3L4-OS-06c-2

位置情報サービスにおけるプライバシ保護技術

A Survey of Privacy Protection Methodologies for Location-Based Services

川本 淳平*1
Junpei Kawamoto

*1筑波大学システム情報系
Faculty of Engineering, Information and Systems, University of Tsukuba

In these days, many services using location data of people are being launched. At the same time, privacy issues of such location data are argued studiously. We introduce, in this talk, some recent privacy-protection methodologies for such location based services.

1. はじめに

近年，モバイル端末等の普及により，位置情報を用いたサービスが数多く提供され始めている．こうしたサービスは，我々の生活を大きく変えつつあると言える．実際，位置情報を用いた地図サービスの普及により，道に迷うという行為は過去の事柄になりつつある．また，位置情報を利用した広告配信や，多人数の位置情報を合わせることで人の流れを解析し公共サービスの適用に用いるなど，位置情報は商用利用だけでなく公共利用にとっても価値ある情報である．

一方，人々の多くは，位置情報が重要なプライベート情報であると考えている．そのため，例えば誰が何時，何処にいたのかといった情報が推測されないようプライバシを保護した状態で利用する必要がある．

本稿では，この位置情報プライバシ保護に関する取り組みを，次の二つのサービス別に紹介する．

- 位置情報データベース (LBS; location-based service)
- 位置情報の公開サービス

位置情報データベースは，利用者が自信の位置情報を問合せとして送信し，その位置の近くにある POI (point of interest) の情報を取得するサービスをいう．例えば，滞在地点の近くにあるレストランを検索するようなサービスである．位置情報データベースにおける，プライバシ保護の目的は，利用者が自分の滞在位置をサービス提供者に知らせることなく目的の POI 情報を取得することにある．次に，位置情報の公開サービスは，位置情報を入手できる通信事業者やカーナビゲーションサービス提供者が，外部の解析機関が位置情報を利用できるように公開するサービスをいう．ここでのプライバシ保護の目的は，公開された位置情報から個々人の行動が特定されないことである．

2. LBS におけるプライバシ保護

前節で述べた様に，LBS では，利用者は自信の滞在位置をサービス提供者にすら知らせることなく目的の POI 情報を取得することを目的とする．この分野では，大きく分けて次の三つの手法が提案されている．

1. 匿名化サーバの利用
2. 問合せの抽象化
3. プライベート問合せの利用

連絡先: 川本淳平，筑波大学システム情報系，茨城県つくば市天王台 1-1-1，029-853-3826，junpei@mdl.cs.tsukuba.ac.jp

匿名化サーバの利用は，利用者が直接 LBS に問合せを行うのではなく，先ず自信の位置情報を信頼できる匿名化サーバに送信する．その後，この匿名化サーバが代理で LBS に問合せを行い問合せ結果を利用者へ転送する．問合せの抽象化は，利用者が自信の位置（点）をそのまま問い合わせるのではなく，滞在位置を含む領域を問い合わせる方法である．サービス提供者は，利用者がある範囲に滞在していることを知ることはできるが，厳密な位置を知ることはできなくなる．最後のプライベート問合せとは，問合せの内容を秘匿したまま目的の情報を取得する手法である．これを用いて，問合せである位置情報を秘匿したまま関連する POI を取得する方法が提案されている．

2.1 匿名化サーバの利用

匿名化サーバを用いた方法 [Gruteser 03, Gedik 05, Mokbel 06, Kalnis 07] では，利用者は自信の位置を信頼できる匿名化サーバに送信する．匿名化サーバは，同じタイミングで問合せを行った利用者の中から距離の近い k 人を選び，その k 人の位置を包含する領域 (CR; cloaked region) を計算する．匿名化サーバはこの CR を LBS に問い合わせる．サービス提供者は，CR を基に近傍検索すなわち CR に近いいくつかの POI を検索し匿名化サーバへ返却する．匿名化サーバは，得られた問合せ結果を各利用者の本当の位置を基に選別し利用者へ通知する．

このようにすることで，サービス提供者は，一つの CR に含まれる利用者 k 人を区別することができない．いわば k-匿名性 [Sweeney 02] の一種を保証している．一方，この方式の問題点は信頼できる匿名化サーバがボトルネックとなることである．また，利用者がある程度密に固まっていることを仮定しており，そうでなければ，広範囲の CR が LBS への問合せに用いられることになる．その結果，求めている位置との関係が低い POI を多く取得することになり通信コストが余計に掛かってしまうことになる．

2.2 問合せの抽象化

問合せの抽象化方法 [Ardagna 07, Cheng 06, Xu 10] では，匿名化サーバは利用せず，利用者のクライアントが単独で自信の位置を CR に変換して問い合わせる．そして，クライアントは LBS から受け取った POI 情報の中から実際の位置に関連するものだけを選び利用者に提供する．この手法の場合，匿

名化サーバのように他の利用者との区別を基に安全性を定義できないため，別の指標を用いることになる．

2.3 プライバシ保護情報検索の利用

[Ghinita 08] では，プライバシ保護情報検索の一つである cPIR (computational Private Information Retrieval) [Kushilevitz 97] を用いている．cPIR は，一次元のデータベースにおける i 番目のデータを取得する一致検索方法であり，計算理論の観点からサービス提供者は利用者の問合せを知ることができないことが保証できる Ghinita らは，二次元の位置情報を一次元に対応付ける方法として kd 木 [Bentley 75] を用いた近似手法とボロノイ図によるグリッド分割を用いた厳密な手法を提案している．一方，彼らの手法は一致検索方法である cPIR を基にしているため，問合せ地点の最近傍検索 1 件しか取得できないという制限がある．

[Wong 09] は暗号化前後で類似度が変化しない特別な暗号方式を用いて安全な k-近傍探索を提案している．一方で，この暗号は共通鍵暗号方式である．そのため，本方式を LBS に用いる場合，サービス提供者に情報の暗号化を任せることはできず，利用者が予め POI に関する情報を暗号化し LBS に登録する必要がある．また，利用者が複数人いる場合，この共通鍵を安全に共有する方法が別途必要になる．まとめると，Wong らの手法を LBS に用いる場合，利用者の位置情報を秘匿しつつ k-近傍探索を実現するが，その利用に際して制限がある．

3. 位置情報公開におけるプライバシ保護

位置情報には，他の多くのマイクロデータとは異なり．スパース性がありプライバシ保護を難しくしている．ここでいう，スパース性とは，長い時間の位置情報を考えた場合，同じような移動を行う人は少なくなってしまうということを意味する．つまり，ある一人の情報に対して少なくとも $k-1$ 人が同じ情報を持っていれば，その人の情報は他の $k-1$ 人と区別できないため安全であるという k-匿名性 [Sweeney 02] を位置情報に当てはめることは難しい．

LKC プライバシ [Fung 09] は，k-匿名性の条件を緩和し，位置情報の長さを L までしか考えないことで，このスパース性に取り組んでいる．また，[Terrovitis 08] でも同様に，長さ m までしか考えず，かつ移動の順序を考えず set-value と見なす，条件を緩和した k-匿名性の一種である k^m-匿名性を提案している．

[Chen 12] では，ある長さ（論文中では長さ 5 程度）以下の移動シーケンスの発生頻度を集計し公開する場合において差分プライバシ [Dwork 06] を満足する方法を提案している．彼らの手法の特徴は，公開された位置情報の利用方法として，頻出パタンマイニングに用いることを想定し，そのため稀なパタンは予め削除している点である．言い換えれば，移動シーケンスの長さを制限し，また稀なシーケンスを削除することでスパース性を回避している．

そもそも，位置情報を人々の行動シーケンスとして扱うのではなく，ある地点に何人滞在していたのかというヒストグラム形式で公開する方法も提案されている．[Qardaji 13] は，2次元平面をグリッドに分割し，各グリッドに何人滞在しているのかを集計し差分プライバシを満たす形で公開する．彼らの手法は，一度きりの公開を仮定しているが，我々[Kawamoto 13] は人々の行動にマルコフ過程を導入し攻撃者の能力にも仮定を置くことでヒストグラムの継続的公開を実現するアドバーザリアルプライバシ [Rastogi 09] を提案している．

4. おわりに

本稿では，位置情報を利用したサービスにおけるプライバシ保護技術として，位置情報データベース (LBS) と位置情報の公開サービスにおける近年の研究動向を紹介した．

位置情報におけるプライバシ保護の必要性は議論されているものの前述のように，位置情報のプライバシ保護は様々な方法論が提案されており未だデファクトスタンダードな手法は定まっていない．また，新しい位置情報の利用方法が考案されれば，新たな保護技術も必要になる．今後も注目の分野であるといえる．

謝辞

著者は，最先端研究開発プログラム (FIRST)「超巨大データベース時代に向けた最高速データベースエンジンの開発と当該エンジンを核とする戦略的社会サービスの実証・評価」の助成を受けている．ここに記して謝意を表します．

参考文献

[Ardagna 07] Ardagna, C. A., Cremonini, M., Damiani, E., Vimercati, di S. D. C., and Samarati, P.: Location privacy protection through obfuscation-based techniques, in *Proc. of the 21st Annual IFIP WG 11.3 Working Conference on Data and Applications Security*, pp. 47–60, Redondo Beach, CA, USA (2007), Springer-Verlag

[Bentley 75] Bentley, J. L.: Multidimensional Binary Search Trees used for Associative Searching, *Communications of the ACM*, Vol. 18, No. 9, pp. 509–517 (1975)

[Chen 12] Chen, R., Acs, G., and Castelluccia, C.: Differentially Private Sequential Data Publication via Variable-Length N-Grams, in *Proc. of the 19th ACM Conference on Computer and Communications Security*, pp. 638–649, Raleigh, NC, USA (2012), ACM Press

[Cheng 06] Cheng, R., Zhang, Y., Bertino, E., and Prabhakar, S.: Privacy Enhancing Technologies, in Danezis, G. and Golle, P. eds., *Proc. of the Sixth International Conference on Privacy Enhancing Technologies*, Vol. 4258 of *Lecture Notes in Computer Science*, pp. 393–412, Cambridge, UK (2006), Springer-Verlag

[Dwork 06] Dwork, C., Mcsherry, F., Nissim, K., and Smith, A.: Calibrating Noise to Sensitivity in Private Data Analysis, in *Proc. of the Third Theory of Cryptography Conference*, pp. 265–284, New York, NY, USA (2006), Springer

[Fung 09] Fung, B. C. M., Cao, M., Desai, B. C., and Xu, H.: Privacy Protection for RFID Data Categories and Subject Descriptors, in *Proc. of the 24th ACM Symposium on Applied Computing*, pp. 1528–1535, Honolulu, HI, USA (2009), ACM Press

[Gedik 05] Gedik, B. and Liu, L.: Location Privacy in Mobile Systems: A Personalized Anonymization Model, in *Proc. of the 25th IEEE International Conference on Distributed Computing Systems*, pp. 620–629, Columbus, OH, USA (2005), IEEE Computer Society

[Ghinita 08] Ghinita, G., Kalnis, P., Khoshgozaran, A., Shahabi, C., and Tan, K.-L.: Private Queries in Location Based Services: Anonymizers are Not Necessary, in *Proc. of the 28th ACM SIGMOD International Conference on Management of Data*, pp. 121–132, Vancouver, BC, Canada (2008), ACM Press

[Gruteser 03] Gruteser, M. and Grunwald, D.: Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking, in *Proc. of the First International Conference on Mobile Systems, Applications and Services*, pp. 31–42, San Francisco, CA, USA (2003), ACM Press

[Kalnis 07] Kalnis, P., Ghinita, G., Mouratidis, K., and Papadias, D.: Preventing Location-Based Identity Inference in Anonymous Spatial Queries, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 19, No. 12, pp. 1719–1733 (2007)

[Kawamoto 13] Kawamoto, J., Fukuchi, K., and Sakuma, J.: マルコフモデルを仮定した位置情報開示のためのアドバーザリアルプライバシ, 人工知能学会全国大会, pp. 3L4–OS–06c–2 (2013)

[Kushilevitz 97] Kushilevitz, E. and Ostrovsky, R.: Replication Is Not Needed: Single Database, Computationally-Private Information Retrieval, in *Proc. of the 38th Annual Symposium on Foundations of Computer Science*, pp. 364–373, Washington, DC, USA (1997), IEEE Computer Society

[Mokbel 06] Mokbel, M. F., Chow, C.-Y., and Aref, W. G.: The new Casper: query processing for location services without compromising privacy, in *Proc. of the 32nd International Conference on Very Large Data Bases*, pp. 763–774, Seoul, Korea (2006), VLDB Endowment

[Qardaji 13] Qardaji, W., Yang, W., and Li, N.: Differentially Private Grids for Geospatial Data, in *Proc. of the 29th IEEE International Conference on Data Engineering*, Brisbane, Australia (2013), IEEE Computer Society

[Rastogi 09] Rastogi, V., Hay, M., Miklau, G., and Suciu, D.: Relationship Privacy: Output Perturbation for Queries with Joins, in *Proc. of the 28th ACM Symposium on Principles of Database Systems*, pp. 107–116, Providence, RI, USA (2009), ACM Press

[Sweeney 02] Sweeney, L.: k-anonymity: a model for protecting privacy, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, Vol. 10, No. 5, pp. 1–14 (2002)

[Terrovitis 08] Terrovitis, M., Mamoulis, N., and Kalnis, P.: Privacy-preserving anonymization of set-valued data, *Proc. of the International Conference on Very Large Databases*, Vol. 1, No. 1, pp. 115–125 (2008)

[Wong 09] Wong, W. K., Cheung, D. W.-L., Kao, B., and Mamoulis, N.: Secure kNN Computation on Encrypted Databases Categories and Subject Descriptors, in *Proc. of the 35th SIGMOD International Conference on Management of Data*, pp. 139–152, Providence, RI, USA (2009), ACM Press

[Xu 10] Xu, J., Tang, X., Hu, H., and Du, J.: Privacy-Conscious Location-Based Queries in Mobile Environments, *IEEE Transactions on Parallel and Distributed Systems*, Vol. 21, No. 3, pp. 313–326 (2010)