

道路状況を考慮した路線バスの遅延時間予測の精緻化

Elaboration of the Route Bus Delay-time Prediction Considering Road Condition

指導教授 轟 朝幸

M5012 横関 敬裕

1. はじめに

都市内の路線バスは、交通状況だけでなく天候や車内の混雑状況など、様々な要因によって遅延が発生している。昨今では、多くのバス事業者がGPSなどを用いてバスの位置や到着時刻を利用者へ配信するバスロケーションシステム（以下、バスロケ）を導入し、利便性の向上を図っている。しかし、現状のバスロケにおける到着予想時刻は、通過停留所の実績時刻に先行バスの実績所要時間を加算する簡易なシステムが多く、実績値とバスロケにより提供される推定値が乖離し、正確な到着時刻を利用者へ提供できていない可能性がある。

そこで本研究では、様々な遅延要因を考慮することで到着予想時刻情報の精緻化を目指し、路線バスの利便性向上に寄与することを目的とする。

2. 先行研究と本研究の位置づけ

バスロケの精度を向上するため、数多くの研究がなされている。例えば、轟ら¹⁾は日時により異なる遅延の途中停留所の停車時間等を考慮した上で、ニューラルネットワーク（以下、NN）を用いて遅延時間予測モデルを構築した。その結果、最も距離が離れている停留所からの遅延時間予測については、±3分範囲で約76%、±1分範囲では約30%の的中率を示している。

しかしながら、一部の停留所区間においては、±3分範囲内の的中率が15%程度の区間が存在する。その理由として、モデルに道路の混雑状況や交差点などの沿道の情報が考慮されていないことが考えられる。渋滞発生時等の大きな遅れは利用者が特に知りたい情報であるため、これらを的確に予測できないと十分なモデルであるとは言い難い。

そこで本研究では、これらの道路状況を代表する変数を考慮したモデルを構築することで、更なる遅延時間の予測精度の向上を目指す。具体的には、曜日別・時間帯別の道路混雑状況や区間内の交差点数、右左折数をモデルに考慮する。

3. 路線バスの遅延予測の方法

3.1 ニューラルネットワーク

遅延時間の予測には、機械学習型の予測手法の1つであるNNを用いてモデル構築を行う。同モデルについては、実データを用いて反復的に学習することで、相関分析や重回帰分析では把握できない様々な要因と遅延時間との非線形な関係性を表現することが可能となる。本研究では、フィードフォワード型のNN（多層パーセプトロン）とバックプロパゲーション学習を用いて学習を行い、遅延時間の予測を行う。具体的なモデルは、図-6に示すようにデータが入力層から出力層への1方向のみに流れ、一つの間層を持つ階層構造になっており、出力層の結果は教師信号と照合される。なお伝達関数については、シグモイド伝達関数を用いて学習を行っている。

現行のバスロケはリアルタイムの予測が主であるのに対して、本研究においては1ヶ月前、1週間前、1時間前といった、過去の実績データに基づく予測を想定しているため、機械学習の一つであるNNモデルを用いた分析を行う。

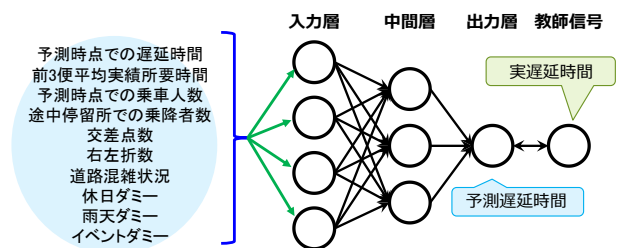


図-1 NNの模式図の例

3.2 モデルに用いる変数

遅延要因としては、先行研究をベースに表-1に示す変数を採用し、モデル分析を行う。なお、本研究において新たに追加した変数を太字および下線で示す。

まず、予測時点での遅延時間を基本的な遅れの情報として、「予測時点での遅れ時間と前3便の平均実績所要時間」のほか「予測時点での乗車人数」,「途中停留所での乗降者数」を考慮する。道路状況として、「交差点数」,「右左折数」,さらに「道路混雑状

況」を得点化した変数を用いる。ここで、道路混雑状況は Google Map にて公開されている曜日別・時間別の混雑状況を、便ごと・区間ごとに0, 1, 2, 3の4段階で得点化することで、モデルに反映させている。

また、「降雨の有無」については、気象庁の AMEDAS データのうち、さいたま市観測所のデータを用いることで、1 mm/日以上降雨があった日を1とするダミー変数として考慮する。また、「イベントの有無」として、浦和美園駅最寄りの埼玉スタジアムでJリーグおよびサッカー日本代表戦が開催された日を1とするダミー変数を導入している。

表ー1 モデルに用いる変数一覧

実績値	ダミー
予測時点の遅延時間	
前3便の平均実績所要時間	
予測時点での乗車人数	休日ダミー
途中停留所での乗降者数	降雨の有無
<u>交差点数</u>	<u>イベントの有無</u>
<u>右左折数</u>	
<u>道路混雑状況</u>	

4. 対象路線と使用データ

本研究では、国際興業株式会社（以下、国際興業バス）が運行する路線において、同社へのヒアリング調査より、比較的遅延が発生しやすい大宮駅東口と浦和美園駅西口間を結ぶ大01系統を対象路線とする（図ー2参照）。遅延時間予測モデルに用いたデータとして、以下に示す4つのデータを用いる。

まず、2014年4月1日から2016年3月31日までの2年間分の各停留所計画到着時刻および実績到着時刻が記録されているバスロケータを用いる。

さらに、道路状況を表すデータとしてGoogle Map にて公開されている曜日別・時間別の道路混雑データ、交差点数、右左折数、気象条件として気象庁のさいたま市降雨データ、またイベントによる多客を



図ー2 大01系統の路線図

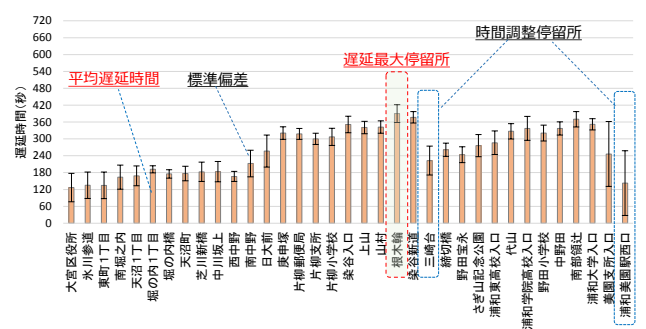
考慮し、浦和美園駅最寄りである埼玉スタジアムでのサッカーの試合日程を用いる。これらのデータの期間はバスロケータと同様である。

5. 基礎集計

5.1 遅延発生状況

図ー3に大宮駅東口から浦和美園駅西口へ向かう系統の停留所別の遅延状況を示す。なおバスロケータは、通信状態の問題により便ごとでデータの取得率が異なる。このため、本研究ではデータ取得率が80%以上のデータ、合わせて約12,000便を分析対象とする。なお、除外されたデータサンプルは全体の0.01%未満のため、これらのサンプルを除外しても予測において大きな影響は出ないものとする。

図ー3より、途中停留所の三崎台と終点である浦和美園駅西口においては、他の停留所と比較して遅延時間は縮小している。これは、遅延の発生を前提にダイヤが設定されているため、これらの停留所では計画所要時間をあえて長く設定することで、遅延が減少しているものと推察される。また、始発停留所の次停留所である大宮区役所停留所に着目すると、2分程度の遅延が発生している。この理由として、ターミナル駅である大宮駅東口付近の道路状況が日常的に混雑していることや信号機が多数設置されていることが要因として考えられる。



図ー3 下り便の停留所別平均遅延時間と標準偏差

また、最も遅延が発生している停留所は根木輪停留所である。その理由の一つとして、交通状況の差異が考えられる。そこで、根木輪停留所の時間帯別の平均遅れ時間を図ー4に示す。

結果に示すように、7時・8時台や16時から19時台といったラッシュ時間帯に遅れが大きくなっていることがわかる。このことから、時間帯ごとに交通状況を考慮することが重要であると考えられる。

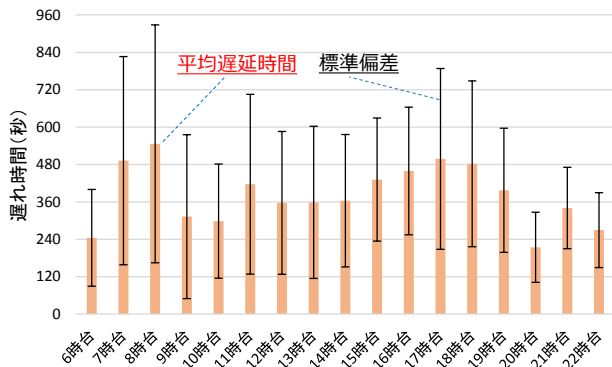


図-4 根木輪の時間帯別平均遅れ時間と標準偏差

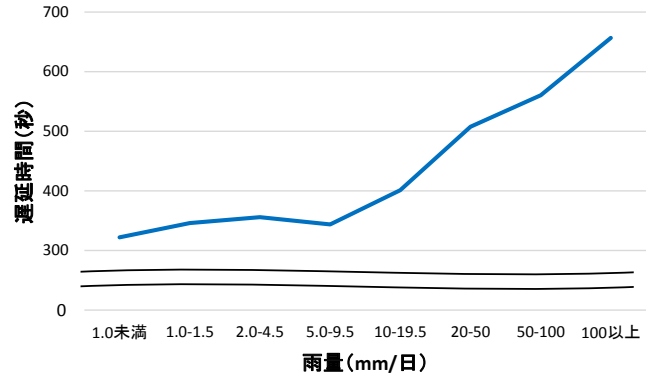


図-6 根木輪停留所の雨量別遅延時間

5. 2 交差点による遅れ

図-5に、平均遅延時間と交差点の関係について示す。ここで、縦の点線が交差点の通過を表しており、線上の数字が交差点数、また矢印が右左折の有無を示している。

また、赤枠で示す西中野停留所から庚申塚停留所間は、3つの交差点を通過する区間であり、この区間だけで150秒近く遅延が増大している。さらに、同区間については、旧道から国道へ合流する地点を左折する必要があり、渋滞が発生している場合などでは、遅延の大きな要因となることが考えられる。

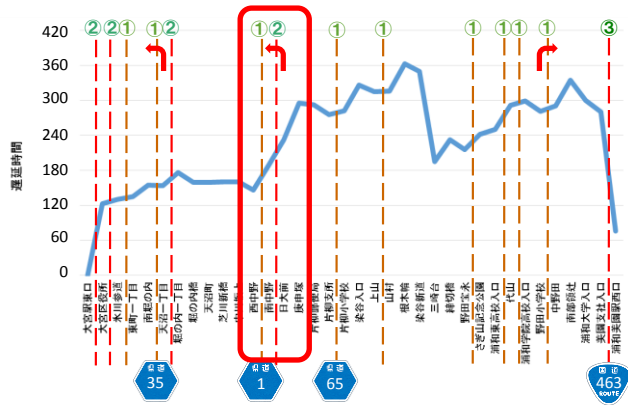


図-5 平均遅延時間と交差点の関係

5. 3 降雨による遅れ

図-6に、根木輪停留所の雨量別の遅延時間を示す。

雨量が大きくなるにつれて遅延時間も大きくなり、日降水量が10mm以上になると、平均でも400秒以上の遅れが発生することがわかる。これは、雨天により道路渋滞が発生することや、普段自転車や徒歩で移動している人々が路線バスを利用し、混雑が発生していることが考えられる。

6. 遅延予測モデル

6. 1 モデルの構築

3章および4章で述べた遅延要因を変数に用い、NNを用いてモデル分析を行う。ここでは、2014年4月から2015年3月までの約6,000便を学習用データとしてモデルを構築し、2015年4月から2016年3月までの約6,000便を推定用データとして、遅延時間の予測に用いる。

6. 2 遅延予測結果

本稿では、最も平均遅延時間の大きい根木輪停留所の遅延時間について、大宮駅東口停留所発車時の予測結果を示す。図-7に、本研究で構築したモデルによる結果を、図-8には参考として、先行研究で構築したモデルを用いた結果を示す。

図-7と図-8を比較すると、決定係数 R^2 が0.58から0.67へ上昇し、ばらつきが少なくなり、予測精度が向上した成果が見られた。また、これまで予測

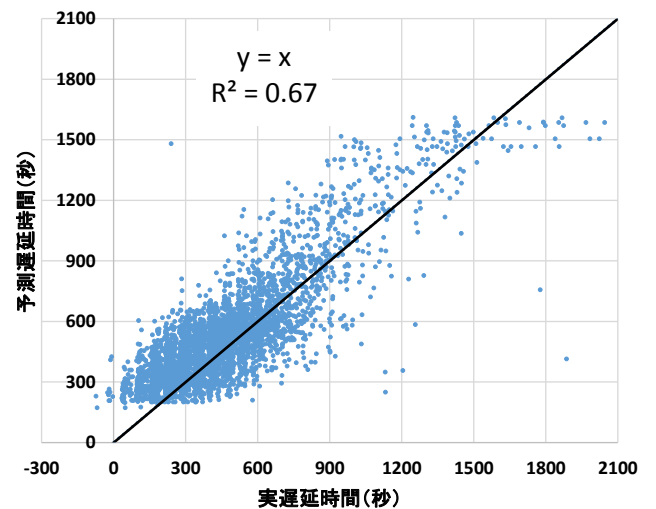
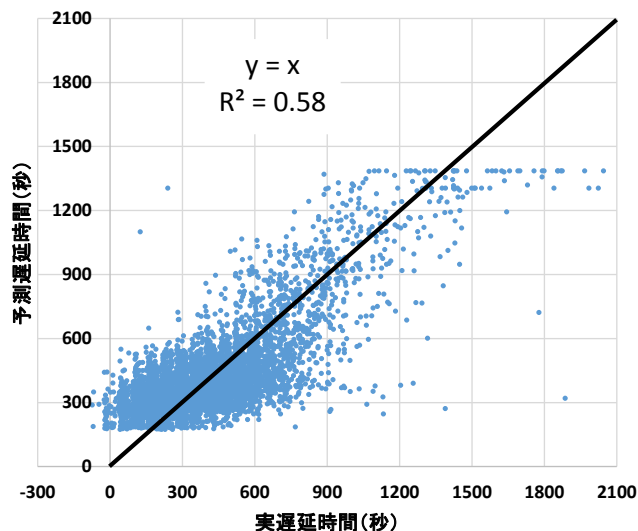


図-7 根木輪停留所の遅延時間予測結果



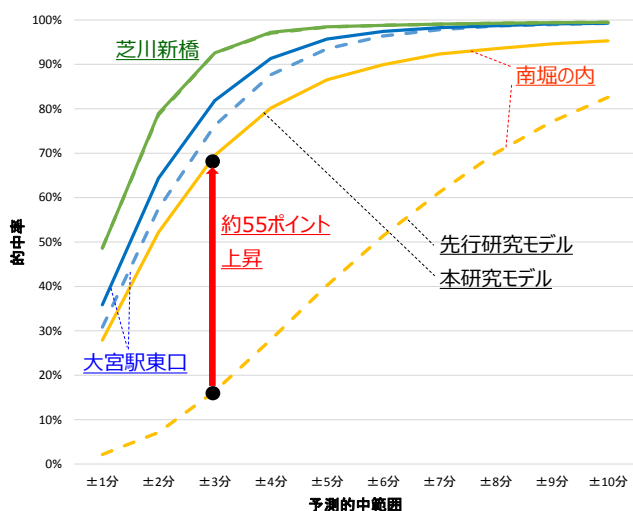
図－8 先行研究のモデルによる予測結果

の最大値が1,400秒程度であったものが、1,500秒以上まで予測することができ、発生頻度が低い「外れ値」と呼ばれる、大きな遅延についても予測ができた。時間帯別の道路混雑状況を反映させることで、より精緻に遅延時間が予測できたものと考えられる。

6. 3 的中率の比較

紙面の都合上、根木輪停留所の遅延時間予測結果のうち、大宮駅東口（約20分後を予測）・南堀の内（約15分後を予測）・芝川新橋（約10分後を予測）の3停留所からの予測の的中率の精度について、既往モデルと本モデルの比較結果を図－9に示す。なお、本モデルの予測結果を実線、先行研究のモデルの予測結果を破線で示している。

図－9より、既往モデルに比べて本モデルの的中率は高く、概ね良好な精度であることがわかる。例



図－9 根木輪停留所遅延予測の的中率の比較

えば、青色で示す大宮駅東口からの予測精度では、±3分範囲内の的中率は、76.2%から81.8%と、既往モデルと比較して約5ポイント上昇している。一方で、黄色で示す南堀の内停留所からの予測精度については、その他の区間に比べて的中率が低下しており、既往モデルについてみると、予測誤差±3分範囲内では14.5%程度の的中率しか得られないとの結果を得た。

しかしながら、本モデルの予測結果をみると、的中率は69.5%まで約55ポイント上昇している。これは南堀の内発車後に国道から旧道へと、交差点通過時に左折するため、その通過時間を過大に推定してしまっただけが要因として考えられる。その点についても、本研究のモデルにおいて大幅に改善できていることを示した。

7. おわりに

本研究では、道路状況を考慮してモデルを構築し、路線バスの遅延時間についてNNを用いて予測し、精度向上の成果が得られた。

具体的には、約20分後の予測において±3分範囲内の的中率がおよそ5ポイント上昇し、80%超という高い精度で予測可能であることを示した。また、25分を越える大きな遅延に対しても、より現実に近い値を示すことができた。さらに、既往モデルで精度が著しく低かった停留所区間についても、±3分範囲内で約55ポイントの予測精度の改善を示した。

今後の課題として、発生頻度が非常に稀な遅延時間帯や、実遅延時間と予測遅延時間が大きく離れた「外れ値」のデータを改善する必要があると考えられる。このため、外れ値が発生する際の状況を、バス事業者へのヒアリング等により遅れの考察を行うことで、モデルに加える変数をより深度化させ、更なる予測精度の向上が期待できる。

謝辞

国際興業株式会社様には、データ提供・ヒアリング調査にご協力頂きました。ここに謝意を表します。

参考文献

- 1) 轟朝幸, 川崎智也, 野村大智, 横関敬裕: ニューラルネットワークを用いた路線バスの遅延時間予測, 交通工学論文集 Vol.3(2017)No.2, pp.A_202-A_207, 2017.