

科学研究費助成事業 研究成果報告書

令和 6 年 6 月 26 日現在

機関番号：12601

研究種目：挑戦的研究（萌芽）

研究期間：2022～2023

課題番号：22K18498

研究課題名（和文）グローバル人流データコモンズ創出のための基礎的研究

研究課題名（英文）Fundamental Research for the Creation of a Global Human Flow Data Commons

研究代表者

関本 義秀（SEKIMOTO, YOSHIHIDE）

東京大学・空間情報科学研究センター・教授

研究者番号：60356087

交付決定額（研究期間全体）：（直接経費） 5,000,000円

研究成果の概要（和文）：本研究では、様々な国や地域の政策現場で汎用的に活用できることを念頭に置いた「全球規模の各国レベルの擬似人流データ基盤とそのエージェントモデル」を構築する事を目標にした基礎的な研究を進めていく事としていた。成果としては世界初となる日本全土をカバーする1.3億人の全国擬似人流データ（3種類のエージェントと7種類の行動）を作成し、所属の空間情報科学研究センターの共同利用データ（JoRAS）から目的地選択候補に数百万の実POIデータを使ったり、交通手段選択についてもエリアごとに過去の統計値に基づきパラメータを変化させ、よりリアリティを持たせるVer2.0のリリースまで行う事ができた。

研究成果の学術的意義や社会的意義

これまで携帯端末による人流データは商業的にも広まってきているものの、プライバシーの関係上、集計データに限られる事が多かったり、それでも値段も大変高価であった。そのような意味で、全国をカバーし、シミュレーションが行いやすいデータ基盤として、都市計画・交通計画等、公益に資する活動のための基礎データの整備が行われたのは大変意義が大きい。

研究成果の概要（英文）：In this study, we were going to conduct basic research with the goal of constructing a “global-scale, country-level pseudo-human flow data infrastructure and its agent models” that can be used universally in various countries and regions for policy making. The results of the research include the world's first creation of nationwide pseudo-human flow data (three types of agents and seven types of actions) for 130 million people covering all of Japan, the use of millions of real POI data for destination selection candidates from the jointly used data (JoRAS) of the Center for Spatial Information Science, to which the company belongs, and the use of parameters for transportation mode selection based on historical statistics for each area. We were also able to release Ver. 2.0, which has more realism by changing the parameters based on past statistics for each area.

研究分野：人間都市情報学

キーワード：擬似人流 デジタルツイン 自動抽出 建物

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

(1). 研究開始当初の背景

近年、デジタル化の急速な進展・高度化が進む中、さまざまなデータを迅速に組合せ、社会課題を解決していく事が期待される。そのため、現実世界をサイバー空間で再現するとともに、個人・民間企業・国家の意思決定と政策立案へ基盤データの構築と整備は不可欠である。空間情報学の分野は、従来からパーソントリップ調査(以降、PT調査と呼ぶ)から得られたトリップデータに対して、時空間内挿入を行うことで生成したデータであり、現在36都市圏、延べ約700万人分の人流データを研究者に提供している。しかしながら、大規模調査であるために多大な費用と労力を要する。また、データセットがカバーする範囲と人口比率が限定されるといった課題も存在している。

一方、近年ではパーソントリップ調査データの他にも、匿名化された携帯電話データの利用可能性が広がってきている。例えば、モバイル空間統計は通信基地局通信履歴(CDR: Call Detail Record)に含まれる位置情報から集計されたデータであり、1時間毎に日本全国のメッシュ別の滞在人数を24時間365日把握することができるデータである。しかし、携帯電話データは非常に高価であるため一般的に入手することは難しい状況にある。さらに、プライバシー保護の観点から、その実データから個人属性や移動属性が提供されることなく、それぞれの移動者の移動目的や地域変容による変化に関する分析は困難である。

(2). 研究の目的

そこで、我々は実世界の人々の動きを俯瞰可能とし、都市計画・交通計画等、公益に資する活動のための基礎データの整備を目標として見据え、安定した精度を持つ「全国擬似人流データ」を作成して提供する。本データセットは下記の特徴を持つ。

- ・ 擬似人流データの開発では、一般的に入手可能なオープンデータとして公開される統計データと既存のPT調査データの集計結果、そして建物データ等の低廉に入手可能なデータのみを用いること。
- ・ 上記の集計レベルの調査データを用いることで、合成的人口を作成し、そして人々の典型的な日常の行動を擬似的に再現し、リアルな人の位置情報ではないため、その結果は研究目的で公開可能。
- ・ 位置情報だけでなく、全人口に対する一人一人の人口属性(世帯・年齢・性別・就業状態)、1日の生活活動、活動に伴う発生した移動トリップ(目的・時刻・起点・終点・交通手段)、また移動軌跡をすべて提供する事。

(3). 研究の方法

本研究では、オープンな調査データと低廉な価格で入手できる商業データから典型的な全国における平日の1日中の擬似人流を再現する。結果として、断片的な位置情報だけでなく、どのような人が、どのような目的で、いつ、どのような交通手段で、どこからどこへ移動するかといった情報を全般的に提供する。そして、今回提供する擬似人流データは擬似人口データ、擬似活動データ、擬似トリップデータ、擬似軌跡データの4種類のデータセットで構成される。ここでは、それぞれのデータの詳細と生成方法について紹介する。

・ 擬似人口の生成

個人の日常的活動(例えば、通勤・通学・食事・買物等)は、人口属性によって影響を受ける。例えば、個人の就業状態によって、昼間に行う生活活動に制限がつき、また、住所によって職場

や生活圏が決められる．そこで，本研究の第一歩としては，国勢調査から提供する集計レベルの人口統計データに基づいて，個人の世帯構成・年齢・性別・就業状況・住所等の情報を推定し，擬似人口データを生成する．具体的には，梶原らの手法により，統計データから世帯単位の人口分布を推定し，そしてゼンリンの建物データに世帯データを割り当てたものであり，建物ごとに世帯数・家族構成・年齢・性別などの属性情報を個人に配分する．

．擬似活動データの生成

地域交通計画策定や将来交通需要予測には，移動軌跡だけでなく，移動の目的，つまり各場所に行く活動についての情報を得る必要がある．一方，近年の非集計の交通行動分析モデルの研究結果により，従来のトリップベースモデルと比べると，アクティビティベースモデルは個人の活動の派生需要と時空間制約の考慮等，時空間の細かな表現ができるメリットがある．そして，擬似人流を再現する際，まず，個々の人間の1日の活動全体を推定する．結果として，擬似活動データは「誰が，いつ，何を」の情報を提供する．

．擬似トリップデータの生成

3.2 で推定した活動内容により，活動は異なる場所に行くと新しいトリップが発生する．本研究では，各トリップに対して，「誰が，いつ，何の目的で，どこからどこへ」の情報を提供する．その内，移動者「誰」，移動目的「活動内容」，出発時間「活動時間」は前節で説明したが，行動先は，2.1 で割り当てた就業状況と移動目的に応じて決定される．

．擬似軌跡データの生成

擬似軌跡データの生成については，交通シミュレータを用い，混雑度や信号制御などを考慮して計算することが最適であるが，既存の交通シミュレータで全国分の人流を対象とすることは難しい．そこで，第一プロトタイプ of 擬似人流の開発では，ダイクストラ法による最短経路探索により移動経路を算出した．交通手段が徒歩，自転車，自動車の場合には DRM の道路ネットワークデータを，鉄道については金杉らが開発した鉄道ネットワークデータを使用した．道路ネットワークのコストとしては，道路リンク毎に設定された制限速度とリンク長から算出した移動時間を，鉄道ネットワークについては鉄道のリンク長(路線の長さ)を設定して計算した．なお，作成した軌跡データは，数秒ごとに移動者の位置を記録し，GPS データと同じように扱うことは可能となる．図1に示すように，擬似軌跡データを道路や鉄道の種別を基準として，ズームレベルに応じた可視・不可視調整を行うことができる．

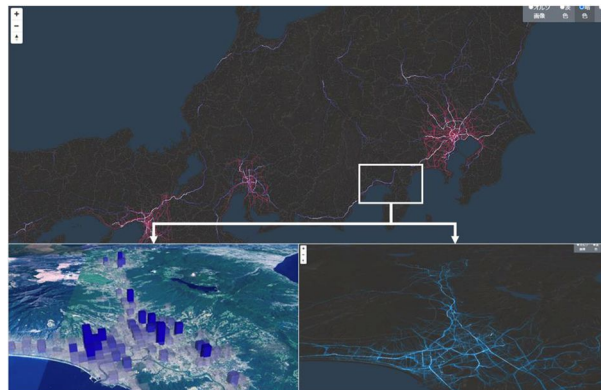


図1 作成した擬似人流の可視化の例．

(上)リンクベース交通量 (下)左：メッシュ滞在人口 右：軌跡データ

(4). 研究成果

・評価方法

本研究では、人口分布・トリップ数・トリップカバー率3つの集計量から作成した擬似人流データを評価する。その中、人口分布に対して富山県と静岡県のコモモビル空間統計データを用いて評価する。トリップ数に対して近畿都市圏 PT データを用いて評価する。トリップカバー率は東駿河湾都市圏 PT データを用いて評価する。

・人口分布に関する評価

図2(a-c)には、コモモビル空間統計(静岡県と富山県)を用いて、12時における人口分布を市区町村・1kmメッシュ・500mメッシュ3つのレベルで評価した。比較結果として、作成した擬似人流データは市区町村レベルにおいて携帯電話データとほぼ同じ人口分布であるが(相関 $R^2 = 0.98$)、空間解像度を上げつつ相関が低くなる傾向がある。現段階の擬似人流データは、1kmメッシュレベルでは実世界の観測データと同じように扱うことが可能と考えられる。

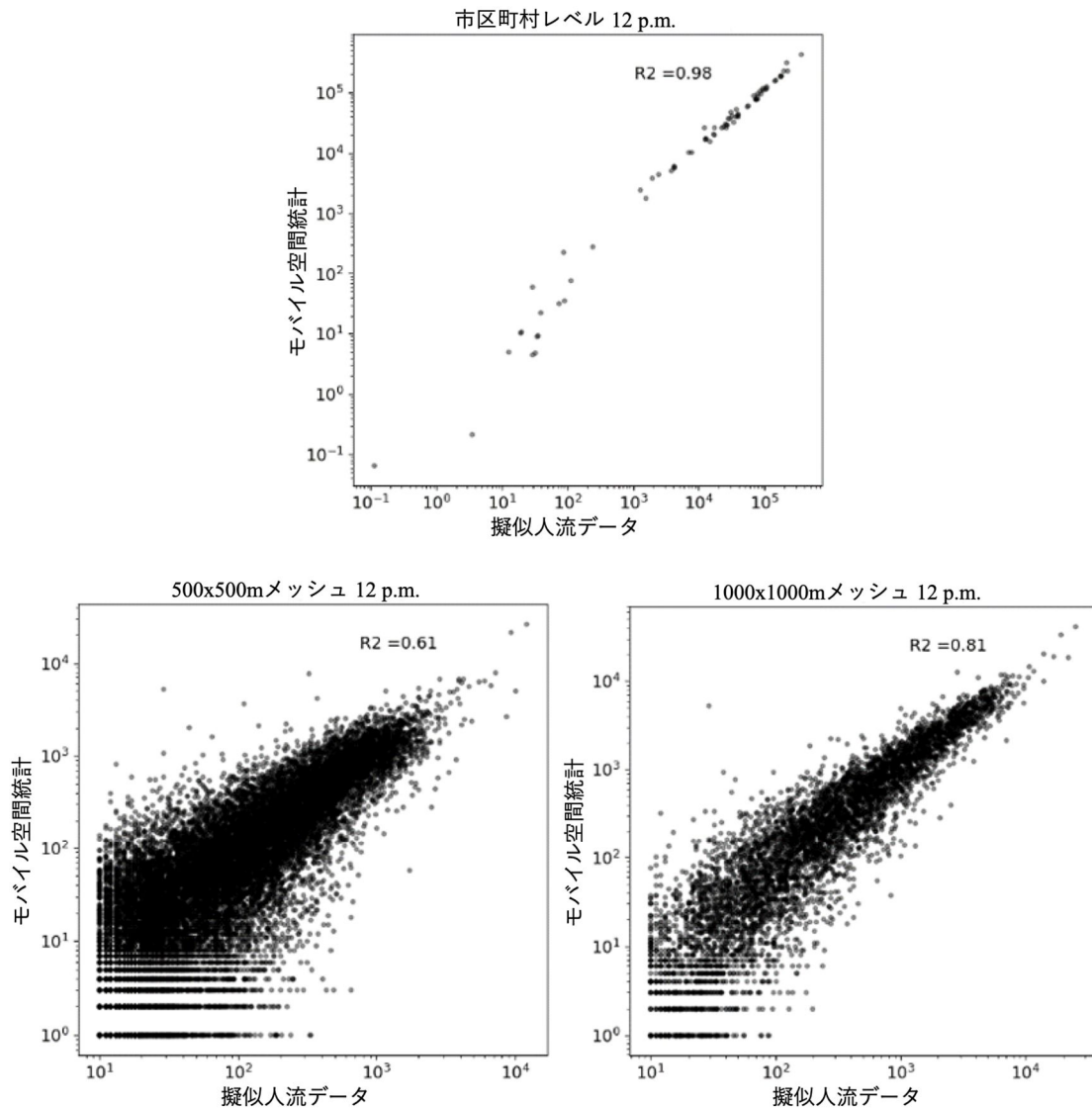


図2 解像度別人口分布の評価結果

((a)市区町村別 (b)1kmメッシュ (c)500mメッシュ)

・トリップ数に関する評価

通勤・通学と私用(買物・通院・食事・その他)トリップのOD量について、大都市圏と地方都市圏別に精度を検証した。2008年東京都市圏を大都市 2011年の中京都市圏 PT 調査と 2016年の東駿河湾 PT 調査と疑似人流の比較を行った(図3)。

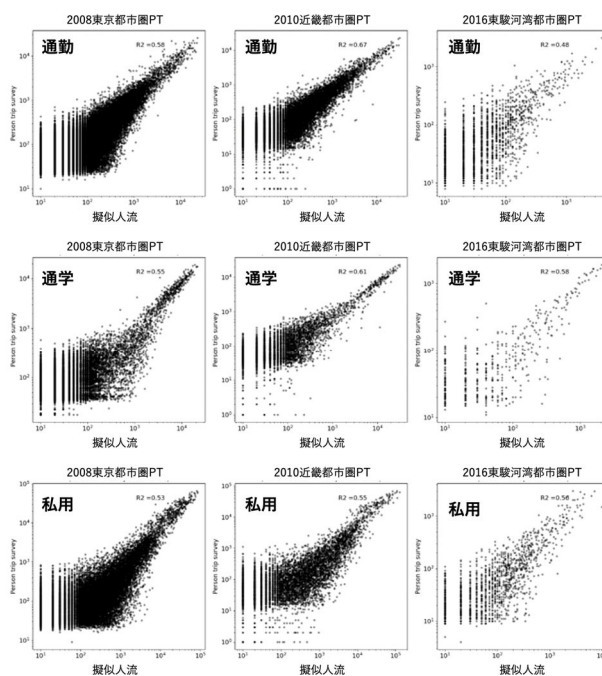


図3 目的別トリップ数の評価結果

・トリップカバー率に関する評価

本研究では、2008年東京都市圏 PT データ・2010年近畿都市圏 PT データ・2016年東駿河湾都市圏 PT データを用いて疑似人流データの生成トリップカバー率を評価した。表1の結果より、東京・近畿都市圏ではほぼ全数を再現できている。一方、東駿河湾都市圏の結果はPT データより10-17%不足していることがわかる。本データセットの行動パターンは大都市圏の調査データから抽出したため、地方の結果でも大都市へ偏り傾向という課題が存在している。

表1 都市圏別目的別トリップカバー率

移動目的	東京都市圏	近畿都市圏	東駿河湾
通勤	95%	105%	83%
通学	104%	113%	102%
娯楽	96%	106%	90%
帰宅	93%	97%	84%

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件/うち国際共著 1件/うちオープンアクセス 2件）

1. 著者名 Chen, S., Ogawa, Y., Zhao, C., & Sekimoto, Y.	4. 巻 195
2. 論文標題 Large-scale individual building extraction from open-source satellite imagery via super-resolution-based instance segmentation approach	5. 発行年 2023年
3. 雑誌名 ISPRS Journal of Photogrammetry and Remote Sensing	6. 最初と最後の頁 129-152
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.isprsjprs.2022.11.006	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Takehiro Kashiyama, Yanbo Pang, Yuya Shibuya, Takahiro Yabe, Yoshihide Sekimoto	4. 巻 -
2. 論文標題 Nationwide Synthetic Human Mobility Dataset Construction from Limited Travel Surveys and Open Data	5. 発行年 2024年
3. 雑誌名 Computer-Aided Civil and Infrastructure Engineering	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する

〔学会発表〕 計4件（うち招待講演 0件/うち国際学会 3件）

1. 発表者名 Chen, S., Ogawa, Y., Zhao, C., and Sekimoto, Y.
2. 発表標題 Large-scale building footprint extraction from open-sourced satellite imagery via instance segmentation approach
3. 学会等名 IGRASS2022（国際学会）
4. 発表年 2022年

1. 発表者名 ホウ岩博, 櫻山武浩, 関本 義秀
2. 発表標題 全国擬似人流データの提供と評価
3. 学会等名 第31回地理情報システム学会
4. 発表年 2022年

1. 発表者名 Zhang, K., Pang, Y., and Sekimoto, Y
2. 発表標題 Deep Learning Approach to Logistics Trips Generation: Enhancing Pseudo People Flow with Agent-based Modeling
3. 学会等名 IEEE ITS Annual Conference (国際学会)
4. 発表年 2023年

1. 発表者名 Yanbo Pang, Pierre Ferry, Kunyi Zhang
2. 発表標題 Synthetic Network Traffic Data Generation using Deep Generative Models
3. 学会等名 Netmob 2023 (国際学会)
4. 発表年 2023年

〔図書〕 計1件

1. 著者名 Yoshihide Sekimoto, Yasuhiro Kawahara	4. 発行年 2023年
2. 出版社 Springer	5. 総ページ数 189
3. 書名 Utilization of Geospatial Information in Daily Life: Expression and Analysis of Dynamic Life Activity	

〔産業財産権〕

〔その他〕

<p>「擬似人流とは」(人の流れプロジェクト内) https://pflow.csis.u-tokyo.ac.jp/data-service/pseudo-pflow/</p>

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	榎山 武浩 (Kashiyama Takehiro) (10611155)	大阪経済大学・経済学部・准教授 (34404)	
研究分担者	矢部 貴大 (Yabe Takahiro) (30940431)	東京大学・空間情報科学研究センター・客員研究員 (12601)	
研究分担者	P a n g Y a n b o (Pang Yanbo) (60870178)	東京大学・空間情報科学研究センター・特任助教 (12601)	
研究分担者	小川 芳樹 (Ogawa Yoshiki) (70794296)	東京大学・空間情報科学研究センター・講師 (12601)	
研究分担者	瀬戸 寿一 (Seto Toshikazu) (80454502)	駒澤大学・文学部・准教授 (32617)	
研究分担者	澁谷 遊野 (Shibuya Yuya) (20847917)	東京大学・空間情報科学研究センター・准教授 (12601)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------