

# 勾配ブースティング決定木を用いた高速道路における突発事象発生時の旅行時間推定

広田 和也<sup>1</sup> 佐藤 雅明<sup>1,2</sup> 中村 修<sup>3</sup>

**概要:** 高速道路では日常的に渋滞が発生し、時間のロス、ドライバーの疲労、そして事故リスクの上昇等を招いている。一方で、現在ドライバーに対して提供されている道路交通情報は、ほとんどが現在か過去の所要時間情報や規制情報である。このためドライバーが最適な経路を決定するには、単に複数経路の所要時間を比較するだけでなく、規制内容や交通量等を参考にして自分が通過する際の影響を各自で推測し判断する必要がある。

本研究では既存の問題を解決する手法として、首都高速の各路線について日時の情報と突発事象を含めた規制情報をもとに、所要時間を推定するシステムを構築した。今回、機械学習を用いた分散型勾配ブースティングフレームワークの LightGBM を用いて、時刻情報と規制情報を学習し、決定木を生成することでその時点での所要時間が推定可能であることを確認した。また、推定に規制情報を活用した場合には活用しなかった場合に比べて推定の精度が向上することを示し、規制情報と所要時間の変動に因果関係があることを確認した。

この検証結果によって、将来的に規制情報や規制の計画が公開された際に、その時間帯に所要時間がどのように変化するかを推定する研究の実現可能性を示した。

## Estimating the OD Travel Time at the Time of Traffic Regulation Using Gradient Boosting Decision Tree

KAZUYA HIROTA<sup>1</sup> MASAACKI SATO<sup>1,2</sup> OSAMU NAKAMURA<sup>3</sup>

### 1. 背景

#### 1.1 渋滞の特徴

高速道路上では日常的に渋滞が発生する。大口 [1] によると渋滞は、その発生する要因によって交通集中渋滞と突発渋滞に大別される。このうち交通集中渋滞は道路・交通条件により生じる一方、突発渋滞は交通事故や車両故障などの突発事象により発生する。この突発事象はいつ、どこで発生するか予見したり、あらかじめ回避をしたりすることが困難である。

渋滞は時間のロス、ドライバーの疲労、事故リスクの上

昇等の損失を招く。このため多くのドライバーは、道路交通情報を収集し、迂回や時間調整などを行い渋滞の発生している箇所を回避しようとする。

#### 1.1.1 渋滞の発生する要因

渋滞は、その発生する要因によって交通集中渋滞と突発渋滞に大別される [1]。また首都高速道路は、主な渋滞発生要因を5つに分類している [2]。このうち道路の構造に由来する「上り勾配」「合流部」「出口付近」といったボトルネックは長期にわたって変化せず、大口 [1] の分類における交通集中渋滞の原因となる。一方で「事故等の異常事象」は予見できない突発事象である。

#### 1.2 首都高速における渋滞

首都高速道路は、東京都、神奈川県、埼玉県、千葉県の

<sup>1</sup> 慶應義塾大学大学院 政策・メディア研究科  
Graduate school of Media and Governance, Keio University

<sup>2</sup> 東海大学 観光学部  
Department of tourism, Tokai University

<sup>3</sup> 慶應義塾大学 環境情報学部  
Faculty of Environment and Information Studies, Keio University

エリアに総延長 327km の路線を有する都市高速道路である。路線の多くが密集市街地の中の公共用地を利用して建設されており、カーブが多く路肩が少ない、左右それぞれからの合流があり加速車線が短いなど、高速自動車国道と比べ制約の多い構造が特徴である。また交通量が多いことから慢性的に混雑していることも特徴である。

### 1.2.1 首都高速における交通規制

突発事象が発生した場合、道路管理者や警察は安全確保、後片付け、現場検証のために現場の走路において一部の車線を走行させないよう規制したり、あるいは全ての車線を封鎖して通行止めとする場合がある。これらの規制により走路上には平常時とは異なるボトルネックが発生することになり、交通を滞留させ渋滞を招く。

また、首都高速は道路上や河川上の高架構造やトンネルなどといった構造物が非常に多く、昼夜を問わず維持管理のための作業が実施されていることも特徴である。これらの工事による車線規制は事故等の突発事象によるものと同様、渋滞の発生する要因となりうる。

## 1.3 本研究の目的

高速道路上で車線規制を伴う渋滞が発生した際、複数存在する経路の中から目的地に最も早く到着するための経路を選択するには、自分がその車線規制等のボトルネックが発生している地点を通過するまでに所要時間がどのように推移するかを推定することが不可欠である。しかし、周辺の地理に明るくない運転者をはじめ多くの運転者にとって難しく、また現状ではこのような推定された道路交通状況は、基本的には運転者へ提供されていない。

また社領ら [3] によると、日常的に利用している道路であっても多くの人が最良の判断ができなかったケースも存在するという。

この問題を解決するには、周期的な交通集中も加味しながら突発事象に対応し、現在当該区間を走行している、あるいはまもなく走行しようとしている利用者にとって必要な時間帯の所要時間を推測するシステムが求められる、

そこで本研究では、このシステムを実現する前段階として、過去の所要時間と規制情報のログデータを学習し、時刻情報と規制情報からその時点の所要時間を推定するシステムを構築する。この推定を高精度に行うことによって利用者の多くが正確な目的地への到着見込み時刻を得ることができるようになり、将来の交通状況を考慮した最善の行動がとれる助けとなる。

## 2. 道路交通情報提供の現状

### 2.1 道路交通情報の概要

本稿では道路交通に関する用語を以下のように定義する。

### 交通状況

現実世界の道路上における、それぞれの車両の位置や速度などを含めた交通流の状況である。

### 道路交通情報

交通状況を計測したデータを加工して一般化し、また道路管理者や警察等が入力した規制状況などとともに取りまとめられ、利用者にとって有益な情報として提供されている道路交通の概況である。

### 突発事象

事故、故障車、落下物等の、道路上で発生する予見が困難な事象である。予め定められた計画的な車線規制や道路工事などは含まない。

### 車線規制

道路管理者の認知している突発事象や道路工事が存在し、1つ以上の車線が閉鎖され走行可能な車線数が減少しているが通行することは可能な状況である。通行止め、出入口や休憩施設の封鎖、JCT 等の本線を除くランプの封鎖は含まない。

## 2.2 交通状況の取得

### 2.2.1 道路会社による所要時間の取得

首都高速では主に車両感知器を用いて交通状況を収集している [4]。これは超音波の反射を利用して設置地点上の車両の存在を感知するセンサーが 5m の間隔において 2 台 1 組で設置されているもので、車両の通行台数や速度などを計測する。首都高速ではこの車両感知器が 300 から 600m 間隔で設置されている。これらの車両感知器によって収集されたデータは 1 分ごとに管制システムに集計され、交通量、平均速度、時間占有率が計算される。

各車両検知器には勢力範囲が設定されている。ある区間の所要時間を算出する際は、その範囲をそれぞれの検知器で計測された速度で通行するのに要する時間の総をその区間の所要時間としている。また、都市高速道路の場合、速度が 20km 以下の場合その勢力範囲は渋滞しており、40km 以下の場合は混雑していると判断されている [1][5]。

### 2.2.2 道路会社による突発事象の情報の収集

電話通報、パトロールで収集された情報、交通管制用テレビカメラの映像は交通管制室に集約される [4]。交通管制室が事故や落下物等の情報を把握した際には、管制システムや道路情報提供装置への情報の入力や、警察や JARTIC をはじめとした関係機関への連絡が行われる。

## 2.3 道路交通情報の提供

日本における道路交通情報の提供方法には、道路事業者の設置する施設を用いた直接的な提供方法と、公益財団法人日本道路交通情報センター (JARTIC) を経由する方法の

2種類に大別される。

日本における道路交通情報の提供主体は都道府県公安委員会であり、公安委員会は道路交通情報の提供に係る事務を JARTIC に委託している。また、道路事業者および道路管理者は、道路の維持、修繕その他の管理のための情報提供を行う。道路交通情報は図 1 に示したように、道路管理者の運用する道路情報提供装置での提供を除くと、全て JARTIC を経由して行われている。

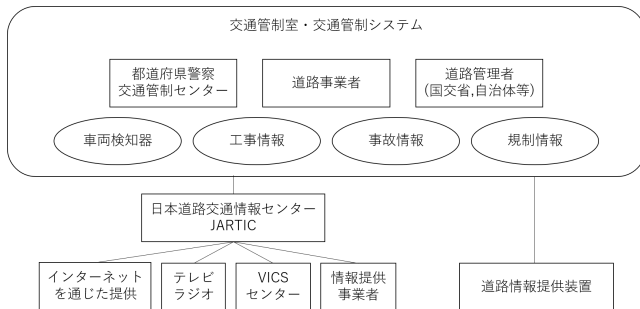


図 1: 道路交通情報提供の流れ

### 2.3.1 道路情報提供装置等での情報提供

首都高速をはじめ道路事業者は、道路利用者に向けて直接的に情報提供を行うために専用の装置を道路上に設置している [4][6]。このうち道路上に設置されているものには、文字や簡易図形を表示する文字情報板、路線図を模して広いエリアの情報提供を行う図形情報板、主要地点までの所要時間の表示に特化した所要時間表示板など複数の種類がある。また休憩施設には「首都高ナビ」等の大型モニターを設置し、幅広い情報提供に利用している。これらは、道路事業者の管制センターから直接入力された情報を使用している。

また、道路上には VICS(道路交通情報通信システム) や ETC2.0 に対応した電波ビーコンが設置されている。これらの発する電波はカーナビゲーションシステム等の車載機器で受信され、それらの車載機器に表示されて活用される。

この他に路側放送(ハイウェイラジオ等)で音声を用いた情報提供も行う場合があるが、首都高速においては現在提供されていない。

### 2.3.2 JARTIC を経由した情報提供

道路事業者や道路管理者が収集した情報は JARTIC が一元的に集約している。JARTIC は J システム [7] を通じて所要時間情報、渋滞情報、突発事象や工事などの事象規制情報を事業者向けに提供すると同時に、電話、ラジオ、テレビ、FM 多重放送、独自の Web サイト等を通して利用者向けに提供している。

また、首都高速など多くの道路事業者は JARTIC の提供する情報を活用し、WEB サイトやスマートフォンアプリケーションなどによって文字や画像を用いた道路交通

情報の提供を行っている。また首都高速においては近年「Twitter」や「LINE」といったソーシャルネットワーキングサービス(SNS)上での半自動的な情報配信も運用している [8]。

## 2.4 まとめ

日本における道路交通情報の提供方法には、道路事業者・道路管理者自身による提供と JARTIC を経由した提供の 2 種類の方法が存在する。ただし、道路事業者による道路情報提供装置を使った直接的な情報提供を受けられる場合は運転中などに限られており、誰でもどこでも手に入れられる情報は JARTIC を経由したものに限定されている。

JARTIC の提供する情報は全国で画一的な形式で配信されており、入手と再利用が容易である特徴がある。一方で通行量などは JARTIC は提供していないことから道路事業者にしかならない情報や、道路情報提供装置等で提供されながらも JARTIC からは提供されない情報なども存在する。

このことは、後述する既存の道路交通情報の推定を一般利用者や中小の主体が再現するにあたって、大きな支障となっている。

## 3. 所要時間推定の手法および事例

### 3.1 所要時間推定の分野における先行研究

既存の将来の所要時間推定の手法に関する研究は、推定手法の違いによって以下の 4 種類に大別することができる。

#### 3.1.1 パターンマッチングによる推定

直近の時間帯における道路上の各区間の断面交通量、車速等の時系列化されたデータの推移を手がかりに、過去のデータの中からこれと類似した推移をもつものを抽出し、得られた断面交通量や車速等の推定値から所要時間を推定するものである。

割田ら [9] は、このうち所要時間データの推移に着目して推定を行っており、首都高速 3 号渋谷線上りでの適用性を確認している一方、事故等の突発事象の発生した際などには推定精度が低下するとしている。

#### 3.1.2 決定木による推定

直近の車速や断面交通量、事故発生状況といった情報を用いて条件分岐を行い、その条件下での典型的な車速、断面交通量、所要時間といった値を採用する。

金澤ら [10] は決定木を利用して、一般道において現在の渋滞発生状況等から将来の所要時間を推定する研究に取り組んでいる。

#### 3.1.3 回帰分析による推定

内村ら [11] は、回帰分析、類似パターン検索、自己回帰移動平均 ARMA モデルによる所要時間推定をそれぞれ行い、比較している。この結果、回帰分析を用いた手法と類

似パターン検索を用いた手法では、ARMA モデルを適用した手法と比べて良好な結果が得られると結論づけている。

また芦田ら [12] は決定木と回帰分析を組み合わせ、決定木により分岐されたそれぞれの条件下で回帰分析を用いた所要時間推定に取り組み、この条件下においてのみ事故渋滞の発生時であっても同様のケースのみを回帰分析で学習することで高精度な推定が可能であるとしている。

### 3.1.4 ニューラルネットワーク (NN) による推定

近年では機械学習の手法である CNN や GNN を用いて過去の交通状況データを学習し推定を行う手法が提案されている。

阪神高速道路の寺前ら [13] は阪神高速道路における交通状況データを利用して所要時間推定に取り組み、CNN を利用した場合と GCN を利用した場合の精度の特性を比較分析している。

## 3.2 既存の交通状況推定の提供事例

交通状況を推定し実際に提供を行っているサービスは、中長期的な推定を行うものと短期的な推定を行うものの2種類に大別できる。ここでは、翌日以降の日毎の交通状況の大きなトレンドを推定するものを中長期的なトレンドの推定、当日から数日以内の時間帯ごとの交通状況を推定するものを短期的な推定とする。

### 3.2.1 年間、月間の日毎の交通状況の推定 (渋滞予想カレンダー)

首都高速をはじめ高速道路事業者は、それぞれ独自に中長期的な渋滞予想を作成し発表している。首都高速の場合は「渋滞予想カレンダー」と題して、毎月月末に翌月の日毎の渋滞の程度を、「特に多い」「多い」「ふつう」「少ない」の4段階で推定し、発表している。

ただし2020年5月頃以降、高速道路事業者各社とも新型コロナウイルス感染症の影響による交通状況の変化を理由として渋滞予想の提供を中止しており、2021年7月現在もこの状況が続いている。

### 3.2.2 繁忙期の交通状況の推定

お盆休みや正月休み、春の大型連休 (ゴールデンウィーク) 等といった繁忙期には、例年、通常の渋滞予想とは別に当該期間の渋滞に関する予想が発表される。これは当該期間のうち特に混雑が予想される日を挙げるとともに、特に大規模な渋滞や混雑が予想される箇所について最大の渋滞等の予想が発表される。

こちらも2020年春の大型連休で発表が見送られていたが、2022年春の大型連休には Nexco3 社および本四高速から発表されている [14]。

### 3.2.3 ドコモらによる「AI 渋滞予知」

株式会社 NTT ドコモは Nexco 東日本および Nexco 中

日本と共同で、「AI 渋滞予知」と題して一部の路線における特定の時間帯の交通状況の推定を行っている [15]。これはドコモの提供するモバイル空間統計情報と過去の渋滞実績情報を学習したモデルを用い、当日正午の人口分布から14時以降の時間帯別の交通状況を推定し、提供するものである [16]。現在、東京湾アクアライン (木更津 JCT-川崎浮島 JCT)、関越自動車道 (沼田 IC-練馬 IC)、東名高速道路 (御殿場 JCT-東京 IC) の3箇所のそれぞれ上り線がサービス対象となっている。

### 3.2.4 ナビタイムの渋滞予測

株式会社ナビタイムジャパンは自社の WEB サービス「Navitime」 [17] 上で、日本国内における各高速道路の交通状況を閲覧日時から1ヶ月先までの期間において1時間ごとに推定したものを、地図上にプロットして公開している。

ナビタイムはこのサービスにおいて推定に利用しているデータについて、自社サービスのユーザーから収集したプローブ情報であるとしているが、これ以上は具体的に使用しているデータセットの詳細、予測情報の更新頻度等は公開されていない。なお、事故や工事による渋滞については考慮されていないという。

## 3.3 まとめ

以上で挙げた所要時間推定手法および提供事例の特徴をまとめると、以下のようになる。なお、ここでいう推定時期の「中期的」は12ヶ月先の推定値まで公開していることを指し、「長期的」はそれ以上のものを指す。

本章で紹介した各種の所要時間推定事例を、表1および表2にまとめた。

突発事象については、それぞれ文献内で突発事象の発生した場合にも良好な精度で推定が可能であると結論づけている場合には○と、突発事象へは対応できないとしている場合は×とした。

またデータの入手性については、無償で一般に公開されているデータを使用している場合は○とすることとしたが、該当するものは無かった。有償契約や複数台の対応車両の用意で得られるデータを使用している場合△と、道路管理者しか得られない情報や大手 IT プラットフォーマーでないと得難い大規模なプローブデータを使用している場合は×とした。

## 4. 提案手法

### 4.1 提案手法の詳細

本研究においては、LightGBM を利用した機械学習手法によって決定木を作成し、近い将来の道路交通状況の推定に利用する。LightGBM は、Microsoft Corporation の開

表 1: 各論文の所要時間推定のカテゴリ

事例	手法	推定時期	突発事象	データ入手
割田ら	パターン	数分後	×	×
金澤ら	決定木	数分後	○	△
内村ら	回帰分析	数時間後	×	△
芦田ら	回帰分析	数時間後	×	×
阪神高速	NN	数分後	×	×

表 2: 各社の交通状況サービスの状況

提供事例	提供者	推定時期	突発事象	データ入手
渋滞予想	各事業者	長期	×	△
繁忙期	各事業者	中長期	×	△
AI 渋滞予知	ドコモら	数時間	×	×
渋滞推定	ナビタイム	数時間 ～中期	×	×

発した分散型勾配ブースティングフレームワークである。決定木ベースの学習アルゴリズムを使用しており、弱学習器を組み合わせ、損失が少なくなるようなモデルを構築する。軽量であり、学習した結果得られた決定木を可視化することができるのが利点である。一方、他の手法と比べ過学習しやすいとされている。

今回所要時間推定に決定木ベースの手法を利用するにあたり、金澤ら [10] の研究によって決定木による渋滞発生・継続の予測の実現可能性が示されている点を参考にした。また LightGBM は XGboost に比べて精度を保ちながら高速化が図られているとされ、この点は本手法について将来的に定期的な学習モデルの再構築を行う際に有益であると考えた。

本研究では学習データとして曜日、時刻、対象となる区間の所要時間、首都高速各路線上の主要区間の車線規制等の時系列データセットを与える。そして、推定対象の時点での曜日、時刻、規制のデータをもとに平均車速を推定させる。

#### 4.2 他手法との比較

提案手法について、既に提案、運用されている所要時間推定手法と比較する。

既存の交通状況推定手法の多くは、道路構造物をモデリングし、また別途その地点に流入する断面交通量を推定することでボトルネックとなる部分での滞留を推定し、渋滞の発生や所要時間の変化を推定する。これらの手法は比較的高精度な推定が得られやすいとされる。一方、対象区間の数だけ道路構造物をモデル化して用意する必要があり、また第 2 章で述べたとおりリアルタイムの断面交通量には基本的に道路事業者や道路管理者しか入手できず、再現することは困難である。

自動車やモバイル端末、専用の GPS 端末を利用して交通

流の中にある車両の座標や移動速度を利用して推定する手法も、対象とする区間の上に十分な数の参加者が必要である。このためドコモ等の大手 IT 企業やナビタイム等の著名カーナビゲーションアプリの配信元、移動体通信会社といったプラットフォーマー以外には運用に至るまでのハードルが非常に高い。

提案手法においては、公開されている道路交通情報のみを使うため、誰でも容易に本手法を活用して検証、活用を行うことができる。また、特に JARTIC の提供する道路交通情報は全国の主要な道路において同一のフォーマット、同程度の品質で配信されており、区間ごとのデータセットを用意することができれば他区間への展開も容易である。さらに機械学習を用いているため、再度の学習も自動化することが可能である。このことから、たとえば昨今の新型コロナウイルス感染症の影響下のように、従来とは大きく異なった交通状況が中長期的に続いた際にも、再度その状況下に特化した推定アルゴリズムを用意し利用することで、高精度な可能であると考えられる。

### 5. 実験システム設計

#### 5.1 システムの概要

本システムの概要を図 2 に示す。本システムは、大きく分けて情報受信部、ログデータの編集部、そしてデータの学習及び推定部の 3 つで構成されている。

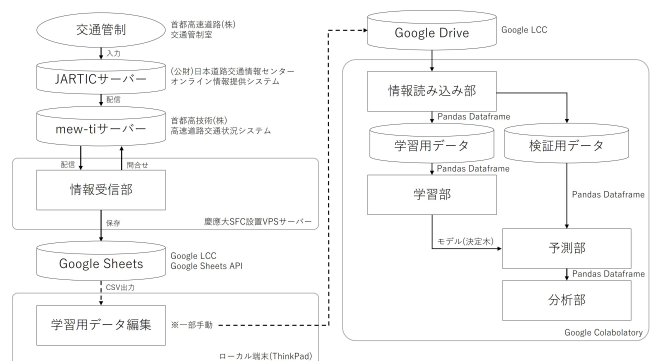


図 2: 本システムの概要

情報受信部では、首都高速の交通管制システムから入力された情報を、JARTIC、mew-ti のシステムを経由して 5 分間隔で取得した。

ログデータ編集部では、Python を利用し、情報受信部で取得した情報を一部編集して学習・推定部に渡すための下処理をした。なおこの操作は本来自動化されるべきところであるが、実験の都合上手元でおこなった。

学習・推定部では後述する Google Colabory を利用し、Jupyter Notebook 上の Python 環境を利用して、LightGBM を用いた学習・推定を行った。

### 5.1.1 情報受信部の実行環境

道路交通情報の収集には、慶應義塾湘南藤沢 ITC の提供する VPS サーバーを利用し、Ubuntu 上で 5 分間隔で定期実行をしいる。

### 5.1.2 ログデータ編集部

収集済みのログデータ (CSV ファイル) に対して、実験に使用する期間のデータの抽出し、日付時刻データの変換といった操作を、ローカルの Ubuntu 端末上で行った。

## 5.2 データの概要

### 5.2.1 Web より取得する情報

「mew-ti」[18] は首都高速道路株式会社のグループ企業である首都高技術株式会社の運営する道路交通情報を提供する WEB サービスである。首都高速の管制システムから JARTIC のサービスを経由して取得した交通情報をもとに交通情報を提供している。

本システムでは、首都高技術株式会社の助言のもと、mew-ti より以下の情報を取得し、Google Sheets に保存している。これを情報受信部の VPS 上で Python を用いてパースし、Google Sheets に保存する。

#### 主要地点間の所要時間情報

今回、首都高速各路線の起終点に近い位置にある出入口を首都高速における主要地点とし、これらの出入口を起終点とした区間の所要時間を取得した。「起終点に近い出入口」は首都高速の LINE Bot[8] にて提供されている路線別所要時間情報の起終点を、標準所要時間は LINEBot 上で首都高速が公式に提供している「渋滞がない場合の所要時間」を採用している。

都心環状線、中央環状線、湾岸線については路線の構造等を考慮し、路線を複数区間に分割して各区間の所要時間を取得した。中央環状線については出入口構造等の問題で内回りと外回りで採用する出入口が異なる箇所がある。八重洲線、8 号線については、路線自体長が極端に短いため、所要時間情報を収集していない。

なお、この所要時間はドライバー向けに公表される名目値であるため、実勢速度が規制速度等の条件を大きく超える場合、その区間を法定速度で走行した場合の速度が所要時間として提示される。この極端に短い所要時間情報は発信されないことになっている。

また、ここで取得する所要時間情報は、起終点間の移動に利用可能なルートのうち利用時間が最短になる経路が選択されている。このため交通状況が順調な場合は最短経路となるルートでの所要時間をとるが、本線通行止、JCT のランプの通行止、大規模な渋滞が発生した場合にはこれ以外の迂回路を利用した場合の所

要時間が提供される。

#### 交通規制情報

所要時間情報と同様に mew-ti サーバーから JSON 形式で配信されている、表 3 に示す交通規制に関する情報を収集している。

表 3: 収集する規制情報データ

項目	データ型
更新時刻	string
路線名称	string
路線名称コード	number
方面	string
開始地点	string
終了地点	string
事由	string
規制内容	string
事由コード	number
規制内容コード	number

#### 取得日時

上記の道路交通情報を取得した日時を記録する。なお本研究においては、上記の交通情報は収集日時の時点における交通状況を反映した真値であるとみなす。

### 5.2.2 学習・推定に与える情報

学習には、それぞれの区間について以下の情報を与える。

#### 目的変数

- 地点間の所要時間情報 (int)  
上記で取得した所要時間情報を与える。

#### 説明変数

- 取得した曜日、時刻 (float)  
取得を行った曜日と時刻が、三角関数を用いてそれぞれ 7 日周期、24 時間周期で循環する数値に変換して与える。
- 路線・方向ごとの道路異常情報件数 (int)  
mew-ti サーバーから取得した規制情報データのうち、その路線上でのあらゆる「規制」に関する情報の件数を与える。この件数にはたとえば路肩に故障者が存在する等、道路上に何らかの異常が場合ものの車線規制は行われていないケースも含まれている。  
所要時間情報と異なり、長大な路線においても区間を分割していない。
- 路線・方向ごとの道路異常情報件数 (int)  
mew-ti サーバーから取得した規制情報データのうち、その路線の本線上での車線規制に関する情報の件数を与える。その車線規制が 1 車線規制であるか

2車線規制であるかは加味しない。これは、首都高における路線のうち湾岸線を除くとほぼ全線が片側2車線である点を考慮し、与えるデータを簡素化するためにこのような仕様とした。また、通行止や出入口閉鎖の情報は含まない。

- 車線規制の新規発生 (bool)
  - 車線規制の件数が0件から1件以上になった場合には True が、それ以外のときには False が与えられる。

### 5.3 データの収集期間およびデータ量

学習・推定に使用するデータは、2020年7月1日から2021年3月31日までに5分間隔で収集した。この総件数は70,586件である。

## 6. 実装

### 6.1 学習・推定システムの実装環境

提案手法におけるログデータの学習、推測には Google LLC の提供する Colaboratory (Colab) を利用した。Colab はクラウド上で提供される Jupyter Notebook 環境であり、ユーザーは Python をはじめとした各種プログラムを自由に実行することができる。また Nvidia CUDA を利用した GPGPU にも対応しているが、今回の検証については GPGPU は使用しなかった。

表 4: 提案手法における学習、推測の実行環境

項目	仕様
名称	Google Colaboratory
CPU	Intel Xeon CPU x86-64Bit
RAM	12GB
ストレージ	70GB
GPU	Nvidia Tesla T4
グラフィックスメモリ	15GB
OS	Ubuntu 18.04.5 LTS
Python	3.7.10
jupyter core	4.5.0

### 6.2 動作フロー

本システム学習・推定部の動作フローを図3に示す。

ログデータ編集部で下処理を行ったログデータは、CSV形式で出力し Google Drive 上に保存する。この CSV ファイルを Colab 上で読み取り、Pandas の DataFrame に格納する。その後 DataFrame 都度の学習や推定実験に使用するデータを抽出し、学習用データと推定実験用のデータに分割する。なお、分割する際は学習用データに 80%、推定実験用データに 20%の行数を与え、抽出はランダムに行

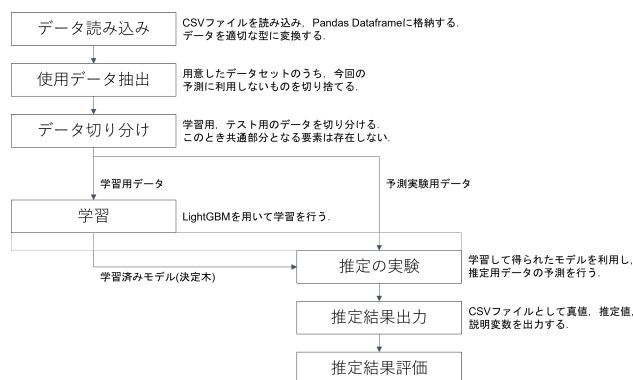


図 3: 学習・推定部の動作フロー\*1

う。このとき、ある時点のデータが両方に共通して与えられることはない。また推定実験用データからは時刻と規制情報のみを説明変数とし、所要時間等の情報は切り捨てている。

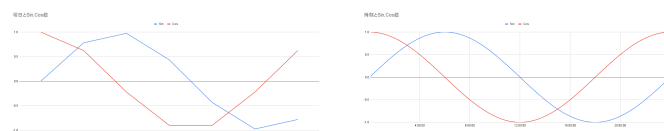
その後、学習用データを用いて LightGBM の学習を行う。この際、学習モデルの評価軸として平均二乗偏差 (RSME) を指定しており、LightGBM は学習用データを分類した際にその推定値と真値の間の RSME が最小となるよう決定木を生成し、採用する。

最後に、用意した評価実験用データについて所要時間等の情報の推定を行い、結果を CSV ファイルとして出力する。

### 6.3 日時データの取り扱い

今回、学習及び推定には日時に関するデータのうち曜日、時刻のみを用いる。これは首都高速「渋滞対応ガイド」[2]には曜日によって渋滞量が変化することが示されていることを参考にしつつ、過去7ヶ月分のみのデータを学習する今回の実験においてこれ以上大きな周期の変動を考慮することは困難だと判断したためである。

月日、曜日、時刻の値については循環する数値として取り扱うために、それぞれ三角関数 (sin,cos) を用いて、一定周期で連続する数値 (-1 以上 1 以下の正弦波) に変換して与えた。



(a) 曜日と Sin,Cos 値 (b) 時刻と Sin,Cos 値

図 4: 曜日、時刻と Sin,Cos 値の関係

## 7. 評価

### 7.1 評価対象路線

前章で挙げた推定対象とする路線のうち、以下の路線のデータを用いて所要時間を推定する実験を行い、提案手法を評価する。

- 3号 渋谷線 上り 用賀-霞が関間

首都高速における典型的な放射路線のひとつであり、渋滞や事故多発地点 [19] が多いため。また接続先の東名高速道路が全国有数の交通量と渋滞多発地点を抱える路線であることから渋谷線を選択した。

#### 7.1.1 学習するデータ

評価実験において、所要時間推定モデルを作成するために学習するデータは以下の通りとする。

- 曜日、時刻  
曜日と時刻を三角関数で変換した数値を与える。
- 交通情報  
学習する交通情報は、第5章で述べた対象全路線の主要地点間の所要時間または、所要時間と道のりから求めた平均車速を与える。
- 規制情報の件数 (件)  
学習する規制情報として、対象全路線における車線規制情報、全ての規制情報それぞれの件数を与える。なお全ての規制情報には通行止めや、道路上に異常が発生しているものの車線規制は行われていない「規制なし」などの情報も含まれている。
- 学習対象とする時間帯  
学習させるデータは、第5章で収集したデータから夜間 (22時~5時) のデータを取り除き、日中のデータのみを用いて学習を行う。
- データ量  
学習させるデータの量は、第5章で取得した7万件から夜間を除き80を抽出した約4万件を学習に使用する。なお各手法の比較時に学習、推定に用いるデータ群はそれぞれ同一である。

#### 7.1.2 推定に利用するデータ

評価実験において、所要時間推定モデルを作成するために学習するデータは以下の通りとする。

- 曜日、時刻  
曜日と時刻を三角関数で変換した数値を与える。
- 規制情報の件数  
推定に使用する規制情報として、各区間における車線規制情報と全ての規制情報それぞれの件数を与える。

以上の要件をまとめて、表5のとおり3つの条件での推定を行うこととする。

表 5: 評価実験の条件

	対象区間	決定木モデルの目的変数	規制件数
提案手法	用賀→霞が関	所要時間	与える
手法 2	用賀→霞が関	所要時間	与えない
手法 3	用賀→霞が関	平均速度	与える

### 7.2 データの外れ値

評価実験において、所要時間の値が全体の1パーセントタイルを下回った行、また取得に失敗している行については外れ値として当該の行のデータを破棄した。今回、この処理によって学習用・推定用データを合わせて49928件中89件 (0.18%) のデータが破棄された。

### 7.3 評価の概要と結果

推定結果についての評価の方針と結果を述べる。

### 7.4 散布図, 箱ひげ図

提案手法における推定値と真値の散布図, 箱ひげ図を図5に示し、推定値の大まかな傾向を確認する。なお真値の最小値, 第1四分位数, 第2四分位数 (中央値) の値は14で一致している。

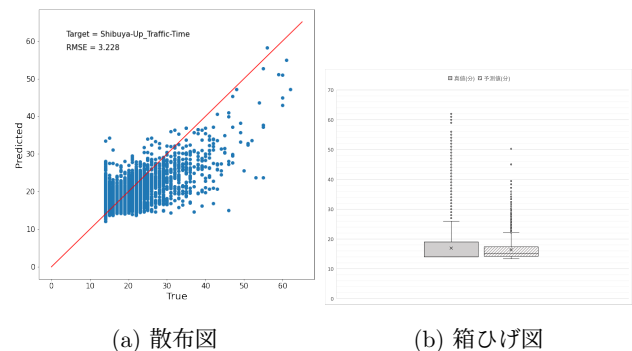


図 5: 提案手法における真値, 予測値の散布図, 箱ひげ図

この散布図から、おおむね真値が大きいときほど予測値が大きくなる傾向が確認できる。

### 7.5 精度

本研究において、精度評価には以下の指標を用いる。なお平均車速ベースの推定においては、路線の長さから推定された平均車速から所要時間を計算し直し、これを評価する。

- 平均値  
推定値の平均値と真値の平均値をそれぞれ算出する。
- 誤差平均  
推定値と真値の差 (絶対値) の平均値。値が小さいほど精度が高い。



- 誤差率平均  
真値に対する誤差の大きさの平均値。値が小さいほど精度が高い。
- 大誤差発生率  
所要時間真値の 25%を超える誤差が発生した割合。値が小さいほど精度が高い。
- 平均二乗偏差 (RMSE)  
推定値と真値の間の RMSE の大きさ。値が小さいほど精度が高い。

実験結果の概要と設定した精度の指標について、結果は表 7.5 のようになった。

表 6: 精度の結果

	真値平均 (分)	推定値平均 (分)	誤差平均 (分)
提案手法	16.94	16.96	1.89
手法 2	16.94	16.98	1.69
手法 3	16.94	16.40	1.81
	誤差率平均 (%)	大誤差発生率 (%)	RMSE (分)
提案手法	10.09	9.89	3.23
手法 2	12.58	14.14	4.32
手法 3	8.79	8.67	3.51

結果について、RMSE、誤差率平均、大誤差発生率は、提案手法が、規制情報を使わなかった手法 2 に比べ良好となった。誤差平均については規制情報を使用しないほうが若干良くなった。これは、データの大部分を占める流れがスムーズなときの予測結果が、規制情報不使用の手法ではより誤差が少なく、平均値を改善させているものと考えられる。

また、LightGBM によって平均車速を推測させた手法 3 の場合、誤差率の平均と大誤差発生率については提案手法に比べ改善した一方で、RMSE については若干悪化した。このうち誤差率の平均については、平均車速が流れの順調なときほど数値が大きくなるものであり、この平均車速についての RMSE が最良となるよう学習を行った結果、流れの悪くなった際の誤差が軽視されたものと考えられる。

## 7.6 ヒストグラム

用意した 3 手法について、真値と予測値の誤差を図 6 のヒストグラムに示し、その傾向を確認する。このヒストグラムにおいて誤差が正の値であるとき、予測値が真値より大きいことを示す。

ヒストグラムから、提案手法および手法 3 では規制情報を使わない手法 2 に比べ、誤差がゼロ付近により集中して

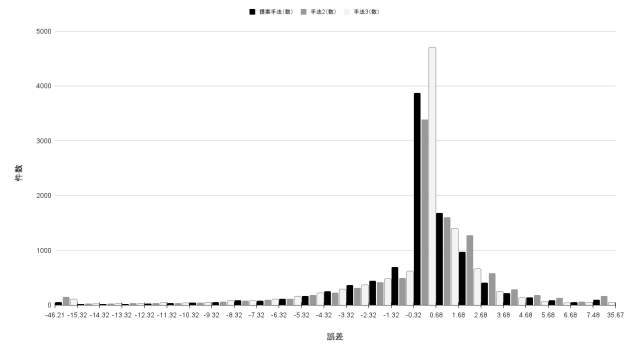


図 6: 3つの手法それぞれで発生する誤差

いることがわかる。また手法 2 では提案手法よりもプラス誤差が出やすくなっている。このことから手法 2 では規制による所要時間増加時にも誤差が少なくなるよう、推定値が大きめにとられている可能性がある。手法 3 では、提案手法に比べると僅かにマイナスの誤差が出やすい。

## 7.7 決定木における重要度の確認

提案手法によって学習して得られた LightGBM の決定木モデルにおける各特徴量の重要度について、上位 10 件を表 7 に再度まとめた。なお、説明変数の重要度の数値は相対値である。

表 7: 提案手法での説明変数の重要度のうち上位 10 件

順位	説明変数	重要度 (相対値)
1	時刻 (sin)	569
2	時刻 (cos)	333
3	曜日 (sin)	257
4	曜日 (cos)	138
5	渋谷線 上り 車線規制件数	117
6	渋谷線 上り すべての規制件数	75
7	晴海線 上り 車線規制の新規発生	73
8	川崎線 下り 車線規制件数	71
9	池袋線 下り 車線規制件数	57
10	目黒線 下り 車線規制件数	45

このうち、とくに時刻、曜日に関する情報が上位を占めた。これらは朝夕の通勤ラッシュ時をはじめ、交通量に大きく影響する変数となっていると考えられ、交通集中渋滞による所要時間への影響を示すものであると考えられる。

また、5,6 位には予測対象とした渋谷線の的上りにおける車線規制、規制全体の件数が位置している。これらは突発渋滞の原因事象となりうるものであり、規制情報、特に車線規制情報が所要時間に影響を与えていることを示している。

一方で、7 位以下には渋谷線に隣接しない各路線に関する

る車線規制情報が挙がった。これらの項目がどのように渋谷線上りの所要時間に影響しているかは説明がつかず、LightGBM が何らかの相関性を発見した、あるいは過学習などが発生した可能性が考えられる。渋谷線の逆方向の規制情報は、上位に入らなかった。

## 8. 結論

本研究では、細やかな交通規制が日常的に多発する首都高速道路に着目し、時刻と JARTIC から得られる規制情報のみを用いて所要時間推定を行うシステムを構築した。また本手法において、特に規制件数情報について所要時間を推定する上で重要な尺度となることを示した。

これらのことから、本研究においては将来的に任意の時点の所要時間について、時刻情報と規制情報をもとに長期間にわたって、あるいはある程度遠い未来であっても所要時間推定を行う研究について、実現可能性が示された。

## 9. 今後の課題と展望

今回は短期間のデータのみを学習し、学習データと同時期の交通状況を予測した。今後はより長期的に所要時間情報を収集し、より長期間にわたって所要時間を予測した際の精度や、学習すべき期間の長さについて分析を行う必要がある。

## 10. 謝辞

本論文の執筆に際してご支援頂いた中村修研究会インターネット自動車研究グループの皆様へ感謝します。また、道路交通情報の収集について助言を頂いた首都高技術株式会社の皆様、ならびに首都高速道路株式会社の皆様へ感謝いたします。

## 参考文献

- [1] 大口敬. 「交通渋滞」徹底解剖. 交通工学研究会, 2005.
- [2] 首都高速道路株式会社. 首都高ドライバーズサイト 各種渋滞情報. <https://www.shutoko.jp/traffic/traffic-info/> 最終確認日 2020 年 06 月 16 日.
- [3] 社領沢, 日下部貴彦, and 朝倉康夫. 都市高速道路における突発事象時の行動調査とその分析. *土木計画学研究・講演集 44 巻*, 2011.
- [4] 首都高速道路株式会社. 首都高の交通管制. [https://www.shutoko.jp/~media/pdf/responsive/customer/use/convenience/infoboard/koutsukansei\\_20160701.pdf](https://www.shutoko.jp/~media/pdf/responsive/customer/use/convenience/infoboard/koutsukansei_20160701.pdf) 最終確認日 2020 年 06 月 16 日.
- [5] 川辺謙一. *首都高速の科学*. ブルーボックス B-1840. 講談社, 2013.
- [6] 首都高速道路株式会社. 首都高の情報提供について. <https://www.shutoko.jp/use/convenience/infoboard/> 最終確認日 2020 年 06 月 16 日.
- [7] 公益財団法人日本道路交通情報センター. 法人向け情報提供サービス j システム. <https://www.jartic.or.jp/s/service/forcorporation/forcorporation01/> 最終

- 確認日 2020 年 07 月 12 日.
- [8] 西田 亘, 佐藤 雅明, 村井 純, 原田 日郎, 金子 聖哉, and 隠田 歩乃加. SNS を用いた道路交通情報逐次配信システムの構築と運用. *マルチメディア, 分散, 協調とモバイル (DICOMO2021) シンポジウム*, 2021.
  - [9] 割田博, 森田紳之, 桑原雅夫, and 田中淳. 統計的手法による所要時間情報提供に関する研究. 2003.
  - [10] 金澤明浩, 杵渕哲也, 毛利仁士, 小川智章, 市河研一, and 荒川賢一. 決定木を利用した交通渋滞予測手法に関する考察. Technical Report 19(2003-ITS-016), 日本電信電話株式会社 NTT サイバースペース研究所, 日本電信電話株式会社 NTT サイバースペース研究所, 東日本電信電話株式会社研究開発センター, 東日本電信電話株式会社神奈川支店, 日本電信電話株式会社 NTT サイバースペース研究所, 日本電信電話株式会社 NTT サイバースペース研究所, mar 2004.
  - [11] 内村圭一, 成松裕介, and 衛藤旭秀. バスロケーション情報を用いたバス停間所要時間予測. *国際交通安全学会誌*, 32(3):64–71, 2007.
  - [12] 芦田優太 and 西岡到. 路側データから生成した交通流モデルによる高速道路の交通状況予測. Technical Report 16, NEC 情報・ナレッジ研究所, NEC 情報・ナレッジ研究所, nov 2015.
  - [13] 寺前智文, 向井理沙, 西海能史, 鈴木健太郎, 小島悠紀子, 廣田敦士, 幡山五郎, 影本義明, 櫻木伸也, and 阿部敦. Graph convolutional networks を用いた阪神高速道路における交通渋滞発生予測. *第 18 回 ITS シンポジウム 2020*, 2020.
  - [14] 東日本高速道路株式会社. ゴールデンウィークの渋滞予報ガイドの掲載について. [https://www.e-nexco.co.jp/news/important\\_info/2022/0415/00011233.html](https://www.e-nexco.co.jp/news/important_info/2022/0415/00011233.html) 最終確認日 2020 年 05 月 20 日.
  - [15] 東日本高速道路株式会社. 高速道路 ai 渋滞予知の実証実験 (nexco 東日本&ntt ドコモ). [https://www.driveplaza.com/trip/area/kanto/traffic/ai\\_traffic\\_prediction.html](https://www.driveplaza.com/trip/area/kanto/traffic/ai_traffic_prediction.html) 最終確認日 2020 年 07 月 12 日.
  - [16] 株式会社 NTT ドコモ. 東京湾アクアラインの渋滞を「ai 渋滞予知」で回避する一携帯電話ネットワークの運用データの統計処理に基づく未来予測の試み一. [https://www.nttdocomo.co.jp/corporate/technology/rd/technical\\_journal/bn/vol127\\_2/006.html](https://www.nttdocomo.co.jp/corporate/technology/rd/technical_journal/bn/vol127_2/006.html) 最終確認日 2020 年 07 月 12 日.
  - [17] 株式会社ナビタイムジャパン. 高速道路の渋滞予測 - navitime. <https://www.navitime.co.jp/highwaycongestion/prediction/> 最終確認日 2020 年 07 月 12 日.
  - [18] 首都高技術株式会社. 首都高の道路交通情報サイト mew-ti. <https://search.shutoko-eng.jp/> 最終確認日 2020 年 07 月 07 日.
  - [19] 首都高速道路株式会社. 首都高ドライバーズサイト. <https://www.shutoko.jp/use/safety/map/> 最終確認日 2020 年 07 月 07 日.