

速度データを用いた歩行者量の推定-熊本県を対象として-

鳥取大学	学生会員	○戸田 雅生
鳥取大学	正会員	長曾我部 まどか
鳥取大学	正会員	谷本 圭志

1. はじめに

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。近年、情報通信端末から人々の位置情報を常時様々な地点において把握できるようになった。しかし、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。そこで本研究では、人々の行動のパターンを分類し歩行者量を推計する手法を開発し検証する。

2. 本研究の考え方

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。そこで、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。移動速度データにトピックモデルを適用し特徴的な行動を抽出した上で、その行動がエリアや時間帯に占める量を推計する。

3. 分析手法

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²⁾。

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。LDAでは文書 d のトピック割合 θ_d とトピック

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。トピック割合 θ_d 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_d 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。

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

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。

キーワード トピックモデル, 速度データ, 歩行者量調査
 連絡先 〒680-8552 鳥取市湖山町南 4-101 鳥取大学工学部社会システム土木系学科 公共システム研究室
 TEL 0857-31-5007

4. 分析結果

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。本研究では、移動速度 v を v_1, v_2, v_3 ($0\text{km/h}, 0\text{km/h} < v \leq 6\text{km/h}, 6\text{km/h} < v$)の3つに区分した。文書 d については、1時間当たりの3次メッシュとした。

(1) 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1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。

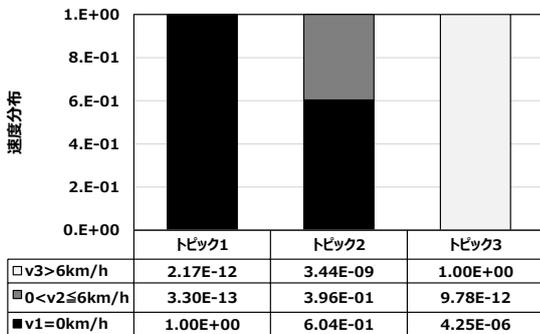


図1 トピックごとの速度分布

(2) 検証結果

推計された歩行量と実際の通行量を比較した結果を図2に示す。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。推計値は、4月1日と8日の結果である。

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、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、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。

(3) 他のエリアへの適用

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3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。

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。

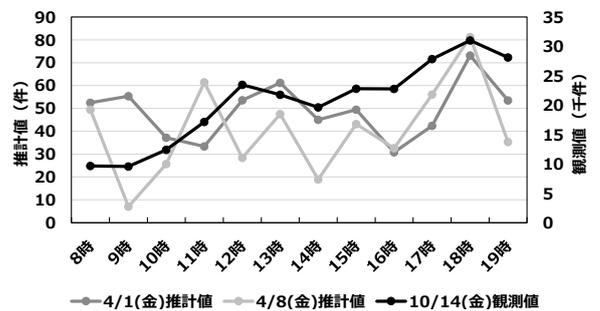


図2 推計値と観測値の比較(銀座通り周辺)

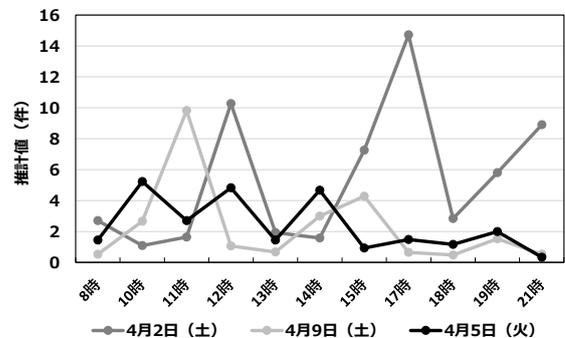


図3 推計値の比較(菊池公園周辺)

5. おわりに

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) 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）
- 2) 岩田具治：トピックモデル，講談社，2016。