Randomized Responseを用いた柔軟な匿名データ収集

清　雄一[†a)]　大須賀昭彦[†b)]

Flexible Anonymized Data Collection with Randomized Response Scheme

Yuichi SEI[†a)] and Akihiko OHSUGA[†b)]

あらまし　ユーザがカテゴリー化された自身のデータを改変してサーバに送信し，サーバは得た情報から統計的な解析を行う，というプライバシー保護モデルを実現するRandomized Responseスキームが提案されている．サーバ側は受け取った情報から，各カテゴリーに属すユーザ数の真の分布を推測する．各ユーザの真のカテゴリーがどのカテゴリーに改変されてサーバへ送信されるかは，あらかじめ設定された確率行列に基づいて決定される．確率行列の値を変更することで，異なるプライバシー保護レベルを実現できる．また，プライバシー保護レベルと，サーバにおける推測誤差とはトレードオフの関係にある．従来は，全ユーザが同一の確率行列を利用する状況のみが想定されており，ユーザごとにプライバシー保護レベルを変えることができないという制約があった．本論文では，ユーザごとに異なる確率行列を利用するモデルを提案する．異なる確率行列が利用される場合，サーバ側において各カテゴリーに属すユーザ数の分布を推測する手法は確立されていない．本論文では推測誤差を定量的に取扱い，最も確からしいユーザ数の分布を推測する手法を提案する．従来手法と比較してサーバ側での推測誤差を70%程度削減できることを，数学的解析及び実データを用いたシミュレーションによって示す．

キーワード　ユビキタスコンピューティング，プライバシー，匿名化

1. まえがき

ユビキタスコンピューティング技術の発展により，様々なユーザ情報を収集しマイニングを行う研究が盛んになっている[1], [2]．しかしユーザ情報をそのまま収集するとプライバシー情報の漏洩リスクがあるため，ユーザの属性データをカテゴリー化し，一定の確率でユーザが属さないカテゴリーからランダムに選択されたカテゴリー情報を収集するRandomized Responseスキーム（RR）が提案されている[3]～[6]．

RRは，取り得るユーザの属性値の範囲が0から99であるとしたとき，0から9までをカテゴリーC_1，10から19までをカテゴリーC_2のように表現する．例えば，あるユーザの属性値が37である場合，カテゴリーはC_4である．このとき，RRではある確率pでC_4をサーバへ報告するが，$1-p$の確率でC_4以外の任意のカテゴリーをサーバへ報告する．サーバへ報告された情報からは，ユーザが属すカテゴリーを正確に判定できないため，プライバシー情報を一定レベルで保護することができる．一方で，多くのユーザから情報を収集することで，各カテゴリーに属すユーザ数の概算値を推測することができる．

RRではこのように，ユーザの真のカテゴリーがC_iであるとき，ある確率$p_{j,i}$でカテゴリーC_jを選択してサーバへ報告する．各i, jに対して$p_{j,i}$の値を設定したものを**確率行列**と呼ぶ．確率行列をどのように設定するかによって，プライバシー保護レベルと各カテゴリーに属すユーザ数の推測誤差が変動する．一般にこれらはトレードオフの関係にある．

従来のRRでは，全ユーザが同じ確率行列を使用する状況が想定されていた．したがって，全ユーザについてプライバシー保護レベルを同一に設定する必要があった．しかし，ユーザによってはプライバシー保護レベルを高めたいという要望もあると考えられ，ユーザごとに異なるプライバシー保護レベルを設定できることが望ましい（プライバシー保護レベルの指標については**3. 2**において述べる）．

従来研究においては，ユーザごとに確率行列が異なる場合において各カテゴリーに属すユーザ数を推測す

† 電気通信大学大学院情報システム学研究科，調布市
Graduate School of Information Systems, The University of Electro-Communications, Chofu-shi, 182–8585 Japan
a) E-mail: sei@is.euc.ac.jp
b) E-mail: ohsuga@euc.ac.jp

る手法が確立されていない．本論文では，このような状況においても推測できる手法を提案する．

本論文の構成を示す．**2.** では，本論文が想定するモデルを述べる．**3.** では，RR に関する既存研究について記述する．**4.** において，本論文が提案する手法を記述する．**5.** では，提案手法のプライバシー及び推測誤差について解析を行う．**6.** では，提案手法と既存手法の比較を，数学的解析及びシミュレーションによって実施する．**7.** において考察を示す．**8.** で本論文のまとめを記す．

2. 想定モデル

2.1 プライバシーモデル

本論文で想定するプライバシーモデルを述べる．各ユーザが自分の情報を開示することによってサーバに与える情報をプライバシー情報と定義する．言い換えると，ユーザが RR に参加しているかどうかにかかわらず，サーバが当該ユーザに関して推測できるような情報はプライバシー情報とはみなさない．

一般的なアンケート調査を例に挙げて説明する．C_1 を 0〜1000 万円，C_2 を 1000 万円〜2000 万円，C_3 を 2000 万円〜3000 万円，のようにカテゴリーを定義して行う給料についてのアンケート調査を考える．あるユーザ A の回答が「C_1 か C_2 のいずれか」であり，ユーザ B は未回答であったとする．ユーザ A としては，「C_1 か C_2 のいずれか」という情報しか開示しておらず，ユーザ B は何の情報も開示していない．しかし，その他ほぼ全てのユーザが「C_1 である」と回答した場合，ユーザ A やユーザ B についてのカテゴリーも「高い確率で C_1 である」と推測することができる．しかしアンケートに回答していないユーザの情報が，その他多くのユーザの回答結果から推測されたとしても，通常はプライバシー情報の漏洩とはみなされないと考えられ，本論文においてはこのようなプライバシーモデルを想定する．

このプライバシーモデルのフォーマルな定義は次のとおりである．これは，Evfimievski ら [7] や Kasiviswanathan ら [8] が想定するモデルと同一であり，本論文では彼らのプライバシーモデルをそのまま採用する．

各ユーザ u の真のカテゴリー C_u は，全てのユーザで共通の確率分布から独立にランダムに選択されたものとみなす．この確率分布を p_c とおくと，この p_c 自体はプライバシー情報ではなく，サーバが p_c を知ることをユーザは許容する．言い換えると，ユーザ u を除く全てのユーザについて真のカテゴリーの情報が得られたとしても，その情報とユーザ u の真のカテゴリー C_u とは独立しているため，C_u に対する推測には何の影響も与えない．

このモデルに基づいて，プライバシー保護レベルを具体的な数値として表すプライバシー指標は，**3.2** において述べる．

2.2 想定環境

ユビキタスコンピューティング環境においてユーザのデータを取得し，取得情報をサーバへ送信する．この情報を基に，あらかじめ設定された各カテゴリーに属すユーザ数を把握することを目的としたサービスを想定する．取得するユーザのデータは，Public Health [9] における年齢，性別，人種，体重や病名，匿名交通モニタリングにおける自動車の速度 [10] 等が考えられる．

また，ユーザは RR に参加することによって，サーバから何らかの特典（ポイント等）を得ることができると想定する．調査に協力することによって特典を得ることができるような想定は多くの研究でも採用されており，ユーザの動機付けに関する研究も盛んに行われている [11], [12]．更に，プライバシー保護レベルを下げると，ユーザの効用は減少すると想定する．一方，より多くの特典を得ると，ユーザの効用が増加すると想定する．各ユーザにおけるこれらの効用関数はそれぞれ異なる．各ユーザは，特典を得ることによる効用の増加と，プライバシー保護レベルを下げることによる効用の減少とのトレードオフを取り，プライバシー保護レベルを決定する．**2.1** で定義したプライバシーモデルの下では，プライバシー保護レベルを下げることによる効用の減少量は，当該ユーザがどれだけ自分の情報を正しくサーバに伝えるかのみに依存する．つまり，サーバにおける，各カテゴリーに属すユーザ数の推測精度には依存しないことに注意されたい．

また RR を実施するサーバ側は，各カテゴリーに属すユーザ数の推測誤差が下がると効用が増加し，ユーザに多くの特典を与えるほど効用が減少すると想定する．そのトレードオフの関係から，「プライバシー保護レベルと特典付与量の関係式」を決定する．これは，ユーザがプライバシー保護レベルをどの程度に設定すると，特典をどれだけ与えるかについての関係を表す式である．ユーザ及びサーバにおける具体的な効用関数やこの関係式の決定方法については将来課題とし，

本論文のスコープ外とする．

RR によってユーザ属性を収集される対象となるユーザは，スマートフォンの電子機器を身に付けており，その電子機器に，ユーザ属性やプライバシーに関するユーザの要望を管理するユーザエージェントが存在すると想定する [13], [14]．このユーザエージェントはサーバから，収集対象の属性名（例えば「年齢」），カテゴリー定義（例えば 0 歳〜9 歳を C_1，10 歳〜19 歳を C_2，等）及び，プライバシー保護レベルと特典付与量の関係式の情報を受け取り，ユーザにとって適切なプライバシー保護レベルを決定することができる．ユーザエージェントはサーバに対し，ユーザ属性を基に算出したカテゴリー ID 及び，利用した確率行列（確率行列の生成方法は **4. 1** において述べる）を通知する．

サーバ側は，各ユーザエージェントに対して，収集対象の属性名，カテゴリー定義及び，プライバシー保護レベルと特典付与量の関係式を通知し，算出されたカテゴリー ID 及び利用した確率行列の通知を受け取ることができる．また，**4. 2** に述べる手法を用いることで，各カテゴリーに属すユーザ数を推測する機能を有する．

3. 関 連 研 究

ユーザが自身のデータを改変してサーバに送信し，サーバは得た情報から解析を行う，というプライバシー保護モデルはローカルモデルと呼ばれる [8]．本章では，ローカルモデルの代表的手法である Randomized Response (RR) 及び，プライバシー指標を記述する．

3. 1 Randomized Response (RR)

3. 1. 1 基本アルゴリズムと推測誤差

カテゴリー数を F とし，それぞれ $C_1, C_2, \ldots, C_F$ と表す．ユーザはいずれかのカテゴリーに属し，これを **True Category (TC)** と呼ぶ．あるユーザの True Category が C_i であるとき，ある確率で C_i 以外のカテゴリーを選択し，サーバへ報告する．サーバへ報告するカテゴリーを **Negative Category (NC)** と呼ぶ．TC が C_i であるとき，C_j を NC として選択する確率を $p_{j,i}$ とし，確率行列をあらかじめ設定しておく．確率行列は，対角成分を $p_{i,i} = p$，対角成分以外の全 $i \neq j$ における各成分を同一の値に設定する Uniform Perturbation が広く利用されている [3], [4], [15], [16]．この確率行列は以下のように表すことができる．

$$M(p) = \begin{pmatrix} p & \frac{1-p}{F-1} & \frac{1-p}{F-1} & \cdots \\ \frac{1-p}{F-1} & p & \frac{1-p}{F-1} & \cdots \\ \frac{1-p}{F-1} & \frac{1-p}{F-1} & p & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{1}$$

TC が実際に C_i であるユーザ数を x_i とおく．これは未知の値であり，サーバ側で推測したい値である．サーバに NC を報告したユーザ数を N とし，NC として C_i が報告された数を y_i とする．TC が C_i であると推測されるユーザ数を $\widehat{x_i}$ と表すと，

$$\begin{aligned} &\hat{x} = M(p)^{-1} y^{\tau} \\ &\text{where } \hat{x} = \{\widehat{x_1}, \ldots, \widehat{x_F}\},\ y = \{y_1, \ldots, y_F\} \end{aligned} \tag{2}$$

と計算することができる．ここで，$M(p)^{-1}$ は行列 $M(p)$ の逆行列，y^{τ} は行列 y の転置行列を表す．また，$\widehat{x_i}$ は x_i の不偏推定量となる [4]．

確率行列 $M(p)$ を利用した際の推測誤差として MSE（Mean Square Error）を計算する．カテゴリー C_i が TC であると推測されるユーザ数の MSE を σ_i^2 で表し，次式で定義する．

$$\sigma_i^2 = (\frac{\widehat{x_i}}{N} - \frac{x_i}{N})^2 \tag{3}$$

TC が C_i であるユーザ数の割合を $P(X = C_i)$，NC が C_i であるユーザ数の割合を $P(Y = C_i)$，式 (2) を用いて TC が C_i であると推測されるユーザの割合を $P(\widehat{X = C_i})$ とすると，C_i における推測誤差の期待値 $E[\sigma_i^2]$ は以下の式で表すことができる [4]．

$$\begin{aligned} E[\sigma_i^2] &= E(P(\widehat{X = C_i}) - P(X = C_i))^2 \\ &= \sum_{j=1}^{F} \left(M(p)_{i,j}^{-1}\right)^2 Var\left(\frac{N_j}{N}\right) \\ &+ \sum_{\substack{j,k \\ j \neq k}}^{F} 2 \cdot M(p)_{i,j}^{-1} M(p)_{i,k}^{-1} Cov\left(\frac{N_j}{N}, \frac{N_k}{N}\right), \end{aligned} \tag{4}$$

where

$$\begin{aligned} Var(\frac{N_j}{N}) &= \frac{1}{N} \cdot P(Y = C_j)\left(1 - P(Y = C_j)\right), \\ Cov(\frac{N_j}{N}, \frac{N_k}{N}) &= -\frac{1}{N} \cdot P(Y = C_j) P(Y = C_k) \end{aligned}$$

ここで，$M(p)_{i,j}^{-1}$ は，行列 $M(p)$ の逆行列 $M(p)^{-1}$ における i 行 j 列目の要素を表す．また，期待値算出時に $P(Y = C_i)$ の値が不明である場合，$P(Y = C_i) =$

$1/F$ $(i=1,\ldots,F)$ と設定する．

3.1.2 RR における既存手法

Huang ら [4] は，推測誤差を小さくできる確率行列を遺伝的アルゴリズムを用いることによって求める手法を提案している．

Agrawal ら [15] は，確率行列の対角成分を各ユーザで異なる値に設定するモデルを提案しており，本論文のモデルと近いが，以下の点で異なる．Agrawal らの手法では，確率行列の対角成分の基準値を全ユーザ共通で p' と設定し，各ユーザにおける確率行列の対角成分を $p'+r$ と設定する．ここで，r はある定数 a に対し，$-a$ から a までの範囲でランダムな値が設定される．各ユーザで r をどのように設定したかサーバに伝えないことにより，各ユーザのプライバシー保護レベルを向上させることを目指している．サーバ側では，全ユーザが同じ確率行列 $M(p')$ を利用していると想定して推測を行っており，ユーザごとにプライバシー保護レベルを変えることはできない．本論文の提案手法を用いることによって，確率行列の対角成分の基準値 p' をプライバシー保護レベルの異なるユーザごとに変えることが可能となる．したがって，Agrawal らの手法と本提案手法を純粋に組み合わせることによって，プライバシー保護レベルと推測誤差のトレードオフをより向上させることができると考えられる．

NC を通知するサーバが複数存在するときに，サーバ同士が共謀した場合でもプライバシーを保護する手法も提案されている [17], [18]．このモデルにおいても，ユーザごとに異なる確率行列が利用されることは想定されていない．

RR の特殊なケースとして，確率行列として $M(0)$ を用いる Negative Survey と呼ばれる手法 [5], [6], [19], [20] も多数提案されている．これらの研究においても，さまざまなプライバシー指標に応じた確率行列が提案されているが，ユーザごとにプライバシー保護レベルを設定可能な手法はない．本論文が提案する手法を併用することによって，ユーザごとにプライバシー保護レベルを変更可能な手法に拡張可能であると考えられる．

3.2 プライバシー指標

ローカルモデルで広く利用されているプライバシー指標に ρ_1-to-ρ_2 プライバシーがある [7]．これはローカルモデルの一つである RR の分野でも広く利用されている [15]〜[17]．RR の分野においては，1 以上の実数 γ を用意し，確率行列 $M(p)$ が以下を満たせば指定された γ に対してプライバシー情報の保護がなされていると定義することができる [15]．

$$\frac{M(p)_{j,i_1}}{M(p)_{j,i_2}} \leq \gamma, \quad \forall i_1, i_2, j \in \{1,\ldots,F\} \tag{5}$$

直観的には，あるユーザの NC から推測される当該ユーザの TC として，最も可能性の高いカテゴリーと最も可能性の低いカテゴリーにおける可能性の比がプライバシー指標となっている．Kasiviswanathan ら [8] は，ρ_1-to-ρ_2 ではなく差分プライバシー [21] に基づいたプライバシー指標を利用しているが，同様の指標である．

以降では，プライバシー指標として γ の値を用い，各ユーザが γ の値を設定し，プライバシー保護レベルを調整することができると想定する．

4. 提案手法

ユーザが設定するプライバシー保護レベルを満たした上で，サーバ側での推測誤差を最小化することが期待される確率行列の構築方法を **4.1** で述べる．次に，各カテゴリーに属すユーザ数の真の分布を推測するためのアルゴリズムを **4.2** で提案する．主に利用されるパラメータを表 1 に示す．

4.1 プライバシー保護レベルに基づいた確率行列の構築

4.1.1 では，ユーザが設定するプライバシー保護レベル γ を満たすための p が取り得る値の範囲を特定し，**4.1.2** では取り得る値の範囲の中で最もサーバ側での推測誤差を小さくすることができる p の値を一意に特定する．

本論文では，確率行列として **3.** で述べた Uniform Perturbation が利用されることを想定して議論を進める．

表 1 Notation
Table 1 Notation.

F	ユーザ属性のカテゴリー数（$2 \leq F$）
N	ユーザ数
C_i	ID が i であるカテゴリー
γ	ユーザが設定可能なプライバシー保護レベル（$1 \leq \gamma$）
x_i	TC が C_i である実際のユーザ数
y_i	NC が C_i であるユーザ数
$\widehat{x_i}$	TC が C_i であると推測されたユーザ数
$M(p)$	対角成分が p，それ以外の成分が $(1-p)/(F-1)$ である行列
$M^{-1}_{s,t}$	行列 M の逆行列における s 行 t 列目の値
δ	全ユーザにおける確率行列の対角成分のユニーク数

4.1.1 γ を満たす p の値の範囲算出

式 (5) より，確率行列の対角成分である p は，プライバシー保護レベル γ が指定された場合，以下の式が満たされる必要がある．

$$\frac{1}{\gamma} \leq \frac{p}{(1-p)/(F-1)} \leq \gamma \tag{6}$$

したがって，p の取り得る値の範囲は，$1 \leq \gamma$，$2 \leq F$ を考慮すると，

$$\frac{1}{1+\gamma(F-1)} \leq p \leq \frac{\gamma}{\gamma+F-1} \tag{7}$$

である．

4.1.2 推測誤差を最小化する p の値算出

式 (7) の範囲で最も推測誤差の期待値を小さくできる p の値を決定する．式 (4) において，$P(Y=C_i)=1/F$ $(i=1,\ldots,F)$ としたときの推測誤差は，後述の **4.2.2** で導出する式 (20) で表される．p について偏微分すると，

$$-\frac{2(F-1)^3(F+1)}{F^2 N_k (F\cdot p-1)^3} \tag{8}$$

となる．したがって，$p < 1/F$ までの範囲では p の値が大きくなるほど推測誤差は増加し，$1/F < p$ の範囲では p の値が大きくなるほど推測誤差は減少する．また，式 (20) に対して，$p = 1/F+\delta$ と設定した場合も，$p = 1/F-\delta$ と設定した場合も，等しい値を取るため(注1)，推測誤差は $p = 1/F$ で対称な関数となる．以上より，式 (7) を満たす範囲の中で，$1/F$ との差の絶対値が最も大きい値を選ぶことになる．これは，$1 \leq \gamma$ かつ $2 \leq F$ の場合は，

$$p = \frac{\gamma}{\gamma+F-1} \tag{9}$$

である．

4.2 各カテゴリーに属すユーザ数の推測

4.2.1 推測アルゴリズム

サーバ側で，収集した NC の情報から各ユーザの真の TC の分布を推測する手法を提案する．推測アルゴリズムの流れを図 1 に示す．各ユーザ t における確率行列の対角成分を p_t とおく．各 p_t $(t=1,\ldots,N)$ が等しいユーザをグループ化し，各グループを $S_1, S_2, \ldots, S_\delta$ とする（図 1 の左端における各ユーザグループ）．また各 S_k $(k=1,\ldots,\delta)$ に属すユーザ数を N_k とおく．

(注1)：具体的には $-(F-1)(F(1+(1+\delta^2-F)F)-1)/(\delta^2 F^5 N_k)$ となる．

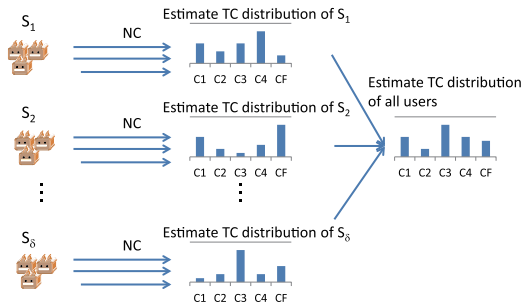


図 1　TC の分布を推測する流れ
Fig. 1　Flow of estimation of TC distribution.

各グループ S_k の中において，TC が C_i である実際のユーザ数を $x_{k,i}$，C_i を NC として報告したユーザ数を $y_{k,i}$，TC が C_i であると推測されるユーザ数を $\widehat{x_{k,i}}$ とおく．

各グループ S_k の中において，式 (2) を用いて $\widehat{x_{k,i}}$ $(i=1,\ldots,F)$ を算出する（図 1 における “Estimate TC distribution of S_k”）．

最後に，あるカテゴリー C_i を対象として算出された各 $\widehat{x_{k,i}}$ $(k=1,\ldots,\delta)$ に注目し，最終的な $\widehat{x_i}$ を算出する（図 1 における “Estimate TC distribution of all users”）．$\widehat{x_{k,i}}$ は，S_k の中において TC が C_i であると推測されたユーザ数であるから，簡単な方法としては全てを単純に合算し，

$$\widehat{x_i} = \sum_{k}^{\delta} \widehat{x_{k,i}} \tag{10}$$

とすることもできる．しかし，各 $\widehat{x_{k,i}}$ はそれぞれ異なる推測誤差の期待値をもつため，推測誤差が小さいものをより重視して足し合わせることで，$\widehat{x_i}$ の最終的な値の誤差を低減することができると考えられる．

S_k のユーザ群において，TC が実際に C_i であるユーザの割合を $P(X_k=C_i)$，NC として C_i を報告したユーザの割合を $P(Y_k=C_i)$，式 (2) を用いて TC が C_i であると推測されたユーザの割合を $P(\widehat{X_k=C_i})$ とおく．TC が C_i であると推測されたユーザ数の推測誤差を $\sigma^2_{k,i}$ とおくと，その期待値は

$$E[\sigma^2_{k,i}] = E[(P(\widehat{X_k=C_i}) - P(X_k=C_i))^2] \tag{11}$$

である．

ここで，$P(\widehat{X_k=C_i})$ は $P(X_k=C_i)$ の不偏推定量であるから [4]，$E[P(\widehat{X_k=C_i})] = P(X_k=C_i)$ である．したがって，$E[\sigma^2_{k,i}]$ は $P(\widehat{X_k=C_i})$ の分散を表

している．

ここで，不偏推定量 $P(\widehat{X_k = C_i})$ と真実の値 $P(X_k = C_i)$ の差が正規分布に従うと仮定すると，$P(\widehat{X_k = C_i}) = \widehat{x_{k,i}}/N_k$ となる確率 $f(\widehat{x_{k,i}})$ は次式で表される，

$$f(\widehat{x_{k,i}}) = \frac{1}{\sqrt{2\pi}\sigma_{k,i}} \exp(-\frac{1}{2\sigma_{k,i}^2}(\frac{\widehat{x_{k,i}}}{N_k} - \frac{x_{k,i}}{N_k})^2) \tag{12}$$

また，式 (12) における $x_{k,i}/N_k$ は，TC が C_i であるユーザの割合を示している．これはどの S_k においても同一であると仮定すると，

$$\frac{x_{k,i}}{N_k} = \frac{x_i}{N} \tag{13}$$

とおくことができる．式 (12)，(13) より，C_i について，一組の $\widehat{x_{k,i}}$ $(k = 1, \ldots, \delta)$ が得られる確率は次式で表される．

$$\begin{aligned} &\prod_{k=1}^{\delta} f(\widehat{x_{k,i}}) \\ &= (\prod_{k=1}^{\delta} \frac{1}{\sqrt{2\pi}\sigma_{k,i}}) \exp(-\sum_{k=1}^{\delta} \frac{1}{2\sigma_{k,i}^2}(\frac{\widehat{x_{k,i}}}{N_k} - \frac{x_i}{N})^2) \end{aligned} \tag{14}$$

式 (14) を最大化する x_i が最も確からしい x_i と言える．

$$\sum_{k=1}^{\delta} \frac{1}{2\sigma_{k,i}^2}(\frac{\widehat{x_{k,i}}}{N_k} - \frac{x_i}{N}) = 0 \tag{15}$$

が成り立つとき式 (14) は最大化される．このときの x_i の値を $\hat{x_i}$ とおくと，

$$\widehat{x_i} = N \cdot \frac{\sum_{k=1}^{\delta}(1/\sigma_{k,i}^2)\widehat{x_{k,i}}/N_k}{\sum_{k=1}^{\delta} 1/\sigma_{k,i}^2} \tag{16}$$

と求めることができる．

4.2.2 推測アルゴリズムの簡素化

TC がカテゴリー C_i であると推測されるユーザ数を計算するための式 (16) は，$\sigma_{k,i}^2$ 及び $\widehat{x_{k,i}}$ の計算が必要である．それぞれ式 (4) 及び式 (2) に基づいて計算することもできるが，計算量が大きいため次のように簡素化された計算式を利用する．

式 (1) より，確率行列の対角成分が p_k である場合，確率行列 $M(p_k)$ の逆行列は

$$M(p_k)^{-1} = \begin{pmatrix} p_1 & p_2 & p_2 & \cdots \\ p_2 & p_1 & p_2 & \cdots \\ p_2 & p_2 & p_1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$
$$\text{where} \quad p_1 = \frac{p_k + F - 2}{F \cdot p_k - 1}, \ p_2 = \frac{p_k - 1}{F \cdot p_k - 1} \tag{17}$$

である．したがって，任意の整数 t $(1 \le t \le F)$ に対して，

$$\begin{aligned} &\sum_{i=1}^{F} (M(p_k)_{t,i}^{-1})^2 = p_1^2 + (F-1)p_2^2 \\ &= \frac{F(F + p_k^2 - 3) - 2p_k + 3}{(F \cdot p_k - 1)^2} \end{aligned} \tag{18}$$

であり，また，

$$\begin{aligned} &\sum_{i,j,\, i \ne j}^{F} M(p_k)_{t,i}^{-1} M(p_k)_{t,j}^{-1} \\ &= 2(F-1)p_1p_2 + (F-1)(F-2)p_2^2 \\ &= \frac{(F-1)(p_k-1)(F \cdot p_k + F - 2)}{(F \cdot p_k - 1)^2} \end{aligned} \tag{19}$$

が成り立つ．

確率行列の対角成分が p_k であるユーザ数を N_k とおく．任意の i, j について $P(Y = C_i) \approx P(Y = C_j)$ と仮定すると，任意の i について $\sigma_{k,i}$ の値は等しい．式 (4)，(18)，(19) より，任意の i について，

$$\sigma_{k,i}^2 = \frac{(F-1)(F^2 + 2p_k - F(1 + p_k^2) - 1)}{F^2 N_k (F \cdot p_k - 1)^2} \tag{20}$$

と表すことができる．

また，式 (2) 及び式 (17) から，$\widehat{x_{k,i}}$ を次式で求めることができる．

$$\begin{aligned} \widehat{x_{k,i}} &= \sum_{j=1}^{F} M(p_k)_{i,j}^{-1} Y_{k,j} \\ &= \frac{p_k + F - 2}{F \cdot p_k - 1} \cdot Y_{k,i} + \frac{p_k - 1}{F \cdot p_k - 1} \cdot (N_k - Y_{k,i}) \end{aligned} \tag{21}$$

これらの結果を用いて，NC から TC の分布を推測するアルゴリズムを Algorithm 1 のように記述することができる．

Algorithm 1 Estimate TC distribution

Input: # of categories F, Users U
Output: TC distribution of all users
Creates Hashtables $users$, $dists$, $mses$
Creates Array $rslts$[]
/*Divides users based on their diagonal elements*/
for $u \in U$ **do**
  $users.put(getP(u), users.get(p) \cup \{getP(u)\})$
end for
/*Estimates TC distribution of each group of users*/
for $p \in users.keys()$ **do**
  $dists.put(p, calcDist(F, p, users.get(p)))$
  $mses.put(p, calcMSE(F, p, users.get(p)))$
end for
/*Estimates TC distribution of all users*/
for $p \in users.keys()$ **do**
  $denom \Leftarrow denom + \frac{1}{mses.get(p)}$
end for
for $i = 1, \ldots, F$ **do**
  for $p \in users.keys()$ **do**
    $rslts[i] \Leftarrow rslts[i] + \frac{1}{mses.get(p)} \times \frac{dists.get(p)[i]}{|users.get(p)|}$
  end for
  $rslts[i] \Leftarrow |U| \times \frac{rslts[i]}{denom}$
end for
return $rslts$

Algorithm 2 calcDist

Input: # of categories F, Diagonal element p, Users S_k
Output: Estimated TC distribution in S_k
Creates Array $rslts_k$[]
for $i = 1, \ldots, F$ **do**
  $rslts_k[i] \Leftarrow$ result calculated by Eq. 21
end for
return $rslts_k$

Algorithm 3 calcMSE

Input: # of categories F, Diagonal element p, Users S_k
Output: Expected MSE in S_k
Creates Variable mse_k
$mse_k \Leftarrow$ result calculated by Eq. 20
return mse_k

ここで，Hashtable について，put(Key, Value) は指定された Key と Value を格納する関数であり，get(Key) は指定された Key に対応する Value が存在する場合はその Value を返し，対応する Value がない場合は空集合を返す関数である．また keys() は格納されている Key の集合を返す関数である．また，$getP(u)$ はユーザ u の確率行列における対角成分 p の値を返す関数である．

5. 推測誤差の数学的解析

提案手法によって推測された各 TC に属すユーザ数と，本来のユーザ数との推測誤差を解析する．

S_k のユーザ群における推測誤差は，式 (20) で表されるとおりである．最終的な推測値を表す式 (16) は，確率行列の対角成分が異なる全ての S_k の結果を合算しているため，推測誤差も合算する必要がある．

異なる大きさの誤差をもつ値を合算した際の最終的な誤差は，Propagation of Error Formula [22] を利用して計算することができる．Propagation of Error Formula によると，$z = f(t_1, t_2, \ldots, t_\delta)$ で表される関数 z があり，各 t_k の分散が $\sigma_{t_k}^2$ で表されるとき，z の分散 σ_z^2 は次の式で表される．

$$\sigma_z^2 = \sum_{k=1}^{\delta} \left(\frac{\partial z}{\partial t_k}\right)^2 \sigma_{t_k}^2 \tag{22}$$

提案手法では，z に相当するものが $\widehat{x_i}/N$ であり，t_i に相当するものが $\widehat{x_{k,i}}/N_k$ である．したがって式 (16)，式 (22) より，カテゴリー C_i に対する推測誤差は以下の式で表される．

$$\sigma_i^2 = \sum_{k=1}^{\delta} \left(\frac{\partial \frac{\widehat{x_i}}{N}}{\partial \frac{\widehat{x_{k,i}}}{N_k}}\right)^2 \sigma_{k,i}^2 \tag{23}$$

ここで，$\widehat{x_i}/N$ を $\widehat{x_{k,i}}/N_k$ で偏微分すると，式 (16) より，

$$\frac{\partial \frac{\widehat{x_i}}{N}}{\partial \frac{\widehat{x_{k,i}}}{N_k}} = \frac{1/\sigma_{k,i}^2}{\sum_{k=1}^{\delta} 1/\sigma_{k,i}^2} \tag{24}$$

である．したがって式 (23)，(24) より，

$$\sigma_i^2 = \sum_{k=1}^{\delta} \left(\frac{1/\sigma_{k,i}^2}{\sum_{k=1}^{\delta} 1/\sigma_{k,i}^2}\right)^2 \sigma_{k,i}^2 = \frac{1}{\sum_{k=1}^{\delta} 1/\sigma_{k,i}^2} \tag{25}$$

となる．

また，同様の議論により，$\widehat{x_i}$ を式 (10) を用いて計算した場合の推測誤差は，

$$\sum_{k=1}^{\delta} N_k^2 \sigma_{k,i}^2 / N^2 \tag{26}$$

となる．

6. 評　　価

式 (25) で定義した推測誤差（MSE）の数学的評価及び，実際に TC から NC を生成して推測誤差を計算するシミュレーション評価を行った．推測誤差の算出に関しては，各カテゴリーにおける推測誤差を計算し，その平均値を取った．

比較対象として次の 2 手法を用意した．比較対象の一つは，RR に参加するユーザの中で最もプライバシー保護レベルが厳しいユーザに合わせて確率行列を作成し，全ユーザでその共通の確率行列を利用する手法である．この手法を **Simple** 手法と呼ぶことにする．もう一つの比較対象は，各ユーザのプライバシー保護レベルに応じて異なった確率行列を利用するが，各カテゴリーに属すユーザの推測時に提案手法である式 (16) ではなく簡易的な式 (10) を利用する手法である．この手法を **NaiveAdding** 手法と呼ぶことにする．

[7] では，ρ_1-to-ρ_2 プライバシーがプライバシー指標として用いられており，γ の値は 9 から 99 までの値を取るよう設定されている．これに基づき本評価においては，最もプライバシー保護レベルが厳しいユーザは $\gamma = 10$，最も寛大なユーザは $\gamma = 100$ とし，各ユーザは γ の値を $10, 20, \ldots, 100$ までの 10 段階で設定できると想定する．RR に参加するユーザが設定する γ の値が，10 から 100 までの間で一様に分布している状況を想定して評価を行った．

6.1 数学的解析による評価

Simple 手法，NaiveAdding 手法，提案手法の各手法で，ユーザ数 $N = 1{,}000$ とし，カテゴリー数 F を 10 から 300 まで変化させて推測誤差を計算した結果を図 2 に示す．Simple 手法は，プライバシー保護レベルが最も厳しいユーザに合わせているため，プライバシー保護レベルを緩和してもよいと考えているユーザのプライバシーを過剰に保護している．したがって，サーバにおける推測誤差の値が大きい．NaiveAdding 手法は，プライバシー保護レベルが異なるユーザごとに分けて TC のユーザ分布を推測しているが，最後に誤差に応じた合算をしていない．そのため，推測誤差が Simple 手法より小さく，提案手法より大きい．

全体として，F の値が増加するほど推測誤差は増加している．これは，式 (9) で算出される確率行列の対角成分 p の値が，F が増加するほど減少し，$1/F$ に近づくためである．式 (8) より，p の値が $1/F$ に近づくほど推測誤差が増加することが分かっている．一方，F が増加すると，式 (16) で表される各 $\widehat{x_i}$ の値が相対的に小さくなる．したがって，F の増加は推測誤差の値が小さくなる方向にも作用している．この二つの影響により，F の値の増加によって推測誤差が増減していると考えられる．

次に，カテゴリー数 $F = 50$，ユーザ数 $N = 100, 1000, 10000$ に設定して推測誤差を計算した結果を図 3 に示す．いずれの手法も N の増加に応じて推測誤差の値が減少していることが分かる．また，どの N をとっても，提案手法の推測誤差が最も小さい値を実現していることが分かる．

6.2 シミュレーション評価

6.2.1 実データを用いた評価

UCI の Adult データセット [23] を用いて評価を行った．本データセットは，匿名化手法における研究分野で広く利用されている [24], [25]．45,222 件のデータのうち，1,000 件を抽出して評価を行った．また Adult データセットの属性のうち，ユビキタスコンピューティング環境で推測可能だと考えられる年齢，人種の 2 属性を利用した．年齢は 15-19，20-24 のように 5 歳ず

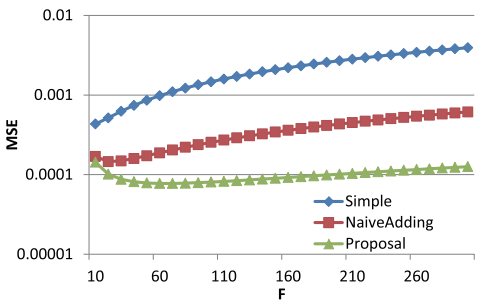


図 2　F を変化させたときの推測誤差
Fig. 2　MSEs with different F.

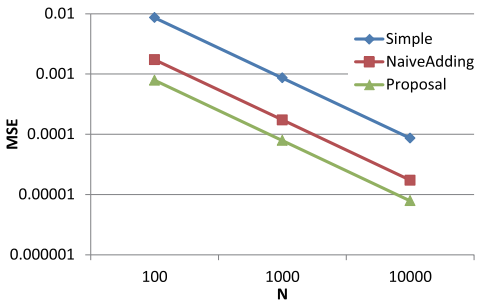
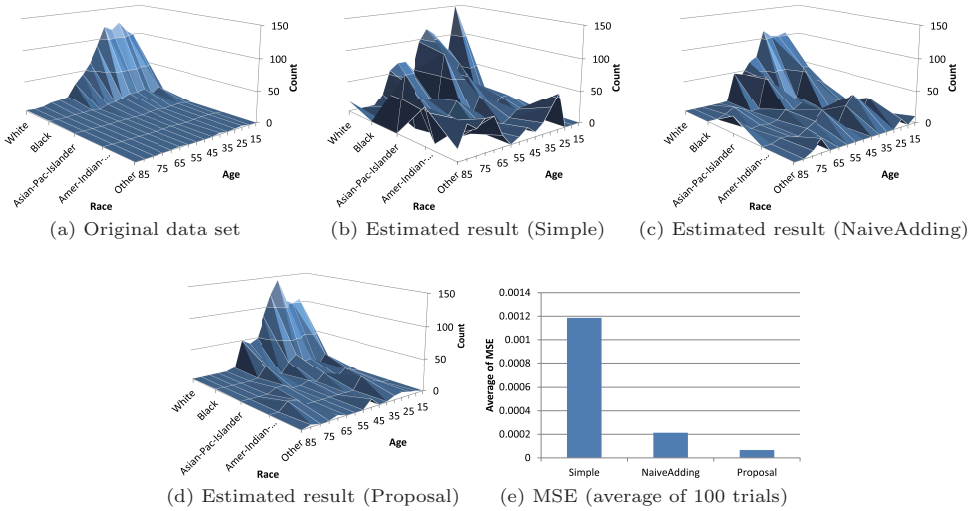


図 3　N を変化させたときの推測誤差
Fig. 3　MSEs with different N.



(a) Original data set (b) Estimated result (Simple) (c) Estimated result (NaiveAdding)

(d) Estimated result (Proposal) (e) MSE (average of 100 trials)

図 4　Adult データセットの分布と推測結果
Fig. 4　Distribution of Adult data set and estimated results.

つカテゴリー化して利用した．抽出されたデータには [15-19] から [90-94] の 16 カテゴリーまで存在し，人種は White, Black, Amer-Indian-Eskimo, Asian-Pac-Islander, Other の 5 種類であった．したがって計 80 の組合せがある．年齢及び人種の組合せにおけるデータ分布を，図 4 (a) に示す．

Simple 手法，NaiveAdding 手法，提案手法を用いて NC の情報から TC の分布を推測した結果をそれぞれ図 4 (b)，図 4 (c)，図 4 (d) に示す．Simple 手法は，元のデータ分布と比べて推測誤差が大きいことが図から分かる．NaiveAdding 手法は，Simple 手法よりは元のデータ分布をよく推測できているが，まだ推測誤差が見受けられる．提案手法においても元のデータ分布を完全に推測することはできていないが，比較対象の手法と比べて最も誤差を少なく推測できていることが図から読み取れる．

Adult データセットに対して，NC の生成から TC の分布を推測する試行を 100 回繰り返してそれぞれ推測誤差を計測し，その平均値を取った結果を図 4 (e) に示す．提案手法における推測誤差は Simple 手法の約 5.6%，NaiveAdding 手法の約 31%の値となっている．また，NaiveAdding の推測誤差は約 0.00021，提案手法の推測誤差は約 0.00007 であり，数学的評価を行った図 2 の結果とほぼ一致していることが確認できる．

6.2.2　計 算 時 間

提案手法において，収集された NC のデータから，TC のデータ分布を求めるために必要な時間を計測した．実験は，OS が Windows 7 Professional 64 bit, CPU が Intel Core i7-3712QM CPU @ (2 CPUs), RAM が 8GB である PC を利用して行った．

カテゴリー数を 10 から 10,000 まで，ユーザ数を 100 から 100,000 まで変動させてシミュレーションを行ったが，いずれも 1 秒未満で TC のデータ分布を推測することができた．

7. 考　　察

各ユーザはプライバシー保護レベルとして γ の値を設定できるとしたが，適切な値の設定方法は将来課題である．

本論文では Uniform Perturbation の確率行列で比較を行ったが，**4.2.1** で提案したサーバにおける推測アルゴリズムは，確率行列に依存せずに利用できる手法である．そのため全く異なる確率行列でも有効な手法であると考えている．今後は，様々な確率行列に対して評価を行う必要がある．ただし，**4.2.2** に示した推測アルゴリズムの簡素化は，Uniform Perturbation を前提としているため，異なる確率行列を利用する場合は簡素化したアルゴリズムを利用することはできない．

また提案手法は，特定のプライバシー指標に依存せず，各カテゴリーに属すユーザ数を推測することができる．他のプライバシー指標を用いた際における，既存手法との比較を行う必要がある．

8. むすび

ユーザが自身のデータを改変してサーバに送信し，サーバは得た情報から解析を行う，というプライバシー保護モデルにおいて広く利用されている Randomized Response (RR) スキームにおいて，各ユーザがプライバシー保護レベルを柔軟に設定可能な手法を提案した．従来手法においては，全ユーザが同一のプライバシー保護レベルを設定する必要があり，異なるプライバシー保護レベルが設定されている状況における，サーバ側での解析手法が確立されていなかった．本論文では，このような状況において，サーバ側で解析できるアルゴリズムを提案した．また，RR の中でも広く用いられている Uniform Perturbation モデルにおいて特に，少ない計算量で実行可能な推測アルゴリズムの提案を行った．比較対象とした従来手法や簡易手法と比べ，サーバ側での推測誤差を 70%程度削減できることを，数学的解析及び実データを用いたシミュレーションによって示した．

謝辞　本研究は JSPS 研究費 24300005，23500039，25730038 の助成を受けたものです．

本研究を遂行するにあたり，研究の機会と議論・研鑽の場を提供して頂き，ご指導頂いた国立情報学研究所/東京大学 本位田 真一教授をはじめ，活発な議論と貴重なご意見を頂いた研究グループの皆様に感謝致します．

文　献

[1] 飯尾　淳，吉田圭吾，小池亜弥，清水浩之，白井康之，桑山晃一，栗山桂一，小浪宏信，高山隼佑，“属性付き位置情報ログが示す行動特性と消費傾向の関係，” 情処学論，vol.52, no.7, pp.2256–2267, 2011.

[2] J. Froehlich, E. Larson, S. Gupta, G. Cohn, M. Reynolds, and S. Patel, “Disaggregated end-use energy sensing for the smart grid,” IEEE Pervasive Computing, vol.10, no.1, pp.28–39, 2011.

[3] S.L. Warner, “Randomized response: A survey technique for eliminating evasive answer bias,” J. American Statistical Association, vol.60, no.309, pp.63–69, 1965.

[4] Z. Huang and W. Du, “OptRR: Optimizing randomized response schemes for privacy-preserving data mining,” Proc. IEEE ICDE, pp.705–714, 2008.

[5] M.M. Groat, B. Edwards, J. Horey, W. He, and S. Forrest, “Application and analysis of multidimensional negative surveys in participatory sensing applications,” Pervasive and Mobile Computing, vol.9, no.9, pp.372–391, Jan. 2013.

[6] Y. Bao, W. Luo, and X. Zhang, “Estimating positive surveys from negative surveys,” Statistics & Probability Letters, vol.83, no.2, pp.551–558, Feb. 2013.

[7] A. Evfimievski, J. Gehrke, and R. Srikant, “Limiting privacy breaches in privacy preserving data mining,” Proc. ACM PODS, pp.211–222, 2003.

[8] S.P. Kasiviswanathan, H.K. Lee, K. Nissim, S. Raskhodnikova, and A. Smith, “What can we learn privately ?,” SIAM Journal on Computing, vol.40, no.3, pp.793–826, 2013.

[9] J. Corburn, “Confronting the challenges in reconnecting urban planning and public health,” American journal of public health, vol.94, no.4, pp.541–546, 2004.

[10] C. Sharp, S. Schaffert, A. Woo, N. Sastry, C. Karlof, S. Sastry, and D. Culler, “Design and implementation of a sensor network system for vehicle tracking and autonomous interception,” Proc. EWSN, pp.93–107, IEEE, 2005.

[11] J.-S. Lee and B. Hoh, “Sell your experiences: A market mechanism based incentive for participatory sensing,” Proc. IEEE Percom, pp.60–68, 2010.

[12] D. Yang, G. Xue, X. Fang, and J. Tang, “Crowdsourcing to smartphones: Incentive mechanism design for mobile phone sensing,” Proc. ACM Mobicom, pp.173–184, 2012.

[13] Y. Tian, B. Song, and E.-N. Huh, “Privacy preserving with a purpose-based privacy data graph,” Trustworthy Ubiquitous Computing, vol.6, pp.249–267, Atlantis Press, 2012.

[14] M. Mutka and L. Ni, “PrudentExposure: A private and user-centric service discovery protocol,” Proc. IEEE PerCom, pp.329–338, 2004.

[15] S. Agrawal and J.R. Haritsa, “A framework for high-accuracy privacy-preserving mining,” Proc. IEEE ICDE, pp.193–204, 2005.

[16] R. Chaytor and K. Wang, “Small domain randomization: Same privacy, more utility,” Proc. VLDB Endow., vol.3, no.1-2, pp.608–618, 2010.

[17] X. Xiao, Y. Tao, and M. Chen, “Optimal random perturbation at multiple privacy levels,” Proc. VLDB, pp.814–825, ACM, 2009.

[18] Y. Li, M. Chen, Q. Li, and W. Zhang, “Enabling multilevel trust in privacy preserving data mining,” IEEE Trans. Knowl. Data Eng., vol.24, no.9, pp.1598–1612, 2012.

[19] F. Esponda, “Negative surveys,” Technical report, arXiv.org, 2006.

[20] J. Horey, M.M. Groat, S. Forrest, and F. Esponda, “Anonymous data collection in sensor networks,” Proc. MobiQuitous, pp.1–8, IEEE, 2007.

[21] C. Dwork, “Differential privacy,” Automata, Languages and Programming, vol.4052, pp.1–12, Lecture Notes in Computer Science, Springer, 2006.

[22] H.H. Ku, “Notes on the use of propagation of error formulas,” Journal of Research of the National Bu-

reau of Standards, vol.70C, no.4, pp.331–341, 1966.
[23] UCI Machine Learning Repository, “Adult Data Set,” http://archive.ics.uci.edu/ml/datasets/Adult
[24] K. LeFevre, D. DeWitt, and R. Ramakrishnan, “Incognito: Efficient full-domain k-anonymity,” Proc. ACM SIGMOD, pp.49–60, 2005.
[25] A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam, “l-diversity: Privacy beyond k-anonymity,” ACM TKDD, vol.1, no.1, article no.3, 2007.

（平成 25 年 7 月 28 日受付，10 月 21 日再受付）

清　　雄一　（正員）

1981 年生．2009 年東京大学大学院情報理工学系研究科博士後期課程修了．同年（株）三菱総合研究所入社．同社情報技術研究センター，金融ソリューション本部等に所属．2013 年より電気通信大学助教，現在に至る．分散コンピューティング，セキュリティ，プライバシー保護技術等の研究に従事．情報処理学会，電子情報通信学会，IEEE Computer Society 各会員．

大須賀昭彦　（正員）

1958 年生．1981 年上智大学理工学部数学科卒．同年 (株) 東芝入社．同社研究開発センター，ソフトウェア技術センター等に所属．1985〜1989 年 (財) 新世代コンピュータ技術開発機構 (ICOT) 出向．2007 年より，電気通信大学大学院情報システム学研究科教授．2012 年より，国立情報学研究所客員教授兼任．工学博士 (早稲田大学)．主としてソフトウェアのためのフォーマルメソッド，エージェント技術の研究に従事．1986 年度情報処理学会論文賞受賞．IEEE Computer Society Japan Chapter Chair，人工知能学会理事，日本ソフトウェア科学会理事を歴任．情報処理学会，電子情報通信学会，人工知能学会，日本ソフトウェア科学会，IEEE Computer Society 各会員．