Vol.2011-MBL-57 No.4
Vol.2011-UBI-29 No.4
2011/3/7

不定期測位位置情報履歴に基づく行動パタン解析手法

中原 史貴[†]　　村上 隆浩[†]

GPS機能付き携帯電話等を用いて記録された不連続な位置情報の履歴から，利用者が頻繁に滞留する場所と時間帯を抽出する手法を提案する．本手法は測位記録の構成要素である位置情報(緯度及び経度)と時刻情報を，緯度-経度-時刻軸の３次Euclid空間上にプロットし測位記録同士をクラスタリングすることで滞留点を発見した．安定した結果を得るために，1日あたり平均10件程度の位置情報を測位している利用者で1人あたり700～1000件程度のデータが必要であることがわかった．

Behavioral Pattern Analysis Using Noncontiguous Position Logs

Fumitaka NAKAHARA[†]　Takahiro MURAKAMI[†]

This paper proposes a method to extract frequently visited locations from noncontiguous position logs recorded by GPS phone. This method is plotting the location data with latitude, longitude and time for Euclid space and clustering theirs. To extract stable frequently visited locations, this method needed about 700-1000 noncontiguous position logs for a person who is recording logs of about 10.

1. はじめに

1.1 背景

近年，GPS機能付の携帯電話やスマートフォンの普及により，利用者のいる場所に応じた情報を提供するサービスや，利用者がいる場所の情報を発信することができるサービスが提供されつつある．サービスとしては位置連動型情報配信，写真共有サービス，情報共有サービス，ゲームなどがある．

位置連動型情報配信としては，NTTドコモのiコンシェル[1)]がある．GPS機能付端末で取得した利用者の現在地情報をもとに，情報サービス事業者が登録した情報を配信するサービスが提供されている．またGPSを利用した位置情報の取得を5分間隔で行うオートGPS機能を利用した場合には，利用者があらかじめ入力しておいた場所と時間に関する内容を設定することでリマインドするサービスや，現在地から自宅最寄り駅までの終電情報を受け取ることも可能となる．また，経済産業省が平成19年度から21年度まで行っていた情報大航海プロジェクトの一つであるマイ・ライフ・アシストサービス[2)]では，10分から30分の間でGPSの測位間隔を設定することで利用者の位置情報を収集し，収集した位置情報を解析することでグルメ情報の推薦や商店街での購買連携サービスを提供する実証実験を行った．位置連動型情報配信の一つである位置連動型広告は，携帯電話の位置情報を利用して利用者の現在位置に近い店舗やイベントなどのローカル広告を配信する．このような広告配信は，ネット広告に出稿している事業者にとって，全国に向けて広告を配信するのではなく，特定のエリアに絞って広告を出稿したほうが少予算で済む為，事業者からのニーズは高い．

また，情報共有サービスではTwitter[3)]の利用者が急増している．Twitterでは利用者の設定によりつぶやき時に位置情報の付加を選択することが可能で，利用者がある日時にどこにいたのか情報を公開することができる．Panoramio[4)]をはじめとした位置情報をベースとした写真共有サービス，コロニーな生活PLUS[5)]を始めとした位置情報連携ゲームも同様に着実に利用者数が伸びている．

このように利用者の位置情報を活用したサービスは次々と登場しているが，一方ではGPS測位によるバッテリー消費問題と利用者のプライバシー意識の問題が懸念されている．

GPS測位によるバッテリー消費問題については，山田ら[6)]による調査によると，5分間隔で測位した場合には携帯電話は2日以上稼働するという結果が出ているが，これは利用者が毎日携帯電話を充電しなければならず，電池の消耗に見合ったサービスを享受できないのであれば，利用者として定期測位はしたくないと考えるものと思わ

† NEC サービスプラットフォーム研究所
Service Platforms Research Laboratory, NEC Corporation

れる．

利用者のプライバシー意識については，利用者が GPS で測位した位置情報をサービス事業者に提供することで，利用者の自宅，勤務先，通学先などの場所，および，その場所にいた時間が事細かに知られてしまうことを快く感じていない．MyVoice が 2010 年 11 月に行った携帯電話の位置情報サービスに関するアンケート調査[7]によると，携帯電話の位置情報サービス・アプリの利用経験は 2 割程度，GPS 機能の位置情報取得に対して抵抗感を感じている利用者は 4 割弱いるという結果が出ており，利用者のプライバシー意識の強さを裏付けされている．

しかし，プライバシー情報である位置情報であっても，毎日位置情報を取得する必要のない不定期測位による位置情報サービスは，定期測位の場合に比べ個人の行動を特定しづらくなる為，プライバシー情報に対する意識が薄れる傾向にある．この為，不定期測位による位置情報サービスを中心に利用者数を伸ばしている傾向にある．

尚，以下の章の説明で測位記録とは，利用者が GPS 測位により得られる情報で，少なくとも緯度情報，経度情報，測位日時を含む情報とする．

1.2 目的

1.1 で位置連動型広告の配信に対する事業者からのニーズは高い旨を述べたが，定期測位による高度の行動解析を行った位置情報連携サービスはプライバシー意識に対する利用者の抵抗感は強い．一方，不定期測位による位置情報連携サービスは GPS 測位時の現在地周辺の広告配信程度しか行われていない．そこで本稿では，不定期測位で測位記録を蓄積し，利用者が頻繁に行くことの多い場所とその時間帯を抽出することで，位置連動型広告などのサービスで利用者にとって，より精度の高い情報配信が可能となることを目的とする．

2. 関連技術

本章では，利用者の位置情報履歴を利用した行動解析技術ついて整理する．

経済産業省が平成 21 年度に調査を行っていた情報大航海プロジェクト(モデルサービスの開発と実証)の事業報告書[8]では，利用者から送られてくる測位記録を利用者が滞留した位置と時間の情報を持つ滞留記録と，利用者が移動した経路と時間の情報を持つ移動記録とに定義した上で行動情報解析を行っている．滞留記録は GPS で測位した測位記録について，利用者が連続３０分以上，半径５００ｍの円形エリア内に滞在した場所の位置情報，滞留開始日時，滞留期間の記録で，時系列順に測位位置を比較することで抽出している．また移動経路は，利用者が２つの滞留記録の間を移動した際の，移動開始時刻，移動終了時刻及び移動中の測位緯度経度と時刻を持つ情報で，移動記録の前後にそれぞれ滞留記録があることが前提である．但し，GPS の測位途切れがあった場合には，測位途切れのあった前後の測位データの場所に大きなずれがない場合には，一定の時間以内であれば滞留であると判定している．またマイ・ライフ・アシストサービスでは将来状態情報提供機能を持ち，指定された時刻，若しくは，時間帯に応じた利用者の将来のスケジュール情報の行動状態が提供されているが，前記滞留記録と移動記録をもとに推測を行っている．

また山田ら[6]による研究では，現在と空間的・時間的に同一となる過去の移動経路を特定することで，訪問地や経由地を予測する手法を提案している．

一般的に，行動パタン解析システムは，数分間隔から 30 分程度の間隔で位置情報が定期測位されることが前提で解析を行われている．

3. 不定期測位位置情報の解析

3.1 滞留点の抽出

本稿の行動情報解析は測位記録を表 1 に示す行動パタンに分けることで，利用者の行動の特徴を抽出することを示す．例えば位置連動型広告配信を行う際には，広告を配信したいエリアに滞留点がある利用者に対してのみ情報配信を行い，同じエリアに測位記録は存在しているが，移動経路と判定された利用者に対しては広告配信を行わないなどの選択が可能となる．また滞留点はどの時間帯にいる可能性が高いかについても抽出することで，事業者にとってメリットの多い時間帯に滞在している利用者に対して広告配信をするという選択が可能となる．

表 1　行動パタンの構成要素

要素	定義
滞留点	利用者が頻繁に滞留していると推測される地点．また,頻繁に滞在している可能性の高い時間帯の情報も持つ．
移動経路	利用者が頻繁に移動していると推測される移動経路．頻繁に前記移動経路を利用している可能性の高い時間帯の情報も持つ．

3.2 定期測位技術を適応した場合の問題点

不定期測位による測位記録は，利用者や取得日時によりその間隔やデータ量には大きなばらつきがある．この為，1 日数件の測位記録では，時系列に並んだ 2 つの測位記録間の時間が開き過ぎるということが多く発生する．図 1 はある利用者のそれぞれ

別の 1 日分の測位記録を示したものである．図 1 を見ると 1 日数件程度の測位記録の量ではどの辺りで頻繁に立ち寄っているのか，移動経路なのか判別することが困難であることがわかる．ここで，2 章で述べたマイ・ライフ・アシストサービスの解析手法を用いたとすると，時系列順に測位記録を比較していくため，測位間隔の時間が長すぎると滞留記録を抽出することができない．

2 章では主に定期測位を前提とした利用者の行動予測について多くの研究がなされていることを述べたが，これらの研究の多くは，利用者が位置情報を測位した時刻に滞留していたのか移動していたのか判別したのち，判別された情報をもとに滞留点を推測している為，定期測位のアルゴリズムがそのまま使えないことが多い．

3.3 不定期測位情報履歴の解析指針

3.2 で述べた通り，不定期測位の場合には測位記録同士の間隔が開けば開くほど，該時間帯の行動の不確定要素が多くなり，2 章で紹介した時系列で連続した測位記録の特徴を利用した解析が困難となる．

そこで本稿では不定期測位により蓄積された測位記録を 24 時間周期で区切り，測位記録に含まれる緯度・経度情報及び日時データのうち時刻データのみを取り出した時刻情報を，緯度-経度-時刻軸とする 3 次元空間上にデータをプロットした．図 2 はある利用者の不定期測位位置履歴情報をプロットした．実線で囲った時刻軸に平行なデータ集合と，破線で囲った緯度-経度平面に平行なデータ集合が確認できる．これはそれぞれ，時刻軸に平行なデータ集合は滞留点，緯度-経度平面に平行なデータ集合は移動経路と考えられる．つまり実線で囲ったデータ集合を抽出できれば，利用者が頻繁に立ち寄る場所を把握することができるようになる．同様に，破線で囲ったデータ集合を抽出できれば，移動経路を把握することができる．

尚，今回 3 次元空間上にデータをプロットした理由は滞留している時間帯について抽出しやすくすることと，その地点が滞留点なのか移動経路であるか区別しやすくする為である．

本稿では表 1 で定義した 2 つの行動パタンのうち，滞留点の抽出手法を提案する．また提案した手法について，滞留点の抽出に必要な利用者の測位記録がどの程度収集する必要があるか示す．

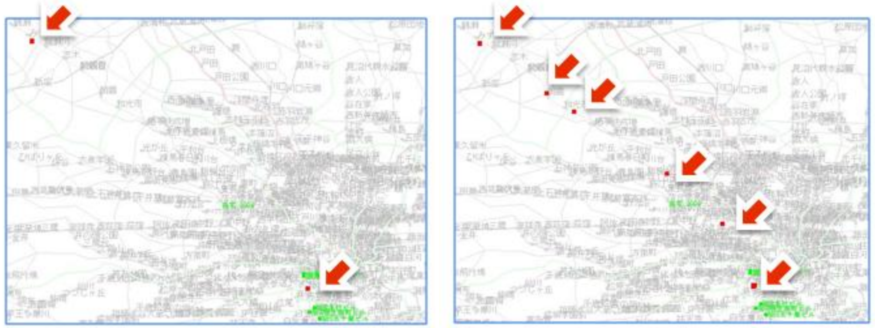

図 1　端末利用者の 1 日あたりの不定期測位位置情報履歴

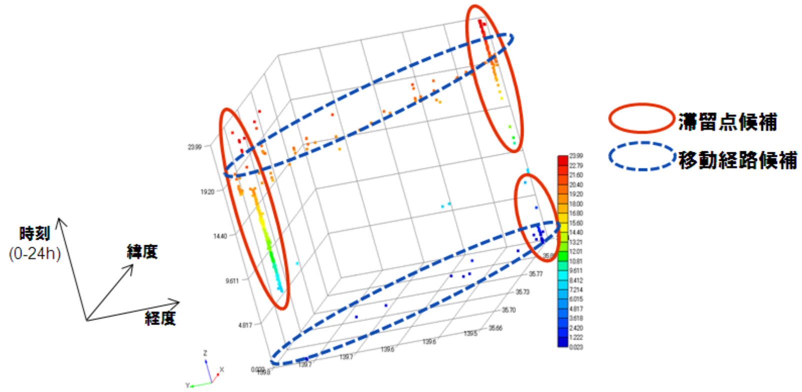

図 2　位置情報履歴の緯度-経度-時刻空間へのプロット

4. 不定期測位位置情報の滞留点抽出手法

4.1 全体構成

本滞留点抽出手法は，図 3 に示した流れで処理を行う．滞留点抽出フローの詳細な処理は 4.2 で述べていく．

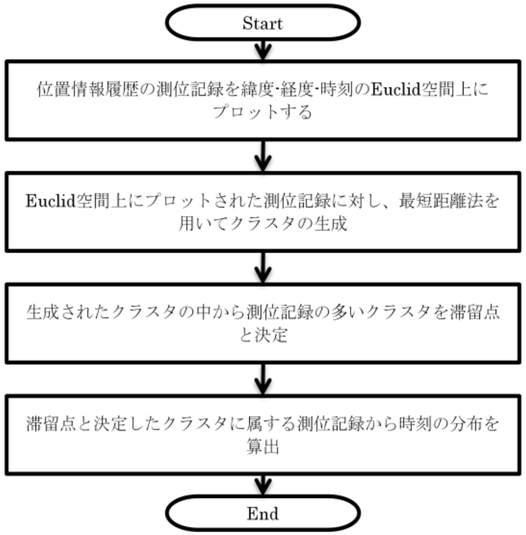

図 3 滞留点抽出フロー

4.2 アルゴリズム

滞留点抽出は，位置情報履歴に含まれる測位記録から滞留点候補となるクラスタを発見し，発見したクラスタから滞留点を特定する．具体的には，測位記録の構成要素である位置情報(緯度及び経度)と時刻情報を,緯度-経度-時刻軸の 3 次 Euclid 空間上にプロットし測位記録同士をクラスタリングすることで滞留点を発見した．以後，クラスタと言った場合には，測位記録同士のクラスタリング演算により求められた滞留点候補クラスタを指すものとする．

滞留点を発見するクラスタリングの計算手法で注意しなければならない点として，位置情報(緯度及び経度)に対し，滞留点と思われるクラスタが生成されやすくするように距離係数 k を定め，時刻に対し距離係数 k を掛け合わせなければならない．測位記録の時刻の単位を時間[h]とした場合の時刻変数をとすると，距離係数 k をかけた後の値$\mathbf{x_{i\,h'}}$は$\mathbf{x_{i\,h'}} = \mathrm{k}\,\mathbf{x_{i\,h}}$で表すことができる．距離係数 k は，滞留点の抽出には，緯度-経度平面に対し時刻軸に対し密となるクラスタが生成されるように値を設定する必要がある．

今回のシステムでは最短距離法を用いたクラスタリング[9]により滞留点の抽出を行った．クラスタリングで利用する測位記録は緯度，経度，及び距離係数を掛け合わせた後の時刻であるので，測位記録変数は$\mathbf{x_i} = (\mathrm{x_{i\,lat}}, \mathrm{x_{i\,lon}}, \mathrm{x_{i\,h'}})$と表すことができる．ここである測位記録を$\mathbf{x_a} = (\mathrm{x_{a,lat}}, \mathrm{x_{a,lon}}, \mathrm{x_{a,h'}})$,$\mathbf{x_b} = (\mathrm{x_{b,lat}}, \mathrm{x_{b,lon}}, \mathrm{x_{b,h'}})$，$\mathbf{x_{a\,h'}} > \mathbf{x_{b\,h'}}$ すると，$\mathbf{x_a}, \mathbf{x_b}$の Euclid 距離は，

$$\mathrm{d}(\mathbf{x_a}, \mathbf{x_b}) = \begin{cases} [(\mathrm{x_{a\,lat}} - \mathrm{x_{b\,lat}})^2 + (\mathrm{x_{a\,lon}} - \mathrm{x_{b\,lon}})^2 + (\mathrm{x_{a\,h'}} - \mathrm{x_{b\,h'}})^2]^{1/2} (\mathrm{x_{a\,h}} - \mathrm{x_{b\,h}} \leq 12) \\ [(\mathrm{x_{a\,lat}} - \mathrm{x_{b\,lat}})^2 + (\mathrm{x_{a\,lon}} - \mathrm{x_{b\,lon}})^2 + (24 - \mathrm{x_{a\,h'}} + \mathrm{x_{b\,h'}})^2]^{1/2} (\mathrm{x_{a\,h}} - \mathrm{x_{b\,h}} > 12) \end{cases}$$

とする．

$\mathrm{x_{a\,h}} - \mathrm{x_{b\,h}}$により条件を分けたのは，深夜 0 時を跨いだクラスタの生成をできるようにする為である．例として $\mathrm{x_{a\,h}} = 1$, $\mathrm{x_{b\,h}} = 23$の値を持つ測位記録の場合，同日の測位記録同士として計算した差分値は 22[h]であるが, 深夜 0 時を跨いだ測位記録同士として計算した差分値 2[h]となる．

またクラスタC_s, C_t間の距離は，

$$\mathrm{d}(C_s, C_t) = \min_{\substack{\mathrm{x_1} \in C_s \\ \mathrm{x_2} \in C_t}} \mathrm{d}(\mathbf{x_1}, \mathbf{x_2})$$

で定義される．クラスタの形成は$\mathrm{d}(C_s, C_t)$の値が小さいものから順にC_sとC_t のクラスタを結合する．測位記録のデータが N 個であるとした時，クラスタが結合される距離を$\mathrm{d^{(1)}}, \mathrm{d^{(2)}}, \cdots, \mathrm{d^{(j)}}, \cdots, \mathrm{d^{(N-1)}}$とすると，

$$\mathrm{d^{(1)}} \leq \mathrm{d^{(2)}} \leq \cdots \leq \mathrm{d^{(j)}} \leq \cdots \leq \mathrm{d^{(N-1)}}$$

の関係が成り立つ．ここでクラスタの結合を N-1 回行うと 1 つのクラスタとなる為，$\mathrm{d^{(j)}}$の値が予め定めたクラスタ閾値$\mathrm{d_{th}}$が$\mathrm{d^{(j)}} > \mathrm{d_{th}}$となった時点でクラスタリングの演算を終了する．ここまでの演算で得ることのできた測位記録の集合は，過去に滞留していたことを表す滞留点となる測位記録の集合と共に，過去に滞留していたとは考えにくい測位記録の集合も含まれる．この為，この測位記録の集合は滞留点候補となるクラスタと言える．以後，クラスタと言った場合には，測位記録同士のクラスタリング演算により求められた滞留点候補クラスタを指すものとする．

そこで, 前記クラスタから滞留点となるクラスタだけを抽出する処理が必要となる．ここではある測位記録数が 1156 件取得済みの利用者の実データを用いて説明を行う．本データを表 2 に示す距離係数 k 及びクラスタ閾値$\mathrm{d_{th}}$でクラスタリングの演算を行ったところ，表 3 のようにクラスタが 99 個生成された．また，図 4 は前記クラスタの演算結果の, 各クラスタに属する測位記録データ数の上位 20 件を示したものである．図 4 を見ると，上位 3 つのクラスタの測位記録データ数が突出して多いことが見て取

れるが，このような上位のクラスタこそが利用者が頻繁に立ち寄る滞留点と言える．この為，上位クラスタのみを抽出する演算が必要となる．本手法では，突出して測位記録データ数の多いクラスタの抽出する方法として，計算の簡略の為，便宜的に各クラスタに属する測位記録データ数は正規分布が成り立つと仮定し，クラスタの個数が予め定めた正規分布の上位の変域を示す滞留点域にあてはまるクラスタを滞留点とする．滞留点域を上位 15%と設定したところ，滞留点域の下限値は 61.60 となったので，この利用者の場合上位 4 つのクラスタを滞留点と見なした．

最後に滞留点と判定されたクラスタにおいて，そのクラスタに従属している測位記録から頻出時間帯を解析する．今回行った頻出時間帯の抽出方法は，滞留点候補のクラスタから滞留点を抽出した方法と同じく，測位記録の時刻分布は正規分布が成り立つとして，正規分布の信頼区間を頻出時間帯として当てはめた．今回は母平均の信頼区間を 60%として計算を行い，信頼区間の下限値と上限値をそれぞれ，頻出滞留開始時刻，頻出滞留終了時刻とした．また滞留点を代表する座標として，クラスタに従属する測位記録の緯度経度の値の平均を取り，滞留点の座標とする．このようにして表 4 のような滞留点情報を得ることができる．

尚，今回のシミュレーションでは計算簡略の為に，緯度変$x_{i\ lat}$と経度変数$x_{i\ lon}$で同距離とした仮定で計算を行っているが，実際には同単位では経度に比べ緯度の方向に距離が長くなる．この為，厳密には緯度変数$x_{i\ lat}$と経度変数$x_{i\ lon}$間に対しても距離係数の定義が必要である．

表 2　測位記録数及び各種設定値

項目	設定値
距離係数 k	4.0×10^{-4}
クラスタ閾値 d_{th}	2.5×10^{-3}
滞留点域	15%

表 3　サンプルユーザの位置情報履歴の演算結果

項目	演算値
生成クラスタ数	99
1 クラスタあたりのデータ数平均	11.677
1 クラスタあたりのデータ数分散	2320.037
滞留点域下限値	61.60

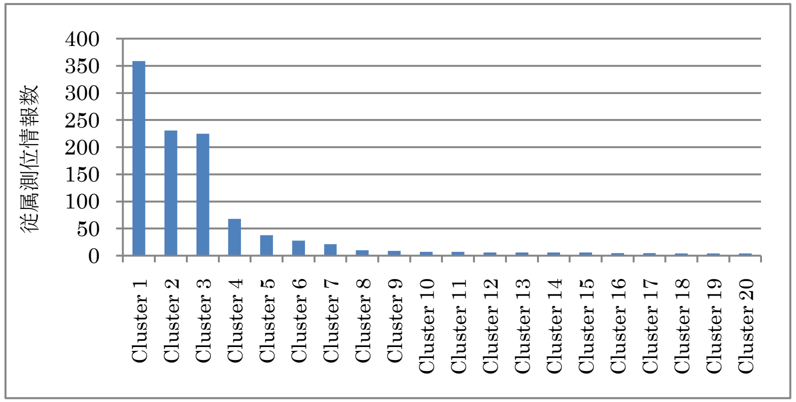

図 4　滞留点候補クラスタ毎の従属測位記録数

表 4 滞留点情報

項目	値
滞留点番号	2
緯度[°]	35.5735
経度[°]	139.6672
頻出開始時刻[h]	8.765
頻出終了時刻[h]	20.745

5. 評価

4 章で示した行動情報解析手法の性能評価を目的として，システムのプロトタイプを実装し評価実験を行った．以下，本手法を評価する為に利用した位置履歴情報の説明及び評価方法を述べた後，その評価結果を示す．

5.1 使用データと評価方法

本手法の評価には，前述の Twitter の「つぶやき」情報を利用した．1 章でも述べたように Twitter は利用者の選択により「つぶやき」には緯度・経度情報を含む位置情報を付与することができる．今回はこの位置情報が付与された「つぶやき」を数多く投稿している利用者のつぶやき日時と場所を抽出し，4 章で述べた滞留点抽出を行った．

今回調査 1 日あたり 1500 件の位置情報付きのつぶやきのあった 74 人分の分析を行った．尚，この 74 人分のデータは明らかに機械的につぶやきが生成されている利用者については予め省いたものである．

滞留点の抽出の評価は 74 人分のつぶやきを用いて，測位記録の量の変化に応じた滞留点の増減により変化量について分析を行い，測位記録の量の個数により滞留点の抽出にどのような現象が起きているのか検討を行った．

5.2 分析結果

本章では 5.1 章で説明した評価方法をもとにデータの分析を行った．

図 5 は利用者の測位記録の収集が進んだ時の滞留点の変化について示したもので，測位記録を 5 つ追加する度に追加前の追加後で滞留点の数の増減がどのように変化したかを示している．増減数は，滞留点の緯度経度の値が追加前と追加後で 500m 以上のずれがなかった場合には変化なし，滞留点の数が増加，減少した場合には増加数及び減少数の値を合わせてカウントした．例えば，追加前と追加後で滞留点の数に変化がないが，滞留点の場所が 500m以上変化した場合には，滞留点の増加数 1 と滞留点の減少数 1 とを足しわせた 2 を変化量してカウントしている．図 5 を見ると測位記録が 200 前後までは滞留点の変化量が極めて大きいことが確認できる．その後測位記録が増加するにつれて，滞留点の変化量が穏やかになることが確認できる．

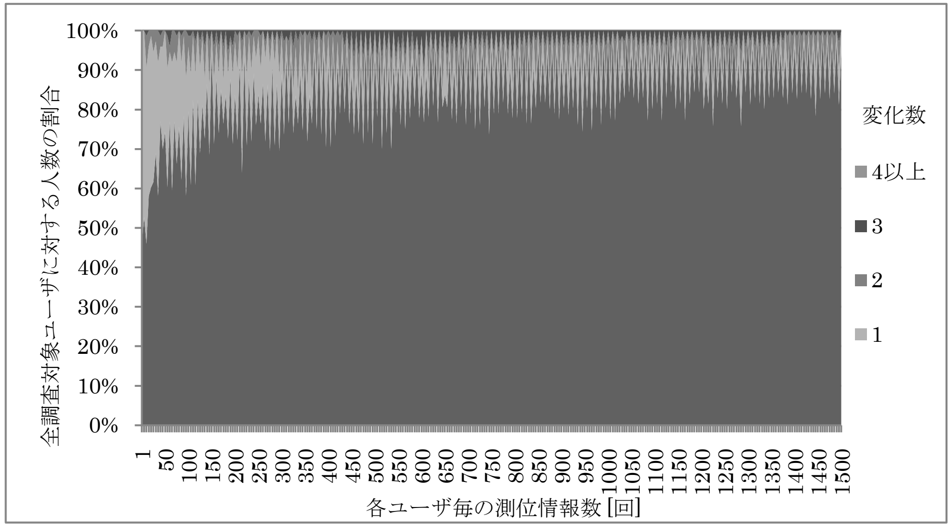

図 5 滞留点抽出位置の測位記録の増加による変化量

5.3 検討

ここからはある一般的な測位記録を提供している利用者の実データを用いて，測位記録の数に応じてどのような滞留点の抽出される傾向があるかについて説明していく．ここでは測位記録を蓄え始めてからの順で滞留点抽出の不安定期，準不安定期，安定期と呼ぶことにして,以下説明をしていく.また抽出された滞留点同士を区別する為，クラスタに属する測位記録数が多い順に，n 番目に多い滞留点を第 n 滞留点と定義する．

○ Phase1 不安定期

図 6 のように，滞留点の抽出数が 0～2 個程度と少なく，滞留点の位置が頻繁に変わる時期を不安定期と呼ぶことにする．不安定期では測位記録が数個追加されただけで滞留点の抽出される位置が変わる傾向にある．また滞留点の抽出数が少ない場合が多いが，これは滞留点の位置決定を統計的予測により算出しているため，わずかな測位記録の追加に敏感な為である．

一般的に測位記録のデータを取得開始してから 150 から 200 程度までが不安定期にあてはまる傾向が強い．

○ Phase2 準不安定期

図 7 のように，第 1 滞留点と第 2 滞留点の位置が測位記録のデータ数に関わらず安定してくる時期を準不安定期と呼ぶことにする．この時期になると滞留点の抽出数が 2 個以上抽出されることが多くなる．その抽出された滞留点のうち，第 1 滞留点と第 2 滞留点の位置が安定してくるが，多くの利用者の場合，自宅及び勤務地/通学先の２つの滞留点が抽出されるためと考えられる．ここで算出された第 1 滞留点と第 2 滞留点の位置は後で説明する安定期に入っても大きくは変化しない．但し，第 1 滞留点と第 2 滞留点の位置が入れ替わることも頻繁に見受けられる．しかし，第 3 滞留点以下の滞留点については測位記録のデータ数が乏しいことが多く，不安定期と同じくわずかな測位記録の追加で滞留点の位置が頻繁に変わることが多い．

また準不安定期のもう一つの特徴として，ほぼ同一の場所で別時間帯の滞留点が抽出されることが多い傾向にある．これは本来，利用者が長時間同一の場所に滞留している傾向があるにも関わらず，測位記録をデータ同士が時刻軸で見た場合に疎らである為に別のクラスタとして抽出される為に生じることが多い．

一般的に測位記録のデータ数が準不安定期に入ってから，700 から 1000 程度までが不安定期にあてはまる傾向が強い．

○ Phase3 安定期

図 8 のように，測位記録を追加しても追加前の滞留点から変化が起きにくくなる

Vol.2011-MBL-57 No.4
Vol.2011-UBI-29 No.4
2011/3/7

時期を安定期と呼ぶことにする．準不安定期に安定していた自宅及び勤務地/通学先と思われる第 1 滞留点と第 2 滞留点の滞留点だけではなく第 3 滞留点以下の滞留点の位置も安定し，その後の測位記録の追加により新たな滞留点が追加されても，利用者が最近よく訪れている為に滞留点が追加されたのだと感じることのできる，利用者の感覚に近い滞留点情報が提供可能になる. 一般的に 700 から 1000 程度の測位記録が蓄積されると不安定期にあてはまる傾向が強い．

5.2 章より，自宅及び勤務先/通学先と思われる第 1 滞留点，第 2 滞留点の抽出が可能となる準不安定期に突入するのはデータ数が 150 から 200 程度であることがわかった．これは 1 日あたり平均 10 件の測位記録を蓄積したとすると，15 日から 20 日程度で自宅と勤務先/通学先が抽出可能となる．同様に自宅と勤務先/通学先以外の滞留点の抽出には 700 から 1000 程度のデータが必要となるので，70 日から 100 日程度を要することなる．但し，本手法は時系列を考慮したアルゴリズムではない為，1 日あたりの測位記録の蓄積量を増やした場合には，滞留点の抽出までの期間はそれだけ短くすることが可能である．但し，測位記録を短い間隔で取得している利用者の場合は 2 章で説明した関連技術で紹介した定期測位を用いたアルゴリズムを利用したほうがよい結果が得られるものと思われる．例えば，マイ・ライフ・アシストサービスの場合 30 分間隔で定期測位を行っているので 1 日あたり 48 回の測位記録が取得できるので，平均 48 回程度以上の不定期測位の場合は, 関連技術で示したような定期測位を前提として行動解析手法を利用したほうが有用な場合が多いと思われる．尚，本手法は測位したことのある位置に対して滞留点を抽出するものであるので，測位したことのない位置を予測することはない．この為，利用者は頻繁に行っているが，その位置で測位をしていなければ, 何件もの測位記録があれども滞留点として抽出されることはない．

また本稿では曜日別の評価を行っていないが，曜日別の利用者行動解析についてもそのままの手法を利用することが可能である．但し，曜日別でない場合に比べ，滞留点抽出に必要となる測位記録の蓄積には時間がかかると想定される．

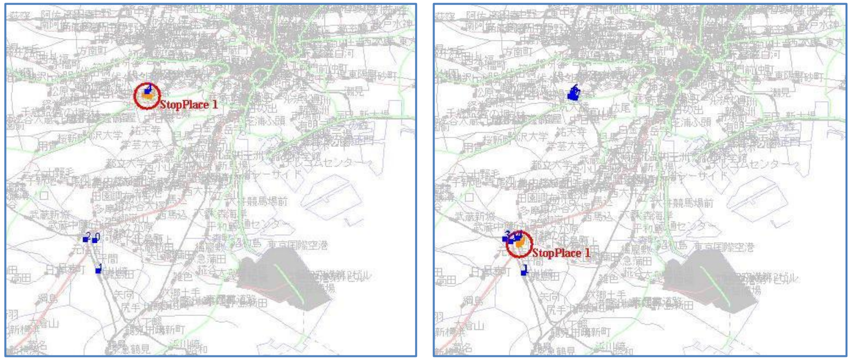

図 6　不安定期の滞留点の抽出（左：測位記録数 20 右：測位記録数 40）

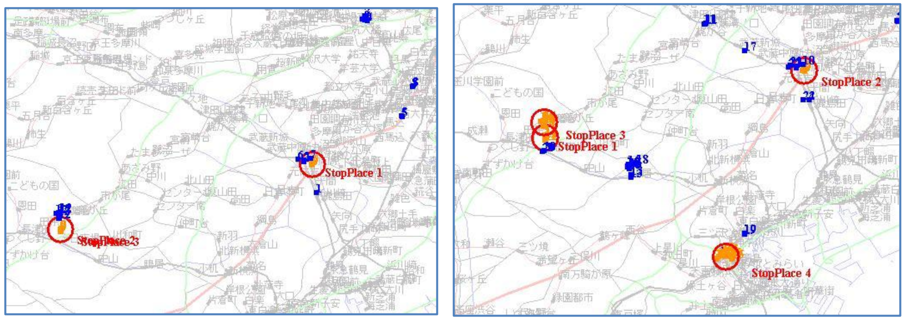

図 7　準不安定期の滞留点の抽出（左：測位記録数 200 右：測位記録数 500）

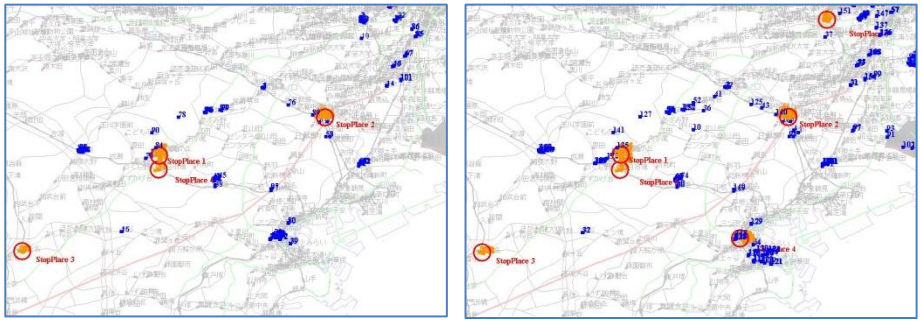

図 8　安定期の滞留点の抽出（左：測位記録数 1000 右：測位記録数 1500）

6. まとめ

本稿では不定期測位による測位記録を利用して利用者が頻繁に訪れる滞留点の抽出が可能であることを示した．また滞留点の抽出に必要となる測位記録のデータ数について，自宅及び勤務先/通学先の判定，それ以外によく訪れる場所の判定に必要となるデータ数の指針についても示すことができた．これにより位置情報の定期測位をすることなくなる為，GPS 測位時のバッテリー残量を気にすることなく端末を利用することができる上，滞留点という頻繁に訪れる場所と時間帯へとデータの抽象化をした為，過去のある時点にどの場所にいたのか晒さずに滞留点情報をサービス事業者に対して提供することで利用者に有用な情報を得ること可能となる．

今後は利用者の移動経路の抽出と，曜日別の滞留点及び移動経路についても抽出する機能について検討を進める予定である．

※謝辞

本研究を進めるにあたり数多くのご助言，ご協力頂いた皆様に謹んで感謝の意を表す．行動情報の解析にあたり可視化ツールである地図ビューアを提供して頂いたNEC大野岳夫氏には更なる感謝の意を表する．

参考文献

1) NTT ドコモ:i コンシェル, http://www.nttdocomo.co.jp/service/customize/iconcier/
2) 経済産業省:情報大航海プロジェクト平成 21 年度の取り組み　マイ・ライフ・アシストサービス, http://www.meti.go.jp/policy/it_policy/daikoukai/igvp/contents/activity09/modelservice/list/personal/ntt.html
3) Twitter, http://twitter.com/
4) Panoramio: http://www.panoramio.com/
5) コロニーな生活 PLUS: http://pc.colopl.jp/pages/wl/welcome.html
6) 山田直治,礒田佳徳,南正輝,森川博之:GPS 搭載携帯電話を用いた移動経路履歴に基づく訪問地・経由地予測システム，情報処理学会研究報告 ユビキタスコンピューティング研究会,”UBI-27-8”(2010)
7) マイボイス株式会社: 携帯電話の位置情報サービスに関するアンケート調査 http://myel.myvoice.jp/products/detail.php?product_id=14813 ,(2010)
8) 平成 21 年情報大航海プロジェクト事業報告書(モデルサービスの開発と実証) http://www.meti.go.jp/policy/it_policy/daikoukai/igvp/index/h22_report/main/model01.pdf
9) 水田正弘,佐藤義治,宮腰政明,河口至商: 最短距離法におけるクラスターの安定性, 北海道大学工学部研究報告第 100 号,pp.99-106(1980).