匿名加工・再識別コンテストIce & Fireの設計

菊池 浩明† 山口 高康‡ 濱田 浩気* 山岡 裕司** 小栗 秀暢¶
佐久間 淳§

†明治大学総合数理学部 ‡(株)NTT ドコモ先進技術研究所
〒164-8525 東京都中野区中野 4-21-1
* NTT セキュアプラットフォーム研究所 ** 株式会社富士通研究所 ¶ニフティ(株)
§筑波大学

あらまし 個人情報の保護とビッグデータの活用を両立させるために，個人情報の匿名加工の法整備が進み，様々な方式が導入されようとしている．その一方，匿名加工の為の方式には様々な提案が行われており，適用する分野やデータの種別に適した方式をどのように選定するか定まっていない．加えて，有用性と安全性を正しく評価する方法が確立していない．そこで，共通の疑似データを用いて，匿名加工とその再識別の技術を競うコンテストを企画する．本稿では，このコンテストの目的，提供する疑似データの選定，サンプルの匿名加工と再識別アルゴリズム，有用性評価方法，安全性評価方法，および，コンテストを支援する評価プラットフォームの設計について述べる．

Ice and Fire: Design of Data Anonymization and De-Anonymization Competition

Hiroaki Kikuchi † Takayasu Yamaguchi ‡ Koki Hamada* Yuji Yamaoka**
Hidenobu Oguri¶ Jun Sakuma§

†School of Interdisciplinary Mathematical Sciences, Meiji University
4-21-1 Nakano, Nakano Ku, Tokyo, 164-8525 Japan
kikn@meiji.ac.jp

‡Research Laboratories, NTT DOCOMO, Inc. ‡NTT Secure Platform Laboratories
‡FUJITSU LABORATORIES LTD. ‡NIFTY Corporation ‡University of Tsukuba

Abstract Data anonymization is ready to go before the big-data business runs successfully while preserving privacy of personal information. While, it is not trivial to choose the best algorithm to make the given data anonymized to be secure for a given particular purpose. To access the risk to be compromised accurately, the data needs to balance the utility and the security. Hence, with common pseudo micro-data, we propose a competition for best anonymization and re-identification algorithm. The paper addresses the aim of the competition, the target micro-data, sample algorithms, utility and security metrics. The design of evaluation platform is also mentioned.

1 はじめに

個人情報保護法が 10 年ぶりに改正されようとしている．改正案は 2015 年 5 月に衆議院を通過し，改正案がこのまま成立すれば，2016 年 1 月から個人情報保護委員会が発足し，2017 年 1 月から全面施行する．オプトアウトの厳格化，グローバル化への対応に並び，大きな柱となっているのが，「匿名加工情報」の新設である．改正法では，「（定められた措置を講じて）特定の個人を識別することができないように個人情報を加工して得られる個人に関する情報

であって，当該個人情報を復元することができないようにしたもの（第二条 9 項)」と定められている．匿名加工情報は，あらかじめ情報の項目と提供方法を公表することで，本人の同意がなくても第三者に提供することが許されており，IoT をはじめとするビッグデータの活用を推進する概念として期待されている．

しかしながら，その加工方法については標準的な定めはなく，業種ごとに定められた認定個人情報保護団体が，消費者や関係者の意見を聴いて，その対象事業に適した個人情報保護指針を定める（第五十三条）こととされている．データの種類や特性が多様で，漏えいのリスクや被害の度合いもまちまちな為に，各業界で特有の安全管理措置と匿名加工方法を定めることとしたものと考えられる．しかし，たとえ業種を特定したとしても加工方法は自明ではない．

世界的にも，技術的に匿名加工の方式を確立する試みが行われている．ISO/IEC で 29151 “Code of practice for PII protection” の標準化が始まったところで，今後もプライバシー強化のための非識別化技術が予定されている．欧州でも，法的なフレームワークと実装可能なプライバシー保護技術や評価指標との間にギャップがあることを指摘し，その溝を埋めるための行為規範が提案されている [5].

我々は，匿名加工の方法が明確に定まっていないことの原因として，匿名加工方法の安全性と有用性を評価する為のテストベッドとなる共通データの欠乏があることと考えている．研究目的として人工的に機械生成したデータには，現実の問題では避けられない極端な偏りや欠損などが排除されていたり，値に意味がないために漏えいの危険性がイメージしにくい．研究者独自に自分たちの提案方式を効果的に評価する個別のデータセットを使うので，方式の公平な評価ができない．再識別のリスクを評価するには，他のリソースを照合して匿名加工データを再識別してみればよい．しかし，改正法では，そういった照合は禁じられている．

そこで，匿名加工技術の開発と再識別に対する公平な安全性評価手法の確立を目的として，教育機関などの演習用として独立行政法人統計センターが作成した疑似ミクロデータ [6] を用いて，匿名加工と再識別のコンテスト “Ice & Fire”[1]を設計した．本データは，平成 16 年全国消費実態調査を基にして生成されており，現実の統計値とほぼ同等の性質を有し，しかも，任意の目的で自由に加工することが許されている．従って，改正法や利用環境に制限されることなく，あらゆる手法で再識別を試みることができる．こうして，様々な匿名加工技術を共通の環境で，公平に定量評価することにより，匿名加工の課題を明らかにし，より高い有用性と安全性を持つ匿名加工方式の開発を促進する．

本稿は，本コンテストで用いる技術の基本定義と有用性と安全性の評価指標を提案する．コンテストの設計においては，考慮した不正行為と対象としたリスクについて述べ，サンプルとするいくつかの再識別アルゴリズムを定義する．更に，競技を円滑に進め，参加者を支援する評価プラットフォームの設計について述べる．

2 匿名加工アルゴリズムと再識別

2.1 基本定義

データセット $\boldsymbol{X}$ は，m 個の属性 $X^1,\ldots,X^m$ を持つ n 個のレコード $\boldsymbol{x}_i=(x_i^1,\ldots,x_i^m)$ $(i=1,\ldots,n)$ からなる．レコード行番号 $I^X=(1,\ldots,n)$ は，各レコードの行番号表す．属性は，連続値，順序属性値，カテゴリ属性値，ブール値などがあり，属性 X の値域を $R(X)$ で表す．データセット $\boldsymbol{X}$ は，特定の個人を識別する[2]情報を含む時，その時に限り個人データとなる．

匿名化 (anonymization) とは，Emam は ISO/TS 25237 を引いて「データ主体とそれを識別するデータの相関を取り除く処理」と定めている [2]. 本稿では，データセット $\boldsymbol{X}$ から処理された匿名加工データを $\boldsymbol{Y}$ で表す．匿名加工データ $\boldsymbol{Y}$ は，$\{X^1,\ldots,X^m\}$ の部分集合の $m'(m'\leq m)$ 個の属性についての $n'(n'\leq n)$ 個のレコード $\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{n'}$ から成る．$\boldsymbol{Y}$ のレコード $\boldsymbol{y}_j=(y_j^1,\ldots,y_j^{m'})$ は，データセット $\boldsymbol{X}$ のあるレコード $\boldsymbol{x}_i$ を基にして処理されており，その関係を $j=\pi(i)$ となる関数 $\pi:\{1,\ldots,n\}\to\{1,\ldots,n'\}$ で表す．π は全射（トップコーディングなどにより識別されやすいレコードを排除することがあるため）や単射（一つのレコードから複数の匿名加工レコードが生成されることとあるため）とは限らないので，厳密な置換 (permutation) とはならないことに注意せよ．匿名加工データにおけるレコードの行番号を関数 π を

[1] Ice は匿名加工処理，Fire は再識別処理を暗示し，両者がそれぞれ技巧を競うことを想定している．

[2] 個人を特定する

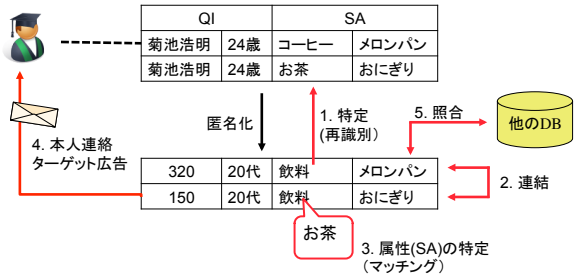

図 1: 匿名加工と再識別

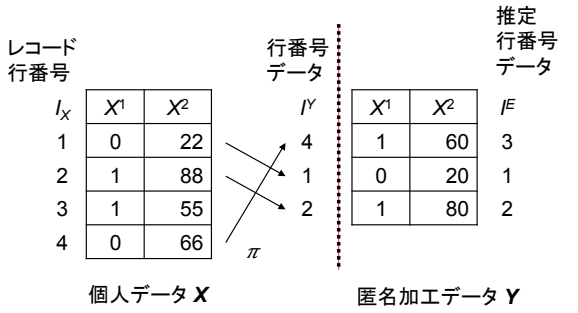

図 2: 行番号データ

用いて $I^Y = (i_1^Y, \ldots, i_{n'}^Y) = (\pi^{-1}(1), \ldots, \pi^{-1}(n'))$ で表す（図 2 参照）.

再識別 (re-identification) は，匿名加工データを分析したり，他のデータセットと照合することにより，匿名加工データのレコード $\boldsymbol{y}_j$ と元の個人データ $\boldsymbol{x}_i$ を結びつける処理を指す．匿名加工データに関わる脅威には，図 1 に示すように，1) 特定 (再識別), 2) 連結, 3) 属性推定（復元），4) 本人への連絡，などのいくつもの種類があるが，本稿では，狭義の再識別として，匿名加工データ $\boldsymbol{Y}$ と照合データから, $\boldsymbol{X}$ のレコード行番号を推定することと定める．すなわち，攻撃アルゴリズム E により，行番号データ I_Y と等しくなる（再識別）推定行番号データ $I^E = (i_1^E, \ldots, i_{n'}^E) \cong I^Y$ を求めることである．以上の行番号データと推定行番号データの関係を図 2 に示す．この例では，$n = 4$ の個人データから，行削除とレコード置換が行われ，$n' = 3$ の匿名加工データが生成されている．推定行番号データは 1 行目が誤っている．

2.2 準識別子 QI と機微属性 SA

属性 X には，マイナンバーの様な直接的な識別子，性別，住所，年齢などの様な，組み合わせることで個人を識別することが可能な間接的な識別子である準識別子 (quasi-idenitifier, 以後 QI)，そして，それ以外の属性がある．それ以外の属性の内，病名や思想信条などの配慮な必要な情報を機微属性 (sensitive information，以後 SA) と呼ぶ．Emam は，QI と SA の例として，それぞれ，学歴，日常言語，学歴，離婚歴などと，遺伝的情報，精神疾患歴，性的情報，生殖情報などを上げている ([2]2 章のコラム).

しかしながら，機微の度合いには個人差が大きく，QI とも SA とも考えられる属性もあり，その定義はしばしば議論になる．そこで本稿では，属性 $X^1, \ldots, X^m$ の内，静的な情報で，しばしばカテゴリー化されている離散値を取る属性を QI とみなし，それ以外をすべて SA と考える．QI と SA の属性を集合 $QI, SA \subset \{1, \ldots, m\}$ で表す．$QI \cup SA = \{1, \ldots, m\}$, $QI \cap SA = \emptyset$ である.

2.3 攻撃者

再識別を行う攻撃者に許されるリソースは，様々な場合が考えられ，一様な仮定は困難である.

本稿でも攻撃者は最大の知識，すなわち，オリジナルの個人データ $\boldsymbol{X}$ と照合して，匿名加工データ $\boldsymbol{Y}$ の再識別を試みる.

2.4 匿名加工アルゴリズム例

匿名加工の技術として，列削除，行削除，一般化，摂動化，リサンプリングがよく知られている．これらの技術を組み合わせて，k-匿名性，ℓ-多様性，t-近似性などの安全性指標を満たすように個人データを匿名加工する.

2.5 再識別アルゴリズム例

再識別アルゴリズムには標準的なものは知られていない．攻撃者に与えられる知識の想定ができないため，共通のアルゴリズムを考えにくいためである．そこで，本稿では，オリジナルの個人データ $\boldsymbol{X}$ を与えて，匿名加工データ $\boldsymbol{Y}$ から行番号データ I^Y を推定する次のアルゴリズムを考える．これらを，コンテストにおける匿名加工データの安全性の初期評価を与える基本アルゴリズムとして用いる.

Sort 小さな摂動化に対しては強力な再識別を実行するが，ソートによって識別をしているためトップコーディングなどのレコード削除に対して弱い．

IdRand k-匿名性が低いデータセットをより高い確率で再識別する．一様な確率で推定するために，候補レコード集合 $C(\boldsymbol{y})$ が小さいほど識別率が上がる．安全性指標 k-anony の平均値が $\overline{k}$ である時，この攻撃アルゴリズムが再識別するレコード数の期待値は $2/\overline{k}$ である．

IdSA IdRand と同様に，候補レコード集合 $C(\boldsymbol{y})$ を用意するが，指定された SA のある属性について $\boldsymbol{y}$ に最も近いレコードを選ぶところが異なる．

Algorithm 1 Sort

1. 入力: 匿名加工データ $\boldsymbol{Y}$, 個人データ $\boldsymbol{X}$,
   出力: 推定行番号データ $I^{Sort} = (i_1^{Sort}, \ldots, i_{n'}^{Sort})$
2. レコード $\boldsymbol{x}_j$ の特徴量を，そのレコードの SA の和で $s_j = \sum_{i \in SA} x_j^i$ と定める．
3. 個人データ $\boldsymbol{X}$ と匿名加工データ $\boldsymbol{Y}$ をそれぞれ，上記の特徴量についてソートした時の順位を $Rank^X, Rank^Y$ とする．
4. $\boldsymbol{y} \in \boldsymbol{Y}$ の各レコードについて，$Rank^X(\boldsymbol{x}_k) = Rank^Y(\boldsymbol{y}_j)$ となる順位 k を推定行番号データ $i_j^{Sort} = k$ とする．

Algorithm 2 IdRand

1. 入力: 匿名加工データ $\boldsymbol{Y}$, 個人データ $\boldsymbol{X}$,
   出力: 推定行番号データ $I^{IR} = (i_1^{IR}, \ldots, i_{n'}^{IR})$
2. レコード $\boldsymbol{y}_i \in \boldsymbol{Y}$ について，QI を共通にする $\boldsymbol{X}$ の候補レコード集合 $C(\boldsymbol{y}_i) = \{\boldsymbol{x} \in \boldsymbol{X} | \boldsymbol{x} \approx^{QI} \boldsymbol{y}_i\}$ を求める．
3. $C(\boldsymbol{y})$ から一様な確率 $p = 1/|C(\boldsymbol{y})|$ でレコード $\boldsymbol{x}_{j*}$ を選び，推定行番号データ $i_i^{IR} = j*$ とする．全ての $\boldsymbol{y}_i$ について Step 2 から繰り返す．

Algorithm 3 アルゴリズム IdSA

1. 入力: 匿名加工データ $\boldsymbol{Y}$, 個人データ $\boldsymbol{X}$, 識別用属性 $X^s \in SA$
   出力: 推定行番号データ $I^{IS} = (i_1^{IS}, \ldots, i_{n'}^{IS})$
2. レコード $\boldsymbol{y}_i \in \boldsymbol{Y}$ について，QI を共通にする $\boldsymbol{X}$ のレコードの集合 $C(\boldsymbol{y}_i) = \{\boldsymbol{x} \in \boldsymbol{X} | \boldsymbol{x} \approx^{QI} \boldsymbol{y}_i\}$ を求める．
3. $C(\boldsymbol{y})$ の各レコード $\boldsymbol{x}_j$ について，属性 X^s について $\boldsymbol{y}_i$ との距離最小となるレコード $\boldsymbol{x}_{j*}$ を
$$j^* = \underset{j \in C(\boldsymbol{y}_i)}{\operatorname{argmin}} |x_j^s - y_i^s|$$
   により定め，推定行番号データ $i_i^{IS} = j*$ とする．

Algorithm 4 アルゴリズム SA21

1. 入力: 匿名加工データ $\boldsymbol{Y}$, 個人データ $\boldsymbol{X}$, 出力: 推定行番号データ $I^{S_{21}} = (i_1^{S_{21}}, \ldots, i_{n'}^{S_{21}})$
2. レコード $\boldsymbol{y}_i \in \boldsymbol{Y}$ の特徴量を $y_j^{S_{21}}$ とする．
3. 個人データ $\boldsymbol{X}$ と匿名加工データ $\boldsymbol{Y}$ をそれぞれ，特徴量 $y_j^{S_{21}}$ についてソートした時の順位を $Rank^X, Rank^Y$ とする．
4. $\boldsymbol{y} \in \boldsymbol{Y}$ の各レコードについて，$Rank^X(\boldsymbol{x}_k) - 1 = \lfloor (Rank^Y(\boldsymbol{y}_j) - 1) \times \frac{n-1}{n'-1} \rfloor$ となる順位 k を推定行番号データ $i_j^{S_{21}} = k$ とする．

3 コンテスト設計

3.1 疑似ミクロデータ

「(教育用) 疑似ミクロデータ（以後，GMD と呼ぶ）は，公的統計のミクロデータの利用を図るため，教育機関などの演習用として独立行政法人統計センターが作成した疑似的なミクロデータである [6]．約 59,400 世帯（単身 5,002 世帯を含む）に対して 3 か月間の家計簿を調べた平成 16 年全国消費実態調査を基に生成されている．この個票データから高次元の集計表を作成し，各セルの量的属性値が多変量（対数）正規分布に従うことを仮定して，多変量正規乱数を作成することで作成されており，元の個票データの特性を保存している．

表 1 に，GMD の基本仕様を示す．簡易データは，

表 1: 疑似ミクロデータ仕様 [6]

データセット	レコード数	QI 数	SA 数	
			支出項目	収入項目
	n	m		
大規模データ	32,027	14	149	34
簡易データ	8,333	14	11	N/A

大規模データの中から世帯人員が 4 名で有業人員が 1 名から 2 名の世帯のみの $n = 8,333$ レコードからなっている．属性の内，世帯主の年齢，住居の種類などの質的属性 14 項目を QI, 食糧，住居，光熱，教育などに分類された消費支出と年間収入などの量的属性を SA とする．簡易データは，消費支出の十大費目のみを有する．

3.2 プレイヤー

本コンテストには次のプレイヤーがある．

1. 匿名加工者（ディフェンス）GMD $\boldsymbol{X}$ を基に，匿名加工処理を行い，匿名加工データ $\boldsymbol{Y}$ と行番号データ I^Y を生成する．再識別者に $\boldsymbol{Y}$ を，審判員に $\boldsymbol{Y}$ と I^Y を提出する．
2. 再識別者（オフェンス）匿名加工データ $\boldsymbol{Y}$ を受け取り，GMD $\boldsymbol{X}$ を参照して，再識別アルゴリズムを実施して推定した推定行番号データ I^E を審判員に提出する．
3. 審判員 (ジャッジ) $\boldsymbol{Y}$ の有用性指標を評価する．I^Y を参照して，$\boldsymbol{Y}$ の安全性指標を評価する．I^E と I^Y を照合して，再識別率を算出する．

最も多くの匿名加工データの再識別を実現した再識別者と，最も有用性評価と安全性評価が高い匿名加工者を勝者とする．

3.3 有用性指標の定義

3.3.1 有用性指標 meanMAE

$SA = \{14, 15, \ldots, 25\}$ についてのデータセット平均値を求め，オリジナル個人データ $\boldsymbol{X}$ と匿名加工データ $\boldsymbol{Y}$ との MAE(平均絶対誤差) で定める．匿名加工データ $\boldsymbol{Y}$ について，

$$\mathrm{meanMAE}(\boldsymbol{X}, \boldsymbol{Y}) = \frac{1}{m} \sum_{i \in SA} |\mu(X^i) - \mu(Y^i)|$$

とする，ただしここで，

$$\mu(X^i) = \frac{1}{n} \sum_{j=1}^{n} x_j^i,\ \mu(Y^i) = \frac{1}{n'} \sum_{j=1}^{n'} y_j^i$$

とする．meanMAE は小さいほど有用性が高い．

3.3.2 有用性指標 cross

属性の部分集合 $A = \{a_1, a_2, \ldots\} \subset QI$ について，可能な値の全ての組合せについて，$B \in SA$ の値の平均 (Mean) と集計数 (Cnt) を求め，オリジナルの個人データ $\boldsymbol{X}$ と匿名加工データ $\boldsymbol{Y}$ の間の MAE で定める．すなわち，

$$\mathrm{crossMean}^{A,B}(\boldsymbol{X}, \boldsymbol{Y}) = \frac{1}{|R(A)|} \sum_{a \in R(A)} |\mu(\boldsymbol{x}^B | \boldsymbol{x}^A = a) - \mu(\boldsymbol{y}^B | \boldsymbol{y}^A = a)|$$

$$\mathrm{crossCnt}^{A,B}(\boldsymbol{X}, \boldsymbol{Y}) = \frac{1}{|R(A)|} \sum_{a \in R(A)} \big| |\{\boldsymbol{x} \in X | \boldsymbol{x}^A = a\}| - |\{\boldsymbol{y} \in Y | \boldsymbol{y}^A = a\}| \big|$$

とする[3]．ここで，$R(A)$ は属性 $X^{a_1}, X^{a_2}, \ldots$ における値域の直積の部分集合であり，$Y^A = a$ を満たすレコードが $\boldsymbol{Y}$ に存在しない時，$\mu(\boldsymbol{y}^B | \boldsymbol{y}^A = a) = 0$, $|\emptyset| = 0$ とする．例えば，$A = \{7, 9\}$ (性別，就業)，$B = 15$ （消費支出）とすると，$R(A) \subset \{(1,1), (1,2), (1,V), (2,1), (2,2), (2,V)\}$ である．$a = (1,2)$ とすると，$\mu(X^B | X^A = a)$ は，$x^7 = 1$ (男)，$x^9 = 2$ (就業) となるレコードだけに制限した X^{15} 消費支出の平均値である．crossMean, crossCnt とも小さいほど有用性が高い．

この指標は A と B によって評価が変わる．年齢と就業の有無に応じて (A)，住居費 (B) がどれ位変動するかなどの評価をモデルしている．

3.3.3 有用性指標: corMAE

SA の属性の任意の二組 X^i, X^j について，ピアソンの相関係数 cor を求め，個人データ $\boldsymbol{X}$ と匿名加工データ $\boldsymbol{Y}$ との間の MAE で定める．

$$\mathrm{corMAE}(\boldsymbol{X}, \boldsymbol{Y}) = \frac{1}{|SA|^2} \sum_{i,j \in SA} \big| \mathrm{cor}(X^i, X^j) - \mathrm{cor}(Y^i, Y^j) \big|$$

[3]外側の | が絶対値，内側の | が集合の要素数であるのに注意せよ

ここで，$\mathrm{cor}(X^i, X^j)$ はピアソンの相関係数である．

corMAE は小さいほど有用性が高い．

3.3.4 有用性指標 nrow

匿名加工アルゴリズムの中には，レコード削除の様に，識別の容易な特異なレコードを削除するものがある．レコードを削除すればするほど再識別のリスクが下がる．しかし，そのことによる有用性の損失を，匿名加工データ $\boldsymbol{Y}$ のレコード数で，

$$\mathrm{nrow}(\boldsymbol{Y}) = |n - |\boldsymbol{Y}|| = |n - n'|$$

と定める．

nrow は大きいほど有用性が高い．

3.4 安全性指標の定義

3.3.1 安全性指標 k-anony

$\boldsymbol{x}_i$ と $\boldsymbol{x}_j$ をデータセット $\boldsymbol{X}$ に属する二つのレコードとする．全ての $X^q \in QI$ について，$x_i^q = x_j^q$ である時，$\boldsymbol{x}_i$ と $\boldsymbol{x}_j$ は QI-同値関係にあると呼び，$\boldsymbol{x}_i \approx^{QI} \boldsymbol{x}_j$ と表す．この同値関係について，データセット $\boldsymbol{X}$ は $R(QI)$ の要素数 $|R(QI)|$ 個の同値類に直和に分割される．ただし，存在しない QI の組は除く．この同値類について，

$$k\text{-anony}(\boldsymbol{X}) = \min_{a \in R(QI)} |\{\boldsymbol{x} \in \boldsymbol{X} | X^{QI} = a\}|$$

$$k\text{-anonyMean}(\boldsymbol{X}) = \frac{1}{|R(QI)|} \sum_{a \in R(QI)} |\{\boldsymbol{x} \in \boldsymbol{X} | X^{QI} = a\}|$$

とする．ここで，$X^{QI} = a$ は，QI の全ての属性 $X^{q_1}, X^{q_2}, \ldots$ の直積が a である様なレコードを表す．すなわち，k-anony は，k-匿名性の k を，k-anonyMean はその平均値を与えている．

例えば，$QI = \{7, 9\}$ （性別，就業）について（疑似ミクロ簡易データ）$\boldsymbol{X}$ のレコードは，表 2 に示される個数だけあったとする．値域はそれぞれ，$\{1, 2\}, \{1, 2, V\}$ である．この時，k-anony $= 9$ ($k = 9$), k-anonyMean $= 1388.8$ である．

k-anony, Mean は大きいほど安全性が高い．

表 2: $QI = \{7, 9\}$ の時の k-匿名性

性別 \ 就業	1	2	V
1	8051	27	9
2	127	101	18

3.4.2 安全性指標 re-id

匿名加工データ $\boldsymbol{Y}$ とその行番号データ $I^Y = (i_1^Y, \ldots, i_{n'}^Y)$ とする．再識別アルゴリズム E によって，$\boldsymbol{Y}$ を分析して推定した推定行番号データを $I^E = (i_1^E, \ldots, i_{n'}^E)$ とする．この時，R による再識別率を，

$$\text{re-id}^E(I^Y, I^E) = \frac{\left|\{j \in \{1, \ldots, n'\} | i_j^Y = i_j^E\}\right|}{n'}$$

と定める．re-id$(I^Y, I^Y) = 1.0$ である．

再識別アルゴリズム Sort, IdRand, IdSA の再識別率をそれぞれ，re-id^{Sort}, re-id^{IdRand}, re-id^{IdSA} とする．

3.5 再識別率の課題「山岡攻撃」

安全性指標に再識別率を用いる場合，レコード順序を変えるだけの匿名加工が優秀と評価されてしまう問題があり，その匿名加工を本稿では「山岡攻撃」と呼ぶ．本コンテストのデータセットはレコード順序に意味はない．よって，前述の各有用性指標もレコード順序に非依存である．一方，再識別率はレコード順序に依存している．そのため，「山岡攻撃」は，有用性を一切下げず，再識別率も普通は 0 にすることができ，実質的に何も匿名加工していないにも関わらず良い評価を得られる．

たとえば，匿名加工データを

$$\boldsymbol{y}_i = \begin{cases} \boldsymbol{x}_n & (i = 1) \\ \boldsymbol{x}_{i-1} & (i > 1) \end{cases}$$
$$I^Y = (1, 2, \ldots, n)$$

とするのが「山岡攻撃」である．この匿名加工データは，普通の再識別アルゴリズムからは推定行番号データは $I^E = (n, 1, 2, \ldots, n-1)$ とされ，再識別率は 0 になる．有用性は，レコード順序を変えただけなので低下していない．

「山岡攻撃」の検知は一見簡単そうだが，摂動化などと組み合わせた拡張攻撃の検知は難しい．単純な攻撃は個人データの全レコードがそのまま残るので検知しやすいが，摂動化などと組み合わせた拡張

攻撃はレコードが変わるため検知が難しくなる．摂動化などをおこなうと多少有用性は下がるが，その程度が小さければ十分良い評価を得られる．

このような理由から，本コンテストでは「山岡攻撃」をルールとして禁止することにした．プレイヤーの良識に任せる方針である．

3.6 有用性指標 IL

Domingo-Ferrer らは，[3] にて匿名加工データの安全性を測る指標について提案している．（おそらく）個人データ $\boldsymbol{X}$ についての匿名加工データ $\boldsymbol{Y}$ とその行番号データ I^Y について，平均移動度を

$$IL(\boldsymbol{X}, \boldsymbol{Y}, I^Y) = \frac{1}{m'n'} \sum_{i=1}^{m'} \sum_{j=1}^{n'} \frac{|x^i_{i^Y_j} - y^i_j|}{\max x^i - \min x^i}$$

と定める．Domingo らの指標は，匿名加工によってレコードが削除することは仮定されていない?ので，変形して用いる．

IL は小さいほど有用性が高い．前述の「山岡攻撃」を考慮して，安全性指標に加える．

4 評価システム

4.1 概要

本コンテストは，データの提出がスケジュールに沿って行われる予備戦 (2015 年 8 月 24 日から 10 月 9 日) と，匿名加工と再識別が無手順に行われる本戦 (10 月 22 日) の 2 回に分けて開催され，両方の順位によって総合順位を決める．

4.2 予備戦システム

予備戦システムの目的は，本戦前に対象データの配布を行い，スケジュールに沿って匿名加工と再識別のアルゴリズムを開発・評価することにより，ルールを周知させ，参加者のアルゴリズムの改良・検討を活性化することである．

処理対象である擬似ミクロデータは利用許可申請が必要であるため，利用者管理を行うための評価システムを必要とする．そのため，ニフティ株式会社の協力のもと，クラウドサーバと評価プラットフォームを利用し，SSL 環境下で参加者ごとに独立した評価環境を構築して安全性を高め，円滑なコンテスト運営を試みる．

表 3: 予備戦システム仕様

サーバ	ニフティクラウド (R) サーバ
システム	ニフティ匿名化処理プラットフォーム
CPU	Intel(R) Xeon(R) 3.00GHz
Memory	4 GB
R 言語	R version 3.2.0
Java	Java(TM) SRE 1.8.0_45
Ruby	ruby 2.2.2p95
Python	Python 2.7.10

また，参加者の開発環境が多様であることが予測されるため，なるべく多くの開発言語に対応したサンプルコードを準備し，同じプログラムを用いて評価する環境を構築した．サンプルコードは，R 言語，Ruby，Python，Java で記述されており，参加者はダウンロードして処理の試行が可能である．表 3 に，本システムの仕様を示す．

4.3 総合評価

匿名加工データの評価方式は数多く提案されており，それぞれが目的に応じた利点を持つ．本コンテストでは 3.3，3.4 で定義した指標を採用し，公平に評価する．だが，これによって匿名加工処理の有益性・安全性の全てを網羅できるわけではない．本コンテストの趣旨は多くの匿名加工手法と再識別手法を掛け合わせ，それぞれ手法の利点/問題点を検討することにある．そのため，評価方式は絶対的な指標の大小ではなく，参加者同士の相対評価である順位 $Rank$ を主に用いる．

総合指標の問題点は準備段階でも多く判明しており，3.5 山岡問題や，データ量を極端に少なくする不正行為など，参加者が指標の問題点をついた対応を行わないよう，ルール上の禁止事項を設けた上で予備戦と本戦を分けることで対応する．

匿名加工コンテストでは，最も有用性が高く，最も安全な匿名加工者を勝者とする．匿名加工データ $\boldsymbol{Y}$ の総合スコア $\mathrm{Score}(\boldsymbol{Y})$ は有用性についての評価値 $\frac{1}{6}\sum_{i=1}^{6} Rank(U_i)$ と安全性についての評価値 $\frac{Rank(S_1)+Rank(S_2)}{4} + \frac{1}{2}Rank(\max_j \text{re-id}^{E_j})$ の和で定める．ここで，表 4 に示す有用性指標 $U_1, \ldots, U_6$，及び安全性指標 $S_1, S_2, E_1 \ldots, E_4$ を用いる．

匿名加工データの安全性は，いかなる攻撃に対しても最低限保証される安全性と解釈し，複数の再識

表 4: 有用性指標と安全性指標

No.	指標	概要	節
U_1	meanMAE	SA 平均絶対誤差	3.3.1
U_2	crossMean	クロス集計値の平均絶対誤差	3.3.2
U_3	crossCnt	クロス集計数の平均絶対誤差	3.3.2
U_4	corMAE	SA の相関係数の平均絶対誤差	3.3.3
U_5	IL	匿名加工データの各値の平均絶対誤差	3.6
U_6	nrow	匿名加工データのレコード数	3.3.4
S_1	k-anony	k-匿名性指標の最小値	3.4.1
S_2	k-anonyMean	k-匿名性指標の平均値	3.4.1
E_1	Sort	SA の総和でソートによる再識別率	2.5
E_2	IdRand	QI からランダムな再識別率	2.5
E_3	SA21	SA21 列について再識別率	2.5
E_4	IdSA	QI から SA15 列による再識別率	2.5

別アルゴリズムの中の最も高い再識別率で評価する．これに対して，再識別アルゴリズムは，再識別に成功したレコード数の相和を採用する．再識別攻撃者の優劣は，データの種類に関わらず，最も多くのユーザを識別した量で評価されるべきと考える．

これらの指標を用いて予備戦を行い，参加者の評価分布や傾向に応じて指標を再検討し，本戦では，対象データの分布変更，指標の重み付け，評価対象列の変更等を実施して公平性を高める．

4.4 本戦システム

本戦システムは，複数のディフェンスとオフェンスの対決を数時間のうちに公平に決着させる．また，この対戦状況を Web 上で自動的に分かりやすく実況中継し，手間をかけずに対戦を観覧するオーディエンスを楽しませる．本戦システムの概観と入出力を図 3 に示す．

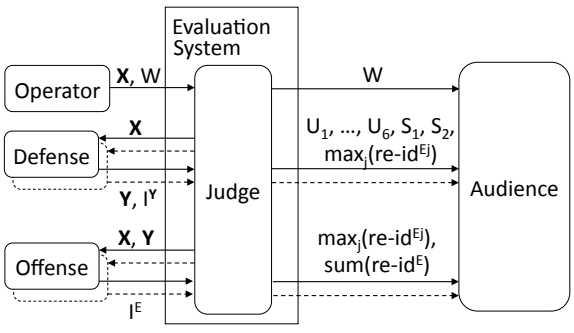

図 3: 本戦システムの概観と入出力

プログラムであるジャッジは，コンテスト運営者であるオペレータ，参加者であるディフェンスとオフェンス，観客であるオーディエンスとやりとりをしながら，対戦を進める．

5 結論

匿名加工と再識別コンテストの設計について報告した．本コンテストで提案した技術の基本定義と有用性と安全性の評価指標は，考慮した不正行為と対象としたリスクを明らかにし，再識別の際に考慮しなくてはならないポイントを提供している．実装した評価プラットフォームは，参加者の評価を支援し，公平な安全性と有用性の評価指標を与えている．

謝辞

本稿の執筆，ならびに，本コンテストの実施には，「擬似ミクロデータ（平成 16 年全国消費実態調査）」（独立行政法人 統計センター）を利用しました．

参考文献

[1] Information Commissioner's Office (ICO), Anonymisation: managing data protection risk code of practice, 2012.

[2] Khaled El Emam, Luk Arbuckle, "Anonymizing Health Data Case Studies and Methods to Get You Started", *O'Reilly*, 2013 (木村による和訳あり).

[3] J. Domingo-Ferrer and V. Torra, "A quantitative comparison of disclosure control methods for microdata", Confidentiality, Disclosure and Data Access: Theory and Practical Applications for Statistical Agencies, pp. 111-133, 2001.

[4] J. Domingo-Ferrer, et al., "Disclosure Risk Assessment via Record Linkage by a Maximum-Knowledge Attacker", 2015 Thirteenth Annual Conference on Privacy, Security and Trust (PST), *IEEE*, 2015.

[5] G. Danezis, et al., "Privacy and Data Protection by Design – from policy to engineering", ENISA, 2014.

[6] 秋山他，"教育用擬似ミクロデータの開発とその利用～平成 16 年全国消費実態調査を例として～", 統計センター製表技術参考資料, 16, pp. 1-43, 2012.