2023 年度年次報告書
基礎理論とシステム基盤技術の融合による Society 5.0 のための基盤ソフトウェアの創出
2021 年度採択研究代表者

田浦 健次朗

東京大学 大学院情報理工学系研究科
教授

実応用に即したプライバシー保護解析とセキュアデータ基盤

主たる共同研究者:
曹 洋（北海道大学 大学院情報科学研究院 准教授）
花岡 昇平（東京大学 医学部附属病院 専任講師）
塙 敏博（東京大学 情報基盤センター 教授）
吉川 正俊（大阪成蹊大学 データサイエンス学部 教授）

研究成果の概要

[1]　プライバシー保護を強制可能なプログラミングのフレームワークとして実応用を重んじて Python をベースにし, numpy, pandas, 機械学習フレームワーク(bintorch)にその機構を導入した。花岡 G との共同でグローバル差分プライバシーを用いた画像分類タスクの実装を行い、同成果を ISEC 研究会論文として発表した。

[2] クロスサイロ連合学習における各サイロの貢献度を評価するために, 効率的で安全なシャープレー値計算手法として 2 サーバプロトコルである SecSV を開発した. SecSV の特徴は, ハイブリッドプライバシー保護方式の利用, 効率的かつ安全な行列乗算法, 評価精度に大きな影響を与えないテストサンプルの戦略的な識別とスキップである.

[3] 時空間連合学習におけるリスクを研究し、防御手法を提案した。時空間連合学習では、サーバーがクライアントの軌跡データを再構築する可能性があり、これは「勾配逆転攻撃」として知られている。既存の手法は時空間データには効果がないため、我々は新しい攻撃手法を提案し、DP を基づく防御手法も評価した。

[4] 差分プライバシーの医用画像への応用について、3 つのジャーナル論文を発表した. すなわち 3) フローベース深層学習生成モデルを用いた医用画像の局所差分プライバシー(LDP)化と, 4) 3D 拡散モデル(DDPM)を用いて差分プライバシーを適用しつつ画像生成を行う研究、5) DP による個人情報保護に資するため、患者の表データと画像データを関連を保ったまま同時生成する研究である.

[5] 公開鍵に基づく暗号化を行うユーザレベルファイルシステム SecFS の基本方針を検討しプロトタイプ実装を行った. SGX を用いることで権限のあるユーザ以外にはファイルへのアクセスもできない上, 他のユーザからプロセスメモリの覗き見も不可能になる. また共有取り消しのための機構として SGX の適用範囲を検討した. また SGX に加えて仮想基盤による隔離を実現する TDX の導入と GPU と組み合わせることで実用的な機械学習プラットフォームへの適用について検討した.

【代表的な原著論文情報】

1) Lele Zheng, Yang Cao, Renhe Jiang, Kenjiro Taura, Yulong Shen, Sheng Li and Masatoshi Yoshikawa. Enhancing Privacy of Spatiotemporal Federated Learning against Gradient Inversion Attacks. Proc. of the 29th International Conference on Database Systems for Advanced Applications (DASFAA), 2024.
2) Shuyuan Zheng, Yang Cao and Masatoshi Yoshikawa. Secure Shapley Value for Cross-Silo Federated Learning. Proc. VLDB Endow. 16(7): 1657-1670, 2023.
3) Hisaichi Shibata, Shouhei Hanaoka, Yang Cao, Masatoshi Yoshikawa, Tomomi Takenaga, Yukihiro Nomura, Naoto Hayashi and Osamu Abe. Local differential privacy image generation

using flow-based deep generative models. Applied Sciences 13 (18), 10132.
4) Hisaichi Shibata, Shouhei Hanaoka, Takahiro Nakao, Tomohiro Kikuchi, Yuta Nakamura, Yukihiro Nomura, Takeharu Yoshikawa and Osamu Abe. Practical Medical Image Generation with Provable Privacy Protection based on Denoising Diffusion Probabilistic Models for High-resolution Volumetric Images. Applied Sciences 14 (8) 3489.
5) Tomohiro Kikuchi, Shouhei Hanaoka, Takahiro Nakao, Tomomi Takenaga, Yukihiro Nomura, Harushi Mori and Takeharu Yoshikawa. Synthesis of Hybrid Data Consisting of Chest Radiographs and Tabular Clinical Records Using Dual Generative Models for COVID-19 Positive Cases. J Imaging Inform Med. 2024 Feb 13.