International Journal of Allied Practice, Research and Review
Website: www.ijaprr.com (ISSN 2350-1294)

Data Integrity Evaluation in Cloud Computing

Shailesh Ingale and Dr. Emmanuel M.
Information Technology Department, PICT, Pune
Information Technology Department, PICT, Pune, India

***Abstract* – The cloud platform has seen considerable amount of growth that has been in the storage as well as the cloud computation sector. The immense growth has been due to the versatility and the increase convenience and ease of use for the users of this platform. Which has been instrumental, lot of individuals and organizations adopt the cloud platform for their storage as well as computational needs. But the problem noticed was that the user had to give up the control of their data to the cloud service provider when transitioning to this platform. This is a very problematic scenario as there is a lack of trust between the public could service provider and the user. Therefore, most of the users decide to encrypt their data to safeguard its contents in the event of a breach or data leakage. But this introduces a problem of searching the encrypted data which makes it very difficult to retrieve. Therefore, this methodology implements an effective approach for performing search over encrypted data through the utilization of the Matrix Transition method along with Reverse Circle Cipher encryption approach and trap door generation. The technique is enhanced through the bilinear pairing and auto key generation to successfully safeguard the integrity of the data against attacks and preserve the privacy.**

Keywords— Cloud computing, Trap Door, Avalanche Effect, Bi-Linear pairing, Data Integrity.

I. INTRODUCTION

Data is one of the most integral aspects that have been instrumental in the growth of various technological advancements that have been performed in recent years. Data has evolved over the years where most of the data in the world used to consist only of books or other types of materials providing information in text format. Nowadays due to large advancements in Technologies and other approaches this type of data has diversified into various different formats that include large amounts of data and other small formats too. Data yah is highly useful for various purposes but one of the central utilizations is to ensure effective and complete transfer of information.

Data is a physical quantity that requires a certain amount of space to be able to store it effectively. This was usually done earlier through the utilization of hard disk in other removable storage devices such as floppy drives and compact discs. There is also been improvements in the storage facilities that have led to the increasing density that is available in magnetic disk drives which have led to an increased amount of storage capacities of these drives. In a normal computer personal computer or a laptop, these devices contain storage facilities such as hard disk drives or other flash-based storage. This type of storage requires regular maintenance and effective backups to remove the redundancy.

This has been done in large organizations that have an extensive amount of storage that is utilized for various different purposes including storing important documents and other data that is useful for the day-to-day functioning of the organization. Earlier all of these facilities were produced in house by the

organization which maintained this type of data extensively over the years. This was a highly time-consuming and expensive procedure that had to be performed to ensure the smooth functioning of the organization.

This was the use case scenario until the development and implementation of the internet platform that allowed for a variety of services to be produced on this platform. One of the most innovative approaches designed through the internet platform was the Cloud Service. The cloud service is referred to as the cloud because it is a misnomer that is attributed to various diagrams to represent the internet as a cloud. This definition stuck therefore it is being used largely to refer to the services that are offered by internet-based approaches. The cloud service is a novel platform that allows for remote storage options that are effective in achieving data retrieval ubiquitously.

The cloud storage platform allows the storage of data onto a remote server that can be accessed effectively from any part of the world utilizing any device that can connect to the internet. This meant that there was a reduction in the maintenance and other drawbacks associated with maintaining local storage. Therefore due to the increased convenience offered by the cloud platform, a large number of users and organizations have opted to completely transition towards a cloud-based storage approach. This has led to the rise of the cloud storage methodology which has been increasing in popularity ever since.

This has led to the creation of a lot of different cloud services such as public clouds that offer storage at a very reasonable price. But these platforms are plagued with various privacy and security concerns as storing the data on a remote server can make it highly susceptible to various forms of attacks and data leakage scenarios. Therefore most of the users of the cloud storage platform encrypt their data before storing it on the public cloud service to reduce the chance of data leakage in the event of a breach. This introduces a new problem that the data becomes very difficult to search and retrieve when it is needed as it is already encrypted.

Therefore to provide a solution to this problem this methodology focuses on the creation of an effective mechanism for performing a search over encrypted data that is stored on public cloud servers. This approach implements the utilization of metric transitions for encrypting the data using the reverse circle cipher approach. The reverse circle cipher approach provides effective encryption with the least amount of computational overhead through the implementation of diffusion and confusion through rotation of the blocks to generate the ciphertext. Which type of encrypted data can be effectively searched through an encrypted query by the application of a trapdoor generation. The methodology proposed in this research article also employs bilinear pairing and auto key generation that is tasked with maintaining the integrity of the data that is being stored along with the privacy effectively. The avalanche effect is used to detect any type of tempering that is being done on the data that is stored on the cloud platform.

This research paper utilizes the section 2 for evaluation of past researches as literature survey, section 3 in depth elaboration the proposed methodology and whereas section 4 provides the assessment of the performance of the system and finally section 5 provides a conclusion to the paper along with future directions of research.

II. LITERATURE SURVEY

Ebenezer R.H.P. Isaac et. al [1] suggested a reverse circle cipher solution based primarily on the principle of circular substitution by reversal transposition. This incorporates the basic cipher character level displacement principle, the vernam polyalphabetic cipher distribution principle, and the transposition cipher diffusion principle for optimal efficiency. Effectively, the cryptosystem used both personal data protection and network security. This now not only optimizes data output in transit but also provides an acceptable degree of data protection. The limitation of the proposed algorithm lies in the selection of the key by the user.

Taek-Young Youn et. al proposed a scheme for maintaining both integrity auditing and safe deduplication in a cloud environment. The presented scheme utilized the BLS signature based HLA

(Homomorphic Linear Authenticator)[2]. To support public integrity auditing authors introduced TPA. The CE-encrypted data is first created in the presented technique and then uploaded to cloud storage by the client to protect confidentiality after verification of the integrity of the outsourced data. Deduplication technology is used to save storage space and costs. CSS performs a PoW protocol during the deduplication process to verify that the client owns a file. Besides, in the context of the integrity audit process, it is necessary to generate and respond to the evidence corresponding to the TPA request. TPA conducts an integrity audit on behalf of clients to reduce production costs for the clients. Rather than the client, the auditor must give the server a challenge to perform an integrity audit protocol regularly. The proposed scheme met the safety objectives and improved the problems of the existing schemes. Besides, it provides better efficiency than existing systems from the point of view of client-side computational overheads.

Bryan H. Wodi et. al proposed a quick search for a privacy-preserving keyword model for encrypted outsourced data. The architecture of the proposed technique consists of the three main entities, data owner, cloud service provider, and data user. The main function of the data owner is to first build an encrypted tree index for all gathered keywords from numerous documents, then generate ciphertext collections for all the documents and outsourced encrypted index and ciphertext collection to the cloud storage server [3]. The cloud service provider saved encrypted index and ciphertext collection to the remote location and performs various search and update operations for various documents provided by the data owner. The data users are usually approved by the data owner to obtain a shared secret key for encrypted user data. Data users can use a search control mechanism to perform a search query using t keywords using a trapdoor TD and retrieve k encrypted documents from a cloud storage server based on a query keyword.

Xuqi Wang et. al introduced the eKAKSE model for data sharing via cloud storage. The architecture of the eKAKSE framework comprises three separate entities: the data owner, the user, and the cloud server. The cloud server creates public parameters using the Setup algorithm. The data owner generates his / her public key and master-secret key using the Keygen algorithm. While the data owner would need the Build Index algorithm to build the public index of every file. The data owner generates an Δi and encrypts keywords for each file using the Encrypt algorithm. When the data owner shares a specified group of files with a user, the Extract algorithm will produce an aggregate key and the aggregate key will be distributed securely to the user. An authenticated user searches for keywords, use his/her aggregate key to construct an aggregate trapdoor through the Trapdoor algorithm, and upload it to the cloud server. When the submitted trapdoor is provided, the cloud server will conduct a keyword search by the Test algorithm according to the user's file collection [4]. The trapdoor attack is the drawback of the proposed technique because it is still a practical problem to resist malicious user and cloud server collusion attacks.

Jianfei Sun et. al proposed a lightweight ciphertext-policy attribute-based keyword search system (CP-ABKS) with policy protection [5], which not only allows the data owner to share light data but also enables data users to retrieve and access shared data in an efficient manner without compromising the privacy of access policies.

The proposed CP-ABKS policy protection system consists of six phases: Initialization, Authorization, Data Outsourcing, Trapdoor Outsourcing, Keyword Retrieval, and Data Access. The algorithms involved in these phases are Setup, KeyGen, Offline / Online. Encrypt, Trapdoor, Search, and Decrypt. The machine initialization is done by the setup algorithm including the public parameter and the master secret key and the KeyGen algorithm performs user authorization producing a secret key for each system user, including a DU. Offline and online algorithms to generate ciphertext, and then outsource CS to ciphertext. When a DU intends to access data of interest, it requires that the Trapdoor algorithm be performed to obtain the search token, the trapdoor outsourcing to CS is subsequently done for retrieval. The CS performs Search algorithms to find the intended ciphertext after receiving the search token and then returns a simple ciphertext to the DU. Eventually; the DU calls the Decrypt algorithm to simulate the phase of accessing data to recover the plaintext message.

Yang Chen et al. introduced a reliable server and a dual application configuration. Based on the concept of the double server they proposed an ABKRS-KGA scheme to help multi-keyword graded search which resists simultaneously within KGA. The model consists of five entities Authority, Server A, Server B, DO (Data Owner), and Users. The role of authority is to generate the users and two servers with a public key system and secret keys. The owner of the data creates the indexes and uploads the indexes to server A and encrypted files to server B Data users use their key to create a trapdoor that contains interesting keywords and apply the trapdoor to server A. Server A stores the data owner-uploaded indexes, the user uploads his / her trapdoor to it, performs part of the search process, and sends the temporary result to server B. Server B is presumed to have ample storage space, it stores encrypted files that users upload. It receives the temporary result from server A and the final search is complete [6]. The authors first presented two types of security models that consider multi-keyword search in the ABKS scheme and the proposed scheme are known to be safe both against CKA and KGA. Besides, the user queries include weighted keywords in the scheme, and the returned files can be rated according to the user's query value, which can dramatically boost user search experience and the efficiency loss caused by incorrect user interest positioning.

Fei-Ju Hsieh et. al presented a multi-keyword semantic search scheme preserved for privacy over encrypted cloud data. The data owner must save the encrypted document in the cloud storage under the proposed scheme. Each document has an encrypted index created by the owner of the data. Users can provide multiple keywords if they wish to find relevant documents inside the dataset. The keywords can vary from the ones used in document indexes. The corresponding encrypted documents will be returned to the user so long as the semantic sense of the query keywords and the document index terms are fairly similar. In the meantime, the application is encrypted and used as a trapdoor for the search to protect user privacy. A secured matching calculation adopted to suit the records so as not to snoop from provider cloud storage. Extended simulations conducted using a sample from a particular database of documents [7]. The limitation of the proposed system is that it failed to build a semantic relationship between different words.

Fan Yin et. al [8] introduced a new privacy-preserving multi-keyword conjunctive query scheme, which can well balance the efficiency and privacy in a query. The authors first build a tree-based index to represent keywords, called the conjunctive tree. The server can perform multi-keyword conjunctive queries efficiently with this conjunctive tree. Simultaneously, they utilized the BGN homomorphic encryption technique to encrypt the conjunctive tree, which may well preserve keyword privacy. A novel wildcard search algorithm was designed to improve the performance of conjunctive queries, replacing a part of computationally costly tree traversing operations with effective string copy operations. The security analysis of the proposed algorithm proved that its efficiently achieve query with small leakage.

Muhammad U. Arshad et. al proposed effective integrity verification schemes. The first scheme is for graphs and query results for graphs that are based on HMACs. The HMACs scheme is for two-party data sharing. Another scheme is for the sharing of data by third parties, which is rewritable HMACs for graphs. The schemes are exploiting the security properties of the XOR operator (especially when used) used with key (k, r), one-way hash functions, and depth-first crossing of graphs) in an efficient yet proven manner [9]. The schemes are based on the local/global integrity verifiers of the vertexes and edges.

Filipe Apolinario et. al [10] has introduced S-AUDIT, a software service that enhances the Shacham-Waters (SW) integrity verification scheme and modifies it to use in cloud storage. S-AUDIT enhances the original SW scheme by providing an overall improvement in efficiency by carefully choosing pair-friendly elliptical curves for parametrization of the SW scheme, and a 50% decrease in storage costs compared to the original SW scheme by using compression point. Besides, it leverages the Function-as-a-Service (FaaS) model to reduce cloud costs by using cloud computing services only when appropriate. These enhancements make S-AUDIT the most cost-effective homomorphic verification method for commercial use. The key contributions of this paper are: the S-AUDIT integrity verification service designed and implementation, a protocol for the verification of data stored in clouds and a proof-of-concept integration of S-AUDIT with a commercial cloud and a cloud-based file system, and an experimental evaluation of the use of this service standalone and integrated with AWS and SCFS.

Bilin Shao et. al presented a dynamic, efficient, and secure data integrity auditing scheme that promotes data privacy protection. This scheme builds the HMBT (Hierarchical Multiple Branches Tree) authentication structures, which not only essentially decreases the height of the authentication structure but also allows tenants to change data with different granularity dynamically. In pre-processing data, this scheme colors the data, preventing the third party auditor from stealing the data privacy of the tenant during the verification process. Finally, for the assessment of the system, the security analysis, and the performance analysis are carried out [11]. The security review shows that this scheme will satisfy the correctness of the audit, help the protection of data privacy, resist forgery attacks and replay attacks. The scheme is contrasted with the current scheme during performance analysis.

Wenting Shen et. al proposed a new idea called identity-based shared data integrity audit with safe hiding of confidential information on the cloud. In this scheme, the confidential and sensitive information can be saved and other information may be released. This makes sharing and use of the file stored in the cloud by others and ensure that confidential information is secured, while remote data integrity audits can still be effectively carried out. The authors also designed the practical identity-based, shared data integrity audit scheme for secure cloud storage with confidential information hiding [12]. A sanitizer is used to sanitize the data blocks which correspond to the file's sensitive information. This system not only provides remote auditing of data integrity but also facilitates data sharing so that confidential information is secured in cloud storage. The safety proof and protection and the experimental review indicate that the proposed scheme achieves desirable safety and efficiency.

Ahmad Alsharif et. al presented a scheme named as EPIC (Efficient and Collusion-Resistant Privacy-Preserving Power Consumption Collection). The main idea is that each SM (Smart Meter) selects several SMs in the network called "proxies" and effectively calculates the mutual hidden masks with each proxy in pairs [13]. Therefore, it would cover its fine grain readings with all the masks exchanged with proxies in the order to cancel all the masks after the inclusion of all the masked readings of all the meters. To effectively mask the fine-grained readings EPIC uses safe and lightweight operations and aggregate the masked readings to permit the utility to gather a fine-grained aggregated reading without leaking sensitive information to consumers. As the reading of meters can be changed during the transmission to the utility, EPIC allows the utility to check the validity of the aggregated reading without requiring access to the individual readings to protect the privacy of consumers. Using homomorphic hash properties, EPIC allows the utility to easily measure electricity bills depend on dynamic pricing, without infringing the privacy of customers.

III. PROPOSED SYSTEM

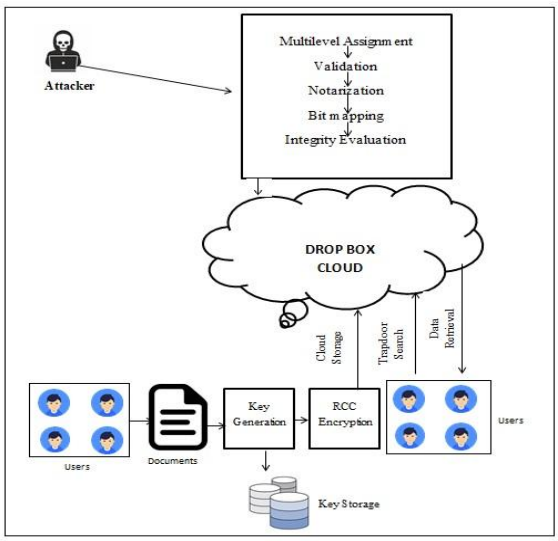

Figure 1: System Overview

The proposed Methodology for implementing an effective search over encrypted data as shown in Figure 1 is detailed below.

Step 1: *User Input* – This is the first step in the presented technique, in this step the user uploads certain data on to the cloud platform. The data that is being uploaded consists of text files which are pre-processed by the model before being uploaded on the cloud. The pre-processing procedure performs stop word removal and stemming that removes the unimportant items in the text, which can significantly reduce the space and time complexity of the system and enhance the search process considerably.

This process of Preprocessing is done by using the following steps

Special Symbol Removal – These are symbols that are utilized for implementing a structure to the spoken English language. These special symbols are not required for any other purpose and are redundant in our methodology. Therefore, all special symbols are removed from the string.

Tokenization – The string obtained from the previous special symbol removal step is provided to this step for tokenization. The tokenization procedure involves the division of the string into smaller parts or tokens. These token are then stored as a well-indexed string for processing in the system further. Tokenization ensures that the string can be easily converted into an array list for further processing.

Stopword Removal – There are certain words in the English language that are purely used for their aesthetic purposes. Some of those words are the stop words that are used to denote pauses and conjunctions of two sentences. These words are only for aesthetic reasons and are otherwise not required by our system. As the stopwords do not provide any additional meaning they are eliminated in this step.

For example, we look at the sentence “going to run” if utilized for the stopword removal step of the preprocessing. The stopword in this sentence is “to” which will be eliminated and the phrase is converted into “going run”. This example showcases that the stopwords removal step does not alter the meaning of the string.

Stemming – There are different versions of a single word in the English language that differ from each other in their timing and tense of the usage. This makes sure that the listener is well aware of the sequence and the language is precise in the conversation. But in our utilization, the preciseness of tense is not necessary. Besides, the stemming procedure does not change the semantics of the word. Therefore, the words in the string are stemmed to their root forms.

For example, “going” will be stemmed to “go” by the removal of the substring “ing” which is restored with an empty character in its position. It can be observed that there is no semantic difference between “going” and “go”. But stemming can have a considerable impact on the time and resources utilized for the processing of the string as it has been shortened which improves the execution time of the system.

Step 2: *Bucket Creation* or Matrix Transition– The data that is processed in the previous step from the user is utilized as an input in this step. The pre-processed input data is eventually coupled together to form a bucket. This separates the words in the text and frames them consecutively into its substring. For Example, the substrings for the word Roman is {“rom”, “roma”, “roman”}. The procedure is detailed in Algorithm 1 given below.

ALGORITHM 1: Bucket / Matrix Translation

// Input : String S_T
// Output : M_L Matrix Translation List
Function : matrixTranslation (**S_T**)
0: Start
1: WRD_{LST}[ ]=Spilt S_T on " "
2: M_L = ∅ [Stopword Text]
3: ***for*** i=2 **to** length of WRD_{LST}
4: WRD= WRD_{LST} [i]
5: ***for*** j=2 **to** length of WRD -1
6: SUB_{WRD}=WRD_{SUB}[0,j]
7: M_L=M_L+ SUB_{WRD}
8: ***end for***
9: ***end for***
10: return M_L
11: Stop

Step 3: *Cloud Storage* – The data buckets formed in the previous step are taken as an input in this step for the bucket members to be subjected to the encryption purposes. The RCC (Reverse circle Cipher) Encryption standard is utilized for this purpose for which the encryption keys are generated.

Key Generation – The MD5 hashing technique is utilized for the creation of the hash key for the data that is being uploaded on to the cloud. Seven characters are selected randomly through the implementation of the continuous rotation process. These characters are provided as an input to the RCC algorithm which generates the keys which are subsequently used to encrypt the data and the user attributes.

Reverse Circle Cipher – Reverse Circle Cipher is one of the most powerful encryption approaches that are capable for implementation on a network. RCC technique executes through the rotation of the input characters in either anti-clockwise or clockwise procedure along with the replacement of the characters. This is achieved in the presented model through the segregation of the data into blocks and performing rotations on the blocks to encrypt them. The encryption is performed accurately by the Reverse Circle Cipher and the encrypted data is then uploaded on to the cloud. The Reverse Circle Cipher is one of the most useful and efficient encryption techniques that is capable of preserving the privacy of the data.

Step 4: *Bilinear Pairing* – Bilinear pairings are generated through the hash keys created in the previous step through the utilization of the MD5 algorithm. To ensure the integrity of the data that is being stored on the platform is maintained the Avalanche effect on the keys is evaluated over extended periods of time.
Avalanche Effect – The definition of Avalanche is the really destructive and sudden fall of accumulated snow over mountain and hilltops in a very quick succession. This is why the term is used to reflect the effect when there is a drastic change in the hash keys that are generated by the MD5 algorithm corresponding small change in the data.

This leads to severe changes over time and can corrupt the hash keys significantly which can trigger an integrity violation on the encrypted data on the cloud. This is a very critical scenario if its integrity is damaged beyond recognition it can make the data unrecoverable.

Step 5: *Searching* – The initial step that is performed for the purpose of enabling the search is the utilization of the trap door. Trap Door – A set of encrypted Queries are used to search the encrypted

entities generated by the bucket formation procedure, this is referred to as the Trapdoor. The procedure of the trap door creation has been elaborated in Algorithm 2 below.

ALGORITHM 2: Trap Door Creation

//Input : Set M_L Matrix Translation List ={M_{L1},M_{L2},M_{L3} ….M_{LN} } a Matrix Translation List for the given Query
// Input: Key_{RCC} Key for RCC Encryption

// Output : T_{RPD} Trap Door(encrypted Query)
Function : trapDoor ($M_{L,}$ Key_{RCC})
1: Start
2: T_{RPD} =**Ø**
3: ***for*** **i=0** to Size of M_L

4: SUB_{STR}=$M_{L[i]}$
5: TDi =encrypt(SUB_{STR}, Key_{RCC})
6: $T_{RPD=}$ T_{RPD}+ T_{Di}
7: ***end for***
8: return T_{RPD}

9: Stop

This trap door is implemented for the search purposes on the cloud platform as the hash keys for both the data and the Query are similar. This improves the accuracy of the searching module and provides effective and complete search results for the particular query.

IV. RESULT AND DISCUSSIONS

The presented technique for performing search over encrypted data and implementing effective data security in the cloud is achieved on a windows based laptop. The laptop consists of an Intel core i5 along with 6 GB of Primary memory. For the deployment of the presented model the Java Programming language is used along with NetBeans as the Integrated Development Environment. The presented Technique has utilized the MySQL for the database requirements. The methodology is deployed using Dropbox a public cloud service. To utilize the drop box system the presented approach uses its authentication key and the respective API. Extensive experimentations are performed to calculate the authenticity of the proposed model is elaborated below.

Precision and Recall – Searching techniques are predominantly evaluated through the execution of Precision and Recall as they are highly precise performance metrics. A collection of relatively sized keywords is used for the query and the search results are listed in Table 1 given below. The average Precision and average Recall are measured. The values are converted into a graph depicting the values in Figure 2.

Experiment No	Relevant File Extracted (A)	Irrelevant File Extracted (B)	Relevant files Not extracted (C)	Precision in % (A/(A+B))*100	Recall in % (A/(A+C))*100
1	16	3	1	84.21052632	94.11764706
2	9	2	0	81.81818182	100
3	6	1	0	85.71428571	100
4	13	2	1	86.66666667	92.85714286
5	9	0	0	100	100
6	15	2	0	88.23529412	100
7	3	0	0	100	100
8	5	0	0	100	100
9	7	1	0	87.5	100
10	8	0	1	100	88.88888889

Table 1: Precision and Recall experiment Results of the proposed model.

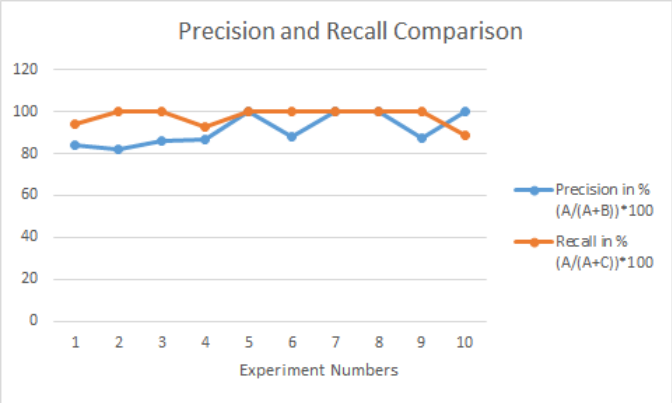

Figure 2: Precision and Recall experiment results of the Proposed Methodology.

These values for the Precision and Recall of the proposed system are compared with that of [14] which implements a model reliant on K Nearest Neighbors clustering technique. These dictate that our technique outperforms the one depicted in [14].

This is owing to the fact that the K Nearest Neighbor algorithm has a high complexity in the resultant system that is observed due to the increase in the number of iterations. The proposed methodology, on the other hand, deploys a variant of the similarity search using matrix transition technique for performing search operations on the files, which provides highly accurate results and also diminishes the complexity of the system as a whole. This can be noticed in Table 2 and the resultant graph plotted for the outcomes recorded in Figure 3.

Methodology	Average Precision	Average Recall
KNN Search	84.5	97
Matrix Transition Search	91.41	97.58

Table 2: Comparative Results of Precision and Recall between KNN Search and Jaccard Search Matrix Transition Search

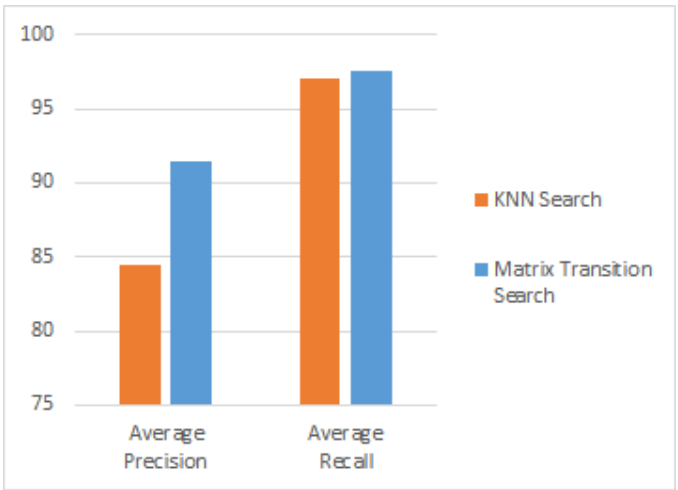

Figure 3: Comparative Results of Precision and Recall, KNN Search V/s Matrix Search Result

V. CONCLUSION AND FUTURESCOPE

There has been a tremendous growth that has been noticed in the cloud platform in the recent years. This is due to the ubiquitous nature of cloud storage that offers increased convenience to its users. This has been one of the most important criterions for the increase in the popularity of the approach which has led to an increased user base of the cloud storage approach. This has been plagued with the security and privacy concerns that have been raised against the remote storage of the data on the cloud platform. This can be highly susceptible to attacks and other data leakages. Therefore, most of the data stored on the cloud platform is encrypted before being uploaded. This makes the data very difficult to search and retrieve. Therefore, the presented technique implements an effective encryption technique utilizing the RCC approach through the assistance of matrix transitions, trapdoor generation and search token approach. The presented approach has been implemented on a public cloud service called Dropbox. The proposed methodology has been evaluated through extensive experimentation and tests that ensures that it is considerably better than the KNN search technique detailed in [14]. This has been confirmed through the implementation of Precision and Recall performance metrics.

This paper also presents a model for the preservation of the integrity of the data as well as the hiding of sensitive data in the cloud platform. This is very critical as there has been an ever-increasing rate of acceptance and integration of the cloud platform due to its increased convenience. For assessment of the presented technique, based on its efficiency, integrity proof generation and auditing time evaluation have been considered.

For the implementation of future research directions, the integrity maintenance can be enhanced to take as input larger files of all formats effectively. And the presented system can be improved to consider storage of different file formats other than text files.

VI. REFERENCES

[1] Ebenezer R.H.P. Isaac, Joseph H.R. Isaac and J. Visumathi, "Reverse Circle Cipher for Personal and Network Security", 2013 International Conference on Information Communication and Embedded Systems (ICICES), IEEE, 29 April 2013.

[2] Taek-Young Youn, Ku-Young Chang, Kyung-Hyune Rhee and Sang UK Shin, "Efficient Client-Side Deduplication of Encrypted Data With Public Auditing in Cloud Storage", IEEE Access (Volume 6), DOI: 10.1109/ACCESS.2018.2836328, 15 May 2018.

[3] Bryan H. Wodi, Carson K. Leung, Alfredo Cuzzocrea and S. Sourav, "Fast Privacy-Preserving Keyword Search on Encrypted Outsourced Data .2019 IEEE International Conference on Big Data (Big Data), DOI: 10.1109/BigData47090.2019.9046058, 26 March 2020.

[4] Xuqi Wang, Yu Xie, Xiangguo Cheng and Zhengtao Jiang, "An Efficient Key-Aggregate Keyword Searchable Encryption for Data Sharing in Cloud Storage" , 2019 IEEE Globecom Workshops (GC Wkshps), DOI: 10.1109/GCWkshps45667.2019.9024540, 9-13 Dec. 2019.

[5] Jianfei Sun , Hu Xiong , Robert H. Deng , Yinghui Zhang , Ximeng Liu and Mingsheng Cao, "Lightweight Attribute-Based Keyword Search with Policy Protection for Cloud-Assisted IoT" , 2019 IEEE Conference on Dependable and Secure Computing (DSC), DOI: 10.1109/DSC47296.2019.8937708, 18-20 Nov. 2019.

[6] Yang Chen, Wenmin Li, Fei Gao, Qiaoyan Wen, Hua Zhang and Huawei Wang, "Practical Attribute-based Multi-Keyword Ranked Search Scheme in Cloud Computing" , IEEE Transactions on Services Computing (Early Access), DOI: 10.1109/TSC.2019.2959306, 18 December 2019.

[7] Fei-Ju Hsieh, Tai-Lin Chin, Chin-Ya Huang, Shan-Hsiang Shen, Chung-An Shen, "Semantic Multi-Keyword Search over Encrypted Cloud Data with Privacy Preservation" , 2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall), 07 November 2019.

[8] Fan Yin, Yandong Zheng, Rongxing Lu, (Senior Member, IEEE), and Xiaohu Tang, " Achieving Efficient and Privacy-Preserving Multi-Keyword Conjunctive Query Over Cloud", IEEE Access (Volume: 7), DOI: 10.1109/ACCESS.2019.2954043, 20 November 2019.

[9] Muhammad U. Arshad, Ashish Kundu, Elisa Bertino, Arif Ghafoor, Chinmay Kundu" Efficient and Scalable Integrity Verification of Data and Query Results for Graph Databases", IEEE Transactions on Knowledge and Data Engineering, Volume: 30 , Issue: 5 , May 1 2018 .

[10] Filipe Apolinario, Miguel L. Pardal, Miguel Correia, "S-Audit: Efficient Data Integrity Verification for Cloud Storage",2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering", DOI: 10.1109/TrustCom/BigDataSE.2018.00073, 06 September 2018 .

[11] Bilin Shao, Genqing Bian Yue Wang, Shenghao Su, Cheng Guo " Dynamic Data Integrity Auditing Method Supporting Privacy Protection in Vehicular Cloud Environment ", IEEE Access (Volume: 6), DOI: 10.1109/ACCESS.2018.2863270, 08 August 2018.

[12] Wenting Shen, Jing Qin, Jia Yu, Rong Hao, and Jiankun Hu, "Enabling Identity-Based Integrity Auditing and Data Sharing With Sensitive Information Hiding for Secure Cloud Storage", IEEE Transactions on Information Forensics and Security (Volume: 14 , Issue: 2 , Feb. 2019), 25 June 2018.

[13] Ahmad Alsharif, Mahmoud Nabil, Samet Tonyali, Hawzhin Mohammed, Mohamed Mahmoud, and Kemal Akkaya, "EPIC: Efficient Privacy-Preserving Scheme with E2E Data Integrity and Authenticity for AMI Networks", IEEE Internet of Things Journal (Volume: 6 , Issue: 2 , April 2019), 21 November 2018.

[14] Cengiz Orencik, Erkay Savasy and Mahmoud Alewiwiz, "A United Framework for Secure Search Over Encrypted Cloud Data ", IACR Cryptology ePrint Archive 2017