

タクシーデータの安全な提供に関する研究

A study on how to provide taxi data safely

橋本真・法制倫理分科会・情報セキュリティ大学院大学

Nowadays, due to the wide usage of low price communication network, more data is available than before. Data can be gathered online from site to remote servers and can be analyzed in very shorter time. We use taxi as the model in our research. A study was made on how to protect the privacy of taxi users and how to protect the secret of taxi companies while collecting/processing the data from taxis.

1. 研究の概要

本研究はタクシーをモデルとし、タクシーで発生したデータを収集・加工する段階でタクシー利用者のプライバシー保護やタクシー業者の秘密を守る方法について検討した。

2. 研究対象としてのタクシー



- 特徴
 - ・移動体である
 - ・無線通信やMobile通信が通信手段として確立されている。
 - ・発生するデータの種類の豊富
 - 取引データ
 - 位置データ
 - 決済手段
 - ...
 - ・中小企業が多く、個人タクシーの場合、全事業者数の8割以上占めている。非常に競争の激しい業界である。

本研究は、タクシーの走行距離と乗車下車のGPS位置に注目

5. G空間情報センター



https://www.geospatial.jp/gp_front/

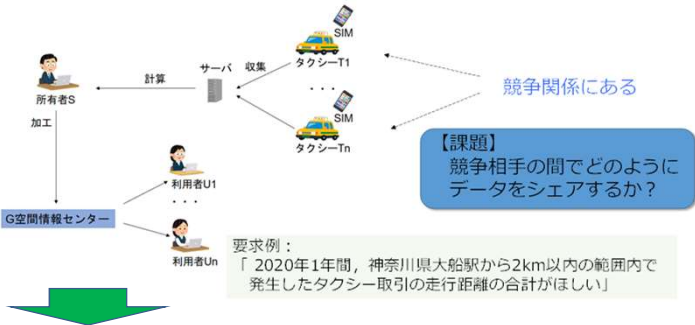
登録されているデータセットは無償のものもあれば、有償のものもある。

有償サービスは今後発展の鍵?

■G空間情報センターは、平成24年3月に政府で閣議決定された地理空間情報活用推進基本計画に基づき、設立されたデータ交換プラットフォーム。一般社団法人社会基盤情報流通推進協議会が運用を行っている。

3. 実現したいシステムのイメージ

タクシーメータからの計量データとGPS位置情報をクラウド上のサーバに集め、データを計算・分析・加工してから販売するという有料サービスを検討してみた。



6. タクシーのGPSデータは個人情報?

- 個人情報の定義(個人情報保護法より)
- 生存する個人に関する情報
- 当該情報に含まれる氏名、生年月日その他の記述等により特定の個人を識別することができるもの
- 二個人識別符号が含まれるもの

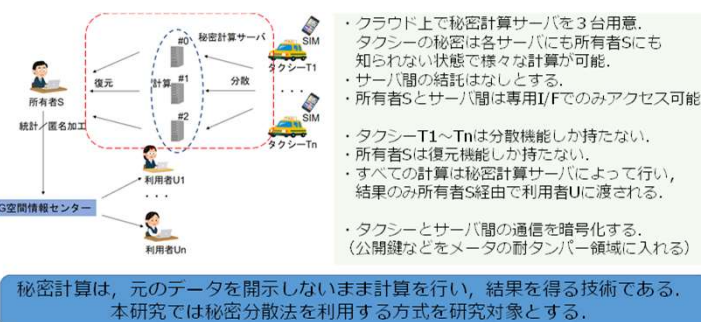
基本的には個人を特定できる情報はすべて個人情報である。

タクシーデータ例	タクシー意味	秘匿
T001	タクシー識別子	×
No.100	通信セッション・カウンター	×
2020/07/20	取引が発生した日付	○
10:00	取引が発生した時刻	○
0352758.020	乗車位置のGPSデータ (北緯: 度分秒.小数点以下)	○
1393721.010	乗車位置のGPSデータ (東経: 度分秒.小数点以下)	○
0352115.070	下車位置のGPSデータ (北緯: 度分秒.小数点以下)	○
1393153.090	下車位置のGPSデータ (東経: 度分秒.小数点以下)	○
12345	走行距離 (m)	○

乗客の個人情報
乗客が身分を明かさず、現金で支払いをしても、システム側の情報でその他の情報と突合せすれば、乗客を特定できてしまう可能性がある。

個人情報保護法の要求により、第三者へ提供する前に、匿名加工を行う必要がある。

4. タクシーデータ提供サービス



7. 仮名加工・匿名加工・統計情報

	個人情報	仮名加工情報	匿名加工情報	統計情報
適性な加工	—	対照表と照合すれば本人が分かる程度まで加工	本人が一切分からない程度まで加工	*
第三者提供時の同意取得	必要	原則第三者提供禁止	同意不要	同意不要

*統計のサンプルが極端に少ない場合、本人を特定してしまう恐れがあるので、要注意。

仮名加工情報も、匿名加工情報も、統計情報も個人情報ではない。

タクシーデータ提供サービスでの対応:	タクシーデータ項目	匿名加工方法
■ 過疎地域などで、タクシーデータのサンプル数が極端に少ない場合、個人情報保護法を違反する恐れがある旨をユーザUIに通知し、データの提供を断る。	タクシー識別子	削除
■ 匿名加工方法	通信セッション・カウンター	削除
	取引が発生した日付	週の最初の日に置き換え
	取引が発生した時刻	削除
	GPSデータ	一定量のノイズを加える
	走行距離 (m)	加工しない

8. まとめ

本研究を通じて、秘密計算技術、特に数種類の計算技術を用いて、統計情報の提供が可能であることを確認した。更に、個人情報保護に関する制度面からの要求についても調査した。これらの知見を今後のシステム作りに生かしたいと考える。