

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

中村 敏和¹・関本 義秀²・薄井 智貴³・柴崎 亮介⁴

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

³正会員 東京大学 空間情報科学研究センター 特任助教 (〒277-8568 千葉県柏市柏の葉5-1-5)

E-mail:usui@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。本研究では、PT調査データそのものを行動パターンとして推定に用いるのではなく、PT調査データとNLモデルを用いて移動モデルを構築することでより多様なパターンを生成し、そこにデータ同化を融合することで推定精度を高める手法について提案する。

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もなっている。

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、

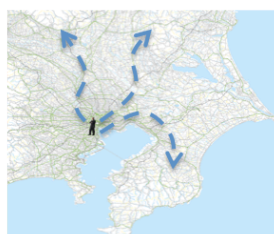
今まではこうした人の流れの把握のために、パーソントリップ (以下PT) 調査データなどで偏りのないサンプルを調査することにより全体を推定するという作業が行われてきた。実際にPT調査は国勢調査と整合をとる形で拡大係数を保持しているため、属性ごとにサンプル率の逆数を拡大係数として算出し、これを各サンプルに与えることで、一人の移動を拡大係数分の人数の移動としている。しかし、さらなる路空間的な詳細化を考える場合、同じ属性を持つ人々が数十人集まり、一日中同じ経路、同じ行動を取るといのはかなり単純化された考え方であるため、様々な観測データを加味し、動的に推定できることが今後必要である。こうした流れは、地球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

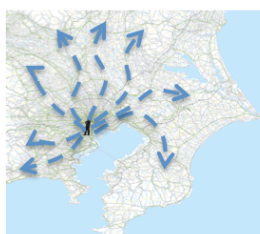
同化することで、より精度の高い推定値を得る手法を提案した。その中で、データ同化手法を用いることで精度を向上させることに成功したが、朝夕のラッシュ時などの誤差にはまだ改善の余地があった。その原因として、パーティクルを生成する際に、PT調査データ自体を行動パターンとして用い、それ以外の行動パターンを生成しなかったために、行動パターンが多様でなかったことが挙げられる。

そこで、本研究では、PT調査データとNLモデルを用いて移動モデルを構築し、多様な行動パターンを生成することで、より精度の高い推定を行うことを目的とする。図1のように、NLモデルを適用することで、PT調査のパターン以外の行動パターンを生成することができると考えられる。移動モデルとして、目的選択及び目的選択を行うモデルをNLモデルの形式で構築し、データ同化手法と融合させることで、都市圏レベルの人の流れを推定する手法を提案する。

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。



PT調査データをパターンとして用いる手法だと、PT調査データに含まれるパターンのパーティクルしか生成できない。



NLモデルを適用することで、より多様なパーティクルを生成することが可能となる。

図1 NLモデルを適用する利点

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、これまでによく行われてきている(例えば、JL.Bowmanら⁴⁾)。本研究では、移動の目的地選択までのモデルを図2に示す2レベルの選択ツリー構造からなるNLモデルで構築する。モデル自体の精度を向上させるには、通学や買物などより詳細な目的分類が望ましく、また、本来移動の推定を行う際は、目的地のみでなく、移動手段及び経路まで生成するモデルを構築するべきである。しかし、本研究では、NLモデルとデータ同化手法の融合手法の提案が主であり、その初期の試行段階として、図2のようなシンプルなモデルを構築し、移動手段は公共交通または自動車が1/2の確率で二択、移動経路は最短経路を用いる。そして、シンプルなモデルの推定値に対し観測値を同化することでその精度を向上させることができることを示す。

また、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た。

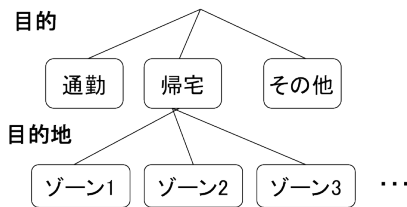


図2 目的・目的地選択のツリー構造

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
目的地選択	$\ln(\text{目的地への通勤者数})$ $\ln(\text{目的地の居住者数})$ 目的地までの距離野逆数

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3)。基本的な枠組みは著者らと同様であるが、本研究では、図3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。

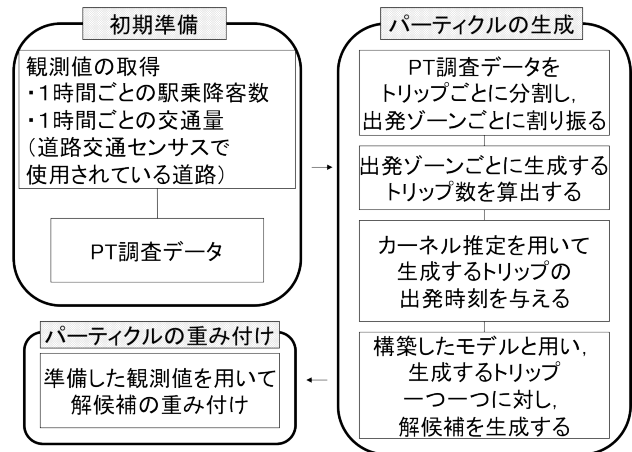


図3 全体フロー

(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4)。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)。

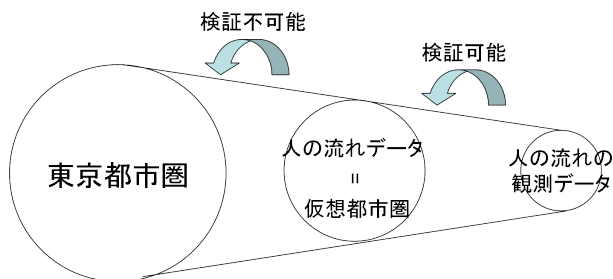


図4 検証方法の概念

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: Core2Duo 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

モデル構築の際のパラメータ推定結果を表3, 表4に示す. ほぼ全てのパラメータが5%有意水準を満たし正負も問題なかった. また, ログサム変数のパラメータも0~1に収まっている.

本研究では, 計算環境の問題から, 3%抽出したサンプル約23000人分のうち, 5000人分のサンプルを用いてパラメータ推定を行った. 人数を増やした場合にパラメータ結果がどう変化するかは将来考察したい課題の一つである.

表3 目的地選択モデルのパラメータ推定結果

変数名	通勤	帰宅	その他
$\ln(\text{目的地への通勤者数})$	1.002**	-	-
$\ln(\text{目的地の居住者数})$	-	0.9908**	-
目的地への距離の逆数	-	-	8.456**
サンプル数	5000		

** : 5%有意

表4 目的選択モデルのパラメータ推定結果

変数名	通勤	帰宅	その他
ログサム変数	0.2758**	0.2601**	0.5746**
定数項	0.2039**	0.5382**	-
6時台	2.4576*	-	-
7時台	1.9242*	-	-
8時台	0.0555*	-	-
15時台	-	1.3970*	-
16時台	-	1.5672*	-
17時台	-	1.9475**	-
18時台	-	2.0587**	-
19時台	-	2.1123**	-
20時台	-	2.6161**	-
21時台	-	3.0416**	-
22時台	-	2.3158**	-
23時台~	-	2.8012**	-
$\ln(\text{出発地への通勤者数})$	-	0.2403**	-
$\ln(\text{出発地の居住者数})$	0.2657**	-	-
サンプル数	5000		

** : 5%有意 * : 10%有意

このように構築したモデルを用いて人の流れを推定した. 本研究ではパーティクルを推定する1トリップにつき10個生成した. これは, PT調査データ自体を行動パターンとしてパーティクルを生成した場合の図5は鉄道駅における乗車客数の時間帯別RMSE, 図6は降車客数の時間帯別RMSEを示したものである. これらから, 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. 特に, 朝夕のラッシュ時において精度が向上している. これは, 今回構築したNLモデルが通勤に重きを置いたものとなっていることも大きいと考えられる. ただ, ラッシュ時の精度は向上しているが, まだ誤差が残っている. これはパーティクル数を増やしたり, モデルをより高精度に構築することで改善できると考えられる. 今後の課題として取り組んでいきたい. また, 図7に道路交通センサ対象道路における自動車通過量の時間帯別RMSEを示す. 道路交通量も朝夕のラッシュ時を中心に, NLモデルを適用することで, 精度の向上が達成できている.

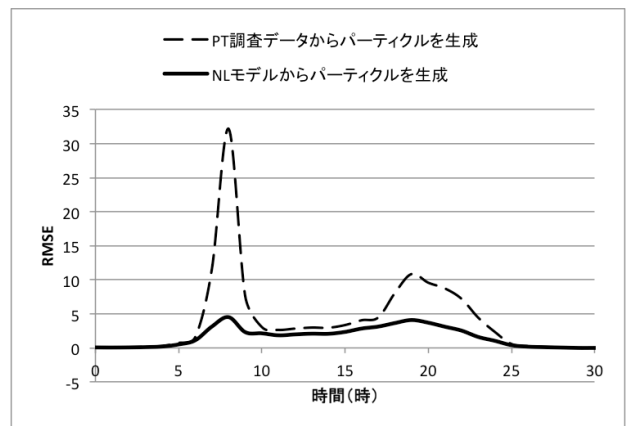


図5 鉄道駅における乗車客数の時間帯別RMSE

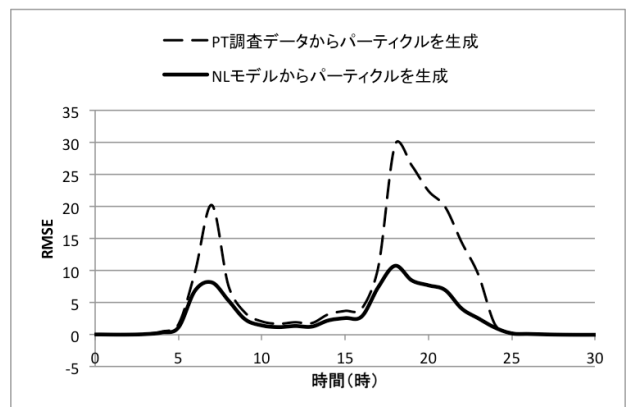


図6 鉄道駅における降車客数の時間帯別RMSE

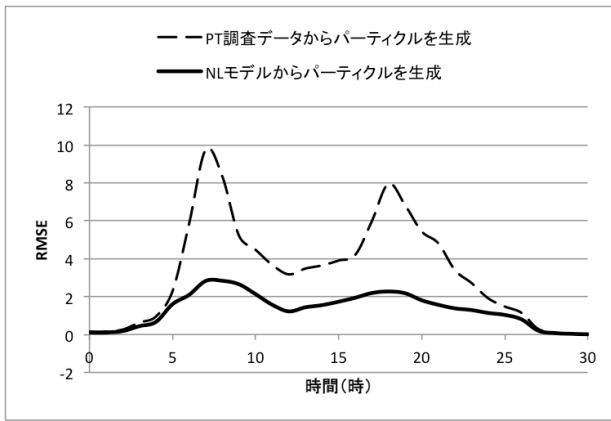


図7 道路交通センサ対象道路における通過量の時間帯別RMSE

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。NLモデルを適用した移動モデルを用いることでPT調査データをそのまま行動パターンとして用いる場合に対し、データ同化手法を融合した際の精度が改善された。特に、朝夕のラッシュ時に大きな改善が見られており、NLモデルを適用した効果が出ていると言える。

今後の課題として、パーティクル数を変化させた場合の精度の考察、モデル自体の精度向上、データ同化に用いる観測値の種類や時間間隔の検討などが考えら

れる。また、本研究では、移動手段及び経路については公共交通と自動車移動を半々で割り当て、最短経路を用いるという方法を用いたが、手段や経路選択はミクロな交通分析には欠かせないものであり、今後研究していく必要が有ると言える。

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, Vol.43, CD-ROM, 2011.
- 4) J.L.Bowman, M.E.Ben-Akiva: Activity-based disaggregate travel demand model system with activity schedules, *Transportation Research Part A* 35, 2000.

(2012.5.7 受付)

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

Toshikazu NAKAMURA, Yoshihide SEKIMOTO,
Tomotaka USUI, 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.