(22) 道路利用者のマイクロブログ分析の可能性評価

藤本　雄紀[1]・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]

[1]非会員　関西大学大学院　総合情報学研究科（〒569-1095 大阪府高槻市霊仙寺町2-1-1）/
国土交通省国土技術政策総合研究所　社会資本マネジメント研究センター　社会資本情報基盤研究室
E-mail: k398881@kansai-u.ac.jp

[2]正会員　東京都市大学准教授　工学部（〒158-8557 東京都世田谷区玉堤1-28-1）
E-mail: imair@tcu.ac.jp

[3]正会員　大阪経済大学准教授　情報社会学部（〒533-8533 大阪市東淀川区大隅2-2-8）
E-mail: k-nakamu@osaka-ue.ac.jp

[4]正会員　関西大学教授　総合情報学部（〒569-1095 大阪府高槻市霊仙寺町2-1-1）
E-mail: tanaka@res.kutc.kansai-u.ac.jp

[5]正会員　阪神高速技研株式会社　システム事業本部（〒550-0011 大阪市西区阿波座1-3-15）
E-mail: nobuhiro-arima@hanshin-tech.co.jp

[6]非会員　阪神高速技研株式会社　システム事業本部（〒550-0011 大阪市西区阿波座1-3-15）
E-mail: takayuki-arakawa@hanshin-tech.co.jp

道路管理では，道路上に設置された機器や車載機器から得られる様々な道路交通データを用いて，渋滞予測や，道路整備や改良の効果計測等の分析が行われている．しかし，これらの道路交通データは，定量的・客観的であるため，道路利用者の所感等の主観的なことは明らかにはできない．

本研究では，道路利用者の移動目的や所感等を含む移動情報をマイクロブログから推定する分析手法を考案し，交通量変化の要因分析等での活用可能性を評価した．

Key Words : *microblog, road user, movement records, text mining*

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，逸早く適切に対応・改善することが求められる．そのため，道路管理者はトラフィックカウンター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．また，車載機器等から得られるプローブデータを用いて，交通流の把握，潜在的事故危険地点の特定や道路整備・改良後の効果計測等が行われている[1)2)]．これらの取り組みにより，交通現象を定量的かつ客観的に捉えることができる．一方，道路利用者の所感等の主観を得ることはできない．

近年広く普及したSNS（Social Networking Service）の一種であるマイクロブログのTwitterに注目すると，日々膨大なつぶやき（以下「ツイート」という．）がインターネット上に投稿されている．これらのツイートの中には，道路利用に係わる内容も含まれており，その内容を解析する既存研究[3)6)]がある．その一例として，ツイートを解析して道路事業に係わるニーズや感想を取得できることを示した研究[3)4)]や，観光地での移動情報を分析した研究[5)6)]がある．これらの研究成果を活用することで，道路利用に係わる所感等の主観的な情報を取得できる．

例えば，交通量等の定量的なデータに対して，道路利用者の意見等の主観的なデータを合わせると，交通量変化の要因分析等の活用に期待できる．著者らは，定量的・主観的データの特長に着目し，現状の道路管理から得られる情報と所感等の道路利用に係わる情報とを関連付ける研究に取り組んでいる[3)]．

本研究の目的は，定量的なデータでは把握ができない通勤や観光等の道路利用者の移動目的や行動をツイート

表-1　ツイートの収集結果

項目	ツイート数（件）	投稿者数（人）	経緯度（Geoタグ）の付与数（付与率）
一般道や高速道路に係わるツイート	2,148,672	614,840	59,160 (2.8%)
阪神高速道路に係わるツイート	22,704	13,457	2,444 (10%)
投稿者のタイムライン	604,780	983	12,851 (2.1%)

から抽出する手法を提案することとした．まず，第2章にて道路利用者のツイートを収集し，その傾向を把握する．次に，第3章にて道路利用者の移動情報の推定手法を考案する．そして，第4章にて，実証実験を行い道路利用者の移動情報の推定可能性を評価する．最後に，第5章で本研究を総括し，今後の展望を述べる．

2. 道路利用者のツイートの調査

(1) 調査目的

本調査では，道路に係わるツイートから道路利用者の行動分析の可能性を検証するため，道路に係わるツイートを収集し，その内容や特性を分析する．

(2) 調査手順

本研究では，既存研究[3]と同様に阪神高速道路に係わるツイートを調査対象とする．まず，著者らが既存研究で作成した一般道，高速道路の路線名，サービスエリア名やインターチェンジ名等の単語を収録した地名辞典を用いて，2016年1月16日から2016年6月18日までに投稿されたツイートをTwitter社提供のSearchAPI[7]を用いて収集する．次に，それらの中から「阪神高速」又は「阪高」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SearchAPI[7]にて収集する．APIを用いると，ツイートの本文，投稿時間や付与されていれば経緯度（Geoタグ）等が取得できる．また，投稿者情報はTwitter上での名前，自己紹介や居住地等が取得できる．さらに，過去に阪神高速道路に係わるツイートを投稿した投稿者は阪神高速道路の利用者である可能性が高いため，投稿者毎にツイートが時系列に並ぶログ（以下「タイムライン」という．）をstatus/user_timelineAPIを用いて収集する．そして，収集したタイムラインを確認し，Geoタグの付与率や位置情報に関する単語が含まれるツイートを確認する．

(3) ツイート収集結果

ツイートの収集結果を表-1に示す．一般道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人であった．このうち，阪神高速道路上の位置を特定できるツイートの投稿者数は983人であり，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．位置情報に関する単語が含まれるツイートは，いずれのツイートにも一定数存在している結果となった．

(4) 道路利用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．

1点目は一般道や高速道路に係わるツイートには多くの位置情報が含まれていた．一般にツイート全体（平均7,000万件/日）に対する経緯度（Geoタグ）の付与率は1%以下と言われている．本調査で収集したツイートを確認すると，一般道や高速道路に係わるツイートの経緯度（Geoタグ）の付与率は2.8%，中でも阪神高速道路に係わるツイートは10%となっている．このことから，道路を利用する投稿者はツイートに経緯度（Geoタグ）を付与する傾向があることがわかった．また，ツイートの本文を確認すると，経緯度は付与していないが，文中に位置情報を特定できる単語が含まれている場合があり，これらも含めると移動情報の推定数が向上すると考えられる．さらに，タイムラインのツイートに着目すると，自身の所在等を投稿する際に利用されている「なう」という語尾を含むツイートや，Twitterと連携した位置情報ゲームアプリから投稿された店舗やランドマークに関するツイート等，位置情報を含むツイートが多く確認できる．このことから，タイムラインのツイートの文章を時系列に分析すると，道路利用者の移動情報が推定できると考えられる．

2点目は交通量増加の要因を把握できる可能性がある．複数の投稿者が同時期に同じ位置情報を含むツイートを投稿する事例が確認できた．例えば，店舗等で開催されるイベントへ向かう内容を複数の投稿者が投稿していた．このことから，交通現象の発生の要因となるイベントの把握に利用でき，道路交通対策への活用が期待できる．

以上の調査結果から，ツイートを投稿の時系列に分析することで，定量的データからは把握できなかった道路利用者の移動目的や所感等を含む移動情報を把握できる可能性のあることがわかった．

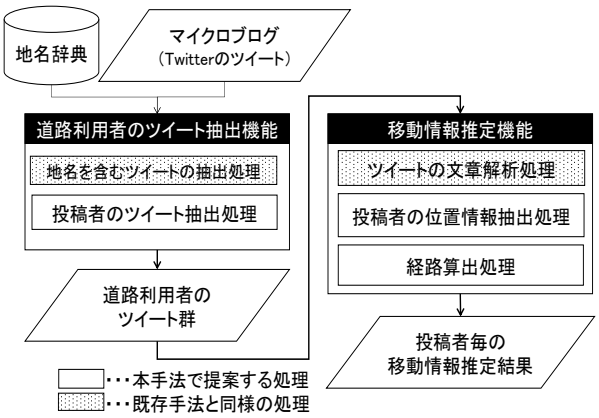

図-1　本手法の手順

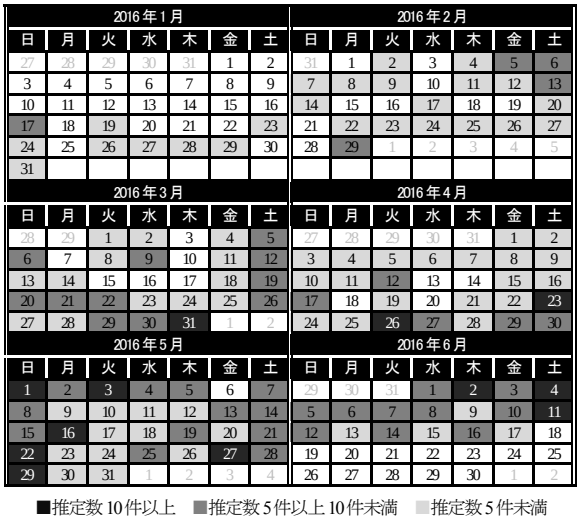

図-2　日別の阪神高速利用者の移動履歴推定結果

3.　道路利用者の移動情報の推定手法の考案

(1)　推定手法の概要

本研究では，道路利用者のツイートを時系列に分析することで，移動情報を推定する手法（以下「本手法」という．）を考案した．本手法の手順は，**図-1**に示すとおりであり，投稿者単位で発着の一連の移動情報を抽出する．**図-1**の網掛け箇所は既存研究と同様の処理であることを指す．本手法は道路利用者のツイート抽出機能および移動情報推定機能から構成される．

(2)　道路利用者のツイート抽出機能

本機能では，道路利用者が投稿したと考えられるツイートを抽出する．まず，既存研究[3]と同様に事前に作成した地名辞典を利用して対象道路のツイートを収集する．そして，そのツイートの各投稿者の前後1日分のツイートを収集して道路利用者のツイート群として出力する．

(3)　移動情報推定機能

本機能では，前述の機能にて収集した道路利用者のツイートを分析して移動情報を推定する．まず，投稿者の日別のツイートを投稿順に取得する．次に，ツイートから位置情報を抽出する．経緯度の位置情報が含まれている場合はそれを優先的に採用する．一方，含まれていない場合はツイートに対し，MeCab[9]を用いて形態素解析を行い，「地域」として判定された形態素を位置情報として抽出する．ここで，一日の中で最初に抽出した位置情報が高速道路上である場合は，投稿者の居住地を出発地として追加する．そして，取得した位置情報をジオコーディングAPI[9]を用いて座標に変換し，それらの地点を通る経路を作成する．

4.　本手法の評価実験

(1)　実験内容

本章では，道路利用者の移動情報の推定可能性を検証するため，本手法を適用して実験する．実験は，ツイートを分析して出発地および到着地を推定し，日別に投稿者数を集計する．そして，最も投稿者数が多い日の結果を地図上で可視化して分析する．実験データは，第2章で収集した2016年1月16日から2016年6月18日の154日間に投稿された阪神高速道路上での位置情報が明らかな投稿者983人とそれらに係わるタイムラインのツイート604,780件を対象とした．

(2)　実験結果

本実験で推定した移動情報を日別に集計した結果を**図-2**に示す．そして，最も推定数が多かった2016年6月11日の18トリップのうち，適切に地図上に描画できた7トリップを**図-3**に示す．**図-2**を確認すると，3月頃から土日に集中して推定数が多くなる傾向があり，観光等の目的での利用者が多いことがわかる．また，**図-3**に示すとおり，高速道路上だけでなく，一般道を含む移動情報を推定できることがわかる．分析したツイートの内容を確認すると，「I'm at EXPOCITY in 吹田市，大阪府」等の詳細な位置情報を示すツイートを抽出ができていた．各トリップの立ち寄り場所に着目すると，トリップ3，4がいずれも同時間帯に万博記念公園周辺を通っている．該当日時のイベントを調査すると，付近のスタジアムでサッカーの試合が開催されていたことがわかった．

一方で，「アメリカンサイズのホットドッグ完食」等，移動情報として関係がないが位置情報（ここでは「アメ

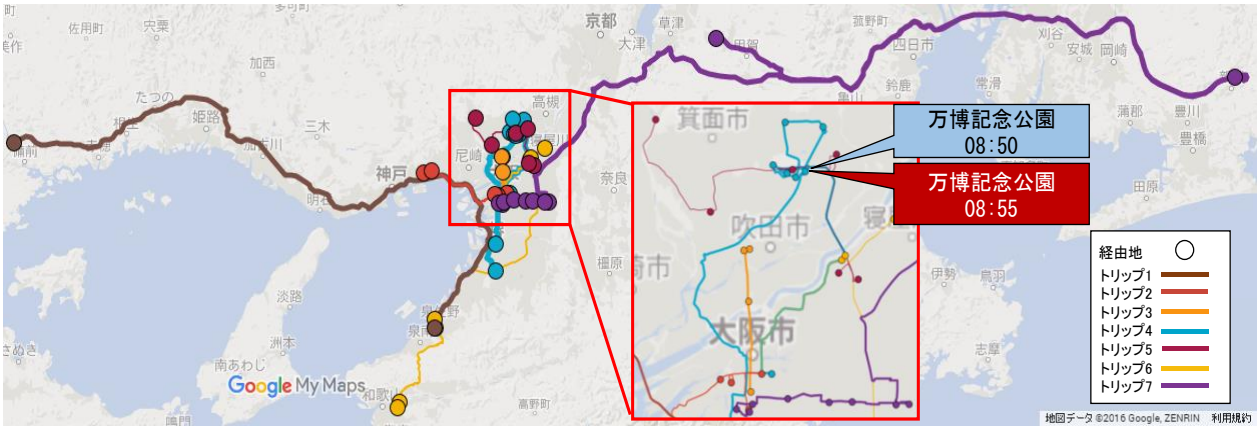

図-3 2016年6月11日の阪神高速道路利用者7トリップの情報（出典：Google Map（2016年6月））

リカ」）として判定されるものが存在しており，その場合は移動情報が正しく推定できない場合があった．

(3) 考察

本手法を用いると，ツイートから道路利用者の移動情報の推定が可能であり，道路管理者が保有している計測機器では把握できなかった一般道を含む実際の出発地および到着地を把握できる可能性のあることがわかった．また，複数の道路利用者が同じ到着地へ向かっていると，その地点のイベント等も推測ができ，交通量の変化要因の把握等に役立つことが示唆される結果を得た．

今後の課題を以下のとおり考察する．まず，本手法では過去に高速道路に係わるツイートを投稿していなければ推定ができない．しかし，それらのツイートを投稿していないが日々道路を利用している投稿者は潜在している可能性があるため，今後はスタジアムでの試合のようなイベントに関するツイートから道路利用者を推定する手法を研究し，移動情報の推定数を増加させる．

道路利用者の移動経路に着目すると，イベント等へ向かう往路での移動中には細かく位置情報を投稿するが，復路での移動中にはツイートが減少する傾向が見られた．そのため，復路が正しく推定できない場合があるが，日々の投稿パターンから時間帯別に位置情報を分析すると，復路を推定できる可能性がある．また，実際の移動情報とは直接関係のない位置情報が取得される問題は，それを除去するクリーニング処理を追加する必要がある．

5. おわりに

本研究では，道路利用者のツイートを時系列に分析することで移動情報を推定する手法を考案した．本手法の評価結果から，ツイートを分析すると主観的なデータの補完になる可能性を確認した．今後はイベントベースでの移動情報の推定や投稿数の少ない復路の移動情報の推定手法を研究し，推定精度を向上させていく予定である．

謝辞：本研究の遂行にあたり，阪神高速道路（株）の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) 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，交通工学研究会，Vol.1, No.2，pp.47-52，2015.
2) 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），土木学会，Vol.70，No.5, pp.I_1159-I_1166，2014.
3) 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），土木学会，Vol.71，No.2，pp.I_188-I_195，2015.
4) 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～，土木計画学研究発表会・講演集，土木学会，Vol.48，2013.
5) 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，研究報告データベースシステム（DBS），情報処理学会，Vol.158，No.28，pp.1-6，2013.
6) 山本理絵，宮部真衣，吉野孝，曽根原登：観光動向把握を目的としたマイクロブログにおける観光情報の特徴分析，情報処理学会関西支部 支部大会 講演論文集，情報処理学会，Vol.158，No.28，pp.1-6，2015.
7) Twitter.Inc.:Twitter Developers,<https://dev.twitter.com/>,(入手 2016.1.18).
8) 工藤拓：MeCab，<http://mecab.sourceforge.jp/>，(入手 2016.1.18).
9) Google Inc.: Google Maps Geocoding API, <https://developers.google.com/maps/documentation/geocoding/intro?hl=ja>, (入手2016.1.18).