

交通ビッグデータの横断的活用による移動パターン把握に関する検討

岐阜大学 学生会員 ○浅井 拓登

岐阜大学 正会員 倉内 文孝

1. 研究の背景・目的

近年、ICT の進歩と普及によりビッグデータが急速に広まっている。交通に関わるビッグデータに着目すると、携帯電話等の移動体の軌跡データ、経路探索アプリから取得される検索ログデータや比較的安価に位置情報を取得できるWi-Fi パケットセンサから得られるデータ等がある。現在、交通ビッグデータを用い、人や車等の移動について分析を試みる研究は様々である¹⁾³⁾等が、これらの研究は単一のビッグデータについて特徴の把握や活用可能性を探ることを目的としているものが多く、横断的な活用に関する研究蓄積は少ない。

本研究ではいくつかの交通ビッグデータを横断的に活用する方法を検討し、研究対象地で生じる移動パターンの把握を目指す。また、他データと組み合わせることにより、研究蓄積の少ない検索ログデータの交通計画への活用可能性を探る。

2. 使用データ

本研究では、京都市内のバス・鉄道を対象とした経路探索アプリ「歩くまち京都アプリ『バス・鉄道の鉄人』」の利用により取得されるGNSS データと検索ログデータを用いる。両データの概要を表-1、表-2 に示す。GNSS データはこのアプリを利用した際に自動的に収集される位置情報のデータである。GNSS データの取得頻度は OS により異なり、アプリ起動時は高頻度であるが、起動していない場合 Android では1時間に1回、iOS では前回取得から500 m 以上離れ、かつ5分以上経過した際に取得されることがわかっている⁴⁾。一方、検索ログデータは、経路探索を利用した際に収集される検索の履歴データである。両データで取得されるユーザーIDは共通のものが使用されており、データの結

表-1 GNSS データの概要

期間	2018年9月1日～2019年2月27日
取得項目	・ユーザーID
	・取得日時
	・緯度経度
	・OS
	・言語+国コード
	・画面タッチ情報

表-2 検索ログデータの概要

期間	2018年9月1日～2019年2月27日
取得項目	・ユーザーID
	・検索実行日時
	・出発地、目的地
	・出発地緯度経度
	・目的地緯度経度
	・検索対象日時
	・出発時刻検索(t or f)
	・OS

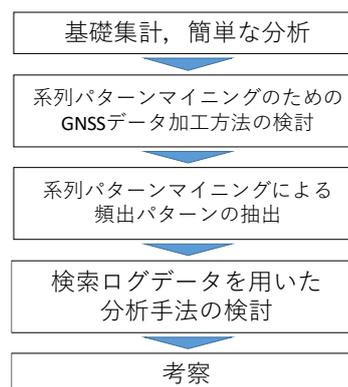


図-1 研究フロー

合が可能である。

3. 研究の流れ

研究フローを図-1 に示す。はじめに各データについて基礎集計や簡単な分析を行い、データの特徴等を把握する。次に、GNSS データについては系列パターンマイニングを採用し、頻出するトリップパターンの抽出を試みる。その前段階として、点データの連続であるGNSS データについて移動判定等の加工やデータのクリーニングをする必要があり、集計結果を参考に外れ値として削除する閾値の設定や移動滞留の判定方法を検討した。頻出パターン抽出の際には、ユ

ーザーの集計値を用いて非階層クラスタリングを行い、各クラスターの属性を推察し、どのように頻出パターンが異なっているのか考察する。

4. ユーザーのクラスタリング

頻出パターンの抽出をするにあたって、非階層クラスタリングによりデータを分類し、歩くまちアプリユーザーの属性を推定する。クラスタリングに用いる集計値として GNSS データのプロット数や検索ログデータの検索数等、47 要因を設定したものを表-3 に示す。ユーザー毎にデータを集計し、OS で分類したのものに対して k-means 法を適用してそれぞれ 4 つのクラスターへ分類した。表-4 に Android のユーザーをクラスタリングした結果を示す。紙面の都合上いくつかの集計値をピックアップし、クラスター毎の平均値を算出している。まず、最もユーザー数の多いクラスター 1 に着目すると、京都市内でのプロットが全体の半分以下であり、土日や昼に多く捕捉されているという傾向がわかる。よって、クラスター 1 は京都市以外からの訪問者であると捉えることができる。一方クラスター 0, 2, 3 は市内でのプロット割合が多く、殆どを京都市内で過ごしている人間だとわかる。特にクラスター 2 は市内でのプロット割合が高く、曜日での変動も少ないため、京都市に在住し学校や仕事等日中の活動も京都市内で行っている層であると推測できる。クラスター 0 はクラスター 0, 2, 3 の中では市内でのプロット数が比較的小さく、ある特定のエリアでのプロットが大きな割合を占めている。また、昼頃のプロット割合が減少する傾向にあるため、労働等の拠点が京都市外に存在する層が多く含まれるのではないかと推測する。

5. おわりに

本研究ではあるくまちアプリデータの基礎的な集計分析を行い、GNSS データから移動滞留を判定する方法を検討した。また、アプリユーザーをクラスタリング分析し、それぞれの属性を推察した。今後は GNSS データについて系列パターンマイニングを適用し代表的な移動パターンの抽出を行い、クラスター間での頻出パターンの違い等について考

表-3 クラスタリングに用いる集計値

集計値	概要
市内プロット数	京都市内で捕捉された回数
市内プロット割合	「市内プロット数」を捕捉回数で除したものの
市内滞在日数	京都市内で捕捉された日数
市内滞在割合	「市内滞在日数」を捕捉日数で除したものの
市内流入数	京都市外から京都市内への流入を行った回数
一日あたり市内流入数	「市内流入数」を「市内滞在日数」で除したものの
捕捉エリア数	京都市内で捕捉されたエリア数
一日あたりプロット数	「市内プロット数」を「市内滞在日数」で除したものの
曜日別プロット割合	京都市内で曜日別に捕捉された回数を「市内プロット数」で除したものの (7要因)
時間帯別プロット割合	京都市内で時間帯別に捕捉された回数を「市内プロット数」で除したものの (3時間ごと8要因)
観光エリアプロット割合	定義した観光エリアでのプロット数を「市内プロット数」で除したものの
国籍	GNSSデータで得られる国コードから東・東南アジアとその他の国に分類 (2要因)
シェア1, 2プロット割合	エリア別の捕捉数で1, 2番目の回数を「市内プロット数」で除したものの
検索数	検索ログデータ数
検索日数	検索対象とした日数
一日あたり検索数	「検索数」を「検索日数」で除したものの
曜日別検索数割合	曜日別検索ログデータ数を「検索数」で除したものの (7要因)
時間帯別検索数割合	時間帯別検索ログデータ数を「検索数」で除したものの (3時間ごと8要因)
検索数と市内プロット数比率	「検索数」を「市内プロット数」で除したものの
検索日数と市内滞在日数比率	「検索日数」を「市内滞在日数」で除したものの

表-4 Android ユーザーのクラスタリング結果

	クラスター0	クラスター1	クラスター2	クラスター3
ユーザー数	750	8148	166	415
市内プロット数	646.120	47.382	2769.307	1474.371
市内プロット割合	0.896	0.463	0.949	0.924
一日あたり市内流入数	0.145	0.491	0.110	0.130
観光エリアプロット割合	0.212	0.332	0.194	0.222
シェア1位プロット割合	0.814	0.606	0.732	0.791
シェア2位プロット割合	0.096	0.154	0.132	0.108
GNSS曜日傾向	金に少ない	土日に多い	変動少ない	金に少ない
時間帯傾向	朝ピーク	昼タピーク	朝ピーク	朝ピーク
一日あたり検索数	2.074	4.270	2.638	2.467

察する。結果に関しては講演時に報告する。

参考文献

- 1) 明光就平：Wi-Fi パケットセンサを用いた観光流動把握のための代表的移動パターン抽出に関する研究，岐阜大学大学院自然科学技術研究科修士論文，2020.
- 2) 香川喬之，桑野将司，福山敬，谷本圭志，川村尚生，菅原一孔：バス経路検索履歴データを用いた移動希望特性の分析，交通工学論文集，2巻，2号，p.A_115-A_124，2016.
- 3) 浅井拓登，杉浦聡志，倉内文孝：OD 交通量・移動滞留データを用いた属性・モード別トリップチェーン推定手法の研究，第60回土木計画学研究発表会・講演集，CD-ROM，2019.
- 4) 山木聡一郎・中西航・杉浦聡志：スマートフォンアプリ GNSS データを用いた京都市内の観光流動分析，土木計画学研究・講演集，Vol.61，CD-ROM，2020.