

路線バスから得られたセンサデータの解析と 車両状態推定に関する検討

米澤 拓也^{1,a)} 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．我々はこの課題を解決し，旅客運送業における業務効率化を実現するために，神戸市内のバス会社が発信しているセンサ情報を用いたバスの状態推定を試みている．センサデータの種類は，緯度，経度，高度，エンジン回転数，車速，総走行距離，パルス値などが挙げられ，各データを一定間隔おき取得している．本稿では，センサデータを利用したバスの状態推定手法の検討結果を報告する．

キーワード：機械学習，データマイニング，センサデータ，ITS

1. はじめに

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．

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人が雇用されており，全運転者中に占める割合は17.4%と運転手の6人に1人以上が満60歳以上であると報告している[1]．また，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，一部のバス事業者の収益は悪化し，運行管理者の人員が削減され，まともな運行管理ができなくなっているのが現状である．また，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もの，法令を承知していても法令遵守の意識が低い事業者が多いことがこのような事故の原因となっていると考えられる．

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．

本研究では，これらの課題を解決するために，路線バスから得られるセンサデータを利用し，路線バスの状態を自動推定することを目的とする．

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

a) yonezawa.takuya.yq2@is.naist.jp

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。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に記録されるデータの例として、車速、時間、距離、エンジン回転数、GPSによる位置情報などが挙げられ、アナログタコグラフでは記録できなかった情報を電子的に記録することが可能となった。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。

- リアルタイム性の欠如

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。しかし、車速やGPS情報、エンジン回転数を生データとして表示するものがほとんどで、運行状態をリアルタイムで推定するものは存在しない。

- 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を利用した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利用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。

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を示す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った。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。また、表1中の運行状態を詳細化したものを表2に示す。運行状態におけるバス停車状態とバス停発車状態は別途詳細な推定を行うものとする。

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) 点検作業後に初めの停留所に向けて移動(配車)
- (3) 停留所においてバスが待機(待機)
- (4) バスが運行状態に移行(運行)
- (5) バスがバス停に停車(バス停停車)

表 1 本研究で推定するバスの状態

状態名	定義
点検	出発前の点検作業中である
配車	点検後に停留所に向けて移動中である
待機	営業所外において停車中である
運行	バスが運行状態である
休憩	4 時間未満の休み
給油	停車していて燃料残量が増加している
洗車	洗車作業中である
帰庫	事業所内において停車中である
回送	運行を終了している

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
	バス停発車	バス停から発車する

- (6) バスがバス停から発車 (バス停発車)
- (7) 再度バスが運行状態に移行 (運行)
- (8) 一定時間の運行後に休憩を行う (休憩)
- (9) バスが運行を終了 (回送)
- (10) バスが事業所内に帰庫 (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。

3.3 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、ワイパー、ブレーキ、左右ウインカーといった自動車の動作に関するデータは CAN (Controller Area Network) を利用して取得し、緯度、経度、標高データは車載 GPS モジュールから取得を行っている。下記では、より詳細な車両状態の把握のために取得を行っているセンサデータについて解説を行う。

- VOC
 VOC (Volatile Organic Compounds) は揮発性有機化合物のことで、トルエン、キシレン、酢酸エチルなどが挙げられる。
- CO₂
 CO₂ は車内の CO₂ 濃度を表す。CO₂ 濃度を利用して車内の混雑度や車内空間の換気量を表す指標として利用する。また、運転手の眠気の推定に利用できる可能性がある。
- 水温
 水温は、エンジンを冷却するために利用している水の温度を示す。
- 油圧
 油圧は、エンジン内部を循環しているエンジンオイルにかかる圧力のことであり、油圧を計測することで、エンジンオイルの劣化具合を把握することが可能と

表 3 路線バスから得られるセンサデータ

データ名	データ取得間隔 [s]
緯度	1
経度	1
誤差	1
標高	1
車速	0.5
エンジン回転数	0.5
総走行距離 (ODO)	0.5
車内温度	1
気圧	0.5
湿度	0.5
VOC	0.5
CO ₂	0.5
燃料残量	0.5
水温	0.5
油圧	0.5
ワイパー	0.5
ブレーキ	0.5
左右ウインカー	0.5
3 軸加速度センサ	0.05

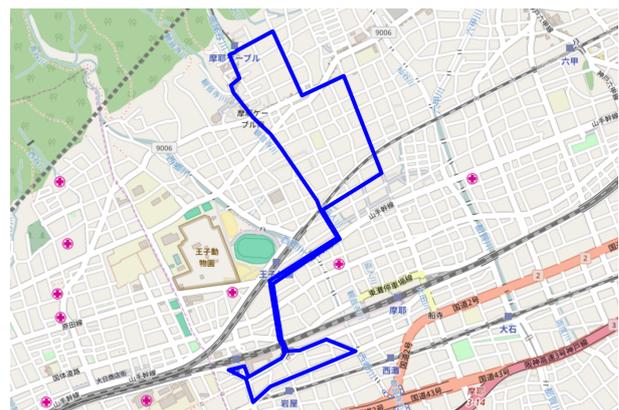


図 1 データ取得対象車両の走行ルート

なる。

- 車内温度
 車内温度は、路線バス内に取り付けられた温度計から得られる車内温度情報である。車内温度を利用して、車内の混雑度や空気質を表す指標として利用する。
 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路線バスから取得したものである。図 1 にデータ取得対象車両の走行ルートを示す。

4. 状態推定手法

4.1 ルールベース手法と機械学習

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 に示した 9 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。

4.2 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 4 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。決定木とはアンサンブル学習における弱学習器の 1 つであり、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 [3]。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

表 4 各分類手法の特徴

	CART	RF	vs-one	vs-rest	Ada
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	×				

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 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 [4]。

- 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、この 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 [4]。

- AdaBoost

AdaBoost とは弱学習器の組み合わせから精度の高い学習器を構築する Boosting の 1 種である。AdaBoost ではサンプルデータに対して重みを導入し、その重み変化させることによって弱学習器毎に異なった入力データを与えることを実現している。しかし、AdaBoost では重みを変化させる際に、重みの値は正解だと小さくなり、不正解だと大きくなるという特徴がある。この特徴により、訓練データにノイズや例外的な事例が含まれる際に、性能が極端に悪化してしまうという欠点が挙げられる [5]。

本研究では、表 4 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う。

5. 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。今後の方針としては、センサデータ取得環境が整い次第、運転手が手

動で記録した各時間における運行状態を教師データとして利用し, Random Forest を利用した多クラス分類器の構築および評価を行う予定である.

本研究の今後の展望として, 路線バスの状態推定を提案したが, 提案した分類手法が有用であれば, 長距離トラックやタクシーといった貨物運送車や他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] Liaw, Andy, and Matthew Wiener: Classification and regression by randomForest. R news 2.3 (2002): 18-22.
- [4] 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.
- [5] 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