7-6　パーソナルデータ利活用に向けたプライバシー保護技術

野島　良　金森祥子　早稲田篤志　江村恵太　林　卓也

平成 26 年 11 月から平成 28 年 3 月の期間、ネットワークセキュリティ研究所セキュリティ基盤研究室において、パーソナルデータ利活用に向けたプライバシー保護技術に関わる研究活動を行った。本稿では、本活動の概要を紹介する。

1　まえがき

人間の日々の活動や社会システムから収集された大量の情報（ビッグデータ）を利活用し、いかに新たな知見やイノベーションを創出するかは競争力の源泉となりつつある。特に、利用価値が高いとされているのは、個人の行動・状態等に関するデータ（パーソナルデータ）であり、個人のプライバシーを守りながらパーソナルデータの利活用を進めることが喫緊の課題となっている。

我が国では、過去においてビッグデータ利活用時にプライバシーへの配慮が不十分であったことにより、社会から批判を受ける事案が発生したこと、プライバシー保護の観点からどのように対処すれば十分であるかが必ずしも明らかではないことから、多くの組織においてビッグデータの利活用を躊躇する事態になっている。このことは、ひいては我が国の産業競争力を低下させる結果にも繋がりかねない。

政府の成長戦略においても、ビッグデータ利活用による経済再生はひとつの柱となっており、パーソナルデータについて事業者の「利用の壁」を取り払い、個人の権利利益侵害を未然に防止しつつ、新産業・新サービスの創出と国民の安全・安心の向上等のための利活用を実現する環境整備を行うことが求められている。また、パーソナルデータの利活用に関する制度見直し方針が決定され、平成 25 年 9 月に内閣官房 IT 総合戦略本部の下に「パーソナルデータに関する検討会」が設置された。平成 26 年 6 月に「パーソナルデータの利活用に関する制度改正大綱」が公表され、平成 27 年 3 月に制度改正の法案が閣議決定、同年 9 月には「個人情報の保護に関する法律及び行政手続における特定の個人を識別するための番号の利用等に関する法律の一部を改正する法律」が公布された。今回の制度改正では、情報通信技術の進歩や個人の主観は時代とともに変動するものであることから、法律では大枠のみ定め、具体的な内容は政省令、規則及びガイドライン並びに民間の自主規制により対応するものとなっている。プライバシー保護の観点から、実際にどのような対処をすればパーソナルデータが利活用できるか、具体的な手法の確立が課題となっている。

この課題に対する検討は、第 3 期中長期計画では実施がうたわれていなかった。しかしながら、上記の動向を踏まえて、平成 26 年 11 月から平成 28 年 3 月の期間においてネットワークセキュリティ研究所セキュリティ基盤研究室では、プライバシー問題に関するワークショップを実施し、有識者へのヒアリングや事例収集を行った。本稿では、本活動を報告する。

2　どのような情報の取り扱いに気をつけるべきか

本活動では、プライバシーに関わる有識者を招へいして、ワークショップを 9 回、ワーキンググループ会合を 3 回開催した。ワークショップでは、セキュリティ技術以外の分野（情報倫理学、教育学、法学、リスク管理、心理学）の有識者による講演と、その分野でプライバシーをどのように取り扱っているか、プライバシーに関してその分野で隆起している問題等に関して、意見交換を実施した。その中で、分野を横断して関心を集め、懸案事項となった話題は、

- どのような情報をプライバシー情報とするか
- どのような情報の取扱いに注意を払うべきか

である。何をプライバシー情報とするかは社会的に大きな関心事ではあると思うが、一般的には、

1．私生活上の事実または私生活上の事実らしく受け取られるおそれのあることがらであること
2．一般人の感受性を基準にして当該私人の立場に立った場合公開を欲しないであろうと認められることがらであること、換言すれば一般人の

感覚を基準として公開されることによって心理的な負担、不安を覚えるであろうと認められることがらであること

3．　一般の人々にいまだに知られていないことがらであること

をプライバシー情報として考えることが多いようである。しかしながら、この解釈は、工学的に非常に扱いづらい。例えば2の「一般人の感覚」など、工学的に扱うためには主観的すぎる。そこで本活動では「一般人の感覚」の工学的な取扱いに焦点を当てることとし、中長期の目標を

- 実験を実施する際、あるいはサービスを展開する際に情報を提供いただく方々にどの程度の抵抗感があるかを推測
- 合意取得のあり方、合意取得の自動化

にとらえて考えることとした。

これらを目標にとらえた場合、重要となるのが、「一般人の感覚」がどのように主観的であるか、すなわち条件によりどのように変化するかを理解することである。最初に、「一般人の感覚」が時代によりどのように変化するのかを知るために、2015年にアンケート調査を実施し、2010年の結果と比較を行った。SNSの普及を考えると、自分の情報を開示することへの抵抗感は和らぐ傾向にあることも想定されたが、2015年の結果では、予想以上に自己情報開示に対する抵抗感が強いことが判明した。また、情報収集者、取得場所、取得期間、取得する情報の条件を変えることにより、どのように抵抗感が変化するかに関し理解するための実験を行った。実験は、いずれも2,000名（男性：1,000名、女性：1,000名、対象年齢：20～60代それぞれ400名ずつ、対象地域：日本全国）を対象（情報提供者）としたアンケート調査がベースである。

本実験の成果の一部として、収集情報や収集者ではなく収集期間を短くすることが、最も情報提供者の情報提供への抵抗感が和らぐといった予想外の結論が得られている（詳細は、[1]を参照のこと）。

平成27年9月に公布された改正個人情報保護法では、個人情報の適正な取得、あらかじめ本人の同意を得ずに、目的外利用の禁止、第三者提供の制限が規定されている。今回の実験結果として、予想以上に情報を提供することに対する抵抗感が高いことから、今後、同意取得の在り方について、情報収集者は再検討をする必要があると考える。情報収集者が違法を恐れるあまり、情報収集をするあらゆる情報に関して合意を得るという現状は、逆に同意取得の形骸化を招く。意味のある同意取得を検討するにあたり、同意取得の自動化もひとつの手法であると考える。同意取得の自動化を実施するために、我々は情報収集のための条件を更に細分化して実施する予定である。

3　プライバシー保護技術

パーソナルデータの利活用に適用できるプライバシー保護技術に関する情報収集を行うことを目的として、既存の方式の解析及び新方式の提案を行った。

3.1　暗号を使ったプライバシー保護技術：期間に依存した匿名性を持つグループ署名とその路車間通信への応用

車が取得したデータ（渋滞情報、道路状況、温度、速度、位置情報など）を路車間通信で報告するシステムにおいて、不正なデータの混在を防ぐために車両の正当性を検証することは重要である。しかしながら、通常の電子署名やメッセージ認証コードを用いると車両を一意に特定してしまうため、その位置情報が公のものとなり、例えば自宅や職場などの情報が漏洩する懸念がある。その対策としてグループ署名を応用した方式が挙げられる。グループ署名とは、あるグループに所属している署名者がグループに所属していることのみを証明可能な署名方式であり、グループ署名を用いて車両のプライバシーを考慮した路車間通信方式が多く提案されている[2][3]。しかしながら、その強い匿名性から経路情報などの"同じ車両が作成した署名であること（リンク付け）"から得られる有用な情報が欠落するという問題点がある、さらに、匿名であるがゆえに、廃車時や署名鍵漏洩に伴う鍵失効処理が非効率であるという問題点も挙げられる。一方で、仮名など常にリンク付けされる場合はプライバシーの観点から問題である。SCMS（Security Credential Management System）[4]では、仮名証明書発行機関が適時証明書を車両に発行することで、プライバシーをある程度担保しつつ経路情報などを得ることが可能であり、証明書発行対象車両を制限することで署名鍵失効も可能ではあるが、証明書更新に関する車両のコスト増が問題となる。

本研究では、期間に依存した匿名性を持つグループ署名を提案した[5]。ある期間に作成した署名はリンク付け可能であるが、期間をまたぐとグループ署名の意味で匿名性を保証する。車両は一切の証明書更新などの手続きを要請されず、また署名鍵の失効に関し署名者が一切の処理を要請されないため、廃車時／署名鍵漏洩時などにも効率的に対応可能である。図1に提案方式の概要を示す。車両が動いた軌跡に合わせて署名を作成し、OpenStreetMap [6]による地図上（小金

井周辺）に色のついた丸で表記する。同期間（図中では色が同じ丸で表記）に作成された署名からは、同じ車両が作成した署名であるが車両の特定まではなされないことを保証する（リンク付け）。これにより、ある期間内における車両の走行軌跡を取得することができる。一方、異なる期間（図中では色の異なる丸の部分）に作成された署名からは、どの車両が作成したのかの情報が一切漏洩しない。すなわちリンク付けされる期間をコントロールすることができ、プライバシーを考慮しつつ、経路情報を収集することができる。また通常の署名（DSA など）と比較しても遜色ない署名生成効率を実現しており、必ずしも計算能力が高くない車両でも署名が生成可能である。

3.2　暗号を使わないプライバシー保護技術：ランダム回答方式とその拡張の差分プライバシーによる評価

昨今のプライバシー意識の高まりにより、個人の識別ができる情報を発信する機器や集められたデータからの個人を識別する方法に対して厳しい目が向けられている。そのため、発信するデータに対してプライバシーをどのように確保するかが重要な問題となっている。このようなプライバシー保護方式のひとつとして、ランダム回答方式がある。ランダム回答方式は Warner により提案された方式 [7] である。ランダム回答方式はプライバシーに差し障りのある情報、例えば逮捕歴があるかというような質問をアンケート方式で集める方法である。しかし、このような問を正直に質問しても、回答者が正直に回答することに抵抗感があることは想像に難くない。そこで、以下のような方法でアンケートを取ることを考える。

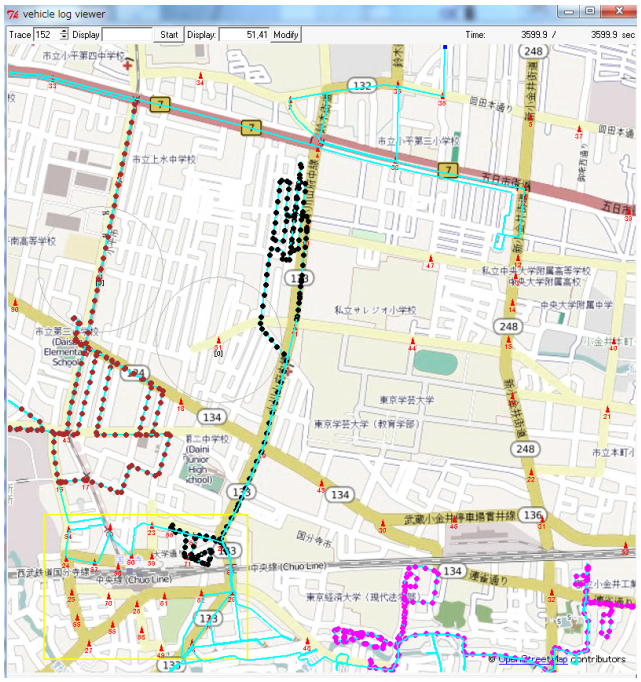

図 1　期間に依存した匿名性を持つグループ署名を利用した路車間通信システムにおける経路情報収集

今、質問者はカードを 3 枚用意する。1 枚目のカードには本来質問者がしたい質問が、2 枚目のカードには「はい、と回答」、3 枚目のカードには「いいえ、と回答」と書かれている。質問者は箱を用意し、3 枚のカードを箱に入れて、回答者に渡す。回答者は箱から 1 枚だけカードを引き、内容を確認し、質問者に見せることなく箱に戻す。最後にそのカードに書かれている内容に従い、質問者に回答する。

このようにすると、たとえ回答者が「はい」と答えたとしても、質問者にとっては 1 枚目のカードを引いて質問に答えたのか、それとも 2 枚目のカードを引いて「はい」と答えたのかを区別することはできない（同様に、「いいえ」という回答を得ても、本来「はい」と答える回答者が 3 枚目のカードを引いた結果「いいえ」と答えているかもしれない）ため、質問に答える抵抗感が低減されることが期待できる。

一方で、質問者にとって、本来の質問に「はい」と答える回答者の割合を見積もることも簡単にできる。例えば全体として N 人からアンケートを行い、そのうち Y 人が「はい」と答えたとする。このとき、本来の質問に「はい」と答える回答者の割合はおおよそ 3 Y/N − 1 と見積もることができる。

このランダム回答方式は単純な方式ではあるが、現在話題となっている匿名化技術の基礎であり、応用例も数多い。しかしながら、ランダム回答方式には多くの拡張方式があり、どの方式が最も優位であるかなどを比較するのが困難である。そこで本活動では、ランダム回答方式とその拡張方式、更には、応用方式の安全性評価を行った。安全性評価として、Dwork が提案した差分プライバシーの概念 [8] を用いることとした。本成果の一部は、下記のとおりである。

- 上記のランダム回答方式を一般化し、その差分プライバシーを導出した。
- Negative Survey[9] とその拡張である Selective Negative Survey[10] について評価を行い、差分プライバシーを満たさないことを示した。
- Negative Survey を 2 回以上繰り返すことで、差分プライバシーを満たすことを示した。

さらには、本成果を用いて、位置情報プライバシーやセンサーで取得された情報の差分プライバシーについても評価を行い、発表した [11]。

4　まとめ

セキュリティ基盤研究室ではワークショップやアンケート調査を使った実験を通じて、プライバシー侵害と感じる事例を収集し、解析を行った。また、プライバシー保護手法について現代暗号技術を用いた方式の提案や差分プライバシーによる評価などを行ってきた。

個人情報保護法の改正などにより、今後はこれまで以上に、プライバシーに関する事案は増えていくものと思われる。セキュリティ基盤研究室ではこれらの研究を発展させ、利用者のプライバシーを守りつつ、情報を利用できる仕組みの構築を目指して研究を続けていく。

【参考文献】

1 S. Kanamori, R. Nojima, H. Sato, and N. Tabata, “A study of willingness for private information providing,” Symposium on Cryptography and Information Security, 2016.
2 J. Guo, J. P. Baugh, and S. Wang, “A Group Signature Based Secure and Privacy-Preserving Vehicular Communication Framework,” in Mobile Networking for Vehicular Environments, 2007, pp.103–108.
3 Q. Wu, J. Domingo-Ferrer, and Ú. González-Nicolás, “Balanced trustworthiness, safety, and privacy in vehicle-to-vehicle communications,” IEEE T. Vehicular Technology, vol.59, no.2, pp.559–573, 2010.
4 W. Whyte, A. Weimerskirch, V. Kumar, and T. Hehn, “A security credential management system for V2V communications,” in IEEE Vehicular Networking Conference, 2013, pp.1–8.
5 K. Emura and T. Hayashi, “A light-weight group signature scheme with time-token dependent linking,” in LightSec, 2015, pp.37–57.
6 OpenStreetMap: https://www.openstreetmap.org/
7 S.L. Warner, “Randomized response: a survey technique for eliminating evasive answer bias,” Journal of the American Statistical Association (Taylor & Francis) 60 (309), pp.63–69, 1965.
8 C. Dwork, “Differential Privacy,” International Colloquium on Automata, Languages and Programming, 2006.
9 F. Esponda and V.M. Guerrero, “Surveys with negative questions for sensitive items,” Statistics & Probability Letters vol.79, Issue 24, 15, pp.2456–2461, 2009.
10 Shunsuke Aoki and Kaoru Sezakim, “Privacy-Preserving Data Mining with Perturbation for Multidimensional Data,” Technical Report of IEICE, vol.114, no.65, pp.143–147, 2014.
11 A. Waseda and R. Nojima, “Evaluation for randomized response techniques using differential privacy,” Symposium on Cryptography and Information Security, 2016.

野島　良　（のじま　りょう）
サイバーセキュリティ研究所
セキュリティ基盤研究室
主任研究員
博士（工学）
暗号理論、暗号プロトコル、情報セキュリティ、プライバシー、セキュリティ

金森祥子　（かなもり　さちこ）
サイバーセキュリティ研究所
セキュリティ基盤研究室
技術員
プライバシー

早稲田篤志　（わせだ　あつし）
サイバーセキュリティ研究所
セキュリティ基盤研究室
主任研究員
博士（情報科学）
情報セキュリティ

江村恵太　（えむら　けいた）
サイバーセキュリティ研究所
セキュリティ基盤研究室
主任研究員
博士（情報科学）
暗号理論

林　卓也　（はやし　たくや）
サイバーセキュリティ研究所
セキュリティ基盤研究室
研究員
博士（機能数理学）
暗号解析、高速実装