Vol.2017-EIP-75 No.3
2017/2/17

匿名加工・再識別コンテストにおける安全性指標の社会実装に向けた検討

黒政 敦史[†1]　小栗 秀暢[†1]　松井 くにお[†1]

概要：個人情報保護法の改正により，「匿名加工情報」という新たな情報の類型が定義された．一方、2015 年よりコンピュータセキュリティシンポジウム（CSS）において匿名加工・再識別コンテスト（PWSCUP）を行い，匿名化データに関する安全性の定量指標による評価の取り組みがなされている．本稿では個人情報保護委員会と経済産業省が作成した匿名加工情報に関する文書を参照しながら，海外の公的文書で用いられている匿名化基準および PWSCUP で使用された安全性指標「再識別率」と比較した上で，匿名加工情報の加工基準，安全性指標について検討する．

キーワード：プライバシー，　個人情報，　個人情報保護法，　匿名加工情報，　統計情報，　***k*-匿名性，**　PWSCUP

1. はじめに

2015 年 9 月に成立し，2017 年 5 月 30 日から施行される個人情報の保護に関する法律の改正法[1]（以後，改正後の同法を「改正法」という）により，匿名加工情報という新たな情報の類型が定義された．匿名加工情報とは「***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をいう***」（改正法 2 条 9 項）．

匿名加工情報は，一定の条件の下で，本人の同意がなくても第三者提供や目的外利用が可能となる．それによって地域や取引先との情報共有やマーケティング分析，機械学習の教師データ等への活用が期待できる．

本稿では，個人情報を含むパーソナルデータに関して，個人の識別可能性などを減少させる技術的な措置を「匿名化処理」，また，匿名化処理されたデータを「匿名化データ」とする．

匿名加工情報の加工基準については，個人情報保護委員会が個人情報保護法ガイドライン匿名加工情報編[2]（以後，「ガイドライン」という），経済産業省が匿名加工情報作成マニュアル[3]（以後，「マニュアル」という）を公開し，今後，業界ごとの認定個人情報保護団体によって指針が整備される見込みである．

また，特定の有用性・安全性の基準や技術に依存せず，匿名化処理とその評価指標に関する研究の伸展のための取り組みとして，2015 年からコンピュータセキュリティシンポジウム（CSS）にて**匿名加工・再識別コンテスト（Privacy Work Shop CUP，以後「PWSCUP」，**2015 年大会[4]は「CUP´15」，2016 年大会[5]は「CUP´16」という）が行われている．

PWSCUP は，匿名化データの有用性と安全性を競うコンテストである．匿名化データのユースケースを想定して，用意されたデータセットに匿名化処理を行うことで，有用性の評価を行う．同時に，その匿名化データをコンテスト参加者に公開し，元データへの再識別処理（以後「再識別攻撃」あるいは「攻撃」という）による安全性評価を行う．

本稿では，匿名化データにおいて元データが再識別される確率指標としての ***k*-匿名性**，及び，PWSCUP における安全性評価について整理する．また，匿名加工情報に隣接する個人情報ならびに，改正法による保護対象外の個人に関する情報に該当しない集計情報等（以後，ガイドラインに記載された「統計情報」を含め「集計情報等」という）などと合わせた，**事業者が実際に匿名化データを使用する上で必要となる加工基準，安全性指標について検討する**．

2. 従来研究

2.1 個人情報保護委員会のガイドライン

匿名加工情報の作成は，改正法 36 条に以下のように定められている．「***個人情報取扱事業者は，匿名加工情報（略）を作成するときは，特定の個人を識別すること及びその作成に用いる個人情報を復元することができないようにするために必要なものとして個人情報保護委員会規則で定める基準に従い，当該個人情報を加工しなければならない***」．匿名加工情報の取扱いに関して，**個人情報保護委員会から改正法に則り作成したガイドライン[2]**が公開された．

ガイドラインにおいて**「特定の個人を識別する」**ことができる基準は**「一般人の判断力又は理解力をもって生存する具体的な人物と情報の間に同一性を認めるに至ることができるかどうか」**と記述されている．また**「一般人及び一般的な事業者の能力，手法等を基準として」**とされていることから，一般人の基準には「再識別能力」と「再識別手法」が存在することがわかるがその詳細は明記されていない．

匿名加工情報の加工に係る手法例として，項目削除／レコード削除／セル削除，一般化，トップ（ボトム）コーデ

†1 ニフティ株式会社 NIFTY Corporation,
〒169-8333 東京都新宿区北新宿 2-21-1　Shinjuku,　Tokyo 169-8333 Japan.

ィング，ミクロアグリゲーション，データ交換（スワップ），ノイズ（誤差）の付加，擬似データ生成，を挙げている．これらは，あくまでも加工手法であり，その加工を通じて作成された匿名化データの安全性指標に関する基準は示されていない．ガイドラインは，元データを加工し匿名化処理を行った事業者が，一般人の能力と手法により再識別できないことを担保した安全性を求めているが，その基準と妥当性の判断は，事業者に委ねられた形になっている．

2.2 経済産業省のマニュアル

経済産業省から匿名加工情報の作成方法に関する参考資料として，匿名加工情報作成マニュアル[3]が公開された．そこには**「匿名加工情報取扱事業者に課される規律（匿名加工情報に係る公表、識別行為の禁止、安全管理措置 等）についての言及は、必要最小限にとどめており、個人情報保護法の法文の解釈を示すものではない」**と記載されており，具体的な安全性の基準にまで踏み込まないこととしている．そのため，検討プロセスや加工手法のみが示されており，加工後の匿名化データの安全性や有用性の指標の記述はされていない．

マニュアルにおいて，「個人識別に係るリスク」と認識されているものとして以下を挙げている．

1： 個人が特定されるリスク
2： データが他の情報と照合されるリスク
3： データを用いて本人へアプローチされるリスク

このうち，2： と 3： のリスクについては，ユースケースに応じて「評価することが望ましい場合もある」とした記述に止まっている．

また，これらのリスクに対する加工方法の検討プロセスとして，以下の段階を上げ，匿名加工の要求レベルは，事業内容や利用形態等によって判断されるべきものとし，具体的な指標は示されていない．

1） ユースケースの明確化
2） *識別子*，*属性*，*履歴*の仕分け
3） 個人識別等に係るリスクの抽出
4） 個人識別等に係るリスクを踏まえた加工方法の検討

1）のユースケースは業務・サービス概要，対象データの項目や規模，利用目的等の整理を求めている．

2）の仕分けにおいて***属性***と***履歴***は共に単体では個人を特定できない情報であると定義されているが，相違点として***履歴***は**「経時的にデータが積み重ねられる」**としている．

3）のリスクの抽出では攻撃者の持つ知識レベルについて**「市販のデータベースなどの照合可能な外部情報」，「提供先が照合可能なデータ」**を対象とし，それらと照合ができないことを要件としているが，具体的な市販のデータベースは例示に止まる．

4）の加工方法は，***属性***と***履歴***に対する再識別性を低減させる手法を紹介している．

属性については**「個人の識別リスクが低減できるレベルまで加工する．この際 *k*-匿名性を考慮する」**と記載され，個人の識別可能性を減少させることを求めている．一方，***履歴***に関しては**「一般的に識別子又は*属性*と組み合わされない限り個人を特定する可能性のない情報が該当するが，*履歴*であっても必ずしも無加工で利用できるものではないことに十分留意する」**とされ，留意点として***履歴***の継続的な提供や位置情報の精度を例示しているが，攻撃者の知識や能力を想定するなどの記述はなされていない．

2.3 匿名化データの定量的な評価指標

匿名化データの安全性や有用性を定量化する指標は多く研究されている．安全性の指標として，***k*-匿名性[6]や，*Pk*-匿名性[7]，*l*-多様性[8]などの，個人が再識別される可能性，又は属性を推定される可能性を計測する指標が多く提案されている．

その中でも ***k*-匿名性は，千田ら[9]が発表した海外事例（以後，「海外事例報告」という）のように，欧州や米国を中心に，匿名化データの再識別リスクの検証のための指標として採用されている．しかし，南らの発表[10]によって，*k*-匿名化処理アルゴリズムの知識を利用したアルゴリズムベースの情報漏洩の問題が指摘されている．

それらの安全性指標群は，パーソナルデータに対する攻撃手法によって区分することができる．Fung らは[11]にて，プライバシー保護データパブリッシング（PPDP : Privacy Preserving Data Publishing）における脅威モデルを定義した．図 1 は，パーソナルデータを匿名化処理してデータ利用者に提供する概念図である．

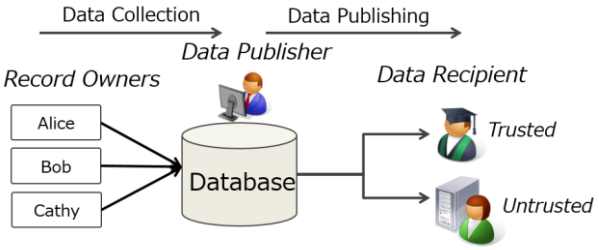

図 1 PPDP への攻撃モデル範囲

このとき，Untrusted な利用者が Record Owner の情報を取得するための攻撃手法をレコード結合（Record linkage），属性結合（Attribute linkage），テーブル結合（Table linkage），確率的攻撃 （Probabilistic Attack）の 4 種類と定義し，それぞれに有効な安全性指標と匿名化処理アルゴリズムを整理した．

しかし，安全性を高める匿名化処理によって，そのデータの有用性を損なうことがある．特に，***k*-匿名化処理は，

元データの属性値に対して一般化や統合，削除などを行うことから，属性が増加するに従って有用性が低下する「次元の呪い[12]」も問題点として指摘されている．

有用性の指標を元データと匿名化データの値における差，と定義する場合，情報量（エントロピー）の比較や，Information Loss[13]などの情報間の距離を計算する方法が提案されている．一方，数値属性だけでなく，カテゴリ属性を用いた場合の指標も提案されており，Prec[14]やDIS[15]等，一般化階層を用いた際の抽象化レベルを計測する方式が提案されている．

2.4 PWSCUP の有用性・安全性指標

前項で検討した有用性・安全性指標は，ユースケースによって適用可能な指標とその定義が異なることから，事業者ごとに作成した匿名化データの定量的な比較が難しいという課題がある．

そこで PWSCUP では，ユースケースに合致した有用性指標を考案し，匿名化データの有用性と安全性を計測した．その際に，安全性指標にはレコード結合による再識別リスクを，参加者同士で検証する形式で実施することで，実際の攻撃者を想定したルールとしている．

表 1 に PWSCUP の概要を示す．

マスターデータはミクロデータとも呼ばれ，名簿のように 1 ユーザの属性を 1 レコードで示すものである．CUP´16 では，1 ユーザが複数回登場するトランザクションデータと組み合わせて競技が行われた．

表 1　PWSCUP 概要

項目	CUP´15	CUP´16
データ形式	マスターデータ型	マスターデータ＋トランザクションデータ型
対象データ	擬似ミクロデータ[16]を利用した家計の年間消費額	UCI データセット[17]を利用した購買履歴
有用性指標	・ある属性の平均絶対誤差 ・クロス集計値の平均絶対誤差 ・クロス集計数の平均絶対誤差 ・SA の相関係数の平均絶対誤差 ・データ各値の平均絶対誤	・RFM 分析の誤差 ・バスケット分析の誤差 ・クロス集計表の誤差
安全性指標	・***k***-匿名性 ・同値類サイズの平均値 ・参加者による再識別数	・参加者による再識別数

PWSCUP における有用性指標は，匿名化データを受けた利用者が対象データを用いて行う分析手法を想定したものとした．

安全性指標については，CUP´15 では実行委員が属性を設定し ***k***-匿名性に相当する指標と再識別数を採用した．CUP´16 では，実行委員が作成したサンプル再識別処理とコンテスト参加者による再識別数を用いて評価した．

この再識別数に関する定義は，*k*-匿名性などの安全性定義と攻撃者の知識レベルが異なる点に特徴がある．そこで，次の章より，まず *k*-匿名性における攻撃者の知識レベルの想定方法について検討し，PWSCUP との比較を行う．

3. *k*-匿名性と再識別率の関係性

3.1 *k*-匿名化処理と全属性同値類化処理

匿名化データの安全性リスクの指標として用いられている *k*-匿名性を満たすための処理「*k*-匿名化処理」について整理する．まずパーソナルデータの用語定義を行う．パーソナルデータは，属性（Attribute）と値（Value）で構成されており，属性は以下の何れかに分類できる．

1.　正識別子（ID）：個人を一意に識別できる属性．
2.　準識別子（QID）：間接的に個人を識別できる属性．
3.　センシティブ属性：正識別子，準識別子以外で，個人のプライバシーに関するもの．
4.　非センシティブ属性：上記以外の属性．

この内，3.と 4. について，本稿ではプライバシー上の軽重を評価しないため**「対象属性（Objective Attribute：OA）」**と両者を統合して呼称する．

k-匿名化処理は，QID について，同じ値を持つユーザが少なくとも ***k*** 人（***k***>1）存在する「同値類（Equivalence Class）」を作成する．その際に OA の値から個人を識別されず，かつデータ分析の目的変数としても利用できることを求め，QID と OA の選定は**匿名化データの提供者が判断する．**

OA における再識別攻撃を考慮せず，QID のみを用いた再識別攻撃が行われると仮定した場合，***k***-匿名化処理された匿名化データの再識別リスクは 1/***k*** 以下となる．

本稿では，この *k*-匿名性における再識別確率を「1/*k* 再識別率」と表記し区別するものとする．

しかし，同一の OA の複数回開示や，時系列で変化する OA を逐次開示することによって，OA から個人が識別されるリスクがあると Sweeny は指摘[8]している．

即ち ***k***-匿名性は QID のみを知る部分知識攻撃者を想定して提案された指標である．そのため，QID と OA には明確な情報としての区分は存在しない．

OA と QID を区分する方法について，米国などでは，識別子（これは組み合わせて利用する準識別子 QID を含む）としてそのまま利用してはいけない属性を定めている．

以下に米国 HIPAA 法（Health Insurance Portability and Accountability Act of 1996；医療保険の携行性と責任に関する法律）によって定められた，個人識別性が高く，公開されることが望ましくない属性群を示す[19].

参考：HIPAA の定める，識別子となりうる 18 属性
1. 氏名　2. 住所　3. 日付（登録日，誕生日等）　4. 電話番号　5. FAX 番号 6. メールアドレス　7. 社会保障番号　8. 医療記録番号　9. 健康保険番号 10. 銀行口座番号　11. 証明書，ライセンス番号　12. 自動車等の免許番号 13. 通信端末番号やシリアル番号　14. Web の URL　15. IP アドレス 16. 生体認証データ　17. 顔等が判別できる写真　18. その他，他者と区別するために作られた識別子全般

表 2 は，QID に対して ***k***-匿名性を用いた一般的な ***k***-匿

名化処理の例である．外見から判明する可能性の高い「性別」属性と，HIPAA の定める 18 属性に含まれる「誕生日」属性から導かれる「年齢」属性を QID として利用し，調査対象である「体重」と「購入品」を OA と設定した．これによって「OA の知識が無い」と判断された部分知識攻撃者に対しては，再識別されないための匿名化処理を正しく行ったと言える．

表 2　*k*-匿名化処理の例

ID	QID		OA			
	性別	年齢	体重	購入品1	購入品2	購入品3
A	男	20代	92kg	マンガ	アダルトビデオ	ビジネス誌
B	男	20代	58kg	アダルトビデオ	ビジネス誌	ゲーム
C	女	10代	48kg	ファッション誌		
D	女	10代	51kg	料理レシピ	マンガ	
E	女	10代	72kg	マンガ	ファッション誌	料理レシピ

しかし「体重」属性や「購入品」属性は，公表されている場合や見た目等で類推できるため，再識別リスクが完全に消失していると定義することが難しい．

例えば，公開された表の中にスポーツ選手や芸能人等，体重や身長等の身体的特徴が公表されている群が含まれている場合，それら外部データとの対照によって個人が識別される可能性がある．それらの外部するデータベースを「照合可能なデータベース」であると定義したとき，攻撃者がどこまで部分的な OA の知識を保持しているか，想定することは困難である．

攻撃者に部分知識が存在することを前提とした場合，大きく 2 つの対応方法がある．***k*-匿名性を強化するために QID と OA を結合させる方法と，*l*-多様性や *t*-近傍性などに代表される Attribute Linkage（属性推定）を防止する方法である．

表 3 に QID と OA を結合して同値類を形成することで，個人識別性を排除した処理の例を示す．本稿ではこれを**「全属性同値類化」**処理とする．

ID	QID		OA			
	性別	年齢	体重	購入品1	購入品2	購入品3
A	男	20代	Avg. 75kg	アダルトビデオ	ビジネス誌	
B	男	20代	Avg. 75kg	アダルトビデオ	ビジネス誌	
C	女	10代	Avg. 57kg	雑誌		
D	女	10代	Avg. 57kg	雑誌		
E	女	10代	Avg. 57kg	雑誌		

表 3　全属性同値類化処理の例

全属性同値類化がされた匿名化データは，個人に対する確定的な識別はできない．しかし，全属性同値類化を行った場合でも，**攻撃者は OA を使った識別ができない**という前提条件が破られたとき，体重や購入品に対する部分知識を用いて個人の属性が推定されるリスクが存在する．

例えば表 3 において，体重をミクロアグリゲーションによって代表値（平均値等）に変化させた場合，ある男性の体重が 58kg であることを知っていた場合，残り 1 名 92kg の男性を識別することができる．

また，女性 2 名が存在し，その 2 名が同じく 50kg 程度であることを知る攻撃者がいた場合，残る 1 名は 70kg 程度の女性である推定が可能となる．

このように，攻撃者が持つ OA に関する知識量を完全に定義することは困難であり，全ユーザの再識別リスクを排除することは難しい．

3.2 *k*-匿名性における再識別リスクの考え方

また，***k*-匿名性は，匿名化データ全体の安全性指標として利用できない点も課題である．*k*-匿名性は「最大」の再識別リスクを持つユーザ群に関する安全性指標であるため，その他のユーザに関する安全性を問わない．

図 2 の例は分布が異なるデータであるが，いずれの系列も ***k*-匿名性では *k*=10 であり，系列 1～3 の 1/*k* 再識別率はすべて 1/10=10%となる．

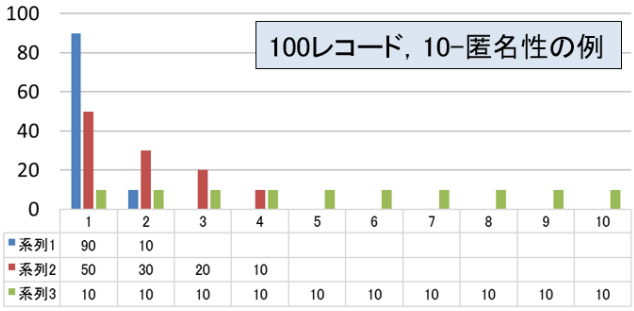

図 2　100 レコード，10-匿名性の分布例

Emam ら[20]によると，このような場合の安全性評価については，***k*-匿名性だけでなく，平均再識別リスクも評価指標として有効である．

例えば，攻撃者が全体の再識別を試みるのではなく，ある 1 人のユーザの再識別を試みる場合は，同値類サイズから導く，平均再識別率によるリスク評価を推奨している．

表 4 に平均再識別率を安全性の基準とする例を示す．ある個人情報を匿名化処理する際，QID として性別（男性，女性）を設定し，等価クラス（A，B）となる匿名化データが作成された．このとき元データを知る攻撃者ならば，少なくとも等価クラス A に所属しているユーザ ID の 1 名を選択し，等価クラスに対して全て同じ ID を入力する再識別攻撃が有効である．例では，***k*-匿名性は”2”であり，***k*-匿名性平均値は”3”である．全体に対して一律値を入力する再識別攻撃を受けた場合の再識別数は”2”，再識別率は”1/3”（33%）となり，全ての等価クラスサイズの平均値の逆数となる．

表 4　*k*-匿名化処理と再識別の例

ID.	性別（QID）	等価クラス	同値化数	再識別攻撃	再識別結果
1	男性	A	2	1	**HIT**
2	男性			1	-
3	女性	B	4	3	**HIT**
4	女性			3	-
5	女性			3	-
6	女性			3	-
安全性指標		***k*-匿名性　2** ***k*-匿名性平均　3**		再識別数 2	

本再識別攻撃手法は，アルゴリズム等を用いることなく，元データを入手できる攻撃者の場合，容易に達成できるものである．そのため，この値を超える範囲で再識別が可能になった場合，効果的な再識別アルゴリズムであると判断できる．

3.3 海外事例報告における運用例

***k*-匿名性を用いた運用例として，カナダの国・地方政府と民間の保健情報を扱う機関 Canadian Institute for Health Information （CIHI） においては，攻撃者となりうるデータ要求者の目的と能力，知識レベルの範囲，及び安全管理措置をマトリクス化してランク付けし，情報の機微性と合わせてリスク評価を行い，提供条件を定める運用を行っている． 図 3 の攻撃者のマトリックス評価を用いて，3 軸の安全性評価を行い，その提供条件を求める．

1) **目的と能力：提供先企業の再識別攻撃可能な知識の想定とそのモチベーションの有無．**
2) **軽減コントロール：提供先企業のデータ保持に関する安全管理体制の評価**
3) **プライバシー侵害の潜在リスク：データの詳細度や属性値の機微性，開示後影響の評価．**

先ずデータ要求者の 1）目的と能力，そして 2）軽減コントロールについてレベル分けを行う．

目的の判断要素として，データ提供者との関係，再識別による財政的な利得，再識別することの非財政的な理由の有無等が挙げられる．能力については，再識別に関する技術専門性，財源，関連性のある他のデータベースへのアクセス可能性等が挙げられる．

軽減コントロールはプライバシーやセキュリティの実践が良ければ High，悪ければ Low となる．そして図 5 左表 に基づき，データ要求者が再識別を試みる確率を Remote（最も低い），Occasional，Probable，Frequent（最も高い） の 4 段階で評価する．

次に，3）プライバシー侵害の潜在リスクを 3 段階にレベル分けする．判断要素として，データの詳細度や機微性，意図的でないまたは許可されていない利用や後続の開示があったときの個人の損害度合い，データ要求者のロケーション（データ提供者との管轄域の差異）等が挙げられる．

最後に図 3 右表 に基づき，***k*-匿名性の *k* の値となるリスクの閾値を 4 段階で求める．

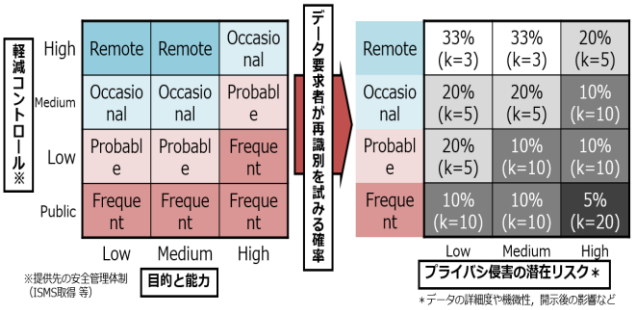

図 3　攻撃者のマトリクス型評価の例

このように，本質的に ***k*-匿名性などの同値類を用いてユーザのプライバシーを侵害させない仕組みの有効性は，再識別を試みる攻撃者の知識レベルと元データの機微性やリスクに依存するため，評価者が正確な攻撃者想定を行えることが前提となっている．**

3.4 CUP´16 における再識別手法

***k*-匿名性を安全性評価として運用する場合の課題として，その評価手法が一定でないことが挙げられる．

まず，QID と OA の区分は，評価者による想定で決まるため，表 3 における「体重」属性が QID になるのか OA になるのかは，評価者によって異なる．加えて，攻撃者の目的と能力を設定する場合でも，攻撃者を一般人と規定するか，再識別のアルゴリズムに関する知識を持っているかによって想定が異なる．

そのため，同じ匿名化データを同じ企業が提供する場合においても，安全性評価の基準は一定しない．そこで，攻撃者の持つ知識の想定を同一にすることで，同基準で評価できる仕組みが求められるようになった．

PWSCUP は，Domingo-Ferrer が想定した最大知識攻撃者モデル[13]を参考に攻撃者を設定し，元データの QID と OA に加えて，その有用性，安全性指標に関する知識を保持した攻撃者モデルを新たに策定した．これを**「PWS 知識」**を持つ攻撃モデルとする．

PWSCUP の想定するデータ提供と攻撃者のユースケースを図 4 に示す．

まず，データ提供者が個人情報 X の匿名加工データ Y を生成し，行番号データ I^Y を廃棄してデータ利用者に提供する．

それに対して，X を入手可能な攻撃者である再識別コンテスト参加者が，Y の再識別をデータ利用者から依頼されたため，X と Y を対照し，かつ，その匿名化処理に用いら

れた安全性や有用性の指標を得た上で再識別を試みるという処理である．

PWSCUP における攻撃者は，元データとなる QID と OA を全て保持した上で再識別攻撃を行うため，OA 値に対して単純な加工を行うだけでは，容易に再識別されることになる．

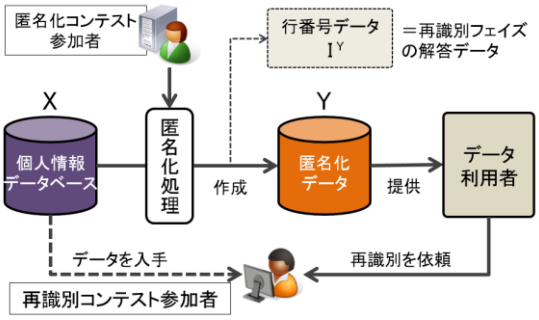

図 4　PWS 知識を持つ攻撃者想定ユースケース

OA を用いた再識別アルゴリズムの例を図 5 に示す．元データ X と匿名化データ Y を比較し，同じ同値類を持つ値を確認した上で，X と Y の OA 値をソートし，その順番に行番号を求めることで再識別を行う．

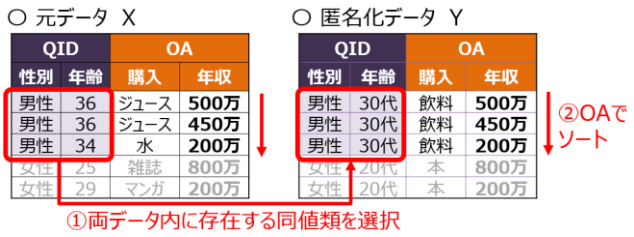

図 5　OA ソート型再識別処理の例

また，PWS 知識攻撃者においては，その処理に必要となる匿名化アルゴリズムも類推できるため，同じアルゴリズムを用いることで容易に元データへの復元が可能な場合もある．

PWSCUP は，このようなユースケースを想定することで，図 4 で示したような攻撃者の想定を最も強く設定した場合の漏洩リスクを評価することが可能となる．

図 6 に，CUP′16 における再識別率の分布を示す．CUP′16 のデータセットとコンテストルール下における再識別率は 22.3%～100%の範囲で分布している．

例えば，CUP′16 において最も再識別率が低いデータは，再識別率 22.3%である．これと，k-匿名性における 4-匿名状態（$k>4$，$1/k$ 再識別率 25%）は攻撃者想定と再識別の定義が異なるため，直接接続ができない．そのため，匿名化データの作成者は，k-匿名化処理だけでなく，更に履歴まで踏み込んだ処理が必要とされる．

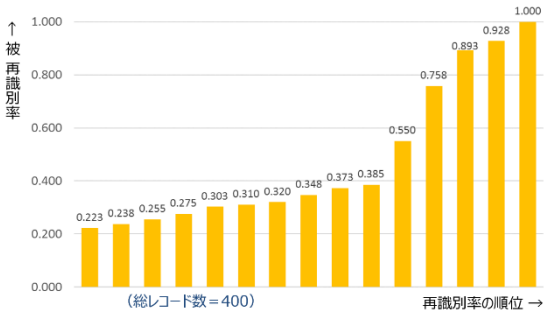

図 6　CUP′16 本戦における再識別率の分布

このような，QID，OA，安全性・有用性評価に関する知識を持つ攻撃者に対抗するには，ID の仮名化，QID の同値類化，OA による再識別を避けるための処理，の 3 種類の加工が適切に行われている必要がある．

表 5 は CUP′16 にて求められる加工の例である．この加工は以下の 3 要素を満たした加工を行っている．

1) ID の仮名化：不可逆ハッシュ処理などによって元 ID に戻せない，かつ ID ソート攻撃にも対処する．

2) QID の同値類化：近い属性値を持つレコード同士について同値類化する．

3) OA の加工：QID の同値類に付属する OA について，数値属性は集計値が近似するようにソートに配慮しつつ配分し，購入品も同値類におけるバスケット分析結果の影響を抑えて配分する．

表 5　CUP′16 にて求められる加工例

ID	QID		OA			
	性別	年齢	体重	購入品1	購入品2	購入品3
XI3	男	20代	60kg	マンガ	アダルトビデオ	ビジネス誌
Z9Q	男	20代	90kg	ビジネス誌	アダルトビデオ	マンガ
J8K	女	10代	53kg	料理レシピ	ファッション誌	
N29	女	10代	62kg	料理レシピ	マンガ	
JTE	女	10代	56kg	マンガ	ファッション誌	

これらの処理によって単純なソート攻撃では再識別されない匿名化処理が実現できる．CUP′16 では，いずれかの要素についての加工が不十分な場合に容易に再識別されるように，再識別プログラムを実装した．

有用性と安全性の両立を目指すため，OA に対して誤差が少ない値の交換（Swapping），レコードの入れ替え処理（Shuffle），履歴の順列変換（Permutation）などの処理が多く用いられた．

その結果として，元データ，安全性・有用性評価という PWS 知識を持つ他の参加者においても，確率的にしか再識別できない匿名化データを生成することができた．それが PWSCUP における「再識別率」の意味である．

表 6 に CUP′16 における元データと，本戦において最も優秀であったデータを有用性指標のクロス集計誤差 U_1-*Cmae1* に関して抽出した結果の比較を例として示す．

表 6　CUP´16 最優秀データと元データの比較

国名	最優秀データ 性別 f	性別 m	平均単価	元データ 性別 f	性別 m	平均単価
Australia	1	26	2.767	5	1	2.767
Austria	2	18	4.154	7	3	4.154
Bahrain	2	6	6.364	1	1	6.363
Belgium	2	2	3.618	17	5	3.619
Brazil	17		3.212	1		3.212
Cyprus	1		2.327	5		2.325
Czech Republic	42		1.232	1		1.233
Denmark	8	1	4.264	5	1	4.263
Finland	1	4	4.807	11	1	4.806
France	5	18	3.525	69	17	3.525
Germany	3	1	3.608	68	17	3.607
Greece	5		2.307	2		2.307
Iceland	6		1.753	1		1.754
Israel	1		1.785	3		1.785
Italy	2	8	4.180	12	1	4.180
Japan	43	7	3.081	6	2	3.269
Lebanon	9		4.388	1		4.388
Netherlands	8	3	2.775	6	2	2.772
Norway	4	7	3.984	8	2	3.984
Poland	8	18	4.553	5	1	4.553
Portugal	5	2	3.551	18	1	3.551
RSA	18		2.856	1		2.856
Saudi Arabia	3		1.824	1		1.824
Singapore		12	4.060		1	4.061
Spain	4	19	5.443	26	2	5.441
Sweden	8		1.940	7		1.941
Switzerland	2	3	3.938	11	5	3.937
United Arab	4		2.084	1		2.084
United Kingdom	15	3	2.872	27	7	2.872
USA	13		1.457	4		1.457
***k*-匿名性**	1			1		
***k*=1 の数**	6			16		
CUP´16 U_1-Cmae1	0.00443			0.00000		
CUP´16 再識別率	22.3%（89 人）			100.0%（400 人）		

U_1-Cmae1 では，国名と性別でクロス集計を行い，平均単価の誤差を求めている．平均単価は（総購入額 / 総購入個数）を示しており，その値の平均絶対誤差 MAE は 0.00443 である．

一方，この両データに対して単純な安全性指標を適用した場合，例えば ***k*-匿名性**を適用した場合は k=1 であり，個人を 1 名まで絞りこむこと（シングルアウト）が可能である．

しかし，最優秀データと元データとのユーザ属性の分布は大きく異なる，例えば Australia の m の出現数は元データ 1 名に対し，最優秀データは 26 名である．対象となる 1 名の存在確率が変化しているため，シングルアウトできることが確定的に判明していない．

CUP´16 の有用性評価には，各クロス集計における属性値の出現数は含まれていないが，その平均単価が含まれている．そのため，属性値の交換を行っただけでは有用性評価が下がらない．そこで平均単価が近いレコードとの値交換を行うことで，有用性を維持している．

値交換によって属性値の分布が変化したことから，属性及び履歴からの再識別も困難となり，安全性が向上している．

元データをそのまま投入した場合，OA ソート型の攻撃によって 100%の再識別率となるが，最優秀データでは，実行委員が作成した再識別プログラムや参加者による再識別攻撃を受けた結果でも，全体の再識別成功率は 22.3%に抑えることができた．

匿名化データの特性に応じた再識別アルゴリズムによる安全性評価は，現状では安全性指標として定式化されておらず，今後の研究が必要な分野である．

3.5 安全性基準のまとめ

・PWSCUP では，元データ及び安全性指標と有用性指標の PWS 知識を持った参加者が，OA を含む全データを用いた再識別攻撃の結果を計測する手法を採用した．参加者は，再識別サンプルプログラムを匿名化アルゴリズムに合わせてチューニングができる能力と手法を持つ技能者．

・ガイドラインは匿名化処理を行う事業者が,「一般人の能力と手法」により再識別できないことを担保した安全性を求めている．

・マニュアルは，***k*-匿名性**を考慮しながら市販のデータベースや提供先が照合可能なデータを用いて再識別ができないことを基準として安全性を求めているが，能力や手法に関する記述はない．

・海外事例報告にあった CIHI では ***k*-匿名性**を用いて許容リスクを判定するが，攻撃者（提供先）の目的と能力，安全管理評価，プライバシー侵害リスクを加味したマトリクス評価のプロセスを経た上で決定している．

表 7 に匿名化データに関する安全性の考え方をまとめた．

表 7　安全性基準の考え方

参照先	攻撃者想定	安全性指標	OA からの攻撃
PWSCUP	PWS 知識を備えた技能者	再識別数	あり
ガイドライン	元データを参照できる一般人	記載なし	記載なし
マニュアル	市販データベース，または提供先でデータを参照できる者	***k*-匿名性を考慮**	履歴であっても必ずしも無加工で利用できるものではない
海外事例報告	攻撃者の知識・能力・モチベーションを評価して ***k*-匿名性を決定**		

事業者は匿名加工情報の作成,利活用にあたりガイドラインに準拠して運用することになるが，安全性基準については PWSCUP または海外事例報告等を参照し，基準に沿った運用プロセスを設計する必要がある．

4. 社会実装に向けた指標提案

海外事例報告の CIHI のプロセスでは，匿名化データの提供にあたってデータ要求案件ごとに攻撃者の知識想定と情報の機微性を合わせたリスク判定を行っている．民間事業者にとって，専門家による合理的なリスク判定を行うための体制整備と運用システムの構築には，匿名化データの利活用に向けた大きな課題である．

攻撃者の知識として PWS 知識を設定することによって，攻撃者の知識想定が必要なくなり，情報の具体的な機微性

（含まれている属性値にセンシティブな要素が無いか, 等）だけを判定することで，全体的な安全性の評価が可能となる．

そこで，我々は PWSCUP で用いられた再識別率を用いて匿名化データの情報類型を定める案をまとめた．

表 8 に匿名化データの情報類型案を示す．項 1～3 は匿名加工情報を中心に，安全管理措置として匿名化処理を行った個人情報，有用性も意識した集計情報等に対して，類型ごとに再識別率の閾値を設定した例である．データセットや利用目的によって閾値を設定し，再識別率をコントロールした匿名化処理を行うことにより，ニーズに沿った匿名化データの生成ができるものと期待できる．

項 4 は，外形的に判断できる十分な匿名化への加工方法と検定の指標案である．

各指標の**閾値** s_1，s_2，s_c の設定にあたり，CIHI 等を参考にしながらリスク評価と運用プロセスについてデータ提供者とのコンセンサス形成を行う必要がある．同時に，検定に用いる再識別率の算定アルゴリズムもユースケースの拡大に合わせて研究が必要となる．

表 8　匿名化データの情報類型案

項	情報類型	指標案	加工方法案	検定案
1	個人情報	再識別率 s_1 **以上**	CUP'16 のような匿名化手法 ***k*-匿名化など**	属性ソート等プログラムによる再識別率算定
2	匿名加工情報	再識別率 s_1 **未満**		
3	集計情報等	再識別率 s_2 **未満**		
4	集計情報等	最小クラスタサイズ s_c 以上	全属性同値類化	最小クラスタサイズ

今後，多くのデータ形式やユースケースに対する考察を深めながら，事業に活かせるデータ活用の議論が進展することを期待する．

参考文献

[1] 個人情報の保護に関する法律及び行政手続における特定の個人を識別するための番号の利用等に関する法律の一部を改正する法律(平成 27 年法律第 65 号)
[2] 個人情報保護委員会, "個人情報の保護に関する法律についてのガイドライン(匿名加工情報編)", [http://www.ppc.go.jp/files/pdf/guidelines04.pdf], (2016).
[3] 経済産業省, "事業者が匿名加工情報の具体的な作成方法を検討するにあたっての参考資料(「匿名加工情報作成マニュアル」)Ver1.0", [http://www.meti.go.jp/press/2016//08/20160808002/20160808002-1.pdf], (2016).
[4] 菊池 浩明,山口 高康,濱田 浩気,山岡 裕司,小栗 秀暢,佐久間 淳, "匿名加工・再識別コンテスト Ice & Fire の設計",コンピュータセキュリティシンポジウム 2015 論文集, 2015(3), pp.363-370,2015-10-14
[5] 菊池 浩明, 小栗 秀暢, 野島 良, 濱田 浩気, 村上 隆夫, 山岡 裕司, 山口 高康, 渡辺 知恵美, "PWSCUP: 履歴データを安全に匿名加工せよ", コンピュータセキュリティシンポジウム 2016 論文集, 2016(2), pp.271-278, (2016).
[6] L.Sweeney, "k-anonymity: a model for protecting privacy", International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, pp.557-570, 2002
[7] 五十嵐 大, 千田 浩司, 高橋 克巳, "k-匿名性の確率的指標への拡張とその適用例", コンピュータセキュリティシンポジウム 2009 (CSS2009) 論文集, 2009 pp.1-6, (2011).
[8] Machanavajjhala, A., Kifer, D., Gehrke, J. and Venkitasubramaniam, M., "l-diversity: Privacy beyond k-anony
[9] 千田 浩司, 吉浦 裕, 島岡 政基, "匿名化基準に関する欧米公的文書 7 選の考察", コンピュータセキュリティシンポジウム 2016 論文集, 2016(2), pp.158-165, (2016).
[10] 南 和宏, 千田 浩司, "再識別リスク評価と匿名化の展望", コンピュータセキュリティシンポジウム 2016 論文集, 2016(2), pp.166-172, (2016).
[11] Fung, B., Wang, K., Chen, R. and Yu, P.S., "Privacy-preserving data publishing: A survey of recent developments", ACM Computing Surveys (CSUR), 42(4), pp.14, (2010).
[12] Aggarwal, C.C., "On k-anonymity and the curse of dimensionality", Proceedings of the 31st international conference on Very large data bases, pp.901-909, (2005).
[13] Domingo-Ferrer, J. and Torra, V., "A quantitative comparison of disclosure control methods for microdata", Confidentiality, disclosure and data access: theory and practical applications for statistical agencies, pp.111-134, (2001).
[14] Sweeney, L., "Guaranteeing anonymity when sharing medical data, the Datafly System.", Proceedings of the AMIA Annual Fall Symposium, pp.51, (1997).
[15] 村本 俊祐, 上土井 陽子, 若林 真一, "データを極小歪曲し k-匿名性を保持したデータに変換するプライバシー保護アルゴリズム", 日本データベース学会 letters, 6(1), pp.97-100, (2007).
[16] 秋山 裕美, 山口 幸三, 伊藤 伸介, 星野 なおみ, 後藤 武彦, "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)
[17] Chen, D., Sain, S.L. and Guo, K., "Data mining for the online retail industry: A case study of RFM model-based customer segmentation using data mining", Journal of Database Marketing & Customer Strategy Management, 19(3), pp.197-208, (2012)
[18] Domingo-Ferrer, J., Ricci, S. and Soria-Comas, J., "Disclosure risk assessment via record linkage by a maximum-knowledge attacker", Privacy, Security and Trust (PST), 2015 13th Annual Conference on, pp.28-35, (2015).
[19] Health Insurance Portability and Accountability Act(HIPAA), "Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule", U.S. Department of Health & Human Services, (2012).
[20] Deng, M., Wuyts, K., Scandariato, R., Preneel, B. and Joosen, W., "A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements", Requirements Engineering, 16(1), pp.3-32, (2011).
[21] El Emam, K. and Arbuckle, L., "Anonymizing health data:case studies and methods to get you started", O'Reilly Media, Inc., (2013)