

商用車プローブデータを用いた 異常な車両挙動の抽出手法に関する検証

古川 知史¹・安藤 宏恵²・柿本 竜治³

¹学生会員 熊本大学 工学部土木建築学科 (〒860-8555 熊本市中央区黒髪2丁目39-1)
E-mail: 184t1230@st.kumamoto-u.ac.jp

²正会員 熊本大学助教 くまもと水循環・減災研究教育センター (〒860-8555 熊本市中央区黒髪2丁目39-1)
E-mail:hiroeando@kumamoto-u.ac.jp

³正会員 熊本大学教授 大学院先端科学研究部 (〒860-8555 熊本市中央区黒髪2丁目39-1)
E-mail:kakimoto@kumamoto-u.ac.jp

自然災害や交通事故等によって引き起こされる異常な挙動を示す車両を検知する手法として、従来の設置型センサーに比べて広範囲かつ道路を特定せずデータを取得できる観点から、プローブデータの活用が注目されている。本研究では、マップマッチングを必要としないドット形式の商用車プローブデータを用いて、移動軌跡データを文字列に変換し類似度を算出することにより、日常と異なる挙動を示す車両の検知法を提案する。異常挙動が発生しやすい状況として実際の高速道路通行止め時と車線規制時に提案手法を適用した結果、5次メッシュの文字列変換に基づくレーベンシュタイン距離による文字列類似度により異常な挙動を高精度で検出できることを示した。

Key Words : probe data, expressway, string similarity, anomaly detection

1. はじめに

自然災害や交通事故、道路工事等によって突発的または計画的に車の流れが妨げられる“交通異常”は度々発生し、それに伴い異常渋滞や日常と異なる傾向の交通状況が引き起こされる。その結果、所要時間の増大による経済損失や環境汚染、緊急車両の通行の妨げ等につながる恐れがある。そのため、異常な挙動を示す車両の存在とその場所をいち早く検知し、何らかの対応を施すことは非常に重要である。

交通異常の状態把握や検知に関する研究は従来より多数行われてきた。Kawasaki et al.¹⁾は、道路にあらかじめ設置された車両感知器等のセンサーを用いて、交通量や密度、車線別の速度差等の統計量に関する平常時と交通障害発生時での差異を分析し、異常を判定する閾値を設定した。設定された閾値をもとに、新たに得られた車両感知器データから異常判定を行い、交通異常検知を試みている。また、川崎・西山²⁾や鈴木・中村³⁾は、映像監視システムの映像に画像処理技術を適用し、交通流解析や事故箇所検知を行っている。しかし、車両感知器や映像監

視システム等の設置型センサーは、特定の道路に関するデータしか得ることができない点や、センサーの設置とその維持に費用と人員が必要であるという問題点を抱えている。これに対し、交通異常を検知する手段として近年注目されているのが、プローブデータである。プローブデータは、車両に搭載されたセンサーから位置情報や速度等の情報を収集したデータであり、必ずしも道路ネットワークデータに吸着させる必要がない。従来の設置型センサーに比べ対象とする道路に制限がなく、広範囲に細街路を通る車両まで捉えることができる。さらに個々の車両ごとに挙動を追うことができるという利点を持つ。そこで本研究では、車両軌跡データを用いて挙動の変化を捉え、交通異常をいち早く検知する手法を提案する。

本稿の構成は以下の通りである。2.ではプローブデータを用いた異常検知に関する既往研究を整理し本研究の位置付けを述べる。3.ではプローブデータによる異常挙動車両抽出のための手法を構築し、4.では用いるデータの概要について説明する。5.では、構築した手法を実際の高速道路通行止め期間の車両に適用し、異常検知の精

度を評価する。6.では、5.における評価結果から最も高精度な検知手法を用いて、車線規制が行われた実際の交通状況に適用し、有効性を検証する。

2. 本研究の位置付け

災害によって引き起こされる異常挙動に関して、吉田ら⁴⁾は、2016年に発生した熊本地震時の道路損傷箇所を対象として、三次元プローブ軌跡データによる交通障害検出手法を提案した。時間的・空間的な差異を検出するため車両軌跡を文字列に変換し、その類似度を評価するレーベンシュタイン距離の算出により、交通障害箇所を検出可能であることを示した。川崎ら⁵⁾は、西日本豪雨時の迂回行動を対象に、商用車プローブデータを用いた機械学習により災害時の迂回路を効率的に検出する手法を提案している。豪雨時の迂回行動を誘因データと併せて分析することにより、効率的でない迂回行動にはリスクが一因となっていることを明らかにした。その他にも、新潟県中越沖地震⁷⁾や東日本大震災⁸⁾、2013年台風18号⁹⁾など具体的な災害時の行動に関するプローブデータを用いた分析も多く行われている。

検知した交通異常を情報提供に活かす検証として、飯島ら¹⁰⁾はプローブデータを利用した「エリア交通流動性情報」の実用化を目指し、エリア混雑指数やエリア変化指数といった指標をもとに異常を定量的かつ視覚的にとらえる方法を提案している。エリア交通流動性情報では異常気象や災害などの非日常的な交通状況が発生した際に、避難や生活物資運搬、被災者支援などのトリップの妨げになるような望ましくないトリップを誘起させないよう、交通利用者に向けてメディアを通じて的確に情報を配信することを主目的としている。

また、全国規模の広域的な分析として日下部⁶⁾は長期的かつ全国的に観測されたデータを用いて、メッシュ単位の旅行速度変動パターンから非日常を示す日を検出する方法を構築した。豪雪などの災害や長期休暇等が及ぼす広域的な影響を検知できている。

Li and McDonald¹¹⁾はプローブ車両の平均旅行時間と隣接する時間間隔の時間差という2つの特徴量を用いた二変量解析モデル (BEAM) により、交通事故等の突発的なアクシデントに対する交通異常検知の手法を提案している。このモデルは、リンク間の旅行時間は需要増加よりも容量低下の影響が大きいという前提に基づいている。交通異常が発生した場合、長いリンクよりも短いリンクほど旅行時間の増加が大きくなる傾向があり、リンク長が検出精度に影響を与えている可能性が示唆されている。Cheu et al.¹²⁾は、高速道路上の事故検出アルゴリズムを開発し、現在の旅行時間と対応する曜日や時間帯における

過去の旅行時間の平均値を比べることで交通異常を検出している。MOSESではプローブ車両混入率を50%とした場合、MLFおよびCPNNという2つのニューラルネットワークと同等の検出率と誤認識率を達成し、平均検出時間が短いという結果が得られた。Asakura et al.¹³⁾はGPS機器を搭載した車両から収集したデータを用いて、交通事故発生時に引き起こされる交通渋滞の発生時刻と位置を予測するアルゴリズムを開発している。さらに交通流シミュレーション解析により、開発アルゴリズムの精度を担保するプローブ車両混入率を明らかにした。

渋滞に着目した分析では、松平ら¹⁴⁾はETC2.0プローブデータを利用し、渋滞先頭における速度回復距離、渋滞先頭加速度、交通相転移時間、下流交通相転移時刻差といった複数の特徴量を組み合わせて交通異常を検知するアルゴリズムを開発した。合計400件の渋滞件数をもとに、1件をテストデータ、残りを学習データとするLeave-one-out法によって検知実験を実施し、適合率81.0%、再現率85.0%と概ね良好な結果が示された。

このように、プローブデータを用いた交通異常検知に関する研究は災害や交通事故などのインシデントに基づくものから、長期的な観測による広域分析など多岐にわたる。しかしながら、データの前処理に時間のかかる場合や、判定基準が感覚的に設定されているなど課題も多い。また、混入率が低いプローブデータの特性上、旅行時間に関する特徴量を用いる場合には、その精度はプローブ車両台数に大きく依存し、特に長距離になればなるほど十分なサンプル数を確保することが困難であることも課題¹⁵⁾となっている。これらの問題に対処するため、自動ナンバープレート認識 (ANPR) 技術を用いて必要なサンプル数を増やすといった方法を取り入れているケース¹⁶⁾もあるが、プローブデータの欠点を補うことができる反面、広範囲のデータを低コストで得られる利点を生かすにくいデメリットもある。

以上の背景を踏まえ本研究では、ネットワークデータを必要とせず、プローブデータから得られるドット形式のみのデータを用いて、災害や交通事故などによって引き起こされる異常挙動車両を抽出する手法を提案する。吉田ら⁴⁾が提案した文字列類似度算出により異常を検知する手法をもとに、平常時に蓄積されたデータを用いて基準を設けることで、災害時のみではなく様々な要因に基づく異常挙動を対象に出来るようにする。既往研究の文字列の類似度は、交通障害が発生する前 (平常時) の車両1台のみを基準に設定し算出され、選定した1台が検出結果に大きな影響を及ぼす恐れがある。これに対し本研究では、平常時の車両軌跡の類似度の分布を調べ、多数の車両挙動を考慮した基準の設定を試みる。これにより、特定の車両軌跡に依存しない異常検知手法の提案を目指す。また、ドットデータを扱うにあたり、空間的メ

ッシュにデータを当てはめる方法では、あらかじめ設定するメッシュサイズが精度に大きな影響を与える。適切なメッシュサイズの検討を行うため、設定するサイズを変化させた際の検知精度を評価する。さらに、自然言語処理分野で発展する文字列類似度指標であるレーベンシュタイン距離¹⁶⁾の指標とジャロ・ウィンクラー距離¹⁷⁾の指標について、それらの指標の交通状況分析に対する適性を検証する。

3. 交通異常の検知法の構築

(1) 文字列の類似度算出による交通異常検知の考え方

交通異常が発生した場合、車両の軌跡に平常時との差異があらわれる。この差異を定量的に評価することができれば、交通異常の発生有無や生じている箇所の検知が可能となる。そこで本研究では、交通障害が発生する前（平常時）と交通障害発生時の車両挙動の軌跡をそれぞれ文字列に変換し、その類似度を算出することで交通異常の検知を試みる。

(2) 使用するプローブデータの概要

本研究で使用したプローブデータは、富士通交通・道路データサービスより提供されたデータであり、事業用トラックに限定し収集されている。その概要¹⁸⁾を図-1に示す。ネットワーク型デジタルタコグラフの情報をもとに道路整備・安全向上や物流の高度化などを目的に、事業者の許諾を得て収集・蓄積し、営業情報・企業情報を秘匿化したものである。本研究では、1秒ごとに記録されたドットデータ形式の軌跡データを使用しており、ネットワークのデータの準備やマップマッチング処理が不要であるため、細街路を含めた全ての車両を対象とした評価の高速化が可能である。また、車両ごとに特定のIDが付いているため、同一車両の重複計算を防ぐことや特定車両の経路変化などを把握することが可能である。車両IDデータに加えて、本研究では記録日時、緯度・経度、車測速度のデータ項目を用いて分析を行う。

(3) 車両挙動の文字列変換方法

設定した対象範囲に対し、1メッシュにつき1文字が付与された四方メッシュと移動軌跡データを重ね合わせる。その結果、1秒ごとの車両ドットデータに、各メッシュに対応した固有の文字が割り当てられ、車両の挙動を文字列に変換することができる。割り当てのために用意した文字は、JIS第1水準¹⁹⁾の漢字2,965字、JIS第2水準の漢字3,390字、ひらがな・カタカナがそれぞれ71字、アルファベットが大文字、小文字それぞれ26字の計6,549字である。これらを分析対象範囲のメッシュの総数に応じて必要な

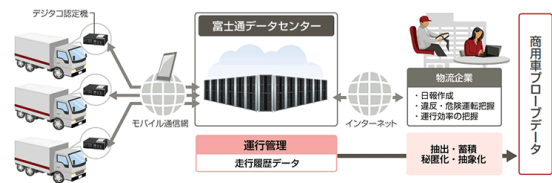
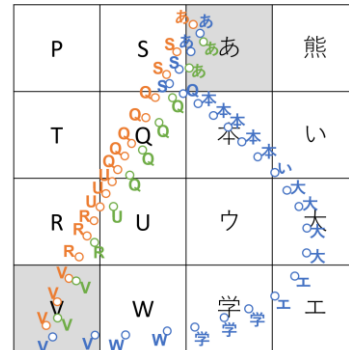


図-1 商用車プローブデータの概要¹⁸⁾

出発地メッシュ：あ
目的地的メッシュ：V



- 平常
ああQQQQURVV
- 速度低下
あSSQQQQUUURRRVVV
- 迂回
ああSSQ本本本本本本大大大大エ工学学学WWWV

図-2 ドットデータの文字列変換イメージ

文字数だけ使用する。類似度算出を行う際に文字化けが発生してしまう恐れがあるため、JIS第3水準、JIS第4水準の使用は避けることにした。また、ひらがな・カタカナについては清音、濁音、半濁音を用いることとし、小文字表記される拗音は用いていない。

図-2に観測間隔を Δt とした場合のプローブ車両軌跡の文字列変換イメージを示す。平常時、速度低下時、迂回時のプローブ車両軌跡の変換イメージを示している。観測間隔 Δt が一定であるため、文字列の長さの違いで時間的差異、文字の種類の違いで空間的差異を表すことができる。

(4) 文字列間の類似度指標

変換した文字列を用いて類似度算出を行う指標について検討する。本研究では「レーベンシュタイン距離」と「ジャロ・ウィンクラー距離」という2種類の文字列類似度指標を用いる。

a) レーベンシュタイン距離の算出方法

レーベンシュタイン距離¹⁶⁾の算出手順は、ある1つの文字列を他の文字列に変換する際に、文字列の「編集」（文字の挿入、削除、置換）を行う度にコスト1を追加する作業を行う。この際、「編集」の回数が最小になるように文字列を変換したときのコストを求める。これが、レーベンシュタイン距離である。本研究では2つの文字列のうち、長い方の文字列の長さで通常のレーベンシュタイン距離を除して「標準化」している。そのため、取

り得る値が0~1となり、0に近づくほど2つの文字列類似度が高いとされる。図-3に文字列 s_1 を文字列 s_2 に変換するレーベンシュタイン距離の算出例を示す。同図の例では、「編集」の最小回数は3であり、標準化されたレーベンシュタイン距離は0.75となる。既往研究²⁰より、レーベンシュタイン距離による軌跡文字列間の類似度算出方法を定式化したものを式(1)に示す。

長さ n, m の2つの軌跡文字列を、ベース $s_1 = \{\alpha_0, \alpha_1, \dots, \alpha_n\}$ 、ターゲット $s_2 = \{\beta_0, \beta_1, \dots, \beta_m\}$ と表す。また、 $\alpha_i (0 \leq i \leq n)$ を s_1 中の i 番目の軌跡文字、 $\beta_j (0 \leq j \leq m)$ を s_2 中の j 番目の軌跡文字とする。2つの軌跡 s_1, s_2 間のレーベンシュタイン距離 $L(m, n)$ は、以下の漸化式で定義される。

$$L(m, n) = \begin{cases} \max(i, j) & \text{if } \min(i, j) = 0 \\ \min \left(L(i-1, j-1) + c(\alpha_i, \beta_j), \right. \\ \left. L(i, j-1) + 1, L(i-1, j) + 1 \right) & \text{otherwise} \end{cases} \quad (1)$$

where,

$$c(\alpha_i, \beta_j) = \begin{cases} 0 & (\alpha_i = \beta_j) \\ 1 & (\alpha_i \neq \beta_j) \end{cases} \quad (2)$$

$0 \leq i \leq n, 0 \leq j \leq m$

ここで、 $c(\alpha_i, \beta_j)$ は文字の置換操作コストである。

b) ジャロ・ウィンクラー距離の算出方法

ジャロ・ウィンクラー距離は、ある文字列と別の文字列の間で一致する文字数と、その文字の置換の要不要から距離を計算する。算出手順¹⁷⁾は、まず式(3)からJaro距離 d_j を計算する。その後、接頭辞を考慮した式(4)によってジャロ・ウィンクラー距離 d_w を算出する。ここで、式(3)の w_1 は文字列 s_1 の重み、 w_2 は文字列 s_2 の重み、 w_t は s_1 を s_2 に置換する際にかかる重みを調整するための定数であり、それぞれ $w_1 = w_2 = w_t = 1/3$ と設定するのが一般的である。 d は文字列 s_1 の文字数、 r は文字列 s_2 の文字数、 c は2つの文字列の一致する文字数、 τ は s_1 を s_2 に編集する際に、置換が必要な文字数を2で割った数値である。また、式(4)における l は2つの文字列間の接頭辞の一致数であり、最大値は4と定められている。取り得る値は0~1で、1に近づくほど2つの文字列類似度は高い。

$$d_j = \begin{cases} 0 \\ w_1 \cdot \frac{c}{d} + w_2 \cdot \frac{c}{r} + w_t \cdot \frac{(c - \tau)}{c} \end{cases} \quad (3)$$

$$d_w = d_j + l \times 0.1(1 - d_j) \quad (4)$$

4. 対象地域と使用データの概要

(1) 分析対象の交通規制

阪神高速道路では、高速道路リニューアルプロジェクトの一環として、2019年11月に4号湾岸線および6号大和

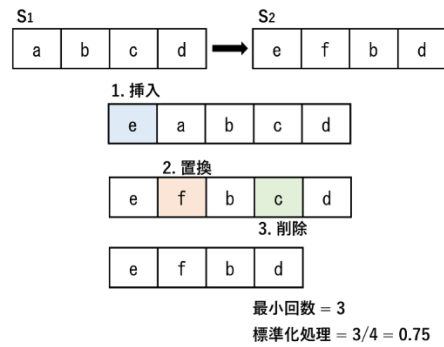


図-3 レーベンシュタイン距離の算出例



図-4 工事の概要²¹⁾

川線において、車線規制と通行止めを伴うリニューアル工事が実施された。湾岸線は、阪神高速道路の1日平均利用台数72万台のうち、10万台が利用²²⁾する重要な役割を担う路線であり、特に、大型車両の利用率が高い道路ということが知られている。工事の概要を図-4に示す。

2019年11月6日(水)~20日(水)までの14日間は4号湾岸線(南港北~三宝)の延長4.7km区間において車線規制が行われた。その後、11月20日(水)~30日(土)までの10日間は4号湾岸線(南港北~大浜)の延長8.8km区間と、6号大和川線(三宝~鉄砲)の延長1.4km区間において終日完全通行止めとなった。車線規制期間は車両の通行が可能であったが、通行止め期間は、区間中の出入口も閉鎖され通行は不可能であった。

(2) 対象ODと分析対象範囲

阪神高速4号湾岸線リニューアル工事による車両軌跡の変化を捉えるため、対象ODペアをO:中島出入口-D:助松JCTに設定し、ピークアワーの出勤時間帯のうち午前8時台に中島出入口の通行記録がある車両を対象とした。通行規制が行われる前の2019年10月24日(木)~30日(水)の対象車両は138台、車線規制期間中の11月7日(木)~13日(水)の対象車両は40台、通行止め期間中の11月21日(木)~27日(水)の対象車両は30台である。比較日数を同一とするため、それぞれ1週間ずつを対象期間とした。分析対象範囲は図-5示すように、通行止め区間および対象ODの各地点が含まれる範囲に設定している。

(3) 使用するメッシュサイズ

メッシュとは地域を網目状で区切ったものであり、国

勢調査や経済センサス調査等では、地域メッシュ²⁾が基準として利用されている。メッシュを用いた集計は地域を隙間なく統一された面積の単位で表現できる。さらに、位置関係の明確な表現、一貫した過去のデータの蓄積ができる長所を持っている。そのため、都市化による地域の人口変化や事業所の変化予測だけでなく、気象データの分析等の様々な分野で活用されている。メッシュを用いた分析では、サイズや集計期間に注意する必要があり、適用メッシュのサイズに依存して結果が変化し得る可能性がある。このような特徴から、本提案手法における適切なメッシュサイズの検討を行うために、1km四方の3次メッシュ、500m四方の4次メッシュ、250m四方の5次メッシュをそれぞれ適用し、その精度を評価する。

5. 通行止め期間における提案手法の適用

本章では、3.で提案した手法を4.で述べた阪神高速4号湾岸線リニューアル工事の対象範囲に適用し、通行止め期間中の異常な挙動を示す車両が検知可能であるかを検証する。この際、提案手法を5次メッシュ、4次メッシュ、3次メッシュの順にそれぞれ適用し、結果の精度を評価する。

(1) 類似度算出のための基準設定

交通異常を検知するにあたり、通行規制が行われる前（平常時）の車両軌跡から異常判定を行うための基準を設定する必要がある。まず、通行規制が行われる前の2019年10月24日（木）～30日（水）の対象車両138台に対して、車両挙動を1台ずつ文字列に変換する。その後、138件の文字列を総当たりで計算し、 ${}_{138}C_2 = 9453$ 件のレーベンシュタイン距離とジャロ・ウィンクラー距離の文字列類似度をそれぞれ算出する。5次メッシュに適用した際の車両挙動の結果の散布図を図-6、図-7にそれぞれ示す。5次メッシュの通行規制前（9453件）のレーベンシュタイン距離の平均値は0.324、標準偏差は0.166、ジャロ・ウィンクラー距離の平均値は0.877、標準偏差は0.240であった。それぞれの確率分布図は図-8に示すようになり、これらを通行規制が行われる前の車両軌跡として、異常判定を行う際の基準に設定した。

(2) 異常判定のための閾値の設定

通行止め期の対象車両30台について1台ずつ異常判定を行う。通行規制前の文字列138件に対して総当たりで計算し、同様にレーベンシュタイン距離とジャロ・ウィンクラー距離の平均値と標準偏差から車両1台あたりの確率分布図を作成し、前節で得た基準の確率分布図と重ね合わせる。任意の1台におけるレーベンシュタイン距



図-5 対象ODと分析対象範囲

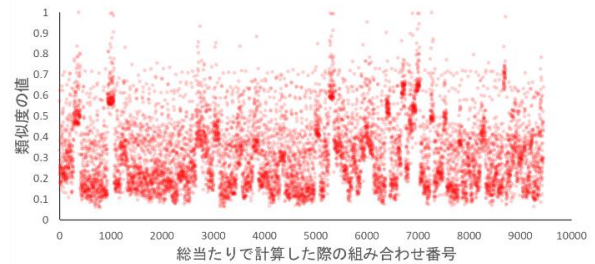


図-6 Le5th 平常時の散布図

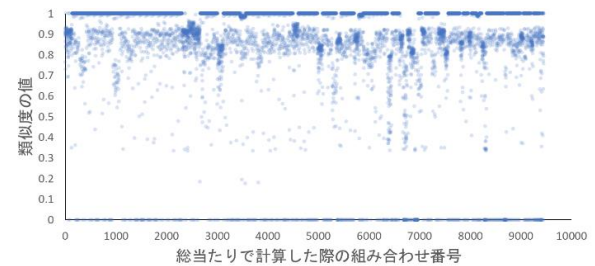


図-7 Ja5th 平常時の散布図

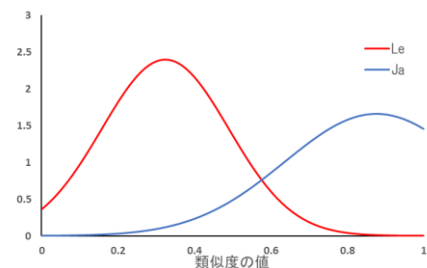


図-8 平常時の確率分布 (Le5th, Ja5th)

離とジャロ・ウィンクラー距離の重ね合わせた例を図-9、図-10にそれぞれ示す。これらの類似度を式(5)、(6)のユークリッド距離によって1台ずつ数値化する。

$$d = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (5)$$

$$d' = \frac{1}{1+d} \quad (6)$$

ユークリッド距離の度合いを調べるために式(6)で逆数を取り、算出される値を0～1とし、分布の完全一致によってユークリッド距離 d が0になる場合を考慮し、分

母に1を加えた形とした。1に近づくほど2つの確率分布図の類似度は高くなる。

これらを踏まえ、通行止め期の車両に対する異常判定の閾値を設定する。本研究では、ユークリッド距離0.10と0.05の2段階の閾値を設定し、閾値以下となる挙動を異常であると判定する。また、ここでは通行規制が行われる前と通行止め期の車両軌跡を比較しているため、通行止めの影響で30台がすべて異常挙動と判定されるという仮定のもと異常検知手法の検証を行う。

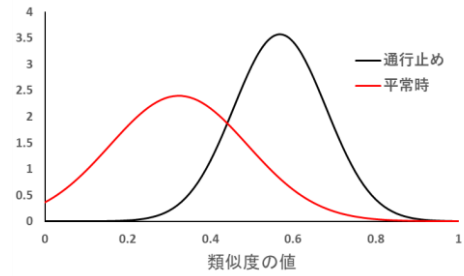


図-9 Le.5th 基準分布と重ね合わせた例

(3) 異常検知のための類似度指標の適性検証

2つの類似度指標を用いて、250m 四方サイズの5次メッシュに対して提案手法を適用し、類似度指標の適性を検証した。ユークリッド距離に基づく閾値別の判定結果を表-1に示す。×が異常判定されたもの、△が異常判定されなかったものになっている。また、その検知台数と検知率をまとめたものを表-2に示す。レーベンシュタイン距離の場合、閾値0.10に関しては分析対象車両30台すべてに対し異常挙動であると判定でき、閾値0.05に関しては、検知台数が22台で検知率は73.33%という結果になった。一方で、ジャロ・ウィングラー距離の場合、閾値0.10に関しては検知台数が2台で6.67%、閾値0.05は検知台数1台で3.33%と数台のみの検知に留まり、異常挙動を示す車両検出のための類似度評価指標として適当ではないことが分かった。

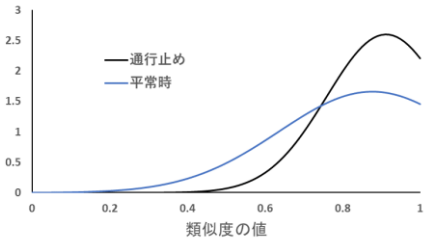


図-10 Ja.5th 基準分布と重ね合わせた例

表-1 5次メッシュに適用した判定結果

No.	Le.Eu.5th	閾値		Ja.Eu.5th	閾値	
		0.10	0.05		0.10	0.05
1	0.0616	×	△	0.1816	△	△
2	0.0619	×	△	0.1650	△	△
3	0.0319	×	×	0.5636	△	△
4	0.0293	×	×	0.0235	×	×
5	0.0451	×	×	0.2482	△	△
6	0.0500	×	△	0.2871	△	△
7	0.0398	×	×	0.7035	△	△
8	0.0374	×	×	0.3756	△	△
9	0.0284	×	×	0.3666	△	△
10	0.0694	×	△	0.1558	△	△
11	0.0398	×	×	0.6303	△	△
12	0.0535	×	△	0.2192	△	△
13	0.0429	×	×	0.3837	△	△
14	0.0288	×	×	0.4559	△	△
15	0.0392	×	×	0.6676	△	△
16	0.0271	×	×	0.4521	△	△
17	0.0385	×	×	0.6815	△	△
18	0.0394	×	×	0.7331	△	△
19	0.0374	×	×	0.6114	△	△
20	0.0442	×	×	0.3044	△	△
21	0.0459	×	×	0.2517	△	△
22	0.0319	×	×	0.5625	△	△
23	0.0394	×	×	0.6884	△	△
24	0.0494	×	×	0.2570	△	△
25	0.0642	×	△	0.1727	△	△
26	0.0505	×	△	0.1850	△	△
27	0.0536	×	△	0.0577	×	△
28	0.0384	×	×	0.5763	△	△
29	0.0372	×	×	0.5177	△	△
30	0.0374	×	×	0.5092	△	△

表-2 5次メッシュにおける検知台数と検知率

		検知台数	検知率
Le.Eu.閾値	0.10	30 /30台	100%
	0.05	22 /30台	73.33%
Ja.Eu.閾値	0.10	2 /30台	6.67%
	0.05	1 /30台	3.33%

本研究では車両軌跡の変化を捉えるために、対象ODを設定した分析を行った。全てのデータが同じOから出発しており文字列における接頭辞が必ず一致するため、接頭辞が大きな影響を持つジャロ・ウィングラー距離指標では、同一OD内の軌跡の違いが上手く表現されず、比較した2つの文字列の類似性が高く評価されたことが、異常判定が数件に留まった要因であろう。

(4) 適用するメッシュサイズの検証

本提案手法は、メッシュサイズが小さいほど細やかな軌跡を表現可能であるため、正確な異常判定につながり、メッシュサイズが大きくなるほど判定精度が粗くなるという特徴を持つ。そこで、適切なメッシュサイズの検討を行うため、5次メッシュより1段階メッシュサイズが大きな500m 四方サイズの4次メッシュに対して、交通異常検知への有効性が確認されたレーベンシュタイン距離指標を適用する。

4次メッシュにおける通行規制が行われる前の車両軌跡138件を同様に文字列に変換し、総当たり（9453通り）で計算した結果、平均値は0.364、標準偏差は0.168となり、平常時の確率分布図を図-11に示す。本章(3)と同様に提案手法を4次メッシュに適用した結果のユークリッド距離を表-3に示す。閾値0.10に関しては、5次メッシュ

と同様に分析対象車両30台全てを異常挙動車両であると判定することができ、閾値0.05に関しては、検知台数が21台で検知率は70%となり、5次メッシュにわずかに劣

る結果となった。閾値0.10の場合、4次メッシュでも十分に異常挙動車両の検知が可能であることが明らかになった。

同様に1km 四方サイズの3次メッシュに対して提案手法を適用する。通行規制前の車両軌跡138件を総当たり9,453通りで計算した結果、平均値は0.348、標準偏差は0.167となり、平常時の確率分布図を図-11に示す。4次メッシュの場合とほとんど変化はない。通行止め期の対象車両30台についても同様の手順で評価した異常判定結果を表-3に示す。閾値0.10に関しては、検知台数が29台で検知率は96.67%、閾値0.05に関しては、検知台数が18台で検知率は60%という結果になり、2つの閾値ともに4次メッシュ、5次メッシュよりも劣ることが分かった。

(5) 通行止め期間の対象車両に手法を適用した考察

吉田らの提案手法⁴⁾において用いられていたレーベンシュタイン距離指標に関しては、本手法においても異常挙動を示す車両検出に対する有効性を示すことができた。一方で、ジャロ・ウインクラー距離指標に関しては、その特性から車両軌跡を表現することに適していないことが分かった。

各メッシュサイズにおける検知台数の割合をまとめたものを表-4に示す。結果として、閾値0.10以下であれば4次メッシュ、5次メッシュがともに望ましく、閾値0.05以下では5次メッシュが最も適切なメッシュサイズであり、メッシュサイズが小さいほど詳細な車両軌跡が表現されていることも示された。図-12に通行規制前の車両軌跡(左図)と通行止め箇所を小さく迂回した車両軌跡(右図)の例を示す。両者の軌跡は非常に類似しているため、メッシュサイズが大きくなるにつれて、図-12中に赤い丸で示した箇所が同一のメッシュに含まれるようになり、類似度の高い文字列として判定された。このような事例が存在するため、メッシュサイズによって検出精度が異なる結果となった。

6. 車線規制期間に提案手法を適用した結果

前章より、レーベンシュタイン距離を5次メッシュに適用することで、より精度の高い交通異常検知に繋がることがわかった。この類似度指標とメッシュサイズの組み合わせを、車線規制期間の対象車両40台に適用する。また車線規制期間の車両は、平常通り通行する車両と渋滞等の交通異常に関与した車両が混在する状況となっており、全ての車両を異常挙動と仮定した通行止め期間とは扱いが異なる。本提案手法によって異常判定された車両に対し、誤検知した割合や取りこぼした割合を明らかにするため、実際に交通異常に含まれた真の異常挙動車

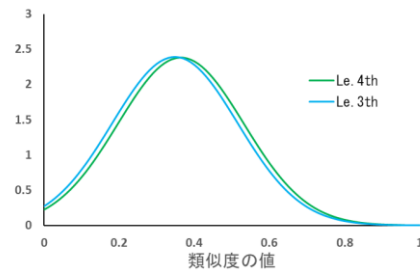


図-11 平常時の確率分布 (Le.4th, Le.3th)

表-3 4次・3次メッシュに適用した判定結果

No.	Le.Eu.4th	閾値		Le.Eu.3th	閾値	
		0.10	0.05		0.10	0.05
1	0.0767	×	△	0.0923	×	△
2	0.0774	×	△	0.0933	×	△
3	0.0372	×	×	0.0377	×	×
4	0.0270	×	×	0.0269	×	×
5	0.0522	×	△	0.0565	×	△
6	0.0614	×	△	0.0715	×	△
7	0.0449	×	×	0.0440	×	×
8	0.0418	×	×	0.0417	×	×
9	0.0349	×	×	0.0350	×	×
10	0.0835	×	△	0.1180	△	△
11	0.0446	×	×	0.0441	×	×
12	0.0647	×	△	0.0750	×	△
13	0.0467	×	×	0.0488	×	×
14	0.0334	×	×	0.0311	×	×
15	0.0425	×	×	0.0421	×	×
16	0.0266	×	×	0.0268	×	×
17	0.0439	×	×	0.0434	×	×
18	0.0443	×	×	0.0444	×	×
19	0.0420	×	×	0.0415	×	×
20	0.0499	×	×	0.0549	×	△
21	0.0497	×	×	0.0541	×	△
22	0.0381	×	×	0.0383	×	×
23	0.0407	×	×	0.0407	×	×
24	0.0542	×	△	0.0623	×	△
25	0.0802	×	△	0.0966	×	△
26	0.0497	×	×	0.0559	×	△
27	0.0505	×	△	0.0575	×	△
28	0.0423	×	×	0.0418	×	×
29	0.0409	×	×	0.0402	×	×
30	0.0412	×	×	0.0406	×	×

表-4 各メッシュサイズの検知率

		3th	4th	5th
		Le.Eu.閾値	0.10	96.67%
	0.05	60%	70%	73.33%

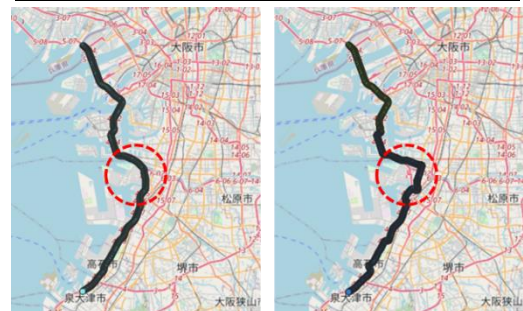


図-12 迂回している様子を示す車両挙動の代表例

両のデータを抽出する必要がある。

(1) 真の異常挙動車両の抽出

阪神高速道路では、速度によって交通状態を下記の3種類に分類している²³⁾。混雑がなく60km/hでの走行が可

能な状態を自由走行，混雑により走行速度が概ね30km/h～60km/h未満の状態を速度低下，混雑により走行速度が概ね30km/h以下の状態を渋滞と定義している。

車線規制は南港北～三宝にかけて行われており，速度低下や渋滞はこの近辺で発生していると考えられる。そこで，前後1つずつのICを含めた区間（天保山～大浜）において渋滞の定義をもとに設定した状態判定閾値を超えた値の車両を，交通異常が発生した真の異常挙動データとする。

次に状態判定閾値の設定方法について述べる。天保山～大浜までの距離は11.3km，所要時間は9分となっている。そこで，所要時間の半分以上である5分（300秒）以上の走行時間において60km/h未満または30km/h以下の速度を記録した車両をそれぞれ速度低下または渋滞の状態に陥った車両とみなす。

(2) 車線規制期間の対象車両に手法を適用した結果

車線規制期間の分析対象車両40台に対し，前章の手法と同様の手順で，ユークリッド距離を1台ずつ算出する。その結果を表-5に示す。1列目が対象車両に割り当てた番号，2列目が算出したユークリッド距離，3～5列目がそれぞれの走行速度範囲内に含まれるドットデータの件数を表している。例として，No.1の車両は天保山～大浜を891秒間30km/h以下で走行したことになる。異常判定のためのユークリッド距離の閾値をそれぞれ0.10，0.11，0.12，0.13に設定した際の40台の判定結果は，×の印が異常判定された車両である。渋滞のみと渋滞と速度低下の両方を対象とした場合の判定を行った。背景色が緑色のものが真の渋滞といえる車両，青色のものが真の渋滞または速度低下といえる車両となっており，異常判定されたものの中で，どちらにも当てはまるものを黄色，一方のみに当てはまるものを該当する色で表している。

これらの判定結果について，適合率と再現率によって精度の評価を行う。適合率とは交通異常と判定したものに對する正解した交通異常件数の割合，再現率とは真の異常データに對する正解した交通異常件数の割合である。渋滞のみに着目した結果と渋滞と速度低下の両方の状態に着目した結果を表-6に示す。渋滞のみに着目した場合，閾値0.11の際に適合率92.86%，再現率100%と最も良い結果となった。また，渋滞と速度低下の状況に着目した場合，閾値0.11の際に適合率が92.86%，再現率81.25%となり，閾値0.13の際に適合率，再現率ともに87.50%と概ね良好な結果を得られた。

(3) 車線規制期間の対象車両に手法を適用した考察

車線規制期間の車両挙動は，多くの車両が通行規制前と似た経路を使用するため，記録される文字の種類も類似する。そのため，経路変更によって文字の種類ごとと変

表-5 車線規制期の対象車両に適用した判定結果

No.	Le.Eu.5th	0-30	30-60	0-60	閾値			
					0.10	0.11	0.12	0.13
1	0.0657	891	207	1098	×	×	×	×
2	0.1569	0	16	16				
3	0.0748	761	217	978	×	×	×	×
4	0.1388	0	35	35				
5	0.1390	0	9	9				
6	0.1403	0	0	0				
7	0.1080	454	224	678		×	×	×
8	0.1734	0	0	0				
9	0.1555	0	114	114				
10	0.1203	0	65	65				×
11	0.0496	934	136	1070	×	×	×	×
12	0.1545	0	95	95				
13	0.1252	281	185	466				×
14	0.1468	0	0	0				
15	0.2170	0	0	0				
16	0.0752	636	251	887	×	×	×	×
17	0.2164	89	121	210				
18	0.0989	387	307	694	×	×	×	×
19	0.1627	0	64	64				
20	0.1992	0	0	0				
21	0.1461	234	200	434				
22	0.1545	0	0	0				
23	0.0543	31	108	139	×	×	×	×
24	0.0809	326	97	423	×	×	×	×
25	0.0898	474	275	749	×	×	×	×
26	0.1541	0	0	0				
27	0.1562	0	0	0				
28	0.1711	169	218	387				
29	0.1010	385	268	653		×	×	×
30	0.1383	0	0	0				
31	0.1066	343	205	548		×	×	×
32	0.0444	489	203	692	×	×	×	×
33	0.1834	0	27	27				
34	0.1890	0	262	262				
35	0.2009	0	0	0				
36	0.1379	0	4	4				
37	0.0646	488	311	799	×	×	×	×
38	0.2274	0	11	11				
39	0.1543	0	0	0				
40	0.0762	627	276	903	×	×	×	×

表-6 検知精度の検証結果

	渋滞のみ			渋滞と速度低下		
	0.10	0.11	0.13	0.10	0.11	0.13
A：交通異常正解数	10	13	13	10	13	14
B：交通異常判定数	11	14	16	11	14	16
C：交通異常件数	13	13	13	16	16	16
適合率 (A/B) (%)	90.91	92.86	81.25	90.91	92.86	87.50
再現率 (A/C) (%)	76.92	100.00	100.00	62.50	81.25	87.50

化する通行止め期間の車両よりも，レーベンシュタイン距離の確率分布図は平常時と類似する傾向となる。そのため，ユークリッド距離による閾値を下げることで，精度よく異常検知が可能な結果を得ることができた。しかしその中で，誤検知した車両と検知漏れした車両が生じている。なぜ，誤検知や取りこぼしが生じたのかを考察する。

a) 誤検知した車両について

No.23のように、異常判定されたにも関わらず、真の異常挙動データには含まれない車両について、誤検知された要因を分析する。図-13にNo.23の車両の軌跡を示す。この車両は、午前8時7分に図-13中に示した地点を出発して中島出入口へと向かい、その後、8時49分に中島出入口から助松JCTへと向かう軌跡が記録されている。本研究では、午前8時台に中島出入口を通過し助松JCTへと向かった記録のあるすべての車両を分析対象としている。中島出入口を通過する以前の経路に対向車線が含まれることを想定していなかったため、同種類のメッシュが重複して記録され、平常時の文字列とは異なる文字列であると判定されたことが要因であると考えられる。



図-13 誤検知した車両の挙動

b) 検知漏れした車両について

30km/h以下の渋滞に着目した場合、閾値を0.11に設定することで完全に検知漏れを防ぐことができた。一方60km/h以下の渋滞と速度低下の状態に着目した場合、No.13, No.21, No.28の3車両で検知漏れが起きている。No.13に関しては、閾値を0.13に設定した際に検知が可能になっている。これら3車両の特徴を相対的に比較すると、30km/h以下の走行時間が300秒の8割～9割に満たない車両は、平常時の車両軌跡との違いがはっきりと表現されず、検知漏れとなる可能性が高いのではないかと推測する。現段階ではサンプル数が不足しているため、分析対象車両を増やすことで、このような結果となる原因を明らかにし、明確な閾値の設定を試みたい。

7. おわりに

本研究は、商用車プローブデータによる交通異常の検知法の提案を行い、2019年11月に行われた阪神高速4号湾岸線リニューアル工事を対象に提案手法の検証を行った。精度の評価では良好な結果を得ることができ、提案手法の有効性が示された。以下が具体的な成果である。

- 本提案手法では、車両の軌跡の時間的・空間的な差異を検出するために、軌跡を文字列に変換し、その類似度を評価する指標を用いて交通異常の検知を行う手法を提案した。文字列の類似度を評価する指標としては、レーベンシュタイン距離とジャロ・ウインクラー距離を用いたが、ジャロ・ウインクラー距離は、交通異常検知には適当でないことがわかった。
- 類似度算出の基準の設定方法に関して、平常時に蓄積されたデータに基づく平常時分布を用いることで、特定の車両に依存しない基準による異常検知手法を提案した。
- 実際の高速道路通行止め期の分析対象車両を用い

て、5次メッシュ、4次メッシュ、3次メッシュに対し提案手法の適用を行い、メッシュごとの検出精度と閾値の関係性を評価した。その結果、5次メッシュに適用した際に、最も精度よく異常挙動車両の検知が可能であることが明らかになった。

- 交通異常に関与した車両とそうでない車両が混在する車線規制時の車両軌跡データに対し、本提案手法を適用した結果、適合率、再現率ともに良好な結果を得られた。ユークリッド距離による閾値を工夫することで、誤検知や取りこぼしを抑えた異常検知が可能となる。
- 通行止め期の車両軌跡を分析することで、経路の違いによる異常を、車線規制期の車両軌跡を分析することで、速度変化の異常を検知可能な手法であることが確認できた。

本研究で提案した手法を適用する上で注意すべき点を以下に示す。

- 使用可能な文字の種類には限りがあるため、分析対象範囲のメッシュの総数を考慮した上で、手法を適用する必要がある。
- 本研究で対象とした阪神高速4号湾岸線リニューアル工事のように、迂回や極端な速度低下が生じる一定程度交通状況に影響を与えるものに対しては有効性が確認されたが、日常の小さな異常に関しての分析は不十分である。
- 本研究は阪神高速道路の中でも主要な4号湾岸線を起点にODを設定したため、平常時における車両軌跡の分析を十分に行うことができたが、安定した基準設定のためにはサンプル数を確保する必要がある。

上記及び本研究全体を踏まえ、今後の展望について以下に示す。

- 通行規制期の車両を分析した際に生じた、異常車両の取りこぼしに関して、明確な閾値を設定するためのサンプル数が十分でないため、複数のODを

対象とした分析が必要である。

- 1メッシュに固有の1文字を付与する手法であるため、文字の種類に応じて対象範囲に限りがあるが、1メッシュに複数の文字を組み合わせたものを付与することで、対象範囲を拡大することが可能となる。そのためには、レーベンシュタイン距離の算出アルゴリズムの改良や、その他の新たな類似度評価指標について分析を進める必要がある。

謝辞：本研究を遂行するにあたって、商用車プローブデータをご提供いただいた富士通交通・道路データサービスに対し、感謝の意を表します。

参考文献

- 1) Kawasaki, Y., Tanaka, A., Goto, H., Takeda, J., Warita, H., Hong, S., Tanaka, S. and Kuwahara, M. : Research on mechanisms to provide attention-attracting information effective in preventing rear-end collisions, 18th ITS World Congress, 2011.
- 2) 川崎紀一, 西山和人 : 画像処理技術を用いた交通流計測と今後の方向性, 土木計画学研究・講演集, Vol.33, CD-ROM, 2006.
- 3) 鈴木一史, 中村英樹, 交通流解析のためのビデオ画像処理システム TrafficAnalyzer の開発と性能検証, 土木学会論文集 D, Vol.62, No.3, pp. 276-287, 2006.
- 4) 吉田政範, 三谷卓摩, 川崎洋輔, 桑原雅夫, 三次元プローブ軌跡データによる交通障害発生時の異常挙動車両検出手法の提案, 土木学会論文集 D3 (土木計画学), Vol.74, No5, p.1149-1158, 2018.
- 5) 川崎洋輔, 梅田祥吾, 桑原雅夫, 機械学習による西日本豪雨時の商用車の迂回の検出および迂回原因の分析, 土木学会論文集 B1 (水工学), Vol.75, No.1, pp.143-154, 2019.
- 6) 日下部貴彦, 全国規模のプローブカーデータを用いた広域的異常事象の検出手法の構築, 土木計画学・研究・講演集, Vol.55, 2017.
- 7) 秦康範, 鈴木猛康, 下羅弘樹, 目黒公郎, 小玉乃理子, 新潟県中越沖地震における通れた道路マップの提供とプローブカー情報の減災利用実現に向けた課題と展望, 日本地震工学会論文集, 第9巻, 第2号, 2009.
- 8) 須藤三十三, 浦川豪, 福重信一郎, 濱本両太, 林春男, 広域的な災害発生後のプローブ情報の活用, -東日本大震災での事例を通して-, 情報システム学会誌, Vol.8, No.1, 2012.
- 9) 早川聡一郎, 塩見康博, プローブデータを用いた豪雨災害時の道路交通状況の分析, 歴史都市防災論文集, Vol.10, pp.145-152, 2016.
- 10) 飯島護久, 堀口良太 : プローブデータに基づくエリア流動性情報提供に関する研究, 第9回シンポジウム, 2010.
- 11) Y, Li. And M, McDonald. : Motorway incident detection using probe vehicles, Proceedings of the Institution of Civil Engineers – Transport, Vol.158, No.1, pp. 11-15, 2005.
- 12) Ruey Long Cheu, Hongtu Qi, and Der-Hong Lee : Mobile Sensor and Sample-Based Algorithm for Freeway Incident Detection, Transportation Research Record: Journal of the Transportation Research Board, 2002.
- 13) Asakura, Y., Kusakabe, T., Nguyen, L. X. and Ushiki, T.: Incident detection methods using probe vehicles with on-board GPS equipment, Transportation Research Part C, Vol.81, pp.330-341, 2017.
- 14) 松平正樹, 藤田幸愛 : ETC2.0 プローブデータによる交通異常検知, 人工知能学会全国大会論文集, 2020.
- 15) 中村俊之, 中嶋康博, 牧村和彦, 井坪慎二 : トラックプローブデータを用いた旅行時間信頼性指標に関する一考察, 土木計画学研究・講演集, Vol.35, CD-ROM, 2007.
- 16) Levenshtein, V. : Binary codes capable of correcting deletions, insertions, and reversals, Soviet Physics Doklady, Vol. 10, No. 8, pp. 707-710, 1966.
- 17) Winkler, William E: String comparator metrics and enhanced decision rules in the Fellegi-Sunter model of record linkage, Proceedings of the Section on Survey Research Methods (American Statistical Association), pp.354-359, 1990.
- 18) 富士通交通・道路データサービスHP : 商用車プローブデータサービス, <https://www.fujitsu.com/jp/group/fitrd/services/commercial-vehicle/probe/> (2022年1月閲覧)
- 19) 漢字辞典 ONLINE, <https://kanji.jitenon.jp/> (2022年1月閲覧)
- 20) 三浦直子, 高木友博 : 語の影響度と編集距離を用いた文の類似度計算手法, 言語処理学会 第21回年次大会 発表論文集, 2015.
- 21) 阪神高速道路湾岸線リニューアル工事HP : <https://hanshin-exp.co.jp/renewal/wangan2019/about/> (2022年1月閲覧)
- 22) 総務省統計局HP : 地域メッシュ統計について, https://www.stat.go.jp/data/mesh/m_tuite.html (2022年1月閲覧)
- 23) 阪神高速道路 HP : 阪神高速道路の渋滞対策, https://hanshin-exp.co.jp/company/torikumi/anzen/ActionP_2015.pdf (2022年1月閲覧)

(2022.3.6 受付)

THE DETECTION METHOD OF ABNORMAL VEHICLE BEHAVIOUR USING COMMERCIAL VEHICLE PROBE DATA

Tomochika FURUKAWA, Hiroe ANDO and Ryuji KAKIMOTO