新任教授研究紹介

# サイバーセキュリティ工学領域

大阪大学大学院工学研究科
電気電子情報工学専攻　教授　宮地　充子

## １．はじめに

本研究室は2015年10月に電気電子情報工学専攻情報通信工学部門通信システム工学の１講座として発足しました。情報セキュリティと暗号理論が研究領域となります。我々を取り巻く情報社会では、多種多様なデータが収集され、その解析結果は医療、産業など様々な分野での利活用が期待されています。しかしながら、意図的に歪められたデータを用いると正しい解析結果を入手できませんし、解析結果も意図的に歪められる可能性もあります。情報セキュリティはデータの秘匿、完全性、可用性を実現し、情報社会に安心と信頼を与える技術です。

2017年４月より、河内亮周准教授、Chen-Mou Chen特任准教授、中正和久助教、Francois Bonnet特任助教、Tung　Chou特任助教、奥村伸也特任助教の体制が整いました。学生に関しては、2017年現在、博士後期課程学生１名、博士前期課程学生８名、学部学生９名、研究生５名で合計23名が本研究室に所属しています。

本研究室で実施している研究テーマは下記となります。

・現在利用されている暗号の解読や効率的な実装方法、次世代の暗号の設計を行う**暗号基盤研究**

・暗号化したまま平均値や分散値などの統計処理や、ソート等のデータ処理を実現するなど、秘匿したディジタルデータの可用性を実現する**情報セキュリティ**の研究

・ビットコインなどディジタルデータを用いた社会システムの実現や、IoT機器、クラウド等のセキュアアプリケーションを実現する**セキュアプロトコル**の研究

本研究室では、数論と計算量理論、情報理論を駆使した暗号理論の研究から、新しい社会システムをモデル化し、それを実現するディジタルプロトコルの構築という基礎から応用に広がる研究を行っています。このため、学生一人一人の研究に必要な知識がかなり違います。そういう異なる知識・研究に触れ合うことで、論理的な思考能力、問題発見能力、そして解決能力を養っていきたいと考えています。

本稿では上記３つの研究分野から具体的な研究として、以下の研究内容について説明させていただきます。

・プライバシを保護した分散多機関データ統合

JST　CRESTプロジェクト『ビッグデータ統合利活用促進のためのセキュリティ基盤技術の体系化』の中での主要研究となります。

・暗号解析

SSL通信や無線LANなどで利用されている暗号方式の解析に関する研究です。

## ２．プライバシを保護した分散多機関データ統合

我々を取り巻く情報社会では、多種多様なデータが多機関で収集されます。例えば、小学校で児童がブランコでけがをした事例を考えます。このとき、事故が起こった遊具に関するデータは学校、病院への救急搬送データは消防署、傷害・後遺症に関する

表１：学校の事故における分散データ管理事例（✓：データ有り、―：データ無し）

| | 生徒名 | 遊具メーカ | 救急搬送データ | 傷害データ | 後遺症データ |
|---|---|---|---|---|---|
| 学校 | ✓ | ✓ | ― | ― | ― |
| 消防署 | ✓ | ― | ✓ | ― | ― |
| 病院 | ✓ | ― | ― | ✓ | ✓ |

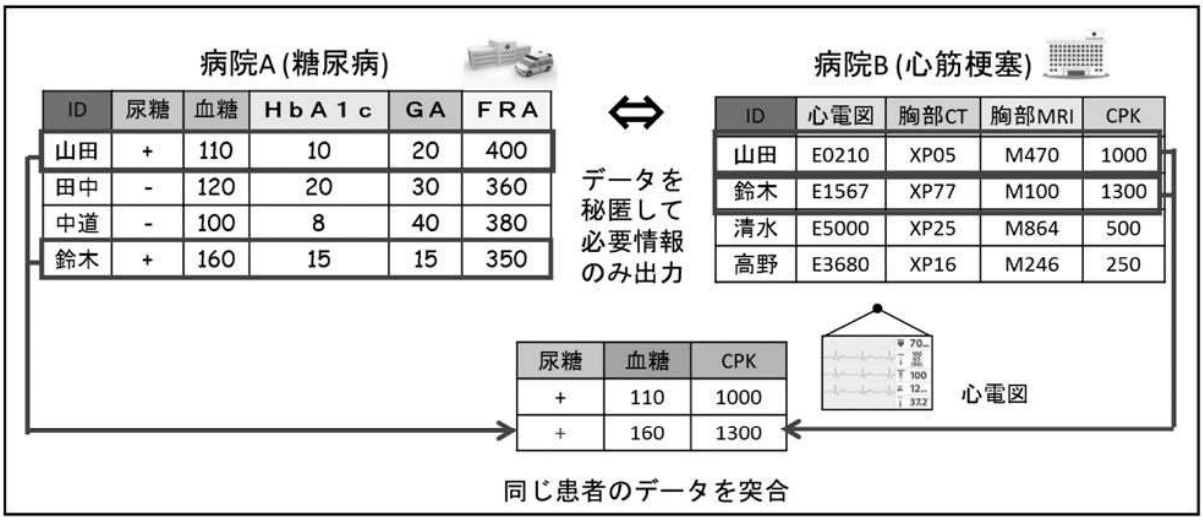



図１：２機関のデータ突合事例

データは病院に管理されます（表１）。つまり、学校での生徒の事故に関する情報では、学校、消防署、病院がそれぞれ同じ事故で異なるデータを管理します。

学校における事故の予防安全の実現には、事故の統計的因果モデルの作成が重要です。これにはこのように異なる機関に分散した関連データの統合的活用が必須です。つまり、異なる機関が独立に収集したデータから生徒の名前などの機微情報は洩れることなく、同じ生徒の事故の情報を突合（**分散多機関データ突合**）できると、事故の詳細なデータの収集が可能になります。

関連データが複数の異なる機関で保管されるケースは医療においても頻繁に起こります。例えば、同じ患者が異なる病気になった場合、複数の病院に通うことが考えられます（図１）。このように独立に２つの医療機関で管理された異なる病気は、それぞれ因果関係がある可能性があります。この時、同一の患者のデータを患者のプライバシを保護しつつ、必要な医療データのみ突合できると、異なる病気の因果関係の詳細なデータの収集が可能になります。

ここで、異なる機関がもつ医療データの突合方法とプライバシの関係について考えます。単純な方法は、１つの医療機関が全データを別の医療機関に渡せば、同じ患者を検索することで、同じ患者のデータを突合することができます。しかし、この場合、本来もつはずでなかった患者の情報である患者の名前、住所などの機微情報を別の医療機関が入手することになります。別の方法として、第３の機関（データ預託機関）にそれぞれの病院が医療情報を渡し、その第３の機関で突合することもできます。しかしこの場合には、第３の機関に患者の機微情報が移動することになります。つまり、単純な突合方法は突合に用いる情報が必要となるため、突合を実施する機関に機微情報が移動し、プライバシ保護を実現することが困難になります。

本研究室で提案された技術は、機微情報を他の機関に移動することなく、データ突合を実現する方式です。次に図２に沿って、プライバシを保護しながら同一のユーザのデータ突合を実現する方法について述べます。

**ステップ①**

各データ集合の突合したいデータx（図２の場合、名前）をハッシュ関数H等[i]で圧縮します。例えば機関１のデータの名前 「竹田」は

竹田→H(竹田)

と一意に不可逆な情報に変換されます。その後、さらに準同型暗号[ii] Eで暗号化します。

H(竹田)→E(H(竹田))

同様の処理を機関2、3で行い、暗号化データを秘匿計算機サーバに送付します。

**ステップ②**

準同型暗号は、暗号文の和が、元の文の和の暗号文に一致する性質を持ちます。この性質を用いて秘匿計算機サーバでは入手した暗号化データをそれぞれ加えることで、各機関のデータの総和の暗号結果を計算できます。図２の場合、下記のように、各機関で計算されたE(H(竹田))のような名前の暗号文は秘

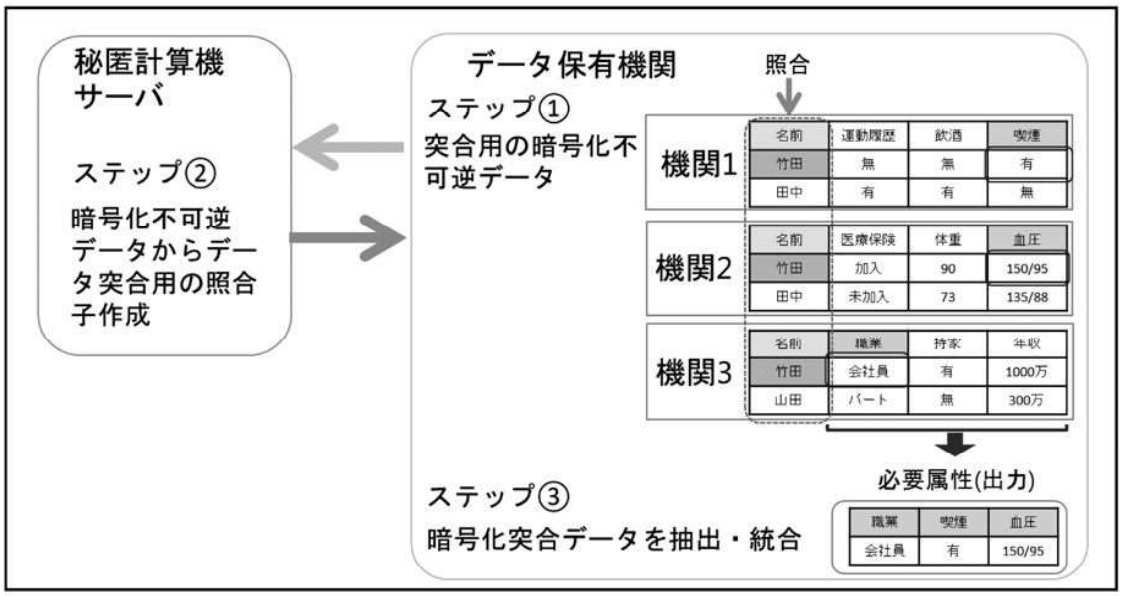



図２：システムモデル

匿計算機サーバで加えられ、名前の和の暗号文となります。

E(H(竹田))＋E(H(田中))＋E(H(竹田))＋E(H(田中))
＋E(H(竹田))＋E(H(山田))

＝E(3H(竹田)＋2H(田中)＋H(山田))

この暗号化データの総和を各機関に送付します。

**ステップ③**

受信した各機関のデータの総和の暗号文を復号して、突合したデータ、図２の場合、竹田さんのデータを出力します。

本提案は以下の特徴を持ちます。

・秘匿計算機サーバには暗号化データのみが送付されるので、データは完全に秘匿されます。

・暗号化不可逆データを用いて、突合が実現されるので機微情報はどの機関にも移動しません。

・各機関の処理時間は機関数に依存しません。

既存研究はデータサイズや機関数に依存する処理時間、通信量が大きな課題であるのみならず、図２の喫煙、職業、血圧等の複数の属性の統合方法がありませんでした。当研究室で考案された方式は、機関数に処理時間が依存せず、さらに複数の属性の統合を可能にします。

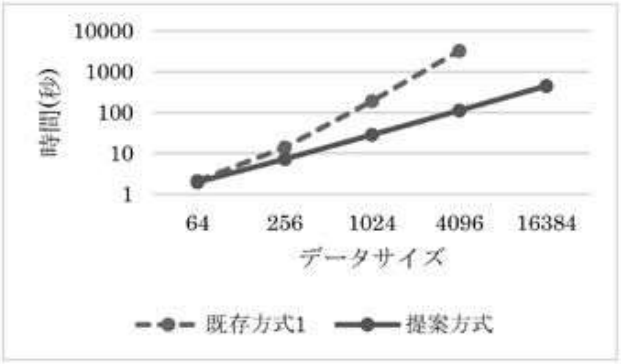
表２に提案方式と既存方式の比較を記載します。提案方式はデータ預託機関が不要で、各機関のデータ数の制限がなく、通信量、計算量を削減した方式となっています。既存方式１と提案方式の処理時間は表３に記載されています。機関数の二乗のオーダで処理時間が掛かる既存方式では、データ数が増えると非常に多くの処理時間がかかり、提案方式の優位性がわかると思われます。

しかしながら、提案方式もデータ数が増えると非常に時間がかかり、リアルタイムでの実用化にはさらなる改良による高速化が必要で、現在、アルゴリズムの改良に取り組んでいます。

表２：提案方式と既存研究の比較

| 方式 | 既存方式１[1] | 既存方式２[2] | 提案方式[3] |
|---|---|---|---|
| データ預託機関 | 不要 | 必要 | 不要 |
| 各機関の計算量（機関数 n） | 機関数 n の２乗の計算量 | 機関数に依存しない | 機関数に依存しない |
| 通信量 | 機関数 n の計算量 | 機関数 n の計算量 | 機関数 n の計算量 |
| データ数の制限 | 全機関が同じデータ数 | 制限なし | 制限なし |
| 秘匿される情報 | データ集合のみ | データ集合とその個数 | データ集合とその個数 |

表３：提案方式と既存方式１との比較
（16 機関、データ数 $2^6$-$2^{14}$）




## ３．暗号解析

上述の章ではセキュリティの応用事例の最新の研究について紹介しましたが、この事例でも利用されているコア技術が暗号です。暗号は今や私たちの生活とは切り離すことはできません。例えば、webサイトを閲覧するときにも、様々な暗号が利用されています。その暗号の安全性解析は、本研究室の重要な研究課題の一つです。私達が日常利用するセキュリティシステムであるインターネットのSecure Socket Layer/Transport Layer Security（以下、SSL/TLS）、無線LANのプロトコルであるWired Equivalent Privacy（以下、WEP）、Wi-Fi Protected Access（以下、WPA）においては各種暗号が利用されています。本研究室では現在利用されている暗号の安全性解析を行うことで、より安全な暗号の提案も行っています。ここでは1987年にRon Rivestによって提案されたRC４と呼ばれる暗号の安全性解析の成果について紹介します。RC４は我々の成果を含む様々な安全性解析の結果、標準暗号としての利用は推奨されていませんが、ダウングレード攻撃により古いバージョンに落として攻撃する手法があり、その安全性解析は現在でも非常に重要です。

一般に安全な暗号とは乱数と区別がつかない暗号です。つまり暗号解析の研究は入力を様々に変えることで、乱数と区別がつくような出力を求め、その偏りを使って、元のメッセージあるいは鍵を復元します。図３はRC４が出力する乱数の分布を表します。RC4では毎回256個の中からランダムに１つの数を出力するので、RC４が乱数の動きと同じであれば、それぞれの数は1/256の確率で出力されます。点線グラフは本来の乱数の分布で、1/256の確率ですべての数が出力することを表します。我々の安全性解析の結果、発見された偏りは２つの直線になります。このうち、凸型の直線はWPAにおける出力、右上がりの直線はRC４における出力です。この実験結果から、RC４が乱数と比べるとかなり偏っていること、WPAにおける不適切なRC４の利用により、さらに偏りが増していることがわかります。我々の研究はまずは偏りを実験的に検知し、次に、アルゴリズムを解析することで、この偏りを出力するアルゴリズムの原因を理論的に証明します。本来、安全な暗号とはランダムな値を出力するように設計される必要があります。つまり、暗号の出力は鍵や他の要素とも相関がないことが必須です。ところが、我々の解析の結果、RC４の鍵の一部であるK[0]、K[1]と出力結果Sに

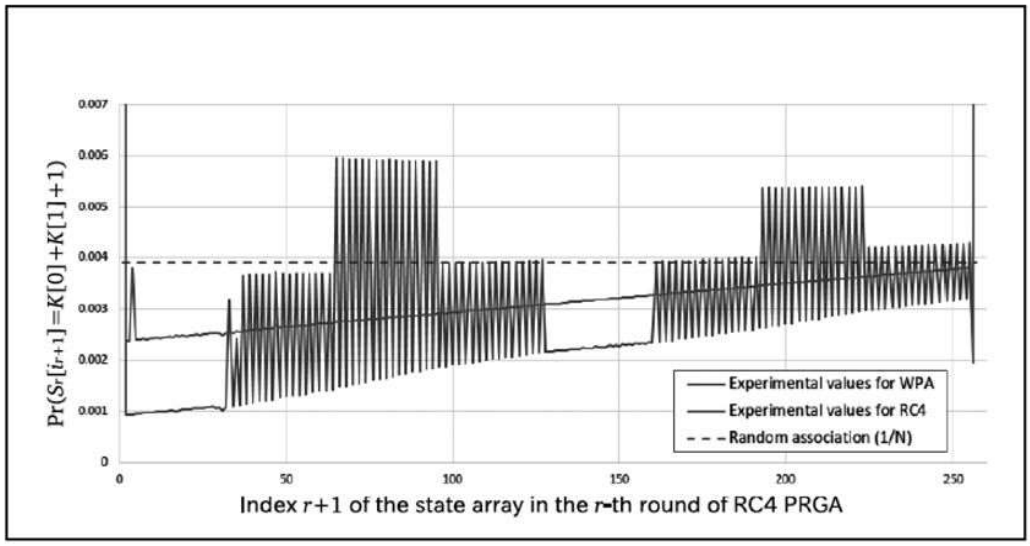



図３：RC４の出力結果の偏り

S = K[0] + K[1] + 1

の関係が成り立つことがわかり、その結果、図３に見られる偏りが発生することを理論的に証明しました[4]。

このように現在利用されている暗号を様々な利用条件において解析し、その結果を用いて新たに安全な暗号の設計に貢献することで、情報社会の安全安心に貢献しております。

## ４．セキュリティ人材教育

現代ではセキュリティ技術は情報インフラストラクチャーを支える必須の技術であるのみならず、医療、教育、公共サービスなど、様々な分野で活躍する人にとっても、切り離せない必須の知識ともいえます。このような背景のもと、文部科学省の高度IT人材を育成する教育プログラムにおいてはセキュリティ人材の育成を進めており、大阪大学工学部は「セキュリティ分野（Basic SecCap）」を運営する連携大学の１つとなっております[5]。具体的には、大阪大学工学部でなされるセキュリティに関する講義は全国の大学に配信され、各大学で受講されています。さらに、実践的なセキュリティ演習も行っており、各大学の学生が大阪大学工学部に集結し、セキュリティの実践演習に参加しております。こうして、先進技術の知識に加え、理解・応用できる実践的能力の開発も含む人材育成を達成する教育を実施しております。なお、本セキュリティプログラムは、大学間連携による教育内容のダイバーシティと、産業界、あるいはセキュリティ関連団体との連携により、実践的人材育成の教育コースを開発していることも特徴です。

## ５．おわりに

本稿では、本研究室で実施している研究について紹介させていただきました。今後も、情報セキュリティと暗号理論に関する先進的な研究テーマに取り組んでいく所存です。大阪大学工業会会員の皆様のご指導、ご支援を賜りますよう、お願い申し上げます。

### 参考文献


1) Kissner and Song, Privacy-Preserving Set Operations, CRYPTO 2005, LNCS 3621, Springer, pp. 241–257, 2005.

2) Many, Burkhart, and Dimitropoulos. Fast private set operations with sepia. Technical Report, 345, 2012.

3) Miyaji, Nakasho, and Nishida, "Privacy-Preserving Integration of Medical Data A Practical Multiparty Private Set Intersection", Journal of Medical Systems, Vol. 41 No. 3, pp. 1-10, (2017).

4) Ito and Miyaji, "New Linear Correlations related to State Information of RC4 PRGA", FSE 2015, LNCS, Springer-Verlag, 2015.

5) https://cy2sec.comm.eng.osaka-u.ac.jp/miyaji-lab/basic-seccap/index-jp.html


（阪大　理学部　平成２年　前期）

i ハッシュ関数とは任意の大きさのデータを固定の値に圧縮する関数である。暗号学的ハッシュ関数では、ハッシュ関数の衝突値（collision）が容易に見つけられない、出力から入力を見つけるのが困難であるなどの条件を満たす必要がある。本提案ではブルームフィルターと呼ばれる要素検証方法を利用しており、ブルー ムフィルターの構成でハッシュ関数を用います。

ii 準同型暗号とは暗号化されたデータを暗号化したまま加法や乗法演算可能となる暗号です。加法準同型暗号では、E(a)+E(b)=E(a+b)、つまり暗号文同士の加法が元の文章の加法a+b の暗号文と一致します。