「ビッグデータ統合利活用のための次世代基盤技術の創出・体系化」
平成 25 年度採択研究代表者

H29 年度
実績報告書

佐久間　淳

筑波大学大学院システム情報工学研究科
教授

自己情報コントロール機構を持つプライバシ保護データ収集・解析基盤の構築と個別化医療・ゲノム疫学への展開

§1. 研究実施体制

(1)佐久間グループ
① 研究代表者:佐久間 淳 （筑波大学大学院システム情報工学研究科　教授）
② 研究項目
・プライバシ保護データ収集・解析基盤の構築

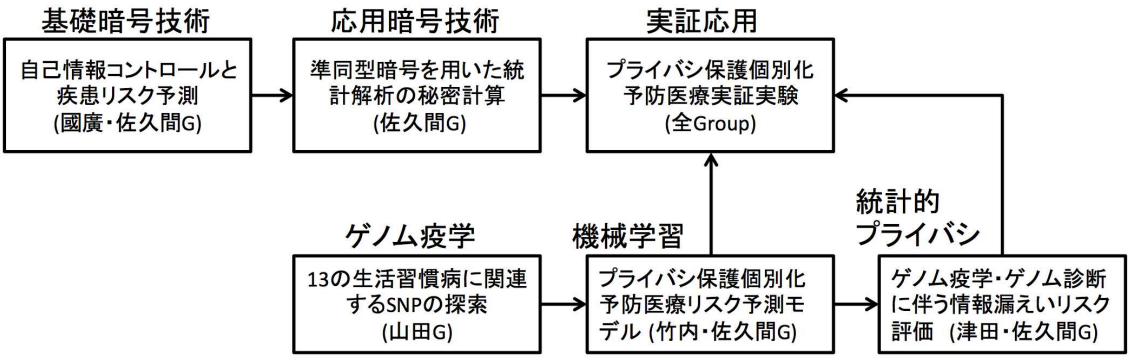
(2)津田グループ
① 主たる共同研究者:津田　宏治 （東京大学大学院新領域創成科学研究科　教授）
② 研究項目
・プライバシ保護ゲノム疫学

(3)竹内グループ
① 主たる共同研究者:竹内　一郎 （名古屋工業大学大学院工学研究科　教授）
② 研究項目
・プライバシ保護データ収集・解析基盤を利用した個別化医療の実証実験

(4)國廣グループ
① 主たる共同研究者:國廣　昇 （東京大学大学院新領域創成科学研究科　准教授）
② 研究項目
・プライバシ保護データ収集・解析基盤の構築

(5)山田グループ
① 主たる共同研究者:山田　芳司 （三重大学先端科学研究支援センター　教授）
② 研究項目
・エクソン全領域関連解析による心筋梗塞発症に関連する機能的遺伝子多型の同定

§2. 研究実施の概要



このプロジェクトでは、個別化ゲノム予防医療の実用化とプライバシ保護を実用レベルで両立させることを目的として、基盤・応用暗号技術、統計的プライバシ保護、機械学習技術をベースとした基礎研究とゲノム疫学研究を実施しつつ、個別化ゲノム予防医療のフィージビリティスタディを推進している。

基盤暗号技術分野では、複数のクライアントが独自の鍵で暗号化を行うことができる Multi-Key 性を持ち，ID に基づく完全準同型暗号方式の提案を行った．多数のクライエントから秘密情報を独立に収集し、秘密計算をサーバーに外部委託するアウトソーシング型秘密計算の実用化には鍵管理が障害となっていたが、それを解消する技術である。

応用暗号技術分野では、準同型暗号を用いた統計解析におけるプライバシ保護において、データそのものの保護のみならず、その情報がどのように利用されるか、その状況をコントロールする技術もあわせて必要である。今年では、(1)複数情報源からえた秘密情報を join する状況において、各区レコードがどの情報源の情報と join されることを許容するか(あるいは許容しないか)をコントロールする技術、(2)準同型暗号そのものにキーワードを埋め込み、指定されたキーワード同士の演算以外は許容しない、誤った演算要求に耐性を持つ新しい準同型暗号の提案、を行った。現実的な状況における秘密計算の実応用に資する技術である。

ゲノム疫学分野では、今までに 15,896 例のエクソームアレイによる SNPs 解析を行い，約 39 億個の SNPs 情報および 13 種類の生活習慣病に関する臨床情報・ライフスタイル情報を包括する大規模なデータベースを構築した.構築したデータベースを用いて 13 種類の生活習慣病の発症に強く関連する遺伝子群および SNPs を特定した.

機械学習分野では、ゲノム疫学の結果として得られた疾患関連遺伝子を用いて、疾患リスク予測モデリングを構築した。今年度は、被験者が入力の一部を秘匿する，あるいは曖昧なまま情報を提供する状況を想定し、その上で、リスク予測結果がどのような範囲で変わり得るかを適切に見積もる技術を開発した。

統計的プライバシ分野では、ゲノム疫学やゲノム疫学の結果として公開される情報から、秘匿されるべきゲノム情報などがどの程度推測されるのかを定量的に評価し、推測される情報を、疫学や診断の目的を損なわない範囲内で制限する方法を研究した。具体的には、(1)二分決定ダイアグラ

ム(BDD)を用いたゲノム診断結果からの個別のゲノムが推測される確率の効率的な推定手法、(2)差分プライバシを保証するカイ二乗検定手法、の二つを開発した。
実証応用として、異なる機関から収集した個人ゲノム情報と医療情報を準同型暗号で暗号化した上で集約し、プライバシを保護しつつ生活習慣病の疾患リスクを評価し、医療機関にレポーティングするシステムを構築し、複数の医療機関でフィージビリティを検証した。

代表的な原著論文

Kazuya Kakizaki, Kazuto Fukuchi, and Jun Sakuma. "Differentially Private Chi-squared Test by Unit Circle Mechanism." Proceedings of the 34th International Conference on Machine Learning, Vol. 70, pp. 1761-1770, 2017.
Kosuke Kusano, Ichiro Takeuchi, and Jun Sakuma. "Privacy-preserving and Optimal Interval Release for Disease Susceptibility." Proceedings of the 2017 ACM Asia Conference on Computer and Communications Security, pp. 532-545, 2017.
Hiroyuki Hanada, Atsushi Shibagaki, Jun Sakuma and Ichiro Takeuchi. "Efficiently Monitoring Small Data Modification Effect for Large-Scale Learning in Changing Environment", Proceedings of The 32nd AAAI Conference on Artificial Intelligence (AAAI 2018),, pp. 1314-1321, 2018.