

Geohash を用いたモバイルセンシングのための データベースシステムの提案

小川 輝樹[†] 伊藤 嘉博[‡] 荒井 健次[‡] 中村 嘉隆[†] 高橋 修[†] 白石 陽[†]

公立はこだて未来大学 システム情報科学部[†]

公立はこだて未来大学大学院システム情報科学研究科[‡]

1 はじめに

近年、モバイルセンシングの研究やサービスが注目されてきている。モバイルセンシングとは、自動車や自転車、人間などの移動体にスマートフォンなどを用いてセンサデバイスを搭載し、センサ自身が移動しながらセンシングを行う技術である。特定の場所に固定して設置した上で環境情報を取得する固定センサとは異なりセンサ自身が移動できるため、センサデバイスを設置できない場所でも細粒度の環境情報を得ることができる。こうした移動センサから、モバイルセンシングデータを収集し、データベースに格納することで、従来の固定センサのセンシングデータと補完的に利用することで様々な用途での利用が期待できる。その活用例として、移動センサの一日の行動履歴や動作パターンをもとにした統計データを利用したサービスが考えられる。

しかし、これらの想定されるモバイルセンシングのサービスにおいては、蓄積されたセンシングデータが膨大な量になることが原因で問い合わせの処理時間が増大するという問題がある。

そこで、Geohash は問い合わせ処理速度向上のために、階層的な空間データ構造である Geohash を用いたモバイルセンシングのためのデータベースシステムを提案する。

2 関連研究

モバイルセンシングの関連研究として、文献[1]ではセンシングデータを利用したライフログ生成及び状況推定を行っている。この研究は、スマートフォン内蔵のセンサを用いてユーザの行動履歴を収集し、得られた情報からライフログデータの生成を行い、端末にデータを記録する。

しかし、この研究では、データの生成処理がすべてアプリケーション側で行われている。そのため、アプリケーション開発のコストやアプリケーションの処理負荷の増大をもたらす。更に、スマートフォンに生成されたデータを蓄積するため記憶容量が制限される。

センサデータベースの関連研究として、モバイルセンサデータベースのためのデータ記録保証機構[2]がある。この研究は、モバイルノード内のデータベースに近傍のセンシングデータを記録していく。その際に、複数のフィルタによりセンシングデータを收拾選択して記録することでセンサデータ品質の保証をしている。

しかし、この研究では、モバイルノードは記憶容量が

少ないため記録できるセンシングデータの量が限られてしまう。

Geohash は、位置情報をハッシュ値で表す階層的な空間データ構造である。緯度経度データをハッシュ値である文字列に変換し、地図などの空間情報データをメッシュ状に分割する技術であり、様々な研究や応用システムで使われている[3,4]。メッシュに対する文字列の割り当てを図 1 に示す。Geohash では、図 1(a)に示すように、矩形領域を 32 (4x8) 個のメッシュに分割し、分割された各メッシュに異なる英数字を割り当てる。図 1(b)は図 1(a)のあるメッシュをさらに分割している様子を表している。文字列の末尾に新たな文字を追加することでより細かいメッシュが表現可能となる。

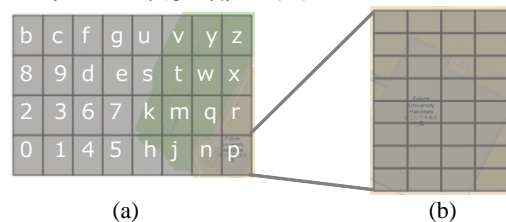


図 1. メッシュ表示 ((a)縮小表現 (b)拡大表現)

3 提案システム

提案システムでは、問い合わせ処理速度の低下を抑制するために Geohash を用いる。緯度経度データを Geohash に変換した文字列にすることにより、指定範囲のデータを容易に絞り込むことができる。それにより、データベースへの問い合わせ処理が従来の緯度経度を元とした検索システムより向上する。更に、分割したメッシュの文字列で検索したデータをモバイルノード別に振り分け一つにまとめる。こうすることで、絞り込んだメッシュ内の人口密度の統計データを取得できる。

(1)システム構成

本システムは、モバイルノードとデータベースシステムから構成される。モバイルノードは位置情報、センシングデータなどをデータベースに提供し、データベースシステムは、位置情報・時刻・モバイル ID・センシングデータを管理処理する。提案システムの構成を図 3 に示す。

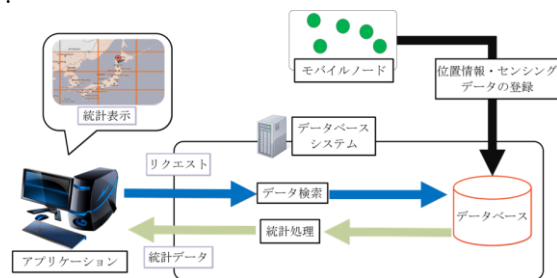


図 3. システム全体の構成図

Proposal of a database system using geohash for mobile sensing

Teruki Ogawa[†], Yoshiro Ito[‡], Kenji Arai[‡], Yoshitaka Nakamura[†], Osamu Takahashi[†], Yoh Shiraishi[†]

[†]School of Systems Information Science, Future University Hakodate

[‡]Graduate School of Systems Information Science, Future University Hakodate

(2)データベース構造

提案システムのデータベースは、経度緯度データとその緯度経度データを Geohash に変換した文字列、センシングデータ、モバイル ID、時刻の 5 つから構成される。

センシングデータは、モバイルノードから取得可能な加速度、角速度等といったセンサデータである。複数の種類のセンシングデータを登録可能だが、指定したメッシュ内の人口密度の統計データ取得するため現段階では加速度と角速度の 2 つを登録している。

時刻はユーザがセンシングデータを取得した日付と時間のデータである。表 1 に詳細なテーブル構造を示す。

表 1. データベースのテーブル構造

geohash	緯度経度	時刻	モバイルID	加速度	角速度	...
wx4eqyuz8dn4e	POINT(116.318417 39.984702)	2008-10-23 02:53:04	aihsuht	0.9417674	9.417674	
wx4eqyux8m4svb	POINT(116.31845 39.984683)	2008-10-23 02:53:10	asdgosk	0.9417674	10.04551	
wx4eqyurxdnds	POINT(116.318417 39.984686)	2008-10-23 02:53:15	aaufrh	0.31392246	9.33919	

(3)データ蓄積

データ蓄積では、モバイルノードから位置情報・センシングデータなどを受け取り(2)で述べた 5 つの属性を持つデータをテーブルに格納する。

(4)データ検索

データ検索では、アプリケーションで指定したメッシュの文字列で蓄積しているデータベースへ前方一致検索で処理を行う。緯度経度データとは異なり、変換した文字列で検索するので前方一致検索が利用できるため、検索処理が高速化する。

(5)統計処理

統計処理では、(4)で検索したモバイルノードがメッシュ内のモバイル ID リストに登録されているかチェックする。モバイル ID リストとは、メッシュ内に存在するモバイル ID を管理するリストである。リストにモバイル ID が存在しなければ登録し、逆にすでに存在した場合は登録しない。この処理を行うことで、モバイル ID の重複を避け、メッシュ内人口密度の統計データを正確に取得できる。

4 実装

提案システムの動作を検証するため、提案する機能及び処理を実装したプロトタイプシステムと評価アプリケーションの構築を行った。

4.1 プロトタイプシステムの構築

データベース管理システム(DBMS)は PostgreSQL を用い、地理情報の取り扱いには拡張機能の PostGIS を利用した。問い合わせや統計処理は Java Servlet で記述した。

使用するデータは GeoLife プロジェクト[5]で収集された GeoLife GPS Trajectories のサンプルデータを用いる。GeoLife GPS Trajectories は 2007 年 4 月から 2011 年 10 月の 178 人、計 17,355 個の軌跡データを収集したデータセットである。

4.2 評価アプリケーションの構築

評価アプリケーションは Java Server Pages を用い、Google Map 上にメッシュごとの人口密度の統計データを表示するシステムを構築した。統計処理を行ったデータ

をアプリケーションに送りメッシュ内のモバイルノード数に応じて色分けをして統計データを地図上に表示する。“wx4ewg” という文字列で前方一致検索を行った時の実行結果を図 4 に示す。図 4 の赤枠が文字列で指定した範囲を表しており、枠内の色が統計処理結果(ここでは、人口密度)を表している。表示されるピン一つ一つがモバイルノードとなる。

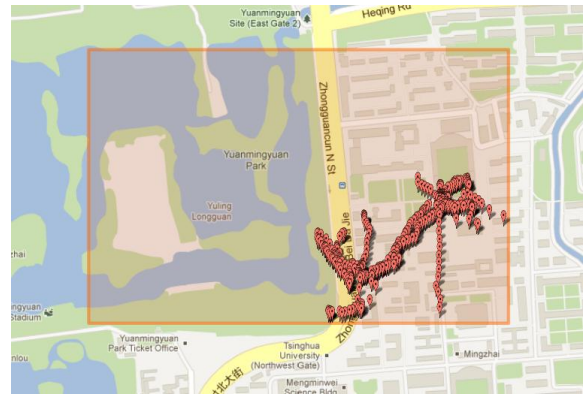


図 4. 統計データの表示画面

5 まとめ

本研究では Geohash を用いたモバイルセンシングのためのデータベースを提案し、基本機能の一つとして、統計処理結果の表示機能の実装を行った。

今後は、複数の種類のセンサデータを考慮したデータの対応付けや、それに伴う統計データの表示方法の改良を進めていく。システムの評価については、提案システムと緯度経度データを元にした検索システムそれぞれに対して処理時間の比較を行う。

更に、検索に応じた統計データが取得できているかについての検証を行う。

図 4 に示したような複数のピンから一つ選択し、そのピンの示すモバイルノードが走っていた、あるいは歩いていた等の状態もセンシングデータから推定できるようにする。

参考文献

- [1] 松浦寛ら：“高機能携帯電話を用いたライフログ収集手法の提案及び状況推測に関する研究”，情報処理学会全国大会講演論文集 2011(1), pp.291-293, 2011.
- [2] 由良淳一ら：“モバイルセンサデータベースのためのデータ記録保証機構”，DBSJ letters 6(4), pp.5-8, 2008.
- [3] 高橋郁久ら：“Wikipedia リンク構造を用いた歴史エンティティの重要度計算”，DBSJ Journal 10(1), pp.25-30, 2011.
- [4] B.Martins, N.Freire, J.Borbinha：“Complex data transformations in digital libraries with spatio-temporal information,” In: Proc. of the 11th International Conference on Asia-Pacific Digital Libraries (ICADL 2008), pp.174-183, 2008.
- [5] Y.Zheng, L.Zhang, X.Xie, W.Y.Ma：“Mining interesting locations and travel sequences from GPS trajectories,” In Proceedings of International conference on World Wild Web (WWW 2009), pp.791-800, 2009..