平成 29 年度

# 修士論文要旨集

高知大学大学院

総合人間自然科学研究科

理学専攻

応用理学コース　情報科学分野

# ペアリングを用いたプロキシ暗号の実装

応用理学コース　情報科学分野　　石 田 裕 貴

クラウドは情報を共有するための極めて便利なサービスである。しかし、サイバーテロなどによりクラウドに預けたデータが漏洩する危険性は否定できない。万が一データが漏洩しても情報を漏洩させないためにはデータの暗号化を行う必要がある。

現在主流の暗号はRSA暗号と呼ばれる公開鍵暗号である。RSA暗号の安全性は素因数分解問題の困難さに基づいているが、コンピュータの性能向上に伴い、解読のリスクを避けるためにはより大きな鍵長が必要とされ、従来の機器では対応できない可能性が出てきた。そのため、RSA暗号に代わる新しい暗号として楕円曲線暗号が注目されている。楕円曲線暗号は楕円曲線上の離散対数問題を利用した公開鍵暗号の総称であり、RSA暗号の10分の1ほどの鍵長で同程度の強度を持っている。

クラウドに置いたデータを複数人で扱う場合、一人対一人にしか対応していない通常の暗号方式では不十分である。例えば第三者が既に暗号化してアップロードされたデータを見たい場合、一度そのデータを所有者がダウンロードし復号化した上で、それをさらに第三者が復号できる形に暗号化してアップロードしなければならない。そこでプロキシ暗号という暗号方式が提案された。プロキシ暗号はデータを複数人で効率よく利用するために考えられた暗号方式であり、暗号化されたデータをさらに第三者が復号して読める形に再暗号化することにより、再ダウンロード・アップロードの手間を省くことができる。

プロキシ暗号は楕円曲線上のペアリングを用いて実現することができる。ペアリングとは楕円曲線上の２個の点を引数としある有限体に値をとる関数であり、双線型性という特性を持っている。この双線型性を利用することによって、三者間鍵共有、IDベース暗号などの暗号技術が実現可能となったが、プロキシ暗号もそういった暗号技術のひとつである。

今回の研究では楕円曲線上のペアリングを用いたプロキシ暗号方式を使い、クラウド上で複数人のユーザーが便利に利用するための暗号化ツールの開発を行った。このツールを使えば、クラウドにファイルをアップロードする際に自動的に楕円曲線暗号による暗号化が行われる。さらにサーバーに再暗号化の命令を送ると、クラウドに預けたデータを元に、グループ内に登録されている自分以外の全てのユーザーに対して再暗号化データが生成される。再暗号化されたデータは各ユーザーが自身の秘密鍵で復号化することができる。これによりグループ内では普通にデータを共有しつつ外部の人間からはファイルの中身を確認されないファイル共有ツールが完成した。

# 推論による情報漏えい防止のための推論関係ハイパーグラフ抽出手法
## ― Twitter データを用いた抽出手法の検討 ―

応用理学コース　　情報科学分野　　長　尾　洸　希

近年，ビッグデータの活用が期待される一方で，個人情報やプライバシーの保護が求められている．例えば，SNS 上の活動記録から自宅を留守にすることが多い時間帯を推測されたり，地震が起きた際に「地震だ」と発言したことで地震観測データと照らしあわされ居場所を推測されたりすることが懸念される．

これらの例のように,複数の既知情報をもとに推論されることによって秘密情報が漏えいしてしまうことを本稿では”推論による情報漏えい”と呼ぶ．従来のセキュリティ技術は，秘密情報そのものの直接的な漏えいを防ぐことを目的としたものが主であったが,ビッグデータ時代においては推論による情報漏えい対策技術も必要となる．

鈴木ら（2012 年）は，どの既知情報の組み合わせからどの情報が推論できるのかを有向ハイパーグラフ（以降，推論関係ハイパーグラフと呼ぶ）で表現することで，推論による情報漏えいをグラフ理論的に分析できるようにモデル化を行った．その後，当研究室では，推論による情報漏えいを防ぐための基礎研究という位置づけで,個々の秘密情報の漏えいリスクを評価する手法の研究やグラフモデルシミュレータの開発を行ってきた．一方，情報漏えいリスクではなく情報漏えいに至る推論経路，すなわち，どの複数情報をもとにどのような推論をたどることで秘密情報へいち早く到達されてしまうかを予測する研究も行われている．

しかし,従来の研究はいずれも推論関係ハイパーグラフがすでに得られたという前提のもとでの研究であり,推論による情報漏えい防止に役立つ推論関係ハイパーグラフを得る手法の研究はされていなかった．

人間は必ずしも演繹的に推論するとは限らない.経験に基づいて帰納的に憶測することも多い.そのような帰納的推論関係を集めてできる推論関係ハイパーグラフを推論に長けた人間が手作業で作成するのではなく,これまでに蓄積された膨大なデータから自動的に抽出できることが望ましい．

そこで，本論文では，相関ルール抽出アルゴリズムを用いて帰納的推論関係を抽出することで推論関係ハイパーグラフを求める手法を提案する．また，相関ルール抽出の代表的なアルゴリズムである Apriori アルゴリズムを改良したアルゴリズムを提案する．

推論による情報漏えい防止に役立つ現実的な推論関係ハイパーグラフを得ることは相当な困難が予想される．その困難さは何に起因するのかを調べるため，提案手法を Twitter データに適用して得られた推論関係ハイパーグラフが推論による情報漏えい防止に役立つものとなるかを考察する．

# 手話トレーニングマシンの開発 – 表情取得について –

応用理学コース　　情報科学分野　　大栗 慶太郎

本研究室ではより良い手話の学習環境の構築を目指し，学習支援システムである「手話トレーニングマシン」の研究・開発を行っている．このシステムは学習者の手話動作を深度センサとグローブを用いて取得し，その動作を判別することで正誤結果を含むフィードバックを提供する．これにより学習者は自身の行う動作が合っているかどうかを判断でき，正しい手話動作の習得につながる．本研究室ではこのシステムの実用化を目標に研究を行っており，判別精度に一定の結果が得られたことから次の段階へ移ることとした．

手話において非手指動作は非常に大きな役割を担っている．この非手指動作とは手指以外の動きを指し，例えば頷きや目の見開き，一瞬の間などがある．手話の自然な表現を習得するには，非手指動作を含めた学習を行う必要がある．既存の手話トレーニングマシンのシステムは主に手指動作のみに着目しており，非手指動作を含めた学習には対応していない．そのため，より自然な手話の学習を支援するにあたって，非手指動作の表現を学べる環境を構築する必要がある．

表情は非手指動作の中で最も重要であり，手話の会話中に頻出する表現の１つである．この手話における表情に着目し，日常的に使われる手話を習得できるよう，システムに表情の要素を導入する．学習者の表情を判別し手指動作と同様に有益なフィードバックの提供を目指す．

筆者の学士研究ではまず既存のデバイスを用いて表情取得を試みた．このデバイスではユーザの顔を認識して表情を６つの感情の強度として取得できる．しかしながら手話に用いられる表情は眉の動き等の微細な動きが多く，感情のみではそれら表情への対応が困難という点が明らかとなった．

本研究では学士研究で得られた課題点を解決するため，手話における表情の新たな取得方法の提案を行い，手話に用いられる表情の取得に特化したシステムを構築し，手本となる表情の教師データを登録する際の簡易性にも着目，手話単語の増加に対応できるようにした．具体的には一般的な表情の取得ではなく，手話の学習に必要な表情の取得に着目し，例えば日本手話において疑問文を表現するとき，顔には目を見開くという特徴が顕著に表出する．これらの表現に対応する表情の取得方法を提案する．提案手法では機械学習とファジィ推論を用いて手話トレーニングマシンに新しい表情取得の仕組みを導入する．KinectV2 深度センサから得られる顔の形状情報を機械学習によって前処理し，その結果を元にファジィ推論を用いて表情に結びつける．センサから得られる顔の形状情報はデータ量が多く複雑であり，人間がルールを設定するのは容易ではないため機械学習を用いている．対して出力に近い部分ではファジィ推論を用いることによって人間が言葉でルールを設定でき，表情を扱いやすくなる．例えば手話単語の辞典に載っている言葉から表情を記述するためのルールを見い出すことができる．これらによって表情を判別するための柔軟な設定が可能になった．

さらに提案手法について評価実験を行い，表情を含めたシステムと学習環境の妥当性に関して検証を行った．

# ビッグデータのヒストグラムと相関係数を再現するデータ生成手法と医療ビッグデータによる検証

応用理学コース　　情報科学分野　　筒　井　真　璃　菜

近年では，スマートフォンなどのデジタル機器の普及により，ビッグデータの収集が幅広く行われている．収集されたビッグデータの利用は，様々な産業で改善や売上げ見積もり等に活用されている．

ビッグデータ利用は，医療情技術の発展でも期待されており，病院など医療機関で蓄積された医療データを用いたビッグデータ分析が試みられている．例えば，様々な合併症を引き起こす慢性腎臓病の診断において，複雑な医療検査値の組み合わせからのビッグデータ分析の有効性が示されている．

しかしながら，ビッグデータ利用に際しては，個人情報保護を守らねばならないものの，完全な技術は確立してない．個人情報保護技術として例えば，匿名化，計測値の一部の乱数化，さらに個人が特定されるような極端な数値データの削除などが挙げられる．しかし，これらデータでは，名寄せ(類似するデータを合わせて個人を特定すること)により個人情報が復元されてしまう危険性がある．そのため，個人情報の復元が不可能なビッグデータ処理技術が求められている．

そこで，著者は，ビッグデータと同じ統計的属性をもち，個人情報を一切含まないデータ生成手法 1，および有意義なビッグデータ分析ができるデータ生成法 2 を提案し，SOM 分析で実証する．

データ生成法 1 は，ビッグデータの分布（ヒストグラム）に比例した確率の乱数でデータを生成（モンテカルロ法）する．これによりビッグデータの分布，平均値，分散等の統計的属性が同等なデータを再現することができる．本手法を慢性腎臓病の疑いのある患者の検査値のヒストグラムに適用し，その分布や最大値・最小値・平均値・分散値が再現されていることを確認している．

次にデータ生成法 2 は，生成法 1 で作成したデータに対して，ビッグデータ分析への影響が大きいと推測されるデータ間の相関係数の再現のため，相関係数の差異の最小化を評価したデータレコードの交換法により実現している．本手法を先の患者の検査値間の相関係数を用いて，生成手法 1 で作成したデータの相関係数の再現を行い，ヒストグラム，相関係数からビッグデータと同等の分布，平均値，分散等の統計的属性に加えて相関係数まで有するデータ生成を行った．

本データ生成法が，ビッグデータ分析で有用か否かを評価するため，代表的な分析方法である SOM(自己組織化マップ)を用いた評価を行う．SOM を利用することで，多数の要素で構成する患者データ（多次元医療データ）から相互関係を可視化できることを利用して，生成データから医療データの SOM 分析と同等な結果が得られるかを確認することができるからである．結果，生成データは元の医療ビッグデータで既に分析された eGFR と尿蛋白の反転関係などを再現していることが確認できた．もちろん，元の医療データを用いずヒストグラムと相関係数のみから得られた結果として個人情報を一切含んでいない．

以上から，本論文で提案するデータ生成法は，個人情報保護が厳しい，特に医療分野におけるビッグデータ分析技術の研究開発に広く貢献するものである．

# 理解しやすさや探求の深さを考慮した<br>学習者に適応した学習コンテンツの推薦手法

応用理学コース　　情報科学分野　　藤崎　優理

初学者が新しい知識を学ぶ時には、参考書を読んだりインターネットで検索したりして欲しい情報を手にいれる。インターネットが普及している現在は後者で情報を求めるケースが多いだろう。しかし、ウェブ検索によって求めている情報を解説しているページを見つけたとしても、読みやすいと思うページだけではなく、読みにくいと思うページもヒットしてしまう。これは、通常の検索サイトでは、学習者の習熟度によって理解しやすい学習コンテンツが異なることや、学習者の探求目的によって求める学習コンテンツが異なることを考慮していないためである。

学習者の習熟度によって理解しやすい学習コンテンツが異なるとは、学習コンテンツには理解するのに前提知識を必要とするものと必要としないものとがあり、それらが理解しやすいかどうかは学習者の前提知識の有無によるということである。例えば、学びたい知識に関連する知識がほとんど無い初学者の場合、前提知識があるものとして説明している学習コンテンツは理解しにくいと感じてしまうが、噛み砕いた表現やとっつきやすい表現を多用して前提知識を必要としない学習コンテンツは理解しやすいと感じるはずである。次に、学習者の探求目的によって求める学習コンテンツが異なるとは、知識を深く学びたい学習者は内容が深く記述されている学習コンテンツを求めるのに対し、概要を掴みたい学習者は要点が簡潔にまとめられている学習コンテンツを求めるという違いのことである。そのため、学習者に適応した学習コンテンツを推薦するためには、学習者の習熟度に応じた理解のしやすさや、学習者の探求目的の深さの度合いを考慮する必要がある。

本研究では、上記をふまえ、学習コンテンツの理解しやすさや学習者の探求目的の深さに基づき、学習者に適応した学習コンテンツを推薦する手法を提案する。本研究で提案する手法は、ウェブ上の数学の学習コンテンツを対象としている。学習コンテンツの理解しやすさはアンダスタンダビリティと定義し、指標として噛み砕き度ととっつきやすさを提案する。ある知識をなるべく易しい同分野の知識のみで表現したものを噛み砕き表現とし、噛み砕き度とは、噛み砕き表現の噛み砕き具合を表したものである。また、とっつきやすさとは、文章中のたとえ話や事例の交わり具合を表したものである。提案手法では、噛み砕き度ととっつきやすさの高い学習コンテンツの方が理解しやすいと感じている習熟度の低い学習者に対しては、噛み砕き度ととっつきやすさが高い学習コンテンツほどアンダスタンダビリティが高くなり、優先的に推薦される。一方、学習者の探求目的の深さに応じて適切な学習コンテンツを推薦するために、学習コンテンツの探求の深さを定義する。提案手法では、知識を深く学ぶのに適した学習コンテンツは探求の深さが深く、概要を掴むのに適した学習コンテンツは探求の深さが浅いとすることにより、学習者の探求目的の深さに応じた推薦が可能となる。

本論文では、アンダスタンダビリティの指標となる学習コンテンツの噛み砕き度やとっつきやすさと、学習コンテンツの探求の深さを推定するアルゴリズムについて提案し、アルゴリズムの実装に必要となる学習コンテンツの特徴量を抽出するために開発した解析・分析プログラムについて述べる。また、推定アルゴリズムに用いる学習コンテンツ収集手法と、推定アルゴリズムの精度を検証するための正解データの生成手法を説明し、収集した数学の学習コンテンツと生成した正解データを用いて精度を検証した結果についても報告する。

# リハーサル差分データの提示にもとづくバックレビュー支援システム

応用理学コース　情報科学分野　　小池　柾伎

プレゼンテーションは，主にスライドと口頭説明をもちいることで，自らが学習した知識を他者へと外化する手段である．また大学の研究室などの教育機関においては，卒業・修士論文発表会や学会発表などのために，発表練習であるプレゼンテーション・リハーサルが行われており，これらは複数回実施されることが多い．プレゼンテーション・リハーサルにおいて，プレゼンタは仲間・同僚であるピアからの指摘をもとに改訂作業を行い，自らの知識の不十分・不適切さに関する気づきを得ることで，プレゼンテーションの改善を図る．したがって，リハーサルにおける改訂作業は，プレゼンテーションの改善を検討すると同時に，知識洗練化を促進させる重要な過程であり，そのためには得られた指摘や議論などを含むリハーサル結果を効果的に活用する必要がある．

筆者が所属する研究室では，これまでにプレゼンテーション・リハーサルを対象としたレビュー支援環境の構築と運用を行っている．本支援環境では，支援における自由な機能拡張や効果検証のための詳細なデータ取得を考慮し，市販のプレゼンテーション・ソフトウェアではなく，独自のデータ形式による発表資料のオーサリングが可能なプレゼンテーション・ツールを開発・利用している．そして本研究では本ツールにもとづき，発表に対するレビュー支援ツールとレビューコメントに対する議論支援ツールを開発し，支援機能を実現した．これによりプレゼンタは，その後の改訂作業に利用するためのリハーサルデータを十分に収集可能となった．しかし，プレゼンタが改訂作業を行う上で，得られたデータ全てを確認しながら修正を行うことは認知負荷が高いため困難である．したがって，改訂作業においては，修正するべき箇所全てを漏れなく改善できているか，またリソースを効果的に参照して各修正箇所を適切に改善できたかなど，プレゼンタが修正状況を常に確認できることが望ましい．そのため本支援環境では，得られたリハーサルデータを整理した状態でプレゼンタにフィードバックし，より効果的な改訂作業によるプレゼンテーション改善の支援を検討する必要がある．

これらの課題を解決する方法として，先行研究では複数のリハーサルデータを比較することで，改訂状態の遷移や他のプレゼンテーションとの違いを確認する手法が提案されている．しかし，リハーサルデータを実際に比較して改訂作業を行うには，さらにプレゼンタが各リハーサルデータの差異を認識し易い状態でプレゼンタに提示する必要がある．そこで本研究では，連続した複数のリハーサルデータの差分を抽出・可視化する表現方法をもちいたプレゼンテーションの比較・検討を行う手法を提案し，独自のプレゼンテーション・ツールとの連携によるバックレビュー支援システムの開発と有効性の検証を行った．