Optimized Paillier Homomorphic Encryption in Federated Learning for Speech Emotion Recognition

Samaneh Mohammadi[1,2], Sima Sinaei[1], Ali Balador[2], and Francesco Flammini[2]
[1] RISE Research Institutes of Sweden, Västerås, Sweden. Email: {samaneh.mohammadi, sima.sinaei}@ri.se
[2] Mälardalen University, Västerås, Sweden. Email: {ali.balador, francesco.flammini}@mdu.se

***Abstract*—*Context:* Federated Learning is an approach to distributed machine learning that enables collaborative model training on end devices. FL enhances privacy as devices only share local model parameters instead of raw data with a central server. However, the central server or eavesdroppers could extract sensitive information from these shared parameters. This issue is crucial in applications like speech emotion recognition (SER) that deal with personal voice data. To address this, we propose Optimized Paillier Homomorphic Encryption (OPHE) for SER applications in FL. Paillier homomorphic encryption enables computations on ciphertext, preserving privacy but with high computation and communication overhead. The proposed OPHE method can reduce this overhead by combing Paillier homomorphic encryption with pruning. So, we employ OPHE in one of the use cases of a large research project (DAIS) funded by the European Commission using a public SER dataset.**

***Index Terms*—Federated Learning, Privacy-preserving Mechanism, Homomorphic Encryption, Speech Emotion Recognition**

I. Introduction

Machine learning (ML) applications often encounter privacy breaches where private data is shared with a central server. Federated learning (FL) has been proposed as a solution to protect privacy, where clients train ML models on their own devices without sharing raw data with the server [1]. FL is particularly useful in applications dealing with private data, such as speech emotion recognition (SER) [2]. By sending local parameters instead of raw speech data, FL can protect clients' confidentiality.

However, when clients participate in collaborative model training and the number of training iterations increases, the FL setup becomes susceptible to various privacy attacks [3]. Furthermore, exchanging local model parameters during the training process enables adversaries to potentially identify neural network parameters, structures, and outputs [4].

To enhance the privacy and confidentiality of clients in FL, employing homomorphic encryption methods such as Paillier homomorphic encryption (PHE) can be a viable solution for SER applications [5]. PHE ensures privacy by performing computations on ciphertext without decryption, keeping sensitive information encrypted and secure during computation and communication. Thus, clients only upload encrypted updates, preventing both the server and third parties from collecting any information during data transmission [6].

The potential of PHE in various applications is promising, but it's important to consider the additional overhead it may introduce in communication and computation between the client and server, especially in scenarios with limited bandwidth and resources [7]. In the case of SER, which includes voice processing, feature extraction, emotion recognition model training, and encryption of model parameters, careful consideration of these challenges is crucial when implementing PHE in FL systems for SER applications.

In this paper, we propose OPHE, a lightweight method for SER applications in FL, designed for resource-constrained edge devices. OPHE combines Paillier homomorphic encryption with pruning to enhance encryption efficiency and scalability. The key contributions of this paper are: (1) Proof-of-concept implementation of Paillier homomorphic encryption for SER in FL, and (2) Propose of OPHE, an algorithm that combines Paillier encryption with pruning to ensure client confidentiality while reducing overhead.

II. Optimized Paillier Homomorphic Encryption

This section describes the threat model we consider in this paper, the proposed OPHE method.

A. Threat Model

In this paper, we consider the honest-but-curious paradigm for the server. This refers to a server that is not malicious and adheres to the FL protocol but maintains a certain level of curiosity about the clients' data and models. This curiosity can potentially compromise the confidentiality of the client's data and models.

B. Proposed Method: OPHE

To address the potential threat of a curious server in FL, we propose a novel approach called optimized Paillier homomorphic encryption (OPHE) in FL for SER applications. Our method focuses on preserving client data confidentiality by developing Paillier homomorphic encryption. The clients encrypt their local SER models using public keys and share them with the server. The Paillier cryptosystem enables the server to process and aggregate model parameters without decryption. Importantly, Paillier homomorphic cryptosystems have been proven to be secure against chosen plaintext attacks, ensuring that ciphertexts do not reveal any information about plaintexts. By employing this encryption scheme, we effectively safeguard clients' data from potential threats, adding an extra layer of protection within the FL system.

In order to optimize our method, we introduce a pruning technique prior to encryption on the client side. This technique

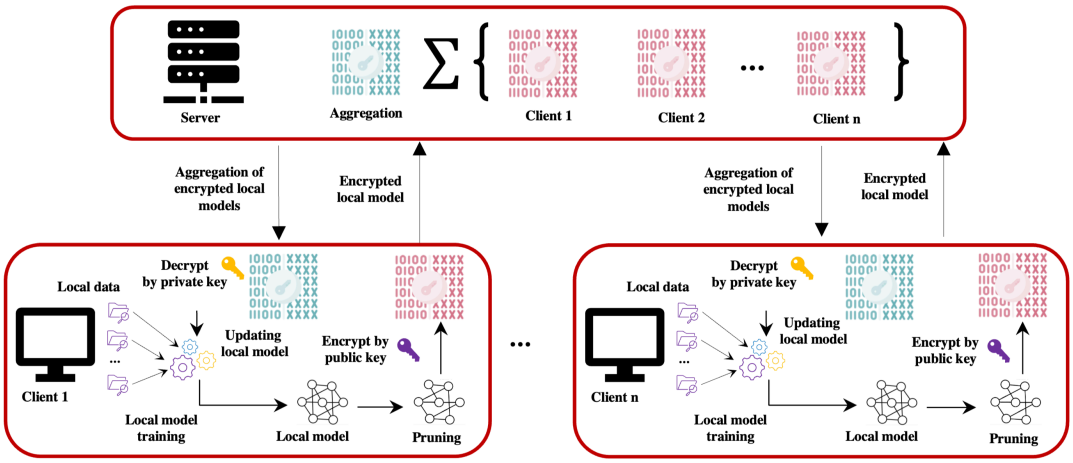

Fig. 1: An overall overview of optimized Paillier homomorphic encryption in federated learning.

selectively eliminates redundant parameters from the client's model, resulting in a more compact model that requires less encryption time. As a result, this approach effectively reduces both computational and communication overheads without compromising the accuracy of the SER model. Figure 1 provides a comprehensive overview of our proposed method.

III. Experimental Results

A. Use Case Description and Simulation Setting

Secure and decentralized media recommendation based on user emotions is an important use case in the digital life category of DAIS project. Achieving this requires a distributed, efficient, secure, and precise SER application. Pursuing these goals has driven the investigation into the potential of OPHE in the field of SER.

In this study, we evaluated OPHE using the CREMA-D SER dataset, which includes 7,442 speech recordings from 91 actors (48 males, 43 females) with various accents. We trained a SER model using the four most common emotion labels (neutral, sad, happy, angry). The model architecture consisted of a multilayer perceptron (MLP) with two dense layers [256, 128], ReLU activation, and a dropout rate of 0.2. In the FL setup, 80% of the data was used for client-side training, while the remaining 20% served for testing and validation purposes.

B. Initial Result

Our initial results demonstrate that OPHE outperforms non-optimized PHE in terms of communication and computation overhead while maintaining a marginal impact on SER accuracy. Specifically, the proposed method achieves an accuracy of 70%, representing a slight 2% reduction compared to the initial model accuracy, while reducing communication and computation overhead by nearly 50%. This makes OPHE an effective solution for SER on resource-constrained edge devices, optimizing resource utilization through a privacy-performance trade-off.

IV. Conclusions and Work in Progress

In this paper, we have introduced a novel and lightweight method named optimized Paillier homomorphic encryption (OPHE) in federated learning for speech emotion recognition. The OPHE method combines Paillier homomorphic encryption with pruning to ensure the client's privacy while enhancing efficiency and scalability.

In this study, we have conducted simulations and implemented OPHE, and obtained preliminary results. Our ongoing work is centered around evaluating its performance in various aspects with more detailed analysis. Additionally, we are continuing to explore and enhance the encryption method's efficiency aspects.

V. Acknowledgement

This work was partially supported by EU ECSEL project DAIS that has received funding from the ECSEL JU under grant agreement No.101007273.

References

[1] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial intelligence and statistics*. PMLR, 2017, pp. 1273–1282.

[2] S. Latif, S. Khalifa, R. Rana, and R. Jurdak, "Federated learning for speech emotion recognition applications," in *2020 19th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*. IEEE, 2020, pp. 341–342.

[3] M. Nasr, R. Shokri, and A. Houmansadr, "Comprehensive privacy analysis of deep learning: Passive and active white-box inference attacks against centralized and federated learning," in *2019 IEEE symposium on security and privacy (SP)*. IEEE, 2019, pp. 739–753.

[4] M. S. Jere, T. Farnan, and F. Koushanfar, "A taxonomy of attacks on federated learning," *IEEE Security & Privacy*, vol. 19, no. 2, pp. 20–28, 2020.

[5] C. Fang, Y. Guo, Y. Hu, B. Ma, L. Feng, and A. Yin, "Privacy-preserving and communication-efficient federated learning in internet of things," *Computers & Security*, vol. 103, p. 102199, 2021.

[6] J. Park and H. Lim, "Privacy-preserving federated learning using homomorphic encryption," *Applied Sciences*, vol. 12, no. 2, p. 734, 2022.

[7] Z. Jiang, W. Wang, and Y. Liu, "Flashe: Additively symmetric homomorphic encryption for cross-silo federated learning," *arXiv preprint arXiv:2109.00675*, 2021.