

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 8 日現在

機関番号：13901

研究種目：挑戦的萌芽研究

研究期間：2013～2014

課題番号：25630215

研究課題名(和文) スマートフォン行動データとコンテキストデータを活用した活動・交通ログ自動生成手法

研究課題名(英文) Study on estimation methods for travel data using GPS and context data

研究代表者

森川 高行 (Morikawa, Takayuki)

名古屋大学・環境学研究科・教授

研究者番号：30166392

交付決定額(研究期間全体)：(直接経費) 3,100,000円

研究成果の概要(和文)：従来行われているパーソントリップ調査では、莫大な費用がかかることや回答率の低下による調査の継続性や精度低下が問題となっている。そこで本研究では、GPSデータから交通データを生成する方法に関する既往文献調査を行い、それを踏まえて、トリップごとに分かれていないGPSデータをトリップごとに分断する方法や、行動目的を推定する手法の検討を行った。その結果、RandomForestモデルにより、行動目的の判別精度は7割を超え、さらに、ある程度の精度を確保するためには必要な調査実施期間は1か月程度であること、被験者数の増加に伴って判別精度は向上することが示された。

研究成果の概要(英文)：Traditional Personal Trip (PT) survey has problems about low accuracy and disability of continuity because of significant high cost and low respondent rate. In order to avoid these disadvantages, GPS data collected by smart phone were utilized in this study to infer trip purpose data. Based on the method of Decision Tree and Random Forest, model for inferring trip purpose was developed and some analyses proceeded subsequently. It is found that the accuracy of trip purpose inference calculated by the model developed in this study is more than 70%. Furthermore, it can be concluded that 1 month continuous GPS data is minimum requirement to get the satisfied results. Also, the results show that the accuracy of trip purpose reference can be improved by increasing the number of respondents.

研究分野：交通計画

キーワード：スマートフォン 位置情報 活動推定 交通調査

1. 研究開始当初の背景

これまで交通行動分析に伝統的に用いられているパーソントリップ調査は 10 年に 1 度大都市圏にて実施されているものであるが、近年アンケート調査にかかる莫大な費用や回答率の低下により、調査の継続やデータの精度等が脅かされている。一方、エネルギー消費、災害時救助、移動サポートの高度化などの社会的要請から、トリップ単位ではなく、人の活動・移動を一体的に調査しモデル化する必要性は高まっている。この相矛盾する要請に答えられる可能性を期待されているのが高度化するモバイル端末である。活動内容や交通手段の把握が重要である人間の行動調査にモバイル端末を用いる試みは、ここ 10 年ほどで取り組まれており(例えば、羽藤ら(2000)) 様々な調査で活用されてきている。GPS 付携帯電話を活用したこのシステムはアンケート調査と比較して、被験者の負荷が少なく長期間の調査も可能であるなど多くのメリットがあるものの、研究代表者らが実際に社会実験等で使用した観点から以下の問題点がある。移動の概念の説明や携帯電話の操作方法、PC での入力方法を説明するための被験者は説明会へ参加することが必須であり、また、調査が長期にわたる場合は、操作が面倒になってしまいいい加減な入力となることもある。そこで、被験者の操作をなくし、鞆などに入れておくだけで自動収集できる Bcals という端末も開発されており(Hato (2010)) 主に自動車利用時の行動観測に使用されているが、移動手段や目的まで完全自動で精度高く把握することはできていない。

2. 研究の目的

GPS データから移動目的や移動手段を判別する手法を開発する。特に移動目的の判別手法の検討、及び、移動目的を判別するために必要なデータ収集期間や被験者数について知見を得ることを目的とする。

3. 研究の方法

(1)使用データ

本研究で使用したデータの概要を表-1 にまとめた。被験者数の合計は 156 名であるが、とに重複した 19 名の被験者がいるため、全被験者数は 137 名である。また、「その他」トリップの全体に占める割合は約 9.5%であり、全体の 1 割に満たないこと、また「その他」トリップを分類することは困難であると考えられるため、「その他」目的以外の 5 つの行動目的を推定するものとし、モデル構築データに「その他」トリップを含めない。ただし、実際の調査において予測するデータから「その他」トリップは除けないため、予測データには含むものとする。よって判別精度の最大値は約 90%となる。

表-1 使用データ概要

項目	内容
調査期間	2010/11/22~12/19
	2008/9/24~10/30
	2008/11/19~12/31
	2011/12/5~2012/2/7
調査方法	プローブパーソン(PP)調査
調査地域	中京都市圏
被験者数	50名, 30名, 50名, 26名
トリップ数	3846, 2568, 4678, 3038
GPS取得間隔	数分から数秒まで様々
行動目的	出勤, 帰宅, 帰社 業務, 自由目的, その他

(2) 自宅と勤務先の位置推定

本研究では出勤・帰宅の判別精度を上げるため、自宅と勤務先の位置を推定する。取得した GPS 情報から各トリップの出発地(O メッシュ)、目的地(D メッシュ)を取得し、1 日の最終トリップの D メッシュのうち、調査期間中の最頻出メッシュを被験者ごとに集計する。また 1 日の最初トリップの D メッシュのうち、調査期間中の最頻出メッシュを被験者ごとに集計する。これらをそれぞれその被験者の自宅、勤務先の存在するメッシュと定義する。

(3) 行動目的推定に関する研究

自宅と勤務先の位置推定を行った後。それらの情報や GPS 情報から行動目的の判別に有効であると思われる説明変数を作成する。本研究では 33 の説明変数を用いた。そしてクロスバリデーション(交差確認・検証法)を用いて。決定木モデルと RF (Random Forest) により判別精度。適切な実施期間。被験者数を検証する。

4. 研究成果

(1)自宅と勤務先の位置推定に関する成果

GPS 取得期間を 1 週間から 5 週間まで増加させたときの自宅、勤務先の位置推定精度の変化を示す。自宅の推定精度は 3 週間から 80%を超える精度となり 4 週間目からは 82.4%の精度で安定し、勤務先の推定精度は 3 週間までは単調に増加し、3 週間目から 56.3%の精度で安定している。よって自宅の位置推定に 4 週間程度の GPS 取得期間があれば約 80%の精度、勤務先(登校先)の位置推定にも 3 週間程度の GPS 取得期間があれば 56.3%の精度となることがいえる。

(2) 行動目的推定に関する成果

決定木モデルと RF それぞれにクロスバリ

デーションを行った結果を図-1に示す。決定木モデルの精度の平均が63.0%、RFの精度の平均が68.3%と、両手法とも60%を超える精度となり、RFの精度が決定木モデルの精度を上回る結果となった。

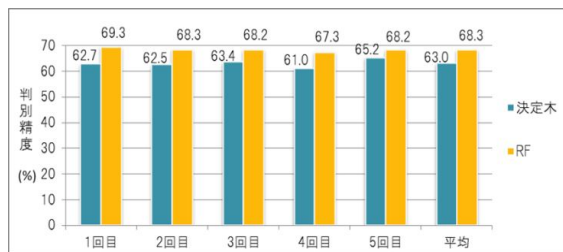


図-1 判別精度

次に、RFにより行動目的別の判別精度の検証を行った。図-2に、クロスバリデーションを行った結果を示す。「帰宅」トリップは91.4%と、非常に高い精度となった。これは自宅の判別の精度が高いことによるものと考えられる。一方、「出勤」トリップは「帰宅」トリップに次いで高い精度ではあるが、76.3%となった。この結果は勤務先の判別精度があまり上がらなかったことが原因であると考えられる。また、「業務」トリップは約60%の精度、「自由目的」トリップは約70%の精度となった。一方、「帰社」トリップに関しては、他4つの行動目的に対して比較的精度が上がらなかった。これは本研究で用いた「帰社」トリップの総数が少ないことがまず考えられる。また「帰社」トリップは全被験者に取得されているトリップではないため、クロスバリデーションを行うためのデータ分割時に「帰社」トリップの少ないグループができてしまい、学習データ数が少なくなることで、そのグループの精度が上がらずに平均精度も下がってしまうと考えられる。

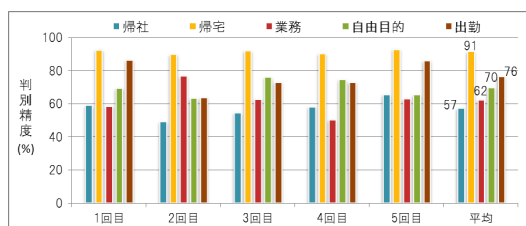


図-2 RFによる行動目的別判別精度

(3) 決定木モデルによる判別過程の可視化

RFを用いることで、高い精度で判別が可能になるが、実際にどのような過程で判別が行われているかは不透明である。そこで、決定木モデルを用いて判別過程を可視化する。ただし、RFと決定木モデルによる分類器は異なるため、大まかな判別過程を可視化するために行うこととする。

以下の図-3にクロスバリデーションの分割のうち、グループ1を例として決定木モデルを示す。なお、判別率は62.7%(1942/3096)である。この分析では5つの変数のみを用いており、他の分析結果からも5~8変数あれば

判別率が6割を超えることが分かった。

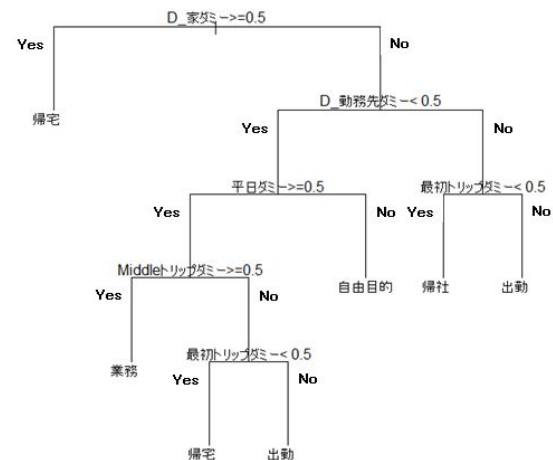


図-3 決定木モデル例(グループ1)

(4) データ収集期間の検討

実際の調査施行時においてどれだけの期間調査を行えばどの程度の判別精度が得られるか検証し、行動目的の判別に適切な調査実施期間を検討する。

今回使用するデータのうち最も長いデータは13週間であり、その被験者数は19名である。学習データとしてこの13週間分のデータを用い、1週間から13週間分までデータ量を増加させ、また予測データをその他118名の取得開始から4週間目までのデータに統一した。最も取得期間の短いデータは4週間であるため、予測データを4週間分とした。そして週を増加させる順序をランダムに入れ替え、この操作を10回繰り返し、その平均を検証精度とした。

図-4に結果を示す。2週間を超えるあたりまで決定木の精度がRFの精度を下回り、実施期間が4~9週間程度までは、決定木とRFの精度にほとんど差がなくなることがわかる。

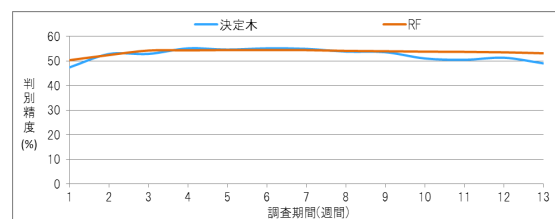


図-4 GPS取得期間による判別精度の推移

(5) 被験者数の検討

被験者数を増加させたときの判別精度の変化を検証する。全被験者137名の取得開始から1週間のデータを学習データとし、10名毎に137名まで増加させ、また予測データを全被験者137名の2週間目以降の3週間分のデータに統一した。GPSデータの長期間の取得による被験者への負担が調査の問題点として挙げられていることを考慮し、できる限り被験者の負担を減らすため、調査期間を

1週間と仮定し、学習データは1週間分とした。調査実施期間の検討と同様に、被験者を増加させる順序をランダムに入れ替え、この操作を5回繰り返し、その平均を検証精度とした。実際の調査施行時においてどれだけの期間調査を行えばどの程度の判別精度が得られるか検証し、行動目的の判別に適切な調査実施期間を検討する。

図-5を見ると、決定木、RFともに被験者数が2~30名を超えたあたりで精度の上昇は緩やかになり、その後、被験者数の増加に伴って精度は向上し続けていることがわかる。決定木モデルでは、被験者数が90名程度で精度は60%を超え、RFでは20名程度で60%、120名程度で70%を超えることがわかる。今回の検証における学習データは全被験者137名の1週間分のデータであることから、学習データのGPS取得期間が1週間程度と短い場合でも、被験者数を増加させることで精度は70%以上を期待できる。本研究ではデータの制約から137名までしか分析できなかったが、さらに被験者数を増加させることで判別精度は向上するものと考えられる。

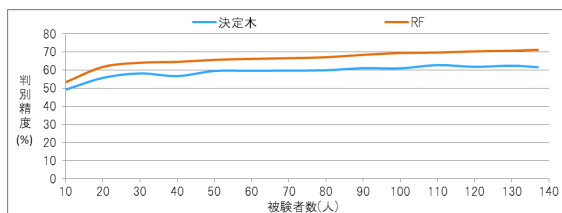


図-5 被験者数による判別精度の推移

(6)まとめと課題

RFを用いて、GPS情報から出勤、帰宅、帰社、業務、自由目的の5行動目的を約75%判別可能であることが分かった。自宅の位置推定精度が80%を超えたことで、帰宅目的のトリップは90%を超える精度で判別可能であったが、勤務先の位置推定精度が高いとはいえないため、出勤、帰社、業務目的のトリップを高い精度で判別することはできなかった。

また、実際の調査施行時においては、約4週間程度の調査実施期間が最も適切であることが分かった。さらに、被験者数の増加に伴って判別精度は向上すること分かった。

今回の結果から、実際に調査時には1週間程度、200名程度のデータを取得してマスターデータを作成し、予測する被験者はGPS情報のみを用いて予測を行う形式が最適であると考えられる。

また、本研究の手法を用いたアプリ等を開発することで、従来のPT調査における若者の回答率低下問題に対し、精度の補完を行うことが可能であると考えられる。このように従来のPT調査と組み合わせ、補完し合うことでより精度の高い調査が行うことが可能であるといえよう。

また、出勤、帰社、業務の判別精度向上の

ためには勤務先の位置推定精度の向上が必要である。

さらに、本研究では被験者数の増加に伴って判別精度は向上したが、どの程度の被験者数で判別精度は最大となるかを検討する必要である。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計1件)

- Gong, L., Morikawa, T., Yamamoto, T., SATO, H.: Derive Personal Trip Data from GPS Data: A Literature Review on the Existing Methodologies, The 9th International Conference on Traffic and Transportation Studies, 2014.【査読あり】

[学会発表](計3件)

- Gong, L., Morikawa, T., Yamamoto, T., SATO, H.: Derive Personal Trip Data from GPS Data: A Literature Review on the Existing Methodologies, Procedia - Social and Behavioral Sciences, Vol. 138, pp. 557-565, 2014年8月, 中国・紹興市.
- Gong, L., Sato, H., Morikawa, T. and Miwa, T.: Stop Point Identification Using Constrained DBSCAN(Density-Based Spatial Clustering of Applications with Noise) Algorithm, 土木計画学研究・講演集, vol. 50. (CD-ROM), 2014年11月, 鳥取大学.
- 塚本健太郎, 佐藤仁美, 森川高行: GPSやコンテキストデータを活用した行動目的の推定手法に関する研究, 土木計画学研究・講演集, vol. 50. (CD-ROM), 2014年11月, 鳥取大学.

[図書](計0件)

[産業財産権]

出願状況(計0件)

取得状況(計0件)

[その他]

特になし

6. 研究組織

(1)研究代表者

森川 高行(MORIKAWA TKAYUKI)
名古屋大学・大学院環境学研究所・教授
研究者番号: 30166392

(2)研究分担者

山本 俊行(YAMAMOTO TOSHIYUKI)
名古屋大学・エコトピア科学研究所・教授
研究者番号: 80273465

三輪 富生 (MIWA TOMIO)
名古屋大学・エコトピア科学研究所・准教授
研究者番号：60422763

薄井 智貴 (USUI TOMOTAKA)
名古屋大学・経済学研究科・特任准教授
研究者番号：20549448
(平成25年3月まで)

(3) 連携研究者

佐藤 仁美 (SATO HITOMI)
名古屋大学・未来社会創造機構・特任講師
研究者番号：00509193