SCIS 2022 2022 Symposium on
Cryptography and Information Security
Osaka, Japan & Online, Jan. 18 – 21, 2022
The Institute of Electronics,
Information and Communication Engineers

属性推定攻撃を考慮する匿名化データの安全性指標の提案
A Privacy Measure of Anonymized Data against Attribute Inference Attack

紀伊 真昇 * 市川 敦謙 * 三浦尭之 * 芝原俊樹 *
Masanobu Kii Atsunori Ichikawa Takayuki Miura Toshiki Shibahara

キーワード 匿名化，匿名データ合成，安全性指標，リスクコミュニケーション，リスクアセスメント，統計的決定理論

あらまし

本稿ではk匿名性 [Fun+10] や差分プライバシー [Dwo06; 寺田 19] とは異なる背景思想をもつ，匿名化技術の安全性を測る数値指標を提案する。提案指標は，安全な匿名化済みデータとは被害者への攻撃行動（差別，偏見等の不当な扱いや何らかの形の暴力）を行う動機を攻撃者に与えないデータである，という思想を出発点としている。ではいつ攻撃行動が行われるかと言えば,「被害者の属性はこれだ（だから攻撃するべきだ）」という確信の度合いと攻撃行動にかかるコストを攻撃者が比較し，攻撃者が前者が後者より大きいと判断したときに攻撃行動が行われる，と考えることが出来る。本研究では「被害者の属性はこれだ」という確信の度合いを統計的決定理論で用いられるベイズ因子を用いて定式化し，安全性指標を構成した。

ベイズ因子を用いて定義されているため，データの形式にかかわらず利用できる。ただしデータが取りうる値が無限に有るような場合は事前分布の選択に注意が必要である。本稿ではこの問題に対するいくつかの対応策を述べる。

提案指標と差分プライバシーの関係として次がわかっている：差分プライバシーと同じ攻撃者の設定のもとで，差分プライバシーを達成する匿名化メカニズムは提案指標での安全性も達成している。この関係は差分プライバシーに新しい解釈を与えるため，重要なものと思われる。

* 日本電信電話株式会社 NTT 社会情報研究所, 〒 180–8585 東京都武蔵野市緑町 3–9–11
NTT Social Informatics Laboratories, 3–9–11 Midori-cho, Musashino-shi, Tokyo 180–8585, Japan.

参考文献

[Dwo06] Cynthia Dwork. “Differential Privacy”. In: *Automata, Languages and Programming*. Ed. by Michele Bugliesi et al. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2006, pp. 1–12. ISBN: 978-3-540-35908-1. DOI: 10.1007/11787006_1.

[Fun+10] Benjamin C. M. Fung et al. “Privacy-Preserving Data Publishing: A Survey of Recent Developments”. In: *ACM Computing Surveys* 42.4 (June 23, 2010), 14:1–14:53. ISSN: 0360-0300. DOI: 10.1145/1749603.1749605. URL: https://doi.org/10.1145/1749603.1749605 (visited on 06/01/2021).

[寺田 19] 寺田 雅之. “差分プライバシーとは何か”. In: **システム／制御／情報** 63.2 (2 2019), pp. 58–63. DOI: 10.11509/isciesci.63.2_58.