「イノベーション創発に資する人工知能基盤技術の創出と統合化」
平成 28 年度採択研究代表者

H28 年度
実績報告書

盛合 志帆

国立研究開発法人情報通信研究機構サイバーセキュリティ研究所
室長

複数組織データ利活用を促進するプライバシー保護データマイニング

§1. 研究実施体制

(1)盛合グループ
① 研究代表者:盛合 志帆 (情報通信研究機構サイバーセキュリティ研究所、室長)
② 研究項目
・プライバシー保護データマイニング手法の開発

(2)小澤グループ
① 主たる共同研究者:小澤 誠一 (神戸大学大学院工学研究科、教授)
② 研究項目
・準同型暗号を用いた暗号化データのロジスティック回帰分析手法の開発
・機械学習ベンチマークデータを用いた性能評価

(3)菅原グループ
① 主たる共同研究者:菅原 貴弘 (株式会社エルテス、代表取締役)
② 研究項目
・利用シーンの想定および収集するログの種類選定

§2. 研究実施の概要

本研究課題では、複数組織での横断的データ利活用を促進するためのプライバシー保護データマイニング、具体的には暗号技術や人工知能技術を活用し、プライバシーやセキュリティを保護した状態で高速にデータ分析や異常検知を行う技術の研究開発を行う。この技術を活用し、金融分野における不正送金検知や顧客に合わせた金利決定の支援といった課題の解決を目指す。

平成 28 年度は、国立研究開発法人情報通信研究機構(NICT)、国立大学法人神戸大学、及び株式会社エルテスの 3 研究グループで連携して本研究課題の立ち上げを行った。プロジェクト発足にあたっての最重要課題は、本研究課題で取り組む金融分野における課題を解決するために必須となる金融取引データ等の入手であった。データを提供頂く金融機関と議論を進め、どのようなデータが課題解決に有効で、どのような形で提供いただくかを検討し、関連する法令・ガイドラインに従って提供を受けることになった。平成 28 年度に実施した研究概要は以下の通りである。

NICT(盛合グループ)では、プライバシー保護データマイニング手法の開発に着手した。具体的には、暗号化データに対して適用できる機械学習アルゴリズムの拡充を目指し、暗号化データ上で実用的に計算できる演算処理を拡充し、効率的な準同型内積演算の一般的構成を提案した。また、多数の参加者が持つデータセットを互いに秘匿したまま深層学習を行うシステムを提案し、上記 2 件について暗号と情報セキュリティシンポジウム 2017 にて発表を行った[1][2]。ビジネスシーンを考慮した適切なセキュリティモデルの設計では、エルテス(菅原グループ)とともに金融機関関係者にヒアリングを行い、データ解析時にどのような脅威を想定して暗号化や匿名化等のデータ保護対策を行うべきか検討した。また、神戸大(小澤グループ)と連携して暗号化データの分類・予測を行う機械学習アルゴリズムの研究開発を行った。

神戸大学(小澤グループ)では、プライバシー保護目的で暗号化された大量のデータを代理計算サーバ(例えば、商用のクラウドサービス)上で高速処理し、個人情報を漏洩させることなく、パターン認識、異常診断、時系列予測など知的データ処理を行うための機械学習アルゴリズムを開発している。平成 28 年度は, NICT の青野らにより提案された加法準同型暗号を用いたロジスティック回帰モデルを拡張し、暗号データに対して、多クラス・パターン認識問題に適用可能なロジスティック回帰分析手法を開発した。また、機械学習ベンチマークデータとして知られる Parkinson データ、Iris データ、Hand-written Digit データを用いた性能評価実験を行い、データを暗号化した状態であっても、一定の条件範囲内では十分なパターン認識の精度が得られることを示した。

エルテス(菅原グループ)では、金融機関とのネットワークを活かして、利用シーンの想定および、収集するデータ選定のために意見交換会を実施した。また、研究材料となる正解データを含む各種金融取引データを受領した。

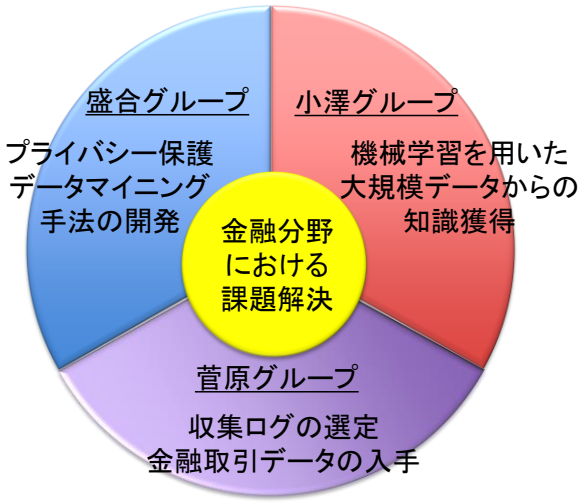

図 1: 平成 28 年度のチーム全体の研究実施概要

[1] 林 卓也, 青野 良範, レ チュウ フォン, 王 立華,「効率的な準同型内積演算の一般的構成」, 暗号と情報セキュリティシンポジウム 2017, 3F2-1, 2017.1.

[2] 青野 良範, 林 卓也, レ チュウ フォン, 王 立華, 盛合 志帆,「加法準同型暗号を用いたプライバシー保護深層学習」, 暗号と情報セキュリティシンポジウム 2017, 1C1-3, 2017.1.