

携帯電話の測位機能を用いた有意位置の学習

遠山 緑生[†] 服部 隆志^{††} 萩野 達也^{††}

位置情報はコンテキストウェアアプリケーションの構築にあたって重要である。本論文では、ユーザの日常活動を記述するのに適した、位置コンテキストを記述するための概念である有意位置の生成手法について述べる。測位機能を持つ携帯電話より取得する測位履歴を対象とし、多段階の半径を用いたクラスタリングにより有意位置を学習する手法を提案する。対象とする測位履歴は、比較的低頻度であり、精度のばらつきが大きいという特徴を持つ。多段階の半径を用いることで、対象機器の特性に適応しながら、利用者の空間的行動のパターンをより適切に記述することが可能となる。この手法の特性と、適切なパラメータの選択について議論する。

Using a Positioning System of Cellular Phone to Learn Significant Locations

NORIO TOYAMA,[†] TAKASHI HATTORI^{††} and TATSUYA HAGINO^{††}

Location information is important to construct context-aware applications. A significant location is a specialized form of location context for expressing a user's daily activity. We propose a method to learn significant locations from positions measured by cellular phones. Cellular phones we used are equipped with a positioning system, where data can be taken in low frequency with wide-varying estimated errors. In order to learn significant locations, our system uses a clustering method which exploits multiple radii for coping with these characteristics and for adapting to a variety of users' spatial behavioral patterns. We also discuss characteristics and appropriate parameters for our method.

1. はじめに

最近の携帯情報機器の高度化とユビキタス情報環境の発展は著しい。本研究の最終的なねらいは、ユビキタス情報環境において携帯情報機器を環境とのインタラクションを行うためのインタフェース装置として活用できるような、コンテキストウェアアプリケーションの作成と利用を可能とすることである。たとえば、「駅の近くで時刻表を調べる」、「部屋に入ったら電気をつける」といったルールの組合せから成るコンテキストウェアアプリケーションを作成する場合を考える。これらのルールは、条件としてのコンテキストと、結果としての操作、両者の組として成り立つ。条件となるコンテキストとしては、現在位置、現在時刻、曜日、近い将来の位置、スケジュールなどが考えられる。結果となる操作としては、携帯電話であればメー

ルの送信、Web ページの操作などが考えられる。

GPS などの測位システムを利用することで、ユーザの位置情報を常時測位し、コンテキストとして利用することが可能になるが、これの取扱いは大きい問題である。測位の結果は緯度と経度からなる空間上の 1 点であるが、ルールにおいて必要となるのはより抽象的な「家にいる」「駅の近くにいる」といった一定の自由度を持つ領域を表現したものである。

本論文では、このようなコンテキストウェアアプリケーションでの利用を想定した、有意位置 (Significant Location)^{1),2)} の概念を用いた位置分析手法について述べる。有意位置は、利用者が滞在する一定範囲の領域であり、中心点の空間座標と半径の組合せによって表現される。学習システムは、測位履歴から利用者が一定時間滞在した領域を判別し、一定半径の円によって有意位置として集約する。有意位置により、単純に緯度・経度による 1 点を指すよりも、より抽象度の高い形で位置を表現できる。

プロトタイプとして、GPS 内蔵の携帯電話の測位機能によって定期的な測位履歴を取得するシステムを実装した。その履歴から有意位置を学習する手法につ

[†] 慶應義塾大学デジタルメディア・コンテンツ統合研究機構
Research Institute for Digital Media and Content, Keio University

^{††} 慶應義塾大学環境情報学部
Faculty of Environmental Information, Keio University

いて検討する．特に2つの問題について議論する．まず第1に，ターゲットとしてGPS内蔵の携帯電話を利用したときに，得られる測位履歴の特性に適應できる手法の提案である．特性は2つあり，どこでも測位できるが精度がばらつくことと，比較的low頻度な位置履歴しか得られないことである．第2に，測位履歴から有意位置を学習する際に，多段階の半径を用いる手法についての議論である．多段階の半径を用いることで，ユーザの空間的行動のスケールに対して多段階に適應できる．たとえば，徒歩での移動と交通機関での移動では，同じ位置と見なすべき範囲の大きさは異なってくるが，これらを段階的に表現することが可能になる．

以下本論文では，まずシステムの概要と特徴を示す．次に3章でこのシステムにおける有意位置の学習手法についてその解決すべき問題点と手法の概要について述べる．4章では被験者の利用結果に基づき，手法の特性と適切なパラメータの選択について議論する．

2. システムの概要

プロトタイプとして，履歴収集クライアントと管理サーバから成るシステムを実装した．一般利用者が特別な訓練なく使えるものを目指し，携帯電話をターゲット機器とした．この機器の測位機能を用い，定期的な測位履歴を記録する．本章では，システムの構成と利用例，その測位履歴の特徴について述べる．

2.1 システムの構成

本システムは履歴収集クライアントと履歴管理サーバから構成され，それぞれに以下に述べる機能を持つ．

2.1.1 履歴収集クライアント

クライアントの実装環境は au (KDDI) の携帯電話，W11K である．この機種は，J2ME⁸⁾ によるアプリケーションの実行機能と，Qualcomm gpsOne テクノロジ⁷⁾ を用いた測位機能を持つ．このターゲット上に，J2ME MIDlet として以下の機能を持つ履歴収集アプリケーションを実装した．

定期的な測位記録 指定した時間間隔で測位を行い，位置を記録する．

操作履歴の記録 電話の発着信，電子メール送受信，ブックマークからの Web ページ操作を記録する．

履歴のサーバへの送信 記録機能により蓄積した履歴をサーバへ送信する．

測位の時間間隔は設定可能だが，2.3.2 項で述べる理由により，10 分ごとに測位を行う．このほか，操作が行われた直後にも測位を行う．履歴個数が 10 個を超えるとサーバへ HTTP 接続し，履歴を送信する．

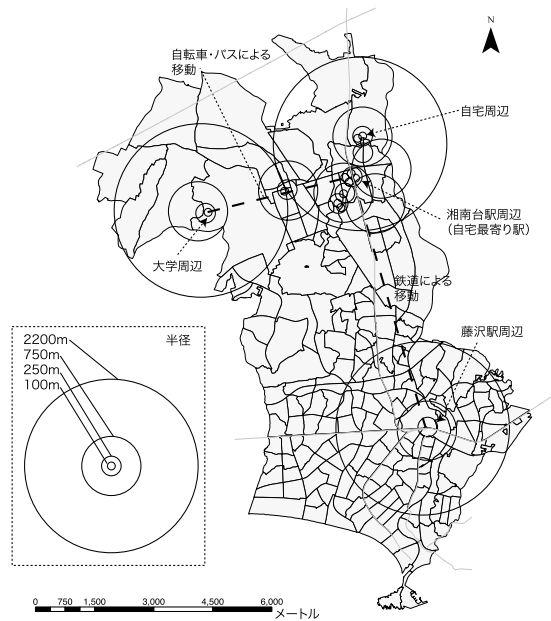


図1 テストユーザの藤沢市内における有意位置
Fig. 1 Significant locations of a user within Fujisawa City.

履歴送信の完了後，クライアント内の履歴を消去する．

2.1.2 履歴管理サーバ

サーバはクライアントから送信された履歴の管理機能と，分析・学習機能を持つ．実装は Linux 上の Perl スクリプトと PostgreSQL データベースシステムによって行った．クライアントからの履歴取得は CGI により行い，データベースに保存する．履歴の分析と学習の機能は，コマンドラインから実行するスクリプトによって行う．3章では，このシステムを用いた履歴からの有意位置の学習手法について議論する．

2.2 利用例

本システムの利用例として，図1に，後述するケース1の被験者の履歴から学習した神奈川県藤沢市内における有意位置を地図上に描画したものを示す．

個々の円が学習された有意位置を示す．このユーザの藤沢市内における行動は大きく3つのエリア，左上の大学近辺，右上の自宅・最寄り駅周辺，右下の藤沢駅周辺から成ることが示されている．この例のように，本システムにより，ユーザの測位履歴から主要な空間的領域を有意位置によって要約することが可能になる．

2.3 本システムの測位履歴の特徴

一般的なGPS測位装置と比べ，携帯電話を用いた本システムは，2つの異なる特性を持つ．まず第1に，どこでも測位できる代わりに精度のばらつきが大きい．第2に，測位頻度が比較的low頻度にならざるをえない．以下にそれぞれの特性について述べる．

表 1 測位頻度とバッテリー持続時間の関係

Table 1 Frequency of location measurements and time to battery running out.

間隔	バッテリー持続時間
10 分	21 時間 20 分
5 分	14 時間 30 分
2 分	9 時間 40 分

2.3.1 測位可能性の高さと精度のばらつき

本システムでは、gpsOne 測位機能により、GPS 信号が受信できない室内であっても、携帯電話基地局の補助を受け、測位することができる。

一方この副作用として、測位精度のばらつきという問題がある。一般的な GPS 専用機器では、機器の固有特性で決まる一定範囲の精度で測位できるのに対し、本システムでは測位モードによって精度が大きくばらつく。具体的には、十分な GPS 衛星が確保できる場合には数 m の誤差で測位できるのに対し、十分な GPS 信号が得られず基地局ベースの測位モードとなる場合など、誤差はときに 2km を超える。

ターゲットシステムの API より測位結果を取得する際、同時に誤差の推定値が楕円形の標準偏差として提供される。これらは、*majoraxiserr*、*minoraxiserr* と *axisangle* の 3 つの値から成る。今回の手法では、単純化のため長辺方向に合わせた正円と見なし、*majoraxiserr* を推定誤差の値として用いた。

2.3.2 低頻度の測位

継続的な測位履歴の蓄積と、実用的なバッテリーの持続時間を両立させるため、本システムでは数分から 10 分程度の低頻度間隔でしか測位を行えない。GPS 専用機器が数秒間隔での測位履歴を 1 日程度にわたって取ることができるのに比べると、非常に低頻度である。

表 1 に、我々のシステムで測位を行った場合の、測位頻度がバッテリー持続時間に与える影響を示す。バッテリー充電なしに履歴を継続的に取得し続けるには、表のような低頻度で記録する必要がある。

3. 有意位置の学習手法

1 章で述べたように、本研究の目的はユーザの携帯電話より得られる測位履歴を集約し、有意位置の学習を行うシステムの実現である。本章ではこの実現手法について述べる。具体的な手法としては、多段階半径を用いた円形領域によるクラスタリングを行う。

本章ではまず、本研究における有意位置の概念を定義し、次に手法の特徴である多段階半径の利用について述べる。次に、元になった Ashbrook らの手法について概要を述べ、これをふまえて本研究における改良

点を述べる。最後に関連研究について触れる。

3.1 有意位置の定義

本研究では、有意位置とは、利用者が滞在する一定範囲の領域であると定義する。

一般には「位置」とは、空間上の 1 点もしくは一定領域であると考えられる。GPS によって得られる測位履歴は、緯度・経度から成る空間上の「点」の羅列である。一方、コンテキストウェアアプリケーションで利用したい位置とは、一定範囲の「領域」である。このような意味での位置を得るには、履歴から得られる点の集合を、一定範囲の領域へと変換する必要がある。

提案手法では「有意位置」は、一定の半径による大きさの円形領域とする。このように、位置を一定の大きさを持つ領域として定義することには、2 つの意味がある。第 1 に、測位誤差への対応であり、もう 1 つは空間上の意味的な指示対象となる実体の持つ範囲の表現である。以下に両者について詳細を述べる。

まず、測位結果は、つねに一定の誤差をとまなう。仮にまったく同一の点から動かずに連続して測位を行ったとしても、測位履歴は一定範囲の誤差を含む近隣点群となる。したがって、誤差を考慮に入れた場合、仮に同じ「点」としての位置を表現する場合であっても、一定範囲の幅を誤差として持つ領域として表現すべきである。

次に、空間上の意味的な指示対象、たとえば「家」や「自宅」などは、一定範囲の大きさを持つ領域である。さらに、「駅の周辺」などを表現したい場合、「周辺」としてまとめた範囲をユーザの行動スケールに応じて異なる大きさを持つ領域として表現する必要がある。

提案手法は、ユーザが明示的な定義を行うことなしに、ユーザの行動履歴である測位履歴のみから、利用者が過去に滞在したことがある一定範囲の領域を有意位置として学習するものである。測位履歴のみを用いて学習を行うため、誤差によって生じる領域の幅と、指示対象が本来持つ領域としての幅をデータのみから判断することは難しい。このため、上に述べたような測位誤差との関係と、指示対象となる領域の大きさという 2 つの側面から、半径の大きさを考慮した円として有意位置を定義する。この際に、本研究では多段階の半径を利用する手法を提案する。本章では、まず、有意位置とする円の多段階半径利用の意義について述べた後、有意位置を学習するためのクラスタリング手法について議論する。

3.2 多段階半径の利用

本研究の手法では、測位履歴の有意位置へのクラス

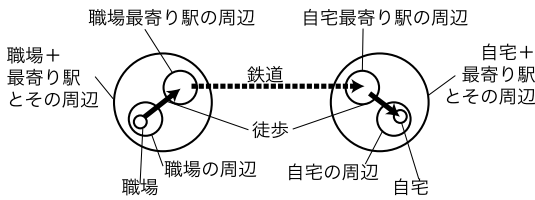


図 2 想定する日常的空间行動のシナリオ
Fig. 2 The typical scenario of daily activity.

タリングによる集約時に、多段階の半径を用いる。サイズについては 4.3 節で議論するが、今回は 100 m, 250 m, 750 m, 2,200 m の 4 段階の半径を用いた。多段階の半径の利用には 2 つの目的があり、これに合致するような半径の選択について考察した。まず第 1 の目的は、空間的行動の多様なスケールを表現できるように、有意位置の半径に段階を設定することである。もう 1 つの目的は、測位精度がばらつく問題に適応させるためである。以下に 2 つの目的について述べる。

3.2.1 空間的行動のスケール

たとえば、「駅の周辺」を表現したい場合に、その位置と見なすべき範囲は、徒歩での移動の場合と、何らかの乗り物による移動の場合とは異なり、前者より後者的の方がより大きい範囲を「周辺」とすべきだと考えられる。本研究では、このように「位置」は、ユーザの空間的行動のスケールに応じて、多段階の大きさを持つと仮定し、このようなスケールの有意位置による表現に多段階半径の組合せを利用する。

図 2 に、ユーザの想定行動シナリオを示す。この例では、大中小の 3 段階の半径を用いた場合を示す。最も大きい単位の位置間において、ユーザは自動車や公共交通機関などに乗車して、数 km 単位の移動を行う。ユーザはどこかの位置で乗り換えなどを行う可能性もある。これらの乗り物を降りた後、ユーザは最終的な目的地に向けて歩いて向かう。たとえば、最寄り駅から自宅に向かうなどの行動が考えられる。この段階の移動と滞在は、徒歩圏内における位置を区別できる中間の大きさによる位置として表現される。最後に、最も小さい単位の位置においてユーザは静止状態となる。たとえば自宅やオフィスなどである。このように、行動に複数の段階的なモードを設けることで、人間の行動スケールの階層的構造に適応させることを目指した。

3.2.2 測位精度のばらつきへの対応

多段階半径利用の 2 番目の目的は、測位精度のばらつきへの対応である。2.3.1 項において述べたように、本システムで得られる測位履歴の誤差は、数 m から 2 km を超すような場合まで、状況によって大きくばらつく。これらの誤差は、どの程度 GPS 信号が受信

できるかなど、位置の特性に応じて変わってくる。たとえば、一般に建物内の位置では GPS の受信感度が低く、屋外に比べると誤差が大きくなる。

4.3.2 項でテストケースの履歴をもとに検討するが、誤差を考慮しないと、誤った有意位置の生成という問題が発生する。たとえば、誤差が 500 m 以上あるような測位結果が累積された状態で、半径 100 m の有意位置を学習させると、本来の有意位置の近くのずれた場所に、不適切な誤った有意位置が生成されてしまう。

多段階の半径の利用により、個々の位置の誤差の傾向に応じて半径を調整することが可能になる。精度の高い測位が可能な場所はより小さい半径の細かい位置として、精度の低い測位しかできない場所はより大きい半径で、それぞれ有意位置を得ることができる。このように多段階の半径を利用するため、本手法では、クラスタリング時に測位誤差に基づく「切り捨て処理」を導入する。この詳細については 3.3.3 項で述べる。

3.3 クラスタリングによる有意位置の学習

本研究の有意位置の学習手法は、Ashbrook ら^{1),2)}による、有意位置を生成するための提案を基本とした。しかし、本システムでユーザの滞在位置を判定する際に元の手法を適用するには、いくつか問題があり、本研究では改良した手法を提案する。本研究における改良点について比較するため、以下ではまず元の手法について概要を述べた後、改良した手法について述べる。

3.3.1 Ashbrook らの手法

Ashbrook らのアルゴリズムは、ユーザがある位置に滞在したことを判定する基準として、次の仮定を用いている。まず、ユーザが滞在するのは室内であると。そして、通常の GPS は、ユーザがビルなど室内に入った場合、GPS 信号を受信できなくなるので、非受信状態が室内での滞在を意味する、という仮定を設定している。以上の仮定に基づき、このアルゴリズムは大きく分けて次の 3 つの段階から成る処理を行う。

- (1) 測位履歴からユーザが滞在した place を判定。
- (2) 距離的に一定範囲内にある複数の place を単一の location にクラスタリング。
- (3) sublocation の分析。

まず最初に、履歴中から GPS 信号を受信できなくなった地点が place として取り出される。各 place は、緯度・経度の座標によって示される点である。

次のクラスタリング段階では、3.1 節で述べたように、ユーザが同一地点へ複数回滞在したとしても、測位結果の座標はある程度ばらつくので、指定した半径内の領域に収まる複数の place を同一の位置としてクラスタリングによって集約するという作業を行う。ク

ラスタリングには k-means クラスタリングアルゴリズムの変種の利用が提案されている。

次に、生成された location を必要に応じてより小さい半径を用いた sublocation に分割する。これにより、キャンパスレベルの領域と其中的建物レベルの領域など、異なるサイズの領域を区別することが可能としている。Ashbrook らはこの段階を利用する提案を行っているが、何段階の sublocation が必要かなどの詳細な議論はされていない。本研究における多段階の半径の利用は、この sublocation の概念と実装を我々の要求に合うように拡張し、より詳細な考察を行ったものであるといえる。

なお上記手法による有意位置の学習後、位置間の移動を予測する手法について議論を行っているが、本論文では予測の問題は扱わないので、ここでは触れない。

3.3.2 滞在位置の判定手法の変更

本システムにおける改良の必要性は、Ashbrook らの研究で利用しているシステムと、我々のシステムの特性の違いからくるものである。我々の提案手法は、2.3.1 項において述べた、本システムの測位可能性の高さと精度のばらつきの特性に適応することを目指している。本システムで元手法を利用する際の問題は大きく 2 つある。まず第 1 に、元手法におけるユーザの滞在判定基準の仮定が適当でないこと、第 2 に測位誤差のばらつきの問題への対処が必要なことである。

まず、第 1 の問題である、ユーザの滞在判定基準の仮定の不適当さについて述べる。前述のように、元手法では GPS 信号を受信できなくなることが、室内での滞在を意味する、という仮定が利用されている。しかし、滞在を室内に限定することはいわば消極的な仮定であり、Ashbrook らの利用した機材の制約を逆説的に利用しているにすぎない。対して、本システムでは gpsOne の測位機能を用いることで、ほぼどこでも測位を行うことができるため、室内においても信号を失って測位できなくなることはなく、この仮定を利用するのは不適当である。

本システムのどこでも測位できる利点を生かすためには、元手法の place の検出に相当する作業を行う別手法を見つける必要があった。単純な代替案として、履歴中の時間的に隣接する 2 回の測位地点の距離を計り、一定距離以下に収まるものを元手法の place に相当するものとして取り出し、それをクラスタリングするという手法を実装してみた。しかし、低頻度の測位と精度がばらつく履歴の特徴のため、測位地点の距離だけから滞在地点を判定するのが難しいことが分かった。これらの特徴を考慮しない単純なクラスタリング

では、第 2 の問題として、3.2.2 項で述べた、測位誤差のばらつきへの対処法が問題となる。

3.3.3 本システムの滞在位置判定手法

3.3.2 項で述べた問題に対処するより適切な手法を検討した結果、本システムでは、手順を次のように変更した。この手順を 3.2 節で述べた多段階の半径について、それぞれの大きさで実行する。

- (1) 履歴中測位地点のクラスタリングによる「有意位置候補」の生成。
- (2) 有意性の基準に基づくフィルタリングによる候補のふるい分け。

我々の手法では、先にクラスタリングを行う。この段階では、履歴中の測位地点をクラスタリングによっていくつかの有意位置候補にまとめ、一度各測位地点をいずれかの候補に所属させる。本システムのクラスタリングモジュールは、Ashbrook らが提案した k-means 法の変種について、そのアルゴリズムの解説に基づき、新たに実装したものである。

クラスタリングの際、3.2.2 項で述べた測位精度のばらつきの問題に対処するため、測位結果の推定誤差に基づく「切り捨て処理」を導入する。この切り捨て処理は、次のように行う。半径 r を用いたクラスタリング時に、履歴上のある測位記録の推定誤差がその半径 r を超える場合、その測位記録はクラスタリング対象から除外する。この推定誤差の値としては、2.3.1 項で述べた $majoraxiserr$ を利用する。履歴中の $majoraxiserr > r$ であるすべての測位地点は、クラスタリング処理時に対象から除外される。

次に、フィルタリングをこれらの候補に対して行い、ユーザが滞在した位置とそれ以外の位置のふるい分けを行う。この段階により、有意位置候補の中に含まれる有意性の低い位置を除去する。

このフィルタリング部分については、2 つの手法を検討した。当初の案は、フィルタリングの有意性基準に、有意位置候補中の測位地点数について、履歴中の全地点数に対する比率を求め、この比率が一定の閾値以下のものを除去するという手法である。この手法については別稿⁶⁾において論じた。

しかし、この測位地点数比率による閾値を用いた手法は、適切な閾値の設定が難しいという問題があった。たとえば、本来の有意位置の周辺などの、滞在位置ではないがよく通過する、といった領域に形成される有意位置候補は、履歴が長くなると通過中にたまたま測位される回数が増えるため、閾値を超えてしまい除去されなくなってしまう。このような状況を避けるには、閾値を適切に設定しなおす必要があるが、比率を用い

た閾値は適切な値の設定が難しかった。

3.3.4 総滞在時間の閾値に基づくフィルタリング

3.3.3 項で述べた問題に対応するため、より履歴の増加に影響されにくく、閾値の設定を行いやすい手法を検討し、総滞在時間の閾値に基づくフィルタリング手法を用いることとした。この手法は、測位地点数比率に基づく手法に比べると若干複雑な処理になるが、計算量は大きく変わらない。この手法の手順を以下に示す。

- (1) クラスタリングによって有意位置候補を生成。
- (2) 測位履歴中において隣接する 2 回の測位地点が、同じ候補の範囲内か判定。
- (3) 2 点が同じ候補の範囲内の場合、両者の時間差を計算し、対象候補の滞在時間カウンタに追加。
- (4) これを全履歴にわたって行い、各候補の総滞在時間を取得。
- (5) 設定した総滞在時間の閾値を下回る候補を除去。

4.2 節では、この手法の性質、特に総滞在時間の閾値について実験結果をもとに議論する。

3.4 関連研究

本論文の有意位置学習の手法は、本章で述べたように、Ashbrook ら^{1),2)}の手法を改良したものである。ほかに、GPS などの測位システムを用いた人間の位置と行動の分析に関する関連研究をいくつか示す。

文献 3) では、各種のベイジアンフィルタを用いて利用者の位置推定を行う手法についてまとめている。特に、ローカルエリアにおいて赤外線や超音波などのセンサを統合して利用する状況について述べている。文献 4), 5) では、GPS 履歴と GIS データに基づき、動的ベイジアンネットワークを用いて精度の高い位置推定と予測を行う手法を提案している。モデルは低抽象度の層と高抽象度の層から構成される。低抽象度層は GPS による位置と速度の履歴から構成され、高抽象度層は目的地と移動手段の種類から成る。これらの研究は、今後本研究によって得られる有意位置の応用として、位置と操作の関連付けのモデルを作成する際などに参考にできると考えられる。

4. 議 論

本章では、本研究の提案手法の 2 つの論点について、被験者の利用履歴からの実験結果に基づいて議論する。第 1 の論点は、有意位置候補のフィルタリング手法に関する問題である。次に、クラスタリング段階において用いる有意位置の半径としてどのような値を使うかという問題について議論する。

4.1 テストケース

2 人の被験者に実装した履歴収集システムを日々持ち歩いてもらい、実験データとして利用履歴を収集した。ケース 1 の利用者からは、約 21 日分に相当する 3,130 点の測位履歴を得た。ケース 2 の利用者からは、約 27 日分に相当する 4,020 点の測位履歴を得た。以下の議論は、この実験データに基づいて行う。

被験者は 2 人とも大学院生であり、生活の基本的なパターンは家と学校との往復で、これに周辺地域や休日などの別行動が混ざる、という行動パターンである。

4.2 有意位置候補の学習

3.3.2 項で述べた改良した手法の性質を検証するため、後述する 4 段階の代表半径である 100, 250, 750, 2,200 m を用いて実行した結果学習される有意位置候補について、その総滞在時間別の個数を示す。ケース 1 の利用者について表 2 に、ケース 2 について表 3 に、それぞれ示す。なお、表中において「X ~ Y 分」と書かれている項目は、X 分以上 Y 分未満を意味する。この結果に基づき、有意位置候補のふるい分けを行う閾値について述べる。

これらの表で一見して分かるように、有意位置候補のうちかなりの割合は総滞在時間 0 分である。つまり、滞在中ではなく移動中に測位された点によって形成された候補であり、除去すべき対象である。このように総滞在時間 0 となる候補がかなり多いのは、総滞在時間に基づくフィルタリング手法が、不要な有意位置候

表 2 ケース 1: 総滞在時間別の有意位置候補の地点数
Table 2 Case1: numbers of significant location candidates and total staying time.

総滞在時間 (分)	100 m	250 m	750 m	2,200 m
0 分	153	102	77	39
1 ~ 15 分	7	4	2	8
15 ~ 25 分	4	5	3	3
25 ~ 35 分	1	2	1	0
35 ~ 45 分	3	1	3	1
45 分以上	9	11	10	9
15 分以上	17	19	17	13

表 3 ケース 2: 総滞在時間別の有意位置候補の地点数
Table 3 Case2: numbers of significant location candidates and total staying time.

総滞在時間 (分)	100 m	250 m	750 m	2,200 m
0 分	220	146	88	41
1 ~ 15 分	13	8	2	5
15 ~ 25 分	6	3	4	1
25 ~ 35 分	1	2	1	2
35 ~ 45 分	0	2	0	0
45 分以上	14	15	17	14
15 分以上	21	22	22	17

補を除去するという目的において有効であることを意味しており、望ましい結果であるといえる。

0分以上一定時間滞在している点についても、適切な閾値となる滞在時間を判断する必要がある。これに関しては、表の候補個数のみからは適正値は判断できないので、生成された個々の有意位置について検証し、実際のユーザの行動を考慮に入れ有意性を考察した。

2.3.2 項で述べたように、測位間隔は基本的に約 10 分である。つまり、2 回の連続した測位にわたって同一地点にいる場合に約 10 分滞在していると判定される。個々の有意位置候補を検証すると、総滞在時間が短い候補の多くは、1 度だけたまたま短時間滞在した地点であり、有意とするには偶然による場合が多かった。総滞在時間 15 分を超えると、一定の有意性が認められた。少なくとも 1 度、3 回の測位分（約 20 分）以上にわたって滞在しているか、もしくは 2 回以上同じ場所で滞在したと判断できる。今回は閾値として 15 分を用いるのが妥当であると判断し、表 2、3 中の最後の行に、この閾値を超えた有意な候補の地点数を示した。

4.3 半径の大きさの選択

次の問題として、有意位置の適切な半径の大きさについて考察する。多段階の半径を用いるため、適切な段階数で代表となる半径を選ぶ必要がある。今回のシステムでは、3.2 節で述べた多段階の有意位置利用の 2 つの目的である、空間的行動の多様なスケールの表現と、測位精度のばらつきへの適応を実現できるように代表半径を選択した。以下にそれぞれの観点からの半径の選択について考察する。

この問題を考えるにあたり、GeoOnion RDF/XML vocabulary⁹⁾ における人間の空間的行動のスケールの設定についての議論を参考にした。この議論では、人間の行動スケールに段階を設ける場合、基準距離の n 乗をスケールとして用いるのが妥当であり、特に 3m の n 乗の距離を単位とするものが利用しやすいとし、1m から 387km に及ぶまでの範囲でスケールとして用いることが提案されている。

実際に我々のテストケースに基づいて、半径の段階について様々なパターンをテストした。等差間隔を用いると、たとえば同じ 250m の間隔でも、250m と 500m の半径で学習した有意位置には明らかな違いがあるのに対し、2,000m と 2,250m の半径では結果にほぼ違いがない。本システムにおいても、等差間隔による段階よりは基準距離の n 乗段階を用いた方が適切であることが分かった。以下の半径に関する議論では、この 3m の n 乗ルールを参考にして考察を行う。

4.3.1 空間的行動パターンに関する考察

まず半径を決めるための第 1 の観点として、空間的行動のスケールについて考察する。利用者の空間的行動パターンに対し、3.2.1 項で述べたように、段階的な行動モードを仮定した。具体的には静止、徒歩、乗車モードの 3 つの行動モードを仮定し、これらの各モードに対して代表する半径を選ぶこととした。モードが後者になるに従い、ユーザの行動範囲がより大きくなるので、有意位置の半径をより大きいものとした。

ここで、各モードにおける半径とその選択基準を考える必要がある。有意位置はユーザが滞在した領域であるから、一定時間内での移動量に基づき滞在の有無を判定することとすると、ある一定時間内での移動が、指定半径以内に収まれば滞在、そうでなければ非滞在、という判断が行える。そこで、実験中に測位の時間間隔として用いた 10 分間での移動距離を、基準単位として考え、この距離に基づいて各モードの半径を選択することとした。測位履歴中の 2 点の移動距離がこの半径の円内に収まれば、滞在していたと判定する。表 4 に、各モードの想定移動速度と、その速度での 10 分間での移動距離の値を示す。

各モードの代表半径として、表 4 に示した d の範囲から適当な値を選ぶ。この際、3m の n 乗ルールを参考にすると、各モードについて 3m の 5, 6, 7 乗に相当する 243m, 729m, 2,187m が利用できる。利便性を考慮して値を 50m 区切りのもので近い値を選ぶと、250m, 750m, 2,200m になる。今回はこれらの値を、それぞれ静止、徒歩、乗車モードにおける代表とした。

4.3.2 切り捨て処理の影響

次に、3.3.3 項で述べた、精度のばらつきに対応するためにクラスタリング時に行う、切り捨て処理の有意位置生成への影響について考察する。

まず、切り捨て処理に利用する推定誤差値の傾向について示す。3.3.3 項で述べたとおり、推定誤差値は、2.3.1 項で述べた *majoraxiserr* を利用する。

図 3 に、テストケースの履歴について、推定誤差値がどの程度に収まるかを示す。横軸は半径の大きさを m 単位で示す。縦軸は、推定誤差の値が横軸にお

表 4 各行動モードと速度、10 分間における移動距離
Table 4 Movement modes, speed and distance in 10 minutes moving.

行動モード	速度	距離/10 分
静止	$s < 3 \text{ km/h}$	$d < 500 \text{ m}$
徒歩	$3 \text{ km/h} < s < 6 \text{ km/h}$	$500 \text{ m} < d < 1 \text{ km}$
乗車	$6 \text{ km/h} < s$	$1 \text{ km} < d$

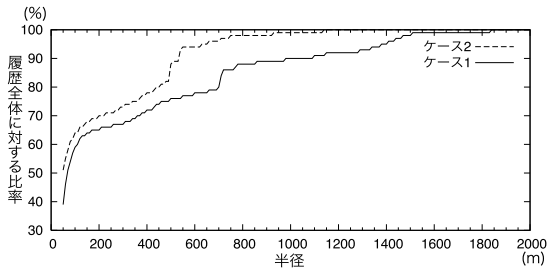


図3 *majoraxiserr* が半径より小さい測位点の比率

Fig. 3 Ratios of measured points which have *majoraxiserr* smaller than radii to the entire logs.

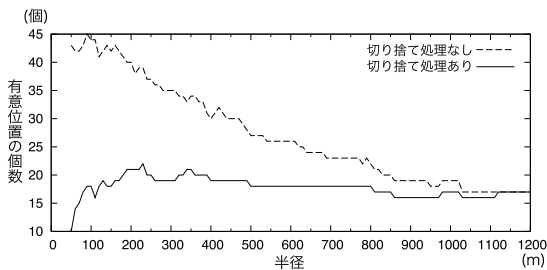


図4 切り捨て処理の有意位置個数への影響

Fig. 4 Effect of cut-off processing for numbers of significant locations.

いて指定された半径の値より小さい測位点の、履歴全体に対する比率を示す。切り捨て処理を行わなければ測位履歴の全地点が対象となるのに対し、切り捨て処理を行うことで縦軸で示される割合の点のみが対象となり、その他の点は除外される。したがって、縦軸で示される割合は、半径 r でのクラスタリング時に、測位履歴中の全地点に対する学習対象となる点の比率を示すものである。なお、最大の推定誤差値はケース1で2.1 km、ケース2で1.5 kmであった。

このような推定誤差値の傾向をふまえ、切り捨て処理が有意位置の個数に与える影響を検討する。図4は、ケース1の履歴に対して、10 m 間隔の各半径においてクラスタリング処理を行い学習された有意位置の個数を示す。切り捨て処理の影響を示すため、この図では切り捨て処理を行った場合と行わなかった場合の2つを比較する。1,200 m を超えると切り捨て処理の有無が有意位置個数にほぼ影響を与えなくなるので、それ以下の半径について示した。

図4において、両者の個数の差となるものは、3.2.2 項に示した精度の問題を主要因として、本来生成されるべき有意位置のずれた周囲に過剰に生成される有意位置である。これらの違いを具体的に評価するには、ユーザ自身が生成される有意位置の有効性を主観的に評価する必要がある。表5に、テストケース1

表5 切り捨て処理と有意位置の有効性の関係

Table 5 Relations of cut-off processing and validness of significant locations.

半径	切り捨て処理あり			切り捨て処理なし		
	有効	過剰	有効率	有効	過剰	有効率
2,200 m	12	1	0.923	12	1	0.923
750 m	17	0	1	17	6	0.739
250 m	19	0	1	23	14	0.622
100 m	14	2	0.875	23	23	0.5

についてユーザ自身がこの評価を行った結果を示す。切り捨て処理の有無、両者の場合について有意位置の学習を行い得られた結果について、実際に必要と考えられる有効な位置と、過剰に生成されたと考えられる過剰な位置に分類し、それぞれの個数を数えた。有効率は有意位置全体に対する有効な位置数の比率である。2,200 m では切り捨て処理がほぼ影響を与えないので、両者の違いはない。しかし、半径が小さくなるにつれて、切り捨てなしの場合の有効率は下がり、過剰に生成される位置が増加している。切り捨て処理の有無、両ケースの有効率の比較により、切り捨て処理が過剰な位置の生成を適切に防いでいることが分かる。

一方、半径250 m、100 m の場合に、有効な位置の個数自体が若干切り捨て処理によって減っており、有意位置生成を過度に防いでいる可能性がある。これらの差となる位置を個々に検証したところ、生成元となる履歴中の測位地点群の推定誤差値が大きいために、小さい半径でのクラスタリング時に切り捨て処理によって捨てられ、有意位置候補として生成されなくなったものであった。つまり、切り捨て処理は、主観的に見て無意味でない位置であっても、測位精度が低い場合はそれらの位置の生成を防いでしまうことがある。しかし、これは測位精度に応じた有意位置の生成を行うという目標の代償であるといえ、ある程度やむをえないと考えられる。

4.3.3 切り捨て処理の効果と代表半径

次に切り捨て処理を適切に機能させる代表半径の選択について考察する。

一般に、小さいサイズの半径はより詳細な位置情報を示せると考えられるが、精度の問題を考慮に入れると、小さければよいというわけではなく、一定の大きさで有意性の限界に達する。

たとえば、今回のケースでは半径50 m は小さすぎる。図4に見られるように、半径50 m から100 m にかけて有意位置の個数が増加し、100 m 前後でいったん落ち着く。実際に生成された個々の有意位置を調べたところ、ほぼすべての50 m 半径の有意位置は、1対1対応する100 m の有意位置を持つことが分かった。

図3に示したように、切り捨て処理を行いながら半径50mを用いた場合、50%以上の測位点が捨てられている。50m程度の半径での有意位置は小さすぎるため学習時にかなり多くの情報が捨てられてしまい、独立した意味のある位置を形成することができないのである。今回のテストケースの履歴では、最小の有意位置半径としては100m程度が限界だと考えられるので、この半径を最小のものとして用いることにした。

一方、半径が大きい側については、半径1,500mを超えた時点で、切り捨てられる点数は全体の1%を下回った。行動モードの代表のうち最大の半径である2,200m半径は、切り捨て処理の影響を受けない程度に大きい半径の代表と考えることもできる。

4.3.4 4段階の代表半径

以上で論じた観点に基づき、今回は100m, 250m, 750m, 2,200mの4段階の半径を用いるのが適当であると判断した。残る問題は、4段階の半径で十分であるか否かである。より多くの段階を導入した場合の効果を見るため、中間になる500mと1,000mの半径で有意位置を生成し、生成された個々の有意位置の性質について検証した。結果として、ほとんどの500mと1,000m半径の有意位置は、対応する750m半径の有意位置を持っていた。この結果から、500m, 1,000mの中間半径を導入する意味は小さいと考えられる。

図1に示した例は、これらの半径の値に基づいて学習したものである。この例で示されるように、これら4段階の半径を用いた有意位置の表現は、半径サイズの階層間と隣接する有意位置間の関係がバランス良く構成されており、3.2.1項で述べた多段階半径の利用目的がよく達成されているといえる。

5. おわりに

本論文ではgpsOne技術による測位機能を持つ携帯電話から得られる測位履歴を対象とした有意位置の学習手法を提案した。コンテキストウェアアプリケーションにおいて利用しやすい、抽象度の高い位置情報の表現である有意位置を、対象とする測位履歴を集約して得る手法について述べた。提案手法では、有意位置候補をクラスタリングによって生成したのち、各候補の総滞在時間に基づくふるい分けを行うことで有意位置を学習する。この際、多段階の半径を用いることで、ユーザの空間的行動のスケールに対して多段階に適応させるとともに、機器から得られる測位履歴の精度の特徴に対応させる。被験者の利用結果に基づき、提案した手法の性質とパラメータについて考察を行った。結果として、有意位置候補のふるい分けの基準

として総滞在時間15分を閾値とし、半径の段階としては100m, 250m, 750m, 2,200mの4段階を用いるのが適切であることを示した。今後、本システムによって学習された有意位置の応用について検討していく。有意位置に基づく各種の確率的モデルを構築することで、有意位置からの関連事象の推測と予想を行うことを考えている。まず、有意位置間の時間的推移関係をモデル化し、近未来の位置の予測を行う手法について検討している。また、本研究の目標である、コンテキストと操作のルール化に向け、有意位置と操作履歴を関連付けるためのモデルについて検討を行う予定である。

謝辞 システムの実装と実験に協力していただいた慶應義塾大学大学院政策・メディア研究科の加藤文彦、豊田陽一、水鳥敬満、太田尚志の各氏と、ご指導いただいた環境情報学部小檜山賢二教授に、つつしんで感謝の意を表する。

参考文献

- 1) Ashbrook, D. and Staner, T.: Learning Significant Locations and Predicting User Movement with GPS, *Proc. IEEE International Symposium on Wearable Computers*, pp.101–108 (2002).
- 2) Ashbrook, D. and Starner, T.: Using GPS to Learn Significant Locations and Predict Movement Across Multiple Users, *Personal and Ubiquitous Computing*, Vol.7, Issue 5. pp.275–286 (2003).
- 3) Fox, D., Hightower, J., Liao, L., Schulz, D. and Borriello, G.: Bayesian Filters for Location Estimation, *IEEE Pervasive Computing Magazine*, Vol.2, Issue 3, pp.24–33 (2003).
- 4) Liao, L., Fox, D. and Kautz, H.: Learning and inferring transportation routines, *Proc. National Conference on Artificial Intelligence*, pp.348–353 (2004).
- 5) Patterson, D., Liao, L., Fox, D. and Kautz, H.: Inferring High-Level Behavior from Low-Level Sensors, *Proc. International Conference on Ubiquitous Computing*, pp.73–89 (2003).
- 6) Toyama, N., Ota, T., Kato, F., Toyota, Y., Hattori, T. and Hagino, T.: Exploiting Multiple Radii to Learn Significant Locations, Location- and Context-Awareness, *Proc. 1st International Workshop, LoCA 2005*, p.157, Springer-Verlag GmbH (2005).
- 7) gpsOne. <http://www.cdmatech.com/>
- 8) Java 2 Platform, Micro Edition (J2ME), Sun Microsystems. <http://java.sun.com/j2me/>
- 9) Brickley, D., et al.: GeoOnion ESW Wiki.

<http://esw.w3.org/topic/GeoOnion>

(平成 17 年 3 月 31 日受付)

(平成 17 年 10 月 11 日採録)



遠山 緑生 (正会員)

1997 年慶應義塾大学環境情報学部卒業。1999 年同大学大学院政策・メディア研究科修士課程修了。2005 年同博士課程単位取得退学。2002 年同大学インフォメーションテクノロ

ジーセンター助手。2005 年より同大学デジタルメディア・コンテンツ統合研究機構講師。修士 (政策・メディア)。Web 技術, システムソフトウェア, コピキタスコンピューティング等に興味を持つ。日本ソフトウェア科学会会員。



服部 隆志 (正会員)

1962 年生。1985 年京都大学理学部卒業。1990 年同大学大学院理学研究科博士課程研究指導認定退学。1990 年同大学数理解析研究所助手。1993 年理学博士号を取得。1994 年慶應義塾大学環境情報学部専任講師。2001 年より同助教授として現在に至る。ビジュアルプログラミング, 実演によるプログラミング等に興味を持つ。ソフトウェア科学会会員。



萩野 達也 (正会員)

1958 年生。1981 年京都大学理学部卒業。1983 年同大学大学院理学研究科修士課程修了。1986 年エジンバラ Ph.D. コース修了。Ph.D. in Computer Science. 1987~1990 年京都大学大型計算機センター助手。1990 年慶應義塾大学環境情報学部専任講師。1992 年同助教授。2001 年同教授, 現在に至る。1997 年より World Wide Web Consortium Deputy Director for Asia。プログラミング言語, システムソフトウェア, 情報技術教育等に興味を持つ。電子情報通信学会, ソフトウェア科学会, ACM 各会員。