



匿名化によるデータセットの情報損失を測る指標

データ間距離を用いた情報損失指標

技術分野分類 22010：土木材料・施工・建設マネジメント

技術キーワード 4：複合材料・新材料

産業分類 E-29：その他の製造業

内 容	概要	データセットに含まれるデータを置き換えることにより失われる情報の量を測る指標。匿名化情報の評価などに応用ができる。																																			
	従来技術・競合技術との比較（優位性）	従来の情報損失指標は、適用できるデータセットの種別が限定的である、置き換えたデータとの差異を表していない、などの問題があった。本指標は、データ間の距離さえ設定すればあらゆるデータセットに対して計算することができる汎用性が大きな利点である。下図のように、数値と非数値の両方の属性を含むデータセットの匿名化へも適用可能である。																																			
	本技術の有用性	匿名化アルゴリズムの評価や、生成された匿名化情報の評価に用いることができる。パーソナル情報に限らず適用できるので、データを置き換える操作に対して広く活用できる。																																			
関連情報 (図・表・写真等)		<table border="1" style="display: inline-table; vertical-align: middle;"> <thead> <tr> <th>年齢</th> <th>職業</th> </tr> </thead> <tbody> <tr><td>24</td><td>小学校教員</td></tr> <tr><td>45</td><td>プログラマ</td></tr> <tr><td>34</td><td>高校教員</td></tr> <tr><td>55</td><td>カフェ店員</td></tr> <tr><td>52</td><td>SE</td></tr> <tr><td>61</td><td>販売員</td></tr> <tr><td>41</td><td>幼稚園教員</td></tr> </tbody> </table> → <table border="1" style="display: inline-table; vertical-align: middle;"> <thead> <tr> <th>年齢</th> <th>職業</th> </tr> </thead> <tbody> <tr><td>20代</td><td>教員</td></tr> <tr><td>40代</td><td>接客業</td></tr> <tr><td>30代</td><td>教員</td></tr> <tr><td>50代</td><td>技術職</td></tr> <tr><td>50代</td><td>技術職</td></tr> <tr><td>60代</td><td>接客業</td></tr> <tr><td>40代</td><td>教員</td></tr> </tbody> </table>	年齢	職業	24	小学校教員	45	プログラマ	34	高校教員	55	カフェ店員	52	SE	61	販売員	41	幼稚園教員	年齢	職業	20代	教員	40代	接客業	30代	教員	50代	技術職	50代	技術職	60代	接客業	40代	教員	データセットをいくつかのグループに分割し、同じグループ内のデータを同一の値に置き換える匿名化手法をマイクロアグリゲーションといい、各グループが k 個以上のデータを含む時、データセットは k -匿名性を持つという。 k -匿名化はビッグデータ活用のため広く研究されている。		
年齢	職業																																				
24	小学校教員																																				
45	プログラマ																																				
34	高校教員																																				
55	カフェ店員																																				
52	SE																																				
61	販売員																																				
41	幼稚園教員																																				
年齢	職業																																				
20代	教員																																				
40代	接客業																																				
30代	教員																																				
50代	技術職																																				
50代	技術職																																				
60代	接客業																																				
40代	教員																																				
適用可能製品		人の位置情報、購買履歴、駅の乗降履歴などのパーソナル情報を用いた個人的及び社会的サービスを提供できる。また、分析・解析により、学術的や産業的に有益な情報を生成できる。																																			
技術シーズ 保有者	氏名 所属・役職	秋山 寛子 電子情報工学科・助教																																			
技術シーズ 照会先	窓口 TEL/FAX e-mail	地域共同テクノセンター 026-295-7117 / 026-295-7124 nrtc71@nagano-nct.ac.jp																																			

■知的財産

■試作品状況

無

提示可

提供可

作成日 2018年11月16日