A novel secure deep ensemble learning protocol based on Conjugacy search problem homomorphic encryption scheme

Tran Anh Tu, Luong The Dzung, Hoang Duc Tho, Nguyen Hoang Anh

***Abstract*—Nowadays, machine learning and deep learning have been widely employed. User privacy is an issue to consider in problems such as medicine, and finance. Machine learning models not only require accurate predictions but also ensure the privacy and security of data for users. In this paper, we propose a method to ensure the privacy for training and using deep learning models that employs a homomorphic encryption scheme based on the conjugate search problem. This method implements encryption on the data before transferring them to a cloud server, which stores local deep learning models from participants to predict the encrypted data, then the encrypted prediction results are sent back to users, and they perform decryption to get the model's prediction result. These results can also be assembled to create a new training dataset for a model from the client. It is evident that our proposed model on the MNIST dataset produces an accuracy over 98% with some very simple network architectures and approximates the accuracy of centralized complex models, which does not ensure privacy.**

***Tóm tắt*— Hiện nay, học máy và học sâu nói chung đã và đang được ứng dụng rất rộng rãi. Tuy nhiên, trong nhiều bài toán như y tế, tài chính, dữ liệu riêng tư của người dùng là một vấn đề cần xem xét. Các mô hình học máy không chỉ yêu cầu dự đoán chính xác mà còn cần đảm bảo được tính riêng tư và bảo mật của dữ liệu cho người dùng. Trong bài báo này, nhóm tác giả trình bày một phương pháp đảm bảo tính riêng tư cho việc huấn luyện và sử dụng các mô hình học máy đặc biệt là học sâu sử dụng hệ mã hóa đồng cấu dựa trên bài toán tìm kiếm liên hợp. Phương pháp pháp mã hóa đồng cấu này thực hiện mã hóa dữ liệu trước khi gửi lên các máy chủ đám mây, nơi lưu trữ mô hình học học sâu cục bộ của các bên tham gia để đưa ra các dự đoán tương ứng trên dữ liệu đầu vào ở dạng mã hóa, sau đó kết quả dự đoán sẽ được trả về người dùng và người dùng thực hiện giả mã để nhận được kết quả dự đoán của mô hình. Các kết quả này cũng có thể được xây dựng thành một bộ dữ liệu huấn luyện để thực hiện quá trình xây dựng và huấn luyện lại một mô hình cho máy khách. Nhóm tác giả chỉ ra rằng, mô hình đề xuất của nhóm tác giả trên bộ dữ liệu chuẩn MNIST cho độ chính xác lên tới gần 99% với kiến trúc mạng rất đơn giản và gần như có độ chính xác xấp xỉ với các mô hình phức tạp tập trung không đảm bảo tính riêng tư cho dữ liệu.**

Keywords*—*Deep learning, Privacy Preserving machine learning, secure multi-participant computation.

Từ khóa*—*Học sâu, đảm bảo tính riêng tư cho học máy, tính toán bảo mật nhiều thành viên

I. INTRODUCTION

Deep learning is one of the advanced approaches to Machine learning and draws more and more attention recently. Deep learning is widely utilized in various areas such as image processing, face recognition, voice identification, medicine prediction... The advantage of the deep learning model is the ability to automatically learn features of the data in order to establish better new features for prediction.

However, the effectiveness of deep learning models remarkably relies on the quality and quantity of the data. Sharing data for fitting an effective deep learning model is naturally necessary. However, there has to be a trade-off between user privacy and accuracy. Data in the training and prediction phase can contain highly sensitive information such as medical, financial or personal data, in such cases, data is confidential. Sharing data between the participants of the collaborative model building may result in hacking, copying, reuse, and information leakage. This poses a significant threat to user privacy, causing serious damage to the reputation, economy, users and organizations.

Different from traditional models, the deep learning model has many special features such as a large number of parameters, complex structure, and nonlinear computations that work on real numbers. As a result, ensuring privacy for deep learning is distinctive. Basically, these approaches can be classified according to the data-sharing model. Based on this classification, research in this field mainly focuses on three approaches. The first one is transforming the input data, ensuring original data is utilized in training, not revealing information, and an accurate prediction. In this model, training and usage of the model are provided by a service server. The data will be uploaded from the user and calculation is carried out by the server. In addition, the type of data shared between participants and servers is the input data. To ensure the privacy of data, this approach is targeted at researching solutions to transform the data. Local inputs from the participants are transformed and forwarded directly to service servers for performing computations and training. According to the approach, original data is often transformed before using the homomorphic cryptographic algorithms, secure multi-party computation protocols, and secret sharing techniques, or noise perturbation. This method is quite classical and traditional, which has been used a lot in previous studies on private data mining, and machine learning with privacy assurance. This method has the advantage that it might be used in both training and predicting phases. On the contrary, there are a lot of problems coming along, especially in terms of performance as well as accuracy. Therefore, in deep learning models, this method is often not employed in the training phase, which requires many repetitive complex calculations.

The second approach is considered as the most useful way of training distributed data models. This is also a method that has many practical applications such as Google Keyboard, IoT, model sharing, or distributed learning with popular representations such as distributed model training, split learning, SGD with large batch size and federated learning, is an efficient approach that allows participants to collaborate together on training an aggregate model based on their local data. In this approach, each participant in the training phase is responsible for training its own local data and sending local intermediate training results (parameter or gradient) to other parties or servers to assemble a global model.

With the methods mentioned above, the model’s architecture and parameters are shared among the participants, along with intermediate training results of the model such as gradients, activation functions, and updated weights as well. Although data is not leaked directly, in [12], the authors have shown that the original raw data can be approximately reconstructed by an attacker, especially in the case when the architecture and parameters of the model are not protected. Additionally, models need to embrace the common architecture, which causes much inconvenience during operation.

To solve these problems, Papernot et al. have come up with a promising method called Private Aggregation of Teacher Ensembles or PATE in short. This is the associative learning approach, in the black box way, multiple models are trained with distinct datasets, such as records from different subsets of users. These models directly are trained on sensitive data, so these models are not made public, but instead used as a "teacher" to train a "student" model. Students learn to predict selected outputs based on the votes of all teachers and are not allowed to directly access every teacher or teacher’s fundamental arguments.

The student model is trained by all the teachers, no teacher or dataset that determines the training process, therefore, the teacher’s data and model are not revealed even if the opponent keeps the student model.

The advantage of associative learning is that participants do not need to agree on the hyper-parameter and model architecture among the participants. However, its disadvantage is that the accuracy will decrease significantly. It also requires teacher models to be of good quality, which is unlikely to happen in practically. Furthermore, when the student model makes queries to the teacher model sufficiently, the results can be used to perform a black-box attack on the teacher’s model in order to replicate its training data. Moreover, the student model must have enough publicly available data transferred to teachers, which is a remarkable risk to privacy, especially in medical or financial issues.

Therefore, this study aims to find ways for protecting and balancing the privacy of training data, availability and maintainability of Deep learning model performance. We propose a solution to increase the security and efficiency of this associative learning model by using homomorphic encryption based on the conjugation analysis problem.

The paper is divided into 6 sections:

Section 1: Presenting approaches of privacy assurance for machine learning models; Section 2: Giving information about the considered model; Section 3: Describing homomorphic encryption based on the problem of Conjugacy search problem; Section 4: Presenting associative learning protocol concealing inputs using homomorphic encryption based on the conjugate search problem; Section 5: Providing several evaluations about the security and performance of the proposed protocol. Finally, a conclusion is made about the results and limitations of the paper.

II. PROBLEM STATEMENTS

In gathering data for model building and evaluation, no matter how we try, the data collected is not always all the available data. Collecting all the data is completely impractical. Therefore, there is no certainty that the prediction model that we build on the collected data always provides good results on the unseen data. On the other hand, the data sample itself contains noise. Therefore, whatever algorithm is used for model building, we also need to complement techniques in order to avoid or reduce the overfitting and improve the generalization of the model. Since each algorithm is built on different approaches, even if training data can be different, for each problem, based on the "large number principle", a combination of results from many different models is likely to yield better results, which is called Ensemble learning.

Consider the following practical problem: A retail company has information about potential customers. This company wants to propose installment packages to customers in order to stimulate revenue. However, the problem is that there is no certainty about the credit history of the customer to ensure that the customer will repay on time. Usually, the way to solve this problem is that the company has to send the banks a list of potential customers for evaluation. This might expose customer information, as well as, difficultly to customize specific requirements such as extending credit for a certain group of clients. The question is how can the retail company build its own credit model based on unlabeled data and ensure the data is not disclosed to anyone, even the banks?

In the proposed model, it is supposed that there are $N + 1$ participants in the training phase, where:

A participant S acts as a student owning the unlabeled dataset *Dunlabeled*. This data is private and the student does not wish to disclose it to any other party. The student wants to make use of the data so as to construct their own model without revealing private information. The difficulty for students is finding a way to label data without disclosing information.

N participants $\{T1, T2, \dots, TN\}$ play a role as teacher models. These participants keep Wi local models trained by their own private data. Teachers provide students with labeling results for student data based on their local models.

In this model, students send unlabeled data to teachers Ti and receive predictions of labels in the data. Based on the result, the student will choose a label for its *Dunlabeled* data and then build a suitable machine learning model based on the data and label received.

The requirement is that a solution needs to be proposed to ensure the privacy of data from a student sent to teacher models so that no participants (including attackers) could get students' private information except themselves, also information about the model of participating teachers.

For simplicity in the evaluation, we assume that all participants (teachers and students) are semi-trusted or honest-but-curious. These members will strictly comply with a designed π protocol but attempt to infer additional information from other parties when implementing the protocol. In other words, parties do not actively interfere with the protocol, but only try to get as much information as possible from the data obtained.

In addition, for the simplicity of evaluation, we suppose that all teachers have the same deep learning model with common hyper-parameters regarding architecture, the number of neurons and similar types. Models embrace different parameters during training because of distinct local datasets.

Based on these assumptions, the paper proposes a protocol that ensures the privacy of data from a student by a homomorphic cryptosystem based on the Conjugacy search problem.

III. HOMOMORPHIC CRYPTOSYSTEM BASED ON THE CONJUGACY SEARCH PROBLEM

A. Conjugacy Search Problem-CSP

The conjugacy search problem is one of the difficult problems (NP-Hard) used in cryptography to build highly secure cryptosystems. The special thing is that cryptosystems based on the problem permit operations on real numbers, which hardly occurs in cryptosystems. Conventional cryptosystems based on problems such as Discrete Logarithm, Prime factorization, LWE...usually require large integers or polynomials; hence, there should be solutions to transform data when the inputs are real numbers. This makes computation extremely large. Cryptosystems based on the Conjugacy search problem are responsible for addressing this problem. The problem is stated as follows:

Given a non-commutative algebraic structure Θ and $\Theta_0 \subseteq \Theta$, $b \in \Theta$, where $b = hah^{-1}$ and $a \in \Gamma_0$, it's hard to find h, $a \in \Theta$.

The difficulty of the CSP problem is even still useful for post-quantum cryptosystems. At present, it is still very difficult to deal with known quantum algorithms. In fact, CSP is a special form of the Group Factorization Problem (GFP). This problem was recently shown to be unsolvable with $d \geq 4$, on a linear group GL$_d$(R). Therefore, the order of the matrix used in our protocol is chosen to be a minimum of 4, in order for protocol security.

B. Homomorphic cryptosystem on Non-commutative ring

Initialization: On a non-commutative ring R, choose 4 elements $h_1, h_2, h_3, h_4 \in R$ so that $H = \begin{pmatrix} h_1 & h_2 \\ h_3 & h_4 \end{pmatrix}$ is reversible, then H is the secret key used for encryption.

Encryption: Provided a message $m \in R$, user randomly selects $r_1, r_2 \in R$ and generates the matrix $M = \begin{pmatrix} m & r_1 \\ 0 & r_2 \end{pmatrix}$. The user encrypts the corresponding message $C = Enc_H(m) = HMH^{-1}$.

Decryption:

Message $m = Dec_H(C) = (H^{-1}CH)_{11}$, where $(H^{-1}CH)_{11}$ signal for the element in the top left corner of the matrix $H^{-1}CH$.

Homomorphism of cryptosystems

Regarding two ciphertexts C1, C2 corresponding to two plaintexts $m1$, $m2$, where $C_1 = HM_1H^{-1}$ and $C_2 = HM_2H^{-1}$.

Considering homomorphism with addition:

$$C_1 + C_2 = H\begin{pmatrix} m_1 & r_1 \\ 0 & r_2 \end{pmatrix}H^{-1} + H\begin{pmatrix} m_2 & r_1' \\ 0 & r_2' \end{pmatrix}H^{-1} = H\begin{pmatrix} m_1 + m_2 & r_1 + r_1' \\ 0 & r_2 + r_2' \end{pmatrix}H^{-1}$$

Then we have $Dec_H(C_1 + C_2) = m_1 + m2$ which means that the cryptosystem is homomorphic for addition. Similarly, we consider homomorphism for multiplication:

$$C_1 \times C_2 = H\begin{pmatrix} m_1 & r_1 \\ 0 & r_2 \end{pmatrix} \times H^{-1} \times H\begin{pmatrix} m_2 & r_1' \\ 0 & r_2' \end{pmatrix} \times H^{-1} = H \times \begin{pmatrix} m_1m_2 & r_1'' \\ 0 & r_2r_2'' \end{pmatrix} \times H^{-1}$$

We also have $Dec_H(C_1 \times C_2) = m_1m2$. Therefore, the cryptosystem is also homomorphic for multiplication.

C. Security of cryptosystem

The security of the above cryptosystem relies on the noncommutative property of the matrix, which makes the encryption scheme one-way. This ensures that the hacker will not be able to obtain the plaintext message from the ciphertext. In fact, the secrecy of the message is based on the difficulty of the conjugacy search and eigenvalues of the ciphertext matrix on a non-commutative ring.

Due to the difficulty of the conjugate search problem, the adversary will not be able to achieve M by splitting the ciphertext $C = HMH^{-1}$. Therefore, message m is not leaked by splitting ciphertext C. According to the encryption algorithm, a plaintext message can be regarded as an eigenvalue of the ciphertext matrix C given:

$$C = HMH^{-1} = H\begin{pmatrix} m_1 & r_1 \\ 0 & r_2 \end{pmatrix} H^{-1}.$$

Since $m, r_1, m_2 \in R$, where R is a non-commutative ring, it is impossible for an attacker to collect m by solving eigenvalues equation of C. This proves that the cryptosystem is secure.

IV. ASSOCIATIVE LEARNING PROTOCOL CONCEALING INPUT USES HOMOMORPHIC CRYPTOSYSTEM BASED ON THE CONJUGACY SEARCH PROBLEM

To ensure that the input forwarded from a student to a teacher is not revealed, the paper proposes the application of a homomorphic cryptosystem based on the conjugacy search problem. In general, a teacher can be any machine learning model, however, for simplicity and consistency in the evaluation, we assume that all teacher models are of deep neural network. As mentioned previously, we can approximate classical models as a deep neuronal models with appropriate architecture.

Data submitted by students will be passed through deep neural network models on teachers to make a prediction based on the respective model of each teacher. The prediction results will be utilized to build a student model. This model is also a deep neural network model. The data sent to the teacher from a student is encrypted using a homomorphic encryption algorithm based on the conjugacy search problem in the previous section.

However, one thing that we need to notice is that the deep learning model needs to have non-linear activation functions, which does not guarantee the essentials of homomorphic encryption, therefore, we need to transform the deep learning model, in another word, we have to transform the activation function of the deep learning model from nonlinear to a proper form available for homomorphic cryptography.

To ensure features of homomorphic cryptosystems, the model should be modified as follows:

Activation Layer: The common activation layer is a nonlinear function such as ReLU, Sigmoid, and Tanh. Therefore, we need to find alternative activation functions to approximately ensure the properties of these functions on usage.

Aggregate sampling class: In a homomorphic cryptosystem, it is impossible to use the max pooling function. Consequently, we make use of the mean pooling function, the average pooling function has only addition so it can be used on homomorphically encrypted data.

Dropout class: Dropout class gets rid of random data during training. So Dropout class cannot be used.

From the above modifications, we can build a non-disclosure association learning protocol using a homomorphic cryptosystem based on a conjugacy search problem like algorithm 1 below.

Algorithm 1: Associative learning protocol concealing input employs homomorphic encryption based on the Conjugacy search problem.

Input: Student S keeps unlabelled data $D_{unlabelled}$, which contains data points $x \in \mathbb{R}^n$ and teachers $\{T_1, T_2, \ldots, T_N\}$, where each teacher T_j owns corresponding W_j model.

Output: Student's W_S associative learning model

While $D_{unlabelled} \neq \varnothing$ do

Student:

- Randomly generate matrix $H = \begin{pmatrix} h_1 & h_2 \\ h_3 & h_4 \end{pmatrix}$ is invertible.

- Use the secret key H to encrypt the x_i component values of the vector $x = (x_1, x_2, \ldots, x_n)$ into $c_i = Enc_H(x_i) = H \cdot X_i \cdot H^{-1}$, where $X_i = \begin{pmatrix} x_i & r_{i1} \\ 0 & r_{i2} \end{pmatrix}$, where r_{i1}, r_{i2} are random numbers.

- Send $c_i, i \in 1, 2, \ldots, n$ to all teachers $T_j, j = 1, 2, \ldots, N$.

Parallel teachers T_j:

- Labelling encrypted data obtained based on the model. The resulting label is R_{cj}. The labels R_{cj} are encrypted and teachers themselves do not acknowledge the actual labeling results.

- Send R_{cj} to Student.

Student:

- Use the key matrix H to decrypt the labelling results located in the top left corner of the matrix $H^{-1} R_{cj} H$;

- Select labels for data point x based on the results obtained from teacher models;

- Remove labeled element from $D_{unlabelled}$.

Student:

- Use labeled data to train the student model.

- Return model W_S.

One thing to notice is that we can extend the size of matrix H and mask matrices X_i with respect to $2m \times 2m$ given $m \geq 1$. However, to reduce communication and computation costs, we choose the minimum size as algorithm above.

V. PROTOCOL EVALUATION

A. *Protocol security analysis*

Based on the security of the cryptosystem and conjugacy search problem, even if an attacker obtains the ciphertext C_i, it is impossible to reverse the corresponding plaintext x_i. Therefore, during data transmission from students to teachers, the protocol is secure. For each data point, a student uses a different randomly generated key, so the security of a data point completely depends only on the key in the transmission phase. So the protocol is not affected by key sharing attacks.

B. *Communication and computation cost analysis*

The protocol implements two-phase connections. In the first one, student computes ciphertexts and sends these to the teacher machines. The number of ciphertexts is n corresponding to the size of the input vector. Each ciphertext has a size of 2×2. Therefore, the student needs to send an amount of data $4n$ to a teacher. The number of teachers is N, the total amount of data to send is $4nN$. In the second phase, teachers compute and return encrypted prediction matrix R_{cj} to the client. The matrix has size of 2×2. As a consequence, the total bandwidth required to carry out the transmission in the phase is $4N$.

The bandwidth required for transmission over the entire protocol is $4nN+4N$ where n is the size of the input vector and N is the number of teachers.

In the protocol, math operations are very simple on low-level matrices, therefore, the speed and performance are quite good. It almost doesn't effect execution performance compared to the case when encryption is not applied.

C. Evaluate accuracy of the training model

To evaluate the proposed model, in this paper, we use the MNIST handwritten dataset. In the experiment, we split the MNIST dataset into 3 subsets. The first consists of 50,000 labeled samples distributed among teachers. The second one has 10,000 unlabeled samples kept by students, representing the unlabelled dataset. Finally, 10,000 samples are reserved for testing the accuracy of the model.

In the first labeled 50,000-sample dataset, we divide it by the number of teachers so as to evaluate the effect of teachers on the protocol. In practice, it is hardly that the number of teachers is huge and each teacher requires acceptable accuracy. So we only choose the evaluation benchmarks with the number of teachers corresponding to 2, 3, 4 and 5. The teacher models will be randomly selected between VGG11 and VGG13 models for simplicity of calculation. The test results with more parameters model might yield better results, however, the goal of the topic is not to built a model with high accuracy but confirming the effectiveness of privacy protection protocol. As a result, we will not use models with complex architectures.

Given 2 teachers, each teacher keeps around 25,000 data in the labeled training dataset, where one model is VGG 11 and the other is of VGG 13 architecture.

With a model of 3 teachers, we choose 2 among them using VGG 11 architecture, and the remaining teacher has VGG 13 architecture with the dataset of each teacher including 17,000, 16000 and 17000 samples, respectively.

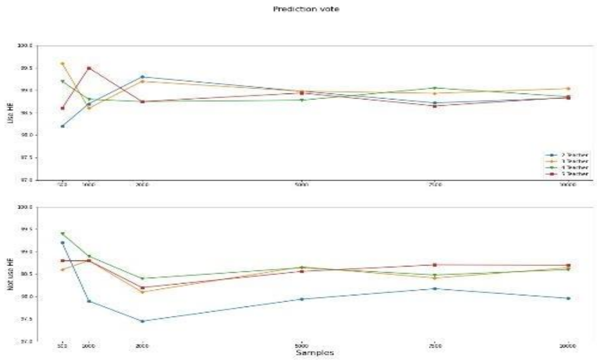
Given a model of 4 teachers, we choose 3 with VGG 11 architecture, and the other has VGG 13 architecture, with datasets of 5,000, 10,000, 15,000 and 20,000, respectively. To ensure the relative randomness and unbalance of the data distribution.

For a model of 5 teachers, we choose 3 models with VGG 11 architecture, and the others use VGG 13 architecture, with the amount of data 10000 evenly distributed among teachers.

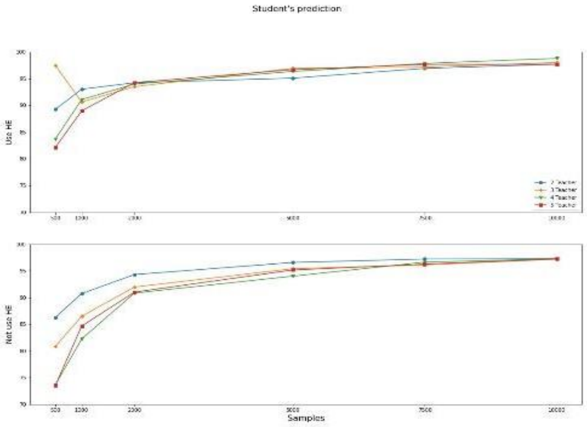
In the first labeled 50,000-sample dataset, we divide it by the number of teachers so as to evaluate the effect of teachers on the protocol. In practice, it is hardly that the number of teachers is huge and each teacher requires acceptable accuracy. So we only choose the evaluation benchmarks with the number of teachers corresponding to 2, 3, 4 and 5. The teacher models will be randomly selected between VGG11 and VGG13 models for simplicity of calculation. The test results with more parameters model might yield better results, however, the goal of the topic is not to built a model with high accuracy but to confirm the effectiveness of privacy protection protocol. As a result, we will not use models with complex architectures.

On the student side, we use the VGG 11 model to retrain the selected and labeled data by the teachers. To evaluate the effect of the number of samples possessed by the student, we divide the unlabeled dataset of 10,000 samples into small subsets and train the student model using 500, 1,000, 2,000, 5,000 and 10,000 samples respectively.

The result of 10000 student's data labeling is shown in the Figure below:



Evaluation results on test set of 10,000 samples on student's trained model with the number of samples are shown in Figure below:



With the evaluation scenarios, we make an evaluation on both models that do not use the proposed protocol and the model with the proposed protocol.

The results indicate that the proposed protocol produces a little high accuracy, close to the model that does not make use of the protocol under the same conditions and is above 98%. Some cases with the proposed protocol are even more accurate than the case without protection protocol.

In general, the proposed protocol gives an accuracy comparable to the application of native models without protection.

In summary, the proposed protocol is feasible in practice and does not affect the efficiency of the training model much.

VI. CONCLUSION

This paper has presented a general approach for ensuring privacy for machine learning models and deep learning based on the ensemble learning model. The article analyzes some advantages and disadvantages of the model, then proposes an improved solution to improve the privacy of the training process of the deep learning networks according to the ensemble learning model. The proposed model uses secure multi-party computation techniques and allows privacy for student data as well as prediction results. The results show that the proposed protocols are quite effective in terms of implementation on ensuring high accuracy while maintaining the privacy of data. The results show that the proposed model is capable of achieving accuracy up to more than 98%, which is almost equivalent to the data-centralized and non-private models. This shows the efficiency of the model.

ACKNOWLEDGMENT

The author remains thankful to lecturers in the Academy of Cryptography Techniques for their useful discussions and suggestions during the preparation of this technical paper.

REFERENCES

[1] C. Aggarwal. Neural Networks and Deep Learning. Springer, Cham, 2018..

[2] C. C. Aggarwal and P. S. Yu, editors. Privacy-Preserving Data Mining - Models and Algorithms, volume 34 of Advances in Database Systems. Springer, 2008

[3] U. M. A¨ıvodji, S. Gambs, and A. Martin. Iotfla: A secured and privacy-preserving smart home architecture implementing federated learning. In 2019 IEEE Security and Privacy Workshops (SPW), pages 175–180. IEEE, 2019.

[4] M. Al-Rubaie and J. M. Chang. Privacy-preserving machine learning: Threats and solutions. IEEE Security Privacy, 17(2):49–58, 2019.

[5] Y. Bengio, I. Goodfellow, and A. Courville. Deep learning, volume 1. MIT press Massachusetts, USA:, 2017.

[6] Boles and P. Rad. Voice biometrics: Deep learning-based voiceprint authentication system. In 2017 12th System of Systems Engineering Conference (SoSE), pages 1–6. IEEE, 2017.

[7] Bu, Y. Ma, Z. Chen, and H. Xu. Privacy preserving backpropagation based on bgv on

cloud. In 2015 IEEE 17th International Conference on High Performance Computing and Communications, 2015 IEEE 7th International Symposium on Cyberspace Safety and Security, and 2015 IEEE 12th International Conference on Embedded Software and Systems, pages 1791–1795, 2015.

[8] J. Chen, X. Pan, R. Monga, S. Bengio, and R. Jozefowicz. Revisiting distributed synchronous sgd. arXiv preprint arXiv:1604.00981, 2016.

[9] Guo and N. Zhang. A survey on deep learning based face recognition. Computer vision and image understanding, 189:102805, 2019.

[10] Gupta and R. Raskar. Distributed learning of deep neural network over multiple agents. Journal of Network and Computer Applications, 116:1 – 8, 2018.

[11] Hard, C. M. Kiddon, D. Ramage, F. Beaufays, H. Eichner, K. Rao, R. Mathews, and S. Augenstein. Federated learning for mobile keyboard prediction, 2018.

[12] Hitaj, G. Ateniese, and F. Perez-Cruz. Deep models under the gan: Information leakage from collaborative deep learning. In Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, CCS '17, page 603–618, New York, NY, USA, 2017. Association for Computing Machinery.

[13] P. Li, J. Li, Z. Huang, T. Li, C.-Z. Gao, S.-M. Yiu, and K. Chen. Multi-key privacy-preserving deep learning in cloud computing. Future Generation Computer Systems, 74:76 – 85, 2017.

[14] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith. Federated learning: Challenges, methods, and future directions. IEEE Signal Processing Magazine, 37(3):50–60, 2020.

[15] L. Lyu, X. He, Y. W. Law, and M. Palaniswami. Privacypreserving collaborative deep learning with application to human activity recognition. In Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, CIKM '17, page 1219–1228, New York, NY, USA, 2017. Association for Computing Machinery.

[16] P. Mohassel and Y. Zhang. Secureml: A system for scalable privacy-preserving machine learning. In 2017 IEEE Symposium on Security and Privacy (SP), pages 19–38, 2017.

[17] N. Papernot, M. Abadi, U. Erlingsson, I. Goodfellow, and K. Talwar. Semi-supervised knowledge transfer for deep learning from private training data. arXiv preprint arXiv:1610.05755, 2016.

[18] L. T. Phong, Y. Aono, T. Hayashi, L. Wang, and S. Moriai. Privacy-preserving deep learning via additively homomorphic encryption. Trans. Info. For. Sec., 13(5):1333–1345, May 2018.

[19] M. I. Razzak, S. Naz, and A. Zaib. Deep learning for medical image processing: Overview, challenges and the future. Classification in BioApps, pages 323–350, 2018.

[20] L. Rokach. Ensemble Learning: Pattern Classification Using Ensemble Methods (Second Edition). World Scientific Publishing Co Pte Ltd, Singapore, 2nd edition, 2019.

[21] R. Shokri and V. Shmatikov. Privacy-preserving deep learning. In Proceedings of the 22nd ACM SIGSAC conference on computer and communications security, pages 1310–1321, 2015.

[22] Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis. Deep learning for computer vision: A brief review. Computational intelligence and neuroscience, 2018.

[23] S. Wagh, D. Gupta, and N. Chandran. Securenn: Efficient and private neural network training. In Privacy Enhancing Technologies Symposium. (PETS 2019), February 2019.

[24] X. Wang, Y. Zhao, and F. Pourpanah. Recent advances in deep learning, 2020.

[25] J. Yuan and S. Yu. Privacy preserving back-propagation neural network learning made practical with cloud computing. IEEE Transactions on Parallel and Distributed Systems, 25(1):212– 221, 2014.

[26] Q. Zhang, L. T. Yang, and Z. Chen. Privacy preserving deep computation model on cloud for big data feature learning. IEEE Trans. Comput., 65(5):1351–1362, May 2016.

ABOUT THE AUTHORS

Tran Anh Tu

Workplace: Academy of Cryptography Techniques

Email: tutran@actvn.edu.vn

Education: Master degree

Recent research: Privacy Preserving Deep Learning, Secure-Multi-Party Computation.

Luong The Dzung

Workplace: Academy of Cryptography Techniques

Email: thedungluong1@gmail.com

Education: Asst. Prof

Recent research: Privacy Preserving Deep Learning, Secure-Multi-Party Computation

Hoang Duc Tho

Workplace: Academy of Cryptography Techniques

Email: thohd80@gmail.com

Education: PhD

Recent research: Cryptography

Nguyen Hoang Anh

Workplace: Civil Cryptography management and products inspection department, Hanoi.

Email: hoanganh@gmail.com

Education background: received the BSc degree from the Academy of Cryptography techniques in 2003;

Received the MSc degree from the Academy of Cryptography techniques in 2008;

Recent research direction: Information security; Cryptography.