位置情報における匿名性・多様性保証とその活用

宮川　伸也† 森　拓也† 佐治　信之† 小林　功‡ 栗山　桂一‡

日本電気株式会社　サービスプラットフォーム研究所†
株式会社エヌ・ティ・ティ・ドコモ 法人事業部‡

1. はじめに

近年，GPS や RFID などのセンシング技術の発達により，位置などのプライバシ情報を活用しやすくなった．プライバシ情報をさまざまな SP(サービスプロバイダ)に提供することによって，多くの便利なサービスを受けられるが，半面、不正利用や情報漏えいの危険性が高くなる．特に，インターネットに流出し，多くの人の目にさらされた場合，その被害は計り知れない．一方，SP はプライバシ情報から個人を特定・推定できる内容を除いても，特徴さえ分かればサービスを提供できる場合が少なくない．そのため，プライバシ情報を「匿名化」して SP に提供することが重要となる．本稿では，特に位置情報について，これまでに行われてきた匿名化の問題点について述べ，多様化による解決方法を提案する．実証実験の結果についても触れる．

2. 位置匿名化の課題

代表的な匿名化の一つである k-匿名化は，同じ準識別情報を持つデータが k 個以上となるようにデータを加工し，データからユーザを特定できないようにする手法である．位置の k-匿名化は，あるユーザの緯度・経度などの詳細な位置を，k 人以上の他ユーザが含まれるエリアなどの抽象化された位置に加工することで実現される[1]．この方法は，現在地や目的地などの一箇所の位置からユーザを特定できないようにする有効な方法であり，終電案内や現在地周辺検索などの多くのサービスに適用できる．

しかし，ユーザの行動範囲や行動傾向を考慮したサービスを想定した場合，一箇所だけなく，これまで行ったことのある場所を含めて複数箇所の位置を SP に提供することが望まれる．このとき，それぞれの位置に対して k-匿名化を適用した場合，次のような問題がある．

(a) 近くにいるユーザが似た行動をする場合

図 1 左のように，ユーザ A の自宅と病院の位置(点)とその間の移動(点線)が，それぞれ k-匿名化(k≧2)によって，円と実線に表すように匿名化されたとする．左の円にユーザ A が居住していることを知る閲覧者は，ユーザ A の自宅の正確な位置を確認できなくても，ユーザ A が病院に通院していることがわかる．

(b) ユーザの複数の位置が知られている場合

図 1 右のように，ユーザ X の自宅，勤務先，病院の位置(点)とその間の移動(点線)がそれぞれ k-匿名化(k≧2)によって円と実線に表すように匿名化されたとする．ユーザ X の自宅と勤務先を知る閲覧者は，自宅と勤務先が合致するユーザ X を特定でき，病院に通院していることがわかる．

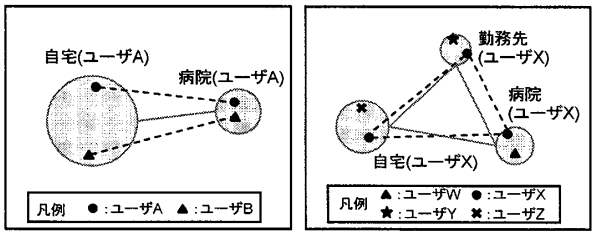


図 1 位置の k-匿名化例

3. 位置多様化の導入

これらを解決するため，ℓ-多様性を満たすように位置を多様化する手法を提案する．ℓ-多様性とは，同じ準識別情報を持つデータセットの非識別情報が少なくとも ℓ 通りの多様性を持つ性質である[2]．ここでは，複数ユーザが含まれるようにユーザの任意の位置を抽象化し，さらに他の位置が ℓ 通りの多様性を有するようにする．ユーザの n 箇所の位置に対して ℓ-多様性を満たす処理を(n, ℓ)-多様化と記す．(1, ℓ)-多様化は k-匿名化(k=ℓ)と同様である．(n, ℓ)-多様化(n>1)は，(n-1, ℓ)-多様化を行った後，任意の n-1 箇所全ての位置周辺に居場所があり，かつ，残り 1 箇所の位置が異なる他ユーザが含まれるように n-1 箇所の位置を抽象化する．

● 例 1: (2, ℓ)-多様化

図 1 左のユーザ A に着目した場合，自宅周辺にいるが病院周辺に行かない他ユーザが含まれるように自宅の位置をぼやかし，病院周辺も同様にぼやかす(図 2)．病院以外の場所に行く他ユーザを含むように自宅周辺(左円)がぼやかされるため，ユーザ A の自宅を知る閲覧

Achieving Anonymity and Diversity for Location-based Services

†Shinya Miyakawa, Takuya Mori and Nobuyuki Saji (Service Platforms Research Laboratories, NEC Corporation)
‡Isao Kobayashi and Keiichi Kuriyama (Corporate Marketing Division, NTT DOCOMO, INC.)

者であっても，ユーザ A が通院することを確信できない．

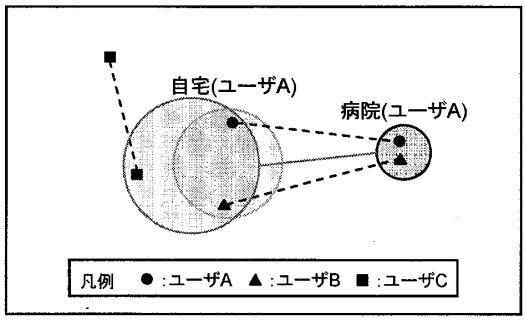


図 2 (2, ℓ)-多様化例

● 例 2: (3, ℓ)-多様化

図 1 右のユーザ X に着目した場合，まず，任意の 2 箇所について(2, ℓ)-多様化を行う．次に，自宅と勤務先の両方の周辺に居場所があり，かつ，病院には行かないユーザ(ユーザ V)が含まれるように，自宅と勤務先の両方をぼやかす(図 3)．同様の処理を勤務先-病院，病院-自宅に対しても行う．ユーザ X の自宅と勤務先を知る閲覧者であっても，ユーザ X が病院に行くユーザなのか他の場所に行くユーザなのか特定できない．

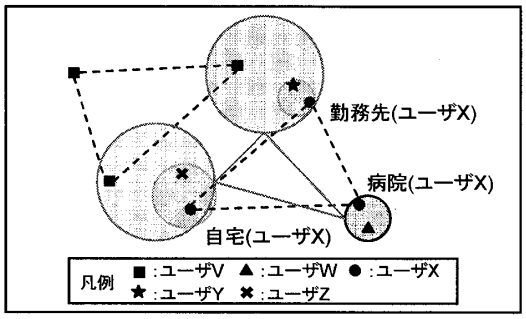


図 3 (3, ℓ)-多様化例

4. モデルサービスへの適用

ユーザの行動範囲周辺のグルメ情報を推薦するモデルサービスに，提案手法を適用した(図 4)．ユーザの位置を携帯電話によって定期的に測位し，プラットフォームの行動情報リポジトリに蓄積する．行動情報管理は，蓄積された緯度・経度形式の位置を分析して、自宅や勤務地などの複数のセンシティブな位置からなる行動範囲をユーザ毎に生成する(行動範囲の位置の数はユーザ毎に異なる．)．そして， 匿名化エンジンは，行動範囲の各位置を地域メッシュコードに変換して多様化し，SP に提供する．ℓ-多様性を満たすまで 250m 四方，500m 四方，1km 四方，2km 四方の順に広げ，それでも満たさない場合は，その位置を SP に提供しない(切り落とす)．SP は多様化された各位置周辺のグルメ店舗情報をユーザに配信する[3]．

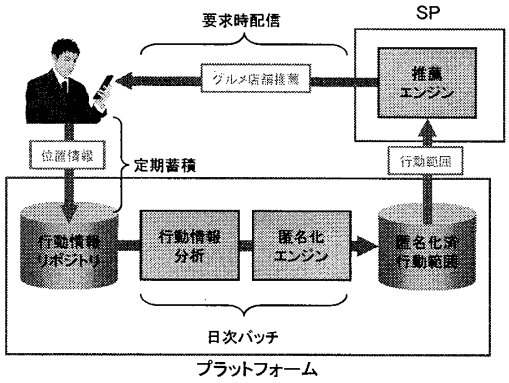
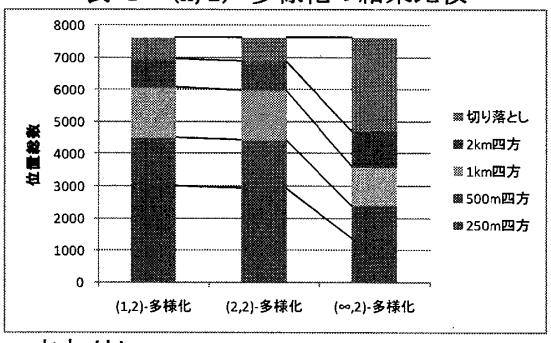


図 4 実証実験システム構成

表 1 は，首都圏のユーザ約 3,000 人の行動範囲の多様化を行った結果であり，生成されたエリアの大きさ毎の数，および，切り落とされた位置の数を表す．(∞, 2)-多様化は，行動範囲の位置の数が m のとき(m, 2)-多様化を行うことを表す．(n, ℓ)-多様化の n を増やすと，エリアが大きくなり，切り落とし数も増加する傾向にあるが，ユーザ単位で集計すると，すべての位置が切り落とされたユーザはほとんどいない．

表 1 (n, 2)-多様化の結果比較



5. おわりに

本稿では，位置に ℓ-多様性の考え方を導入した(n, ℓ)-多様化を提案し，実証実験の結果について述べた．今回対象とした全ユーザの多様化処理時間は数分程度であったが，ユーザ数の増加に伴い性能改善が課題となる．また，元データから歪曲度を考慮するなどの拡張が考えられる．

本研究は，経済産業省「平成 21 年度情報大航海プロジェクト(モデルサービスの開発と実証)」における「マイ・ライフ・アシストサービス」実証実験の一環として実施された．

参考文献

[1] M. Gruteser etc, *Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking*, In MobiSys, 2003.

[2] A. Machanavajjhala etc, *l-Diversity: Privacy Beyond k-Anonymity*, In ICDE, 2006.

[3] 桐越孝之他，"「マルチモードレコメンド基盤」のコンテキストアウェア拡張方式"，第 72 回情報処理学会全国大会，1C-5 (発表予定)