

オンデマンドバスにおける 運用コスト削減のための乗車リクエスト予測

木田 雄也[†] 堀 磨伊也^{††} 高野 茂[‡]
九州大学[†] 九州大学^{††} 九州大学[‡]

1. はじめに

過疎化や高齢化が進む地方では、生活に必要な施設への移動手段としてバスなどの交通サービスが必要不可欠である。しかし、従来の定時路線型のバスでは需給バランスの調整が難しい。そこで、アプリケーションなどにより乗車のリクエストを行うことで利用可能なオンデマンドバスの導入が期待される[1]。

オンデマンドバスにおいて解決すべき課題として、人件費や燃料費などのコストの問題や乗車時間などの快適性の問題がある。その要因として時間や日によって利用者数にばらつきがあり配車数が最適ではないこと、乗車リクエストが入るたびに運行ルートが変更されることがある。

本研究では、上記の問題を解決するために乗車リクエスト数を予測し、その結果を用いて運行ルートの最適化や配車数の決定を行う。運行ルートの最適化には短期先の予測、配車数の決定には長期先の予測を用いること[2]が有効だと考えられる。長期先の予測と短期先の予測では予測の手法や用いるデータなどが異なるため本報告では短期先の予測に焦点を当てる。

2. 周辺環境データを用いた時系列モデリングによる乗車リクエスト予測

本研究では、運行ルートの最適化を行うために短期先の乗車リクエスト数の予測を行う。本報告では、Long Short-Term Memory (LSTM) を予測モデルに用いる。ただし、過去の乗車リクエスト数データのみでは予測が困難であると考えられるため、周辺環境データを用いるとともに入力データのパターンに応じて複数のモデルを作ることで高精度の予測を行う。

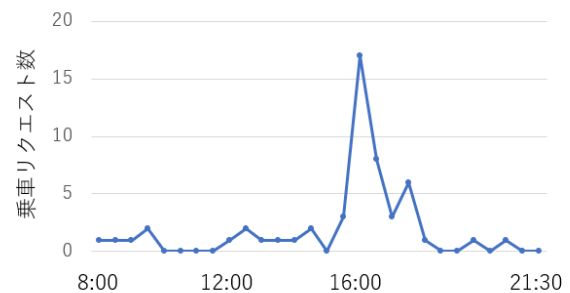


図1 ある一日の乗車リクエスト数の推移

2.1 予測に用いる入力データ

研究の対象としたオンデマンドバスの運行時間は 8:00 から 22:00 であり、乗車リクエスト数は 30 分間隔で集計される。実際に観測されたデータの例を図1に示す。時間によりリクエスト数のばらつきがあることがわかる。

周辺環境データとしては、気象データ（日照時間、気温）、予測対象地点の周辺の停留所の乗車リクエスト数、曜日データをそれぞれ用いる。事前の分析により、天気によって利用者数に偏りがあること、周辺の停留所間の乗車リクエスト数に相関があること、曜日により利用者数に偏りがあることが分かったため各データを利用する。曜日データはダミー変数に変換して利用する。

予測に用いるデータは対象地点の乗車リクエスト数に周辺環境データを加えベクトル化したものである。連続した $n+1$ 個のデータを1セットとし、その n 個目までを入力データにし最後の1つを教師データにする。ここでは日付が変わる際の不連続な部分はこの操作は行わない。

2.2 入力データのパターンごとのモデル構築

本研究では、現在の段階で得られるデータ数が少ないため、入力データの時系列パターンごとにクラス分類を行い同じパターンのデータのみでモデルを生成する。クラス分類には k 平均法を用いる。本研究では、入力データのうちの対象地点の乗車リクエスト数のみでクラス分類を行う。これにより、直近の時系列パターンに近い

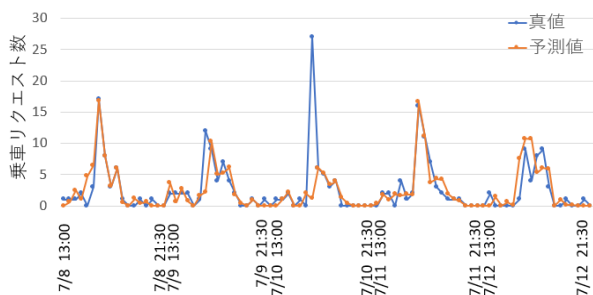


図 2 提案手法による乗車リクエスト数の予測値と真値の比較

もの同士で分類され予測の精度の向上が見込める。

本研究で扱う乗車リクエスト数は、時系列データであるため LSTM を用いてモデルを生成する。クラス分類により k 個に分類されたデータ群それぞれに対し LSTM でモデルを生成し予測値を求める。

3. 実験

実験対象は九州大学内を運行中のオンデマンドバスとし、その中の1つの停留所を予測対象とする。実験データは2019年6月1日から7月31日までの平日の乗車リクエスト数とした。クラス数は4 ($k = 4$) でそれに合わせモデルも4個生成した。本実験では、5時間前 ($n = 10$) までのデータを予測の入力とし次の30分間の予測値を出力するモデルを生成した。本実験では予測精度の検証の際に5分割交差検証を行った。

提案手法による予測値と真値の比較を図2に示す。乗車リクエスト数が少ない時間帯の予測精度は高いがリクエスト数が大きく上昇している部分の予測精度が低いことがわかる。これは何らかの特別な要因でリクエスト数が急に上昇したことで学習データのパターンにないデータとなってしまったためであると考えられる。

提案手法との精度の比較のため、対象地点の乗車推移データのみで LSTM モデルを作ったもの(モデル A)、対象地点の乗車推移データに周辺環境データを加えて LSTM でモデルを作ったもの(モデル B)、対象地点の乗車推移データを用いクラス分類を行い LSTM モデルを作ったもの(モデル C)を用意し、これに提案手法を加えた4つのモデルで、モデルが出力した予測値と真値の平均絶対誤差 (MAE) と平均平方二乗誤差 (RMSE) を求めた。その結果を表1に示す。MAE に関しては、提案手法が最も精度が高かった。RMSE に関

表 1 提案手法とその他のモデルの平均絶対誤差 (MAE) と平均平方二乗誤差 (RMSE) の比較

	MAE	RMSE
A) 対象地点の乗車数推移データのみ	1.78	2.65
B) 対象地点の乗車数推移データ + クラスターリング	1.69	2.47
C) 対象地点の乗車数推移データ + 異種データ	1.30	2.13
D) 対象地点の乗車数推移データ + 異種データ + クラスターリング (提案手法)	1.12	2.19

しては、モデルCよりも少し大きくなっているが急激に乗車リクエスト数が変化している部分の予測の精度が低いことが原因であると考えられる。

4. まとめ

本報告では、周辺環境データを用いるとともに入力データのパターンに応じて複数のモデルを生成することで高精度の乗車リクエスト予測を行う手法を提案した。実験では学習データの過去のパターンにない急激な乗車リクエスト数の変化に弱いことが分かった。また、他のモデルとの比較では、対象地点の乗車データ、周辺環境データ、クラスターリングのすべてを用いた提案手法が最も精度が高かった。今後の課題として、別の停留所や別の時期にこの手法を用いた予測が行えるかの検証が挙げられる。また本実験では短期先の予測を行ったが、配車数の決定のためには長期先の予測も必要になると考えられる。

謝辞

本研究は、国立研究開発法人科学技術振興機構のセンター・オブ・イノベーション (COI) プログラム、JPMJCE1318、株式会社イマーゴ及び株式会社 NTT ドコモの支援を受けたものである。

参考文献

- [1] 坪内孝太, 大和祐幸, 稗方和夫: “過疎地における時間指定のできるオンデマンドバスシステムの効果”, 日本ロボット学会誌, Vol. 27, No. 1, pp. 115-121, 2009.
- [2] 入本勇次, 上田弘樹, 伊藤秀将: “乗合オンデマンド交通の AI 需要予測システム”, 東芝レビュー, Vol. 74, No. 1, 2019.