

路線バスから得られたセンサデータを利用した 運行状態分類モデルの評価

米澤 拓也¹ 新井 イスマイル² 藤川 和利²

概要：旅客運送業において運行管理者が安全管理、運行管理の観点から、運行中、回送中といった車両の状態を把握することは重要である。現在、運行管理者が車両の運行状態を把握する際には、運転者が手動で状態をリアルタイムに伝え、また日報として自動記録しており、この操作は運転者にとって大きな負担となっている。本研究では、上記の課題を解決し、運行管理の効率化を実現するために、路線バスから得られるセンサデータを利用した運行状態の自動推定を目的とする。1日分の運行データセットを訓練データとし、RandomForestを利用した分類器を構築した。分類器と同一路線データセットでの正答率は、イレギュラーな運行でなければ0.97以上と高い精度を示し、提案手法の有効性を示した。

Evaluation of an operation state classification model using sensor data obtained from buses

TAKUYA YONEZAWA¹ ISMAIL ARAI² KAZUTOSHI FUJIKAWA²

1. はじめに

2016年1月15日、長野県北佐久郡軽井沢町の国道18号碓氷バイパス付近で、定員45人の観光バスがガードレールをなぎ倒し、道路脇に転落するという事故が発生した。本事故において、乗員2人・乗客39人中15人が死亡、生存者も全員が重軽傷を負い、バス事故としては過去30年の中で、最多の死者が出る事故となった。本事故の原因として、下記の2つが挙げられる。

- (1) 旅客運送業の人手不足による労働環境の悪化
- (2) 運行管理者の管理不足

(1) について、日本バス協会は、満60歳以上の高齢運転者を雇用している事業者は648社で、13908人が雇用されている[1]。また、満60歳以上の運転手が占める割合は全体の17.4%と、運転手の6人に1人以上であると報告している。[1]。国土交通省はバス事業者数について、2000年の規制緩和を受け、事業者数は2000年度の3308社か

ら、2014年度には6648社の約2倍に増加したと報告している[2]。事業者数増加による人手不足の影響から運転手の高齢化が進み、さらに、1人あたりの総走行距離の増加といった労働環境の悪化が本事故に要因になったと指摘されている。

(2) について、国土交通省は、バスの運行会社に対し特別監査を行った結果、連続運転時間など、道路運送法の基準に違反する記録があったと報告した。この理由として、バス事業者数の増加による収益の低下や、ノウハウの不足が挙げられる。市場競争の激化により、多くのバス事業者の収益は悪化し、運行管理者の負担が増加している。規制緩和後に新規参入した運送事業経験のない事業者や、小規模事業者には、運行管理のノウハウがないもの、法令を承知していない事業者が多いことが運行管理不足の原因になっていると考えられる[3]。また、本事故を受けて、運輸規則の改正が実施され、運行管理者の態勢・責任は大幅に強化された[4]。このような背景から、運行管理者の負担は今後も増加していくと考えられる。

上記の背景より、旅客運送業において、運行管理者が安全管理、運行管理の観点から、運行中、回送中といった車両の状態を正確かつリアルタイムに把握することは重要で

¹ 奈良先端科学技術大学院大学情報科学研究科
Graduate School of Information Science, Nara Institute of Science and Technology

² 奈良先端科学技術大学院大学総合情報基盤センター
Information Initiative Center, Nara Institute of Science and Technology

ある。また、車両の運行状態をリアルタイムに把握することで、運行管理者側からの車両のマネジメントが可能となり、運行管理を効率化することが可能になると考えられる。

現在、運行管理者が車両の運行状態を把握する際には、運転者が手動で手元端末を操作し、状態をリアルタイムに伝え、運行状態の記録を行っている。しかし、運行状態が変化するたびに端末を操作するという煩雑さや、端末の誤操作による誤った運行状態の記録が大きな課題となっている。上記の課題を解決し、正確な運行状態の把握を実現するために、路線バスから得られるセンサデータを利用した運行状態推定を本研究の目的とする。

本研究では、運行中や回送中といった11個の状態を推定するために、ルールを必要としない機械学習を利用した車両の状態推定を行った。同一路線に対する分類性能評価では、各路線ごとに、Random Forestを利用した分類モデルを作成し、分類を行った。本研究で利用するデータセットは、2つの路線をそれぞれ3日間走行した路線バスのログデータであり、1日目の走行データをモデル生成に利用し、残りの2日間の走行データをテストデータとして用いた。同一路線データセットでの正答率は、イレギュラーな運行でなければ0.9以上となっており、高い分類精度を示した。

2. 関連技術および課題

2.1 関連技術

近年、車両の運行状態を管理するにあたり、タコグラフと呼ばれる運行記録計が広く利用されている。

2015年に国土交通省は「貨物自動車運送事業輸送安全規則」を改正し、運行記録計による記録および当該記録の保存を義務付ける対象を拡大した[4]。従来では車両総重量8トン以上の事業用トラックが記録保存の対象であったが、上記の規則改正により、車両総重量7トン以上の事業用トラックが対象となった。このような背景から、運行記録計の装着義務は今後も拡大していくと考えられる。

運行記録計はアナログタコグラフとデジタルタコグラフの2種類に大別され、近年はタコグラフの各要素を数値化し、電子的に記録するデジタルタコグラフが多く採用されている。

従来のアナログタコグラフでは記録紙を利用していたが、デジタルタコグラフにおいては記録紙に代わり、メモリーカードといった記録媒体やクラウドサービスが利用されている。デジタルタコグラフでは、アナログタコグラフで記録していた時間、車速、エンジン回転数といった情報に加え、GPSによる位置情報や総走行距離といった情報を記録することが可能になった。デジタルタコグラフを利用して情報を記録することにより、運行状態の解析作業の高速化や正確性の向上、法定速度や休憩時間を遵守しているかを容易に確認し、業務効率の改善も可能となる。

多くの貨物自動車運送事業や旅客運送業において、運行状態ログの利便性の観点から、デジタルタコグラフを利用した運行管理が行われている。

2.2 関連技術の課題

デジタルタコグラフを利用した運行管理[5][6]では、複数の課題が存在するため、リアルタイムに正確な車両の運行状態を把握することは困難である。デジタルタコグラフを利用した運行管理における課題を下記に示す。

- リアルタイムに運行状態を把握できない
運行管理者がリアルタイムで車両の状態を把握することは重要である。デジタルタコグラフを利用した運行管理では、運行記録をメモリーカードといった記録媒体やクラウド上に保存し、データの読み出しを行っている。メモリーカード型のデジタルタコグラフは、運行を終了してからメモリーカードの運行記録を専用端末で読み出す必要があり、リアルタイム性に欠ける。クラウド型のデジタルタコグラフでは、現在地情報、車速、エンジン回転数をリアルタイムに表示することは可能であるが、運行状態をリアルタイムで把握できる製品は下記に示す通り、運転者の端末操作に頼っている。
- 運転者の端末操作の負担
デジタルタコグラフを利用した運行管理では、車両の運行状態管理を専用端末を利用して行っている。しかし、運行状態が変化する度に端末を操作し、運行情報を更新する必要があるため、ドライバーにとって端末の操作による運行状態の記録は大きな負担となっている。
- 誤った運行状態の記録
デジタルタコグラフを利用した車両の運行管理では、ドライバーが端末を操作して運行状態の登録作業を行っている。一般的なデジタルタコグラフではメニュー画面に運行状態の一覧が表示され、ドライバーが一覧の中から現在の運行状態を選択し、登録を行っている。しかし、端末の操作ミスや運行状態の変更忘れといった不注意により、誤った運行状態が記録されてしまう場合がある。国土交通省は旅客自動車運送事業運転規則において、「一般乗合旅客自動車運送事業者および特定旅客自動車運送事業者は、事業用自動車の運転者が乗務した時は、運行状態を運転者ごとに記録させ、その記録を1年間保存しなければならない」と定めている。上記の規則や、運行状態のログの活用の観点から、誤った運行状態の記録は大きな課題である。
上記規則では、休憩を行った地点及び日時、乗務開始時及び終了時における総走行距離など、様々な事項の記録保存を義務付けている。今後、より詳細な運行記録が要求される可能性や、自主的に取得したデータを利用した運行改

善の実現という観点から、本研究では、表 1 に示す、みなと観光バス株式会社が実業務で記録している運行状態の自動推定を行うものとする。

3. 提案手法

本研究では、リアルタイム性、ドライバーの負担削減、正確性の向上を実現するため、兵庫県神戸市内を走行する路線バスから得られたセンサデータのみを利用して路線バスの状態の自動推定を行う。

3.1 IoT の普及

近年、センサやアクチュエータといった小型のデバイスをインターネットでつなぐ Internet of Things (IoT) が注目を集めている。IoT 技術の発展に伴い、実環境に分散、配置された多数の IoT デバイスから実環境のあらゆる状況の推定を補助するセンサデータが取得可能となっており、離れたところにあるモノの状態をデータを通じてリアルタイムに知ることが可能となりつつある。また、センサの小型化、高精度化、低価格化も進んでおり、広範囲に適用可能となっている。

既存のネットワーク対応デジタルタコグラフは、GPS 情報や車載カメラの映像をクラウド上に転送しているが、本研究においては、車両に取り付けられたセンサ群から得られたセンサデータのみをサーバに転送する。このセンサデータを利用してバスの状態推定を行うものとする。

3.2 推定するバスの状態

本研究において、推定するバスの状態を表 1 に示す。運行状態は、バス停において停車している運行状態、バス停付近で発着している運行状態、走行している運行状態が存在する。これらの詳細な運行状態は、配車状態や、回送状態と競合してしまう可能性が存在するため、運行状態をさらに 3 状態に細分化した。表 1 中の「運行」状態を詳細化したものを表 2 に示す。

下記に、一般的なバス運行における始業から終業までの流れを示す。

- (1) 出発前に点検作業を行う (点検)
- (2) 出発前に洗車作業を行う (洗車)
- (3) 点検作業後に初めの停留所に向けて移動 (配車)
- (4) 初めの停留所においてバスが待機 (待機)
- (5) 運行状態に移行 (運行)
- (6) 走行している (運行-走行)
- (7) バス停に停車 (運行-バス停停車)
- (8) バス停付近で発着中 (運行-バス停発着)
- (9) 一定時間の運行後に休憩を行う (休憩)
- (10) 再度バスが運行状態に移行 (運行)
- (11) 運行を終了 (回送)
- (12) ガソリンスタンドにおいて給油 (給油)

表 1 本研究で推定するバスの状態

状態名	定義
点検	運行開始前の点検作業中である
洗車	洗車作業中である
配車	点検後に停留所に向けて移動中である
待機	初めの停留所においてエンジンを OFF
運行	バスが運行中である
休憩	4 時間未満の休み
回送	運行を終了し、事業所へ移動している
給油	ガソリンスタンド内において停車中である
帰庫	事業所内において停車中である

表 2 「運行」状態の詳細

状態名	詳細	定義
運行	バス停停車	バス停において停車中である
	バス停発着中	バス停付近で発着中である
	走行中	バスが走行中である

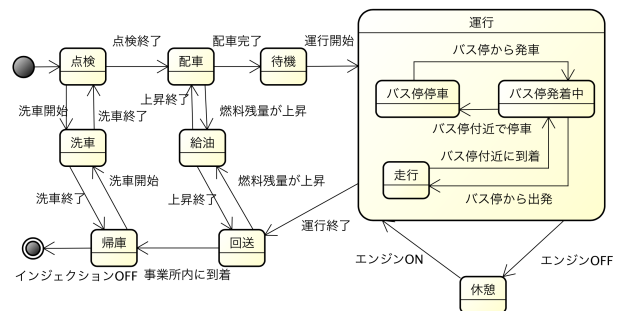


図 1 乗合バスの業務の流れ

(13) 事業所内に帰庫 (帰庫)

また、洗車、給油なども運転者が運行時間外に業務として行うため、表 1 内で定義を行っている。バス業務の流れを図 1 に示す。

4. 利用する特徴量について

4.1 得られるセンサデータについて

路線バスから得られるセンサデータの種類を表 3 に示す。エンジン回転数、車速、エンジン回転のパルス値といった自動車の動作に関するデータは自動車に備え付けた機器から観測される電気信号を解析することで取得している。また、緯度、経度、標高データは後付した車載 GPS モジュールから取得している。本研究で解析に利用するセンサデータは兵庫県神戸市内灘区を走行する、みなと観光バス株式会社の路線バスから 1 秒おきに取得したものである。図 2、図 3 にデータ取得対象車両の走行路線を示す。

4.2 新たに定義する特徴量について

本研究では、センサデータのみを特徴量とする運行状態分類と比較し、分類精度向上を実現するために下記に示す 2 つの特徴量を新たに追加し、分類を行う。

表 3 路線バスから得られるセンサーデータ

データ名
緯度
経度
高度
車速
エンジン回転数
総走行距離 (ODO)
パルス値

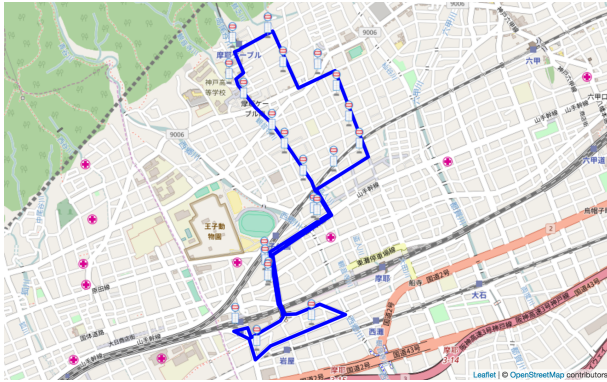


図 2 データ取得対象車両の走行路線 (路線 S)

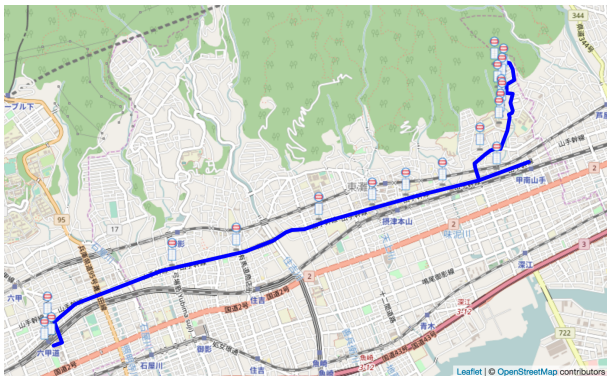


図 3 データ取得対象車両の走行路線 (路線 D)

- 車速の変化量
車速の変化量については、得られた車速データから、1秒前の車速と現在の車速の変化量を計算し、定義するものとする。
- 最寄りのバス停までのユークリッド距離
最寄りのバス停までのユークリッド距離については、それぞれの路線に属するバス停の座標データと車両の座標データからユークリッド距離を計算し、定義している。なお、各路線におけるバス停の数は、S路線は22個、D路線は19個である。

本研究では上記の2つの特徴量を加え、表4に示す計9つの特徴量を用いて状態分類モデルの構築を行った。なお、これらの特徴量は得られるセンサーデータを元に1秒毎に定義している。

表 4 状態分類に用いる特徴量

特徴量
緯度
経度
高度
車速
エンジン回転数
総走行距離 (ODO)
パルス値
車速の変化量
最寄りのバス停までの距離

5. 状態推定手法

路線バスの運行は、図1に示すような状態遷移によって実現されている。この観点から、状態推定手法としてHMM(Hidden Markov Model)を利用したモデル化が有用であると考えられる。HMMを本研究に適用するためには、状態遷移を明確に定義し、モデル化する必要がある。しかし、これらの作業は一定の運行ノウハウを必要とし、モデルを構築するためのデータ加工が必要なため、規制緩和後に新規参入した運送事業経験のない事業者にとって大きな負担となると考えられる。上記より、本研究では、ルールベース手法と機械学習を利用した状態推定を検討する。

ルールベース手法では、分類を行うために複数のIF-THEN形式のルールを利用する。しかし、ルールベース手法は、ルールに適合しない状態を処理できないという欠点が挙げられる。また、状態推定を行うためにセンサーデータのパターンを全てルール化することの困難さや、車両状態の追加や削除に柔軟に対応できないという課題が存在する。このような背景から、ルールを必要としない機械学習に着目し、機械学習を利用して車両の状態推定を行う。本研究では表1と表2に示した11個の車両状態を推定するために、教師あり学習を利用した多クラス分類を行う。

5.1 多クラス分類手法の比較

本節では、教師あり学習を利用した多クラス分類手法の比較を行う。また、表5に各多クラス分類手法の比較を示す。

- 決定木 (CART)
決定木はデータの特徴量を用いた簡単なルールで分岐を作り、特徴空間を分割することを通じて判別や回帰を行うモデルのことである。CARTモデルの流れを下記に示す。

(1) 木の構築

何らかの基準を満たすまで、予め定義しておいたコストに基づいて特徴空間を2分割する作業を繰り返す。

(2) 剪定

構築された木の深さが深いほど、複雑なデータを扱うことが可能になるが、過学習の可能性が存在する。剪定では過学習を防ぐために予め定めておいたパラメータによってモデルの複雑度を制御する。

決定木の特徴として、高次元の判別が容易に視覚的に確認できるという点が挙げられる。また、決定木の問題点としては、判別結果の分散が大きく、データが少し変わっただけで木の構造や判別ルールが大きく変わってしまう点が挙げられる。

● Random Forest

Random Forest は、複数の決定木 (decision tree) を利用した多クラス分類手法である。決定木をアンサンブル学習における弱学習器の1つとして取り扱い、個々の学習器としての精度は高くないが、複数の決定木の結果を統合し、評価することによって高い予測性能を得ることが可能となる [7]。また、Random Forest の特徴を下記に示す。

- 大きいデータセットにも効率的に動作する
- 変数の数が大きい場合でも安定して動作する。
- 識別に用いる変数の重要度を算出することが可能
- 欠損値を含むデータについても学習・識別が可能

● SVM (Support Vector Machine)

SVM は、教師あり学習を用いる 2 クラス分類器の 1 つである。SVM では、カーネル関数を用いて与えられたデータを高次元へと写像し、写像した空間において、2 クラス間のマージンが最大となる識別境界を求める。これらから、SVM は高い汎化能力をもち、大きな次元を持つ学習データを利用して過学習を起こすことが少ないと言われている。通常の 2 クラス SVM を複数組み合わせることで、多クラス分類器を実現する手法が提案されている。

- one-versus-one 方式
one-versus-one は 1 対 1 方式と呼ばれている。K 個のクラスの組み合わせについて 2 クラス SVM を学習し、その結果得られた $K(K-1)/2$ 個の分類器を適用して、最も多くの分類器が正例として投票したクラスを分類結果とする方法である。しかし、この方式では分類クラスが一意に定まらない可能性や、K の値が大きい場合、予測にかかる計算時間も大きいという欠点がある [8]。
- one-versus-rest 方式
one-versus-rest は 1 対他方式と呼ばれている。この方式は K このクラスがあるときにあるクラス C_k に属するデータを正例、それ以外のデータを負例として K 個の別々の SVM $y_k(x)$ を学習する方法である。この方式では、個々の SVM による予測が矛盾し、1

表 5 各分類手法の特徴

	CART	RF	vs-one	vs-rest
計算時間	×	○	×	○
クラスの一意性	○	○	×	×
ノイズへの耐性	×	○	○	○
欠損値への耐性	×	○	○	○
判別の安定性	×	○	○	○

表 6 S 路線の分類に用いるデータセット

	時間	レコード数	欠損数
1 日目	06:28:00~21:33:00	54280	20
2 日目	05:55:00~21:35:00	56366	34
3 日目	05:50:00~21:27:00	56187	33

表 7 D 路線の分類に用いるデータセット

	時間	レコード数	欠損数
1 日目	05:26:00~21:59:00	58486	1094
2 日目	05:34:00~22:19:00	58253	2047
3 日目	05:34:00~22:04:00	58092	1308

つの入力に同時に複数のクラスが割り当てられる可能性がある [8]。

本研究では、表 5 より、Random Forest を利用して分類器を構築し、バスの状態推定を行った。評価実験に用いたデータセットは、異なる 2 つの S 路線と D 路線を走行した 2 車両の 3 日分の、合計 6 運行分のデータである。本研究では、構築したモデルに対し、同一路線に対する分類性能、他路線に対する分類性能の評価を行った。

6. 同一路線に対する分類性能評価

同一路線に対する分類性能評価では、各路線ごとに、Random Forest を利用してモデルを作成し、テストデータを適用し、分類を行った。

センサデータの送信には LTE 回線を利用しており、トンネル内や高架下などといった環境下において通信状態が不安定になる。そのような環境下での TCP を利用したセンサデータ送信は、再送処理が繰り返されてしまい、帯域の圧迫に繋がる可能性がある。この観点から、センサデータは路線バスから UDP パケットを利用して学内のデータベースに送信している。UDP は高速性を重視しており、パケット再送処理を行わないため、通信環境が不安定な環境下において一部データに欠損が生じている。

表 6、表 7 に、S 路線および D 路線の分類において用いるデータセットの概要を示す。

6.1 S 路線モデル-S 路線データ

S 路線の分類性能評価では、S 路線 1 日目のデータセットを訓練データとして分類モデル構築に利用した。また、2 日目、3 日目の S 路線データセットをテストデータとし、

表 8 S 路線モデル-S 路線データの分類結果

	S 路線 2 日目		S 路線 3 日目	
	総データ数	正答率	総データ数	正答率
OFF	52	1	123	0.268
点検	998	0	249	0
洗車	485	0	659	0.196
配車	2474	0.987	3196	0.977
待機	1535	0.001	1345	0
運行-走行	37249	0.988	38740	0.989
運行-発着	3939	0.908	3884	0.920
運行-停車	5037	0.892	3992	0.948
休憩	3409	0.999	3034	0.998
回送	872	0.487	673	0.539
給油	67	1	99	1
帰庫	249	0.879	194	0.639
全体	56366	0.977	56187	0.963

表 9 D 路線モデル-D 路線データの分類結果

	D 路線 2 日目		D 路線 3 日目	
	総データ数	正答率	総データ数	正答率
OFF	107	1	376	1
点検	168	1	230	0.535
洗車	1996	0	1788	0
配車	1475	0.985	3481	0.416
待機	123	0	1265	0.070
運行-走行	27023	0.989	25158	0.913
運行-発着	3627	0.934	3209	0.860
運行-停車	1847	0.949	2267	0.957
休憩	19474	1	16619	0.931
回送	1329	0.971	2987	0.669
給油	104	1	117	1
帰庫	980	1	595	0.375
全体	58253	0.974	58092	0.849

モデルの分類性能評価に用いる。

表 8 に S 路線モデルによる S 路線データセットの分類結果を示す。

6.2 D 路線モデル-D 路線データ

D 路線の分類性能評価では、D 路線 1 日目のデータセットを訓練データとして分類モデル構築に利用した。また、2 日目、3 日目の D 路線データセットをテストデータとし、モデルの分類性能評価に用いる。

表 9 に、D 路線モデルによる D 路線データセットの分類結果を示す。

6.3 考察

表 8 における S 路線 2 日目、S 路線 3 日目、表 9 における D 路線 2 日目から、正答率が 0.97 前後と非常に高い分類精度を示していることが分かる。

しかし、表 9 における D 路線 3 日目に着目すると、正答

率が 0.849 と、他の分類結果より低い精度となっていることが分かる。この理由として、D 路線 3 日目の走行データがイレギュラーな運行によるものであったことが挙げられる。D 路線 3 日目の運行の流れは、事業所を出発、配車、一定時間運行、事業所に帰庫、配車、一定時間運行、事業所に帰庫となっており、平常時の運行フローと異なるものとなっていた。表 9 の 3 日目の配車状態の正答率は 0.416 であるのに対し、同表の 2 日目、表 8 の 2 日目、3 日目の配車状態の正答率はおよそ 0.98 前後となっており、大きな差があることが分かる。この理由として、平常運行における配車状態の判別の大きな要因となる特徴量が総走行距離であるという点が挙げられる。平常運行時における配車状態への遷移は点検状態終了後に 1 度遷移するだけであり、総走行距離からの判別が実現できる。しかし、D 路線 3 日目においては配車状態への遷移が 2 度行われているため、配車状態への遷移が 1 度である正常運行データを用いたモデルでは判別ができなかったと考えられる。このようなイレギュラーな運行が行われなければ、同一路線分類モデルは極めて高い分類精度を示し、高精度な運行状態推定を行うことが可能になると考えられる。

表 8、表 9 において、点検、洗車、待機の 3 状態が低い分類精度を示していることが分かる。点検、洗車の 2 状態は、どちらも事業所内でエンジンを稼働させながら行うものとなっており、2 状態の判別は GPS 情報のみに依存している。GPS 情報のみではこの類似した 2 状態の分類は困難であり、明確な分類器を構築できなかったことが原因であると考えられる。待機状態は、テストデータの大半が、休憩状態・運行-走行状態・運行-発着状態のいずれかに誤分類されていた。運行-走行状態・運行-発着状態に誤分類された原因として、運転手がバスのエンジンを切らずに待機状態に遷移していたことが挙げられる。エンジンが稼働していて、かつバス停近辺に停車している場合には、センサーデータのみからでは待機状態と運行状態を区別することができないため、誤分類の要因になったと考えられる。また、エンジンを稼働させずに待機している場合であっても、休憩状態と待機状態の定義が同等になってしまうため、誤分類の要因になったと考えられる。

7. 他路線に対する分類性能評価

他路線に対する分類性能評価では、各路線ごとに、1 日目の走行履歴を Random Forest を利用してモデル化し、他の路線を走行した 2 日分のデータセットを適用し、分類を行った。

表 10 に、本節での評価に用いるデータセットの概要を示す。また、利用するデータセットは表 6、表 7 と同様である。

表 10 他路線に対する分類性能評価に用いるデータセット

評価対象	訓練データ	テストデータ
S 路線モデル	S 路線 1 日目	D 路線 2 日目・3 日目
D 路線モデル	D 路線 1 日目	S 路線 2 日目・3 日目

表 11 S 路線モデル-D 路線データの分類結果

	D 路線 2 日目		D 路線 3 日目	
	総データ数	正答率	総データ数	正答率
OFF	107	0	376	0
点検	168	0	230	0
洗車	1996	0	1788	0
配車	1475	0.718	3481	0.644
待機	123	0	1265	0
運行-走行	27023	0.946	25158	0.875
運行-発着	3627	0.799	3209	0.797
運行-停車	1847	0.866	2267	0.849
休憩	19474	1	16619	0.932
回送	1329	0.393	2987	0.101
給油	104	0	117	0
帰庫	980	0.135	595	0.334
全体	58253	0.880	58092	0.825

表 12 D 路線モデル-S 路線データの分類結果

	S 路線 2 日目		S 路線 3 日目	
	総データ数	正答率	総データ数	正答率
OFF	52	1	123	1
点検	998	1	249	0.996
洗車	485	0	659	0
配車	2474	0.992	3196	0.95
待機	1535	0	1345	0
運行-走行	37249	0.983	38740	0.983
運行-発着	3939	0.601	3884	0.597
運行-停車	5037	0.861	3991	0.832
休憩	3409	0	3034	0
回送	872	0.859	673	0.854
給油	67	0	99	0
帰庫	249	0	194	0
全体	56366	0.844	56187	0.849

7.1 S 路線モデル-D 路線データ

S 路線モデルの評価では、S 路線の 1 日目のデータセットを訓練データとしてモデルを構築し、D 路線の 2 日目と 3 日目のデータセットをテストデータとして評価を行った。表 11 に S 路線モデルによる分類結果を示す。

7.2 D 路線モデル-S 路線データ

D 路線モデルの評価では、D 路線の 1 日目のデータセットを訓練データとしてモデルを構築し、S 路線の 2 日目と 3 日目のデータセットをテストデータとして評価を行った。表 12 に D 路線モデルによる分類結果を示す。

7.3 考察

表 11, 表 12 より、全体の正答率は各日において 0.85 前後を示していることが分かる。

運行-走行状態は 0.875~0.983 と非常に高い正答率を示しており、運行-停車状態も 0.85 前後と、高い正答率を示している。運行-発着状態については、表 11 では 0.8 未満、表 12 では 0.6 前後と、運行-走行状態・運行-停車状態と比較すると低い正答率を示していることが分かる。この理由として、路線ごとにバス停データを利用し、最寄りのバス停までのユークリッド距離を特徴量として用いたため、他の路線の運行データに適応できなかったことが考えられる。また、給油状態についても、路線ごとに利用するガソリンスタンドが異なるため、他路線のデータセットに対してモデルを適用すると、正答率が下がってしまうという結果となった。

他路線のデータセットに対する汎用性という観点では、運行-走行状態・運行-停車状態といった、センサデータを利用して明確な分類が可能な状態以外は総じて低い正答率を示した。以上から、他路線のデータセットに対して汎用性を持たせることは、本研究の提案手法では困難であると考えられる。

また、本稿における汎用性の評価は、1 路線の走行データを利用した路線ごとのモデルについて評価を行った。モデルの汎用性向上の観点から、複数路線の走行データを利用した複数路線型分類モデルの構築を行い、評価することが今後の課題として考えられる。

8. おわりに

本稿では、近年の旅客運送業における運行管理の背景から、車載センサから得られたセンサデータのみを利用して車両の状態を推定する手法について提案した。同一路線に対する分類では、正常な運行であれば、正答率は 0.97 前後と高い分類率を示した。また、イレギュラーな運行であっても、正答率は 0.85 と高い分類精度を示した。他路線に対する汎用性の評価については、全体の正答率は 0.85 程度であったが、いくつかの状態の正答率が 0 であり、正確な状態分類は実現できなかった。以上から、実業務中で状態分類モデルを運用するためには、各路線ごとにセンサデータを取得し、路線分類モデルを構築する必要があると考えられる。

精度向上の観点から、信頼性の高いセンサデータのみを利用した分類モデル構築、複数路線の走行データを利用した分類モデルの評価、新たな特徴量の定義などが今後の課題として挙げられる。

また、本稿では路線バスの状態推定を提案したが、車両状態分類モデルの汎用性の向上を実現することができれば、長距離トラックといった貨物運送車や、タクシーや貸切バスなどの旅客運送車の運行状態も推定可能になると考

えられる。

謝辞 本研究の一部は JSPS 科研費 16K00147 の助成によるものである。

参考文献

- [1] 公益社団法人 日本バス協会:2014年版 日本のバス事業,入手先 http://www.bus.or.jp/about/pdf/h26_busjigyo.pdf
- [2] 国土交通省:自動車関係統計データ,入手先 <http://www.mlit.go.jp/common/000117167.pdf>
- [3] 公益社団法人 日本バス協会:バス事業の現状と取り組みについて入手先 <https://www.mlit.go.jp/common/001127098.pdf>
- [4] 国土交通省:貨物自動車運送事業輸送安全規則,入手先 <http://law.e-gov.go.jp/htmldata/H02/H02F03901000022.html>
- [5] 株式会社デンソーセールス:DN-magic PREMIUM 入手先 <https://www.denso-sales.co.jp/d-navi/product/dn-magic.premium/>
- [6] 株式会社トランストロン:ネットワーク型車載情報システム DTS-D1D 入手先 <http://www.transtron.com/products/dts-d1d.html>
- [7] Liaw, Andy, and Matthew Wiener: Classification and regression by randomForest. R news 2.3 (2002): 18-22.
- [8] Hsu, Chih-Wei, and Chih-Jen Lin: A comparison of methods for multiclass support vector machines. IEEE transactions on Neural Networks 13.2 (2002): 415-425.
- [9] Freund Yoav, Schapire Robert, 安倍直樹: ブースティング入門, 人工知能学会誌 Vol.14,771-780,1999