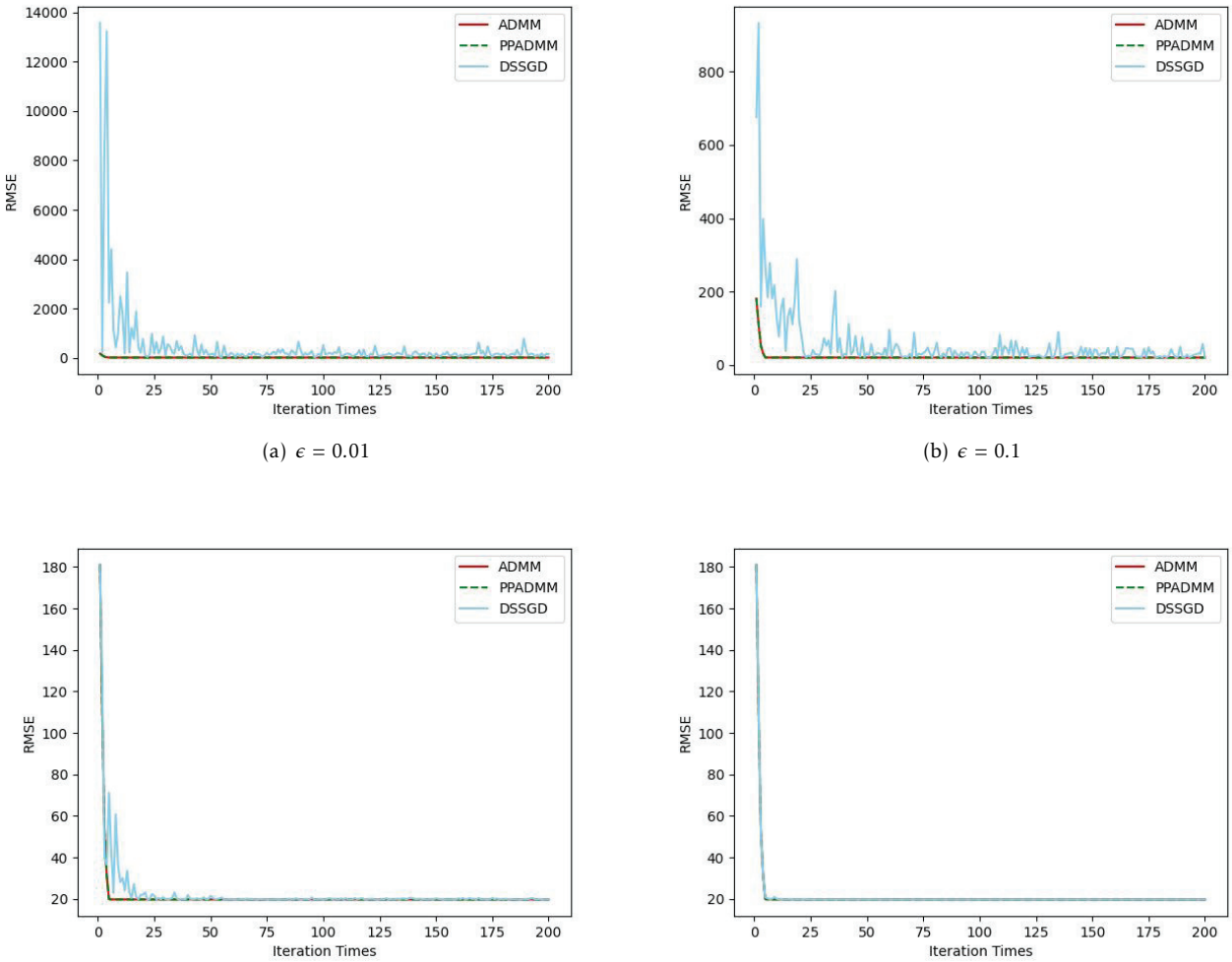


(a) $\epsilon = 0.01$ (b) $\epsilon = 0.1$

(c) $\epsilon = 1$ (d) $\epsilon = 10$

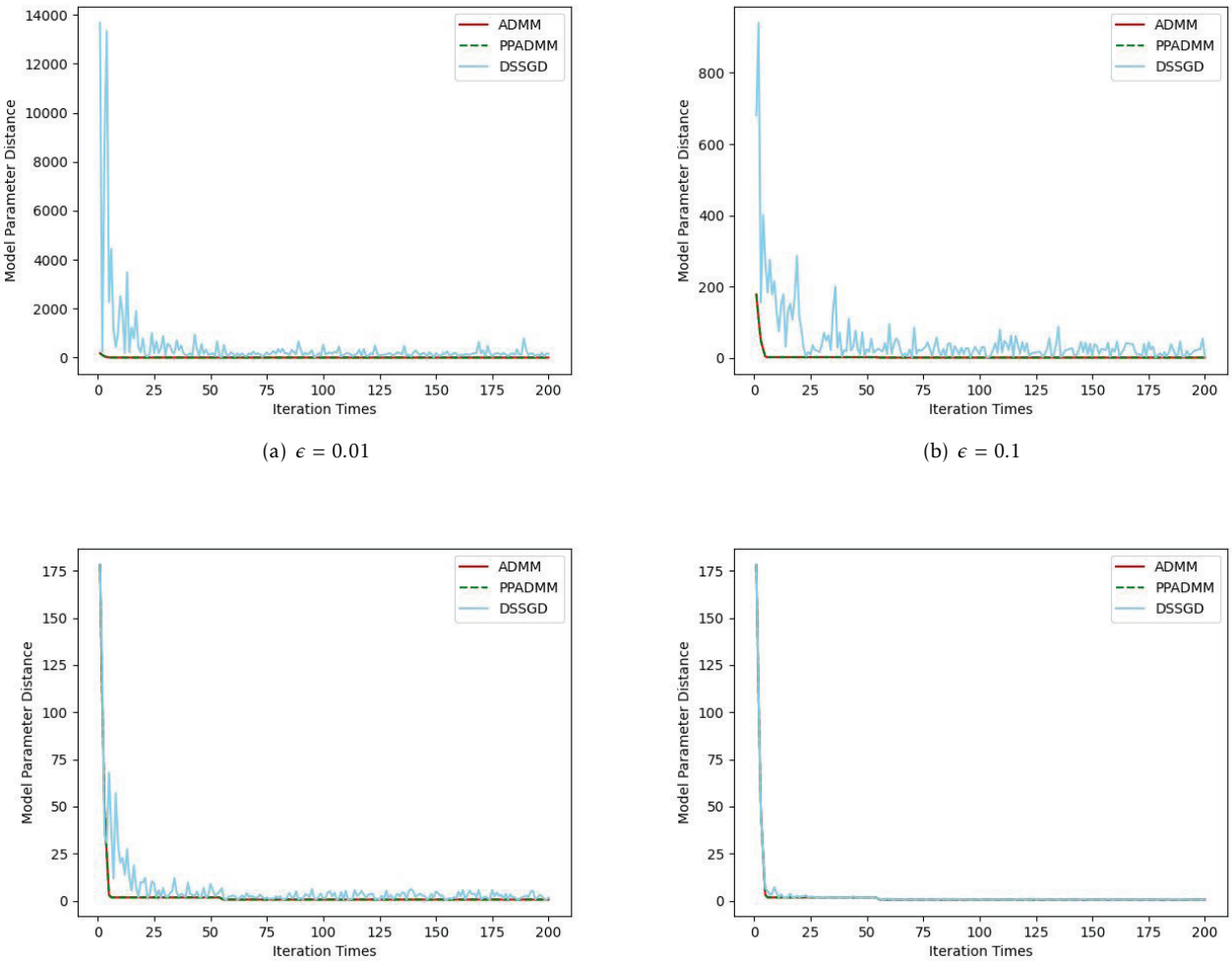
Figure 3. Relationship between RMSE and the number of iteration of ADMM, PPADMM, and DSSGD.

iteration number. We consider the actual running time of the algorithm as the computation cost in this experiment.

5.3. Experimental Results

We compare three algorithms: the original ADMM-Lasso, our algorithm PPADMM, and a classic DP-based privacy preserving collaborative learning algorithm, DSSGD [12], with respect to their performance on model accuracy and computational efficiency using the metrics in Section 5.2.

We measure both RMSE and MPED to evaluate the accuracy of the regression models trained by all these three algorithms with different privacy parameters $\epsilon = 0.01, 0.1, 1, 10$, respectively. The result is shown in Fig. 3. We can see that our algorithm PPADMM has the same RMSE as the original ADMM-Lasso algorithm. This is because during the secure aggregation process of PPADMM, the primary server S_1 removes the aggregated noise with the help of the auxiliary server S_2 to obtain the exact original global model parameters, thus the final model produced by PPADMM and ADMM-Lasso are identical. Since in DSSGD the participants perturbs the local parameters with Laplace noise in every iteration, it results in a model that predicts labels of data with a much higher RMSE compared with PPADMM and ADMM-Lasso. For example, RMSE of DSSGD can still be larger than 100 when $\epsilon = 0.1$ after it converges, while PPADMM and ADMM-Lasso has a RMSE close to 20, which is the standard deviation of the gaussian noise of the dataset. The convergence speed of PPADMM and ADMM-Lasso is also faster than that of DSSGD, according to Fig. 3. We can also see that there is an inherent trade-off between



(a) $\epsilon = 0.01$

(b) $\epsilon = 0.1$

(c) $\epsilon = 1$

(d) $\epsilon = 10$

Figure 4. Relationship between MPED and iteration number under ADMM, PPADMM, and DSSGD

the data privacy and model utility in a DP-based mechanism. The RMSE of predicted labels generated by DSSGD decreases as the privacy parameter ϵ increases, indicating that the accuracy of the model increases as the privacy guarantee decreases. Such trade-off does not appear in PPADMM and ADMM-Lasso, as RMSE of labels predicted in both mechanisms are independent of the privacy parameter ϵ.

We also measure how Model Parameter Euclidean Distance (MPED) of three algorithms (see Eq. (10)) change with the iteration number as the other criteria for model accuracy. The result is shown in Fig. 4. We can find that similar to the measurement of RMSE, PPADMM and ADMM-Lasso generate smaller MPED under different privacy parameters ϵ compared to DSSGD, which means our algorithm trains a model closer to the true distribution of original dataset.

Fig. 5 shows the computation time of three algorithms. The x-axis represents the iteration number that the algorithm runs, and the y-axis represents the real-world computation time of the algorithm spends with such number of iterations. We can see that PPADMM has a higher computation cost compared with ADMM-Lasso and DSSGD. The main reason that PPADMM incurs higher computational cost is that even though we apply SHA-256 as our cryptographic hash function, and it is more efficient than cryptographic techniques such as Yao's Garbled Circuit and homomorphic encryption used in other existing cryptographic-based privacy preserving collaborative algorithms, SHA-256 itself is still more time-consuming compared with the generation of Laplace noise in DSSGD. Although our algorithm has a higher computation cost than DSSGD, we can see that the difference between these two is not

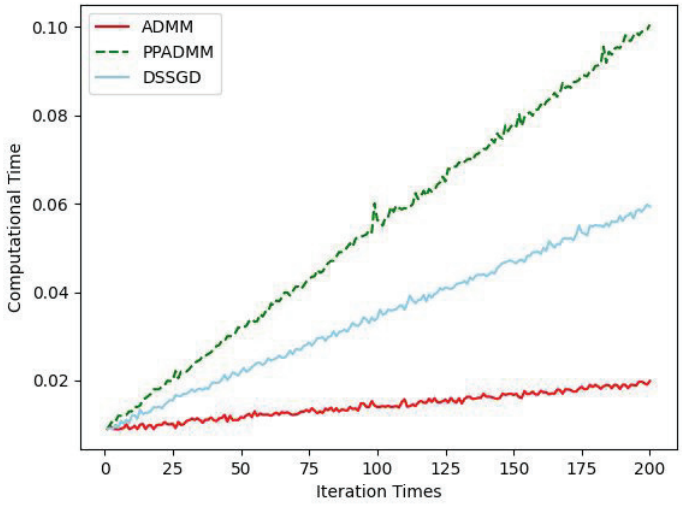


Figure 5. Comparison of the computation time of ADMM, PPADMM, and DSSGD

unacceptably large, and unlike other cryptographic-based mechanisms [6, 8, 11], our algorithm results in a linear-growth computational cost. We let participants and S_2 apply SHA-256 to generate secure hash digests at each iteration in our implementation, allowing S_2 to change the secret key shared with a participant during the training process. An more computationally efficient approach for participants and S_2 is that they can pre-compute the hash digest before the collaborative learning starts, if they agree on not changing the shared secret key in the middle of the training.

Since the information leakage of participants' local datasets caused by exchanged intermediates in collaborative learning is fairly complex, and it varies corresponding to different types of inference attacks that the adversary launches, there is yet no universal criterion that quantitatively measures such privacy leakage in a collaborative learning system. We argue that our algorithm provides stronger privacy guarantee than DP-based mechanisms, since the intermediate local parameters exchanged in PPADMM is merely the cryptographic hash digest of the original ones. It is computationally infeasible for the adversary to acquire the original local parameters if he does not know the secret key shared between participants and S_2, due to the one-way property and collision-free property of the cryptographic hash function. On the other hand, DP-based mechanisms provide privacy protection by adding noise generated from a certain distribution such as Laplace distribution. Such distribution can be estimated through multiple rounds of observation on the perturbed intermediates, and such estimation could be used to reduce the impact of perturbation on these intermediates and infer the distribution of original intermediates.

6. Conclusions and Future Work

In this paper, we introduce a novel privacy-preserving collaborative learning mechanism based on secure SUM aggregation via two non-colluding servers. Our solution allows the server to receive accurate aggregated local model update in each iteration without learning any individual participant's local model update and can achieve the same level of accuracy of standard collaborative learning mechanisms. Built upon efficient cryptographic primitives, the computation cost of our mechanism is also orders of magnitude lower than existing encryption-based solution. We have confirmed the efficacy and efficacy of our mechanism through experiment studies.

References

[1] R. Shokri, M. Stronati, C. Song, and V. Shmatikov, "Membership inference attacks against machine learning models," in *2017 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2017, pp. 3–18.

[2] B. 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*, 2017, pp. 603–618.

[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] L. Melis, C. Song, E. De Cristofaro, and V. Shmatikov, "Exploiting unintended feature leakage in collaborative

learning," in *2019 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2019, pp. 691–706.

[5] J. Geiping, H. Bauermeister, H. Dröge, and M. Moeller, "Inverting gradients - how easy is it to break privacy in federated learning?" in *Advances in Neural Information Processing Systems*, vol. 33, 2020.

[6] A. Gascón, P. Schoppmann, B. Balle, M. Raykova, J. Doerner, S. Zahur, and D. Evans, "Secure linear regression on vertically partitioned datasets." *IACR Cryptol. ePrint Arch.*, vol. 2016, p. 892, 2016.

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

[8] C. Zhang, M. Ahmad, and Y. Wang, "Admm based privacy-preserving decentralized optimization," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 3, pp. 565–580, 2018.

[9] V. Nikolaenko, U. Weinsberg, S. Ioannidis, M. Joye, D. Boneh, and N. Taft, "Privacy-preserving ridge regression on hundreds of millions of records," in *2013 IEEE Symposium on Security and Privacy*. IEEE, 2013, pp. 334–348.

[10] G. Danner and M. Jelasity, "Fully distributed privacy preserving mini-batch gradient descent learning," in *IFIP International Conference on Distributed Applications and Interoperable Systems*. Springer, 2015, pp. 30–44.

[11] K. Bonawitz, V. Ivanov, B. Kreuter, A. Marcedone, H. B. McMahan, S. Patel, D. Ramage, A. Segal, and K. Seth, "Practical secure aggregation for privacy-preserving machine learning," in *proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, 2017, pp. 1175–1191.

[12] R. Shokri and V. Shmatikov, "Privacy-preserving deep learning," in *Proceedings of the 22nd ACM SIGSAC conference on computer and communications security*, 2015, pp. 1310–1321.

[13] J. Hamm, A. C. Champion, G. Chen, M. Belkin, and D. Xuan, "Crowd-ml: A privacy-preserving learning framework for a crowd of smart devices," in *2015 IEEE 35th International Conference on Distributed Computing Systems*. IEEE, 2015, pp. 11–20.

[14] T. Zhang and Q. Zhu, "Dynamic differential privacy for admm-based distributed classification learning," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 1, pp. 172–187, 2016.

[15] E. Nozari, P. Tallapragada, and J. Cortés, "Differentially private distributed convex optimization via functional perturbation," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 1, pp. 395–408, 2016.

[16] Y. Wang, M. Hale, M. Egerstedt, and G. E. Dullerud, "Differentially private objective functions in distributed cloud-based optimization," in *2016 IEEE 55th Conference on Decision and Control (CDC)*. IEEE, 2016, pp. 3688–3694.

[17] B. Cyphers and K. Veeramachaneni, "Anonml: Locally private machine learning over a network of peers," in *2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*. IEEE, 2017, pp. 549–560.

[18] S. Boyd, N. Parikh, and E. Chu, *Distributed optimization and statistical learning via the alternating direction method of multipliers*. Now Publishers Inc, 2011.

[19] J. Liu, J. Yang, L. Xiong, and J. Pei, "Secure skyline queries on cloud platform," in *2017 IEEE 33rd international conference on data engineering (ICDE)*. IEEE, 2017, pp. 633–644.