

# 都営バスのオープンデータによる渋滞検知の精度向上のための運行特性を考慮した手法の提案

畠中 希<sup>†</sup> 青柳 宏紀<sup>‡</sup> 藤田 智也<sup>‡</sup> 山名 早人<sup>‡</sup> 小口正人<sup>†</sup>

<sup>†</sup>お茶の水女子大学 〒112-8610 東京都文京区大塚 2-1-1

<sup>‡</sup>早稲田大学 〒169-8555 東京都新宿区大久保 3-4-1

E-mail: <sup>†</sup> {hatanaka.nozomi, oguchi}@is.ocha.ac.jp, <sup>‡</sup> {aoyagih, tfujita, yamana}@yama.info.waseda.ac.jp

**あらまし** 交通渋滞の解消は喫緊の課題である。渋滞情報は道路上に設置されている感知器等でリアルタイムに収集されているため感知器が設置されていない道路では渋滞を検知できない。他方で路線バスは地域住民の移動手段であるため、感知器がない道路も走行している。よって、本研究では網羅的に渋滞を検知するためオープンデータとして提供されている都営バスのデータを用いる。

また、バスには様々な運行特性がある。例えば、バスが定刻より早く運行している際、速度を落として運行する可能性がある。このような運行特性を考慮し特徴量の追加やデータの前処理を行うことで渋滞検知の精度の向上を検討する。

**キーワード** 渋滞、オープンデータ、バス、機械学習

## 1. はじめに

交通渋滞は様々な悪影響をもたらしており、国土交通省によると日本における渋滞の経済的損失は年間12兆円、時間的損失は1人あたり年間30時間に上ると試算されている。したがって、このような損失を抑えるため渋滞情報をリアルタイムで取得し、渋滞を回避することは重要である。

従来の渋滞情報は、道路上に設置されてある車両感知器や光ビーコン等から収集されている。しかし、道路によっては感知器等が設置されていない場合もあるため、全ての道路が網羅されている訳ではない。

他方で、公共交通機関の1つである路線バスは、地域住民の移動手段であるため、主要でない道路を走行する場合もある。そのため、感知器等がない道路も走行している。また、官民データ活用推進基本法（平成28年法律第103号）において、国及び地方公共団体はオープンデータに取り組むことが義務付けられた。そのため、近年では公共交通機関に関するデータもオープンデータとして公開されており、路線バスのデータも含まれている。

さらにバスには定刻より早く運行している際、速度を落として運行するなどの運行特性があると考えられる。よって、本研究では、路線バスの1種である都営バスのオープンデータを用いて、運行特性を考慮した渋滞検知手法を提案する。

て提案した。一般に、バスの運行頻度が少ない時間帯は速度推定の精度は低くなるという問題点がある。そこで、Samal らは、気象データ等の外部要因のデータを利用し精度の向上を目指した。k-mean 法により、類似した運行データについてクラスタリングを行い、予測精度を改善した。交通速度の RMSE が、クラスタリングを適用しない場合は毎時 4.0~4.5 マイルであったのに対し、クラスタリング適用後は毎時 2.9~3.3 マイルとなることを示した。

Kyaw ら[2]は、バスの GPS データに加え特徴量として、道路の混雑要因となる、道路付近のレストランやショッピングモールといった POI(Place of Interests)の数を追加し、交通速度の推定手法を提案した。

Gu ら[3]は、2020 年、バスの GPS 軌道データをもとに、自己組織化写像を用いて、バス停間の道路セグメントの交通渋滞を推定する手法を提案した。Gu らは、バスの停留所での停車時間、バスの平均速度、乗客のバス停間の旅行時間を入力としたニューラルネットワークを用いる手法を提案した。本手法によって、推定した交通速度と交通速度を代表するタクシー速度との相関係数が 0.94 隣、両者の間に高い相関があることを示した。

## 2.2 プローブ車両の GPS データを用いた渋滞検知

Xu ら[4]は、2012 年、複数台のバスの平均移動時間を算出し渋滞検知を行う手法を提案した。T 分間に渋滞検知対象道路を通過したバスの平均移動時間を算出する「T-window Average」と、渋滞検知対象道路を通過した、直近の N 台のバスの走行時間を記録し、平均走行時間を算出する「N-window Average」を定義し、

## 2. 関連研究

### 2.1. プローブ車両を用いた交通速度推定

Samal ら[1]は、2017 年、バスのリアルタイム GPS 軌道データを利用した交通速度推定を行う手法につい

それぞれが個別に設定した閾値を超えた場合に、渋滞が発生していることをシミュレーションにて示した。

Wang ら[5]は、2013 年、1 台の車両からのデータを用い、機械学習にて交通状態を分類する手法を提案した。学習アルゴリズムとしてランダムフォレスト、Adaboost, SVM を用いて行い、それぞれ 91.59%, 89.43%, 87.86%の精度で検知できることを示した。

Carli ら[6]は、2015 年、バスの GPS プローブデータによる効率的な自動渋滞検知手法を提案した。バスの平均速度や、バスが閾値を超えた速度で走行した時間、といった指標を計算し、指標の以上を検知することで、信号機故障の異常を検出した。

### 3. 先行研究

青柳ら[7]は都営バスのリアルタイム運行データと機械学習を組み合わせた渋滞検知手法を提案した。渋滞検知の対象区間を渋谷から池袋間の明治通りとし、停留所ごとの発車時刻のデータから走行速度を算出し、一定の時間帯を「渋滞」と「非渋滞」とする 2 値分類を行った。全停留所に対して学習器を 1 つ用意した場合の F1 スコアは 0.306、停留所区間ごとに個別の学習器を用意した場合の F1 スコアは平均 0.399、特定の区間だと 0.675 となった。また、一定の時間に複数のバスが走行する場合のみのデータを使用することで精度向上を目指した。その結果、特定の区間において F1 スコアが 0.742 へ向上した。

## 4. 提案手法

### 4.1 概要

本研究では青柳ら[7]の提案手法をもとにバスの運行特性を考慮した手法を提案する。

一般に渋滞検知では、渋滞を「時速 10km 未満で走行している状態」と定義する。本研究においても、先述した定義を利用するため連続する 2 つの停留所の発車時刻の差異を停留所区間の距離で割ることで走行速度を求める。また、バスが定刻より早く走行している場合、運転手はわざと速度を落として運行する可能性があり、誤って渋滞であると判断される場合がある。したがって、定刻との差異を特徴量に含めることで先行研究の精度向上を目指す。

### 4.2 特徴量抽出

提案手法で用いる記号を表 4.1 に示す

表 4.1 特徴量抽出に用いる記号の定義表

記号	定義
$b_i$	$i$ 番目のバス
$p_j$	$j$ 番目のバス停
$s_j$	$j$ 番目の区間
$l_j$	区間 $s_j$ の距離
$td_{ij}$	$b_i$ の $p_j$ の出発時刻
$\Delta td_{ij}$	区間 $s_j$ にかかる所要時間
$v_{ij}$	$b_i$ の $s_j$ での平均速度
$T_{ij}$	$b_i$ が $p_j$ を出発する時間を 20 分ごとに分類した指標
isFast	$b_i$ が定刻より早く走行しているか

本手法で用いるデータは、バス  $b_i$  が  $p_j$  を  $td_{ij}$  に出発したというデータである。まず、 $b_i$  の出発時刻のデータを区間  $s_j$  の順番になるよう並び替え、バス  $b_i$  のバス停  $p_j$  の出発時刻  $td_{ij}$  と、バス  $b_i$  のバス停  $p_{(j+1)}$  の出発時刻  $td_{i(j+1)}$  を抽出する。 $\Delta td_{ij}$  は出発時刻  $td_{ij}$  と  $td_{i(j+1)}$  から次式(1)で表される。

$$\Delta td_{ij} = td_{i(j+1)} - td_{ij} \quad (1)$$

次に、バス  $b_i$  の区間  $s_j$  における平均速度  $v_{ij}$  を算出する。平均速度  $v_{ij}$  は次式で表される。

$$v_{ij} = \frac{l_j}{\Delta td_{ij}}$$

次に、バス  $b_i$  のバス停  $p_j$  の出発時刻  $td_{ij}$  を 20 分ごとの時間帯  $T_{ij}$  に分類する。 $T_{ij}$  の値は、0 時 0 分 0 秒から 0 時 19 分 59 秒を 0 として、以後 20 分おきに 1, 2, 3, ... と定義する。

次に isFast はバス  $b_i$  が定刻より早く走行しているか否かを表している。定刻より早く走行している場合は True, そうでない場合は False とする。

### 4.3 分類手法

バス  $b_i$  が走行した、バス停区間  $s_j$  において「渋滞」「非渋滞」の二値分類を行う。学習器は、ランダムフォレスト、AdaBoostXGBoost, サポートベクターマシン(SVM)を用いる。

## 5. 評価実験

### 5.1 渋滞検知の対象区間

本実験で渋滞検知を行う対象区間を図 5.1 に示す

本実験で対象とする区間は、池 86 系統の「池袋駅東口」～「渋谷駅東口」間とする。オープンデータとして得られるデータが、バス停出発時刻のデータのみであるため、終点のバス停については、出発時刻を得ることができない。そのため、本実験では終点区間に対しては、渋滞検知を行わないこととする。

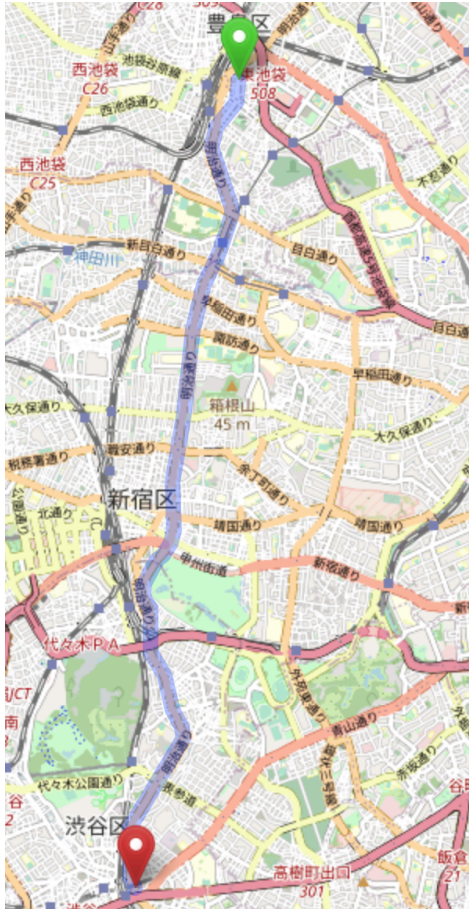


図 5.1 渋滞検知対象区間(OpenStreetMap より)

## 5.2 データセット

### 5.2.1 バスの運行データの取得方法

都営バスの運行データは、「公共交通オープンデータセンター」の API から取得した。本 API は、URL を指定しリクエストを送ると、公共機関に関するデータを JSON 形式で返す。次の表 4.1 に取得したバス運行データの一部を示す。表 4.1 における、「Date」が日付、「BusNumber」がバスの運行番号、「PassPole」が出発したバス停の識別文字列、「PassTime」が「PassPole」のバス停を出発した時刻である。

表 5.2 取得したバス運行データ(一部)

Date	BusNumber	PassPole	PassTime
2022-11-30	C220	SodaiRiko.1051.2	10:23:49
2022-11-30	E394	JingumaeRokuchome.760.1	10:22:40
2022-11-30	E397	ShinjukuYoncho.714.8	10:21:58

### 5.2.2 渋滞ラベルの取得方法

渋滞ラベルの正解データは、Google Maps Platform の Directions API にリクエストを送ることで取得した。データの収集期間は 11 月 30 日から 12 月 13 日の二週間と 12 月 17 日から 12 月 31 日の二週間、頻度は 5 分おきで取得した。Directions API では、地図上の出発地点と到着地点を緯度経度により指定すると、2 点間の道のり(m)と、リアルタイムの交通状況を踏まえた所要時間(s)を含むデータを返す。出発地点と到着地点に、それぞれバス停  $b_i$  とバス停  $b_{i+1}$  の緯度経度データを入力することによって、バス停間の道のりとリアルタイムの所要時間を取得することができる。道のりを所要時間で割ることによって、その区間での交通速度が分かる。交通速度が渋滞の定義である、時速 10km 未満であれば、渋滞と判断し、そうでなければ、非渋滞と判断する。取得した渋滞ラベルデータの一部を表 4.2 に示す。表 4.2 における「Date」が日付、「Time」が時間、「From」が区間の始点のバス停、「To」が区間の終点のバス停、「Duration」が区間を一般車が通過するのにかかる時間、「Distance」が区間の道のりである。

表 5.3 取得した渋滞ラベルデータ(一部)

Date	Time	From	To	Duration[s]	Distance[m]
2022/1/1/30	9:0:0	池袋駅東口	南池袋一丁目	158	615
2022/1/1/30	9:0:0	北参道	新宿四丁目	190	1,127
2022/1/1/30	9:0:0	新宿四丁目	日清食品前	159	787

## 5.3 実験結果

### 5.3.1 定刻より早く運行している場合

詳細と結果は予稿集論文までに記載する

### 5.3.2 定刻より遅く運行している場合

詳細と結果は予稿集論文までに記載する。

## 6. まとめ

本研究についてまとめる。詳細は予稿集論文までに記載する。

## 謝 辞

本研究は、令和 4 年度「東京都と大学との共同事業」として東京都政策企画局からの交付を受けています。

## 参 考 文 献

- [1] C. Samal, F. Sun and A. Dubey, "SpeedPro: A Predictive Multi-Model Approach for Urban Traffic Speed Estimation," 2017 IEEE International

Conference on Smart Computing (SMARTCOMP), 2017, pp. 1-6, doi: 10.1109/SMARTCOMP.2017.7947048.

- [2] C. Kyaw, N. N. Oo and W. Zaw, "Estimating Travel Speed of Yangon Road Network Using GPS Data and Machine Learning Techniques," 2018 15th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications, and Information Technology (ECTI-CON), 2018, pp. 102-105, doi: 10.1109/ECTICon.2018.8619908.
- [3] T. Kyaw, N. N. Oo and W. Zaw, "Estimating Travel Speed of Yangon Road Network Using GPS Data and Machine Learning Techniques," 2018 15th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications, and Information Technology (ECTI-CON), 2018, pp. 102-105, doi: 10.1109/ECTICon.2018.8619908.
- [4] Y. Gu, Y. Wang, and S. Dong, "Public Traffic Congestion Estimation Using an Artificial Neural Network," ISPRS International Journal of Geo - Information, vol. 9, no. 3, p. 152, 2020
- [5] Y. Xu, Y. Wu, J. Xu, and L. Sun, "Efficient Detection Scheme for Urban Traffic Congestion Using Buses," 2012 26th International Conference on Advanced Information Networking and Applications Workshops, 2012, pp. 287-293, doi: 10.1109/WAINA.2012.62.
- [6] C. Wang and H. Tsai, "Detecting urban traffic congestion with single vehicle," 2013 International Conference on Connected Vehicles and Expo (ICCVE), 2013, pp. 233-240, doi: 10.1109/ICCVE.2013.6799799.
- [7] 青柳宏紀, 岡田一洸, 山名早人. 都バスのリアルタイム運行データを用いた渋滞検知. DEIM2022 第 14 回データ工学と情報マネジメントに関するフォーラム. 2022