Volume 6, Issue 4, October 2019 ISSN: 2348-8565 (Online)

International Journal of Modern Engineering and Research Technology

Website: http://www.ijmert.org Email: editor.ijmert@gmail.com

# A Theoretical Framework for Data Sharing Over Cloud Using Matrix Factorization Technique

**Roshni Mishra**
*M.Tech. Research Scholar*
*Shri Ram Group of Institution*
*Jabalpur, (M.P.) [INDIA]*
*Email: kanakmishra2011@gmail.com*

**Sapna Choudhary**
*Associate Professor*
*Department of Computer Science & Engineering*
*Shri Ram Group of Institution*
*Jabalpur, (M.P.) [INDIA]*
*Email: choudharysapnajain@gmail.com*

## ABSTRACT

*In this work we have developed an efficient theoretical framework for cloud data sharing scheme using matrix factorization technique. The proposed theoretical framework is simple and robust. Also it is able maintain integrity of the data at different geographical locations.*

***Keywords:***—*Matrix Factorization, Non Negative Matrix Factorization, Singular Value Decomposition, Cloud Data Security, Theoretical Framework.*

## I. INTRODUCTION

While cloud storage is advantageous and gives workers access to their information anyplace, whenever, on almost any gadget, cloud storage security is a best worry for associations' IT and security divisions. The advantages brought by cloud storage - from versatility and openness to diminished IT overhead – are driving quick selection at ventures far and wide, and there are steps that organizations should take to enhance cloud storage security and keep touchy information sheltered and secure in the cloud.

Organizations and ventures utilize cloud administrations since they give financially savvy and adaptable options in contrast to costly, privately executed equipment. In any case, leading business in the cloud implies that classified records and delicate information are presented to new dangers, as cloud-put away information dwells outside of the breaking points of numerous shields used to ensure touchy information hung on-preface. All things considered, undertakings must take extra measures to anchor cloud storage past the occasionally essential assurances offered by suppliers.

The ascent of Internet of Things (IoT) innovation and the associated office has likewise made undertakings more dependent on cloud innovation, but while driving security dangers. Indeed, even brilliant printers have been discovered powerless against information spillage, and as more corporate gadgets move toward becoming web associated, the potential for bargain or unintended spillage increments.

As undertakings move further along the cloud selection bend, cloud storage security is turning into a best need – both in endeavors' IT design and data security methodologies. Organizations currently perceive that it's basic to secure delicate information while empowering representatives to appreciate the execution and adaptability of the cloud.

Cloud storage suppliers and undertakings share obligation regarding cloud storage

security. cloud storage suppliers execute standard assurances for their stages and the information they process, such verification, get to control, and encryption. From that point, most endeavors supplement these insurances with included safety efforts of their own to support cloud information assurance and fix access to delicate data in the cloud.

## II. Related Work

In the past decade a number of cloud data security schemes have been proposed by various group of researchers [1]-[5].

Manivannan et al **[1]**, have proposed an secured mechanism for encryption of database known as transposition, substitution, collapsing, and shifting (TSFS) method with three key. At the point when the quantity of keys expanded, at that point the handling and method may likewise increment. Encryption of database especially the information put away in the database should be scrambled which improves the protection and security of information put away in cloud. For shared transmission, synchronizer used to assemble all the keys and the customer framework gets the key from the synchronizer which decode the conveyed information that is as of now encoded.

A three-layer framework structure is proposed by N. Jose et al **[2]** in which each layer plays out its very own obligation to guarantee the information security of cloud layers. Verification of client is done in first layer and the information encryption by utilizing AES method is done in second layer. First layer utilizes get to control apparatuses to check for approved client and to limit unapproved access to clients information. The third layer underpins the quicker client information recuperation by utilizing Byzantine adaptation to non-critical failure method techniques.

In these model, an encryption and unscrambling method for give security to that information is actualized by G. Vandana et al **[3]**. Additionally give an additional security layer we utilize the client area and land position. To give this, we require Anti-parody GPS which is gives exceptionally precise area of the client for getting to information and it can give us the scope, longitude and elevation precisely This technique can be valuable for some applications, for example, banks, huge organizations, foundations, and so forth.

S. Ajoudanian et al **[4]**, proposes structure comprising of four layers. First layer is virtual machine layer which is the fundamental security layer. cloud storage is the second layer of cloud computing and it has a capacity structure which joins different assets from numerous outside cloud specialist co-ops, consequently it can fabricate a substantial virtual capacity framework. The fourth layer is virtual system screen layer which is a blend of equipment and programming segments in the virtual frameworks. The cloud security collusion supplier ought to clarify the systems and ought to likewise guarantee that approved clients can just access the mists.

As detailed by A.E. Youssef et al **[5]**, proposed a security demonstrate for cloud computing which enhances the protection issues in cloud and it can likewise shield cloud from vulnerabilities. It comprises of different units, for example, check and approval, benefit control, information assurance, assault recognition/aversion unit. The working of first layer is to approve the information precision, uprightness and shared assets in the cloud however it performs verification of cloud clients. Benefit control security unit is important to control cloud utilization by Different people and associations. It secures client's protection and guarantees information

trustworthiness and classification by applying a gathering of guidelines and strategies that control that has the specialist to do what on the cloud. The last most layers in this model is identification and aversion unit which recognize assaults and vindictive insiders for information get to and furthermore avert undesirable modules being introduced without the learning of specialist co-op. It expands the security framework inside the cloud condition.

## III. The Proposed Theoretical Framework

In this work we have proposed a new theoretical framework for cloud data security. The proposed theoretical framework is based on matrix factorization technique. The proposed theoretical framework is divided into several steps and different types of security mechanism is provided in each step for robust data security.

For simplicity we have divided the proposed theoretical framework into two parts (a) the forward method (FM) and (b) the backward method (BM). In the FM (see Figure 1) the given file is processed in the way that its various sub-parts are generated and distributed in the different geo-graphical locations. While in the BM (see Figure 2) the various parts are taken from different geo-graphical locations and reverse processing is applied to form the original image.

### *A. The Forward Method*

1. The MATLAB provides different ways of reading different types of files. So in the proposed FM we assume that the data has already been read and it is placed in the clipboard of the computer.

2. In the next step the data is fetched from clipboard and it is converted into byte array. Also information about file is stored in the database.

3. In this step the byte array obtained in the previous step is encrypted and then converted into the matrix. At this point we have many choice of binary encryption algorithms. We have selected the Binary RSA algorithm proposed in **[14]**. Now we convert the encrypted binary array to a equivalent matrix.

4. In this step the matrix is factored into three sub-matrix using the singular value decomposition technique **[13]**. Now the sub-matrices are stored in locations

5. At last the system displays the unique ID of the file for future use.

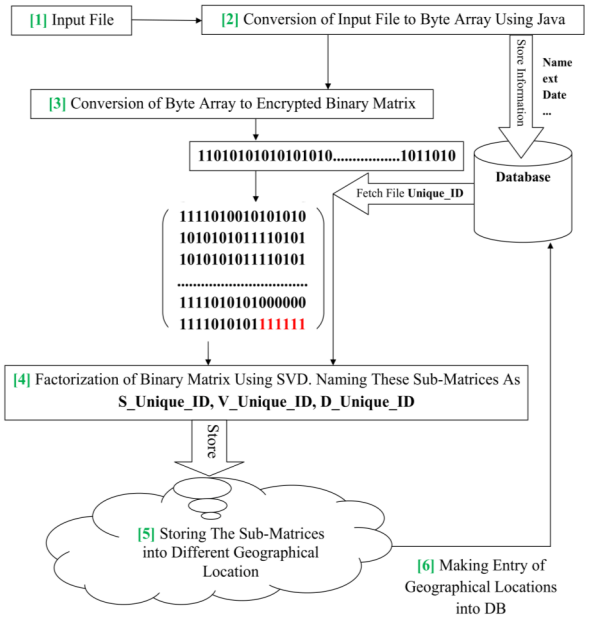


*Figure 1: Flowchart of the Forward Method.*

### *B. The Backward Method*

This method contains all the working of the FM method but in reverse direction. In first step the client sends the unique-ID of the file to the method. For this the client needs to copy the unique ID in the computer. The

algorithm will fetch the unique ID from clipboard.

1. Next the cloud server finds the information about the files from the saved locations.
2. Now these sub-matrix are converted to the MATLAB readable files and then the reverse SVD method is applied to form the original binary encrypted matrix from the sub-matrices S, V, and D.
3. Now the binary matrix is processed to form the binary array. After that the Binary RSA method is applied to decrypt the binary matrix.
4. At last the decrypted binary array is converted into equivalent text, and the text is placed in the clipboard.

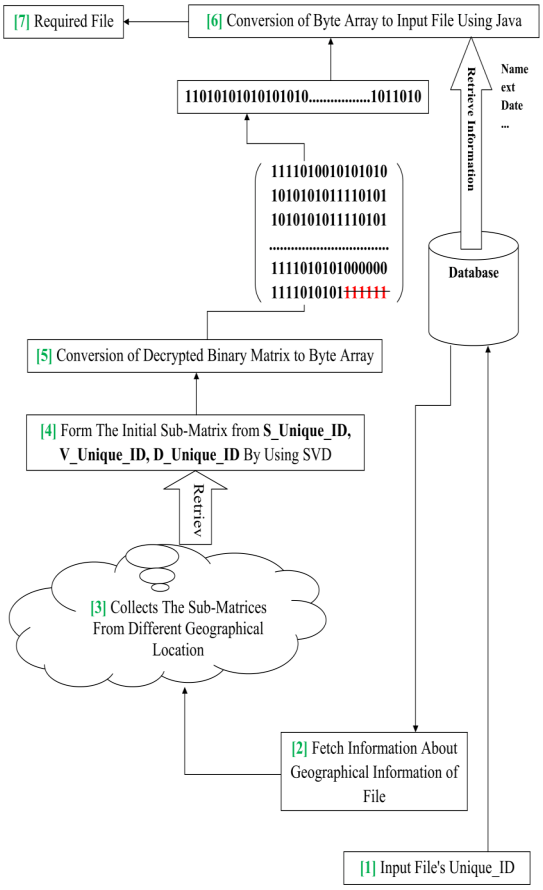


*Figure 2: Flowchart of the Backward Method.*

## IV. Security Analysis of The Proposed Theoretical Framework

By the basic assumption that the client does not provide his/her login information and file's unique ID to anyone else. We are going to analyze the performance of the proposed theoretical framework.

1. If someone is able to access one part of file which is placed in some geo-graphical location then he cannot recreate the file. This is because formation of other two factored matrix in the SVD matrix factorization method is not an easy task.
2. In the worst case if someone has access of all the parts of the file (this situation is itself next to impossible). Then also he/she will face trouble in the formation of decrypted binary array from the encrypted matrix. Formation of array by discarding the last row and accessing meaning full contents from the second last row is not an easy task at all. In this stage one wrong step may corrupt the whole information.
3. At last step for creating file from the binary array the user must have knowledge about extension of the file so that the data may be written in the file efficiently. It is noted that binary array contains header information of the file and storage method for this header information is different for different extensions.

## V. Simulation

We have developed a theoretical framework for cloud data sharing. For simplicity the proposed theoretical framework is divided into two sub methods namely FM and BM. Here we have also explained this framework in detailed and we have shown the necessary programming steps for

helping the developers in the implementation of the proposed theoretical framework. The robustness of the proposed theoretical framework is also discussed in this paper. We invite developers from various research field to implement and use the efficiency of the proposed theoretical framework in the practical scenarios.

As it is mentioned several times that we have proposed a theoretical framework. However just to check the robustness of the proposed theoretical framework in this paper we are performing a small simulation. In this paper we have used the proposed theoretical framework and then we have implemented in the MATLAB. As it is implemented in MATLAB hence we are assuming that the data is fetched from the file and placed in the clipboard of the system. If the same algorithm is implemented in the other programming language then the code must be done in such a way it should be capable in reading the data directly from file.

***Directory Structure***

The directory structure contains four sub-folders namely:

- code_image: this contains MATLAB code of the proposed theoretical framework for images.
- code_textfile: this contains MATLAB code of the proposed theoretical framework for text.
- data: this contains the input data for the proposed theoretical framework
- save_data: this will contain the results generated by the proposed theoretical framework.

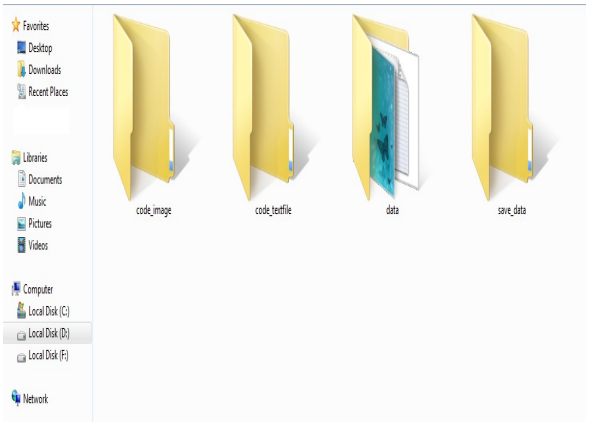


*Figure 3: Directory structure for the proposed theoretical framework.*

***B. For Text***

As mentioned earlier that we have developed two different version of the proposed method (text and image). At first we are going to discuss the text implementation of the proposed theoretical framework. The MATLAB does not provide a single way to read data from all type of text formats. Hence for simplicity the proposed framework is designed by taking an assumption that the user opens the file (see Figure 3) and copies all the text that is contained in the file. Now this copied text is placed in the clipboard by the operating system. From this point processing of the proposed text algorithm starts. Now we are going to understand this through simple steps.

1. The data is fetched from the clipboard and stored in a variable **str**.
2. Next the data present in variable **str** is pre-processed so that it can be easily processed by the MATLAB environment. The results are stored in a variable **text**.
3. Next the forward method **forward_method_text** is called and the variable **text** is passed in it for the processing.

4. Inside the function **forward_method_text** the text data present in the variable **text** is converted into binary form and then stored in a vector variable named **binary_text**.

5. Length of data present in the variable **binary_text** is calculated and then it is processed in such a way that the vector **binary_text** can be converted into equivalent matrix.

6. Next this matrix is converted in the form of three factored matrices using **singular value decomposition** method.

7. Now these three matrices are stored in some **pre**-defined locations in .txt file (see Figure 4).

8. Next the current date and time string is fetched using MATLAB's built-in function and then this is processed so that the results can be used as a key value.

9. The key value is displayed in the screen so that the user can note it for future use.

10. In the backward method the user simply copies the key value and then runs the method run_code_b_text.m.

11. The run_code_b_text.m method fetches the key value from clipboard and then fetches the shares from the desired locations and process them to produce the text.

12. After the algorithm finishes its working, the text is displayed in the MATLAB's command window.

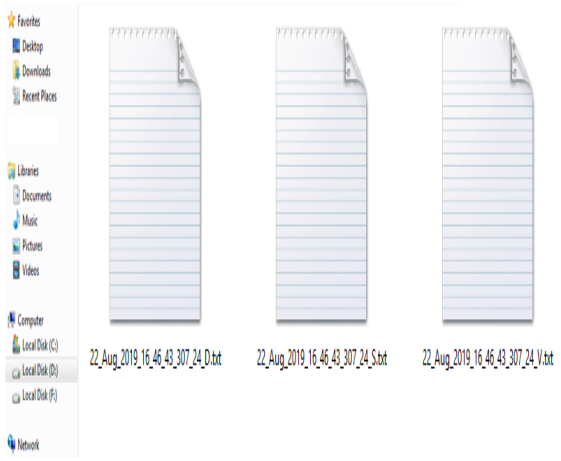


*Figure 4: The shares generated by the proposed theoretical framework [text method].*

### *C. For Image*

As mentioned earlier that we have developed two different version of the proposed method (text and image). Now we are going to discuss the image implementation of the proposed theoretical framework. The MATLAB provides many ways to read data from all type of image file formats. The image implementation of the proposed theoretical framework is very simple. Here just the image (already in matrix form) is processed using **singular value decomposition**. The results are stored in the MATLAB's defined format **.mat** file.

## VI. Conclusion

We have developed a theoretical framework for cloud data sharing. For simplicity the proposed theoretical framework is divided into two sub methods namely FM and BM. Here we have also explained this framework in detailed and we have shown the necessary programming steps for helping the developers in the implementation of the proposed theoretical framework. The robustness of the proposed theoretical framework is also discussed in this chapter. We invite developers from various research field to implement and use the efficiency of the proposed theoretical framework in the practical scenarios.

**REFERENCES:**

[1] D. Manivannan and R. Sujarani, "Light weight and secure database encryption using TSFS Algorithm," Proceedings of the International Conference on Computing Communication and Networking Technologies, IEEE, 2010.

[2] N. Jose, C. Kanmani, "Data security mode enhancement in cloud environment," Journal of Computer Engineering Volume 10, Issue 2 (Mar. - Apr. 2013).

[3] G. Vandana et al, "Improve security of data access in cloud computing using location," International Journal of Computer Science and Mobile Computing, Vol.4 Issue.2,February-2015.

[4] S. Ajoudanian and M.R. Ahmadi, "A Novel data security model for cloud computing," International Journal of Engineering and Technology, Vol.4, No.3, June 2012.

[5] A.E. Youssef and M. Alageel, "A framework for secure cloud computing," International Journal of Computer Science Issues, Vol. 9, Issue 4, No 3, July 2012.

[6] H. Bensmail, G. Celeux, "Regularized gaussian discriminant analysis through eigenvalue decomposition." J. Am. Stat. Assoc. 91(436), 1743–1748 (1996).

[7] P.M. Aiswarya, A. Raj, D. John, L. Martin and G. Sreenu, "Binary RSA encryption algorithm," 2016 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), Kumaracoil, 2016, pp. 178-181.

[8] JAMA The Package, Version 1.0.3 (November 9, 2012), https://math.nist.gov/javanumerics/jama/doc/

[9] Jama/SingularValueDecomposition.java at master fiji/Jama GitHub https://github.com/fiji/Jama/blob/master/src/main/java/Jama/SingularValueDecomposition.java

[10] https://www.online-toolz.com/tools/text-binary-convertor.php

[11] http://www.bluebit.gr/matrix-calculator/calculate.aspx

* * * * *