

jh221003

## エージェントモデルと統計データを用いた全国規模の擬似人流データの開発

関本義秀（東京大学）

近年、人間の移動に関連する広範な地理情報データセットの急増により、個人および集団レベルでの日常の移動パターンのメカニズムを解明する機会が提供されています。このような分析は、交通予測、病気の拡散、都市計画、および汚染などの社会問題を解決するために不可欠です。しかし、データ収集の対象となったユーザーのプライバシーに関する懸念から、そのようなデータの公開は制限されています。この課題に対処するために、我々は社会の基礎データの整備を目標として見据え、エージェントベースモデリングとシミュレーション手法を用いて、全国の人口に対して安定した精度を持つ「全国擬似人流データ」を作成し、提供しています。

### 1. 共同研究に関する情報

- (1) 共同利用・共同研究を実施している拠点名（該当するものを残す）

東京大学 情報基盤センター  
mdx

- (2) 課題分野（該当するものを残す）

データ科学・データ利活用課題分野

- (3) 共同研究分野（HPCI 資源を利用している研究課題のみ、該当するものを残す）

- (4) 参加研究者の役割分担

関本 義秀	研究総括
檜山 武浩	コード開発・検証
小川 芳樹	コード開発・検証
龐 岩博	コード開発・検証
鈴木 豊太郎	コード検証
小林 博樹	コード検証

### 2. 研究の目的と意義

近年、ビッグデータと関連技術の進化に伴い、都市開発をデータ駆動型に基づいてよりよく理解し、推進することが可能となりました。とりわけ、モビリティデータは、都市のダイナミクスのさまざまな指標を提供し、これによりスマートおよびデジタルシティの開発を、インテリジェントトラフィックシステム、都市計画、流行病の予防、商業などの複数の学問領域の観点から容易にします。

しかし、携帯電話などから得られた生データはプライバシーに関する懸念から、データの公開は制限されています。そのため、合成データセットの利用は、軌跡データのジオプライバシーを保護の問題を克服する有望なアプローチです。このような合成データは、本質的な統計的特性を保持しながら、広範な応用で実際のデータと同等の性能を発揮することが可能です。しかしながら、これらのデータセットやモデルの複雑さと、これらを運用するために必要な計算量が多すぎるため、多くの現行の運用モデリングシステムは比較的低い時空間分解能での位置と時間の選択を実装しています。

そして、本研究の目的は、「全国規模擬似人流データ」(Pseudo-PFLOW)を開発することで、全国規模での典型的な人間の日常活動と移

動をシームレスに捉えることです。Pseudo-PFLOW は、限定的な都市圏で収集された旅行調査データ、およびオープンデータとして利用可能な国家統計および空間データを使用して開発され、実際の移動に基づく人々の移動の表現ではなく、むしろ典型的な日常行動の擬似的な表現として機能します。

加えて、Pseudo-PFLOW を利用することで、都市の空間移動のトレンドを特定し、都市計画、商業開発、感染症の伝播のシミュレーションなど、広範な領域に適用することができます。

そして、この研究の目的と意義は、人々の動きと行動を理解し、都市の開発、交通計画、公衆衛生対策など、社会の幅広い領域に対するデータ駆動型の洞察を提供するための、合成移動データの利用と開発にあると言えます。

### 3. 当拠点の公募型研究として実施した意義

本研究が「学際大規模情報基盤共同利用・共同研究拠点」(通称: JHPCN) の公募型研究として行われた意義は以下の通りです。

1. 高度な計算資源へのアクセス: JHPCN は、スーパーコンピュータやデータ基盤、ストレージ、可視化装置などを含む高度な計算機資源を提供しています。これにより、本研究に必要な大量のデータの処理や高度な分析が可能となりました。また、このような高度な資源は、擬似データの生成や、広範囲での人口動態のモデリングなど、本研究の規模と複雑さを支える上で必須です。

2. 研究者間の協働: JHPCN のフレームワークは、各構成拠点に所属する研究者との共同研究体制を促進します。これにより、計算科学、データ科学、計算機科学の専門家との協働が可能となり、本研究のより深い洞察と高度な分析を可能にしました。さらに、これらの研究者のネットワークは、新たなアイデア

や手法を共有し、問題を解決するための追加的なリソースを提供します。

3. 研究ネットワークと公共貢献: JHPCN は、研究者が自身の研究成果を拠点シンポジウムで発表し、国外で発表したり出版したりするための支援も提供します。これにより、本研究の成果が広く共有され、その影響力が拡大します。また、このような公開的な研究プラットフォームを通じて、本研究は公共の利益に貢献し、都市開発や公衆衛生などの重要な社会問題に対する解決策を提供することができます。

したがって、本研究が JHPCN の公募型研究として行われたことは、その質と影響力を高め、広範で深い洞察を提供する上で非常に重要な意義を持ちます。

### 4. 前年度までに得られた研究成果の概要

前年度までに、本研究は日本全国 47 都道府県 1724 市区町村、約 1 億 3 千万の人口を対象として、擬似人流という合成の人間移動を表すデータセットを作成しました。このデータセットは、公開統計データと限定的な移動調査を利用し、個々の人々の日常の移動行動の時空間パターンを捉える 7 つの基本的な日常活動に従う 3 つのエージェントモデルを含む擬似人流ジェネレータを構築し、大規模計算シミュレーションを用いて擬似人口データ・擬似活動データ・擬似トリップデータ・擬似軌跡データ 4 種類のデータセットを作成しました。作成したデータセットは、既存の地上の真実データセットと比較され、時間ごとの人口分布、旅行量、旅行範囲を含む統計的特性を再構築し、決定係数が 0.5 から 0.98 の範囲であることが示されています。

なお、本データセットは東京大学 CSIS が整備する研究用空間データ基盤 JoRAS を経由で提供している。本年度までに、12 研究グループにデータを提供し、共同研究を行なって

います。

## 5. 今年度の研究成果の詳細

今年度の研究成果では、擬似人流データの生成に新たなアプローチを2つ組み込むことで、移動データの精度と現実性を一段と向上させることができました。これらのアプローチは、

1. 道路網や公共交通の時間表による混雑を考慮した交通手段選択と移動経路の生成

裾野市を対象としたこの研究では、GTFS データを用いて道路網や公共交通の時間表に基づいて混雑状況を考慮し、それによって交通手段の選択と移動経路の生成を行いました。これにより、合成的な人間の移動データに混雑状況に対する応答性を持たせることができ、データの現実性を向上させることができました。また、地元の旅行調査データと比較して、新たなアプローチがどの程度現実の旅行行動を反映できているかを評価しました。

2. ミクロ交通シミュレータを組み込むことで、実際の道路網上の混雑状況を再現

千葉県を対象としたこの研究では、ミクロ交通シミュレータである Multi-Agent Transport Simulation (MATSim) を組み込み、実際の道路網上の混雑状況を再現しました。MATSim を用いることで、各個人の移動行動や交通量が道路網上でどのように相互作用するかを詳細にシミュレートすることが可能となり、これにより更に現実に近い人間の移動データを生成することができました。また、実際の交通量データやパーソントリップ調査データと比較して、シミュレーション結果の妥当性を評価しました。

これらの研究成果は、合成的な人間の移動データの生成における新たな進歩を示しており、このようなデータを用いた研究や応用の可能性を広げています。また、これらのア

プローチは、さまざまな地域や状況に応じてカスタマイズや適用が可能であり、その汎用性も示しています。今後は、これらのアプローチをさらに改善し、研究の精度を更に高めることを目指します。

## 6. 進捗状況の自己評価と今後の展望

この研究では、オープンな統計データと限定的な旅行調査を用いて、"Pseudo-PFLOW" と呼ばれる合成的な人間の移動データを生成するアプローチが提案されています。評価は携帯電話のデータと都市圏の旅行調査データを用いて行われ、その結果、生成されたデータセットは実際のデータの統計的な特性を以下の3つの観点で再現することができました：人口分布、旅行量、旅行範囲。しかしながら、いくつかの限界も明らかにされており、その中で最も重要なものは、精密な時空間レベルでの精度の確保と、道路網や公共交通の時間表による混雑を考慮しない結果の偏りです。

この研究の自己評価に基づくと、以下のような点が挙げられます：

1. 進捗状況：提案したアプローチは、大規模な人間の移動データを生成する能力があり、その精度は各都市圏の規模と旅行の目的によって異なるという結果が得られました。この結果は、この研究の目標であった人間の移動データの合成に成功したことを示しています。一方で、この研究はまた、提案したアプローチのいくつかの限界を明らかにしており、これらの問題に対処するための改善策を開発する必要があります。

2. 今後の展望：この研究の今後の展望としては、まず、高度な時空間レベルでの精度の確保に取り組むことが挙げられます。これは、活動の選択や目的地の選択などの個々の行動モデルの改善によって実現可能です。また、既存のパーソントリップ調査などの人の

行動に関する調査データには、物資の流動やビジネスストリップなどが十分に捉えられていないという限界があります。そのため、これらのデータを基に生成される人の流動や交通流は、実際の状況よりも過小評価される傾向があります。これらの改善を行うことで、更に現実に近い人間の移動データを生成することが期待されます。

(preprint) Jue Ma, Yuya Shibuya, Yanbo Pang, Hiroshi Omata, Sekimoto Yoshihide. A Cost-and-Effect Simulation Model for Multiple Urban Planning Scenarios Using Pseudo-People-Flow Data. (2023). Preprint. DOI: 10.2139/ssrn.4396161  
(preprint) Takehiro Kashiyama, Yanbo Pang, Yoshihide Sekimoto, Takahiro Yabe. (2022). arXiv preprint arXiv:2205.00657

## 7. 研究業績

### (1) 学術論文 (査読あり)

なし

### (2) 国際会議プロシーディングス (査読あり)

Aayush Tewari, Yanbo Pang, Yoshihide Sekimoto (2022). Uncertainty of Traffic Congestion Estimation using Nationwide Pseudo Trip Data and Agent-based Simulation. *IEEE BigData '22*

### (3) 国際会議発表 (査読なし)

なし

### (4) 国内会議発表 (査読なし)

Yanbo Pang, 榎山武浩, 関本義秀. 強化学習を用いた都市圏レベルの人の流れの再現手法の構築. 第29回地理情報システム学会

Yanbo Pang, 榎山武浩, 関本義秀. 全国擬似人流データの提供と評価. 第31回地理情報システム学会  
澁谷 遊野, Yanbo Pang, 関本 義秀. シナリオに基づく擬似人流生成のためのシミュレーション基盤の構築. 第31回地理情報システム学会

笠原 有貴, 関本 義秀, 龐 岩博, 榎山 武浩, 瀬崎 薫. 擬似人流データにおける時刻表を考慮した自治体全域の交通手段の推計—静岡県裾野市を対象に. 第31回地理情報システム学会

### (5) 公開したライブラリなど

今研究で生成した擬似人流データセットは東京大学 CSIS が整備する研究用空間データ基盤 JoRAS を経由で提供しています。一方、擬似人流データの生成コードは、Github 上に公開しています。 (<https://github.com/sekilab/DSPFlow>)

### (6) その他 (特許, プレスリリース, 著書等)