

# 新型コロナの流行による鉄道の混雑状況の変化についての研究

## The research of a change of congestion on railroads since covid-19 outbreak in Japan

今井青葉<sup>†</sup> 松田賢史<sup>†</sup> 小島一輝<sup>†</sup> 斉藤潤<sup>†</sup>  
茂木修也<sup>†</sup> 西田龍史<sup>†</sup> 池辺正典<sup>†</sup>  
Aoba Imai<sup>†</sup> Kenji Matsuda<sup>†</sup> Itsuki Kojima<sup>†</sup> Jun Saito<sup>†</sup>  
Shuya Mogi<sup>†</sup> Tatsufumi Nishida<sup>†</sup> Masanori Ikebe<sup>†</sup>

<sup>†</sup> 文教大学 情報学部

<sup>†</sup> Faculty of Information and Communications, Bunkyo University.

### 要旨

コロナ渦の中で発表された GoTo キャンペーンにより、外出をする人が増え移動間の感染リスクも高まっている。特に鉄道での三密は避けることは難しく、より一層の感染対策が必要である。この研究では Twitter における鉄道の混雑状況のツイートを収集し、それを利用した混雑状況の確認するための研究を行う。

## 1. はじめに

### 1.1. 背景

現在新型コロナが猛威を振るっている中で、GoTo キャンペーンが開始された。外出の際には鉄道等の交通機関を利用することが多く、感染情報やそのリスクを予め把握することが可能であれば有効であると考えられる。このため、本研究では鉄道の混雑状況を Twitter から収集、可視化する。

### 1.2. 鉄道について

本研究の対象となる鉄道会社は JR 東日本とする。JR 東日本は日本国内では東北地方、関東地方、甲信越地方を中心とした地域に鉄道路線を保有する。その中で本研究においては山手線と東海道本線（上野東京ライン）の混雑状況についての収集と分析を行う。山手線は日本の首都である東京の都心部で環状運転を行う路線で、2015 年以降の朝ラッシュ時混雑率は外回り、内回りとも 170%になる。上野東京ラインは、東京都から神奈川県を経由し、熱海駅がある静岡県までをつなぐ路線である。これらの路線は混雑することも多いことから混雑情報の収集に需要があると考えられる。

### 1.3. ツイッターについて

Twitter とは短文のテキスト情報を投稿することができる SNS でプログラミング開発環境向けに提供されている Twitter API を利用することで容易に目的としたツイートの取得が可能である。本研究では、Twitter API から検索機能を利用することで路線に関連するキーワード検索による情報収集から混雑情報を収集する。

## 2. 研究の概要

### 2.1. 既存研究

既存の研究では、Twitter から鉄道の運行情報や混雑に関する情報を収集する研究[1][2]が見られる。また、混雑の情報を可視化するための研究[3][4][5]が行われていることから、混雑情報に関する需要が高いことが分かる。

### 2.2. 研究の目的

本研究では、Twitter 上の投稿データを元にした、新型コロナウイルス流行に伴う鉄道機関における混雑状況の取得と可視化を目的としている。鉄道会社の発信する運行情報に加え、Twitter 上の情報を可視化することで、混雑に関する情報の把握を助けることが本研究の特徴である。

### 2.3. 処理方法

本研究では Twitter 上の情報を収集し、そのデータを反映させ混雑情報を可視化する。収集するツイートは Twitter API から検索機能を用いて、混雑に関する情報のツイートを収集する。また、路線図を可視化することで、ツイートの分布の把握を容易にすることで視覚的に混雑の情報を可視化することを目指す。

す。混雑に関する情報については、混雑は赤、空いている場合には青のように、色による直感的な把握が可能な可視化によって情報の把握を助ける。このような方法を対象である山手線と東海道線を対象として行うことで、Twitter から収集した混雑情報の可視化を行う。

本研究で開発するシステム上での地図表現は、オープンソースソフトウェアである Leaflet.js を使用する。また、地図上における駅や路線図などのデータには GeoJSON による定義を行った。駅はマーカー、路線図はラインデータとして地図上に表現し、混雑状況を色で切り替える。これにより混雑状況が視覚的に認識可能となる。また、混雑状況に関するツイートの収集は Python と Twitter API を用いる。

## 2.4. ツイートの収集処理

Python と Twitter API を用いて、鉄道の混雑状況を収集する。山手線全駅と上野東京ライン東京駅～熱海駅間の混雑状況をキーワード検索により収集する。収集の条件は「検索対象となる駅名」、「混雑状況の推定が可能な単語」の OR 検索で行う。

## 2.5. 分析処理について

対象区間および各駅の混雑状況に関するツイートを Python と Twitter API を用いて収集し、収集したツイートのうち、混雑を示すものとそうでないものに分類する。集められたツイートは python によって 1 つの csv データとして書き出し、このデータから一週間の大きな混雑状況を Leaflet.js を用いてヒートマップとして地図上に可視化する。集められたツイートはマーカーのクリックによってポップアップで画面上に表示する。駅間のラインデータに関しても同様に、収集されたツイートを表示するポップアップ画面を設定する。また、これらのツイートとの関係性の分析のために各駅の時刻表も表示する。これは、どの車両が混雑しているかを把握するためである。以上の処理を通じて可視化されたデータをもとに、Twitter 上から混雑状況の取得が可能であるか検証を行う。

## 3. おわりに

本研究では Twitter API と Python を用いて混雑情報のツイートを収集し、集められたツイートを Python により 1 つの csv データ化し、混雑情報を地図上に可視化する処理を行った。鉄道の混雑情報を集める際、「全アカウント」、「公式アカウント」、「特定アカウント」の 3 パターンで行った。その結果、全アカウントからツイートを収集することが最も効果的な方法であることが分かった。表記ゆれに対応した混雑状況推定の単語を設定した場合、公式アカウントや特定アカウントのツイートも同時に収集でき、有益な情報を多く収集することが可能になった。今回の研究から、利用者の多い路線における混雑状況を把握において、Twitter 上の投稿の利用を利用する方法は一定の効果があることがわかった。しかし、上野東京ラインにおける戸塚駅～熱海駅間では、混雑状況の推定が可能なツイートの件数が少なく、理世社の少ない地域の情報収集には課題があると考えられる。

### 参考文献

- [1] 堀江岳, 関谷聡大, 金子雄一郎: 公式 twitter の基づく都市鉄道の運転見合わせの発生状況と情報提供の実態分析, 土木学会論文集 F3, 土木情報学, Vol71, No.2, pp.II28-33, 2015.
- [2] 亓理湧, 豊田哲也, 大原剛三: 鉄道の混雑検出センサとして機能する Twitter ユーザの推定, 第 79 回全国大会講演論文集, 情報処理学会, Vol.2017, No.1, pp.521-522, 2017.
- [3] 鈴木涼平, 藤代一成: 鉄道車両における混雑度の画像計測と可視化～シミュレーションによる予備検討～, 第 76 回全国大会講演論文集, 情報処理学会, Vol.2014, No.1, pp.319-320, 2014.
- [4] 志小田雄宇, 伊藤和敬, 坂入整, 中村真純: 列車プローブデータを活用した鉄道混雑可視化システムの開発, 電子情報通信学会総合大会講演論文集, 電子情報通信学会, Vol.2016, No.1, pp.149, 2016-03-01.
- [5] 佐原亭, 坂入整: 在来線混雑可視化システムの開発, 日本鉄道技術協会誌(JREA), 日本鉄道技術協会, Vol.2019, No.1, pp.43268-43271, 2019.