

非日常交通における車両感知器データ欠測を伴う路線の旅行時間予測 —機械学習によるアプローチ—

織田 利彦* (一般財団法人道路交通情報通信システムセンター)

A Machine Learning Based Approach for Travel Time Prediction on Roadways with Missing Vehicle Detector Data
in Unusual Traffic Flows

Toshihiko Oda* (Vehicle Information & Communication System Center)

VICS (Vehicle Information & Communication System) Center has provided drivers with useful traffic information. The travel time is usually generated based on the traffic flow data collected via vehicle detectors, and this service has offered excellent assistance and great convenience for drivers. Thus, the detectors are indispensable to observe and manage the traffic flows. However, owing to recent fiscal constraints, installation of detectors has been reduced significantly. Moreover, the renewal, repair and upgrade of the infrastructure has been delayed, which has led to degradation and malfunctioning of detectors. The degraded detectors have resulted in missing traffic flow data and lowered the quality of the service. Meanwhile, unusual traffic flows caused by large scale and special events have often occurred, and unexpected heavy congestion can be generated on the surrounding roads of events sites. To overcome these problems, we propose an innovative approach for compensation of missing detector data and prediction of travel time under unusual traffic flows by using a machine learning algorithm. Under the framework of the proposed approach, the training datasets are formed by sliding windows and generated by k-means clustering analysis, and a random forest model is applied to the prediction. In order to verify the effectiveness of the approach, we conducted an experimental study on arterial roads around Nakayama Racecourse on the days of the horse racing. The results show that our approach can predict travel time relatively accurately, which will lend strong support to enhance the information service provided by VICS Center.

キーワード : VICS センター, 車両感知器機能劣化, ランダムフォレスト, スライディングウィンドウ, k-means クラスタリング, コサイン類似度)

(VICS Center, degradation and malfunctioning of detectors, random forest, sliding window, k-means clustering, cosine similarity)

1. はじめに

道路交通情報通信システムセンター (Vehicle Information & Communication System Center, 以下 VICS センター) ⁽¹⁾はこれまで 20 年以上にわたり, 道路利用者に対する情報提供サービスを展開し, ドライバーの利便性向上に大きく寄与してきた。このサービスのなかで旅行時間ももっとも利用価値の高い情報の一つに挙げられ, 経路案内をはじめとしてその活用に対する要望は非常に高い。

これら旅行時間の多くは路上の車両感知器から収集される交通流データにもとづいて生成され⁽²⁾, 交通管制センターを通じて VICS センターに送信される。車両感知器は交通流監視を担い, 都市交通管理を支えるうえできわめて重要な役割を果たす機器で, 交通情報の収集拡大に向け, さらな

る配備が望まれるところであるが, 昨今の財政事情の悪化に伴い, 整備縮減, 更新遅延といった深刻な事態を迎えている。とりわけ, 更新遅延による交通流監視機能の劣化が危惧され, それに起因した交通流データの欠測, さらに付随して旅行時間情報の品質低下も懸念される⁽³⁾。

一方, 昨今の交通事情に目を向けると, 都市部を中心に大規模なイベント開催が著しく, 普段は混雑がみられない道路で交通集中による重渋滞を呈するという非日常的な事象が頻発している^{(4),(5)}。上述したように車両感知器の機能劣化による交通流データ欠測という深刻な事態が潜行するなか, 特異な交通事象における情報提供サービスへの期待が高まっているが, データ欠損時での対応は容易なことではない。

筆者は上述の社会インフラの機能劣化および頻発化する

非日常交通という切実な課題認識に立脚し、これまで船橋市の中山競馬場で開催される G1 (Grade 1) レース時を取り上げ、機械学習を適用した旅行時間の推計および予測方法を提案してきた^{(4)・(6)}。機械学習を扱うには学習データが必要となるが、定期的な変動が繰り返される日常交通流に比べ、競馬開催という特異日での交通流データはきわめて少なく、かつその変動もレースによって異なる。従前の研究では、G1 以外の競馬開催日の交通流データを学習に用いて検討を進めてきたが、両者に変動パターンの違いがみられ、良好な結果が得られていないケースも散見される。本論文では、これに対してスライディングウィンドウ法、クラスタリングによる学習データの抽出を試み、その効果を確認する。さらに、このアプローチにもとづき、車両感知器データがすべて欠測を伴うという路線での旅行時間を予測するとともに、その有効性を検証する。

2. 中山競馬開催および周辺路線概況

〈2・1〉 中山競馬開催状況 中山競馬場は、図 1 に示すように北方十字路交差点南側に位置し、周辺には競馬開催に備えた駐車場が散在している。競馬開催は 2017 年および 2018 年を例にとると、いずれも年間 41 日行われ、曜日には平日 1 日、土曜 20 日、日曜・祝日 (以下休日) 20 日である。また、開催期間は 1 月、2 月下旬から 4 月、9 月から 10 月初旬、12 月と年 4 期にわたる。このなかで、最高格付けとされる G1 レース (以下 G1) は 6 日開かれている。表 1 は 2017 年および 2018 年の G1 の開催日で、年末に開催されるホープフルステークス (レース 6) を除き、土曜、休日に行われている。表中、2017 年中山大障害および 2018 年有馬記念レースは祝日と重っている。

〈2・2〉 路線概況 中山競馬場周辺には、松戸市と市川市を南北につなぐ千葉県道 180 号 (以下松戸原木線) と、葛南地域と北総地域を東西に結ぶ同 59 号 (以下市川印西線) があり、両者は北方十字路交差点で交差している (図

表 1 G1 レース (2017 年・2018 年)

Table 1. Grade 1 horse racing 2017 and 2018.

Racing		Date		Day of the week
R1	Nakayama Grand Jump	2017	April 15	Saturday
		2018	April 14	Saturday
R2	Satsuki Sho	2017	April 16	Sunday
		2018	April 15	Sunday
R3	Sprinters Stakes	2017	October 1	Sunday
		2018	September 30	Sunday
R4	Nakayama Daishogai	2017	December 23	Saturday*
		2018	December 22	Saturday
R5	Arima Kinen	2017	December 24	Sunday
		2018	December 23	Sunday*
R6	Hopeful Stakes	2017	December 28	Thursday
		2018	December 28	Friday

* : National holiday

1)。本研究では交通量の多い前者を取り上げ、本郷町交差点から大柏橋交差点の区間を対象とする。この路線では 6 リンク (VICS リンク)、片側 1 車線で構成され、区間長は 3.8km である。路線の特徴として、北方十字路交差点以南を除いて幅員が狭く、さらに、近隣に京葉道路が控えていることから、大型車両の通行がきわめて多いことが挙げられる。とりわけ、北方十字路交差点では松戸原木線北進から市川印西線東進への流出および市川印西線西進から松戸原木線南進への流入交通が顕著である。

3. 競馬開催日における交通流変動

機械学習を用いた特異日における旅行時間予測に向け、学習データが必要である。学習データについては、交通管制センターで得られる交通流データ、すなわち旅行時間および車両感知器データが適切であるが、表 1 に示したように、G1 開催は年間 6 日で、かつ曜日もさまざまである。日常交通は基本的に定期的な変動のため、機械学習の適用時での学習データ選択は比較的容易である。これに対し、G1 開催時は日常交通と異なるため、対応したデータを抽出することは容易なことではない。筆者はこれまで G1 開催時における旅行時間が、G1 以外の競馬開催時 (以下 HR) および非開催時 (NHR) とどの程度異なるか、2017 年のデータを用い、同一曜日を対象に比較するとともに、その特徴を調べてきた^{(4)・(6)}。

図 2 は年間の HR、NHR における旅行時間 (5 分間値) について最大値および平均値の比較である⁽⁶⁾。図中、土曜日は 1~4 月、また、休日(1)、休日(2)はそれぞれ 1~4 月、9~12 月での値である。図から平均値は 5~15% の増加がみられる一方、最大値については違いが目立ち、HR では NHR の 1.6~1.8 倍に達している。後述するように、これは 16:00~17:00 付近での急増による。また、図 3、図 4 では G1、

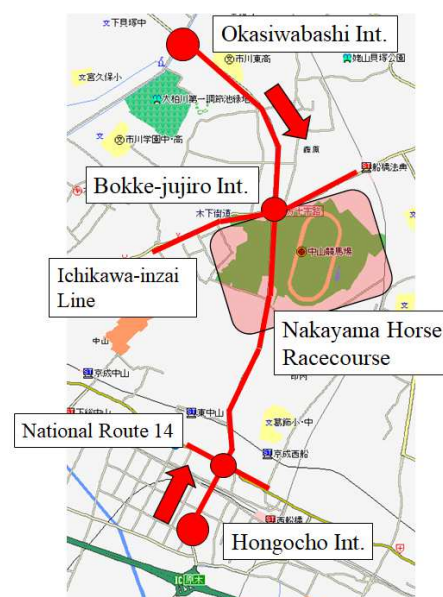


図 1 実験路線

Fig.1. Experimental roadway

HR, NHR の間で差異が顕著な R4, R5, R6 を取り上げた。このなかで唯一平日に開催された R6 については、比較対象となる HR が存在しない。このため、開催日を冬期休暇中

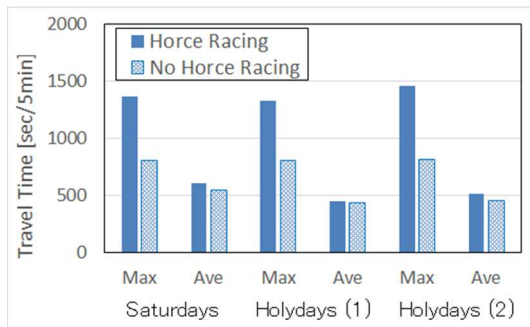


図 2 競馬開催・非開催日における最大・平均旅行時間比較

Fig. 2. Comparison of maximum and average travel time on Saturdays and holidays between days with and without racing events.

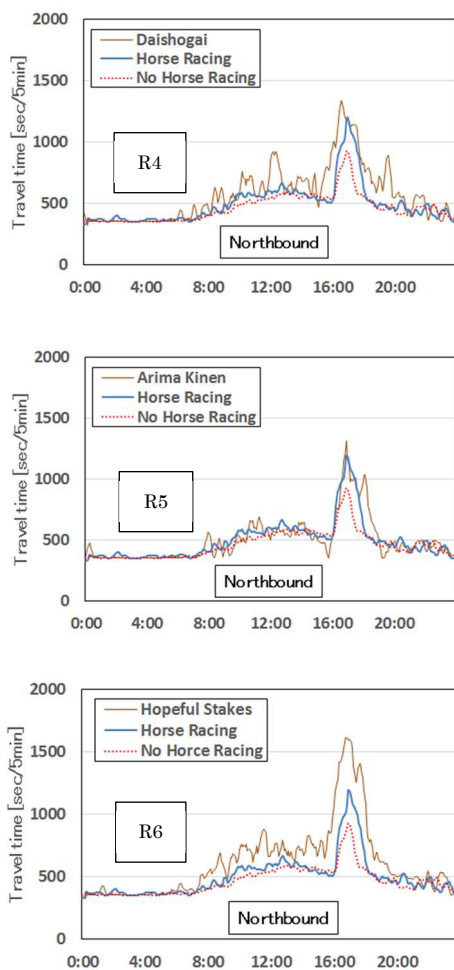


図 3 旅行時間比較 (北進)

Fig. 3. Comparison of travel time in the northbound on holidays between days with and without racing events.

とみなし、休日での HR データを用いた。以下に変動に関し、判明した点および推察を記す。

- (a) G1 および HR では競馬終了後の 16:00 過ぎに旅行時間が急増し、17:00 前後で最大値に達する。NHR にはこのような急激な変動はみられず、旅行時間の最大値も G1, HR に及ばない。
- (b) 南進では R4, R5, R6 と HR との変動パターンの乖離が著しい。これは R4, R5 は休日開催であるものの、曜日はそれぞれ土曜 (祝日), 日曜, R6 は木曜 (冬期休暇) で、日曜開催の HR での変動とは異なると推察される。なお、紙面の関係で掲載していないが, R1, R2, R3 での変動パターンは同一曜日開催の HR とおおむね一致している。

4. 実験

〈4・1〉概要 実験では特異交通事象の発生を G1 開催時とし、G1 以外では車両感知器データ、旅行時間が得られているとして検討を進める。車両感知器は北進、南進でそれぞれ 17 基、12 基設置され、これらから計測される交通量、占

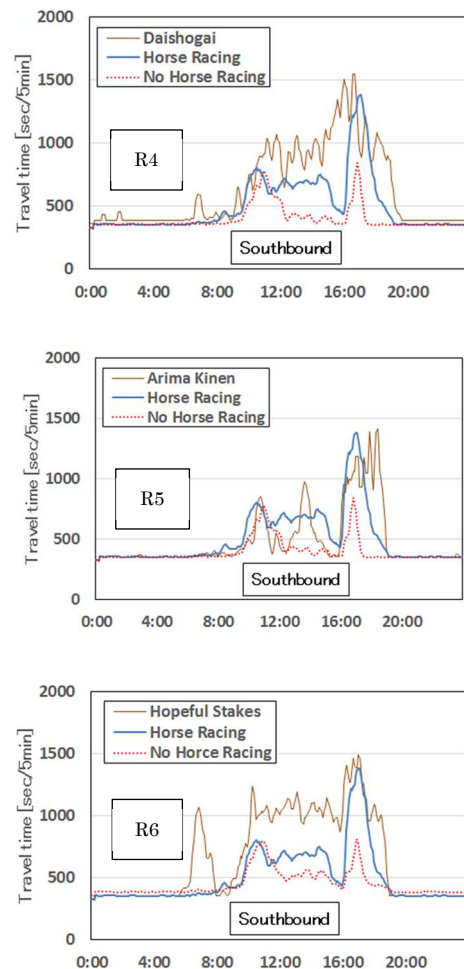


図 4 旅行時間比較 (南進)

Fig. 4. Comparison of travel time in the southbound on holidays between days with and without racing events.

有率にもとづいて旅行時間が生成されている。本研究では、G1 開催時にこの区間における車両感知器データがすべて欠損していると仮定する。これに対し、何らかの補填をし、それにもとづいて旅行時間推計を試みる。

いま、路線区間に隣接するリンクの交通流が区間内の交通流と密接に関連すると考える。すでに 2.2 で述べたように、区間への流入出交通が激しい交差点として北方十字路が挙げられる。この交差点は市川印西線と交差し、西進からの流入および東進への流出交通が顕著で、路線区間の交通流との関連性がきわめて高い。こうした実情を考慮し、大柏橋、本郷町の各交差点とともに、北方十字路交差点と接続する VICS リンク上の車両感知器から得られる交通流データを抽出する (図 5)。ここで、各流入出リンクに設置されている車両感知器は、路線区間北端の大柏交差点で流入 1 基、流出 3 基、南端の本郷交差点では流入 2 基、流出 3 基、さらに、北方十字路における市川印西線からの流入 3 基、市川印西線への流出 3 基で、実験ではこれらから得られる交通量および占有時間を入力変数として用い、旅行時間との関係づけを図る。

〈4.2〉学習データの抽出 競馬開催のような非日常時での交通流は対象となるデータ量が限られ、機械学習の適用にあたり、いかなるデータで学習させるか、その選択は重要な課題となる。筆者はこれまでレース開催と同一曜日の HR における終日の車両感知器データ (交通量, 占有率) を学習対象としたが、前掲の図 3, 図 4 からわかるように、旅行時間の変動パターンが異なる時間帯が散見され、機械学習を適用したものの、必ずしも良好な結果が得られていない。本研究では、この点に改良の余地があると考え、図 6 に示すようなスライディングウィンドウ法を適用し、旅行時間変動の特徴を抽出する。図中、 p , q はそれぞれウィンドウ幅, スライド幅である。以下にその処理を記す。なお、旅行時間, 交通量, 占有率はいずれも 5 分間値である。

- (a) 当該 G1 レース r の前日までの競馬開催日 d_i ($i = 1, 2, \dots, m$) の旅行時間データ π^{d_i} について、変動が伴う時間帯 (6:00~21:00) に対し、スライディングウィンドウによって分割し、ウィンドウ t における旅行時

間データ $\pi_t^{d_i}$ を取り出す。(図 7)。上記以外の時間帯 00:00~6:00, 21:00~24:00 についてはそれぞれ一つのウィンドウとみなす。

- (b) 抽出した旅行時間 $\pi_t^{d_i}$ についてウィンドウ単位に k-means 法を用いてクラスタリングを行うとともに、クラスター k の重心ベクトル π_t^k を取り出す。
- (c) 当日 r について、旅行時間生成が完了した時間帯における旅行時間 π_t^r を抽出し、 π_t^r と (c) で得られた π_t^k とのコサイン類似度を算出する。
- (d) (c) で得たコサイン類似度が最大となるクラスターに対応する日付 d_t^k を取り出す。さらに、 d_t^k における交通量, 占有時間を学習データとする。

上記処理をすべてのウィンドウに対して実施し、開催日 r における学習データを生成する。

〈4.3〉ランダムフォレストの適用 本研究では機械学習の手法として複数の決定木 (Decision Tree) から構成されるランダムフォレスト (Random Forest, 以下 RF) (7) を適用する。決定木は条件分岐を繰り返しながら樹形図に展開し、識別等を行う方法であるのに対し、RF ではこれを寄せ集め、各決定木の結果にもとづいて代表値を出力するもので、ここでは代表値として平均値を扱う。

実験では RF に対して 4.2 で得られた交通量, 占有率を

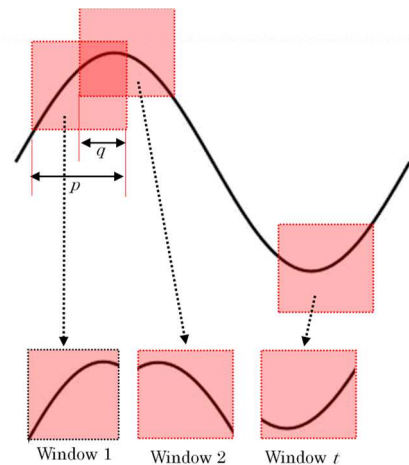


図 6 スライディングウィンドウ法
Fig. 6. Sliding window method.

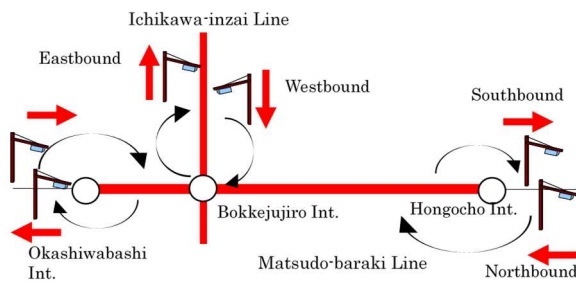


図 5 機械学習への入力に用いる流入出交通流
Fig. 5. Inflow and outflow traffic used for input of machine learning.

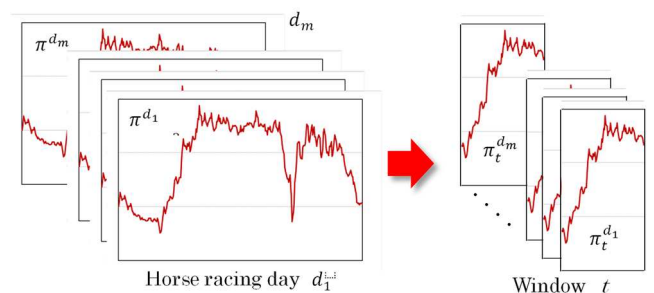


図 7 スライディングウィンドウによるデータ抽出
Fig. 7. Extraction of data with sliding windows.

入力データ，それに対応した旅行時間を出力データとして与え，双方を関連づけるとともに，G1 当日の交通量，占有率から旅行時間を推計および予測する。

〈4・4〉 実験結果 実験にあたり，第 1 ステップとして学習データの違いによる旅行時間の推計効果を検証した。図 8 は前述の G4, G5, G6 に関する推計結果である。図中，データセット A は従前の方法による学習データで，同一曜日開催の HR を取り出し，その平均値（5 分単位）としている。これに対し，データセット B は 4・2 に記した方法にもとづいて抽出したデータで，各々 RF を通じて学習し，旅行時間を推計した。評価にあたり，平方平均自乗誤差率（Root Mean Squared Error Ratio, 以下 RMSE 率）および相関係数は，VICIS センターから道路利用者に情報提供された旅行時間を実測値として算出している。この図から，データセット B にもとづいた推計によって RMSE 率は 20.4% 以下，相関係数は 0.93 以上となり，北進 R5 を除き，大幅に改善されていることがわかる。とりわけ，北進 R6，南進 R4 で著しい。また，図 9 は G1 のなかで唯一平日開催レース R6 について，推計値の違いを時系列変動で比較したものである。図中，推計旅行時間 A, B はそれぞれデータセット A, B を用いて算出した結果である。R6 は前述の通り，学習データの抽出が難しいが，データセット B によってレース終了直後の旅行時間の急増におおむね対応できていることがわかる。本アプローチによって RMSE は北進 13.5%，南進 8.7%，相関係数についてもそれぞれ 0.12, 0.06 の向上が確認された。

第 2 ステップでは効果が確認されたデータセット B を用

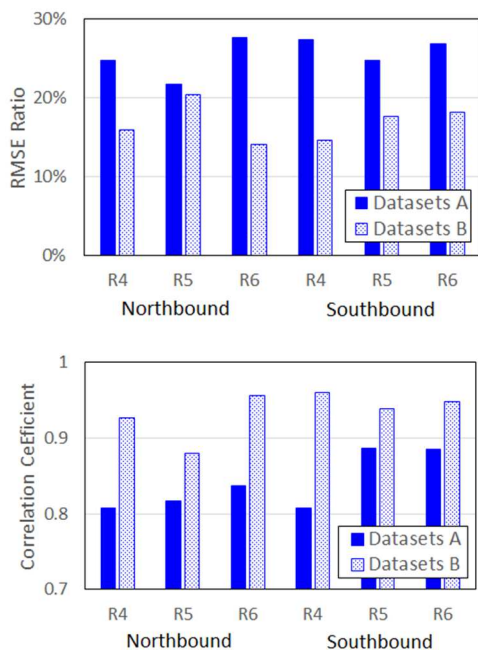


図 8 学習データの違いによる推計旅行時間の比較
Fig.8. Comparison of estimated travel times based on two different training datasets.

い，15 分先の旅行時間予測を試みた。方法は第 1 ステップと同様で，得られた予測値と実測値との比較を行った。結果

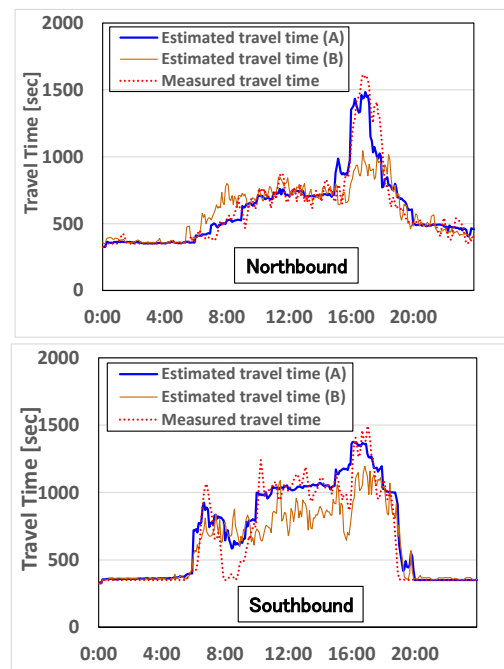


図 9 学習データの違いによる推計旅行時間の比較 (R6 開催時)

Fig.9. Comparison of estimated travel times on the day of Hopeful Stakes Horse Racing based on two different training datasets.

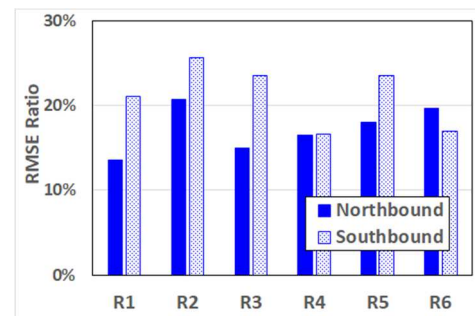


図 10 予測旅行時間の RMSE 率
Fig.10. RMSE ratio of predicted travel time.

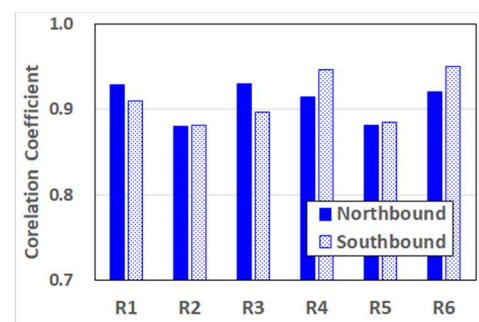


図 11 予測旅行時間の相関係数
Fig.11. Correlation coefficient of predicted travel time.

を図 10, 図 11 に示す。RMSE 率は北進 13.5~20.7%, 南進 16.6~25.7%, 相関係数は北進 0.88~0.93, 南進 0.88~0.95 という比較的良好な結果を得た。参考までに, データセット A を用いた場合と比較した場合, RMSE 率でみると, 北進 8.3~14.1%, 南進 0.8~11.2%の精度向上が確認され, とりわけ, R6 では両方向とも改善がもっとも大きい。

一方, 推計時を含め, RMSE 率 20%以上となったケースとして南進 R2, R3 が挙げられる。図 12, 図 13 に予測旅行時間の変動を示す。両者とも祝日と重ならない日曜開催のレースで, 図から見てわかるように, その特徴は交通集中が発生する時間帯 16:00~18:00 で旅行時間(実測値)が急増し, その直後に急減する点にある。HR を含め, 他レースにおいてもこうした傾向がみられるものの(図 3, 図 4 参照), R2, R3 ではその差が顕著で, いずれも交通集中時間帯の前後 30 分間で 1,000 秒以上の旅行時間増減が発生する。予測ではこの激しい変動に対する追従がやや不十分で, ピーク時で実測値との乖離が生じている。今回は学習データセットについて新たな生成方法を提示し, それにもとづいて推計および予測を行った結果, 南進 R2, R3 を除く他のレースでは精度が向上し, 好ましい対処ができたものの, 南進 R2, R3 でみられる増減が激しい特異な時間帯への対策につ

いては検討の余地を残しており, 今後の課題としたい。

4. おわりに

一般道路における旅行時間予測は 30 年以上前から検討されてきたテーマ^{(2),(4)}で, 研究者とともに道路利用者からの関心もいまなお高い。しかしながら, この間, 研究を取り巻く環境は大きく変化した。昨今, 財政事情の悪化を背景に交通インフラの機能劣化の進行という深刻な事態に陥り, その一方で大規模イベント等に伴う特異な交通事象が頻発するという課題を抱えている。本研究ではこうした憂慮すべき実態を背景として論及したもので, 機械学習を適用し, 旅行時間予測について検討を行った。実験で用いたデータは 2017 年に限られ, 必ずしも十分とは言えないが, 本アプローチによって上述の状況下における旅行時間予測の可能性をおおむね見出すことができた。今後の課題として南進 R2, R3 で観測されるように, 短時間で旅行時間変動幅がきわめて大きいケースへの対処が挙げられる。

最後に, 本研究にあたり, 機械学習に関して(株)応用数理研究所 代表取締役 佐々木俊久氏に多大な協力をいただいた。また, 千葉県警察本部交通部交通規制課管理官 鶴澤政幸氏からは交通情報収集および提供に関し, 有益なご意見を賜った。ここに記して深甚なる謝意を表します。

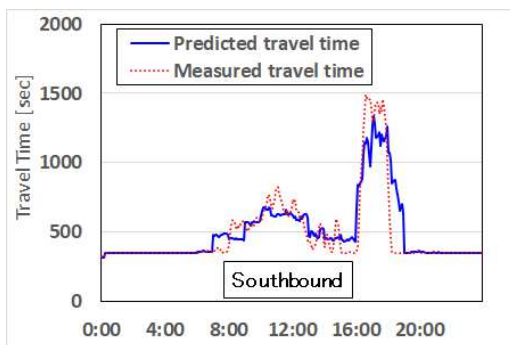


図 12 R2 における予測旅行時間変動(南進)

Fig. 12. Fluctuation of predicted travel time in the southbound on the day of Satsuki-sho Horse Racing.

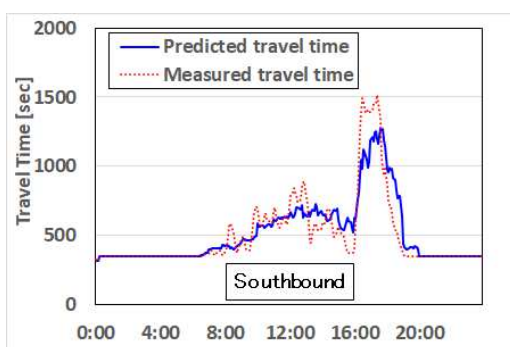


図 13 R3 における予測旅行時間変動(南進)

Fig. 13. Fluctuation of predicted travel time in the southbound on the day of Sprinters Stakes Horse Racing.

文 献

- (1) <https://www.vics.or.jp/>
- (2) T. Oda, "Prediction of Travel Time on Arterial Roads Using Vehicle Detector Information", *Trans. of ISCIE*, Vol. 10, No. 12, pp. 647-655 (1997) (in Japanese)
織田利彦: 車両感知器情報を用いた幹線道路における旅行時間予測, システム制御情報学会論文誌, Vol. 10, No. 12, pp. 647-655 (1997)
- (3) T. Oda, "A Machine Learning Based Approach for Complementing Missing Traffic Flow Data in an Urban Roadway Network", *The Paper of Technical Meeting on ITS, IEE Japan*, ITS-18-2, pp.5-10 (2018) (in Japanese)
織田利彦: 機械学習による交通流データ欠測時における補填, 電気学会, 電気学会研究会資料, ITS 研究会, ITS-18-2, pp.5-10 (2018)
- (4) T. Oda, "Travel Time Prediction for a Large-scale Event Based on Machine Learning", *The Paper of Technical Meeting on ITS, IEE Japan*, ITS-18-14, pp. 15-20 (2018) (in Japanese)
織田利彦: 機械学習による大規模イベント開催時における旅行時間予測, 電気学会, 電気学会研究会資料, ITS 研究会, ITS-18-14, pp. 15-20 (2018)
- (5) T. Oda, "A Machine Learning Based Approach for Prediction of Travel Time in Traffic Singularity Event in an Urban Roadway Network", *The Paper of Technical Meeting on ITS, IEE Japan*, ITS-19-5, pp. 17-22 (2019) (in Japanese)
織田利彦: 機械学習による特異な交通事象における旅行時間予測, 電気学会, 電気学会研究会資料, ITS 研究会, ITS-19-5, pp. 17-22 (2019)
- (6) T. Oda, "An Ensemble Learning Based Approach for Travel Time Estimation in Roadways with Missing Detector Data in a Traffic Singularity Event", *The Paper of Technical Meeting on ITS, IEE Japan*, ITS-19-20, pp. 43-48 (2019) (in Japanese)
織田利彦: 特異な交通事象における車両感知器データ欠測に対する旅行時間推計 - アンサンブル学習によるアプローチ -, 電気学会, 電気学会研究会資料, ITS 研究会, ITS-19-20, pp. 43-48 (2019)
- (7) 平井 有三: 「はじめてのパターン認識」, 森北出版 (2012)