サンプリングを用いた際の個人識別リスクの評価

菊池 亮[1]

概要：様々な主体からデータを集約し分析する場合，個々のデータには個人のプライバシ情報が含まれるため，データをそのまま他者に渡すと個人のプライバシ情報が漏えいするリスクがある．そのため，データをうまく加工することで，プライバシ情報の漏洩リスクを低減する匿名化が研究されている．匿名化のデータ加工方法には一般化やノイズ付加の他にも，公的統計などで用いられるサンプリングがある．しかし，サンプリングにプライバシ保護の効果があることは直観的にわかるものの，どのような攻撃者に対して，どの程度個人識別のリスクを低減できるのかは明らかでない．そこで本論文では，サンプリングを使用した際の個人識別リスクを評価し，適切に制限された背景知識を持つ攻撃者に対して個人識別リスクを低減できること，および任意の背景知識を持つ攻撃者に対しては Post Randomization Method と組み合わせることで個人識別リスクを低減できることを示す．

キーワード：サンプリング，k-匿名性，PRAM，統計的開示制御/制限，差分プライバシ

How much can the sampling reduce the identification risk?

Ryo Kikuchi[1]

Abstract: There have been many studies in data anonymization. k-anonymity and (ϵ, δ)-differential privacy are the most popular measures in the area of data anonymizaion, and several techniques such as suppression, generalization, and noise addition are used to satisfy those measures. However, the sampling, a popular anonymization method in official statistics, have not been paid attention compared to other methods. Although several works estimated the effect of sampling by (ϵ, δ)-differential privacy, it is still unknown that how much the sampling can reduce the identification risk. In the paper, we point out that one of the main reasons of that is strong background knowledge: An adversary may know the universe. We show that the sampling can reduce the identification risk against an adversary with reasonable background knowledge. In addition, we show how to combine the sampling with post-randomization method (PRAM) for an adversary with arbitrary background knowledge.

Keywords: Sampling, k-anonymity, PRAM, Statistical disclosure control/limitation, Background knowledge

1. はじめに

様々な主体からデータを集約し分析する場合を考える．このとき，データには個人のプライバシ情報が含まれるため，データをそのまま他者に渡すと，個人のプライバシ情報が漏えいする危険性がある．そのため，プライバシ情報が漏えいする危険が低減するようにデータを加工する方法として，匿名化が研究されている．

よく知られた匿名化の安全性基準としては，自身と同じ値を持つレコードが少なくとも $k-1$ 個存在するという k-匿名性 [17] やそれに類する基準群 [2,7,11]，及び “隣り合う”2 つのデータベースを利用した場合に，ある出力となる確率の比が一定に抑えられるという差分プライバシ [3–5] がある．

k-匿名性を満たす方法には，主に一般化やレコード削除が知られており，差分プライバシを満たす方法としては，主にラプラス分布に従うノイズの付加が行われている．

[1] NTT セキュアプラットフォーム研究所

一方で，公的統計の分野では，個人のプライバシを保護するための統計的開示抑制/制御と呼ばれるデータ加工基準・方法が研究されてきている．伝統的な方法として，トップ/ボトムコーディングやリコーディングと共に，**サンプリング**と呼ばれるレコード抽出が行われている．これはレコードの公開範囲を一部に限定することでプライバシを保護するもので，例えばセンサスデータの匿名化では英国，カナダ，アメリカ等で 1%サンプリングや 3%サンプリングが行われている．

1.1 サンプリングを用いた際の個人識別リスクの評価

直観的には，サンプリングを行うとその際に抽出されなかった個人の情報は開示されないため，一定のプライバシ保護効果があると考えられる．しかしながら，その効果の定量化は難しい．

統計的開示制御/制限では，サンプリングされた標本に対してモデルを適用し，母集団の分布を推定し母集団の中での一意性を評価することが一般的に行われてきた（例えば [6]）．しかし，母集団一意性は個人識別リスクと非常に関連が深いと考えられるものの，どのような知識を持った攻撃者に対して，個人識別のリスクが具体的にどの程度低減しているかは非自明である．

サンプリングは差分プライバシでおいても研究されてきている．単純なサンプリングのみの場合差分プライバシを満たすことは難しいが，差分プライバシを満たすメカニズムにサンプリングを組み合わせることによって安全性を強化できることや [15]，Post RAmdomization Method (PRAM) との組み合わせ [10]，k-匿名化との組み合わせ [9]，外れ値が無い前提を置く [1] などで差分プライバシを満たせることが知られている．しかしながら，差分プライバシを満たしたとしても，それがどの程度個人識別のリスクを低減できているかは非自明である．例えば，ある ϵ について差分プライバシを満たしたとしても，個人識別のリスクは，それ以外の要素，例えば計算する統計値や母集団の分布に強く依存してしまうことが知られている [8]．

1.2 本論文の成果

本論文は，サンプリングを用いた場合，どのような攻撃者に対してどの程度個人識別リスクを低減できるのかを評価するのを目的とする．本論文では，個人識別のリスクを Pk-匿名性 [7] という指標を基に評価する．この指標は，匿名化されたテーブルの中のどのレコードも，元テーブルのどのレコードであったかを当てられる確率が高々 $1/k$ である，ということを保証する指標である．

本論文では，まずサンプリングの評価が難しい理由を，攻撃者の背景知識の観点から考察する．特に，任意の背景知識を仮定した場合，サンプリングのみでは個人識別リスクの低減が難しいことを指摘する．故に，サンプリングが個人識別リスクを低減し得るのは，**攻撃者の背景知識が制限できる**とき，もしくはサンプリングと**他の方法を組み合わせる**場合である．

次に，適切に背景知識を制限した攻撃者に対してであれば，サンプリングを行うことで個人識別リスクが低減できることを示す．本論文で想定する攻撃者は，元テーブルのうち一部レコードを知っているような攻撃者である．例えば都道府県毎に集めたデータを集約する際，攻撃者は住んでいる都道府県についてのデータは知っているが，他の都道府県のデータは知らないといった状況に相当する．この結果を用いれば，上記のような状況において所望の個人識別リスクを満たすためにはどの程度サンプリングすればよいかが明らかにできる．

さらに，背景知識の制限が難しい場合，サンプリングと PRAM と呼ばれるランダム化を組み合わせることによって，任意の背景知識を持つ攻撃者に対して個人識別のリスクを低減できる手法を提案する．PRAM を用いることで個人の識別リスクが低減できることは知られているが，サンプリングと組み合わせることで，PRAM を単独で用いる場合に比べ PRAM で加えるノイズを小さくすることができ，サンプリングが個人識別リスクの低減に寄与していることを示す．

1.3 関連研究

δ-presence [12,13] は一般化とサンプリングを組み合わせることで，ある個人が δ 以上の確率で匿名化に含まれているかどうかがわからないという指標である．Probabilistic k-anonymity [16] は k-匿名性の確率的な拡張であり，サンプリングも扱っているが，攻撃者から見てどのレコードも等確率に生起することを仮定している．

2. 準備

集合 $\mathbb{A}$ について $|\mathbb{A}|$ は集合の要素数とし，$\lfloor n \rfloor$ は n を超えない最大の整数とする．本論文での匿名化の入出力はテーブルである．すなわち，ある個人に対応するレコードが並んだものである．簡単のため，複数の属性を纏めて 1 種類と見て議論を進める．特に明記しない限り，本論文におけるサンプリングとは，n レコードのテーブルから $\lfloor pn \rfloor$ レコード（ただし $p < 1$）をランダムに抜き出す操作である．

テーブルに対する匿名化を形式的に議論するための定義を与える．

T, τ:匿名化前のテーブルの確率変数及びそのインスタンス
T', τ':匿名化テーブルの確率変数及びそのインスタンス
$\mathcal{R}, \mathcal{R}'$:匿名化前/後のテーブルのレコード集合
$\mathcal{V}, \mathcal{V}'$:匿名化前/後のレコードが取りうる属性値の組み合わせの集合．
Δ, δ:匿名化処理の確率変数及びそのインスタンス
f_X: 確率変数 X に関する確率密度関数

Π, π:$\mathcal{R} \to \mathcal{R}' \cup \{\phi\}$，ただし $|\mathcal{R}'| \leq |\mathcal{R}|$ であるようなサンプリングの確率変数及びそのインスタンスで，ϕ はサンプリングされなかったことを示す特別な記号とする．

ここで τ, τ', δ はそれぞれ $\tau : \mathcal{R} \to \mathcal{V}$, $\tau' : \mathcal{R}' \to \mathcal{V}'$, $\delta : (\mathcal{R} \to \mathcal{V}) \to (\mathcal{R} \to \mathcal{V}')$ であるような関数であり，π, δ, τ, τ' の間には $\delta(\tau) = \tau' \circ \pi$ が成り立つ．すなわち，元テーブルにデータ保護処理を施してレコードのシャッフルを行ったものが匿名化テーブルだということである．

定義 2.1 π を $\pi : \mathcal{R} \to \mathcal{R}'$ であるような写像としたとき，匿名化アルゴリズム δ は以下を満たす．

$$\delta(\tau) = \tau' \circ \pi.$$

2.1 PRAM

PRAM [14] とは Kooiman らによって提案された匿名化手法で，遷移確率行列と呼ばれる行列に基づき，各レコードの値を確率的に書きかえる方法である．例えば遷移確率行列を A とすると，$A_{u,v}$ は元テーブルで属性値が $u \in \mathcal{V}$ であったときに $v \in \mathcal{V}'$ に変化する確率を表しており，この確率に従ってテーブルの値を 1 つずつ書き換えていく．

2.2 *Pk*-匿名性

Pk-匿名性とは，*k*-匿名性をランダム化にも適用できるよう確率空間に拡張したもので，“匿名化テーブルの任意のレコードを，誰のものか $1/k$ 以上の確率で当てることができない” ことを保証する．

まず，攻撃者による攻撃成功の確率を定義する．ある τ', $r \in \mathcal{R}$, $r' \in \mathcal{R}$, $\Delta(T) = T' \circ \Pi$ について，

$$\Pr[\Pi(r) = r' \mid T' = \tau']$$

が攻撃成功の確率である．直観的にこの確率が表しているのは，どのレコードがサンプリングされたのかは知らない攻撃者に対して，元テーブルのレコード r が匿名化されて匿名化テーブル τ' のレコード r' となる確率である．

以上の攻撃成功の確率を用いて，*Pk*-匿名性は以下のように定義される．

定義 2.2 ([7]) 匿名化処理 Δ と匿名化テーブル τ' が，任意の攻撃者 f_T，任意のテーブル τ，任意のテーブル τ に含まれるレコード $r \in \mathcal{R}$，任意の匿名化テーブル τ' に含まれるレコード $r' \in \mathcal{R}$ について，

$$\Pr[\Pi(r) = r' \mid T' = \tau'] \leq \frac{1}{k}$$

であるならば，(Δ, τ') は *Pk*-匿名性を満たすという．

3. 考察：背景知識によるサンプリングの効果の違い

まず，サンプリングの評価が難しい理由について考察する．公的統計においてはサンプリングが主要な方法となっているのに比べ，パーソナルデータの匿名化においてはそれほど主要な匿名化手法となっていない．この一つの理由は，攻撃者の背景知識の設定にあると考えられる．

公的統計においては，元テーブル（母集団）全てを知っているような攻撃者はあまり現実的ではないとされている．例えば攻撃者がある個人のレコードの属性値（の一部）を知っていて，その個人を匿名化テーブルから特定しようとする場合を考える．このとき，もしサンプリングされた後の匿名化テーブルに，攻撃者の知っている属性値を持つレコードが 1 つしか無かったとしても，そのレコードは元テーブルでは複数存在する（母集団一意ではない）可能性がある．そのため，攻撃者から見て，サンプリングしていない場合と比較してその 1 つしか無いレコードが攻撃者が特定しようとしている個人かどうかわかりづらくなっているため，サンプリングが個人識別のリスクを低減していると言える．

一方パーソナルデータの匿名化では，一般に攻撃者は元テーブルの値全て，もしくは元テーブルの準識別子全てを知っていることが想定されている．そのような攻撃者が，例えば元テーブルで 1 つしかない属性値を持つ個人を特定したい場合，その個人がサンプリングされると，確実に匿名化テーブルのうちどのレコードが当該の個人であるか特定できてしまう．このように，背景知識として元テーブルの情報を知っている攻撃者を想定した場合，サンプリングでは個人の特定が可能な匿名化テーブルを出力してしまうため，パーソナルデータの匿名化の設定ではあまりサンプリングが用いられていないと考えられる．

以上の考察を纏めると，サンプリングが個人識別リスクを低減できるのは，攻撃者から見て，サンプリングする元テーブルに不確実性がある場合と考えられる．言い換えれば，サンプリングする元テーブルに不確実性があれば，匿名化テーブルにある属性値を持つレコードが 1 つしか無かったとしても，元テーブルには複数あった可能性があり，個人識別は難しくなる．このような不確実性が生じる状況は，以下の 2 つが考えられる．

(1) 攻撃者の元テーブルに関する背景知識が制限されている
(2) サンプリングをする前に（攻撃者の知らない）処理が行われる

次章からは，この両者の場合について，それぞれサンプリングがどの程度リスクを低減できるかを評価する．

4. 攻撃者の背景知識を制限した場合

本章では，攻撃者の背景知識を制限した場合に，どの程度サンプリングが個人識別リスクを低減できるか評価する．

4.1 想定

攻撃者がテーブルの一部レコードを知っている場合を考

える．例えば，各都道府県の患者情報を集めて匿名化する場合，神奈川県の病院の勤務者は神奈川県の患者情報に詳しいが，全国の情報については未知である場合が相当する．

4.2 個人識別のリスク指標

確率的な匿名化を行った際の個人識別リスクの評価指標として Pk-匿名性があるが，オリジナルの Pk-匿名性は任意の攻撃者の知識を想定している．そのため，前章で議論した通り，サンプリングのみではオリジナルの Pk-匿名性を満たすことはできない．そこで，ここでは Pk-匿名性に倣った，攻撃者の知識を制限した個人識別リスクの評価指標を用いる．

定義 4.1 匿名化処理 Δ と匿名化テーブル τ' が，任意のテーブル τ，任意のテーブル τ に含まれるレコード $r \in \mathcal{R}$，任意の匿名化テーブル τ' に含まれるレコード $r' \in \mathcal{R}$ が，**ある攻撃者** f_T に対して

$$\Pr[\Pi(r) = r' \mid T' = \tau'] \leq \xi$$

であるならば，f_T **に対して**個人識別リスクが ξ であるという．

4.3 攻撃者モデル

本論文では，以下のような攻撃者を想定する．

(1) 全 n レコードのテーブル τ のうち，その一部である n_0 個のレコードのテーブル τ_0 の値を知っている．

(2) 攻撃者は自身の知らないレコードの値の分布を，現在自身が知っているレコードの分布と同じだと推測する．

2 について，攻撃者は n_0 個のレコード τ_0 以外の知識がないため，知らないレコードの値を一様と推定することに比べても，この推測は現実的と考えられる．

上記のモデルを定式化すると，$\bar{\mathcal{R}}$ を攻撃者から未知のレコードの集合としたとき，攻撃者の知識 f_T は任意の $\tau \in \mathcal{T}$ について

$$f_T(\tau) := \prod_{r \in \bar{\mathcal{R}}} \frac{|\{r' \mid r_0 \in \mathcal{R}_0, \tau(r) = \tau(r_0)\}|}{n_0} \qquad (1)$$

となる．

4.4 個人識別リスクの評価

匿名化手法としてサンプリング，すなわち全体のレコード数を n としたとき，サンプリング後のレコード数が $\lfloor pn \rfloor$ となるように，レコードの中から一様ランダムに抜き出す場合を考える．

このとき，攻撃者モデルで定義した背景知識を持つ攻撃者に対して，サンプリングがどの程度個人識別リスクが低減されるかを評価する．簡単のため，$pn = \lfloor pn \rfloor$ と仮定する．また，$\widehat{n} := n - n_0 - (pn - 1)$ とし，テーブル τ のうち，属性値 v であるレコード数を $\#_\tau(v) := \left|\tau^{-1}(\{v\})\right|$ と書く．

主張 4.1 全レコード数が n であるような元テーブル τ から，pn レコードをサンプリングした場合を考える．このとき，テーブル τ に対してサンプリングを行った場合，任意の τ_0, τ' について，式 1 に対して個人識別リスクは

$$\sum_{i<\widehat{n}} \binom{\widehat{n}}{i-1} \frac{1}{\sum_{j<\widehat{n}} j\binom{\widehat{n}}{j-1}(n_0-1)^{i-j}}$$

となる．ただし $\binom{\widehat{n}}{0} = 1$ とする．

[**証明のスケッチ**] 個人推定の確率の最大値を得るために，まず想定する攻撃者にとっての最適な元テーブルの知識 τ_0 と匿名化テーブル τ' を考える．あるレコード r の特定を行うときに攻撃者にとって最も有利になる τ_0, τ' として

- $\#_{\tau_0}(\tau_0(r)) = \#_{\tau'}(\tau_0(r)) = 1$，すなわち τ_0 でユニークだったレコードが，匿名化テーブルでもユニークなレコードとなる場合で，
- $|\{r' \mid r' \in \mathcal{R}', \tau'(r) \notin \{\tau_0(r_0) \mid r_0 \in \mathcal{R}_0\}\}| = pn - 1$，すなわち匿名化テーブルのユニークなレコード以外は，τ_0 に含まれない，

という場合を考える．前者はユニークであれば出現確率が小さくなることから，後者は攻撃者にとって未知なレコードの個数を最小化することからくる．このとき，攻撃者にとって未知なレコード数が $\widehat{n}$ であることに注意すると，$\binom{\widehat{n}}{0} = 1$ としたとき，

$$\begin{aligned}
&\Pr[\Pi(r) = r' \mid T' = \tau'] \\
&= \sum_{i<n} \Pr[\Pi(r) = r' \mid T' = \tau', T \in \mathcal{T}_i] \Pr[T \in \mathcal{T}_i \mid T' = \tau'] \\
&\leq \sum_{i<\widehat{n}} \frac{1}{i} \Pr[T \in \mathcal{T}_i \mid T' = \tau'] \\
&= \sum_{i<\widehat{n}} \frac{1}{i} \frac{\Pr[T' = \tau' \mid T \in \mathcal{T}_i] \Pr[T \in \mathcal{T}_i]}{\Pr[T' = \tau']} \\
&= \sum_{i<\widehat{n}} \frac{1}{i} \frac{\frac{i}{\binom{n}{pn}}(pn)!\binom{\widehat{n}}{i-1}\left(\frac{1}{n_0}\right)^{i-1}\left(1-\frac{1}{n_0}\right)^{\widehat{n}-(i-1)}}{\sum_{j<\widehat{n}} \frac{j}{\binom{n}{pn}}(pn)!\binom{\widehat{n}}{j-1}\left(\frac{1}{n_0}\right)^{j-1}\left(1-\frac{1}{n_0}\right)^{\widehat{n}-(j-1)}} \\
&= \sum_{i<\widehat{n}} \frac{1}{i} \frac{\frac{i}{\binom{n}{pn}}\binom{\widehat{n}}{i-1}\left(\frac{1}{n_0}\right)^{i-1}\left(1-\frac{1}{n_0}\right)^{\widehat{n}-(i-1)}}{\sum_{j<\widehat{n}} \frac{j}{\binom{n}{pn}}\binom{\widehat{n}}{j-1}\left(\frac{1}{n_0}\right)^{j-1}\left(1-\frac{1}{n_0}\right)^{\widehat{n}-(j-1)}} \\
&= \sum_{i<\widehat{n}} \frac{\binom{\widehat{n}}{i-1}\left(\frac{1}{n_0}\right)^{i-1}\left(1-\frac{1}{n_0}\right)^{\widehat{n}-(i-1)}}{\sum_{j<\widehat{n}} j\binom{\widehat{n}}{j-1}\left(\frac{1}{n_0}\right)^{j-1}\left(1-\frac{1}{n_0}\right)^{\widehat{n}-(j-1)}} \\
&= \sum_{i<\widehat{n}} \frac{\binom{\widehat{n}}{i-1}(n_0-1)^{\widehat{n}-i+1}}{\sum_{j<\widehat{n}} j\binom{\widehat{n}}{j-1}(n_0-1)^{\widehat{n}-j+1}} \\
&= \sum_{i<\widehat{n}} \binom{\widehat{n}}{i-1} \frac{1}{\sum_{j<\widehat{n}} j\binom{\widehat{n}}{j-1}(n_0-1)^{i-j}}.
\end{aligned}$$

となる． □

4.5 効果と議論

本章での背景知識の制限は，テーブルのうち一部しか知らないということである．この制限は既知のレコード数に対してはロバストである．例えば，各都道府県の患者情報を集めて匿名化する際に，神奈川県の患者情報に詳しい攻撃者を考えたとき，東京の数人について知っていたとしても，本章で導出した個人推定リスクはそれほど増えない．

しかし，神奈川の患者情報の中に，明らかに「絶対に全国に 1 人しかいない病歴」を持つ人がいた場合，本論文の指標の前提を満たすことはできない．そのため，現実にこの指標を使用する際は，トップ/ボトムコーディング等を用いた外れ値の削除と組み合わせることを想定すべきである．

4.6 出力のレコード数を保証しないサンプリングの安全性について

サンプリングには，全体のレコード数を n としたとき $\lfloor pn \rfloor$ 個のレコード数を出力するような，出力のレコード数が決まっているサンプリングの他にも，レコード毎に独立に確率 p でサンプリングする方法が考えられる．しかしながら，レコード数を保証しないサンプリングは，少なくとも本論文での評価基準では個人識別のリスクを低減できない．

本論文ではサンプリングの評価を行う際に任意の匿名化テーブル τ' を想定している．すなわち，出力し得る匿名化テーブルのうち，1 つでも個人識別のリスクを低減できない場合があれば「個人識別のリスクは低減できていない」と判断される．一定の確率での匿名化の失敗を許容する定義も考えられるが，例えば k-匿名化のアルゴリズムは常に k-匿名性を満たすデータを出力することを考えれば，サンプリングを用いた際の個人識別リスクの評価の第一歩として自然であると考える．

さて，レコード数を保証しないサンプリングの場合，攻撃者の知らないレコード数が 0 になってしまう匿名化テーブル τ を出力する，すなわち

$$\#_{\tau'}\left(\tau'(r') \notin \{\tau_0(r_0) \mid r_0 \in \mathcal{R}_0\}\right) \geq n - n_0$$

となる可能性がある．そのため，レコード数を保証しないサンプリングは，少なくとも本論文の評価方法では個人識別リスクを低減できない．これは，直観的には，レコード数を保証しないサンプリングの場合，元データをそのまま出力する可能性があり，その場合匿名性が保証できないということである．

5. 他手法と組み合わせた場合: サンプリングと PRAM

これまで見てきた通り，サンプリングは任意の背景知識を持っている場合は個人識別リスクの低減は難しいが，背景知識を限定し，攻撃者から見た元テーブルの分布に不確実性がある場合は，個人識別のリスクを低減することができる．

一方，アプリケーションによっては妥当な背景知識の定義が難しく，任意の背景知識に対して個人識別リスクを下げる方法も必要となる．しかし，これまで見てきた通り，サンプリングは任意の背景知識に対しては安全でない．そこで，サンプリングをする前に確率的な処理である PRAM を適用し，個人識別リスクを低減させることを考える．

直観的には，もし攻撃者が任意の背景知識を持っている，すなわち元のテーブルを知っているとしても，まず PRAM を行うことによって，サンプリングの対象となるテーブルに不確実性が発生するため，サンプリングを行って個人識別のリスクを低減することができる．

5.1 PRAM とサンプリングを組み合わせた際の個人識別リスク

任意の背景知識を持つ攻撃者に対して，PRAM とサンプリングを組み合わせることで，どの程度個人識別リスクが低減できるのかを評価する．

定理 5.1 遷移確率行列 A に基づき PRAM を適用した後，全 n レコードのうち pn レコードをサンプリングした場合，

$$k \geq 1 + (n - pn) \min_{\substack{u,v \in \mathcal{V} \\ u' \in \mathcal{V}'}} \frac{A_{u,u'}}{A_{v,u'}} + (pn - 1) \min_{\substack{u,v \in \mathcal{V} \\ u',v' \in \mathcal{V}'}} \frac{A_{u,v'} A_{v,u'}}{A_{u,u'} A_{v,v'}}$$

であるような Pk-匿名性を満たす．

[証明]

Pk-匿名性の定義から，個人識別の確率 $\Pr[\Pi(r) = r' \mid T' = \tau']$ を見ていく．

$$\begin{aligned}
&\Pr[\Pi(r) = r' \mid T' = \tau'] \\
&= \frac{\Pr[\Pi(r) = r' \wedge T' = \tau']}{\Pr[T' = \tau']} \\
&= \frac{\Pr[\Pi(r) = r' \wedge \Delta(T) = \tau' \circ \Pi]}{\Pr[\Delta(T) = \tau' \circ \Pi]} \\
&= \frac{\sum_{\tau \in \mathcal{T}} f_T(\tau) \Pr[\Pi(r) = r' \wedge \Delta(\tau) = \tau' \circ \Pi]}{\sum_{\tau \in \mathcal{T}} f_T(\tau) \Pr[\Delta(\tau) = \tau' \circ \Pi]}
\end{aligned}$$

最後の等式は T が Π, Δ と独立であることによる．上の式は [7] の lemma 3 より，任意の f_T では元テーブルを知っているような攻撃者が識別確率の最大値を取る，すなわち $f_T(\tau) = 1$ かつ $\bar{\tau} \in \mathcal{T} \setminus \{\tau\}$ について $f(\bar{\tau}) = 0$ が最大となるため，

$$\begin{aligned}
&\leq \frac{\Pr[\Pi(r) = r' \wedge \Delta(\tau) = \tau' \circ \Pi]}{\Pr[\Delta(\tau) = \tau' \circ \Pi]} \\
&= \frac{\frac{1}{\binom{n}{pn}(pn)!} \sum_{\pi(r) = r'} \Pr[\Delta(\tau) = \tau' \circ \pi]}{\frac{1}{\binom{n}{pn}(pn)!} \sum_{\pi \in \mathcal{P}} \Pr[\Delta(\tau) = \tau' \circ \pi]} \\
&= \frac{\sum_{\pi(r) = r'} \Pr[\Delta(\tau) = \tau' \circ \pi]}{\sum_{\pi \in \mathcal{P}} \Pr[\Delta(\tau) = \tau' \circ \pi]}
\end{aligned}$$

ここで簡単のため上の式の逆数をとる．

$$\frac{\sum_{\pi\in\mathcal{P}}\Pr[\Delta(\tau)=\tau'\circ\pi]}{\sum_{\pi(r)=r'}\Pr[\Delta(\tau)=\tau'\circ\pi]} = 1+\frac{\sum_{\pi(r)\neq r'}\Pr[\Delta(\tau)=\tau'\circ\pi]}{\sum_{\pi(r)=r'}\Pr[\Delta(\tau)=\tau'\circ\pi]}$$

サンプリング π によって抽出されるレコードの集合を $\widetilde{\mathcal{R}}_\pi := \{r \mid r\in\mathcal{R}, \pi(r)\neq\phi\}$ と定義し，ある π，$s\in\mathcal{R}$，$s'\in\mathcal{R}'$ について

$$A^{\tau,\tau'}_{s,s'} := \Pr[(\Delta(\tau))(s)=\tau'(s')]$$

と定義する．ただし，便宜上 $A^{\tau,\tau'}_{s,\phi}=1$ とする．

このとき，PRAM の処理はレコード毎であることに注意すると，先ほどの式は，

$$\begin{aligned}
&= 1+\frac{\sum_{\pi(r)\neq r'}\prod_{s\in\mathcal{R}}A^{\tau,\tau'}_{s,\pi(s)}}{\sum_{\pi(r)=r'}\prod_{s\in\mathcal{R}}A^{\tau,\tau'}_{s,\pi(s)}}\\
&= 1+\frac{\sum_{\pi(r)=\phi}\prod_{s\in\mathcal{R}}A^{\tau,\tau'}_{s,\pi(s)}+\sum_{\substack{\pi(r)\in\widetilde{\mathcal{R}}_\pi\\ \pi(r)\neq r'}}\prod_{s\in\mathcal{R}}A^{\tau,\tau'}_{s,\pi(s)}}{\sum_{\pi(r)=r'}\prod_{s\in\mathcal{R}}A^{\tau,\tau'}_{s,\pi(s)}}\\
&= 1+\frac{\sum_{t\neq r}A^{\tau,\tau'}_{t,r'}\sum_{\substack{\pi(r)=\phi,\\ \pi(t)=r'}}\prod_{s\neq t}A^{\tau,\tau'}_{s,\pi(s)}}{\sum_{\pi(r)=r'}A^{\tau,\tau'}_{r,r'}\prod_{s\neq r}A^{\tau,\tau'}_{s,\pi(s)}}\\
&\quad+\frac{\sum_{\substack{t\neq r\\ t'\neq r'}}A^{\tau,\tau'}_{t,r'}A^{\tau,\tau'}_{r,t'}\sum_{\substack{\pi(t)=r'\\ \pi(r)=t'}}\prod_{s\neq t,r}A^{\tau,\tau'}_{s,\pi(s)}}{\sum_{\pi(r)=r'}A^{\tau,\tau'}_{r,r'}\prod_{s\neq r}A^{\tau,\tau'}_{s,\pi(s)}}\\
&= 1+\frac{\sum_{t\neq r}A^{\tau,\tau'}_{t,r'}\sum_{\substack{\pi(t)=\phi,\\ \pi(r)=r'}}\prod_{s\neq r}A^{\tau,\tau'}_{s,\pi(s)}}{\sum_{\pi(r)=r'}A^{\tau,\tau'}_{r,r'}\prod_{s\neq r}A^{\tau,\tau'}_{s,\pi(s)}}\\
&\quad+\frac{\sum_{\substack{t\neq r\\ t'\neq r'}}A^{\tau,\tau'}_{t,r'}A^{\tau,\tau'}_{r,t'}\sum_{\substack{\pi(r)=r'\\ \pi(t)=t'}}\prod_{s\neq t,r}A^{\tau,\tau'}_{s,\pi(s)}}{\sum_{\pi(r)=r'}A^{\tau,\tau'}_{r,r'}\prod_{s\neq r}A^{\tau,\tau'}_{s,\pi(s)}}\\
&= 1+\frac{\sum_{t\neq r}A^{\tau,\tau'}_{t,r'}\sum_{\substack{\pi(t)=\phi,\\ \pi(r)=r'}}A^{\tau,\tau}_{r,\pi(t)}\frac{\prod_{s\neq r}A^{\tau,\tau'}_{s,\pi(s)}}{A^{\tau,\tau'}_{t,\pi(t)}}}{\sum_{\pi(r)=r'}A^{\tau,\tau'}_{r,r'}\prod_{s\neq r}A^{\tau,\tau'}_{s,\pi(s)}}\\
&\quad+\frac{\sum_{t\neq r}A^{\tau,\tau'}_{t,r'}\sum_{\substack{\pi(r)=r'\\ \pi(t)\neq\phi}}A^{\tau,\tau'}_{r,\pi(t)}\frac{\prod_{s\neq r}A^{\tau,\tau'}_{s,\pi(s)}}{A^{\tau,\tau'}_{t,\pi(t)}}}{\sum_{\pi(r)=r'}A^{\tau,\tau'}_{r,r'}\prod_{s\neq r}A^{\tau,\tau'}_{s,\pi(s)}}\\
&= 1+\frac{\sum_{t\neq r}A^{\tau,\tau'}_{t,r'}\sum_{\pi(r)=r'}A^{\tau,\tau}_{r,\pi(t)}\frac{\prod_{s\neq r}A^{\tau,\tau'}_{s,\pi(s)}}{A^{\tau,\tau'}_{t,\pi(t)}}}{\sum_{\pi(r)=r'}A^{\tau,\tau'}_{r,r'}\prod_{s\neq r}A^{\tau,\tau'}_{s,\pi(s)}}\\
&= 1+\frac{\sum_{\pi(r)=r'}\sum_{t\neq r}\frac{A^{\tau,\tau'}_{t,r'}A^{\tau,\tau}_{r,\pi(t)}}{A^{\tau,\tau'}_{t,\pi(t)}}\prod_{s\neq r}A^{\tau,\tau'}_{s,\pi(s)}}{\sum_{\pi(r)=r'}A^{\tau,\tau'}_{r,r'}\prod_{s\neq r}A^{\tau,\tau'}_{s,\pi(s)}}
\end{aligned}$$

ここで文献 [7] の lemma 4 より，任意の π ただし $\pi(r)=r'$ について

$$\frac{1}{A^{\tau,\tau'}_{r,r'}}\sum_{t\neq r}\frac{A^{\tau,\tau'}_{t,r'}A^{\tau,\tau'}_{r,\pi(t)}}{A^{\tau,\tau'}_{t,\pi(t)}} \tag{2}$$

を最小化するような $A^{\tau,\tau'}$ が存在すれば，

$$\frac{\sum_{\pi(r)=r'}\sum_{t\neq r}\frac{A^{\tau,\tau'}_{t,r'}A^{\tau,\tau}_{r,\pi(t)}}{A^{\tau,\tau'}_{t,\pi(t)}}\prod_{s\neq r}A^{\tau,\tau'}_{s,\pi(s)}}{\sum_{\pi(r)=r'}A^{\tau,\tau'}_{r,r'}\prod_{s\neq r}A^{\tau,\tau'}_{s,\pi(s)}}$$

は式 2 で最小値となる．そのような $A^{\tau,\tau'}$ は任意の $s\neq r$ について $\tau(s)=v, v\neq\tau(r)$ であり，任意の $s'\neq r'$ について $\tau(s')=v', v'\neq\tau(r')$ である．また，π はサンプリングのため，pn 個のみサンプリングされることに注意すると，先ほどまでの式は，

$$\geq 1+(n-pn)\min_{\substack{u,v\in\mathcal{V}\\ u'\in\mathcal{V}'}}\frac{A_{u,u'}}{A_{v,u'}}+(pn-1)\min_{\substack{u,v\in\mathcal{V}\\ u',v'\in\mathcal{V}'}}\frac{A_{u,v'}A_{u',v}}{A_{u',v}A_{u,v'}}$$

となる． □

5.2 サンプリングの効果

サンプリングと PRAM の組み合わせの場合，個人特定の確率 $\frac{1}{k}$ の k は

$$k\geq 1+(n-pn)\min_{\substack{u,v\in\mathcal{V}\\ u'\in\mathcal{V}'}}\frac{A_{u,u'}}{A_{v,u'}}+(pn-1)\min_{\substack{u,v\in\mathcal{V}\\ u',v'\in\mathcal{V}'}}\frac{A_{u,v'}A_{u',v}}{A_{u',v}A_{u,v'}}$$

であるのに対し，PRAM のみの場合 [7] は，

$$k\geq 1+(n-1)\min_{\substack{u,v\in\mathcal{V}\\ u',v'\in\mathcal{V}'}}\frac{A_{u,v'}A_{u',v}}{A_{u',v}A_{u,v'}}$$

となる．ここで，

$$\min_{\substack{u,v\in\mathcal{V}\\ u'\in\mathcal{V}'}}\frac{A_{u,u'}}{A_{v,u'}}>\min_{\substack{u,v\in\mathcal{V}\\ u',v'\in\mathcal{V}'}}\frac{A_{u,v'}A_{u',v}}{A_{u',v}A_{u,v'}}$$

であるから，サンプリングが個人識別リスクの低減に役立っていることがわかる．

なお，出力のレコード数を保証しないサンプリングは，PRAM との組み合わせにおいても個人識別リスクを低減しない．理由は攻撃者の背景知識を制限した時と同様，サンプリングするテーブルをそのまま出力する可能性があるからである．

6. 結果と今後の課題

本論文では，サンプリングがどの程度個人識別リスクを低減できるかを明らかにするために，背景知識によっては

サンプリングのみでは低減できないこと，適切に背景知識を制限できる場合には個人識別リスクを低減できること，さらに PRAM と組み合わせることで任意の背景知識に対して識別リスクを低減できることを示した．

これにより，妥当な範囲で背景知識が制限する場合，例えば全国からデータを持ち寄り，一部の地方の情報のみを知っている攻撃者を想定する場合，どの程度のレコード数をサンプリングすれば良いかの基準として用いることができる．

また，背景知識の制限が現実的に難しい場合は，サンプリングに PRAM を組み合わせることで，個人識別のリスクを低減できることを示した．この方法はこれまで知られていた PRAM のみの方法に比べ加えるノイズを少なくできている．

層別抽出などの複数の重みを用いたサンプリング方法への対応は，今後の課題である．

参考文献

[1] K. Chaudhuri, E. Halperin, S. Rao, and S. Zhou. A rigorous analysis of population stratification with limited data. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2007, New Orleans, Louisiana, USA, January 7-9, 2007*, pp. 1046–1055, 2007.

[2] B. Chen, D. Kifer, K. LeFevre, and A. Machanavajjhala. Privacy-preserving data publishing. *Foundations and Trends in Databases*, 2(1-2):1–167, 2009.

[3] C. Dwork. Differential privacy. In *Automata, Languages and Programming, 33rd International Colloquium, ICALP 2006, Venice, Italy, July 10-14, 2006, Proceedings, Part II*, pp. 1–12, 2006.

[4] C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, and M. Naor. Our data, ourselves: Privacy via distributed noise generation. In *Advances in Cryptology - EUROCRYPT 2006, 25th Annual International Conference on the Theory and Applications of Cryptographic Techniques, St. Petersburg, Russia, May 28 - June 1, 2006, Proceedings*, pp. 486–503, 2006.

[5] C. Dwork, F. McSherry, K. Nissim, and A. D. Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography, Third Theory of Cryptography Conference, TCC 2006, New York, NY, USA, March 4-7, 2006, Proceedings*, pp. 265–284, 2006.

[6] L. Franconi and S. Polettini. Individual risk estimation in μ-Argus: A review. In *Privacy in Statistical Databases: CASC Project International Workshop, PSD 2004, Barcelona, Spain, June 9-11, 2004. Proceedings*, pp. 262–272, 2004.

[7] D. Ikarashi, R. Kikuchi, K. Chida, and K. Takahashi. k-anonymous microdata release via post randomisation method. In *Advances in Information and Computer Security - 10th International Workshop on Security, IWSEC 2015, Nara, Japan, August 26-28, 2015, Proceedings*, pp. 225–241, 2015.

[8] J. Lee and C. Clifton. How much is enough? choosing ϵ for differential privacy. In *Information Security, 14th International Conference, ISC 2011, Xi'an, China, October 26-29, 2011. Proceedings*, pp. 325–340, 2011.

[9] N. Li, W. H. Qardaji, and D. Su. On sampling, anonymization, and differential privacy or, k-anonymization meets differential privacy. In *7th ACM Symposium on Information, Compuer and Communications Security, ASIACCS '12, Seoul, Korea, May 2-4, 2012*, pp. 32–33, 2012.

[10] B. Lin, Y. Wang, and S. Rane. On the benefits of sampling in privacy preserving statistical analysis on distributed databases. *CoRR*, abs/1304.4613, 2013.

[11] A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam. L-diversity: Privacy beyond k-anonymity. *TKDD*, 1(1), 2007.

[12] M. E. Nergiz, M. Atzori, and C. Clifton. Hiding the presence of individuals from shared databases. In *Proceedings of the ACM SIGMOD International Conference on Management of Data, Beijing, China, June 12-14, 2007*, pp. 665–676, 2007.

[13] M. E. Nergiz and C. W. Clifton. d-presence without complete world knowledge. *IEEE Trans. Knowl. Data Eng.*, 22(6):868–883, 2010.

[14] K. Peter, W. Leon, and G. Jose. *PRAM: a Method for Disclosure Limitation of Microdata.* Research paper. CBS, 1997.

[15] A. Smith. Differential privacy and the secrecy of the sample, 2009. `https://adamdsmith.wordpress.com/2009/09/02/sample-secrecy/`.

[16] J. Soria-Comas and J. Domingo-Ferrer. Probabilistic k-anonymity through microaggregation and data swapping. In *FUZZ-IEEE 2012, IEEE International Conference on Fuzzy Systems, Brisbane, Australia, June 10-15, 2012, Proceedings.*, pp. 1–8, 2012.

[17] L. Sweeney. k-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(5):557–570, 2002.