

Twitter データによる北海道新幹線開業前後の 地域拠点ワードの共起パターン比較

A comparison of cooccurrence pattern of regional words before and after
the opening of Hokkaido Shinkansen using Twitter data

室蘭工業大学 建築社会基盤系学科

株式会社 HRS 環境防災部

室蘭工業大学大学院工学研究科 暮らし環境系領域

室蘭工業大学大学院工学研究科 暮らし環境系領域

○学生員 中田竣也 (Syunya Nakata)

正員 山本岳 (Gaku Yamamoto)

正員 浅田拓海 (Takumi Asada)

正員 有村幹治 (Mikiharu Arimura)

1. はじめに

2016年3月26日、北海道新幹線が開業し、九州、本州、そして北海道がつながり、終着駅のある函館エリアでは、観光客の入り込みが増加傾向にあることが報じられているところである。地域としては、その動向を定量的かつリアルタイムに把握し、観光拠点計画等による受け入れ体制を整えていくことが重要となる。そのためには、まず、開業前後で観光客の挙動や地域、拠点に対する関心がどのように変化したのかを明らかにする必要がある。

観光に関する調査は、古くからアンケート形式のものが採用されてきたが、迅速性に欠ける、収集範囲が限られるなどの問題が指摘されている。そこで、最近では、SNS データの活用が着目されており、任意期間における全国のユーザーのコメント（ツイート）を収集し、特定ワードの発生数などから各観光地への関心度などが分析されている^{1),2)}。しかしながら、関心先となる各地域・拠点の「つながり」について分析を試みた事例は少ない。地域や観光拠点は、単独ではなく複数が結びついて関心が持たれる場合も多い。このような、単独発生数では埋もれてしまうが、他の多くの拠点と共にコメントされる拠点を抽出することができれば、それらを含んだ新たな観光拠点計画や周遊ルートの設定などに役立つ情報となり得る。

そこで、本研究では、Twitter の投稿データから抽出した地域や観光拠点に関するキーワード（以下、地域拠点ワード）をユーザー毎に関連付けし、単独発生とともに、他の多くと共に発生（共起）されるワードを抽出する。さらに、これを2015年と2016年で比較することで、北海道新幹線開業がもたらした地域拠点への関心パターンの変化を明らかにする。

2. データ

本研究では、函館周辺への観光地に関心があるコメントを検索することを目的として、道南、道央における主要地域（函館、木古内、福島町、上ノ国、江差、厚沢部、乙部、七飯、北斗市、森町、鹿部町、知内、八雲、長万部、奥尻、せたな、瀬棚、室蘭、苫小牧、登別、壮瞥、白老、ニセコ、真狩、留寿都、ルスツ、喜茂別）を対象に、これらの27ワードを設定し、以下の期間においてツイートを検出した。なお、これらは、全国で地名に重複が無いことを条件に選んだ。検索対象期間は、2015年の4～6月（開業前）および2016年の4～6月（開業後）である。

3. 地域拠点ワードの階層別発生数

上記の27ワードで検索したツイートには、他の地域や、より詳細な観光拠点が共にコメントされている。そこで、全国の地域名に加え、道内の観光拠点をWEB

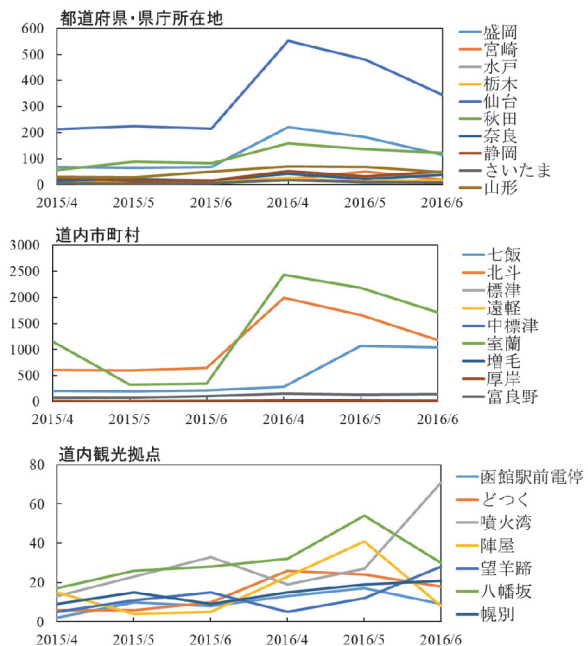


図-1 地域拠点ワードを発生させたユーザー数

サービス等で収集し、これらのワード（地域拠点ワード）を以下の3つの階層に区分した。なお、検出の重複を可能な限り少なくするため、事前にその可能性のある地名は除外した。

地域拠点ワードを発生させたユーザー数を階層別に集計した結果を図-1に示す。ここでは、全期間の発生ユーザーの合計が50人未満のワードは除外し、2015年合計値から2016年合計値への増加率が1.5以上のワードのみをプロットした。

都道府県・県庁所在地では、盛岡、仙台などの新幹線ルート上にある東北地域では開業後4月に大きな増加を見せている。秋田や山形などは、ルート上にはないが新幹線開業による移動時間短縮（50分程度）から、函館、北海道への期待が大きくなったと思われる。道内市町村では、函館近隣の七飯町や北斗市に加え、中標津や増毛などの道東地域も開業後に増加となった。観光拠点の階層になると、月別の発生ユーザー数は50人程度までに減少するが、開業後に増加するワードも幾つか見られる。また、開業後4月に大きな増加はなく、その後に次第に増加する傾向がある。

4. 地域拠点ワードの共起パターン分析

4-1 ユーザーの類型化

上記で対象としたツイートは、北海道新幹線開業に関係がないものも含まれる。そこで次に、ユーザーの特性について分析し、その類型化を行った。

まず、地域拠点ワードを函館周辺エリア（函館市、七飯町、北斗市）とその他の道内地域に関するワードに区分し、それぞれユーザー毎に開業後のワード発生数の増加量（2016 年合計数-2015 年合計数）を求めた（図-2）。開業後に両ワードの発生が増加したグループ 1 と 2 は、全体の 11% となり、その内 60% を占めるグループ 1 は、道内ワードよりも函館ワードの増加数が多い。

本研究では、北海道新幹線開業後に函館エリアの地域や観光拠点に関心度が高まった。潜在的な観光需要に着目するために、以下では上記のグループ 1 のユーザー（2,379 名）を分析の対象とする。

4-2 地域拠点ワードの中心性分析

図-1 に示したように、単独の発生数だけでは、ワード間のつながりは見えてこない。そこで、ユーザー毎に、期間中（2015 年と 2016 年）に発生した複数の地域拠点ワードを全組み合わせで結び、これを全ユーザーで束ねた共起マトリクスを作成した。図-3 はそれを可視化した例であり、2 対のワード（ノード）を結びリンクはユーザー数の情報を持っている。なお、以下では、ユーザー数が 10 未満のリンクは、分析から除外する。図-3 の場合は、2015 年に A が一定数のユーザー（10 人以上）によって B や D と共にコメントされており、2016 年にはさらに E が加わり結びつくワードが増加する。このようにノードから出る線の数、次数中心性といい、ネットワークにおける居所的な中心性を表す指標として用いられる。

地域拠点ワードを発生させたユーザー数と上記の次数中心性についてそれぞれ 2015 年から 2016 年への増加量を求め、両者の関係をプロットした（図-4）。発生ユーザー数の増加が大きいワードは、函館、札幌、小樽などの全国的にも有名な地域である（図中 A）。一方、発生ユーザー数の増加は少ないが、次数中心性の増加が大きいワードも見られた。例えば、図中 B にプロットされた、美利河、函館どつく、聖ヨハネ教会、八幡坂などの函館エリアの観光拠点がある。これらを含むツイートを調べたところ、美利河はダムや温泉などへの関心として、聖ヨハネ教会や八幡坂では函館の名所として記念に訪れたり、写真などが投稿されていた。これらは、北海道新幹線開業によって、他の多くのワードと共に関心を持たれるようになり、周遊ポイントの一つとしての存在が大きくなったと言える。

5. まとめ

本研究は、北海道新幹線の開業による各地域や拠点への関心変化を明らかにするため、開業前後の地域拠点ワードの共起パターンについて比較分析を行った。

まず、北海道新幹線開業後に発生ユーザー数が増加した地域拠点ワードを抽出した。道内では特に、北海道新幹線の駅がある北斗市や函館市などの周辺地域において、増加が顕著であった。道外では、新幹線ルート上にある東北地域において増加が目立った。次に、ユーザー毎に対象期間内でコメントされた複数のワードを関連付けて共起ネットワークを作成し、各ワードの次数中心性を算出した。発生ユーザー数の開業後増加は少ないが、次数中心性（共にコメントされた他のワードの数）の増加が大きいものとして、函館エリアの幾つかの観光拠点が抽出された。

今後の課題を以下に示す。まず、本研究では質的な分析は行っていない。すなわち、地域拠点に対する肯定的、否定的なコメントなどについて分類し、それら

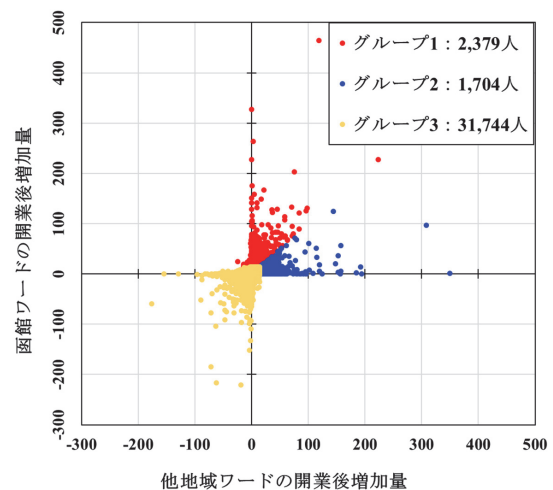
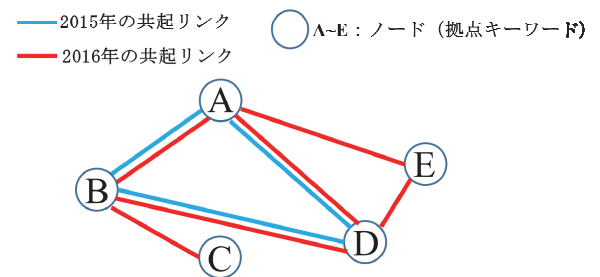


図-2 他地域ワードと函館ワードの発生増加数の関係



10人以上のユーザーが共起させた2対のみ分析対象とする。

図-3 地域拠点ワードの共起パターンと次数中心性

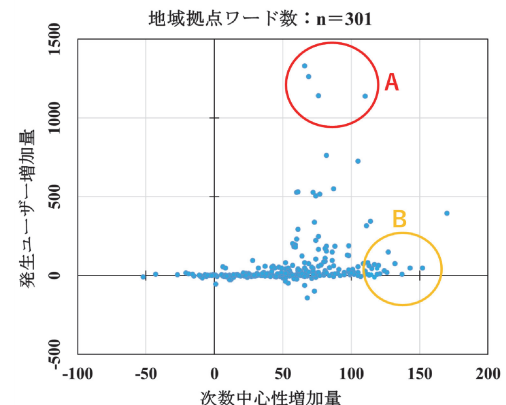


図-4 次数中心性及び発生ユーザー数の増加

を考慮した分析が必要と考える。また、ユーザーの居住地が対象地域の内外どちらなのかについて、ユーザーのコメント履歴から判定する手法を開発したい。様々な分析と照らし合わせて信頼性の検証も必要である。その上で、本手法をベースに、リアルタイムかつ広域・大規模なデータによる観光分析手法の確立を目指したい。

参考文献

- 1) 谷口守, 星野奈月, 富永透見: 自治体の観光資源に対する認識と SNS 上の「話題」とのギャップ分析, 土木計画学研究・講演集, Vol.50, P35, 2014.
- 2) 山本岳, 浅田拓海, 鈴木貴文, 有村幹治: Twitter データを用いた地域拠点ワードの共起パターン分析, 土木計画学研究・講演集, Vol.53, pp.1444-1448, 2016.