

氏 名 町田 史門

学位(専攻分野) 博士(情報学)

学位記番号 総研大甲第 1883 号

学位授与の日付 平成28年9月28日

学位授与の要件 複合科学研究科 情報学専攻  
学位規則第6条第1項該当

学位論文題目 SNS におけるユーザの投稿内容に応じたプライバシー保護に関する研究

論文審査委員 主 査 教授 越前 功  
教授 曾根原 登  
准教授 岡田 仁志  
准教授 相原 健郎  
教授 小館 亮之 津田塾大学

論文内容の要旨  
Summary of thesis contents

Twitter や Facebook に代表される Social Networking Service (SNS) は、世代を問わず人々の日常生活に溶け込み、オンラインコミュニケーションツールとして欠かせないものとなっている。その一方、SNS ユーザの不用意な投稿により、個人に関するセンシティブデータの漏洩が問題となっている。投稿の結果、自身や他人のセンシティブデータが友人関係性の弱い知人レベルの友人や面識のない他人にまで拡散し、友人関係のもつれや職を失うユーザが出るなど、何気ない投稿が予期しないトラブルへと発展している。このような漏洩の原因の 1 つに SNS ユーザが持つ投稿に対する主観的な判断基準がある。各ユーザの判断基準は、個人の主観に依存しているため、その投稿判断を誤った場合に、センシティブデータの漏洩が発生する可能性がある。そのため、SNS ユーザにとって、どのような情報がセンシティブデータとなりえるのかを明確にし、統一的な判断基準のもとに投稿判断が行えることが理想である。また、SNS 提供者側では、ユーザのセンシティブデータ漏洩を防ぐために、投稿に対する公開ポリシーや公開対象者を選択する公開範囲の設定機能などのプライバシーポリシー管理機能を提供している。しかし、SNS におけるプライバシーポリシー管理は複雑性があるだけでなく、その維持にも多くの労力が必要とされるため、簡易に対処する手段が必要である。

本論文では、センシティブデータの漏洩を無理に防止するのではなく、SNS の特徴である他ユーザとのオンラインコミュニケーションの楽しさを極力維持しながら、SNS 投稿ユーザの主観的な判断基準のみに依存せず、客観的なセンシティブデータの判断基準、および投稿内容に含まれる他ユーザのプライバシーポリシーに基づいて、投稿メッセージを適応的に保護する手法の確立を目的とする。具体的には、SNS 投稿時に、テキスト、写真を含む投稿予定メッセージにセンシティブデータを含むか否かを検知し、その検知結果により、限定的な公開範囲・公開対象者を自動提案する。さらに、投稿される写真に写る/写り込む被写体のコミュニティや状況に応じたプライバシーポリシーに基づき、当該人物の顔領域を適応的に保護する。

本目的に向けて、1 つめの課題として、「SNS におけるセンシティブデータの客観的な判断基準の定義」に取り組む。SNS ユーザの主観的な判断基準に依存した SNS 投稿により発生するセンシティブデータ漏洩を防止するために、ユーザにとって、どのような情報がセンシティブデータとなりえるのかを明確化し、客観的な判断基準を定義する。本課題では、最初に従来研究で言及されるセンシティブデータの漏洩が実際に発生していることを確認するため、過去 1 年間分の Twitter アーカイブを用いて、センシティブデータ漏洩の発生有無を試算評価するとともに、投稿予定メッセージからセンシティブデータを検知するための手法を検討した。次に、この評価結果を受け、SNS 投稿前の投稿予定メッセージに含まれるセンシティブデータの自動検知を実現するために、当該メッセージにセンシティブデータを含むか否かの客観的な判断基準の定義の第一歩として、公文書におけるプライバシーの取り組みを参照し、非公開とすべき情報の内容分類と公開範囲を表す開示レベルを対応付けた、「SNS におけるプライバシー侵害情報分類表」を提案した。さらに、本分類表のシステム適用・実装に向けて、分類表の妥当性を評価調査により示した。本分類表をシステムに適用することにより、SNS ユーザは、投稿を行う前に客観的な判断基準のもと、投稿判断を可能とする第一歩を示した。

(別紙様式 2)  
(Separate Form 2)

次に 2 つめの課題として、「センシティブデータの漏洩検知に基づき投稿ユーザへ指摘・通知」に取り組む。投稿ユーザが投稿メッセージにセンシティブデータを含んでいることに気付かずに投稿してしまい、その結果、ユーザがその投稿を後悔している課題に対して、1 つ目の課題で提案した分類表が持つ、SNS 投稿メッセージに含まれるセンシティブデータの有無とその開示レベルを客観的に判断可能とする特性を活かし、センシティブデータの漏洩有無の検知・通知と、その情報の重要度に沿った公開範囲の自動設定を提供する“センシティブデータの漏洩検知に基づく公開範囲の設定方式”を提案した。また、その実現化として、アプリケーション：Adaptive Disclosure Controller for Facebook を実装し、検知したセンシティブデータの重要度に応じた開示レベルを投稿ユーザへ視覚的に示し、センシティブデータが含まれていることを認識できるよう通知した。さらに、SNS ユーザが持つ、投稿内容に応じた特定の個人や興味を持つコミュニティへのメッセージ公開要望に対して、容易に任意の公開範囲・公開対象者を定義可能とすることで、特定コミュニティ・グループにとらわれることなく、投稿メッセージを公開できることを示した。

最後に 3 つめの課題として、「コミュニティに応じた被写体のプライバシーポリシーの反映」に取り組む。投稿ユーザは写真に写る・写り込む被写体の公開・非公開などのプライバシーポリシーが分からないため、自身の主観的な判断基準に依存した SNS 投稿を行わざるを得ない課題に対し、投稿ユーザに被写体のコミュニティや状況に応じたプライバシーポリシーを参照させる手法として、“被写体のコミュニティベース・プライバシーポリシーの設定方式”を提案した。本方式では、被写体が属するコミュニティ内外におけるプライバシーの振る舞いをポリシーとして埋め込んだタグ：PrivacyTag を用いて、コミュニティ内外で当該人物の顔領域を適応的に保護する。また、本方式の実現化として、被写体から PrivacyTag の検知・解析を可能としたアプリケーション：Photo Privacy Realizer for Facebook の実装を行った。これらにより、SNS 投稿ユーザの主観的な判断基準のみに依存せずに、被写体のコミュニティや状況に応じたプライバシーポリシーを反映させることを可能とした。

本論文では、SNS ユーザの主観的な判断基準に依存した投稿により発生するセンシティブデータ漏洩を防止するために、客観的な判断基準の定義の第一歩として、SNS におけるセンシティブデータの分類を提案・評価した。そして、本分類表をベースに、投稿前のユーザへ視覚的に指摘・通知をする設定方式を提案・実装した。また、軽視されやすい写真に写る被写体のプライバシー保護のために、被写体のポリシーを撮影者・投稿ユーザに参照させる方式を提案・実装した。本論文で提案したこれらの方式は、現在問題となっている SNS ユーザのセンシティブデータの漏洩に対して、その防止に寄与するものと考えている。

博士論文の審査結果の要旨  
Summary of the results of the doctoral thesis screening

出願者、町田史門氏は、「SNSにおけるユーザの投稿内容に応じたプライバシー保護に関する研究」と題する論文を提出し、この論文およびその内容に基づく研究発表に基づき博士論文の審査が行われた。本論文の貢献は、SNSユーザの主観的な判断基準に依存した投稿により発生するセンシティブデータ漏洩を防止するために、客観的な判断基準として、SNSにおけるセンシティブデータの分類を提案・評価し、さらに本分類表をベースに、投稿前のユーザへ視覚的に指摘・通知をする設定方法を提案・実装したことである。さらに、軽視されやすい写真の被写体のプライバシー保護のために、被写体のポリシーを撮影者・投稿ユーザに参照させる方法を提案・実装したことである。

本論文は6章から構成される。第1章では、本研究の背景、目的、本目的に向けた課題について述べ、第2章「SNSユーザのプライバシー保護」では、SNSユーザのプライバシーを保護するためにSNS側から提供されるプライバシー管理機能の1つである、投稿メッセージの公開範囲設定機能とその課題について述べ、その課題により引き起こされるSNSユーザの投稿後の後悔について説明している。そして、このようなユーザの後悔を防ぐために従来研究で提案されているSNSユーザのプライバシー保護手法として、SNS投稿ユーザを保護主体とした手法、写真の被写体を保護主体とした手法について述べながら、これら従来手法の課題と本研究の位置付けを述べている。

第3章「SNSにおけるセンシティブデータの分類」では、最初に、従来研究で言及されるセンシティブデータの漏洩が実際にSNSで発生していることを確認するため、過去1年間分の1億4,000万件のTwitterアーカイブを用いて、センシティブデータ漏洩の発生有無を試算評価するとともに、投稿予定メッセージからセンシティブデータを検知するための手法を検討している。次に、この評価結果を受け、SNS投稿前の投稿予定メッセージに含まれるセンシティブデータの自動検知を実現するために、当該メッセージにセンシティブデータを含むか否かの客観的な判断基準の定義の第一歩として、非公開とすべき情報の内容分類と公開範囲を表す開示レベルを対応付けた、“SNSにおけるプライバシー侵害情報分類表”を提案している。さらに、本分類表のシステム適用・実装に向けて、分類表の妥当性を評価調査により示している。

第4章「センシティブデータの漏洩検知に基づく公開範囲の設定方式」では、3章で提案したプライバシー侵害情報分類表が持つ、SNS投稿メッセージの内容からセンシティブデータ漏洩の有無と開示レベルの客観的な判断を可能とする特性を活かし、SNS投稿予定メッセージからセンシティブデータの漏洩検知とその公開範囲・公開対象者の自動設定を提供する手法を提案している。そして、本提案手法の実現化として、Facebookを対象SNSとした“Adaptive Disclosure Controller for Facebook”の実装について述べている。

第5章「被写体のコミュニティベース・プライバシーポリシーの設定方式」では、3章で提案したSNSにおけるプライバシー侵害情報分類表において、投稿の公開範囲を表現した開示レベル2に該当する情報：“個人特定可能な写真”に焦点を当て、写真に写る/写り込む人物のプライバシーに注目する。写真の被写体のプライバシーを保護するために、当該人物が属するコミュニティ内外におけるプライバシーの振る舞いをポリシーとして埋め込んだタグ

“PrivacyTag”を用いて、コミュニティ内外で当該人物の顔領域を適応的に保護する手法を提案する。また、このPrivacyTagを定義するために、まず予備評価を実施し、その評価結果を基にタグのデザイン検討を行う。そして既存手法で用いられるQRコードベースタグ

(別紙様式 3)

(Separate Form 3)

と提案タグの評価実験を行う。さらに、本提案手法の実現化として、Facebook を対象 SNS とした “Photo Privacy Realizer for Facebook” の実装について述べている。

本研究の成果は、査読付学術雑誌論文 1 篇、査読付国際会議 3 篇、その他学会発表 5 篇（いずれも出願者が主著）として発表されており、研究内容が国内外で認められていることを示している。また、本研究の成果により、情報処理学会山下記念研究賞をはじめとした学会賞を 3 件受賞している。以上の学術的貢献を総合的に判断して、本博士論文は学位を授与するに値すると判断した。