Computer Security Symposium 2018
22 - 25 October 2018

属性推定攻撃を抑止可能なプログラム送付型匿名化方式の提案

前田 若菜[1]　山岡 裕司[1]

概要：匿名化はプライバシー侵害が起きないように元のデータを加工するため，生成された匿名データは元データよりも有用性が低下する．匿名データを受領する受領者にとっての有用性を高める方法として，受領者自身の有用性の考え方を示した匿名化意向を提供者に提案する方式が存在する．しかし，受領者が元データを分析せずに，匿名化前後で有用性の低下を抑えるような匿名化意向を作成するのは難しいという課題があった．そこで，元データを分析するプログラムによって匿名化意向を生成し匿名化を行うプログラム送付型匿名化方式を提案する．単純方式の場合，データ中に一人しかいないような特定個人について，属性の値を推定する属性推定攻撃の脅威がある．提案方式では，複数のサンプリングによって一人しかいないような特定個人がデータに含まれないようにし，含まれた場合でも当該データの分析結果が匿名化意向として採択されにくくすることで攻撃成功確率を低減する．

キーワード：
PPDP, プライバシー保護, 匿名化, 属性推定攻撃, オーダーメード集計

Wakana Maeda[1]　Yuji Yamaoka[1]

1. はじめに

パーソナルデータの活用法として，保有するパーソナルデータを第三者に提供する方法がある．しかし，パーソナルデータをそのまま提供するとプライバシー侵害の問題がある．そのため，個人のプライバシーを保護しながら，データの有用性を保つようにデータを加工して公開するPPDP（Privacy-preserving data publishing）という考え方がある [1][2]．基本モデルとして，データを加工して提供する提供者と，加工されたデータを受け取る受領者から成り立つ．

PPDP の実現方法の一つに匿名化がある．匿名化されたデータである匿名データは，個票形式のデータである．匿名データはプライバシー侵害が起きないように加工されているため，元データとは同一ではない．そのため，匿名データを受け取る受領者にとってのデータの有用性が減少しないように匿名化することが求められる．

受領者にとっての有用性は受領者の目的によって異なる．例えば，収入，年齢，居住地から成るデータを使って収入との関係性を分析したい受領者を仮定する．このとき，受領者が年齢に基づいて収入分析をしたいなら，居住地データを加工するより年齢データを加工する方が有用性の減少は大きい．一方，受領者が居住地に基づいて収入分析をしたいなら，年齢データを加工するより居住地データを加工する方が有用性の減少は大きい．このように，データをどう扱うかによって有用性の考え方が異なる．受領者の考える有用性を実現するためには，受領者がどのデータは加工してほしくないのか，又は加工してもいいかという意向を匿名化に反映させればよい．そのため，受領者の意向を提供者に伝える方式をとるのが有効である．この意向を匿名化意向とする．

しかし，どのデータは加工していいか，加工しないほうがいいかという受領者の意向がデータ受領前には明確になっていない場合がある．例えば，受領者が収入予測モデルを作成するときに，収入と相関が強い情報が重要だと考えるなら，相関が強い情報を加工するのは有用性減少に影響するだろう．逆に，相関の弱い情報を加工するのは，相関が強い情報を加工するよりも有用性減少への影響が小さいだろう．相関の高さは実際にデータを分析すればわかるが，受領者は加工前のデータを受領して分析することはで

[1] 株式会社富士通研究所
FUJITSU LABORATORIES LTD.

きない．なぜなら前述したように，プライバシー保護がなされていない加工前のパーソナルデータを受領することができないからである．このように，受領者は匿名化に対して意向を決めるのが難しいという課題がある．単純に相関を求めることは提供者側で可能だが，相関以外の分析手法やデータの丸め方，どの尺度でデータを扱うかなど受領者の複雑な要求を可能にする仕様を提供者側で用意をするのは難しい．

受領者が直接パーソナルデータを見ずにパーソナルデータを分析する方法として，プログラム送付型集計がある [3]．これは，政府統計データ提供 [4][5] におけるリモートアクセス集計の一つであり，受領者がプログラムを提供者に送付し，集計表を得るものである．例えば，ルクセンブルク所得研究では，受領者は統計ソフトパッケージ用のスクリプトを提供者に送付し，集計結果として年次ごとのジニ係数などを得られる．

このプログラム送付型集計では，受領者は集計表を得ることができるが，集計の元となる個票形式のデータを得ることができない．しかし，受領者が個票形式のデータを必要とする場合がある．例えば，機械学習を行う場合には，学習データとして必要なのは個票形式のデータである．このように個票形式のデータが必要な場合は，集計表ではなく，個票形式のデータである匿名データを受領する必要がある．しかし，このプログラム送付型集計では匿名データを受領することはできない．

本稿では，受領者が送付したプログラムの出力に基づいて匿名化する方式を提案する．この方式のことをプログラム送付型匿名化と呼ぶことにする．この方式では，受領者はプログラムを使ってパーソナルデータを直接見ずに分析し，その分析結果をふまえて匿名化意向を生成することができる．そして，提供者はこの匿名化意向に基づいて匿名データを生成し，受領者に提供する．なお，プログラムを統計ソフト用のスクリプトには限定せず，自由な記述を受け付けることとする．

プログラム送付型匿名化には，属性推定攻撃への対策を行う必要がある．本稿における属性推定攻撃とは，受領データをもとに特定個人の知られたくない属性の値を推定する攻撃である．プログラム送付型匿名化では，攻撃者がプログラムを工夫することで，特定個人の属性の値を推定可能な匿名化意向を生成することが可能である．提供者がこの匿名化意向にしたがってパーソナルデータを匿名化し攻撃者に提供したとする．このとき，攻撃者は受領した匿名データから匿名化意向を推測し，結果として特定個人の属性情報を推定することができる．

本稿では，属性推定攻撃を抑止可能なプログラム送付型匿名化方式を提案する．提案方式は次の二つのアイデアによって成っている．一つは，パーソナルデータからサンプリングした複数の部分データについてプログラムを実行し，最大出現頻度の匿名化意向を採択する．もう一つはプログラムが出力できる匿名化意向の種類数を制限する．これらは，一人しかいないような特定個人がデータに含まれないようにし，含まれた場合でも分析結果が匿名化意向として採択されにくくすることで，攻撃成功確率を低減するものである．

(a)

識別子	準識別子			センシティブ属性
name	gender	age	occupation	Income
Alice	male	31	teacher	4M
Bob	male	31	teacher	4M
Carol	female	33	teacher	4M

(b)

準識別子			センシティブ属性
gender	age	occupation	Income
-	-	teacher	4M
-	-	teacher	4M
-	-	teacher	4M

(c)

準識別子			センシティブ属性
gender	age	occupation	Income
male	31	teacher	4M
male	31	teacher	4M
-	-	teacher	4M

(d)

準識別子			センシティブ属性
gender	age	occupation	Income
male	31	teacher	4M
male	31	teacher	4M
-	-	-	-

表 1 パーソナルデータ例と抑制例

本稿の貢献は，次のとおりである．

- 受領者の意向を，受領者が作成したプログラムの分析結果によって生成し，その意向に基づいて匿名化を行うプログラム送付型匿名化を提案した．
- プログラム送付型匿名化を行った際におこりうる属性推定攻撃を明らかにし，さらにこれを抑止する方式を提案した．

本稿の構成は次のとおりである．2 章では関連研究を記述，3 章でプログラム送付型匿名化とそれへの攻撃モデルを定義する．4 章で攻撃を抑止する提案方式を説明し，5 章で提案方式が 3 章で定義した攻撃を抑止する様相を詳説する．最後に 6 章にて本稿をまとめる．

2. 関連研究

パーソナルデータは，個人に関するデータであり，表形式で表現される．各個人のデータは，表の 1 行として表現される．本稿ではこれをレコードと呼ぶ．表の列をなす属性として，識別子，準識別子，センシティブ属性がある．識別子は，ID や氏名など単体で個人を直接識別できる属性である．準識別子は，単体で個人を識別できないが，他の準識別子と組み合わさることで個人を識別できる属性をさす．センシティブ属性は，受領者が分析対象としている属性，並びに，攻撃者にとって未知の属性をさす．複数のレコードにおいて，準識別子について同じ属性値をもつグループを同値類と呼ぶ．例えば，表 1 の (a) における背景色がついた準識別子の組み合わせ {male,31,teacher} がひとつの同値類，色のついていない {female,33,teacher} がもうひとつの同値類である．

匿名化に用いられる代表的な指標として，k-anonymity がある [6]．これは，データ中に同値類が k 個以上存在することを示す指標である．言い換えれば，ある準識別子の組み合わせから個人を k 人未満に絞り込めないことをしめ

す指標である．k-anonymity を満たすように匿名化することを k-匿名化という．

k-anonymity を満たすために使われる方法として一般化（generalization）と抑制（suppression）がある [6][7]．一般化は，値を抽象的な値又は上位概念に置換する方法である．例えば，数値データであれば，28 歳を 20 代にまるめたり，カテゴリカルデータであれば，川崎市を神奈川県と上位の概念に置換したりする．抑制は，値そのものを開示しない，つまり削除する方法である．属性抑制，レコード抑制，セル抑制などがある．属性抑制とは，データが k-anonymity を満たさない原因となる準識別子の任意の属性を全て削除する方法である．レコード抑制とは，k-anonymity を満たさない準識別子の組み合わせを持つレコードを削除する方法である．セル抑制は，準識別子の組み合わせのうち k-anonymity を満たさない値を削除する方法である．例えば，あるレコードにおいて年齢は削除されるが，他のレコードでは年齢は削除されない．セル単位で削除を行う．表 1 において，(a) を 2-匿名化するために抑制を行った結果が (b) (d) である．(b) が属性抑制，(c) がセル抑制，(d) がレコード抑制を示したものである．

匿名化は，プライバシー保護だけでなく，有用性の点で元のパーソナルデータと匿名データとの分析結果の差異を小さくすることをめざしている．Tian[8] はその差異を小さくする方式として，受領者の匿名化意向に従って匿名化する方式を提案している．この方式は，受領者が提供者に匿名化意向として有用性設計（Utility Specification）を送付するものである．この有用性設計には，加工してほしくない属性を優先順位をつけて定めることができる．作成に当たっては，受領者が自身の意向を決めることが難しいという課題がある．1 章であげた例のように，センシティブ属性と相関が強い属性を加工したくないと考えた場合，元のパーソナルデータから相関が強い属性を明らかにしなければ有用性設計を作成することができない．

3. 定義

本章では，3.1 節で定義のための準備，3.2 節でプログラム送付型匿名化の単純方式を定義する．そして，3.3 節で攻撃モデルについて定義する．

3.1 準備

パーソナルデータ D の識別子属性を除いた $j+1$ 列のデータについて匿名化を行う．匿名化されたデータを匿名データ D' とする．列は，準識別子属性 $A_1, A_2 ... A_j$，センシティブ属性 $S = \{d_1, d_2, ... d_{|S|}\}$ から成っている．ただし $|S|$ はセンシティブ属性 S の要素数である．任意の同値類を e_x，同値類 e_x を構成するレコード数を $|e_x|$ とする．同値類 e_x がもつセンシティブ属性の値の集合を Y_{e_x} とし，センシティブ値集合と呼ぶこととする．

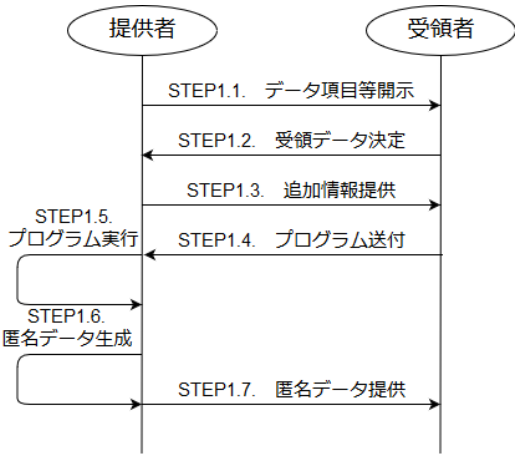

図 1 受領者の意向を踏まえた第三者提供の手続き

匿名化意向は順序付きリスト $r = [a_1, a_2, ... a_{|j+1|}]$ で表される．a は準識別子属性，又はセンシティブ属性から抽出される．なお，本稿における匿名化意向とは，Tian[8] と同じく，なるべく加工しないでほしい属性の優先順位を定めたものである．ただし，本稿では簡略化のために，Tian と異なり加工を避けてほしい値の指定は受け入れない．

なお，本稿では匿名化方法としてセル抑制による k-匿名化 [9] を用いた．

3.2 プログラム送付型匿名化の単純方式

パーソナルデータを保有する提供者と，パーソナルデータを加工した匿名データを受領する受領者がいる．図 1 に示すような匿名データ受領までの手続きを以下に記述する．なお，一連の手続きは一つのデータについて一度限り行うことができ，2 回以上行うことができないとする．

手続き

Step1.1. 提供者による，受領者へ属性等を開示

受領者が求めるデータを探せるように，提供者は提供可能なデータをまとめたデータカタログのようなものを開示する．開示内容として，データに含まれている属性情報などがある．

Step1.2. 受領者による，受領データの決定

提供者が提供できるデータから，求めるデータを決定し，提供者に伝える．このとき，受領者がプログラム実装するうえで更なる情報が必要であれば，受領者は追加情報を請求する．例えばテストデータなどが考えられる．

Step1.3. 提供者による，受領者へ追加情報提供

提供者は，受領者から追加情報請求があれば，それに応じて追加情報を提供する．

Step1.4. 受領者による，提供者へ匿名化意向を出力するプログラム送付

受領者はプログラムとして，パーソナルデータを分析し，分析結果をふまえて匿名化意向を生成するものを作成し，

提供者に送付する．

Step1.5. 提供者による，プログラム実行及び匿名化意向を生成

提供者は，パーソナルデータを受領者から送付されたプログラムに入力し，プログラムを実行する．プログラムは実行後，出力として匿名化意向を生成する．

Step1.6. 提供者による，生成された匿名化意向に基づいた匿名データ作成

提供者は送付されたプログラムの実行結果である匿名化意向に基づいて，求められたパーソナルデータの匿名化を行って匿名データを生成する．

Step1.7. 提供者による，受領者へ匿名データ提供

本ステップをもって，匿名データ提供手続きを終了する．

3.3 属性推定攻撃モデル

提供者はパーソナルデータ D 中には，ただ一人しかいない準識別子の組み合わせをもつ人物 X が含まれている．この人物 X が攻撃対象である．このうえで，攻撃モデルを記述する．

攻撃者がパーソナルデータ D 中に人物 X が含まれていることを知らない場合，攻撃者の意図は，提供者が保有するパーソナルデータ D 中に人物 X が存在するかを知ることである．一方，攻撃者がパーソナルデータ D 中に人物 X が含まれていることを知っているとき，攻撃者の意図は人物 X のセンシティブ属性の値を推定することである．

攻撃者が提供者に対してできることは，匿名化意向を生成するプログラムを送付することである．また，攻撃者が提供者から受け取れる情報は，匿名化意向に基づいて作成された匿名データ D' である．

送付するプログラムを工夫することで，攻撃者は人物 X のセンシティブ属性の値に応じて特定の匿名化意向を生成することができる．そして，特定の匿名化意向を生成することで，特定の匿名データ D' を受領することが可能である．攻撃者は，受領した匿名データ D' から，どの条件によって匿名化意向が出力されたかを推測し，人物 X のセンシティブ属性の値を推定することができる．

手順

図 2 に攻撃の手順の概念図を示す．人物 X が属する同値類を $e_x(|e_x| = 1)$，センシティブ属性 S の値を d_x とする．攻撃者は，同値類 e_x がもつセンシティブ属性の値の集合である Y_{e_x} の要素に応じて匿名化意向 r を出力するようなプログラムを作成する．同値類 e_x のセンシティブ値集合 $Y_{e_x} = \{d_1\}$ ならば匿名化意向を $r_1, Y_{e_x} = \{d_2\}$ ならば $r_2, ..., Y_{e_x} = \{d_{|S|}\}$ ならば $r_{|S|}$ といったものである．提供者は，送付されたプログラムにパーソナルデータを入力し，プログラムを実行する．出力された匿名化意向に従って匿名データ D' を作成し，これを攻撃者に提供する．

攻撃者は匿名データ D' の加工の傾向から，匿名化意向

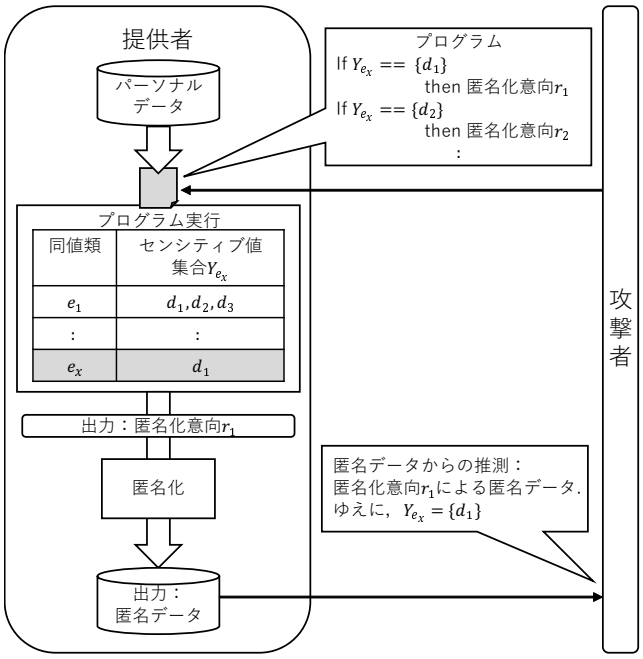

図 2 属性推定攻撃の概念図

r_x が出力されたと推測し，人物 X のセンシティブ属性が d_x であると推定できる．このように，プログラムの出力結果を反映した匿名データの傾向から，当該人物に関するセンシティブ属性の値に対して推定攻撃が可能である．

攻撃例

パーソナルデータ D は，性別属性 $A_1 = \{$ 女性，男性 $\}$，職業属性 $A_2 = \{$ 会社員，看護師 $\}$，センシティブ属性として婚姻状態 $S = \{$ 未婚，既婚 $\}$ から構成されている．攻撃者はパーソナルデータ D には，ただ一人しかいない同値類 $e_x = ($ 女性 , 看護師 $)(|e_x| = 1)$ をもつ Alice が存在しているのを知っている．攻撃者は，Alice について知らないセンシティブ属性である婚姻状態について推定したいと考えている．

そこで，攻撃者は提供者に送付するプログラムの実行結果に次のようなことを考える．同値類が $e_x = ($ 女性 , 看護師 $)$ のセンシティブ値集合 Y_{e_x} の要素によって，プログラムが出力する匿名化意向を変えよう，というものである．同値類 e_x のレコードの同値類 e_x がもつセンシティブ値集合が $Y_{e_x} = \{$ 未婚 $\}$ の場合は，匿名化意向は $r_{unmarried} = [S, A_1, A_2]$，$Y_{e_x} = \{$ 既婚 $\}$ の場合は，匿名化意向は $r_{married} = [S, A_2, A_1]$，同値類 e_x のレコードがなかったなどのそれ以外の場合は $r_{null} = [A_1, A_2, S]$ と出力するようにプログラムを作成した．攻撃者はこのプログラムを提供者に送付する．

提供者は，送付されたプログラムにパーソナルデータ D を入力し，プログラムを実行する．出力された匿名化意向 $r_{married} = [S, A_2, A_1]$ に従って匿名データ D' を作成し提供する．匿名データ D' を受領した攻撃者は，匿名データ D' が S, A_2, A_1 の順で保護されていることから匿名化

意向が $r_{married}$ であったと推測する．匿名化意向 $r_{married}$ は，$Y_{e_x} = \{$ 既婚 $\}$ のときに出力される．ゆえに，攻撃者は Alice の婚姻状態が既婚であると推定できる．

4. 提案方式

本章では，3.1 節で示した STEP1.5 についてどのように 3.2 節の属性推定攻撃を抑止するかというアイデアとそのアルゴリズムを詳説する．

4.1 アイデア

パーソナルデータ中の $|e_x| = 1$ となるような特定個人のレコードを反映した匿名化意向を r_{target}，それ以外の匿名化意向を r_{other} とする．また，匿名化意向 r の出現頻度を $|r|$ とする．提案方式は，以下の二つのアイデアを用いて属性推定攻撃を抑止する．

(1) パーソナルデータからサンプリングした複数の部分データについてプログラムを実行する．そして，得られた複数の匿名化意向の中から，最大出現頻度の匿名化意向を採択する．

サンプリングで得られたデータを用いることで，匿名化意向 r_{target} が必ずしも生成されないようにする．パーソナルデータ全体を入力にプログラムを実行する場合，パーソナルデータ中に $|e_x| = 1$ となるような特定個人のレコードが必ず含まれるため，匿名化意向 r_{target} が生成される．しかし，サンプリングで得た部分データについてプログラムを実行した場合，部分データ中に $|e_x| = 1$ となるような特定個人のレコードが含まれてなければ，匿名化意向 r_{target} は生成されない．これにより，確率的に匿名化意向 r_{target} が生成されるのを抑制することができ，攻撃成功確率を低減できる．

また，複数の匿名化意向の中から最大出現頻度の匿名化意向を採択するため，匿名化意向 r_{target} が一回以上生成されたとしても，それ以外の匿名化意向 r_{other} が多ければ，匿名化意向 r_{target} は採択されない．このように攻撃成功確率を低減できる．また，複数行うことで，一回きりの偏ったサンプリング結果が最終的な匿名化意向として採択される可能性を抑制できる．

(2) プログラムが出力できる匿名化意向の種類数に上限をもうける．

アイデア (1) に対し，攻撃者は匿名化意向 r_{target} が最大出現頻度となるように，それぞれ異なる匿名化意向 r_{other} を大量に生成することが考えられる．そのため，大量のそれぞれ異なる匿名化意向 r_{other} が生成されるのを抑制するために，匿名化意向の種類数に上限をもうける．種類数が上限以上のものについては，処理を終了し，匿名データを提供しない．

例えば，5 つの部分データを使って匿名化意向を得るとする．このとき，匿名化意向 r_{target} の出現頻度が $|r_{target}| = 2$ だとする．残り 3 つの部分データについて，それぞれ異なる匿名化意向 $r(|r| = 1)$ を出力すれば，最大出現頻度をもつ匿名化意向は r_{tagert} となる．r_{tagert} が採択されると攻撃は成功する．しかし，匿名化意向の種類数の上限を 3 にした場合，この攻撃は上限を超えているため処理が終了する．攻撃者は匿名データを受領できなくなり，推定を行うことができない．

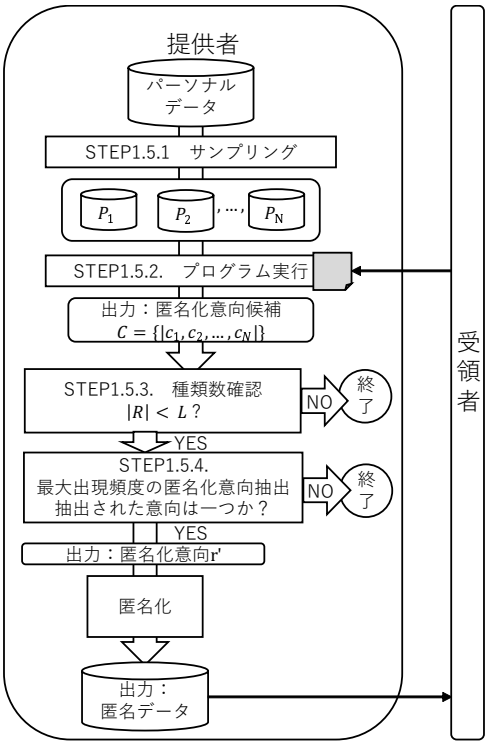

図 3　提案アルゴリズムの概念図

4.2 アルゴリズム

基本的には 3.1 節の STEP1.1〜1.7 の手続きを踏襲する．提案方式では，STEP1.5「提供者による，プログラム実行及び匿名化意向を生成」について，プログラムの実行から匿名化意向生成までの処理に対し，属性推定攻撃を抑止するアイデアを取り入れた．図 3 は提案方式の処理手続きを示した概念図である．以下に STEP1.5 にアイデアを盛り込んだアルゴリズムの説明をする．なお，これらはすべて提供者側で行われる処理である．

Step1.5.1 サンプリングによる部分データの生成

パーソナルデータ D からサンプリング率 s，サンプリング回数 N 回により，N 個の部分データ $P_n \in P$ を作成する．P は部分データの集合である．

Step1.5.2. 部分データ P_n を入力にプログラム実行

部分データ P_n を入力に，STEP1.4 で送付されたプログラムを実行する．部分データ P_n による出力結果として匿名化意向候補 c_n を得る．匿名化意向候補の多重集合を $\mathcal{C} \ni c_n$ とする．

Step1.5.3. 匿名化意向の種類数確認

匿名化意向候補多重集合 C には複数の匿名化意向が存在する．この匿名化意向は，3.1 節で定義した r_z で表すことができる．匿名化意向集合を $R = \{r_1, r_2, ...r_z\}$ とし，種類数を $|R|$ とおく．このとき，$|R|$ が提供者の設定した上限数 L を下回っているかを確認する．$|R| < L$ ならば，STEP1.5.4. に進む．そうでなければ，処理を終了する．

Step1.5.4. 最大出現頻度の匿名化意向抽出

得られた匿名化意向候補多重集合 $\mathcal{C}$ について，最大出現頻度の匿名化意向が複数あれば，出力せずに処理を終了する．最大出現頻度である匿名化意向が一つならば，それを匿名化意向 r' として出力する．

5. 提案方式による属性推定攻撃抑止

提案方式が 3.3 節で定義したような属性推定攻撃を抑止する様相を詳説する．

同値類 $|e_x| = 1$ となる特定のレコードが，部分データ集合 P のいずれかの部分データに含まれる個数を $f(0 \leq f \leq N)$ とする．例えば，部分データ集合 P 中の P_1, P_2 だけに含まれていた場合，$f = 2$ である．すべての部分データのうち，f 個に含まれる確率は $p(f) = {}_N C_f \, s^f \, (1-s)^{N-f}$ である．次に，人物 X が含まれない部分データに対する匿名化意向における最大出現頻度のものを r_{max} とする．

パーソナルデータ D は，属性 $A_1, A_2, ...A_j$ とセンシティブ属性 $S = \{d_1, d_2, ...d_{|S|}\}$ から構成されている．攻撃者の目的は，特定の同値類 e_x（ただし $|e_x| = 1$）であるような人物 X のセンシティブ属性の値の推定である．なお，人物 X のセンシティブ属性の値は d_1 とする．

5.1 属性推定攻撃抑止　一般化

攻撃者は，同値類 e_x のセンシティブ値集合が $Y_{e_x} = \{d_1\}$ ならば匿名化意向を $r_1, Y_{e_x} = \{d_2\}$ ならば r_2,..., 同値類 e_x のレコードが存在しないなどそれ以外の場合は $r_{|S|+1}, ..., r_{z-1}, r_z$ のいずれかを一様分布で出力するプログラムを提供者に送付する．

提供者はパーソナルデータ D を N 回サンプリングし，N 個の部分データを作成する．この部分データを入力に，送付されたプログラムを実行する．人物 X が含まれる部分データ数は f，含まれない部分データ数は $N - f$ 個である．人物 X が含まれない部分データ数に対しては，プログラムに従って一様分布で $r_{|S|+1}, ..., r_{z-1}, r_z (z \leq (j+1)!)$ のいずれかが出力される．

人物 X のセンシティブ属性の値を推定可能な匿名化意向 r_1 の出現頻度は，$|r_1| = f$ である．出現頻度 $|r_{max}|$ について，場合分けして考える．

(1) $f < |r_{max}| \leq N - f$ のとき

最大出現頻度の匿名化意向は r_{max} である．そのため，$|r_1|$ が匿名化意向 r' として出力されることはない．ゆえに攻撃者は，人物 X のセンシティブ属性の値が d_1 であることを推定することができない．

(2) $|r_{max}| = f$ のとき

最大出現頻度の匿名化意向は r_1, r_{max} の二つであるため，処理を終了する．そのため，攻撃者は，人物 X のセンシティブ属性の値が d_1 であることを推定することができない．

(3) $|r_{max}| < f$ のとき

$|r_1|$ が最大出現頻度となる．このときの取りうる $|R|$ の最小値は，次の t の式の解を使って得ることができる．

$N - f = qt + u \ (1 \leq q < f, u < q, 1 \leq t)$

$u = 0$ のとき，$|R| = t + 1$，$u > 0$ のとき $|R| = t + 2$ である．ただし，q, t, u は整数である．

このとき，$|R| \geq L$ であれば，処理を終了する．そのため，攻撃者は，人物 X のセンシティブ属性の値が d_1 であることを推定することができない．

f	$p(f)$	$p(f)$の累積確率
0	0.107	0.107
1	0.268	0.376
2	0.302	0.678
3	0.201	0.879
4	0.088	0.967
5	0.026	0.994
6	0.006	0.999
7	0.001	1.000
8	0.000	1.000
9	0.000	1.000
10	0.000	1.000

表 2 f の値における $p(f)$ と $p(f)$ の累積確率．ただし，$N = 10, s = 0.2$ のとき．

5.2 属性推定攻撃抑止 具体例

具体的に，サンプリング率 $s = 0.2$, サンプリング回数 $N = 10$，匿名化意向の種類の上限 $L = 4$ のときを考える．パーソナルデータ D は，属性 A_1, A_2 とセンシティブ属性 $S = \{d_1, d_2\}$ から構成されているとする．このとき，$|e_x| = 1$ となるようなレコードをもつ人物 X が部分データ P_n に含まれる確率は，サンプリング率 $s = 0.2$ より 0.2 である．f の値ごとの $p(f)$ と $p(f)$ の累積確率を表 2 に示す．

最も起こりやすい $f = 2$ のときを想定して詳説する．なお，表 2 より $p(f = 2) = 0.302$ である．

攻撃者は，同値類 e_x のセンシティブ値集合が $Y_{e_x} = \{d_1\}$ ならば匿名化意向を $r_1, Y_{e_x} = \{d_2\}$ ならば r_2, 同値類 e_x のレコードが存在しないなどそれ以外の場合は一様分布で $r_3, r_4, ..r_{12}$ のいずれかを出力するプログラムを提供者に送付する．

提供者はパーソナルデータ D を N 回サンプリングし，N 個の部分データを作成する．この部分データを入力に，送付されたプログラムを実行する．$f = 2$ により，人物 X が含まれる部分データ数は 2，含まれない部分データ数は 8 である．

(1) $f < |r_{max}| \leq N - f$ のとき

つまり，$2 < |r_{max}| \leq 8$ のとき，$3 \leq |r_{max}| \leq 8$ である．

例えば，$|r_{max}| = 8$ のとき，存在しえる匿名化意向は r_1 と r_{max} である．ゆえに $|R| = 2$ である．$|R| < L$ であるため，処理を続行する．最大出現頻度の匿名化意向は r_{max} であるため，そのため，$|r_1|$ が匿名化意向 r' として出力されることはない．ゆえに攻撃者は，人物 X のセンシティブ属性の値が d_1 であることを推定することができない．

$|r_{max}| = 6$ の場合は，r_1, r_{max} 以外にもう一つか二つの匿名化意向が出力される．なぜなら匿名化意向候補は N 個出力されなければならない，すなわち匿名化意向の出現頻度の合計が N にならなければならないからである．このとき，$|R| = 4$ であれば，種類数の上限以上のため，処理が終了する．攻撃者は匿名データ D' を受け取ることができないため，人物 X のセンシティブ属性の値を推定することができない．

(2) $|r_{max}| = f$ のとき

つまり，$|r_{max}| = 2$ である．$N - f - |r_{max}| = 6$ より，少なくともう三つ他の匿名化意向が出力される．このとき存在しえる匿名化意向の種類数は $|R| \geq 5$ である．

$|R| \geq L$ であるため，処理は終了する．攻撃者は匿名データ D' を受け取ることができないため，人物 X のセンシティブ属性の値を推定することができない．

仮に，上限を下回っていれば，最大出現頻度の匿名化意向が少なくとも r_1, r_{max} の二つ以上であるため，処理は終了する．同様に，攻撃者は推定することができない．

(3) $|r_{max}| < f$ のとき

つまり，$|r_{max}| < 2$ のとき，すなわち $|r_{max}| = 1$ である．そのため，匿名化意向として，r_1 以外に 8 個の匿名化意向が出力される必要がある．このとき，$|R| = 9$ である．$|R| \geq L$ であるため，処理は終了する．

6. おわりに

本稿では，受領者が自身の目的に沿った有用性の高い匿名データを受領するための，プログラム送付型匿名化方式を提案した．これにより，受領者は元データの分析結果を用いて匿名化意向を生成できるため，従来より有用性の高い匿名データを受領可能である．さらに，提案方式ではプログラム送付型匿名化において発生しうる属性推定攻撃の成功確率を低減し，安全性を向上した．本稿では具体的にセル抑制による匿名化を例に，その抑止の仕組みについて詳説した．一般化の匿名化でも同様の抑止が可能である．

今後の課題として，次の二つがある．

- 属性の値の分布傾向をふまえた場合の攻撃と抑止方法の検討
- スワッピングやノイズなど他の匿名化手法を用いた場合の攻撃と抑止方法の検討

参考文献

[1] Fung, Benjamin C M. Wang, K. Chen, R. Yu, Philip S.: "Privacy-preserving data publishing : A survey of recent developments." In: ACM Computing Surveys, Vol. 42, No. 4, 2010.

[2] 南 和宏:プライバシー保護データパブリッシング, 情報処理, 54(9), pp.938–946, 2013.

[3] 小林良行:公的統計ミクロデータ提供の現状と展望——一橋大学での取り組みをもとに，日本統計学会誌，41(2),401–420, 2016.

[4] 伊藤伸介：諸外国における政府統計データの提供の動向について，中央大学経済研究所 Discussion Paper No. 267,2016.

[5] UNECE.: "Managing statistical confidentiality and microdata access: Principles and guidelines of good practice." 2007.

[6] Latanya Sweeney.: "k-anonymity: A model for protecting privacy." Interna- tional Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, Vol. 10, No. 05, pp. 557-570, 2002.

[7] Samarati, Pierangela and Sweeney, Latanya.: "Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and suppression." Technical report, SRI International,1998

[8] Tian, Hongwei, and Weining Zhang.: "Privacy-Preserving Data Publishing Based on Utility Specification." Social Computing (SocialCom), 2013 International Conference on. IEEE, 2013.

[9] Yamaoka, Yuji, and Kouichi Itoh. "k-presence-secrecy: Practical privacy model as extension of k-anonymity." IEICE TRANSACTIONS on Information and Systems 100(4) ,pp.730–740, 2017.