Journal of Information Systems Engineering and Management
2025, 10(24s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Homomorphic Encryption-driven AI through Text Mining in The Cloud

Qays Jabbar Abed[1], Dr. Rami Tawil[2]
[1]*Faculty of Science Lebanese University, Lebanon*
[2]*Faculty of Science Lebanese University, Lebanon*

ARTICLE INFO

Received: 25 Dec 2024

Revised: 12 Feb 2025

Accepted: 24 Feb 2025

ABSTRACT

In the current era, there is increasing interest in data security, especially in cloud computing. Homomorphic Encryption (HE) supported by Artificial Intelligence (AI) technology offers a promising solution in this field. Homomorphic encryption ensures that computations are performed on encrypted data without decryption, thus ensuring privacy. However, the integration of AI algorithms and text mining in the cloud environment is still a challenging topic. The study aims to develop a framework for partial homomorphic encryption combined with a deep learning algorithm for text mining in the cloud environment. The aim of the proposed approach is to evaluate the trade-offs between security and computational performance through deep learning to ensure the highest accuracy.The proposed method uses frequency coding and combines it with the developed deep learning algorithm, which is based on the dynamic change of the weights accompanying the neural network. The text mining model is integrated by multiplying the encrypted frequency by the factor derived from the weight in the neural network iterations. The model was trained on data in two standard datasets and the model was tested afterwards. The computational overheads were evaluated as the text size before and after encryption, the use of computing resources, and the amount of noise generated. Using HE allowed for successful text mining on encrypted data, with minimal impact on accuracy. The ciphertext size was 3.3x larger than plaintext, with increased computational overhead. The computational resource utilization was balanced in an acceptable manner for cloud storage, with noise growth not exceeding 31% while accuracy remained at 98%. In this study, the feasibility of using homomorphic encryption on texts supported by deep learning technology in a cloud environment was concluded. This provides a solution for computational operations on data while preserving privacy. The framework provides a balance between security and computational efficiency and is important for applications that require high levels of security, despite some challenges that may be solved in the future by machine learning and working on larger texts.

Keywords: Homomorphic Encryption; Deep Learning, Text Mining; Cloud environment, Weight Extraction.

INTRODUCTION

The rapid development of information technology and cloud computing has led to a major revolution in the field of data processing and various applications, including artificial intelligence. With the rapid growth of technology came increasing concerns about data security, and one of the most important data is sensitive text data stored in the cloud [1]. Information security used to rely on traditional encryption methods to maintain data confidentiality, in both cases, whether data is static or transmitted between devices through computer networks. In different environments such as the cloud environment, traditional encryption methods fail to maintain data security, as decryption in such cases leads to disasters. Cloud computing has gained popularity in recent times due to its multiple benefits of cost savings, high scalability and optimal accessibility. Many companies and applications have been attracted to the cloud due to its ability to provide the resources needed by the beneficiary at a lower cost and time, on demand and without infrastructure. However, the increasing reliance on cloud services has raised concerns, as users are relinquishing direct control over their sensitive data [2]. This leads to security issues as analysis and processing are far from the beneficiary.

From this standpoint, interest in Homomorphic Encryption (HE) came as an approach to data encryption, where calculations can be performed on encrypted data directly, without decrypting it [3]. This enhances data security and by maintaining the function of artificial intelligence. This is important in many fields, most notably health care, financial and legal services. Symmetric encryption (SE) and asymmetric encryption (AE) are traditional methods that have been considered as a cornerstone of data security for a period of time [4]. These methods are good at storing or transmitting data confidentially, however, encryption in the cloud has another dimension with limitations to traditional methods. One of the biggest challenges is performing mathematical operations on the data directly on the data after encryption. For this, it must be decrypted and then the mathematical operations must be performed and then it is stored in the cloud. For this reason, homomorphic encryption came about, which works to perform mathematical operations through encrypted data without decrypting it. There are two types of symmetric encryption: Partially Homomorphic Encryption (PHE) which includes addition and multiplication in isolation, and Fully Homomorphic Encryption (FHE) which is more advanced than the first as it can support both addition and multiplication, so any arithmetic function can be evaluated [5].

In this study, homomorphic encryption is integrated with text mining in the context of artificial intelligence that processes data in the cloud environment. The symmetric encryption approach processes text data after encryption in a way that preserves data privacy. In this study, the efficiency of data encryption in cloud computing and the impact of artificial intelligence on data security are explored [6]. Through objective evaluation on real data, the proposed approach proves its worth in encrypting and analyzing text securely in the cloud.

The study aims to secure data in cloud storage using one of the artificial intelligence techniques, which is deep learning. Maintaining data security is one of the priorities we take into account when choosing cloud computing, so we can clarify the following objectives:

- To develop a hybrid model for secure text mining when storing in the cloud. To ensure that the data is secure, we use partial homomorphic encryption based on deep learning in choosing randomness and the primary key.
- To evaluate the computational efficiency of AI-based homomorphic encryption and text mining resulting from the performance of the proposed method with traditional methods.

LITERATURE REVIEW

Symmetric encryption is very important in text encryption because of its ability to perform computational operations without the need for decryption. The concept was first introduced by [7] carrying the idea of computational operations for the privacy of confidential data. Fully symmetric encryption was developed by [8] where a feasible and applicable scheme was first proposed. Later, efforts were made to improve approaches and methods based on partially symmetric encryption, especially some of the symmetric techniques supporting it such as BFV, CKKS and TFHE which were proposed by [9]. Several studies have considered symmetric encryption in applications that support cloud computing. For information security, artificial intelligence-enhanced methods were used. Artificial intelligence was used in encrypted data on the cloud to address the challenges faced by data security in a study proposed by [10]. Algorithms that address the computational cost of hardware acceleration are a priority and are used to enhance the capacity and efficiency of models.

Text mining is one of the leading fields that are enhanced by artificial intelligence, and texts are extracted by generating random data by predicting new patterns in each cycle [11]. Traditional methods relied on statistical data and fixed rules, but through the development of deep learning, the performance was improved by regularly predicting the features extracted from the text [12]. Models that improved natural language processing (NLP) added high accuracy in classifying the extracted texts and logical analysis [13]. When storing information in the cloud, information security is a major challenge, for this reason, secure multi-party models (MPC) were developed, although there are limitations in terms of accuracy on these models at the expense of security [14]. Cloud computing is one of the advanced models that have made a difference in terms of efficiency. Especially those that depend on artificial intelligence, but they include challenges in terms of data security. Cloud data security is a controversial topic for many researchers, and eventually, studies that included symmetric encryption proved the worthiness of how to encrypt using artificial intelligence [15]. Many studies have dealt with the integration of deep learning with symmetric encryption and concluded many cases that contributed to the development of the method. Some studies have introduced neural networks with the topic of symmetric encryption to reduce computational cost and increase performance [16].

PREPOSED METHODS

In this study, the idea is to develop a hybrid framework that combines homomorphic encryption and deep learning text mining in the cloud environment. The goal is to perform computational operations such as addition and multiplication on cloud data without decrypting it in order to maintain information security. The main modules of the work are as follows:

First: Homomorphic encryption where the text is encrypted before cloud storage.

Second: Encrypted text mining using deep learning and based on variable weights in the hidden layers, this is to support the encryption process of generating random numbers and the key.

Third: Decryption and then interpreting the data by evaluating the results.

The methodology focuses on the mathematical formula in the encryption process as the arithmetic operations do not affect the encrypted data, and the reliance on deep learning to improve mining tasks.

We use here partial homomorphic encryption because it supports encrypted integers and is suitable for text representations as well. It is computationally feasible if we add some AI data.

The plaintext value consider as $x \in \mathbb{Z}_q$ such as $\mathbb{Z}_q$ring of q integrate model) while the homomorphic encryptionis x belong public key pk and can represent as:

$$Enc_{pk}(x) = (c_1, c_2) \text{ such as } c_1 = E(x) \qquad (1)$$

Where E consider the encryption function and (c_1, c_2) represent the operations like addition or multiplication within ciphertext. And the poerations supported are:

Addition with two encrypted values $Enc_{pk}(x)$and $Enc_{pk}(y)$ where the homorphic sum is:

$$Enc_{pk}(x) \oplus Enc_{pk}(y) = Enc_{pk}(x + y) \qquad (2)$$

And the other operation is multiplication for two encrypted values $Enc_{pk}(x)$and $Enc_{pk}(y)$then produce:

$$Enc_{pk}(x) \otimes Enc_{pk}(y) = Enc_{pk}(x \times y) \qquad (3)$$

Both operation driven buy deep learning during basic arithmetic operations. Deep learning model adapt and support encrypted data, the tokenizing input text through embedding in cipher text representing as $t = [t_1, t_2, \dots, t_n]$ where t_i is mapped in the embedding of e_1, by using the function of f, such as:

$$e_i = f(t_i) \qquad (4)$$

For secure process we make embedding convert to integer through token:

$$e_i' = tokenize(e_i) \qquad (5)$$

Through obtaining token e_i' then generate the embedding and apply homomorphic encryption to it $Enc_{pk}(e_i')$. Our proposed method is based on assigning a dynamic weight derived from the layers in the neural network. Which comes from the feedback and how far it is from the input layer, also the weight depends on which node the features output and which node and layer they will go to. In the training process, the weight is calculated accurately during each cycle and from it the appropriate randomness can be calculated to encrypt the text. Mechanism of transferring data within hidden layers as:

$$class(W, K, V) = softmax \frac{WK^T}{\sqrt{d_k}} V \qquad (6)$$

Where *W,K and V* Consider weight, key, and data value, and d_k represents the key vector dimension. While the softmax function compute the weight for each iteration, all data will modify during training until we reach the optimal value for encryption as shown in Figure 1.

In homomorphic encryption, the multiplication can perform the same as addition with some modification due to the multiplication consider repeated additions.

$$Enc_{pk}(W) \otimes Enc_{pk}(K) = Enc_{pk}(WK^{T}) \quad (7)$$

The feedback controls the multiplication in the neural network, and the number of layers with corresponding nodes helps in computing activities, then encrypted data effectively

RESULTS AND DISCUSSION

In this section, we present the experimental results obtained from applying the proposed model. The results focus on accuracy, computational efficiency, and security. Experiments were conducted on a standard dataset to standardize with previous studies such as the 20 Newsgroups dataset (for text classification). The performance of the proposed method based on deep learning will be compared with previous conventional models, whether in traditional AI or traditional encryption approaches, and then we prove the merit and effectiveness of the proposed method.

In order to measure accuracy, a set of experiments was designed to evaluate the AI approach on encrypted data. The model was trained on 80% of the data in 20 Newsgroups and the text was classified sequentially, through the measurement the feasibility of the approach was determined. As shown in Table 1.

To measure accuracy, deep learning works on two approaches, the first without extracting dynamic weights and the second with considering variable weights. The application was applied to two types of datasets and training on 80% of the data, and the text was classified sequentially during processing, storage and analysis.

Another criterion that was taken into consideration was measuring the computational cost through homomorphic encryption. This is done by measuring the access time in addition to the time needed for processing, for the same dataset using the deep learning model for data in the cloud environment. Taking into account the computing resources that are provided by a 10th generation CPU core i7 and GPU 2G. As shown in Table 2.

The Table 2 shows that there is a significant increase in inference time when we apply the proposed model. The increase was a noticeable 128.28% in access time, and for the other dataset it was a 135.85% increase. This is according to variables such as text size and algorithm capacity. Also, computational resources also have an impact on the resulting value, so it is necessary to pay attention to them.

One of the most important criteria to consider is security. In homomorphic encryption, several metrics are used that are important for text mining:

First: The size of the encrypted text, which has an impact on security and the cost of cloud storage. Larger texts are more secure, but they cost storage space.

Second: The growth of noise, which is added during encryption and subsequent operations. Too much noise helps security, but leads to performance degradation.

Third: The failure rate of decryption, which is due to too much noise or insufficient freshness of the encrypted text.

Fourth: The time taken for encryption or decryption, which measures the efficiency of encryption.

The relation between these metrics can be illustrated in Table 3.

Table 3 shows how the ciphertext size increased by 3.3-3.2× when compared with text before encryption. This is because of the padding and metadata needed for encryption. Then the noise growth during 15 iterations is still under control and increases by just 30.6% for the first dataset, and for the other datasets, it increased by 31%. The decryption failure rate is still low indicating good performance with the proposed model.

CONCLUSION

In this study, we present a new approach that includes integrating homomorphic encryption with one of the artificial intelligence algorithms for text mining in cloud storage. The proposed approach works on processing and encrypting data securely to maintain privacy by allowing mathematical operations such as addition and multiplication on the encrypted data. The proposed approach ensures that the data is returned after decryption without changing even after performing mathematical operations on it. The proposed method ensures that the data is secure and robust against cryptographic attacks. In this study, a deep learning method based on the recurrent neural network was used with the development of dynamic weights. The dynamic weights are extracted from the neural network environment as the length of the feedback and its bias to any node in the hidden layer as well as the difference between its distance from the input and output layers. The extracted weight works as a factor that multiplies the frequency of the letter in the encrypted letter table. The study contributes to maintaining the security of sensitive encrypted data after performing computational operations on them, in various accounting sectors such as healthcare, accounting, finance and legal. The study provides a balance between computational expenses such as increasing the size of the encrypted text and noise in the cloud environment using computational resources and the efficiency and accuracy of data security. This approach provided promising results, proving the worthiness of the proposed approach, in terms of accuracy (98%) and latency increase (135%). However, challenges such as computational costs and scalability still exist, and future work may include integrating hybrid algorithms with machine learning or working on computational operations other than addition and multiplication.

REFRENCES

[1] Gudimetla, S. R. (2024). Data encryption in cloud storage. International Research Journal of Modernization in Engineering Technology and Science, 6, 2582-5208.

[2] Dewangan, R. R., Soni, S., & Mishal, A. (2025). An approach of privacy preservation and data security in cloud computing for secured data sharing. Recent Advances in Electrical & Electronic Engineering, 18(2), 176-195.

[3] Liu, W., You, L., Shao, Y., Shen, X., Hu, G., Shi, J., & Gao, S. (2025). From accuracy to approximation: A survey on approximate homomorphic encryption and its applications. Computer Science Review, 55, 100689.

[4] Yang, Y., Fan, H., Zhang, J., Li, B., Ma, H., & Gu, X. (2025). SMSSE: Size-pattern Mitigation Searchable Symmetric Encryption. IEEE Transactions on Information Forensics and Security.

[5] Ci, S., Hu, S., Guan, D., & Koç, Ç. K. (2025). Privacy-preserving word vectors learning using partially homomorphic encryption. Journal of Information Security and Applications, 89, 103999.

[6] Dhal, S. B., & Kar, D. (2025). Leveraging artificial intelligence and advanced food processing techniques for enhanced food safety, quality, and security: a comprehensive review. Discover Applied Sciences, 7(1), 1-46.

[7] Yuan, J., Liu, W., Shi, J., & Li, Q. (2025). Approximate homomorphic encryption based privacy-preserving machine learning: a survey. Artificial Intelligence Review, 58(3), 82.

[8] Gentry, C. (2009, May). Fully homomorphic encryption using ideal lattices. In Proceedings of the forty-first annual ACM symposium on Theory of computing (pp. 169-178).

[9] Clet, P. E., Stan, O., & Zuber, M. (2021). BFV, CKKS, TFHE: Which one is the best for a secure neural network evaluation in the cloud?. In Applied Cryptography and Network Security Workshops: ACNS 2021 Satellite Workshops, AIBlock, AIHWS, AIoTS, CIMSS, Cloud S&P, SCI, SecMT, and SiMLA, Kamakura, Japan, June 21–24, 2021, Proceedings (pp. 279-300). Springer International Publishing.

[10] Devi, T. A., & Jain, A. (2024, May). Enhancing Cloud Security with Deep Learning-Based Intrusion Detection in Cloud Computing Environments. In 2024 2nd International Conference on Advancement in Computation & Computer Technologies (InCACCT) (pp. 541-546). IEEE.

[11] Fan, G. F., Han, Y. Y., Li, J. W., Peng, L. L., Yeh, Y. H., & Hong, W. C. (2024). A hybrid model for deep learning short-term power load forecasting based on feature extraction statistics techniques. Expert Systems with Applications, 238, 122012.

[12] Bendaouia, A., Qassimi, S., Boussetta, A., Benzakour, I., Benhayoun, A., Amar, O., ... & Hasidi, O. (2024). Hybrid features extraction for the online mineral grades determination in the flotation froth using Deep Learning. Engineering Applications of Artificial Intelligence, 129, 107680.

[13] Shamshiri, A., Ryu, K. R., & Park, J. Y. (2024). Text mining and natural language processing in construction. Automation in Construction, 158, 105200.

[14] Harinath, D., Bandi, M., Patil, A., Murthy, M. R., & Raju, A. V. S. (2024). Enhanced Data Security and Privacy in IoT devices using Blockchain Technology and Quantum Cryptography. Journal of Systems Engineering and Electronics (ISSN NO: 1671-1793), 34(6).

[15] Li, B., Feng, Y., Xiong, Z., Yang, W., & Liu, G. (2021). Research on AI security enhanced encryption algorithm of autonomous IoT systems. Information sciences, 575, 379-398.

[16] Wu, J., Zhang, K., Wei, L., Gong, J., & Ning, J. (2024). Practical Searchable Symmetric Encryption for Arbitrary Boolean Query-Join in Cloud Storage. IEEE Transactions on Information Forensics and Security.