

# Shamir's Key Based Confidentiality on Cloud Data Storage

Kamalraj Durai

Department of Computer Science, Bharathiar University, India

**Abstract:** Cloud computing is a flexible, cost effective and proven delivery platform for providing business or consumer services over the Internet. Cloud computing supports distributed service over the Internet as service oriented architecture, multi-user, and multi-domain administrative infrastructure, hence it is more easily affected by security threats and vulnerabilities. Cloud computing acts as a new paradigm where it provides a dynamic environment for end users and also guarantees Quality of Service (QoS) on data confidentiality. Trusted Third Party ensures the authentication, integrity and confidentiality of involved data and communications but fails on maintain the higher percentage of confidential rate on the horizontal level of privacy cloud services. TrustedDB on the cloud privacy preservation fails to secure the query parsers result for generating efficient query plans. To generate efficient privacy preserving query plans on the cloud data, we propose Shamir's Key Distribution based Confidentiality (SKDC) Scheme to achieve a higher percentage of confidentiality by residing the cloud data with polynomial interpolation. The SKDC scheme creates a polynomial of degree with the secret as the first coefficient and the remaining coefficients picked up at random to improve the privacy preserving level on the cloud infrastructure. The experimental evaluation using SKDC is carried out on the factors such as system execution time, confidentiality rate and query processing rate, which improves the efficiency of confidentiality rate and query processing while storing and retrieving in cloud.

**Keywords:** Confidentiality, privacy, cloud computing, SKDC, privacy preserving and polynomial interpolation.

Received April 25, 2015; accepted January 28, 2016

## 1. Introduction

Cloud users remotely store the user data and attain a high quality of services on various cloud applications. Cloud computing acts as a computing paradigm where it provides a dynamic environment for end users and also guarantee QoS on data confidentiality and integrity. The services share a pool of configurable resources and data with higher confidentiality and integrity level [11, 15, 16]. Cloud Service Providers (CSP) are part of administrative entities where the data outsourcing is really relinquishing the user's decisive control over the fate of the cloud data. Cloud storage is gaining higher popularity for the outsourcing of day-to-day data management. Confidentiality is used on accessing the set of cloud database information with a high security level [7, 14]. Integrity monitoring of data in Cloud Storages (CDs) is as essential for any data center, to avoid any data corruption. Cloud computing offers cloud based data privacy preservation on the different applications [1, 2, 8, 12, 19]. This research work is carried out on cloud data privacy where preservation is undertaken under the wide range of both interior and exterior threats. This paper is divided into four sections, Section 1 offers a basic introduction about cloud technology and issues in cloud. The reviews of the related work are discussed in section 2 towards confidentiality and give the limitations of the conventional methods. Section 3 introduces the architecture of the proposed work and discusses the

components of implementing a secured cloud. The results and analyses are discussed in section 4 and the evaluation of the implementation using the simulator CloudSim. Section 5 summarises the conclusion and the advantages and provides suggestions for future work.

## 2. Related Work

Cloud users remotely store the user data and attain high quality services on various cloud applications. The services share a pool of configurable resources and data with a higher confidentiality level. An existing trusted third party within a cloud environment as presented in [3, 5, 6, 9, 10, 13, 21] specifically designs the Public Key Infrastructure. The trusted third party ensures the authentication, integrity and confidentiality of the involved data and communications but fails in maintaining a higher percentage of confidentiality rate on the horizontal level of privacy cloud services. Trusted hardware based database with privacy and data confidentiality (TrustedDB) Scheme [4, 19] inbuilt and runs on actual hardware for different stages of query processing and achieves a higher confidentiality rate. TrustedDB on the cloud privacy preservation fails to secure the query parsers result for generating efficient query plans [4].

Data confidentiality using fragmentation as demonstrated in [12] applied relational databases to process the independent fragments. The fragmentation

technique efficiently stores the data on the cloud service providing servers but has not been tested practically on the cloud environment. The secure cloud storage system [18] supports privacy-preserving public auditing on a single user setting. The work is not further implemented on a multi-user setting. Privacy-preserving public auditing on the cloud environment uses the aggregation and algebraic properties of the confidentiality but has not benefited from the batch auditing.

To generate the efficient privacy preserving query plans on the cloud data, Shamir's Key Distribution based Confidentiality (SKDC) scheme is presented. SKDC scheme handles the horizontal and vertical levels of privacy cloud data services using the vertical proxy query processing operations. It ensures query efficiency using matrix (row (horizontal) x column (vertical)) accesses. Matrix form query processing in SKDC ensures a higher confidentiality rate by the cloud service providers. Vertical proxy query processing handles practically all different types of queries among the varying groups of clients.

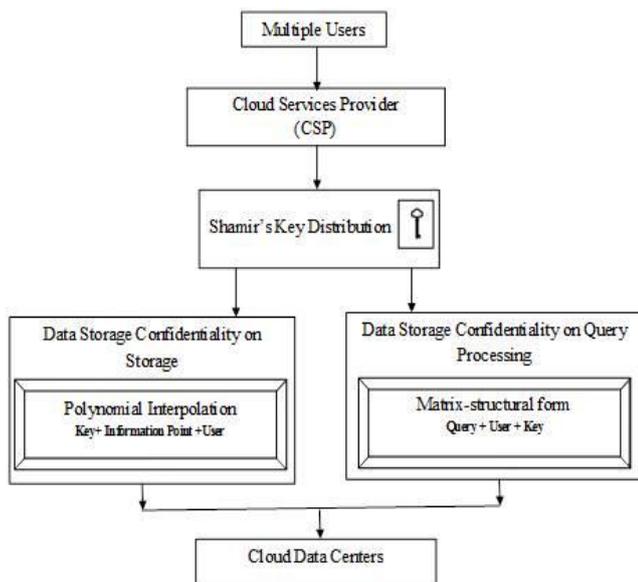


Figure 1. Shamir's key distribution based confidentiality processing.

Shamir's Key distribution achieves a higher percentage of confidentiality by residing the cloud data with polynomial interpolation. The SKDC scheme creates a polynomial with the secret as the first coefficient and the remaining coefficients picked at random to improve the privacy preserving the level on a cloud infrastructure. The 'k' is the key hidden from the public cloud users to improve the confidentiality rate. Key distribution supports batch auditing where multiple users' request for data auditing is held concurrently at a higher rate of confidentiality.

### 3. Confidentiality on Cloud Data

In cloud computing, the data needs to be stored

securely, since the data are outsourced to a third party provider where it provides a dynamic environment for end users and also guarantees QoS on data confidentiality [16]. Thus, to provide an efficient platform in cloud with privacy preserving query plans SKDC scheme is presented. Shamir's key distribution achieves a higher percentage of confidentiality by residing the cloud data with polynomial interpolation. The SKDC scheme creates a polynomial of degree with the secret as the first coefficient. Then, the remaining coefficients are picked up randomly to improve the privacy preserving level on the cloud provider. The 'k' is the key which is secretly maintained from the public cloud users to improve the confidentiality rate [17]. By using Shamir's key distribution we proposed a model for confidentiality in cloud data storage.

#### 3.1. Shamir's Key Distribution Based Confidentiality on Cloud Data Storage and Query Processing

In this section, we formulate the problem of SKDC scheme to maintain high confidentiality and to provide privacy to cloud users and also provide the necessary background regarding confidentiality maintenance and the proposed method SKDC scheme. The two components involved in the design of the SKDC scheme are:

1. Cloud storage data confidentiality.
2. Privacy preserving on query processing [15, 21].

Privacy level maintenance on cloud data storage and query processing is the main objective in the proposed work. The cloud data storage in our model stores the data and maintains the security using the Shamir's key distribution model. Maintaining a higher confidentiality level is a serious concern in cloud which is addressed through Shamir's key distribution model. Data confidentiality is one of the most significant security (i.e., privacy) concerns that is achieved in our proposed work that helps in safeguarding the client's information from attackers. The steps involved in the design of confidentiality via Shamir's Key Distribution are shown in Figure 1.

Shamir's key distribution implements a higher percentage of confidentiality on the data by residing in the cloud. The SKDC scheme uses polynomial interpolation and the matrix-structural form to provide data storage confidentiality on storage and data storage confidentiality on query processing respectively. The SKDC scheme creates polynomial degree coefficients to improve the security (i.e., privacy preserving) level. In the matrix-structural representation, the horizontal and the vertical forms are used to easily access the pattern and to fetch the result for the user query from the cloud infrastructure.

The horizontal row form information is used to

improve the quality of services on processing the query. The vertical form handles different types of queries among varying groups of clients simultaneously, to fetch the accurate query result with in a minimal processing time.

The proposed SKDC architecture for different cloud users is described in Figure 2. Cloud computing is a model for enabling convenience on-demand network access for cloud users. Initially, the user requested information for cloud data storage is carried out using polynomial interpolation. The SKDC scheme creates a polynomial of degree with the secret key as the first coefficient and the remaining coefficients to improve the privacy preserving level on cloud infrastructure. The ‘k’ is the key hidden from the public cloud users that helps in improving the confidentiality rate. Shamir’s key distribution supports batch auditing where multiple user requests for data auditing is placed concurrently with the objective of providing a higher rate of confidentiality.

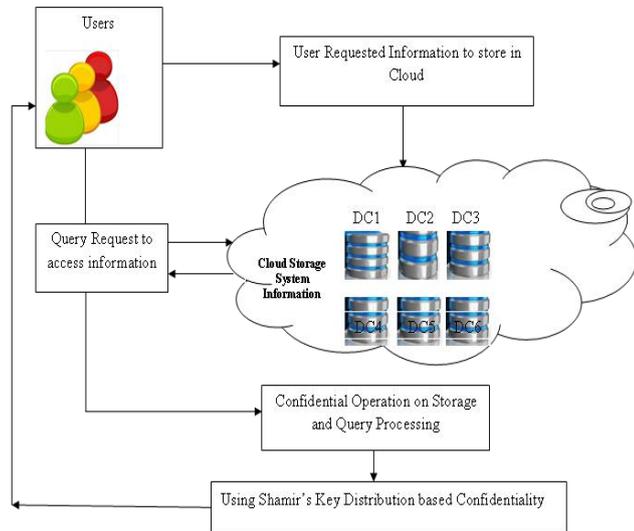


Figure 2. Proposed shamir’s key distribution based confidentiality architecture.

The SKDC scheme handles query processing using the matrix-structure form. The horizontal and the vertical levels of privacy cloud data service query processing uses the matrix form. The SKDC scheme ensures query efficiency using matrix. As a result, the SKDC scheme ensures high confidentiality on cloud data storage query processing. The horizontal row is used to improve the privacy preserving; whereas, the vertical column accesses multiple user queries [7].

### 3.2. Cloud Storage Data Confidentiality

The first component i.e., cloud storage data confidentiality with the objective of achieving here higher confidentiality is discussed in detail. The design of cloud storage data confidentiality using the SDKC scheme proceeds in such a way that if a client wants to store the information in the cloud, a distributed key is sent to the cloud server. The storage of information on

the cloud server with the user’s authentication supports the effective data forwarding of the same. Once the system has been designed, the next step is to convert the designed form into the actual implementation in the SKDC scheme using polynomial interpolation.

### 3.3. Polynomial Interpolation

SKDC on cloud data storage uses the polynomial interpolation points. The purpose of using polynomial is to obtain the different curves (i.e., sizes) of the user information in the cloud server. The cloud storage zone uses the lookup table and interpolates it between those information points of different users. Let us assume a set of information points,  $P$ , obtained from multiple cloud users,  $U$ , to store in the cloud with the polynomial property is given Equation (1).

Polynomial Property

$$(U_i) = U_1 [P_1], U_2 [P_2], U_3 [P_3], \dots, U_m [P_m] \quad (1)$$

The polynomial property  $U_i$  with different information points is then stored in the cloud infrastructure. With the application of the polynomial interpolation,  $PI$ , in the SKDC scheme even complicated information is approximated and stored in the cloud infrastructure. The results of the cloud storage using the polynomial interpolation are obtained significantly with a minimum processing time. The construction of the interpolation for all the cloud users using linear information is given in Equation (2),

$$\begin{bmatrix} u_1^n & u_1^{n+1} & u_1^{n+x} \\ u_2^n & u_2^{n+1} & u_2^{n+x} \\ u_i^n & u_i^{n+1} & u_i^{n+x} \end{bmatrix} + \begin{bmatrix} c_n \\ c_{n-1} \\ c_0 \end{bmatrix} = \begin{bmatrix} u_1 [PI_i] \\ u_2 [PI_i] \\ u_i [PI_i] \end{bmatrix} \quad (2)$$

The  $U_i$  user’s information from ‘n’ to ‘n+x’ points are stored in the cloud zone. The coefficient is used to embed the Shamir distributed key with the information points to improve the confidentiality result. The SKDC scheme on cloud storage creates a polynomial degree with a secret key as a first coefficient  $C_n$  and the remaining coefficients with size  $C_{n-1}$ . Multiple storage information from multiple users is also provided with a high privacy rate in the proposed method. Multiple users’ information  $U_i$  is embedded with the polynomial property. Shamir key is the hidden key embedded with the cloud storage information and it is shown as,

$$K = \sum_{i=1}^m U_i \text{ mod } n \quad (3)$$

The Shamir key ‘K’ is embedded with the information points by the user to improve the confidentiality rate.  $U_i$  is the information points after applying the polynomial interpolation form. Shamir’s key distribution is used on the information point’s storage to improve the confidentiality level. Algorithm 1 gives a outline of the process about the cloud data storage for the SKDC scheme.

Algorithm 1: Cloud Data Storage for the SKDC

*Input:* Cloud Users  $\{U_i = U_1, U_2, \dots, U_m\}$ , Polynomial Interpolation points  $\{PI_j = PI_1, PI_2, \dots, PI_n\}$ , Information Points  $\{P_a = P_1, P_2, \dots, P_b\}, i, j, a, m, n, K$

*Output:* Higher confidentiality maintained on cloud storage zone

```

1:  $U_i$ 
2:  $\{ P_a$ 
3:  $\{$  apply polynomial degree of interpolates
4: Polynomial Property  $= U_1 * [P_1] + U_2 * [P_2] +$ 
 $U_3 * [P_3] + \dots + U_m * [P_b]$ 
5: use look up table on information point storage from (2)
6: embed Shamir key distribution
7:  $K = U_i \text{ mod } n$ 
8:  $\}$ 
9:  $\}$ 

```

In the SKDC cloud infrastructure, the data storage contains the entities, such as cloud user and cloud service provider. With the application of Shamir's key distribution, the cloud user stores a large amount of data in the cloud server with lesser time taken for storage. The look up table is supported with an array indexing operation to store the information points in the cloud infrastructure. The lookup table in the SKDC scheme is used extensively in processing the user query based on the cloud storage information. The time being saved in terms of processing is the significant part in the proposed method.

### 3.4. Privacy Preserving on Query Processing

In this section, the second component privacy preserving on query processing is designed with the objective of increasing the confidentiality rate on query processing. The results of the query retrieval process in the SKDC scheme retrieve the result from the cloud infrastructure by using the look up table information. The original information is retrieved and provided to the authenticated users, provided they submit the accurate private key to the cloud server using the matrix-structure form. The resulting symbols and the coefficients are checked on the cloud zone to evaluate the privacy level measure. The cloud server heavily loaded with the multiple users is easily processed using this matrix-structure form.

The SKDC scheme is used to extract the result for a range of the user queries with a high privacy level. The tuple updates are performed in the look up table using the insertion and deletion operations. The user  $U$ , collects the query result from the cloud zone using the matrix-structure query processing or matrix-structure form which is detailed in the following section.

### 3.5. Matrix-Structure Form

The matrix structure form in the SKDC scheme involves information retrieval in the proposed work with fast access of data. The matrix-structure form ensures query efficiency using a matrix form with row (i.e.) horizontal and column (i.e.) vertical accesses. Figure 3 shows the matrix-structure form that contains

certain user information.

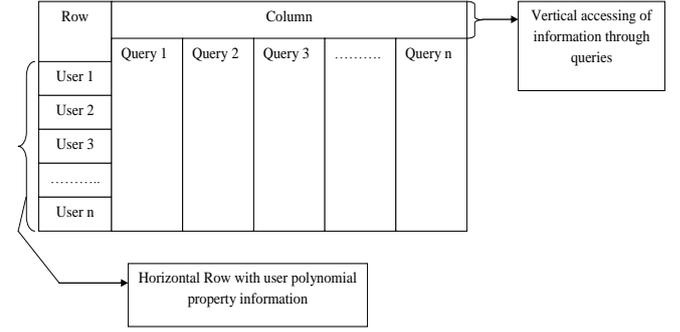


Figure 3. Design of the matrix-structure form.

The matrix-structure form consists of the information points and also the query terms with higher a confidentiality rate using the SKDC scheme. Each row with the SKDC scheme consists of the user's information points, whereas each column vector defines the query terms requested from the client side. SKDC scheme ensures query efficiency using matrix-structure form. The query processing is improved by privacy rate in the proposed method by using the query processing in Equation (4):

$$q(U) = K + \sum_{i=1}^m c^n U_i \text{ mod } n \quad (4)$$

The private key ' $K$ ' is embedded with the query ' $Q$ ' to fetch the result for the particular user ' $U$ '. The coefficient points of the polynomial interpolation property are used to improve the privacy level. Algorithm 2 describes the matrix-structure form. The user embeds the distributed key with the query to the cloud server to fetch the result from the stored information. The stored information points are processed and produce higher confidentiality results from the cloud infrastructure. This proposed SKDC algorithm is efficient and strong in attaining confidentiality while storing and retrieving the results through query processing.

*Algorithm 2: Matrix-Structure Form for the SKDC Scheme*

*Input:* Cloud Users  $\{U_i = U_1, U_2, \dots, U_m\}$ , Rows  $\{R_s = R_1, R_2, \dots, R_n\}$ , Columns  $\{C_t = C_1, C_2, \dots, C_n\}$

*Output:* Process the query with high confidentiality rate

```

1: for each  $U_i$ 
2: accept query terms with Shamir key distribution ' $K$ '
3:  $K = U_i \text{ mod } n$ 
4: for each  $R_s$ 
5: updated cloud user information
6: for each  $C_t$ 
7: update queries on each term
8: process the queries
9: end for
10: end for
11: end for

```

## 4. Experimental Evaluation

The SKDC scheme is developed to improve the

privacy level using the Amazon Simple Storage Service (Amazon S3) dataset. This dataset based on confidentiality maintenance is implemented in JAVA. The coding is done in JAVA with the cloudSim platform easily identifies the confidentiality level before implementing it in the real world scenario. Amazon S3 is a warehouse of data, such as images, files, and other types of useful information. The Amazon S3 is a reliable, fast, inexpensive data storage infrastructure for efficient query processing. Amazon S3 stores data objects redundantly on multiple devices diagonally on multiple services and permits simultaneous read and write access. The read-write access to these data objects helps to easily recover the needed information.

Amazon S3 based storage of files are discussed and used in our experimental discussions to identify the result percentage. The SKDC scheme compares the existing work with the Trusted Third Party Model (TTPM) [21] tasked with assuring specific security characteristics of TrustedDB [4]. The experiment is conducted on factors, such as confidentiality level, quality of service on cloud data storage zone, and query processing efficiency rate on dealing with cloud information.

**4.1. Results Analysis Of The SKDC Scheme**

To evaluate the confidentiality and privacy performance with SKDC scheme, two well-known privacy schemes are compared which are the TTPM [21] and TrustedDB [4, 5, 9, 20].

**4.1.1. Impact of Confidentiality Level**

The confidentiality level using the SKDC scheme is the amount of confidentiality provided by the cloud users to the cloud using Shamir’s key distribution model through polynomial interpolation points. From Table 1, it is clear that, the higher the level of confidentiality, the more effective the scheme.

Table 1. Tabulation for confidentiality level.

No. of Users (U)	Confidentiality level (%)		
	SKDC	TTPM	TrustedDB
3	58.35	46.32	38.28
6	61.45	49.42	41.38
9	68.33	56.30	48.26
12	62.89	50.86	42.82
15	71.35	59.32	51.28
18	69.88	57.85	49.81
21	80.25	68.22	60.182

**4.1.2. Impact of Quality of Service on the Cloud Data Storage Zone**

From Table 2, it is clear that, the quality of service on the cloud data storage zone using the SKDC scheme is the amount of privacy handled by the cloud. The quality of service in terms of privacy using the SKDC scheme is the product of horizontal level of privacy and the vertical level of privacy provided to the cloud

users by the cloud. It is measured in terms of percentage (%).

$$QoS_p = \sum_{i,j=1}^n Hor Privacy Query_i \cup Ver Privacy_j \quad (5)$$

Table 2. Tabulation for quality of service.

No. of Users (U)	Quality of Service (%)		
	SKDC	TTPM	TrustedDB
3	49.72	44.71	36.67
6	51.33	46.32	38.28
9	58.39	53.38	45.34
12	55.77	50.76	42.72
15	61.35	56.34	48.30
18	60.42	55.41	47.37
21	65.99	60.98	52.94

**4.1.3. Impact of Query Processing Efficiency**

The query processing efficiency rate when dealing with cloud information handles the number of successful queries handled by the cloud. From Table 3, it is clear that, the query processing efficiency  $QP_{Eff}$  using SKDC scheme is the ratio of queries successfully addressed as  $Queries_A$  by the cloud to the number of queries provided by the user  $Queries_P$  to the cloud. It is measured in terms of percentage (%).

$$QP_{Eff} = \frac{Queries_A}{Queries_P} \quad (6)$$

Table 3. Tabulation for query processing efficiency.

No. of Queries	Query Processing Efficiency (%)		
	SKDC	TTPM	TrustedDB
5	0.6	0.55	0.48
10	0.7	0.62	0.52
15	0.86	0.78	0.71
20	0.8	0.79	0.73
25	0.84	0.81	0.78
30	0.93	0.83	0.80
35	0.91	0.85	0.81

**5. Conclusions**

The SKDC Scheme processes the queries and maintains the privacy preserving confidential data services. Column proxy query processing operations are also carried out in the SKDC scheme ensuring query processing effectiveness using the matrix form. The SKDC scheme creates a polynomial of degree with the secret as the first coefficient and the remaining coefficients picked up at random to improve the privacy preserving level on the cloud infrastructure. The key is hidden from the public cloud users to improve the confidentiality rate. Shamir’s key distribution supports batch auditing where multiple user requests for data auditing is held concurrently at a higher confidentiality rate. Thus, in future, we can improve the key as a complex while it may increase the computational process.

**References**

[1] Abbadi M. and Ruan A., “Towards Trustworthy

- Resource Scheduling in Clouds,” *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 6, pp. 973-984, 2013.
- [2] Foto N., Borkar V., Carey M., Polyzotis N., and Jeffrey D., “Map-Reduce Extensions and Recursive Queries,” in *Proceedings of the 14<sup>th</sup> International Conference on Extending Database Technology*, Uppsala, pp. 1-8, 2011.
- [3] Ateniese G., Burns R., Curtmola R., Herring J., Kissner L., Peterson Z., and Song D., “Provable Data Possession at Untrusted Stores,” in *Proceedings of the 14<sup>th</sup> ACM Conference on Computer and Communications Security*, Alexandria, pp. 598-609, 2007.
- [4] Bajaj S. and Sion R., “TrustedDB: A Trusted Hardware-Based Database with Privacy and Data Confidentiality,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 3, pp. 752-765, 2014.
- [5] Bajaj S. and Sion R., “TrustedDB: A Trusted Hardware-Based Database with Privacy and Data Confidentiality,” in *Proceedings of the ACM SIGMOD International Conference on Management of Data*, Athens, pp. 205-216, 2011.
- [6] Castell S., “Codeo Practice and Management Guidelines for Trusted Third Party Services,” INFOSEC Project Report, 1993.
- [7] Ciriani V., Vimercati S., Foresti S., Jajodia S., Paraboschi S., and Samarati P., “Combining Fragmentation and Encryption to Protect Privacy in Data Storage,” *ACM Transactions on Information and System Security*, vol. 13, no. 3, pp. 13-22, 2010.
- [8] Damiani E., Vimercati S., Jajodia S., Paraboschi S., and Samarati P., “Balancing Confidentiality and Efficiency in Untrusted Relational DBMSs,” in *Proceedings of the 10<sup>th</sup> ACM Conference on Computer and Communications Security*, Washington, pp. 93-102, 2003.
- [9] Hashizume K., Rosado D., Fernández-Medina E., and Fernandez E., “An Analysis of Security Issues for Cloud Computing,” *Journal of Internet Services and Applications*, vol. 4, no. 1, pp. 38-46, 2013.
- [10] Hubbard D. and Sutton M., *Top Threats to Cloud Computing v1. 0*, Cloud Security Alliance, 2010.
- [11] Hudic A., Islam S., Kieseberg P., Rennert S., and Weippl E., “Data Confidentiality Using Fragmentation in Cloud Computing,” *International Journal of Communication Networks and Distributed Systems*, vol. 1, no. 3-4, pp. 325-329, 2012.
- [12] Sakthivel A., “Enhancing Cloud Security Based on Group Signature,” *The International Arab Journal of Information Technology*, vol. 14, no. 6, pp. 923-929, 2017.
- [13] Shacham H. and Waters B., “Compact Proofs of Retrievability,” in *Proceedings of International Conference on the Theory and Application of Cryptology and Information Security*, Melbourne pp. 90-107, 2008.
- [14] Sugumaran M., Murugan B., and Kamalraj D., “An Architecture for Data Security in Cloud Computing,” in *Proceedings of World Congress on Computing and Communication Technologies*, Trichirappalli, pp. 252-255, 2014.
- [15] Sugumaran M., Murugan B., and Kamalraj D., “An Architecture for Data Security in Cloud Computing,” in *Proceedings of the International Conference on Information Technology and Applications*, IEEE Computer Society, Trichirappalli, pp. 252-255, 2013.
- [16] Wang C., Cao N., Ren K., and Lou W., “Enabling Secure and Efficient Ranked Keyword Search Over Outsourced Cloud Data,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 8, pp. 1467-1479, 2012.
- [17] Wang C., Chow S., Wand Q., Ren K., and Lou W., “Privacy-Preserving Public Auditing for Secure Cloud Storage,” *IEEE Transactions on Computers*, vol. 62, no. 2, pp. 362-375, 2013.
- [18] Wang Q., Wang C., Ren K., Lou W., and Li J., “Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 22, no. 5, pp. 847-859, 2011.
- [19] Wu S., Li F., Mehrotra S., and Ooi B., “Query Optimization for Massively Parallel Data Processing,” in *Proceedings of the 2<sup>nd</sup> ACM Symposium on Cloud Computing*, Cascais, 2011.
- [20] Zheng Q., Xu S., and Ateniese G., “Efficient Query Integrity for Outsourced Dynamic Databases,” in *Proceedings of the ACM Workshop on Cloud Computing Security*, Raleigh, pp. 71-82, 2012.
- [21] Zissis D. and Lekkas D., “Addressing Cloud Computing Security Issue,” *Future Generation Computer Systems*, vol. 28, no. 3, pp. 583-592, 2012.



and database.

**Kamalraj Durai** received his MCA degree from St. Joseph's College, Trichy during 2007 and currently a research scholar in Bharathiar University, Coimbatore, India. His research interest includes parallel and distributed computing,