PWS Cup 2019 サンプルプログラムについて

PWS Cup 実行委員会

2019 年 8 月 16 日（2019 年 9 月 19 日更新）

1　はじめに

本資料では PWSCup2019 の匿名加工，ID 識別，トレース推定のサンプルプログラムを簡単に説明する．尚，本資料における記号は，次の文献にて与えられる．

- CSS2019 に投稿された PWS Cup 2019 のルール論文 [1].

2　PWSCup2019 サンプルプログラム

2.1　匿名加工サンプルプログラム

2.1.1　A1-none.py

元トレースをそのまま出力する．即ち，位置情報の加工は何もしない（ただの仮名化に相当）．以後，このアルゴリズムを A1-none と呼ぶ．

2.1.2　A2-MRLH.py

MRLH（Merging Regions and Location Hiding 或いは Precision Reducing and Location Hiding）[2] に基づいて，元トレースに対して領域統合と位置情報の削除を行う．具体的には，各位置情報（計 nt 個）に対し，x 軸，y 軸方向の ID（バイナリ系列）のそれぞれ下位 μ_x，μ_y-bit を落とすことで一般化し，確率 λ で削除する．以後，このアルゴリズムを A2-MRLH(μ_x, μ_y, λ) と呼ぶ．A2-MRLH$(1, 1, 0.5)$ の例を図 1 に示す．

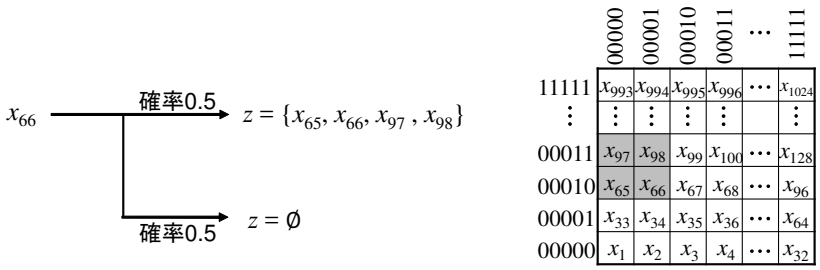

図 1　A2-MRLH$(1, 1, 0.5)$ の例

2.1.3 A3-kRR.py

k-RR（k-ary Randomized Response）[3] に基づいて，元トレースに対してノイズ付与を行う．具体的には，各位置情報（計 nt 個）に対して，確率 $\frac{e^\epsilon}{m-1+e^\epsilon}$ で元の領域 ID をそのまま出力し，残りの確率で他の領域 ID をランダムに出力する．このアルゴリズムは，各位置情報に対して ϵ-differential privacy を保証する（詳細は文献 [3] を参照）．以後，このアルゴリズムを A3-kRR(ϵ) と呼ぶ．

2.1.4 A4-PL.py

PL（Planar Laplace）メカニズム [4] に基づいて，元トレースに対してノイズ付与を行う．具体的には，各位置情報（計 nt 個）に対して，2 次元ラプラス分布に従うノイズを付与して領域 ID に離散化する．このアルゴリズムは，半径 r km 以内において l-differential privacy を保証する ϵ-geo-indistiguishability（但し，$\epsilon = l/r$）を満たす（詳細は文献 [4] を参照）．ϵ が小さいほど，ラプラス分布の分散が大きくなる．以後，このアルゴリズムを A4-PL(l, r) と呼ぶ．

2.1.5 A5-YA.py

山岡匿名化（シャッフル匿名化）．全ユーザ（計 $n = 2000$ 人）のうち，最初の割合 p（$0 \leq p \leq 1$）のユーザ（例えば，$p = 0.1$ のときはユーザ ID が 1 から 200 のユーザ）を部分集合として選び，その部分集合の中でランダムにトレースをシャッフルする．p が小さいほど有用性が大きく，$p = 0$ のときは A1-none.py と等価である．以後，このアルゴリズムを A5-YA(p) と呼ぶ．

2.2 ID 識別サンプルプログラム

2.2.1 I1-rand.py

非復元ランダムに ID 識別を行う．即ち，1, 2, ..., n の permutation をユーザ ID の推定値として出力する．

2.2.2 I2-VisitProb.py

領域滞在分布に基づいて ID 識別を行う．具体的には，まず参照トレースからユーザ毎に，各領域に滞在する確率で構成される「領域滞在分布」を最尤推定により学習する．この際，確率 $= 0$ となっているところについては，小さい正の値（$= 10^{-8}$）を代入する（これは後述する尤度が 0 ならないようにするため）．次に，公開加工トレースに対して，各ユーザに対する尤度（計 t 個の各位置情報に対する尤度の積）を計算し，尤度の最も大きなユーザ ID として識別する．この際，一般化に対しては尤度の平均をとり，削除に対しては尤度を更新しない．このアルゴリズムに基づいて ID 識別を行う様子を図 2 に示す．

尚，本アルゴリズムは高速化のため，公開加工トレースの各位置情報（計 t 個）に対して確率 10% で尤度を更新する（確率 90% で尤度を更新しない）ようにし，10 個以上の領域からなる一般化に対してはランダムに 10 個領域を選ぶようにしている．

2.2.3 I3-HomeProb.py

ユーザは 8 時台に自身の家の領域にいる確率が高いことを利用して ID 識別を行う（8 時台の位置情報のみを用いるように，I2-VisitProb.py を修正したもの）．具体的には，まず 8 時台の位置情報のみを用いて領域滞在分布を学習する．次に，公開加工トレースにおける 8 時台の位置情報に対してのみ尤度を計算し，尤度の最

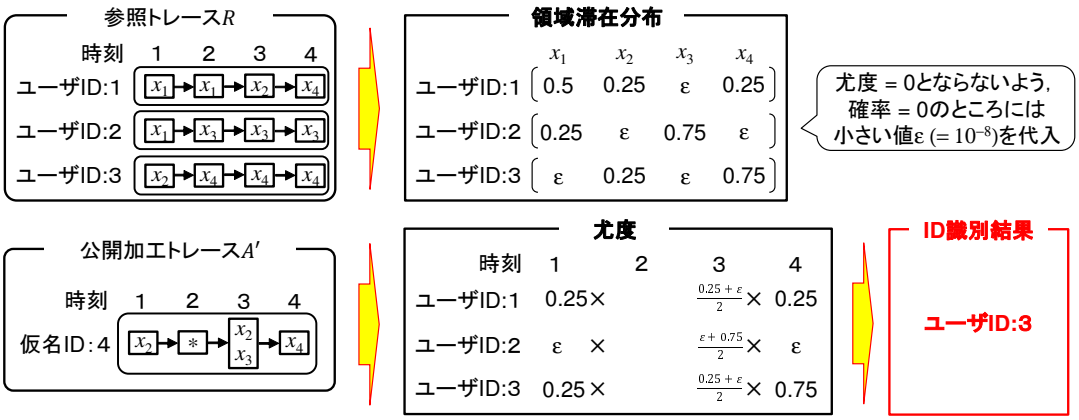

図 2　I2-VisitProb.py に基づく ID 識別の例

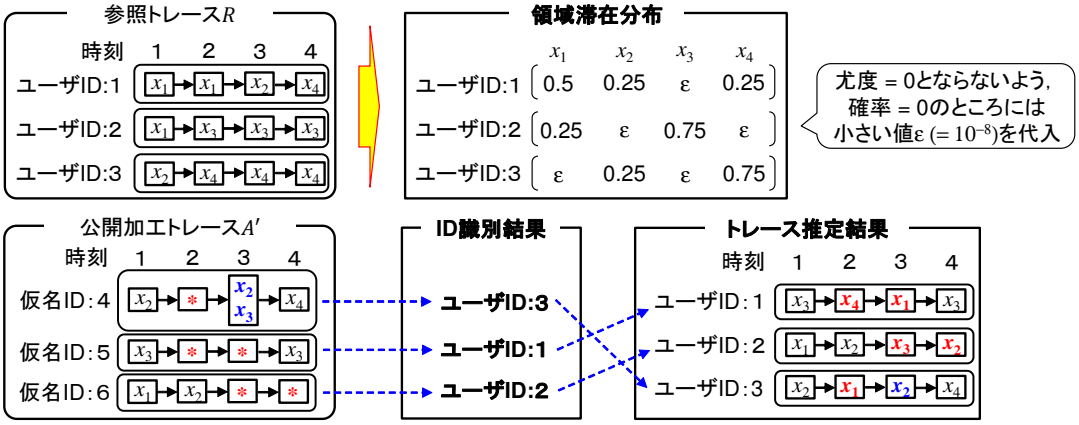

図 3　T2-VisitProb.py に基づくトレース推定の例

も大きなユーザ ID として識別する．

2.3　トレース推定サンプルプログラム

2.3.1　T1-rand.py

ランダムにトレース推定を行う．即ち，各位置情報（計 nt 個）に対して 1〜m をランダムに選んで領域 ID の推定値として出力する．

2.3.2　T2-VisitProb.py

領域滞在分布に基づく ID 識別を行った後，トレース推定を行う．まず，I1-VisitProb.py を用いて，仮名 ID が 1, ..., n の順に ID 識別を行う（但し，識別したユーザ ID は以後選ばないようにすることで，ユーザ ID の重複がないようにする）．次に，ID 識別を行った各トレースに対して，元の位置情報を推定する．具体的には，ノイズに対しては位置情報をそのまま出力し，一般化に対しては一般化された領域の中からランダムに位置情報を出力し，削除に対しては全領域の中からランダムに位置情報を出力する．このアルゴリズムに基づいて ID 識別を行う様子を図 3 に示す．

2.3.3 T3-HomeProb.py

ID 識別の際に 8 時台の位置情報のみを用いるように（即ち，I3-HomeProb.py を用いるように）T2-VisitProb.py を修正したもの．

3 サンプルプログラムを用いた評価実験（予備戦用）

3.1 実験条件

疑似人流データ [5] から大阪（緯度：34.65〜354.74，経度：135.44〜135.56）に対して，（東京と同じように）32×32 の領域に分割して生成モデルを学習し，2 チーム分の参照トレース $R^{(i,j)}$，元トレース $O^{(i,j)}$（どちらも 8:00〜17:59 の 2 日分）を生成した（$1 \leq i \leq 2$，$1 \leq j \leq 2$）．このうち，$R^{(1,1)}$ と $O^{(1,1)}$ を用いて，以下のような実験を行った．

まず，元トレース $O^{(1,1)}$ に対して，匿名加工サンプルプログラムを用いて匿名加工を行った．このとき用いた匿名加工アルゴリズムは，以下のとおりである．

- A1-none.
- A2-MRLH(μ_x, μ_y, λ)．但し，(μ_x, μ_y, λ) = (0,0,0.1), (0,0,0.2), (0,0,0.5), (0,0,0.8), (1,1,0), (1,1,0.1), (1,1,0.2), (1,1,0.5), or (1,1,0.8).
- A3-kRR(ϵ)．但し，ϵ = 0.1, 1, 2, 4, 6, 8, 10, 12, or 14.
- A4-PL(l,r)．但し，(l, r) = (1,1), (2,1), (3,1), (4,1), (5,1), (6,1), or (7,1).
- A5-YA(p)．但し，p = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, or 1.

これらによって得られた各公開加工トレースに対して，有用性 s_U を求めた．

その後，3 つの ID 識別サンプルプログラム（I1-rand.py，I2-VisitProb.py，I3-HomeProb.py）と，3 つのトレース推定サンプルプログラム（T1-rand.py，T2-VisitProb.py，T3-HomeProb.py）を用いて ID 識別安全性の最小値 $s_{I,min}$ と，トレース推定安全性の最小値 $s_{T,min}$ を求めた．

3.2 実験結果

評価実験の結果を図 4 に示す．図 4(i) は有用性の要求値を $s_{req} \geq 0$ としたとき（即ち，全公開加工トレースを有効としたとき），図 4(ii) は有用性の要求値を $s_{req} \geq 0.7$ としたときの結果である．

ID 識別安全性とトレース推定安全性の間にはある程度の相関が見られるものの，ID 識別に強い匿名加工が必ずしもトレース推定に対して強くないことが分かる．特に，図 4(i) より，A5-YA(1)（即ち，全ユーザ内でランダムにトレースをシャッフルする山岡匿名化）は ID 識別安全性が非常に高いものの，トレース推定安全性に対しては何も加工しない A1-none とほぼ同じであることが分かる．これは，攻撃者が A5-YA(1) に対しては，ID 識別を行わずに（即ち，各匿名加工トレースに対して元のユーザ ID を正しく推定することなく），トレース推定を行っていることを意味している．

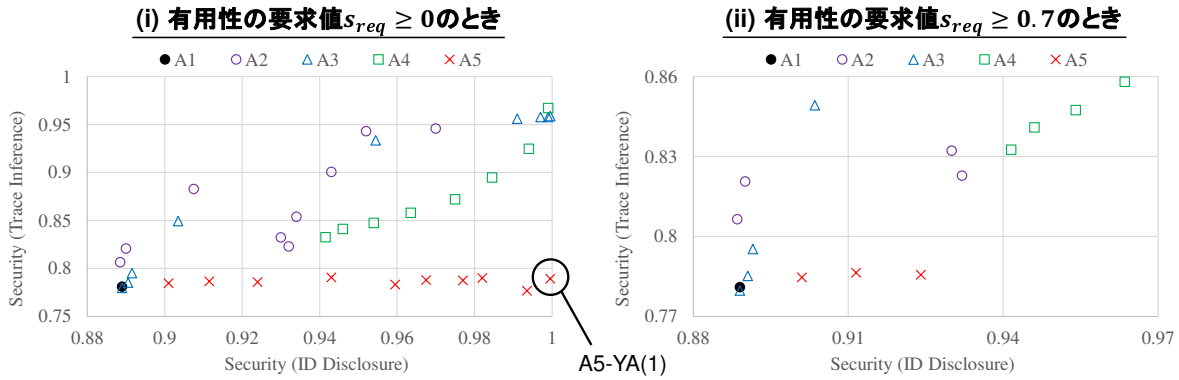

図 4　2 日ずつの参照トレース・元トレースに対してサンプルプログラムを用いた評価実験結果（横軸：ID 識別安全性 $s_{I,min}$，縦軸：トレース推定安全性 $s_{T,min}$）

4　サンプルプログラムを用いた評価実験（本戦用）

4.1　実験条件

3 章の評価実験において，参照トレースと元トレースの長さを 2 日分ずつから様々な長さに変化させたときの ID 識別安全性とトレース推定安全性を調べた．具体的には，まず参照トレースと元トレースの長さを 2，5，10，15，20，25，30 日分ずつと変化させ，匿名加工サンプルプログラムとして A1-none を用いたときの（即ち，仮名化のみを施したときの）安全性を評価した．次に，参照トレースと元トレースの長さを 20 日分ずつに固定し，匿名加工サンプルプログラムとしては 3 章と同じものを用いた場合の安全性を評価した．

尚，ここでの参照トレース・元トレースは，（平日 5 日の後に土日が続くような）連続した日々から構成されたものではなく，散発的にサンプリングされた日々から構成されたものであることに注意されたい．

4.2　実験結果

参照トレースと元トレースの長さを 2 日分ずつから 30 日分ずつまで変化させたときの A1-none（仮名化トレース）の安全性を図 5 に示す．ここでは，3 つの ID 識別サンプルプログラム（I1-rand，I2-VisitProb，I3-HomeProb）と 3 つのトレース推定サンプルプログラム（T1-rand，T2-VisitProb，T3-HomeProb）のぞれぞれに対する安全性を示している．

図 5 より，トレースが短いときは ID 識別安全性とトレース推定安全性が高いが，トレースが長くなると ID 識別安全性とトレース推定安全性が大幅に下がることが分かる．また，トレースが 15 日ずつ以上のときに，I2/T2（トレース全体にわたる領域滞在分布を用いた ID 識別／トレース推定）が I3/T3（8 時台の領域滞在分布を用いた ID 識別／トレース推定）より安全性を下げることに成功している．これは，同一ユーザ内でも日ごとのばらつきが大きいものの，トレースが長いとユーザごとのパターンが大分明確になり，ユーザごとのパターンを（トレース全体にわたる）領域滞在分布として捉えることができるようになったため，と考えている．尚，I2，I3，T2，T3 では領域滞在分布を用いているが，ユーザごとのパターンをより的確に捉えることのできる特徴量が存在する可能性もあり，検討の余地が残されている．

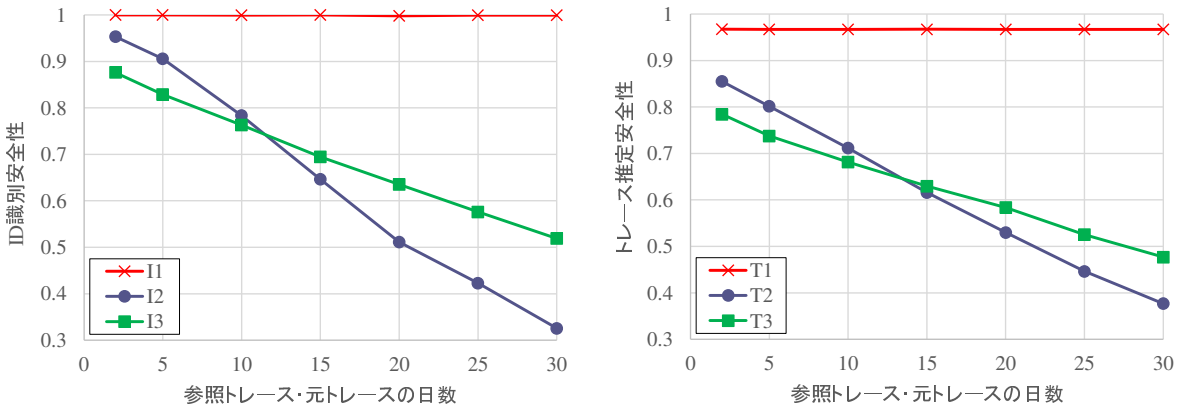

図 5　参照トレース・元トレースの長さを変化させたときの A1-none（仮名化トレース）の ID 識別安全性とトレース推定安全性

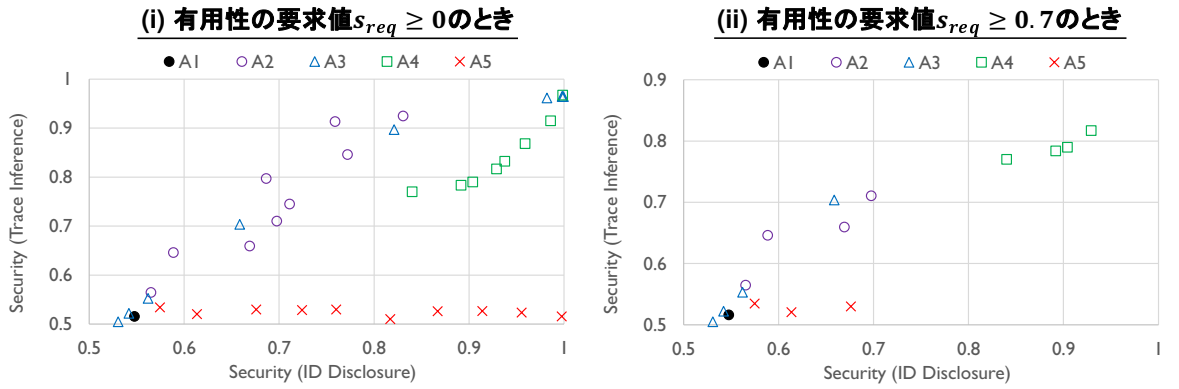

図 6　20 日ずつの参照トレース・元トレースに対してサンプルプログラムを用いた評価実験結果（横軸：ID 識別安全性 $s_{I,min}$，縦軸：トレース推定安全性 $s_{T,min}$）

次に，20 日ずつの参照トレース・元トレースに対して，3 章と同じ匿名加工サンプルプログラムを用いた場合の安全性を図 6 に示す．ここでは，3 つの ID 識別サンプルプログラム（I1-rand，I2-VisitProb，I3-HomeProb）と 3 つのトレース推定サンプルプログラム（T1-rand，T2-VisitProb，T3-HomeProb）を用いたときの ID 識別安全性の最小値 $s_{I,min}$ と，トレース推定安全性の最小値 $s_{T,min}$ を示している．

図 6 より，A4-PL（2 次元ラプラスメカニズム）がサンプルの ID 識別／トレース推定に対して一番高い安全性を実現していることが分かる．尚，A1-none（仮名化トレース）よりやや安全性が低い加工トレースがあるのは，I2-VisitProb が高速化のために，公開加工トレースの各位置情報に対して確率 10% で尤度を更新する（確率 90% で更新しない）ようにしており，そのランダム性による影響と考えている．

参考文献

[1] 村上隆夫，荒井ひろみ，井口誠，小栗秀暢，菊池浩明，黒政敦史，中川裕志，中村優一，西山賢志郎，野島良，波多野卓磨，濱田浩気，山岡裕司，山口高康，山田明，渡辺知恵美，“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”，CSS2019.

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

[3] P. Kairouz *et al.*, “Discrete Distribution Estimation under Local Privacy,” Proc. ICML, pp.2436–2444, 2016.

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

[5] 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/`