

道路利用の実態を投稿するマイクロブログユーザの信頼性分析手法に関する研究

藤本 雄紀^{1,a)} 今井 龍一² 中村 健二³ 田中 成典⁴ 有馬 伸広⁵ 荒川 貴之⁶

受付日 2017年12月25日, 採録日 2018年7月10日

概要: 道路管理では、交通事故や渋滞等の様々な現象を検知し、機動的に交通規制等の措置を講ずる必要がある。道路管理者は道路上に設置されたカメラやパトロール等により道路交通の状況を日々監視しているが、大規模災害時にはこれらの機器が故障し状況を把握できない可能性がある。その対応策の取り組みとして、近年、道路管理者が SNS に投稿されたリアルタイムな交通現象に係わる情報を分析する事例が見受けられる。SNS の投稿内容は現状の監視方法や統計調査では収集しにくい移動目的や所感が含まれている。このことをふまえると、信憑性の高い投稿内容の活用方策が確立できれば、道路管理の高度化が期待できる。本研究では、道路利用者が SNS に投稿した特定道路の投稿数と、過去の投稿傾向から分析できる習慣行動を用いて、投稿者を四象限に分類して信頼性を分析する手法を考案した。この結果、本手法で信頼性が高いと判定した投稿者は、交通現象に係わる有用な情報を投稿する傾向があることが分かった。そして、道路管理における SNS 活用の可能性を本研究の成果を用いて考察した。

キーワード: 道路管理, 道路利用, マイクロブログ, テキストマイニング, 信頼性分析

Research for Reliability Analysis Concerning Microbloggers Posting Actual Situations of Road Use to SNS

YUKI FUJIMOTO^{1,a)} RYUICHI IMAI² KENJI NAKAMURA³ SHIGENORI TANAKA⁴ NOBUHIRO ARIMA⁵
TAKAYUKI ARAKAWA⁶

Received: December 25, 2017, Accepted: July 10, 2018

Abstract: In road management, it is necessary to detect diverse phenomena such as traffic accidents and congestion and take measures including traffic regulations flexibly. Road administrators have been monitoring road traffic conditions daily, for example, through cameras installed on the roads or by patrols. However, there is a possibility that they may not be able to grasp the conditions during large-scale disasters as such equipment may break down. As a measure to deal with the situation, there are cases in recent years that road administrators analyze real-time information related to traffic phenomena posted to SNS (social network services). The contents posted to SNS contain microbloggers' purposes of travel or thoughts and impressions that are hard to collect with the present monitoring methods or statistical survey. Based on this, if a measure to utilize highly reliable post contents is established, advancement of road management can be expected. In this study, a method for analyzing reliability of microbloggers by classifying them into four quadrants is devised using the number of SNS entries posted by road users regarding a specific road and habitual behavior that can be analyzed from their past posting tendencies. The result shows that the microbloggers who are judged as reliable by the proposed method have a tendency to post useful information related to traffic phenomena. Then potential for making good use of SNS in road management is discussed using the outcome of the proposed method.

Keywords: road management, road use, microblog, text mining, reliability analysis

1. はじめに

道路管理では、事故や渋滞等の交通現象に対し、機動的かつ適切に交通規制や情報発信等の対応が求められる。このため、道路管理者は、トラフィックカウンタや監視カメラ等を道路上に設置し、道路状況を常時監視している。一方、道路交通調査や分析の動向に着目すると、近年、携帯電話や車載機器等の多様な媒体から得られる交通関連ビッグデータの活用に向けた取り組みが活発化している [1], [2], [3], [4], [5], [6], [7], [8]。たとえば、パーソントリップ調査や道路交通センサ等の統計調査への補完 [4], [5], [6], 交通流の把握 [7] や潜在的事故危険箇所の特定 [8] 等があげられる。現行の交通関連ビッグデータは、移動者の移動目的（通勤や帰宅等）や移動中の所感の把握が難しいという特徴がある [2], [3]。そのため、既存研究 [9] では、移動目的の推定等が試みられている。このデータの特徴をふまえると、科学的根拠となる定量的な交通関連ビッグデータに加えて、移動目的や所感等の定性的なデータも収集ができると、実態に即した道路交通サービスの提供に寄与できると考えられる。

定性的な様々な意見を収集するための一方策として、SNS (Social Networking Service) の一種であるマイクロブログの Twitter を用いた既存研究 [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20] がある。道路管理を対象にした既存研究に着目すると、膨大なつぶやき（以下「ツイート」という）から道路利用に係わるツイートの抽出や [10], [11], [12], 収集したツイートの道路交通分析への適用可能性が示唆 [10], [13] されている。

既存研究の成果から、Twitter は道路利用に関する様々な意見（ツイート）を収集できる。図 1 は、道路管理者の情報収集方法を示している。狭域を詳細に把握する手段の整備は進んでいる一方、広域を一様に把握する手段は十分に整備されていない。また、大規模な災害が発生すると、一般通報、報道やパトロール車による情報収集以外は機能

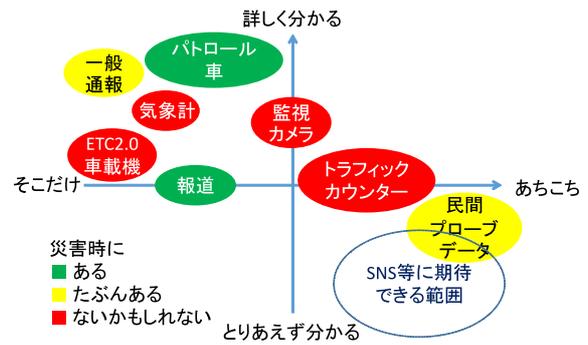


図 1 道路管理者の情報収集方法

Fig. 1 Media for gathering information by road administrators.

しなくなる可能性がある。ここに SNS の情報を活用できると、広域の情報収集手段となりうることを期待できる。過去の災害時における SNS の活用事例 [14] もふまえると、道路管理でも平常時・異常時の両方を対象にした活用方策の確立は有用である。しかし、ツイートは膨大かつ種々雑多であるため、道路管理に利用するには、交通事象や道路整備等の各事案に係わる信憑性の高い意見を速やかに抽出する必要がある。その一方策として、著者らは、特定道路を利用していることが明らかな投稿者に着目すると、信憑性の高い情報を得られる可能性があると考えた。ツイートの信憑性を分析する既存研究は、特に災害時の流言に着目して分析する研究が多数を占めている。主な研究として、ツイートに対する周囲の反応から信憑性を評価する研究 [14], [15] 等があげられるが、投稿後の拡散状況を分析するため、即座に信憑性の高い投稿であるかを判断できない。また、コンテンツの内容が実体験に即した内容であるかを判断するため、体験情報を抽出する研究 [16] がある。この研究から、文章中に含まれる単語等から実際に経験している内容であるかを判断できることが分かる。しかし、Twitter のような文字数が少ないコンテンツでは正確な分析が難しくその他の手法で経験の有無を推定する手法の考案が必要である。そこで、Twitter のコンテンツの特徴でもある位置情報に着目すると、ジオタグと呼ばれる経緯度の付与されたツイートから観光経路を生成する研究 [17] やツイートの文章に含まれる地名を用い、ジオコーディングして位置を推定する研究 [18], [19] がある。これらの既存研究は、ツイートを時系列に分析するため、ツイートの投稿間隔が空いた場合は経路や位置を適切に推定できない。このため、推定精度を確保するには補完が必要となり、その一案として、上述の特定道路のツイートの投稿頻度の高い投稿者の習慣的な行動を蓄積・活用することがあげられる。以上から、ツイートの内容に着目するのではなく、特定道路の利用頻度や投稿傾向等の投稿者の属性に着目して事前に信頼性を分析する。信頼性を分析できると、図 1 に示すように、SNS の情報を活用して有事の際に道路管理者が

1 関西大学大学院総合情報学研究所
Graduate School of Informatics, Kansai University, Takatsuki, Osaka 569-1095, Japan
2 東京都市大学工学部
Faculty of Engineering, Tokyo City University, Setagaya, Tokyo 158-8557, Japan
3 大阪経済大学情報社会学部
Faculty of Information Technology and Social Sciences, Osaka University of Economics, Osaka 533-8533, Japan
4 関西大学総合情報学部
Faculty of Informatics, Kansai University, Takatsuki, Osaka 569-1095, Japan
5 阪神高速株式会社計画部
Planning Department, Hanshin Expressway Company Limited, Osaka 541-0056, Japan
6 阪神高速技研株式会社システム事業本部
System and Data Solution Department, Hanshin Expressway R&D Company Limited, Osaka 530-6123, Japan
a) k398881@kansai-u.ac.jp

信頼できる投稿者と双方向コミュニケーションをとりながら情報収集できる仕組みが構築できる。しかしながら、特定道路のツイートの投稿頻度の高い投稿者の有無、さらにツイートの内容の特性は解明されていない。また、ツイートをを用いた移動目的の抽出や投稿間隔に依存しない経路推定の手法も確立されていない。

この状況に基づいて、本研究の目的を特定の投稿者の日々の道路利用に関するツイートから抽出した習慣行動および特定道路の利用状況に基づく投稿者の信頼性分析手法の開発とした。また、実際に道路管理者と議論し、活用事例をとりまとめ、本研究の意義を明らかにする。なお、特定道路のツイートの抽出や習慣行動の推定アルゴリズムは既存手法に基づくが、本論文の主題である移動経路の生成や信頼性を分析するアルゴリズムは新規の知見である。

本論文では、まず、2章で道路利用に関するツイートを収集し、投稿内容の特性を分析する。次に、3章でツイート投稿者の信頼性分析手法の概要を述べる。4章で同手法のアルゴリズムを詳述する。そして、5章で実証実験を行い信頼性の分析可能性を評価し、6章では本研究の成果が道路交通分析の高度化に寄与できるかを考察する。最後に、7章で本研究を総括し、今後の展望を述べる。

なお、本研究におけるツイートの取扱いは、個人情報やプライバシーに抵触しないよう、Twitterのプライバシーポリシー [21] に準じている。

2. 投稿者のツイートの調査

2.1 調査目的

本研究では、道路に係わるツイートおよび投稿者の傾向を調査するため、ツイートを収集し、その内容や特性を分析する。

2.2 調査手順

ツイートの調査対象は、既存研究 [11] と同様に阪神高速道路に係わるツイートとする。

まず、Twitter社提供のSearchAPI [22] を用いて、2016年1月16日から6月18日までの155日間に投稿されたツイートのうち、著者らが既存研究 [11] で作成した地名辞典を用いてツイートを収集する。地名辞典は、道路管理者が公表資料等で道路利用者に発信する際に使用している単語や、SNSを含むWebサイトで扱われている単語を手作業で収集した。その結果、表1に示すような高速道路や一般道路の路線名、サービスエリア名やインターチェンジ名等の正式名称、略語、通称を含む971語を収録した。これらの語句を含むツイートには、「阪神高速の京橋PAはいつも流行ってるね」等の道路利用の所感や、「阪神高速3号神戸線月見山から渋滞中」等の交通現象に係わる情報が含まれている。

次に、収集したツイートから「阪神高速」または「阪

表1 地名辞典の一例

Table 1 Examples of geographical dictionary.

カテゴリ	地名辞典の用語
高速道路名	阪神高速, 阪高 (略称)
路線名	三号神戸線, 環状線
一般道路名	国道1号線, 明姫幹線, 明幹 (略称), 近江グリーンロード (通称)
SA/PA名	湊町SA, 中島PA
IC名	京橋, 鳴尾浜

表2 ツイートの収集結果

Table 2 Results of tweets collected.

項目	ツイート数 (件)	投稿者数 (名)
一般道や高速道路に係わるツイート	2,148,672	614,840
阪神高速道路に係わるツイート	22,704	13,457
投稿者のタイムライン	604,780	983

高」のキーワードを含むツイートを抽出し、そのツイートの投稿者情報をAPIで収集する。SearchAPIを用いると、投稿者情報、ツイートの本文やジオタグ等が収集できる。投稿者情報は、Twitter上での名前、自己紹介や居住地等が収集できる。そして、抽出された投稿者ごとの過去のツイートのログ (以下「タイムライン」という) をstatus/user_timelineAPIを用いて収集する。最後に、投稿者ごとに阪神高速道路に係わるツイートを集計し、傾向を確認する。

2.3 ツイートの収集結果

ツイートの収集結果を表2に示す。一般道路や高速道路に係わるツイート2,148,672件のうち、阪神高速道路に係わるツイートは22,704件、投稿者数は13,457名であった。このうち、阪神高速道路上の位置を推定できるツイートの投稿者数は983名であり、その投稿者に係わるタイムラインのツイート数は604,780件であった。

2.4 投稿者のツイートの分析結果

収集したツイートを分析した結果、次の2点の知見が得られた。

2.4.1 投稿者のTwitterの利用実態

分析結果から、日常的に阪神高速道路に係わるツイートを投稿している投稿者が多く存在することを確認できた。調査した155日間の投稿者ごとの投稿件数を分析すると、10件以上投稿している投稿者が141名存在することが分

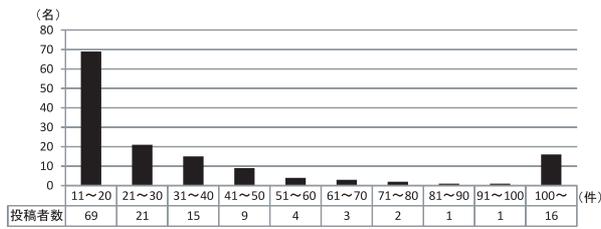


図 2 阪神高速道路に係わるツイート数別投稿者分布

Fig. 2 Distribution of microbloggers calculated by the number of tweets related to Hanshin Expressway.

かった (図 2 参照). その詳細を確認すると, 投稿件数の多い投稿者で最大 344 件となっており, 高い頻度で阪神高速道路に係わるツイートを投稿している投稿者が存在することが分かる. また, それらのツイートを目視で確認すると, 「阪神高速 3 号神戸線. 若宮あたりで渋滞. 事故?」等の交通現象が発生したことを周知させる内容や, 事故の状況を把握できる写真や動画付きのツイートも確認できた.

この結果から, 特定道路の情報を投稿する投稿者が存在しており, 交通現象の情報を収集できることが分かる. しかしながら, これらの情報の信憑性は担保されていない. このことはツイートを大規模災害時の情報収集ツールとして活用する際の弊害 [11] となっている. そのため, 膨大なツイートから信憑性の高い情報を体系的に集計できる必要があるといえる.

2.4.2 信憑性分析の可能性

ツイート内容の信憑性を分析するには, ツイートに対する反応や訂正情報を用いる研究 [14] がある. この研究では, 信憑性を分析したいツイートの反応に対し, 誤りを指摘する表現の有無によって判定している. また, リツイートやツイートに付与された URL ソース等から信憑性を分析する研究 [15] があるが, ツイートに対する反応が現れるまで判定できない点が共通課題としてあげられる.

そこで本研究では, その現場に居合わせた当事者が投稿したツイートこそが信憑性の高いツイートであると考え, 投稿者が実際に経験しているかという点に着目する. また, 平常時から特定道路の話題を投稿する投稿者はその道路に関心があると考えられるため, 過去の特定道路の投稿数にも着目する.

2.5 投稿者の分類

図 3 に示すとおり, 著者らの仮定に基づき, 特定道路の投稿数と, 特定道路の利用経験という尺度から, 投稿者を 4 つに分類し定義する. そして, この 2 つの尺度が高い投稿者が信頼性の高い投稿者であると判断する.

まず, 信頼性が高い第一象限 (以下「信頼性最高象限」という), 信頼性が比較的高い第四象限 (以下「信頼性高象限」という) の投稿者に着目する. 信頼性最高象限の投稿者は, 特定道路の投稿が多く特定道路の利用頻度が高い

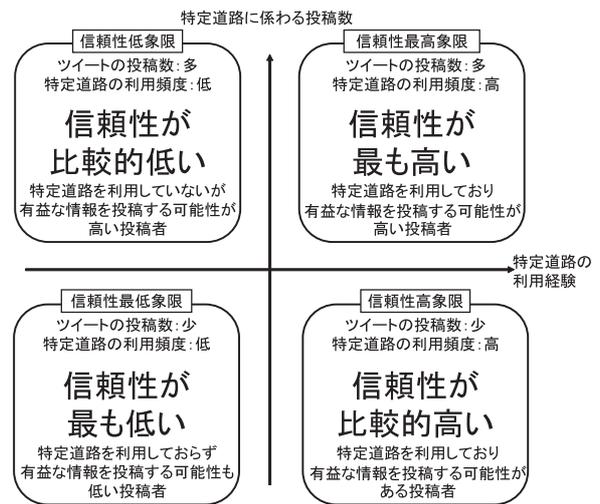


図 3 投稿者の信頼性の分類

Fig. 3 Classification of microbloggers based on their reliability.

投稿者である. 日々特定道路を利用しており, ツイート数から特定道路に関心があるといえるため, 最も信頼性が高い. 信頼性高象限の投稿者は, 投稿数は少ないものの, 利用頻度が高いため, 有事の際には状況等を報告する可能性が高い. この 2 つの象限の投稿者から得られることが期待されるツイートの例をあげると, 「阪神高速 3 号神戸線にてトラックが玉突き事故」というような交通現象の発生を検知できる情報である. 次に, 信頼性が比較的低い第二象限 (以下「信頼性低象限」という), 信頼性が低い第三象限 (以下「信頼性最低象限」という) の投稿者に着目する. 信頼性低象限の投稿者は, 利用頻度が低いものの, 特定道路のツイート数が多く関心が高いため, 交通現象等の情報を投稿する可能性がある. 一方, 経験情報に基づかない投稿も含まれているため, 信頼性は比較的低い. 信頼性最低象限の投稿者は, 自分自身とは関係がないため, 「阪神高速で事故があつて怖いな」等の単なる感想や信頼性最高象限, 信頼性高象限の投稿者の情報をリツイートするのみにとどまる可能性が高いため, 信頼性が最も低い. 以上の分類結果から, 本研究では, 信憑性の高いツイートを投稿しているであろう信頼性最高象限, 信頼性高象限の投稿者を推定する手法を考案する.

3. 投稿者の信頼性の分析手法の考案

本章では, 前章の結果をふまえ, 課題を次に示すとおり設定し, その解決策となる手法の手順を概説する.

1 点目の課題は, 「道路交通分析において, 適切にツイートの信憑性を分析する手法が確立されていないこと」である. これに対し, ツイートの投稿者の過去の投稿から得られる特定道路の投稿数と利用頻度に基づき, 投稿者の信頼性を分析することに対応する.

2 点目の課題は, 「交通現象等の情報を提供する有用な投稿者を推定できないこと」である. これに対し, 考案した

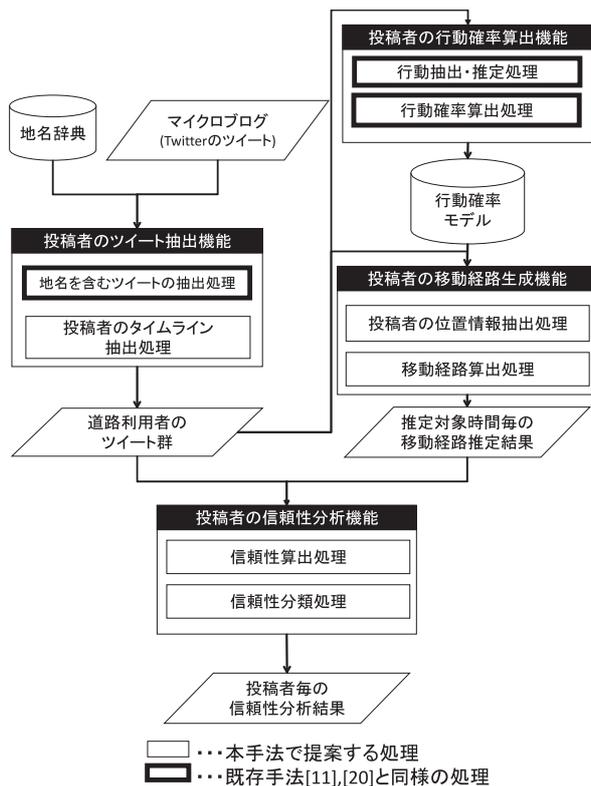


図 4 本手法の手順

Fig. 4 Process of proposed method.

信頼性を分析する手法を用いて、前章で定義した信頼性最高象限、信頼性高象限の有用な投稿者を抽出することで対応する。

3.1 本手法の手順

解決策を反映した信頼性の高い投稿者の抽出手法（以下「本手法」という）の手順を図 4 に示す。本手法は投稿者のツイート抽出機能、投稿者の行動確率算出機能、投稿者の移動経路生成機能および投稿者の信頼性分析機能から構成される。投稿者の信頼性は前章で述べたとおり、投稿者の特定道路に係わる投稿数と経験情報を用いて分析する。

特定道路に係わる投稿数は投稿者のツイート抽出機能を用いて、各投稿者のタイムラインの収集後、特定道路の単語を含むツイートを抽出し、集計する。

特定道路に係わる経験情報は、投稿者の移動経路生成機能により、投稿者の居住地と、日々の移動目的地との組合せから経路を推定して対象道路を通過していれば利用したことがあると判断する。そのために必要な居住地と移動目的地は市区町村レベルの推定詳細度であれば、経験情報として活用できると仮定する。まず、既存手法 [20] に基づいた投稿者の行動確率算出機能を用いて習慣行動を分析する。次に、特定の行動時に投稿されたツイートに含まれる地理識別子をジオコーディング API [23] を用いて座標に変換し市区町村レベルで出現回数を集計する。最後に、出現回数の多い地名を居住地および目的地として設定する。ただ

し、居住地はプロフィールに設定されている場合にのみ、その地名を優先する。

以上の機能で抽出した特定道路に係わる投稿数および経験情報を用いて、投稿者の信頼性分析機能で投稿者の信頼性を分析する。本手法の詳細な処理手順は次章に示す。

4. 投稿者の信頼性の分析アルゴリズム

4.1 投稿者のツイート抽出機能

本機能は地名を含むツイートの抽出処理および投稿者のツイート抽出処理から構成される。地名を含むツイートを投稿した投稿者は、その地名の投稿者であると判断し、その投稿者のタイムラインを取得する。

4.1.1 地名を含むツイートの抽出処理

本処理は、2章で述べた手順と同様、地名辞典 $GD = \{gd_1, gd_2, gd_3, \dots, gd_i\}$ を作成し (i は語句の件数を指す)、その地名を含むツイート群 $TW = \{tw_1, tw_2, tw_3, \dots, tw_j\}$ を Twitter 社提供の SearchAPI を用いて収集する (j はツイートの件数を指す)。

4.1.2 投稿者のタイムライン抽出処理

本処理は既存研究に加えて新規開発した処理である。収集したツイート群 TW から、投稿者 $TU = \{tu_1, tu_2, tu_3, \dots, tu_k\}$ を取得し (k は投稿者の件数を指す)、その投稿者のタイムラインを収集して道路利用者 tu_k のツイート群 $TLtu_k = \{tl_1, tl_2, tl_3, \dots, tl_l\}$ を status/user.timelineAPI を利用して取得する (l はツイートの件数を指す)。ただし、この API では過去 3,200 件までしか取得できないため、対象となる投稿者に対し日々 API を実行する必要がある。

4.2 投稿者の行動確率算出機能

本機能は既存手法 [20] の処理を活用する。本機能は行動抽出・推定処理および行動確率算出処理より構成される。投稿者のタイムライン $TLtu_k$ のうち、出勤中、勤務中、帰宅中および睡眠中の四種類の行動 beh_m (m は各行動を指す) に関する単語を含む投稿を各曜日の 1 時間ごと $sun_0, sun_1, \dots, sat_{23}$ に集計する。そして、全体の投稿数で正規化した結果 $Post$ を用いて、行動確率 $V(tu_k, beh_m) = \{Post_{beh_m:sun_0}, Post_{beh_m:sun_1}, Post_{beh_m:sun_2}, \dots, Post_{beh_m:sat_{23}}\}$ を算出し、行動確率モデルに格納する。そして、各曜日の時間帯ごとに行動確率が最も高い行動をその時間帯の行動とする。なお、投稿数が少ない場合や、行動 beh_m に係わる単語を含む投稿が存在しない場合には行動を不明 $beh_{unknown}$ として出力する。既存手法 [20] では、4 種類の行動に加えて、食事中およびその他の行動も推定するが、本手法では、居住地か移動目的地のどちらに在るのかを推定することを重要視するため、居住地か移動目的地を一意に特定することが困難な食事中とその他の行動は除外する。そして、睡眠中と出勤中の投

稿に含まれる位置情報から居住地を推定し、勤務中と帰宅中の投稿に含まれる位置情報から移動目的地を推定する。一般的に睡眠中は投稿数が少なくなるが、なかには「家に帰るお風呂あがって僕も寝ますのよおやすみなさい」のような居住地に居ることが推定できる投稿や「おやすみなさい。@六甲道,新在家」のような居住地の位置情報を含む投稿も存在するため推定対象とした。

4.3 投稿者の移動経路生成機能

本機能は本研究で新規開発した機能である。投稿者の位置情報抽出処理および経路算出処理から構成される。行動確率 $V(tu_k, beh_m)$ に基づいて、投稿者の行動を分類する。分類した結果を用いて、ツイートから位置情報を抽出し、移動経路を推定する。

4.3.1 投稿者の位置情報抽出処理

本処理はツイートを各行動に分類し、その行動ごとにツイートの含まれるジオタグまたは投稿に含まれる座標から居住地の位置情報と、主要な目的地の位置情報を抽出する。本処理の手順を次に示す。

STEP 1 では、得られた行動確率 $V(tu_k, beh_m)$ に基づいて、行動を補完する。行動確率が勤務中の場合は、その前後1時間をそれぞれ出勤中および帰宅中として行動確率を補完する。行動確率が出勤中および帰宅中の場合はその前後1時間を同様の行動として行動確率を補完する。

STEP 2 では、STEP 1 で得た補完後の行動確率 $V_{lifecycle}(tu_k, beh_m)'$ を基に、ツイート $TLtu_k$ を行動 beh_m ごとにツイート群 $Tweetgroup(tu_k, beh_m)$ としてまとめる。

STEP 3 では、ツイート群 $Tweetgroup(tu_k, beh_m)$ のツイートのジオタグが含まれているかを確認し、含まれている場合は座標を抽出する。含まれていない場合は、ツイートを MeCab [24] を用いて形態素解析した後、ツイートの含まれる副詞および助詞を除いた単語群を抽出し、ジオコーディング API を用いて地域に係わる単語を座標に変換する。このとき、地域に係わる単語が複数存在した場合もすべての単語をジオコーディングする。そして、座標群 $Geogroup(tu_k, beh_m) = \{geo_{km0}, geo_{km1}, geo_{km2}, \dots, geo_{kmm}\}$ に時系列に追加する (n は座標の件数を指す)。以上の STEP 1~3 を投稿者 tu_k ごとに繰り返す。

4.3.2 移動経路算出処理

本処理は行動ごとのツイート群に含まれる座標群 $Geogroup(tu_k, beh_m)$ を用いて投稿者の居住地から勤務地の移動経路を算出する。

まず、居住地および移動目的地を設定する。居住地は投稿者のプロフィールから取得する。プロフィールに居住地が含まれない場合は睡眠中または出勤中の座標が最も多く含まれる市区町村を出発地とする。移動目的地は日々変化することが考えられるため、複数の移動目的地を候補地と

して抽出する。勤務中または帰宅中の座標群の中で頻出するものを移動目的地の候補とする。

そして、居住地から各勤務地を通る移動経路を Google Directions API [25] を用いて作成し、投稿者の移動経路 $MRTu_k = \{r_1, r_2, r_3, \dots, r_o\}$ とする (r は移動経路を指し、 o は移動経路の件数を指す)。

4.4 投稿者の信頼性分析機能

本機能は算出した投稿者の移動経路 $MRTu_k$ と、道路利用者のツイート群 $TLtu_k$ を入力として投稿者 tu_k の信頼性を分析する。

4.4.1 信頼性算出処理

本処理は、道路利用者 tu_k のツイート群 $TLtu_k$ から、特定道路の単語が含まれるツイートを抽出し、その件数を $SRCount(TLtu_k)$ とする。そして、移動経路算出処理で算出した移動経路 $MRTu_k$ のうち、特定道路を経由する移動経路の件数を $SRCount(MRTu_k)$ として集計する。

4.4.2 信頼性分類処理

本処理では、特定道路の単語が含まれるツイート件数の閾値 α 、特定道路を経由する移動経路の件数の閾値 β を用いて投稿者 tu_k の信頼性 $R(tu_k)$ を算出する。信頼性 $R(tu_k)$ には図 3 の 4 つの分類結果である $Quad_1$ (信頼性最高象限)、 $Quad_2$ (信頼性低象限)、 $Quad_3$ (信頼性最低象限)、 $Quad_4$ (信頼性高象限) が格納される。これを用いて投稿者の信頼性を分析する。

$$R(tu_k) =$$

$$\begin{cases} Quad_1 & \text{if } SRCount(TLtu_k) > \alpha \cap SRCount(MRTu_k) > \beta \\ Quad_2 & \text{if } SRCount(TLtu_k) > \alpha \cap SRCount(MRTu_k) < \beta \\ Quad_3 & \text{if } SRCount(TLtu_k) < \alpha \cap SRCount(MRTu_k) < \beta \\ Quad_4 & \text{if } SRCount(TLtu_k) < \alpha \cap SRCount(MRTu_k) > \beta \end{cases}$$

5. 本手法の有用性の検証

5.1 検証方法

本研究では、3章であげた課題の解決策として本手法が有用であるかを確認する。課題 1 に対しては、「検証 1: 特定道路を日常利用する投稿者の抽出」を実施し、経験情報を取得できるかを検証する。課題 2 に対しては「検証 2: 抽出した投稿者の信頼性の評価」を実施し、信頼性の高い投稿者を抽出できているかを確認する。なお、ツイートは 2章で収集したツイートに加えて、その後も継続して収集したツイートも対象とする。事前検証として、一般的に投稿者の位置情報を正確に把握することは難しい [26] とされているため、投稿者の位置情報をどの程度の詳細度であれば推定でき、その結果が移動経路の推定で活用できるかを確認する。

検証 1 では、本手法を適用して阪神高速道路を日常利用している可能性が高い投稿者を抽出できるかを確認する。

まず、投稿者のタイムラインを取得して習慣行動を分析する。本検証では、習慣行動を出勤中、勤務中、帰宅中および睡眠中の4つとする。そして、出勤中および帰宅中の時間帯に投稿された阪神高速道路に係わるツイート数と、阪神高速道路を利用しているかという観点から、阪神高速道路を日常利用している可能性があるかを判定する。投稿者が阪神高速道路の利用者であるかは、車両の移動軌跡のプロブデータ等の他データとの組合せ分析により解明できる可能性があるが、プロブデータと投稿者との対応付け手法が未確立のため、本研究では今後の議論とする。

検証2では、検証1で抽出した投稿者を図3に示す4つの象限に分類した結果と、一般に投稿者の属性を分析する既存研究[27]で用いられるフォロー数とタイムラインの投稿数に着目し、フォロー数の多い投稿者上位100名を抽出した場合（以下「既存研究1」という）と、タイムラインの投稿数の多い投稿者上位100名で試行した場合（以下「既存研究2」という）とで事故や渋滞といった交通現象に係わるツイート数や投稿内容から傾向を分析する。信頼性を分析する手法には、大別してツイートの内容や、やりとりを分析する手法[14]と、ツイート内容に加えてユーザの属性を分析する手法[15]がある。本手法は投稿者の属性から信頼性を分析する手法であるため、ツイートを分析する手法との比較は検証の対象外とする。この既存研究との比較により、フォロー数やタイムラインの投稿数といったデータだけではなく、投稿傾向や習慣行動に基づく体験情報を考慮することの有効性を評価する。

5.2 検証対象データ

本検証では、2章で収集した2016年1月16日から2016年6月18日のツイートに加え、2016年10月16日までの274日間に投稿されたツイート2,941,157件を用いる。このうち、阪神高速道路上での位置情報が明らかな投稿者2,058人とそれに係わるタイムラインのツイート3,473,261件を対象とした。

5.3 事前検証：位置情報の推定詳細度の分析結果と考察

5.3.1 検証の概要

本検証では、位置情報をツイートからどの程度の詳細度で推定でき、その結果が移動経路算出処理に活用可能かを3つの検証により確認する。まず、検証条件のツイートのうち、ジオタグが付与されたツイート100件を無作為に抽出し、ジオタグの位置情報を正解として、文章に含まれる地名に係わる語句からどの程度の詳細度で推定できるかを検証する。単一のツイート中に複数の地名が含まれる場合は、いずれかの地名の位置情報がジオタグの位置情報と一致していれば正解とする。次に、移動経路算出処理でプロフィールの位置情報を優先することの妥当性を評価する。プロフィールに正解となる位置情報を含む投稿者を人手で

表3 位置情報の推定精度

Table 3 Estimated accuracy of location information.

項目	都道府県	市区町村
対象データ数 (件)	100	
正解数 (件)	89	72
不正解数 (件)	11	28
正解率	89%	72%

表4 投稿者の居住地の推定精度

Table 4 Estimated accuracy of place of residence each user.

項目	プロフィール	行動推定
対象データ数 (件)	76	
正解数 (件)	69	51
不正解数 (件)	7	25
正解率	91%	67%

抽出し、プロフィールから居住地を推定した結果と、投稿者の行動から居住地を推定した結果とで精度を比較する。最後に、今回の検証対象である阪神高速道路の利用傾向や路線の特徴を調査し、位置情報の推定詳細度に問題がないかを確認する。

5.3.2 位置情報の推定詳細度の検証結果

位置情報の推定詳細度は、都道府県レベルでは高い精度で推定できるとされている[28]。しかし、本研究では投稿者の居住地、移動目的地を分析する必要があるため、市区町村レベルで推定ができるかを検証した。結果を表3に示す。都道府県レベルでは正解率89%、市区町村レベルでは正解率72%の精度で推定でき、都道府県レベルに比べると精度は落ちるが、市区町村レベルの生活範囲をある程度把握することは可能であるといえる。特定に失敗したツイートは、ジオタグが付与されているものの、文章中の地名に係わる語句が「Osaka」等のローマ字表記の場合に推定できないことが確認できた。これに対しては、語句を日本語に変換して処理すること等が考えられるため、今後より詳細な位置情報の推定詳細度が必要となる場合にはこの処理を追加する。

5.3.3 プロフィールに含まれる位置情報の妥当性の検証結果

投稿者の居住地の推定結果を表4に示す。プロフィールから居住地を推定した結果は正解率91%、投稿者の行動から居住地を推定した結果は正解率67%と、前者の方が正確に把握できるため、本研究ではプロフィールに含まれる位置情報を優先することとする。居住地推定に失敗したプロフィールは、居住地欄に英数字や記号等のノイズとなる語句が含まれていた。そのため、プロフィールに含まれる不

表 5 本手法による阪神高速道路の利用者抽出結果

Table 5 Results of extracting users of Hanshin Expressway with proposed method.

阪神高速道路に係わるツイート件数	道路利用者数 (名)	阪神高速道路の利用者数 (名)
10 件以上	9	5(55.6%)
9 件	2	0 (0.0%)
8 件	2	0 (0.0%)
7 件	2	0 (0.0%)
6 件	6	1(16.7%)
5 件	10	1(10.0%)
4 件	10	4(40.0%)
3 件	29	16(55.1%)
2 件	78	31(39.7%)
1 件	325	105(32.3%)

要な語句を取り除くクリーニング処理を施すとより精度高く推定できる可能性がある。

5.3.4 考察

以上の検証結果から、位置情報は市区町村レベルの詳細度で推定でき、プロフィールに含まれる位置情報を優先した方が高精度に推定できることが明らかとなった。これらの結果が移動経路算出処理に活用可能かを確認するため、阪神高速道路の利用状況を調査した。国土交通省の報道発表資料 [29] によると、阪神高速道路は 12km~18km の距離の利用者が最も多いとされている。一方、市区町村ごとに阪神高速道路の総延長を集計したところ、神戸市北区を除きいずれも 12km 以下であった。これらのことから、市区町村内の移動を目的として阪神高速道路を利用する可能性が低いと、市区町村レベルで居住地や移動目的地の推定ができれば、移動経路の推定には十分活用できるといえる。

5.4 検証 1：特定道路を日常利用する投稿者の抽出結果と考察

表 5 は検証 1 の結果を示しており、表中の道路利用者数は、阪神高速道路に係わるツイート件数ごとに集計したユニークユーザ数である。これらのうち、出勤中および帰宅中に阪神高速道路に係わるツイートを投稿する投稿者 473 名を抽出できた。これらの結果を確認すると、10 件以上ツイートを投稿している投稿者が 9 名存在しており、5 名 (55.6%) が阪神高速道路を 1 カ所以上通っている。一方、10 件未満の投稿者の中にも、日常的に利用している可能性が高い投稿者が存在していることが分かる。得られた知見を以下に示す。

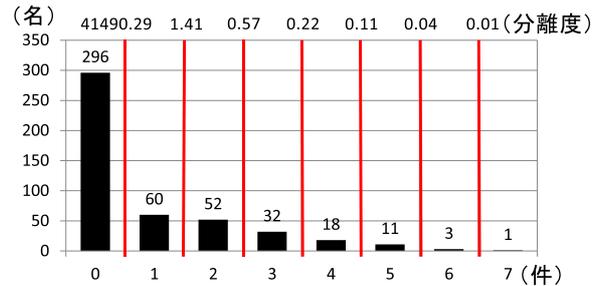


図 5 阪神高速道路を通る移動経路数のヒストグラム

Fig. 5 Histogram of the number of routes moving on Hanshin expressway.

5.4.1 特定道路の経験情報の抽出

出勤中および帰宅中に特定道路に係わるツイートを多く投稿する投稿者は、その道路の利用頻度が高い傾向があることが分かった。一方、特定道路に係わるツイート件数が少ないにもかかわらず、移動経路に特定道路を利用している可能性が高い投稿者も多数存在していた。これらの投稿者は災害時等の有事の際には経験情報に基づく信憑性の高い情報を提供する可能性があるため、重要な投稿者であるといえる。この結果から、阪神高速道路に係わるツイートを 10 件以上投稿している投稿者は特定道路のツイート数が多いという判断ができる。そして、阪神高速道路を通る移動経路を推定した結果を図 5 に示す。縦軸は投稿者数、横軸は移動経路数を示す。この結果に対し、何件以上移動経路が存在していれば阪神高速道路の利用者として判断できるかを検証するため、大津の二値化処理 [30] を適用する。この手法は、分離度という値が最大となる閾値を求め、自動的に二値化を行う手法である。分離度の算出結果は赤線の縦軸に示している。この結果を確認すると、分離度が最大となった閾値は 1 であったため、特定道路を経由する移動経路の件数の閾値 β は 1 が妥当であるといえる。

また、実際に習慣的に特定道路を使用していることが明らかでない投稿者のツイートの例を表 6 に示す。ツイートの内容を確認すると、「おはようございます」や「帰る」といった行動に関するツイートがある。そのため、投稿者のタイムラインを分析すると、日々の習慣的な行動を抽出でき、特定道路を日常利用しているか等の経験情報を抽出できる。

以上から、特定道路の投稿数と利用頻度の尺度がいずれも高い投稿者を抽出することで課題 1 である「道路交通分析において、適切にツイートの信憑性を分析する手法が確立されていないこと」は対応できる。

5.4.2 行動推定の精度

今回収集した投稿者 1 人あたりのタイムラインの数は平均すると 1,838 件であった。しかし、投稿者によってはほとんど投稿がなく、タイムラインを十分に収集できていないものも存在した。そのため、投稿者によって行動推定数が低下する等、推定数のばらつきが確認できた。これは、

既存研究 [20] でもあげられているが、投稿者の行動は1日の平均投稿数が少なければ適切に推定できない。そのため、投稿数に依存しない行動推定手法の考案が必要であることが分かった。

5.5 検証 2:抽出した投稿者の信頼性の評価結果と考察

検証1の結果から、本手法の閾値は、特定道路の単語が含まれるツイート件数の閾値 α を10、特定道路を経由する移動経路の件数の閾値 β を1とすることが妥当であると

表 6 行動の分析可能なツイートの例

Table 6 Examples of analyzable tweets concerning microbloggers' behavior.

ツイート内容	投稿日時
おはようございます(^-^)/ 平日最終日な金曜日、今日も無理せず乗りきりましょう☆	2016/2/5 6:53
付近を通過ー！ 日中は寒さまだマシンやな☆ (@ 阪神高速 池田木部第一出入口 in 池田市, 大阪府)	2016/2/5 11:35
伊丹まで南下 (@ 伊丹1交差点)	2016/2/5 13:21
兵庫県→大阪府へと！ (@ 軍行橋 in 伊丹市, 兵庫県)	2016/2/5 14:38
ちよいとまったり♪ (@ 五月山公園 売店)	2016/2/5 15:51
まったりしたし、そろそろ帰るかな！ (@ 五月山公園 in 池田市, 大阪府)	2016/2/5 16:13

判断した。抽出した投稿者を分類した結果を表 7 に示し、得られた知見を以下に示す。

5.5.1 経験情報を考慮することの有用性

各手法で抽出した投稿者のツイート内容を表 7 に示す。まず、“ニュースを引用したツイート件数”に着目すると、本手法の信頼性最高象限の結果が0件であった。このツイートは交通現象の検知の観点から重要性が低いと考えられるため、有用なツイートのみを適切に抽出できている。次に、状況把握に活用できると考えられる“動画像付きのツイート数”に着目すると、いずれの既存手法も取得ができていないが、本手法では信頼性最高象限で1件、信頼性高象限で2件取得できている。最後に、“1人あたりの交通現象に係わる有用なツイート数”に着目する。この項目はニュースを引用したツイートを除く1人あたりの交通現象に係わるツイート数である。本手法の信頼性最高象限のみ、1人あたり2件以上も投稿しており、有益な情報を提供する可能性が高い投稿者が抽出できている。この結果から、本手法を用いると信頼性の高い投稿者を抽出できることが分かる。

本手法の各象限の結果を確認すると、信頼性最高象限の投稿者が5名で人数が少ないにもかかわらず、事故や渋滞といった交通現象に係わるツイートを14件も投稿していた。また、信頼性最高象限の投稿者の移動経路を可視化した結果(図 6 参照)を確認すると、京都、兵庫や和歌山等から阪神高速道路を利用して様々な目的地に移動しており、経験情報を抽出できていることが分かる。信頼性低象限の投稿者は信頼性最高象限に次いで1人あたりの交通現象に係わるツイート数が0.20と高いが、体験情報に基づかない情報であるため、信頼性はやや低い。ツイート内容を確認すると、例に示しているようなニュース記事を引用したのも7件のうち2件(29%)含まれ、興味本位で投稿している事例も散見された。信頼性最低象限の投稿者は最も

表 7 各手法の信頼性の分析結果

Table 7 Results of reliability analysis of contents posted by miclobloggers each method.

手法	象限	人数(名)	交通現象に係わるツイート数(件)	ニュースを引用したツイート数(件)	動画像付きのツイート数(件)	一人当たりの交通現象に係わる有用なツイート数(件)	交通現象に係わるツイート例
本手法	最高	5	14	0(0%)	1	2.8	宝塚一神戸三田今で渋滞してるから伸びるようなら神戸三田で降りて六甲北六甲トンネル経由して阪神高速の湾岸線乗る方が速いかも？六甲トンネルまっすぐ下って湾岸線住吉浜やし
	低	24	7	2(29%)	0	0.20	渋滞車列にトラック、2人死亡…阪神高速湾岸線
	最低	286	48	23(48%)	0	0.08	阪神高速の事故怖いなー中身グチャグチャやん
	高	158	44	6(14%)	2	0.24	阪神高速3号神戸線の大阪方面月見山付近で5台の玉突き事故で渋滞発生中です
既存手法1	-	100	67	56(84%)	0	0.11	阪神高速玉突き事故で逮捕 (読売テレビ)
既存手法2	-	100	24	10(42%)	0	0.14	阪神高速湾岸線で多重事故 2人死亡 神戸市消防局 (テレビ朝日)

1人あたりの交通現象に関わるツイートの投稿数が0.08と低く、有益な情報提供者が少ないことが分かる。また、ツイートの内容を確認すると、48件中23件(48%)がニュース記事の引用ツイートや感想のツイートで、信頼性低象限よりも多く既知の情報を拡散目的で投稿している投稿者が多く抽出されていることが分かる。最後に、信頼性高象限の投稿者は信頼性低象限と同程度の割合で1人あたりの交通現象に係わるツイートを投稿しており、その内容は自身が渋滞に巻き込まれた状況を写真付きで投稿している事例も確認できた。そのため、投稿者自身が何らかの交通現象に巻き込まれた際には情報を提供する可能性がある。

以上から、課題2である「交通現象等の情報を提供する有用な投稿者を推定できないこと」は本手法の信頼性高象限、信頼性高象限の投稿者を抽出することで解消できる。また、特定道路の利用に関する経験情報を考慮する方がより信頼性の高い投稿者を取得できる。

5.5.2 有益な情報の抽出漏れ

本検証の抽出結果を確認すると、3章で述べた傾向に近

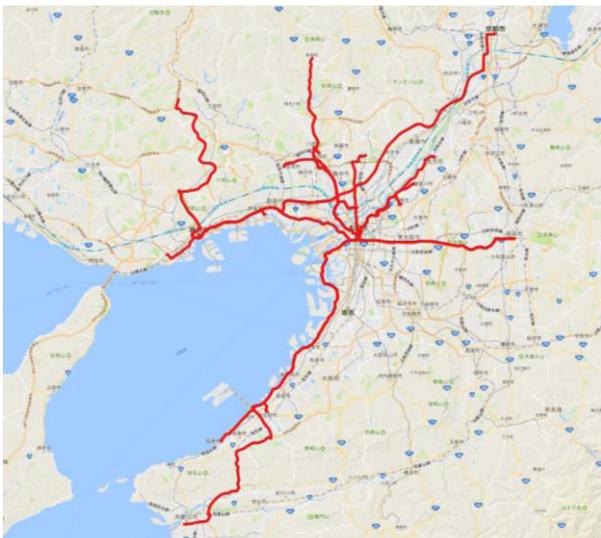


図6 投稿者の移動経路可視化結果

Fig. 6 Visualization of moving routes of microbloggers. (出典: Google Map (2017年12月))

い結果となったが、例外も確認された。具体的には、信頼性最低象限の投稿者でも「吹田経由で行くと渋滞だ、池田から入ろうとしても渋滞だ。」というような有用なツイートをしている投稿者も10名存在していた。これらの投稿者は取得できているタイムラインが少なく、適切に経験情報を分析できていなかった。そのため、今後は漏れなく正確に信頼性を分析するために、興味関心の分析等、より詳細な投稿者の属性分析や投稿者間の関係性の分析も視野に入れる。

6. 道路管理に対するツイートの活用方策

前章までの研究成果により、ツイートから道路管理に有益な交通行動の所感の抽出が可能であることが示唆された。本章では、本研究の成果およびTwitterのツイートが道路管理の高度化に寄与できる可能性を考察する。1章で述べたとおり、道路交通分析では、プローブデータ等の定量的なデータを用いた分析は非常に活発である。一方、定性的な分析はアンケート調査等の方法が一般的であるが、頻繁に実施するのが難しい。その現状をふまえ、道路交通分析の支援方策として、本研究で得られた投稿者の移動経路やその間に投稿された意見の活用方法を道路管理者や道路交通分析者と議論を重ねて表8に示すとおりまとめた。

6.1 突発的な交通現象発生の検知

ツイートに含まれる情報によってはリアルタイム性が高いため、事故や交通規制等の交通現象の発生の検知に活用できる。図7は、阪神高速道路の高速隊が事故を認知した時間よりも先に情報が投稿された事例で、現場の写真とともに「阪神高速、環状線から長田へのところで事故ホヤホヤ、渋滞長くなることでしょー」と投稿されていた。この情報は2016年4月13日に13号東大阪線で発生した事故に関する情報であるが、阪神高速道路が保有する交通統計データでは高速隊が18時18分に認知したと記録されているのに対して、ツイートは18時16分に投稿されていた。このことから、ツイートを分析するとリアルタイムで交通

表8 現行の道路管理におけるツイートの活用方法例

Table 8 Examples of tweets utilized by road management in current.

ユースケース名	活用データ	ツイートを組み合わせた分析内容
突発的な交通現象発生の検知	道路交通情報 (事故発生箇所等)	・交通現象のリアルタイムな検知
交通現象の事後分析	道路交通情報 (事故発生箇所等)	・動画や画像を用いた事故等の交通現象の要因分析
道路利用者の移動目的の分析	プローブデータ 交通量	・高速道路 IC 通過者の目的地分析 ・高速道路 IC の利用者数分析
SA/PA や道の駅等のマーケティング	プローブデータ 交通量	・SA/PA や道の駅の利用実態分析 ・SA/PA や道の駅の利用所感分析



図 7 リアルタイムな投稿に付与された写真
 Fig. 7 Photos attached into real-time post.
 (出典：Twitter (2017 年 12 月))



図 8 交通現象に関する投稿に付与された動画
 Fig. 8 Video attached into post related to traffic phenomenon.
 (出典：Twitter (2017 年 12 月))

現象を検知できる可能性がある。

6.2 交通事故等の事後分析

ツイートには写真や動画といったメディア情報を付与できるため、これらの中には交通現象に関する情報も含まれている。図 8 はドライブレコーダで撮影された交通事故発生の映像が付与されている事例である。この事例のように、ツイートには道路管理者が入手困難な道路利用者視点での映像が含まれる場合があるため、交通現象の事後分析等に活用できると考えられる。

6.3 道路利用者の移動目的の分析

本研究の手法を用いて移動経路を分析すると、投稿者の移動の起終点や移動目的を把握できる可能性がある。ただし、交通量と比してツイート数の方が少ないため、道路利用者の意見としての代表性は確保されていないことに注意が必要であるが、たとえば、平日日時間帯別の高速道路の IC ごとの利用特性の分析材料になる。高速道路の IC ごとに混雑が予想される時間帯や突発的なイベントに応じて変動する料金の導入や、事前の情報提供等により、渋滞緩和

や安全性向上等の検討への活用が期待できる。

6.4 SA/PA や道の駅等のマーケティング

2005 年の道路公団の民営化以降、高速道路の SA/PA は単なる休憩施設にとどまらず、様々なサービスが展開されており、SA/PA に行くことを目的としたドライブの特集が雑誌等で散見されるようになってきた。また、道の駅は地域活性化の拠点として、特産物の販売等を行っている。これらの施設の利用者特性や利用目的の分析はアンケート調査で行われていることが多い。

本研究の成果を活用すると、目的地に到着するまでに立ち寄った地点やそれに対する意見も取得できるため、SA/PA や道の駅の利用目的や利用所感を得ることができる。それらに応じた改善を図ることで SA/PA の収益の増収や道の駅の活性化の一助になると考えられる。

7. おわりに

本研究では、投稿者のツイートのタイムラインから行動を推定して移動経路を抽出した結果および特定道路のツイート数から投稿者の信頼性を分析する手法を考案した。本手法の検証結果から、投稿者の習慣的な道路の利用実態を推定し、投稿者に対する信頼性を判定できる有用性を示した。そして、本研究の成果を活用して収集したツイートの活用方策を考察した。

今後の課題として、投稿者の 1 日の平均投稿数によって行動推定精度が低下する課題や、漏れなく正確な信頼性を分析する手法の考案等があげられる。また、今回は阪神高速道路を対象として検証を実施したが、著者らはその他の高速道路、一般道のキーワードを含むツイートも並行して収集している。高速道路、一般道に関わる情報提供者という領域に限定すると、ツイート数や位置情報の推定詳細度には留意が必要であるが、キーワードを変更するのみで阪神高速道路と同様に信頼性を分析できる可能性がある。今後、他の地域でも同様の結果が得られるかを検証する。そして、今回の研究成果を活用して、災害時を念頭に置いた情報収集システムの開発にも取り組む。また、事故や渋滞だけでなく大雨や台風の頻繁に発生する自然現象等、様々な現象に関するデータや、道路管理者が保有する道路交通に関する統計データ等とツイートとの相関を分析し新たな知見を得る研究を進めていく。

謝辞 本研究の遂行にあたり、阪神高速道路株式会社の安東千夏氏、岡山真人氏から貴重なご意見を賜った。ここに記して感謝の意を表する。

参考文献

- [1] 今井龍一, 深田雅之, 重高浩一, 矢部 努, 牧村和彦, 足立龍太郎: 多様な動線データの組合せ分析による都市交通計画への適用可能性に関する考察, 土木計画学研究・講

- 演集, Vol.48, No.134, pp.1-9, 土木学会 (2013).
- [2] 井上 直, 石神孝裕, 石井良治, 中野敦, 菊池雅彦, 前川敦: 交通関連ビッグデータを踏まえた総合都市交通体系調査のあり方, 土木計画学研究・講演集, Vol.53, No.15-10, pp.2142-2148, 土木学会 (2016).
- [3] 中野 敦, 森尾 淳, 菊池雅彦, 井上 直: 交通関連ビッグデータとパーソントリップ調査の特徴に関する一考察, 土木計画学研究・講演集, Vol.53, No.15-01, pp.2077-2082, 土木学会 (2016).
- [4] 新階寛恭, 今井龍一, 池田大造, 永田智大, 森尾 淳, 矢部努, 重高浩一, 橋本浩良, 柴崎亮介, 関本義秀: 携帯電話網運用データに基づく人口流動統計とパーソントリップ調査手法との比較による活用可能性に関する研究, 土木計画学研究・講演集, Vol.53, No.15-03, pp.2083-2094, 土木学会 (2016).
- [5] 中矢昌希, 白水靖郎, 松島敏和, 田中文彬, 立川太一, 池田大造, 永田智大, 新階寛恭, 今井龍一: 都市交通分野における人口流動統計データの活用に向けた一考察—近畿パーソントリップ調査との比較によるデータの特長と課題に関する分析, 土木計画学研究・講演集, Vol.53, No.15-04, pp.2095-2103, 土木学会 (2016).
- [6] 今井龍一, 池田大造, 永田智大, 福手亜弥, 金田穂高, 重高浩一, 鳥海大輔, 廣川和希: 携帯電話網の運用データを用いた人口流動統計から算出した自動車 OD 量と道路交通センサとの比較分析—道路交通分野へのモバイル空間統計の適用可能性, 土木計画学研究・講演集, Vol.53, No.23-02, pp.619-627, 土木学会 (2016).
- [7] 鈴木彰一, 田中良寛, 佐治秀剛, 牧野浩志: ITS スポットを用いた特殊車両の走行状況確認方法の提案と実験データによる検証, 交通工学論文集, Vol.1, No.2, pp.47-52, 交通工学研究会 (2015).
- [8] 橋本浩良, 水木智英, 高宮 進: プローブデータを利用したボトルネック交差点とその影響範囲の特定方法, 土木学会論文集 D3 (土木計画学), Vol.70, No.5, pp.L1159-L1166, 土木学会 (2014).
- [9] 新階寛恭, 池田大造, 永田智大, 森尾 淳, 石井良治, 今井龍一: 携帯電話網の運用データに基づく人口流動統計におけるトリップ目的推定手法に関する研究, 土木計画学研究・講演集, Vol.55, No.26-07, pp.1-10, 土木学会 (2017).
- [10] 矢野晋哉, 伊藤秀昭, 安田幸司: 道路開通に関するツイッター情報の分析事例, 土木計画学研究・講演集, Vol.48, No.203, pp.1-5, 土木学会 (2013).
- [11] 今井龍一, 中村健二, 田中成典, 藤本雄紀: マイクロブログを用いた道路事業に係わる意見の抽出と位置情報推定に関する研究, 土木学会論文集 F3 (土木情報学), Vol.71, No.2, pp.L188-L195, 土木学会 (2015).
- [12] 今井龍一, 高橋哲朗, 田嶋聡司, 山影 譲, 重高浩一: 道路事業に係わる行政相談資料及び Twitter のつぶやきに対するテキストマイニング技術の適用—道路事業評価の高度化支援に向けた一考察, 土木計画学研究・講演集, Vol.48, No.205, pp.1-8, 土木学会 (2013).
- [13] 今井龍一, 西澤一希: 道路交通分析におけるマイクロブログの適用可能性の一考察, 土木学会関東支部技術研究発表会講演概要集, Vol.43, No.IV-21, pp.1-2, 土木学会 (2016).
- [14] 梅島彩奈, 宮部真衣, 荒牧英治, 灘本明代: 災害時と平常時 Twitter におけるデマとデマ訂正ツイートの特徴分析, 電子情報通信学会技術研究報告, Vol.111, No.361, pp.59-64, 電子情報通信学会 (2011).
- [15] Castillo, C., Mendoza, M. and Poblete, B.: Information Credibility on Twitter, *Proc. 20th International Conference on World Wide Web*, pp.675-684, ACM (2011).
- [16] 池田佳代, 田邊勝義, 奥田英範, 奥雅 博: Blog からの体験情報抽出, 情報処理学会論文誌, Vol.49, No.2, pp.838-848, 情報処理学会 (2008).
- [17] 中嶋勇人, 新妻弘崇, 太田 学: 位置情報付きツイートを利用した観光ルート推薦, データベースシステム研究会研究報告, Vol.158, No.28, pp.1-6, 情報処理学会 (2013).
- [18] 杉谷卓哉, 白川真澄, 原 隆浩, 西尾章治郎: 教師あり機械学習を用いたツイート投稿時のユーザ位置推定手法, データベースシステム研究会研究報告, Vol.158, No.26, pp.1-8, 情報処理学会 (2013).
- [19] Schulz, A., Hadjakos, A., Paulheim, H., Nachtwey, J. and Mühlhäuser, M.: A Multi-Indicator Approach for Geolocalization of Tweets, *Proc. 7th International AAAI Conference on Weblogs and Social Media*, pp.573-582, AAAI (2013).
- [20] 田中成典, 中村健二, 寺口敏生, 中本聖也, 加藤 諒: マイクロブログから抽出したユーザの習慣に基づく行動推定に関する研究, 情報処理学会論文誌: データベース, Vol.6, No.3, pp.73-89, 情報処理学会 (2013).
- [21] Twitter 社: Twitter Privacy Policy, 入手先 (<https://twitter.com/privacy?lang=en>) (参照 2017-12-22).
- [22] Twitter 社: Twitter Developers, 入手先 (<https://dev.twitter.com/>) (参照 2017-12-22).
- [23] Google 社: Google Maps Geocoding API, 入手先 (<https://developers.google.com/maps/documentation/geocoding/intro?hl=ja>) (参照 2017-12-22).
- [24] 工藤 拓: MeCab, 入手先 (<http://mecab.sourceforge.jp/>) (参照 2017-12-22).
- [25] Google 社: Google Maps Directions API, 入手先 (<https://developers.google.com/maps/documentation/directions/?hl=ja>) (参照 2017-12-22).
- [26] Cheng, Z., Caverlee, J. and Lee, K.: You are Where You Tweet: A Content-based Approach to Geo-locating Twitter Users, *Proc. 19th ACM International Conference on Information and Knowledge Management*, pp.759-768, ACM (2010).
- [27] Pennacchiotti, M. and Popescu, A.: Democrats, Republicans and Starbucks Afficionados: User Classification in Twitter, *Proc. 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.430-438, ACM (2011).
- [28] 高橋哲朗: 発言同期を用いたマイクロブログ著者の位置推定, 自然言語処理研究会研究報告, Vol.214, No.17, pp.1-7, 情報処理学会 (2013).
- [29] 国土交通省: 近畿圏の新たな高速道路料金導入後の交通状況について, 入手先 (<http://www.mlit.go.jp/common/001210996.pdf>) (参照 2018-04-24).
- [30] 大津展之: 判別および最小 2 乗規準に基づく自動閾値選定法, 電子情報通信学会論文誌, Vol.J63-D, No.4, pp.349-356, 電子情報通信学会 (1980).



藤本 雄紀 (学生会員)

1991 年生. 2014 年関西大学総合情報学部総合情報学卒業. 2016 年同大学大学院総合情報学研究科知識情報学専攻博士課程前期課程修了. 修士 (情報学). 現在, 同大学院総合情報学研究科総合情報学専攻博士課程後期課程在学中. 2012 年 (株) 関西総合情報研究所入社, 現在に至る. ビッグデータ解析, システム設計等の研究開発に従事.



今井 龍一 (正会員)

1975年生。2000年関西大学大学院工学研究科土木工学専攻博士課程前期課程修了。2000年日本工営(株)入社。2009年3月博士(工学)東京大学。2010年8月国土交通省国土技術政策総合研究所。2015年4月から京都市大学准教授、現在に至る。2016年度文部科学大臣表彰科学技術賞「科学技術振興部門」受賞。



中村 健二 (正会員)

1981年生。2009年関西大学大学院総合情報学研究科総合情報学専攻博士課程後期課程修了。博士(情報学)。現在、大阪経済大学情報社会学部教授。2016年度文部科学大臣表彰科学技術賞「科学技術振興部門」受賞。



田中 成典 (正会員)

1963年生。1988年関西大学大学院工学研究科土木工学専攻博士課程前期課程修了。博士(工学)。現在、関西大学総合情報学部教授および社会空間情報科学研究センター長。2016年度文部科学大臣表彰科学技術賞「科学技術

振興部門」受賞。



有馬 伸広

1967年生。1992年北海道大学大学院工学研究科土木工学専攻修士課程修了。現在、阪神高速道路株式会社計画部調査課長。交通調査、ITSその他交通技術に関する企画開発を担当。



荒川 貴之

1975年生。1997年和歌山工業高等専門学校土木工学科卒業。2011年阪神高速技研株式会社入社。現在、阪神高速技研株式会社システム事業本部。