Journal of Discrete Mathematical Sciences & Cryptography
ISSN 0972-0529 (Print), ISSN 2169-0065 (Online)
Vol. 28 (2025), No. 5-A, pp. 1497–1504
DOI : 10.47974/JDMSC-2148

Secure SVM training using privacy preserving isomorphic encryption algorithm

Rakesh Kumar Godi [†]
Department of Computer Science
Central University of Karnataka
Kalaburagi 585367
Karnataka
India

Vijay Shankar Sharma [*]
Sandeep Kumar Sharma [§]
Amit Chaurasia [‡]
Department of Computer and Communication Engineering
Manipal University Jaipur
Jaipur 303007
Rajasthan
India

Atul Srivastava [@]
Department of Computer Science and Engineering
Amity School of Engineering and Technology
Amity University
Lucknow 226028
Uttar Pradesh
India

Anuradha Pillai [#]
Department of Computer Science and Engineering
Symbiosis Institute of Technology
Pune 412115
Maharastra
India

[†] *E-mail:* rakeshgodi@cuk.ac.in
[*] *E-mail:* vijayshankar.sharma@jaipur.manipal.edu (Corresponding Author)
[§] *E-mail:* sandeep.sharma@jaipur.manipal.edu
[‡] *E-mail:* amit.chaursia@jaipur.manipal.edu
[@] *E-mail:* atul.nd2@gmail.com
[#] *E-mail:* anuradha.pillai@sitpune.edu.in

Abstract

Machine Learning algorithm such as Support Vector Machine (/SVM) are identified to be significant for prediting the variables related to the pre-defined output in real world applications. This machine learning model when imposed over the encrypted data is highly indispensable for protecting data and model information against malicious attackers. These malicious adversaries launch the attack over the data either during the phase of prediction or training. Torus-based Fast Fully Homomorphic Encryption (TFFHE) scheme is identified to facilitate potential evaluations of encrypted data related to real numbers, this merits of this TFFHE motivated the option of implementing a privacy preserving machining algorithm which can be utilized during the process of training. In this paper, Secure SVM Training using Privacy Preserving Homomorphic Encryption Algorithm using TFFHE is proposed for preventing inefficient operations and numeric stability during the phase of training in the encrypted domain. This TFFHE is proposed based on the improvement of GSW and its associated ring variants. With respect to real world datasets, this TFFHE-based SVM confirmed better performance on par with the state of the art FHE and logistic regression classifiers used for comparison. This study to the best of the knowledge is one of the few practical algorithms which could be used for SVM model training with the integration of FHE.

Subject Classification: *Primary 68P25, Secondary 68P27.*

Keywords: *Machine learning, Torus-based fast fully homomorphic encryption (TFFHE), Support vector machine (SVM), Training phase, Privacy preservation.*

1. Introduction

The utilization of different data analytic application tasks such as financial services, healthcare, and marketing in the real world has the completely increased the attention over machine learning [1]. Due to the General Data Protection Regulation (GDPR), the privacy preservation of data is more crucial since the process of unauthorized data access related to data owners is illegal and immoral [2], thus machine learning approaches has increased the demands of the services which are offered with maximized data privacy. In specific, Support Vector Machine (SVM) being one of the machine learning algorithms is determined to be more suitable for data classification [3]. This SVM despite of extensive popularity of deep learning approaches is still more used since it requires less amount of data, and it also possesses the merit of working over medium sized data and kernel works. It operates through the employment of kernel trick-based learning of complex non-linear data patterns [4]. Moreover, the SVM model in the phase of training need to solve the problem of convex optimization problem for finding out the optimal parameters on par with the deep learning models that need to handle the problem of non- convex optimization [5-7]. However, training data and model parameters need to be protected during the application of SVM models in the real world for the purpose of guaranteeing privacy preservation [8-9]. This privacy preservation of data during the process of SVM training phase need to be achieved using

homomorphic encryption algorithms since it is identified to be more ideal for the tasks related to machine learning [10].

In specific, Torus-based Fast Fully Homomorphic Encryption (TFFHE) scheme being implemented for attaining approximate arithmetic among real number data is confirmed to be more suitable for machine learning tasks since it helps in performing homomorphic computations over the real number data. In this paper, Secure SVM Training using Privacy Preserving Homomorphic Encryption Algorithm using TFFHE is proposed for facilitating private training and prediction using the meris of homomorphic encryptions and different secure protocols used for performing secure multi-party computation. This TFFHE-SVM approach is proposed not only for employing privacy preservation, but also is useful during the inference phase. It is developed as a secure training algorithm for the SVM model which provides protection to both training data and model information without facilitating access to the parties interacting in the computing scenario. It initially adopted a least square SVM algorithm for substituting the problem of constrained optimization since its necessities impotent operations to be performed in the domain of the encryption. It further used the method of gradient descent for preserving the training algorithm of FHE-based SBM training model. Finally, the proposed TFFHE-SVM approach when compared with the baseline training algorithms using different real-world datasets confirmed better performance in terms of Mean Time of Iteration (seconds) and classification accuracy aligning with prior privacy-preserving SVM applications in healthcare [11].

2. Related works

Guo et al. [12] proposed an approach using zero knowledge proof and SVM-based privacy preservation scheme for authenticating users who are cooperating with respect to different service providers. This privacy preservation approach using the merits of polynomial for improving its efficiency during its implementation process. It was proposed for handling the conflict between precision degree necessitated by the cryptographic algorithms using the use of fuzziness which provides uncertainty. In addition, Park et al. [13] proposed a homomorphic encryption algorithm-based SVM training model which handled the issue of numerical instability and impotent operations in the encrypted domain. This HFE-SVM model was proposed for both training and validation phases. The degree of privacy imposed by FHE was comparatively better than the baseline approaches such as logistic regression under real world datasets. Moreover, the results confirmed better privacy preservation during its support over the development of practical applications.

3. Secure SVM Training using Privacy Preserving Homomorphic Encryption Algorithm using TFFHE

In this section, a scalable and secure SVM training model using Torus-based Fast Fully Homomorphic Encryption (TFFHE) scheme is presented.

Let us consider the training samples represented using $\{(x_1, y_1), (x_2, y_2), \ldots\ldots, (x_n, y_n)\ \}$ such that the domain of operational space is given by$R^d \in [-1, +1]$. Then the problem of optimization with respect to this SVM training model is given by Equation (1)

$$Ob_p = \min_{\varepsilon} \frac{1}{n} \sum_{i=1}^{n} \varepsilon_i + \lambda \|w_i\|^2 \tag{1}$$

Subject to $y_i(\omega x_i + b) = 1 - \varepsilon_i$, $\varepsilon_i \geqslant 0$ and $1 \leq \forall i \leq n$.

But, this problem of optimization contains a function such as which is having unfriendly functions with respect to homomorphic encryption process.

3.1 *Process of TFFHE*

This TFFHE mechanism pertains to the type of fast fully homomorphic encryption mechanism in which each of the level 'l' needs a defied modulus m_l, where $0 \leq l \leq d - 1$. The factors associated with TFFHE scheme adopts a decreasing modulus ladder with d as the used for determining the depth of the circuit. In this context, determining suitable modulus value between m_d and m_0 is more challenging, especially when considering practical FPGA-based implementations requiring low-latency and area efficiency [14], since the magnitude of completely depends on the parameter of security termed α. In specific, the different magnitude value of m_d helps in determining circuits of different depths. Moreover, the potentialiy of this TFFHE mechanism completely depends on the circuit depth (l) and parameter of security (d), which can be further improved using advanced homomorphic evaluation techniques developed in [15]. The steos involved in the process of implementing TFFHE mechanism is listed as follows.

TFFHE. Setup: Determine the modulus of monotonically decreasing ladder using the circuit depth (l) and parameter of security (d) using Equation (2)

$$\text{TFFHE.Setup} = \mathrm{q_{MF9l)}}(\mathrm{l, d, n}, \chi) \tag{2}$$

TFFHE. KeyGen: For every $l = m_l$ decreasing to 0 do the following

Execute $SC_{Key(i)}$ depending on Equation (3)

$$SC_{Key(i)} \longleftarrow Encrypt(Key_{Gen-secret}(l^*)) \text{ with } SC_{Key(i)} = \{sc_i\} \tag{3}$$

Execute $PU_{Key(i)}$ depending on Equation (4)

$$PU_{Key(i)} \longleftarrow Encrypt(Key_{Gen-Public}\,(sc_i)) \text{ with } SC_{Key(i)} = \{AK_i\} \tag{4}$$

Initialize the value of $SC_{Key(i)} \longleftarrow SC_{Key(i)} \otimes SC_{Key(j)}$ where $SC_{Key(j)} \in Z_q^{n-1}$ (5)

Execute the key switching modules depending on Equation (6)

$$\alpha_{SC_{Key(i)}=\{sc_i\}} \longleftarrow Key_{Gen}(PU_{Key(i)}, AK_i) \tag{6}$$

When $PU_{Key(i)} = \alpha_{SC_{Key(i)}=\{sc_i\}}$ with $0 \leq l \leq d - 1$.

Then generate the key based on $PU_{Key(i)}$ and $SC_{Key(i)}$ (7)

TFFHE. Encr ($i_m, PU_{Key(i)}$): Use the input message i_m and encrypt based on Equation (8)

$$\text{Encr}\ (i_m, PU_{Key(i)} = \alpha_{SC_{Key(i)}=\{sc_i\}}(PU_{Key(i)}, i_m\ ) \tag{8}$$

TFFHE. Decr ($C_m, SK_{Key(i)}$): Use the ciphertext C_m and decrypt based on Equation (9)

$$\text{Decr}(C_m, SK_{Key(i)}) = \alpha_{SC_{Key(i)}=\{sc_i\}}(SK_{Key(i)}, C_m\ ) \tag{9}$$

TFFHE.Add ($\boldsymbol{C_{m(1)}, C_{m(2)}, PU_{Key(i)}}$): When there are two ciphertexts $\boldsymbol{C_{m(1)}}$ and $\boldsymbol{C_{m(2)}}$ under the impact of the secret key $SK_{Key(i)}$. Then the function of Refresh. FHE can be applied over the considered ciphertexts for the objective of computing two identical ciphetexts for attaining the computation of $C_{M(t)} = C_{m(1)} \otimes C_{m(1)}$ with the constraints presented in Equation (10)

$$C_{M(t)} = \left[\frac{2}{m}\left(C_{m(1)} \otimes C_{m(1)}\right)\right] \tag{10}$$

3.2 *Preparation of data*

In this phase, TFFHE method is responsible for supporting the process of encrypting a vector of multiple number of plaintexts into a single ciphertext that possesses slot-wise operations. But plaintexts need to be packed essentially from a given database for attaining efficient computation process. This data preparation phase adopted two different encoding methods depending on the data size. Let us consider there are n instances such as $x_1, x_2, x_3, , \ldots . x_k$ with k number of features and their associated labels $y_1, y_2, y_3, , \ldots . y_k$ represented as $y_n \in [-1,1]$. Then the input matrix $X \in R^{d\times(n+1)}$is determined using zero columsn as specified in Equation (11)

$$X = [0 \quad x_1 \quad x_2 \quad x_3 \quad \cdots \quad x_k] \tag{11}$$

3.3 *Phase of training*

In the phase of training the least square approaches are used for handling the linear system such that distinct solutions are defined when $\mathrm{X^T X}$ is a positively defined. But it is numerically unstable and more expensive to determine the exact least square solution since it necessitates computation depending on $(\mathrm{X^T X})^{-1}$. Hence the method of gradient descent is used for training the secure SVM model, this innovativeness is mainly due to the computation facility offered by the TFFHE scheme which addresses the problem of approximation erros. Thus finally, a method of gradient descent with TFFHE approach is used for securing the training and inference phase of the SVM model.

4. Results and Discussion

The experiments of this proposed TFFHE-SVM model and the baseline approaches are evaluated using the machine that possesses a configuration of an Intel Xeon CPU E5-2660 v3 @ 2.60GHz. which is conducted for 10 number of iterations with the help of gradient descent employed during the implementation

of both the models used for comparison. In specific, Table 1 & 2 and Figure 1,2,3 & 4 presents the performance comparison results of the proposed TFFHE-SVM model with the baseline approaches used for evaluation on the Mean Time of Iteration (Sec), Accuracy (%), Computation Time (Sec.), Privacy Preservation Degree (%) parameters.

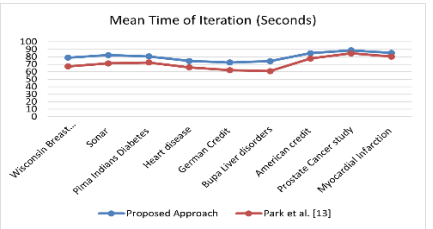

Figure 1
Performance Comparison of TFFHE-SVM in terms of Mean Time

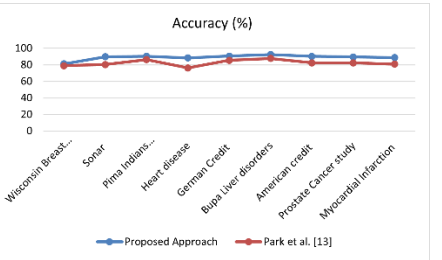

Figure 2
Performance Comparison of TFFHE-SVM in terms of Accuracy (%)

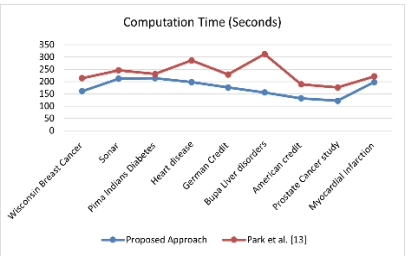

Figure 3
Performance Comparison of TFFHE-SVM in terms of Comp. Time (Sec.)

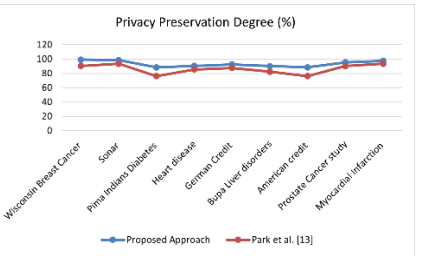

Figure 4
Performance Comparison of TFFHE-SVM in terms of Privacy Preservation Degree (%)

Table 1
Performance evaluation using mean time of iterations and accuracy

Dataset	Mean Time of Iteration (seconds)		Accuracy (in %)	
	Proposed Approach	Park et al. [13]	Proposed Approach	Park et al. [13]
Wisconsin Breast Cancer	78.93	67.21	81.26	78.95
Sonar	82.38	71.26	89.76	80.32
Pima Indians Diabetes	80.54	72.34	90.42	86.32
Heart disease	74.56	65.98	88.46	76.24
German Credit	72.42	62.12	90.54	85.48
Bupa Liver disorders	74.21	60.94	92.56	87.62
American credit	84.84	77.82	90.26	82.38
Prostate Cancer study	88.62	84.56	89.54	82.41
Myocardial Infarction	85.24	80.42	88.76	80.92

Table 2
Performance evaluation using computation and privacy preservation degree

Dataset	Computation Time (seconds)		Privacy preservation degree (in %)	
	Proposed	Park et al. [13]	Proposed	Park et al. [13]
Wisconsin Breast Cancer	161	214	99.21	90.42
Sonar	212	246	98.46	93.56
Pima Indians Diabetes	214	231	88.46	76.24
Heart disease	198	286	90.54	85.48
German Credit	176	229	92.56	87.62
Bupa Liver disorders	156	312	90.26	82.38
American credit	132	189	88.46	76.24
Prostate Cancer study	122	176	95.48	90.21
Myocardial Infarction	198	221	97.54	93.46

5. Conclusion

Secure SVM Training using Privacy Preserving Homomorphic Encryption Algorithm using TFFHE is proposed for preventing inefficient operations and numeric stability during the phase of training in the encrypted domain. This TFFHE is proposed based on the improvement of GSW and its associated ring variants. With respect to real world datasets, this TFFHE-based SVM confirmed better performance on par with the state-of-the-art FHE and logistic regression classifiers used for comparison. This study to the best of the knowledge is one of the few practical algorithms which could be used for SVM model training with the integration of FHE.

References

[1] C. Song and X. Shi, "ReActHE: A homomorphic encryption friendly deep neural network for privacy-preserving biomedical prediction," *Smart Health,* vol. 32, pp. 100469 (2024).

[2] A. M. Ibrahim, M. Farouk, and M. W. Fakhr, "Privacy preserving image retrieval using multi-key random projection encryption and machine learning decryption," *J. Adv. Res. Appl. Sci. Eng. Technol.*, vol. 42, no. 2, pp. 155–174 (2024).

[3] A. Hangan, D. Lazea, and T. Cioara, "Privacy preserving anomaly detection on homomorphic encrypted data from IoT sensors," arXiv preprint, arXiv:2403.09322 (2024).

[4] U. Mehta, J. Vekariya, M. Mehta, H. Kaur, and Y. Kumar, "A review of privacy-preserving machine learning algorithms and systems," *Appl. Data Sci. Smart Syst.*, pp. 220–225 (2025).

[5] M. K. Morampudi, N. Gonthina, S. Bojjagani, N. K. Sharma, and D. Veeraiah, "Reliable and privacy-preserving multi-instance iris verification using Paillier homomorphic encryption and one-digit checksum," *Signal, Image Video Process.,* vol. 18, no. 4, pp. 3723–3735 (2024).

[6] B. Mi, J. Zhou, D. Huang, and Y. Weng, "Privacy-preserving data processing method for IoV based on homomorphic conjugacy search problem," *IEEE Trans. Intell. Transp. Syst.* (2024).

[7] H. Lee, "Blind evaluation framework for fully homomorphic encryption and privacy-preserving machine learning," arXiv preprint (2024).

[8] H. Shin, J. Choi, D. Lee, K. Kim, and Y. Lee, "Fully homomorphic training and inference on binary decision tree and random forest," Cryptology ePrint Archive (2024).

[9] J. Yan, Z. Zhang, Y. Wang, H. Li, X. Liu, and X. Zhang, "Verifiable and privacy-preserving online diagnosis based on multiclass SVM and CKKS leveled homomorphic encryption," in Proc. 27th Int. Conf. Comput. Supported Cooperative Work Des. (CSCWD), pp. 2794–2799 (May 2024).

[10] A. Falcetta and M. Roveri, "EVAD: Encrypted vibrational anomaly detection with homomorphic encryption," *Neural Comput. Appl.,* vol. 36, no. 13, pp. 7359–7372 (2024).

[11] C. Guo, Y. Liu, H. Zhang, L. Wang, and J. Chen, "A novel biometric authentication scheme with privacy protection based on SVM and ZKP," *Comput. Secur.,* pp. 103995 (2024).

[12] R. Wu, B. Wang, and Z. Zhao, "Privacy-preserving medical diagnosis system with Gaussian kernel-based support vector machine," *Peer-to-Peer Netw. Appl.,* pp. 1–16 (2024).

[13] S. Park, J. Byun, J. Lee, J. H. Cheon, and J. Lee, "HE-friendly algorithm for privacy-preserving SVM training," *IEEE Access,* vol. 8, pp. 57414–57425 (2020).

[14] X. Hu, Z. Li, Z. Wang, and X. Lu, "ALT: Area-efficient and low-latency FPGA design for torus fully homomorphic encryption," *IEEE Trans. Very Large Scale Integr.* (VLSI) Syst. (2024).

[15] B. Wei and X. Lu, "Improved homomorphic evaluation for hash function based on TFHE," *Cybersecurity*, vol. 7, no. 1, pp. 14 (2024).

Received November, 2024