情報通信研究機構

NICT 先端研究 ⑬

機械学習 暗号化のまま実現

プライバシー保護データ分析技術の研究開発の目的は、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ーを保護した状態で高速にビッグデータ向けの分析を行うことである。本研究で開発した技術を使えば、データ処理を行う第三者にデータの内容を開示せずに済み、プライバシーを保護できるようになると期待される。

現在、データマイニングサービスの実施を否定する動きもあり、どのように解決すべきか課題となっている。

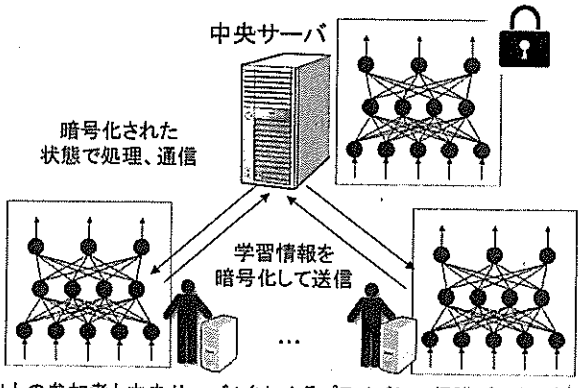
私たちは、その課題を解決するために、2015年からプライバシー保護データ分析の技術を開発してきた。15年に、暗号化したままデータを処理する「準同型暗号」において、暗号化したデータのセキュリティーレベルを上げる技術（SPHERE、スフィア）を開発した。この技術により、プライバシーに関わるデータを個人の寿命よりも長い期間にわたって安全に利活用することが可能になる。

翌年、SPHERE方式を用い、データを暗号化した状態でロジスティック回帰分析を高速に行う手法を開発した。この技術により、大量のデータを暗号化したまま複数のグループに分類することが可能になり、シミュレーションによって、サーバー上で1億件のデータを30分以内で分析可能であることが確認できた。

現在は、データを開示せずに深層学習が可能となる技術の開発に取り組んでいる。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い結果を得ることが可能なプライバシー保護深層学習システムを提案した（図）。効率性を検討するために、金融データに関わる28万件強の取引レコードを用いて実験し、数分程度で提案システムの学習および推測が完了することを確認した。また、各レコードを約1ミリ秒で推測できた。

今後、提案した技術を社会実装していくとともに、金融分野のデータを含め、多種多様の実データに適用する予定である。

（火曜日に掲載）



N人の参加者と中央サーバ1台によるプライバシー保護データ分析（分散協調学習）

プライバシー保護深層学習のイメージ

サイバーセキュリティ研究所セキュリティ基盤研究室主任研究員 レ・チュウ・フォン

大学院博士課程修了後、東京工業大学研究員を経て、09年11月、NICT入所。暗号プリミティブの設計、プライバシー保護データ分析に関する研究開発に従事。