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

東京都市大学　学生会員　○西澤　一希
東京都市大学　正 会 員　　今井　龍一

1. はじめに

道路交通分析は，路線の将来需要推計，交通渋滞の改善，事故対策，道路整備・改良の効果計測および交通状況の把握をはじめ多岐にわたる．各分析は，交通量常時観測，一般交通量調査や旅行速度調査等の交通データが用いられている．この科学的根拠に基づいた分析結果に対して，道路利用者や周辺住民の意見を常時観測して抽出し，反映することができると，道路交通分析のさらなる高度化が期待できる．

近年，意見収集や情報提供の方法のひとつとして，SNS（Social Networking Service）の活用が注目されている．SNS の一種であるマイクロブログの Twitter のつぶやき（以下，「ツイート」という．）の投稿数は 1 日 7,000 万件とも言われており，主観性のある内容が多い [1]．Twitter の活用例としては，自社製品の宣伝，商品に対するニーズの把握，マーケティング分析，遅延や事故等の交通情報の共有が挙げられる．Twitter は携帯端末からもアクセスが可能であり，何処でも誰でもリアルタイムに投稿できる．

道路交通分野における Twitter を用いた既往研究として，既存の道路交通データとツイートとを組み合わせた道路事業に係わるツイート抽出のテキストマイニング手法を考案し，その有用性を明らかにしている [1)-2)]．また，抽出したツイートを既存の道路交通分析に組み合わせることによる有効性の研究が今後の課題として述べられている．

本稿は，マイクロブログの Twitter からテキストマイニング手法により抽出した国民意見を詳細に分析した結果を報告する．

2. 研究方法

本研究の実施手順を**図 1** に示す．本研究では，まず，道路に関するツイートを収集し，実態を調査する．次に，ツイートの出現単語を類型化する．その結果を基にして道路交通分析へのツイート適用可能性を考察する．

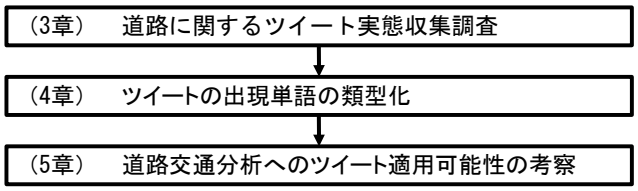

図 1　本研究の実施手順

3.　道路に関するツイート実態収集調査

本研究では，既往研究のツイート抽出のテキストマイニング手法 [1)]を用いて道路に関する意見を収集した．対象路線は近畿・関東圏の高速道路および国道とした．

ツイートの収集結果を**表 1** に示す．対象路線名および言い回しを含めた 1,201 の用語を検索ワードとしてツイートを収集し，リツイートは除去した．

関東圏で最も多く収集できた道路は，東名高速道路の検索ワード「東名」で 427 件/日であった．一方，検索ワード「東名高速道路」は 5 件/日であった．近畿圏は，阪神高速道路の検索ワード「阪神高速」で 256 件/日であった．一方，検索ワード「阪神高速道路」は 39 件/日であった．調査結果から，正式な路線名よりも略語の方が多くのツイートを収集できるため，略語を収録した地名辞典を構築し，検索性を高めていくことが今後の課題として挙げられる．

表 1　ツイートの収集結果

条件等	結果
収集期間	2014年12月25日～2015年11月16日
検索ワード	1,201
収集件数	2,971,598
処理後	2,064,686

4.　ツイートの出現単語の類型化

（1）形態素解析

本研究では，ツイートの出現単語を類型化するために，前章にて収集したツイートの内，近畿圏および関東圏の高速道路名を含むツイートに対して文章を品詞単位に分解する形態素解析を MeCab[3)]を利用して実施した．**表 2** は関東圏，**表 3** は近畿圏の解析結果の一部を示している．関東圏は，116,261 件のツイートが 1,081,865 語に分解されて 22,310 種類にグループ化された．近畿圏は，115,251 件のツイートが 1,035,987 語に分解されて 20,133 種類にグループ化された．

キーワード　マイクロブログ　道路交通分析　ツイート
連絡先　〒158-8557 東京都世田谷区玉堤 1-28-1 東京都市大学 TEL 03-5707-0104　E-Mail：g1218063tcu.ac.jp

表 2 関東圏の解析結果

順位	品詞名	出現数
1	東名	44,876
2	道	39,562
3	た	10,197
4	高速	8,490
5	圏	7,949
6	自動車	7,949
7	ます	7,853
8	現在	7,686
9	下り	7,637
10	央	7,582

表 3 近畿圏の解析結果

順位	品詞名	出現数
1	高速	58,060
2	阪神	51,603
3	た	31,045
4	道	15,423
5	だ	14,501
6	道路	13,732
7	線	12,023
8	する	11,917
9	時	11,898
10	いる	11,337

(2) ツイート出現単語の類型化

次に，本研究では前節の解析結果を用いて類型化した．主観性のある意見を抽出するために，各語を道路状態と利用感想とに大別した．道路状態は「事故」や「工事」等の道路上で起こる事象を表す単語とした．利用感想は道路利用者の感情を表す単語とした．大別した単語に対して，類語辞典を参考に異音同義語を追加した．最後に，道路交通分析の項目に関わる単語を追加した.これら追加した単語に対して，前章の解析結果の単語を分類し，**表 4** および**表 5** に示す結果を得た．括弧内は関東圏の単語の出現数を表す．近畿圏に比べて関東圏は「開通」の件数が多かった．これは，平成 27 年 10 月 31 日に桶川北本～白岡菖蒲間が開通しており，ツイートを参照すると「繋がってくれてありがとう圏央道」や「祝圏央道開通！」等の圏央道開通に関する話題があがっていたことが要因であると考えられる．また,「渋滞」や「事故」等も比較的多く，ツイートを参照すると「東名事故渋滞で」等の事象発生に関する話題が多く見受けられた．道路状態と利用感想の単語出現数を比較すると，近畿圏および関東圏ともに，道路利用者の多くは利用感想よりも道路状態に関心を示している特徴があった．

5. 道路交通分析へのツイート適用可能性の考察

前章では，ツイートの出現単語を道路状態と利用感想とに分類した．実際にこれらの単語を含むツイートを参照すると，例えば「大阪で一番怖い高速道路は 1 号環状線から土佐堀に降りようとすると，左に 4 回車線移動しないといけない阪神高速」や「阪神高速環状線の車線変更難易度は異常」等の主観性のある内容が確認できた．このような内容を抽出できると，道路利用者の視点からの道路構造や交通特性等の意見を道路交通分析に反映できる可能性があり，分析の高度化へ寄与することが考えられる．

表 4 ツイートの出現単語の類型結果（利用感想）

グループ名	利用感想
要求	期待：20（786）件 希望：57（21）件　欲しい：107（47）件
恐怖	危険：230（96）件 恐い：22（11）件　恐怖：53（23）件
不安	不安：37（117）件 泣く：43（15）件　心配：53（31）件
不快	疲れる：123（132）件 難しい：79（79）件　怒り：28（79）件
安心	安全：288（101）件 無事：117（113）件
好感	楽しい：206（115）件 大好き：63（46）件　綺麗：157（127）件
快適	快適：107（58）件 良い：517（211）件　速い：64（37）件
感謝	ありがとう：146（377）件 感謝：48（77）件

表 5 ツイートの出現単語の類型結果(道路状態)

グループ名	道路状態
混雑	混む：1003（519）件 混雑：70（105）件
渋滞	渋滞：7503（6328）件 停止69（31）件　止まる：289（134）件
開通	開通：417（1300）件 つながる：76（987）件
トラブル	事故：5119（3300）件 追突：1135（436）件　衝突918（530）件
工事	工事：1895（1350）件 舗装：18（13）件
規制	通行止め：5813（566）件 規制：214（320）件

6. おわりに

本研究では，道路交通分析へのマイクロブログの組み合わせ手法を考案するため，道路に関するツイートを収集し，高速道路に係わるツイートを類型化して特徴を分析した．

今後は，収集したツイートと道路交通分析との特徴を踏まえた上で，組み合わせの分析手法を考案して，適用可能性を検証する．

参考文献

1) 今井龍一，中村健二，田中成典，藤本雄紀：マイクロブログの解析による道路事業評価に係わる情報抽出の研究，土木情報学シンポジウム講演集，Vol.40，土木学会，2015.10.
2) 今井龍一，高橋哲郎，田嶋聡司，山影護，重高浩一:道路事業に関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，Vol.48，土木学会，2013.11.
3) 工藤拓：Mecab<MeCabhttp://mecab.sourceforge.jp>，（参照 2015.11.14）