

パーソントリップ調査の行動データによる 個人属性逆推定手法の研究

ブンポン 健人¹・奥村 誠²・大窪 和明³

¹学生会員 東北大学大学院 工学研究科土木工学専攻 (〒980-8577 宮城県仙台市青葉区片平2-1-1)
E-mail:bunpon@cneas.tohoku.ac.jp

²正会員 東北大学教授 災害科学国際研究所 (〒980-8577 宮城県仙台市青葉区片平2-1-1)

³正会員 埼玉大学助教 理工学研究科 (〒338-8570 埼玉県さいたま市桜区下大久保255)

ICT技術の進展に伴い、様々な場面で人々の行動ログデータが採取されるようになった。例えばNTTドコモやゼンリンデータコムは携帯電話ユーザの位置情報の活用に向けて取組を開始している。現在こうしたデータは、個人情報保護の観点から{住所、勤務先}などの個人属性を非識別化し、数100m単位のメッシュ毎の人口データに加工した後に公開されている。現在の方法とは異なる方法で加工することによって、新しい活用方法が見つかる可能性があるが、そのような検討はこれまで行われていない。加工前データでは①各サンプルの位置情報が数分～数時間の時間間隔を持って把握できるが、②個人属性{住所、勤務先…}は把握できない。そこで本分析では、①から把握できる1日以上期間の「移動状態{移動、滞在}」を想定し、PT調査のマスターデータを用いて「①各サンプルの移動状態{移動、滞在}」から、「②個人属性{勤務の有無、自動車利用の有無…}」を推定する方法を検討する。

Key Words : *person trip survey, decision tree, mobile phone data, GPS*

1. はじめに

(1) 正確な需要分析の必要性

人口が減少トレンドに変化するなかで、日本の社会資本整備は転換期にある。人口が増加し経済成長を続けていた時には、建設時点で利用者が少なかったとしても(供給超過)、いずれ人口に伴って需要が増加するため、社会インフラの需要予測には正確な精度が求められていなかった。しかし、人口減少社会を迎え財政力が縮小する中、不要な社会インフラに過剰な投資をすることを避けるため、変化する需要の動向をいち早く、詳細に、しかも正確に把握することが求められている。

(2) 既存の社会調査データ(質問・記述式)の限界

例えば都市内の交通計画を考えれば、従来交通需要を把握するために、都市内の旅客交通流動を総合的に捉えるためのパーソントリップ調査(以下、「PT調査」と呼ぶ)や、自動車を対象とする道路交通センサス、鉄道、バスを対象とする大都市交通センサスなどの調査が行われてきた。これらは、いずれも調査費用が莫大にかかる上に、プライバシー意識の高まりから回答率が低下し、

さらにアンケート調査では時刻などが丸められて回答されるなどの問題があり、実態に即した結果が得られているとは限らない。また調査スパンも数年と長いため、実際の交通施設整備計画には用いにくいのが現状である。

2. ICT技術の進展とデータオリエンテッドな計画アプローチの登場

(1) 観測・計測式データの登場

情報通信技術の高度化により、従来はアンケート等で調査しなければ得られなかったような人々の行動に関する情報が、ICT機器の相互通信によって得られるようになった。その最たる例が携帯電話やPHSの位置特定である。1990年後期には基地局と端末間の相互通信によって特定した端末の位置情報を、福祉・保安目的で提供するサービスが提供されはじめた²⁾。近年では、基地局と端末間の相互通信によって得られる利用者の時々刻々の位置情報(以下、「携帯位置情報」と呼ぶ)を都市計画に活用しようという動きが活発化している。

例えば株式会社NTTドコモは、携帯位置情報を用いて

時々刻々の人口動態を推計し、これを「モバイル空間統計」と呼び、都市計画・防災計画に役立てることを目指している³⁾。また、株式会社ゼンリンデータコムも自社サービスの利用者から許諾を得て取得した位置情報（携帯電話内蔵のGPSから採取されるもの）を用いて時々刻々の人口動態を推計し、これを「流動人口推計データ」と呼び、マーケティング・データとしての販売を行っている⁴⁾。

(2) 携帯位置情報の活用の潮流

これらの携帯位置情報データの活用を念頭に置いた研究も既に行われている。清家ら⁵⁾は株式会社NTTドコモとの共同研究として「モバイル空間統計」データについて空間解像度・時間解像度・人口構成・移動人口などの信頼性の検証を行っている。また関本ら⁶⁾は株式会社ゼンリンデータコムの保有するデータを用いて、東日本大震災における首都圏の帰宅困難者数の推計に取り組んでいる。

これらの携帯位置情報は任意時点での位置情報を把握できる点に優位性があり、震災下のような調査困難な状況下における被災地域の実態把握や、調査スパンの長いPT調査などの既存調査を補完・代替するデータソースとしての活用が期待されている。

しかし、こうした携帯位置情報の活用にあたっては個人情報保護等の観点から、情報を保有する携帯キャリアでさえも個人属性（年齢、職業…）を取り扱うことができない。また位置情報も、一般的に外部へ提供される際には同一個人の時々刻々の位置情報を辿ることができないように加工されている。株式会社ゼンリンデータコムが販売している「流動人口推計データ」も250mや500m単位のメッシュデータとして集計されたデータである。

(3) 本研究の目的

携帯位置情報を社会資本整備などの都市計画に活かそうと考えた際に、メッシュやエリア毎に集計されることによる情報量の損失の影響は非常に大きい。具体的には時間の経過に伴う同一個人の状態の変化を把握できないため、施策の実施によってその人の行動が本当に変化したかどうか判定できない。

こうした状況の中で、携帯位置情報を都市計画に活かすためのアプローチとして以下の2点が考えられる。

◆個人情報に抵触しないような形で非集計データを扱う

非集計の携帯位置情報が個人情報保護に抵触する理由は、時々刻々の位置情報を辿ることで自宅や職場位置を推定でき、個人の特定につながるからだと考えられる。そこで携帯位置情報から地理的情報を取り除き、移動状

態（移動しているのか、同一地点に留まっているのか）に関する時間のデータに変換すれば、個人の同定はできなくなり、個人情報に抵触せずに行動の変化の把握に活用できる可能性がある。

◆データの集計方法を工夫してもらう

地域毎に単純集計した人口動態ではなく、特定の属性を持つサンプルのみを抽出した上で集計を行うことで、都市計画上有益な情報が得られる可能性がある。例えば、東日本大震災の前後で「勤務先に通勤できている人」を抽出すれば、震災からの産業の回復具合を見ることがができる。しかし、携帯位置情報データからは個人属性を把握できないので、何らかの形で個人属性を推定する必要がある。

そこで本研究ではPT調査を用いて、「移動状態（移動しているのか、同一地点に留まっているのか）」と「個人属性」の関係性について調べる。ここで取り扱う「個人属性」とは、上述の「移動状態」に対して影響を与えるような、性別や労働状態、就学状態、交通手段などを意味している。つまり本研究では、「個人属性」を記述するような「移動状態」に関するルールを構築し、どのような「移動状態」のグループが、どのような「個人属性」を持つのかを推定することを目指す。

(4) 分析手法

本研究では非線形判別分析のための手法の1つである決定木モデルを用いて分析を行う。決定木モデルでは、説明変数がある基準で分岐させ、目的変数の判別・予測のためのモデルを構築する。分岐の基準がIF-THENのように簡潔に表現でき、分岐過程を木構造で視覚的に表現できることなどから近年利用が進んでいる。決定木モデルには大きく分けて、CHAID、C4.5/C5.0/See5、CARTの3種類をベースとしたアルゴリズムが存在するが、本研究ではオープンソースの統計解析ソフトウェアであるRにおいて、CARTアルゴリズムに基づく決定木分析を実行できるパッケージ `mypart` を用いて分析を進める。CARTアルゴリズムでは説明変数を2進木に分岐させ、分岐の基準としてジニ係数もしくはエントロピー指標を用いることができるが、本研究ではジニ係数を用いて分析を行う。なお本研究では、前節で述べた個人属性に基づく分類分けを考え（3章(1)）、これを「個人属性分類」と呼んで目的変数とする。そして移動状態から構築される説明変数（以降、「説明変数（移動状態）」と呼ぶ）を用いて決定木分析を行い、個人属性分類に基づきサンプルを上手く分類するための分岐ルールを構築する。なお島崎ら⁸⁾は、PTデータに対して決定木分析を行

って単一トリップの位置情報から個人属性を逆推定している。本研究では同一個人の日々の全トリップに関する移動状態から個人属性分類を逆推定するが、地理的情報は用いない点で分析内容は大きく異なっている。

3. PT調査マスターデータと分析の準備

本研究では2002年に実施された第4回仙台都市圏パーソントリップ調査のマスターデータ（以下、「PTマスターデータ」と呼ぶ）を用いる。なお本研究では移動状態を用いた分析を行うため、トリップなしのサンプルについては扱わない。

(1) 個人属性分類の構築

まず本研究の分析の目的変数となる個人属性分類の基盤としてPT調査のフェイス調査項目から表-1の6種類の属性を選択した。これらの属性は2通りもしくは3通りの値をとり、その値の組み合わせによって、図-1に示す96種類の個人属性分類が考えられる。それぞれの個人属性

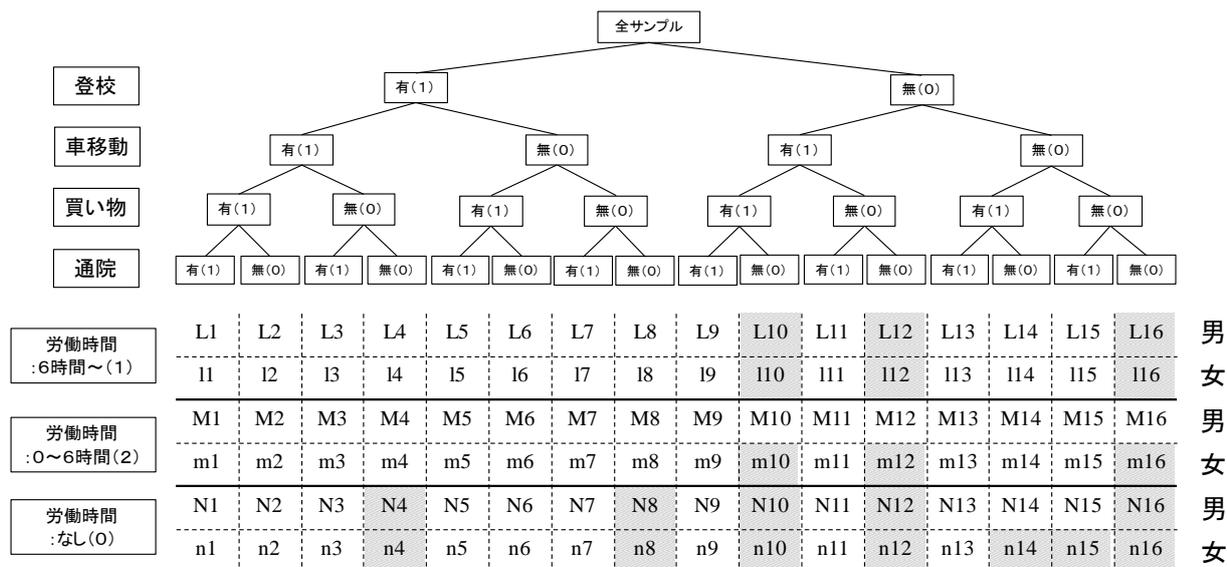
分類のPTデータにおけるサンプル数を図-2に示す。本研究では96種類の個人属性分類のうちサンプル数が1000以上のものを用いて分析を行うこととし、対象となる21種類の個人属性分類とその属性値の一覧を表-2に示す。なお以降では表-2にも示したように、労働時間が6時間以上のサンプルを「正規雇用者」、労働時間が0~6時間の場合を「非正規雇用者」、登校ありの場合を「学生」、労働時間、登校いずれもない場合を「専業主婦・無職」とみなす。

(2) 説明変数（移動状態）の構築

次に、前節で構築した個人属性分類を説明するための説明変数（移動状態）を、PT調査のトリップ調査項目を用いて構築する。トリップの所要時間は移動にかかる時間と考えることができ、前のトリップの到着から次のトリップの開始までの時間は、目的をもって1つの地点に滞在（活動）している時間と考えることができる。このような考え方にに基づき、本研究では表-3に示した24種類の変数を構築し、説明変数（移動状態）として用いる。

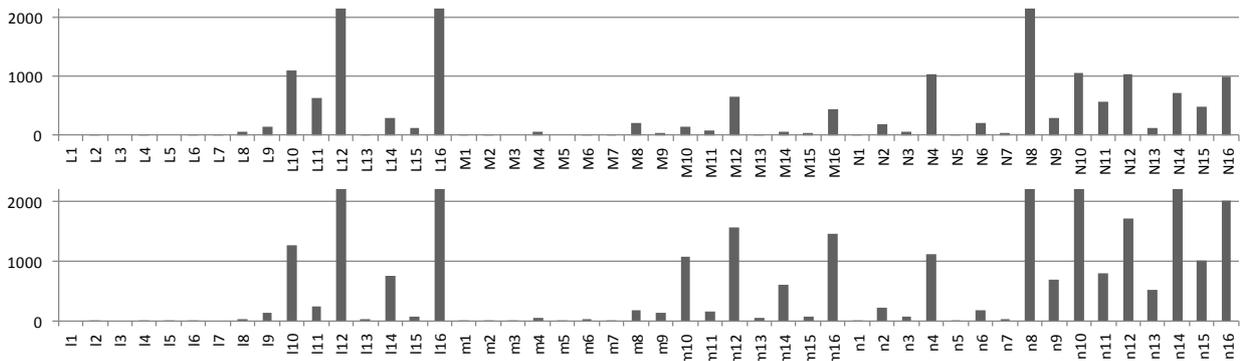
表-1 個人属性分類を構築するための属性一覧

属性名	属性値	内容
労働時間	1:6時間~, 2:0~6時間, 0:なし	普段の労働時間
登校	1:あり, 0:なし	調査日の登校状況
車移動	1:あり, 0:なし	調査日の車利用トリップの有無
買い物	1:あり, 0:なし	調査日の「買い物」目的トリップの有無
通院	1:あり, 0:なし	調査日の「医療・厚生・福祉施設」への移動の有無
性別	1:男性, 2:女性	-



注1: 網掛け部分は実際に本研究で使用する変数

図-1 個人属性分類の候補一覧



注1：サンプル数2000以上の部分は示していない

図-2 個人属性分類候補のサンプル数

表-2 分析に使用する個人属性分類とその属性値

個人属性分類	性別	労働時間	登校	車移動	買い物	通院	
L10	女	6時間～	なし	有	有	なし	正規雇用
L10	男	6時間～	なし	有	有	なし	
L12	女	6時間～	なし	有	なし	なし	
L12	男	6時間～	なし	有	なし	なし	
L16	女	6時間～	なし	なし	なし	なし	
L16	男	6時間～	なし	なし	なし	なし	
m10	女	0～6時間	なし	有	有	なし	非正規雇用
m12	女	0～6時間	なし	有	なし	なし	
m16	女	0～6時間	なし	なし	なし	なし	
n4	女	なし	有	有	なし	なし	学生
N4	男	なし	有	有	なし	なし	
n8	女	なし	有	なし	なし	なし	
N8	男	なし	有	なし	なし	なし	
n10	女	なし	なし	有	有	なし	専業主婦・無職
N10	男	なし	なし	有	有	なし	
n12	女	なし	なし	有	なし	なし	
N12	男	なし	なし	有	なし	なし	
n14	女	なし	なし	なし	有	なし	
n15	女	なし	なし	なし	なし	有	
n16	女	なし	なし	なし	なし	なし	
N16	男	なし	なし	なし	なし	なし	

4. 結果と考察

(1) 決定木分析による分岐モデル構築

3章で構築した個人属性分類と説明変数（移動状態）を用いて、統計解析ソフトウェアRのパッケージmvpartを用いて決定木分析を行う。なお、使用するデータセットとしては、表-2の21種類それぞれの個人属性分類を持つサンプルをそれぞれ500サンプルずつランダム抽出したものを2組用意し、1組は決定木分析によるモデル構築用（以後、「モデル構築データ」と呼ぶ）、もう1組は構築モデルの検証用（以後、「モデル検証データ」と呼ぶ）として用いる。モデル構築データを用いて決定木分析を行った結果の構築モデルを図-3に示す。図-3中の分岐ルールに従い説明変数（移動状態）を分岐させることでデータセットは①～⑧の8種類の分岐先に分類される。

構築モデルを「モデル構築データ」に対して適用した結果を表-4(ア)に、「モデル検証データ」に適用した結果を表-4(イ)に示し、表中ではそれぞれの個人属性分類を持つサンプルが①～⑧の分岐先に何サンプル分類されるかを表している。

表-3 分析に使用する説明変数（移動状態）の候補一覧

説明変数	説明
最初移動	最初のトリップの開始時間
最終移動	最後のトリップの終了時間
外出時間	「最終移動」-「最初移動」
総移動時間	トリップの所要時間の総和
午前移動	「総移動時間」の内で0～12時のもの
午後移動	「総移動時間」の内で12時～のもの
0-6移動	「総移動時間」の内で0～6時のもの
6-12移動	「総移動時間」の内で6～12時のもの
12-18移動	「総移動時間」の内で12～18時のもの
18-移動	「総移動時間」の内で18時～のもの
総滞在時間	「前トリップ到着時間」から「次トリップ開始時間」のまでの時間間隔の和
午前滞在	「総滞在時間」の内で0～12時のもの
午後滞在	「総滞在時間」の内で12時～のもの
0-6滞在	「総滞在時間」の内で0～6時のもの
6-12滞在	「総滞在時間」の内で6～12時のもの
12-18滞在	「総滞在時間」の内で12～18時のもの
18-滞在	「総滞在時間」の内で18時～のもの
トリップ数	1日の総トリップ数
午前トリップ数	「トリップ数」の内で0～12時のもの
午後トリップ数	「トリップ数」の内で12時～のもの
0-6トリップ数	「トリップ数」の内で0～6時のもの
6-12トリップ数	「トリップ数」の内で6～12時のもの
12-18トリップ数	「トリップ数」の内で12～18時のもの
18-トリップ数	「トリップ数」の内で18時～のもの

ここで表-4中の網掛け部分は、分岐先(①~⑧)に「多く」分類された個人属性分類を表している。具体的には、分岐先に全ての個人属性分類がランダムに分類されている場合(ある個人属性分類が確率1/21で分類され、それ以外の個人属性分類が確率20/21で分類される二項分布)の99%有意水準以上のサンプル数が分類された場合を表している。表-4(ア)(イ)それぞれの網掛け部分を見比べると、いずれかで網掛けとなる部分が46箇所、一方で網掛けであるが他方では網掛けでない部分が6箇所であり(表-4中の太枠部分)、その割合は6/46≒13%であり、今回の構築モデルは他のデータセット(モデル検証データ)に対してもある程度当てはまるモデルであるといえる。構築モデルの一般性が確認されたため、今後の考察については、表-4(ア)の「モデル構築データ」の結果をベースに行っていく。

(2) ①~⑧の分岐先の特徴に関する考察

①~⑧のそれぞれの分岐先に対してどの個人属性分類をもつサンプルが分類されたのかを図-4に示す。ここで図の縦軸はそれぞれの個人属性分類の500サンプルのうち、図の分岐先に分類された割合を表している。また図中に各分岐先に分類されたサンプルの総数を併せて示す。網掛け部分については図中の説明文の通りである。

ここでもう一度図-3をみると、分岐ルールとして「総滞在時間」が多く採用されている。「総滞在時間」とは表-3に示した通り、前トリップの到着から次トリップの開始までの時間間隔の総和であり、1日のうちの自宅以外での滞在(活動)時間を表している。総滞在時間の値によって分岐先は4つのグループに分けられる。以下では「滞在時間」を「活動時間」と読み替えて、各グループを解釈していく。

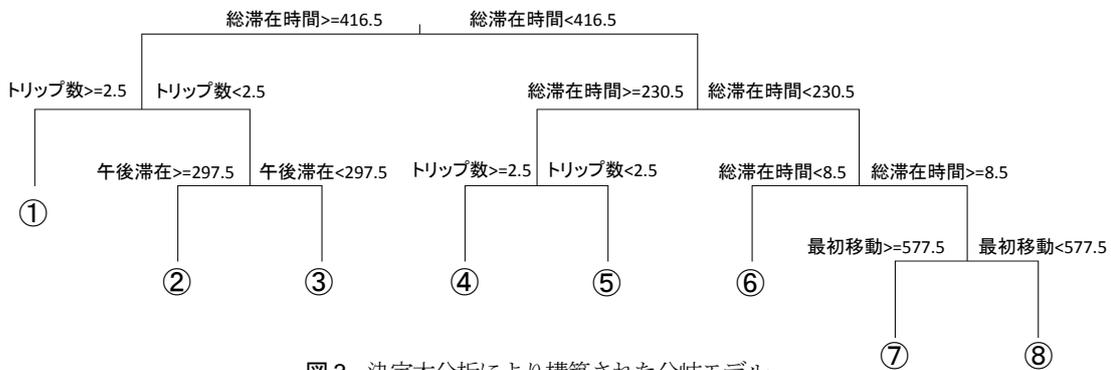


図-3 決定木分析により構築された分岐モデル

表-4 構築した分岐ルールの適用結果

分岐先にランダムに分類される場合の99%有意水準以上のサンプルが分類されている
 一方で網掛けであるが他方では網掛けでない部分

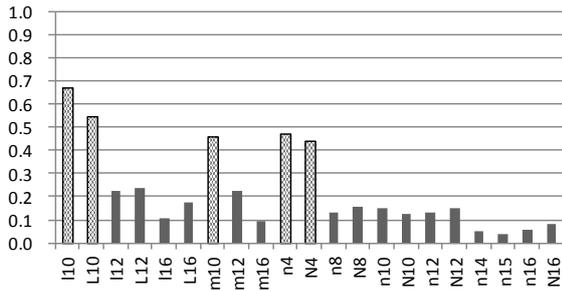
(ア) モデル構築データに適用

	I10	L10	I12	L12	I16	L16	m10	m12	m16	n4	N4	n8	N8	n10	N10	n12	N12	n14	n15	n16	N16
①	337	273	112	119	54	88	231	111	46	238	220	66	78	74	63	65	75	24	19	28	40
②	1	2	288	263	365	305	1	37	40	88	104	117	116	1	1	12	16	1	8	6	17
③	0	0	20	14	24	15	0	27	39	78	58	190	165	0	0	4	9	0	6	8	9
④	51	91	13	18	4	14	142	70	27	22	18	4	3	100	103	101	77	70	56	64	44
⑤	2	4	13	9	17	12	2	154	265	54	70	104	106	5	3	60	43	3	84	55	46
⑥	4	7	37	59	21	45	8	28	19	7	17	7	21	8	22	63	87	7	7	98	107
⑦	100	107	16	12	14	11	113	49	50	6	6	4	2	297	292	173	155	390	220	200	177
⑧	5	16	1	6	1	10	3	24	14	7	7	8	9	15	16	22	38	5	100	41	60

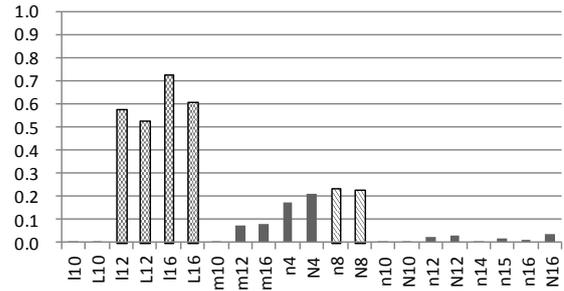
(イ) モデル検証データに適用

	I10	L10	I12	L12	I16	L16	m10	m12	m16	n4	N4	n8	N8	n10	N10	n12	N12	n14	n15	n16	N16
①	338	283	115	130	66	85	231	117	54	222	211	59	64	60	49	71	66	14	18	30	37
②	2	6	288	261	348	317	0	36	40	116	95	128	101	0	3	14	16	5	4	11	19
③	0	0	24	10	25	15	0	31	39	74	63	178	187	0	0	4	9	0	12	1	12
④	53	57	12	17	5	12	150	48	30	15	20	7	6	112	119	112	65	81	45	80	37
⑤	3	3	12	4	16	3	0	161	261	52	83	102	114	4	5	38	53	8	71	52	48
⑥	2	9	31	64	23	49	4	42	24	11	17	15	14	11	11	68	96	10	15	82	109
⑦	98	130	16	6	13	11	108	44	35	7	5	3	11	300	296	167	151	372	216	208	183
⑧	4	12	2	8	4	8	7	21	17	3	6	8	3	13	17	26	44	10	119	36	55

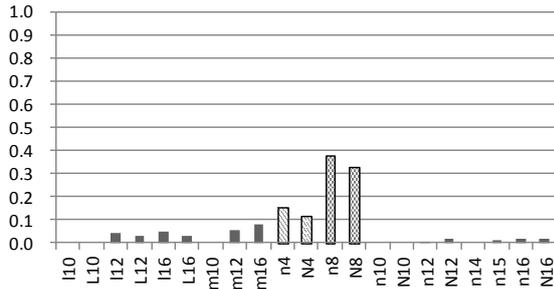
①: 分類サンプル数=2361



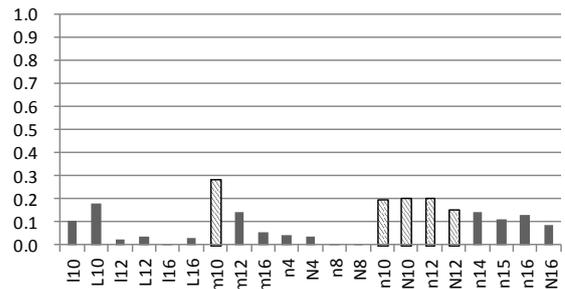
②: 分類サンプル数=1789



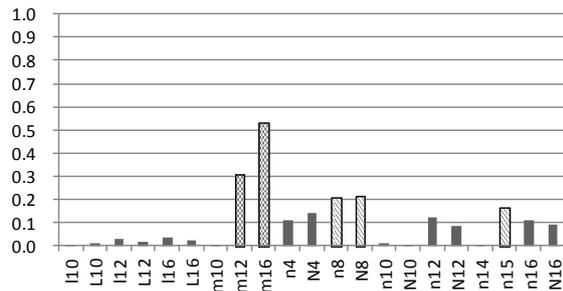
③: 分類サンプル数=666



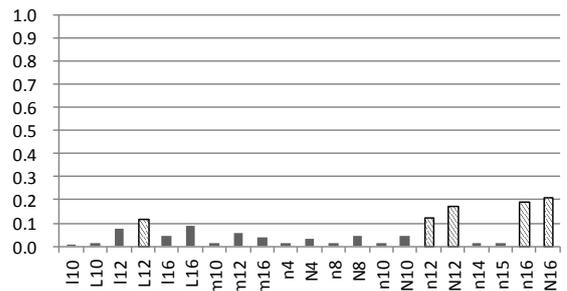
④: 分類サンプル数=1092



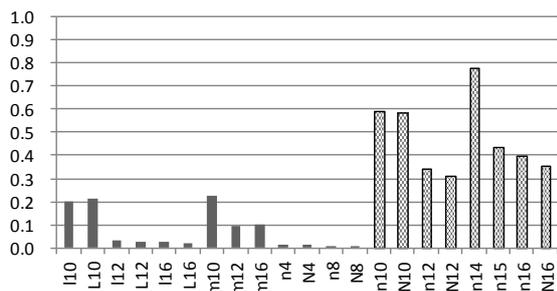
⑤: 分類サンプル数=1111



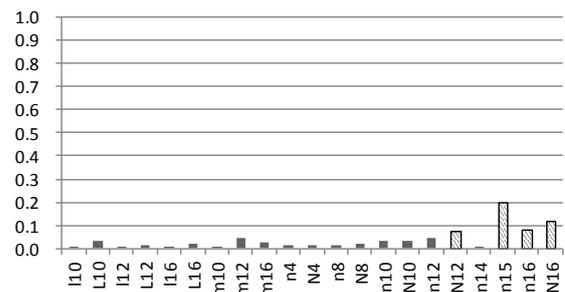
⑥: 分類サンプル数=679



⑦: 分類サンプル数=2394



⑧: 分類サンプル数=408



注 1: 各図中の縦軸はそれぞれの個人属性分類の分類数を、それぞれの個人属性分類のサンプル数 (=500) で基準化したものである。例えば、個人属性分類 l10 に関して 8 つ全ての図での縦軸値の和をとると 1 になる。



網掛けパターン 1: それぞれの個人属性分類の中で最大の分岐先



網掛けパターン 2: 分岐先にランダムに分類される場合の 99% 有意水準以上のサンプルが分類されている

図-4 それぞれの個人属性分類の分岐先割合

◆第1グループ：416.5≤総滞在時間（分岐先：①②③）

活動時間の長いサンプルが分類される。「正規雇用者」「学生」は主に第1グループに属している。また「非正規雇用者」については、「車移動」があり「買い物」トリップをしているようなサンプルが第1グループに属している。

①と②③には分岐ルール上の大きな違いとしてトリップ数の条件があり（①はトリップ数 ≥ 2.5 ，②③はトリップ数 < 2.5 ），これによって分類される個人属性分類の傾向が大きく異なる。各分岐先で網掛けとなっている個人属性分類を見ると，①ではいずれも「車移動」があり，②③ではいずれも「車移動」がない。このことから自動車の利用可能性によってトリップ数が多くなることがわかる。

具体的には，各分岐先に以下のような個人属性分類が分類されている。

分岐先①

- 網掛けパターン1の個人属性分類
：「I10, L10→正規雇用」「m10→非正規雇用」「n4, N4→学生」
- 網掛けパターン2の個人属性分類
：なし

分岐先②

- 網掛けパターン1の個人属性分類
：「I12, L12, I16, L16→正規雇用」
- 網掛けパターン2の個人属性分類
：「n8, N8→学生」

分岐先③

- 網掛けパターン1の個人属性分類
：「n8, N8→学生」
- 網掛けパターン2の個人属性分類
：「n4, N4→学生」

◆第2グループ：230.5≤総滞在時間<416.5（分岐先：④⑤）

活動時間が中程度のサンプルが分類される。「非正規雇用者」の内で買い物トリップがない仕事以外の活動時間が短いサンプルや「専業主婦，無職」の内で「車移動」有り，活動時間の長いサンプルが分類されている。

④と⑤の分岐ルール上の違いもトリップ数の条件であり（④はトリップ数 ≥ 2.5 ，⑤はトリップ数 < 2.5 ），その違いが，分類された個人属性分類の「車移動」の有無として表れている。

具体的には，各分岐先に以下のような個人属性分類が

分類されている。

分岐先④

- 網掛けパターン1の個人属性分類
：なし
- 網掛けパターン2の個人属性分類
：「m10→非正規雇用」「n10, N10, n12, N12→専業主婦，無職」

分岐先⑤

- 網掛けパターン1の個人属性分類
：「m12, m16→非正規雇用」
- 網掛けパターン2の個人属性分類
：「n8, N8→学生」「n15→専業主婦，無職」

◆第3グループ：8.5≤総滞在時間<230.5（分岐先：⑦⑧）

活動時間の短いサンプルが分類される。⑦⑧共に「専業主婦，無職」のサンプルが多く分類されている。

⑦と⑧の分岐ルール上の違いは，「最初移動」の条件である（⑦は最初移動が577.5分 \div 午前9時37分より遅く，⑧は最初移動が午前9時37分より早い）。⑧を見ると，「通院」で移動する個人属性分類のn15が多く比較的多く分類されており，n15には「車移動」もないことから，通院のために朝早く家を出ていることがわかる。

具体的には，各分岐先に以下のような個人属性分類が分類されている。

分岐先⑦

- 網掛けパターン1の個人属性分類
：「n10, N10, n12, N12, n14, n15, n16, N16→専業主婦，無職」
- 網掛けパターン2の個人属性分類
：なし

分岐先⑧

- 網掛けパターン1の個人属性分類
：なし
- 網掛けパターン2の個人属性分類
：「N12, n15, n16, N16→専業主婦，無職」

◆第4グループ：総滞在時間<8.5（分岐先：⑥）

自宅外での活動時間が極めて短いサンプルが分類され，主に「専業主婦，無職」が分類されている。

具体的には，各分岐先に以下のような個人属性分類が分類されている。

分岐先⑥

□網掛けパターン1の個人属性分類

: なし

□網掛けパターン2の個人属性分類

: 「L12→正規雇用」「n12, n16, N16→専業主婦, 無職」

(3) 小結

最後に本研究の成果を用いた携帯位置情報の活用可能性について述べる。携帯位置情報の1日以内における同一ユーザの情報をを用いれば、表-3の本研究で用いた説明変数(移動状態)に相当する情報を各ユーザに対して構築できる。

そして、今回構築した①～⑧のルールを適用することで個人属性に基づいて限定されたユーザのみを抽出できる。適用に当たっては、なるべく個人属性がシャープに分かれて(ある特定の属性を持つユーザだけを抽出できる)、かつ抽出されるユーザ数が多いような分岐ルールを適用することが望ましい。図-4を見ると、分岐ルール①②⑦等は、比較的個人属性がシャープに分かれており抽出数も多い。例えば、②を適用すれば「車利用」のない正規雇用者を多く抽出でき、⑦を適用すれば「専業主婦・無職」を多く抽出できる。そして複数地域での②⑦の適用結果を比較すれば、地域間の就業状態の違いを見ることができる。

4. おわりに

(1) 本研究の成果

本研究では、PT マスターデータに対して決定木分析を行い説明変数(移動状態)を分岐してサンプルを個人属性分類毎に分けるためのルールを構築し、そのルールによって実際にどのような属性のサンプルを抽出できるのかを確かめた。

そして本成果を携帯位置情報に適用すれば、3章(3)の小結で述べたように就業状態などの地域の活動レベルを表現でき、地域間での比較や、同一地域での時系列での比較に用いることができる。例えば、2011年に発生した東北地方太平洋沖地震に伴う東日本大震災の被災地のデータに対して適用できれば、震災からの復興度合いなどを表現できる可能性がある。

(2) 今後の課題

本研究で用いた第4回仙台都市圏PT調査は2002年に実施されたものであり、2013年現在では人々の移動状態と個人属性の関係性も変化している可能性もあり、他の年次の調査や、他都市の調査でどのような結果が得られるかを確かめる必要がある。

また、本成果を実際の携帯位置情報に適用するためには、個人情報への抵触可能性など多くの問題が山積しており、今後1つ1つ検討していきたい。

謝辞: 本研究では分析のために第4回仙台都市圏パーソントリップ調査のマスターデータを用いました。調査関係者に謝意を示します。

参考文献

- 1) 土木学会: 道路交通需要予測の理論と適用, pp.13-22, 2003.
- 2) 小林潔司, 朝倉康夫, 山崎朗: これからの都市・地域政策, pp.83-98, 中央出版社, 2005.
- 3) NTT ドコモ: モバイル空間統計に関する情報 (http://www.nttdocomo.co.jp/corporate/disclosure/mobile_spatial_statistics/)
- 4) ゼンリンデータコム: 流動人口推計データ (<http://www.zenrin-datacom.net/business/other/#statistic>)
- 5) 清家剛, 三牧浩也, 原裕介, 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, 2011
- 6) 関本義秀, 中村敏和, 増田祐介, 金杉洋: 大規模なGPS 情報をもとにした東京都市圏における震災時の行動分析, 都市計画学研究・講演集 Vol.45, 2012
- 7) 金明哲: R によるデータサイエンス, pp.229-246, 森北出版, 2007.
- 8) 島崎康信, 関本義秀, 柴崎亮介: 個人属性の情報量に応じたトリップ目的の判別精度に関する研究—パーソントリップ調査の時空間内挿データと決定木分析を用いて—, 都市計画論文集 N0.45-3, 2010.
- 9) 岡村敏之, 藤原章正, 張峻屹: データマイニング手法を用いた都市公共交通のマーケティングに関する基礎的分析, 土木計画学研究・講演集 Vol28, 2003.

(2013.5.7 受付)

A study on reverse expectation of personal attributes from the activity data in person trip survey

Kento BUNPHONG, Makoto OKUMURA and Kazuaki OKUBO