

位置情報プライバシーの統計的安全性

統計数理研究所 南和宏

要旨

スマートフォンの普及に伴い、我々の位置情報の取得が容易になり、多くのユーザーの移動履歴は、交通情報の提供、都市設計といった社会サービス、また商圈分析等の商用ビジネスにも活用されている。その一方、位置情報から、個人のプライバシーに関する情報が漏洩する危険性が懸念されている。したがって位置情報の安全な2次利用には、匿名化と呼ばれる個人の識別情報を取り除くデータ加工が不可欠である。

しかし位置情報の匿名化処理では、氏名等の個人の識別子を削除するのみでは不十分であり、一般的には位置情報から k 未満のユーザーに絞り込むことを防ぐ k -匿名化処理を行う。ただし位置情報の時系列データには、個人の行動習慣、移動経路の制約等を反映した時空間の相関性が存在し、 k -匿名化した位置情報から元の軌跡の復元されるリスクが存在する。

本発表では、一般に広く普及している k -匿名化技術を位置情報に適用した場合の問題点を指摘し、ユーザーの移動パターンをマルコフチェーンでモデル化し、匿名化技術の安全性を状態空間モデルにおける内部状態の推定問題として評価する手法を紹介する。統計的な安全性の実現には、匿名化加工処理におけるランダム性の導入が不可欠であり、位置情報の匿名化加工における今後の技術的課題を議論する。