6W-03

匿名加工における攻撃者知識と安全性に関する評価

林　弘悦†　　西山　賢志郎‡　　土井　洋‡　　趙　晋輝††
中央大学大学院†　　情報セキュリティ大学院大学‡　　中央大学††

1. はじめに

個人の情報を含むデータを加工することでプライバシーを侵害することなく分析，利用可能な形式に変換する処理を匿名加工と呼ぶ．匿名加工されたデータの安全性を評価するためには，適切な攻撃者モデルの設定は不可欠である．攻撃者モデルとしては，例えば最大知識攻撃者モデル[1]が使われる場合がある．一方，攻撃者モデルを変えた場合は安全性への影響が生じるはずで，それを評価しておく必要がある．本稿では，匿名加工データにおける安全性と攻撃者モデルの関係性について，検討結果と実験評価を示す．

2. 匿名加工と攻撃者モデル

本稿では，有限個のレコード(行に対応)と属性(列に対応)からなるテーブル形式のDBを扱う．テーブルの各セルに格納されている値は属性値と呼ばれる．属性には有限種類の属性値分類を格納するカテゴリ属性や，整数や実数値などを格納する数値属性がある．

最大知識攻撃者モデル[1]では，匿名加工データの安全性を評価する際に，攻撃者がオリジナルテーブルT全体と匿名加工テーブルT′全体を知識として持ち，2つのテーブルの個体同士を対応づける(マッチング)攻撃を想定しており，最も強い攻撃者モデルである．

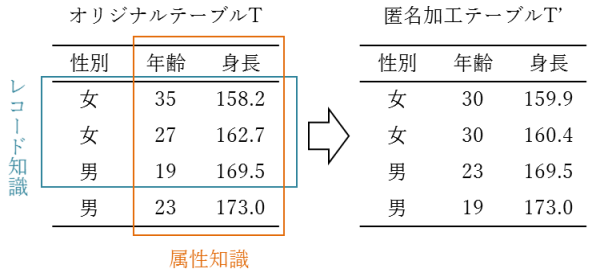

図1 テーブル匿名加工の例

この攻撃者モデルは最悪の場合を想定している．しかしながら，現実では攻撃者がすべての情報を有した状態で対応づけを図ることは考えにくい．よって，最大知識攻撃者だけを想定するのではなく，攻撃者の持つ知識を変化させ，実際の攻撃に対する安全性を測定することが望ましい．

本稿では，最大知識攻撃者モデルにおける攻撃者の知識を弱めた以下3種類の攻撃者知識モデルを設ける．

[部分レコード知識モデル]
攻撃者はTの全て(N_0個)のレコードのうちN_0個のレコードのみを知識として持つ．

これはテーブル内に攻撃者の親しい知人の情報が含まれている場合などに当てはまる．例えば図1のTにおける全てのレコード($N = 4$)のうち，先頭3レコード($N_0 = 3$)を知識として持つことが考えられる．

[部分属性知識モデル]
攻撃者はTのM個の属性のうちM_0個の属性のみを知識として持つ．

これはテーブル内の攻撃者にとって入手しやすい属性情報と入手しづらい属性情報が明確である場合に当てはまる．例えば属性総数が3個である図1のTでは，{年齢，身長}の2属性を知識として持つことが考えられる．

[部分テーブル知識モデル]
Tの内のNレコード×M属性のうち部分集合$N_0 \times M_0$個のレコード，属性のみを知識として持つ．

これは部分レコード知識モデル，部分属性知識モデル双方よりも弱い攻撃者を想定している．

あまり親しくない知人やweb上から収集した他人の情報を知識として持つ攻撃者に当てはまる．例えば図1のTでは{年齢，身長}の2属性について先頭3レコードを知識として持つことが考えられる．

3. 評価方法

2章で示した知識モデルにおいて，攻撃者の知識が弱まるほど攻撃者がマッチングに成功する確率が減り，安全性が高まることが期待できる．しかし，知識制限と安全性について定量評価は行われていない．本稿では，背景知識が変化した場合の安全性について実験を行う．モデルによる差異を定量的に評価することが目的の一つであるため，匿名化手法とマッチング手法として簡潔なものを採用する．

[実験の流れ]Tに匿名加工を行い，T′を得る．T′とTの部分知識を用いてマッチングを行う．マッチングの正解率に基づく指標を評価する．

[実験データ]UCI Machine Learning Repositoryで公開されるDiabetic Retinopathy Debrecen Data Setの属性番号1〜20の内，無相関な4つの数値属性(属性番号3, 9, 17, 18)を用いる．

[匿名加工]Tの標準偏差σに対し，平均0，標準偏差$p \times \sigma$となる正規分布に従うランダムノイズ加算を用いる．本稿では$p = 0.05, 0.1$とする．

[マッチング方法]攻撃者が知識として有するレコード(Tの一部)を匿名加工テーブルT′のレコードへ対応付けを行う以下の2つの方法を用いる．

1. ユークリッド距離が最小なレコード同士を対応づける手法(距離ベースマッチング)．
2. T, T′内で各属性を降順でソートし，同じ順位のレ

A study on the relationship between attacker's background knowledge and privacy for anonymization
†Hiroyoshi HAYASHI, Graduate school of Science and Engineering, Chuo Graduate School
‡Kenshiro NISHIYAMA, Hiroshi DOI, Institute of Information Security University
††Jihui CHAO, Chuo University

コードを対応づける手法(順位ベースマッチング).

[評価指標]本稿では[3]を参考にした以下の2つの指標を用いる.

$$\text{entire} = \frac{\text{マッチング正解数}}{\text{T}'\text{全体のレコード数}}$$

$$\text{restricted} = \frac{\text{マッチング正解数}}{\text{知識レコード数}}$$

これらの指標は，0に近いほど安全な匿名加工がなされていると解釈できる． 実際，TとT′の個体同士の対応付けに失敗しているからである.

- 部分レコード知識モデルでは$N_0 = N \times r$として，rの値を50%~100%まで10%刻みで試行した.
- 部分属性知識モデルでは4つの数値属性3，9,17,18について全組み合わせを試行した.

4. 評価

実験結果を図2,3,4,5及び表1に示す． 図2,3,4,5における高さは指標値である． 横軸は知識属性の組合せであり，左の方ほど知識属性数が大きい．奥行きは知識レコード数であり，奥の方ほど知識レコード数が大きい.

(1) 評価指標entireについての評価

距離ベースマッチングについて，指標entireは知識レコード数が大きければ大きく，小さければ小さくなると予想していた.

知識属性数が4の場合は予想通りである． 知識属性数が小さいときは知識レコード数の影響を受けにくいことが今回の実験データの特徴である(図2の知識属性数が1の場合，図3の知識属性数が1または2の場合).

実験データに依存するが,知識属性数が大きい場合はノイズ加算だけでなく，置換やレコーディングなどの他の匿名加工手法を用いる必要がある.

順位ベースマッチングはpを0.05や0.1にしてもマッチングは成功しなかった． 更にノイズを小さく，具体的には$p = 0.01$とした結果を表1に示す． 単一属性での結果から，今回実験で用いたデータは少しでもノイズを与えると順位が変動するデータであると言える.

(2) 評価指標restrictedについての評価

知識レコード数が変化しても，指標値は変化しないと予想した． しかし，実験結果では知識レコード数の増加に伴ってrestricted指標値が単調に少しずつ増加している． 知識レコード数が増えるほど正確なマッチングができていた.

5. まとめ

本稿では，従来の攻撃者モデルよりもより現実的な想定を行い，攻撃者の有する知識を制限したモデルを提案した． 評価実験により提案モデルでの安全性の性質を得た． より詳細な実験と，安全性と有用性の関係性についての調査が今後の課題である.

6. 参考文献

[1] J. Domingo-Ferrer and K. Muralidhar. New directions in anonymization: permutation paradigm, verifiability by subjects and intruders, transparency to users. CoRR, abs/1501.04186, 2015.

[2] 伊藤，村田，高野． ミクロデータにおける匿名化技法の適用可能性の検証―全国消費実態調査と家計調査を用いて―，統計研究彙報，第71号，pp.83-124, 2014.

[3] 菊池，小栗，野島，濱田，村上，山岡，渡辺． PWSCUP: 履歴データを安全に匿名加工せよ．CSS2016集，pp.271-278, 2016.

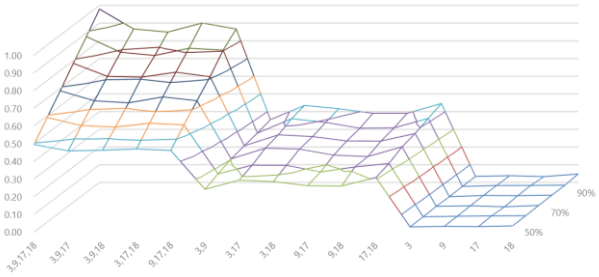

図2 指標entire，$p = 0.05$，距離ベースマッチング

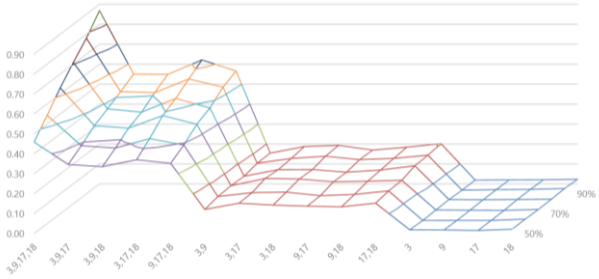

図3 指標entire $p = 0.1$，距離ベースマッチング

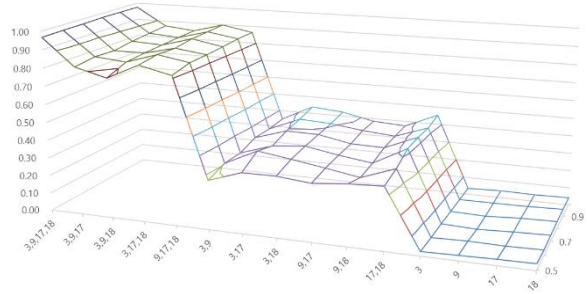

図4 restricted，$p = 0.05$，距離ベースマッチング

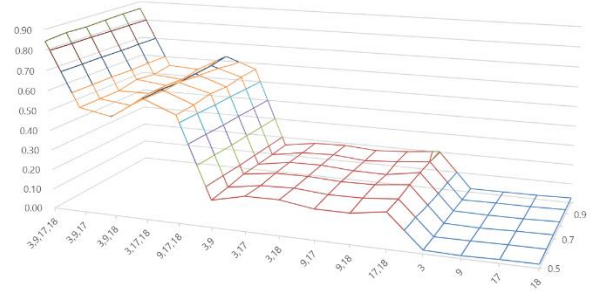

図5 restricted，$p = 0.1$，距離ベースマッチング

表1 entire，$p = 0.01$，順位ベースマッチング

	知識属性数														
	4	3				2						1			
	3,9,17,18	3,9,17	3,9,18	3,17,18	9,17,18	3,9	3,17	3,18	9,17	9,18	17,18	3	9	17	18
ノイズなし	0.99	0.99	0.99	0.99	0.98	0.99	0.99	0.99	0.98	0.98	0.98	0.05	0.98	0.97	0.93
ノイズ p=1%	0.12	0.15	0.14	0.10	0.14	0.14	0.09	0.09	0.14	0.14	0.15	0.10	0.14	0.14	0.17