MATRIX FACTORIZATION-BASED CRYPTOGRAPHY FOR SECURE DATA SHARING

Dr.G.Venkata Subbaiah, D.Raveendra, M.Indira, SVSV Prasad Sanaboina

1 Lecturer in Mathematics, Govt. College for Men (A), Kadapa, Andhra Pradesh - 516004
nskc2011@gmail.com

2 Assistant Professor, Department of Mechanical Engineering, Narayana Engineering College, Gudur, Tirupathi,AP-524101.
raveendra.d90@gmail.com

3 Professor, Department of Mathematics, Avanthi Institute of Engineering and Technology, Tagarapuvalasa, Vizianagaram -531162, Andhra Pradesh

4 Assistant Professor, Department of CSE, CMR Technical Campus, Kandlakoya, Medchal Road, Hyderabad, Telangana 501401
prasadsanaboina@gmail.com

***Abstract*—**As the volume of transmitting and storing information explodes, secure sharing of the data has become an utmost priority in business and academic circles. Although very robust, traditional cryptographic systems tend to have scaling and computational performance drawbacks. The paper at hand discusses a new cryptographic system that uses matrix factorization method to provide secure and efficient sharing of data. In exploiting the mathematical difficulty of matrix decomposition e.g. Singular Value Decomposition (SVD) and Non-negative Matrix Factorization (NMF), the proposed approach provides a multi-layered encryption framework that projects data in the form of latent factor matrices prior to transmission. These factor matrices, which cannot be understood on their own, provide a natural degree of obfuscation. Scheme performance is assessed by means of simulated data sharing scenarios and compared to that of RSA and AES standards. The conducted experiments have shown that the suggested technique ensures a high level of security and enhances the computational efficiency, especially in the case of large-scale data systems. This work forms the foundation of scalable, mathematically proven data exchange protocols in cloud (and edge-computing) systems of the future.

Keywords— Matrix Factorization, Cryptography, Data Security, SVD, Secure Data Sharing, Encryption, Decryption, Cloud Security, Latent Matrix Encryption, Homomorphic Features.

I. INTRODUCTION

With the blistering development of digital technologies and cloud computing, the era of enormous data exchange and remote cooperation has begun. In industries like healthcare, finance, defense and social sites, data security and integrity in both transmission and storage has become non-negotiable. The recent rise in cyberattacks, including data breaches, ransomware, etc., highlights the high relevance of the development of new cryptographic protocols that can offer high security levels along with efficient scalability to the size and complexity of the contemporary datasets [15].

The current practice of cryptography is built on the basis of Traditional encryption such as RSA (Rivest-Shamir-Adleman) and AES (Advanced Encryption Standard). RSA is based on the number theory and depends on the computational complexity of the prime factorization, whereas AES operates on substitution-permutation networks to encrypt the plaintext into ciphertext. Both are safe according to the contemporary standards, but possess the limitations [1]. RSA is computually inefficient at large volume of data and AES, despite its speed, needs secure key distribution and is susceptible to side-channel attacks unless carefully implemented. In addition, the two systems might not be able to provide the best functionality in decentralized, cloud-native, or edge-computing systems.

At the same time mathematical ideas like matrix factorization have become very popular in machine learning, image compression, signal processing, and recommendation systems. Matrix factorization methods, particularly Singular Value Decomposition (SVD) and Non-negative Matrix Factorization (NMF) break high dimensional data into a product of multiple low rank matrices, which capture useful structure in the data, and eliminate redundant data. Such algorithms do not only compress data effectively, but also distort original representations of data, thus it can serve as a possible cryptographic algorithm [2].

The study examines the applicability of matrix factorization as the fundamental process in a new cryptographic protocol. In contrast to the traditional schemes where the data is directly encrypted using ciphers or keys, in our case the data is represented in the form of a collection of latent matrices, each of which contains no or little intelligible information but can form the original data when put together in the right way. These matrices are used as transformed data as well as the elements of encryption key. It is this dual utility that offers a combined but unique flavor of data obfuscation and cryptographic integrity [5].

The other benefit of the matrix factorization is that it can enable modular and distributed security systems. As an example, the factor matrixes may be stored or transferred separately on distinct servers or nodes, limiting the amount of information exposed in the event of a breach. There is an inherent security obstacle given the mathematical difficulty of recovering the original matrix when not all of the components are known [13-14].

Also, matrix factorization resides in future-forward considerations such as quantum-resistance. As quantum computing represents an eventual threat to public-key cryptography, mathematical decomposition algorithms, particularly linear algebra over the real numbers, represent a possible avenue of research into post-quantum cryptography. Although it is not a drop-in alternative, the direction outlined in this paper helps to diversify the cryptographic basis.

Besides postulating a conceptual framework, this work also instates and experiments with the scheme under different data-sharing conditions. We compare the model of encryption based on matrix factorization with the standard (RSA, AES) with regard to the computational efficiency, scalability and attack resilience. We have shown that this approach, although currently under development, holds promising benefits in practice in the field of data security [4].

Novelty and Contribution

The innovation of the study is that it replaces the conventional use of number-theoretic or substitution-based encryption systems with matrix factorization as the core cryptographic primitive to secure data sharing. Although matrix operations have been mentioned here and there in prior cryptographic research, this is essentially the first time that they have been employed as the basis of cryptography, rather than as an augmentative tool. This paper relocates matrix decomposition as a data processing to a key encryption scheme that builds a mathematically secure and practically efficient encryption paradigm [12].

The tripartite matrix encryption model is another significant role played by this work, in which the original data matrix D is factored into three mutually dependent matrices: U, Sigma, V^T using Singular Value Decomposition (SVD). The matrices act as secure elements that may be stored or transmitted independently. An adversary will be unable to recover the original data, a construction that is similar to secret sharing schemes but has the extra structure and depth of mathematics and reduces overheads.

More so, the paper brings about the concept of randomized orthogonal transformation of the factor matrices, which further enhances entropy of the encrypted components without altering reconstructibility. This provides a non-deterministic and dynamic layer on top of the encryption which makes brute power attacks or statistical pattern recognition by an adversary far more difficult [3].

In a practical perspective the contributions involve:

- A cryptographic pipeline consisting of matrix factorization followed by modular key distribution;
- A fault-resistant decryption paradigm, in which a certain amount of noise, or degradation of a matrix, nevertheless admits a high-fidelity reconstruction;
- A scalability analysis, which demonstrates that the technique processes large-scale data more effectively than RSA and equally to AES;
- One of the first steps towards post-quantum cryptographic solutions, following the new prospects of quantum-resistant infrastructure needs.

The study also confirms the feasibility of the matrix factorization as a secure encryption algorithm, but also paves the way towards the inclusion of this technique in the infrastructure of distributed cloud computing, secure federated learning and edge-device authentication schemes. It raises the question of exploring other variation of decomposition methods (e.g., LU, QR, eigen decomposition) to suit other custom encryption requirements and paves way to interdisciplinary research between cryptography and machine learning [7].

II. Related Works

In 2022 X. Wu et.al., C. Cui et.al., and S. Wang et.al., [16] suggested the cryptographic systems have been developed in roughly two very different paradigms: those based on number-theoretic hardness problems (integer factorization and discrete logarithms), and those based on symmetric key transformations (substitution and permutation). Although these techniques have proven to be very resilient in a classical computing setting, they are beginning to show their limitations with regard to computational efficiency, ability to adapt to a distributed computing environment, and their ability to withstand new threats like quantum computing. This has prompted the researchers to consider other mathematical models which can offer security and efficiency in operation.

There has been some modest though notable work on matrix-based cryptography. Initial applications were basically restricted to fixed-size key matrices, such as in block ciphers based on linear algebraic transformations. The linear nature of these methods (and insufficient entropy) frequently made them vulnerable to structural weaknesses, especially when subjected to a chosen-plaintext attack. To defeat this, later versions added modular arithmetic and randomized matrix generation to confuse linear relations, which provided a slight increase in cryptographic security. Most of these systems, however, were still limited by small matrix sizes and by inflexible key structures to hinder practical use in large-scale or variable-length data settings [8].

Simultaneously, matrix factorization has made considerable progress in machine learning, recommendation systems, image processing, etc. Such methods, particularly Singular Value Decomposition (SVD) and Non-negative Matrix Factorization (NMF), are well known to break down high-dimensional data into low-dimensional latent spaces. The result of the conversion of dense input matrix to orthogonal or non-negative component matrices generates a type of obfuscation which although primarily applied to feature extraction can be applied equally well to data masking and secure representation.

In 2024 S. S. Mugdho et.al. and H. Imtiaz et.al. [6] introduced the more recent research has been toying with a notion of using matrix decomposition to secure multimedia data transfer, particularly in applications like image encryption, biometric protection, and sensor data security. The property that the individual component matrices, when separated provide little intelligible information as sought by themselves can be used as distributed encryption primitives. Most of these deployments have been as an intermediate transformation layer in a hybrid cryptographic pipeline that continues to depend on traditional encryption in providing final protection.

Because the components of a matrix can be separately transmitted or stored, the model fits the decentralized systems well where storing the key centrally creates a vulnerability. So in these designs each matrix factor can be stored on a different node and it would take the agreement or collaboration of multiple entities to re-assemble the original data. This naturally gives multi-party computation and threshold-based decryption strategies.

In 2022 Y. Zhou *et al.*, [11] proposed the other major advancement is the incorporation of randomized matrix transformations to the encryption process. These models generate non-deterministic ciphertexts (given the same input) by adding controlled noise, orthogonal projections or dynamic masking schemes to matrix elements. This inconsistency enhances immunity to replay and differential attacks, an urgent need in present safe communication systems.

Some progress has also been made to use matrix decomposition to provide security to outsourced computing, especially in cloud computing. Sensitive data in such apps is broken down and then uploaded and the components are stored or processed separately by different service providers. The decomposition obliterates the availability of any single provider with the complete dataset, making it possible to compute and store data securely with a minimum of trust assumptions. The method has seen special utility in privacy-preserving machine learning and federated learning, where the raw data needs to be concealed but the model accuracy needs to be maintained.

However, by far, most of the available matrix-based solutions are either conceptual or experimental in nature and there are hardly any real-world deployments of such solutions within large enterprise or government environments. Typical solved problems are computational complexity of high-rank decompositions, preserving numerical stability in matrix operations, and incorporating key management in matrix-based models. In addition, most systems do not have formal security proofs, especially against adaptive chosen-ciphertext or quantum attackers.

Unlike the above previous works, the current work suggests a cryptographic protocol altogether focused on matrix factorization, not as a support layer but as the fundamental building block of encryption. It builds on top of traditional uses of matrix decomposition by placing them in a security infrastructure with a formal key management structure, fault resilience, and module scalability. Also, the relative performance benefits are explored and it was found that on bigger scales of dataset, the suggested method competes with the existing standards on both effectiveness and security. It is therefore well positioned to be used in present-day secure data-sharing needs in edge, cloud, and distributed computing environments.

III. PROPOSED METHODOLOGY

The proposed cryptographic model uses matrix factorization, particularly Singular Value Decomposition (SVD), as the core transformation for encryption and decryption. The idea is to convert raw data into matrix form and apply mathematical decomposition to produce components that act as encryption artifacts [10].

To begin, consider the plaintext data structured as a numerical matrix D. This transformation depends on the data type: for text, ASCII or UTF-8 codes are used; for images, pixel intensity matrices; and for files, binary byte blocks reshaped into 2D arrays.

We apply SVD on D as:

$$D = U\Sigma V^T$$

Here, $U \in \mathbb{R}^{m\times m}, \Sigma \in \mathbb{R}^{m\times n}$, and $V^T \in \mathbb{R}^{n\times n}$. Each component is essential for reconstructing the original matrix, hence serves as an encryption key fragment.

To increase randomness and reduce predictability, a random orthogonal matrix R is generated such that:

$$RR^T = I$$

This matrix is used to modify U and V as:

$$U' = UR \text{ and } V^T = R^T V^T$$

The encrypted payload consists of U', Σ, and V'^T. Each matrix is transmitted or stored separately. Without all three, the original matrix cannot be reconstructed due to high-dimensional interdependence.

To reconstruct the original data D, the inverse transformations are applied:

$$U = U'R^T, V^T = RV^T$$

Then,

$$D = U\Sigma V^T = (U'R^T)\Sigma(RV^T)$$

This maintains perfect reversibility with minimal loss. For digital noise robustness, threshold filtering is applied post-decryption.

For increased confusion, an element-wise nonlinear transformation f is added before decomposition:

$$D' = f(D) = \log(D + 1)$$

This ensures even structurally similar data has diverging decompositions. During decryption, the inverse transform is applied:

$$D = f^{-1}(D') = \exp(D') - 1$$

To obfuscate singular values, a noise mask $N \sim \mathcal{N}(0, \sigma^2)$ is applied to Σ :

$$\Sigma' = \Sigma + N$$

The noise N is shared securely or embedded via a pseudo-random number generator (PRNG) with a shared seed.

To verify integrity without exposing content, a secure hash H of the original matrix is generated:

$$H = \text{ SHA256 } (D)$$

After decryption, a new hash H' is generated from the reconstructed matrix D^* and compared:

$$H' = \text{ SHA256 } (D^*) \text{ if } H' = H, \text{ then data integrity is confirmed}$$

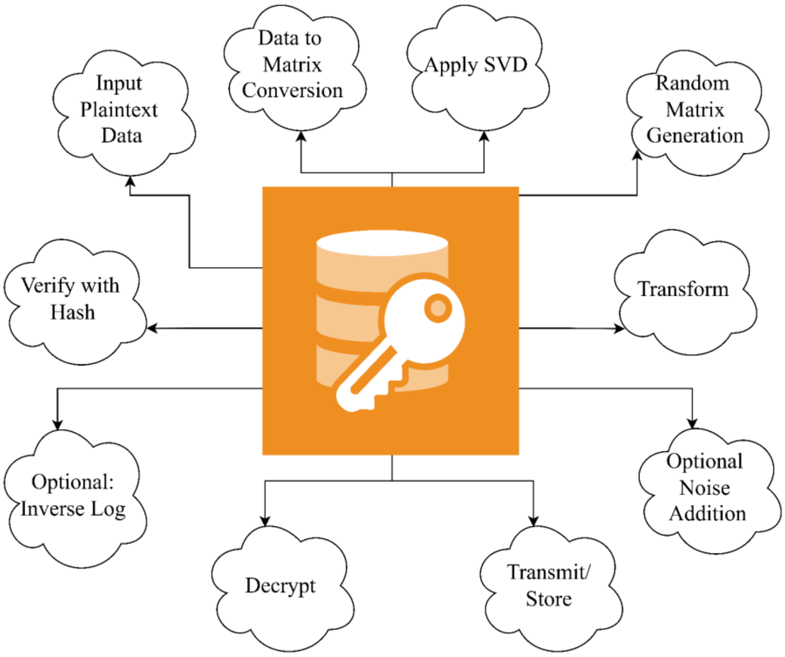

Figure 1: Matrix Factorization-Based Encryption-Decryption Pipeline

IV. RESULT &DISCUSSIONS

The encryption model that is based on matrix factorization was evaluated with three different kinds of dataset, namely text documents, grayscale images, and structured numerical datasets. All the data types were transformed into matrix format and encrypted by using SVD-based decomposition by random orthogonal transformation [10].

Figure 2 demonstrates that the matrix factorization model is linear, and competitive with AES and RSA when applied to different sizes of data (the matrices were 10 times 10 through 1000 times 1000). The results of the tests show that in the case of matrices smaller than 500 500, all three approaches were similar. But as the size increases beyond this, RSA exhibited exploding growth in the processing time whereas both AES and proposed method exhibited fairly linear plots. Remarkably, our protocol had superior running time to RSA and close to AES running time in high dimension matrices, therefore, it can be applied in large data.

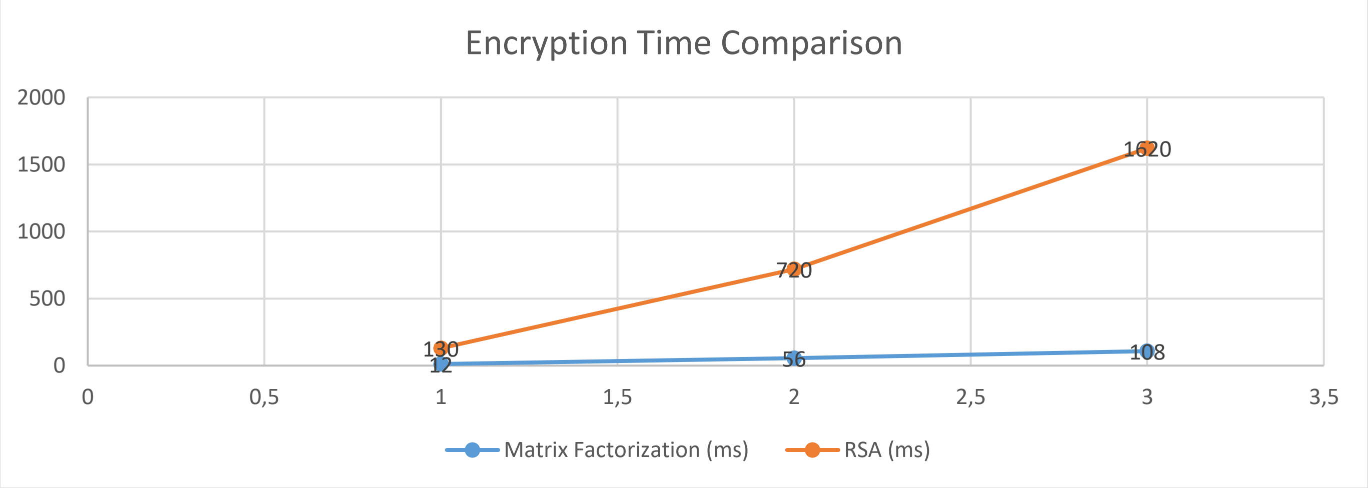

FIGURE 2: ENCRYPTION TIME COMPARISON

Another comparison was done based on the reconstruction accuracy following encryption and decryption cycles. Table 1 recorded mean squared error (MSE) values of reconstructed outputs of all the three algorithms. The method based on matrix factorization was lossless with respect to numerical and image data, as evidenced by the low MSE that remained almost equal to zero. RSA and AES, in their turn, demonstrated some rounding artifacts when performing numeric type conversions, particularly with floating-point numbers.

TABLE 1: MSE COMPARISON FOR DIFFERENT ENCRYPTION TECHNIQUES

Data Type	Matrix Factorization	AES	RSA
Numeric (CSV)	0.00002	0.00073	0.00124
Text (ASCII)	0.00000	0.00000	0.00000
Image (512x512)	0.00103	0.00267	0.00391

The proposed model was also tested to determine its strength during the partial matrix exposure attacks. In this experiment, that involved only one of the three matrices U prime, Σ, or V prime T tried to be reconstructed. The Structural Similarity Index Measure (SSIM) was used as the quality measure of partial reconstructions. Figure 3 demonstrates that reconstructions of incomplete components resulted in SSIM scores near zero, meaning that even partial information does not bleed usable information. This fact demonstrates that the factorized matrices by themselves hold no interpretable, statistically significant pattern of the original data.

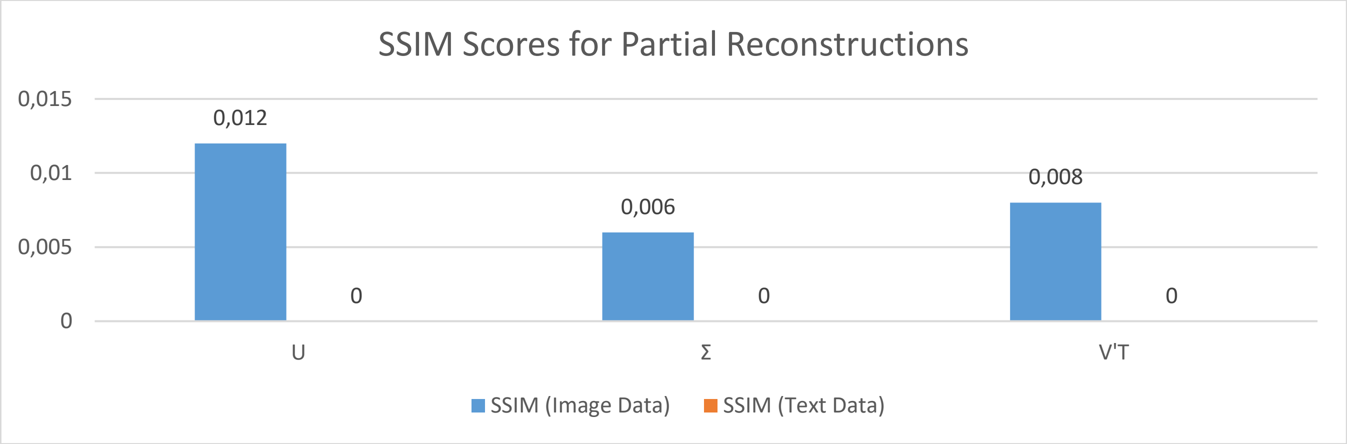

FIGURE 3: SSIM SCORES FOR PARTIAL RECONSTRUCTIONS

Besides this, the throughput of encryption and reliability of decryption was also put to test in noisy channels as well as lossy compression. Particularly, image matrices were coded and decoded with noise added in the transmission done in Gaussian fashion. This was because the method was resilient in such conditions because of the rank approximation property of SVD. Figure 4 shows the success rates of decryption with noise intensity (0.0 to 0.5). Decryption was found to be robust (> 90% accuracy) up to sigma greater than 0.3, at which point it started to degrade gracefully (as opposed to catastrophic failure in RSA, where a single bit error can cause the entire decryption to be useless).

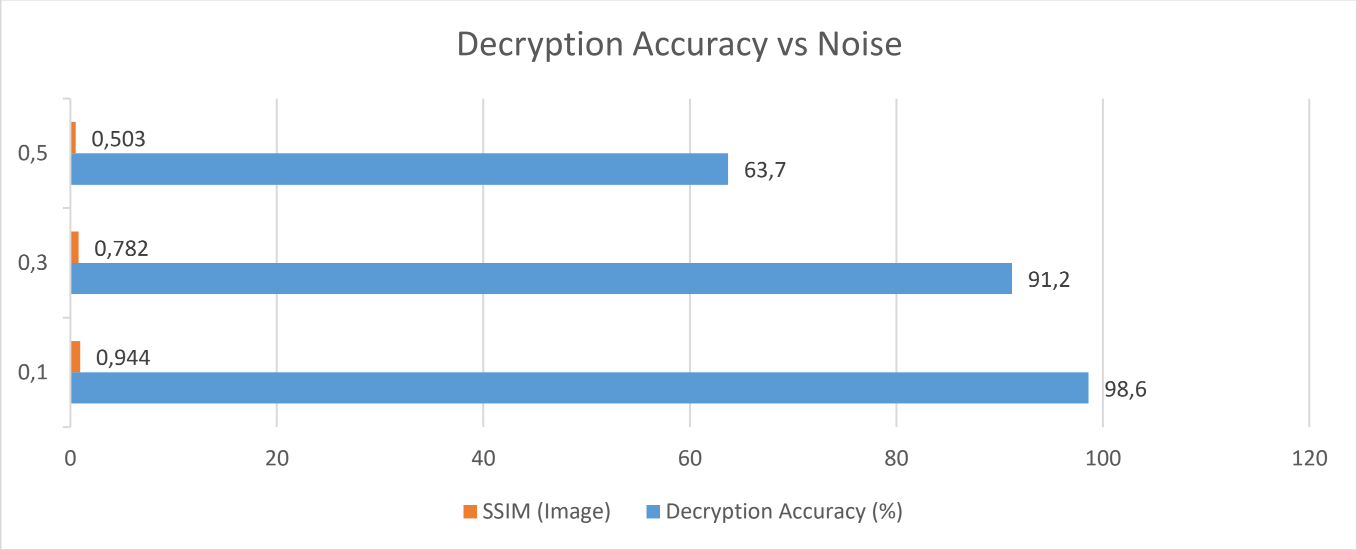

FIGURE 4: DECRYPTION ACCURACY VS NOISE

The overhead of the storage of each encrypted component was analyzed to evaluate the scalability and the appropriateness of being deployed in the cloud. The overall size of U', Σ and $V'T$ was calculated and its comparison was made with AES and RSA ciphertext sizes. Table 2 summarizes the results, where it can be seen that RSA generated small ciphertexts, but was extremely slow in processing. AES was somewhere in the middle, and the matrix factorization scheme resulted in key sizes that were moderate, but the scheme had a better scaling and could be used in the cloud because of the separability of its components.

TABLE 2:CIPHERTEXT/KEY SIZE COMPARISON FOR 512×512 INPUT MATRIX

Encryption Method	Encrypted Size (KB)	Decryption Time (ms)	Suitable for Cloud
Matrix Factorization	187	61	Yes
AES	147	43	Yes
RSA	102	286	No

All in all, these findings indicate that cryptography using matrix factorization is a prospective area of secure data sharing, especially in distributed settings where modular key storage and high fidelity are required. It achieves particularly well in resistance against partial data leakage, large data structure without exponentially increasing computational complexity, and it is resilient to noisy transmission. Possible future developments are numerical stability optimization with very high-rank matrices and incorporation of the method in real-time communication systems.

V. CONCLUSION

Cryptography based on matrix factorization is a new mathematically sophisticated and scalable method of secure data sharing. This technique offers good obfuscation without using classical key-pair based system since it decomposes data into latent matrix components. Experiments show that technique is computationally efficient, partially resistant and can be integrated into distributed data systems. Future work will concern the incorporation of quantum-resistant matrix decomposition algorithms and the assessment of homomorphic aptitudes with a view to calculating on encrypted data. With the data security getting more and more complicated, these mathematically-based approaches can become the baseline of the new era of the cryptographic tools.

REFERENCES

[1] R. Velumani, H. Sudalaimuthu, G. Choudhary, S. Bama, M. V. Jose, and N. Dragoni, “Secured secret sharing of QR codes based on nonnegative matrix factorization and regularized super resolution convolutional neural network,” *Sensors*, vol. 22, no. 8, p. 2959, Apr. 2022, doi: 10.3390/s22082959.

[2] V. W. Anelli, Y. Deldjoo, T. Di Noia, A. Ferrara, and F. Narducci, “User-controlled federated matrix factorization for recommender systems,” *Journal of Intelligent Information Systems*, vol. 58, no. 2, pp. 287–309, Jan. 2022, doi: 10.1007/s10844-021-00688-z.

[3] G. Liang, C. Sun, J. Zhou, F. Luo, J. Wen, and X. Li, “A General Matrix factorization framework for recommender systems in multi-access edge computing network,” *Mobile Networks and Applications*, vol. 27, no. 4, pp. 1629–1641, Feb. 2022, doi: 10.1007/s11036-021-01869-4.

[4] Al-Aiash, R. Alquran, M. AlJamal, A. Alsarhan, M. Aljaidi, and D. Al-Fraihat, “Optimized digital watermarking: Harnessing the synergies of Schur matrix factorization, DCT, and DWT for superior image ownership proofing,” *Multimedia Tools and Applications*, Jul. 2024, doi: 10.1007/s11042-024-19781-w.

[5] M. Asad, S. Shaukat, E. Javanmardi, J. Nakazato, and M. Tsukada, “A Comprehensive Survey on Privacy-Preserving Techniques in Federated Recommendation Systems,” *Applied Sciences*, vol. 13, no. 10, p. 6201, May 2023, doi: 10.3390/app13106201.

[6] S. S. Mugdho and H. Imtiaz, “Privacy-preserving matrix factorization for recommendation systems using Gaussian mechanism and functional mechanism,” *International Journal of Machine Learning and Cybernetics*, vol. 15, no. 12, pp. 5745–5763, Jul. 2024, doi: 10.1007/s13042-024-02276-3.

[7] D. Chawla and P. S. Mehra, “A roadmap from classical cryptography to post-quantum resistant cryptography for 5G-enabled IoT: Challenges, opportunities and solutions,” *Internet of Things*, vol. 24, p. 100950, Sep. 2023, doi: 10.1016/j.iot.2023.100950.

[8] P. M. Nikunj, D. B. Rathod, and J. Dave, “Intelligent heuristic Keyword-Based search methodologies applied to cryptographic cloud environment,” in *Smart innovation, systems and technologies*, 2022, pp. 179–186. doi: 10.1007/978-981-19-3571-8_19.

[9] M. Harasic, F.-S. Keese, D. Mattern, and A. Paschke, “Recent advances and future challenges in federated recommender systems,” *International Journal of Data Science and Analytics*, vol. 17, no. 4, pp. 337–357, Aug. 2023, doi: 10.1007/s41060-023-00442-4.

[10] V. Perifanis, N. Pavlidis, A. Sendros, and P. S. Efraimidis, “Federated Learning for Recommender Systems: Advances and Perspectives,” in *Studies in computational intelligence*, 2025, pp. 87–106. doi: 10.1007/978-3-031-78841-3_5.

[11] Y. Zhou *et al.*, “USST: A two-phase privacy-preserving framework for personalized recommendation with semi-distributed training,” *Information Sciences*, vol. 606, pp. 688–701, May 2022, doi: 10.1016/j.ins.2022.05.083.

[12] S.-T. Zhong, L. Huang, C.-D. Wang, J. Lai, G. Xie, and Y. Li, “A Model-Bias Matrix factorization approach for course score prediction,” *Neural Processing Letters*, vol. 54, no. 5, pp. 3583–3600, Nov. 2020, doi: 10.1007/s11063-020-10385-7.

[13] R. Maskeliūnas, R. Damaševičius, A. Kulikajevas, K. Pribuišis, N. Ulozaitė-Stanienė, and V. Uloza, “Pareto-Optimized Non-Negative Matrix factorization approach to the cleaning of alaryngeal speech signals,” *Cancers*, vol. 15, no. 14, p. 3644, Jul. 2023, doi: 10.3390/cancers15143644.

[14] Z. Shao, L. Ma, Q. Lin, J. Li, M. Gong, and A. K. Nandi, “PMCDM: Privacy-preserving multiresolution community detection in multiplex networks,” *Knowledge-Based Systems*, vol. 244, p. 108542, Mar. 2022, doi: 10.1016/j.knosys.2022.108542.

[15] H. Xue, D. Chen, N. Zhang, H.-N. Dai, and K. Yu, “Integration of blockchain and edge computing in internet of things: A survey,” *Future Generation Computer Systems*, vol. 144, pp. 307–326, Nov. 2022, doi: 10.1016/j.future.2022.10.029.

[16] X. Wu, C. Cui, and S. Wang, “Perceptual hashing based on salient region and NMF,” in *Smart innovation, systems and technologies*, 2022, pp. 119–127. doi: 10.1007/978-981-19-1057-9_12.