拡大係数再付与による PT トリップチェイン情報の更新について

井戸川　昌樹[1]・Khin　INGYIN[2]・塚井　誠人[3]

[1]学生会員　広島大学大学院　先進理工系科学研究科（〒739-8527 広島県東広島市鏡山 1-4-1）
E-mail: m234136@hiroshima-u.ac.jp

[2]学生会員　広島大学大学院　先進理工系科学研究科（〒739-8527 広島県東広島市鏡山 1-4-1）
E-mail: m211389@hiroshima-u.ac.jp

[3]正会員　広島大学大学院　先進理工系科学研究科（〒739-8527 広島県東広島市鏡山 1-4-1）
E-mail: mtukai@hiroshima-u.ac.jp

都市交通計画の基幹調査であるパーソントリップ調査（PT）には，新型コロナウイルスなどの交通需要の急な変化を適切に把握できないという課題がある．近年注目されているモバイル空間統計（MSS）には高い時空間解像度と速報性があるが，実務的な信頼性に問題がある．本研究では，熊本都市圏を対象に，PTとMSSを組み合わせてトリップチェインの拡大係数を更新する方法を試みる．具体的には，PTから得られるトリップチェインを時間帯別，地点別，滞在数として記録されるMSSの形式に変換し，Lasso回帰を用いて拡大係数を算出する．滞在人口モデルの決定係数が比較的高かったことより，過去（2012年）のPTトリップパターンによって最近のMSSの滞在人数をほぼ表現することができた．

Key Words : *Mobile Spatial Statistics, data fusion, lasso regression*

1. はじめに

都市交通計画の基幹調査であるパーソントリップ調査（以下，PT）は，抽出率が低く調査間隔が長いため，新型コロナウイルスなどによる急激な交通需要の変化がを把握できない．そこでこれを補完する交通調査として近年，時空間解像度と抽出率が高く，速報性を有する携帯電話位置情報を利用したモバイル空間統計（以下，MSS）が注目を集めている．MSSにはいくつかバリュエーションがあるが，地点別滞在人数の速報性が最も高い．よってこの形式の MSS と PT の長所を生かして，相補的な利用法を検討する意義は大きい．

本研究では，熊本都市圏を研究対象地域として，PT から得られるトリップチェインに着目して MSS で取得されている周辺分布を参照して各トリップの拡大係数を再付与，つまり更新する手法を検討する．具体的には，PT から抽出したトリップチェインを，MSS の時間帯，地点別の滞在記録形式に変換して MSS データに回帰することによって，各トリップチェインの拡大係数を算出する．なお回帰モデルには，制約付き重回帰分析および観測データへの過剰適合を防ぐために，Lasso 回帰を用いる．以上の分析について，モデルの精度と課題について考察する．

本論文の構成は，以下のとおりである．2. では，GPS を用いた交通調査に関する研究，都市交通計画におけるモバイル空間統計の活用可能性に関する研究，および PT から得られる OD 表更新に関する既往研究を整理して，本研究の位置づけを示す．3. では，データ概要と分析手法について説明する．4. では，制約付き重回帰分析と Lasso 回帰の概要について説明する．5. では，両モデルの精度に関する結果と課題について述べる． 6. では，結論および考察を述べる．

2. 既往研究の整理と本研究の位置づけ

(1) GPS を用いた交通調査に関する研究

PT では調査対象者の個人属性，出発／到着時間，移

動手段や活動目的などの多岐にわたる項目が調査される．これらの項目は都市交通計画のための基礎情報と位置付けられている．しかし，調査間隔が数年以上となる PT では，都市活動の短期的な変化を捉えることが難しい．他方で PT は調査コストも高いため，調査間隔の短縮も困難である．さらに回答者の記憶に頼る PT では，詳細な移動経路や乗り換え時刻を含めて調査することは容易ではない．そこでGPS携帯電話などの観測機器を対象者に貸与して，彼らの詳細な位置・時刻情報を取得するプローブパーソン調査（以下，PP）が開発された．

円山ら[1]はPPの課題として，①機器の貸与を前提としているため対象者が限定され，大規模調査の実施が容易でないこと，②通信費，機器レンタル費がかかることなどを指摘した上で，参加者自身が保有するスマートフォンを利用して，PT 調査と連携した大規模調査を行った．これは当時は世界的にも例が少ない貴重な調査であり，約 1000 人が参加した．スマートフォンの所有率が低いと予想される年齢層のデータを取得するため，希望者にタブレット端末を貸し出すことによって，60 歳以上のサンプルを全体の 11%確保した．円山らは，一連の調査手法によりサンプルの偏りを抑えて，比較的大規模な調査が可能なことを示した．ただし円山の手法は，自ら述べた①と②の課題に対しては調査用のソフトウェアを自作することで部分的に対処したのみであり，調査コストの問題が解消したわけではない．さらに調査への参加が自発的なことも，調査結果のバイアスに留意すべき理由となっている．

参加者のスマホを調査に利用する際に生じるバッテリー消費の問題に関して，野原ら [2]は，Bluetooth 通信を用いた手法では，従来の PP よりも消費電力が減少したと述べた．山崎ら[3]は，つくば市で実施したPP被験者に事後アンケート結果を実施した．この結果から，持続的な調査実施のために，被験者の募集時に生じるサンプルの偏りや実質的なサンプル数が少ないことへの対処，および調査時のスマートフォンの電源の入れ忘れ防止などの，改善点を指摘した．

GPS 調査は，従来の PT よりも動的なデータが大量に得られる一方，約 1 分ごとに位置情報を取得するため見かけ上のデータ数は膨大でも，トリップベースに加工すると実質的なサンプル数が少ない．他方でGPS調査は，プライバシーへの懸念から調査の許諾数を増やしにくい．これらの点を考慮すると，GPS 調査には詳細なトリップが把握できるメリットがある一方で，直ちに PT の代替手法とすることは，現実的に困難と考えられる．

⑵　モバイル空間統計の活用可能性に関する研究

NTT ドコモ社が提供するモバイル空間統計（以下，MSS）には，発着地が判明する人口流動統計と滞在地のみ判明する人口分布統計の二種類存在する．

寺田ら[4]は，都市交通計画におけるMSSの実用化に向けて，空間解像度の信頼性を検証した．その結果，空間解像度は，携帯電話基地局の設置密度と強い関係があり，対象となるエリアが広いほど，また人口密度が高いところほど信頼性が高いことを示した．

清家ら[5]は，都市交通計画におけるMSSの人口分布統計の時間解像度の信頼性を PT との比較によって検証した．その結果，両データの滞在人口がほぼ一致したゾーンが見られるので，一定の信頼性が得られるとした．一方で，人口分布統計の滞在人口は滞在者と移動者を含むため，昼間の時間帯の人口分布統計は，PT の滞在人口よりも 20%多く推計されるゾーンがあった．当該ゾーンは，多くの移動者の存在が想定される幹線道路沿いを含んでいる．データの観測特性に留意すべきと指摘した．

室井ら[6]は，都市間旅客交通データとしてのMSSの活用可能性を調査するために，MSS の空間網羅性や速報性を国勢調査および PT と比較して，それぞれの優位点と留意点をまとめた．空間網羅性に関して MSS は，広範囲なエリアにおける調査が比較的容易であること，都道府県単位や市区町村単位での滞在人口の信頼性は高いと結論付けた．その一方で，国勢調査と比較すると小地域の推計人口の精度には限界があることを指摘した．速報性に関して，MSS はデータ取得からデータ活用可能までの期間（タイムラグ）が非常に短く，優れていることがわかった．ただし，移動手段および経路，活動目的は取得できないことを指摘した．

時空間解像度や速報性に優れる MSS は都市交通計画での活用可能性が高いことが判明した．一方，MSS の滞在人口統計値は，道路や交通路上での移動者を含むことや移動手段，および移動経路が取得できないという課題があると指摘されている．

⑶　PT から得られる OD 表の更新に関する研究

今井ら [7]は，MSS の人口流動統計（OD 量）と都市交通計画の基幹調査である PT に基づき集計された OD 量を，空間解像度に着目して比較検証した．その結果，市区間をまたぐ広域の移動の相関が高いが，分析対象エリアが詳細の移動の相関が低いことを示した．

国土交通省が 5 年に一度全国で実施している全国都市交通特性調査（以下，全国 PT）は，全国の都市における人の行動を統計的に把握する調査であり，都市規模別の生成原単位や交通分担率等の傾向把握を目的としている．ただしこの調査は，各都市で 30 地区の調査区を選定して行われるため，全地域の居住者のサンプルは得られないという限界がある．

菊池ら[8]は 2015 年に取得された全国 PT 調査，PT および MSS の人口流動統計を補完的に使用する方法について検討した．具体的には，全国 PT から部分的に推計した OD 交通量と MSS の人口流動統計から得られる OD 表を統合して得られる都市圏全体の推計データを，PT から得た OD 表との差が小さくなるように補正した．推計した OD 交通量を PT から得た OD 表と比較した結果，ゾーン内外では相関係数は高いが，ゾーン内々では両データの相関が低いことが確認された．

越智ら[9]は，PT から一部抽出した OD 表と人口流動統計を用いて推計した OD 表と PT の OD 表を比較した．目的別手段別 OD 量の相関係数は 0.8 以上と高かったものの，一部の目的別 OD 量の推計結果では相関係数が低く，推計精度の向上が求められる，と指摘した．

⑶ 本研究の位置づけ

PT は調査費用が高いことやリアルタイムデータが得られない等の課題がある．アプリを使った GPS 調査では，携帯端末を保有しない者を捕捉するのに機器を貸与するための追加費用がかさむ．また GPS はプライバシーの問題への懸念から多くの調査協力者が得にくい．以上より，GPS のみでは PT の課題解決は難しい．

MSS の活用可能性に関する研究では，交通ゾーン単位での時空間の信頼性の検証が進められる一方で，その利用に多くの留意点があることがわかった．また MSS の人口流動統計を用いて PT の OD 表を更新する手法では，一部のエリアでは高い精度で推計できていた．一層の精度を求めるならば，トリップチェイン情報を参照した OD 表更新手法の開発が考えられる．なおトリップチェインが把握できる人口流動統計は費用が高く，計算処理に時間がかかるため，実務で活用する上では，困難な点が残る．一方でコストが低く速報性が高い MSS の人口分布統計と PT のトリップチェインを補完的に活用する研究例は，これまでは見られなかった．

本研究では，取得年次が異なる PT と MSS を相補的に利用する．具体的には，PT から取得されるトリップチェインの拡大係数を，MSS で取得される周辺分布に合致するように算出することで，両データの補完的な利用方法を検討する．一連の分析から，過去の PT から得たトリップチェインによって表現可能な最近の MSS のトリップパターンを明らかにできる．またモデルで表現できない残差を明らかにすることによって，新たなトリップパターンの出現を検討する．本研究の新規性は，年次の異なる両データを相補的に用いて推計することで，PT の更新の可能性を検討することである．



図 1　分析対象地域（熊本都市圏）

3. データと分析手法

⑴ 使用データの概要

図 1 に分析対象地を示す．使用データは，2012 年（平成 24 年）に実施された第 4 回熊本都市圏 PT 調査と，携帯電話利用者の位置情報から各地点の滞在者数の時空間分布を推定した MSS である．後者のデータとして，人口分布統計を用いる．

PT の主な調査項目は，対象者の 1 日の移動と活動目的である． PT のゾーン単位で編成される出発/到着地と移動交通手段，および各滞在地における主な活動目的などを時系列で並べたトリップチェインを得る．さらに，対象者の性別，年齢，通勤・通学先などの個人情報も付帯する．また今回使用した PT は，熊本県が平成 24 年に実施した熊本都市圏の C ゾーン間流動である．この調査では，約 12 万世帯に調査票を配布し，有効票として約 4.3 万世帯（約 9.7 万人分）のサンプルを取得した．抽出率は 6.8～8.6%である．調査後に都市圏単位の OD 表を推計するために，補正処理を経て各サンプルに拡大係数が付される．サンプルの抽出率がほぼ各層で等しいなら，各サンプルの拡大係数の期待値は抽出率の逆数であり，だいたい 11.63～14.71 と見込まれる．

MSS は，2013 年に株式会社 NTT ドコモ（以下，ドコモ）が事業化した人口統計情報である．これは，ドコモの携帯電話の各基地局下に存在する携帯電話在圏情報から作成される人口統計情報である．1 時間ごとに滞在人数を，24 時間 365 日把握できる．その空間解像度は 4 次メッシュ（500m）であり，時間解像度は 1 時間である．調査対象は 15～79 歳までの男女であり，ダブルカウントを避けるため，法人名義の契約データ等は除去されている．また個人情報保護の観点から，推計人口が少人数（9 人以下）のエリアについては，「秘匿処理」が行われる．MSS が網羅する国内携帯保持者は約 8000 万台（2020 年 3 月現在）であり，同データに含まれる訪日外国人が国内滞在中にローミングで接続する携帯電話の延べ台数は，年間約 1200 万台（2019 年実績）である．以上の携帯保持者の規模が示すように，MSS は国内最大

級の人口統計情報である．

⑵　分析手法

PT と MSS を照合して，両データの異同を明らかにするには，両データに共通する年齢および性別ごとのデータの全てについて，時空間解像度をそろえる必要がある．MSS は携帯電話の普及率が低い 80 歳以上や，サービスの契約ができない 14 歳以下の滞在人数が推計対象外である．そこで本研究の対象年齢層は 15～79 歳とする．15～19 歳は 1 区分として，それ以外の 20～79 歳は 10 歳階級とすると，性別と合わせて個人属性は 14 区分とする．時間解像度は MSS フォーマットに合わせて，1 時間とする．

データ取得月は，PT が 2012 年 10 月，MSS が 2019 年 10 月として，季節ごとの交通需要の違いによる影響を最小限にする．両データの異同比較を目的とするなら，PT と MSS の取得年次は一致している方が望ましいが，熊本都市圏 PT は 2012 年 10 月分実施の一方で，MSS が最初に整備された年次は 2014 年のため，同時点のデータは入手できない．本アプローチでは，過去の PT から得たトリップチェインによって表現可能な最近の MSS のトリップチェインを明らかにする．

調査時間は人の一日を把握するため，0:00～23:59 に統一する．なお PT の調査時間は 3:00～27:00 である．MSS と統合利用するために，0:00～3:00 の滞在場所について 0:00～第一トリップの開始時刻までと最終トリップ到着時刻から 23:59 までを自宅滞在と仮定する． また自宅のあるゾーンと第一トリップで出発したゾーンや最終トリップで到着したゾーンが異なっていると，滞在場所を正確に把握できない．そこで以下の分析では，調査時間補完対象トリップは，自宅に発着するホームベースのトリップチェインのみを扱う．

空間解像度は PT フォーマットで最小の空間集計単位である C ゾーンを用いる．このため， 1 辺 500m の 4 次メッシュが利用可能な最小単位である MSS を，PT フォーマットに合わせて C ゾーンに集計し直す．一般に C ゾーンは 4 次メッシュよりも面積が大きく，C ゾーン境界は複数の 4 次メッシュにまたがるため，C ゾーンと 4 次メッシュは 1 対 1 で対応しない．そこで簡単のため， C ゾーン境界上の 4 次メッシュとの重なり面積が最大の C ゾーンに，そのメッシュ全体を割り当てることとした．具体例を，図 2 に示す．

本研究で用いる PT からは，97,109 サンプルが得られた．そのうちトリップ回数が 0 回の者は 11,255 サンプルなので，それらを除くと 85,854 サンプルが得られる．さらに簡単のため外出者（トリップ回数 1 回以上）の 85.8%を占めるトリップ回数 2～4 のサンプル（73,630 サンプル）を抽出し，ここからホームベースのラウンドトリップを行った者（42,848 サンプル）に焦点を当てる．次に PT から抽出するトリップチェインについて述べる．MSS でカウントされている滞在人数の一部には移動中の人が含まれている．一方で，PT から得られるトリップチェインは移動と滞在が交互に現れるが，PT フォーマットを以下の手順で集計しても移動の経路情報は処理されない．これは経路上の各地点の移動時刻は記録されていないためである．今回は簡略化のため，移動時間は滞在時間よりも十分短いと仮定して，移動中のゾーン情報を無視する．滞在ゾーンにのみ着目した PT からのトリップチェインの抽出について，図 3 と図 4 を基に示す．図 3 では，トリップ回数が 3 回のときに PT から得られる情報（4 地点の移動開始～終了時刻）と滞在ゾーンのみに着目したトリップチェインの情報（4 地点の滞在開始～終了時刻）を示している．以下の図に示す記号を，次のように定義しておく．

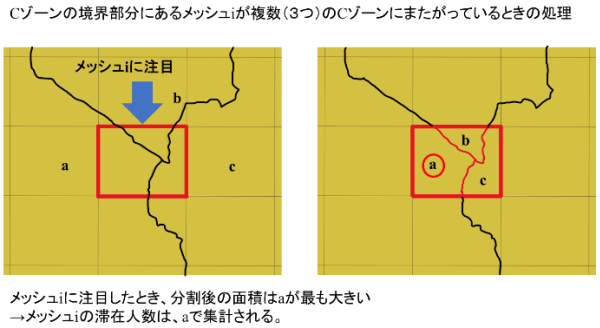


図 2　メッシュi と C ゾーンの対応

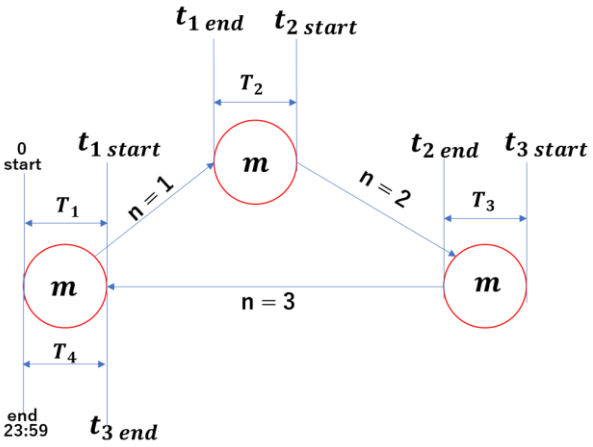


図 3　トリップチェイン抽出の模式図

n：トリップ回数
$t_{n\,start}$：n 番目の移動開始時刻
$t_{n\,end}$：n 番目の移動終了時刻
$T_N = t_{n+1\,end} - t_{n\,start}$：N 番目の滞在時間
m：滞在ゾーン
k：1 時間ごとの時間帯

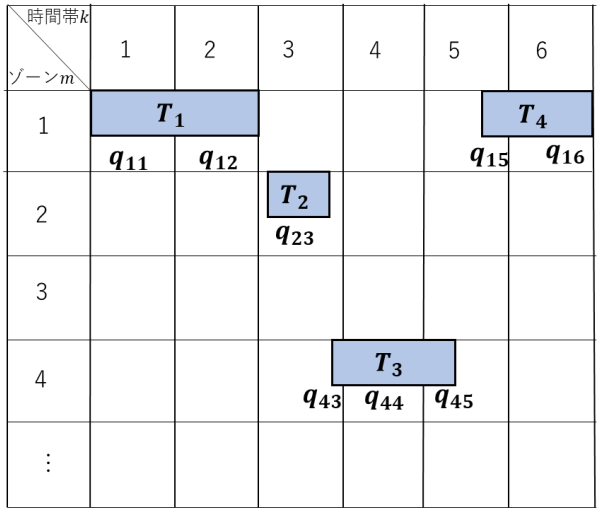


図 4　T_N とmからq_{mk}への変換

q_{mk}：滞在ゾーンmで，滞在時間T_N の部分トリップを滞在時間帯kを考慮して算出したゾーン別，時間帯別，滞在人数

p：トリップチェイン

$G^p_{mk} = q(p,m,k)\ U\ q(p,m',k')$

図 4 では，PT フォーマットのT_N とmを用いて，MSS フォーマットのq_{mk}に変換するフォーマット編成について述べる．滞在時間を 1 時間当たりに分割する．

上記の変換が示すように，滞在のないmとkの組み合わせ(m',k')は，すべて$q_{m'k'}=0$である．さらに，後述する重回帰分析の入力変数を作成する際の簡単を図るため，同時間帯区分kで，複数ゾーンmにまたがって滞在しているときは，時間帯区分kの中で最も滞在時間が長いq_{mk}の値のみ保持し，残りは 0 とする．その後，同一のトリップチェインについて算出した滞在地点と時刻を合成する．

以上の手順を全サンプルに適用した結果，時間と空間の集約効果のため，G^p_{mk}は同一または類似したトリップチェインpを多く含むようになる．その場合，個人属性区分lが同一のとき， MSS データでは同一の説明変数として観測されるため，それらの識別はできなくなる．

そのようなG^p_{mk}は，後述する分析では多重共線性の原因となるため，除外する必要がある．そこで個人属性区分lごとに説明変数間のコサイン類似度を計測して，同一または類似したトリップチェインpを除外する．式（1）にコサイン類似度を示す．

$$\cos(\boldsymbol{A},\boldsymbol{B}) = \frac{AB}{\|A\|\|B\|} \qquad (1)$$

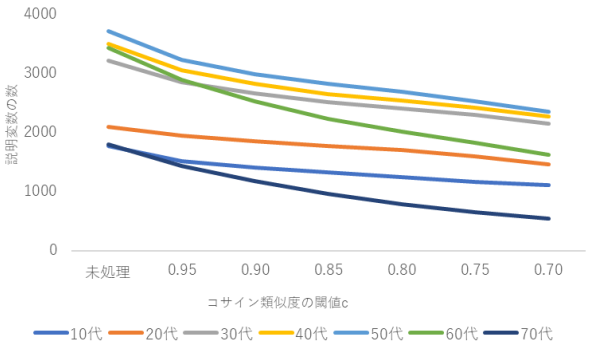


図 5　閾値 c 別，説明変数の数の推移（男性）

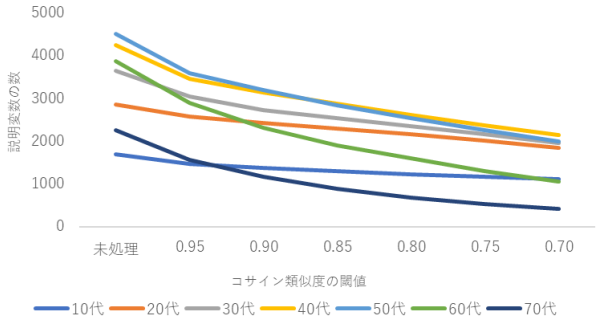


図 6　閾値 c 別，説明変数の数の推移（女性）

式（1）において A，B はトリップチェインG^p_{mk}を同一フォーマットでベクトル化した変数をあらわす．コサイン類似度の値が閾値cよりも大きい場合に，２つのトリップチェインは類似していると判断して，一方を説明変数から削除する．その後，相互に類似度の低いトリップチェインが得られる．男性および女性について非類似の説明変数の数の推移を図 5,6 にそれぞれ示す．閾値cは 0.95 から 0.05 刻みで，0.70 まで設定した．これらの図より，閾値を低くすると類似するトリップチェインが増加し，モデルに投入する説明変数が減少することが確認できる．

さらに MSS についても，空間集計単位を C ゾーンに変換して目的変数データS^l_{mk}を求めた．

4.　制約付き重回帰分析と Lasso 回帰の概要

本研究では，制約付き重回帰分析と Lasso 回帰を用いる．抽出するトリップチェインを比較するため，これらの説明変数と目的変数は共通である．説明変数は PT から得られた個人属性区分lごとの類似度の低いトリップチェインr，時間帯k，ゾーンgで観測される滞在人数に変換したベクトルG^p_{mk}であり，目的変数は，MSS から得た，時間帯k，ゾーンgで観測される滞在人数S^l_{mk}を用いる．また，PT と MSS を対応させる分析の目的を明確

にするため，今後は，パラメータβ^rを拡大係数と呼ぶ．

全ての滞在人数はトリップチェインに由来すると仮定するため，拡大係数に非負制約，切片に 0 制約を加えた重回帰分析を用いる．式（2）に制約付き重回帰分析，式（3）に制約条件を示す．

本研究で抽出したトリップパターンは目的変数の観測時点とは異なるが，都市圏の大まかな土地利用や交通網に違いはない．このため，仮に MSS と同時点の PT トリップチェインが観測できたとしても，過去のトリップパターンとある程度類似した観測になると考えられる．そこで，以下の分析では，説明変数の代表性を留保して，より汎化性能の高い拡大係数を求めるために，Lasso 回帰を用いる．Lasso 回帰では，係数の絶対値の和をペナルティ項として加算して式（4）を最小化する．式（5）に重回帰分析と同じ制約条件を示す．Lasso 回帰ではペナルティ項の負荷量を定めるαは推計できないため，その値は，探索的に求めなくてはならない．

$$\hat{\beta}_{ols}^{r} = arg\,min\Sigma^{g,k}\left(S_{gk} - \hat{S}_{gk}\right)^2 \tag{2}$$

$$\beta^r \geqq 0, \qquad \beta^0 = 0 \tag{3}$$

$$\hat{\beta}_{lasso}^{r} = arg\,min\Sigma^{r}\left(S_{rk} - \hat{S}_{rk}\right)^2 + \alpha\Sigma^{r}\left|\beta^r\right| \tag{4}$$

$$\beta^r \geqq 0, \qquad \beta^0 = 0 \tag{5}$$

5. 分析結果と考察

本章では，MSSにPTから抽出したトリップチェインを回帰する両モデルの時間帯別市町村別（地区別）に算出した決定係数と，拡大係数についてまとめる．

(1) 重回帰分析後のトリップチェイン閾値別決定係数

表 1,2 に男性および女性のトリップチェイン閾値c別，および個人属性区分l別の決定係数をそれぞれ示す．説明変数間のコサイン類似度の閾値cが 0.90 のとき，全ての個人属性区分lの決定係数が 0.9 以上となった．これらの表より，男女間の決定係数の差がほぼないこと，および 70 代男女では閾値cが小さくなると決定係数がより低下することがわかる．

いずれの個人属性区分でもおおむね0.9以上の決定係数が得られたことは，過去（2012年次）のトリップパターンで最近（2019年次）のMSS滞在人数をほぼ表現可能なことを示している．以上の結果を参照してLasso回帰では，コサイン類似度の閾値cを0.95と設定したトリップチェインを用いる．

表 1　トリップチェイン類似度別の決定係数（男性）

	0.95	0.90	0.85	0.80	0.75	0.70
10代	0.95	0.95	0.95	0.95	0.94	0.94
20代	0.97	0.97	0.96	0.96	0.96	0.95
30代	0.97	0.97	0.96	0.96	0.96	0.96
40代	0.98	0.98	0.98	0.97	0.97	0.97
50代	0.98	0.98	0.98	0.97	0.96	0.97
60代	0.97	0.96	0.95	0.94	0.92	0.91
70代	0.93	0.90	0.87	0.83	0.78	0.71

表 2　トリップチェイン類似度別の決定係数（女性）

	0.95	0.90	0.85	0.80	0.75	0.70
10代	0.94	0.94	0.94	0.94	0.94	0.93
20代	0.98	0.98	0.98	0.98	0.98	0.98
30代	0.99	0.99	0.99	0.98	0.98	0.97
40代	0.99	0.99	0.99	0.99	0.98	0.98
50代	0.97	0.97	0.97	0.97	0.96	0.96
60代	0.97	0.96	0.95	0.93	0.91	0.87
70代	0.96	0.94	0.90	0.84	0.77	0.72

表 3　α別，個人属性別の決定係数（男性）

	0.0001	0.001	0.01	0.1	0.2	0.3	0.4	1
10代	0.95	0.95	0.95	0.95	0.92	0.90	0.86	0.63
20代	0.97	0.97	0.97	0.96	0.94	0.91	0.87	0.54
30代	-	-	0.97	0.96	0.93	0.90	0.85	0.48
40代	-	0.98	0.98	0.97	0.95	0.92	0.88	0.54
50代	-	-	0.98	0.97	0.94	0.90	0.84	0.39
60代	-	0.98	0.98	0.97	0.94	0.89	0.84	0.36
70代	-	0.96	0.96	0.94	0.90	0.83	0.75	0.13

（モデルが収束しないで，決定係数が得られなかった値には「-（ハイフン）」を入れる）

表 4　α別，個人属性別の決定係数（女性）

	0.0001	0.001	0.01	0.1	0.2	0.3	0.4	1
10代	0.94	0.94	0.94	0.93	0.91	0.88	0.84	0.59
20代	-	0.98	0.98	0.97	0.96	0.93	0.90	0.61
30代	-	-	-	0.98	0.95	0.92	0.88	0.55
40代	-	-	-	-	-	0.93	0.90	0.63
50代	-	-	-	-	0.94	0.90	0.86	0.51
60代	-	-	-	-	0.94	0.90	0.86	0.47
70代	-	0.98	0.98	0.96	0.94	0.90	0.86	0.45

（モデルが収束せず，決定係数が得られない値には「-（ハイフン）」を入れる）

(2) Lasso回帰のハイパーパラメータ別決定係数

Lasso 回帰のハイパラメータα別個人属性区分l別の決定係数を表 3,4 に示す．これらの決定係数は，モデル推定に使ったデータセットではなく，モデルの汎化性能を確認するために用いたテストデータの 5 回分の推計値の平均である．

これらの表より，男女とも，αが大きくなると決定係数は低下することがわかる．なおαが小さいと，モデルが収束せず，決定係数が得られない．表 3,4 を制約付き重回帰分析の結果（$\alpha = 0.95$）と比較すると，Lasso 回帰の同一年代のデータの決定係数が小さい．つまり，Lasso 回帰は過学習の防止に寄与していると考えられる．男性では，$\alpha \geq 0.01$で全年代で，モデルが収束して，決定係数が得られる一方，女性のモデルが同様となるの

表 5　時間別決定係数分類表

分類	開始	終了	継続時間
朝（家）	3	6	4
出勤時間	7	10	4
帰宅時間	17	20	4
夜（家）	20	23	4
仕事	9	18	9

表 6　熊本市内５か所の空間別決定係数（男性）

市区町村	10代	20代	30代	40代	50代	60代	70代
中央区	0.96	0.96	0.90	0.95	0.96	0.96	0.91
東区	0.94	0.97	0.97	0.96	0.97	0.98	0.95
西区	0.93	0.97	0.96	0.98	0.97	0.98	0.95
南区	0.78	0.92	0.94	0.96	0.96	0.95	0.94
北区	0.77	0.94	0.94	0.95	0.95	0.95	0.93

表 7　熊本市内５か所の空間別決定係数（女性）

市区町村	10代	20代	30代	40代	50代	60代	70代
中央区	0.90	0.94	0.91	0.94	0.90	0.88	0.89
東区	0.81	0.90	0.92	0.90	0.88	0.90	0.89
西区	0.68	0.88	0.90	0.93	0.90	0.94	0.94
南区	0.64	0.83	0.87	0.89	0.86	0.86	0.86
北区	0.73	0.81	0.82	0.90	0.85	0.83	0.88

表 8　空間別決定係数（男性）

市区町村	10代	20代	30代	40代	50代	60代	70代
熊本市	0.95	0.97	0.95	0.97	0.97	0.97	0.95
合志市	0.89	0.94	0.96	0.97	0.95	0.96	0.92
菊池市	0.74	0.90	0.87	0.86	0.82	0.71	0.63
菊陽町	0.44	0.93	0.99	0.99	0.98	0.97	0.96
大津町	0.93	0.97	0.98	0.97	0.97	0.95	0.95
益城町	0.86	0.91	0.96	0.98	0.96	0.97	0.97
西原村	-	0.80	0.87	0.80	0.71	0.73	-
御船町	-	0.56	0.71	0.83	0.90	0.75	0.68
嘉島町	-	0.71	0.59	0.47	0.50	0.65	-
甲佐町	-	0.39	0.58	0.66	0.80	0.76	0.78
宇城市	0.23	0.52	0.92	0.85	0.87	0.85	0.83
宇土市	0.35	0.82	0.93	0.90	0.90	0.86	0.83

-：決定係数が負であるため表示せず.

表 9　空間別決定係数（女性）

市区町村	10代	20代	30代	40代	50代	60代	70代
熊本市	0.88	0.93	0.92	0.93	0.90	0.90	0.91
合志市	0.82	0.90	0.95	0.95	0.89	0.91	0.90
菊池市	-	0.39	0.61	0.66	0.09	0.11	0.40
菊陽町	0.57	0.92	0.96	0.94	0.90	0.91	0.88
大津町	0.67	0.91	0.95	0.94	0.92	0.93	0.87
益城町	0.61	0.77	0.91	0.94	0.87	0.90	0.91
西原村	-	-	-	-	-	-	-
御船町	-	0.07	-	0.25	0.42	0.31	0.16
嘉島町	-	0.66	0.56	-	-	-	-
甲佐町	-	-	-	-	0.32	0.54	0.53
宇城市	-	0.44	0.63	0.63	0.48	0.55	0.58
宇土市	-	0.46	0.74	0.79	0.63	0.65	0.55

-：決定係数が負であるため表示せず.

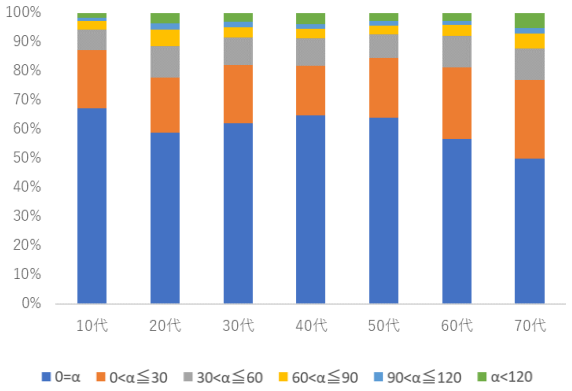


図 7　重回帰後の拡大係数出現割合（%）（男性）

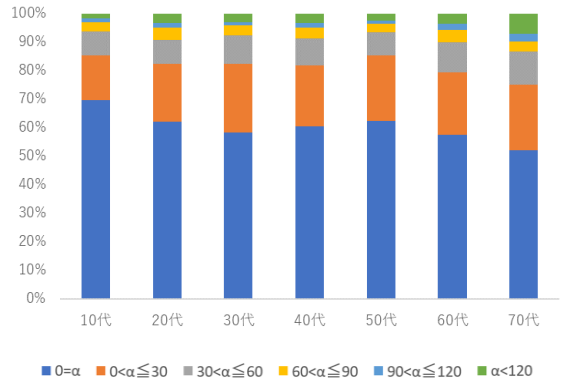


図 8　重回帰後の拡大係数出現割合（%）（女性）

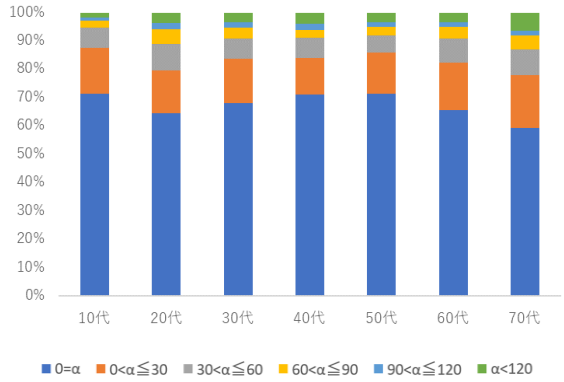


図 9　Lasoo 後の拡大係数出現割合（%）（男性）

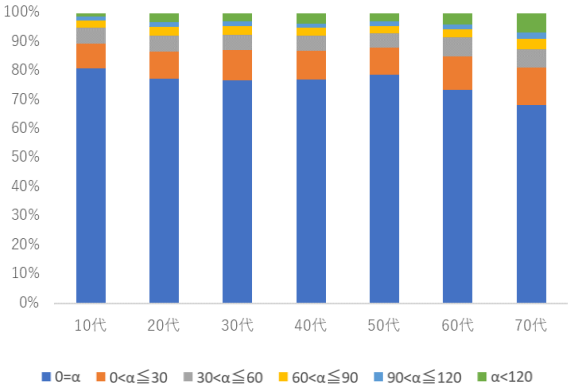


図 10　Lasoo 後の拡大係数出現割合（%）（女性）

表 10　時間別決定係数（男）　(α=0.1)

	朝	出勤	帰宅	夜	仕事
10代	0.97	0.94	0.93	0.94	0.95
20代	0.97	0.95	0.96	0.95	0.97
30代	0.97	0.94	0.95	0.94	0.98
40代	0.99	0.95	0.94	0.96	0.97
50代	0.98	0.94	0.95	0.97	0.97
60代	0.99	0.95	0.96	0.98	0.95
70代	0.97	0.91	0.95	0.96	0.90

表 11　時間別決定係数（女）　(α=0.3)

	朝	出勤	帰宅	夜	仕事
10代	0.89	0.86	0.87	0.83	0.89
20代	0.94	0.90	0.92	0.93	0.94
30代	0.94	0.90	0.89	0.93	0.92
40代	0.96	0.91	0.89	0.94	0.93
50代	0.93	0.89	0.85	0.91	0.90
60代	0.94	0.87	0.90	0.93	0.87
70代	0.93	0.86	0.92	0.93	0.87

は$\alpha \geq 0.3$のときである．男性は全年代で，αが0.01～0.1のときの決定係数の差が小さいことに注意して，以下では男性のαは 0.1，女性のαは 0.3 のモデルに着目する．これらの推定値に基づいて時空間別決定係数を算出して，推計したモデルの再現性能を確認する．

(3)　Lasso回帰から得た滞在人数の時空間決定係数

時間別決定係数は，表 5 のように 1 日を 5 区分して算出した．Lasso 回帰（男性α=0.1，女性α=0.3）から推定した C ゾーン別，時間帯別，滞在人数と観測値の決定係数を表 6～9 に，それぞれ示す．熊本市および市内（中央区，東区，西区，南区，北区）の決定係数は比較的高い．しかし一部の市区町村の決定係数は低いか，または負の値となった．時間帯別の決定係数は比較的高く，0.83 以上となった．特に朝時間帯の値が最も高かった．

(4)　拡大係数の階級別分布

重回帰，および Lasso 回帰によって得た拡大係数の階級別分布を個人属性別に図 7～10 にそれぞれ示す．全年代で拡大係数が 0 の割合が最も高い．重回帰では PT トリップパターンの 50～70%，Lasso 回帰では，60～80%に拡大係数が付与されない結果となった．重回帰よりも Lasso 回帰の方が，拡大係数の値が 0 の割合が高く，女性ではその傾向が顕著である．これは Lasso 回帰による説明変数の抑制効果のためである．また拡大係数が大きなトリップチェインの出現回数の割合は低い．なお年代の傾向は両図で共通しており，40,50 代を除いて，高年代にほど拡大係数が 0 の割合が低下する．

6.　おわりに

本研究の成果は，主に以下の 3 点である．1 点目は，提案した滞在人口モデルの決定係数が比較的高かったことである．したがって過去（2012 年）の PT トリップパターンで最近の MSS の滞在人数をほぼ表現できるため提案手法によってある程度はトリップチェインの時点間移転が可能と考えられる．2 点目は，決定係数を空間および時間で分解して，詳細なモデル性を検討したことである．その結果，都市部と郊外部で決定係数に大きな違いがあること，時間帯別では朝（3～6 時）の決定係数が最も高いこと，1 日の中での差は小さいこと，および空間別決定係数のばらつきは時間帯別決定係数のばらつきよりも大きなことがわかった．3 点目は，PT トリップチェインの半数程度には，拡大係数が付与されなかったことである．

今後検討すべき課題をまとめる．3 点目の知見に関連して PT トリップチェインとの比較を行って，拡大係数の値が 0 となったトリップの特徴を分析する必要がある．またこの間の土地利用の変化を確認して新しいトリップパターンが発生した可能性を検討する必要がある．

参考文献

1) 円山琢也，佐藤嘉洋，野原浩大郎：熊本都市圏 PT 調査と連携したスマートフォン型交通調査の活動紹介, 2014
2) 野原浩大郎，円山琢也：スマートフォン型交通調査の開発・試行・改良，土木計画学研究・講演集，Vol.48，#132, 2013.
3) 山崎 恭彦，橋本浩良，高宮進，矢部努，今井龍一，塚田幸広，山王一郎，石田東生：スマートフォンアプリを活用した交通行動調査 手法に関する基礎的研究 ～つくば市におけるプローブパーソン調査を通して～，土木計画学研究・講演集 Vol.49，CD-ROM, 2014.
4) 寺田雅之，川上博，岡島一郎，篠崎俊哉，坂下昭宏：モバイル空間統計の実用化に向けた取り組み，情報処理学会デジタルプラクティス, Vol6, No.1, 2015.
5) 清家剛，三牧浩也，原祐介，小田原亨，永田智大，寺田雅之：まちづくり分野におけるモバイル空間統計の活用可能性に係る研究，都市計画論文集，Vol.46，No.3, pp.451-456, 2011.
6) 室井寿明，磯野文暁，鈴木俊博：モバイル・ビッグデータを用いた都市間旅客交通への活用に関する研究，第 51 回土木計画学研究発表会・講演集，Vol.51, CD-ROM, 2015.
7) 今井龍一，藤岡啓太郎，新階寛恭，池田大造，永田 智大，矢部努，重孝浩一，橋本浩良，柴崎亮介，関本義秀：携帯電話網の運用データを用いた人口流動統計の都市交通

分野への適用に関する研究，土木計 画学研究・講演集, Vol.52, 2015.
8) 菊池雅彦，岩舘慶多，羽藤英二，茂木渉，森尾淳：全国 PT 調査データと携帯電話基地局データを用いた地方都市での OD 表の実務的推計，土木学会論文集, Vol.74, No.5, pp.677-691, 2018.
9) 越智 健吾，関信郎，岩舘慶多，石神孝裕，若井亮太，石井良治，杉田渓：PT 調査データとビッグデータを用いた実務的 な詳細スケール目的別手段別 OD 表の作成手法, 土木学会論文集, Vol.75, No.5, pp.709-717, 2019.

Updating PT trip chain information by reassigning expansion coefficient

Masaki IDOGAWA, Khin INGYIN, Makoto TSUKAI

The Person Trip (PT) survey, a key survey for urban transportation planning, has the problem of not being able to respond quickly to changes such as a new coronavirus. Mobile Spatial Statistics (MSS), which has been attracting attention in recent years, has high spatio-temporal resolution and rapid reporting capability, but has practical reliability problems in urban transportation planning. In this study, a method to update the expansion coefficient of the trip chain by combining PT and MSS is tested for the Kumamoto metropolitan area. Specifically, the trip chain obtained from the PT is converted into the MSS format, which records the number of trips by time of day, by location, and by number of stays, and the expansion coefficients are calculated using Lasso regression. The relatively high coefficient of determination of the population stay model allowed us to represent the number of recent MSS stays with the PT trip patterns in the past (2012).