SCIS 2022 2022 Symposium on
Cryptography and Information Security
Osaka, Japan & Online, Jan. 18 – 21, 2022
The Institute of Electronics,
Information and Communication Engineers

プライバシを保護したRNNによる非対話型文書分類
One round classification of encrypted text data using RNN

齋藤 拓巳 *† 岡 響 † 中橋 彬 † 尾形 わかは *
Takumi Saito Hibiki Oka Akira Nakahashi Wakaha Ogata

キーワード 秘密計算, 完全準同型暗号, 機械学習, 自然言語処理, リカレントニューラルネットワーク

近年, 機械学習が急速に発展し, 人間の意思決定やその補助をするシステムとして, 様々な分野で広く採用されている. 大量のデータを処理する機械学習では, メモリや速度の面で高性能なコンピュータを要求される. そうした観点から, クラウド上で機械学習を提供するサービス (Machine Learning as a Service, MLaaS) が登場している. MLaaS では, 学習済みのモデルをサービス提供者のクラウドが持ち, クライアントがデータを送信し, そのデータをサービス提供者が処理する. しかし, クライアントがサービス提供者を信用することが前提となっており, クライアントが送信するデータに機密性やプライバシ上の懸念がある場合, サービスを利用することができない. そこで, プライバシを守りながら MLaaS を行う研究が行われており, Secure Multiparty Computation(SMC) や Homomorphic Encryption(HE) などを用いた手法が提案されている. HE は暗号化した状態で計算が可能な暗号方式であり, これを用いて画像分類を行った研究が知られている [3].

画像と同様, 文書も機密性やプライバシ上の懸念がある場合が多い. HE を利用して自然言語処理 (NLP) タスクを行う研究では, リカレントニューラルネットワーク (RNN) を利用しない方式 [1] と, サーバとクライアント間での対話によって RNN を実現した方式 [4] が知られている. 連続的なデータを処理することに適している RNN は, より高度な処理が可能な Long Short Term Memory(LSTM), Gated Recurrent Unit(GRU) などの発展を持つ, NLP の重要なモデルである. そのため, RNN を用いる方式は今後の発展性を期待できる. しかし, 後者の方式では, 1) サーバが途中結果の暗号文をクライントに送信, 2) クライアントは秘密鍵を用いてこれを復号し再暗号化してサーバに返送, という複数回の通信を行わなければならず, 本来の MLaaS とは異なっている.

本論文では, クライアントとサーバの通信が 1 回のみの非対話型で, プライバシを保護した RNN による文書分類を考える. 非対話で長い連続データを扱うためには, HE に基づくノイズを Bootstrapping によって制御する必要がある. 本研究で利用する HE は TFHE[2] に基づくものであり, 任意の関数を Lookup-Table(LUT) として評価可能な Bootstrapping (Programmable Bootstrapping, PBT) を利用する. PBT を持たない HE を利用する機械学習の活性化関数は多項式で近似することが一般的であるが, 本研究では PBT を活用した活性化関数にて精度を保つアプローチを示す.

* 東京工業大学, 〒 152-8552 東京都目黒区大岡山 2-12-1, Tokyo Institute of Technology, 2-12-1 O-okayama Meguro-ku Tokyo, 152-8550 Japan {saito.t.bi, ogata.w.aa}@m.titech.ac.jp
† アイマトリックス研究所株式会社, 〒 216-0004 神奈川県川崎市宮前区鷺沼 3-2-6 鷺沼センタービル 4F, imatrix laboratory corp, Saginuma Center Building 4F 3-2-6 Saginuma Miyamae-ku Kawasaki-shi Kanagawa, 216-0004 Japan {hibiki.oka, akira.nakahashi}@imatrix.co.jp

参考文献

[1] Ahmad Al Badawi, Louie Hoang, Chan Fook Mun, Kim Laine, and Khin Mi Mi Aung. Privft: Private and fast text classification with homomorphic encryption. *IEEE Access*, 8:226544–226556, 2020.

[2] Ilaria Chillotti, Marc Joye, and Pascal Paillier. Programmable bootstrapping enables efficient homomorphic inference of deep neural networks. *IACR Cryptol. ePrint Arch.*, 2021:91, 2021.

[3] Ran Gilad-Bachrach, Nathan Dowlin, Kim Laine, Kristin Lauter, Michael Naehrig, and John Wernsing. Cryptonets: Applying neural networks to encrypted data with high throughput and accuracy. In *International conference on machine learning*, pages 201–210. PMLR, 2016.

[4] Robert Podschwadt and Daniel Takabi. Classification of encrypted word embeddings using recurrent neural networks. In *PrivateNLP@ WSDM*, pages 27–31, 2020.