プレスリリース
2018 年 7 月 18 日

国立研究開発法人情報通信研究機構
国立大学法人筑波大学
国立研究開発法人科学技術振興機構

プライバシーを保護したまま医療データを解析する暗号方式を実証

～中身を見なくても誤データ混入防止、医療ビッグデータの安全な利活用へ～

【ポイント】

- 暗号化した医療データの中身を見ることなく、解析対象外データの混入を防ぐ解析手法を開発
- 個人の遺伝情報と病気の罹患情報との統計的な関連性を、暗号化したまま安全に解析
- プライバシーを保護した安全なビッグデータ解析への応用に期待

国立研究開発法人情報通信研究機構（NICT、理事長: 徳田 英幸）セキュリティ基盤研究室と国立大学法人筑波大学（筑波大学、学長: 永田 恭介）は、国立大学法人三重大学 山田芳司教授のご協力の下、このたび、医療データを暗号化したまま解析することに成功し、開発した暗号方式の性能を実証しました。4,500 名程度の暗号化された医療データに対し 1 分弱で、病気の罹患情報と個人の遺伝情報との統計的な関連性を、各個人の病気の有無や遺伝情報を知ることなく、安全性を確保したままで解析できます。

また、本暗号方式では、解析中にデータの中身を見ることが許されない医療データに対して、解析対象外のデータが混在した場合でも高速に検出することができ、その解析結果が正当であることを暗号理論的に証明できました。これにより、個人のプライバシーを保護して情報漏えいを防ぎながら、医療ビッグデータを安全に利活用できるようになり、新たな診断方法や治療法の開発につながることが期待されます。

本成果は、電子情報通信学会 情報セキュリティ研究会（ISEC）（2018 年 7 月 25 日（水）～26 日（木）、札幌）での発表に先立ち、2018 年 7 月 18 日（水）に講演論文がオンライン公開されます。

【背景】

近年、医療ビッグデータ法[*1] が整備されるなど、プライバシーを保護したまま医療データを安全に活用し、新たな治療法の開発等に役立てようという動きが盛んになっています。医療データの情報漏えい等に対する安全策として、暗号化は有効であり、暗号化したままデータに関する演算が可能な暗号方式である準同型暗号[*2]を用いたプライバシー保護データ解析の研究が進められています。暗号文からはデータに関する情報が漏れないため、データを明かすことなく第三者に解析処理を依頼することや、データそのものを組織間で受け渡すことが難しい医療分野や金融分野における安全な統計処理など、様々な応用が考えられています。

一方で、医療データを暗号化すると、暗号化されているがゆえに、解析対象のデータかどうかを判定することができません。そのため、対象外のデータが統計処理に使用された場合でも検出できずに解析がそのまま行われ、誤った統計値が出力される懸念がありました。解析前に一度暗号文を復号し、解析対象データであることを確認する場合、データ解析を行う第三者にデータの内容を開示する必要があり、プライバシー上の懸念事項となります。

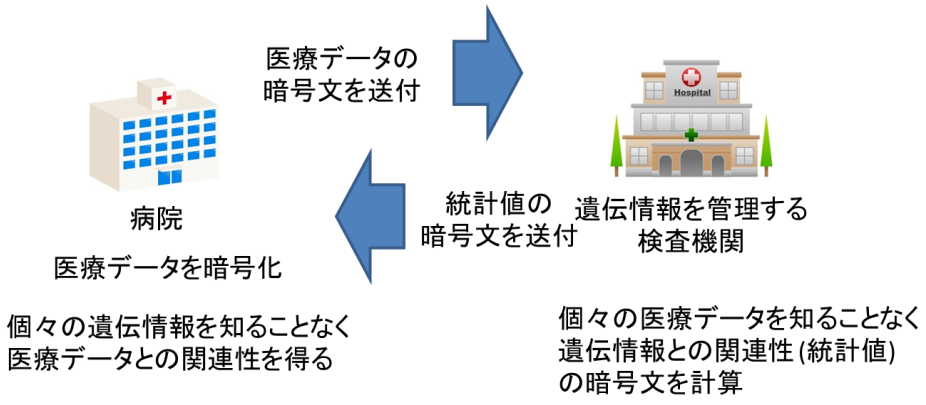

図 1. 想定シナリオ

【今回の成果】

NICT が中心となり開発した誤データ混入防止機能を持つ準同型暗号方式「まぜるな危険準同型暗号[*3]」を用

いて、今回、実際の医療データに対する解析を行いました。誤データ混入防止機能により、解析時に中身を見ることが許されない医療データに対し、データの中身を見ることなく解析対象の医療データであるかどうかを判定できます。

本実証実験では、病気の罹患情報と遺伝情報を解析対象データとし、「ある病気を罹患していること」と「ある遺伝的特徴を持つこと」との統計的な関連性を解析するシナリオを想定しました（図 1 参照）。具体的には、病院が病気の有無に関するデータを暗号化し、遺伝情報を管理する検査機関に暗号文を送付、検査機関が遺伝情報との統計的な関連性を計算することを想定しました。この際、検査機関は各個人の病気の有無を知ることなく、病院側も各個人の遺伝的特徴を知ることはありません。また、仮に別の病気の医療データの暗号文が混在した場合でも、検査機関で検出できます。さらに、検査機関は統計値の暗号文を計算するのみであり、統計値そのものも検査機関に知られることはありません。より詳細な流れは補足資料を参照ください。

本実験においては、解析値として「遺伝的特徴を持ち、かつ、病気を罹患している人の数の暗号文」を計算しました。今回、4,500 名程度のデータに対し、1 分弱で暗号化及び解析が完了すること、また、異なる病気の医療データの暗号文が混在した場合でも数ミリ秒程度で検出できることを確認しました。

本実験は、医療の発展目的への使用に関する患者の同意を得て三重大学病院が収集した匿名化された医療データを用い、三重大学内の外部のネットワークからはアクセスできない環境にて行いました。

なお、NICT は暗号化データ解析手法の研究開発及び暗号化データ解析ツールの開発を担当、筑波大学は暗号化データ解析手法の研究開発及び医療データ検定方法の検討を担当、三重大学は臨床情報及び遺伝情報をデータベース化し、本研究に提供いただきました。

【今後の展望】

本技術により、医療分野において、多くの被験者から収集したデータを、プライバシーを保護したまま解析することが可能になります。さらに、その解析結果に対象外のデータが混入していないことを暗号理論的に証明することで、解析結果の妥当性を向上させることが可能になります。これにより、新たな診断方法や治療法の早期かつ効率的な発見につながることが期待されます。

本成果は、2018 年 7 月 25 日（水）～26 日（木）に、札幌コンベンションセンターで開催される電子情報通信学会情報セキュリティ研究会（ISEC）で発表します。

論文情報

「電子情報通信学会技術研究報告 情報セキュリティ」にて、2018 年 7 月 18 日（水）に論文が公開されます。
https://www.ieice.org/ken/index/ieice-techrep-2018.html
なお、閲覧には電子情報通信学会にて技術研究報告ダウンロード権を購入する必要がありますので、ご注意ください。

江村恵太（NICT/JST），林卓也（NICT/JST），陸文傑（筑波大/JST），盛合志帆（NICT/JST），佐久間淳（筑波大/JST/理研），山田芳司（三重大/JST），まぜるな危険準同型暗号を用いた医療データに対する χ^2 独立性検定，情報セキュリティ研究会（ISEC），電子情報通信学会

＜ 本件に関する問い合わせ先 ＞
情報通信研究機構
サイバーセキュリティ研究所
セキュリティ基盤研究室
江村 恵太
Tel: 042-327-5690
E-mail: security@ml.nict.go.jp

筑波大学
コンピュータサイエンス専攻
情報学群情報科学類
佐久間 淳
E-mail: jun@cs.tsukuba.ac.jp

＜ JST 事業に関すること ＞
科学技術振興機構
戦略研究推進部
松尾 浩司
Tel: 03-3512-3526, Fax: 03-3222-2066
E-mail: crest@jst.go.jp

＜ 広報 ＞
情報通信研究機構
広報部 報道室
廣田 幸子
Tel: 042-327-6923
Fax: 042-327-7587
E-mail: publicity@nict.go.jp

筑波大学
広報
E-mail: kohositu@un.tsukuba.ac.jp

科学技術振興機構
広報課
Tel: 03-5214-8404
Fax: 03-5214-8432
E-mail: jstkoho@jst.go.jp

＜用語解説＞

***1 医療ビッグデータ法**

「医療分野の研究開発に資するための匿名加工医療情報に関する法律」。健康・医療に関する先端的研究開発及び新産業創出を促進し、健康長寿社会の形成に資することを目的として、2017 年 5 月 12 日に公布された。

***2 準同型暗号**

暗号化されたデータに対して加算と乗算を行うことができる技術。加算・乗算を組み合わせることで様々な計算を暗号化したまま行えるため、プライバシー保護データマイニングへの応用が期待されている。

***3 まぜるな危険準同型暗号**

2016 年に NICT が中心となり東京大学、筑波大学と共同で開発した誤データ混入防止機能を持つ準同型暗号方式

【関連情報】NICT お知らせ（2016 年 11 月 7 日）
「プライバシー保護データ処理技術に対して「第 19 回コンピュータセキュリティシンポジウム 2016（CSS 2016）」最優秀論文賞を受賞」http://www.nict.go.jp/info/topics/2016/11/161107-1.html

- 江村 恵太, 林 卓也, 國廣 昇, 佐久間 淳. まぜるな危険準同型暗号. コンピュータセキュリティシンポジウム（CSS）2016.
- Keita Emura, Takuya Hayashi, Noboru Kunihiro, Jun Sakuma: Mis-operation Resistant Searchable Homomorphic Encryption. AsiaCCS 2017: 215-229

補足資料

今回の実証実験の詳細

「まぜるな危険準同型暗号」では、図 2 に示すように、データとは別にキーワードも暗号化し、暗号文が同じキーワードに関連しているか否かをキーワードそのものを知ることなく判定できます。また、図 3 のように、異なるキーワードに関連した暗号文に対して準同型演算を行った場合に警告を発するとともに、警告を無視し無理やり準同型演算を行った場合でも復号時に判定可能となる機能を実現しています。さらに、今回暗号化したまま加算を行う機能をサポートしています。

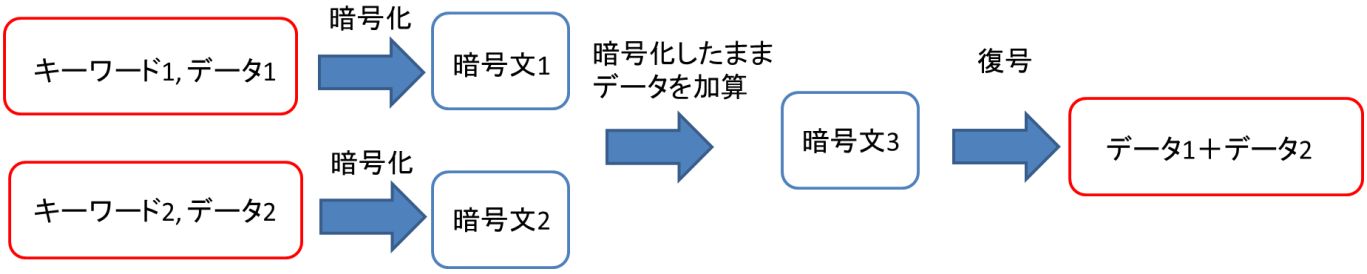

図 2. 同じキーワード（キーワード 1 = キーワード 2）に対する準同型演算

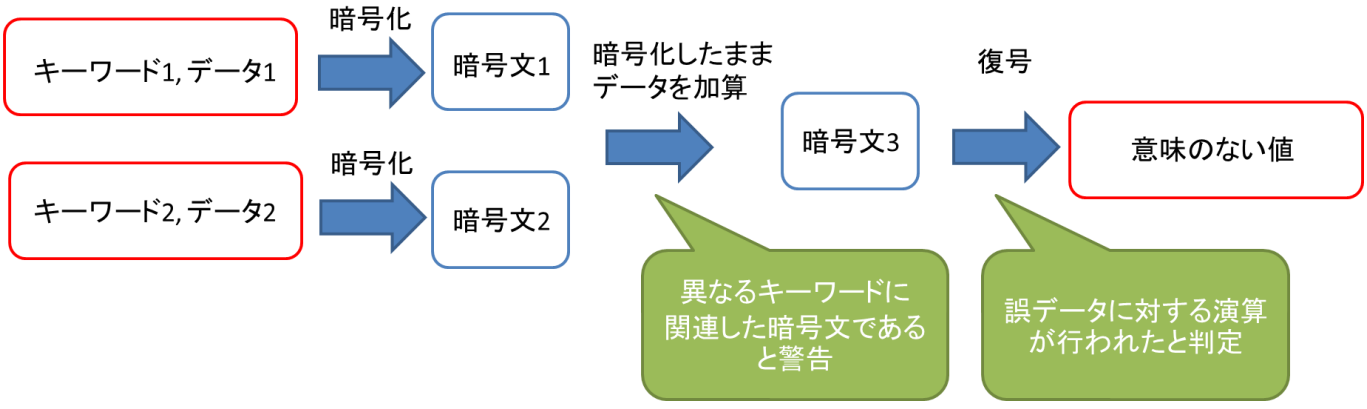

図 3. 異なるキーワード（キーワード 1 ≠ キーワード 2）に対する準同型演算

想定シナリオの詳細な流れを図４に示します。病院と検査機関がそれぞれ医療データと遺伝データを管理し、それぞれのレコードが同じ患者（図の A、B、C）を指していると仮定します。ここで、データは匿名化されているため、実際の患者にひも付いたデータではないことに注意してください。

最初に、病院が管理する医療データを病院の公開鍵で暗号化します。このとき、病名をキーワードとして、病気を罹患しているかどうかをデータとして暗号化します。この暗号文を検査機関に送付します。検査機関では、ある遺伝的特徴を持つ患者の暗号文に対し暗号化したまま加算を行うことで、この遺伝的特徴を持ち、かつ、病気を罹患している人の数の暗号文を計算します。

なお、暗号化されているため、個人の病気の罹患有無を知ることはできません。また、ここで仮に他の病気に関する暗号文が誤って混在したとしても検出できます。この暗号文を病院に送付し、病院は復号することで解析値を得ます。この際、検査機関で異なる病名に対して計算を行っていないことを病院は確認できます。

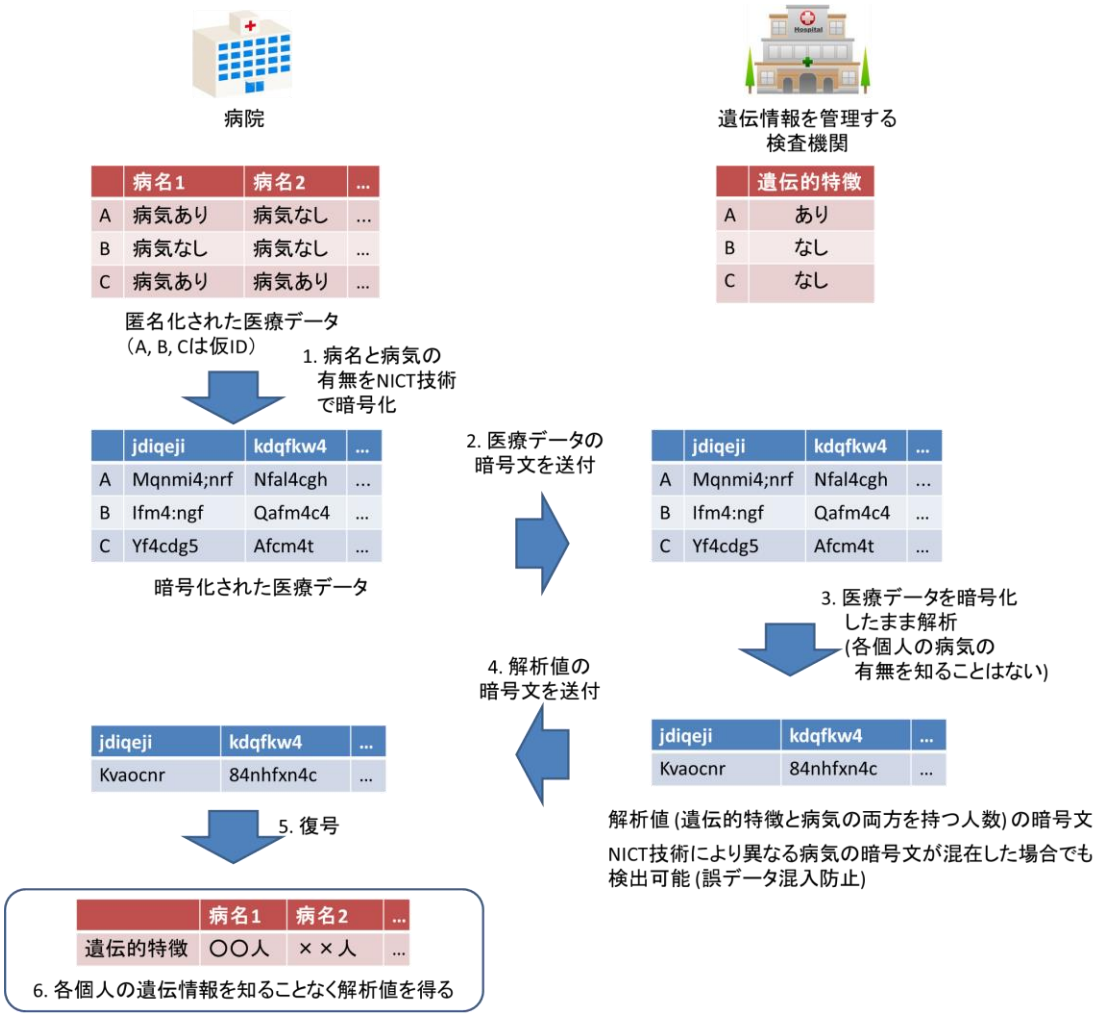

図 4. 医療データを暗号化したままでの統計値計算

本研究の一部は国立研究開発法人科学技術振興機構（JST）の戦略的創造研究推進事業 CREST「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（研究総括：栄藤 稔）※」における研究課題「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（課題番号 JPMJCR168A 研究代表者：NICT セキュリティ基盤研究室 盛合志帆 室長）」及び「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（研究総括：喜連川 優）※」における研究課題「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（課題番号 JPMJCR1302 研究代表者：筑波大学 佐久間 淳 教授）」の連携の下で行われました。

※文部科学省の人工知能／ビッグデータ／IoT／サイバーセキュリティ統合プロジェクト（AIP プロジェクト）の一環として運営