

速度データを用いた歩行量の推定に関する研究

戸田 雅生¹・長曾我部 まどか²・谷本 圭志³

¹学生会員 鳥取大学大学院持続性社会創生科学研究科 (〒680-8552 鳥取市湖山町南4-101)

²正会員 鳥取大学助教 工学部社会システム土木系学科 (〒680-8552 鳥取市湖山町南4-101)
E-mail:mchoso@tottori-u.ac.jp

³正会員 鳥取大学教授 工学部社会システム土木系学科 (〒680-8552 鳥取市湖山町南4-101)
E-mail: tanimoto@tottori-u.ac.jp

従来、商店街の歩行量や道路の通行量は調査員により計測されてきたが、限られた地点や日時しか実態を把握できなかった。しかし近年では、携帯電話等の情報通信端末から得られる位置情報によって、人々の行動を常時様々な地点において把握できるようになった。そこで本研究では、速度データに対しトピックモデルを適用することで、様々なエリアにおいて歩行量を推計する方法を開発し検証した。推計結果と実際の通行量調査の結果を比較し、歩行量を概ね推計できることを確認した。さらに、通行量調査が実施されていないエリアにおいて歩行量を推計できる可能性を示した。

Key Words : *Volume of Pedestrian Traffic, Velocity Data, Topic Model*

1. はじめに

新しい店舗の設置や営業時間を検討するためには、人口のみならず、歩行量や自動車の通行量といった人々の動きも把握することが重要である。自動車や歩行者の通行量を把握する方法としては、特定の地点において調査員が計測器を用いて測る方法が一般的である¹⁾。しかしながら、調査には多くの費用と時間を要するため、数多くの地点で頻繁に実施することは難しい。また、実施する時間や地域に限られることから、イベントの有無による影響を受けやすく、日常的な通行量よりも過大に観測されることもあり得る。また、通行量は日によって異なるため、限られた期間において調査を行う場合、正確な情報を入手することは困難である。以上のことから、連続して通行量を把握でき、調査費用や調査回数を抑えた方法が必要である。

近年、スマートフォンなどの携帯電話等の情報通信端末の普及により、端末利用者の位置情報を常に把握することが可能となった。人々の位置情報が蓄積されることで、人々の動きを詳細にかつ長期的に把握することも可能である。すなわち、実際に見て測るカウント調査を実施することなく、人々の位置情報から人々の通行量を把握できることが期待される。

しかしながら、人々は自動車による移動、歩行、睡眠などといったように様々な速度で行動しているため、位

置情報のみでは、その人がその地点でどのような行動を取っているのかは明らかでない。したがって、通行量を把握したい場合、取得されたデータから人々の行動を区別する必要がある。

そこで本研究は、行動ごとの移動速度の分布に着目する。例えば、自動車の移動は、信号では停車し、渋滞では低速で移動し、自動車専用道路では高速で移動しており、行動に応じた速度分布があると考えられる。また、時間帯やエリアによって、行動の種類や量が異なると考えられる。すなわち、行動ごとの速度の分布と時間帯やエリアごとの行動の分布を推計する必要がある。

以上より、本研究はポイント型流動人口データの速度データを用いてエリアにおける人々の行動のパターンを抽出する手法を開発し検証する。自然言語処理の分野で用いられるトピックモデルの手法を用いることで、ある日時、エリアの行動を推定するアプローチをとる。その上で、推定した結果を実測データと比較し、推計の妥当性を実証的に検証する。

2. 本研究の位置づけ

通行量の把握に関する研究として、庭川ら²⁾は、監視カメラの映像を使用して、人々の混雑度を映像解析により分析した。長田ら³⁾は通行量の連続的な調査にあたり

受動赤外線型自動計測器を設置し、カウント値の妥当性を検証した上で通行量の変動を分析した。これらの研究では一部のエリアや調査の範囲が限られているため、歩行量など人々の動きを十分に把握することが難しい。

近年、情報通信端末から得られた位置情報を用いて、人々の行動を把握している研究も多く見られる。佐藤・丸山⁹⁾は、スマホアプリから得られた位置情報に対し、カーネル密度推定法を適用し、簡易的に人々の回遊行動圏を推定する方法を提案している。また筆者ら⁵⁾は、移動速度データから被災前後の人々の移動速度を推計し、被害が地域において移動速度が低下すること、移動速度の回復過程には自治体ごと異なることを明らかにした。しかし、人々のどのような行動に影響を与えたのかは明らかになっていない。

トピックモデルを用いた研究として、塚井・椎野⁶⁾は、公共交通に関する議事録にモデルを適用し、比較的語数の少ない文章においてもモデルが適用できることを示した。抽出されたトピックの元文書の討議内容との対応は良好であり、細かいトピック推移の抽出が可能であると述べている。さらに塚井・塚野⁷⁾は、3次元メッシュ多属性データに因子分析とトピックモデルを適用して、詳細地理特性の抽出性能を比較し検証している。このように文書解析やそれ以外の分野へ適用して研究していることから、トピックモデルは文書以外の分析においても適用可能であることが分かる。

そこで本研究では、人々の行動を把握するために人々の移動速度データに着目し、速度データにトピックモデルを適用することで、特徴的な行動を抽出し、その行動がエリアや時間帯に占める量を推計する。推計した結果を実測データと比較し、推計の妥当性を検証する。

3. 分析手法

(1) 分析の流れ

本研究ではトピックモデルを用いて人々の移動速度のデータから行動を分類し、ある時間帯のあるエリアにおける歩行量を推計する。トピックモデルは、文書を分析する手法であり、文書のトピック分布とトピックの単語分布を確率的に求めるモデルである⁸⁾。分析の流れを図1に示す。まず、速度データにトピックモデルを適用し、3つの行動に分類する。次に、推計された歩行量と実際の通行量データを比較し、モデルの妥当性を検証する。最後に、通行量調査が行われていないエリアにおいて歩行量の推計を行い、平日と休日、イベントの有無による歩行量の差異を明らかにする。

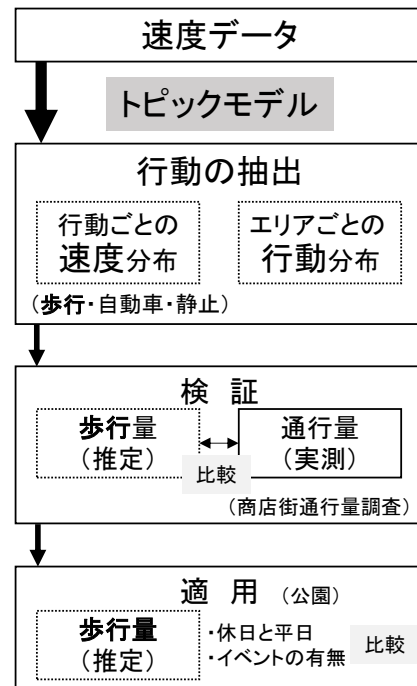


図1 分析の流れ

(2) トピックモデル

トピックモデルを推定するにあたり、潜在ディリクレ配分モデル(Latent Dirichlet allocation, LDA)を用いる^{9),10)}。LDAでは文書 d のトピック割合 θ_k とトピック k における単語の出現確率 ϕ_k をディリクレ分布によって求める。ここで、ディリクレ分布は、 $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_k)$ ($\alpha_k > 0$)をパラメータとして式(1)と定義される。また、 $\Gamma()$ はガンマ関数である。

$$p(\theta_d | \alpha) = \frac{\Gamma(\sum_{i=1}^k \alpha_i)}{\prod_{i=1}^k \Gamma(\alpha_i)} \prod_{k=1}^K \theta_k^{\alpha_k - 1} \quad (1)$$

同様に、単語の出現確率 ϕ_k は文書 d の単語数 N_d を用いて、 $\beta(k \times N_d)$ をパラメータにもつディリクレ分布から求めることができる。LDAでは、トピック割合 θ_k に応じて文書 d における1つの単語 w_{dn} に対して、1つの潜在トピック z_{dn} が割り当てられる。この潜在トピック z_{dn} とトピック k の単語の出現確率 ϕ_k により、単語 w_{dn} が生成される。パラメータ α, β が与えられたときのトピック混合分布 θ_k の同時分布と N 個のトピック z の集合と N_d 個の単語 w の集合を式(2)に表す。

$$p(\theta_d, z, w | \alpha, \beta) = p(\theta_d | \alpha) \prod_{n=1}^{N_d} p(z_{dn} | \theta) p(w_{dn} | z_{dn}, \beta) \quad (2)$$

式(3)は式(2)を θ_d で積分し、 z_{dn} で和をとったもので、これが文書 d の周辺分布となる。

$$p(w|\alpha, \beta) = \int p(\theta_d|\alpha) \left(\prod_{n=1}^{N_d} \sum_{z_{dn}} p(z_{dn}|\theta_d) p(w_{dn}|z_{dn}, \beta) \right) d\theta_d \quad (3)$$

式(3)においては、変分ベイズ推定を用いて近似計算を行う。

4. 分析結果

(1) データの概要

株式会社Agoopの「ポイント型流動人口データ」¹¹⁾を用いる。本データは、Agoopが提供するスマートフォン向けアプリケーションのユーザーのうち、アクセス権限を得たユーザーのGPS位置情報により作成されたものである。データには、時間、緯度・経度、移動速度、移動方向などが含まれる。人々の詳細な行動データが膨大にあることから、人々の行動の傾向を把握することに適している。

分析対象は、2016年4月に熊本県で取得されたデータとする。データ件数は1日あたり約10万件~13万件である。ここで、トピックモデルで用いる単語データが離散量であるのに対して、移動速度データは連続量であることから、移動速度 v を v_1 , v_2 , v_3 の3つに分類した。それぞれ、 v_1 は0km/h、 v_2 は0km/hより大きく6km/h以下、 v_3 は6km/hより大きい。歩行者の移動速度の平均が約4km/hから7km/hであることから v_2 と v_3 を区分する値を6km/hとし、 v_2 と v_3 で歩行者と自家用車を区別できる値とした。文書 d については、1時間当たりの3次メッシュとした。

また、推計値と比較するデータとして、熊本商工会議所の商店街通行量調査¹²⁾のデータを使用した。本調査は、毎年10月に熊本市内37地点の日曜と平日の通行量を計測している。調査の概要を表1に示す。

(2) 推計結果

トピック数 $K=3$ として30日間分の推計を行った。4月8日(金)におけるトピックごとの速度分布を図2に示す。トピック1は v_1 が多くを占めており、つまりほとんど動いていない行動を示す。トピック2は v_1 と v_2 が多く、速度0と低速で構成される行動である。トピック3は v_3 の割合が大きいことから、歩行速度以上の速度で動いている行動である。これらの結果からトピック1, トピック2, トピック3をそれぞれ「静止」、「歩行者」、「自動車」と解釈した。

表 1 通行量調査の概要¹²⁾

調査日時	平成 28 年 10 月 14 日 (金) 10 月 16 日 (日)
調査時間	午前 8 時~午後 8 時 (1 時間毎)
調査地点	熊本市内商店街 37 地点 (中心市街地 26 地点、熊本駅と周辺部 11 地点)
調査対象	商店街内の歩行者及び自転車通行者
実施主体	熊本商工会議所・熊本市

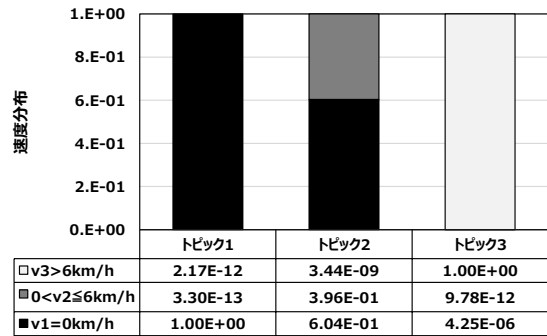


図 2 トピックごとの速度分布

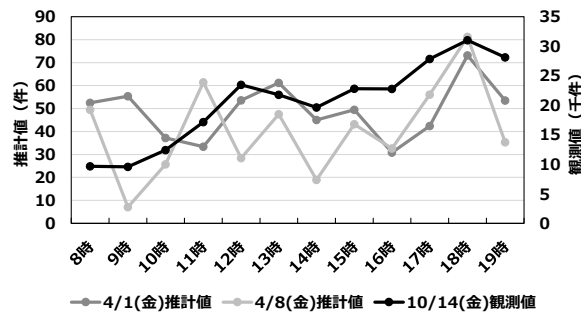


図 3 推計値と観測値の比較 (銀座通り周辺)

と解釈した。

(3) 検証結果

トピックモデルにより推計された歩行量と実際の通行量を比較した結果を示す。ここでは、熊本商工会議所の通行量調査における銀座通り周辺を対象とする。比較するにあたり1) 通行量調査では歩行者だけではなく自転車通行者も対象とされている点、2) 推計値は銀座通り周辺を含む3次メッシュにおける歩行量である点に留意する必要がある。推計値は取得されたデータ数にトピックモデルより推計された歩行割合をかけた件数を示す。

銀座通り周辺の金曜日の歩行量を比較した結果を図3示す。観測値は、8時から12時にかけて通行量が増加しており、一度減少した後に、16時から18時にかけて再び増加していることが見てとれる。推計値に着目すると、4月1日については、11時から13時と16時から18時にかけて歩行量が増加しており、4月8日については、9時から11時と16時から18時にかけて増加していることが分かる。全体的な傾向として、朝から昼にかけて歩行量は一度増

加し、その後夕方まで減少した後、再び夜にかけて増加していることが見てとれる。観測値と同様に、推計値もこの傾向が見られることから、速度データより歩行量を推計できたと考えられる。

(4) 他のエリアへの適用

イベントや季節の影響によって、歩行量が変化する可能性が高い地域を選定し、歩行量の推計を行った。桜の名所である菊池公園を対象とし、平日と土曜日、土曜日の1週間前後を比較する。図4に結果を示す。まず、平日（4月5日）と休日（4月2日と9日）を比較すると、休日は10時から12時の歩行量が多いことが見てとれる。次に、休日の1週間前後（4月2日と9日）を比較すると、4月2日は15時から21時の歩行量が多いことが見てとれる。これは平日と比較しても多い。2016年における熊本県の桜の開花日は3月22日、満開日は4月2日であり¹³⁾、かつ満開期間は5日～6日と言われていることから、4月2日（土）と9日（土）には桜の開花というイベントの有無により歩行者量に変化があったと推察される。つまり夜間に桜を見に訪れる人を把握することができたと考えられる。

以上より、平日では一日を通して、大きな歩行量の変化は見られなかった。また、休日間でも日にちによって大きな変化が見られた。イベントの有無により平日と休日、休日間とで歩行量に変化があることが、トピックモデルの推計により確認できた。

5. おわりに

本研究は、情報通信端末から取得されるポイント型流動人口型データにおける移動速度データに対し、トピックモデルを適用することで歩行量を推計した。推計結果と実際の通行量調査結果との比較により、おおむね歩行量を推計できたことが確認された。さらに、本手法において、歩行量調査が実施されないエリアにおいて、歩行量を推計できる可能性を示した。

参考文献

- 1) 鈴木英之：中心市街地活性化指標の検討-通行量調査を代替もしくは補完する新しい方法について-，日本経営診断学会論集，Vol.15, pp. 16-22, 2015.
- 2) 庭川誠，中村英樹：交通流解析のためのビデオ画像処理システム Traffic Analyzer の開発と性能検証，電気学会論文集 D（産業応用部門誌），Vol.124, No.10,

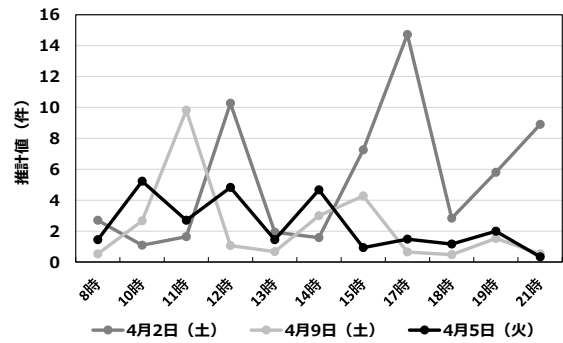


図4 推計値の比較（菊池公園周辺）

- pp.1060-1066, 2004.
- 3) 長田哲平，加納壮貴，大森宜暁，古池弘隆：中心都市街地における受動赤外線型自動計測器を用いた歩行者通行量の分析，交通工学論文集，Vol.4, No.1, pp.B_38-B45, 2018.
- 4) 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, No.2, pp. 192-199, 2016.
- 5) 長曾我部まどか，谷本圭志，前田夏輝：被災地における住民行動の回復過程に関する分析一人々の移動速度に着目して一，土木学会論文集 F6（安全問題），Vol.74, No.2, pp. I_101-I_109, 2018.
- 6) 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, pp. I_341-I_352, 2016.
- 7) 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.
- 8) 岩田具治：トピックモデル，講談社，2016.
- 9) 佐藤一誠：トピックモデルによる統計的潜在意味解析，コロナ社，2015.
- 10) David M. Blei, Andrew Y. Ng and Michael I. Jordan: Latent Dirichlet Allocation, Journal of Machine Learning Research, Vol.3, pp. 993-1022, 2003.
- 11) 株式会社 Agoop 社：流動人口データ，<https://www.agoop.co.jp/floating-population/>（最終アクセス：2020年1月30日）
- 12) 熊本商工会議所：商店街通行量調査，<https://www.kmt-cci.or.jp/investigation/shopstreet.php>（最終アクセス：2020年1月30日）
- 13) 気象庁：生物季節観測の情報（2011～2019），開花日 http://www.data.jma.go.jp/sakura/data/sakura003_06.html，満開日 http://www.data.jma.go.jp/sakura/data/sakura004_06.html（最終アクセス：2020年1月30日）

(2020.3.8 受付)

STUDY ON ESTIMATION OF THE VOLUME OF PEDESTRIAN TRAFFIC BY USING VELOCITY DATA

Masaki TODA, Madoka CHOSOKABE and Keishi TANIMOTO