

運行実績データに基づくバス到着時刻予測モデルの開発

佐藤 孝大† 大場 みち子†

公立ほこだて未来大学システム情報科学部†

1. 背景

近年、自家用車の利用数増加や、人口の減少により、地方圏の公共交通機関利用者が減少している[1]。利用者の減少により、交通事業者の収益が低下し、路線の廃止・減便が発生している。路線の廃止・減便により利便性が低下することから、さらに利用者の減少に繋がるという悪循環が生じている。この悪循環の中で、路線バスに生じている利便性低下には、対策の一つにバスロケーションシステムがある。バスロケーションシステムとは、バスに搭載したGPS(Global Positioning System)から現在の位置情報を取得し、バスの現在位置や到着予測時刻を利用者に提供するシステムである。どの路線のバスがいつどの停留所に到着するかを把握できるため、利便性向上の一助になる。しかし、問題点として、到着予定時刻の予測精度が不十分であることが挙げられる。予測結果が実際の到着時刻と異なる場合、待機時間の増加やバスに乗り損ねる事態が予想され、かえって利便性が低下するといえる。遅延要因には、道路状況や天候、乗車人数等様々な要因が考えられ、予測精度向上にはこれらを考慮する必要がある。

本研究では、バスの到着時刻予測精度の向上を目的に、バスの複合的な遅延要因を推定した上で、遅延要因を考慮したバス到着時刻予測手法を提案する。

2. 先行研究 / 関連研究

藤原らはバスの遅延要因の特定を目的に、運行実績データと道路データの分析をしている[2]。分析の結果、信号数等の道路特性の違いが、遅延要因に関係する可能性が示された。しかし、左折数や信号数等の道路特性が遅延要因としてどれほど遅延に影響し、どれほど遅延が発生するかまでは十分評価できていない。本研究はこの先行研究を踏まえて、他の遅延要因を検討する。

バスの到着時刻予測については、天気、月、曜日などの質的データを用いた所要時間予測[3]や、バスの乗降者数データを用いた予測[4]が提案されている。前者については、大きくても誤差が10分以内の予測を行えているが、始点から終点までの総所要時間を対象としているため、任意の停留所間で算出することができない。後者については、通過する停留所が多くなるほど実遅延時間との差が増加しており、適切な予測モデルが設定できていない。

関連研究では、天気や乗降者数といったデータに着目し、それらの遅延要因としての有効性が示されている。しかし、予測モデルの精度や、適応できる区間に課題があり、現在運用されているバスロケーションシステムには適用できない。また、遅延には道路特性、天気、乗降者数など、多面的な要因が関連していると考えられる。そのため、ある要因にのみ特化した予測モデルではなく、複数の要因に対応することで、精度向上が見込めると考えられる。

3. 提案手法

過去のバス運行実績データの分析から遅延要因を推定する。その結果から重回帰分析を用いることで、現在走行中のバス到着時刻予測モデルを開発する。

3. 1. 対象データと遅延要因の推定

バスの運行実績データを分析し、遅延の特徴と遅延要因と考えられるデータを比較することで遅延要因を推定する。運行実績データとは、バスロケーションシステムが自動収集している各停留所で生じた遅延時間と到着時刻のデータである。遅延要因の候補には、乗降者数データ、天気データ、道路データがある。天気データとは、気温や湿度、積雪量等の天気に関するデータ全般である。道路データとは、停留所間距離、信号数、車線数、右左折数のことである。

遅延要因の分析を行う上で、運転手の経験的な知識や思考も考慮するため、函館のバス運転手2名に対し、ヒアリングを行った。結果として、バスの遅延には乗降者数や停留所と交差点の位置関係に影響を受けたことがあるとの回答

Development of Prediction Model for Bus Arrival Time
Based on Bus Traffic Record Data

†Kodai Sato †Miciko Oba

†School of System Information Science, Future University
Hakodate

が得られた。よって、停留所の交差点までの距離も道路データとして扱い、乗降者数を中心とした分析を行う。

3. 2. 予測モデル作成

本研究では重回帰分析を用いて回帰式を作成することで到着時刻予測を行う。目的変数には停留所間遅延時間を設定し、推定した遅延要因を説明変数に設定する。モデルの評価には重相関係数を用いることで説明変数の当てはまりの良さを見る。当てはまりが悪い場合は説明変数を再設定する。本研究では現在走行中のバスに対して到着時刻予測を行うが、現時点ではバスの運行情報をリアルタイムで取得することができない。そのため、取得できないデータに関しては、過去の平均値などの停留所特徴を説明変数に適用することで対応する。

4. 予測モデルの作成と考察

函館バスのある路線の一部区間に対し、重回帰分析を行った。この区間は函館の利用者の多い停留所を複数通過しており、主要な区間の一つである。分析は2015年1月全便往路についての4530件のデータを対象に行った。目的変数と説明変数を表1に示す。

表1. 目的変数と説明変数の一覧

目的変数	停留所間の遅延時間(秒)
説明変数	月・曜日別過去遅延時間平均(秒)
	乗車数平均(人)
	降車数平均(人)
	通過数平均(人)
	降水量(mm)
	積雪(cm)

目的変数は停留所間の遅延時間とした。説明変数は予測対象日と同一の月・曜日の過去遅延時間平均、乗車数平均、降車数平均、通過数平均、降水量、積雪とした。重回帰分析の結果、重相関係数が0.541となり、ある程度の相関関係を示した。この結果から作成した予測モデルを用いて、2017年1月3日の遅延予測を行った。対象区間は重回帰分析と同様の区間とし、全便を対象とした。天候は曇りで降雪・積雪は無かった。予測の結果、予測結果と実績値の誤差平均は-89.59秒となり、あまり良好な結果を得られなかった。そのうちの最も差が大きくなった10時台のデータを図1に示す。全体的に実際の遅延時間とは大きく異なり、時刻表の時刻よりも非常に早く通過するという予測結果となった。原因として、予測対象日には積雪や降水が観測されておらず、天候の影響が少なく、実際よりも

到着時刻が早く予測されたと考えられる。そのため、天気データを説明変数に加えるだけではなく、積雪の有無や量によって、細かく予測モデルを検討する必要があることが示唆された。

予測結果の停留所番号12から15に着目すると、他と比べ精度が良いといえる。理由として、この区間は他の停留所と比べて利用率が極端に少ない区間であるため、人の乗降者の影響を受けず、過去の遅延時間平均との差が生じにくいためだと考えられる。しかし、重回帰分析の乗降者数におけるt値に着目すると、どちらも絶対値が2以下となっており、予測モデルには影響が少ない結果となった。そのため、説明変数に乗降者数は適していない場合があり、各停留所の乗降者数の傾向の違いごとに、説明変数や補正を検討する必要があることが示唆された。

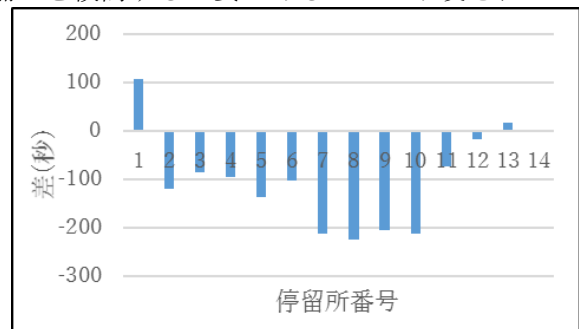


図1 10時台における予測値と実績値の差

5. おわりに

本研究では、バスの到着時刻予測精度の向上を目的に、バスの運行実績に基づいたバス到着時刻の予測手法を提案した。結果として、天候や停留所の利用頻度により、予測モデルをそれぞれ作成する必要があることが示唆された。

今後は、天候と停留所の利用頻度によって説明変数を再調整し、複数の予測モデルを設定することで精度向上を目指す。

参考文献

- [1] “地域公共交通の現状.”, 国土交通省, <http://www.tb.mlit.go.jp/kinki/kansai/program/02.pdf>, (2016. 7. 24 アクセス).
- [2] 藤原由美恵, 白石陽, 「道路データとバス運行データを用いた路線バスの遅延要因分析」, 第78回全国大会講演論文集, vol. 2016, no. 1, pp. 395-396, 3月2016.
- [3] 辰巳浩, 大野雄作, 「バスプローブデータを用いた路線バスの予想所要時間に関する基礎的研究」, 都市政策研究, no. 9, pp. 79-86, 2010.
- [4] 前川裕一, 中島秀之, 白石陽, 「乗降者数データと運行実績データを用いたバス到着時刻予測手法の提案」, 第76回全国大会講演論文集, vol. 2014, no. 1, pp. 157-158, 2014.