

NLモデルとパーティクルフィルタ法を融合した 都市圏レベルの人の流れの推定手法の 精度向上に向けた検討

中村 敏和¹・関本 義秀²・柴崎 亮介³

¹ 学生会員 東京大学大学院 新領域創成科学研究科 (〒277-8568 千葉県柏市柏の葉 5-1-5 総合研究棟 435 号室)

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

² 正会員 東京大学 空間情報科学研究センター 特任准教授 (〒277-8568 千葉県柏市柏の葉5-1-5)

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

⁴ 正会員 東京大学 空間情報科学研究センター 教授 (〒277-8568 千葉県柏市柏の葉5-1-5)

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

近年、地震や火災発生あるいは大規模イベント開催における混乱に伴う二次的災害や、ターミナル駅等の交通結節点における混雑などを解消する必要性から、ダイナミックに時々刻々と変動する多くの人々の流動を日常的に把握する必要性が出てきている。一方で、ICカード、ETC、スマートフォンなどの普及により、より多様な観測値の取得が可能となっている。このような背景のもと、著者らは、東京都市圏などの広範囲における人の流れを推定する際に、数%のサンプルを抜き出して行われるPT調査データからNLモデルを構築し、パーティクルフィルタ法を用いてそれらの観測値を同化することで、より精度の高い推定値を得る手法を提案した。それを踏まえ、本研究では、構築するNLモデルの精度や観測値の種類と同化後の精度の関連性について検討し、より効果の高い推定手法の構築を目指す。

Key Words : *traffic flow analysis , data assimilation, nested logit model, person trip data*

1. はじめに

地震や火災発生あるいは大規模イベント開催における混乱に伴う二次的災害や、ターミナル駅等の交通結節点における混雑などを解消する必要性から、ダイナミックに時々刻々と変動する多くの人々の流動を日常的に把握する必要性が出てきている。さらに、近年ではインフルエンザなど感染症のパンデミックが問題視されている中、さらなる感染拡大を防ぐため感染経路のシミュレーションに用いられる¹⁾など、様々な分野に応用され、その価値はさらに高まっている。

また、その一方で、近年では、各種計測技術の発展により、GPSを用いた個人の移動経路、CCTVカメラを用いた面的な人数、ICタグを用いた自動改札による駅の乗降客数、携帯電話基地局等への端末登録数、あるいはデパートの時間帯別来場者数など、様々な切り口で人の移動について計測できるようになってきている。

しかし、都市圏などの広範囲においてすべての人の流れを計測することは不可能に近いことである。そのため、今まではこうした人の流れの把握のために、パーソント

リップ（以下PT）調査データなどで偏りのないサンプルを調査することにより全体を推定するという作業が行われてきた。実際にPT調査は国勢調査と整合をとる形で拡大係数を保持しているため、属性ごとにサンプル率の逆数を拡大係数として算出し、これを各サンプルに与えることで、一人の移動を拡大係数分の人数の移動としている。しかし、さらなる路空間的な詳細化を考える場合、同じ属性を持つ人々が数十人集まり、一日中同じ経路、同じ行動を取るといのはかなり単純化された考え方であるため、様々な観測データを加味し、動的に推定できることが今後必要である。こうした流れは、地球科学などの分野では、現象のモデルに対して断片的な観測値を融合する手法として、データ同化という手法が使われており²⁾、参考になる可能性がある。

そうした背景のもと、著者ら³⁾は、東京都市圏などの広範囲における人の流れを推定する際に、数%のサンプルを抜き出して行われるPT調査データを行動パターンとして用い、パーティクルフィルタ法を用いて観測値を同化することで、より精度の高い推定値を得る手法を提案した。その中で、データ同化手法を用いることで精度

を向上させることに成功したが、朝夕のラッシュ時などの誤差にはまだ改善の余地があった。その原因として、パーティクルを生成する際に、PT調査データ自体を行動パターンとして用い、それ以外の行動パターンを生成しなかったために、行動パターンが多様でなかったことが挙げられた。

さらに、著者ら³⁴はその課題を改善するために、人の流れの推定手法としてNLモデルを構築し、多様かつ現実 に即した行動パターンの生成を行い、精度向上に成功した。ただし、構築したNLモデルは、移動の中でも目的選択と目的地選択のみを推定し、交通手段選択確率は一定であると仮定したシンプルなものであるため、推定モデル自体の精度向上には余地があると言える。

そこで、本研究では、同様にNLモデルを用いて移動モデルを構築する上で、交通手段選択まで含めることでより精緻なモデルを構築し、その効果がどれだけあるかの検討を行った。また、提案する手法では、観測値による推定値の補正が重要な要素の一つであるため、観測地点の数や観測値の質の影響について簡単に実験及び考察を行った。

具体的な構成としては、まず、第2章で、PT調査データとNLモデルを用いた移動モデルの構築について述べる。次に、第3章では、NLモデルとデータ同化手法を融合させた推定手法について述べる。そして、第4章では約80万人存在する都市を想定し、その人の流れとそれを断片的に観測したデータをシミュレートし、その観測データから推定するケーススタディを行い、手法の検証を行うとともに、第5章で本研究の結論および今後の課題についてまとめる。

2. PT調査データとモデルの構築

本章では、まず、モデル構築に用いるPT調査データについて簡単に触れ、その後、NLモデルを用いた移動モデルの構築について述べる。

(1) PT調査データについて

PT調査とは、人の一日の行動を「目的を有する移動＝トリップ」ごとに分割し、数%のサンプルを抽出しアンケート方式で調査したものである。「パーソントリップ」は、一人の人間がある目的を持って出発地から到着地までに要する移動の単位のことであり、アンケート用紙に記載し郵送回答してもらうことによって4個程度のトリップで構成される人間の一日の行動移動を知ることができる。アンケートには性別・年齢などの個人情報も記載されることで、個人属性ごとのデータの抽出も可能である。

またパーソントリップ調査において、1トリップは1～複数個の「サブトリップ」で構成されている。トリップは「出勤」「営業」などのように目的を持つトリップであるが、例えば、ある「出勤」トリップを構成する移動が「自転車・徒歩・鉄道・徒歩」という4つに細分化される時、このトリップは4つのサブトリップから構成されていることになる。

また、PT調査は対象都市圏で数%のサンプルを抽出して行うため、年代性別嗜好などの偏りなく人の行動を調査できる点で、他の行動調査データに比べて利用価値が高いと言える。

(2) 移動モデルの構築

NLモデルを用いた移動モデルの構築は、これまでに多く行われてきている（例えば、J.L.Bowmanら⁴）。本研究では、移動の目的地選択までのモデルを図-1に示す3レベルの選択ツリー構造からなるNLモデルで構築する。また、移動経路は最短経路を用いる。

モデルの説明変数リストを表1に示す。国勢調査などを利用することで、ゾーンごとの居住者数及び通勤者数の情報を取得できると考え、目的選択や目的地選択の変数に居住者数や通勤者数を用いた。また、目的として、「通勤」「帰宅」「その他」を考え、交通手段として、「鉄道」「自動車」「自転車」「その他」を選択肢とした。また、目的地はPT調査のゾーンを用いた。

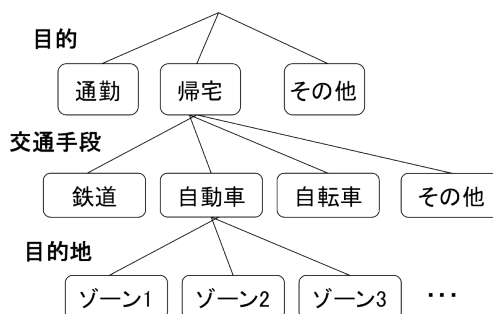


図-1 目的・目的地選択のツリー構造

表1 説明変数リスト

選択レベル	説明変数
目的選択	出発時刻ダミー $\ln(\text{出発地の居住者数})$ $\ln(\text{出発地への通勤者数})$ ログサム変数
交通手段選択	所要時間
目的地選択	$\ln(\text{目的地への通勤者数})$ $\ln(\text{目的地の居住者数})$ 目的地までの距離の逆数

3. データ同化手法を融合した推定手法

本章では、本研究で提案する NL モデルとデータ同化手法を融合した人の流れ推定の手順について述べる。データ同化手法及びパーティクルフィルタ法については、著者ら³⁾及びその参考文献に詳しく述べてあるため、ここではその適用方法について述べる。

(1) 全体フロー

まず、推定の全体フローについて述べる（図 2）。基本的な枠組みは著者らと同様であるが、本研究では、図 2 の右下で生成するトリップに対し解候補を生成する際に、構築した移動モデルを用いる。初期準備として、データ同化に用いる観測値と PT 調査データを準備する。次に、PT 調査データをトリップごとに分割し、出発ゾーンごとに振り分ける。この振り分けたトリップ集合から各ゾーンを出発ゾーンとするトリップ数とその出発時刻を計算する。そして、生成するトリップに対し、構築した移動モデルを用いて、いくつかの解候補（パーティクル）を生成し、観測値を用いてそれら解候補を評価することで推定を行う。

パーティクルの生成以外の、前処理、生成するトリップの出発時間の算出、パーティクルの重み付け部分に着いては、著者ら³⁾と同様の手法を用いる。以下、その内容について簡単に触れるが、詳しくは著者ら³⁾を参照されたい。

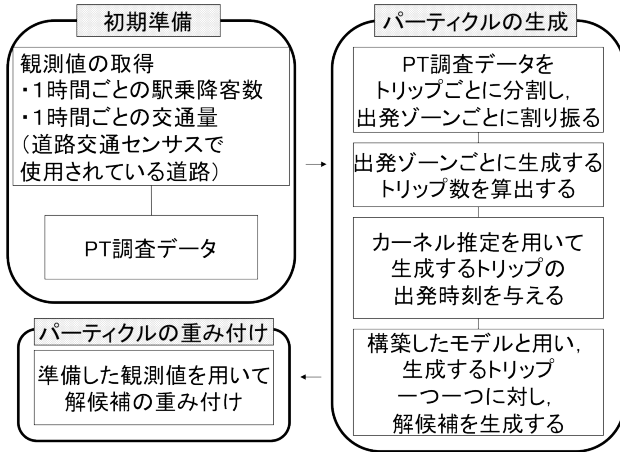


図2 全体フロー

(2) パーティクルの生成

本節では、データ同化を用いた推定手法のうち、パーティクルを生成するモデルについて述べる。本研究では、PT調査データとNLモデルから構築した移動モデルを用いて、トリップベースで確率的にパーティクルを生成するモデルを構築した。出発ゾーンごとのトリップ発生数

は、各ゾーンの人の流れデータに占めるトリップ数を抽出率で割ったものとし、出発時間をカーネル密度推定を用いて分散させる。最後に、出発ゾーンごとに発生させるトリップとして、構築した移動モデルを用いて確率的にパーティクルを生成する。以下に、その手順を詳しく述べる。

Step1 ゾーンごとにトリップ集合を生成

まず、人の流れデータをトリップごとに分割し、出発ゾーンごとに集約し、トリップ集合を生成する。例えば、あるゾーンのトリップ数が全約 200 万トリップのうち 300 トリップ存在する場合、そのゾーンには、トリップ数 300 の集合が生成される。

Step2 出発時間の分散

次に、発生させるトリップの出発時間を分散させる。人の流れデータから得られたトリップ集合の出発時間に対してカーネル密度推定を行い、発生させるトリップ数を分配する。カーネル関数として、本研究では、渡邊らが用いているように、Gaussian 関数と Silverman によるバンド幅算出方法を用いる。

Step3 パーティクル生成

各出発ゾーンごとに Step2 で分散させた出発時間ごとにトリップの解候補（パーティクル）を生成する。ここで、構築した移動モデルから生成し、多様なパーティクルを生成することで、よりデータ同化手法の効果を高める。

(3) パーティクルの重み付け

本節では、人の流れの推定モデルによって得られたパーティクルに対して、観測値を同化させる手法について述べる。推定モデルによって得られた各パーティクルに重みを与え、観測値を得るごとに重みを更新していくことで、同化していく。例えば、各トリップごとにパーティクルを 100 個生成するとすると、初期値として重み 1/100 を与え、その重みを観測値を得るごとに更新していくことになる。観測値を得たときの重みの更新は以下の式を用いて行う。

$$\hat{w}(i) = \frac{w(i) \times n(t)}{\sum_{j=1}^n w(j)} \quad (1)$$

ここで、 n は観測されたパーティクルの数、 $w(i)$ は i 番目の観測されたパーティクルの重みで、 $n(t)$ は観測値である。例えば、ある駅で 1 時間ごとに改札を出て行く人の

数を観測しているとする、 $n(t)$ は時間 $t-1$ から t までの観測値、 $w(i)$ はその間にその駅で改札を出ることになるパーティクルの重みとなる。そして、時間 $t-1$ から t までにその駅で改札を出ることになる全てのパーティクルの重みに対して、(1)式を用いて更新を行う。

4. ケーススタディ

(1) 実験条件

東京都市圏の人の流れの全体を実際に計測することは難しいため、人の流れデータから東京都市圏の人の流れを推定しても検証することが難しい。しかしその一方で、完全に仮想の都市空間で仮想的な人の流れや観測データを想定しても現実への展開性が見込みにくい。そこで、著者ら³⁾と同様に、本研究では提案する手法の検証を行うために、PT 調査データを全体の真値とするような仮想都市圏を想定した。

具体的には、推定すべき仮想都市圏約 80 万人分の人の流れに対して、仮想都市圏から 3%のサンプルを抽出し、非集計の観測データとする。そして、非集計の観測データから仮想都市圏を推定し、その推定値と仮想都市圏の人の流れ (PT 調査データ自体) を比較することで検証を行う (図-3)。また、データ同化に用いる集計の観測データとしては、東京都市圏の鉄道駅における 1 時間ごとの乗降客数と、道路交通センサスの観測地点における 1 時間あたりの交通量を得ることができると仮定して用いる (表 2)。

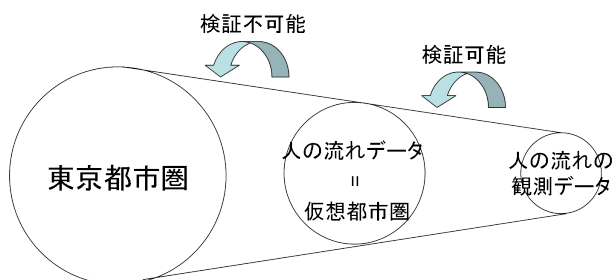


図-3 検証方法の概念

表 2 実験条件

種別		対象データ	説明
データ設定	真値	80 万人の 1 分ごとの人の時空間位置データ	東京 PT 調査データ(1998)を利用
	観測データ	道路通過交通量	道路交通センサスの観測地点を利用：2105 箇所（1 時間あたりの通過交通量）
		駅の乗降客数	首都圏の全駅（約 1400）の乗降客数（1 時間あたり）
		非集計調査データ	東京 PT 調査データ(1998)から 3%を抽出
計算環境	ハードウェア		CPU：Core 2 Duo 3GHz、メモリ：3GB
	プログラミング言語		Java
	パーティクルフィルター		パーティクル数：約 2000 万、繰り返し計算数：約 200 回、計算時間：約 2 日

(2) 推定結果

東京 PT 調査のデータから最尤法によってパラメータを算出し構築したモデルとデータ同化手法を融合した提案手法を用いて、人の流れを推定した。本研究では、モデル構築の際に交通手段推定を組み込んだ点が違いの一つであるため、まずはパーティクル生成手法の違いによる精度の変化について考察する。その際に、PT 調査自体の行動パターンをそのまま用いてパーティクルを生成する場合、目的と目的地を NL モデルで記述したモデルを用いてパーティクルを生成する場合と、本研究で構築した、目的と目的地だけでなく交通手段も含めた NL モデルを用いてパーティクルを生成する場合の 3 パターンを用いて比較を行った。本研究ではパーティクルを推定する 1 トリップにつき 5 個生成し

た。図 4 は鉄道駅における乗車客数の時間帯別 RMSE、図 5 は降車客数の時間帯別 RMSE、図 6 は道路交通センサス対象道路における自動車通過量の時間帯別 RMSE を示したものである。NL モデルを用いてパーティクルを生成することで、より観測値に即した推定結果を得ることができていることが分かるが、本研究で組み込んだ交通手段選択の部分は大きな改善につながらない。今回の実験では、パーティクルを 1 トリップにつき 5 個生成したが、その数が十分でなかったため、モデルの精度向上の効果が出にくいと考えられる。そのため、ある程度特徴をつかんだモデルを構築する事ができたら、十分な数のパーティクルを生成する事を目指すべきである。ただし、東京都市圏のような広範囲での推定を行う場合、計算リソースとの兼ね合いが出て来てしまうため、その点が課題に挙げられる。

最後に、観測値の質とその効果について検討を行う。

今回は鉄道駅の観測値、乗車客数又は降車客数について考察を行う。鉄道駅 i の観測値の質を表す指標 I_i として、以下の値を用いる。

$$I_i = -\sum_j \log(P_{ij}) \quad (2)$$

ここで、 j は観測値が乗車客数の場合はその降車駅、降車客数の場合はその乗車駅を表し、 P_{ij} は観測値全体に対する、鉄道駅 i (j) で乗車し鉄道駅 j (i) で降車した人数の割合である。目的駅や出発駅がばらけているほど I_i は大きい値となり、多く他の駅と関係する駅ほど大きい値を取る事になる。例えば、単線駅は利用のされ方の種類が少ないため小さい値に、乗換駅は利用のされ方が様々なので大きい値になりやすい。

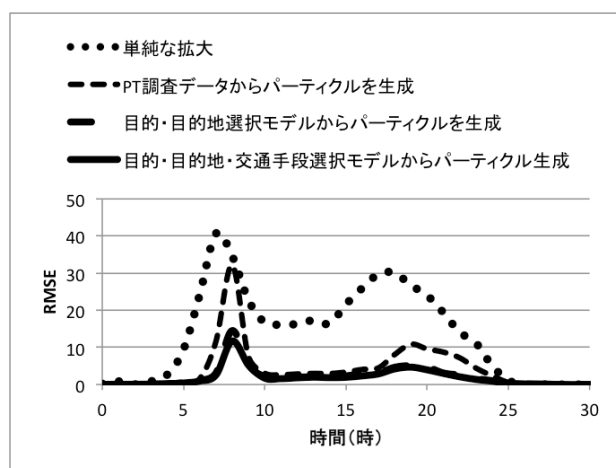


図4 鉄道駅における乗車客数の時間帯別 RMSE

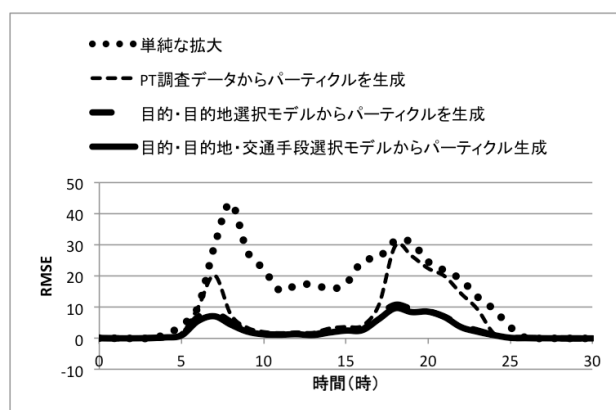


図5 鉄道駅における降車客数の時間帯別 RMSE

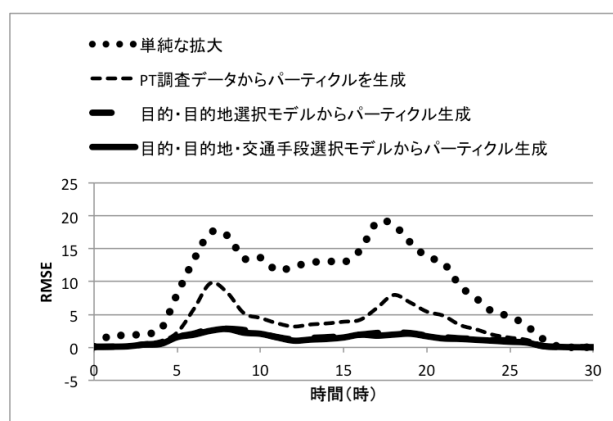


図6 道路交通センサス対象道路における通過量の時間帯別 RMSE

5. 結論

本研究では、都市圏レベルの人の流れの推定手法として、NL モデルを適用した移動モデルの構築とデータ同化手法の融合による手法を提案し、簡単な検証を行った。また、観測値の質や量と精度の関係についても簡単な検証を行った。

結果として、NL モデルを用いる事で精度向上につながっているが、NL モデル自体の精緻化は大きな改善にはつながらなかった。これはパーティクル生成数が十分でなかったことが要因として挙げられる。また、多く他の駅と関係する駅での観測値が、データ同化による効果が高い事が分かった。

今後の課題として、まずパーティクル数を変化させた場合の精度の考察が考えられる。ある程度特徴をつかんだモデルを構築できたら、観測値との同化の際に十分な数のパーティクルが必要であると考えられる。一方でパーティクルを増やすにはその分計算リソースが必要となるため、パーティクル数と精度の関係を検証することはかなり意義のある事だと言える。

謝辞：

本研究は、文部科学省のグリーン・ネットワーク・オブ・エクセレンス（GRENE）事業 環境情報分野の一環として実施したものである。

参考文献

- 1) Ohkusa Y., & Sunagawa, T.: Application of an individual-based model with real data for transportation mode and location to pandemic influenza, *Journal of Infect Chemother*, 13, 380-389, 2007.
- 2) 淡路敏之, 蒲地政文, 池田元美, 石川洋一, データ同化：観測・実験とモデルを融合するイノベーション, 京都大学学術出版会, 2009.

- 3) 中村敏和, 関本義秀, 薄井智貴, 柴崎亮介「パーティクルフィルタを用いた都市圏レベルの人の推定手法の構築」, 土木計画学研究・講演集, Vol43, CD-ROM, 2011.
 - 4) 中村敏和, 関本義秀, 薄井智貴, 柴崎亮介「NL モデルとパーティクルフィルタ法を融合した都市圏レベルの人の流れの推定手法の構築」, 土木計画学研究・講演集, Vol45, CD-ROM, 2012.
 - 5) J.L.Bowman, M.E.Ben-Akiva: Activity-based disaggregate travel demand model system with activity schedules, Transportation Research Part A 35, 2000.
- (2013.5.7 受付)

Accuracy Analysis of Estimation of People Flow in a Metropolitan Area Using Nested Logit Model and Particle Filter

Toshikazu NAKAMURA, Yoshihide SEKIMOTO, Ryosuke SHIBASAKI

Recently, people flow information has become necessary to mitigate secondary disasters following earthquakes, fires, or other major events, and to improve congestion at railway stations, roads, and public spaces. With the fast development of information technologies, nowadays the collection of people flow data becomes much easier and we can have different kinds of measurement data, such as train use data gotten by IC card, high way use data gotten by Electronic Toll Collection System, and so on. However, most of them have been used separately. Moreover, to estimate people flow, many movement models have been developed. Nested Logic Model is one of the most useful models. Based on these backgrounds, we are trying to estimate people flow in a metropolitan area by combining the movement model using Nested Logit Model and these different kinds of observation data together to make a more accurate estimation about people, based on data assimilation techniques. We propose an algorithm using Particle Filters for data assimilation of people flow data and estimate people flow in Tokyo metropolitan area, assuming that we can get the number of people who ride or drop trains at each station as observations and the number of people who use each main road in Tokyo metropolitan area.