

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

山田 直 治^{†1,†2} 磯 田 佳 徳^{†1}
南 正 輝^{†2} 森 川 博 之^{†2}

本稿では、測位間隔が長く、少量の GPS 位置情報履歴から、ユーザの訪問地・経由地を予測する手法を示す。筆者らは、測位間隔が長く、少量の GPS 位置情報から、移動経路を逐次的に精練する検討を行ってきた。本稿では、ユーザの反復的な行動に着目し、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について述べる。被験者 6 名の 7ヶ月間の GPS 位置情報履歴を用いた評価実験により、68%の精度で訪問地を予測することが可能であることが示される。これにより、携帯電話の電力消費量を抑制しつつ、訪問地・経由地に関する情報を配信することが可能となる。

Movement Prediction based on Route histories for GPS enabled Cellular Phones

NAOHARU YAMADA,^{†1,†2} YOSHINORI ISODA,^{†1}
MASATERU MINAMI^{†2} and HIROYUKI MORIKAWA^{†2}

This paper proposes the movement prediction from sparse and few GPS locations acquired by the GPS enabled cellular phones. The authors have developed the incremental route refinement system from the sparse and few GPS locations. Focusing on the human's repetitive behavior, this paper identifies the past route that is spatially and temporally same as the current movement. An experiment with 6 subjects over periods of 7 months demonstrates that the proposed system can predict the user's destination by 68%. Owing to this system, the user can acquire the information on the predicted area with low power consumption of cellular phones.

1. はじめに

近年、いわゆる日本版 e911¹⁴⁾ の施行により、ほぼ全ての携帯電話に GPS が搭載されており、GPS で取得した位置情報を用いたサービスが注目されている。例えば、米国では、位置情報を用いたサービスが 7000 種類以上となり、今後も増加傾向にある¹¹⁾。

一方、携帯電話に搭載された GPS モジュールの省電力化により、ユーザの位置を定期的に取得することが可能になりつつある。例えば、NTT ドコモでは、約 5 分間隔でユーザの位置情報を取得するオート GPS を提供している¹²⁾。そのため、今後は、定期的に取得された位置情報を用いたサービスが多く出てくることが予想される。

定期的に取得された位置情報を用いたサービスの一つとして、ユーザがこれから通過・滞在する場所を予測し、それらの場所に関する情報を提供するサービスが考えられる。例えば、ユーザが次に滞在する場所（以下、訪問地と記載）に関する天気予報やイベント情報、レストラン情報を配信したり、ユーザがこの後通過する場所（以下、経由地と記載）で利用する電車の乗換案内情報・遅延情報や、道路の渋滞情報・事故情報を配信する。これにより、ユーザの行動を支援することが可能となる。

実用的な訪問地・経由地予測システムの実現に際しては、ユーザに負担をかけずに訪問地・経由地を正確に予測することが重要となる。これまでも訪問地・経由地予測システムが数多く提案されてきたが、要件を満たすことができていない。

実用的な訪問地・経由地予測システムを実現するためには、以下の 3 つの要件を満たす必要がある。

- (1) 携帯電話の利用 GPS 位置情報を取得する手段として、GPS ロガーなどの専用デバイスがある。しかしながら、ユーザは専用デバイスを常時携帯する必要があり、ユーザにとって負担となる。また、置き忘れや付け忘れなどが発生する可能性も高い。したがって、携帯電話に搭載された GPS を利用することが望ましい。
- (2) 測位間隔が長くても予測可能 訪問地・経由地予測に基づく情報配信を行う上では、システムはユーザの位置情報を定期的に取得し、訪問地・経由地を予測する必要がある。

†1 NTT ドコモ サービス&ソリューション開発部
Service & Solution Development Department, NTT DOCOMO, INC.

†2 東京大学 先端科学技術研究センター
Research Center for Advanced Science and Technology, The University of Tokyo

従来研究では、GPS 位置情報を高頻度で取得して訪問地予測を行っているが、携帯電話のバッテリー容量は限られている。そこで、通話やメールといった通常の携帯電話の利用に影響を与えないために、携帯電話の電力消費量を抑えなければならない。そのためには、電力を消費する GPS 測位の間隔を長くしなければならない。

- (3) 少ない位置情報履歴からでも予測可能 ユーザは、サービスを利用し始めて早い段階でサービスの価値を享受できないと使い続けてくれない。そのためには、GPS 測位履歴が少ない状況でも、ユーザの訪問地・経由地を予測できなければならない。

本稿では、上述した要件を満たす上で課題となる、測位間隔が長く、かつ少量の GPS 位置情報履歴からの訪問地・経由地予測に対して、逐次的に推定・精練した移動経路履歴に基づいて予測を行うことで、上記要件を満たす実用的なシステムを示す。まず (1) の要件に対しては、本研究では携帯電話に搭載された GPS を用いたユーザの位置情報を取得する。

次に (2) の要件に対しては、GPS 位置情報の取得間隔を 5 分間隔とする。5 分間隔で取得した場合の携帯電話の電力消費量に関する事前実験では、携帯電話は 2 日以上駆動し続けており、ユーザは基本的に毎日充電することを考えると、通常の携帯電話の利用に大きな影響を与えないと考えられる。

最後に (3) の要件に対しては、移動中に取得した連続する 2 つの位置情報を包含する矩形 (Minimum Bounding Rectangle, 以降、MBR と記載) によって位置情報間の移動経路を補間することで、ユーザが一度でも通過した移動経路を特定することができる。そして、過去と同一の移動経路を通過した場合には、新たに取得した位置情報を移動経路に追加することで、移動経路をより正確に特定することができる。

その上で、逐次的に精練された移動経路履歴に基づいて、訪問地・経由地の予測を行う。具体的には、我々の日々の行動の多くは、過去の行動の反復であることに着目し、現在と同じ曜日・時間帯に現在地を通過した過去の移動経路履歴に基づいて、ユーザの訪問地・経由地を予測する。

筆者らは、測位間隔が長く少量の位置情報履歴から移動経路を正確に推定する手法について検討を進めてきた⁸⁾。本稿では、推定した移動経路履歴に基づいて、ユーザの訪問地・経由地を予測する手法について述べる。

本稿の構成は以下の通りである。まず 2. で関連研究について整理し、本研究の位置づけを明確化する。次いで 3. で訪問地・経由地予測システムを示す。4. では 3. で示した訪問地・経由地予測システムの実装と評価について述べ、5. でまとめとする。

2. 関連研究

本章では、従来の訪問地・経由地予測システムについて整理し、移動経路に基づいて訪問地・経由地を予測することの重要性について述べる。次に、過去の移動経路に基づく訪問地・経由地予測システムを整理した上で、本研究で示す訪問地・経由地予測システムの差異を明確化する。

訪問地・経由地予測システムとして、従来の多くの研究では、ユーザが滞在した場所 (以下、滞在地と記載) を特定し、滞在地間の遷移関係に基づいて訪問地・経由地を予測する手法を提案している。Marmasse ら⁹⁾ はユーザが入った施設が滞在地であると考え、ユーザが施設に入ると GPS 衛星の信号が取得できなくなる特性を利用して滞在地を特定している。つまり、システムが定期的に GPS 位置を取得し、GPS 衛星の信号を取得できなかった場所を滞在地として特定する。一方、Ashbrook ら³⁾ や Kang ら⁴⁾、西野ら¹⁵⁾ は、ユーザが長時間いた場所が滞在地であると考え、Ashbrook ら³⁾ は K-means 法により、Kang ら⁴⁾ は time-based clustering により、西野ら¹⁵⁾ は DBSCAN⁷⁾ により、ユーザが長時間いた場所を滞在地として特定している。そして、Ashbrook ら³⁾ や Kang ら⁴⁾ はマルコフモデルにより、西野ら¹⁵⁾ は系列マイニング¹⁰⁾ により、滞在地間の遷移関係をモデル化している。これらの手法により、システムは滞在地間の遷移関係に基づいて、現在の滞在地から、次の訪問地を予測することが可能となる。

しかしながら、これらの手法には 3 つの問題がある。第 1 に、自宅やオフィスなどといった行動の起点となる滞在地では、常に遷移確率が最大の滞在地が訪問地予測結果となってしまう。その他の多くの滞在地について誤った予測を行ってしまう。例えば、自宅からの遷移確率最大の滞在地が会社であるとする、自宅にいる場合の次の訪問地は必ず会社になってしまう。第 2 に、これらの手法は滞在地のみに着目しているため、経由地の予測ができない。そのため、例えば、訪問地へ向かう電車の乗換案内情報や遅延情報などといった情報を提供することができない。第 3 に、ユーザが初めて訪れた滞在地では、他の滞在地との遷移関係が存在しないため、次の訪問地の予測ができない。例えば、ユーザが帰宅途中に、これまで入ったことのない飲食店で食事をしてから、自宅に帰ったとする。この場合、この飲食店から他のエリアへの遷移関係が存在しないため、次の訪問地として自宅を予測することができない。

これらの問題を解決するために、いくつかの研究では移動経路履歴に基づいて訪問地・経由地を予測する手法を提案している。Zhou ら¹⁾²⁾ は density and join-based clustering に

より移動経路を特定している。Liaoら⁵⁾⁶⁾は階層型マルコフモデル、階層型 Conditional Random Fields により移動経路を特定している。移動経路履歴を利用することで、現在的位置を過去に通過した移動経路を特定し、当該移動経路上で一定時間滞在した場所(以下、終着地と記載)、およびそこへ至るまでの経由地を予測結果とすることができる。これにより、上述したような、訪問地へ向かう電車の乗換案内情報や遅延情報などといった情報を提供することが可能となる。また、初めて訪れた滞在地からの訪問地・経由地についても、既知の移動経路上に復帰した時点で、訪問地・経由地を予測することができる。これにより、帰宅途中に、これまで滞在したことのない場所で滞在した場合でも、移動開始後に、通勤経路上を自宅の方向へ移動していることを検知して、自宅を次の訪問地と予測することが可能となる。

しかしながら、これらの手法には2つの問題がある。第1に、これらの手法は高頻度にGPS位置情報を取得することを前提としているため、携帯電話の電力消費量が増大する。これにより、携帯電話が頻繁にバッテリー切れとなってしまう、通話やメールといった携帯電話の通常の利用の妨げとなる。第2に、クラスタリング手法を用いる場合には、事前に大量のGPS位置履歴が必要となる。特に移動中はGPS位置情報がまばらになるため、十分なGPS位置履歴を蓄積するまでに長期間かかるが、それまでの間、訪問地・経由地を予測することができない。

このように、1.で示した要件、すなわち、測位間隔が長く、少量の位置情報履歴でも訪問地・経由地を正確に予測する手法は確立されていない。これに対し、本研究では、測位間隔が長い場合でも、過去と同一の移動経路を通過したことを検知して、新たに取得した位置情報を用いて過去の移動経路を精錬することで、正確な移動経路履歴に基づいて、訪問地・経由地を予測することができる。

3. 移動経路履歴に基づく訪問地・経由地予測手法

一般に、我々の日々の行動は、過去の行動の反復であることが多い。例えば、平日の朝は、いつもと同じ路線や道路を通過して出勤する。たとえ、出張先のように普段訪れない場所でも、そこへ行くために利用する路線や道路は同じであることが多い。そこで本研究では、現在地を同じ曜日・時間帯に通過した過去の移動経路を特定し、当該移動経路上で現在地以降に通過する場所を経由地として、また、当該移動経路の終着地を訪問地として予測する。

3.1 システム構成

訪問地・経由地予測システムの構成図を図1に示す。本システムは、GPS搭載携帯電話と

訪問地・経由地予測サーバから構成される。携帯電話上のGPSモジュールは定期的にGPS位置情報として、4つ組データ(緯度、経度、推定測位誤差、測位時刻)を取得する。訪問地・経由地予測サーバは、移動経路表示機能、同一経路判定機能、移動経路更新機能、空間的同一経路判定機能、時間的同一経路判定機能、通過確率算出機能の6つの機能から構成される。

図2に訪問地・経由地予測処理のフローを示す。まず、GPSモジュールが定期的にGPS位置情報を取得する。携帯電話は取得したGPS位置情報(緯度、経度、推定測位誤差、測位時刻)を、携帯電話網を介して訪問地・経由地予測サーバへリアルタイムに送信する。

次に、訪問地・経由地予測サーバの空間的同一経路判定機能は、現在地を通過する過去の移動経路を特定する。具体的には、現在の位置情報と空間的に重複する過去の移動経路を特定する。

現在地を空間的に含む移動経路が1つ以上存在すれば、時間的同一経路判定機能は、当該移動経路を通過した回数のうち、今回と同じ曜日、時間帯に通過した回数を特定する。

その上で、空間的に同一と判定された各移動経路について、通過確率算出機能は、今回と同じ時間帯に通過した回数と、異なる時間帯に通過した回数から、当該移動経路の通過確率を算出する。そして、通過確率に基づいて、訪問地・経由地を予測する。

その後、ユーザが一定時間滞在地にとどまったことを検知した場合に、移動経路表示機能は、移動中に取得したGPS位置情報を用いて、今回移動した経路を特定する。続いて、同一経路判定機能は、今回の移動経路と空間的に重複する過去の移動経路を特定する。そのような移動経路が存在すれば、移動経路更新機能が、今回新たに取得した位置情報を用いて、過去の移動経路を精錬する。

以降の節では、GPS位置情報と移動経路情報の表現形式について説明した上で、訪問地・経由地予測サーバの主要な機能である3つの機能:空間的同一経路判定機能、時間的同一経路判定機能、通過確率算出機能について詳細に述べる。その他の機能の詳細については⁸⁾を参照されたい。

3.2 GPS位置情報と移動経路情報の表現

本研究では、GPS位置情報を、位置 p_j と測位時刻 t_j で表す。位置 p_j は緯度・経度を中心とし、1辺が推定測位誤差の2倍であるような正方形で表現する(図3)。推定測位誤差はGPS測位時に算出される水平方向の誤差(HDOP: Horizontal Dilution of Precision)を利用する。

移動経路情報は、移動経路 r_i と当該経路を通過した回数 f_i (以下、通過回数と記載)で

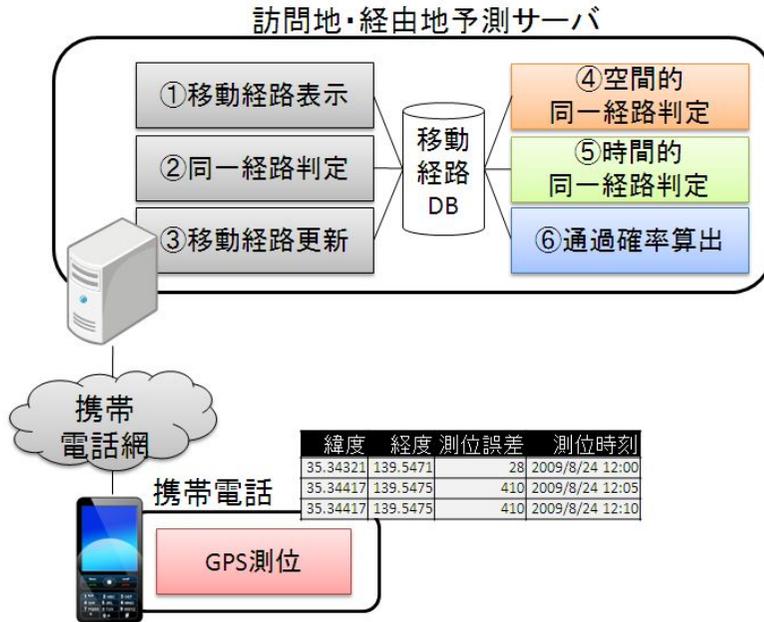


図 1 訪問地・経路地予測システム

表す。移動経路 $r_i (1 \leq i \leq I)$ は、ユーザが移動中に取得した位置 p_j を測位時刻 t_j 順に並べた上で、隣接する 2 つの位置（すなわち 2 つの正方形）を包含する最小矩形（MBR: Minimum Bounding Rectangle）で表す。移動経路 r_i の定義を式 (1) に示す。

$$r_i = \{((^i p_j, ^i t_j), (^i p_{j+1}, ^i t_{j+1}), m_{j+1}) | 1 \leq j \leq I\} \quad (1)$$

ここで、 m_{j+1} は、 $^i p_j$ と $^i p_{j+1}$ を包含する MBR を示す。また、 $\{^i p_j | 1 \leq j \leq J-1\}$ は、移動経路 r_i 上を移動中に取得した位置情報である。

ユーザの滞在・移動状態は、連続する N_{m-s} 個の GPS 位置情報が空間的に重複するか否かで判定する。連続する N_{m-s} 個の GPS 位置情報が空間的に重複する場合に、ユーザは滞在中であると判定し、そうでない場合に、ユーザが移動中と判定する。

GPS 位置情報系列 $\{(^1 p_j, ^1 t_j) | 1 \leq j \leq 7\} (N_{m-s} = 2)$ から推定した移動経路 r_1 を図 4 に示す。システムは、まず、空間的に重複する位置情報 $^1 p_1, ^1 p_2, ^1 p_3$ のうち、推定測位誤差が最も小さい $^1 p_1$ を出発地として特定する。次に、 $(^1 p_4, ^1 t_4)$ を

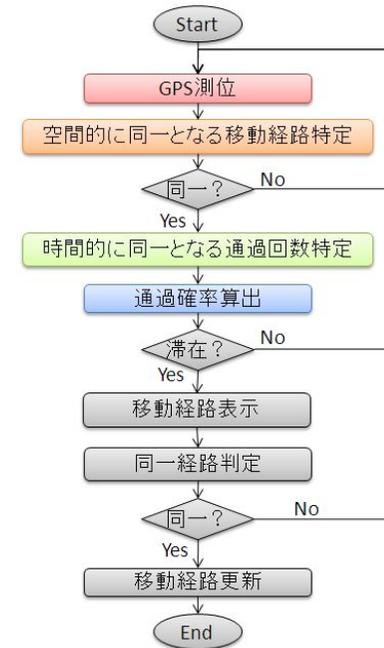


図 2 訪問地・経路地予測フロー

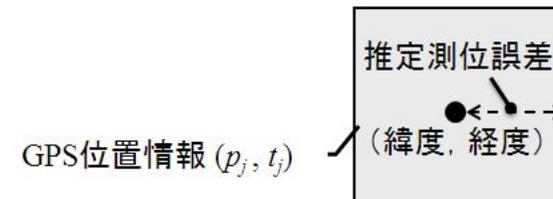


図 3 GPS 位置情報

取得し、システムは $((^1 p_1, ^1 t_1), (^1 p_4, ^1 t_4), m_{14})$ を移動経路 r_1 として推定する。次に、 $(^1 p_5, ^1 t_5), (^1 p_6, ^1 t_6), (^1 p_7, ^1 t_7)$ を取得し、システムは空間的に重複する 3 つの位置情報の中で最も推定測位誤差の小さい $(^1 p_5, ^1 t_5)$ を終着地に設定し、 $((^1 p_1, ^1 t_1), (^1 p_4, ^1 t_4), (^1 p_5, ^1 t_5), m_{14}, m_{45})$ を移動経路 r_1 として推定する。

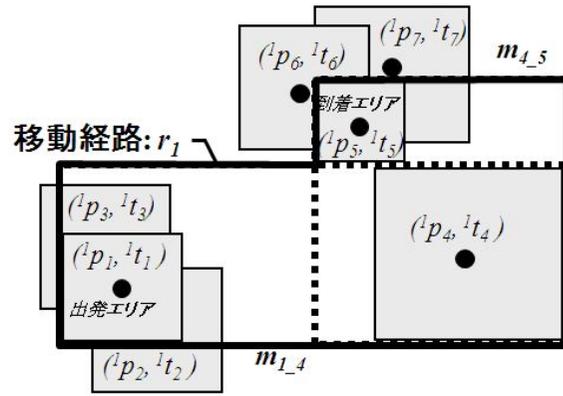


図 4 移動経路表現

3.3 空間的同一経路判定

空間的同一判定機能では、現在地までの移動経路と空間的に同一となる過去の移動経路情報 $(r_i, f_i) (1 \leq i \leq I)$ を特定する。具体的には、移動を開始してから現在までに取得した N 個の位置 $\{p_j | n - N + 1 \leq j \leq n\}$ と、過去の移動経路 r_i とのユークリッド距離の累積値が閾値 TH_d m 以下の場合に空間的同一と判定し、当該移動経路の通過回数 f_i を取得する。

空間的同一経路判定の例を図 5 に示す。(a)において、 p_1 および p_2 は、過去の移動経路 r_1 と空間的に重複しており、累積距離が 0m であることから、現在の移動経路は r_1 と空間的に同一であると判定する。

一方、(b)において、 p_3, p_4 および p_5 と過去の移動経路 r_1 との距離は、それぞれ 0m, 100m, 200m である。そのため、 p_4 までの現在の移動経路は r_1 と空間的に同一と判定されるものの、 p_5 の時点で累積距離は 300m となり、閾値 $TH_d = 200m$ 以上となることから、空間的に同一でないと判定する。

3.4 時間的同一経路判定

時間的同一経路判定機能では、空間的に同一と判定された過去の移動経路情報 $(r_i, f_i) (1 \leq i \leq I)$ について、同じ曜日・時間帯に現在の位置情報 (p_n, t_n) を通過した回数 f'_i を特定する。具体的には、まず r_i について、当該経路を m 回目 $(1 \leq m \leq f_i)$ に通過した際に取得した位置 ${}^{i-m}p_1, \dots, {}^{i-m}p_K$ を抽出する。次に、 p_n と重複する MBR を構成する位置情報の組 $({}^{i-m}p_{k1}, {}^{i-m}t_{k1}), ({}^{i-m}p_{k2}, {}^{i-m}t_{k2})$ を特定する。そして、曜日(平日, 休日)が同一で ${}^{i-m}t_{k1} - T \leq t_n \leq {}^{i-m}t_{k2} + T$ を満たす場合に、時間的に同一であると判定する。ここで

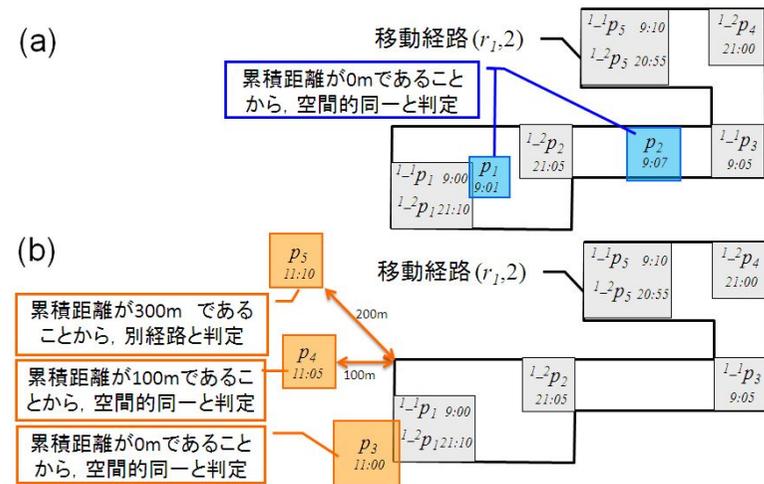


図 5 過去の移動経路 r_1 に対する空間的同一経路判定例 ($TH_d = 200m$, ${}^{i-m}p_k$: 移動経路 r_i を構成する k 番目の位置を表す)

T (分) は通過時刻の多少の揺れ(電車の遅延や、少し早めの出勤など)を吸収するための定数である。

時間的同一経路判定の例を図 6 に示す。 r_1 の通過回数 f_1 は 2 回とし、簡単のため、2 回とも現在と同じ曜日に通過したものとし、 $T = 0$ とする。1 回目の r_1 通過時に取得した位置は ${}^{1-1}p_1, {}^{1-1}p_3, {}^{1-1}p_5$ であり、現在の位置情報 $(p_1, 9:01)$ と重複する MBR を構成する位置情報が $({}^{1-1}p_1, 9:00), ({}^{1-1}p_3, 9:05)$ であることから、 $9:00 \leq 9:01 \leq 9:05$ を満たすため、時間的に同一と判定する。次に、2 回目の r_1 通過時に取得した位置は ${}^{1-2}p_1, {}^{1-2}p_2, {}^{1-2}p_4, {}^{1-2}p_5$ であり、 $(p_1, 9:01)$ と重複する MBR を構成する位置情報が $({}^{1-2}p_2, 21:05), ({}^{1-2}p_1, 21:10)$ であることから、 $21:05 \leq 9:01 \leq 21:10$ を満たさないため、時間的に同一でないと判定する。その結果、時間的に同一となる通過回数 f'_1 は 1 回となる。

3.5 通過確率算出

通過確率算出機能では、空間的・時間的に同一と判定された過去の移動経路に対して、通過確率を算出し、確率最大の終着地、およびそこへ至る経由地を特定する。

過去の移動経路から、ユーザが現在通過している移動経路を特定する上で、時間的に同一であることは重要である。なぜならば、我々の日々の反復行動の多くは、時間帯が決まって

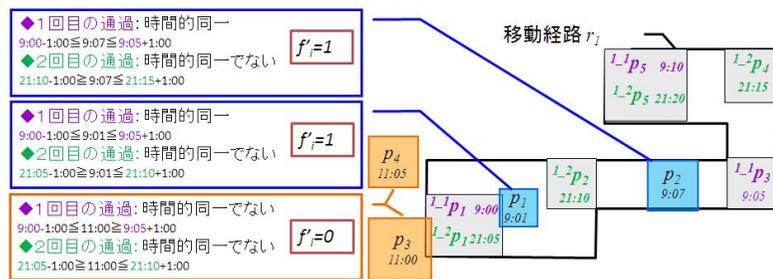


図 6 r_1 (2 回通過) に対する時間的同一経路判定例 ($T=60$)

いるからである。例えば、通勤経路上を自宅から会社へ移動している場合、平日の朝であれば、高い確率で会社へ向かっている。しかし、平日の夜や休日であれば、会社へ向かう可能性は低く、会社方面にあるショッピングセンタやレストラン等へ向かっている可能性が高い。

そこで、本稿では、空間的かつ時間的に同一と判定された移動経路の通過回数に重みを付けて、各経路の通過確率を算出する。具体的には、まず現在地を含めた過去 N 個の位置 $\{p_j | n - N + 1 \leq j \leq n\}$ について、空間的同一経路判定機能において空間的に同一と判定した移動経路 r_i の通過回数 $f_i(j)$ と、時間的同一経路判定機能において時間的に同一と判定した $f'_i(j)$ から、式 (2) により r_i の通過確率を算出する。ここで、 $I(j)$ は空間的同一経路判定機能において p_j と空間的に同一と判定された移動経路数、 α は重み係数を示す。なお、 $I(j) = 0$ となった場合には、再度空間的に同一となる過去の移動経路が出現した時点で、訪問地・経由地予測処理を再開する。

次に、各移動経路の通過確率を終着地毎に集約し、通過確率最大の終着地を訪問地予測結果として特定する。さらに、現在地から当該終着地へ至る移動経路のうち、通過確率最大の移動経路を経由地予測結果として特定する。

$$p(r_i, N) = \frac{\sum_{j=n-N+1}^n \{\alpha f'_i(j) + (1 - \alpha)(f_i(j) - f'_i(j))\}}{\sum_{j=n-N+1}^n [\sum_{i=0}^{I(j)} \{\alpha f'_i(j) + (1 - \alpha)(f_i(j) - f'_i(j))\}]} \quad (2)$$

ただし、 $f_i(j) : p_j$ と空間的に同一と判定した r_i の通過回数、 $f'_i(j) : p_j$ と空間的・時間的に同一と判定した r_i の通過回数、 $I(j) : p_j$ と空間的に同一と判定された移動経路数、 α : 重み係数

4. 実装と性能評価

3 で示した訪問地・経由地予測システムの性能を評価する目的で、システムのプロトタイプ実装と評価実験を実施した。評価実験では、訪問地予測に関する研究で最も多く引用されている Ashbrook ら³⁾ の手法と訪問地予測精度の比較実験を行った。また、訪問地予測精度の時間変化に関する評価実験を行った。

4.1 実装

提案手法の性能を検証するために、我々はプロトタイプシステムを開発した。プロトタイプシステムは、携帯電話として、市販の NTT ドコモの携帯電話を利用した。また、移動経路推定・予測サーバとして、Core2 Duo プロセッサ、12GB のメモリを搭載したサーバを利用した。訪問地・経由地予測サーバの 6 つの機能は、Java で実装した。

なお、携帯電話上の GPS 測位機能を実現するために、NTT ドコモのビジネス Mopera GPS ロケーションサービス¹³⁾ を利用した。Java アプリケーションを携帯電話上に常駐させることで、携帯電話上に GPS 測位機能を実現することも可能だが、この手法には 2 つの問題がある。第 1 に、Java アプリケーションを常駐させることによる電力消費量が多い。第 2 に、Java アプリケーションから GPS 測位を行う場合、GPS モジュールが算出した推定測位誤差は 3 段階に集約されてしまう。

これに対して、ビジネス Mopera GPS ロケーションを利用することで、携帯電話上に GPS 測位機能のためのアプリケーションを常駐させる必要がなく、消費電力を大幅に削減することが可能となる。また、GPS モジュールが算出されたメートル単位の推定測位誤差を取得することが可能となる。

4.2 訪問地予測の性能評価

訪問地予測の性能について、2 つの観点で評価を実施した。第 1 に、訪問地予測に利用する学習データの日数に対する訪問地予測精度を評価した。第 2 に、移動を開始してからの経過時間に対する訪問地予測精度を評価した。

性能評価を行うために、被験者 6 名の位置情報を 5 分間隔で、7ヶ月間 (210 日間) 収集した。そして、それらを 7 つの期間に分割し、そのうち 1 つ (30 日間) をテストデータに、残りの 6 つ (180 日間) を学習データとした。テストデータの選択方法によって 7 パターンのデータセットができるため、それぞれのデータセットについて訪問地予測精度を算出し、7 パターンのデータセットの平均値を最終的な訪問地予測精度とした。訪問地予測精度は式 (3) により算出した。なお、評価に用いた諸元を表 1 に示す。

$$\text{訪問地予測精度} = \frac{\text{実際の訪問地と予測結果が一致した回数}}{\text{訪問地予測回数}} \quad (3)$$

実験結果を図7に示す。今回の実験では、従来手法である滞在地間の遷移モデルによる予測では、訪問地予測精度が平均38.7%であったのに対し、提案方式は68.2%と約30%程度予測精度を向上させることができた。これは、従来手法が滞在地によって次の訪問地が決定してしまうのに対して、提案手法では、ユーザの移動経路に基づいて次の訪問地を決定することにより、移動中に取得した位置情報を用いて訪問地の予測精度を向上できるためである。

学習データ日数については、30日間程度で予測精度が飽和している。これは、30日間程度で普段よく訪れる場所が網羅されていることを示している。

なお、ユーザによって、訪問地予測精度や、精度が飽和する学習データ日数に差が生じている。これは、初めて訪問する場所、すなわち訪問地予測不可能な場所の数がユーザによって異なるためである。例えば、被験者Aは内勤者であり、出張がほとんどないため、初めて訪問する場所もほとんどない。そのため、少量のGPS位置情報履歴でも、訪問地を高精度に予測することができた。一方、ユーザDは長期間の出張や旅行へ行っており、初めて滞在する場所が多かったため、訪問地の精度が低下した。

次に、ユーザが移動を開始してからの経過時間に応じた訪問地予測精度について分析した。具体的には、上述した訪問地予測結果において、普段利用している自宅（鎌倉）から会社（横須賀）への訪問地予測結果と、時々利用している自宅から出張先（東京）への訪問地予測結果を抽出した。なお、鎌倉は横須賀と東京の間に位置している。

自宅から移動後の訪問地予測の時間変化を図8に示す。実験結果より、自宅から出た直後から、会社（横須賀）へ訪れることを高い精度で予測していることが分かる。これは従来の滞在地間の遷移関係に基づく手法も同様に高い精度で予測できる。

一方、出張先（東京）へ向かう際には、移動を開始して2回目の測位までは、会社（横須賀）へ訪れると予測されてしまうため、非常に低い予測精度となっている。これは従来の滞在地間の遷移関係に基づく手法も同様に低い精度となってしまふ。

表1 訪問地予測性能評価実験諸元

空間的同一判定閾値 THd	200m
時間的同一判定時間閾値 T	60分
時間的同一通過回数重み	0.99

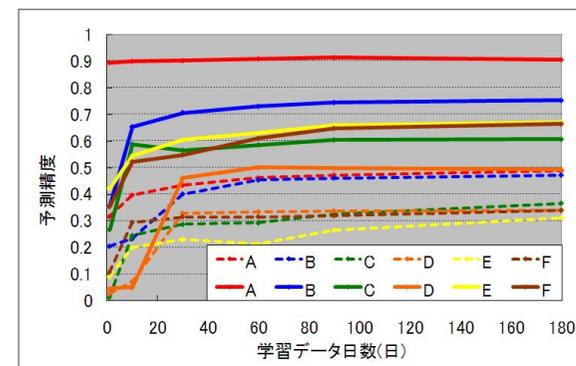


図7 学習データ量に対する訪問地予測精度（実線：提案方式，点線：Ashbrookらの方式）

しかしながら、提案手法では移動を開始して3回目の測位以降は非常に高い精度で訪問地を予測できている。これは、3回目の測位の時点で、被験者は最寄り駅から東京方面へ移動したため、通常の場合（横須賀）への移動経路を逸脱したためである。さらに、この際、自宅から出張先への過去の移動経路を特定できており、出張先へ至る経由地も特定できていた。

このように、普段は通らない移動経路でも、ユーザが移動することで、過去の移動経路を絞り込むことができ、訪問地・経由地の予測精度を向上することができる。

5. おわりに

本稿では、逐次的に精練された移動経路履歴に基づいて、訪問地・経由地を予測するシステムを示した。本システムでは、低頻度で取得したGPS位置情報を用いて逐次的に精練された過去の移動経路のうち、空間的に同一となる移動経路の通過回数と、空間的・時間的に同一となる移動経路の通過回数に基づいて、訪問地・経由地を予測している。

6人の被験者による7ヶ月間の実験により、従来の滞在地間の遷移モデルよりも高精度に訪問地・経由地を予測することができることを確認した。

今後は、訪問地・経由地予測結果に対する信頼性を推定する機構を入れることで、信頼性の高い予測結果に連動した情報を配信する機能について検討を進める予定である。

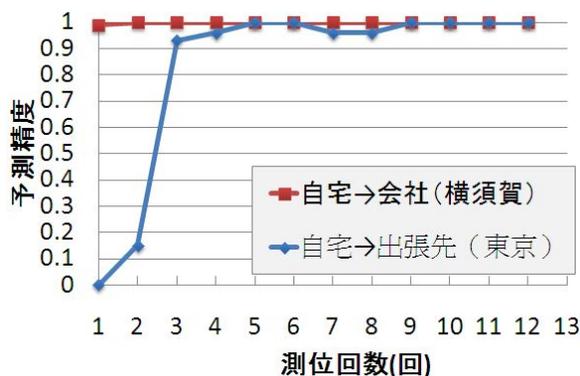


図 8 移動後の測位回数に対する訪問地予測精度の評価

参 考 文 献

- 1) C. Zhou, S. Shekar, and L. Treveen: Discovering Personal Paths from Sparse GPS Traces, JCIS 2005 Workshop on Data Mining, 2005.
- 2) C. Zhou, D. Frankowski, and P. Ludford and S. Shekhar and L. Terveen: Discovering personally meaningful places: An interactive clustering approach, ACM Trans. Inf. Syst. vol.25 no.3, 2007.
- 3) D. Ashbrook and T. Starner: Using gps to learn significant locations and predict movement across multiple users, Personal and Ubiquitous Computing Vol. 7 Issue. 5, pp.275-286, 2003.
- 4) J. H. Kang, W. Welbourne, and B. Stewart and G. Borriello: Extracting places from traces of locations, 2nd ACM International Workshop on Wireless Mobile Applications and Services on WLAN Hotspots(WMASH), pp.110-118, 2004.
- 5) L. Liao, D. Fox, and H. Kautz: Learning and Inferring Transportation Routines, AAAI-04, 2004.
- 6) L. Liao, D. Fox, and H. Kautz: Extracting Places and Activities from GPS Traces Using Hierarchical Conditional Random Fields, Intl. Journal of Robotics Res. 26 - 1, pp.119-134, 2007.
- 7) M. Easter, H.P. Kriegel, J. Sander, and X. Xu: A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise, KDD '96, pp.226-231, 1996.
- 8) N. Yamada, Y. Isoda, M. Minami, and Hiroyuki Morikawa: Incremental Route Refinement for GPS-enabled Cellular Phones, The Fifth International Conference

- on Mobile Computing and Ubiquitous Networking (ICMU 2010), 2010.
- 9) N. Marmasse and C. Schmandt, Location-aware information delivery with com-motion, HUC, pp.157-171, 2000,
- 10) R. Agrawal, and R. Srikant: Mining Sequential Patterns, ICDE'95, pp/3-14, 1995.
- 11) Skyhook Wireless Location Aware App Report: <http://www.locationrevolution.com/stats/skyhookfebfereport.pdf>
- 12) オート GPS: http://www.nttdocomo.co.jp/service/information/auto_gps/index.html.
- 13) ビジネス Mopera GPS ロケーション: <http://www.docomo.biz/html/service/gpslocation/>.
- 14) 総務省 事業用電気通信設備規則の一部改正について: http://www.soumu.go.jp/main_sosiki/joho_tsusin/policyreports/joho_tsusin/denki_bukai/pdf/051025_3.s01.pdf
- 15) 西野正彬, 山田智広, 瀬古俊一, 茂木学, 武藤伸洋, 阿部匡伸: 時刻情報を含む特徴滞在パターンを用いた行動予測方式の検討, 2009 年電子情報通信学会総合大会, 2009 .