

バス運行実績データの分析に基づく 到着時刻予測モデルの提案と精度検証

今井 瞳^{1,†1} 廣井 慧^{1,2,a)} 河川 信夫^{1,2,3}

受付日 2018年4月12日, 採録日 2018年10月2日

概要: 情報通信技術の発展により, 公共バスの位置情報やバス停出発時刻など, バスに関する様々な情報が容易に取得可能となった. しかし, バスは交通や天候などの様々な要因に影響を受け, 運行状況が変化しやすく, 正確な到着時刻予測が難しい. 本論文では, 重回帰分析とカルマンフィルタを併用し, バスの進行に合わせてリアルタイムに運行状況を考慮し予測精度を更新する到着時刻予測モデルを提案する. 提案モデルは, 過去の運行実績データをもとに算出した重回帰分析を用いた予測結果を, カルマンフィルタの初期値に入力する「MRKF モデル」と, MRKF モデルを改良しバスの進行に合わせて重回帰分析を用いて予測結果を更新する「EMRF モデル」である. 予測結果をもとに, 到着予測時刻とその誤差を提示し, バス停での待ち時間における利用者の不安や不満を解消し, 利用者の行動選択肢を広げ, よりバスを利用しやすい社会の実現を目的とする. 提案モデルについて評価を行い, 平均予測誤差が従来手法の約 186 秒から約 17 秒となった. さらに, 到着予測時刻と誤差の提示方法について Web 調査を行い, 169 名の回答結果により, 利用者のバス遅延や予測誤差に対する意識を調べ, 本提案モデルの有効性を確認した.

キーワード: バスロケーションシステム, 重回帰分析, カルマンフィルタ, 遅延予測

Arrival Prediction Model and Precision Analysis based on Bus Traffic Data

HITOMI IMAI^{1,†1} KEI HIROI^{1,2,a)} NOBUO KAWAGUCHI^{1,2,3}

Received: April 12, 2018, Accepted: October 2, 2018

Abstract: With the development of information communication technology, it is possible to easily acquire various kinds of information on a bus, such as bus location and departure/arrival time. However, bus traffic is influenced by various factors such as traffic and weather, thus the operation conditions are likely to change, and accurate arrival time prediction is difficult. This paper proposes arrival time prediction models that update the prediction accuracy in real time using the multiple regression model and the Kalman filter simultaneously. The proposed models are an MRKF model that inputs prediction results based on multiple regression models calculated based on past operation record data into the initial value of the Kalman filter, and an EMRF model that updates predictions by multiple regression models according to the bus location data. We evaluated the proposed models based on the analysis of actual bus operation data using prediction error. Our models resulted that the average prediction error from about 186 seconds to about 17 seconds. Furthermore, we conducted a Web survey for 169 answers on the method of presenting the predicted arrival time with errors.

Keywords: bus arrival information system, multiple regression model, kalman filter

¹ 名古屋大学大学院工学研究科
Graduate School of Engineering, Nagoya University, Nagoya,
Aichi 464-8603, Japan

² NPO 法人位置情報サービス研究機構 (Lisra)
Location Information Service Research Agency (Lisra),
Nagoya, Aichi 460-0002, Japan

³ 名古屋大学未来社会創造機構
Institutes of Innovation for Future Society, Nagoya University,
Nagoya, Aichi 464-8603, Japan

^{†1} 現在, ヤフー株式会社

^{a)} k.hiroi@ucl.nuee.nagoya-u.ac.jp

1. はじめに

バスロケーションシステムは、無線通信や GPS などを用いてバスの位置情報を取得し、バス停の表示板やインターネットを用いた情報の提示により、運行管理を効率的に行い、利用者サービスを向上するシステムである。近年のバスロケーションシステムの発展により、バス停での到着・出発情報に加え、走行位置（緯度・経度）など、様々なデータが動的に取得可能となった [1]。またバスロケーションデータに加え、IC カードの普及や様々なデータを収集できる車載器の発展により、乗降者人数 [2]、[3] や運転手の行動などのデータ化とその分析ができるようになってきた。

バスは、多くの人に日常的に利用される主要交通機関の 1 つであり、今後も重要な交通手段であり続ける。そのため、バスロケーションシステムを利用し、よりバスを利用しやすくするツールが次々と導入されている。利用者にとってバスの不便な点として、渋滞 [4] や天候など [5] の影響による遅延があげられる。この遅延に対しバスの利便性を向上させるために、バスロケーションデータを利用した、到着時刻を提示するサービスがすでに国内外の様々な交通機関で導入され、利用されている。

京都市営バスのバス到着案内システム [6] に代表される、各バス停の案内板で接近案内表示システムでは、バスの位置情報として「1 つ前/2 つ前のバス停に近づいています」、「まもなくきます (Approaching)」などといった情報提示を行っている。バス停の表示板だけではなく、スマートフォンアプリとして、バスの接近情報を提供しているサービスも多数の地域で見られる。しかし、これらのサービスはバスが付近のバス停を通過した情報をもとに作られている。つまり、これらは利用者の乗りたいバスが、バス停の 1 つ前もしくは 2 つ前のバス停を通過した時点で到着時刻が分かるサービスである。

我々は、このような利用者が乗る直前の到着予測を情報提示する従来型のサービスではなく、利用者がバスに乗る 1 時間前、30 分前など事前に自分が乗車するバス停にバスが到着する時刻、および目的地のバス停に到着する時刻を精度良く情報提示するサービスの構築を目指している。バス運行では、当日の天候、交通状況などによってバス停区間ごとに遅延が変動し、到着予測時刻に影響を与える。バス停区間ごとにどの程度遅延が蓄積し、また解消するのか、正確な把握は難しく、したがって到着時刻予測を正確に行うことも困難である。

既存研究では、バス遅延の原因となる要素の分析が進められており、たとえば、大野らによる研究 [7]、[8] では、実績データとダイヤの比較から運行実態を把握し、所要時間予測に関する基礎的研究として、運行状況に影響する要素ごと（天気、台風の有無、気温、月、曜日など）の分析が行われている。また、前川らによる研究 [2] では、到着時

刻予測を目的としたデータ分析が行われており、直前に通過したバス停での遅延時間が所要時間に与える影響や、乗降者人数と遅延時間の関係について分析している。

さらに、到着時刻や遅延を高精度で予測するために、従来研究では、重回帰分析やカルマンフィルタを用いた到着時刻予測が行われている。Patnaik らによるバスの到着時刻予測を行った研究 [9] では、バス停間の所要時間や距離、乗降者数、ドアの開閉時間を用いて重回帰分析による予測を行っており、高い精度での予測を可能にしている。Chen らの研究 [10] では、バスがあるバス停区間を通過した際に、バス停区間での所要時間と出発地点からの累積情報をもとに、カルマンフィルタを用いて、次のバス停区間での所要時間をリアルタイムに予測している。しかし、重回帰分析は、過去の運行データをもとにした静的な予測であり、雨天時やバス停近隣でのイベント開催などによる乗客増加や交通状況の変化などの要因による遅延を考慮していない。また、カルマンフィルタは誤差を加味した、動的な予測が可能であるが、他のルートと比べ著しく遅延が生じるバス停など特徴のあるバス停区間の影響を大きく受けてしまい、すべての運行ルートで正確な予測を行うことが困難である。

そこで本論文では、バス事業者より提供された運行実績データを用いて統計的な分析を行い、動的に入力する到着時刻予測モデルを提案する。具体的には、重回帰分析とカルマンフィルタを併用した MRKF モデルと、その改良モデルである EMRF モデルの 2 つの到着時刻予測モデルを提案し、予測精度について実際のバスロケーションデータを用いて比較した。MRKF モデルは、過去の運行実績データをもとに算出した重回帰分析による予測結果を、カルマンフィルタの初期値に入力するモデルである [11]。EMRF モデルは、MRKF モデルを改良した到着時刻予測モデルであり、重回帰分析をカルマンフィルタ内に組み込むことにより、系統ごとの傾向を再予測するモデルである。MRKF モデルでは重回帰分析による予測結果を初期値入力のみを使用したため、事前予測傾向と大きく異なる運行をした場合に対応できなかったが、EMRF モデルではバスの進行に合わせて重回帰分析による予測を更新し、さらなる精度向上を目的としている。EMRF モデルの特徴として、バスが目的地に近づくにつれて、その予測精度が向上することがあげられる。これらのモデルについて従来手法との比較を行い、精度向上を確認したのち、予測結果および誤差について提示方法の Web 調査を行った。調査結果から、利用者が遅延時間や予測誤差についてどのような意識を持っているか確認し、予測結果の精度基準を検討している。

従来のバス到着時刻予測は、バス停区間での所要時間やバスの接近を予測に着目した研究が多く、得られた結果をどのように評価し、利用者に提示するべきかといったことに着目した研究は少なかった。本論文では、バスが「あと

何分でバス停に到着するか」という具体的な到着時刻の高精度な予測とその推定精度を提示できるバス到着時刻予測を目指す。到着予測時刻と同時に、運行データの統計解析結果を用いた誤差範囲の利用者への提示を想定しており、到着予測時刻がどの程度の推定精度であるかを同時に提示し、待ち時間における利用者が選択する行動の幅を広げ、バスの利便性向上を目指している。利用者のバス停での待ち時間における不安 [12], [13] を解消し、より快適にバスを利用できる社会の実現が本論文の目的である。

2. 先行研究

2.1 バスロケーションシステムを用いた情報提供事例

利用者に対して、主に時刻表に基づいた到着時間を提示しているサービスの事例を述べる。京都市営バスのバス到着案内システム [6] では、各バス停の案内板で接近案内表示システムを用いて、バスの位置情報として、「1つ前/2つ前のバス停に近づいています」、「まもなくきます (Approaching)」といった情報提示を行っている。名古屋市交通局のなごや乗換ナビでは、スマートフォンアプリを通じて、利用者にバスの接近情報を提供している。ドイツ鉄道では、スマートフォンアプリ「DB Navigator」やホームページ [14] を通じてバスの遅延時間を提供する。これらのサービスは、図 1 の「従来の表示方法」に示すように、バス停を出発したこと、該当のバス停への接近について情報提示するが、バス停とバス停の間のどの位置にあるかを知ることはできない。また、現在の遅延時間は分かるが、自分が乗車するバス停や目的地においてどれくらい遅延するのか、回復する見込みはあるのかなどといった予測時間を知るのは難しい。つまり、提示される接近情報や遅延時刻は、1つ前もしくは2つ前のバス停通過に基づいて予測されており、当日の天候、交通状況などによって区間ごとに変動する到着時刻の事前の提示は難しい。

2.2 運行状況分析

接近情報だけでは、正確な予測が難しい到着時刻や遅延について、バスロケーションデータを用いてバスの運行状況を分析し、その関係性を見出す研究も多数行われている (表 1)。たとえば、大野らによる研究 [7], [8] では、実績データとダイヤの比較から運行実態を把握し、所要時間予

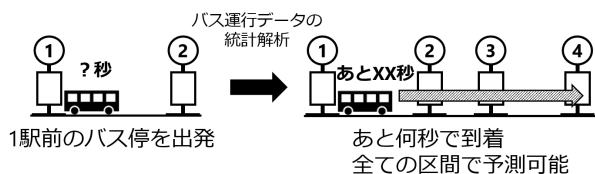


図 1 従来手法との比較

Fig. 1 Comparison with conventional methods.

測に関する基礎的研究として運行状況に影響する要素ごとの分析が行われている。天気、台風の有無、気温、月、曜日などの要素ごとに平均所要時間と標準偏差を算出し、それぞれの傾向について分析している。また、前川らによる研究 [2] では、到着時刻予測を目的としたデータ分析が行われている。直前に通過したバス停での遅延時間が所要時間に与える影響や、乗降者人数と遅延時間の関係について分析している。

2.3 重回帰分析を用いたバス到着時刻予測

上記のようなバスの接近情報、運行状況分析だけではなく、高精度にバスの到着時刻を予測する研究として重回帰分析を用いた予測手法がある。重回帰分析は、式 (1) に示すように、予測の対象となる目的変数 Y を複数の説明変数 $X_i (i = 0, 1, \dots)$ を用いて統計的に予測する手法である。定数 a_i は偏回帰係数と呼ばれ、説明変数ごとに算出される。

$$Y = a_0 + a_1 X_1 + \dots + a_i X_i \quad (1)$$

Patnaik らによるバスの到着時刻予測を行った研究 [9] では、バス停間の所要時間や距離、乗降者数、ドアの開閉時間などを用いて予測を行っており、高い精度での予測を可能にしている。しかし、重回帰分析は、遅延に影響する様々な要素を考慮した予測が可能であるが、過去の運行データをもとにした静的な予測であり、雨天時やバス停近隣でのイベント開催などによる乗客増加や交通状況の変化を考慮していない。したがって、このようなリアルタイムに変化する環境に対応し利用者に対してバスの到着時刻を提示するのは難しい。

表 1 運行状況分析に関する研究

Table 1 Previous study on bus operation analysis.

研究	地域	要素	手法
文献 [2], [3]	北海道函館	通過バス数, 乗降者人数	通過バス数, および平均乗降者数と遅延時間の比較
文献 [15]	東京都市圏	曜日, 時間帯, 天候, 五・十日	各要素と遅延時間について数量化 I 類を用いて分析
文献 [7], [8]	福岡都市圏	ダイヤ, 天気, 台風, 気温, 月, 曜日, 時間帯	運行実績データとダイヤとの乖離, および各要素と平均所要時間の関係性を分析
文献 [16]	北海道函館	乗降者数, 通過数, 降水量, 降雪量, 積雪量, 平均遅延時間, 累積遅延時間	重回帰分析を用いて各要素とバス停間の遅延時間の関係を分析

2.4 カルマンフィルタを用いたバス到着時刻予測

カルマンフィルタは、直前までの情報と現在取得したデータをもとに、最も最適なシステムの状態を推定する手法である。Chen らの研究 [10] では、バスがあるバス停区間を通過した際に、バス停区間での所要時間と出発地点からの累積情報をもとに、次のバス停区間での所要時間をリアルタイムに予測している。カルマンフィルタは誤差を加味した、動的な予測が可能であるが、ある特徴のあるバス停区間、その影響を大きく受けてしまい（たとえば他のルートと比べ著しく遅延が生じるバス停が存在した場合など）、正確な予測を行うことが困難である。

3. 提案モデル

3.1 提案モデルの概要：重回帰分析とカルマンフィルタの併用

本論文では、重回帰分析とカルマンフィルタを併用した予測手法により、リアルタイムに運行状況を考慮し、バスの進行に合わせて予測精度を更新するモデルを提案する。重回帰分析を用いたバス到着時刻予測では、過去の運行実績データから各路線・バス停区間ごとの特色（たとえば、遅れが大きく生じるバス停区間、雨の日利用者が増加する路線など）を算出し、統計的に予測できるが、交通状況や天候などの影響により運行状況が変化し、遅延が生じるバスの走行に合わせた動的な予測には適さない。一方カルマンフィルタを用いた予測では、リアルタイムに取得した運行実績データを用いて動的な予測ができるが、直前までの情報をもとに予測するため、入力急激な変化に対して高精度な予測を行うのは難しい。

そこで本論文では、図 2 に示すように、重回帰分析とカルマンフィルタを合わせて用いて互いの欠点を補完し合う予測方法を提案する。具体的には、まず重回帰分析により過去のデータから時間帯、曜日などを説明変数としてバス停区間の所要時間を統計的に予測し、その結果をカルマンフィルタにより動的に更新して、バスの進行に合わせて予測精度を向上させる。本モデルにより、重回帰分析の欠点である動的な予測をカルマンフィルタによって補い、カルマンフィルタの欠点である入力急激な変化を重回帰分

析によって事前に予測し、より高精度な予測がリアルタイムに可能となり、環境の影響を受け変化しやすいバスの運行状況に対応できると考える。なお、重回帰分析で利用できる説明変数は、各事業者で記録しているデータの種類によって異なる。本論文で説明変数の算出に利用したデータと算出手順について 3.2 節に詳述する。

3.2 重回帰分析で用いる説明変数

本論文における重回帰分析では、目的変数 Y をバス停区間の所要時間（バス停ごとの出発時刻の差）とし、説明変数は記録するデータの種類の異なるため、事業者ごとに調整した。表 3 から、名古屋市営バス、名鉄バスともに、バスごとの ID、系統情報、バス停の発着、バス停情報が記録されている。これらのデータから、名古屋市営バス、名鉄バスに共通する基本的な説明変数 X としてバス停区間、曜日、時間帯を算出した。さらに、これらの要素に加え、名古屋市営バスでは、バス停での到着時刻と発車時刻の差分から停車時間を算出し、あるバスの 30 秒ごとの現在位置から停止回数を算出した。名古屋市営バスでは、所定発時刻についてデータとして記録はされていないため、同年月日に該当するバス停の時刻表を用いて、各運行バスの所定発時刻、遅れとして算出した。名鉄バスでは、所定発時刻、利用者数を用いて、所定所要時間、移動人数、 n 駅前における遅れ ($n \geq 1$) を算出した。

バス停区間は、各系統のバス停に対し、各バス停ごとに順番に番号を振り、出発駅から到着駅までを区分した。始発駅を 1 とし、始発駅から 2 番目の駅の区間を 1~2、2 番目の駅から 3 番目の駅の区間を 2~3 として区分した。到着駅までバス停が N 個あった場合、区分の個数は $N - 1$ 個となる。

時間帯は、Patnaik らによる研究 [9] をもとに、5時から8時までを「Early Morning」、朝の通勤時間帯である8時から10時までを「Late Morning」、10時から13時までを「Early Noon」、13時から17時までを「Late Noon」、帰宅時間帯である17時から19時までを「Evening」、19時以降を「Night」と定義した。曜日については、対象のバスの運行した曜日を「Monday」、「Tuesday」、「Wednesday」、「Thursday」、「Friday」、「Saturday」、「Sunday」の7個に区分した。

名古屋市営バス固有の説明変数について、遅れは実所要時間と所定所要時間の差分とし、所定所要時間よりも早い場合を負、遅い場合を正と定義した。停止回数は、名古屋市営バスで記録されている30秒おきのバスの現在位置のデータを利用し、現在位置と30秒前の位置の2点間のユークリッド距離を緯度・経度から算出し、移動距離とした。さらにこの現在位置と30秒前の位置が記録された時刻の差分から経過時間を算出し、移動距離と経過時間から速度を算出した。この速度が時速5km未満になった場合を遅

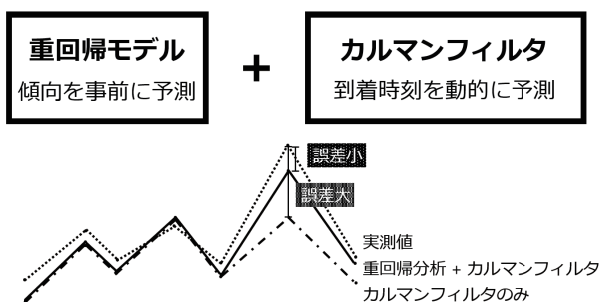


図 2 提案モデル

Fig. 2 Proposed model.

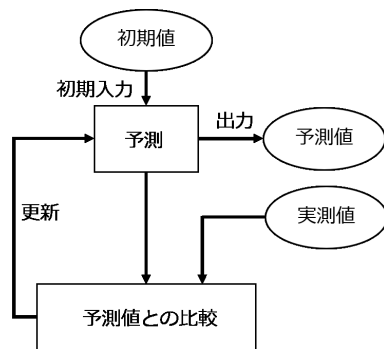


図 3 MRKF モデルの概念図

Fig. 3 Conceptual diagram of MRKF model.

れに影響するものとみなし、その回数の合計を停止回数とした。停車時間はバス停に到着した時刻と出発した時刻の差分から算出した。停止回数と停車時間はバス区間ごとに同時間帯の運行バスの平均値を算出し利用した。

名鉄バス固有の説明変数について、移動人数は、名鉄バスで記録されている各バス停での乗車人数、降車人数をもとに、乗車人数と降車人数を比較し、多い方の人数とした。 n 駅前における遅れは、並び順をもとに定義した。たとえば、バス停区間 3~4 における所要時間を予測する場合、バス停 2 における遅れが 1 駅前の遅れ、バス停 1 における遅れが 2 駅前の遅れとなる。

3.3 Multiple Regression Kalman Filter モデル

本論文では、はじめにバスの到着時刻予測を行うひとつめのモデルとして、Multiple Regression Kalman Filter モデル (以下、MRKF モデル) を提案する。MRKF モデルとは、過去の運行実績データをもとに算出した重回帰分析による予測結果を、カルマンフィルタの初期値に入力するモデルである。Chen らによる研究 [10] では初期値として時刻表の所要時間 (以下、所定所要時間) を入力するため、予定通りに運行していない場合に精度が低下する。そこで MRKF モデルでは、バスの運行状況が曜日や時間帯などに応じた特徴を持つことを利用し、重回帰分析を用いて各系統の傾向に合わせて算出した事前予測所要時間を初期値として用いることにより、時刻表通りに運行していない場合にも対応したより高精度な予測が可能であると考えた。

次に重回帰分析の結果をもとに、図 3 に示すカルマンフィルタを用いた予測について説明する。本論文では、終点をバス停 N とし、任意のバス停 k において、バス停区間 $k \sim (k+1)$, $k \sim (k+2)$, \dots , $k \sim N$ の所要時間を予測した。システムの初期状態として重回帰分析による予測値を入力し、あらかじめ著しく遅延が生じる区間などに対応できるようにした。バスが出発した後は、バス停に到着するごとにシステムの状態、および予測値の更新を行い、終点到着するまでこれを繰り返す。更新を繰り返すことで、重回帰分析による予測値が実測値と異なっていた場合

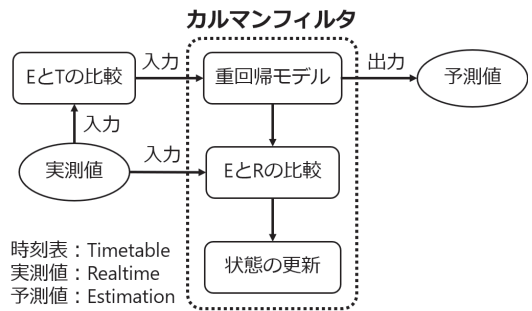


図 4 EMRF モデルの概念図

Fig. 4 Conceptual diagram of EMRF model.

にも修正を行うことができ、またバスが終点に近づくにつれて精度を向上させられると考える。

カルマンフィルタでは一般的に、任意のバス停区間 $(k+1) \sim j$ におけるシステムの状態は、1 つ前の状態をもとに式 (2) で表される状態方程式を用いて求められる。

$$x_{k+1,j} = \Phi_{k+1}x_{k,j} + u_k + W_{k,j} \quad (2)$$

Φ_{k+1} はシステムの状態遷移に関わる線形モデル、 u_k は状態ベクトル、 $W_{k,j}$ はノイズを表す。また、観測量 z_k と状態変数 $x_{k,j}$ は式 (3) で表される観測方程式の関係を持つ。

$$z_k = H_k x_{k,j} + v_{k,j} \quad (3)$$

H_k は観測モデル、 $v_{k,j}$ はノイズを表す。状態変数 $x_{k,j}$ を式 (4) のように所定所要時間 $T_{k,j}$ と実所要時間 R_k を用いて定義した。

$$x_{k,j} = (T_{k,j}, R_k) \quad (4)$$

ただし、 $T_{k,j}$ は任意のバス停 k からバス停 j までの合計値、 R_k は始発駅からバス停 k までの合計値とした。

3.4 Extended Multiple Regression Filter モデル

次にバスの到着時刻予測を行う 2 つめのモデルとして、Extended Multiple Regression Filter モデル (以下、EMRF モデル) を提案する。EMRF モデルとは、3.3 節で算出した重回帰分析をカルマンフィルタ内に組み込むことにより、系統ごとの傾向を再予測するモデルである。MRKF モデルでは重回帰分析による予測結果を初期値入力のみで使用したため、事前予測傾向と大きく異なる運行をした場合に対応できなかった。EMRF モデルではバスの進行に合わせて重回帰分析を用いた予測の更新も行われるため、事前予測傾向と異なる運行にも対応でき、より精度が向上すると考える。

EMRF モデルの概念図を図 4 に示す。基本的な変数の定義、および挙動は 3.3 節で説明した MRKF モデルと同様であるため、改良点について説明する。入力値について、 $k = 1$ では、同様にシステムの初期状態として過去デー

タを用いた予測値を入力するが、 $k > 1$ では、バス停区間 $(k-1) \sim k$ の予測値と時刻表から、リアルタイムに遅れを算出し、重回帰分析によって再予測した結果を入力した。次に定義式について、式 (2) における状態ベクトル u_k の代わりに重回帰式 (式 (1)) を用いた。また、式 (4) で定義した状態変数 $x_{k,j}$ について、 $T_{k,j}$ を重回帰分析による予測所要時間 $E_{k,j}$ に変更した (式 (5))。

$$x_{k,j} = (E_{k,j}, R_k) \tag{5}$$

4. 使用データ

本論文では、NPO 法人位置情報サービス研究機構 (Lisra) を通じて提供された名古屋市交通局のバスロケーションデータ：2014 年 12 月 13 日 (土) から 12 月 19 日 (金)、および名鉄バス株式会社から提供されたバスロケーションデータ：2016 年 7 月 1 日 (金) から 15 日 (金)、2017 年 1 月から 3 月を使用した。各データの詳細について、以下にまとめる。

4.1 名古屋市営バス運行データ

提供されたデータには名古屋市内のバス 1,030 両、バス停 3,784 駅、系統 664 ルートが含まれており、1 日あたり約 220 万レコードのデータ量である。表 3 に名古屋市営バス運行データとして記録されるデータを示す。各運行バスには車載機番号としてバスごとにユニークな番号が振られており、車載器番号ごとに、系統 ID やイベント発生時刻などが記録される。イベントとは、バス停到着出発を示しており、イベント種別 (出発, 到着, その他) とイベント発生時刻および対象のバス停 ID が記録される。名古屋市営バス運行データでは、所定発時刻が記録されないため、名古屋市交通局ホームページから該当日時の時刻表を入手し、所定発時刻として利用する。また、名古屋市営バスでは、バスに搭載された GPS に基づき 30 秒ごとにバスの現在位置として緯度・経度を取得している。

これらの 664 ルートのうち、本論文は表 2 に示す 8 系統について 2014 年 12 月 13 日～20 日、22 日の運行データを対象に分析を行った。これらの 8 系統は、シンプルなルートをしており、かつ 1 日を通して運行本数の多い系統である。そのため、バス停区間ごとの遅延原因を分析しやすく、運行本数の多さからデータ量を確保でき解析のしやすい系統であると考えた。8 系統には、名古屋駅、栄など主要駅のバス停を含む系統 ([基幹 1] 栄 - 笠寺駅, [名駅 19] 港区役所 - 名古屋駅) と、JR 駅や地下鉄駅と住宅地を結ぶ系統 ([平針 12] 地下鉄原 - 地下鉄平針, [東巡回] 大曽根 - 茶屋ヶ坂) を含ませた。また、欠損値を含むデータについては除去を行っている。

表 2 使用した名古屋市営バスおよび名鉄バスデータ
Table 2 Nagoya city bus and Meitetsu bus data for analysis.

事業者	系統 ID	系統名
名古屋市営バス	8415	[基幹 1] 栄 - 笠寺駅
	8471	[名駅 19] 港区役所 - 名古屋駅
	8784	[黒川 14] 安井町西 - 黒川
	8921	[平針 12] 地下鉄原 - 地下鉄平針
	8939	[鳴子 14] 地下鉄鳴子北 - 大高駅
	8990	[東巡回] 大曽根 - 茶屋ヶ坂
	9014	[富田巡回] 戸田 - 戸田荘
	9015	[富田巡回] 戸田荘 - 戸田
名鉄バス	10430	JR～イオンモール岡崎～東岡崎

表 3 各事業者における運行実績データの記録内容

Table 3 Description of bus operation data.

種別	名古屋市営バス	名鉄バス
バスごとの ID	車載機番号	ダイヤ ID
系統情報	系統 ID	系統 ID, 系統表示名
バス停の発着	イベント種別 (出発, 到着, その他) とイベント発生時刻	実績発時刻
バス停情報	系統上のバス停 ID	停留所コード, 停留所名称, 並び順
所定発時刻	-	所定発時刻
バスの現在位置	30 秒ごとの緯度・経度	-
利用者数	-	乗車人数・降車人数

4.2 名鉄バス運行データ

提供されたデータには愛知県内のバス 710 両、バス停 1,539 駅、系統 523 ルートが含まれており、約 1 GByte のデータ量であった。名鉄バスでは、データの記録はバス停出発時にのみ行われており、運行バスごとにユニークなダイヤ ID が振られている (表 3)。さらにダイヤ ID ごとに、系統 ID, 系統表示名, ダイヤ ID, 停留所名称, 実績発時刻, 所定発時刻が記録される。また、名鉄バスでは利用者数として乗車人数, 降車人数が記録されている。今回の検証では、系統 523 ルートのうち、岡崎市内を走る系統のみを扱い、1 日に 1 ダイヤしかないものや、欠損を除く、67 系統を利用した。この 67 系統の分析には、2017 年 3 月 1 日～31 日の運行データを利用している。また、5.3 節ではデータ量を比較するため、2016 年 7 月 1 日～14 日、2017 年 1 月 1 日～3 月 31 日の運行データを利用した。本論文では、主に系統：JR～イオンモール岡崎～東岡崎 (図 5) を扱う。選定した系統：JR～イオンモール岡崎～東岡崎は岡崎市を運行する名鉄バスのうち平日朝夕方の通勤時間帯、休日の利用者数が多く、また遅延時間の大きいバス停である東岡崎駅、明大寺前のバス停を含む。到着時刻が時間帯によって変動するため、本論文の提案モデルによる遅延予測の効果が見えやすいと考え選定した。



図 5 [名鉄バス] 系統：JR～イオンモール岡崎～東岡崎 (決定係数 0.77)

Fig. 5 [Meitetsu Bus] Route: JR ~ Aeon Mall Okazaki ~ Higashi Okazaki (Coefficient of determination 0.77).

5. 系統ごとの特徴に関する検証

5.1 重回帰分析による決定係数

名鉄バスの 2017 年 3 月 1 日から 31 日までのデータに対して、重回帰分析を行った結果を図 6 に示す。対象の 67 系統に対し、決定係数を算出し、降順に並べたところ、決定係数は 0.90 から 0.46 となった。また、名古屋市営バスのデータの決定係数を図 7 にまとめる。

決定係数とは、説明変数が目的変数をどれくらい説明できるかを表す指標である。0 から 1 までの値をとり、1 に近いほど相対的な誤差が少ないと考えられる。名鉄バスの決定係数の平均は 0.69 となった。最も決定係数が示す値が大きかったのは、系統：本宿～富田病院～東部地域福祉セ～美合であった。しかし、係数の値を見ると 5.86×10^{11} といった不自然に大きい値になっており、重回帰分析が正しく行われていなかったことが分かった。これは、この系統にはバス停が 25 駅あり、説明変数が多すぎたためだと考える。一方、最も小さかったのは、系統：奥殿陣屋～康生町～東岡崎であり、これはバス停区間：岩津百々～青木町における不規則な渋滞が原因ではないかと考えられる。

名古屋市営バスの平均値は 0.76 であり、最も決定係数が高くなったのは、系統：[富田巡回] 戸田荘 - 戸田で 0.88 となった。また、最も低い値を示したのは、系統：[平針 12] 地下鉄原 - 地下鉄平針で、0.58 となった。系統：[富田巡回] 戸田荘 - 戸田では、バス停区間のうち、約 6 割の説明変数で t 値が 2 以上となったが、系統：[平針 12] 地下鉄原 - 地下鉄平針では、 t 値が 2 以上となったバス停区間の説明変数は約 4 割だった。 t 値とは重回帰分析の優位性を示す値であり、 t 値が大きいほどその説明変数が目的変数の予測に有効であることを表す。この説明変数の t 値の違いによって系統ごとの差が生じ、系統：[平針 12] 地下

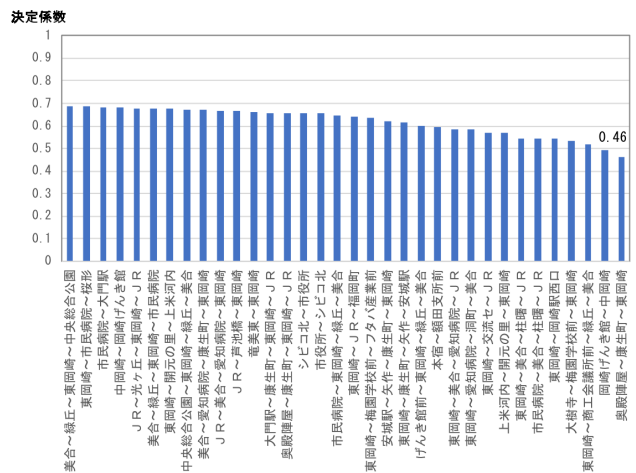
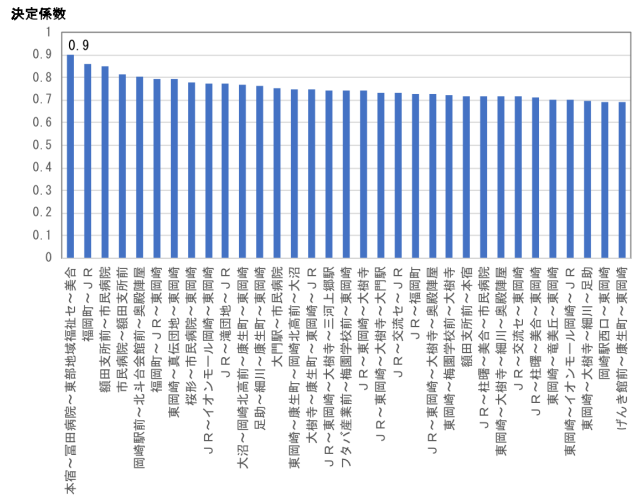


図 6 名鉄バスにおける各系統の決定係数
Fig. 6 Coefficient of determination (Meitetsu bus).

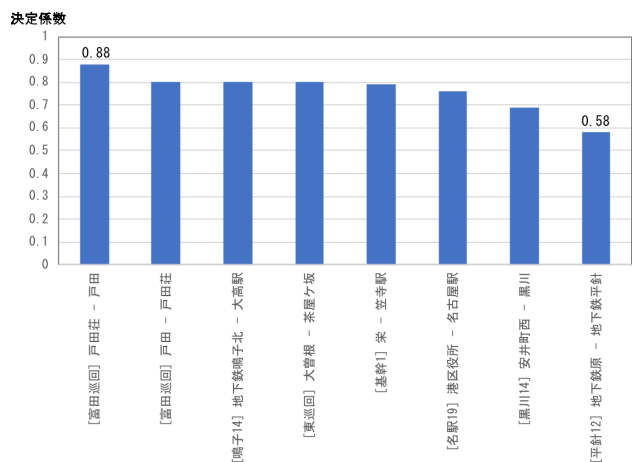


図 7 名古屋市営バスにおける各系統の決定係数
Fig. 7 Coefficient of determination (Nagoya city bus).

鉄原 - 地下鉄平針の決定係数が低くなったと考えられる。ただし、決定係数はいずれも 0.5 以上となりやや高い値を示しており、本論文の遅延予測には十分な精度であると考えられる。

R^2 値とバス停数の関係を図 8 に示す。 R^2 値とは、デー

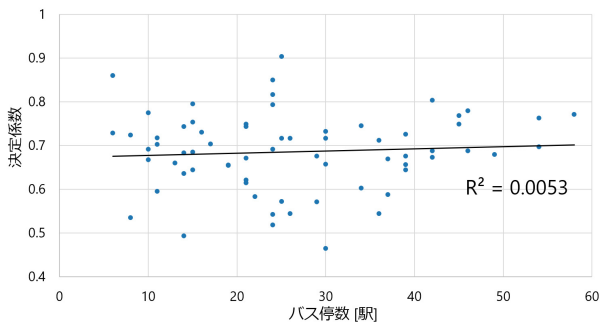


図 8 名鉄バスにおける R^2 値とバス停数の関係

Fig. 8 Relationship between the R^2 value and the number of bus stops (Meitetsu bus).

タをモデルでどれくらい説明できたかを表す値で、1 に近いほど良いと考えられる。決定係数は 0.0053 となり、相関は見られなかった。

5.2 重回帰分析による係数計算

名古屋市営バスにおける系統 8415 : [基幹 1] 栄 - 笠寺駅において重回帰分析を行った結果を表 4 に示す。学習には 2014 年 12 月 13 日から 19 日までの 7 日間の運行実績データを用いた。また、名鉄バスにおける系統 : JR ~ イオンモール岡崎 ~ 東岡崎において重回帰分析を行った結果を、表 5 に示す。学習には 2017 年 3 月 1 日から 30 日までの 30 日間を用いた。それぞれ「バス停区間 1-2」, 「Early Morning」, 「Monday」をそれぞれ基準 0 として係数を算出した。p 値とは、計算結果が統計的に有意かどうかを判断するためのものであり、重回帰分析の場合「その要素が他の要素に比べて重要ではない」という仮定が正しい確率を表す。p 値 < 0.05 であるものを統計的に有意であるとし、アスタリスクを表記した。ただし名古屋市営バスでは、バス停 1-2 間での所要時間は、始発駅バス停 1 における運転手による出発記録時刻が正確ではないため、考慮しないものとする。

5.2.1 バス停区間

表 4(a) に示す名古屋市営バスのバス停区間に関しては、大きく遅れるところや大きく遅れを回復する区間など、区間ごとの傾向の違いが確認できた。ほとんどのバス停区間で p 値は 0.05 以下であり、バス停区間は遅延に関して重要な要素であることが分かった。同様に、表 5(a) に示す名鉄バスのバス停区間に関しても、区間ごとに係数が変化し遅れの大きいバス停区間と小さい区間の傾向が見られる。

5.2.2 時間帯

表 4(b) に示す名古屋市営バスの時間帯に関しては、p 値が 0.05 より小さくなるものはなく、そのほかのパラメータより遅延への影響が少ないが、「Early Morning」の係数の値が最も大きく、遅延が生じやすいという結果になった。また、曜日に関しては、Friday と Saturday の影響が大き

表 4 [名古屋市営バス] 系統 : [基幹 1] 栄 - 笠寺駅における係数
Table 4 [Nagoya city bus] Coefficient of Route: [Kikan 1] Sakae - Kasadera St.

(a) バス停区間

バス停区間	名称	係数	p 値
2~3	栄~矢場町	0	-
3~4	矢場町~丸田町	-12.8	0.02 *
4~5	丸田町~鶴舞公園前	-121.4	0.00 *
5~6	鶴舞公園前~東郊通三丁目	8.9	0.07
6~7	東郊通三丁目~高辻	25.8	0.00 *
7~8	高辻~高辻	-50.4	0.00 *
8~9	高辻~雁道	30.6	0.00 *
9~10	雁道~堀田通五丁目	11.3	0.02 *
10~11	堀田通五丁目~牛巻	-19.8	0.00 *
11~12	牛巻~名鉄堀田	16.0	0.00 *
12~13	名鉄堀田~地下鉄堀田	-111.4	0.00 *
13~14	地下鉄堀田~呼続大橋	4.3	0.4
14~15	呼続大橋~千竈通二丁目	13.0	0.00 *

(b) その他の要素

時間帯	係数	p 値
Early Morning	0	-
Late Morning	-4.2	0.2
Early Noon	-4.2	0.2
Late Noon	-2.8	0.3
Evening	-6.5	0.05
Night	-3.4	0.4
曜日	係数	p 値
Monday	0	-
Tuesday	1.4	0.7
Wednesday	0.3	0.9
Thursday	4.5	0.2
Friday	11.0	0.00 *
Saturday	8.1	0.03 *
Sunday	6.9	0.06
その他	係数	p 値
遅れ	-1.0	0.00 *
平均停車時間	1.0	0.00 *
平均停止回数	38.5	0.00 *
切片	-53.4	0.00 *

いことが分かった。名鉄バスの、表 5(c) の時間帯に関しては、Late Noon と Evening が、曜日に関しては、Friday, Saturday, Sunday といった週末の影響が大きいと分かった。これは、この系統にはイオンモール岡崎というエンターテイメントモールがあり、朝より昼過ぎ~夕方、平日より休日の方が利用者が多くなるためだと考えられる。実際、図 9 に示すように、Friday, Saturday, Sunday の乗車人数は他の日に比べて多くなっている。

5.2.3 名古屋市営バス固有の説明変数

名古屋市営バスのみで示される遅れ、平均停車時間、平均停止回数に関しては (表 4(b)), p 値はすべて 0.05 より小さくなり、遅延に関して重要な要素であることが分かつ

表 5 [名鉄バス] 系統：JR～イオンモール岡崎～東岡崎における係数

Table 5 [Meitetsu bus] Coefficient of Route: JR ~ Aeon Mall Okazaki ~ Higashi Okazaki.

(a) バス停区間

バス停区間	名称	係数	p 値
1~2	岡崎駅前～岡崎市シビックセンター	0	-
2~3	岡崎市シビックセンター～戸崎口	-118.58	0.00 *
3~4	戸崎口～イオンモール岡崎	91.98	0.00 *
4~5	イオンモール岡崎～岡崎警察署	50.50	0.00 *
5~6	岡崎警察署～岡崎警察署前	-75.24	0.00 *
6~7	岡崎警察署前～芦池橋	-112.34	0.00 *
7~8	芦池橋～国立研究所下	-132.71	0.00 *
8~9	国立研究所下～明大寺町	107.95	0.00 *
9~10	明大寺町～東岡崎	-173.07	0.00 *

(b) n 駅前における遅延

n 駅前	係数	p 値
1 駅前	-0.12	0.00 *
2 駅前	0.09	0.00 *
3 駅前	0.02	0.21
4 駅前	-0.00	0.84
5 駅前	-0.01	0.47
6 駅前	-0.01	0.60
7 駅前	0.02	0.27
8 駅前	0.00	0.95
9 駅前	0.00	0.94

(c) その他の要素

時間帯	係数	p 値
Early Morning	0	-
Late Morning	-9.42	0.00 *
Early Noon	-0.44	0.00 *
Late Noon	11.95	0.00 *
Evening	9.53	0.00 *
Night	-11.62	0.00 *
曜日	係数	p 値
Monday	0	-
Tuesday	1.86	0.49
Wednesday	-0.10	0.97
Thursday	2.63	0.30
Friday	6.58	0.014 *
Saturday	24.92	0.00 *
Sunday	19.64	0.00 *
その他	係数	p 値
所定所要時間	0.37	0.00 *
移動人数	7.91	0.00 *
切片	123.18	0.00 *

た。それぞれについて具体的に見てみると、遅れに関しては負の係数として出ているため、1つ前の時点で時刻表よりも遅ければ早めようと、早ければ遅めようと、時刻表との差を埋めようとする運転手の働き [1] が表れているものと考えられる。平均停車時間に関しては、係数が 1.0 となり、バス停での停車時間がそのまま遅延につながる結果となった。平均停止回数に関しては、1 回の停止によって、約 40 秒の遅延が生じるという結果となった。

5.2.4 名鉄バス固有の説明変数

表 5 (b) および表 5 (c) に示す名鉄バス固有の説明変数について、2 駅前までの遅延、時間帯、金、土、日、所定所要時間、移動人数の影響が、他の要素に比べて大きいことが分かった。表 5 (b) では、2 駅より前のバス停が与える影響は小さく、係数として考慮する必要はないと考えられる。

5.2.5 p 値に関する考察

名古屋市営バスと名鉄バスに共通する説明変数であるバス停区間に関しては、p 値はおおむね 0.05 以下となり、遅延に関して大きな影響を与えていると考えられる。また、時間帯については名古屋市営バスでは比較的 p 値が大きくなったが、名鉄バスではすべての時間帯で p 値が大きくなった。これは名古屋市営バスで使用した系統がバスレーンを利用する基幹バスであり、混雑時間帯にバス専用規制が行われているため、時間帯ごとの道路の混雑状況にはさほど左右されないと考えられる。曜日については、Friday, Saturday が名古屋市営バス、名鉄バスで共通して遅延に影響した。

また、名古屋市営バス固有の説明変数である遅れ、停車時間、名鉄バス固有の説明変数である所定所要時間、遅延、移動人数についてはいずれも p 値が 0.05 以下であり、遅延に影響を与えると考えられる。しかし、係数についてはバス停区間の係数と比較し小さい係数となっており、影響は与えるものの大きな遅延にはつながりにくいと考えられる。ただし、停止回数については係数が比較的大きい。これは信号で停止した回数が遅延につながったと考えられる。

名古屋市営バスについて、基幹バス路線以外の 7 系統でもバス停止区間、遅れ、停車時間、停止回数の説明変数はほとんどで p 値が 0.05 以下となり、時間帯、曜日についてはほとんどの系統で p 値が 0.05 以上であった。ただし、停車時間については 0.05 以下となる系統と 0.05 以上となる系統があった。これは、バス停ごとの停車時間の差が大きい系統で、停車時間が遅延に影響する説明変数となったと考えられる。

以上のことから、バス停区間は提案手法の重回帰分析を行ううえで、重要な説明変数であり、時間帯についても基幹バス路線でない通常の系統では遅延に影響を及ぼす。また、曜日はバスの系統によって重要な説明変数となる系統とならない系統がみられた。遅れ、停車時間、停止回数、遅延、移動人数については、データの入手は事業者ごとの

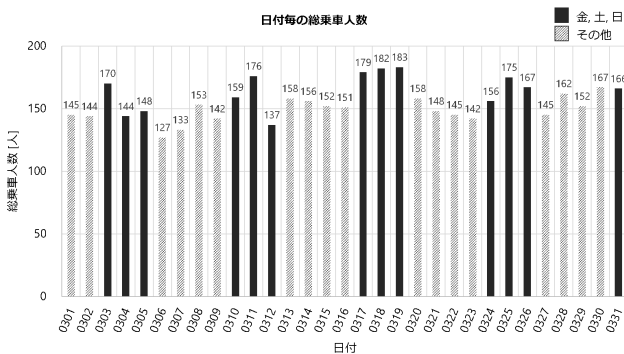


図 9 [名鉄バス] 系統：JR～イオンモール岡崎～東岡崎における日付ごとの乗車人数

Fig. 9 [Meitetsu bus] Number of Passengers of Route: JR ~ Aeon Mall Okazaki ~ Higashi Okazaki.

表 6 [名鉄バス] データ量の変化による誤差の変異

Table 6 [Meitetsu bus] Variation of error due to change in data.

	14 日分	101 日分	外れ値除去後
平均誤差 [秒]	36.4	38.0	28.6
最大誤差 [秒]	327.4	254.4	119.0
最小誤差 [秒]	0.1	0.4	0.18
標準偏差	45.4	36.6	23.2
R^2 値	0.30	0.42	0.53

記録データに依存するものの、それぞれ遅延に影響する説明変数であり、提案手法に組み込むことでより精度の高い予測が可能になると考えられる。

5.3 データ量の変化による重回帰分析の精度検証

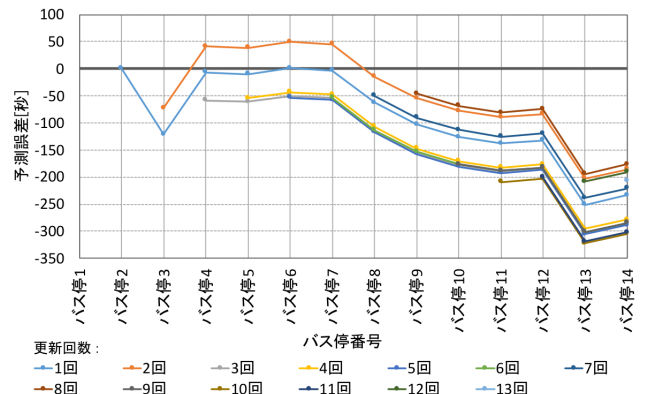
名鉄バスのデータに対して、データ量の変化による重回帰分析の予測精度への影響を検証した。2016年7月1日～14日までの14日間、7月に加え1月1日～3月31日までのデータを追加した101日間、および101日間のデータに対して外れ値を除去し、比較を行った。検証日はすべて2016年7月15日とする。系統：JR～イオンモール岡崎～東岡崎に対して、重回帰分析による予測の検証を行った結果を表6に示す。14日分のデータと101日分のデータでは誤差の平均値に大きな変化はないが、最大値や標準偏差はデータ量の増加により小さくなった。外れ値を除くと、平均値、最大値、標準偏差の値がさらに小さくなり、より予測精度が向上したと考えられる。一方、最小値に関しては14日分のデータでもかなり小さくなっており、データ量の増加による大きな変化は見られなかった。

R^2 値は、データをモデルでどれくらい説明できたかを表す値で、1に近いほど良いと考えられる。14日分のデータを用いた場合の R^2 値が0.30であったのに対し、101日分のデータでは0.42となり、データ量の増加による精度の向上を確認できた。また、外れ値を除いた場合の R^2 値は0.53となり、突発的なイベントや記録の不備などによる外れ値の除去により、精度はさらに向上した。以上のことから、重回帰分析に用いるデータ量の増加により、予測精度は向上すると考えられる。

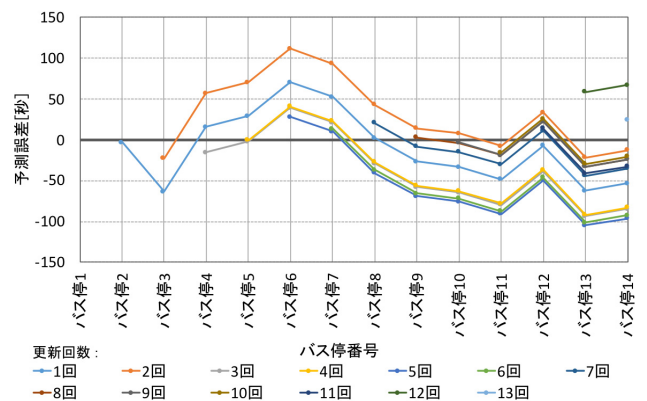
6. 提案モデルを用いたバス到着時刻予測に関する評価

6.1 予測誤差を用いた MRKF モデルの精度評価

MRKF モデルによる予測精度の向上を確認するため、Chen らによる研究 [10] と同様にカルマンフィルタの初期入力を所定所要時間とした場合との比較を行った。評価には名古屋市営バスの系統 8415：[基幹 1] 栄～笠寺駅における2014年12月15日（日）のデータを用いて、同年12月22日（日）の始発ダイヤが各バス停に到着するまでの所要時間を予測した結果を用いた。各モデルによる予測



(a) 初期入力が所定所要時間の場合



(b) MRKF モデルを用いた場合

図 10 [名古屋市長営バス] 予測誤差を用いた先行研究との精度比較
Fig. 10 [Nagoya city bus] Comparison of accuracy with previous research using prediction error.

値と運行データによる実測値との差を予測誤差とし、実測値より長く予測した場合は正、短く予測した場合は負の値をとる。バスの進行に合わせて予測の更新を行った結果を図10にまとめる。縦軸に予測誤差、横軸にバス停番号をとっており、予測誤差が0に近いほど予測結果が良いことを表す。たとえば図10(b)において、更新1回の予測ではバス停3における予測誤差が約-60秒ほどであったが、更新2回の予測では約-25秒ほどになっており、0秒に近づくように改善されているのが分かる。その後も同様に、全体的に更新回数が増えるにつれて予測誤差が小さくなるように改善されている。

図10(a)より図10(b)の方が全体的に予測誤差は0に近くなっており、MRKFモデルによる予測精度の向上が確認できた。また、図10(b)では、終点に近づくにつれ予測誤差が0に近づき、予測結果が修正されていることが読み取れる。しかし、MRKFモデルでは重回帰分析による予測の更新を行わないため、グラフに見られるようにバス停間の変化が等しくなった。

表2に示す8系統について、MRKFモデルによる平均予測誤差を図11に示す。縦軸は平均予測誤差を、横軸は、バス停の何駅前でいった予測であるかを表している。全体

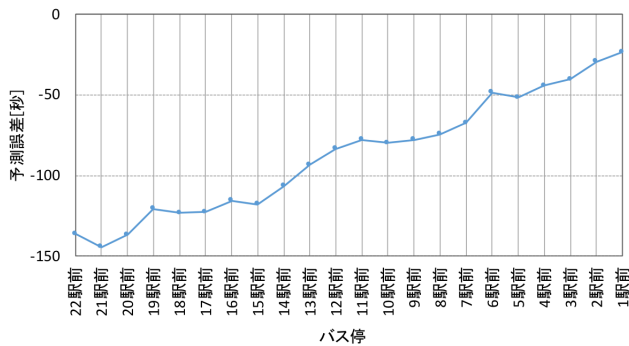


図 11 [名古屋市営バス] 平均予測誤差

Fig. 11 [Nagoya city bus] Average prediction error.

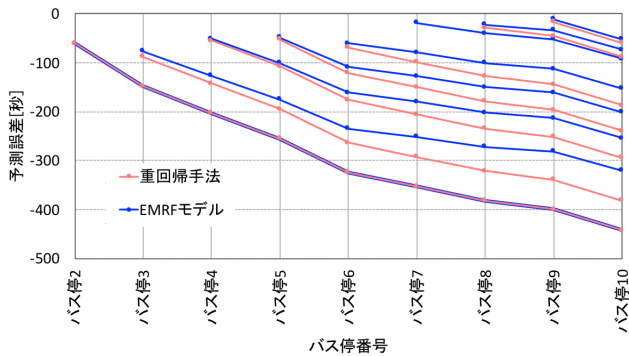


図 12 [名鉄バス] 予測誤差を用いた重回帰との精度比較

Fig. 12 [Meitetsu bus] Comparison of accuracy with multiple regression using prediction error.

として実所要時間よりも短く予測しているという結果となった。また、予測するバス停に近いほど予測誤差が小さくなる傾向を確認できた。10 駅前における予測で約 80 秒の誤差となっており、高い予測精度を実現できたと考える。

6.2 予測誤差を用いた EMRF モデルの精度評価

EMRF モデルによる予測精度の向上を確認するため、重回帰分析による事前予測結果との比較を行った。評価には、名鉄バスの系統：JR～イオンモール岡崎～東岡崎における 2017 年 3 月 1 日～30 日のデータを分析用、同年 3 月 31 日（金）のデータを検証用として用いた。ただし、検証日における移動人数と n 駅前で遅延には、30 日分の平均値を用いた。系統：JR～イオンモール岡崎～東岡崎における評価結果を図 12 に示す。図 10 と同様に、縦軸に予測誤差、横軸にバス停番号をとっており、予測誤差が 0 に近いほど予測結果が良いことを表す。横軸のバス停の並び順は、どの地点で予測を開始したかを表しており、EMRF モデルではバス停 1 については重回帰分析による事前予測の結果を用いている。図 12 より、重回帰分析により事前予測結果に比べ、EMRF モデルでは予測誤差が小さくなるように修正が行われていることが確認できる。

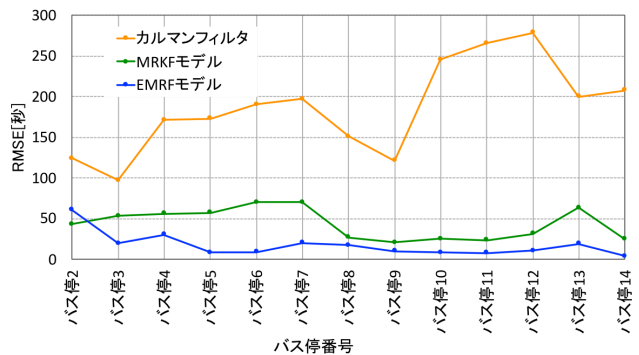


図 13 [名古屋市営バス] 各モデルの RMSE 値による予測精度比較

Fig. 13 [Nagoya city bus] Comparison of prediction accuracy by RMSE value of each model.

表 7 [名古屋市営バス] 各モデルの予測誤差

Table 7 [Nagoya city bus] Prediction error of each model.

	先行研究	MRKF	EMRF
最大値 [秒]	278.8	70.3	60.7
最小値 [秒]	97.0	20.6	3.9
平均値 [秒]	186.3	43.3	17.0
標準偏差 [秒]	53.3	18.2	14.3

6.3 RMSE 値による予測モデルの精度比較

MRKF モデルおよび EMRF モデルを、RMSE (Root Mean Squared Error) (式 (6)) を用いて評価する。

$$RMSE = \sqrt{\frac{1}{N-1} \sum_{k=1}^{N-1} (R_{k,k+1} - E_{k,k+1})^2} \text{ [秒]} \quad (6)$$

N はバス停の数であり、 $N-1$ はバス停の区間数を表す。また、 R_{k_1,k_2} 、 E_{k_1,k_2} はそれぞれバス停区間 $k_1 \sim k_2$ における実所要時間、予測所要時間を表す。RMSE は、予測値 (E_{k_1,k_2}) が実測値 (R_{k_1,k_2}) からどれくらい離れているかを示し、0 に近いほど予測が正確であるといえる。はじめに、名古屋市交通局の系統 8415：[基幹 1] 栄～笠寺駅を用いて評価を行った。Chen らによる研究 [10] をもとにした初期状態に所定所要時間を用いるモデル、MRKF モデル、および EMRF モデルによる予測精度を比較した。検証日は 2014 年 12 月 20 日（金）とし、13 日（金）から 19 日（木）までの 1 週間分のデータで各モデルを作成した。各 RMSE 値を図 13、各予測誤差の最大値、最小値、平均値、標準偏差を表 7 にまとめる。

図 13 において、先行研究に比べて MRKF モデルの方が RMSE 値が非常に小さいことから、初期状態に所定所要時間を用いるより重回帰分析による事前予測所要時間を用いた方が予測精度が高いという結果が得られた。これにより、名古屋市営バスに関しては、曜日や時間帯などにより運行状況に特徴的な傾向があり、過去データから事前予測を行うことで、より精度の高い予測が可能であると考えられる。また、バス停 1 における予測を除いて MRKF モデ

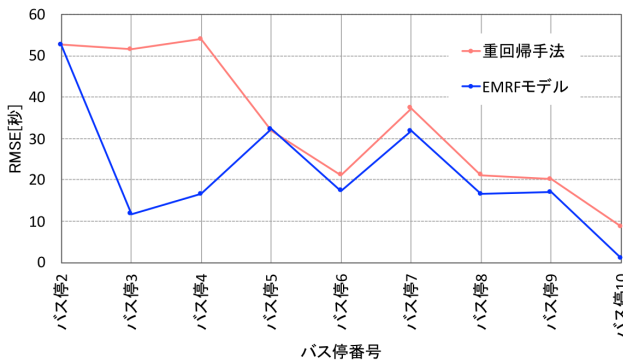


図 14 [名鉄バス] 系統：JR～イオンモール岡崎～東岡崎における RMSE 値の比較

Fig. 14 [Meitetsu] Comparison of RMSE values (Route: JR ~ Aeon Mall Okazaki ~ Higashi Okazaki).

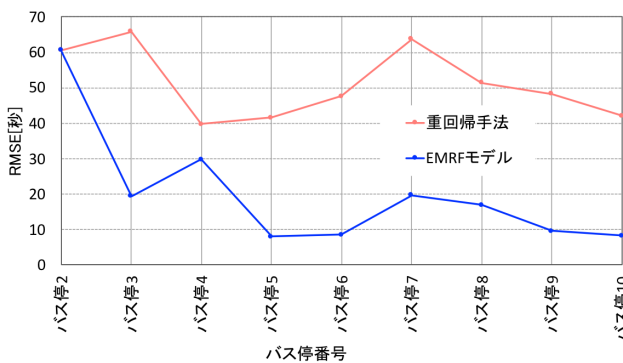


図 15 [名鉄バス] RMSE 値の平均値の比較

Fig. 15 [Meitetsu] Comparison of average RMSE values.

ルより EMRF モデルの方が RMSE 値が小さいことから、カルマンフィルタ内で重回帰分析による再予測を行うことで、予測精度が向上するという結果が得られた。

バス停 1 において MRKF モデルの方が予測精度が高かったのは、EMRF モデルでは事前予測結果をそのまま使用しているのに対し、MRKF モデルではバス停 1 における予測をカルマンフィルタを用いて予測し直しているためであると考えられる。これにより、運転手が時刻表の到着時間に近づけようと回復運転を行うことで、所要時間が状況により大きく変化した。事前予測のみを用いるよりもリアルタイムに再予測した結果を用いる方が、より高い精度で予測できると考えられる。MRKF モデルと EMRF モデルについては終点に近づくほど RMSE 値が小さくなる傾向にあり、バスの進行に合わせた精度の向上も確認できた。

表 7 に示す予測誤差の最大値について、先行研究より MRKF モデル、および EMRF モデルの方が値がかなり小さくなった。最小値については、先行研究、MRKF モデル、EMRF モデルの順でかなり小さくなった。EMRF モデルでは、カルマンフィルタ内で重回帰分析の再予測を繰り返すことにより予測誤差が改善され、最小値が MRKF モデルよりも小さくできたと考える。平均値についても最小値と同様に先行研究、MRKF モデル、EMRF モデルの順に

小さくなった。標準偏差については、先行研究より MRKF モデル、および EMRF モデルの方が小さくなっており、誤差のばらつきが小さくなり予測の安定性が増したと考えられる。以上のことから、先行研究と比較し、MRKF モデル、MRKF モデルより EMRF モデルの方が誤差が小さくなっており、提案モデルによる精度向上、および安定性の向上が確認できた。

次に、名鉄バスの系統：JR～イオンモール岡崎～東岡崎を用いて評価を行った。モデルの作成、および検証には、6.2 節と同様のデータを用いた。系統：JR～イオンモール岡崎～東岡崎について、重回帰分析による事前予測と EMRF モデルによる予測の各 RMSE 値を算出した結果を図 14 に示す。横軸がどのバス停で行った予測であるか、縦軸が RMSE の値を表している。図 14 より、全体的に重回帰分析による事前予測よりも EMRF モデルの方が RMSE 値が小さくなっており、特にバス停 2 において予測が大きく修正されていることが確認できた。また、検証日を 3 月内で 1 日から 31 日まで変化させて予測を行い算出した RMSE 値の平均値を図 15 に示す。図 15 より、1 カ月分のデータによる平均値で見ても EMRF モデルの方が RMSE 値が小さくなっており、予測精度を向上できたと考える。

7. 到着時刻予測の情報提示方法の検討

7.1 精度提示方法の概要

提案手法による予測結果を用いて、式 (7) により標準偏差 (SD: Standard Deviation) を用いて幅を持たせた予測所要時間 E' を算出し、約 95% の予測精度で利用者に提示する。 E は提案モデルを用いた予測所要時間、 R は実所要時間を表す。最短到着時刻の提示により、「早くてもこの時間なら、コンビニに寄ろう」、「急ぐ必要はないからゆっくり歩こう」など、利用者の行動選択肢を広げることができると考える [12]。また、最大到着時刻の提示には、バス停での待ち時間における「あとどれくらいで到着するのだろう」という不安を軽減する効果があると考えられる。提案モデルでは全バス停区間で予測が可能のため、降車駅への予測到着時刻も提示可能であり、利用者は目的地への到着予定時刻を知ることができる。さらに、提案モデルによりバスの進行に合わせて予測が更新されるため、渋滞などによる遅延の影響がリアルタイムに反映される (図 16)。このように具体的な到着予測時刻の提示により、利用者はより正確にバスの運行状況を把握でき、行動しやすくなると考えられる。

$$E' = E \pm 2SD(E - R) \tag{7}$$

7.2 提示方法に関する Web 調査

実際に利用者がバスの遅延や提示方法についてどのよう

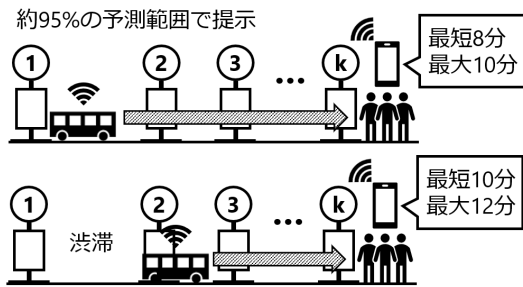


図 16 利用者への提示方法

Fig. 16 Method of presenting the predicted arrival time.

表 8 回答者の属性

Table 8 Attribution of subjects.

(a) 年齢 (人)		(b) 職業 (人)		(c) 性別 (人)	
10代	6	学生	78	男性	113
20代	112	社会人	99	女性	71
30代	24	主婦	2		
40代	27	その他	5		
50代	12				
60代	2				
その他	1				

な意識・意見を持っているのか確認するために、バス到着時刻予測に関する Web 調査を行った。調査項目は、利用目的や頻度、バスの遅延に関する意見、到着予測に関する意見など全 12 項目とした。集計期間は、2018 年 1 月 29~30 日であり、Web ページ上にアンケートを用意し、Facebook や Twitter などの SNS を通じてアンケート回答依頼の旨を拡散した。総回答数は 184 名で、そのうち有効回答数は 169 名であった。回答者属性について、表 8 にまとめる。在住の地域については愛知県が 107 名、東京都が 25 名を占めた。

7.3 調査項目

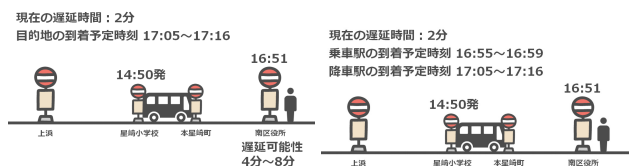
バスの遅延と到着予測に関する意見を尋ねた 9 項目について、調査項目および結果を述べる。調査項目は下記の 9 項目であり、質問 1-4 では「1 分以内」、「3 分以内」、「5 分以内」、「10 分以内」、「それ以上」の 5 つの選択肢を用意し、質問 5-8 では「1: 参考にしない」、「2: あまり参考にしない」、「3: どちらでもない」、「4: 少し参考にする」、「5: 参考にする」の 5 段階評定とした。質問 9 では、図 17 に示すような提示画面を作成した。スマートフォン画面に表示されることを前提に、遅延時間、予測到着時刻、残り時間、グラフ表示の 4 種類の提示画面を作成し、最も好ましい提示方法を選択してもらった。

質問 1 通勤・通学時、バスの遅延はどれくらいまで許せますか。

質問 2 買い物などで目的地へ向かうとき、バスの遅延はどれくらいまで許せますか。

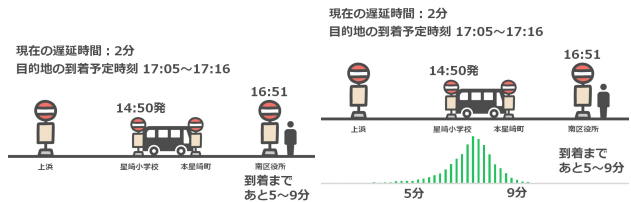


(a) スマートフォンへの予測表示



(b) 「遅延時間」

(c) 「予測到着時刻」



(d) 「残り時間」

(e) 「グラフ表示」

図 17 表示画面

Fig. 17 Example screen.

質問 3 帰宅時、バスの遅延はどれくらいまで許せますか。

質問 4 予測誤差はどれくらいまで許せますか。

質問 5 予測誤差が 0 分~1 分程度生じる可能性がある場合、どれくらい参考にしますか。

質問 6 予測誤差が 1 分~5 分程度生じる可能性がある場合、どれくらい参考にしますか。

質問 7 予測誤差が 5 分~10 分程度生じる可能性がある場合、どれくらい参考にしますか。

質問 8 予測誤差が 10 分以上生じる可能性がある場合、どれくらい参考にしますか。

質問 9 スマホ画面の「?」の部分 (図 17(a)) に下記の予測画面 (図 17(b), (c), (d), (e)) が表示されます。スマートフォンを用いたバスの到着予測についてどの表示方法が 1 番好ましいですか。

7.4 調査結果と考察

質問 1-3 までの回答結果を図 18 に示す。質問 1 の通勤・通学時のバスの遅延について「5 分以内」が最も多く、半数以上の割合を占めた。次に「3 分以内」、「10 分以内」と続いた。質問 2 の買い物などで目的地へ向かうときのバスの遅延について、「10 分以内」が最も多く、約半数の

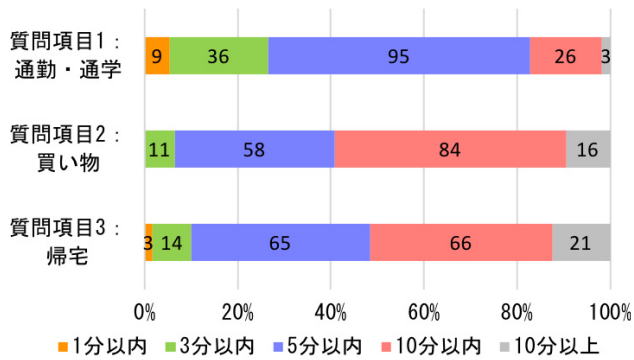


図 18 質問 3-5: 遅延についての意識

Fig. 18 Question3-5: Point of view about delay.

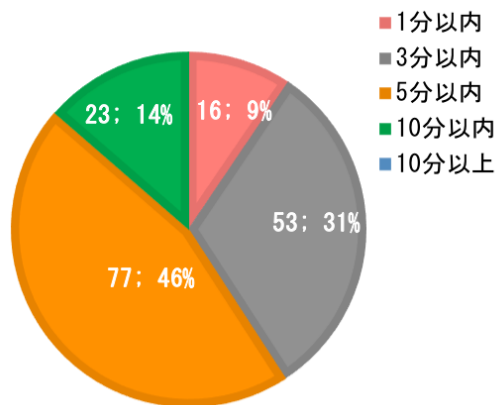


図 19 質問 4: 予測誤差についての意識

Fig. 19 Question4: Point of view about prediction error.

割合を占めた。次に「5分以内」の回答者が58名であり、30%程度を占めた。質問3の帰宅時のバスの遅延について、「10分以内」と「5分以内」がそれぞれ66名、65名であり、共に約40%を占めた。バスの遅延時間については、利用目的に応じて異なる特徴が見られた。通勤・通学時については、「5分以内」が約半数を占めたが、買い物や帰宅時については「5分以内」に加えて「10分以内」も多く回答されており、通勤・通学時に比べて利用者の行動に余裕があると考えられる。通勤・通学、買い物、帰宅時のいずれの場合も、1分以内の遅延のみ許容できるという回答は非常に少なかった。

予測誤差について質問4の結果を図19に示す。「5分以内」が46%、「3分以内」が31%と続いた。次に多かったのは「10分以上」で、14%であった。その詳細として、質問5-8までの結果を図20に示す。質問5の予測誤差が0分~1分程度生じる可能性がある場合では、「5:参考にする」が150名と最も多く、全体の90%近くを占めた。質問6の予測誤差が1分~5分程度生じる可能性がある場合では、「4:少し参考にする」が最も多く、「5:参考にする」、「3:どちらでもない」と続いた。この3つの評価はそれぞれ、70名、58名、38名となり全体の90%以上を占めた。

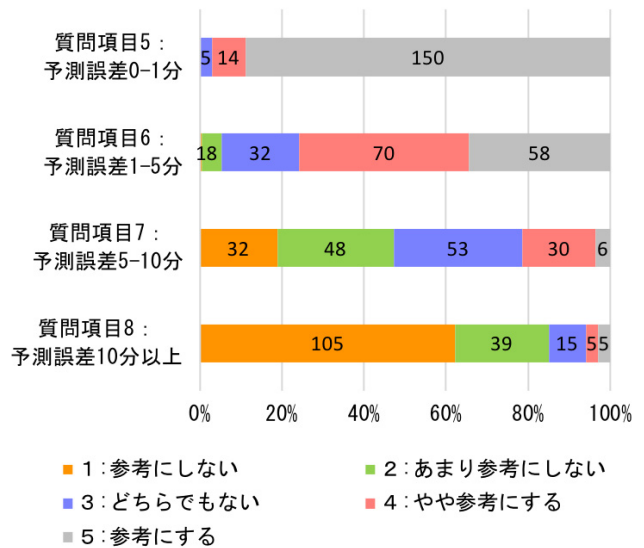


図 20 質問 5-8: 予測誤差についての意識

Fig. 20 Question5-8: Point of view about prediction error.

質問7の予測誤差が5分~10分程度生じる可能性がある場合では、「3:どちらでもない」が最も多く、「2:あまり参考にしない」、「1:参考にしない」と続いた。この3つの評価はそれぞれ、53名、48名、30名となり全体の約80%を占めた。質問8の予測誤差が10分以上生じる可能性がある場合では、「1:参考にしない」が105名と最も多く、全体の約60%を占めた。

予測誤差については、質問4で「3分以内」と「5分以内」の回答が70%以上を占めており、予測モデルの精度基準として「予測誤差3分以内」が考えられる。これは、本予測モデルを用いた情報提示をバス利用者の利便性向上に役立てるための1つの基準であり、本論文で提案したEMRFモデルは予測誤差の平均値17.0秒、最大値60.7秒を達成していることから(図7)、我々が目的とする到着時刻予測の情報提示に資する精度となっていると考えられる。

さらに質問5-8の予測誤差をどれくらい参考にするかについては、5分以内であれば90%以上が「3:どちらでもない」より上の評価を回答しており、予測誤差5分が1つの目安になると考える。一方、予測誤差が5分~10分程度になると、評価は大きく下がり、半数近くの利用者が「1:参考にしない」、「2:あまり参考にしない」を回答した。10分以上になると約80%の利用者が「1:参考にしない」、「2:あまり参考にしない」を回答しており、現実的に運用不可能であると考えられる。本論文では、雨天時やバス停近隣でのイベント開催などによる乗客増加や交通状況の変化など突発的な遅延要因に特化した分析は対象としていないが、精度低下が考えられるこのような状況においては、5分以内の予測誤差を精度目標とすることで情報提示の円滑な提供が行えるという1つの指針が得られた。

今回の調査では、予測誤差について利用シーンを特定せず、許容できる誤差を尋ねたが実際には、確実に利用者が

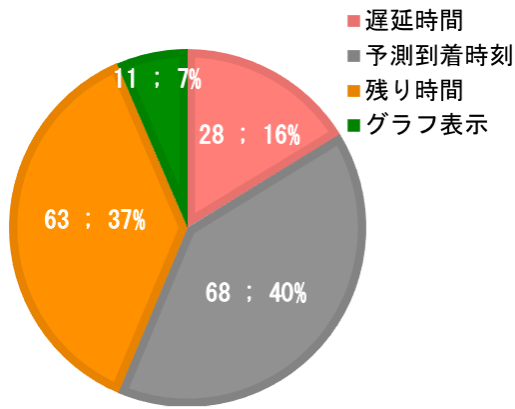


図 21 質問 9：予測の提示方法

Fig. 21 Question9: About predicted arrival time.

目的のバスに乗らなければならない状況と、乗り遅れた場合は次のバスに乗ればよい状況で許容できる予測誤差は異なると考えられる。確実に目的のバスに乗らなければならない状況では、予測誤差が0分~1分程度生じる可能性がある場合には、利用者は到着予測時刻の1分前程度からバス停でバスを待ち、遅くとも2分程度バスを待てば、乗車できるが、到着予測時刻に5分から10分の予測誤差がある場合、利用者は到着予測時刻の10分前からバス停で待ち、長い場合には到着予測時刻の10分後まで20分間バスを待たなければならない。ただし、バスの時刻表上で所定の到着時刻が到着予測時刻の10分前より遅い場合は所定時刻からがバスの待ち時間となる。しかし、乗り遅れより待ち時間を減らしたい状況では、到着が予測される時刻にバス停に向かうことで待ち時間を削減することができる。今後は、今回の調査で得た精度基準をもとに、許容できるバスの遅延時間とあわせて様々な状況におけるより詳細な精度目標を調査する必要がある。

最後に、質問9の予測の提示方法についての結果を、図21に示す。「予測到着時刻」が約40%で最も多く、次いで「残り時間」、「遅延時間表示」と続いた。「予測到着時刻」と「残り時間」が全体の約80%を占めた。利用者は「遅延時間」よりも、実際にバス停に到着する時刻や時間を示す「予測到着時刻」と「残り時間」を知りたいと分かった。今回の調査では、シンプルな提示方法が好まれたが、実際には予測が外れ利用者がバスに乗れない状況を避けるため、バスがどの程度の確率で予測より早く来るか遅く来るかを合わせて提示する必要があると考えている。今回の調査で、予測誤差を同時に示す提示方法を好んだ回答者は少なかったことから、予測誤差についてシンプルな提示方法を検討する必要がある。

8. まとめと今後の課題

本論文では、複数事業者の運行実績データに対して、重回帰分析とカルマンフィルタを併用したリアルタイムバス

到着時刻予測モデルを提案した。提案モデルは、重回帰を用いて算出した予測結果をカルマンフィルタの初期値に入力するMRKFモデルと、その改良モデルであるカルマンフィルタの中で重回帰分析の再予測を行い、その予測値と時刻表を比較するEMRFモデルである。事業者・系統ごとに重回帰分析を行い、事業者・系統ごとの特徴や重回帰を用いたデータ量の影響と傾向を検証した。

提案モデルについて予測誤差とRMSE値を用いた評価を行い、従来研究に比した予測精度の向上、および終点に近づくにつれて修正が行われたことを確認した。特に、EMRFモデルでは平均予測誤差が従来手法の約186秒から約17秒まで向上した。また、予測結果およびその誤差提示する方法を提案し、Web調査を行った。調査結果から、バス利用者が遅延時間や予測誤差についてどのような意識を持っているか確認した。予測誤差の利用者の許容範囲についての調査結果から、予測モデルの精度基準として「予測誤差3分以内」が考えられ、本論文で提案したEMRFモデルは予測誤差の平均値17.0秒を達成しており、我々が目的とする到着時刻予測の情報提示に十分な精度となった。また、本論文で用いたバスロケーションデータには含まれない突発的な遅延が発生する場合であっても、5分以内の予測誤差が精度目標として得られ、情報提示サービスを構築するうえでの1つの指針が得られた。

今後の課題について以下に示す。

- 提案モデルの改善

本研究では、ダイヤごとに1つ前のバス停における情報を用いて重回帰分析による予測結果を更新した。一方、前ダイヤの実績データは、同じ路線を走る最新の情報であり、予測の更新に最適であると考えられる[24]。そこで、本研究の更新に加え、前ダイヤの情報を用いた更新を行うことにより、リアルタイムな運行状況に対応した予測結果をより早く・正確に提供できる可能性が高い。また、本論文では名古屋市営バスおよび名鉄バスを利用して重回帰分析およびカルマンフィルタを用いた遅延予測を行った。利用した運行データは、名古屋市営バスが約1週間、名鉄バスが3カ月程度であり、4月の通勤・通学での利用者の増加や、8月の夏休み期間による通学での利用者の減少など、季節性を鑑みたモデルの改善も重要であると考えられる。さらに、今回の分析では、名古屋市・岡崎市地域のバス運行データを用いたため、比較的1日の運行本数の多いシステムを用いた予測であった。今後は、運行本数が少ない地域など他地域の運行データを取り入れ、地域性を考慮した精度検証も必要である。

- 実証実験とサービス化

提案モデルを用いた到着時刻予測について、実証実験を行い利用者の声を取り入れた改良が必要である。実証実験により利用者の意見を反映させ、提案モデルを

サービス化し、利用者がより快適にバスを利用できる社会が実現されるであろう。

謝辞 バス運行データを提供いただいた名鉄バス株式会社、および株式会社メイツコム、名古屋市営バスに感謝します。また、本研究はJST OPERAの支援および総務省SCOPE(受付番号172106102)の委託を受けて実施した。

参考文献

- [1] 吉田長裕, 内田 敬, 日野泰雄: GPSを活用したバスの定時運行評価に関する事例分析, 土木計画学研究・講演集, Vol.27, XII(172) (2003).
- [2] 前川裕一, 林 夏美, 牧野友哉, 白石 陽: バス到着時刻におけるバス運行所要時間データと乗降客数データの活用, マルチメディア通信と分散処理ワークショップ論文集, No.6, pp.165-171 (2013).
- [3] 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, 情報処理学会第76回全国大会, 2V-3, pp.157-158 (2014).
- [4] 頭川正信, 高野伸栄, 萩原 亨: バス運行データを用いた路線バスの運行評価に関する研究, 土木計画学研究・講演集, pp.285-288 (1998).
- [5] 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, 情報処理学会第78回全国大会, 4U-07 (2016).
- [6] クラウド Watch: 京都市営バスがiBeaconで通知, 停留所の画面に「まもなく到着」, 入手先 (http://cloud.watch.impress.co.jp/docs/news/20141104_ghg674245.html) (参照 2016-05).
- [7] 大野雄作, 辰巳 浩: 路線バスの運行所要時間予測に関する基礎的研究, 土木計画学研究・講演集, ROMBUNNO.104, Vol.38 (2008).
- [8] 大野雄作, 辰巳 浩: バスプローブデータを用いたバス運行実態の解析手法に関する一考察, 土木学会西部支部研究発表会講演概要集, IV-013, pp.593-594 (2008).
- [9] Patnaik, J., Chien, S. and Bladikas, A.: Estimation of Bus Arrival Times Using APC Data, *Journal of Public Transportation*, Vol.7, No.1, pp.1-20 (2004).
- [10] Chen, M. Liu, X. and Xia, J.: Dynamic Prediction Method with Schedule Recovery Impact for Bus Arrival Time, *Transportation Research Record Journal of the Transportation Research Board*, pp.208-217 (2005).
- [11] 今井 瞳, 廣井 慧, 河口信夫: 路線バスの運行データ分析に基づく到着時刻予測と精度解析, マルチメディア, 分散協調とモバイルシンポジウム 2016 論文集, pp.631-637 (2016).
- [12] 鴨原育子, 山田 稔, 齊藤 修, 兼子恭平: 利用者位置から検索するバスナビゲーションシステムに関する研究, 土木学会論文集 F3 (土木情報学), Vol.70, No.2, pp.293-302 (2014).
- [13] 井関崇之, 佐藤誠治, 小林祐司: 大分都市圏におけるバス交通利便性評価と整備課題に関する研究, 日本建築学会九州支部研究報告, Vol.50, pp.373-376 (2011).
- [14] ドイツ鉄道ホームページ, 入手先 (<http://www.deutschebahn.com/de/start/>) (参照 2018-01).
- [15] 中村文彦, 新谷洋三, 太田勝敏: バス運行実績データを用いたバス運行状況に関する研究, 土木学会年次学術講演会講演概要集第4部, Vol.41, pp.247-248 (1986).
- [16] 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, 情報処理学会第78回全国大会, 4U-07, pp.395-396 (2016).
- [17] 工藤 宏, 故 振程, 内村圭一: 路線バス運行情報を用いたバス停留所要時間予測, 電気関係学会九州支部連合

大会 (2004).

- [18] 元田良孝, 宇佐美誠史, 金澤 崇: バス乗降時間の要因に関する基礎研究, 交通工学研究発表会論文報告集, Vol.25, pp.269-272 (2005).
- [19] 桜井博隆, 廣島康裕: 地方都市における公共交通サービスに対する利用者満足度と路線収支に関する分析, 土木学会第58回年次学術講演会, IV-419, pp.837-838 (2003).
- [20] 伊藤昌毅, 川村尚生, 菅原一孔: スマートフォンを利用したバスロケーションシステムの開発, 電子情報通信学会論文集, pp.2327-2339 (2013).
- [21] 内村圭一, 成松裕介, 衛藤旭秀, 胡 振程: バスロケーションシステム情報を用いたバス停留所要時間予測, 国際交通安全学会誌, Vol.32, No.3, pp.224-231 (2007).
- [22] 安田幸司, 塩土圭介, 吉村英二, 市川晴雄, 浅井加寿彦, 中川真治: プローブカーデータを用いたバス運行状況の評価に関する実証的分析, 土木計画学研究・講演集, Vol.28, IV(111) (2003).
- [23] 工藤宏之, 故 振程, 内村圭一: 路線バス運行情報を用いたバス停留所要時間予測, 電気関係学会九州支部連合大会, 10-1P-06, 526 page (2004).
- [24] 坂本邦弘, 菅野光司, 久保田尚, 谷島 賢: 利用者ニーズに対応した到着時刻予測に基づくバス情報システムの実用的開発, 土木計画学研究・論文集, No.20, pp.857-864 (2003).
- [25] Nagatani, T.: Chaos control and schedule of shuttle buses, *Physica A: Statistical Mechanics and its Applications*, Vol.371, No.2, pp.683-691 (2006).
- [26] Zhang, C. and Teng, J.: Bus Dwell Time Estimation and Prediction: A Study Case in Shanghai-China, *Proc. Social and Behavioral Sciences*, Vol.96, pp.1329-1340 (2013).



今井 瞳

2016年名古屋大学工学部電気電子・情報工学科卒業。2018年同大学大学院修士課程修了。



廣井 慧 (正会員)

2004年東北大学工学部電子工学専攻卒業。同年東日本電信電話株式会社入社。2011年慶應義塾大学大学院メディアデザイン研究科修士課程修了。2014年慶應義塾大学大学院メディアデザイン研究科博士課程修了。博士(メディアデザイン学)。同年名古屋大学未来社会創造機構特任助教を経て、2018年から同大学工学研究科助教。災害情報通信、センサネットワークの研究に従事。



河口 信夫 (正会員)

1990年名古屋大学工学部電気電子工学科卒業。1995年同大学大学院工学研究科情報工学専攻博士課程満了。同年同大学工学部助手。同大学講師，准教授を経て，2009年より同大学大学院工学研究科教授。NPO 位置情報サー

ビス研究機構 Lisra 代表理事。モバイルコミュニケーション，ユビキタスコンピューティング，行動センシングの研究に従事。博士（工学）。ACM, IEEE, 人工知能学会，日本ソフトウェア科学会，電子情報通信学会，日本音響学会各会員。