プライバシーリスク分析手法の提案

三本 知明† バス オニバン† 清本 晋作†

†株式会社 KDDI 研究所
356-8502 埼玉県ふじみ野市大原 2-1-15
{to-mimoto, basu, kiyomoto}@kddilabs.jp

あらまし

パーソナルデータを有効に活用するにあたり，プライバシーは重要な問題であり，k-匿名化，l-多様化といった手法により，データを加工することでプライバシーの保護を実現している．匿名化はプライバシーを保護する一方で，過度な匿名化は匿名化されたデータの有用性を低下させるため，適切なリスク分析を行い，最適な匿名化を行う必要がある．本研究では既存の攻撃者の知識を仮定したリスク分析手法を応用し，より現実的なパーソナルデータのリスク分析が可能な分析手法を提案する．分析手法は，リスクに基づいた開示可能な知識の推定や，攻撃者のリソースを考慮したリスク分析など，状況に応じた複数の手法を提案する．

A proposal for privacy risk analysis methods

Tomoaki Mimoto† Anirban Basu† Shinsaku Kiyomoto†

†KDDI R&D Laboratories Inc.
2-1-15 Ohara, Fujimino, Saitama 356-8502, JAPAN
{to-mimoto, basu, kiyomoto}@kddilabs.jp

Abstract

Personal data contain private and sensitive information. When data owners publish such data, they use a well-known technique called k-anonymity and achieve privacy protection. This process reduces the probability of re-identification of the records, but on the other hand the data availability decrease. In this paper, we apply an existing risk analysis method which can quantify risk of re-identification by assuming the attacker knowladge and propose more practical methods. We can decide scope of disclosure from the risk and infer the attacker knowladge from his resource.

1 はじめに

ICT 分野の発展に伴い，ビッグデータの利活用に注目が集まり，現在数多くの研究が行われている．データを二次利用するにあたり，プライバシーの保護を考えることは必要不可欠であり，その方法の一つとして匿名化がよく知られている。しかしデータの匿名性と有用性はトレードオフの関係にあるため，過剰な匿名化はデータの有用性を下げることとなる．また，あらゆる状況において最適な匿名化手法というものは存在せず，データの種類やそのデータの使い道など様々な状況の違いに応じて匿名化の手法を考える必要がある．加えて，従来の匿名性の評価においては，攻撃者があらゆる知識を持つという仮定が前提としてあったが，現実にはそのよ

うな状況は考えにくく，攻撃者の保持する背景知識について考察する必要もある．

本研究では既存のアルゴリズムを応用し，複数の状況に応じたプライバシリスクの分析手法を提案する．リスク分析を行うことにより，データセットのリスクの原因となるレコード，属性を分析・評価することで，適切な匿名化手法を決定するための一つの指標となる．我々が提案するリスク分析手法は大きく三通りあり，状況に応じてデータセットを追加，変更することでより細かいリスク分析を行うことが可能である．また，データセットに各属性の取得コストを付与することで，攻撃者のリソースを元に攻撃者がどのような背景知識を持ちうるかを推定する手法や，データセットの各属性の匿名化レベルを変更することで，最適なデータセットを推定する手法も本稿に含める．

提案手法の一つでは攻撃者の背景知識を仮定し，その仮定のもとリスク分析を行う．この手法と既存のアルゴリズムとの違いは出力部分にあり，これは追加，変更したデータセットに対してもリスク分析を行えるよう単純な構成とした．また，もう一つの提案手法では，許容されるリスクを仮定し，そのリスクを上回らないような属性の組合せを推定，例示することが可能である．これは外部に公開する基準として，個人が再識別されるリスクの上限が決まっている場合などに，データセットをどのように加工するかを決定するための手助けとなる．三つ目の提案手法では，単独では脅威とならないが，別の属性と組み合わせることで個人識別リスクが大幅に上がる恐れのある属性の推定，例示が可能である．これにより，脅威となる属性の組合せを同時に公開しないことで潜在的なリスクを上げないような施策をとることが可能となる．これらのリスク分析手法に加えてデータセットの追加，変更を行うことで，その場合のリスクの変化や，各属性の取得にかかるコストを考慮した場合の攻撃者の動作の変化を知ることが可能である．

以下では本稿の構成を説明する．本稿では 2 章で今回扱う表記や攻撃者モデルなどの前提，および関連研究について，続く 3 章で今回扱うアルゴリズムの元となる既存アルゴリズムについて述べる．4 章で提案手法の詳細を説明し，5 章でそれらの評価を行う．そして 6 章で本稿のまとめを述べる．

2 準備

2.1 表記

ここでは本稿で扱う記号について説明する．

- D: 属性数が u，レコード数が v のデータセットをあらわす．本稿では k-匿名化されたデータセットとして扱う．
- d: D と同じ属性を持つレコード数 w のデータセットをあらわす．
- D_d: D に d を追加したデータセットをあらわす．
- D_C: D の各属性を取得する際に必要なコストを追加したデータセットをあらわす．
- F : D の各属性のデータの丸め度合い (匿名化のレベル) を定義したファイルをあらわす．
- D_F : F により D をの一部，あるいは全部の属性の匿名化レベルを変更したデータセットをあらわす．
- $ATTR = \{attr_1, ..., attr_u\}$: データセットの属性の集合をあらわす．
- $ATTR_a = \{attr'_1, ..., attr'_l\} (l \leq m)$: データセットの属性の部分集合をあらわす．本稿では攻撃者の背景知識をあらわす．
- $r_p (1 \leq p \leq v)$: データセットにおける p 行目のレコードをあらわす．D_d の場合は $1 \leq p \leq v + w$ となる．
- $r_p^{attr_q}$: データセットのレコード r_p における属性 $attr_q$ のカラムをあらわす．
- C_{attr_q}: $attr_q$ を取得する際に必要なコストをあらわす．

2.2 攻撃者モデル

提案手法において，攻撃者は外部情報から背景知識として一つ，あるいはそれ以上の評価対象のデータセットの属性 $ATTR_a \subseteq ATTR$ を保有するものとする．攻撃者はこの背景知識を用いて各レコードの背景知識に相当する属性のカラムを検査し，各属性のカラムが同一のレコードが存在しなければ，個人を再識別できたこととなる．また同一のレコードが k 個存在する場合，個人を再識別できる確率，すなわちリスクは $\frac{1}{k}$ となる．

2.3 関連研究

匿名化とはプライバシー情報を含むデータを加工することでユーザを特定できないようにする処理である．ユーザがどの程度特定できないかという度合いを示す指標は様々あり，よく知られているものとして k-匿名性 [1] がある．匿名化処理にあたり，データセットの属性は，識別子，準識別子，機密属性に分類することができる．識別子とは，名前や国民 ID などそれだけで個人を特定できる属性であり，準識別子とは，それだけでは個人を特定できるとは限らないが，他のデータセットにあらわれる可能性があり，それらを紐付けると個人を特定できる可能性がある属性である．また，機密属性とは病名や購買履歴などユーザが個人と紐付けられた状態で公開されたくないような属性のことである．k-匿名性とは，任意の一つ以上の準識別子を選択した時に，同じ準識別子の組合せを持つレコードがデータセットに k 個以上存在することを保証する指標である．匿名性の指標は，k-匿名性以外にも，l-多様性 [2] や t-近似性 [3]，m-不変性 [4] などが知られている．

k-匿名化では上述の通り，同じ準識別子の組合せを持つレコードがデータセットに k 個以上存在することを保証するが，実際の攻撃においては各属性を知識として得るにはコストがかかり，攻撃者がすべての準識別子を背景知識として持つとは考えにくい．そのため，実際の攻撃に対しては，過度の匿名化がされている場合がある [5][6].

3 既存アルゴリズム

ここでは 4 章で利用するアルゴリズムについて述べる．

3.1 indexRepeats()

アルゴリズム 1 では，入力としてデータセット D，及び攻撃者の背景知識 $ATTR_a$ を与える．その後 HashMap を用いて，$\forall r_p \in D$ に対し，$r_p^{attr'_q}(q = 1, ..., l)$ を連結した文字列 K_p をキーとして出現回数を保存する．同様のアルゴリズムはハッシュテーブルを用いても実装可能だが，衝突処理について考慮の必要がなく，検索，挿入にかかる計算量も $O(K_p)$ と比較的小さい Radix tree を採用している．

Algorithm 1 `indexRepeats`$(D, ATTR_a)$: indexing repeats in attribute values.

Require: The $k-$anonymised dataset D, the set of known attributes $ATTR_a$

1: Repeat detector $RT \leftarrow$ empty radix tree.
2: **for** $\forall r_p (p = 1, ..., n) \in D$ **do**
3: $K_p \leftarrow r_p^{attr'_1} || r_p^{attr'_2} || ... || r_p^{attr'_l}$.
4: **if** $K_p \in RT$ **then**
5: RT.put$(K_p, 1 + RT$.get$(K_p))$.
6: **else**
7: RT.put$(K_p, 1)$.
8: **end if**
9: **end for**
10: **return** RT

3.2 Attack simulation algorithm

アルゴリズム 2 では，攻撃者の背景知識 $ATTR_a$ を選び，`indexRepeats`$(D, ATTR_a)$ を用いて統計的なリスク分析を行う．様々な攻撃者の背景知識を仮定して個人の識別されるリスクを評価することで，実際の攻撃におけるリスクがどの程度あるかを定量的に判断することが可能である．

既存研究 [6] では，匿名化処理を行った UCI Adult dataset[8] に対して，攻撃者の背景知識

を変更しながらリスク分析，および評価を行っている．ここで扱うデータセットは全部で 13 の属性を持ち，準識別子として age, workclass, education, marital-status など 8 つの属性を持つものしている．図 1 では，攻撃者がすべての準識別子を背景知識として保有すると仮定した場合のアルゴリズム 2 の処理結果を表している．この結果から，すべての攻撃に対して個人を再特定できるリスクは 0.02 以下であるが（k-匿名性），そのうちの 8 割においては，実際のリスクは 0.01 以下しかないということが読み取ることができる．

Algorithm 2 Attack simulation with arbitrary attributes.

Require: The k-anonymised dataset D, an arbitrary set of attributes $ATTR_a$

1: $RT \leftarrow$ `indexRepeats`$(D, ATTR_a)$.
2: **for** $\forall r_p \in D$ **do**
3: $n \leftarrow RT$. get (K_p).
4: $\Pr(re\_id|K_p) \leftarrow \frac{1}{n}$.
5: **end for**
6: **return** Cumulative Distribution of $\Pr(re\_id|K_p)$ for all $ATTR_a$.

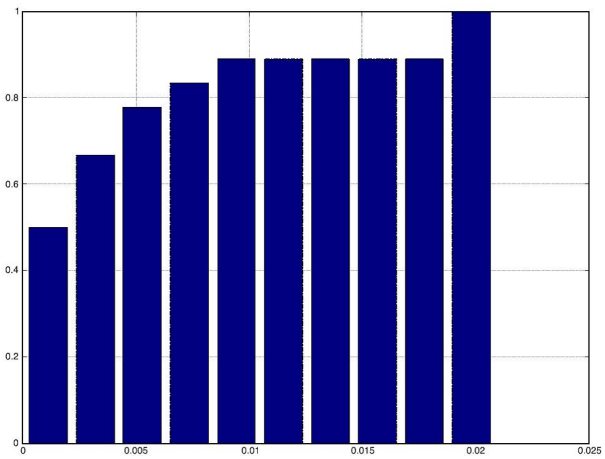

図 1: Risk as probability of re-identification versus attack success with the attacker's knowledge of all the quasiidentifiers.

4 提案手法

本研究では既存研究同様 `indexRepeats`$(D, ATTR_a)$ を用いて，データセットを追加した場合のリスク変化の評価や，データセットの各属性の取得コストを考慮した場合における攻撃者の背景知識の推定，許容可能なリスクを仮定した場合の最適なデータセットの推定など，複数のリスク分析・評価手法を提案する．各分析・評価手法の詳細を以下に述べる．

4.1 想定された攻撃に対するリスク評価手法

想定された攻撃に対するリスク評価手法は，既存研究と同等の評価方法であるが，データセットの追加，変更に対応するため出力を一部変更する．本手法では，入力としてデータセット D(あるいは D_d, D_C, D_F) と攻撃者の背景知識 $ATTR_a$ を与え，すべての $r_p^{attr_q}(attr_q \in ATTR_a)$ が一致するレコード数の逆数をリスクとして出力する．データセットが D_d の場合は，入力を D，D_d と分けてそれぞれの場合で評価を行い，そのリスクの差分を明示して出力する．また，D_C の場合はその所要コストを含めて出力し，D_F の場合は D_d の場合と同様，入力を D，D_F と分けて評価を行い，そのリスクの差分を明示して出力する．詳細はアルゴリズム 3 の通り．

Algorithm 3 Attack simulation with arbitrary attributes.

Require: The k-anonymised dataset D, an arbitrary set of attributes $ATTR_a$

1: $RT \leftarrow$ `indexRepeats`$(D, ATTR_a)$.
2: **for** $\forall r_p \in D$ **do**
3: $n \leftarrow RT$. get (K_p).
4: $\Pr(re\_id|K_p) \leftarrow \frac{1}{n}$.
5: **end for**
6: **return** $p, \Pr(re\_id|K_p)$ for all K_p.

4.2 リスクに基づく開示可能な知識推定手法

リスクに基づく開示可能な知識推定手法では，入力としてデータセット D(あるいは D_d , D_C, D_F)，攻撃者が少なくとも持つ背景知識 $ATTR_a$，及び許容可能なリスク N を与える．なお，$ATTR_a = \phi$ でも構わない．本手法では，網羅的にすべての属性の組み合わせに対して想定された攻撃に対するリスク評価手法にてリスクを算出し，そのうちの入力で与えた許容可能なリスクを上回らないような属性の組み合わせを出力する．データセットが D_d の場合は，入力を D，D_d と分けてそれぞれの場合で評価を行い，属性の組み合せの差分を明示して出力する．属性の組み合せの差分とは，例えば d を加えることでリスクが変化した $ATTR_a$ や，リスクが変化したことにより，出力対象となった $ATTR_a$ をあらわす．また，D_C の場合はその所要コストを含めて出力し，D_F の場合は D_d の場合と同様，入力を D，D_F と分けて評価を行い，属性の組合せの差分を出力する．詳細はアルゴリズム 4 の通り．

Algorithm 4 Attributes analysis based on the risk

Require: The k-anonymised dataset D, an arbitrary set of attributes $ATTR_a$, allowable risk N

1: **repeat**
2: **while** $N \leq n$ **do**
3: $RT \leftarrow$ `indexRepeats`$(D, ATTR_a)$
4: $n \leftarrow RT$. get (K_p).
5: $\Pr(re\_id|K_p) \leftarrow \frac{1}{n}$
6: $ATTR_a \leftarrow ATTR_a \cup \exists attr_q \in ATTR$
7: **end while**
8: **return** $ATTR_a, p, \Pr(re\_id|K_p)$ for all K_p
9: **until** all combinations of $ATTR_a$ are calculated

4.3 リスクが高い知識の組み合わせ推定手法

リスクが高い知識の組み合わせ推定手法では，入力としてデータセット D(あるいは D_d, D_C , D_F)，組み合わせる知識数 $x(\geq 2)$，及びリスク幅の基準値 y を与える．本手法ではまず $|ATTR_a| = x$ となるような $ATTR_a$ を選択する．その $ATTR_a$ とデータセットに対して，想定された攻撃に対するリスク評価手法を用いてリスクを算出し，またすべての $attr_q \in ATTR_a$ とデータセットに対しても同様にリスクを算出する．その後属性を組み合わせる前 $attr_q$ と後 $ATTR_a$ でどの程度リスクが変化したか（リスク幅）を計算し，リスクの基準値 y を超える組合せを出力する．データセットが D_d の場合は，入力を D，D_d と分けてそれぞれの場合で評価を行い，リスク幅の差分を明示して出力する．また，D_C の場合はその所要コストを含めて出力し，D_F の場合は D_d の場合と同様，入力を D，D_F と分けて評価を行い，リスク幅の差分を明示して出力する．詳細はアルゴリズム 5 の通り．

4.4 コストを考慮した攻撃者知識推定手法

コストを考慮した攻撃者知識推定手法では，入力としてデータセット D の各属性にコストを与えたデータセット D_C，攻撃者が少なくとも持つ背景知識 $ATTR_a$，及び攻撃者のリソース R_a を与える．なお，$ATTR_a = \phi$ でも構わない．本手法では，$attr_q \notin ATTR_a$ を $ATTR_a$ に追加し，想定された攻撃に対するリスク評価手法を用いてリスクを算出するという動作を所要コストが R_a を超えるまで繰り返す．最終的に所要コストが R_a 以下となる属性の組み合わせを出力する．詳細はアルゴリズム 6 の通り．

なお，各属性の所要コストは属性ごとに独立ではない場合があり，また，時間や環境によって変動するため，その決定は容易ではないが，本研究では各属性の所要コストは既知として議論しないこととする．

Algorithm 5 High risk combination of attributes analysis

Require: The k-anonymised dataset D, the number of the knowladges x, the criterion of the risk y

repeat
  for x times **do**
    $ATTR_a \leftarrow ATTR_a \cup \exists attr_q \in ATTR$
    $RT \leftarrow$ `indexRepeats`$(D, attr_q)$
    $n_q \leftarrow RT$. get $(attr_q)$.
    $\Pr(re\_id|attr_q) \leftarrow \frac{1}{n_q}$
  end for
  $RT \leftarrow$ `indexRepeats`$(D, ATTR_a)$
  for $\forall r_p \in D$ **do**
    $n \leftarrow RT$. get (K_p).
    $\Pr(re\_id|K_p) \leftarrow \frac{1}{n}$
  end for
  $z \leftarrow \frac{\Pr(re\_id|K_p)}{min(\Pr(re\_id|attr_q))}$
  if $z > y$ **then**
    return $ATTR_a, p, \Pr(re\_id|K_p), z$
  end if
until all combinations of $ATTR_a$ are calculated

Algorithm 6 Attributes analysis based on the cost

Require: The k-anonymised dataset D_c, an arbitrary set of attributes $ATTR_a$, the resourse of the attacker R_a

repeat
  for $R_a \geq C_{ATTR_a}$ **do**
    $ATTR_a \leftarrow ATTR_a \cup \exists attr_q \in ATTR$
  end for
  $RT \leftarrow$ `indexRepeats`$(D_c, ATTR_a)$.
  for $\forall r_p \in D_c$ **do**
    $n \leftarrow RT$. get (K_p).
    $\Pr(re\_id|K_p) \leftarrow \frac{1}{n}$.
  end for
until all combinations of $ATTR_a$ are calculated
return $ATTR_a, p, \Pr(re\_id|K_p)$ for all K_p.

4.5 最適匿名化レベル推定手法

最適匿名化レベル推定手法では，入力として D，匿名化レベルを定義したファイル F，攻撃者が持つ背景知識 $ATTR_a$，及び許容可能なリスク N を与える．F は，例えば住所であれば level1 は丁目・番・号単位，level2 は町域単位，level3 は市区町村単位，level4 は都道府県単位というものを想定する．本手法では，F を用いて元のデータセット D の属性 $ATTR_a$ の匿名化のレベルを変更しながら，想定された攻撃に対するリスク評価手法を用いてリスクを分析し，許容可能なリスクに収まるようなデータセットを出力する．詳細はアルゴリズム 7 の通り．

Algorithm 7 Optimization dataset analysis

Require: The k-anonymised dataset D, the anonymous definition file F_D an arbitrary set of attributes $ATTR_a$, allowable risk N

repeat
  while $N \leq n$ **do**
    $D_F \leftarrow$ change the level of $attr_q \in ATTR_a$
    $RT \leftarrow$ `indexRepeats`$(D_F, ATTR_a)$
    $n \leftarrow RT$. get (K_p).
    $\Pr(re\_id|K_p) \leftarrow \frac{1}{n}$
  end while
  return $D_F, p, \Pr(re\_id|K_p)$ for all K_p
until all combinations of the level of $attr_q \in ATTR_a$ are calculated

5 評価

ここまで複数のリスク分析手法を提案したが，その多くが繰り返しの試行が必要である．繰り返しの回数はデータセットの属性数 u に対して高々$\sum_{i=1}^{u} {}_uC_i$ であり，属性数に依存する．また上記のアルゴリズムでは，木の構築は攻撃者の背景知識を決定した後に行われるため，繰り返しの回数だけ木の構築を行う必要がある．Radix tree ではデータの挿入に $O(K_p)$ かかるため，属性数が多い場合は，網羅的にすべての属性の組み合わせに対してリスク分析を行うことが困難

となる．したがって，その場合はすべての組合せに対してリスク評価を行うのではなく，ランダムに選んだ複数の組み合せパターンに対してリスク評価を行い，統計的に結果を得る必要がある．そのため属性数に関わらず，網羅的にすべての属性の組み合せに対してリスク分析を行うには，アルゴリズムを変更して木の構築回数を減らす手法を検討する必要がある．

また最適匿名化レベル推定手法に関しては，$|ATTR_a| = l$ とし，匿名化レベルの段階数をそれぞれ S とすると，S^l 回データセットを変更する必要がある．したがって本手法についても，攻撃者の背景知識，匿名化レベルの段階数によっては他の手法同様網羅的なリスク分析ではなく，ランダムに選んだ複数の組み合わせパターンに対してリスク評価を行い，統計的に結果を得る必要がある．

6 おわりに

本稿では，既存のアルゴリズムを用いた複数のリスク分析手法を提案した．提案したリスク分析手法を用いることで，様々な状況に応じた現実的に考えられるリスクの原因となるレコードや属性の推定が可能となり，より適切な匿名化を行うための一つの指標となる．また，最適匿名化レベル推定手法においては，許容リスクを超えないようなデータセットを生成することで，最適な匿名化の提案が可能である．今後の展望としては，木の構築回数を減らし，高速処理できるようなアルゴリズムの検討を進めるとともに，更に他の状況を想定したリスク分析手法を提案を進める．

謝辞

本研究は CREST-JST の支援の下に行われたものである。

参考文献

[1] L. Sweeney. ”k-anonymity: a model for protecting privacy” *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, 10 (5), pp. 557-570, 2002.

[2] A. Machanavajjhala, D. Kifer, J. Gehrke and M. Venkitasubramaniam. ”l-Diversity: Privacy Beyond k-Anonymity” *ACM Transactions on Knowledge Discovery from Data(TKDD)*, vol.1, no.1, p.3, 2007

[3] N. Li, T. Li, and S. Venkatasubramanian. ”t-closeness: Privacy beyond k-anonymity and l-diversity” *ICDE 2007*, pp.106-115, 2007.

[4] X. Xiao and Y. Tao. ”m-invariance: towards privacy preserving re-publication of dynamic datasets” *SIGMOD2007*, pp.689-700, 2007.

[5] A. Basu, A. Monreale, J. C. Corena, F. Giannotti, D. Pedreschi, S. Kiyomoto, Y. Miyake, T. Yanagihara, and R. Trasarti, ”A privacy risk model for trajectory data” in *Proceedings of the IFIP WG 11.11 International Conference on Trust Management(IFIPTM)*, Singapore, 2014.

[6] A. Basu, T. Nakamura, S. Hidano and S. Kiyomoto, ”k-anonymity: risks and the reality, Accepted for publication” in *the IEEE International Symposium on Recent Advances of Trust, Security and Privacy in Computing and Communications (RATSP, collocated with the IEEE TrustCom)*, Helsinki, 2015.

[7] S.Kiyomoto, Y.Miyake and T. Tanaka, ”Privacy Frost: A User-Oriented Data Anonymization Tool” in *Sixth International Conference on Availability, Reliability and Security, ARES 2011*, Vienna, 2011.

[8] M. Lichman, “UCI machine learning repository,” 2013. [Online]. Available: http://archive.ics.uci.edu/ml