

メッッシュ統計人口と個人移動軌跡データの統合による 建物内滞在人数の推定

河森 大樹[†] 西村 拓哉[†] 赤木 康紀^{††} 谷 宣廣^{†††} 松園 崇市^{†††}

[†] 株式会社 NTT データ 技術革新統括本部 〒135-8671 東京都江東区豊洲3-3-9

^{††} 日本電信電話株式会社 NTT 人間情報研究所 〒239-0847 神奈川県横須賀市光の丘1-1

^{†††} 株式会社 NTT データ コンサルティング&ソリューション事業本部 〒135-8671 東京都江東区豊洲3-3-9

E-mail: [†]{Daiki.Kawamori,Takuya.Nishimura}@nttdata.com, ^{††}yasunori.akagi.cu@hco.ntt.co.jp,
^{†††}{Nobuhiro.Tani,Soichi.Matsuzono}@nttdata.com

あらまし 建物内に滞在する人数の把握は交通や都市の計画、小売業の出店計画において重要である。例えば小売業では、ビル内への新規出店を検討する際に、あらかじめビル内に滞在する人数が把握できていれば利益の大きいビルへの出店が可能となる。先行研究は人口統計情報に土地の性質などの補助情報を組み合わせることで、空間的に高解像度な領域の人数を推定しているが、建物単位での推定は実現できていなかった。本研究では、メッッシュ統計人口と個人移動軌跡データを入力として用いた。それらの関係性を分析した後、それらを組み合わせることで建物単位の全滞在人数を推定する手法を提案した。また、実データを用いてベースライン手法よりも高精度に推定できることを確かめた。

キーワード 滞在人数推定、地理データ、デジタルツイン

1 はじめに

近年、society5.0に代表されるように現実空間から収集したデータを仮想空間上で分析・活用する取り組みが注目を集めている。仮想空間上に集積したデータに基づく予測やシミュレーションの結果を現実世界にフィードバックすることで、経済発展や社会的課題の解決を実現することが期待されている。このようなデータの中でも特に注目を集めているのが人流データである。人流データとは、人間の移動や滞留といった行動を観測したデータを指す。代表的なものには、人口統計情報や個人移動軌跡データがある。人口統計情報とは各エリア・各時刻におけるエリア内の人口を集計・統計化したデータを指す。従来から存在していた人口統計情報としては国勢調査などがある。これは行政地区単位でエリア内の全人口を把握可能なデータであるが、一年に一度程度しか観測できないという問題があった。国勢調査の他に近年では、携帯電話をはじめとする情報機器の進歩により、よりリアルタイムな人流を観測可能な「メッッシュ統計人口」が注目を集めている。これは携帯電話のつながる仕組みをもとに作成されており、一辺が数百メートル程度の四角形の領域内の全人口を、数時間ごとに観測可能なデータである。代表的なデータとしては、株式会社ドコモ・インサイトマーケティングが収集・販売している「モバイル空間統計」[4]が挙げられる。このデータは商圈分析や物流の効率化に活用されている。個人移動軌跡データとは特定のユーザのある時刻における位置を観測したデータである。代表的なデータには、ピーコン、GPSなどが存在する。これらのデータは、タクシーやバスのための需要予測やマーケティング、防災計画への活用が検討さ

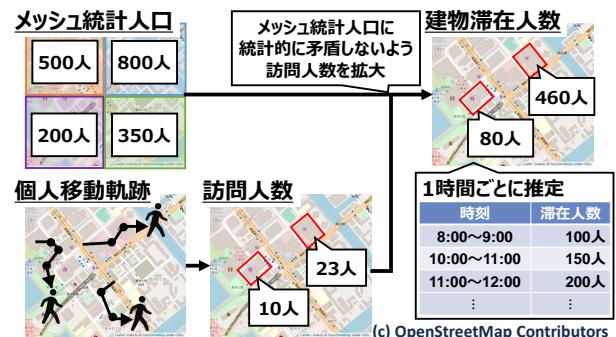


図1 本研究では、メッッシュ統計人口を個人移動軌跡から抽出した訪問人数に応じて各建物に割り当てることで建物滞在人数を推定する。本研究の貢献は以下の二つである。(1) 実データに基づきメッッシュ統計人口と個人移動軌跡データの関係性を明らかにした。(2) 個人移動軌跡から訪問人数を抽出することで現実の人流を正確に捉えた、拡大推計手法の提案。本論文での地図描画にはOpenStreetMap [11] を用いた。

れている。

このように様々な用途への活用が期待されている人流データであるが、交通の需要予測やマーケティングに利用する際にはいつ・どの建物に・どのくらいの人がいたかを知る必要がある。以降ではこのような情報のことを建物滞在人数と呼ぶ。任意の建物内に滞在する全人数が分かれれば、利用者数が多そうな建物の周辺にタクシーを配車する、小売業において新規店舗の出店を検討する際に来店者数をあらかじめ予測する、といった施策の実施が可能になる。しかし、既存の人流データからは建物滞在人数を取得することができない。例えば人口統計情報は、エリア内の全人数を取得することが可能なデータであるが、建物

単位の人数は把握できず、個人移動軌跡は建物レベルの解像度で位置情報を取得できるが、全人数を観測することができない。

そこで本研究では、複数の人流データを組み合わせることで、任意の建物内に滞在する全人数を推定するタスクに取り組む。建物単位のようなごく狭い領域における滞在人数の推定を図る先行研究には様々なものが存在する[2][7]。これらは人口統計情報における全人数を、エリアの性質（市街地・居住地など）や建物の密集度等の補助情報に基づいて、より詳細な領域にマッピングするものである。しかし先行研究には以下のようないくつかの課題がある。

- 補助情報の持つ空間解像度以上に詳細な領域に対しては妥当な割り当てを行うことができない。
- 詳細な領域における全人数を推定したとしても、それを検証するための建物ないしはエリア単位の真の人数を取得できているものは少ない。

そこで本研究では、人口統計情報としてモバイル空間統計を選択し、さらに組み合わせる補助データに個人移動軌跡を利用する手法を提案する。本手法における入出力と全体の概要を図1に示す。個人移動軌跡は観測可能なユーザ数こそ少ないが、建物レベルの空間解像度を持つ人流データである。これを用いることで建物単位の空間解像度であっても妥当な重みを割り当てることが可能になる。建物滞在人数を推定するまでの具体的な手順は、まず個人移動軌跡データが取得可能なユーザの中から、全人数を推定したい建物に訪問したユーザの人数を取得し、これを人口統計情報に矛盾しない形で拡大推計することで、建物単位の全人数を推定するというものである。

また本稿では、実際に収集したメッシュ統計人口と個人移動軌跡データを用いてこれらの関係性分析を行う。この分析により、人口統計情報と個人移動軌跡の間に相関関係は存在するのか、また平日休日・時間帯・エリアといった軸でデータを区切った際に、人口統計情報と個人移動軌跡の間の関係性がどのように変化するかといった実データでしか知り得ない知見を示す。加えて、本稿では提案手法の有効性を示すために、弊社ビルの入退館記録を用いて作成した実際の滞在人数を用いた。これは建物単位で全人数を取得することの困難さから、先行研究ではほぼなされておらず、本稿では実データを正解として用いた分析・評価であるからこそ見えた課題の抽出結果を示す。

本研究の貢献を以下に示す。

- (1) 実データに基づくメッシュ統計人口と個人移動軌跡データの関係性の分析。
- (2) 個人移動軌跡の中から実際に建物に訪問したユーザを抽出することで、現実の人流をより正確に反映した拡大推計手法を提案。
- (3) 実際のビル内滞在人数を用いた提案手法の有効性検証。

本論文の構成は以下の通りである。2章にて関連研究、4章に本稿の提案手法、5章で実験結果を示した後、6章で結論を述べる。

2 関連研究

2.1 動的人口推定に関する研究

動的人口推定とは任意のエリア内の定住者及び通勤・通学などで領域外から流入する人間も含めて推定するタスクを指す。本タスクでは數十分、数時間単位の人数というように、比較的高い時間解像度での推定結果が求められる場合が多い。動的人口推定は、近年の情報技術の発達により、携帯電話の位置情報、SNSの投稿位置やICカードの利用履歴情報といった時間解像度の高いデータが取得可能になったことで注目を集めた分野である。Devilleら[3]は携帯電話の基地局情報をもとに推定した動的人口を国勢調査等の既存の人口統計情報と比較することで、携帯電話の位置情報活用の可能性について述べた。しかし上記のような複数データの活用においては、各データの時間解像度がデータソースによって異なるという点が課題となっている[8]。

本研究では時間解像度の低いモバイル空間統計に合わせて、個人移動軌跡データから取得した訪問人数を一時間単位に集計することで、データごとの時間解像度の違いに対処している。

2.2 人口統計情報の空間的な高解像度化に関する研究

人口統計情報の高解像度化を図る既存研究における基本的な手順は、人口統計情報をグリッドや建物のようなさらに小さな領域（小領域）に分割し、この領域内に滞在する全人数を推定するというものである。これらの研究はアプローチの違いによって大きく二つに分けられる[1]。一つ目がボトムアップ型である。これは行政地区単位といったエリア内部の、一部の小領域に関して全人数を観測したデータが存在することを前提とする。このアプローチでは、全人数を観測できている小領域および、小領域ごとの土地特性、隣接する小領域間の関係性などを用いて、人口統計情報に矛盾しない形で未知の小領域内の人数を推定する手法が取られる[9]。二つ目がトップダウン型である。このアプローチではまず、人口統計情報と補助情報を組み合わせ、補助情報に応じて人口統計情報を取得したエリア内の小領域に重み付けを行う。その後、人口統計情報を重みに応じてエリア内の各小領域に割り当てる。最も単純な手法は面積按分法[10]であり、人口統計情報を小領域の面積に応じて割り当てるものである。人口統計情報と組み合わせる補助情報として、都市部や居住地といった小領域の性質[5]、人工衛星で撮影した夜間照明から推定した人口密度[6]などが用いられる。また近年では、複数の補助情報を同時に考慮した手法も提案されている。例えばChengら[7]は、先に述べたような補助情報を複数利用し、深層学習モデルを用いて重みを計算することで高精度な推定を実現した。

本研究はトップダウン型の手法を用いて、動的な建物滞在人数を推定する研究に分類される。トップダウン型の手法は原理上、補助情報として用いるデータの持つ空間解像度よりもさらに詳細な小領域には妥当な推定ができない。また先行研究では、建物単位のような空間的に極めて高解像度な補助情報を取得で

表 1 本論文で用いる記号

記号	定義
M	地域メッシュ全体の集合
U	GPS ユーザ全体の集合
A	ユーザの性年代別属性全体の集合
B_m	メッシュ m における建物全体の集合
I_u	ユーザ u の一日分の測位点系列
i_u	ユーザ u の任意の測位点
N_u	ユーザ u の一日分の測位点系列の長さ
$p_{m,t,a}$	時刻 t ・メッシュ m 内の、性年代属性が a である全人口
$v_{m,t,a}$	時刻 t ・メッシュ m 内の、性年代属性が a である GPS ユーザ数
$p_{b,t,a}$	時刻 t ・建物 b 内の、性年代属性が a である滞在人数
$v_{b,t,a}$	時刻 t ・建物 b 内の、性年代属性が a である GPS ユーザ数
$c_{t,a}$	時刻 t ・性年代属性 a の拡大推計パラメータ

きていたために、建物単位での推定はほとんど行われていない。本研究は先行研究の手法を踏襲しつつ、補助情報として空間的に極めて高解像度なデータである個人移動軌跡データを用いることで、建物滞在人数の推定を実現する。

3 予備分析

本章では、本研究で扱うメッシュ統計人口データと個人移動軌跡データについて、その概要を述べた後にそれらの関係性について分析する。またデータの説明にあたり、本稿で用いる記号を表 1 に示す。

3.1 研究に用いたデータ

メッシュ統計人口

メッシュ統計人口とは、一辺が数十から数百メートルの四角形の領域内の全人口を数時間おきに取得することができるデータである。例えば「午前 7 時から午前 8 時にかけてエリア B には 300 人の人間がいた」というものである。これらの多くは携帯電話の位置情報から作成されており、代表的なデータとしてドコモ・インサイトマーケティング社の「モバイル空間統計」などがある。本研究では、メッシュ統計人口として「モバイル空間統計」を用いた。このデータは携帯電話のつながる仕組みをもとに作成されており、一辺が 500 メートル程度の領域単位で領域内の一時間ごと・性年代別の全人数を把握可能である。また、モバイル空間統計の集計単位は日本政府が定めた 2 分の 1 地域メッシュ[13] にならって定義されている。よって以後はこの領域を 2 分の 1 地域メッシュと呼ぶ。

個人移動軌跡データ

個人移動軌跡データとは時系列順に整列した測位点の集合を指す。測位点とは緯度・経度とそれが記録された時刻から成る 3 つ組のデータである。本研究では個人移動軌跡データとして、国内の各種アプリケーションから収集された GPS データを用いた。また、個人移動軌跡はユーザごとに一意な ID を割り当てられており、個人を識別することが可能なデータである。本稿では、任意のユーザ u の一日分の移動軌跡を I_u 、この中の測位点の数を N_u とする。また、任意の測位点 i_u は、測位時刻及び緯度・経度情報 p_i の情報を持つ。さらに、一つ前の測位点

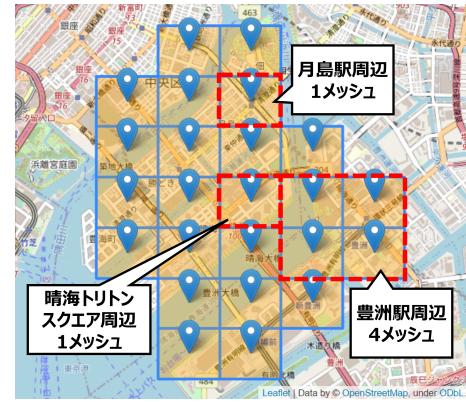


図 2 実験対象エリア

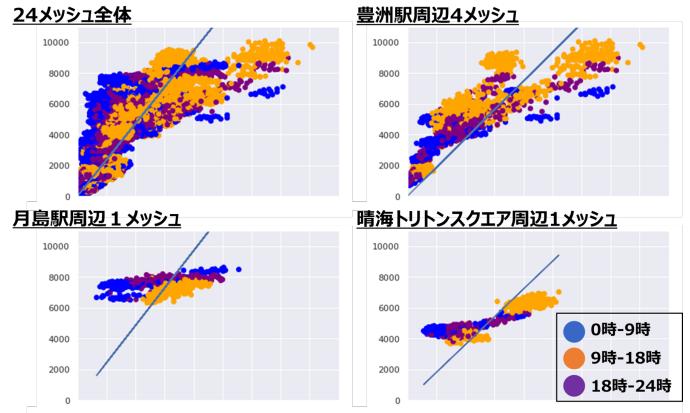


図 3 メッシュ統計人口と GPS ユーザ数の相関関係を分析。

ただし図中の回帰直線について、二つのデータが比例関係にあるかを分析するために切片は 0 とした。

の測位時刻からの経過時間を測位点 i_u における測位間隔と呼ぶ。また、個人移動軌跡はメッシュ内的一部のユーザのみから取得可能であり、これだけではメッシュ内の全人数を観測することはできない。以降では、個人移動軌跡が観測可能なユーザの数のことを GPS ユーザ数と呼ぶ。

データ収集期間・エリア

本稿における実験では、2021 年 1 月の一ヶ月間に取得された個人移動軌跡データ、モバイル空間統計データを用いた。これらのデータは図 2 に示す 24 メッシュ分の領域から取得したものである。

3.2 2 分の 1 地域メッシュ単位でのメッシュ統計人口と GPS ユーザ数の相関の調査

実際のデータにおける、メッシュ統計人口と個人移動軌跡間の関係性を明らかにするため、2 分の 1 地域メッシュ単位でメッシュ統計人口と GPS ユーザ数の間に成り立つ関係を分析する。

2 分の 1 地域メッシュレベルでメッシュ統計人口と GPS ユーザ数の相関を確認した結果を図 3 に示す。一つのデータ点はある一時間における 2 分の 1 地域メッシュ内のメッシュ統計人口と GPS ユーザ数を表しており、縦軸が GPS ユーザ数、横軸がメッシュ統計人口である。ただし、GPS ユーザ数に関し

ては機密保護の観点からマスク処理を施しており、縦軸と横軸のスケールは二つのグラフで統一した。また、時間帯ごとの傾向を確認するため、データを青(0~9時)、橙(9~18時)、紫(18~24時)に塗り分けた。

図3の左上図が2021年1月分のデータを24の2分の1地域メッシュぶん全てプロットしたもの、右上図が図3上に示した豊洲駅周辺の4メッシュについてプロットしたものである。またこの4メッシュの内部には、後に手法の評価に用いる建物の全滞在人数を観測したビル(評価用ビル)が含まれる。

左上図について、メッシュ統計人口とGPSユーザ数間の相関係数は0.775であり、これらのデータ間には強い相関があることが確認できる。また、ややはらけてはいるもののメッシュ統計人口とGPSユーザ数は概ね比例関係にあることが確認できる。次に右上図について、0~9時の時間帯にやや相関が崩れている様子が確認できるものの、こちらの場合もメッシュ統計人口とGPSユーザ数は概ね比例関係にある。これらの二つのグラフを比較した場合、24メッシュ全てを利用した場合に比べて評価用ビル周辺のメッシュを利用した方がグラフの傾きが小さいことが分かる。この結果から、24メッシュ全てを利用して求めた拡大推計パラメータをもとに拡大推計を行った場合、滞在人数の推定値が正解より多く推定される可能性が高いことが分かる。

次に図3の下段の図について述べる。左下図、および右下図はそれぞれ図2に示す月島駅周辺メッシュ、晴海トリトンスクエア周辺メッシュ内のメッシュ統計人口とGPSユーザ数の相関を分析したものである。まず左下図について確認すると、24メッシュ全体を確認したときの傾向とは異なり、月島駅周辺メッシュにおいてはGPSユーザ数とは関係なくメッシュ統計人口がほぼ一定であることがわかる。月島駅周辺メッシュの特徴として、「月島もんじゃストリート」という観光地が含まれていることがある。橙色の点が図のやや右寄りに固まっていることから、深夜帯に月島駅周辺メッシュに滞在していたユーザに入れ替わりで、日中に観光目的で月島駅周辺メッシュに訪問するユーザが流入しているため、メッシュ統計人口は変動しないと推測できる。次に右下図について、こちらは深夜帯は人が少なく、日中は人が多いという24メッシュ全体の傾向から外れてはいないが、切片が見られるケースである。この原因として、晴海トリトンスクエア周辺メッシュ内には大規模なオフィスビルと居住地が存在することが挙げられる。日中は通勤者によりメッシュ統計人口が増加するが、深夜帯にも居住者が一定数存在するためにメッシュ統計人口が一定以下に減少しないと考えられる。

このように、24メッシュ全体を見ればメッシュ統計人口とGPSユーザ数は概ね比例関係にあり、豊洲駅周辺のように比較的オフィスビルの占める割合が高いメッシュも同様の傾向を示すことが確認できる。一方で、観光地が含まれるメッシュや居住地の割合が高いメッシュについては、メッシュ統計人口とGPSユーザ数が比例関係にあるとは言えない場合がある。このことは、メッシュの性質ごとに適切な拡大推計パラメータの算出方法を変える必要があることを示唆する。

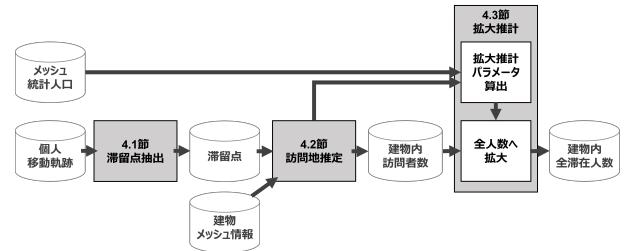


図4 概要図

4 提案手法

メッシュ統計人口と個人移動軌跡データを入力として、建物内の滞在人数を推定する手法を提案する。提案手法の概要を図4に示す。また、提案手法の各処理の詳細について以降の節で順に述べる。

4.1 滞留点抽出

滞留点とは、居住地 → 職場 → 商業施設 → 居住地などといったユーザの一連の移動の中で、長時間特定の箇所に留まる測位点の集合を指す。本研究では、滞留点抽出手法として西田ら[12]が提唱したD-Starを用いる。

D-Starのアルゴリズムについて述べる。まず、ユーザ u の一日分の移動軌跡を I_u 中の各測位点 i_u に対し、移動窓 q 内かつ距離が ϵ 内にある測位点を滞留候補点として取得する。各測位点ごとに得られた滞留候補点に対し、滞留候補点中の先頭の測位点から末尾の測位点までの間の経過時間がクラスタ形成判定閾値 δ 以上であれば、その測位点をコアポイントとして採用する。また、コアポイントとして採用された測位点に紐づく滞留候補点を、コアポイントを中心とした一つの滞留クラスタとして定義する。次に別の移動窓において採用された滞留クラスタ同士が同一の測位点を含む場合、これらのクラスタをマージする。これにより、コアポイントごとに定義された滞留クラスタが結合され、大きな滞留クラスタが生成される。最後にマージ後のクラスタ内における経過時間が滞留判定時間閾値を超える場合、これを滞留点として採用する。こうして取得された滞留点は、ユーザID、滞留点の重心、滞留開始・終了時刻を持つ。

しかし、D-Starを用いて滞留点抽出を行う場合にうまく滞留点が取得することができないユーザが存在する。そこで、このようなユーザに対しても滞留点を適切に取得するための前処理について述べる。

滞留点が抽出できないのは、以下の三つのパターンのいずれかに該当するユーザである。

- 観測できた測位点が1点のみのユーザ
- 測位点集合 I_u 内の経過時間がクラスタ形成判定閾値 δ 未満のユーザ
- 測位間隔が極端に短いユーザ

この三つのうち、前の二つのパターンに属するユーザからはD-Starではどうしても滞留点を抽出することができない。よつ

てこのようなユーザはノイズデータとしてあらかじめ削除する。三つ目について、D-Star は固定の移動窓 q 内の経過時間が、クラスタ形成判定閾値 δ を越えた場合に滞留判定を行うため、測位間隔が極端に短い場合は滞留点が抽出されづらくなる。そこで本研究では、測位点を一定間隔ごとにサンプリングすることで、測位間隔のばらつきを小さくすることで対処することとした。

4.2 訪問建物推定

訪問建物推定のアルゴリズムについて述べる。訪問建物推定とは、4.1 節で取得した滞留点と建物を紐づける処理である。滞留位置をもとに滞留点を建物に紐づけることで、どの人物が、どの建物に、いつからいつまで滞留したかという情報を取得することができる。訪問建物推定には様々な先行研究 [14] [15] があるが、本研究ではシンプルなアルゴリズムでの推定を実施した。前提として、全ての建物は自身の位置および、平面方向の形状を情報として持つものとする。以降ではこの情報を建物メッシュと呼ぶ。まず、滞留点の座標を中心として、半径 20 メートルの円を定義する。次に、この円と建物メッシュが重なる建物を訪問建物候補として取得する。最後に、訪問建物候補の中で最も滞留点の座標と距離が近い建物を、滞留点に訪問建物として紐づける。この手法では、半径 20 メートル以内に建物が存在しない場合は上記のアルゴリズムでは訪問建物を紐づけることができない。このような滞留点に対しては訪問建物不明 (unknown) ラベルを割り当てる。

4.3 拡大推計

拡大推計とは、訪問建物推定結果、地域メッシュ内の個人移動軌跡データ、地域メッシュ統計人口の三つを入力として、ある建物内の一時間ごとの全人数を推定する処理を指す。拡大推計は、拡大推計パラメータの算出、人数を推定したい建物内の GPS ユーザ数の算出、全人数への拡大の 3 ステップから成る。

拡大推計パラメータの算出

拡大推計パラメータとは、GPS ユーザ数を何倍すればメッシュ統計人口と等しくなるかの係数を指す。時刻 t におけるメッシュ m 内の性年代属性が a であるメッシュ統計人口を $p_{m,t,a}$ 、同じく時刻 t におけるメッシュ m 内の性年代属性が a である GPS ユーザ数を $v_{m,t,a}$ とおくと、拡大推計パラメータは以下の式により定義される。

$$c_{t,a} = \sum_{m \in M} p_{m,t,a} / v_{m,t,a} \quad (1)$$

建物内の GPS ユーザ数の算出

推定対象の建物に滞在した性年代ごとの GPS ユーザ数 $v_{b,t,a}$ を取得する方法について述べる。4.2 節で述べたように、訪問建物推定結果は、滞留開始・終了時刻、及び訪問建物を持つ。まず訪問建物推定結果の中から、推定対象の建物と紐づいたデータのみを抽出する。その後、滞留開始時刻から滞留終了時刻まで、一時間ごとに建物内の滞留人数に一人を加算する処理を抽出した全ての訪問建物推定結果に対して行うことで、建物内に

滞在する性年代別・一時間ごとの人数を取得する。また、一時間のウインドウごとに集計した際の、ウインドウ内の滞留時間が 60 分未満の場合は、そのウインドウ内の滞留人数は（ウインドウ内の滞留時間 [分]) / 60[人] として計算する。

全人数への拡大

最後に、拡大推計パラメータと建物内に滞在した GPS ユーザ数を突合し、集計時間、性年代が合致するデータ同士を乗算した後に性年代について集計することによって、建物内の滞在人口を推定する。ビル内の一時間ごとの全滞在人数 $p_{b,t}$ は以下のように性年代別に定義した拡大推計パラメータと同じく性年代別のビル内 GPS ユーザ数 $v_{b,t,a}$ の積を、性年代別に集計した和となる。

$$p_{b,t} = \sum_{a \in A} c_{t,a} \times v_{b,t,a}$$

5 実験

本章では建物滞在人数を推定し、推定結果と正解データを比較することで手法の有効性を検証する。また、4.1 節で述べた個人移動軌跡データの前処理の有無、および 4.3 節で述べた拡大推計パラメータ算出時に使用する 4 次標準メッシュの数を変化させることによる、推定結果の変化についても検証する。

5.1 実験設定

入力データ

図 3 上に示した豊洲エリア周辺 4 メッシュ内のモバイル空間統計および個人移動軌跡を用いた。ただし、5.4 節の実験のみ、図 3 に示す 24 メッシュ分すべてのモバイル空間統計・個人移動軌跡を用いた。

正解データ

評価用ビルの入退館データをもとに 1 時間ごとの建物内滞在人数データを作成し、正解データとして用いた。

評価手法

評価指標として、推定誤差を評価する一般的な指標である MAE (Mean Absolute Error) および MAPE (Mean Absolute Percentage Error) を用いた。これらの指標は値が小さいほど推定結果が正しいことを示す。MAE は絶対的な誤差の平均を示し、MAPE は正解値に対する誤差の割合の平均を表す。この二つの指標を比較すると、例えば人数が少ない時間は MAE が小さくとも、MAPE が大きな値を示す傾向がある。

ベースライン手法

提案手法と比較するベースラインについて述べる。本稿ではベースライン手法として、2 章で述べたトップダウン型のアプローチで用いられることがある、メッシュ内の小領域の面積に基づいてメッシュ統計人口を按分する手法（面積按分法）[10] を採用した。具体的には、ある一時間におけるメッシュ統計人口を建物ごとの延べ床面積（床面積と階数の積）で按分するこ

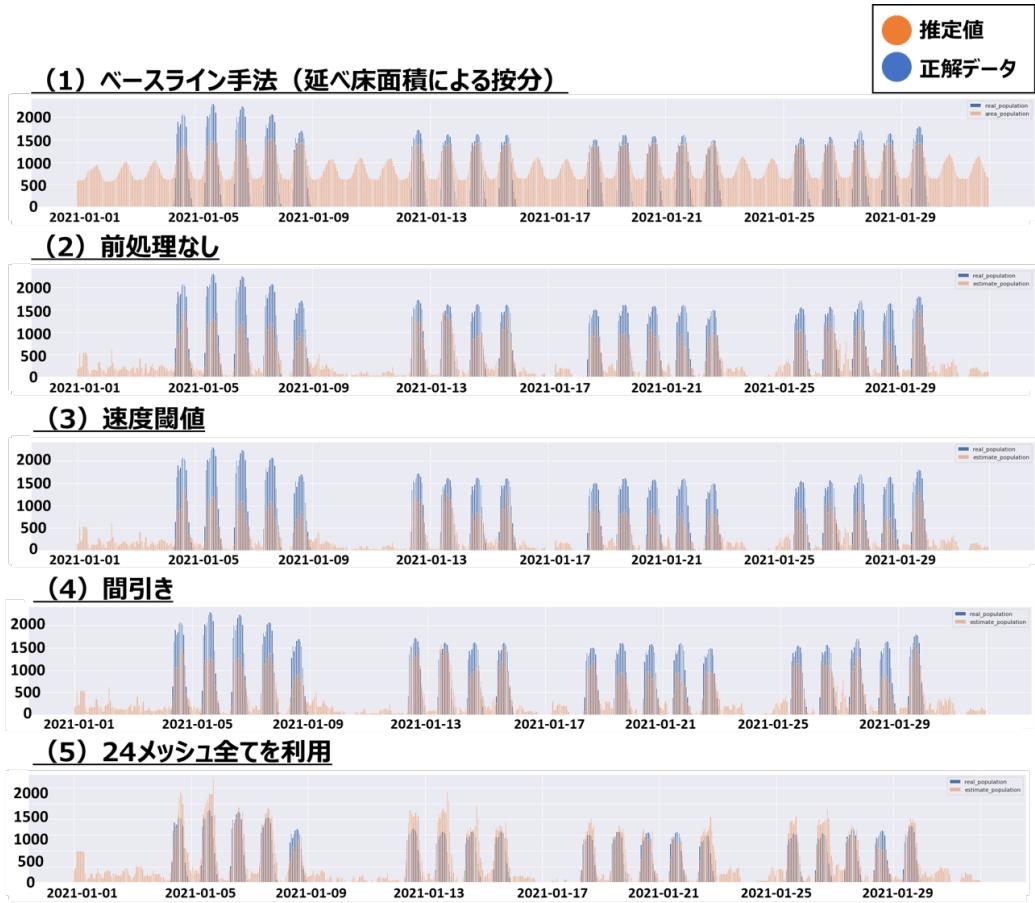


図 5 拡大推計結果の比較

表 2 個人移動軌跡データの前処理方法による定量評価結果の比較

	ベースライン		前処理なし		速度閾値		間引き		24 メッシュ	
時間帯	MAE	MAPE	MAE	MAPE	MAE	MAPE	MAE	MAPE	MAE	MAPE
平日全体	454.58	223.78	330.95	80.75	376.39	76.09	313.23	83.17	323.63	135.99
平日 (0–9 時)	654.06	938.74	150.58	154.24	143.07	135.44	155.76	166.56	186.10	335.46
平日 (9–18 時)	236.47	13.29	598.56	37.19	712.10	44.44	545.53	33.94	309.97	19.90
平日 (18–24 時)	494.78	424.16	170.22	141.21	182.24	118.52	170.72	151.45	513.95	296.27

とで、評価用ビル内に滞在する一時間ごとの全人数を推定した。

5.2 ベースライン手法との比較

モバイル空間統計を建物ごとの延べ床面積で按分するベースライン手法（図 5(1)）と個人移動軌跡から訪問建物を推定し、拡大推計を行った手法（図 5(2)）の比較を行った。ベースライン手法と比べて提案手法は、休日・夜間帯の推定精度が大きく改善したことが分かる。これは個人移動軌跡から建物への訪問人数を推定したこと、現実の人流をより正確に反映した推定ができたためと考えられる。

5.3 個人移動軌跡データの前処理による滞留点抽出結果の変化

個人移動軌跡データから滞留点を正しく抽出することを目的として、前処理による拡大推計結果の変化について検証を行った。ただし、測位点数が一定以下の場合はそもそも D-Star による滞留点の抽出が困難であるため、測位点数がある程度多い

ユーザに絞って検証を行った。

4.1 節で述べたように、測位間隔が極端に短い場合、D-Star では滞留の抽出が困難になる。この問題に対処する方法としては、以下の 2 つのアプローチが挙げられる。

(1) クラスタ形成閾値を時間閾値から速度閾値に変更

(2) 個人移動軌跡データに対して間引きを行う

前者の手法は、測位間隔に応じてクラスタ形成閾値を動的に変更できれば、測位間隔が短い場合でも滞留を抽出できるという考えに基づく。クラスタ形成閾値を経過時間から速度に変更することで、経過時間が短くとも移動距離が短ければ滞留クラスタ化するようになると考えた。後者の手法は 4.1 節に述べたとおりである。

前処理を行わなかった場合および、上記の二つの手法をそれぞれ適用して拡大推計を行った結果を図 5 (2)(3)(4) に示す。また、推定結果を平日・休日、および時間帯 (0–9 時、9–18 時、18–24 時) に分割し、定量評価を行った。この中でも商圏分析

などにおいて重視される平日の評価結果を表2に示す。

図5に示す通り、定性的には大きな変化を確認することはできない。一方で、表2で定量評価結果を確認すると、平日全体で見た場合はMAE, MAPEともに間引きを行った場合が最も優れていることが分かる。また、間引きによる結果の中でも特に、平日(9~18時)の時間帯における推定精度向上が著しい。速度閾値を用いた場合は、早朝・夜間の時間帯は最も推定精度が高いが、平日全体で見るとやや推定精度が悪化した。この原因を考えると、まず速度閾値を用いた場合は、測位位置に誤差が生じると移動速度が閾値を超えてしまい滞留クラスタ化しやすくなる傾向があり、一方で早朝・夜間は端末の移動が少ないために測位位置の誤差が生じにくく、結果として推定精度が向上したためであると考えられる。よって以降の実験では、個人移動軌跡データに対して間引きを適用したものを入力として用いた。

5.4 拡大推計パラメータの算出方法による推定結果の変化

最後に、拡大推計パラメータの算出時に参照するメッシュの数を変化させた場合の拡大推計結果の変化について検証した。入力として豊洲駅周辺の4メッシュのみを利用したものを図5(4), 24メッシュ全てを利用したものを図5(5)に示す。

これらを比較すると、図5(5)の方が推定精度が高い。3章におけるモバイル空間統計人口とGPSユーザ数の関係性分析から、メッシュごとに適切な拡大推計パラメータが異なることが予想される。これを踏まえると、人数を推定したい建物が含まれる豊洲駅周辺のメッシュを用いて拡大推計パラメータを算出したほうが妥当であるように思われるが、今回は逆の結果となった。今回正解データが取得できた評価用ビルでのみ、偶然推定がうまくいった可能性もあるが、現状では、正解として利用できる建物内滞在人数データが一種類であるため、検証することができない。

一方で、図3を参照する限り、図5(4)の推定値は図5(5)の8割程度になるのが妥当である。しかし、結果は6割強程度の推定値となっている。これは、評価用ビル周辺のメッシュのみを用いて計算した拡大推計パラメータより、評価用ビルにおける拡大推計パラメータが大きいことを示す。この結果は、建物ごとに拡大推計パラメータが大きく異なる可能性を示唆する。

6 おわりに

本稿では、モバイル空間統計と個人移動軌跡を組合わせ、建物内の全滞在人数を推定する手法を提案した。面積按分法にもとづくベースライン手法と比較して、提案手法は平日休日や時間帯ごとの傾向をうまくとらえることができることを示した。また、実験により拡大推計パラメータの算出方法による推定結果の変化について検証を行い、拡大推計パラメータが建物ごとに異なる可能性について述べた。

一方で、今回は正解データとなる建物内の滞在人数データを一件分しか収集できなかったために、提案手法が建物でも有効に機能するのかを検証することができていない。また、建物ご

とに拡大推計パラメータが異なる可能性について言及したが、こちらも建物内の滞在人数データが一つしかないため検証できていない。今後は正解データの収集をすすめ複数の建物においても適切な拡大推計が行われるかを調査し、手法の頑健性の向上を図る。

文 献

- [1] N.Wardrop, W.Jochem, T.Bird, H.Chamberlain, D.Clarke, D.Kerr, L.Bengtsson, S.Juran, V.Seaman and A.Tatem "Spatially Disaggregated Population Estimates in the Absence of National Population and Housing Census Data." Proceedings of the National Academy of Sciences Vol.115, No.14, pp.3529–3537, 2018.
- [2] S. Wang, R. Li, J. Jiang, and Y. Meng. "Fine-Scale Population Estimation Based on Building Classifications: A Case Study in Wuhan." Future Internet, Vol. 13(10), 251, 2021.
- [3] P. Deville, C. Linard, S. Martin, et al. "Dynamic Population Mapping using Mobile Phone Data." Proceedings of the National Academy of Sciences, Vol. 111, No. 45, pp. 15888–15893, 2014.
- [4] M. Terada, T. Nagata, and M. Kobayashi. "Population Estimation Technology for Mobile Spatial Statistics." NTT DOCOMO Technical Journal, Vol. 14, No. 3, pp. 10–15, 2013.
- [5] H. Bast, S.Sabine, and W.Simon, "Fine-grained population estimation." Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems, 2015.
- [6] S. Amaral, G.Camara, A. Monteiro, J. Quintanilha and C.Elvidge , "Estimating Population and Energy Consumption in Brazilian Amazonia using DMSP Night-Time Satellite Data." Jurnal of Computers, Environment and Urban Systems, Vol.29, No.2, pp.179–195, 2005.
- [7] L. Cheng, L. Wang, R. Feng and J. Yan. "Remote Sensing and Social Sensing Data Fusion for Fine-Resolution Population Mapping with a Multimodel Neural Network." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Vol. 14, pp.5973–5987, 2021.
- [8] L. Xiang and P. Pöllmann. "Dynamic Population Estimation Using Anonymized Mobility Data." Proceedings of the 28th International Conference on Advances in Geographic Information Systems. 2020.
- [9] E.Weber, V.Seaman, R.Stewart, T.Bird, A.Tatem, J.McKee, B.Bhaduri, J.Moehl and A.Reith. "Census-independent population mapping in northern Nigeria." Journal of Remote Sensing of Environment, Vol.204 ,pp.786–798, 2018.
- [10] 貞広幸雄. "空間集計データにおける面補間法の推定精度評価." Journal of Mathematica japonicae, Vol. 52, No.3 pp. 359–376, 2000.
- [11] OpenStreetMap Japan. available from <https://openstreetmap.jp/>, (Accessed 2022-02-10).
- [12] K. Nishida, H. Toda, and Y. Koike. "Extracting Arbitrary-Shaped Stay Regions from Geospatial Trajectories with Outliers and Missing Points." Proceedings of the 8th ACM SIGSPATIAL International Workshop on Computational Transportation Science, 2015.
- [13] 総務省統計局地域メッシュ統計. available from <http://www.stat.go.jp/data/mesh/>, (Accessed 2022-02-09).
- [14] 西田 京介, 戸田 浩之, 倉島 健, 内山 匠. "確率的訪問POI分析: 時空間行動軌跡からのユーザモデリング." マルチメディア, 分散協調とモバイルシンポジウム 2013 論文集, Vol.2013, pp-334-345, 2013.
- [15] D.Lian and X.Xie. "Learning Location Naming from User Check-in Histories." Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic

