差分プライバシの実現手法と利用についての調査

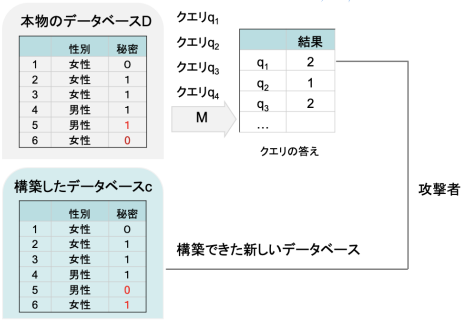
A Survey of Mechanisms and Applications on Differential Privacy

胡若寧・ネットワーク分科会・情報セキュリティ大学院大学

In some cases, the original database information can be constructed from publicly available statistical information. This creates a risk of personal information being leaked when disclosing or analyzing statistical information. In fact, the U.S. Census has attempted this kind of rebuilding attack and found it to be risky. Protecting the privacy of individuals has become an issue in utilizing data. Differential privacy techniques can reduce the risk of information leakage by adding design noise to statistical data. We started researching what differential privacy is, what the mechanisms are, what the benefits of differential privacy are, and how to assess the security and usefulness of differential privacy.

1、研究背景

公開されている統計情報から、元々のデータベース情報を構築できることがあります。



これにより、統計情報を開示または分析する際に個人情報が漏洩するリスクが生じます。実際に、米国国勢調査はこのような再建攻撃を試み、その危険性を示しました

アメリカセンサスは2018年に内部の実験を実施し、結果的に2010年の国勢調査において、約5200万人の情報が正しく特定されていたことが明らかになった。

国勢調査には、年齢、性別、人種などの情報が含まれており、攻撃者は公開された統計情報をもとに、もとのデータベースの年齢や人種などの情報を再構築することができた。

2010年国勢調査データにおいて約5, 200万人、すなわち米国の総人口の約17%が正しく再特定された

2020年のアメリカセンサスでは差分プライバシーが採用された。

2、先行研究

差分プライバシの技術を用いることで、統計データに設計ノイズを加えることによって、情報漏えいのリスクを軽減することができます。差分プライバシは Dwork で提案された[1], ランダム化関数M: $\mathcal{D} \rightarrow \mathcal{R}$ が下の式を満たすとき，Mは ε - 差分プライバシを満たします。

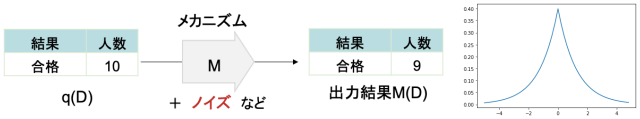
$$Pr[M(\mathbf{D}) \in S] \leq e^{\varepsilon} \cdot Pr\,[M(\mathbf{D}') \in S]$$

D, D′ ∈ $\mathcal{D}$, S は Mの出力集合$\mathcal{R}$ の任意の部分集合

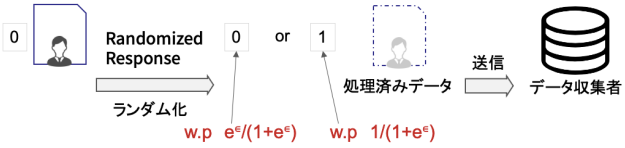
[1]Dwork, C. : Differential Privacy, Proc. 33rd Intl. Conf. Automata, Languages and Programming - Volume PartII,LNCS, Vol.4052, pp.1-12 (2006)

差分プライバシとその実現手法について説明し、ラプラスメカニズムとランダムレスポンスメカニズムについても検討し、有用性を評価します。[2]

1)ラプラスメカニズム



2)ランダムレスポンスメカニズム



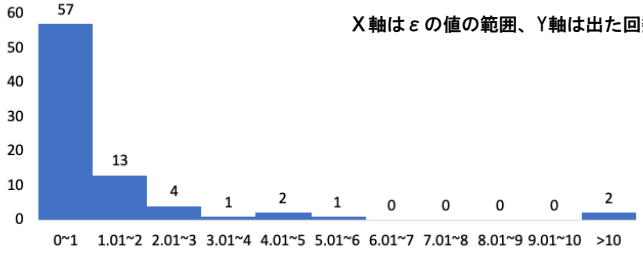
3) ノイズの評価

ε=0.01		ε=0.5		ε=2		ε=10	
results	noises	results	noises	results	noises	results	noises
-77	-87	10	0	11	1	10	0

1000回実行し、出力結果の分析は表に示される

安全性良くなる	epsilon	クエリの値	尺度(b)	出力の平均	ノイズの標準偏差	有用性良くなる
	0.01	10	100.0	4.047	137.922679	
	0.50	10	2.0	9.889	2.774595	
	2.000	10	0.5	9.956	0.740684	
	10.000	10	0.1	10.002	0.077473	

また、個人情報の保護指標である ε について、企業や研究で実際に使用されている ε の値や、 ε を決定する方法についても調査し、結果を報告します。



ε の設定のまとめ

[2] Cynthia Dwork, Aaron Roth: The Algorithmic Foundations of Differential Privacy; Foundations and Trends in Theoretical Computer Science, Vol.9, Nos3-4,PP211-407 (2014)