IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 6, June 2017

TSDT: TPA Based Secure Data Transmission of Cloud Data Storage System

Sandhya Dwivedi[1], Juhi Kanungo[2]

Dept. of Computer Science & Engineering, AITR, Indore, M.P[1,2]

Abstract: Data integrity and storage efficiency are two important requirements for cloud storage. Cloud computing has emerged as a long-dreamt vision of the utility computing paradigm that provides reliable and resilient infrastructure for users to remotely store data and use on-demand applications and services. Currently, many individuals and organizations mitigate the burden of local data storage and reduce the maintenance cost by outsourcing data to the cloud. However, the outsourced data is not always trustworthy due to the loss of physical control and possession over the data. In this context a storage service provider is demonstrated who offers storage services to their clients additionally the intermediate service distributors are also demonstrated that re-distribute the service of primary service providers. In this research work we are presented a novel approach i.e. TPA based Secure Data Transmission (TSDT) for the cloud data security using successful data transmission. Due to this for securing the data on the cloud storage and unsecured network transmission the cryptographic solution is proposed. Our experimental results prove the effectiveness and efficiency of our mechanism when evaluating shared data integrity.

Index: Data Security, Encryption, Decryption, AES, Cloud Computing, Cryptography, TPA, Deduplication, MD5.

I. INTRODUCTION

Data integrity and storage efficiency are two important requirements for cloud storage. Outsourcing data to cloud service for storage becomes an important trend, which benefits in sparing efforts on heavy data maintenance and management. The outsourced cloud storage is not fully trustworthy, it raises security concerns on how to realize data deduplication in cloud while getting integrity auditing [1]. Nowadays, organizations produce a huge amount of sensitive data, such as personal information, financial data, and electronic health records. Consequently, the amount of digital data produced has increased correspondingly and often overwhelmed the data storage capacity of many organizations. The management of such a large amount of data in local storage system is difficult and incurs high expenses because of high-capacity storage systems needed and the expert personnel to manage them. Although the cost of storage hardware has tremendously decreased in recent years, about 75% of the total ownership cost is still attributed to management data storage [2].

A. Cloud Data Storage

In the last decade, the demand of outsourcing data is greatly increased. Data storage and high performance computation are the main needs which have to be fulfilled. These services are provided by many cloud computing service providers like Drop box, Google App Engine, Amazon Simple Storage Service (AmazonS3), etc. The advantage of storing data in cloud servers is that the data owners can reduce the overhead of buying extra strong servers and also avoid hiring of server management engineers. The technology used for internet based development is nothing but cloud computing. Cloud provider offers one of the most fundamental services that is data storage. Data encryption is a basic solution to maintain security of data and the encrypted data is uploaded into the cloud. Depending on the possibility to identify privacy and security users cannot join the cloud computing systems [3].

B. Features of Cloud Storage

Most of the services are free up to certain number of gigabytes, and storage also. All the cloud provider provides all the features to the end user like drag and drop, syncing files and folder in your desktop, mobile device and soon [4].

- Usability – the cloud provider are always usable at any time.
- Bandwidth – avoid of sending the files to individual instead of send a web link to the end user through email it.
- Accessibility – stored files and folder are accessible from any were in cloud platform
- Disaster Recovery – back up facility should be there in cloud storage which is helpful for the businesses. Back up the data should be more important in electronic medium. Backup files are remotely stored and access through internet connection
- Cost Savings – In cloud storage 3percent gigabyte are there to store data internally. Businesses and organizations reduce the cost by using cloud storage. Data are stored remotely on it.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 6, June 2017

C. Secure Cloud Storage Auditing

Outsourcing storage into the cloud is economically attractive for the cost and complexity of long-term large-scale data storage. At the same time, though, such a service is also eliminating data owners' ultimate control over the fate of their data, which data owners with high service-level requirements have traditionally anticipated. As owners no longer physically possess their cloud data, previous cryptographic primitives for the purpose of storage correctness protection cannot be adopted, due to their requirement of local data copy for the integrity verification. Besides, the large amount of cloud data and owner's constrained computing capabilities further makes the task of data correctness auditing in a cloud environment expensive and even formidable for individual cloud customers. Therefore, enabling public auditability [5] for cloud storage is of critical importance so that owners can resort to a specialized third party auditor (TPA) to audit cloud storage services and maintain strong storage correctness guarantee, while saving their own precious computing resources.

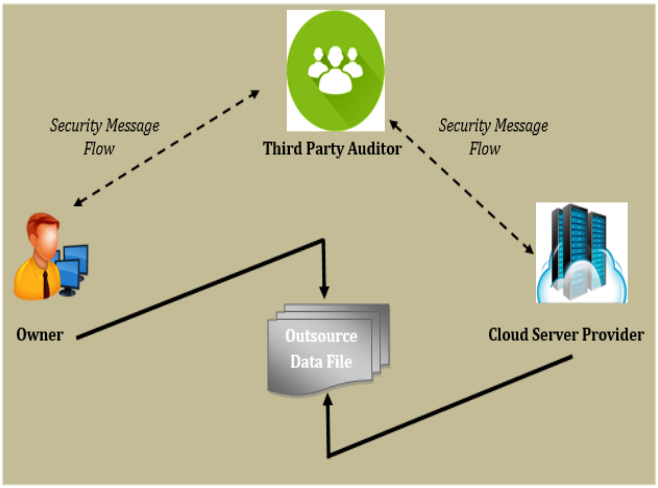

Figure 1: Cloud Auditing System

Considering TPA might learn unauthorized information through the auditing process, especially from owners' unencrypted cloud data, new privacy-preserving storage auditing solutions are further entailed in the cloud to eliminate such new data privacy vulnerabilities. Moreover, for practical service deployment, secure cloud storage auditing should maintain the same level of data correctness assurance even under the condition that data is dynamically changing and/or multiple auditing request are performed simultaneously for improved efficiency. Techniques we are investigating/developing for these research tasks include proof of storage, random-masking sampling, sequence-enforced Merkle Hash Tree, and their various extensions/novel combinations [6].

II. LITERATURE SURVEY

Cloud computing faces many problems on integrity and privacy of user's data stored in the cloud. Hence it requires some secure and efficient methods which can ensure the integrity and privacy of data stored in the cloud.

In this paper **Monir Azraoui et al. [7]** presents StealthGuard, an efficient and provably secure proof of retrievabillity (POR) scheme. StealthGuard makes use of a privacy preserving word search (WS) algorithm to search, as part of a POR query, for randomly-valued blocks called watchdogs that are inserted in the file before outsourcing. Thanks to the privacy-preserving features of the WS, neither the cloud provider nor a third party intruder can guess which watchdog is queried in each POR query. Similarly, the responses to POR queries are also obfuscated. Hence to answer correctly to every new set of POR queries, the cloud provider has to retain the file in its entirety. StealthGuard stands out from the earlier sentinelbased POR scheme proposed by Juels and Kaliski (JK), due to the use of WS and the support for an unlimited number of queries by StealthGuard. The paper also presents a formal security analysis of the protocol.

Yunhong gu et al. [8] proposed Sector which enables users to work with large datasets stored over multiple distributed nodes as if the files were on their local disk. Users do not need to locate data, manage data across multiple nodes, back up data, and manage the addition of new nodes or the deletion of existing nodes to the system.

Trust and security have prevented businesses from fully accepting cloud platforms. To protect clouds, providers must first secure virtualized datacenter resources, uphold user privacy, and preserve data integrity. **Kai Hwang et al. [9]** suggest using a trust-overlay network over multiple data centers to implement a reputation system for establishing trust between service providers and data owners. Data coloring and software watermarking techniques protect shared data objects and massively distributed software modules.

These techniques safeguard multi-way authentications, enable single sign-on in the cloud, and tighten access control for sensitive data in both public and private clouds.
Cloud computing enables highly scalable services to be easily consumed over the Internet on an as-needed basis. While cloud computing is expanding rapidly and used by many individuals and organizations internationally, data protection issues in the cloud have not been carefully addressed at current stage. Users' fear of confidential data (particularly financial and health data) leakage and loss of privacy in the cloud may become a significant barrier to the wide adoption of cloud services. In this paper, **Anna Squicciarini et al. [10]** explore a newly emerging problem of information leakage caused by indexing in the cloud. Authors design a three-tier data protection architecture to accommodate various levels of privacy concerns by users. According to the architecture, they develop a novel portable data binding technique to ensure strong enforcement of users' privacy requirements at server side.

III. PROPOSED WORK

This section provides the detailed understanding about the proposed technique for secure cloud storage in cryptographic aspects of cloud data in publicly dispersed environment. Therefore the detailed description of the benefits and the solution details are reported.

A. System Overview

Data security has consistently been a major issue in information technology. In the cloud computing environment, it becomes particularly serious because the data is located in different places even in the entire globe. Data security and privacy protection are the two main factors of user's concerns about the cloud technology. Though many techniques on the topics in cloud computing have been investigated in both academics and industries, data security and privacy protection are becoming more important for the future development of cloud computing technology in government, industry, and business. Data security and privacy protection issues are relevant to both hardware and software in the cloud architecture.

The invention of cloud computing offers the user to experience the efficient computing and scalable storage solutions. Therefore a number of individual users and organizations are preferred to use the cloud servers as the data hosting services. Cloud computing is an answer of new generation computational and storage requirements. Now in these days every application are designed in such manner by which more and more traffic is collected using the applications. In addition of that the reliable and long term services are also deployed using the SaaS (software as a service) concepts. The key reason behind development of SaaS, these applications are never compromises with the performance of applications additionally the cloud platform provides a secure and scalable storage solutions for the applications.
The proposed work is reports the investigation about the outsourced data and their sensitivity and security issues. In this we developed a mechanism for cloud security by means of TSDT i.e. TPA based Secure Data Transmission in public cloud environment. Additionally for providing end user trust and security management the upload, download and search space services are provided

B. Problem Domain

The cloud environment provides support for efficient computing and enables to provide the efficient computing and storage solutions at the remote end. In this presented work the main aim to address the following issues in the existing cloud storage:

1. **Data Security:** the data is placed on the cloud which is not much secured due to third party access and treads therefore the data security in cloud storage is required
2. **Data Owner and Client Privacy Management:** the data owner and client in not distinguishable using the data additionally the privacy on such data is access is required.
3. **Searchable Data Space:** the cryptographic manner of data security converts the formats and not a bit of data recovered during the information retrieval.

In this section the issues and challenges are provided in the next section the involved solution steps are reported.

C. Solution domain

In order to provide end to end solution for the cloud storage the following solution steps are included.

1. **Authentication Management:** In authentication management the system and user attributes are recovered additionally the one time password is included to manage the secure authentication.
2. **Cryptographic Data Security:** In this phase the MD5 and AES based hybrid cryptographic algorithm is consumed for providing the security.
3. **Providing the Search Solution over the Encrypted Data:** The keyword based search system is provided for identifying the user and their data during different data retrieval operations.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 6, June 2017

D. Methodology

For the advancement of the current security scenario here we present a proposed TSDT scheme for ensure security and privacy enhancement of the system using figure 2. According to the given diagram the proposed working model contains the three key roles for the system. First the client who wants to host their data or retrieve the data from the cloud storage, second the server where the data is stored in the cryptographic manner and finally the TPA (third party auditor). In first the client initiates the connection request to the cloud storage. In the next step as the request is arrived on the cloud server then the cloud server initiate the authentication process using the TPA.

In this phase the TPA verifies the end client using the random question answering. In the response the user submit the answers of the asked questions. In this phase the server (TPA) verifies the answers and also provides the acknowledgment to the user for authentication success or failure. If the user is verified through the server then the user can utilize the application for data hosting or recovery. During any process when the request arrived to the server. The server pass the message to the TPA and TPA is then handle the process of security. Therefore the hybrid cryptographic process is proposed that accepts the TPA generated keys for encryption and recovery. The server distributed key only works a single time and then the key is not effective for other sessions.

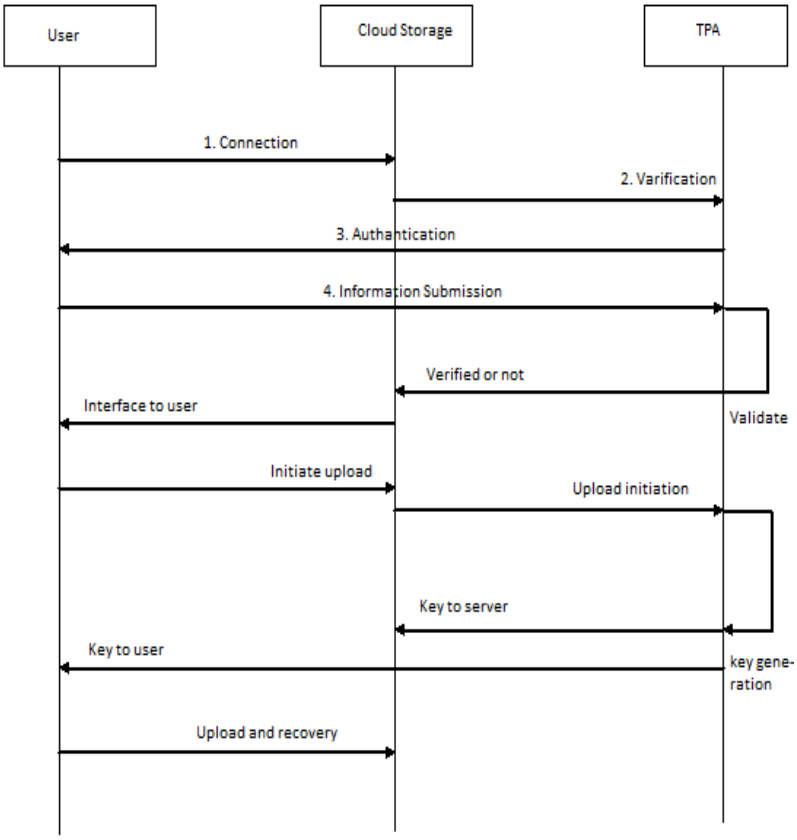

Figure 2 Proposed Communication Model

Therefore the proposed work involves two major contributions:

1. A secure cryptographic technique for cloud between client and server using TPA
2. Implementation of secure cryptographic data search

a. Proposed Cryptography

This section introduces the summarized steps of the proposed encryption technique proposed for securing the data during network file exchange in untrusted environment is given in three main parts.

1. Key generation: in this phase the authentication process of the user is performed, after authentication TPA server generates the key for utilizing in both the ends client and server.
2. Encryption: using the obtained key through the server the encryption is performed and transmitted to the server.
3. Decryption: during the data request of the user the same key is utilized and the data is recovered by the user.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 6, June 2017

Table 1 key generation

Input : User credentials U_s, Verification answers A_v Output: generated Key K_s
Process:
1. $if(Server.checkCredential(U_s) == true)$
a. $Q = TPA.AskQuery()$
b. $if(TPA.Verify(Q, A) == true)$
i. $K_s = Server.GenrateKey()$
c. $end\ if$
2. Else
a. Send error message
3. End if
4. Return K_s

Table 2 encryption process

Input: input file to upload F, TPA based key K_s Output: generated cipher text C_t
Process:
1. $H_s = MD5.genrateHash(F)$
2. $C_t = AES.encrypt(F, K_s)$
3. $SendtoTPA(H_s)$
4. $SendtoServer(C_t)$

Table 3 decryption process

Input: received cipher C_t, server key K_s Output: original file F
Process:
1. $F = AES.decrypt(C_t, K_s)$
2. $H_r = MD5.genrateHash(F)$
3. $SendtoTPA(H_r)$
4. $if(TPA.compare(H_r, H_s) == true)$
a. Receive file
5. Else
a. Reject file
6. End if

b. Implementation of secure cryptographic data search

In order to keep secure the data search and also the keywords by which a file can be identified a KNN based process is used. The proposed step process is defined in two different modules namely extraction and preservation of keywords from the uploaded file, and secondly the query processing.

Table 4 keyword extraction and preservation

Input: input file F Output: keyword H_k
Process:
1. $R = readInputFile(F)$
2. $P_d = preProcessData(R)$
3. $for\ (i = 0; i \leq P_d.length; i++)$
a. $T_f^i = \frac{number\ \ of\ times\ \ word\ \ appear\ ed}{total\ \ words}$
4. End for
5. $keywords[50] = selectTopfifty(T_f^i)$
6. $H_k = MD5.GenrateHash(keywords[50])$
7. D ← Preserve_to_database(H_k)

DOI10.17148/IJARCCE.2017.6688

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 6, June 2017

Table 5 keyword based search

Input: Database D, user query Q Output : query results R
Process: 1. $T[k] = tokenizer.createToken(Q)$ 2. $for(i = 0; i \leq k; i++)$ a. $H_i = MD5.genrateHash(T[i])$ 3. End for 4. $R = KNN.seach(D, H_i)$ 5. return R

IV. RESULT ANALYSIS

A. Precision

In any data retrieval or search applications the precision is a fraction of search results which is most relevant to the input data query. The provided precision of the proposed TPA based Secure Data Transmission are given using figure 5.1. This can be evaluated using the user feedback basis and can be evaluated by the following formula.

$$\text{Precision} = \frac{\text{Relevant Documnet} \cap \text{Retrieved Documnet}}{\text{Retrieved Documnet}}$$

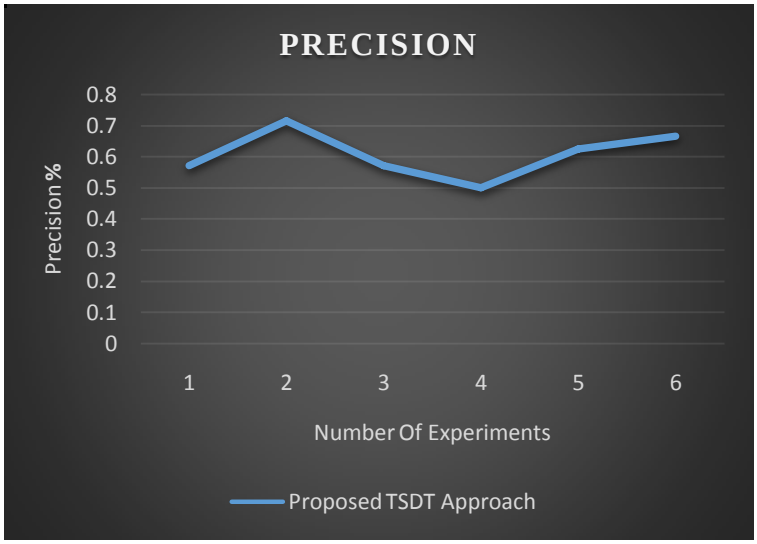

Figure 3 Precision Rate

The precision rate of the implemented system is described in the figure 3, the computed precision values are demonstrated using the Y axis of the given figure and the X axis shows the link data. According to the obtained results the performance of the proposed system is showing efficient result when user queries are retrieved. In addition of the precision rate is growing continuously as the similar kinds input data dependency.

B. Recall Rate

In data retrieval application or the search application recall values are measured for accuracy measurement in terms of relevant document retrieved or relevant data obtained according to the input user query. This can be evaluated using the following formula.

$$\text{Recall} = \frac{\text{Relevant Documnet} \cap \text{Retrieved Documnet}}{\text{Relevant Documnet}}$$

The figure 4 show the recall values of the proposed TSTD System for the secure communication. In order to represent the performance of the proposed work, the X axis contains the number of experiments to run the input user query and the Y axis reports the obtained recall rate of the implemented system. According to the obtained results the performance of the proposed system is enhances as if we increase the experiment length. Proposed concept is adoptable for the cloud storage security concern. Therefore the performance of the proposed system is much efficient for the publicly dispersed cloud environment.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 6, June 2017

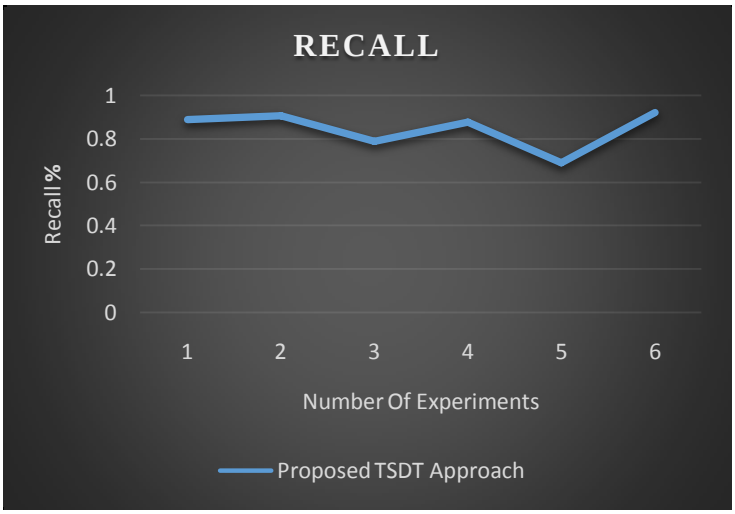

Figure 4 Recall Rate

C. F-Measures

The f-measures of the system demonstrate the fluctuation in the computed performance in terms of precision and recall rates. The f-measures of the system can be approximated using the following formula.

$$\text{F} - \text{Measures} = 2.\frac{\text{Precision X Recall}}{\text{Precision} + \text{Recall}}$$

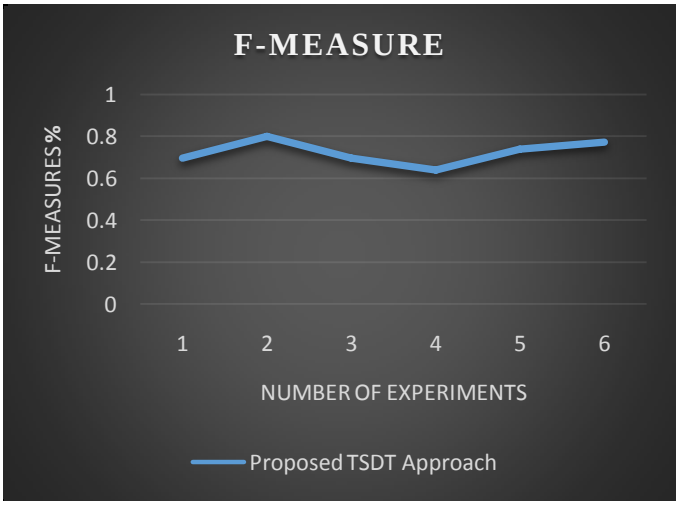

Figure 5 F-Measures

The figure 5 show the performance of proposed systems in terms of f-measures. To demonstrate the performance of the system the X axis shows the number of experiments and the Y axis shows the obtained performance in terms of f-measures. According to the obtained results the performance of the proposed system is much stable and provide ease of use of security. Data protection and user authentication is much require for the organization that can be retrieve data using cloud storage. In addition of that the results are in more progressive manner as if we increase number of experiments. Thus the obtained results are adoptable and efficient for secure communication of data transmission using cryptography technique.

D. Memory Consumption

The amount of main memory required to execute the algorithm with the input amount of data is known as the memory requirement of the system. The total memory consumption of the algorithm is computed using the following formula.

$$\text{Consumed Memory} = \text{Total Memory} - \text{Free Memory}$$

The figure 6 show the memory consumption of the system. In this diagram the amount of main memory consumed is given in Y axis and the number of experiments which are reported in X axis. According to the obtained results the proposed algorithm consumes fewer resources as we seen during the execution of algorithm.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 6, June 2017

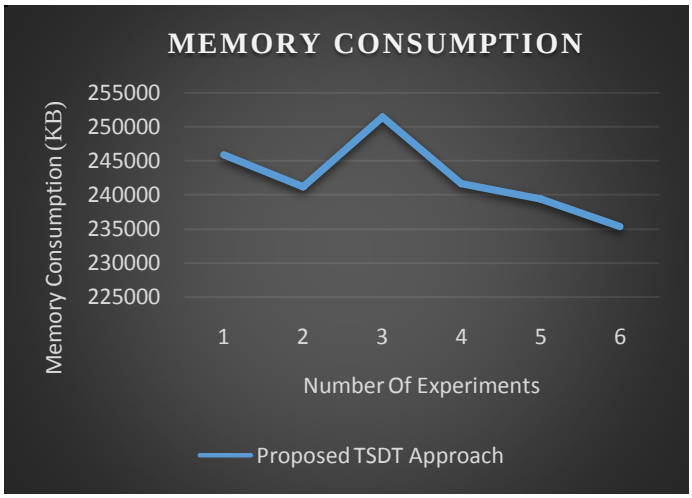

Figure 6 Memory Consumption

E. Server Response Time

The amount of time required to produce the outcome after making the request from the server is termed as the server response time. The response time not included the encryption or decryption activity during these measurements. The computed response time of the proposed technique for cloud based secure communication is demonstrated using the figure 7. X axis of this diagram contains the amount of experiments performed using the system and the Y axis shows the amount of time required for generating the response through the server. According to the computed results the response time is not depends on the amount of file size or other parameters. That is directly depends on the amount of work load on the target server where the data is stored or the application is hosted.

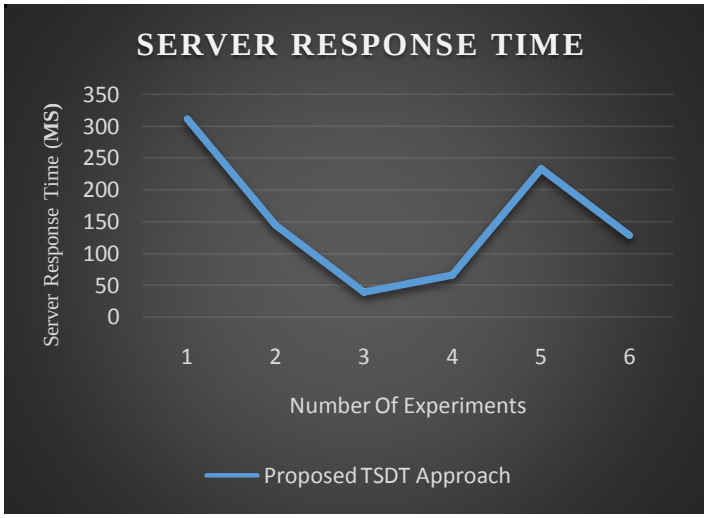

Figure 7 Response Time

V. CONCLUSION

Cloud computing is a promising and emerging technology for the next generation of IT applications. The barrier and hurdles toward the rapid growth of cloud computing are data security and privacy issues. Reducing data storage and processing cost is a mandatory requirement of any organization, while analysis of data and information is always the most important tasks in all the organizations for decision making. We believe that data storage security in Cloud Computing is an emerging computing paradigm, allows users to share resources and information from a pool of distributed computing as a service over Internet. Cloud storage is much more beneficial and advantageous than the earlier traditional storage systems especially in scalability, cost reduction, portability and functionality requirements. In this work we developed TPA based Secure Data Transmission i.e. **"TSDT"** is initiated in a dispersed environment where the single storage is secure way of authenticity and storage or hosting services. In addition of that for preventing the unauthorized access to the system a strong user authentication technique using the normal credential and OTP is prepared. Furthermore for securing the data in storage and untrusted network an AES and MD5 based cryptographic

technique is implemented. This technique also checks the communicated files integrity for finding the authenticity of the data transmission during the network file exchange.

REFERENCES

[1] Mr. Satish Shelar and Prof. S. Y. Raut, "Review On Deduplicating Data and Secure Auditing in Cloud", International Research Journal of Engineering and Technology (IRJET), Volume: 02 Issue: 09, Dec-2015
[2] M. Ali, S.U. Khan, A.V. Vasilakos, "Security in cloud computing: opportunities and challenges", Inf. Sci. 305 (2015) 357–383.
[3] Preeti Gulab Sonar and Pratibha Dattu Shinde, "A Novel Approach for Secure Group Sharing in Public Cloud Computing", International Journal of Computer Applications (IJCA), Volume 127 – No.11, October 2015
[4] Moritz Borgmann and Tobias Hahn, "On the Security of Cloud Storage Services", SIT Technical Reports, March 2012.
[5] Cong Wang, Sherman S.M. Chow, Qian Wang, Kui Ren, and Wenjing Lou, "Privacy-Preserving Public Auditing for Secure Cloud Storage," IEEE Transactions on Computers (TC), 2011, (INFOCOM'10).
[6] Cong Wang, Kui Ren, Wenjing Lou, and Jin Li, "Towards Publicly Auditable Secure Cloud Data Storage Services", IEEE Network Magazine, Vol. 24, No. 4, pp. 19-24, July/August 2010.
[7] Azraoui, Monir, et al. "Stealthguard: Proofs of retrievability with hidden watchdogs." European Symposium on Research in Computer Security. Springer International Publishing, 2014.
[8] Yunhong gu, Robert L. 2009. Grossman. Sector: high performance wide area community data storage and sharing system. Future Generation Computer Systems, 20 May 2009
[9] Hwang, Kai, and Deyi Li. "Trusted cloud computing with secure resources and data coloring." IEEE Internet Computing 14.5 (2010): 14-22.
[10] Squicciarini, Anna, Smitha Sundareswaran, and Dan Lin. "Preventing information leakage from indexing in the cloud." Cloud Computing (CLOUD), 2010 IEEE 3rd International Conference on. IEEE, 2010.