

NAIST-IS-MT1651127

修士論文

車両から得られたセンサデータを利用した 路線バス運行状態分類手法の提案

米澤 拓也

2018 年 3 月 15 日

奈良先端科学技術大学院大学
情報科学研究科

本論文は奈良先端科学技術大学院大学情報科学研究科に
修士（工学）授与の要件として提出した修士論文である。

米澤 拓也

審査委員：

藤川 和利 教授 （主指導教員）

安本 慶一 教授 （副指導教員）

新井 イスマイル 准教授 （副指導教員）

車両から得られたセンサデータを利用した 路線バス運行状態分類手法の提案*

米澤 拓也

内容梗概

旅客運送業において運行管理者が、安全管理，効率的な運行管理の観点から車両の運行状態を把握することは重要である．現在，運行管理者が運行中や休憩中といった運行状態を把握する際には，運転者がデジタルタコグラフとよばれる記録計を手動で操作することにより，運行状態を記録している．しかし，デジタルタコグラフを利用した運行管理では以下の課題が存在する．(1) リアルタイムに運行状態を把握することができない，(2) 運転者の端末操作の負担が発生，(3) 誤った運行状態の記録，(4) 運行データの二次利用が困難といった課題が存在する．本研究ではこれらの課題を解決し，運行の効率化を実現するために，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分類を実現することを目的とする．本研究では，分類の高速性や汎用性の観点から，Random Forest を利用した運行状態分類モデルを構築した．神戸市内を走行する路線バスの実運行データを利用し，運転手がデジタルタコグラフを操作して記録した運行状態と，構築した運行状態分類モデルが出力した運行状態を比較し，提案手法の検証を行った複数の運行状態分類モデルを構築し，評価を行った結果，同一路線に対する分類ではモデルの正答率は 0.9 前後となった．また，運転手が手動で運行状態を入力した結果と比較すると，運行状態分類モデルの正答率が手動入力の結果を上回るという結果が得られた．

キーワード

IoT, ITS, 機械学習, センサデータ

* 奈良先端科学技術大学院大学 情報科学研究科 修士論文, NAIST-IS-MT1651127, 2018 年 3 月 15 日.

Proposal of Bus Operation States Classification Method Using Vehicle Sensor Data*

Takuya Yonezawa

Abstract

In bus companies, it is important for an operation manager to grasp operation states of vehicles from a viewpoint of safety management and improving an operation efficiency. Currently, for allowing operation managers to grasp operation states of vehicles, drivers should record operation states by manually operating a recorder called "Digital-tachograph." However, the following issues exist in the operation management using the digital tachograph. (1) it is impossible to grasp operation states in real time, (2) burden of terminal operation by the drivers, (3) erroneous operation states are recorded (4) secondary usage of the operation data is difficult In order to solve these problems and to realize efficient operation, we propose a method for automatic classification of operation states using sensor data obtained from buses. We implemented a classifier using Random Forest with the sensor data. As a result of implementation and evaluating operation states classification models, the correct answer rate of the model was around 0.9 in same route classification. In addition, when compared with the result that drivers manually input operation states, the correct answer rate of the operation state classification models exceeded manual input results.

Keywords:

*Master's Thesis, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT1651127, March 15, 2018.

IoT, ITS, Machine Learning, Sensor data

目次

図目次	vi
表目次	vii
第 1 章 序論	1
第 2 章 旅客運送業の現状	4
第 3 章 既存の運行管理の課題	7
3.1 アナログタコグラフ	7
3.2 デジタルタコグラフ	8
3.3 タコグラフを利用した運行管理の課題	10
第 4 章 運行状態自動分類の提案	12
4.1 本研究で分類するバスの運行状態	13
4.2 得られるセンサデータについて	13
4.3 状態分類に利用する特徴量について	15
4.4 状態推定手法の考察	17
4.5 自動分類モデルのシステム化	21
第 5 章 分類結果	24
5.1 同一路線モデルによる分類 (路線 B)	26
5.2 他路線モデルによる分類	27
5.3 複数路線モデルによる分類	29
5.4 運転手の手動入力との比較	30
第 6 章 考察および今後の展望	33
6.1 考察	33
6.2 今後の展望	35
第 7 章 結論	38

謝辞	39
参考文献	40
発表リスト	42

図目次

1.1	本研究におけるゴール	2
2.1	バス事業における平均年齢および高齢運転手割合の遷移	4
2.2	バス事業者の事業者数の遷移	5
3.1	アナログタコグラフによる運行管理の例	8
3.2	デジタルタコグラフによる運行管理の例	9
4.1	本研究で実現するリアルタイム運行状態分類の例	12
4.2	乗合バスの業務の流れ	14
4.3	リアルタイム分類実現のためのアーキテクチャ	22
5.1	データ取得対象路線 (路線 A)	24
5.2	データ取得対象路線 (路線 B)	25
5.3	正答ラベル作成に用いた車載カメラ映像の例	27

表目次

4.1	本研究で推定するバスの状態	14
4.2	「運行」状態の詳細	14
4.3	路線バスから得られるセンサデータ	15
4.4	状態分類に用いる特徴量	18
4.5	各分類手法の特徴	21
5.1	路線 A から得られたデータセットの概要	25
5.2	路線 B から得られたデータセットの概要	25
5.3	各車両における運行状態の正答データ総数	26
5.4	全評価結果の概要	27
5.5	同一路線モデルによる分類結果 (路線 B)	28
5.6	他路線モデルによる分類結果 (路線 A モデル)	29
5.7	他路線モデルによる分類結果 (路線 B モデル)	30
5.8	複数路線モデルによる分類結果 (路線 A+B モデル)	31
5.9	手動入力と状態分類モデルの正答率の比較	31
6.1	各分類モデルにおける特徴量の重要度	33

第 1 章 序論

近年，IoT (Internet of Things) の概念の普及により，様々なモノがインターネットに接続され，モノの状況の把握，モニタリングなどが実現されている．また，コネクテッドカーの概念も普及してきており，自動車の IT 化により，車両内部のセンサ群やクラウドを利用した様々な安全運転支援ソリューションや，テレマティクス保険，危険運転検知システムなどが提案されている．

旅客運送業において，規制緩和により事業者数が急増し，業界における競争はさらに加速している．それに伴う新規参入事業者のノウハウ不足や，業績悪化により，許可区運送会社は新技術や新製品に対応できず，必要十分な運行管理が提供できていない．車両の運行計画の作成およびドライバーへの安全運転指導を行う運行管理者にとって，運行最適化やドライバー指導のために，正確な車両の運行状況や運行履歴を把握することは重要である．しかし，既存のタコグラフと呼ばれる記録計を用いた運行管理方法では，ドライバー，運行管理者の両者に対して負担が大きくなっており，IoT 化，コネクテッドカー化による負担削減が実現できていないのが現状である．また，IoT 化，コネクテッドカー化を実現するための製品は高額であり，経営状況が悪い事業者においては，ドライバーおよび運行管理者への負担がさらに増加していく傾向にある．

近年の旅客運送業における課題をステークホルダー毎にまとめると下記のとおりとなる．

- ドライバー
 - 車両内の限られたインタフェース下における端末操作の負担が発生
 - 運転業務に加え，運行終了後に，担当した運行の報告業務
- 運行管理者
 - ドライバーが記録する運行履歴の正当性の確保義務
 - 特殊な記録紙の解析ノウハウが必要
 - 車両の運行状況をリアルタイムに把握することが困難
 - 法令を遵守していないドライバーが存在しても把握が困難

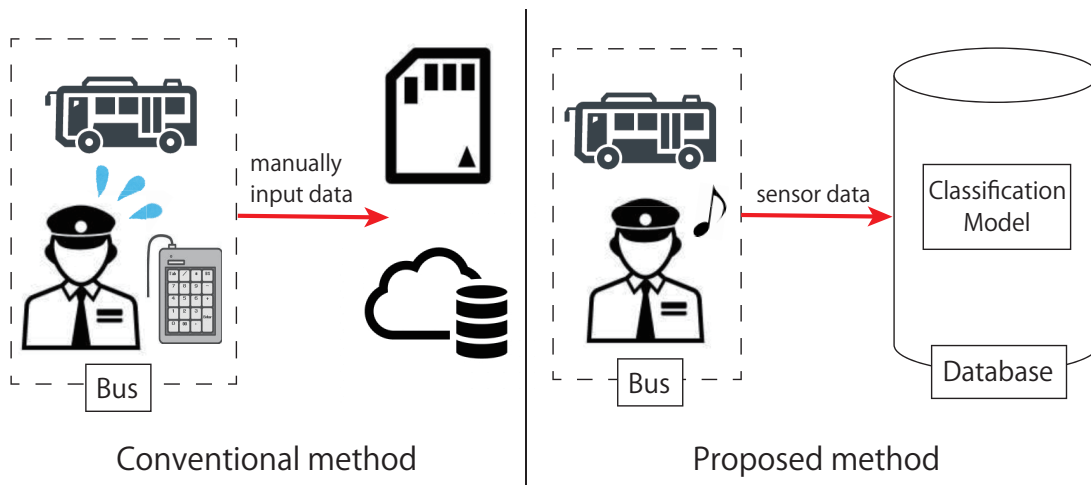


図 1.1 本研究におけるゴール

- 経営者

IoT 化，コネクテッドカー化を実現するためには，新機器を導入する必要があり，高コスト

我々は，各ステークホルダー毎の課題を解決するために，神戸市内のバス会社と共同で安価な車載 IoT モジュールを利用した車両センサデータの取得および解析を行っている．

本研究では，特に，ドライバーと運行管理者の負担削減に重点を置き，車両センサデータを利用した運行状態の自動分類を試みる．本研究のゴールのイメージ図を図 1.1 に示す．具体的には，ドライバーが入力した運行状態記録と，構築した運行状態分類モデルが出力した結果を比較し，ドライバーが手入力した運行状態記録より正確な運行状態分類を目的とする．また，車両センサデータのみを利用することにより，ドライバーの運転中における端末操作の負担を解消することも可能となる．

本稿では，まず近年の旅客運送業における現状および背景を解説する．その後，既存の運行管理方法の紹介と，未解決となっている運行管理の課題について解説する．そして，課題解決を実現するための手法を提案し，検証条件や利用す

る特徴量，提案手法実現のためのシステムアーキテクチャについて解説する．その後，従来の運行管理手法による結果と，提案手法による結果を比較し，評価を行う．最後に提案手法の結果の考察と，今後の展望について述べる．

本稿の構成は以下のとおりである．第 2 章では，近年の旅客運送業における現状について述べ，それらの要因および本研究のモチベーションについて述べる．第 3 章では，旅客運送業において一般的に利用されている運行効率化方法の概要，利点，課題について解説し，本研究の新規性について述べる．第 4 章では，上記で述べた既存の運行管理方法の課題を解決するための手法および実現方法を提案する．第 5 章では，第 4 章で提案した手法を実装および複数の検証項目を設定し，評価実験を行う．第 6 章では，評価実験の結果から，評価結果の考察，提案手法の妥当性および今後の展望について議論する．第 7 章では，本稿のまとめを行う．

第 2 章 旅客運送業の現状

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 14 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 [1]．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 歳以上の高齢運転手を雇用している事業者は 648 社で，13908 人が雇用されていると報告している [2]．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 [2]．図 2.1 にバス事業における平均年齢および高齢運転手割合の遷移を示す．

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

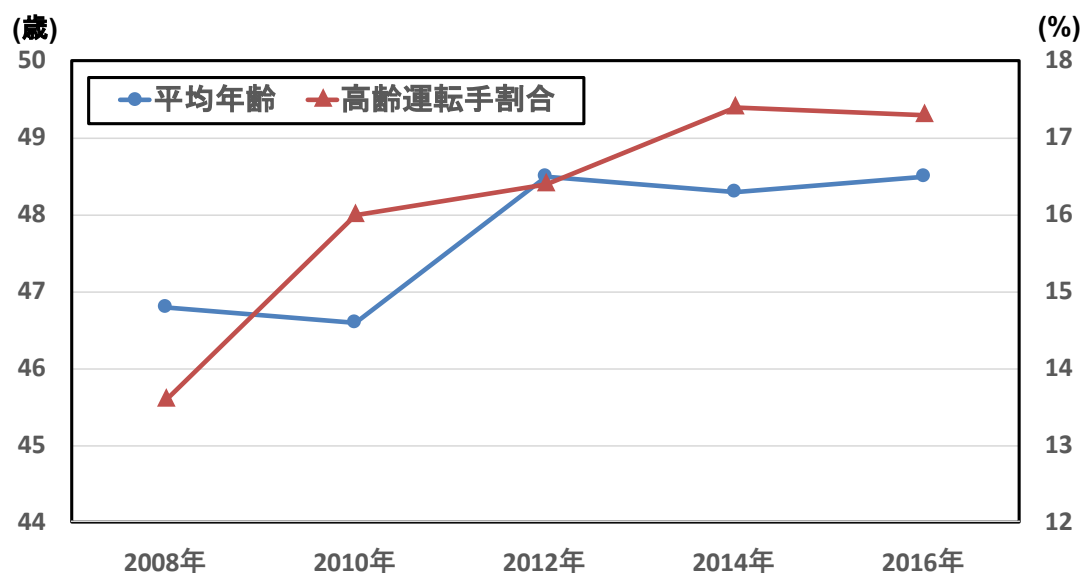


図 2.1 バス事業における平均年齢および高齢運転手割合の遷移

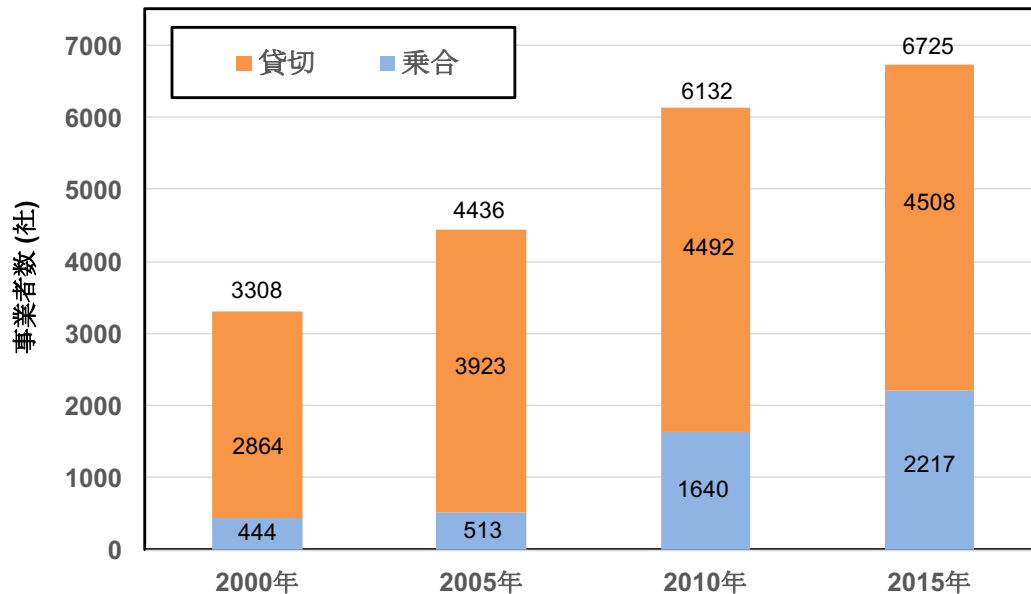


図 2.2 バス事業者の事業者数の遷移

2000 年度の 3308 社から、2015 年度には 6725 社の約 2 倍に増加したと報告している [3].

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 [4]。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 [5]。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、運行中や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。

第 3 章 既存の運行管理の課題

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 [5]。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。

タコグラフは、アナログタコグラフとデジタルタコグラフの 2 種類に大別される。

3.1 アナログタコグラフ

アナログタコグラフでは、速度、走行距離、走行時間が専用の用紙（チャート紙）に時系列の折れ線グラフで記録される。このチャート紙の上に、ドライバーが運行終了後に運行内容を手動で記録することにより、運行の把握や管理を行っている [6]。運行管理者はチャート紙に記録されたデータとドライバーの記載データを読み取ることにより、運行指導や安全運転指導が可能となる。図 3.1 にアナログタコグラフを利用した運行管理の例を示す。

アナログタコグラフは、比較的安価であり、導入のコストが低いという利点があるが、下記の欠点が存在する。

- チャート紙の解析に習熟が必要
- チャート紙に記録されたデータの解析に時間を要する
- 取得可能なデータが限られている
- 運行終了後でなければ、当日の運行状況を把握することが不可能
- ドライバーが当日の運行状況を偽装して報告することが可能
- ドライバー、運行管理者への負担が発生する

近年は従来のアナログタコグラフに記録されているデータを、CAN ネット

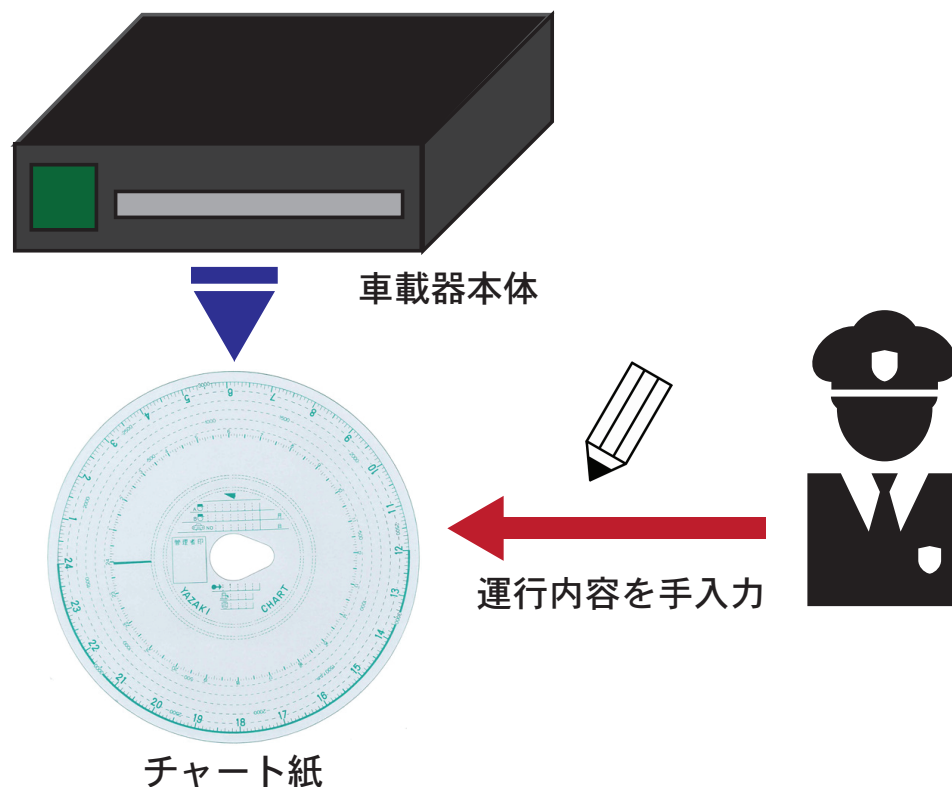


図 3.1 アナログタコグラフによる運行管理の例

ワークからデジタルで取得可能にしたデジタルタコグラフが登場したことにより、アナログタコグラフを利用する企業は減少している。

3.2 デジタルタコグラフ

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 [7]。また、アナログタコグラフでは、ドライバーが運行終了後にチャート紙に洗車、点検、休憩などといった運行状態を記

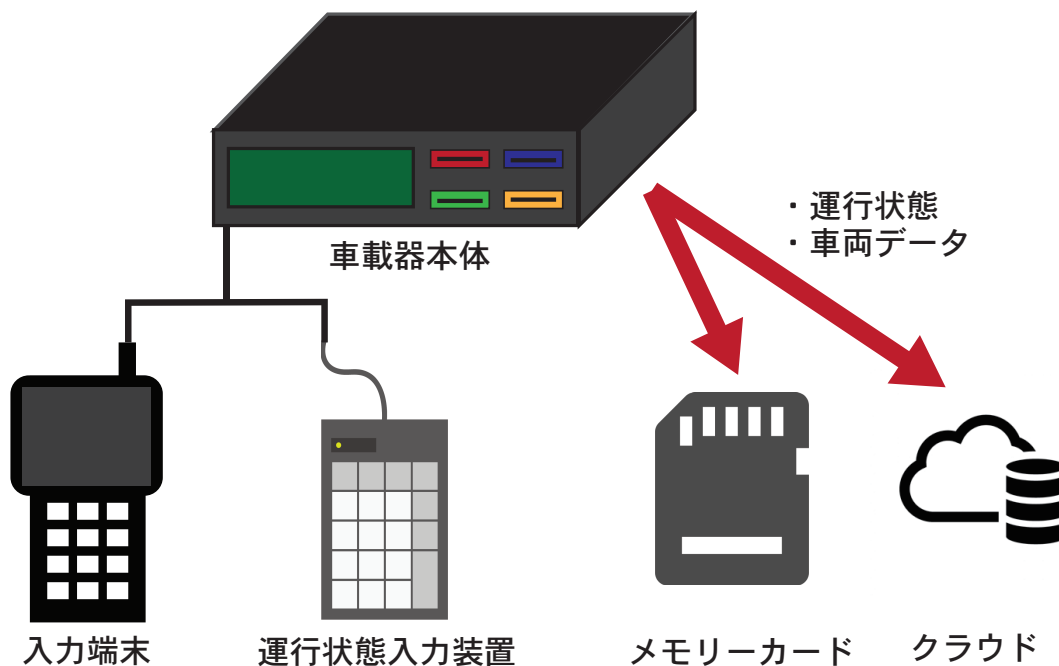


図 3.2 デジタルタコグラフによる運行管理の例

録していた。デジタルタコグラフでは、車載器に付随するテンキーあるいは運行状態入力用端末を利用して、運行状態の記録を随時行っている。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。図 3.2 にデジタルタコグラフを利用した運行管理の例を示す。

アナログタコグラフと比較し、デジタルタコグラフは利点が多いが、下記の欠点が存在する。

- 機器が高価であり、導入のコストが高い
- 取得可能なデータがメーカー依存であり、データの二次利用が困難
- クラウド対応型の製品も存在するが、多くの製品は記録媒体を採用している

- ドライバーが当日の運行状況を偽装して報告することが可能
- ドライバーへの端末操作の負担が発生する

3.3 タコグラフを利用した運行管理の課題

アナログタコグラフを利用した運行管理と、デジタルタコグラフを利用した運行管理 [8][9] では、下記の共通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し、運行データを有効活用することは困難である。タコグラフを利用した運行管理における共通の課題を下記に示す。

リアルタイムに運行状態を把握できない

運行管理者が、車両運行の効率化を実現するために、リアルタイムに車両の状態を把握することは重要である。アナログタコグラフでは、業務を終了しなければ、該当記録を確認することができない。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に把握できる製品は、運転者の端末操作に頼っているのが現状である。

運転者の端末操作の負担

アナログタコグラフを利用した運行管理では、業務終了後のチャート紙への記入、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、運行終了後に該当の運行記録をまとめて記入する必要がある。デジタルタコグラフの場合においては、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。ドライバーにとってチャート紙への記入や、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。アナログタコグラフでは運行終了後に

タコチャート紙へ手動で記録している。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、記入漏れ、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。

データの二次利用が困難

アナログタコグラフではチャート紙に記録するため、該当の運行記録をデジタル化して再利用することが困難である。また、デジタルタコグラフを用いた場合では、データフォーマットや利用可能なデータがメーカー依存であるといった課題が存在する。旅客運送業者にとって、データの二次利用は安全な運行を実現する上では非常に重要である。データの二次利用により、車速データを利用して速度違反や運行遅延が頻発する箇所を特定したり、運転履歴を利用したドライバーへの安全運転指導や運行ルートの最適化などが期待される。これらの観点から、車両の走行データの汎用性は確保されるべきであるといえる。

国土交通省は、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推奨している。今後、より詳細な運行記録が要求される可能性や、上に示す4つの課題を解決し、ドライバーおよび運行管理者の負担削減および正確な運行状態の把握を実現するために、センサデータを利用した運行状態自動分類を試みる。本研究では、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を自動分類するものとする。

第4章 運行状態自動分類の提案

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。

本研究で実現するリアルタイム運行状態分類の例を図 4.1 に示す。通常のバス業務において、運行管理者は複数の車両のマネジメントを行う。



図 4.1 本研究で実現するリアルタイム運行状態分類の例

4.1 本研究で分類するバスの運行状態

本研究において、推定するバスの状態を表 4.1 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 4.1 中の「運行」状態を詳細化したものを表 4.2 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）
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 4.1 内で定義を行っている。バス業務の流れを図 4.2 に示す。

4.2 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 4.3 に示す。本研究で利用するセンサデータは検証用に利用している独自開発のセンサノードを利用して取得している。エンジン回転数・車速・総走行距離は、車両内の CAN (Controller

表 4.1 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が OFF
運行	バスが運行中である
休憩	4 時間未満の休みでありエンジンが OFF
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 4.2 「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

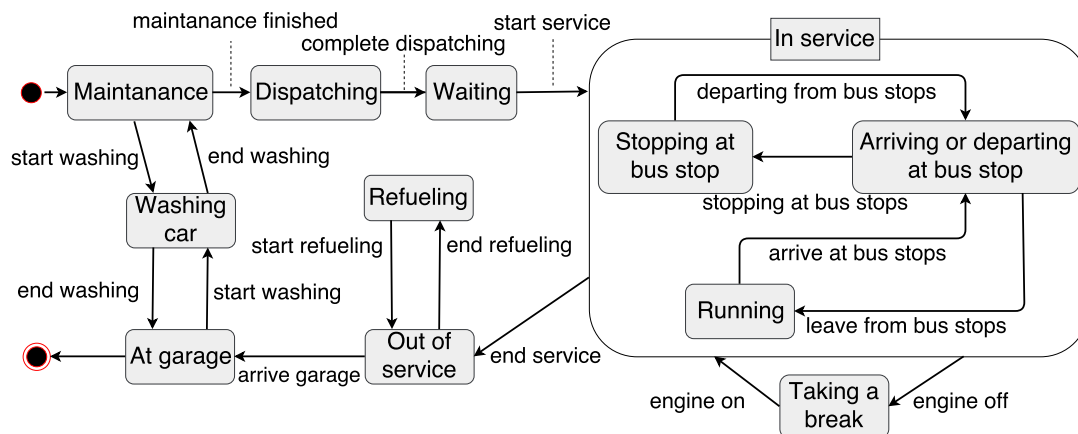


図 4.2 乗合バスの業務の流れ

表 4.3 路線バスから得られるセンサデータ

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）
方位
大気圧
気温
湿度

Area Network) から得られるエンジン回転のパルス値を元に算出している。また、緯度・経度・高度・方位はセンサモジュールに取り付けた GPS モジュールより、湿度・大気圧はセンサモジュール内に埋め込まれたセンサより、気温はセンサモジュールに外付けした気温センサより取得している。

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 0.5 秒おきに取得したものである。

なお、汎用性の観点から、デジタルタコグラフで一般的に利用可能なセンサデータ（緯度、経度、高度、車速、エンジン回転数、総走行距離、方位）のみを状態分類の特徴量として採用した。

4.3 状態分類に利用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 5 つの特徴量を新たに追加し、分類を行う。

- 車速の変化量

得られた車速データから、0.5 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。車速の変化量を定義することで、車両の挙動をより詳細に把握することが可能となる。特に"運行"の3状態の分類精度向上につながると考えられるため、新たな特徴量として定義した。

- 最寄りのバス停までの距離

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直線距離を計算し、定義している。最寄りのバス停までの距離を定義することにより、"運行-バス停停車"、"運行-バス停発着中"の2状態をより高精度に分類できると考えられるため、定義した。

- 事業所までの距離

事業所の座標データと車両の座標データから直線距離を計算し、定義している。"待機"、"点検"、"給油"の3状態は、いずれも車両が停止している状態において行われるものであり、非常に類似している。しかし、"待機状態"は1つ目のバス停付近、"点検"状態は事業所内、"給油"状態はガソリンスタンドあるいは事業所内において行われる。事業所までの距離を、既存のGPS情報に加え、定義することにより、これらの3状態の差異化を実現できると考えられるため、定義した。

- 洗車場までの距離

洗車場の座標データと車両の座標データから直線距離を計算し、定義している。"点検"、"洗車"の2つの状態はどちらも事業所内において、車両が停車した状態で行われるものであり、非常に類似している。また、洗車場は事業所の敷地内に位置しており、GPS情報のみで分類することは難しいと考えられる。洗車場までの直線距離を定義することにより、事業所内における上記の2状態の分類精度向上に寄与すると考えたため、新たな特徴量として定義した。

- 事業所フラグ

事業所における4点の座標データと車両の座標データから、定義している。事業所内において行われる状態は"点検"、"洗車"、"給油"、"帰庫"の4つの状態である。上記の特徴量と併用することにより、上記の4つの状態をより正確に分類するため、新たな特徴量として定義した。

なお、上記の特徴量における2地点間の距離は下記に示すヒュベニの公式に基

づき算出している.

$$2 \text{ 地点間の距離 } D = \sqrt{(D_{lng}M)^2 + (D_{lat}N\cos P)^2} \quad (4.3.1)$$

D_{lng} : 2 地点の緯度差

D_{lat} : 2 地点の経度差

P : 2 地点の緯度の平均

$M = \frac{R_x(1 - E^2)}{W^3}$: 子午線曲率半径

$N = \frac{R_x}{W}$: 卯酉線曲率半径

$W = \sqrt{1 - E^2 \sin^2 P}$

$E = \sqrt{\frac{R_x^2 - R_y^2}{R_x^2}}$: 離心率

R_x : 長半径 (WGS84 に基づく)

R_y : 短半径 (WGS84 に基づく)

本研究では上記の 5 つの特徴量を加え, 表 4.4 に示す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データを元に 0.5 秒毎に算出している.

4.4 状態推定手法の考察

本研究では表 4.1 と表 4.2 に示した 11 個の車両状態を推定するために, 多クラス分類を行う. 下記において, 他クラス分類手法として, 決定木 (CART), Random Forest, SVM, HMM, ルールベース手法, CNN, LSTM の比較および検討を行う. また, 各種法における評価項目の検証結果を表 4.5 に示す.

決定木 (CART)

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である. CART モデルの流れを下記に示す.

表 4.4 状態分類に用いる特徴量

特徴量	データ元
緯度	GPS
経度	GPS
高度	GPS
車速	センサデータ
エンジン回転数	センサデータ
総走行距離 (ODO)	センサデータ
方位	GPS
車速の変化量	車速
最寄りのバス停までの距離	緯度, 経度
事業所までの距離	緯度, 経度
洗車場までの距離	緯度, 経度
事業所フラグ	緯度, 経度

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

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 つとして取り扱う。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 [10]。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

Random Forest では複数の決定木を作成し、バギングを行うという過程が存在するため、データ数が少ないとうまく学習ができないという欠点が挙げられる。

Support Vector Machine (SVM)

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も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 (1 対 1) 方式
 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 [11]。
- one-versus-rest (1 対多) 方式
 K 個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 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 [11]。

Hidden Markov Model (HMM)

HMM は、隠れ変数のあるマルコフモデルのことであり、音声認識や自然言語処理といった時系列データの解析に広く用いられている。路線バスの運行は、図 4.2 に示す状態遷移によって実現されているため、状態の遷移を隠れ変数で表し、現在の運行状態を出力状態とすることが可能である。この観点から、状態推定手法として HMM を利用したモデル化が有用であると考えられる。また、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。

ルールベース手法 (IF-THEN)

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。

Convolutional Neural Network (CNN)

ニューラルネットワークはノード同士の重みを変化させることによって出力を得る機械学習手法の 1 つである。畳み込みニューラルネットワーク (Convolutional Neural Network: CNN) は、手書き数字認識のために提案されたものである。CNN では、畳み込み層とプーリング層が交互に繰り返された後に全結合した多層パーセプトロンが配置されている。畳み込み層では複数のフィルタを用いて入力に対する特徴の抽出を行い、プーリング層では畳み込み層で抽出された特徴に対して間引きを行い、特徴ベクトルを生成する [12]。CNN では、畳み込み層におけるフィルタ生成、プーリング層における特徴ベクトル生成に多量の計算時間およびリソースが必要となるといったデメリットが存在する。

Long short-term memory (LSTM)

LSTM は RNN (Recurrent Neural Network) の拡張として低減されたモデルのことである。LSTM は RNN の中間層のユニットを lstm block と呼ばれるメモリと 3 つのゲートを持つブロックに置き換えることにより実現されている [13]。LSTM では過去に学習したデータを忘却する手段が存在しないため、状況

表 4.5 各分類手法の特徴

	CART	RF	SVM	HMM	IF-THEN	CNN	LSTM
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	×	○	○	×	○	○	×
構築の容易さ	○	○	○	○	×	○	○

に合わせた対応が不可能であった。近年では、上記の課題を解決するために、忘却機能をもたせた、忘却機能付き LSTM が提案されている [14]。忘却機能付き LSTM は状況変化を自動で判別し、記憶を初期化可能となるため、モデルの汎用性を向上させることが可能となった。しかし LSTM では、過去の入力データをさかのぼり、隠れ層の出力値を計算し、伝搬していくため計算量が多くなってしまふという欠点が挙げられる。この計算処理の直列性から、2 Hz で取得している車両センサデータを利用した運行状態分類は、計算資源の観点から困難であると考えられる。

本研究では、表 4.5 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。Random Forest ではデータ数が少ないと上手に学習できない恐れがあるが、本研究に用いるデータセットは車両 4 台から 2 Hz で取得したものであるため、この課題は克服できると考えられる。

4.5 自動分類モデルのシステム化

本節では、リアルタイム分類を実現するためのアーキテクチャについて述べる。

本研究では、自動分類モデルのシステム化を実現するにあたり、IoT のデータ特性に対応した通信プロトコルとして、Message Queuing Telemetry Transport (MQTT) を採用し、図 4.3 に示すアーキテクチャで実装を行う。

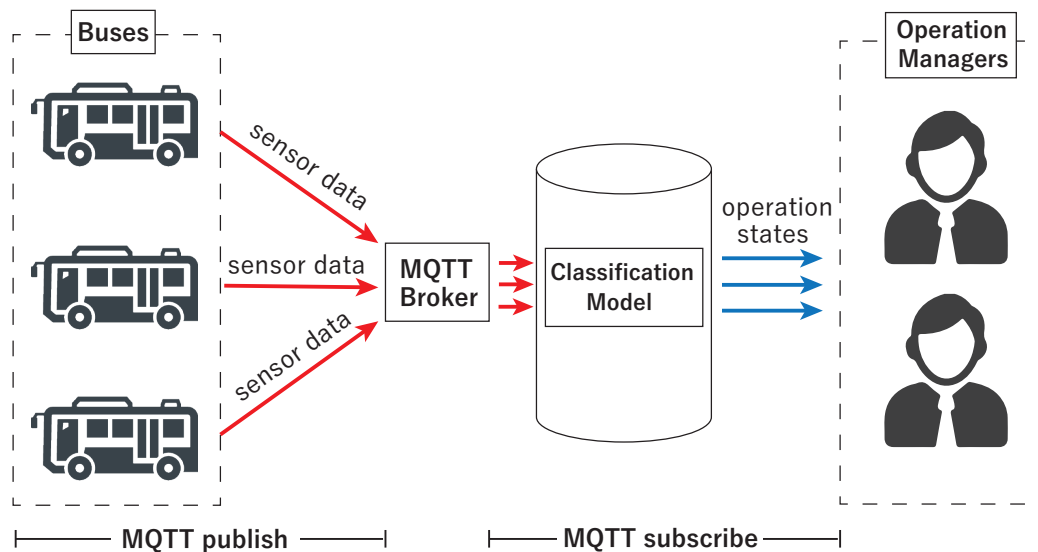


図 4.3 リアルタイム分類実現のためのアーキテクチャ

アーキテクチャの状態分類処理フローは下記のとおりである。

1. MQTT-publisher である車両が，MQTT-broker にセンサデータを publish する
2. 認証ユーザが MQTT-broker を subscribe する
3. ユーザが subscribe したデータを元に，ユーザ帰属の状態分類モデルが状態を自動判定する
4. 状態分類モデルが出力した結果およびセンサデータをユーザに通知する

図 4.3 に示すアーキテクチャにおいて MQTT を採用する理由として，下記の 3 点が挙げられる。

1. 軽量プロトコルであり，狭帯域においても実現が可能
2. 有限リソース下でも安定して動作する
3. Publish/Subscribe モデルを採用しており，柔軟な構築が可能

1 点目について，走行する車両からセンサデータを取得するに際に課題となるのは，狭帯域下におけるデータ送信である．例えば，トンネル内や高架下な

どにおいては、通信の切断や狭帯域化が予測される。MQTT では、メッセージサイズが最小で 2 バイトとなる点や、HTTP と比較した際の通信トラフィックが 1/10 から 1/100 になるなど、非常に軽量な通信プロトコルであり、安定したデータ送信が可能となる [15]。

2 点目について、本研究では車載センサモジュールから 24 時間、2 Hz でセンサデータを取得し続けている。また、車両に取り付けられたセンサモジュールは常時給電のアクセサリソケットから電力を取得している。MQTT では電源の消費量が大幅に削減されており、1 点目と合わせて、使用する計算リソースおよび車載バッテリーの観点から、本研究のアーキテクチャにおけるプロトコルとして適していると言える。

3 点目について、旅客運送業において、様々な運行形態 (貸切バス、路線バス、長距離バスなど) が存在する。運行形態の中には、運行管理者からの密な運行指示が必要のない形態も存在すると考えられるため、全ての車両の運行状態を予測し、運行管理者に送信する必要はないと考えられる。Publish/Subscribe モデルでは、データ取得の際にトピック名を指定できるため、路線バストピックを持つ車両の運行状態のみを通知するなどといった柔軟な実現が可能である。また、Publish/Subscribe モデルの採用により、他者がデータを利用可能となり、危険運転事例や安全運転指導に役立てることが可能となる。

第 5 章 分類結果

評価実験に用いたデータセットは、異なる 2 つの A 路線と B 路線を走行した計 4 車両、合計 8 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、複合路線モデルの分類性能の評価を行った。図 5.1 および図 5.2 にデータ取得対象の 2 路線の路線図を示す。なお、バス停の数はそれぞれ 20 個、11 個となっている。

また、表 5.1 および表 5.2 に取得した運行データにおける業務ごとの開始時間および終了時間、レコード数を示す。車両 No.5016 においては、異なる運転手 3 人が業務を 1 回ずつ行い、合計 3 運行分のデータセットとなっている。車両 No.1758, 2769, 2770 においては、異なる運転手 5 人が業務を 1 回ずつ行い、合計 5 運行分のデータセットとなっている。

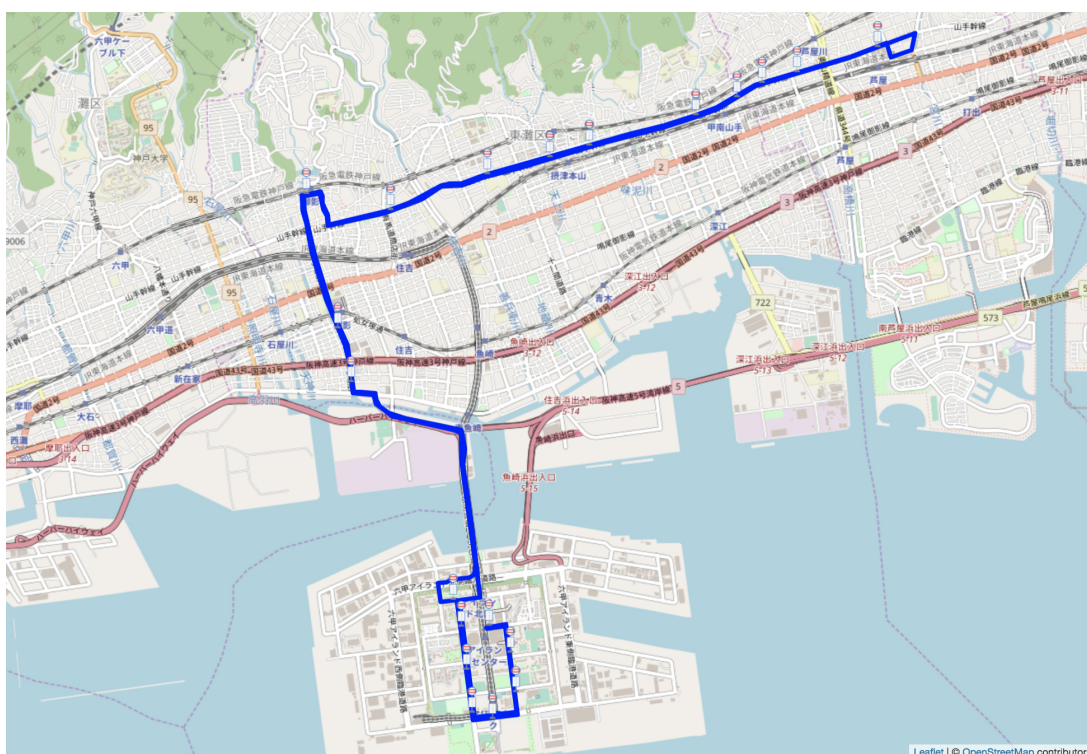


図 5.1 データ取得対象路線 (路線 A)

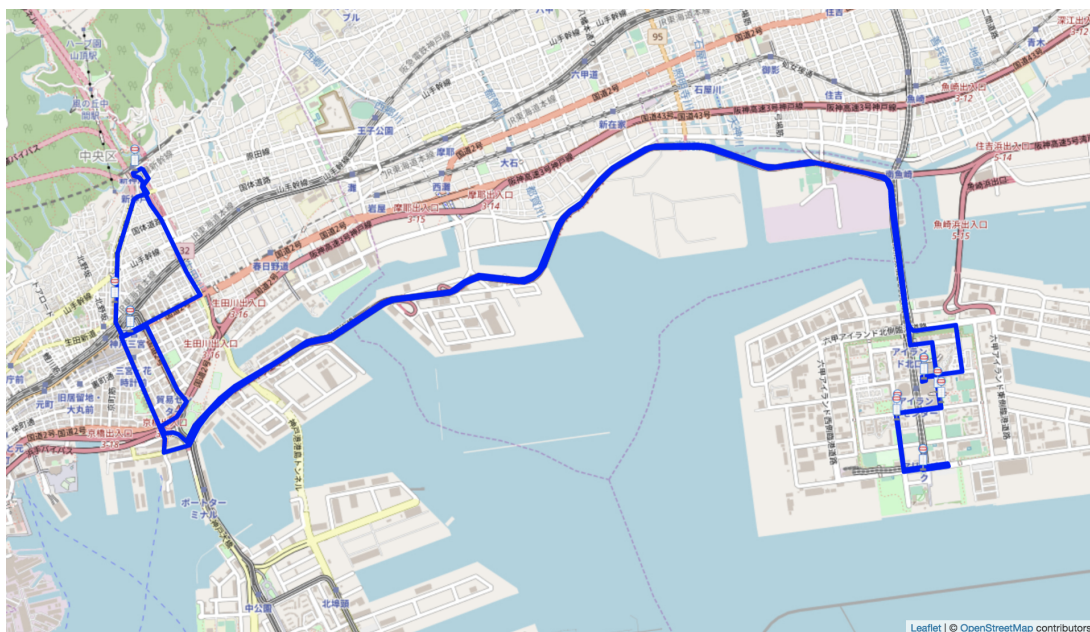


図 5.2 データ取得対象路線 (路線 B)

表 5.1 路線 A から得られたデータセットの概要

車両	時間	レコード数
車両 No.5016	06:41:32-12:25:45	41215
	12:36:08-18:20:56	41266
	18:25:11-22:41:13	30567

表 5.2 路線 B から得られたデータセットの概要

車両	時間	レコード数
車両 No.1758	06:50:43-12:15:53	38971
	13:29:11-21:22:55	56495
車両 No.2769	06:05:00-14:12:23	58463
	15:01:19-21:50:57	49079
車両 No.2770	06:44:51-13:16:28	46920

表 5.3 各車両における運行状態の正答データ総数

	路線 A	路線 B		
	No.5016	No.1758	No.2769	No.2770
点検	6283	6756	7677	2060
配車	4986	1629	1518	765
待機	4586	353	883	325
運行-走行	66642	65335	73071	30886
運行-発着	3058	1501	1930	880
運行-停車	4925	2661	2907	1396
休憩	18901	13817	17304	9046
回送	1760	1933	1485	1024
給油	670	174	256	167
帰庫	144	574	219	371
洗車	1093	733	292	0
合計	113048	95466	107542	46920

また、表 5.3 に各車両における運行状態の正答データ総数を示す。車両 No.5016, 1758, 2768 では図 4.2 における各運行状態全てに 1 度以上遷移しているが、車両 No.2770 では、洗車状態への遷移が一度も行われなかった。

これらのデータセットにおける正答ラベルは、車両に取り付けられた車載カメラの映像および時刻表、車両の走行軌跡をもとに作成した。なお、特徴量は 2 Hz 毎に定義している。正答ラベル作成に用いた車載カメラから得られる映像の例を図 5.3 に示す。

また、本章における評価項目の結果の概略を表 5.4 に示す。

5.1 同一路線モデルによる分類 (路線 B)

本節では、構築した運行状態分類モデルに対し、同一の路線を走行した車両のテストデータを適用した結果を示す。評価では、路線 B において最もレコード数



図 5.3 正答ラベル作成に用いた車載カメラ映像の例

表 5.4 全評価結果の概要

	路線 A	路線 B		
	No.5016	No.1758	No.2769	No.2770
A 路線モデル (No.5016)	-	0.764	0.747	0.681
B 路線モデル (No.2769)	0.661	0.929	-	0.906
A+B 路線モデル (No.5016, 2769)	-	0.919	-	0.873

が多い車両 No.2769 の 2 運行分のデータをモデル構築に利用し，車両 No.1758 の 2 運行分と車両 No.2770 のデータを評価用データセットとして利用した．路線 B における同一路線分類モデルによる分類結果を表 5.5 に示す．

表 5.5 より，両車両において，点検状態，運行-走行状態，運行-停車状態，休憩状態が 0.8 以上と高い正答率を示した．また，待機状態と回送状態については両車両とも正答率が 0.7 以下という結果となった．配車状態，待機状態，給油状態，帰庫状態については，2 車両間において正答率に大きな差が見られた．

5.2 他路線モデルによる分類

本節では，構築した運行状態分類モデルに対し，異なる路線を走行した車両のテストデータを適用した結果を示す．本研究では下記の 2 パターンについて評価を行った．

表 5.5 同一路線モデルによる分類結果 (路線 B)

状態名	車両 No.1758		車両 No.2770	
	正答率	レコード数	正答率	レコード数
点検	1.000	6756	1.000	2060
配車	0.952	1629	0.018	765
待機	0.705	353	0.000	325
運行-走行	0.992	65335	0.990	30886
運行-発着	0.662	1501	0.906	880
運行-停車	0.943	2661	0.974	1396
休憩	0.822	13817	0.801	9046
回送	0.135	1933	0.083	1024
給油	0.000	174	0.887	371
帰庫	0.943	574	0.275	167
洗車	0.000	733	-	-
合計	0.929	95466	0.906	46920

- 路線 A モデル-路線 B データセット
- 路線 B モデル-路線 A データセット

路線 A モデル-路線 B データセット

路線 A モデルの構築には、車両 No.5016 の 3 運行分のデータを利用し、テストデータには車両 No.1758, 2769, 2770 の全ての運行データを利用した。表 5.6 に分類結果を示す。

表 5.6 より、運行-走行状態は 3 車両において、正答率が最低でも 0.937 と、高い結果を示していることが分かる。しかし、待機状態、休憩状態、帰庫状態については、どの車両においても正答率は 0 という結果が得られた。また、回送状態、給油状態、洗車状態の 3 つの状態は、各車両間において、正答率に大きな揺らぎが確認できる。

路線 B モデル-路線 A データセット

路線 B モデルの構築には車両 No.2769 の 2 運行分のデータセットを利用し、

表 5.6 他路線モデルによる分類結果 (路線 A モデル)

状態名	車両 No.1758		車両 No.2769		車両 No.2770	
	正答率	レコード数	正答率	レコード数	正答率	レコード数
点検	0.979	6756	0.817	7677	0.050	2060
配車	0.548	1629	0.582	1518	0.498	765
待機	0.000	353	0.000	883	0.000	325
運行-走行	0.937	65335	0.949	73071	0.959	30886
運行-発着	0.524	1501	0.677	1930	0.669	880
運行-停車	0.734	2661	0.603	2907	0.824	1396
休憩	0.000	13817	0.000	17304	0.000	9046
回送	0.621	1933	0.345	1485	0.107	1024
給油	0.724	174	0.184	256	0.000	167
帰庫	0.000	574	0.000	219	0.000	371
洗車	0.151	733	0.747	292	-	0
合計	0.764	95466	0.747	107542	0.681	46920

テストデータには車両 No.5016 の 3 運行分のデータセットを利用した。表 5.7 に分類結果を示す。

表 5.7 より，点検状態，運行-走行状態，帰庫状態については，正答率が 0.9 以上と高い結果を示していることが分かる。また，待機状態，休憩状態，回送状態，給油状態，洗車状態については，正答率がほぼ 0 という結果となった。

5.3 複数路線モデルによる分類

教師データに複数の路線データを利用した場合の結果を示す。本節では，図 5.1 と図 5.2 に示した 2 つの路線を統合した運行状態分類モデルに対し，路線 A と路線 B を走行した車両データを適用した結果を示す。分類モデルの構築には，車両 No.5016 (A 路線) と車両 No.2769 (B 路線) の全ての運行データを利用し，テストデータには B 路線 (車両 No.1758, 2770) の運行データを利用した。

表 5.7 他路線モデルによる分類結果 (路線 B モデル)

状態名	車両 No.5016	
	正答率	レコード数
点検	0.939	6283
配車	0.248	4986
待機	0.000	4586
運行-走行	0.985	66642
運行-発着	0.485	3058
運行-停車	0.069	4925
休憩	0.000	18901
回送	0.016	1760
給油	0.000	670
帰庫	0.896	144
洗車	0.000	1093
合計	0.661	113048

表 5.8 に分類結果を示す.

表 5.8 より, 点検状態, 運行-走行状態, 運行-停車状態, については, 正答率が 1 近くという非常に高い結果を示した. 回送状態, 給油状態, 洗車状態については, 2 車両とも正答率は 0.3 近く, あるいは 0 という結果となった. 配車状態, 待機状態, 帰庫状態については, 2 車両間に大きな正答率の差が見られたが, 3 状態とも車両 No.1758 の方が非常に高い正答率を示している.

5.4 運転手の手動入力との比較

表 5.4 より, 同一路線モデルの分類精度が最も高精度という結果が得られたため, 本節では路線 B における同一路線分類モデルと運転手の手動入力結果の比較を行った. なお, 比較対象は, 路線 B モデル構築に利用した No.2769 を除く, No.1758 の 2 運行と No.2770 の 1 運行であり, 合計 3 運行, 3 ドライバー

表 5.8 複数路線モデルによる分類結果 (路線 A+B モデル)

状態名	車両 No.1758		車両 No.2770	
	正答率	レコード数	正答率	レコード数
点検	1.000	6756	1.000	2060
配車	0.941	1629	0.051	765
待機	0.992	353	0.000	325
運行-走行	0.980	65335	0.983	30886
運行-発着	0.653	1501	0.876	880
運行-停車	0.952	2661	0.988	1396
休憩	0.779	13817	0.678	9046
回送	0.328	1933	0.217	1024
給油	0.000	174	0.000	371
帰庫	0.937	574	0.018	167
洗車	0.000	733	-	-
合計	0.919	95466	0.873	46920

表 5.9 手動入力と状態分類モデルの正答率の比較

	車両 No.1758		車両 No.2770
	運行 1	運行 2	運行 1
運転手手動入力	0.719	0.812	0.885
状態分類モデル	0.855	0.931	0.906

のデータセットとした。本研究では 0.5 秒おきに運行状態を定義しているが、ドライバーが実業務中に同様の粒度で運行状態を登録することは困難であると考えられる。よって、評価には 10 秒毎、つまり 20 レコード毎に記録されている状態の正誤を判定し、評価結果とした。

表 5.9 に運転手の手動入力による正答率と、運行状態分類モデルが出力した正答率の比較を示す。

表 5.9 より、各運行において、運転手の手動入力結果よりも運行状態分類モデ

ルの方が高い正答率を示していることが分かる。運転手の手動記録において、運行状態から休憩状態への切り替え忘れが非常に多く見られた。しかし表 5.3 に示した通り、1 業務において、運行の 3 状態（運行-走行，運行-発着，運行-停車）が占める割合が非常に高かったため，上記の切り替え忘れが発生しても，運転手の記録結果の正答率が極度に低くなるということではなかった。

第 6 章 考察および今後の展望

6.1 考察

本研究において構築した A 路線モデル, B 路線モデル, A+B 路線モデルにおける特徴量の重要度の上位 5 項目を表 6.1 に示す.

各分類モデルにおいて, 新たに定義した特徴量である最寄りのバス停までの距離, 事業所までの距離, 洗車場までの距離が上位 3 項目を占めていることが分かる. 特に, A 路線モデルと A+B 路線モデルにおいては最寄りのバス停までの距離の貢献度が 0.25 弱となっており, この特徴量が大きく貢献していることが分かる. 下記において, 構築した 3 つの状態分類モデルの評価結果について, 考察を行う.

同一路線分類モデル (路線 B)

点検状態の正答率は車両 No.1758, 2770 とも 1 となっている. 事業所内において停車している際に行われる他の状態として, 洗車状態および車両 No.1759 における給油状態が存在する. しかし, これらの 2 状態のほとんどが点検状態に誤分類されていた. 事業所内において停車している場合のレコードのほとんどが点検状態に分類され, 結果として点検状態の正答率が 1 となった.

運行-停車状態については, 表 6.1 に示したとおり, 最寄りのバス停までの距離を新たに定義し, 特徴量として貢献していると考えられるため, 正答率が 0.95 前後と非常に高い精度を示している.

休憩状態については, どの車両も同じ地点において休憩状態に遷移する. 休憩

表 6.1 各分類モデルにおける特徴量の重要度

A 路線モデル		B 路線モデル		A+B 路線モデル	
最寄バス停距離	0.227	事業所距離	0.197	最寄バス停距離	0.237
事業所距離	0.172	最寄バス停距離	0.145	洗車場距離	0.141
洗車場距離	0.110	洗車場距離	0.125	事業所距離	0.111
総走行距離	0.094	緯度	0.107	総走行距離	0.101
経度	0.093	エンジン回転数	0.092	緯度	0.071

状態の定義は"4 時間未満の休みでありエンジンが OFF"となっていたが、データセット中のエンジン回転数を確認すると、休憩状態に遷移しているにもかかわらずエンジン回転数が 0 になっていない、つまりエンジンを OFF にしていない事例が数多く見られた。この要因としては、当日の交通状況によるものが大きいと考えられる。路線 A では、片道約 35 分の経路をピストン運行する路線となっている。しかし、データ取得日が休日だったことも影響し、次の運行開始までの時間が短い場合にはエンジンをかけたまま休憩状態に遷移し、そのまま運行状態に遷移というパターンが見られた。このような定義通りの休憩を行っていない場合において誤分類が発生している。

給油状態については、ガソリンの給油場所が異なるため、分類結果に大きな差が発生したと考えられる。モデル構築に利用した車両 No.2769 および No.1758 は、運行終了後に近隣のガソリンスタンドにおいて給油した後に事業所に帰庫する。しかし車両 No.2770 は事業所内に設置されているガソリン給油装置を利用して給油しており、事業所内において行われる他の状態である点検状態に誤分類されていた。

洗車状態および点検状態はどちらも事業所内において行われる状態となっている。本研究では、これらの 2 状態の差別化を図るために新たな特徴量として、事業所までの距離および洗車場までの距離を定義した。しかし、洗車状態終了後に車両の移動を行わず、点検状態に遷移する例が多く、データセット上ではこれらの 2 状態の区分がほぼ不可能となっていたと考えられる。洗車状態のほとんどが点検状態に誤分類されており、このような結果となった。

他路線分類モデル（路線 A モデル、路線 B モデル）

点検状態については、同一路線分類モデルと同様に、事業所内において停車している場合にはレコードのほとんどが点検状態に誤分類されており、結果として点検状態の正答率が高くなった。

待機状態、休憩状態については、路線 A モデルおよび路線 B モデルにおいて、正答率が全て 0 という結果となった。A 路線および B 路線において、待機と休憩を行う地点がほぼ固定である。よって、別の路線で構築したモデルでは分類が困難という結果となった。

運行-停車状態については、同一路線分類モデルでは、非常に高い正答率を示

していたが、他路線分類モデルにおける正答率の平均は 0.7 程度と大きく低下していることが分かる。この理由として、新たに定義した特徴量である、最寄りのバス停までの距離が原因であると考えられる。図 5.1 に示す路線 A と図 5.2 に示す路線 B を比較すると、路線前半では島内において共通するバス停を経由しているが、路線後半では、それぞれ東西に走行していき、まったく異なるバス停を経由していることが分かる。異なるバス停を経由するため、バス停間の距離間隔が変化し、別路線のデータセットの分類が困難になったと考えられる。また、運行-発着状態についても上記の理由により正答率が低くなったと考えられる。

給油状態については、車両 No.5016 と No.1758 が同一のガソリンスタンドにおいて給油を行っていた。よって、車両 No.5016 から構築した路線 B モデルを車両 No.1758 に適用した際にも、正答率が極度に低下するということとはなかったと考えられる。

複数路線分類モデル

点検状態については、同一路線分類モデル、他路線分類モデルと傾向が同様になっており、正答率が 1 となっている。

休憩状態、運行-発着状態、運行-停車状態は GPS 情報に大きく依存していると考えられ、A 路線と B 路線を統合したモデルでは、同一路線分類モデルより正答率が低く、他路線分類モデルより高い正答率が得られた。

給油状態については、モデル構築に使用した 2 車両分のデータセット両方においてガソリンスタンドにおいて給油を行っているため、正答率が 0 となっている。

本研究では 2 路線を統合した分類モデルであり、全体での正答率は同一路線分類モデルと比べると低い結果となっている。今後、新たな路線のデータセットを追加し、正答率の変化を検証する必要があると考えられる。

6.2 今後の展望

今後検証すべき課題として、下記の内容が挙げられる。

- モデルの分類の安定性の検証
- 別の路線を走行した車両のデータセットの採用

- 正答率が低かった状態の分類精度向上
- 他の分類手法の検証

1 つ目について、本研究では A 路線で 1 車両、B 路線で 3 車両のデータセットを用いて検証を行った。表 5.5 では、同一路線の走行データからモデルを構築したにもかかわらず、車両 No.1758 と No.2770 の正答率に大きな揺らぎが確認できる。例えば、配車状態においては、車両 No.1758 での正答率は 0.952 であるが、車両 No.2770 では 0.05 と大きな差が確認できる。このような分類のバイアスの原因を検証するために、運行データをさらに増やし、モデルの分類精度をより詳細に検証する必要があると考えられる。

2 つ目について、本研究では A 路線と B 路線の 2 路線を対象として検証を行った。表 5.4 では、複数のモデルのうち、同一路線分類モデルの分類精度が最も高いという結果となった。しかし、複数路線分類モデルにおいては、同一路線分類モデルでは正答率が低かった待機状態および回送状態の正答率が向上している。これは、路線を統合したモデルを構築したことにより、汎用性が向上したことと起因する可能性がある。よって、より多くの路線を統合した複数路線分類モデルを構築し、各運行状態の正答率の変化を検証する必要がある。

3 つ目について、分類精度が最も高かった表 5.5 に示す同一路線分類モデルにおいて、回送状態、給油状態、洗車状態の正答率が非常に低いことが確認できる。例えば、回送状態は一定時間の業務が終了した後に行われる状態であるにもかかわらず、運行前に行われる配車状態に誤分類されていた。この要因は、本研究で提案した状態分類モデルが状態遷移を考慮していないことに起因すると考えられる。Random Forest が出力した結果に対して、簡単なルールベース手法を利用した状態遷移の検証を行うことで、出力結果の補正や状態遷移の正当性の検証が可能になると考えられる。

4 つ目について、本研究ではモデル化の容易さおよび分類の高速性を考慮し、Random Forest を採用したが、上記で述べたように状態遷移の観点から誤った遷移をしている出力結果が存在した。本研究では採用しなかった HMM は状態遷移を反映可能な分類モデルとなっており、分類精度が向上する可能性がある。HMM を用いたリアルタイム運行状態分類が可能か否かを検証するために、現行のデータ取得間隔で HMM の検証および、サーバの負荷検証を行う必要がある。

と考えられる.

第 7 章 結論

本稿では、まず近年の旅客運送業における現状について述べ、それらの要因及び本研究のモチベーションについて述べた。次に、旅客運送業において一般的に利用されている運行管理方法の概要、利点、課題について解説し、本研究の新規性について述べた。既存の運行管理方法では、ドライバーに対する端末操作の負担、運行データの二次利用の困難さ、専用端末の導入コストといった複数の課題が存在する。そこで、本研究では、これらの課題を解決するために、Random Forest を利用した運行状態分類モデルの提案を行い、システムを実現するための MQTT ベースのアーキテクチャを提案した。また、センサモジュールから得られるセンサデータおよび、それらから計算可能な特徴量を新たに定義し、運行状態分類モデルの高精度化を試みた。評価を行うにあたり、神戸市内のみなど観光バス株式会社協力の下、車両 4 台の 1 日分の実運行データおよび車載カメラの映像を利用して正解データを作成した。次に評価項目として、(1) 同一路線分類モデル、(2) 他路線分類モデル、(3) 複数路線分類モデル、を設定し、構築したモデルの評価および、運転手が手動で入力した運行記録との比較を行った。評価の結果、同一路線分類モデルと複数路線分類モデルは、一部の運行状態の分類精度は低いものの、全体としては 0.9 前後の正答率を示した。また、運転手が手動入力した運行記録と比較すると、同一路線分類モデルモデルの正答率が手動入力を上回るという結果が得られた。汎用性という観点では、本提案の実現をするにあたり、静的に必要なデータは路線に属するバス停の緯度・経度情報のみとなっている。各バス会社においてバス停の緯度・経度情報を管理することは一般的であるため、5 章の結果より、本研究で提案した手法は他の路線・バス会社においても有用であると考えられる。今後は、正答率が低い運行状態の分類を実現するために、ルールベース手法を利用した状態遷移の正当性判定や、他路線のデータセットを増やし、分類モデルの追加検証が必要となる。

謝辞

主指導教員であり，適切な研究指導をしていただくなど様々な側面から研究のサポートをしてくださいました本学情報基盤システム学研究室の藤川和利教授に心から感謝致します．副指導教員であり，研究の方向性についての的確な助言をくださいました本学ユビキタスコンピューティングシステム研究室の安本慶一教授に心から感謝致します．副指導教員であり，研究指導，論文執筆指導，そして多くの研究発表の機会を与えてくださいました本学情報基盤システム学研究室の新井イスマイル准教授に心から感謝致します．研究だけでなく，WIDE プロジェクトの活動にも積極的に助言を下さいました，本学情報基盤システム学研究室の垣内正年助教に心から感謝致します．研究活動や論文執筆に関してご指導を戴きました，本学情報基盤システム学研究室の油谷暁助教に心から感謝致します．研究に関して的確なご指摘，ご指導を戴きました，京都産業大学コンピュータ理工学部の秋山豊和准教授に心から感謝致します．研究活動を行うにあたり，データ提供および実験に快く協力してくださいました，みなと観光バス株式会社の皆様に心から感謝いたします．研究面や生活面において，多くの支援をしていただいた本学博士後期課程修了生であり現東京工業大学学術国際情報センター助教の石井将大氏に心から感謝致します．様々な面から研究活動を支援してくださいました本学総合情報基盤センターの鷺尾多佳子女史，辻元理恵女史，中野彩子女史に心より感謝致します．様々な面から研究活動を支え，共に切磋琢磨しあった本学情報基盤システム学研究室の同期とはじめとした友人の皆様に心より感謝致します．最後に私の意思を尊重し，研究活動に関する理解を示すとともに，経済面や生活面において多大な支援を頂き最後まで見守ってくださいました家族に心から感謝致します．

参考文献

- [1] The Japan Times, “Overnight bus crashes in Karuizawa, killing 14,” (<https://www.japantimes.co.jp/news/2016/01/15/national/night-ski-bus-flips-karuizawa-three-killed-11-critical-condition/#.WgT4Fx00MWp>)(accessed 2017-11-10)
- [2] 公益社団法人 日本バス協会, “2016 年版 日本のバス事業” (http://www.bus.or.jp/about/pdf/h28_busjigyo.pdf) (accessed 2017-12-9)
- [3] 国土交通省, “自動車関係統計データ” (<http://www.mlit.go.jp/common/000117167.pdf>) (accessed 2017-12-9)
- [4] 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” (<https://www.mlit.go.jp/common/001127098.pdf>) (accessed 2017-12-9)
- [5] 国土交通省, “貨物自動車運送事業輸送安全規則”
- [6] 厚生労働省 安全衛生関係リーフレット, “交通労働災害防止のための新しい安全衛生管理手法のすすめ” (<http://www.mhlw.go.jp/new-info/kobetu/roudou/gyousei/anzen/dl/0912-2d.pdf>) (accessed 2017-12-9)
- [7] M. Koike and M. Isogai, “Telematics Service for Commercial Vehicles to Realize Safe Traffic Society” Fujitsu Scientific and Technical Journal, vol.51, no.4, pp.43–47, Oct. 2015.
- [8] 株式会社デンソーセールス, “DN-magic PREMIUM” (https://www.denso-sales.co.jp/d-navi/product/dn-magic_premium/) (accessed 2017-12-9)
- [9] 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” (<http://www.transtron.com/products/dts-d1d.html>) (accessed 2017-12-9)
- [10] Liaw, Andy, and Matthew Wiener, “Classification and regression by randomForest.” R News, vol.2, no.3, pp.18–22, Dec. 2002.
- [11] Hsu, Chih-Wei, and Chih-Jen Lin, “A comparison of methods for multi-

- class support vector machines.” IEEE transactions on Neural Networks, vol.13, no.2, pp.415–425, 2002.
- [12] Y.LeCun, L.Bottou, Y.Bengio and P.Haffner, “Gradient-based learning applied to document recognition.” Proceedings of the IEEE vol.86, no.11, pp.2278–2324, 1988.
 - [13] Felix A.Gers, Nicol N.Schraudolph and Jürgen Schmidhuber, “Learning Precise Timing with LSTM Recurrent Networks,” Journal of machine learning research, pp.115–143, Aug. 2002.
 - [14] Felix.A.Gers, J.Schmidhuber and F.Cummins, “Learning to forget: continual prediction with LSTM,” Neural Computation, vol.12, no.10, pp.2451–2471, Oct. 2000.
 - [15] MQTT.ORG, MQTT, (<http://mqtt.org/>) (accessed 2017-11-10)

発表リスト

国際会議

Takuya Yonezawa, Ismail Arai, Toyokazu Akiyama, Kazutoshi Fujikawa, “Random Forest Based Bus Operation States Classification Using Vehicle Sensor Data,” International Workshop on Pervasive Flow of Things (PerFoT 2018), Athens, Greece, Mar. 2018. (Accepted)

Takuya Yonezawa, Ismail Arai, Toyokazu Akiyama, Kazutoshi Fujikawa, “Proposal of classification method of bus operation states using sensor data,” Proceedings of 2017 IEEE International Conference on Big Data, pp.4781–4783, Dec. 2017.

国内会議

米澤 拓也, 新井 イスマイル, 藤川 和利, “路線バスから得られたセンサーデータを利用した運行状態分類モデルの評価,” マルチメディア, 分散, 協調とモバイル (DICOMO2017) シンポジウム 論文集, pp.65–72, Jun. 2017.

米澤 拓也, 新井 イスマイル, 藤川 和利, “路線バスから得られたセンサーデータの解析と車両状態推定に関する検討,” 研究報告モバイルコンピューティングとパーベシブシステム (MBL) , Vol.2016-MBL-81, No.3, pp.1–5, Dec. 2016.