

交通機関の利用履歴の履歴長による本人特定性の検証

疋田 敏朗† 山口 利恵†

† 東京大学情報理工学系研究科
113-8654 東京都文京区本郷 7-3-1

toshi@yamagula.ic.i.u-tokyo.ac.jp, yamaguchi.rie@i.u-tokyo.ac.jp

あらまし 個人情報保護においてプライバシー保護は重要であると言われており、利用履歴を匿名化したとしてもその個人特有の利用履歴がある場合には本人の特定がしやすいと一般的に言われている。この本人特定リスクは利用履歴や乗降履歴の長さが長くなるにつれ、高くなるということが知られている。

本研究では大規模な移動履歴である「人の流れプロジェクト」の移動履歴から乗降履歴を生成し、実際の利用傾向による履歴の分布について考察を行った上で、履歴長と本人特定性の関係についての検討を行い、移動履歴のほとんどが二点間の移動であること、またほぼすべての履歴が一意であることを示した。

Personal Re-Identification study from Trajectory

Toshiro HIKITA† Rie Shigetomi Yamaguchi†

†The University of Tokyo
7-3-1 Hongo, Bunkyo, Tokyo 113-8654, JAPAN
toshi@yamagula.ic.i.u-tokyo.ac.jp, yamaguchi.rie@i.u-tokyo.ac.jp

Abstract Privacy Protection of Trajectories are getting important in now a days. Anonymized trajectories have a risk of Re-Identification, when the link length of trajectory get long.

In this paper, we will discuss uniqueness and Re-Identification risk of single trajectory from real large scale trajectory database, called People Flow. Our experiment shows about 98% of those trajectories are unique and most of trajectories are simple pattern. Also discuss about characteristic of Trajectory data like public transport IC card.

1 はじめに

近年、GPSを始めとする測位デバイスが携帯機器の機能の一部として搭載されるようになり、さらに携帯網の発展により携帯機器が測位した位置情報をセンター上のサーバに送信し、蓄積することが現実的になっている。さらにスマートフォンとそのアプリケーションや各種IoTデバイスの普及によって、一般のユーザから大量に位置情報を収集・蓄積することができるよう

になっており、蓄積した位置情報を活用することで従来は困難であった新たなサービスが次々と生まれるようになってきている。

また鉄道乗車券のICカードが進展し、これらの乗車券の利用履歴を用いることでIoTデバイスの移動履歴と同様に利用者の移動履歴を得ることも難しくなくなった。

これらの各種IoTデバイスやICカードから得られる移動履歴情報を活用すれば移動履歴の推定ならびに予測を高精度に行うことができる

と考えられる。

しかしながら、IoT デバイスや IC カードから発生する移動履歴情報はその内容に位置情報やその移動履歴を含むことから、個人を特定しうる情報やセンシティブな情報が含まれることからそのまま利活用することはできない。

すでに、鉄道乗車に関する IC カードの移動履歴が事業者間で提供されたり [1]、個人の移動履歴に関するプライバシー保護に関する懸念が示されるなどの事例も発生している。それらの懸念に対応するためには履歴を匿名化することによって対応を行うことが考えられる。すでに携帯事業者では位置情報の履歴を匿名化してプライバシー保護を行いつつ統計を行う [2] という事業が始まっている。

交通履歴データに関しても同様に移動履歴はプライバシーを保護し、個人を特定しないまま移動履歴を利用するための変換手法が望まれている。我々は都市開発においては最も必要と考えられる移動者の出発地と目的地の履歴に着目して匿名化手法を検討 [3] した。次に必要になるのは鉄道 IC カードのような移動履歴の匿名化である。

IC カードの利用履歴に関しては数学モデルによるプライバシー検討は行われていたが、大規模な実データを用いた乗降履歴の分布と個人特定性に関する議論は従来行われてこなかった。

本論文では、まず 2 章でまず位置情報と移動履歴に関する匿名化に関して紹介を行う、その上で従来の移動履歴情報の匿名化に関する研究はデータ提供者側のプライバシー保護に着目しており、移動履歴に対してノイズの形でダミーの移動履歴を加えるなどの加工を施すなど、データ利用の観点が存在しないという問題があることを示す。

次に 3 章では移動履歴データとして利用する「人の流れプロジェクト」の「2008 年東京圏人の流れデータセット（空間配分版）」の説明を行う。さらに 4 章ではこのデータを元に移動履歴の形に変換を行い、移動履歴の傾向に傾向と個人特定性がどの程度かという検討を行う。5 章では、4 章の実験結果について、データの特性も含めた議論を行う。

2 従来の研究

まず一般的な履歴データの匿名化について説明を行う。

個人を直接的かつ一意的に識別する属性、たとえば氏名¹、個人番号²などを示し、これを**個体識別属性**と呼ぶ

個人を一意的に識別できないとしても複数の属性を組み合わせると個人を一意的に識別できるものもある。たとえば性別、生年月日、住所などが該当する。これらの属性を**疑似識別属性** (Quasi Identifier, 以下 **QID**) と呼ぶ。

あるデータ T から個人が特定できないようなデータ T' を生成する変換作業を匿名化と呼ぶ。匿名化の手法としては k -匿名化 [4] が有名である。 k -匿名化は概念で同一の疑似識別属性に対して、最低でも $n \geq k$ のデータが存在するように、疑似識別属性を曖昧化する。例えば氏名情報のみを削除し、会員番号のみを利用する方法は一般的には仮名化と呼ばれる。仮名化を行っても個体識別属性が残っていると一意に識別できるため、仮名化は厳密な意味での匿名化ではない [5]。また、 k -匿名化の情報では、匿名化として不十分として、データの種別を定量的に計る手法 l -diversity [6] やデータの全体の割合傾向を計る手法 t -closness といった手法も提案されている [7]。

次に匿名化の位置情報への拡張について述べる。位置情報について k -匿名化を行った例 [8] は 2003 年に Gruteser らによって報告されている。この例では地点をグリッドごとに区切り、それぞれの地点情報をもとに k -匿名化が行われている。Gkoulalas-Divanis らによるまとめ [9] によれば、 k -匿名化の手法は一般的に今あるデータを中心とした区切り方と地形情報を活用したグリッドベースの区切り方の 2 種類に分けることができると主張している。

k -匿名化の他の匿名化手法としてはノイズを混入するという手法が挙げられる。文献 [10]

¹厳密には氏名だけでは同姓同名の個人が複数存在する可能性があるが、社会通念では個体識別属性とみなされている

²各個人に一意に割り当てられている番号、例えば日本でいえばマイナンバー、米国でいえばソーシャルセキュリティナンバー。

[11]では実際の位置情報の他に複数のダミーの位置情報を挿入させることでデータ自体の匿名性を担保する手法について記述されている。またダミーデータの混入手法についてはより高度な手法が提案されている。Niuら提案[12]によればダミーデータの配置場所を統計的に検討することで、ダミーユーザの現実的な配置が可能になり、より強固な配置が可能になるとされる。

移動履歴に関してもダミーデータを加えて匿名化するという手法が提案[13]されている。この手法はランダムにダミーデータを加えた移動履歴情報を生成することで、リアルユーザのデータを秘匿化する。しかしながらダミーを利用する方法では受領した位置情報にダミー情報がかかりの確率で紛れ込むため位置情報の利用者側から見るとデータが使いにくいという問題が発生する。例えば実際の情報に4倍のダミーデータを混入した場合、位置情報を的中させることができる確率を20%近くに低下させることができるが、利用者から見ると1/5でしか正確なデータが存在しないということになる。これは特にビッグデータ処理を前提とした場合にデータ自体の信頼性がなくなることを意味しているため、データの利用目的によってはこの手法は使えない。

また移動履歴をグリッド化して k -匿名化する方法はいくつか提案されている山口[14]の手法では単一のグリッドで k -匿名化を実施するという手法が提案されており、著者ら[3]は可変グリッドを利用した単体移動履歴の匿名化を提案している。

一方で乗車のような履歴の匿名化については菊池ら[15]が数学的モデルにより、鉄道駅の乗降客数データの分布から類推した移動履歴の匿名性に関する検討を行っている。

しかしながら大規模な実データを用いた乗降履歴の分布と個人特定性に関する議論は従来行われてこなかった。

3 個人特定性と移動履歴の匿名化要件

本論文の目的はIoTデバイスやスマートフォンの移動履歴を想定した実際の移動履歴を前提にどの程度の履歴データであれば個人特定性が存在するかを検討することである。

本章では上記目的を達成するために必要な手法について我が国の個人情報保護法に照らし合わせて、個人特定性を失わせることについて検討し、その上で移動履歴における個人特定性について検討する。

まず評価データについて説明を行う。その上で評価実験の結果について説明をする。

3.1 個人情報保護法と個人の特定

2015年8月現在における我が国の個人情報保護法においては、個人情報とは個人が特定できるような情報のほかに、『(他の情報と容易に照合することができ、それにより特定の個人を識別することができることとなるものを含む。)]』という形で他の情報と照合することで個人が特定できる情報もまた個人情報であるとされている。ここで識別とはそれが誰だかわからないが特有の1名に分離できるということであり、特定とはそれが固有の1名を示すこととされる。

ここであるデータ T が存在した場合にそのリストの全項目をQIDとして k -匿名化を行ったデータ T' については、1つの情報について少なくとも2つ以上の列が該当することから一意に特定ができないことが知られている。すなわちデータ T の全項目をQIDとして k -匿名化による変換を行うことができればその情報について個人を特定することはできないとすることができる。

3.2 移動履歴の匿名化のために必要な要件

昨今、FoursquareやFacebookなどのSNSへの投稿は位置情報を付加することが可能であったり、地点情報を付加してチェックインすることができるため、別の手段で収集された履歴情

報や位置情報とデータ処理を行う履歴情報を照合することで個人を特定することも難しくはなくなりつつある。そのため個々の履歴情報に関して、他の情報を用いた場合でも個人を一意特定しうる状態ではないことが必要ということになる。

本研究では上記の事情を鑑みたと移動履歴の長さが個人特定性にどのような影響をあたえるのかを検討することとする。移動履歴から個人が特定されるケースを列挙すると以下のようなケースが考えられる。

1. その情報自体が個人を特定できる情報を含む場合
2. 位置情報自体が自宅などを指し示す場合
3. 位置情報と時刻の組み合わせにより個人が特定される場合
4. 履歴が固有のために個人が特定される場合

今回はこのうち4について検討をすることとする。これは組み合わせ情報が固有であるために個人が特定できることである。履歴情報は同じIDの履歴をリンクさせることで個人が特定可能であると指摘をされている。

そこで実際の移動履歴を元にどの程度の履歴長であれば、個人が特定をされるのかについて実際のデータを元に検討することとする。

3.3 使用データについて

今回提案手法の実験を実施するためには現実的な移動履歴を持ち、人数が多く、なおかつユーザからの同意または個人情報保護上問題がないデータを利用する必要がある。そこで上記の条件を満たすデータとして、東京大学空間情報科学研究センターの「人の流れプロジェクト」[16]の「2008年東京都市圏 人の流れデータセット（空間配分版）」のデータを利用することとした。

今回の実験データは東京都市圏交通計画協議会が収集したパーソントリップ調査によるデータを元にしている。実験データとしては関しては元のパーソントリップ調査のデータを用いて、

住所詳細を記載していないものをベースとし、以下に示す空間配分を行った空間配分版とした。空間配分とはゾーンごとにまとめられた地点情報について、個々人の位置情報をゾーン範囲内の建物の分布に合わせて詳細位置に確率的に再配分し、現実のデータに近づける処理のことである。

表1に、人の流れプロジェクトのデータにおける位置履歴情報定義を利用した位置履歴情報の例を引用する。このデータ定義については実際のものだが、データ自体に関しては定義に合わせて著者が作成したダミーデータとなっている。

この例では20-25歳の学生でかつ女性であるユーザ12345は、東京大学構内から徒歩で本郷三丁目駅に移動し、本郷三丁目駅から新宿駅まで移動をした後に、新宿駅から高尾山口駅まで鉄道で移動し、最後に高尾山山頂まで徒歩で移動をした、この移動の目的はレジャーであったことがわかる。

今回利用したデータセットには本データセットには特定日付の586882ユーザ分の移動履歴が含まれている。この移動履歴のうち、鉄道乗車券のデータを念頭におき、一日のうち一回でも鉄道を利用した188275人の履歴を用いることにした。

4 実移動データを用いた移動履歴データと個人特定性の検討

本章では3章で示したデータを用いて移動履歴と個人特定性の関係について検討する。まず実際のトリップ情報から移動履歴を抜き出す手法を説明し、その上で移動履歴データの分布について検討を行う。さらに移動履歴と個人特定性の関係について考察を行う。

4.1 実験用の移動履歴データの生成

本節では移動履歴データの生成方法を示す。1のデータを用いる。

まず移動データのうち交通手段が鉄道すなわち”12”であるものを抜き出す。今回であればTrip1のSubTrip2が該当である。それぞれの始

表 1: 人の流れプロジェクトのデータ例 (データはダミーデータ)

ID	番号	サブ	日時	経度	緯度	性別	年齢	職業	目的	手段
12345	1	1	2014/12/10 10:05	139.7619	35.7143	2	4	13	99	1
12345	1	1	2014/12/10 10:20	139.7605	35.7075	2	4	13	99	1
12345	1	2	2014/12/10 10:20	139.7605	35.7075	2	4	13	99	12
12345	1	2	2014/12/10 10:40	139.7001	35.6909	2	4	13	99	12
12345	1	2	2014/12/10 12:00	139.2696	35.6321	2	4	13	99	12
12345	1	3	2014/12/10 12:00	139.2696	35.6321	2	4	13	99	1
12345	1	3	2014/12/10 13:30	139.2436	35.6251	2	4	13	99	1

点と終点 [本郷三丁目] と [高尾山口] を抜き出す。鉄道乗車券であれば [本郷三丁目] と [新宿], [新宿] と [高尾山口] という履歴が残るはずであるが、人の流れプロジェクトのデータが始点と終点のみが性格に抜き出せる仕様であるため、乗換駅での乗降履歴については検討に入れることができなかった。

抜き出したデータを各ユーザごとに整理を行う。この場合であれば「12345:[本郷三丁目, 高尾山口]」であり、このあと渋谷に向かったとすれば「12345:[本郷三丁目, 高尾山口, 渋谷]」となる。

ここで駅数を履歴に関与する駅の数、履歴超を移動区間の数と定義する。先ほどの [本郷三丁目],[高尾山口],[渋谷] の3駅が履歴にある場合は移動履歴における駅数は3であり、履歴長は2となる。今後は履歴長を中心に特定性との関係を見ていく。

このように整理した各ユーザの1日分のデータを元にその移動履歴が他のユーザの1日の移動履歴と重複しているかどうか?について検討する。

履歴において個人特定性があるとは、その履歴が単独であるということであり、同じ履歴が複数存在すれば個人特定性はないこととなる。

4.2 実データからの移動履歴データの生成

まず各ユーザの履歴を履歴長ごとに整理する。鉄道を利用した移動について1日のデータを履歴長とその数で整理すると表2と図1に示すようになる。

履歴の総数はユーザ数にして188275ユーザ、各ユーザごとの1日の履歴を記録しているため、2区間の履歴が最も多く、82%の履歴が2区間の利用履歴となっている。このうち146321件94%の履歴が自宅最寄り駅→出先最寄り駅→自宅最寄り駅という経路である。

次に移動履歴のうち、移動履歴が一意であるものを抽出する。ここでは一日分のデータを蓄積し、時間については考慮に入れないこととした。また、履歴は履歴自体でのマッチングとし、部分的な履歴が一致するパターン、すなわち [A] → [B] と [A] → [B] → [C] についてはそれぞれを一意とし、重複していないこととする。

この場合に一意である履歴は184573件であり全履歴の98%が他の履歴と重複しない一意なものであることがわかる。

すなわち乗車カードやIoTによる移動履歴単体をそのまま利用した場合は時刻情報を削ったとしても日単位で98%の移動履歴は単一であって、他の履歴と区分されるため、ユーザを外部観察できていれば個人を特定することができることになる。

4.3 データ量を半減させた場合の移動履歴データ

??節で生成した移動履歴データは一意のデータが98%となっていた。一意のデータがお多いのは移動履歴のユーザ数が少なすぎるためなのか?それともユーザ数とは関係がないのか?前記の疑問を調査するためにユーザ数を半減させた実験データを生成して比較することとした。

表 2: 移動履歴数と履歴長

履歴長	履歴数	一意履歴数	個人特定率
1	8000	7410	92.6
2	154997	152122	98.1
3	17396	17188	98.8
4	6257	6229	99.6
5	1132	1131	99.9
6	370	370	100
7	79	79	100
8	37	37	100
9	5	5	100
10	1	1	100
11	1	1	100

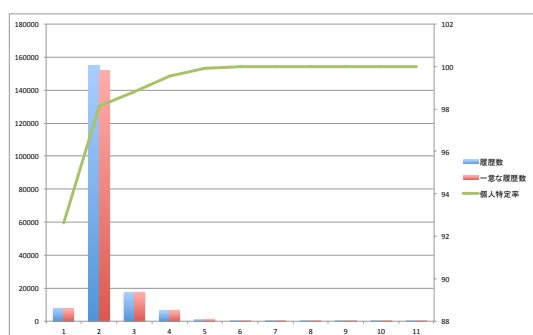


図 1: 移動履歴数と履歴長

半減したデータを生成するにあたってはランダムに生成することとし、ユーザ ID を奇数/偶数で分けるようにした。それ以外の方法は移動履歴データと同じである。

今回はユーザ ID が奇数のユーザを抽出し、生成した履歴データの特性を表 3 と図 2 に示す。ユーザ数は 93994 人であり、ユーザー数の 49.9% をカバーしている。

全数の移動履歴に比べて、一位履歴の数が増え、個人特定率が上がっていることがわかる。場合の数が増加するためであり、結果としては妥当であることがわかる。

5 移動履歴と個人特定性に関する考察

本実験結果を受けて、移動履歴データと個人特定性に関して以下の三点で考察を行う。

表 3: データ量を半減させた場合の移動履歴数と履歴長

履歴長	履歴数	一意履歴数	個人特定率
1	3985	3972	99.7
2	77332	77197	99.8
3	8770	8765	99.9
4	3090	3087	99.9
5	554	554	100
6	196	196	100
7	45	45	100
8	19	19	100
9	1	1	100
10	1	1	100
11	1	1	100

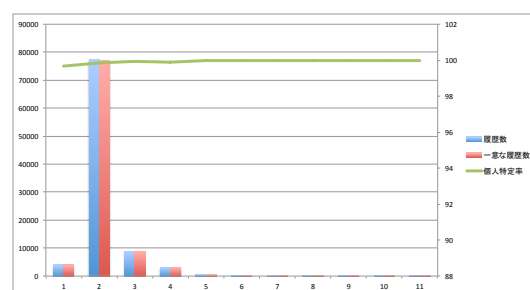


図 2: データ量を半減させた場合の移動履歴数と履歴長

- 移動履歴データの履歴と個人特定性について
- 履歴数と個人特定性の関係について
- 履歴長と個人特定性の関係について

5.1 移動履歴データの履歴と個人特定性について

本説では、生成した移動履歴データから履歴の一意性と個人特定性に関して議論を行う。

今回の履歴生成では一日分の移動履歴をまとめて、履歴としている。その上で移動履歴自体が一意であるかどうかを判断した。この方法には二点ほどの問題がある。ひとつは短い履歴がより長い履歴の部分を構成した場合である。もう一つは乗換駅の存在である。

履歴が部分一致した場合、今回は異なる履歴として扱っている。部分一致した部分については一意でないと判断するほうが自然ではあるが、その場合は一致していない部分を前の履歴と当区分するのか？という問題が発生する。履歴としての linkability を考慮しない場合には部分一致した履歴と、残りの部分履歴を分離する形になる。そのような分離した履歴で問題がないかどうか？についての検討が必要である。

また今回調査した結果では、90%近いの移動履歴が自宅最寄り駅、目的地の往復、または自宅最寄り駅から複数の目的地、最後に自宅最寄り駅という形での移動であった。このタイプの移動履歴に関しては履歴長3以上の移動履歴が履歴長2の移動履歴に部分一致する可能性は低いことを考慮する必要がある。

もちろん両方共、一意の履歴であることから2つの履歴が削除されるのであれば、部分一致した部分だけでも残すべきという考え方は理解できる。個人特定性も考慮した上で必要な区間を抽出する方法については議論が必要である。

次に今回の履歴独特の制限として、一つの移動目的自体の発駅と着駅でしか評価ができないという問題があげられる。これは人の流れプロジェクトのデータと大本のパーソントリップデータが乗換駅を考慮していないことが原因である。対策としてはデータ生成時に可能性のある経路を探索する方法が考えられる。

例えば[千歳烏山]から[東京テレポート]という経路の場合は[千歳烏山]から[京王新宿]、[新宿]から[大崎]、[大崎]から[東京テレポート]という履歴を生成する必要がある。

ただし、鉄道乗車券の経路をエミュレートする場合は直通運転または改札内(ラッチ内)乗り換えのような乗車券に履歴が残らない乗換方法についても考慮する必要がある。先の例で言えば[大崎]に関しては直通運転かつ改札内乗り換えであるため、乗り換えの履歴には入れないような処理を行う必要がある。

このような場合に問題なく、乗車カードの履歴同様の経路を抜き出す方法については検討が必要であると考えられる。

5.2 履歴数と個人特定性の関係について

次に履歴のデータ数に関してである。今回はユーザ数が60万人弱、鉄道利用者数が19万人弱のデータを用いたがそのほとんどの利用履歴が一意であるという結果であった。ユーザー数を9万人強とした場合はさらにそれ以上の率で一意であった。

この点から明らかなのは一日の総利用者数が10万人オーダではほぼすべての利用履歴が一意であるということである。特にパーソントリップ調査の場合はある程度特定の集団に調査票を配布しているという事実(すなわちより重複は起こりやすい)を考慮に入れても、この数では足りないとかんがえられる。

実際の移動履歴をそのまま利用し、IDを仮名化するだけで個人が特定されないというには現在のデータ量では望み薄であり、なんらかのデータ加工が必要ということを示唆している。データの増加量を考えると少なくとも百万人オーダ、できれば千万人オーダが必要と考えられる。

ただし広域東京圏の人口は四千万人程度、毎日鉄道を利用するユーザの数を考えると非常に非現実的な想定となる。すなわち何らかの形で匿名化かそれ以外の手法が必要とかんがえられる。

5.3 履歴長と個人特定性の関係について

著者が行っている検討[3]によれば、すべての移動が履歴数1であるとした場合にはエリア広域化の手法を用いずとも6-7割程度の移動についてはk匿名化ができることを確認している。

また著者らのグループはlinkabilityを失わせた履歴長1の移動履歴については個人特定率は25%程度であること調べている。すなわち一日の移動履歴すべてではなく、特定の履歴部分を抽出すれば匿名性を保つことは可能である。

一方、購買や移動に関してはユーザの志向を理解するためにもある程度の履歴長が履歴長が必要であると言われている。しかしながら18万人のデータでなおかつ履歴長が2であるデータ

であっても重複しない可能性が極めて高いことをこの実験データは示している。

6 まとめ

本論文では近年の ICT 技術の発展に応じて、スマートフォンのアプリケーションや IoT デバイス、特に IC カードの移動履歴に関してその個人特定性の検討を行うことを目的に、個人が特定されないという個人情報保護法への対応に関する検討を行った。

そのうえで移動履歴から個人が特定されないために必要な項目を検討し、とくに履歴の長さに関する検討が必要であることを説明した。

さらに東京大学空間情報科学研究センターの「人の流れプロジェクト」の「2008 年東京圏人の流れデータセット（空間配分版）」のデータを利用して、鉄道の利用履歴のデータを生成し、履歴のほとんどが自宅最寄り駅→目的地→自宅最寄り駅の 2 点の往復であること、また 98% のデータが一意であり、履歴データをそのまま加盟化して用いた場合は個人特定がなされやすい状態であることを示した。

さらに移動履歴のに関する部分一致の可能性について検討し、履歴長が 2 以上である移動履歴に関して、一意性を減少させる方法に関する検討が必要であることを示した。今後は履歴長を維持しつつ、個人特定性の少ない匿名化手法に関する研究を進めていきたい。

なお、本研究は東京大学空間情報科学研究センターの「人の流れプロジェクト」との共同研究である。データの整備並びに提供を行っていただいた空間情報科学研究センター各位に感謝する。

参考文献

- [1] ”Suica に関するデータの社外への提供についての有識者会議”：“Suica に関するデータの社外への提供について” (2014).
- [2] 寺田雅之：“モバイル空間統計：携帯電話ネットワークを活用した人口推計技術とその応用（ビッグデータ特別セッション）”, pp. 63–66 (2014).
- [3] 疋田敏朗, 山口利恵：“階層化符号表現を利用した移動履歴の匿名化手法”, マルチメディア、分散、協調とモバイル (DICOMO2015) シンポジウム 2015 情報処理学会 (2015).

- [4] L. Sweeney: “k-anonymity: a model for protecting privacy”, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, **10**, 5, pp. 557–570 (2002).
- [5] 板倉陽一郎 伊藤孝一 菊池浩明 高木浩光 高橋克巳 中川裕志 疋田敏朗 廣田啓一 山口利恵：“「完全な匿名化」幻想を超えて”, 暗号と情報セキュリティシンポジウム 2014 電子情報通信学会 (2014).
- [6] A. Machanavajjhala, D. Kifer, J. Gehrke and M. Venkatasubramanian: “l-diversity: Privacy beyond k-anonymity”, *ACM Transactions on Knowledge Discovery from Data (TKDD)*, **1**, 1, p. 3 (2007).
- [7] D. Rebollo-Monedero, J. Forné and J. Domingo-Ferrer: “From t-closeness to PRAM and noise addition via information theory”, *Privacy in Statistical Databases* Springer, pp. 100–112 (2008).
- [8] M. Gruteser and D. Grunwald: “Anonymous usage of location-based services through spatial and temporal cloaking”, *Proceedings of the 1st International Conference on Mobile Systems, Applications and Services, MobiSys '03*, New York, NY, USA, ACM, pp. 31–42 (2003).
- [9] A. Gkoulalas-Divanis, P. Kalnis and V. S. Verykios: “Providing k-anonymity in location based services”, *SIGKDD Explor. Newsl.*, **12**, 1, pp. 3–10 (2010).
- [10] H. Lu, C. S. Jensen and M. L. Yiu: “Pad: privacy-area aware, dummy-based location privacy in mobile services”, *Proceedings of the Seventh ACM International Workshop on Data Engineering for Wireless and Mobile Access* ACM, pp. 16–23 (2008).
- [11] H. Kido, Y. Yanagisawa and T. Satoh: “An anonymous communication technique using dummies for location-based services”, *Pervasive Services, 2005. ICPS'05. Proceedings. International Conference on IEEE*, pp. 88–97 (2005).
- [12] B. Niu, Q. Li, X. Zhu, G. Cao and H. Li: “Achieving k-anonymity in privacy-aware location-based services”, *Proc. IEEE INFOCOM* (2014).
- [13] P. Shankar, V. Ganapathy and L. Iftode: “Privately querying location-based services with sybil-query”, *Proceedings of the 11th international conference on Ubiquitous computing* ACM, pp. 31–40 (2009).
- [14] R. S. Yamaguchi, K. Hirota, K. Hamada, K. Takahashi, K. Matsuzaki, J. Sakuma and Y. Shirai: “Applicability of existing anonymization methods to large location history data in urban travel”, *Systems, Man, and Cybernetics (SMC), 2012 IEEE International Conference on IEEE*, pp. 997–1004 (2012).
- [15] 菊池浩明, 高橋克巳：“乗降履歴データの安全な匿名化は可能か?”, 暗号と情報セキュリティシンポジウム 2014 電子情報通信学会 (2014).
- [16] Y. Sekimoto, R. Shibasaki, H. Kanasugi, T. Usui and Y. Shimazaki: “Pflow: Reconstructing people flow recycling large-scale social survey data”, *IEEE Pervasive Computing*, **10**, 4, pp. 0027–35 (2011).