

# 道路データとバス運行データを用いた バスの遅延要因分析

藤原 由美恵<sup>†</sup> 花田 智<sup>‡</sup> 白石 陽<sup>†</sup>

公立はこだて未来大学システム情報科学部<sup>†</sup> 公立はこだて未来大学大学院システム情報科学研究科<sup>‡</sup>

## 1. はじめに

路線バスは渋滞や積雪による路面状況の悪化、バス利用者の乗降に要する時間などの影響を受けやすく、計画された到着時刻通りに運行することが難しい。そのためバス利用者の待ち時間が長くなり、利便性の低下につながっている。そこで路線バスの利便性向上を目的として、バスロケーションシステムが多くの自治体で導入されている。バスロケーションシステムとは、GPS(Global Positioning System)を利用してバスの位置情報を収集し、バスの現在地やバス停への到着予定時刻をユーザに提示するシステムである。

しかし、問題としてバスロケーションシステムで提示されているバス停への到着予定時刻が不正確なことが挙げられる。その原因としてバスの遅延が十分に予測されていないことが考えられる。バスの遅延を予測するためには、バスの遅延に影響している要因を明らかにする必要がある。そこで本研究では、バスの遅延要因分析を行う。

## 2. 関連研究

バスの遅延予測に関する研究として、文献[1]では、天気や月、曜日などの質的データに着目してバスの到着時刻予測を行っている。文献[2]では、過去の運行実績データに着目してバスの到着時刻予測を行っている。文献[3]では、乗降客数データに着目してバスの到着時刻予測を行っている。

これらの既存研究ではバスの走行に影響すると考えられるバス停間の道路特性については考慮されていない。ここで本研究におけるバス停間の道路特性とは、バス停間の距離や信号数、右左折の回数などバス停間の道路に関する情報のことである。道路特性に関する研究として、文献[4]では道路の混雑度と自動車の走行速度の関係を分析している。この研究では、信号交差点密度が低い範囲では混雑度が大きくなるほど平均走行速度が低下する傾向を示している。また、信号交差点密度が高い範囲ではその傾向が弱くなり、混雑度と平均走行速度の相関が低いという傾向を示している。

したがって道路特性は路線バスの遅延時間にも影響を与えると考えられる。例えば、信号交差点密度が低い範囲では混雑度が大きいほど遅延時間が大きくなる可能性がある。バス停間の道路特性と既存研究で着目していたデータ(質的データ、運行実績データ、乗降客数データ)を組み合わせて分析することにより、遅延時間の予測に役立てることが可能であると考えられる。

### Analyzing Delay Factors of Local Buses by Using Road Attributes and Bus Operation Data

<sup>†</sup>Yumie Fujiwara <sup>‡</sup>Yoh Shirashi

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

<sup>‡</sup>Satoru Hanada

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

## 3. 提案手法

### 3.1 対象データ

本研究では、道路データとバス運行データ、天気データを用いた遅延要因の分析を研究目的とする。特に道路データとバス運行データを主な分析対象とする。本研究において道路データは、前述のバス停間の道路特性に関するデータのことであり、道路データとしては、バス停間の距離、信号数、右折数、左折数を対象とする。また、バス運行データは、バスの運行に関わるデータを指し、運行実績データと乗降客数データからなる。運行実績データとしては、各バス停間の遅延時間とバス路線の累積遅延時間を対象とする。各バス停間の遅延時間は、実際のバス停間の所要時間からバス停間の計画所要時間を減算した時間である。バス路線の累積遅延時間は、実際の到着時刻から計画到着時刻を減算した時間である。乗降客数データとしては、バス停での乗車数、降車数、バスに乗っている人数(以下、通過数)を対象とする。天気データとしては、降水量、降雪量、積雪量を対象とする。

### 3.2 データベース

本節では、遅延要因の分析におけるデータ管理の方法を説明する。運行実績データと乗降客数データはデータの管理方法が異なる。したがって、分析に用いるバス運行データを作成するために運行実績データと乗降客数データの対応付けが必要となる。そこで、Javaを用いてそれぞれのExcelデータを対応付けるプログラムを作成し、PostgreSQL上でデータベース化を行った。運行実績データは実際のバス到着時刻が取得されず、無効データとなる場合がある。その場合は前後のデータを用いて無効データの補間を行った。

### 3.3 分析手法

研究目的を達成するために、バスの遅延時間と複数の要因の関係を分析する。要因分析手法の一つに、重回帰分析がある。重回帰分析とは、ある変数(以下、目的変数)に影響を与えている要因(以下、説明変数)を抽出し、説明変数が目的変数に対する影響度を求める手法である。本研究ではバスの遅延時間に対して道路データとバス運行データがどの程度影響しているのか求めるために、重回帰分析を用いて遅延要因分析を行う。

## 4. 道路データとバス運行データの分析

### 4.1 バス停間の特徴

函館バスのある路線(以下、バス路線A)のバス停間の遅延時間について分析を行った。 $X_i$ はバス路線Aがi番目に通過するバス停を表す。 $L_{i,j}$ はバス停 $X_i$ を始点、バス停 $X_j$ を終点とするバス停間を表す。本稿では、4つのバス停間 $L_{2,3}$ と $L_{3,4}$ と $L_{16,17}$ と $L_{17,18}$ に着目し、その分析結果を4.2節で述べる。バス停 $X_2$ は鉄道の主要駅付近にあり、 $X_{17}$ は函館市内の生活中心地にある。なお、 $X_0$ はバス路線Aの起点バス停である。バス運行データに関する各バス停間の特

徴を表1に示す。L<sub>2,3</sub>とL<sub>3,4</sub>は乗降客数が少ない。一方で、X<sub>17</sub>は生活中心地にあるため乗降客数が多い。また、L<sub>2,3</sub>は平均遅延時間が大きい、L<sub>3,4</sub>は平均遅延時間が負の値をとっており、バス停間の計画所要時間より短く運行していることが多い。遅延時間の標準偏差はL<sub>17,18</sub>は他のバス停間より小さい。また、道路データに関する各バス停間の特徴を表2に示す。信号密度は信号数をバス停間の距離で除算したものである。L<sub>3,4</sub>は最も信号密度が高いバス停間である。L<sub>17,18</sub>はバス停間に信号が存在しないため信号密度が0のバス停間である。

表1 バス停間の特徴 (バス運行データ)

	乗車数 (人)	降車数 (人)	通過数 (人)	遅延時間 (秒)	
	平均	平均	平均	平均	標準 偏差
L <sub>2,3</sub>	0.6	0.05	5.0	134.87	61.44
L <sub>3,4</sub>	0.4	0	5.6	-40.27	60.46
L <sub>16,17</sub>	5.1	4.9	11.7	68.65	63.05
L <sub>17,18</sub>	2.4	1.3	11.8	7.12	37.85

表2 バス停間の特徴 (道路データ)

	バス停間 の距離 (m)	信号数 (個)	信号密度 (個/km)
L <sub>2,3</sub>	672	4	6.0
L <sub>3,4</sub>	231	2	8.7
L <sub>16,17</sub>	585	2	3.4
L <sub>17,18</sub>	423	0	0

0.441 と低い値を示した。ここで、各バス停間の特徴と重相関係数の相関を求めた。その結果、信号密度と重相関係数の相関が0.942であり、非常に強い相関を示した。ゆえに、信号密度が高いほど本稿で用いた説明変数(乗車数、降車数、通過数、降水量、平均遅延時間、累積遅延時間)が有効であることが示唆された。重相関係数が高い原因として、信号密度が多いほど走行速度が一定になり、交通量による走行時間の変動が小さくなることが考えられる。そのため、乗降客数(乗車数、降車数、通過数)と天気(降水量)についての変動のみとなり、説明変数に用いたデータで良い傾向がみられたと考えられる。したがって、信号密度に応じて説明変数を検討する必要がある。

すべてのバス停間で平均遅延時間に着目すると、t値が2以上であることから、影響度が大きい説明変数と考えられる。つまり、その路線の過去の遅延時間はバスの遅延を予測する際に有効である可能性がある。また、それぞれのバス停間に着目するとL<sub>3,4</sub>とL<sub>16,17</sub>では累積遅延時間が遅延時間への影響度が大きく、L<sub>17,18</sub>では乗車数の影響度が大きいという結果であった。よって、各バス停間で遅延時間に影響する要因は異なるということが示唆された。

今後は他のバス路線でも分析を行い、バス停間によって説明変数に用いるデータを検討していく。また、道路データと遅延時間の関係を分析していく。

表4 各バス停間の重相関係数

バス停間	重相関係数
L <sub>2,3</sub>	0.592
L <sub>3,4</sub>	0.834
L <sub>16,17</sub>	0.541
L <sub>17,18</sub>	0.441

## 4.2 分析結果

4.1節で述べたバス停間の遅延時間について重回帰分析を行った。分析は2014年2月、6月、10月の平日のデータを対象に行った。重回帰分析の目的変数と説明変数を表3に示す。

表3 目的変数と説明変数の一覧

目的変数	バス停間の遅延時間 (秒)
説明変数	乗車数 (人)
	降車数 (人)
	通過数 (人)
	降水量 (mm)
	降雪量 (cm)
	積雪量 (cm)
	平均遅延時間 (秒)
累積遅延時間 (秒)	

目的変数はバス停間の遅延時間とした。説明変数は乗車数、降車数、通過数、降水量、降雪量、積雪量、平均遅延時間、累積遅延時間とした。平均遅延時間は便ごとに平均値を計算した。そのうち2014年10月のデータを用いた分析結果について述べる。なお10月は降雪が無かったため、降雪量、積雪量を説明変数から除いた。また、X<sub>4</sub>の降車数が0であったため、分析するにはL<sub>3,4</sub>のみ降車数を説明変数から除いた。バス路線Aの各バス停間について重回帰分析を行った結果を表4に示す。L<sub>3,4</sub>の重相関係数が0.834と高い値を示した。しかし、L<sub>17,18</sub>の重相関係数は

## 5. おわりに

本研究では、道路データとバス運行データを用いてバスの遅延要因分析を行った。分析結果より、道路データによって変動要因を変えて分析を行う必要があることが示唆された。また、過去の遅延時間は現在の遅延時間に影響している傾向があった。今後は分析する路線を増やして説明変数に用いるデータを検討する。また、道路データと遅延時間の関係を分析していく。

## 謝辞

本研究を進めるにあたり、遅延要因の分析に用いた運行実績データや乗降客数データは函館バス株式会社の協力によるものである。ここに深く感謝の意を表す。

## 参考文献

- [1]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).
- [2]内村圭一, 成松裕介, 衛藤旭秀, 胡振程, バスロケーション情報を用いたバス停間所要時間予測, 国際交通安全学会誌, Vol.32, No.3, pp.64-71 (2007).
- [3]前川裕一, 中島秀之, 白石陽, 乗降客数データと運行実績データを用いたバス到着時刻予測, 情報処理学会第76回全国大会講演論文集, Vol.2014, No.1, pp.157-158 (2014).
- [4]小塚清, 上坂克巳, 青木亮二, 道路交通センサデータから見た混雑度と旅行速度の関係に関する一考察, 土木計画学研究発表会, Vol.46, pp.1-6 (2012).