

シェアサイクルにおけるメッシュごとの需要タイプ分類

諸田健太郎[†] 佐藤彰洋[‡]横浜市立大学データサイエンス学部データサイエンス学科^{†‡}

1. 研究の背景と目的

地球規模の環境問題に対処するため、自動車から公共交通や自転車への移行が促進されている。また、自転車は都心の効率的な移動手段として注目され、シェアサイクルとして各国で普及しつつある[1]。特に、シェアサイクル事業の事業性を高めるためには、需要に合わせた機体の配置を行うことが必要であり、効率的な配置を行うための需要予測が求められる。

従来研究として、Yuら(2020)[2]は人流データを用いて、メッシュレベルでシェアサイクルの潜在的な需要を把握した上で潜在的な使用パターンと潜在的な使用方法の分析を行っている。

そこで、本研究ではメッシュごとに自転車数の変動の様子に特徴があることに注目し、昼と夜の自転車数の違いから需要タイプの分類し、昼と夜の自転車数の差を定量的に予測する方法を提案する。

本研究の目的は、昼と夜の需要タイプの分類によって自転車を配備するエリアの組み合わせを定める、進出エリアの優先度を定めることである。特に、ステーションデータからシェアサイクルの需要の多い地域を特定する手法と、メッシュエリアの特性から分類と需要の予測をする手法を提案する。

2. データ

2.1 データ

使用するデータは、一般社団法人公共交通オープンデータセンター[3]が提供する、2023/06/01~2023/06/30の東京都全域（一部神奈川県を含む）におけるステーションの情報である。このデータの更新頻度は5分ごとの時系列データであり、JSONファイル形式で提供されている。天気データ、メッシュに関する人口や世帯数、人口密度、コンビニの数やATMの数などはMESHSTATS Web API[4]を使用した。駅の数には駅データ[5]から、全国の駅名と位置情報が付随しているデータを用いた。また、位置情報に紐づくデータは3次メッシュコード[6]に変換して使用している。

2.2 分析手法

分析手法については大きく分けて3つのステッ

プに分けられる。1つ目は自転車数の推移から、メッシュにt検定を用いて、「昼間集中型メッシュ」か「夜間集中型メッシュ」か「その他」に分類する。2つ目は、自転車数の推移から分類された結果を被説明変数と見立てて、メッシュごとの地理情報や人口のデータを用いて、ロジスティック回帰で予測分類を行うことである。3つ目は、ロジスティック回帰で用いたような説明変数を用いた上で、メッシュごとの昼夜の自転車数の差を重回帰分析で求めることである。

3. 研究結果

3.1 t検定を用いたメッシュの需要タイプ分類

まず昼の時間を10:00~17:59、夜の時間を21:00~4:59として自転車台数を抽出する。その中でも1ヶ月間で平日かつ天気が晴れかつ曇りの日のデータを用いた。その上で、昼間と夜間の自転車台数の標本平均の差の検定を行う。有意水準5%で棄却されたメッシュを対象にして、昼の方が平均の高いメッシュを「昼間集中型メッシュ」とし、夜の方が平均の高いメッシュを「夜間集中型メッシュ」として分類する。その結果、分類されたものは表1の通りとなる。

表1: 2023年6月1カ月分の3次メッシュごとのt検定の結果

昼間集中型メッシュ	夜間集中型メッシュ	棄却されず	合計
118	206	12	336

3.2 ロジスティック回帰モデルを用いた昼夜の需要タイプ分類予測モデルの作成

説明変数としては、MESHSTATS APIから人口や世帯数、人口密度などの99種類のデータとOpenStreetMap APIからポイントデータを取得して集計を行った、コンビニの数やATMの数などの10種類のデータと、駅の数とを合わせた110列を標準化して使用した。

ロジスティック回帰の計算方法については以下のように定義する。世界メッシュコード m における説明変数ベクトル \mathbf{x}_m と重みベクトル \mathbf{w}_m を用いて、 y_m が昼間集中型メッシュになる確率を p_m とすると y_m は(2)式のように求める。

$$p_m = \frac{1}{1+e^{-\mathbf{w}_m^T \mathbf{x}_m}}, y_m = \begin{cases} 1 & (p_m \geq 0.5) \\ 0 & (p_m < 0.5) \end{cases} \quad (2)$$

ここで、 $\mathbf{x}_m = (a_m, x_{1m}, x_{2m}, \dots, x_{110m})^T$ とする。ここでTは転置を意味する。重みの更新についてはニュートン-ラフソン法を用いた。定義は(3)式の通りになる。

$$\mathbf{w}_{new} = \mathbf{w}_{old} - \mathbf{H}^{-1}\nabla E(\mathbf{w}_{old}) \quad (3)$$

ここで、 \mathbf{H} はヘッセ行列、 $\nabla E(\mathbf{w})$ はスコア関数とする。変数選択はステップワイズ法を用いた。

今回、学習データは東京都に存在するメッシュ294個で、評価データは神奈川県に存在するメッシュ21個である。

ここで、 z は係数を標準誤差で割ったものとし、 $P|z|$ は「推定値が0である」ことを帰無仮説としたときのP値とする。VIFは定数項以外の説明変数でVIF統計量を計算した値である。推定される係数、p値、VIFを表2に示す。

表 2: ロジスティック回帰の結果

変数名	係数	標準偏差	Z	P> z	VIF
定数項	-0.991	0.176	-5.643	0.000	-
変数1	-1.808	0.230	-7.854	0.000	1.022
変数2	0.794	0.230	3.255	0.001	1.227
変数3	0.354	0.175	2.021	0.043	1.232
変数4	-0.143	0.173	-0.830	0.406	1.026

変数1は「2015年0~14歳男性人口(人)」、変数2は「医療機関数」、変数3は「ATM数」、変数4は「最低標高(m)」とする。このモデルの学習データのAccuracyは0.843で、評価データのAccuracyは0.681であった。AICは243.40で、BICは261.81となった。

「最低標高(m)」以外の説明変数に関しては、係数が0であるという帰無仮説を有意水準5%で棄却できる。また、どの変数もVIFが10を下回っているため、多重共線性の疑いも少ないと言える。

3.3 重回帰分析での昼夜の自転車数の差分

変数選択ではステップワイズ法と、VIF統計量が10を超えないような変数を選んで用いた。ここで、 t は係数を標準誤差で割ったものとし、 $P|t|$ は「推定値が0である」ことを帰無仮説としたときのp値とする。

表 3 重回帰分析の結果

変数名	係数	標準偏差	T	P> t	VIF
定数項	13.932	4.622	3.014	0.003	-
変数3	2.841	1.843	1.542	0.124	1.712
変数5	0.025	0.000	7.319	0.000	7.580
変数6	0.025	0.003	8.707	0.000	9.680
変数7	0.032	0.009	3.697	0.000	9.639
変数8	0.010	0.003	3.256	0.001	9.824

変数5は「2014年労働者数(人)」、変数6は「0歳以上14歳以下2050年将来推計人口」、変数7は「2020年世帯数(世帯)」、変数8は「2014年事

業所数」とする。

このモデルの決定係数は0.471で、AICは2844.689となった。学習データのRMSEは29.920で、評価データは18.715となった。変数3の「ATM数」以外の説明変数に関しては、「係数が0」という帰無仮説を有意水準5%で棄却できる。また、どの変数もVIFが10を下回っているため、多重共線性の疑いも少ない。

4. 結論

本研究では、メッシュごとに昼と夜の需要タイプの分類を行い、t検定を用いた分類、ロジスティック回帰での複数の説明変数からメッシュ特徴を分類するモデルの作成、重回帰分析を用いた昼夜の平均の差の予測を行った。

今後の課題としては、新たな土地へのメッシュの適合度合いを検証することと、メッシュごとの需要をより高い精度で予測することである。

参考文献

- [1] 国土交通省, 環境にやさしい「自転車社会」の未来地図, GOOD CYCLE JAPAN, 2023/04/05, <https://www.mlit.go.jp/road/bicycleuse/good-cycle-japan/environment/>
- [2] Q. Yu, H. Zhang, W. Li, Y. Sui, X. Song, D. Yang, R. Shibasaki, W. Jiang, Mobile phone data in urban bicycle-sharing: Market-oriented sub-area division and spatial analysis on emission reduction potentials, Journal of Cleaner Production, Vol. 254 (2020) 119974
- [3] 公共交通オープンデータセンター, ドコモ・バイクシェア バイクシェア関連情報, https://ckan.odpt.org/dataset/c_bikeshare_gbfs-d-bikeshare, 2023/04/10
- [4] MESHSTATS, MESHSTATS Web API, <https://www.meshstats.org/meshstats>, 2023/11/24
- [5] 駅データ.jp, 駅データ, <https://ekidata.jp/>
- [6] 一般社団法人世界メッシュ研究所, オープンライブラリ, https://www.fttsus.org/worldgrids/ja/our_library/, 2023/11/24

Classification of Demand Types and Demand Forecasting by Mesh in Shared Bicycles

† Kentaro Morota • Yokohama City University

‡ Aki-Hiro Sato • Yokohama City University