

長期に渡る携帯ナビ移動履歴を用いたユーザの生活パターン推定手法の検討

柴崎真理子・藤田秀之・木實新一・有川正俊

Estimation of Life Patterns using Long-Term Trajectory History

Mariko SHIBASAKI, Hideyuki FUJITA, Shin'ichi KONOMI

and Masatoshi ARIKAWA

Abstract: In recent years, researches to estimate the context of user's activity using trajectory history of user's mobile phone and purchase history are increasing to develop better information services and commercial activities typically marketing. Though there are many researches analyzing trajectory history to estimate user's activity context, analysis of long-term trajectory of a user remains relatively unexplored. This research examines how to estimate and extract user's life pattern for estimating user's context using trajectory history for the long-term.

Keywords: 意味情報付きの移動履歴 (semantic trajectory), 全地球測位システム (GPS), 生活パターン (life pattern)

1. はじめに

GPS 機能付き携帯端末の普及により, 人やモノの位置情報が記録されるようになった. 近年では, 長期間に渡る位置情報の蓄積が可能となり, ユーザの活動パターンをより深く分析できる可能性が出て来た. それに伴い, ポイントの羅列である大量の移動履歴からユーザの行動文脈を抽出する手法について研究が行われて来た.

移動履歴に意味情報を付与し, ユーザの行動文脈を抽出する方法として, Guc et al. (2008) はユーザからのインタビュー調査から得られた移動目的と移動手段を移動履歴にラベリングして分析する枠組みを提案したが, 大量の移動履歴の分析ではインタビュー調査は不向きである.

Spaccapietra et al. (2008) は, 移動履歴の各計測点

を停留時間に応じて, 停留点と移動点に分類し, 停留点の地図上の属性情報 (レストラン, 公園など) を移動履歴に付与することでユーザの行動文脈を抽出する手法を提案している. Ying et al.

(2010) は, その連続的なパターンからユーザの類似性を算出する手法を提案した.

しかしながら, 移動履歴から長期に渡ってユーザの生活パターンを分析し, その視点から行動文脈推定を試みた研究は少ない. 生活に関する調査では, NHK国民生活時間調査報告書(NHK放送文化研究所, 2010) がある. しかし, この調査では, 行動を「必需行動」, 「拘束行動」, 「自由行動」の3つに分類し, 日常的に繰り返される行動について調査しているため, 「帰省」, 「旅行」や「出張」などの突発的な生活イベントへの調査はなされていない.

本研究では, 移動履歴から長期渡ってユーザの生活パターンを分析し, その視点からの行動文脈の推定を最終目標としている. 本稿では, 携帯端末の1年間に渡る移動履歴を用いて月日, 時間帯,

柴崎真理子 〒277-8568 千葉県柏市柏の葉 5-1-5

東京大学新領域創成科学研究科社会文化環境学専攻

E-mail: ms@csis.u-tokyo.ac.jp

位置の3つの枠組みから、集団と個人それぞれに着目して可視化し、長期に渡る移動履歴から生活パターンを抽出するために、ユーザの行動パターンの有無および長期的な移動履歴からどのような生活パターンを見い出すことができるかを検討した。

2. 分析手法

2.1 位置データ取得方法

本研究で用いられた移動履歴は、携帯端末に付属している GPS レシーバから取得されている。この GPS レシーバでは 5 分ごとに位置情報を自動計測している。また、同付属の加速センサーを用いて携帯端末の動きを加速度から検知し、動き始めると自動的に計測を始める。本論文では、位置データが取得されている間は、ユーザは物理的な移動や運動を伴う活動をしていると推測した。したがって、計測回数はユーザの活動の有無およびその激しさを表すと考える。次に、位置データが取得されたユーザの属性について述べる。対象ユーザは、携帯端末のナビゲーションアプリを使用しているユーザに限定され、それ以外の個人属性は全くない。

本研究では、東京都を通過したユーザのうち 100 人をランダムに算出し、2010 年 8 月 1 日から 2011 年 7 月 31 日の 1 年間の移動履歴を対象に分析を行う。

2.2 分析の枠組み

収集された移動履歴は、携帯端末の GPS 機能によるものであるため、常には位置情報を計測しておらず、位置情報が収集されていない多くの期間と時間帯が点在している。したがって、断片的な位置情報のみから、生活パターンがどの程度抽出されるかという問題点がある。

2.1 節で述べた GPS レシーバの稼働の特性から、本論文では位置情報が取得されている間は、ユーザは物理的な移動や運動を伴う活動していると推測した。したがって、計測回数はユーザの活動の有無およびその激しさを表すと考える。

したがって、1 年間分の移動履歴の計測回数分布を日付単位と時間単位にそれぞれ可視化し、長期的な移動履歴から、ユーザの活動の有無で表現される生活パターンが抽出できるか検討した。同時に、取得された位置情報が休日か平日かに影響を受けているか時系列的な可視化で明らかにし、日常的に繰り返されていない行動パターンが抽出できるか検討した。

2.3 月単位での計測データ分布分析

まず 1 日あたり位置情報が何回計測されているか算出し、1 ヶ月ごとのその平均値をとった。各ユーザごとに、月単位の 1 日あたりの平均計測回数の中央値をとり、1 日の行動を何回計測しているか算出した。次に、1 ヶ月ごとに何日間位置情報が計測されているか算出した。

2.4 時間単位での計測データ分布分析

時間帯によって活動の有無が分かれるか検討するために、1 時間あたりの計測回数を算出し、その分散値を算出した。

2.5 休日平日単位の移動履歴の可視化

各ユーザの移動履歴を、休日平日ごとの計測地点を時系列可視化を行うシステムを実装した。

3. 結果

3.1 月単位での計測データ分布分析

図-1 は、1 ヶ月ごとの 1 日あたりの平均計測回数の変動を表す。

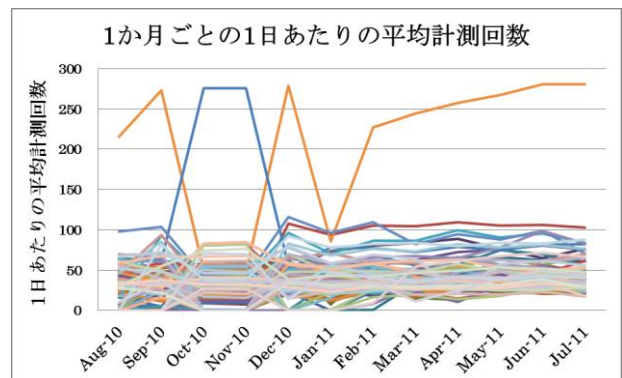


図-1 1 ヶ月ごとの 1 日あたりの平均計測回数

1 ヶ月平均計測回数の中央値は 40 回から 50 回に

集中的に分布しており、その平均値は 40.6 回であった。つまり、1 日に約 40.6 回の GPS 測位があると考えられる。

次に、1 ヶ月ごとに何日計測があるか算出した。図- 1 から明らかなように、ユーザによっては計測がない月があるため、中央値を用いた。1 ヶ月ごとの測定のある日数の中央値の平均値を算出した結果、1 ヶ月に平均 28.4 日間、GPS 測位が行われていることがわかる。

3.2 時間単位での計測データ分布分析

まず、時間帯ごとに1年間に渡る計測データがどのように分布しているか可視化した。

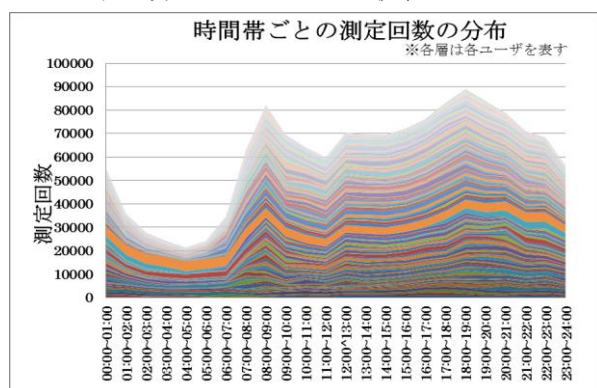


図- 2 時間帯ごとの測定回数の分布

図- 2 より、午前3時から4時は活動が少ないことから、ユーザが就寝している可能性を示している。午前8時から9時と18時から19時が活動のピークがきて、深夜になるにつれ減少していることがわかる。次に、各ユーザごとの時間帯ごとの計測回数の変動を図- 3 にて可視化した。

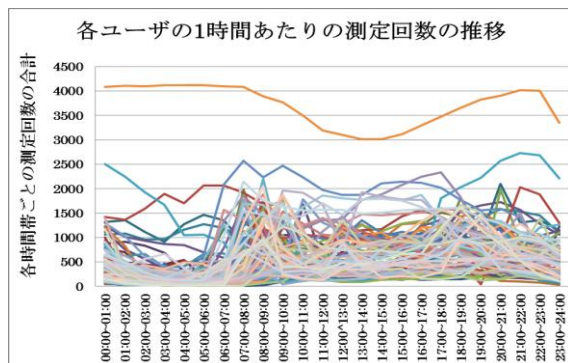


図- 3 各ユーザの1時間あたりの測定回数の推移

次に、各ユーザの1時間ごとの計測回数値を母集団とした分散を算出する。

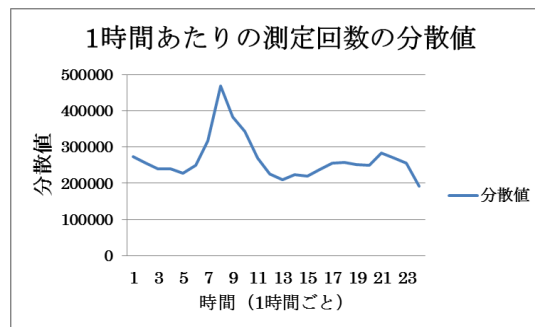


図- 4 1時間あたりの計測回数の分散値

1 時間あたりの各ユーザごとの計測回数値の分散値が大きいということは、活動を行い GPS 測位が行われたユーザやそうではないユーザが混在していたため、計測回数値にばらつきがでたと考えられる。分散値は午前8時付近に大きくなっているため、ユーザによって活動を始める時間帯が異なっていることが分かる。

3.3 休日平日単位の移動履歴の可視化の結果

各ユーザの休日と平日ごとの移動履歴を可視化した。

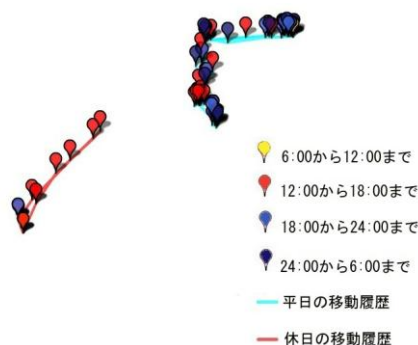


図- 5 ユーザの一週間の移動履歴

この可視化より、平日はある一定の行動圏域内で生活しているが、休日では遠出をするなど突発的な行動の変化あることがわかる。また、平日の移動履歴を細かく分析する。

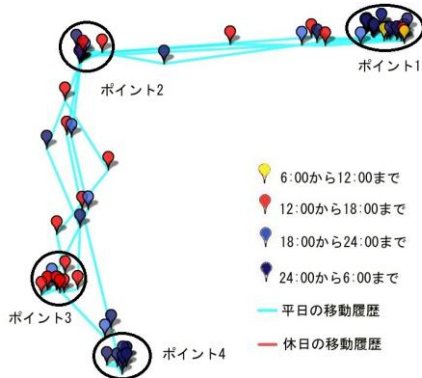


図-6 平日の移動履歴

平日の移動履歴を4つのポイントに分けて特徴付けることができる。午前6時から12時までの計測点があるポイント1と12時から18時までの計測点が集中しているポイント3、24時から午前6時までの計測点がポイント4にある。計測点の時間と計測回数および滞留時間から、その計測点が移動する途中に計測されたものなのか、滞留している間に計測されたか推測し、その地点の位置属性（公園、歓楽街など）を合わせることで、各ユーザの生活パターンを推測することができると考えられる。

3.4 考察

長期的な移動履歴には、計測値が無い月および時間帯があるが、その積み重ねにより、活動時間帯や停留点とその停留時間および公園や学校などの位置属性から、休日の娯楽活動から出張や旅行などの突発的なイベントを含む生活パターンを推察できることがわかる。計測地点の位置属性を組み合わせることで、ユーザの職業などの個人属性を含む生活パターンの分析および推定が可能となることがわかる。

そのためにも、計測点にどのような位置属性をラベル付けするのが重要となる。

4. おわりに

周囲の人々の大まかな生活パターンを把握など有用な知見を抽出することで、地域の防災計画策定に役立てると考える。今後は長期に渡る生活

パターンの抽出のためのラベルのスキームについて検討する。

謝辞

本研究は、株式会社ゼンリンデータコムが提供する混雑統計を利用いたしました。ここに記して、謝意を表明いたします。

参考文献

- NHK 放送文化研究所 (2010):「日本人の生活・2010—NHK 国民生活時間調査」, NHK 出版.
- Josh, Jia-Ching. Ying., Eric, Hsueh-Chan. Lu., Wang-Chien, Lee., Tz-Chiao, Weng. and Vincent, S. Tseng., 2010. *Mining user similarity from semantic trajectories*. *ACM LBSN' 10*.
- Baris, Guc., Michael, May., Yucel, Saygin. and Christine, Korner., 2008. *Semantic annotation of GPS trajectories*, *11th AGILE International Conference on Geographical Information Science 2008*.
- Spaccapietra, S., Parent, C., Damiani, M. L., De Macedo, J. A., Proto, F. and Vangenot, C., 2008. *A conceptual view on trajectories* *Data and Knowledge Engineering*, 65, 126-46.