

トピックモデルを用いた スマホ型回遊調査データの基礎分析

川野 倫輝¹・円山 琢也²

¹ 学生会員 熊本大学大学院自然科学研究科社会環境工学専攻 (〒860-8555 熊本県熊本市中央区黒髪 2-39-1)

E-mail:178d8811@st.kumamoto-u.ac.jp

² 正会員 熊本大学准教授 くまもと水循環・減災研究教育センター

(〒860-8555 熊本県熊本市中央区黒髪 2-39-1)

E-mail:takumaru@kumamoto-u.ac.jp

地方都市都心部の活力衰退の対策の検討には GPS による人の回遊行動データの取得・分析が有効である。回遊行動データの分析においては、回遊行動パターン抽出が主な関心となることも多く、トピックモデルなど抽出手法の検討も進んでいる。しかし、商業地域のような比較的狭い範囲で取得されたデータを対象としてパターン抽出を行う研究は数少なく、トピックモデルにおいてもこのようなデータに対する適用実績は存在しない。本研究では、市街地レベルの範囲で取得されたデータに対するトピックモデルの有用性の検証を念頭に、2013 年秋に熊本市都心部で実施された回遊調査のデータにトピックモデルを適用した。抽出された回遊行動パターンの特徴は、トリップ属性の集計結果の解釈と概ね一致しており、その有用性が確認できた。

Key Words : travel behavior survey, GPS, smart phone, topic model, pattern mining

1. はじめに

(1) 研究の背景・目的

地方都市が共通して抱える課題に都心部の活力の衰微があり、その対策の検討のためには、人の回遊行動データの取得と分析が有効である。回遊行動データ取得には、調査参加者の記憶に頼らずに詳細なデータの把握が可能という点から GPS 機器貸与型の回遊調査が効果的である。一方で、貸与できる機器数の制約により大量のサンプルの取得が困難という課題があった。この課題に対し、近年では、スマートフォン(スマホ)の GPS や Wi-Fi パケットセンサを利用した回遊・周遊調査が注目を集めている。このうち、スマホ型調査とは、専用アプリを調査参加者自身が保有するスマホに配布し、GPS 機能で移動軌跡などを収集するものである。専用アプリ内において個人属性の登録を求めることで、個人属性と移動軌跡を紐づけた分析が可能となる。スマホ型調査の実施例として、2013 年秋に熊本市都心部で実施された回遊調査がある。佐藤・円山¹⁾は、この調査を対象にカーネル密度推定法を用いた回遊圏域の可視化手法を提示している。これらの手法は、移動滞在判別を行うことなく回遊圏域、特に滞在・滞留地点を簡易に推定することを可能にしている。

回遊行動データの分析においては、滞在地点のパターン分析が主な関心となることが多い。しかしながら、佐藤・円山¹⁾の研究においては、回遊行動パターン抽出は行われていない。このような商業地域等の比較的狭い範囲で取得された回遊行動データに対してパターン抽出を適用する研究は、例えば矢部²⁾の研究が挙げられるが、適用事例が限られており、手法が確立されているとは言い難い。膨大なデータの分析に適する機械学習的手法による分析などの検討が望まれる。

一方で、広域の観光周遊行動に関しては、観光振興に向けた知見の発見を狙って観光客の観光地訪問パターンを抽出する研究事例は数多く存在する。さらにデータマイニング手法の発展と共にパターン抽出の手法も多様になってきている。これらの手法のひとつにトピックモデルがある。トピックモデルとは、文書データの解析手法として開発された機械学習の一手法であり、教師データなしでデータ中からトピック(潜在変数)を推定することが可能である。文書データの解析手法として開発されながらも、数多くの分野で応用され、古屋ら³⁾の研究のように回遊行動データへの適用もある。しかし、適用事例が少ないため、他の回遊行動データ、特に先述の通り比較的狭い範囲で得られた回遊行動データへの適用事例

の蓄積が求められる。

以上の背景を踏まえ、本研究では2013年秋に熊本市都心部で実施された回遊調査のデータを対象に、トピックモデルを用いた回遊行動パターン抽出とその結果に基づく分析を行う。適用事例の少なかった位置情報、中でも市街地という比較的狭い範囲にトピックモデルを適用し、その有用性を検証することを目的とする。

(2) 既存研究のレビュー

本章ではトピックモデルに焦点を当て、既往研究を整理する。トピックモデルは、先述の通り、国内外の土木計画・都市計画分野において、幅広い適用・応用事例が見られるようになってきている手法である。文書データ分析に用いたものとしては、学術論文の研究トピックの分析として、Sun and Yin⁹⁾は、交通研究分野の22の国際誌に掲載された論文のアブストラクトを対象にトピックモデルを適用し、研究トピックの増減傾向や研究トピック間の関連を把握している。レポート会議やワークショップなどの発言データの解析として、塚井ら⁹⁾は、地域公共交通会議の討議録にトピックモデルを適用し、討議中のトピック抽出やトピック推移の把握を通して同手法の適用可能性を検証している。また、アンケート調査の自由記述分析として、Kuhn⁷⁾は、航空安全報告システム中のインシデントレポートに構造トピックモデルを適用している。構造トピックモデルはトピックモデルの拡張モデルであり、抽出されたトピック間の相関を把握することが可能である。川野ら⁸⁾は聞き取り調査の自由意見にトピックモデルと離散連続モデルを適用し、トピックと個人属性の関係を統計的に分析する手法を提案している。文書データ以外の解析への応用事例では、神谷ら⁹⁾は、モバイル空間統計データに、文書をメッシュ、滞在者の居住区を単語として、トピックモデルを適用し、地域別人口特性の解釈を行っている。塚井ら¹⁰⁾は、詳細地理情報の解析へのトピックモデル適用の妥当性を議論している。そして、位置情報に適用した例としては、古屋ら⁴⁾が訪日外国人のGPSデータにトピックモデルを適用し、外国人の観光地の訪問パターンを分析している。しかし、1節でも述べた通り、トピックモデルの応用はこのように幅広い一方で、回遊行動データへの適用例は少ない。また、古屋ら⁴⁾の研究は、日本全国を対象としており、市街地レベルの回遊行動データへトピックモデルを適用する研究は筆者の知る限り存在しない。本研究は、このようなデータに対する回遊行動パターンの抽出に関する研究の基礎的なものと位置づけられる。

2. 分析方法

(1) トピックモデルの概要

行動パターンの抽出には、トピックモデルの一種であるLatent Dirichlet Allocation¹¹⁾(LDA)を用いる。ここで、LDAの概要について説明する。なお、LDAについての解説は参考文献¹²⁾に基づいて示す。

LDAはBag of Words (BoW) 表現された文書集合を生成するための確率モデルである。BoW表現とは、文章中に現れる単語のベクトル表現である。また、BoW表現は文章の構造は無視しており、単語の出現回数と共起性を表している。LDAは、BoWから得られる単語の共起性を用いて単語と文書をクラスタリングする手法として用いられる。

LDAでは、文書中の各単語に、BoWからは直接得ることのできない潜在変数(トピック)を仮定する。また、LDAの特徴として、文書は複数のトピックから構成され、トピックの構成比としての確率分布をもつ。具体的には、文書 d の i 番目の単語を $w_{d,i}$ として、対応する潜在変数を $z_{d,i}$ と定義する。ここで、トピック数を K とし、 $\theta_{d,k}$ ($k=1, 2, \dots, K$)を文章 d でトピック k が出現する確率とする。トピック分布は $\theta_d = (\theta_{d,1}, \dots, \theta_{d,K})$ となる。また、各トピックはそれぞれに対応した単語の出現分布 ϕ_k ($k=1, 2, \dots, K$)を有している。文書数を D 、文書 d の文章長(総単語数)を N_d とする。 $\phi_{k,v}$ をトピック k における単語 v の出現確率とし、単語の出現分布を $\phi_k = (\phi_{k,1}, \dots, \phi_{k,V})$ とする。 θ_d や ϕ_k はDirichlet分布による生成を仮定するので、以下のように整理できる。

$$\theta_d \sim \text{Dir}(\alpha), d=1, \dots, M \quad (1)$$

$$\phi_k \sim \text{Dir}(\beta), k=1, \dots, K \quad (2)$$

ここで、ハイパーパラメータ α 、 β はそれぞれトピック数 K 、単語数 V の次元をもつ。潜在トピックと各単語は以下のように生成される。

$$z_{d,i} \sim \text{Multi}(\theta_d), i=1, \dots, N_d \quad (3)$$

$$w_{d,i} \sim \text{Multi}(\phi_{z_{d,i}}), i=1, \dots, N_d \quad (4)$$

(2) 回遊行動データへのトピックモデルの適用

回遊行動データにトピックモデルを適用するにあたり、本研究では分析対象エリアをメッシュ化したうえで、文書 d をトリップチェーン、単語 w をメッシュ、文書 d における単語 w の出現回数 $n_{w,d}$ をメッシュ内で測位されるポイント数と考える。図-1にGPSより得られる回遊行動データをBoW表現化する際のイメージ図を示す。即ちここでは、 w はメッシュ、 d は回遊行動データ、 D は総トリップチェーン数、 V は分析対象地域内の総メッシュ数である。この考え方は、古屋ら⁴⁾の手法に従ってはいるが、古屋らが滞在状態のデータのみを用いたのに対して、本

研究においては移動状態のデータもトピックモデルに入力することとする。

なお、LDAではトピック(行動パターン)間の独立を前提としている。実際の行動パターン間には空間的な相関が存在すると思われるが、本研究においては行動パターン間には相関はないものと仮定して分析を行う。

3. 分析の対象

分析の対象とするのは、2013年11月から12月の土・日曜の計6日間に熊本市都心部で実施された、スマホ・アプリを用いた回遊調査である。参加者は回遊を開始するまでに専用アプリを個人所有のスマホにインストールする。回遊開始とともにアプリを起動することで位置情報を取得し、回遊軌跡を記録する。その間、参加者は中心市街地を普段通りに回遊し、帰宅前にポートと呼ぶ受付(対象地に4箇所設置)でアンケートに回答し、調査参加の謝礼(商品券500円)を受け取る。調査で得られる情報として、GPSの軌跡に加えて、アプリ利用の登録時の情報から、調査参加者の性別、年齢、就業状況、居住地、アンケート調査からは、来街目的、来街交通手段、来街場所、同行者等がある。参加者の属性分布などを含めた調査の詳細は別途¹³⁾¹⁴⁾を参照されたい。

本調査に参加し、ポートでのアンケートに回答したのは6日間で延べ1,086サンプルであった。このうち、分析に用いるのはAndroid端末かつ分析対象エリア内で1ポイント以上の位置情報が測位された705サンプルである。Android端末のGPSが10秒毎に位置情報を測位するのにに対し、iOS端末のGPSは10mの移動毎に位置情報を測位する。本研究におけるトピックモデルでは、位置情報の共起性と観測回数に基づいて行動パターンの抽出を行う。この場合、回遊行動パターン中で滞在・滞留地点を表現するには、端末が静止していても位置情報が測位されることが望ましい。このため、本研究では、分析対象をAndroid端末のみに限定して分析を行った。

図2に分析対象エリアを示す。本研究では、分析対象エリアを50mメッシュに分割した。熊本市都心部は2核3モジュールで構成される。2核は大型商業施設が集中する

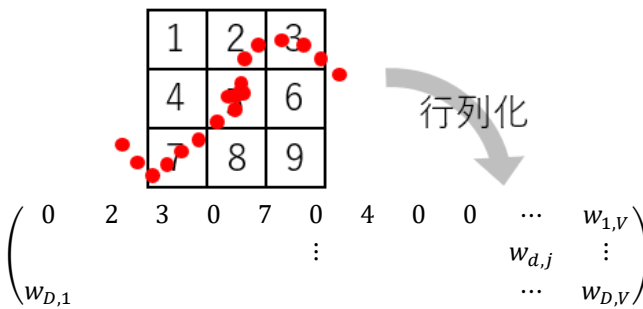


図-1 回遊行動データのBoW表現

商業地区の通町筋地区と、バスターミナルの交通センターを中心とした桜町・花畑地区である。3モジュールは、上通、下通、新市街のアーケード街で構成され、その沿道には飲食店や買い物施設が多く立地している。

4. 分析結果と考察

(1) トピック(回遊パターン)の抽出結果

サンプリング方法には、崩壊型ギブスサンプリングを用いた。また、ハイパーパラメータは一律に $\alpha=(0.1, \dots, 0.1)$ 、 $\beta=(0.1, \dots, 0.1)$ 、サンプリング数は1000回とし、トピック(回遊行動パターン)数は10とした。図-3から図-13に回遊行動パターンの抽出結果を示す。なお、各メッシュは $\phi_k=(\phi_{k,1}, \dots, \phi_{k,V})$ に基づき色付けされている。 $\phi_{k,l}$ はトピック k における単語 v の出現確率を示す。回遊行動データの分析においては、回遊行動パターン k に対するメッシュの寄与率と解釈できる。この寄与率が高いと滞在・滞留、寄与率が低いと移動に近い行動であると考えられる。

図-3では、交通センターと新市街を中心に回遊行動パターンが抽出されている。特に交通センターの正面と新市街アーケード内で寄与率が高くなっている。交通セン

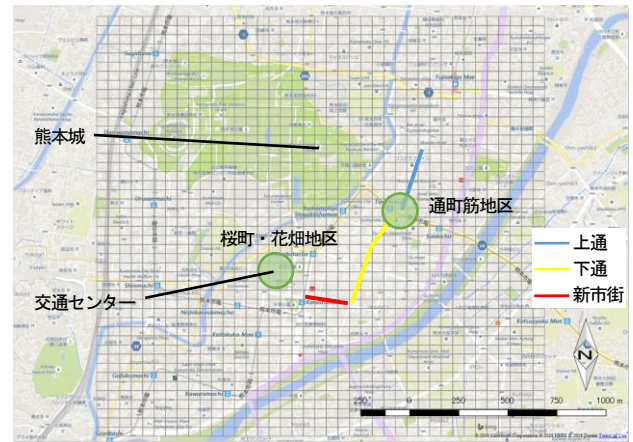


図-2 分析対象エリア

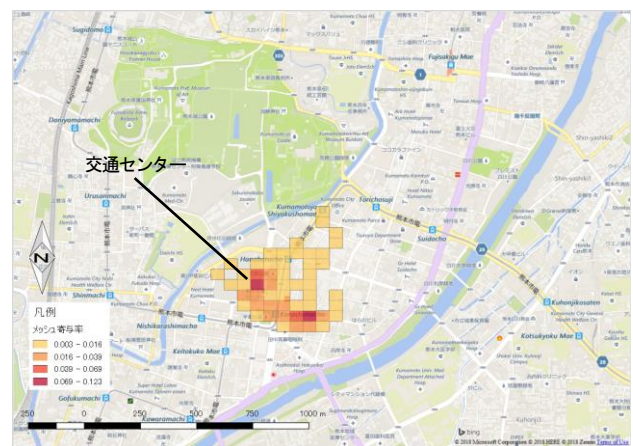


図-3 回遊行動パターン(1)

ターは熊本市都心部のバスターミナルである。よって、この回遊パターンは、バス等によって来街、その後交通センターから新市街へ移動することを示していると推察される。

図-4は、下通アーケードから新市街アーケード、また下通南側も広がる回遊行動パターンである。最も寄与率が高いメッシュは旧ダイエー前のメッシュとなっている。これは旧ダイエー前が待ち合わせ場所としてよく利用されることが原因と考えられる。また、アーケード内の通りの一部も比較的高い寄与率となっていることから、ウインドウショッピングのような低速の移動があったことが考えられる。

図-5では、国道3号線に沿った移動や熊本市都心部から外れた滞在・滞留が含まれていることが特徴的な回遊パターンである。都心部内では、図-4と同様に旧ダイエー前での滞在が確認できる。

図-6では、熊本城を含む回遊行動パターンが抽出できた。寄与率が高いメッシュとして、熊本城天守閣付近、城彩苑、交通センターなどが挙げられる。城彩苑を含むメッシュの寄与率が最も高くなっているのは、同施設で

は食事も可能なため、他と比較して長時間の滞在があったことが理由と考えられる。交通センターから城彩苑や熊本城への移動が確認できるため、このパターンに分類されるサンプルは来街にバスを利用したことが推察される。また、熊本市内における代表的な観光ルートでもあるため、熊本県外出身者も多いことも予想される。

図-7は、上通と下通の両方を含むが、回遊範囲は比較的狭い回遊行動パターンとなっている。回遊範囲こそ小さいが、この範囲には大型の商業施設が集中しているため、買い物や食事目的の長時間の滞在があることが推察される。

図-8は、移動と考えられるメッシュが多く含まれたパターンである。しかし、交通センターや新市街より南のメッシュでは滞在・滞留と考えられるメッシュも存在する。例えば、熊本市電辛島町駅が含まれており、これは電車待ちを表現していると考えられる。この電停から西のメッシュで寄与率が高くなっているが、ここには調査当時百貨店が存在していた。従って、このメッシュでは、百貨店での買い物による滞在が表現されているものと推察される。また、辛島町駅から南に寄与率が比較的高い

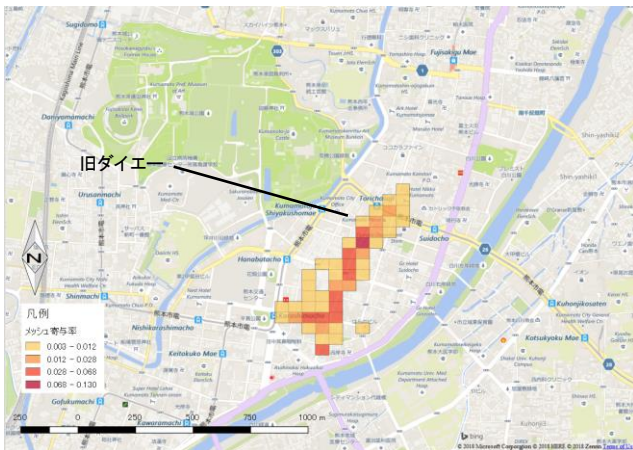


図-4 回遊行動パターン(2)

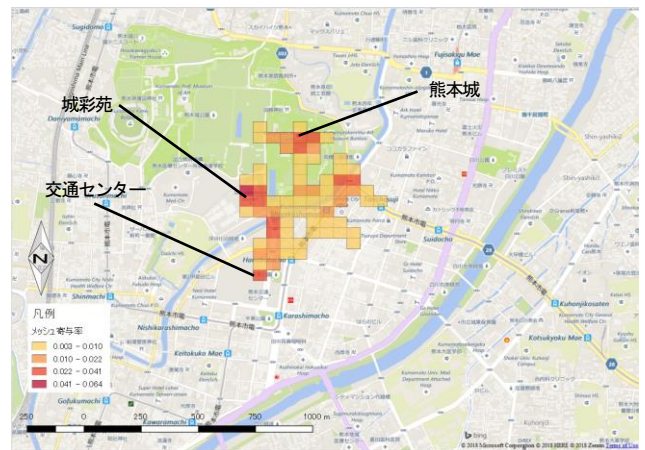


図-6 回遊行動パターン(4)

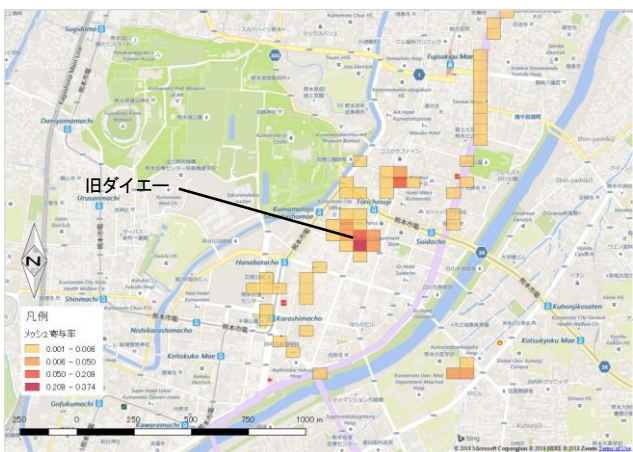


図-5 回遊行動パターン(3)



図-7 回遊行動パターン(5)

メッシュがある。これは、国道 3 号線の三叉路を含んでいるため、信号待ちによる滞在と考えられる。回遊行動データのパターン抽出の際には、滞在の状態を都市や観光地の魅力の尺度として解釈することが多い。しかし、本研究では、滞在状態のデータのみではなく、移動の状態のデータも分析に用いている。よって、信号待ちのように本来改善されるべき滞留の把握ができた。

図9は、通町筋スクランブル交差点と水通町駅付近を中心に滞在・滞留が確認できるパターンである。前者は信号待ち、後者は、熊本市電の路線が敷かれている熊本県道 28 号線上のメッシュも抽出されていることから、電停における電車待ちや乗降と予想される。水通町駅付近には国道 3 号線と県道 28 号線の交差点もあるので、信号待ちによる滞留とも考えられる。大型商業施設が集中する上通・下通両アーケード入り口付近にも滞在と考えられるメッシュが広がっていることから、買い物や食事目的の来街が抽出されていると推察される。

図-10 は、下通の北側を中心とした滞在と見られるパターンである。図4や図5と同じく待ち合わせ場所としてよく利用される旧ダイエー前の他、その付近の商業施設への滞在が考えられるが、カラオケ店やゲームセンタ

一等も立地するエリアであるため、娯楽目的の来街とも考えられる。

図-11 は上通アーケード内の一部のメッシュが抽出されたが、極めて狭い範囲に限られている。その他では、抽出されたメッシュが広範囲に点在していることが特徴的である。

図-12 では、上通アーケード全域とその周辺に回遊行動が現れたパターンである。特にアーケードの北側では、



図-10 回遊行動パターン (8)

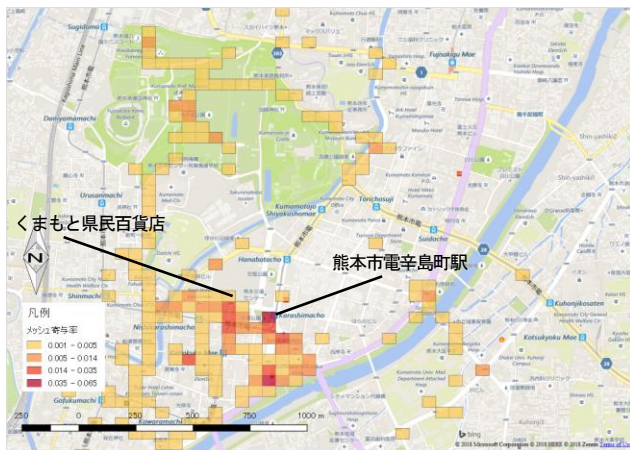


図-8 回遊行動パターン (6)

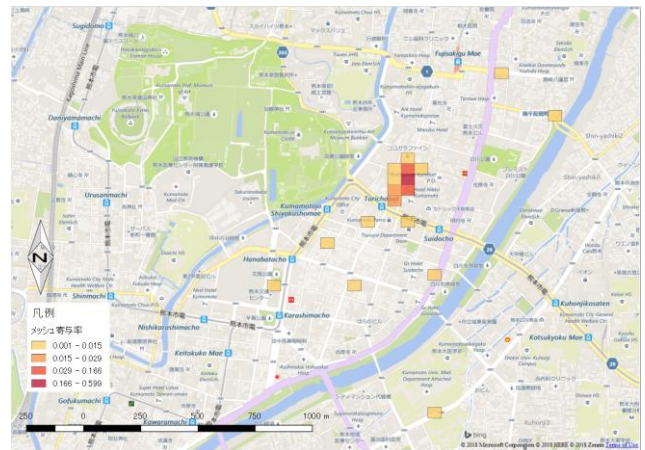


図-11 回遊行動パターン (9)



図-9 回遊行動パターン (7)

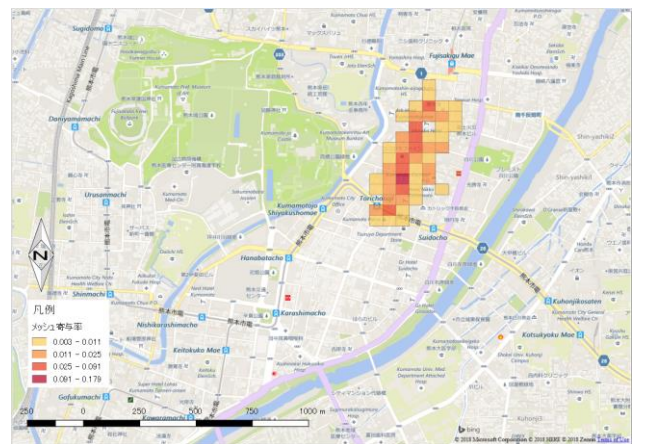


図-12 回遊行動パターン (10)

既往研究⁹⁾より、散歩・散策目的での回遊行動がある都指摘されている。さらに、このパターンでは沿道の施設への立ち寄りとして取れるため、多様な目的での来街が考えられる。

(2) トピック (回遊パターン) 別の個人・トリップ属性分析

本スマホ調査の特徴として個人・トリップ属性が取得されていることがある。そこで、この個人・トリップ属性とトピック (回遊パターン) の関係を分析する。文章 d でトピック k が出現する確率 $\theta_{d,k}$ は、本研究において、サンプル d が回遊行動パターン k に分類される確率と捉えることが出来る。ここでは、 $\theta_d = (\theta_{d,1}, \dots, \theta_{d,k})$ のうち、最も確率 $\theta_{d,k}$ が高い回遊行動パターン k ($\theta_{d,k} \geq 50\%$ となる k) を優勢パターンと設定し、この優勢パターンと個人・トリップ属性の関係を分析する。なお、705サンプル中3サンプルでは、 $\theta_{d,k} = 50\%$ となるパターンが2つ存在している。このため、この3サンプルでは、 $\theta_{d,k} = 50\%$ の2パターンをそれぞれ優勢パターンとした。よって、以後は断りのない限り708サンプルとして分析を行っている。

図-13に回遊行動パターン別の来街手段を示す。交通遮断別に見ると、パターン (1) やパターン (4)、パターン (6) でバスの割合が高くなっている。図-3、図-4、図-8を見ると、どのパターンも交通センターを含むメッシュの寄与率が高くなっていることから、これらのパターンではバスで来街し、交通センターを起点に回遊行動を行ったことが改めて示唆された。熊本市電に着目すると、熊本市電による来街が多いと予想されたパターン (7) (図-9) で市電による来街割合が低くなっており、さらにパターン (6) (図-8) に関しては、市電で来街するサンプルは存在しなかった。本節の分析では優勢パターンを設定しているため、 $0 < \theta_{d,6} < 50\%$ となるサンプルは全てパターン (6) 以外のパターンに集計されている。従って、 $0 < \theta_{d,6} < 50\%$ となるサンプルに熊本市電により来街したサンプルが含まれることが考えられる。

図-14には、回遊行動パターン別の来街目的を示した。なお、来街目的は1サンプルあたり最大3つまで得ている。来街目的が2以上のサンプルは、別サンプルとして集計してある。この結果サンプル数は $N=786$ となった。図-14を見ると、パターン (4) 以外では、買い物が主要な目的となっている。続いて割合が大きい目的は、食事と娯楽である。娯楽に関して見ると、パターン (10) において最も割合が高くなっている。これは、パターン (10) (図-12) が娯楽施設が集中するエリアでの回遊と考えられるという考察に対応している。また、パターン (4) では他のパターンと比較して観光が特徴的となっている。これは図-6に示した回遊パターンが観光ルートと考えられるという考察に整合している。

図-15は回遊行動パターン別の居住地である。パターン (4) では、熊本県外からの来街者割合が他のパターンと比較して大きくなっている。これも図-14の考察と同じく、図-6に関する考察に合致していることがわかる。

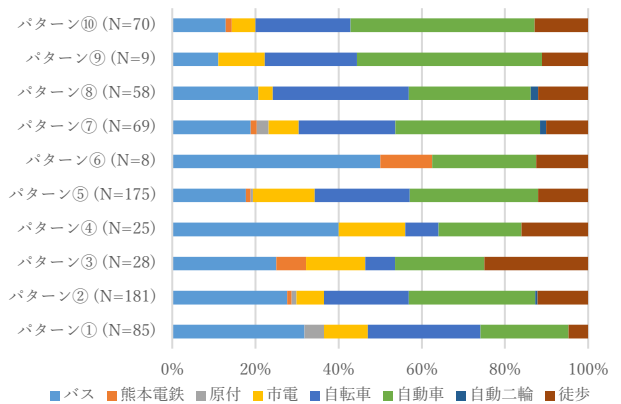


図-13 回遊行動パターン別の来街手段

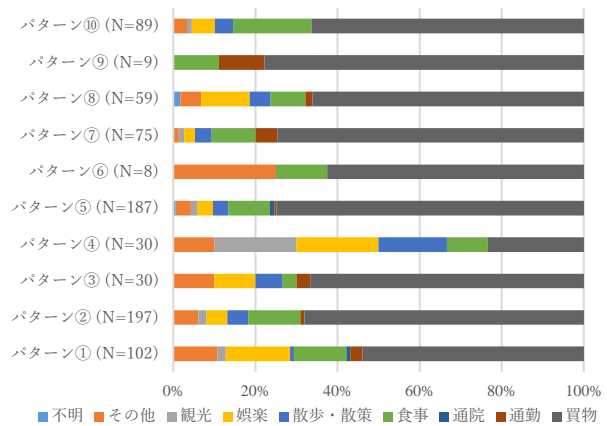


図-14 回遊行動パターン別の来街目的

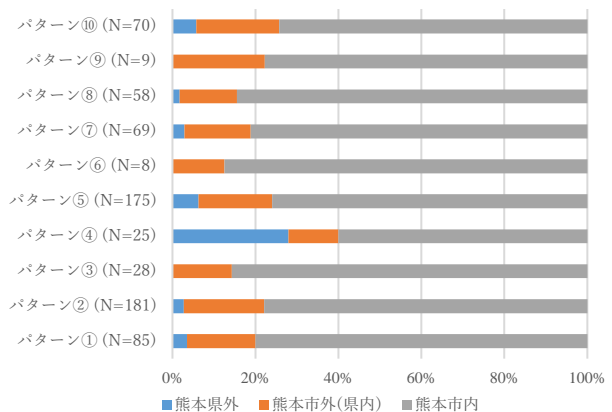


図-15 回遊行動パターン別の居住地

5. おわりに

本研究では、熊本市都心部で実施されたスマホアプリ

型回遊調査のデータを対象にトピックモデルを用いた分析を行った。具体的には、テキスト分析における文書と単語のそれぞれをトリップチェーンとメッシュと捉えてトピックモデルを適用し、10の回遊行動パターンを抽出して解釈を行った。この回遊パターンと個人・トリップ属性の関係を分析すると、抽出された回遊パターンの特徴は個人・トリップ属性の集計結果の解釈と概ね整合しており、比較的狭い範囲で取得された回遊行動データに対してもトピックモデルの有用性が確認できた。

本研究では個人・トリップ属性が判明しているデータを対象に分析を行ったが、トピックモデルに入力した情報はGPSにより得られた位置情報のみである。本研究の結果は、個人属性が判明していないデータにおいて、モデルより得られたトピック分布 $\theta_{a,k}$ (サンプル d が回遊行動パターン k に分類される確率)より、パターン抽出と同時に個人属性が推察できる可能性を示しているとも考えられる。Wi-Fiパケットセンサを始め、個人属性を取得しないデータも多く、個人情報保護の点からも個人属性の取得は困難になってきている。このような状況を踏まえると、本研究から得られた示唆は交通行動分析において重要なものであると考えられる。

今後の課題としては、最適なメッシュサイズの検討や測位誤差の影響の考慮などが挙げられる。さらなる展開として、tf-idf値のようなテキストマイニング分野で適用実績のある尺度を導入することも興味深い。また、本研究では二次元平面で回遊行動パターンを示したが、位置情報の測位時刻を用いることで、時空間の3次元のデータから回遊行動パターンを抽出することも考えられる。

参考文献

1) 佐藤貴大, 円山琢也: スマホ・アプリ型回遊調査データによる熊本都心部回遊行動圏の分析, 都市計画

- 論文集, Vol. 50-3, pp. 345-351, 2015.
- 2) 佐藤貴大, 円山琢也: カーネル密度推定法を応用したスマホ型回遊調査データの時空間分析, 都市計画論文集, Vol. 51-2, pp. 192-199, 2016.
 - 3) 矢部直人: GPS データに対する配列解析の援用, 地理情報システム学会講演論文集, Vol.19, pp.4F-3, 2010.
 - 4) 古屋秀樹, 岡本直久, 野津直樹: GPS ログデータを用いた訪日外国人旅行者の訪問パターンの分析手法の開発, 運輸政策研究, Vol. 20, pp. 20-29, 2017.
 - 5) Sun, L. and Yin, Y.: Discovering themes and trends in transportation research using topic modeling, *Transportation Research Part C: Emerging Technologies*, Vol. 77, pp. 49-66, 2017.
 - 6) Kuhn, K.D.: Using structural topic modeling to identify latent topics and trends in aviation incident reports, *Transportation Research Part C: Emerging Technologies*, Vol. 87, pp. 105-122, 2018.
 - 7) 塚井誠人, 椎野創介: 討議録に対するトピックモデルの適用, 土木学会論文集 D3 (土木計画学), Vol. 72, No. 5, 2016.
 - 8) 川野倫輝, 佐藤嘉洋, 円山琢也: トピックモデルと離散連続モデルを用いた自由記述の量的分析法, 土木学会論文集 D3 (土木計画学), Vol. 72, No. 5, 掲載予定, 2018.
 - 9) 神谷啓太, 布施孝志: トピックモデルを利用した地域別人口特性の把握手法の提案, 第 55 回土木計画学研究発表会・講演集, Vol. 55, 2017.
 - 10) 塚井誠人, 塚野裕太: トピックモデルによる詳細地理情報分析, 土木学会論文集 D3 (土木計画学), Vol. 74, No. 2, pp. 111-124, 2018.
 - 11) Blei, D.M., Ng, A.Y. and Jordan, M.I.: Latent dirichlet allocation, *Journal of Machine Learning Research*, Vol.3, pp. 993-1022, 2003.
 - 12) 岩田具治: トピックモデル, 講談社, 2015.
 - 13) 野原浩大朗, 福所誠也, 井村祥太郎, 円山琢也: スマホ・アプリを利用した熊本都心部回遊調査の分析, 第 49 回土木計画学研究発表会・講演集, Vol. 49, 2014.
 - 14) 石野祐希, 円山琢也, 溝上章志: インタビュー型回遊調査とスマホ型回遊調査の参加者属性に着目した比較分析, 都市計画論文集, Vol. 50, No. 3, pp. 331-336, 2015.

TOPIC MODELING APPROACH TO ANALYZE VISITORS' ACTIVITIES ON DOWNTOWN USING SMARTPHONE-BASED TRAVEL-SURVEY DATA

Tomoki KAWANO and Takuya MARUYAMA

It is effective to analyze GPS-based tracking data for exploring measures to revitalize the downtowns. Existing studies tried to find travel pattern in the GPS-based data and some studies applied topic modeling to this problem. However, the application of this method in small-scale area is limited. In this study, we apply topic modeling to smartphone-based travel survey data in downtown Kumamoto in Autumn 2013, and validate the topic modeling in small-scale area. We successfully found travel patterns by topic modeling and compared it with individual attributes.