

# バイクシェアの需要予測に向けたバッテリー残電量情報の活用

野村 歩夢<sup>†</sup> 白石 陽<sup>†</sup>

<sup>†</sup> 公立ほこだて未来大学システム情報科学部

## 1 はじめに

近年、普及している次世代モビリティの1つとして、バイクシェアサービスがある。バイクシェアサービスとは、共用の電動自転車を利用して移動し、バイクステーションで自転車を貸出・返却するサービスである。

しかし、現行のバイクシェアサービスでは、バイクステーションにおける車両偏在問題が発生している [1]。車両偏在問題とは、バイクステーションで自転車が集中・不足する現象を指す。この問題により、歩行者の進路妨害、ユーザのサービス満足度の低下に繋がる可能性がある。そのため、バイクシェアサービスの運営者は、バイクステーションにおける貸出台数を正確に予測する必要がある。

本研究では、各バイクステーションにおける貸出台数の予測を行う。予測にあたり、自転車に搭載されたバッテリー残電量情報に着目する。文献 [2] によると、バッテリー残電量情報に関する電気自転車運転者の行動変容が報告されており、バイクシェアユーザでも同様の行動変容が期待できる。バイクステーションに配置された、全ての自転車のバッテリー残電量が高い傾向を示す場合、走行可能距離に関する不安感が低下し、そのバイクステーションの利用頻度が増えるという行動変容が発生すると考えられる。

そこで本研究では、バッテリー残電量情報に関する特徴量を用いたバイクステーションの需要予測手法を提案する。本稿では、バイクステーションにおける貸出台数を予測し、その結果からバッテリー残電量情報を活用する有効性を検証する。

## 2 関連研究

バイクステーションの需要予測の関連研究として、文献 [3-5] の研究がある。文献 [3] では、天候データ、人口統計データ、運行データ等を特徴量とした予測モデル

を構築し、予測を行っている。しかし、使用データは全て外的要因であり、内的要因であるユーザの嗜好を予測モデルに反映していない。文献 [4] では、時系列モデルを用いて、予測を行っている。しかし、時系列的な特徴が出現しにくいバイクステーションを予測対象にした場合、時系列モデルを用いた予測手法は適さない。文献 [5] では、複数のバイクステーションをグルーピングし、エリア単位での予測を行っている。しかし、グルーピングにより予測対象の粒度が粗くなるため、個々のバイクステーションの車両偏在問題を解決できない。

## 3 提案手法

### 3.1 需要予測の流れ

図1に提案手法の概要を示す。

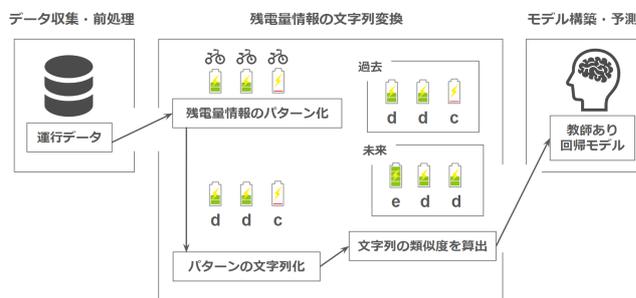


図1: 需要予測の概要

提案手法の流れとして、(1) バッテリー残電量情報を含む運行データの収集・前処理、(2) 一定期間内に、バイクステーションに配置されていた全ての自転車のバッテリー残電量情報のセットをパターン化、(3) Levenshtein距離を用いた非類似度スコアの算出、(4) 非類似度スコアの重み付け、(5) 残電量パターン類似度の算出、(6) 教師あり回帰モデル構築、(7) 貸出台数予測の手順で行う。

### 3.2 残電量パターン類似度の算出

(2) では、各バッテリーの残電量 (1~100%) を対応する文字 (アルファベット) に置換することで、パターンを文字列として表現する。文字置換の対応範囲とし

Utilization of Remaining Battery Information for Demand Prediction in Bike-Sharing

Ayumu Nomura<sup>†</sup>, Yoh Shiraishi<sup>†</sup>

<sup>†</sup>School of Systems Information Science, Future University Hakodate

て、1~15%の場合「a」、16~30%の場合「b」、31%以上は10%間隔で「c」以降のアルファベットに順次変換する。文字列化を用いて数値情報を量子化することにより、貸出時のバッテリー残電量情報の傾向を掴みやすくなる。

(3)では、予測時を表す文字列*i*と、蓄積された貸出時を表す全ての文字列 $j_n (0 \leq n \leq \infty)$ との非類似度スコアを各々算出し、そのなかで、編集距離が最短になる文字列の非類似度スコアを $\min_{i,j} \delta(i,j)$ と定義する。

(4)、(5)では、式(1)~式(3)に示す重み付け方針に従い $\min_{i,j} \delta(i,j)$ を重み付けし、残電量パターン類似度を算出する。 $TF_1$ 、 $TF_2$ は文字列の出現割合、 $F$ は予測時を表す文字列の総数、 $L$ は貸出時を表す文字列の総数、 $N$ は貸出時+非貸出時を表す文字列の総数、 $Sim$ は残電量パターン類似度を表す。

$$Sim = \min_{i,j} \delta(i,j) + TF_1 \times TF_2 \quad (1)$$

$$TF_1 = 1 - \frac{F}{L} \quad (2)$$

$$TF_2 = 1 - \frac{F}{N} \quad (3)$$

式(2)、式(3)は重み付け部分に該当し、式(2)は予測時を表す文字列が、貸出時を表す文字列の中で何回出現したかを表し、式(3)は予測時を表す文字列が、貸出時+非貸出時を表す文字列の中で何回出現したかを表す。この重み付けにより、残電量情報の類似性に加えて、バイクステーションから自転車が貸出される可能性を考慮できると考える。

### 3.3 モデル構築

モデル構築にあたり、機械学習アルゴリズムとして勾配ブースティング手法の一つであるXGBoostを採用する。モデルの評価指標として、RMSE (Root Mean Square Error) を採用する。

## 4 評価実験と考察

### 4.1 実験目的

本実験の目的は、特徴量重要度評価により、バイクステーションの需要予測にバッテリー残電量情報を活用する有効性を評価する。

### 4.2 実験条件

本実験の実験条件を示す。2023年10月17日~12月17日までのCapitalBikeshare社の運行データを使用し予測モデルを構築する。使用する特徴量は、曜日、利用可能台数、時間帯、残電量パターン類似度を用いる。目的変数は、単位時間(30分)あたりの一つのス

テーションにおける貸出台数とする。評価指標として、5分割交差検証を用いて平均RMSEを算出した。

### 4.3 実験結果・考察

図2に予測モデルにおける特徴量重要度を示す。図2が示す通り、バッテリー残電量情報に関する特徴量である $Sim$ が、使用した特徴量のなかで2番目に重要であった。平均RMSEは類似度なしモデルで0.63、類似度ありモデルで0.77であった。これらの実験結果から、バッテリー残電量情報を特徴量に用いる有効性が示唆された。

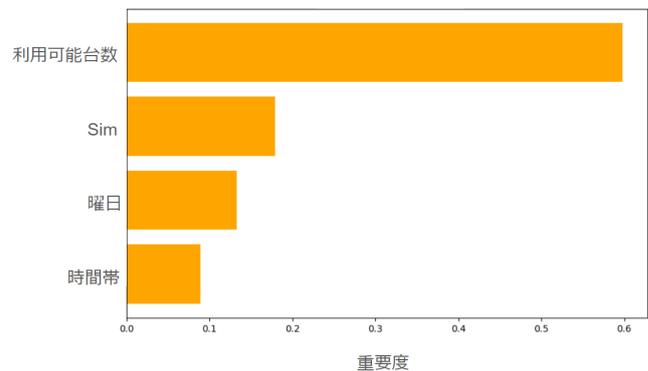


図2: 特徴量重要度

## 5 おわりに

本稿では、自転車のバッテリー残電量情報に関する特徴量を用いた需要予測手法を提案し、需要予測にバッテリー残電量情報を活用する有効性を検証した。実施した実験結果から、バッテリー残電量情報を特徴量に用いる有効性が示唆された。今後は、天候情報や返却台数などを考慮して予測手法を改良し、予測結果に基づいて自転車の再配置シミュレーションを行う。

## 参考文献

- [1] 国土交通省. シェアサイクルに関する現状と課題. <https://www.mlit.go.jp/road/ir/ir-council/sharecycle/pdf01/03.pdf>, (最終アクセス:2023/06/06)
- [2] 上田 嘉紀, 太田 豊, “電気自動車ユーザーはどのようなタイミングで航統距離不安を感じ充電したいと思うか?”, エネルギー・資源学会論文誌, 43(3), pp.84-93 (2022).
- [3] 三村 知洋, 石黒 慎, 鈴木 喬, 山田 暁, “深層生成モデルを用いたバイクシェア需要予測に関する研究”, 情報処理学会研究報告, 2023-MBL-107(8), pp.1-6 (2023).
- [4] M. He, X. Ma and Y. Jin, “Station Importance Evaluation in Dynamic Bike-Sharing Rebalancing Optimization Using an Entropy-Based TOPSIS Approach,” IEEE Access, Vol. 9, pp.38119-38131 (2021).
- [5] Yiming Xu, Xiang Yan, Virginia P. Sisiopiku, Louis A. Merlin, Fangzhou Xing, Xilei Zhao, “Micromobility Trip Origin and Destination Inference Using General Bike-share Feed Specification Data,” Transportation Research Record 2676, pp.223-238 (2020).