

人々の行動軌跡と空間イメージの把握におけるビッグデータ利用の可能性

相 尚寿

Tracking travel trajectory or compassing image on places by big data: a review

AI Hisatoshi

Abstract: This paper overviews several studies focusing on the use of big data to track people's travel trajectory or to compass people's image on places. GPS logs collected from mobile devices such as smart phones and cell phones are one of useful data source to track people's movement. Several methods have been proposed (1) to find dwelling places or trip ends, (2) to presume travel routes, (3) to guess transportation modes. Mobile phone BTS connection logs, keywords used on transfer guidance websites, Tweets on Twitter are also used as data source in some studies. Blog articles are used to promote estimation model of apartment rents by evaluation of places by people. Harsh braking locations tracked from number of vehicles are used to find unsafe road intersection or curves and to evaluate the impact of measures for safety.

Keywords: ビッグデータ (Big data)、GPS ログ (GPS log)、マイクロブログ (Micro blog)、行動軌跡 (travel trajectory)、空間イメージ (image on places)

1. はじめに

高速通信技術の発達と通信エリアの拡大、GPS 機能を搭載したスマートフォンやタブレットなどの携帯型情報端末の普及に伴い、近年は携帯型端末の位置情報や通信機能を用いた様々なウェブサイトおよびアプリケーションが提供されている。例えば携帯型端末を用いた歩行者ナビゲーションや近隣施設検索サービス、携帯型端末からも更新・投稿が可能なブログや掲示板などが挙げられる。これらの機器およびサービスの普及に伴い、端末の位置情報やブログ投稿など空間に関するテキスト情報が日々大量に蓄積され、ビッグデータとして利活用が注目されている。本稿では、特に人々の行動軌跡および地域や空間に対するイメージおよび認識の把握の分野における、これらビッグデータの利活用事例を概観する。データは主として携帯型端末の位置情報と空間に関するテキスト情報を取り上げる。

2. ビッグデータ利用の利点と欠点

初めに、ビッグデータ利用の利点と欠点を概観したい。ビッグデータ活用の利点は、調査対象者への負担軽減、情報収集コストの低減、リアルタイムな情報収集などが挙げられる。従来の行動調査はアクティビティダイヤリへの記入や専用 GPS 端末の携帯など、調査対象者の協力が必要であったものの、日常的に携帯する端末の位置情報を収集する方法であれば対象者の負担は大きく軽減される。

地域イメージの把握も従来はアンケート調査やインタビューが中心で、対象者への負担、調査期間の限定などの課題が存在した。ブログ投稿などの情報を利用することで、これらの課題は大きく改善される。また、アンケート調査とは異なり回答者に調査であることが明示的に伝わらないため、例えばアンケート実施者との対面による否定的評価の回答の躊躇といったバイアスの除去効果もあると考えられる。

一方、これらのビッグデータを利用するには、調査対象者が携帯型端末を所持した当該サービスの利用者あるいはブログなどを開設している情報

発信者に限定されていることに注意する必要があるなどデータの信頼性や再現性に課題も残る。

3. 行動軌跡把握の事例

本節では、GPS ログを中心としたビッグデータを活用した人々の行動軌跡把握の事例を紹介する。これまでに実施されている人々の行動軌跡把握を意図した調査としては、パーソントリップ調査や大都市交通センサスが代表的である。調査対象者に質問票への記入を依頼し、移動の出発地や目的地、交通手段や移動経路などが調査されている。調査期日が限定され、調査間隔も前者が首都圏では約 10 年、後者が 5 年となっている。これに対して、目黒(2008)が携帯電話は行動分析調査用端末が備えるべき各種要件を満たすと指摘したように、携帯電話の GPS データを利用した行動履歴把握が注目されている。携帯電話を利用すれば、調査票や専用端末の準備が不要となりコスト低減が図れるほか、行動履歴を対象者が記録する際に発生する短小トリップの忘失や省略、時刻の丸め込み、誤入力などを防ぐことにもなる。従来の方法は一定の準備期間を設け、調査期間を限定して行っていたものの、ビッグデータであれば常に新しいデータが蓄積されるため、最新情報の利用や必要な時期のみを抽出した利用が可能である。

しかし、取得される GPS データ単独では複数の点の羅列でしかないため、ここから行動軌跡を抽出し、その特性を把握する必要がある。その手順として代表的なものが、(1)出発地と目的地すなわちトリップエンドの抽出、(2)出発地から目的地までの移動すなわちトリップの抽出、(3)トリップ内における移動経路や交通手段の推定である(図 1)。

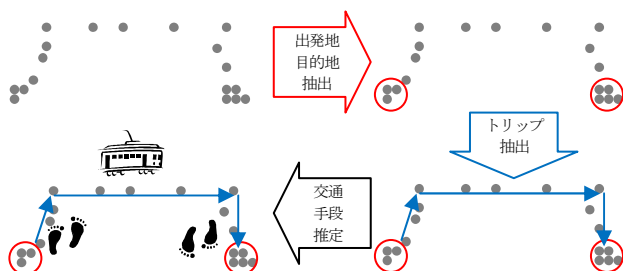


図 1 行動把握における GPS データ処理フローの例

3.1 トリップエンドの抽出

トリップエンドの抽出に関連しては、大野ら(2012)は、一定の時間にわたり一定の距離内に集中している GPS ログを抽出し、複数日に渡るデータにおけるトリップエンドの空間座標にクラスタリング手法を適用してステイエリアを抽出した。堀口ら(2006)は、上記と類似した手法で抽出した滞在地点について、滞在時間帯をもとに居住地や勤務地を判別し、トリップ目的「通勤」「帰宅」の付与を試みた。羽田野ら(2012)は、自宅、職場のほか GPS ログと商業集積地データとを重ね合わせ、商業集積地での滞在を判別し、勤務時間帯や通勤途中での商業地への立ち寄りを含めた行動パターンの分類を行った。

これらの研究では、携帯型端末の電池残量を維持し長時間の記録を得るため、測位間隔が 30 秒や 5 分など長めに設定された事例が多いことから、1 秒間隔の GPS ログを 5 分間隔に加工して、測位間隔を延ばすことによる空間精度の低下を検証した山口(2013)もデータ信頼性の関連研究として挙げられる。

なお、移動軌跡の把握では、GPS データを用いたもののほか、金杉ら(2012)などのように携帯電話基地局を利用したものがある。基地局を用いる方法は GPS 測位と比較して空間分解能が粗くなるため、事前に同一対象者の GPS による移動履歴を取得して移動経路候補を作成しておき、基地局との通信履歴から GPS 履歴とは異なる日の移動経路推定を試みた。長距離移動における推定の有効性、事前の移動経路候補にない特殊な経路は推定できない課題などが議論された。また、見生ら(2012)は乗換案内サイトの検索履歴から、例えば駅前バス停が朝は出発地、夕方には目的地となり、高校や業務地域付近のバス停が朝は目的地、夕方が出発地となる傾向や商業施設に近いバス停は休日に検索が増える傾向などを示した。他には、Twitter 投稿にジオタグとして付加された位置情報を用いた酒巻ら(2011)では、大学および自宅周辺の投稿内容には各々大学と自宅に関連する単語の出現頻度が高いことを示した。また、日常的に訪問する場所の抽出を試み、調査対象者へのアン

ケートでは忘失されていた地点の抽出にも成功した。

3.2 トリップの抽出と移動経路や交通手段の推定

前節で紹介したようにトリップエンドが抽出されれば、時系列的にそれらの間に存在する GPS ログはトリップ中であると解釈できる。これらから移動経路や交通手段を推定する研究も広く行われている。

前出の大野ら(2012)は同一 OD 間の複数日の GPS ログを可視化することで移動経路を視覚的に明らかにした。山田ら(2009)は、GPS ログの測位誤差を考慮してログを正方形で表現し、2 つのログ間の経路を 2 つの正方形を内包する最小矩形で表現した。同一経路に関する GPS ログが繰り返し蓄積されることにより、同一地点のログが高精度のログに置き換えられたり、従来のログ間に新たなログが挿入されて補間されたりすることで、経路推定の精度向上が可能である(図 2)。また、山田ら(2010)では、上記の経路推定を応用して、ある地点を同一曜日の同一時間帯に通過した過去の経路履歴から、現在の移動の目的地を予測する提案を行った。

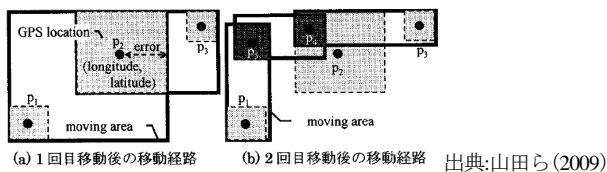


図 2 山田ら(2009)での経路の矩形表現と精度向上

小林ら(2009)は GPS ログのほか、携帯電話に装備された加速度センサーとマイクの情報も活用し、利用者が停止、走行、歩行、自転車、自動車、バス、電車のいずれの状態であるか判別を行った。堀口ら(2006)は、前後の GPS ログから速度および方位角速度を算出し、それらの分布や標準偏差の分布などから移動手段ごとの特徴量を見出し、移動手段判定ルールの設定を試みた。Gong et.al (2011)は米国ニューヨーク市で収集した GPS ログから徒歩、自動車、バス、通勤鉄道、地下鉄の交通手段判定を試みた。道路網、バス路線網とバス停、鉄道駅と出入口の空間情報と重ね合わせることで推定精度の向上が図られ

ている。また、渋滞中は自動車とバスの旅行速度に差がなくなり、さらに徒歩との判別も困難になることから、中間に徒歩を挟み前後が双方とも自動車かバスの移動については全体を自動車かバスとして判定するなどの工夫を提案した。

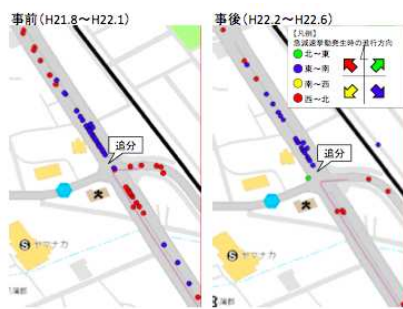
4. 地域や空間に対する認識の把握の事例

本節では地域に対するイメージや空間に対する認識の把握を試みている研究事例を紹介する。

地域に対するイメージや認識が具現化する例の一つが地価や賃料であろう。地価や賃料の推定は、面積や築年数あるいは権利関係などの物件に関する特性と、駅や公園などの利便施設や工場などの迷惑施設までの距離、傾斜度などの地形といった空間的要素をもとに推定モデルを作成する手法が知られている。三浦・浅見(2010)はこの手法の推定精度向上のために、説明変数の一つに地域に対する人々の評価を含めることを試みた。東京 23 区内の地名についてウェブ検索を行い、当該地名を含むブログ記事を抽出した。その投稿内容に対してテキストマイニングを行って抽出された、地名の指す地域に対する評価表現を「高級感」「安全性」などのグループに分類した。賃料推定物件の立地区区ごとおよびグループごとに評価表現の出現頻度を算出し、それを説明変数としてモデルに加えた。抽出する評価表現やそのグループの再構成による賃料推定モデルの決定係数上昇についても議論している。

交通事故対策へ応用した研究も見られる。交通事故発生の要因には例えば曲線や道路周辺景観による見通しの悪さや坂道による運転速度の錯覚など、人間の心理面に関わる要素も多く、アンケートなどから危険個所を抽出するには限界がある。菊地ら(2011)は、自動車急減速データを用いて事故発生の危険性がある地点の抽出と対策について論じた。事故多発地点での対策は急務ながら、その効果の検証には一定期間を経過してからの事故件数を従前と比較する手法を採るため、効果検証に時間を要する。また、事故には至らないものの危険個所の対策も必

要である。急減速は交通事故よりもはるかに多くの事例が収集でき、かつ危険個所に多く分布すると考えられる。多数の自動車の急減速データを分析することで、事故発生の危険性が高い地点の検出と現地調査による対策の立案が可能となり、対策後の効果測定も短時間で可能となる。菊地ら(2011)では、実際に愛知県内で危険な交差点を抽出し、路面への注意喚起や看板の設置などの効果を議論している(図3)。また、本研究が岡田ら(2011)や絹田ら(2008)による、利用データと一般交通状況の関連やヒヤリハット多発個所と事故多発個所の関連の検証に基づき、利用データによる現実状況の再現性や分析対象の妥当性を担保している点も特筆される。



出典: 菊地ら(2011)

図3 交差点周辺の急減速: 対策の事前と事後の比較

5. おわりに

本稿では、携帯電話 GPS ログやブログなどのウェブ投稿を中心に、ビッグデータを活用した人々の行動軌跡把握、地域や空間に対するイメージや認識の把握に関する研究を俯瞰した。いずれも現状はデータの利用可能性が注目され、様々な分析法の提案とそれらの利点と欠点が議論されている状況といえる。本稿で紹介した手法もそれら手法の一部である。

従来のような全数調査やサンプル調査とは異なり、ビッグデータは抽出されるサンプルによる母集団の再現性やデータの信頼性の検証が不足している面がある。全員が携帯型端末を所持してはならず、ブログ投稿も必ず真意や真実を反映するわけではないことに留意すべきであろう。山口(2013)、岡田ら(2011)、絹田ら(2008)などのような信頼性、再現性に関する研究の深度化が待たれる。また、本稿では紹介しな

かったものの、時間および空間分解能が高く個人情報分析も可能なデータを抽出しつつ、個人情報を確実に保護できる秘匿処理技術の開発も待たれる。

参考文献

- 大野夏海・関本義秀・中村敏和・Horanont Teerayut・柴崎亮介(2012): 東京都市圏における長期の GPS データを用いた移動経路の推定に関する研究, 第 21 回地理情報システム学会講演論文集, CD-ROM.
- 岡田朝男・水野裕彰・中村俊之・絹田裕一(2011), 道路交通安全における交通事故とヒヤリハットの関係性に関する基礎的研究, 第 31 回交通工学研究発表会論文報告集, 2011 年 9 月.
- 金杉洋・関本義秀・黒川茂莉・渡邊孝文・村松茂樹・柴崎亮介(2012): 携帯電話基地局通信履歴に基づく人の移動行動の推定可能性に関する研究, 第 21 回地理情報システム学会講演論文集, CD-ROM.
- 菊地春海・岡田朝男・水野裕彰・絹田裕一・中村俊之・萩原剛・牧村和彦(2012): 道路交通安全対策事業における急減速挙動データの活用可能性に関する研究, 土木学会論文集, 68(5), 1193-1204.
- 絹田裕一・北村清州・中村俊之・中嶋康博・牧村和彦・高橋誠・森川高行(2008), 道路交通安全対策の効果計測におけるプローブカーデータの適用可能性に関する検討, 第 7 回 ITS シンポジウム, 2008 年 12 月.
- 見生元気・伊藤昌毅・川村尚生・菅原一孔(2012): 公共交通乗換案内サービスを用いた利用者行動解析, 第 21 回地理情報システム学会講演論文集, CD-ROM.
- 小林亜令・岩本健嗣・西山智(2009): 釈迦: 携帯電話を用いたユーザ移動状態推定方式, 情報処理学会論文誌, 50(1), 193-208.
- 酒巻智宏・岩井将行・瀬崎薫(2011): マイクロブログのジオタグを用いたユーザの行動パターンの推定に関する研究, 電子情報通信学会技術研究報告, 110(400), 37-42.
- 羽田野真由美・上山智史・秋山祐樹・Teerayut Horanont・柴崎亮介(2012): GPS データを用いた商業集積地来訪者の行動パターン抽出方法の検討, 第 21 回地理情報システム学会講演論文集, CD-ROM.
- 堀口良太・長岡亨・畑成年(2006): GPS 携帯電話による大規模パーソンプローブ調査のためのトリップ情報抽出手法に関する研究, 土木計画学研究・講演集, CD-ROM.
- 三浦隆史・浅見泰司(2010): WEB 情報を活用したマンション賃料推定モデルの精度向上, 第 22 回資産評価政策学会研究発表.
- 目黒浩一郎(2008): GPS 携帯電話を活用した交通行動データ収集処理手法の開発, 情報処理学会研究報告, 2008(57), 47-54.
- 山口ともみ(2013): 低頻度 GPS 測位データからの移動経路の同定可能性の検討, 首都大学東京都市環境学部自然・文化ツーリズムコース卒業論文.
- Gong, H., Chen, C., Bialostozky, E., Lawson, C.T. (2011): A GPS/GIS method for travel mode detection in New York City, Computers, Environment and Urban Systems, 36, 131-139.