for ensuring data confidentiality while still allowing data management and query processing on the protected data in the cloud are to be known . To protect the confidentiality of sensitive private data stored in the cloud, encryption is a widely accepted standard technique. Different encryption schemes can be used on different granularities of the relational data. Encrypting the data, however, makes it difficult for the cloud to process queries on the data on behalf of the users, thus various techniques have been proposed for querying encrypted data. These techniques make different trade-offs between two conflicting goals, data confidentiality and query efficiency. The perfect state of strong confidentiality and high efficiency is unlikely to be achievable under today's encryption schemes. Alternatively, have exploring trusted computing instead of encryption and querying on encrypted data.

(i) Querying Encrypted Data

considers encrypted data in the cloud and the data should not be disclosed to adversaries during query processing, while the adversaries could launch statistical analysis and inference attacks to infer the data contents. Start with the early work of querying encrypted data using keyword search on encrypted texts [7], and then focus on providing a survey of the techniques for processing various database queries such as range queries and aggregation queries on encrypted data. Processing range queries requires the ability to compare a ciphertext data value with the encrypted query range boundary values. This can be achieved by providing the cloud provider with some rough information about the data [8]–[11] or building obfuscated index structures in the cloud [12][15]. Addressing how each of the techniques achieves the ability of ciphertext comparison, and discuss their advantages and limitations in terms of preserving data confidentiality, query efficiency, the overheads associated with the cloud and the roles of the cloud and the query client in the query processing. Processing aggregation

queries can be achieved using a special encryption scheme called homomorphic encryption that allows addition and multiplication on ciphertexts without the need for decryption [11], [16], [17]. About these computation properties of homomorphic encryption and further delve into a recent hot topic about fully homomorphic encryption [18]–[20]. Continuing the discussion about fully homomorphic encryption, exploration of some open questions and challenges: Is there a single technique or system that can process all the common database queries on encrypted data [11], [15], [19]?

(ii) Trusted Computing

An alternative to data encryption and querying on encrypted data is to store the encrypted data in the cloud but decrypt and process the plaintext data in a secure trusted container in the cloud [21]–[24]. present this idea of trusted computing as the last part of data confidentiality. Trying to address the following questions: What is trusted computing? What is a trusted hardware? Why and how can it be trusted by users? Finally compare the previous approaches, i.e. data encryption and querying encrypted data, with trusted computing from the three angles of security, performance and database functionalities.

C. Access Privacy

Ensuring data confidentiality is not enough to safeguard data in the cloud. When data is being queried, queries may reveal partial information about the data. Even if both the data and the queries are encrypted, partial information about the data may still be inferred by monitoring users' query access patterns and by analyzing users' accessed positions on the encrypted data. Hence, ensuring access privacy is also needed. The most representative cryptographic protocol for protecting access privacy in general, Private Information Retrieval [25], a special memory structure that obfuscates query access patterns over encrypted data, Oblivious RAM [26], and practical alternative techniques for protecting access privacy. For completeness, the data to which accesses are protected can be plaintext private, public data or ciphertext private data.

Private Information Retrieval (PIR) solves the problem of privately retrieving a data item from a remote database server without revealing to the server which item is retrieved [25]. There are PIR solutions that only use one server [27] and solutions that rely on multiple servers [25]. As earlier single server PIR solutions have been criticized for incurring expensive computation costs and being impractical [28], a recent single server PIR construction was proposed as a fast PIR solution [29]. The discussion of whether PIR is practical is on-going [30]. Goal here is to understand the basic rational of PIR protocols and clarify whether these PIR proposals are practical in terms of computation and communication costs as well as from the cloud service point of view.

Oblivious RAM: One way to make PIR more practical, as proposed in [31], [32], is to employ an oblivious RAM [26] on the cloud server. The basic idea of oblivious RAM is to shuffle and re-sort data items in the RAM during data accesses. Implementing this idea and making it practical is not trivial. Illustration of a recent practical implementation [35] to end the discussion on oblivious RAM.

Practical Alternative Techniques: PIR and oblivious RAM provide strong access privacy, but from a query processing point of view presentation of significant performance overheads. Therefore, some alternative techniques, covered search and index shuffling for protecting accesses to encrypted index [34], and hybrid approaches that apply PIR like cryptographic operations on selected partial data [35]–[37].

II EXISTING AND RELATED WORK

More organizations and individuals would like to store their data in the cloud. However, the data stored in the cloud might be corrupted or lost due to the inevitable software bugs, hardware faults and human errors in the cloud. The sensitive information should not be exposed to others when the cloud file is shared.

Encrypting the whole shared file can realize the sensitive information hiding, but will make the shared file unable to be used by others. If the file has been hacked by hacker it only recovers the file.

In order to verify the integrity of the data stored in the cloud, many remote data integrity auditing schemes have been proposed. To reduce the computation burden on the user side, a Third Party Auditor (TPA) is introduced to periodically verify the integrity of the cloud data on behalf of user. Ateniese et al. [35] firstly proposed a notion of Provable Data Possession (PDP) to ensure the data possession on the untrusted cloud.

In their proposed scheme, homomorphic authenticators and random sampling strategies are used to achieve blockless verification and reduce I/O costs. Juels and Kaliski defined a model named as Proof of Retrievability (PoR) and proposed a practical scheme. In the scheme, the data stored in the cloud can be retrieved and the integrity of these data can be ensured. Based on pseudorandom function and BLS signature, Shacham and Waters proposed a private remote data integrity auditing scheme and a public remote data integrity auditing scheme.

In order to protect the data privacy, Wang et al. proposed a privacy-preserving remote data integrity auditing scheme with the employment of a random masking technique.Solomon et al. utilized a different random masking technique to further construct a remote data integrity auditing scheme supporting data privacy protection. This scheme achieves better efficiency compared with the scheme in . To reduce the burden of signature generation on the user side, Guan et al.designed a remote data integrity auditing scheme based on the indistinguishability obfuscation technique. Shen et al. introduced a Third Party Medium (TPM) to design a light-weight remote data integrity auditing scheme. TPM helps user generate signatures on the condition that data privacy can be protected. In order to support data dynamics, Ateniese et al. [34] firstly proposed a partially dynamic PDP scheme. Erway et al. used a skip list to construct a fully data dynamic auditing scheme. Wang et al. proposed another remote data integrity checking scheme supporting full data dynamics by utilizing Merkle Hash Tree.

To reduce users' key exposure, Yu et al. proposed key-exposure resilient remote data integrity auditing schemes based on key update technique The data sharing is an important application in cloud storage scenarios. To protect the identity privacy of user, Wang et al. designed a privacy-preserving shared data integrity auditing scheme by modifying the ring signature for secure cloud storage. Yang et al. constructed an efficient shared data integrity auditing scheme, which not only supports the identity privacy but only achieves the identity traceability of users. Fu et al. designed a scheme by exploiting a homomorphic verifiable group signature. In order to support efficient user revocation, Wang et al. proposed a shared data integrity auditing

scheme with user revocation by using the proxy resignature. With the employment of the Shamir secret sharing technique, Luo et al. constructed a shared data integrity auditing scheme supporting user revocation. The aforementioned schemes all rely on Public Key Infrastructure (PKI), which incurs the considerable overheads from the complicated certificate management. To simplify certificate management, Wang et al. proposed an identity-based remote data integrity auditing scheme in multicloud storage. The scheme used the user's identity information such as user's name or e-mail address to replace the public key. Wang et al. designed a proxy-oriented remote data integrity auditing scheme based on identity introducing a proxy to process data for users. Yu et al. constructed a remote data integrity auditing scheme with perfect data privacy preserving in identity-based cryptosystems. Wang et al. proposed an identity-based data integrity auditing scheme satisfying unconditional anonymity and incentive. Zhang et al. proposed an identity-based remote data integrity auditing scheme for shared data supporting real efficient user revocation. Other aspects, such as privacy-preserving authenticators and data deduplication in remote data integrity auditing have also been explored. However, existing remote auditing data integrity schemes cannot support data sharing with sensitive information hiding. How to achieve data sharing with sensitive information hiding in identity-based integrity auditing for secure cloud storage.

III PROPOSED SYSTEM

Remote data integrity auditing is future to guarantee the integrity of the data stored in the cloud. In some common cloud storage systems such as the Electronic Health Records system, the cloud file strength contains some sensitive information.

Exploring how to achieve data sharing with sensitive information hiding in identity-based integrity auditing for secure cloud storage. A sanitizer is used to sanitize the data blocks corresponding to the sensitive information of the file and transforms these data blocks' signatures into valid ones for the sanitized file.

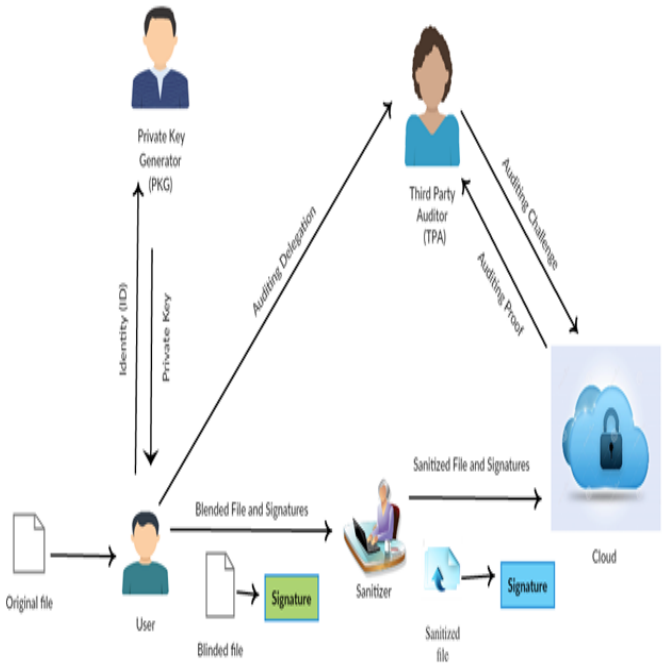


Figure 1: Architecture diagram

In the phase of integrity auditing, signatures are used to verify the integrity of the sanitized file. As a result, the file stored in the cloud able to be shared and used by others on the condition that the sensitive information is hidden, while the data integrity auditing is still able to be efficiently executed in remote.

IV METHODOLOGY

investigate how to achieve data sharing with sensitive information hiding in remote data integrity auditing, and propose a new concept called identity-based shared data integrity auditing with sensitive information hiding for secure cloud storage.The sanitizer can be viewed as the administrator of the information system in a hospital. The personal sensitive information should not be exposed to the sanitary.

To preserve the privacy of patient from the sanitizer, the medical doctor will blind the patient's sensitive information of each record before sending this record to the sanitizer. The medical doctor then generates signatures for this blinded record and sends them to the sanitizer. The sanitizer stores these messages into record information system.

V MODULE DESCRIPTION

A.FILE UPLOADING AND ACTIVATION

The data owner firstly needs to generate signatures for data blocks before uploading them to the cloud. These signatures are used to prove the cloud truly possesses these data blocks in the phase of integrity auditing. And then the data owner uploads these data blocks along with their corresponding signatures to the cloud. The data stored in the cloud is often shared across multiple users in many cloud storage applications. The data owner activate the file to check whether the uploaded file is appropriate or not then the Proxy also activate the file to check the file is Good.

B.DATA INTEGRITY AUDITING

Data integrity auditing scheme that realizes data sharing with sensitive information hiding. However, the data stored in the cloud might be corrupted or lost. Data integrity auditing on the condition that the sensitive information of shared data is protected.

C.SENSITIVE INFORMATION SHARING

Sensitive information hiding to ensure that the personal sensitive information of the file is not exposed to the hacker, and all of the private information of the file is not to be known to the cloud and the shared users. This method not only realizes the remote data integrity auditing, but also supports the data sharing on the condition that sensitive information is protected in cloud storage.

D.GENERATING KEY SIGNATURE

A potential method of solving this problem is to encrypt the whole shared file before sending it to the cloud, and then generate the signatures used to verify the integrity of this encrypted file, finally upload this encrypted file and its corresponding signatures to the cloud. This method can realize the sensitive information hiding since only the data owner can decrypt this file. However, it will make the whole shared file unable to be used by others.

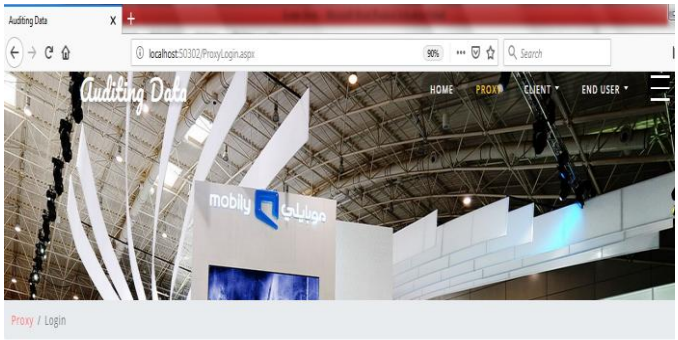
E.FILE SECURITY AND RECOVERY

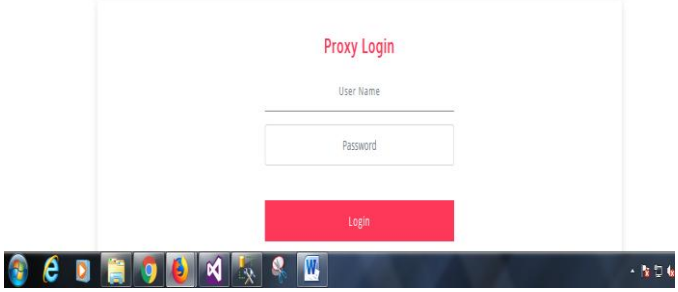
If a file has been partially overwritten or otherwise compromised, the chances of any usable recovery are low, even with the best recovery software in the existing system. In our proposed work, user can easily recover the file while deleted files are inaccessible and are in danger of being overwritten, they can often be recovered**.**

VI RESULTS AND VALIDATION

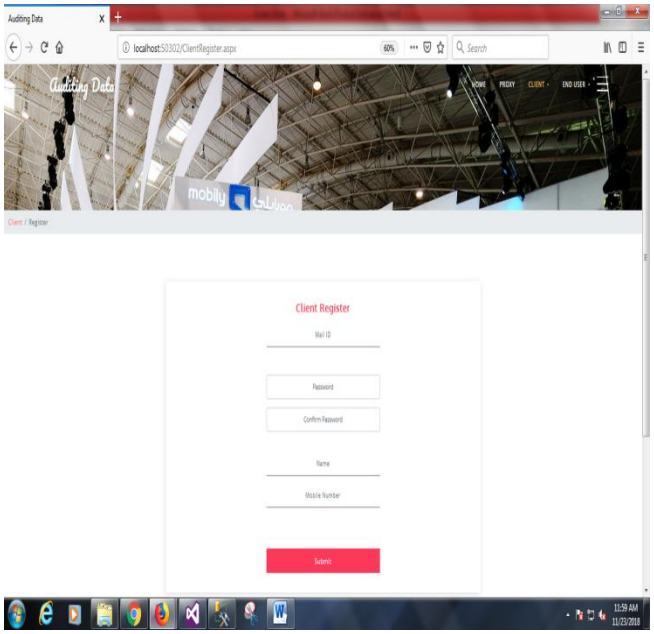
HOME:

PROXY LOGIN:





CLIENT REGISTER:



VII RESULT AND CONCLUSION

Proposed a character based information respectability reviewing plan for secure distributed storage, which bolsters information offering to delicate data covering up. In our plan, the record put away in the cloud can be shared and utilized by others depending on the prerequisite that the touchy data of the document is ensured.

Moreover, the remote information honesty examining is still ready to be proficiently executed. The security evidence and the exploratory investigation exhibit that the proposed plot accomplishes attractive security and productivity.

FUTURE WORK

i) All possible various encryption algorithms which best suitable for cloud computing environment can be implemented

ii) Encryption and decryption option may be given to different user based on sensitivity level of data their sharing

iii) Feature of encrypting the key added

iii) Authentication can be added

REFERENCES

[1] "AWS Solution Providers," http://aws.amazon.com/solutions/solution-providers/, 2009.

[2] T. Ristenpart, E. Tromer, H. Shacham, and S. Savage, "Hey, you, get off of my cloud: exploring information leakage in third-party compute clouds," in ACM Conference on Computer and Communications Security, 2009, pp. 199–212.

[3] J. Somorovsky, M. Heiderich, M. Jensen, J. Schwenk, N. Gruschka, and L. Lo Iacono, "All your clouds are belong to us: security analysis of cloud management interfaces," in Proceedings of the 3rd ACM workshop on Cloud computing security workshop, ser. CCSW '11. New York, NY, USA: ACM, 2011, pp. 3–14.

[4] D. McCullagh, "Privacy leaks hit facebook, google, at&t,"http://news.cnet.com/2702-1009 3-986.html, 2010.

[5] M. J. Schwartz, "Twitter finalizes ftc security settlement," http://www.informationweek.com/news/security/attacks/229301037, 2011.
[6] D. X. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in SP '00: Proceedings of the 2000 IEEE Symposium on Security and Privacy, 2000, pp. 44–55.
[7] H. Hacigumus, B. R. Iyer, C. Li, and S. Mehrotra, "ExecutingSQL over encrypted data in the database service provider model,"in Proceedings of the ACM SIGMOD international conferenceon Management of data, 2002, pp. 216–227. [Online]. Available:citeseer.ist.psu.edu/hacigumus02executing.html 1270
[9] B. Hore, S. Mehrotra, and G. Tsudik, "A privacy-preserving index for range queries," in Proc. of the 30th Int'l Conference on Very Large Databases VLDB, 2004, pp. 720–731.
[10] R. Agrawal, J. Kiernan, R. Srikant, and Y. Xu, "Order preserving encryption for numeric data," in SIGMOD '04: Proceedings of the 2004 ACM SIGMOD international conference on Management of data, 2004, pp. 563–574.
[11] R. A. Popa, C. M. S. Redfield, N. Zeldovich, and H. Balakrishnan, "CryptDB: protecting confidentiality with encrypted query processing," in Proceedings of the Twenty-Third ACM Symposium on Operating Systems Principles, ser. SOSP '11, 2011, pp. 85–100.
[12] E. Damiani, S. D. C. di Vimercati, S. Jajodia, S. Paraboschi, and P. Samarati, "Balancing confidentiality and efficiency in untrusted relational dbmss," in ACM Conference on Computer and Communications Security, 2003, pp. 93–102.
.[15] S. Wang, D. Agrawal, and A. El Abbadi, "A comprehensive framework for secure query processing on relational data in the cloud," in Proceedings of the 8th VLDB international conference on Secure data management, ser. SDM'11. Berlin, Heidelberg: Springer-Verlag, 2011,pp. 52–69.
[16] P. Paillier, "Public-key cryptosystems based on composite degree residuosity classes," in Advances in Cryptology - EUROCRYPT '99, ser.Lecture Notes in Computer Science, vol. 1592. Springer Berlin /Heidelberg, 1999, pp. 223–238.
[17] T. Ge and S. B. Zdonik, "Answering aggregation queries in a secure system model," in Proceedings of the 33rd International Conference on Very Large Data Bases, 2007, pp. 519–530.
[18] C. Gentry, "Fully homomorphic encryption using ideal lattices," in STOC '09: Proceedings of the 41st annual ACM symposium on Theory of computing, 2009, pp. 169–178.
[19] "Computing arbitrary functions of encrypted data," Commun. ACM, vol. 53, pp. 97–105, 2010.
[20] M. Naehrig, K. Lauter, and V. Vaikuntanathan, "Can homomorphic encryption be practical?" in Proceedings of the 3rd ACM workshop on Cloud computing security workshop, ser. CCSW '11. New York, NY USA: ACM, 2011, pp. 113–124.
[21] S. Bajaj and R. Sion, "TrustedDB: a trusted hardware based database with privacy and data confidentiality," in Proceedings of the 2011 international conference on Management of data, ser. SIGMOD '11 2011, pp. 205–216.
[22] D. Song, E. Shi, I. Fischer, and U. Shankar, "Cloud data protection for the masses," IEEE Computer, vol. 45, no. 1, pp. 39–45, 2012.
[23] N. Santos, K. P. Gummadi, and R. Rodrigues, "Towards trusted cloud computing," in Proceedings of the 2009 conference on Hot topics in cloud computing, ser. HotCloud'09. Berkeley, CA, USA: USENIX Association, 2009.
[24] A. Arasu, S. Blanas, K. Eguro, R. Kaushik, D. Kossmann, R. Ramamurthy,and R. Venkatesan, "Orthogonal security with cipherbase," in CIDR.
[25] B. Chor, E. Kushilevitz, O. Goldreich, and M. Sudan, "Private information retrieval," J. ACM, vol. 45, no. 6, pp. 965–981, 1998.
[26] R. Ostrovsky, "Efficient computation on oblivious RAMs," in Proceedings of the twenty-second annual ACM symposium on Theory of computing, ser. STOC '90, 1990, pp. 514–523.
[27] E. Kushilevitz and R. Ostrovsky, "Replication is not needed: Single database, computationally-private information retrieval," in Proceedings of the Anuual Symposium on Foundations of Computer Science (FOCS), 1997, pp. 364–373.
[28] R. Sion and B. Carbunar, "On the computational practicality of private information retrieval," in Network and Distributed System Security Symposium, 2007.
[29] C. A. Melchor and P. Gaborit, "A fast private information retrieval protocol," in IEEE Internal Symposium on Information Theory, 2008, pp. 1848–1852.
[30] F. G. Olumofin and I. Goldberg, "Revisiting the computational practicality of private information retrieval," in Proceedings of the 15th international conference on Financial Cryptography and Data Security, 2011, pp. 158–172.
[31] P. Williams and R. Sion, "Usable private information retrieval," in Network and Distributed System Security Symposium, 2008.
[32] P. Williams, R. Sion, and B. Carbunar, "Building castles out of mud: practical access pattern privacy and correctness on untrusted storage," in ACM Conference on Computer and Communications Security, 2008, pp. 139–148.
[33] E. Stefanov, E. Shi, and D. Song, "Towards practical oblivious RAM," CoRR, vol. abs/1106.3652, 2011.
[34] S. De Capitani di Vimercati, S. Foresti, S. Paraboschi, G. Pelosi, and P. Samarati, "Efficient and private access to outsourced data," in Proc. of the 31st International Conference on Distributed Computing Systems (ICDCS 2011), June 2011, pp. 710–719.
[35] Ateniese .G, Burns .R, Curtmola .R, Herring .J, Kissner .L ,Peterson .Z, and Song .D, (2007)"Provable data possession at untrusted stores," in Proceedings of the 14th ACMConference on Computer and Communications Security,ser. CCS '07, pp. 598–609.

AUTHORS PROFILE

Ms. K.vidhyalakshmi completed her UG, B.Tech (IT) from Anna university , with First Class in 2007. She also completed her PG M.E (CSE) from Anna University, Chennai with first class.she is currently working as a Assistant Professor in the department of IT in Veltech Multitech Dr.Rangarajan Dr.Sakunthala Engineering College and has a teaching experience of 8 years and has handled both UG and PG programs.. She is currently pursuing her Ph.D (ICE) in Anna University, Chennai. Her Research interests include , Cloud Security, Distributed and Parallel Computing ,Cryptography and Network security,Machine learning,IOT and Big data analytics. She has attended many workshops & FDPs sponsored by AICTE related to her area of interest. She is the life member of ISTE.

Dr. S. Thanga Ramya, B.E, M.S(by Res), Ph.D, is an Associate Professor in the Department of Information Technology, since June 2008. She obtained her B.E., (CSE) from Dr.Sivanthi Aditanar College of Engineering and M.S by Research (ICE) from Anna University,Chennai. She has obtained her Ph.D in Information and Communication Engineering from Anna University, Chennai, in 2017. She has been in the teaching profession for the past 18 years and has handled both UG and PG programs. Her areas of interest include programming languages, database management and data mining. She has published 9 papers in various International Journals and Conferences. She has attended many workshops & FDPs sponsored by AICTE related to her area of interest. She has also published 4 books. She is the life member of ISTE.