

生活サービスの利用時間の推計手法に関する研究

公共システム研究室 布野ちなみ

1. はじめに

地方では急速な人口減少の進行により、撤退や縮小せざるを得ないサービスが増え、生活に必要なサービスの入手が困難になることが懸念されている。これらのサービスの持続可能性を確保するためには、サービスの供給に関する効率性を高めることが有効な取り組みとなる。その一つとして、顧客がサービスを必要とする時間に応じてサービスの供給体制を見直すことが考えられる。この時間を把握するために、従来はパーソントリップ調査などのアンケート調査で断片的に把握せざるを得なかった。しかし、近年、携帯端末の位置情報により、人々の行動が容易に把握できるようになってきた。そこで本研究では、モバイル空間統計というビッグデータを用いて、生活サービスの中でも重要なサービスである買い物に着目し、利用時間を推計する手法を検討する。

2. 本研究の基本的な考え方

トピックモデルを用いてモバイル空間統計から人々の行動パターンを抽出する。その上で、各メッシュにおける行動パターンの混入比率を被説明変数とし、生活サービスの立地を含め、メッシュでの行動に特徴をもたらすいくつかの要因を説明変数として、これらの関係を回帰分析により明らかにする。次いで、導出した回帰式を用いて、スーパーの有無別の行動パターンを再現し、その差に着目して買い物の利用時間を取り出すアプローチをとる。また、トピックモデルの適用の有無に応じた推計結果を比較し、トピックモデルを用いることの有効性を検証する。

3. 分析手法

トピックモデルでは文書ごとにトピックの分布 $\theta_d = (\theta_{d1}, \dots, \theta_{dK})$ がある。ここに、 $\theta_{dk} = p(k|\theta_d)$ は文書 d の単語にトピック $k (1 \leq k \leq K)$ が割り当てられる確率である。また、トピックごとの単語の分布のパラメータは $\phi = (\phi_1, \dots, \phi_K)$ で表され、 $\phi_k = (\phi_{k1}, \dots, \phi_{kV}) \geq 0$ はトピック k で語彙 $v (1 \leq v \leq V)$ が生成される確率である。トピック分布 θ_d と単語分布集合が与えられたときの文書 w_d の確率は式(1)のように定式化される。なお、 N_d は文書 d に含まれる単語の数である。

$$p(w_d | \theta_d, \Phi) = \prod_{n=1}^{N_d} \sum_{k=1}^K \theta_{dk} \phi_{k w_{dn}} \quad (1)$$

EM アルゴリズムを用いてパラメータを推計する。本研究では、文書 d をメッシュ、語彙 v を時刻、トピック k を行動パターンと置き換えた上でトピックモデルを適用し、 ϕ_{kv} 、 θ_{dk} を求め、行動パターンの特徴を明らかにする。

4. 分析結果

鳥取市における秋のある1日を対象として、買い物の利用時間を推計した。その結果、図 4.1 を得た。この結果の妥当性を図 4.2 の平成 28 年社会生活基本調査における買い物の行動者率と比較して検証する。トピックモデルを用いない場合はピークが3箇所あるのに対し、トピックモデルを用いた場合はピークが2箇所存在する。時間帯に多少の違いはあるが、ピークが午前、午後のみられるトピックモデルを用いた場合のほうが行動者率に近い分布となる。よって、モバイル空間統計からサービスの利用時間を推計する際にトピックモデルの適用は有効であるといえる。

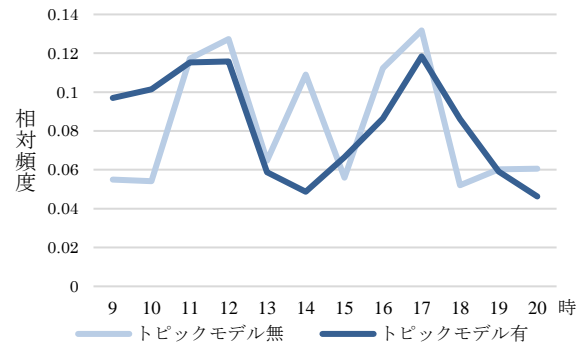


図 4.1 買い物の利用時間の推計結果

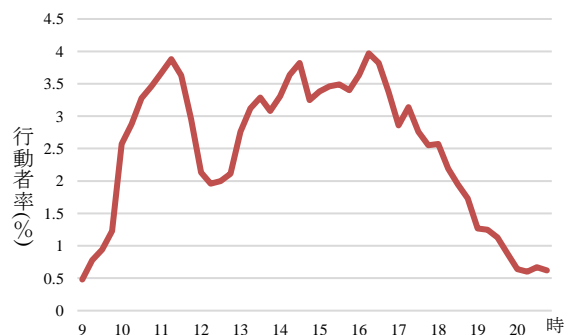


図 4.2 買い物の行動者率