# PWS Cup 2019: ID識別・トレース推定に強い位置情報の匿名加工技術を競う

村上 隆夫[1,a)]　荒井 ひろみ[2]　井口 誠[3]　小栗 秀暢[4]　菊池 浩明[5]　黒政 敦史[6]　中川 裕志[2]
中村 優一[7]　西山 賢志郎[8]　野島 良[9]　波多野 卓磨[10]　濱田 浩気[11]　山岡 裕司[4]　山口 高康[12]
山田 明[13]　渡辺 知恵美[14]

**概要**：2017 年 5 月に改正個人情報保護法が施行され，パーソナルデータは匿名加工情報に加工することで，本人の同意なしに第三者提供ができるようになった．一方，標準的な匿名加工の方法が定まっておらず，パーソナルデータの利活用に向けて，優れた匿名加工の方法を明確にすることが重要課題となっている．我々はこの課題を解決するため，匿名加工データの有用性と安全性を競い合うコンテストを毎年実施している．これまでに，疑似ミクロデータ（全国消費実態調査）や購買履歴の匿名加工を対象としたが，本年度は「位置情報」の匿名加工を対象とする．本稿ではその内容を説明する．

**キーワード**：位置情報プライバシー，匿名加工，ID 識別，トレース推定

# PWS Cup 2019: Location Data Anonymization Competition

Takao Murakami[1,a)]　Hiromi Arai[2]　Makoto Iguchi[3]　Hidenobu Oguri[4]　Hiroaki Kikuchi[5]
Atsushi Kuromasa[6]　Hiroshi Nakagawa[2]　Yuichi Nakamura[7]　Kenshiro Nishiyama[8]
Ryo Nojima[9]　Takuma Hatano[10]　Koki Hamada[11]　Yuji Yamaoka[4]　Takayasu Yamaguchi[12]
Akira Yamada[13]　Chiemi Watanabe[14]

***Abstract:*** The amended act on the protection of personal information, which has been enforced since May 2017, states that personal data can be provided to a third party without users' consent if the data are anonymized as "anonymously processed information." However, anonymization methods are not clear, and hence we annually hold PWS Cup to clarify secure and appropriate anonymization methods. This year, we focus on location data, and hold location data anonymization competition. This paper describes its contents.

***Keywords:*** location privacy, anonymization, ID-disclosure, trace inference

## 1. はじめに

2017 年 5 月 30 日に改正個人情報保護法が施行され，パー

[1] 国立研究開発法人 産業技術総合研究所
AIST
[2] 国立研究開発法人 理化学研究所
RIKEN
[3] Kii 株式会社
Kii Corporation
[4] 株式会社富士通研究所
Fujitsu Laboratories Ltd.
[5] 明治大学
Meiji University
[6] 富士通クラウドテクノロジーズ株式会社
FUJITSU CLOUD TECHNOLOGIES LIMITED
[7] 早稲田大学
Waseda University
[8] 株式会社ビズリーチ
BizReach, Inc.
[9] 国立研究開発法人 情報通信研究機構
NICT
[10] 日鉄ソリューションズ株式会社
NS Solutions Corporation
[11] NTT セキュアプラットフォーム研究所
NTT Secure Platform Laboratories
[12] 株式会社 NTT ドコモ
NTT DOCOMO, Inc.
[13] 株式会社 KDDI 総合研究所
KDDI Research, Inc.
[14] 筑波技術大学
Tsukuba University of Technology
[a)] takao-murakami@aist.go.jp

ソナルデータを適切な匿名加工（ノイズ付与，一般化，削除などのデータ加工，および仮名化）を施すことで，本人の同意なしに第三者提供をすることができるようになった．しかし，適切な匿名加工の方法は，自明ではない．匿名加工の方法や安全性基準は古くから数多く提案されており [1]，「どのような匿名加工が適切か」については個人情報保護委員会が匿名加工に関するガイドライン [2] などに示された情報だけでは不十分である．

EU データ保護指令の作業部会である第 29 条作業部会（Article 29 Working Party）が公表した「匿名化技術に関する意見書」（Opinion 05/2014 on Anonymisation Techniques）[3] では，データが匿名加工されているかどうかを判断する際に考慮すべきリスクとして，以下の 3 つが定義されている．

**Singling out:** レコードが識別（single out）されるリスク

**Linkability:**（同一，あるいは 2 つの異なるデータベースから）個人に関する複数のレコードが紐付けされるリスク

**Inference:** 個人の属性が高い確率で推定されるリスク

これらのリスクに対して安全であることが望ましいとしつつも，各匿名加工技術はそれぞれ長所・短所があり，「どのような匿名加工技術が適切か」についてはケースバイケースと述べるに留まっている．

優れた匿名加工の方法を明確にするため，我々は匿名加工データの有用性と安全性を競い合うコンテスト「PWS Cup」を 2015 年度から実施している．2015 年度には疑似ミクロデータ（全国消費実態調査）を対象とし，2016〜2018 年度には購買履歴を対象としたコンテストを実施した．

## 1.1 PWS Cup 2019

2019 年度は，「位置情報」の匿名加工を対象とするコンテストを実施する[*1]．以下，コンテストの特徴を述べる．

### 1.1.1 位置情報コンテスト

我々の知る限り，位置情報のコンテストは世界的に見ても本コンテストが初である．位置情報コンテストを実施する背景としては，位置情報データの利活用に対する期待が挙げられる．近年，周辺の飲食店などの POI（Point of Interest）検索や，目的地への経路検索といった位置情報サービス（LBS: Location-based Service）が広く利用され，その結果として大量の「トレース」（位置情報を時系列に並べた移動履歴）[*2]が LBS プロバイダーに蓄積されている．この大量のトレースは位置情報ビッグデータとも呼ばれ，人気スポットの分析 [4]，POI に対する（飲食店，ホテルなどの）カテゴリーの自動タグ付け [5]，外国人観光客の動態分析 [6] など，様々な用途に利活用できる．

その一方で，トレースから自宅や通院している病院などが特定される恐れがあり，秘密にしておきたい交友関係なども推測される恐れがある．また，仮名化されたトレースからも高い確率で元のユーザ ID が識別される恐れがあり [7]，仮名化による対策だけでは不十分であることも知られている．従って，トレースを第三者提供する前に匿名加工が必要となるが，トレースに対する適切な匿名加工の方法を見つけるのは容易ではない．例えば，匿名加工技術として「差分プライバシー」と呼ばれる安全性基準を満たす技術が近年特に注目されているが，長いトレースのような時系列データには向いていない可能性がある．具体的には，トレースが長くなるほど privacy budget $\epsilon$ と呼ばれるパラメータが大きくなり，安全性が低下する恐れがある [8]．

トレースに対して，高い有用性と安全性を実現する匿名加工技術はまだよく知られておらず，これを明確にするための一手段として，位置情報コンテストが有効だと考える．

### 1.1.2 部分知識攻撃者モデル

安全性を評価する上で，攻撃者の背景知識を考える必要がある．攻撃者の背景知識に関するモデルとしては，攻撃者が匿名加工前の元データを知っている「最大知識攻撃者モデル」と，匿名加工前の元データを知らず，他のデータに関する知識を部分的に持っている「部分知識攻撃者モデル」の 2 つが知られている．前者は攻撃者として匿名加工データの提供元事業者，後者は提供先事業者を仮定している．

位置情報は特異性が高いことが知られており，例えば文献 [9] では，4 個の位置からなるトレースが unique である（具体的には，そのトレースを持つ人が 150 万人中 1 人である）割合が約 95% であることを示している．従って，最大知識攻撃者モデルでは，元トレースに対してほとんどの位置情報を削除するなどの大幅な加工をしないと，匿名加工トレースから元のユーザ ID が識別される恐れがある．しかし，最大知識モデルでは，そもそも攻撃者が元トレースを全て知っているため，ID 識別をする必要はない（元トレースに関するプライバシーは既に完全に破れている）．即ち，最大知識モデルは攻撃者の背景知識として最悪ケースを考慮した安全性評価ができるという利点はあるものの，現実からは乖離したモデルになっている．

従って，本コンテストでは部分知識攻撃者モデルの下で安全性を評価する．即ち，攻撃者は元トレースを全く知らず，他のトレースに関する部分知識を参考にしながら，匿名加工トレースに対して ID 識別などの攻撃を試みるものとする．本稿では，この部分知識として用いる他のトレースを「参照トレース」と呼ぶ．攻撃者が入手可能な参照トレースとして，例えば個人が SNS などで公開している位置情報などがあるが，個人が普段から公開している位置情報は一般には多くない．従って，参照トレースは現実には少量と考えられる．

[*1] 尚，改正個人情報保護法における匿名加工情報の加工基準と，本コンテストにおける匿名加工の安全性基準は異なることに注意する．具体的には，前者では加工前の元データを保有する事業者（即ち，最大知識攻撃者モデル）において特定の個人を識別できないように加工することを求める一方で，その識別能力と手法は一般人基準であることをガイドライン [2] で示している．後者では第 1.1.2 節で述べるように，攻撃者として提供先事業者（即ち，部分知識攻撃者モデル）を仮定し，その識別能力と手法は専門家（研究者）レベルである．本コンテストでの攻撃に耐えうる匿名加工技術と，一般人基準の最大知識攻撃者に耐えうる匿名加工技術の関係の明確化については今後の課題である．また，本コンテストで得られる知見が，第 1.1.3 節で述べる「ID 識別とトレース推定の 2 軸での評価」も合わせて，将来的な法制度の在り方を議論する上での参考になればと考えている．

[*2] 英語では trace あるいは trajectory と呼ばれる．

本コンテストでは，このことを考慮して参照トレースの長さを設定する（詳細は第 3.1 節を参照）．

#### 1.1.3 ID 識別とトレース推定の 2 軸での評価

本コンテストでは，評価する具体的なリスクとして「ID 識別」と「トレース推定」の 2 つを考える．ID 識別は各匿名加工トレースに対して，元のユーザ ID を推定する攻撃であり，トレース推定は元トレースを完全に復元しようとする（例えば，ユーザ数が $n$ 人，各ユーザの位置情報が $t$ 個あったときに，計 $nt$ 個の位置情報を推定する）攻撃である．ID 識別は再識別とも呼ばれ，トレース推定はトラッキング攻撃 [10] とも呼ばれる．

改正個人情報保護法は匿名加工情報に対するリスクとして ID 識別を考えている．しかし，ID 識別はノイズ付与や一般化などの加工が施されたトレースに対しては，ユーザと元の位置情報の対応付けをしない（ユーザ ID のみ推定する）のに対して，トレース推定はユーザと元の位置情報の対応付けを行う（元トレースを完全に復元する）ため，トレース推定への対策も重要と考えられる．匿名加工トレースは通常，ユーザ ID 単位でのシャッフルに相当する仮名化が施されているため，トレース推定を行うには ID 識別をある程度行う必要があるが，ID 識別に対して安全な匿名加工がトレース推定に対して安全とは必ずしも言い切れない．例えば，$k$-匿名化は ID 識別率が $1/k$ 以下になることを理論的に保証するが，攻撃者に ID 識別を行うことなく属性推定されるリスクが残る [11]．同様に，攻撃者が ID 識別を行うことなく，トレース推定を行う例も存在する（[12] の第 3 章参照）．このように，ID 識別に対する安全性とトレース推定に対する安全性の相関関係は自明ではない．

以上を踏まえて，本コンテストでは匿名加工トレースに対して ID 識別とトレース推定の 2 軸での評価を行う．具体的には，各チームに「ID 識別対策用」と「トレース推定対策用」の 2 つのデータセットを配布する．前者，後者に対して，それぞれ ID 識別，トレース推定に強くなるように元トレースを匿名加工したチームを評価する．一方，攻撃は全匿名加工トレースに対して，ID 識別とトレース推定の両方を行い，ID 識別安全性，トレース推定安全性を下げることに成功したチームを評価する．このようにすることで，各匿名加工トレースに対して，ID 識別安全性とトレース推定安全性の相関関係を明らかにする（詳細は第 2.3 節を参照）．

尚，有用性に関しては，本コンテストでは「要求値」を設け，その要求値を下回る匿名加工トレースは「無効」とする．要求値以上の匿名加工トレースを「有効」とし，それに対して ID 識別とトレース推定の 2 軸での評価を行う．

本コンテストで得られる知見が，将来的な法制度の在り方を議論する上での参考となれば幸いである．

## 2. コンテストの概要

### 2.1 記号の表記

以下，本稿で用いる記号を定義する．自然数の集合を $\mathbb{N}$ とし，非負の実数の集合を $\mathbb{R}_{\geq 0}$ とする．また，$a \in \mathbb{N}$ に対して $[a] = \{1, \ldots, a\}$ と表記する．

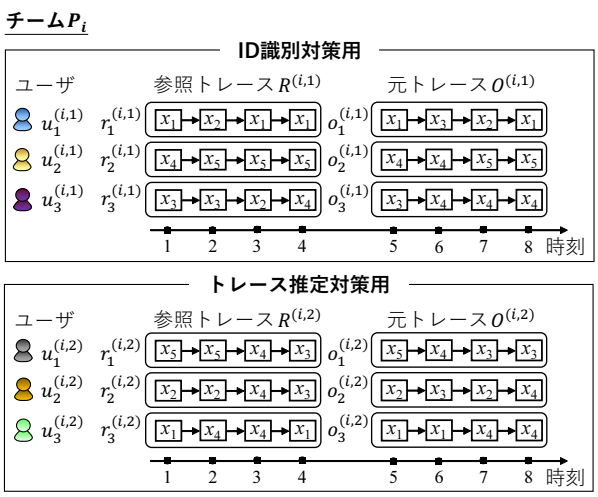


**図 1** 位置情報データセットの例（$i \in [z]$，$n = 3$，$m = 5$，$t = 4$）

本コンテストに参加するチーム数を $z \in \mathbb{N}$ とし，チーム集合を $\mathcal{P} = \{P_1, \ldots, P_z\}$ とする．各チームには，予め位置情報データセットを 2 つ配布する．1 番目のデータセットは ID 識別対策用であり，2 番目のデータセットはトレース推定対策用である．また，各位置情報データセットにおいて，ユーザ数は同じだが，ユーザ集合は異なる（詳細は第 3.1 節を参照）．ユーザ数を $n \in \mathbb{N}$ とし，$i \in [z]$ 番目のチーム $P_i$ に配布する $j \in \{1, 2\}$ 番目の位置情報データセットにおけるユーザ集合を $\mathcal{U}^{(i,j)} = \{u_1^{(i,j)}, \ldots, u_n^{(i,j)}\}$ とする．ユーザ $u_k^{(i,j)}$ の右下の数字 $k \in [n]$ を「ユーザ ID」と呼ぶ．即ち，ユーザ ID は 1 から $n$ までの自然数である．

位置情報は，対象とするエリアを小さい領域に分割することで離散化する（本コンテストでは，東京中心部を均等に $32 \times 32$ 個の領域に分割する）．離散化された位置情報の数を $m \in \mathbb{N}$ とし，位置情報の集合を $\mathcal{X} = \{x_1, \ldots, x_m\}$ とする．位置情報 $x_k$ の右下の数字 $k \in [m]$ を「領域 ID」と呼ぶ．即ち，領域 ID は 1 から $m$ までの自然数である．時刻についても，一定時間おきに区切ることで離散化する（本コンテストでは 30 分おきに区切って離散化する）．各時刻は自然数で表す．

各位置情報データセットは，配布先のチームが匿名加工を施す「元トレース」と，他のチームが ID 識別／トレース推定を行う際の参考情報である「参照トレース」から構成される．参照トレースは時刻 1 から時刻 $t \in \mathbb{N}$ までのトレースであり，元トレースは時刻 $t + 1$ から時刻 $2t$ までのトレースである．$k \in [n]$ に対してユーザ $u_k^{(i,j)}$ の参照トレースを $r_k^{(i,j)} \in \mathcal{X}^t$，元トレースを $o_k^{(i,j)} \in \mathcal{X}^t$ とし，$R^{(i,j)} = (r_1^{(i,j)}, \ldots, r_n^{(i,j)})$，$O^{(i,j)} = (o_1^{(i,j)}, \ldots, o_n^{(i,j)})$ とする．$R^{(i,j)}$ と $O^{(i,j)}$ はそれぞれ，$i \in [z]$ 番目のチーム $P_i$ に配布する $j \in \{1, 2\}$ 番目の位置情報データセットにおける参照トレース，元トレースをユーザ ID 順に並べたものである．$R^{(i,j)}$ も $O^{(i,j)}$ も，$nt$ 個の位置情報からなる．

チーム $P_i$ に配布する位置情報データセットの例を図 1 に示す．この例では，ユーザ数 $n$ は $n = 3$，領域数 $m$ は $m = 5$，時刻 $t$ は $t = 4$ であり，$o_1^{(i,1)} = (x_1, x_3, x_2, x_1)$，

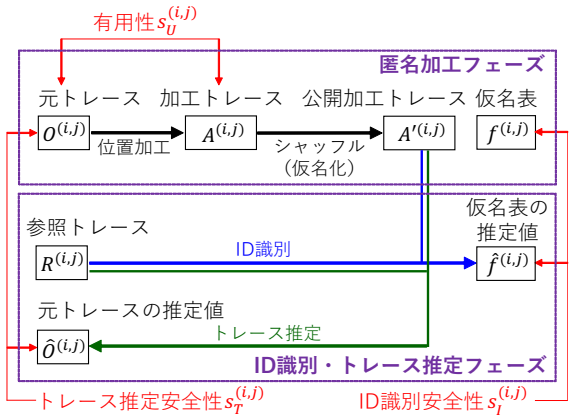


**図 2**　コンテストの流れ（$i \in [z]$，$j \in \{1,2\}$）

$O^{(i,1)} = (o_1^{(i,1)}, o_2^{(i,1)}, o_3^{(i,1)})$ である．

### 2.2　コンテストの流れ

コンテストの流れを図 2 に示す．本コンテストはチーム $P_1, \ldots, P_z$ と審判 $Q$ によって行われる．コンテストは**匿名加工フェーズ，有用性評価フェーズ，ID 識別・トレース推定フェーズ，安全性評価フェーズ**の 4 つのフェーズにより構成され，順に実施される．以下，各フェーズを簡単に説明する．

#### 2.2.1　匿名加工フェーズ

本フェーズでは，位置情報の加工とトレースのシャッフル（仮名化）により匿名加工を行う．

まず，審判 $Q$ が各チーム $P_i$（$i \in [z]$）に対して $j \in \{1,2\}$ 番目の元トレース $O^{(i,j)}$ を配布する．尚，参照トレース $R^{(i,j)}$ については，このフェーズでは配布しない（ID 識別・トレース推定フェーズで配布する）．

次に，チーム $P_i$ が元トレース $O^{(i,j)}$ の各位置情報（計 $nt$ 個）を加工することで，ユーザ ID 順に加工位置情報を並べた加工トレース $A^{(i,j)}$ を作成する．$P_i$ は，加工トレース $A^{(i,j)}$ を $Q$ に提出する．ここで，$P_i$ は 2 つの元トレース $O^{(i,1)}$，$O^{(i,2)}$ の両方を加工しても良いし，いずれか一方のみを加工しても良い．また，一つも加工しなくても，失格にはならない．

加工トレース $A^{(i,j)}$ が提出された後，審判 $Q$ が $A^{(i,j)}$ に対して，$n$ 個のトレースをランダムにシャッフルし，順に $n+1$ から $2n$ までの仮名 ID を付与し，仮名 ID 順に加工位置情報を並べた公開加工トレース $A'^{(i,j)}$ と，元のユーザ ID と仮名 ID の対応表である仮名表 $f^{(i,j)}$ を作成する．尚，公開加工トレース $A'^{(i,j)}$ と参照トレース $R^{(i,j)}$ は ID 識別・トレース推定フェーズにおいて全チームに公開するが，仮名表 $f^{(i,j)}$ は $Q$ だけが参照できるよう秘密にする．

匿名加工（即ち，位置加工とシャッフル）の方法，および加工トレース $A^{(i,j)}$，公開加工トレース $A'^{(i,j)}$，仮名表 $f^{(i,j)}$ の詳細は，第 3.2 節で述べる．

#### 2.2.2　有用性評価フェーズ

審判 $Q$ が，提出された各加工トレース $A^{(i,j)}$ に対して，対応する元トレース $O^{(i,j)}$ を用いて有用性 $s_U^{(i,j)} \in [0,1]$ を算出する．$s_U^{(i,j)}$ は値が大きいほど有用性が高く，小さいほど有用性が低い．有用性指標の詳細は，第 3.4 節で述べる．

審判 $Q$ は，有用性 $s_U^{(i,j)}$ を要求値 $s_{req} \in [0,1]$ と比較し，$s_U^{(i,j)} \geq s_{req}$ の場合には $A^{(i,j)}$ を「有効」と判定し，$s_U^{(i,j)} < s_{req}$ の場合には $A^{(i,j)}$ を「無効」と判定する．有用性の要求値 $s_{req}$ については，$Q$ が予め公開しておく．

尚，有用性 $s_U^{(i,j)}$ は元トレース $O^{(i,j)}$ と加工トレース $A^{(i,j)}$ を用いて計算できるため，各チーム $P_i$ は匿名加工フェーズにおいて，自身の加工トレース $A^{(i,j)}$ が有効かどうか（即ち，有用性 $s_U^{(i,j)}$ が要求値 $s_{req}$ 以上になっているかどうか）のチェックを提出前に行うことができる．有用性算出プログラムも，$Q$ が予め公開しておく．

#### 2.2.3　ID 識別・トレース推定フェーズ

まず，審判 $Q$ が全チームの「有効」な公開加工トレースと参照トレースを公開する（即ち，全チームに配布する）．次に，各チームが他のチームの公開加工トレースに対して，参照トレースを参考にしながら，ID 識別やトレース推定を行う．ここで，各公開加工トレースに対して，ID 識別とトレース推定の両方を行っても良いし，どちらか一方のみを行ってもよい．また，他のチームのトレースを一切攻撃しなくても，失格にはならない．

以下，チーム $P_i$ 以外のチーム $P_h$（$h \neq i$）が，チーム $P_i$ の公開加工トレース $A'^{(i,j)}$ に対して，ID 識別とトレース推定を行う場合について説明する．ID 識別を行う場合は，参照トレース $R^{(i,j)}$ を参考にしながら，$A'^{(i,j)}$ の各仮名 ID に対してユーザ ID の推定値を記した「仮名表の推定値」$\hat{f}^{(i,j)}$ を作成し，$Q$ に提出する．トレース推定を行う場合は，$R^{(i,j)}$ を参考にしながら，$A'^{(i,j)}$ に対応する元トレース $O^{(i,j)}$ における各位置情報（計 $nt$ 個）の推定値を記した「元トレースの推定値」$\hat{O}^{(i,j)}$ を作成し，$Q$ に提出する．チーム $P_h$ が $P_i$ に対して提出できる $\hat{f}^{(i,j)}$ と $\hat{O}^{(i,j)}$ は，それぞれ高々「1 つ」とする（即ち，他の各チームに対して行える ID 識別とトレース推定の回数は，それぞれ最大 1 回）．

仮名表の推定値 $\hat{f}^{(i,j)}$，元トレースの推定値 $\hat{O}^{(i,j)}$ の詳細は，第 3.3 節で述べる．

#### 2.2.4　安全性評価フェーズ

審判 $Q$ が，公開加工トレース $A'^{(i,j)}$ の ID 識別安全性と，トレース推定安全性を評価する．

具体的には，仮名表の推定値 $\hat{f}^{(i,j)}$ と仮名表 $f^{(i,j)}$ を比較することで ID 識別安全性 $s_I^{(i,j)} \in [0,1]$ を算出し，元トレースの推定値 $\hat{O}^{(i,j)}$ と元トレース $O^{(i,j)}$ を比較することでトレース推定安全性 $s_T^{(i,j)} \in [0,1]$ を評価する．$s_I^{(i,j)}$ と $s_T^{(i,j)}$ はどちらも値が大きいほど安全性が高く，小さいほど安全性が低い．ID 識別安全性とトレース推定安全性の指標については，第 3.5 節で詳述する．

各公開加工トレース $A'^{(i,j)}$ は，他のチームによって最大で $z-1$ 回の ID 識別攻撃と，最大で $z-1$ 回のトレース推定攻撃を受ける（ID 識別対策用の公開加工トレース $A'^{(i,1)}$ もトレース推定対策用の公開加工トレース $A'^{(i,2)}$ も，ID 識別とトレース推定の両攻撃を受けることに注意する）．

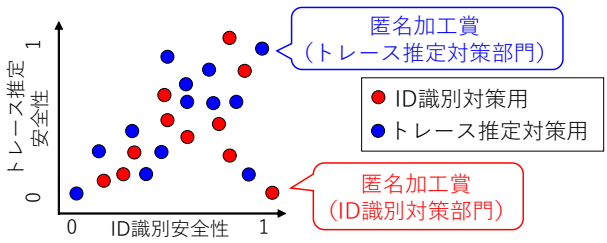


**図 3**　2 つの匿名加工賞の例（赤丸と青丸は，それぞれ ID 識別対策用，トレース推定対策用の公開加工トレースを表す）

また，サンプルプログラム [12] を用いた ID 識別攻撃とトレース推定攻撃も受ける．これらの各々について，ID 識別安全性とトレース推定安全性を算出し，ID 識別安全性の最小値 $s_{I,min}^{(i,j)} \in [0,1]$ と，トレース推定安全性の最小値 $s_{T,min}^{(i,j)} \in [0,1]$ を求める．$s_{I,min}^{(i,j)}$ と $s_{T,min}^{(i,j)}$ を，それぞれ公開加工トレース $A'^{(i,j)}$ の最終的な ID 識別安全性，トレース推定安全性とする（即ち，$A'^{(i,j)}$ にとって最も強力な攻撃を受けたときの安全性を採用する）．

### 2.3　賞

本コンテストでは，以下の賞を設ける．尚，ID 識別安全性とトレース推定安全性は，それぞれ予備戦の値と本戦の値を 1:9 の割合で合計した上で，受賞チームを選定する．

**総合優勝・総合 2 位・総合 3 位**：ID 識別対策とトレース推定対策の両方で優れた成績を修めたチームに賞を与える．具体的には，ID 識別対策用の公開加工トレース $A'^{(i,1)}$ に対する ID 識別安全性 $s_{I,min}^{(i,1)}$ と，トレース推定対策用の公開加工トレース $A'^{(i,2)}$ に対するトレース推定安全性 $s_{T,min}^{(i,2)}$ の総和 $s_{I,min}^{(i,1)} + s_{T,min}^{(i,2)}$ が最も大きい上位 3 チームに賞を与える．尚，無効な加工トレースや未提出の加工トレースに対応する安全性は 0 と見なす．

**匿名加工賞（ID 識別対策部門）**：ID 識別対策用の公開加工トレース $A'^{(i,1)}$ に対する ID 識別安全性 $s_{I,min}^{(i,1)}$ が最も大きい 1 チームに賞を与える．

**匿名加工賞（トレース推定対策部門）**：トレース推定対策用の公開加工トレース $A'^{(i,2)}$ に対するトレース推定安全性 $s_{T,min}^{(i,2)}$ が最も大きい 1 チームに賞を与える．

ID 識別対策部門とトレース推定対策部門の 2 つの匿名加工賞の例を図 3 に示す．

**リスク評価賞（ID 識別部門）**：ID 識別安全性を最も下げた 1 チームに賞を与える（より詳細は，競技ルール [13] に記載）．

**リスク評価賞（トレース推定部門）**：トレース推定安全性を最も下げた 1 チームに賞を与える（より詳細は，競技ルール [13] に記載）．

**プレゼンテーション賞**：各チーム $P_i$ は，どのような匿名加工，ID 識別，トレース推定を行ったかを明確にするため，CSS2019 の会場において，匿名加工，ID 識別，トレース推定のアルゴリズムの概要を発表する．その後，優れた発表をした 1 チームに賞を与える（より詳細は，競技ルール [13] に記載）．

上記の賞のうち，総合優勝・総合 2 位・総合 3 位，および 2 つの匿名加工賞は，ID 識別対策とトレース推定対策のそれぞれで，優れた加工データを集めるためである．また，各チームにおける 1 番目，2 番目のデータセットをそれぞれ ID 識別対策用，トレース推定対策用とし，各部門の匿名加工賞を対応するデータセットの中から選定しているのは，各匿名加工に対して「ID 識別対策として行ったのか」，「トレース推定対策として行ったのか」，その意図を明確にさせるためである．

一方，リスク評価では，ID 識別対策用とトレース推定対策用を合わせた全公開加工トレースに対して ID 識別とトレース推定の両攻撃を行い，ID 識別安全性とトレース推定安全性を最も下げることに成功したチームにそれぞれ賞を与える．これは，各公開加工トレースに対して（図 3 に示されるように）ID 識別とトレース推定の 2 軸で評価し，両者に対する安全性の相関関係を明確にするためである．

但し，各チームへの負担が大きくならないように，加工するデータセットは 1 つでも（或いはなくても）良く，ID 識別とトレース推定のどちらか一方のみを行っても（或いは一切攻撃しなくても）良いようにしている．各チームは自身の都合に合わせて，狙うべき賞を絞る（例えば，リスク評価賞に絞る），といった戦略を練ることができる．

## 3.　コンテストの詳細

### 3.1　データセット

本コンテストではデータセットとして，オープンな公開データセットである疑似人流データ [14] を基に新たに作成した人工データ（以後，PWSCup2019 用人工データ）を使用する．

疑似人流データは，実データを基に作成した 6 日間（2013 年の 7/1, 7/7, 10/7, 10/13, 12/16, 12/22）にわたる東京近郊（首都圏）の人工的なトレースの公開データセットである（尚，容易に照合する情報は公開されていない）．本コンテストでは，東京中心部（緯度：35.65〜35.75，経度：139.68〜139.8）に対して，図 4 のように均等に $32 \times 32 = 1024$ 個の領域に分割し，左下から右上にかけて領域 ID を割り当てる．各領域の大きさは，緯度 1 度あたり 111km，（東京での）経度 1 度あたり 91km として，縦 347m × 横 341m である．位置情報（領域）の数は $m = 1024$ である．

その後，東京中心部における位置情報が 10 個以上あるユーザ（計 10181 名）のトレースを抽出し，これらを学習データとして，マルコフモデルに基づく生成モデルを学習する．その生成モデルから，チーム番号 $i \in [z]$ およびデータセット番号 $j \in \{1,2\}$ 毎に異なる $n = 2000$ 名のユーザ集合 $\mathcal{U}^{(i,j)} = \{u_1^{(i,j)}, \ldots, u_{2000}^{(i,j)}\}$ の参照トレース $R^{(i,j)} = (r_1^{(i,j)}, \ldots, r_{2000}^{(i,j)})$ と元トレース $O^{(i,j)} = (o_1^{(i,j)}, \ldots, o_{2000}^{(i,j)})$ を生成する．尚，生成モデルは各ユーザ $u_k^{(i,j)}$（$1 \le k \le 2000$）の特徴量を保持しており，それを基に参照トレース $r_k^{(i,j)}$ と元トレース $o_k^{(i,j)}$ を生成す

表 1　PWSCup2019 用人工データ

| | |
|---|---|
| ユーザ数 | $n = 2000$ |
| 対象エリア | 緯度：35.65〜35.75，経度：139.68〜139.8 |
| 位置情報数 | $m = 1024$（$32 \times 32$ 個の領域に分割） |
| トレースの長さ | 予備戦では $t = 40$（8 : 00〜17 : 59 の 2 日分，30 分おき）．本戦では日数変更の可能性あり． |

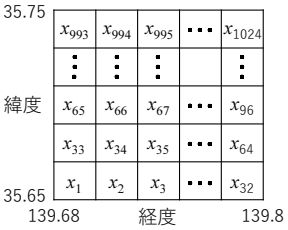


図 4　本コンテストにおける位置情報

る．従って，$r_k^{(i,j)}$ と $o_k^{(i,j)}$ は高い相関を持っており，参照トレースを参考にしながら公開加工トレースに対して ID 識別やトレース推定を行うことが可能である．

時刻については，8 時から 17 時 59 分までを 30 分おきに区切って離散化する．予備戦では，1 日目と 2 日目のトレースを参照トレースに，3 日目と 4 日目のトレースを元トレースとして用いる．即ち，各トレースの長さは $t = 40$ である（時刻 1 は 1 日目の 8 時，時刻 40 は 2 日目の 17 時 30 分，時刻 41 は 3 日目の 8 時，時刻 80 は 4 日目の 17 時 30 分）．但し，本戦では参照トレースと元トレースの日数を（2 日分から）変更する可能性がある．PWSCup2019 用人工データの概要を表 1 に示す．

尚，生成モデルの詳細は非公開とするが，PWSCup2019 用人工データの生成法は以下のような特徴を持っている．

(1) **人口分布の保存**：6 時台から 17 時台までの 1 時間毎の人口分布（$m = 1024$ 個の領域にわたる確率分布）が，元の疑似人流データのそれと近くなるように，人工データを生成する．

(2) **遷移行列の保存**：マルコフモデルの遷移行列（$1024 \times 1024$ の行列）が，元の疑似人流データのそれと近くなるように，人工データを生成する．

(3) **家のモデル化**：各ユーザは朝に高い確率で（6-7 時台は約 95%，8 時台は約 30% の確率で）自身の家の領域にいるように，人工データを生成する（家は，人口分布が保存されるようにしつつ，ユーザ毎にランダムに割り当てる）．但し，6-7 時台の位置情報まで参照・元トレースに含めると，攻撃者はほぼ全ユーザの家の領域を知ることになり，最大知識モデルのように仮定が強すぎるため，8 時以降の位置情報のみを用いる．

また，ユーザ数 $n$ を増やすことで，チーム毎，データセット毎の有用性・安全性のばらつきを少なくすることができる（即ち，公平性の問題を緩和できる）が，$n$ を大きくしすぎると，匿名加工・ID 識別・トレース推定の計算量が大きくなり，各チームの負担が大きくなる．公平性と各チームへの負担のバランスを考慮し，ユーザ数 $n$ は $n = 2000$ と設定する．PWSCup2019 用人工データに関するより詳細な情報については，文献 [15] を参照されたい．

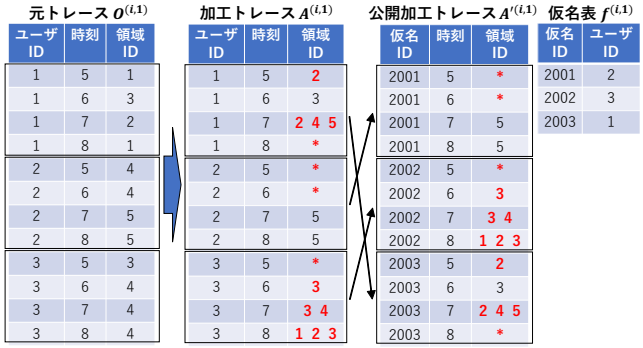


図 5　匿名加工の例（元トレース $O^{(i,1)}$ は図 1 のものと同じ）．ここでは各トレースをユーザ ID（あるいは仮名 ID），時刻，領域 ID の表形式で表現している．また，一般化，削除はそれぞれ領域 ID のリスト（空白区切り），アスタリスク（*）で表している．例えば，「2 4 5」は $\{x_2, x_4, x_5\}$ を意味する．

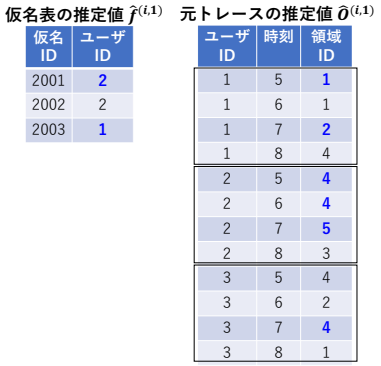


図 6　図 5 の公開加工トレース $A'^{(i,1)}$ に対する ID 識別・トレース推定結果の例（青字：元データと完全に一致しているユーザ ID ／領域 ID）．

### 3.2　匿名加工

匿名加工は，位置情報の加工とシャッフル（仮名化）によって行われる．匿名加工の例を図 5 に示す．

**位置情報の加工**：まず，各チーム $P_i$ が元トレース $O^{(i,j)}$ における各位置情報（計 $nt$ 個）を加工する．本コンテストでは，各位置情報の加工の方法としては「加工なし」，「ノイズ付与」，「一般化」，「削除」の 4 種類を考える．

(1) **加工なし**：元の位置情報を加工せずにそのまま出力する．例えば，$x_1 \to x_1$ とするものである．

(2) **ノイズ付与**：元の位置情報を別の位置情報（集合 $\mathcal{X}$ の別の要素）に変換する．例えば，$x_1 \to x_3$ と変換する．

(3) **一般化**：元の位置情報を 2 つ以上の位置情報からなる集合に変換する．元の位置情報は含めても含めなくてもよい．例えば，$x_1 \to \{x_1, x_3\}$，あるいは $x_1 \to \{x_2, x_3, x_5\}$ と変換する．前者は元の位置情報を含む一般化，後者は元の位置情報を含まない一般化である．

(4) **削除**：元の位置情報を空集合 $\emptyset$ に変換する．例えば，$x_1 \to \emptyset$ と変換する．

以上より，加工後の位置情報の取り得る値の集合を $\mathcal{Y}$ とすると，これは $\mathcal{X}$ のべき集合 $\mathcal{Y} = 2^{\mathcal{X}}$ と表すことができる．即ち，本コンテストでは，各位置情報の加工の方法としてあらゆる方法を許容する．

**シャッフル（仮名化）**：位置情報の加工後，審判 $Q$ が加工トレース $A^{(i,j)}$ をシャッフルすることで仮名化する．具体的には，ユーザ ID $1,2,\ldots,n$ をランダムに置換して，順に $n+1$ から $2n$ までの仮名 ID を付与することで仮名化を行う．例えば，図 5 では，仮名 ID $2001, 2002, 2003$ は，それぞれユーザ ID $2,3,1$ に対応している．

**加工トレース $A^{(i,j)}$・公開加工トレース $A'^{(i,j)}$**：$A^{(i,j)}$ におけるユーザ ID $k \in [n]$ の加工トレースを $a_k^{(i,j)} \in \mathcal{Y}^t$ とし，$A'^{(i,j)}$ における仮名 ID $k \in \{n+1,\ldots,2n\}$ の公開加工トレースを $a'^{(i,j)}_k \in \mathcal{Y}^t$ とする．即ち，$A^{(i,j)} = (a_1^{(i,j)},\ldots,a_n^{(i,j)})$，$A'^{(i,j)} = (a'^{(i,j)}_1,\ldots,a'^{(i,j)}_n)$ である．例えば，図 5 では $a_1^{(i,1)} = (x_2, x_3, \{x_2, x_4, x_5\}, \emptyset)$，$A^{(i,1)} = (a_1^{(i,1)}, a_2^{(i,1)}, a_3^{(i,1)})$，$a'^{(i,1)}_{2001} = (\emptyset, \emptyset, x_5, x_5)$，$A'^{(i,1)} = (a'^{(i,1)}_{2001}, a'^{(i,1)}_{2002}, a'^{(i,1)}_{2003})$ である．

**仮名表 $f^{(i,j)}$**：仮名表 $f^{(i,j)}$ は仮名 ID とユーザ ID のペアの集合で表現することにする．例えば，図 5 では $f^{(i,j)} = \{(2001,2),(2002,3),(2003,1)\}$ である．

### 3.3 ID 識別・トレース推定

ID 識別，トレース推定では，公開加工トレース $A'^{(i,j)}$ を基に，参照トレース $R^{(i,j)}$ を参考にしながらそれぞれ仮名表の推定値 $\hat{f}^{(i,j)}$，元トレースの推定値 $\hat{O}^{(i,j)}$ を出力する．ID 識別，トレース推定の例を図 6 に示す．

**仮名表の推定値 $\hat{f}^{(i,j)}$**：$\hat{f}^{(i,j)}$ は各仮名 ID に対してユーザ ID の推定値を記したものであり，仮名表 $f^{(i,j)}$ と同様に仮名 ID とユーザ ID のペアの集合で表現する．但し，仮名表 $f^{(i,j)}$ では各ユーザ ID は 1 回のみ現れるのに対して，仮名表の推定値 $\hat{f}^{(i,j)}$ では同じユーザ ID が 2 回以上現れてもよい（即ち，ユーザ ID の重複があってもよい）．例えば，図 6 では $\hat{f}^{(i,j)} = \{(2001,2),(2002,2),(2003,1)\}$ であり，ユーザ ID「2」が 2 回出現している．

**元トレースの推定値 $\hat{O}^{(i,j)}$**：$\hat{O}^{(i,j)}$ はトレースの推定値をユーザ ID 順に並べたものである．$\hat{O}^{(i,j)}$ におけるユーザ ID $k \in [n]$ のトレースの推定値を $\hat{o}_k^{(i,j)} \in \mathcal{X}^t$ とする．即ち，$\hat{O}^{(i,j)} = (\hat{o}_1^{(i,j)},\ldots,\hat{o}_n^{(i,j)})$ である．例えば，図 6 では $\hat{o}_1^{(i,1)} = (x_1, x_1, x_2, x_4)$，$\hat{O}^{(i,1)} = (\hat{o}_1^{(i,1)},\ldots,\hat{o}_3^{(i,1)})$ である．

### 3.4 有用性指標

有用性 $s_U^{(i,j)} \in [0,1]$ は，元トレース $O^{(i,j)}$ と加工トレース $A^{(i,j)}$ を用いて算出される．本コンテストでは，加工トレースが様々な用途に使われる可能性を考慮し，汎用性の高い有用性指標を導入する．

具体的には，加工トレースは人気スポットの分析 [4]，POI カテゴリーの自動タグ付け [5]，外国人観光客の動態分析 [6] など，様々なデータ分析に利用可能である．さらには，LBS プロバイダーが位置情報を第三者提供する際に加工するのではなく，ユーザが（LBS プロバイダーを信用せず）自身の位置情報を加工して LBS プロバイダーに送信するモデルも考えられる．例えば，ユーザが加工済み位置情報を LBS プロバイダーに送り，（周辺の飲食店などの）

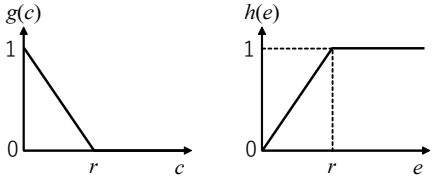

**図 7** 区分線形関数 $g$ と $h$

POI 検索結果を受け取る応用例 [8] が考えられる．

これらの応用例では，位置情報を加工すればするほど有用性が損なわれ，ある一定レベル以上加工すると有用性が完全に損なわれる，と考えられる．例えば，POI 検索において，位置情報にある一定距離以上（例えば 5km 以上）のノイズを加える，あるいは削除する，といった加工を施すと，元の位置情報周辺の POI は全く検索できなくなる．

このことを考慮して，有用性の指標を導入する．まず，$d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_{\geq 0}$ を，2 つの位置情報 $x_k, x_l \in \mathcal{X}$ を入力として，それらのユークリッド距離 $d(x_k, x_l) \in \mathbb{R}_{\geq 0}$ を出力する関数とする．本コンテストでは，$x_k, x_l$ は図 4 のような領域であるため，$d(x_k, x_l)$ を $x_k$ の中心点と $x_l$ の中心点のユークリッド距離とする．各領域の大きさは縦 347m × 横 341m であるため，例えば $d(x_1, x_1) = 0$m，$d(x_1, x_2) = 341$m，$d(x_1, x_{33}) = 347$m である．

次に，元トレース $O^{(i,j)}$ の各位置情報（計 $nt$ 個）と，対応する加工トレース $A^{(i,j)}$ の各加工済み位置情報（計 $nt$ 個）の「ユークリッド距離の平均」を計算する（削除に対しては，$r \in \mathbb{R}_{\geq 0}$ とする）．ユーザ ID $k \in [n]$ における $l \in [t]$ 番目の元の位置情報と加工済み位置情報のユークリッド距離の平均を $c_{k,l} \in \mathbb{R}_{\geq 0}$ とする．例えば，図 5 では $c_{1,1} = d(x_1, x_2)$，$c_{1,2} = d(x_3, x_3)$，$c_{1,3} = \frac{d(x_2,x_2)+d(x_2,x_4)+d(x_2,x_5)}{3}$，$c_{1,4} = r$ である．その後，各 $c_{k,l}$ に対して，図 7 の左に示される区分線形関数を用いて 0 から 1 までのスコア値に変換し（削除に対しては，スコア値は 0），$nt$ 個のスコア値の平均を有用性 $s_U^{(i,j)}$ とする．

即ち，$g : \mathbb{R}_{\geq 0} \to [0,1]$ を，$c \in \mathbb{R}_{\geq 0}$ を入力として，

$$g(c) = \begin{cases} 1 - \frac{c}{r} & (x < r \text{ のとき}) \\ 0 & (x \geq r \text{ のとき}) \end{cases} \tag{1}$$

で表されるスコア値 $g(c)$ を出力する関数とし，

$$s_U^{(i,j)} = \frac{1}{nt} \sum_{k=1}^{n} \sum_{l=1}^{t} g(c_{k,l}) \tag{2}$$

と計算することで，有用性 $s_U^{(i,j)}$ を求める．

区分線形関数 $g$ は，各位置情報に対してユークリッド距離の平均が $r$ 以上になるように加工する（あるいは削除する）とスコア値が 0 になる（即ち，有用性が完全に損なわれる）ように設計されている．本コンテストでは，$r$ の値としては $r = 2$km と設定する．このときの有用性 $s_U^{(i,j)}$ が前述した POI 検索の精度と相関が高いことを（相関係数が 0.9 以上），実験的に確認している（詳細は割愛）．

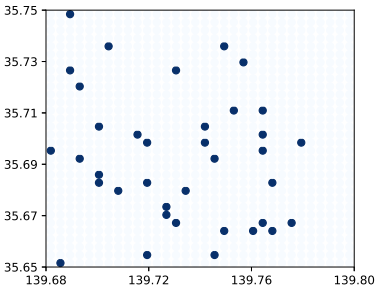


**図 8**　病院カテゴリーの POI を含む領域（青丸，計 37 個）

### 3.5　安全性指標

**ID 識別安全性**：ID 識別安全性 $s_I^{(i,j)} \in [0,1]$ は，仮名表の推定値 $\hat{f}^{(i,j)}$ と仮名表 $f^{(i,j)}$ を比較することで算出される．

本コンテストでは，$s_I^{(i,j)}$ を ID 識別率を 1 から引いたものとして定義する．即ち，$\hat{f}^{(i,j)}$ を $f^{(i,j)}$ と比較することで求まる ID 識別率を $\alpha^{(i,j)} \in [0,1]$ としたときに，

$$s_I^{(i,j)} = 1 - \alpha^{(i,j)} \tag{3}$$

と計算する．図 5 と図 6 の例では，$\alpha^{(i,j)} = \frac{2}{3}$ であり，$s_I^{(i,j)} = 1 - \frac{2}{3} = \frac{1}{3}$ である．

**トレース推定安全性**：トレース推定安全性 $s_T^{(i,j)} \in [0,1]$ は，元トレース $O^{(i,j)}$ と元トレースの推定値 $\hat{O}^{(i,j)}$ をを比較することで算出される．

本コンテストでは，元トレース $O^{(i,j)}$ の各位置情報（計 $nt$ 個）と，対応する元トレースの推定値 $\hat{O}^{(i,j)}$ の各位置情報（計 $nt$ 個）のユークリッド距離を求め，それを基にしたトレース推定安全性の指標を導入する．

ユーザ ID $k \in [n]$ における $l \in [t]$ 番目の元の位置情報とその推定値のユークリッド距離を $e_{k,l} \in \mathbb{R}_{\geq 0}$ とする．例えば，図 5 と図 6 では $e_{1,1} = d(x_1, x_1)$，$e_{1,2} = d(x_3, x_1)$，$e_{1,3} = d(x_2, x_2)$，$e_{1,4} = d(x_1, x_4)$ である．

ここで，元の位置情報とその推定値のユークリッド距離が大きいほど安全性が高くなり，ある一定値以上になれば（例えば，5km 以上）推定は完全に失敗したものと考えることができる．このことを考慮して，図 7 の右に示される区分線形関数 $h$ を用いて，各ユークリッド距離 $e_{k,l}$ を 0 から 1 までのスコア値に変換する．

また，本コンテストでは，トレースの中には通院している病院など，特にセンシティブな位置情報が含まれる可能性があることを考慮する．具体的には，$nt$ 個のスコア値を求めた後，病院の POI を含む領域（以後，病院領域）に対する重みが 10 倍になるように重み付け平均をとったものを，トレース推定安全性 $s_T^{(i,j)}$ とする．疑似人流データ [14] の POI データから「病院」カテゴリーの POI を含む領域を抽出したものを図 8 に示す．この青丸で示されている計 37 個の領域を，全て病院領域とする．

以下，トレース推定安全性 $s_T^{(i,j)}$ の算出方法を定式化する．$h : \mathbb{R}_{\geq 0} \to [0,1]$ を，$e \in \mathbb{R}_{\geq 0}$ を入力として，

$$h(e) = \begin{cases} \frac{e}{r} & (x < r \text{ のとき}) \\ 1 & (x \geq r \text{ のとき}) \end{cases} \tag{4}$$

で表されるスコア値 $h(e)$ を出力する関数とする．また，ユーザ ID $k \in [n]$ における $l \in [t]$ 番目の元トレース $O^{(i,j)}$ の位置情報が病院領域であるときに 10，それ以外の領域であるときに 1 をとる重み変数を $w_{k,l} \in \{1, 10\}$ とする．これらを用いて，

$$s_T^{(i,j)} = \frac{\sum_{k=1}^{n} \sum_{l=1}^{t} w_{k,l} h(e_{k,l})}{\sum_{k=1}^{n} \sum_{l=1}^{t} w_{k,l}} \tag{5}$$

と計算することで，トレース推定安全性 $s_T^{(i,j)}$ を求める．

## 4.　まとめ

本稿では，位置情報の匿名加工を対象としたコンテスト PWS Cup 2019 の内容を説明した．

**謝辞**　本研究は JSPS 科研費 18H04099，19H04113 の助成を受けたものである．

### 参考文献

[1] C. C. Aggarwal, P. S. Yu, Privacy-Preserving Data Mining, Springer, 2008. Springer

[2] 個人情報保護委員会，個人情報の保護に関する法律についてのガイドライン（匿名加工情報編）：https://www.ppc.go.jp/files/pdf/guidelines04.pdf.

[3] Article 29 Data Protection Working Party, “Opinion 05/2014 on Anonymisation Techniques,” WP 216, 2014.

[4] Y. Zheng *et al.*, “Mining interesting locations and travel sequences from GPS trajectories,” Proc. WWW’09, pp.791–800, 2009.

[5] M. Ye *et al.*, “On the Semantic Annotation of Places in Location-Based Social Networks,” Proc. KDD’11, pp.520–528, 2011.

[6] 平成 28 年度北海道外国人観光客再訪促進事業（北海道 LOVERS 拡大促進事業）調査報告書：https://www.visit-hokkaido.jp/company/material/detail/44

[7] S. Gambs *et al.*, “De-anonymization attack on geolocated data,” Journal of Computer and System Sciences, vol.80, no.8, pp.1597-1614, 2014.

[8] M. E. Andrés *et al.*, “Geo-Indistinguishability: Differential Privacy for Location-based Systems,” Proc. CCS’13, pp.901–914, 2013.

[9] Y.-A. Montjoye *et al.*, “Unique in the Crowd: The privacy bounds of human mobility,” Scientific Reports, vol.3, no.1376, pp.1–5, 2013.

[10] R. Shokri *et al.*, “Quantifying location privacy,” Proc. IEEE S&P’11, pp.247–262, 2011.

[11] A. Machanavajjhala *et al.*, “L-diversity: privacy beyond k-anonymity,” Proc. ICDE’06, pp.24–35, 2006.

[12] PWS Cup 2019 サンプルプログラムについて：https://www.iwsec.org/pws/2019/cup19-sample.pdf

[13] PWS Cup 2019 競技ルール：https://www.iwsec.org/pws/2019/cup19-rule.pdf

[14] ナイトレイ，東京大学空間情報科学研究センター (CSIS)，疑似人流データ: https://nightley.jp/archives/1954/

[15] PWS Cup 2019 データセットについて：https://www.iwsec.org/pws/2019/cup19-dataset.pdf