

Wi-Fi パケットセンサを用いた 代表的観光流動パターン抽出の検討

岐阜大学 学生会員 明光 就平
岐阜大学 正会員 倉内 文孝

1. はじめに

近年我が国における観光流動はインバウンド観光を中心に増加の一途をたどっており、それに対応する交通施策が求められている。しかし、季節変動や経年変化が大きい観光流動をアンケート調査等の従来の調査方法で把握することには限界がある。一方、近年では情報通信技術の発達により、交通の分野においてもビッグデータの蓄積が始まっている。交通ビッグデータは従来の交通流動調査では収集が困難であった詳細かつ長期間の移動データを収集可能であり、注目が高まっている。本研究では、これらの交通ビッグデータのうち、Wi-Fi パケットセンサデータに着目する。

本研究では観光地において多くの観光客が利用する経路(ゴールデンルート)を特定することを目標として、従来手法では調査が困難であった移動のつながり(トリップチェーン)の抽出方法の開発を試みる。

2. Wi-Fi パケットセンサデータ

Wi-Fi パケットセンサは Wi-Fi 通信機器がアクセスポイント探索の際に発信する MAC アドレスを収集し取得時刻と共に記録するものである。センサを複数地点に設置すれば、設置位置と取得時刻を紐づけることで移動軌跡の収集が可能となる。しかし、あくまでも Wi-Fi 通信機器の移動軌跡であるため、実交通量と整合したデータであるとはいえ、また移動方法、移動目的、移動者属性などの情報は収集不可能である。

Wi-Fi パケットセンサを用いた移動データ収集は近年各地で行われているが、本研究では、京都市が観光流動の把握を目的として 2018 年 11 月より計測した観光流動調査データを用いる。本調査では、京都駅を中心に主要観光地や鉄道駅等に 39 個の Wi-Fi パケットセンサを設置し観測が行われた。詳細な設置位置に関しては図 1 に示すとおりである。

3. データ整理

3.1 クラスタ分析

Wi-Fi パケットセンサデータは匿名であることから、目的別の交通流動の把握を実施するためには得られた情報を用いた分類などの工夫が求められる。そこで、移



図 1 Wi-Fi パケットセンサ設置位置

表 1 データの集計項目

| 分類要因 | 概要 |
|----------|---|
| 観測時間 | 同一アドレスについて最後に観測された時刻と最初に観測された時刻の差 |
| 総観測数 | 同一アドレスが捕捉された回数 |
| 捕捉日数 | 同一アドレスが捕捉された日数 |
| 捕捉センサ数 | 捕捉されたセンサの数 |
| 第一観測時 | 最初に観測された時刻 |
| 最終観測時 | 最後に観測された時刻 |
| 時間帯別観測割合 | ある時間帯の観測回数を総観測数で除したもの 1時間ごとに 24 要因 |
| 曜日別観測割合 | ある曜日の観測回数を総観測数で除したもの 各曜日別に 7 要因 |
| センサ別観測割合 | あるセンサにおける観測回数を総観測数で除したもの センサ設置地点とし 20 要因 |

動傾向より調査対象者の属性を推定する手法として Wi-Fi パケットセンサデータの特性を踏まえた上で非階層クラスタリングによる分類を行う。データは 1 パケットごとに観測されているためそのままではクラスタリングを実施することができない。そこでアドレスごとに観測された時間、地点(捕捉されたセンサ)などの情報を元に表 1 に示した項目で集計し、クラスタリングを行った。

分類結果の詳細については、発表時に報告する。

3.2 グルーピング

本研究で取り扱う Wi-Fi パケットセンサの設置位置はまばらであり一様ではない。近接し合うセンサ間程、双方で捕捉される確率は高くなるため広域な観光流動よりも狭域での移動が相対的に多く見えてしまう。この課題に対し本研究では、センサの設置密度の及ぼす影響を少なくするため、近接しているセンサ間をグルーピングしたうえで、エリア間のトリップチェーンを抽出することを試みる。グルーピング手法として、Community Detection¹⁾を採用する。この手法によるグルーピング結果には物理的距離は考慮されておらず、内々での移動が多くなるようにノード群が形成される。この手法を適用した結果は図2に示すとおりであり、この16に分けられたエリア間の移動について分析を行う。

4. 系列パターンマイニング

系列パターンマイニングとは、時間順に並べられたアイテム順列の中から頻出するものを抽出する手法である。ここでは、一つ一つの Wi-Fi 通信機器が捕捉されたセンサ ID を順列とする。今、あるアドレス r のデータ列を $T_r = \{t_{1,r}, t_{2,r}, \dots, t_{n,r}\}$ 、データベース全体を $\mathbf{D} = \{T_1, T_2, \dots, T_m\}$ とする。このデータパターンについて、前提部 X をある訪問列の順列、帰結部 Y を前提部 X の後に捕捉された地点と定義する。このような定義のもと、相関ルールとして以下の三つの指標を用いる²⁾。

- ・ 支持度 (Support)

全体パターンのうち、前提部 X 、かつ帰結部 Y を含むパターンの割合を示す。1 に近いほど頻出するパターンである。

$$\text{Support}(X, Y) = \frac{X \text{ かつ } Y \text{ を満たす訪問パターン数}}{\text{全体訪問パターン数}}$$

- ・ 確信度 (Confidence)

X が観測された際に Y が発生する条件付き確率。1 に近いほど X と Y が同時に発生、つまり X を満たすパターンは、 Y に向かう可能性が高い。

$$\text{Confidence}(X, Y) = \frac{X \text{ かつ } Y \text{ を満たす訪問パターン数}}{X \text{ を満たす訪問パターン数}}$$

- ・ リフト値 (Lift)

X かつ Y の確信度を Y の発生確率で割った値である。前提部の影響により帰結部の発生率がどの程度増加する

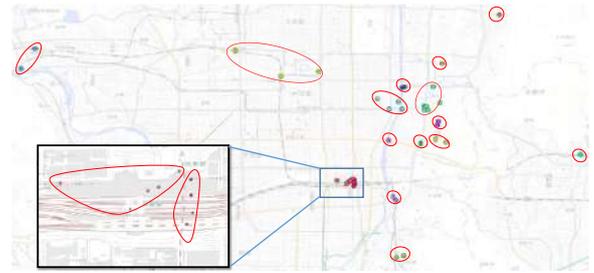


図2 グルーピング結果

のかを示しており、1 以上であれば X の影響により Y が発生しやすくなったといえる。

$$\text{Lift}(X, Y) = \frac{\text{Confidence}(X, Y)}{Y \text{ の発生確率}}$$

これらの指標を試算したうえで、三指標に対し主成分分析をおこない、三つの指標から総合的に見て代表的と考えられるパターンを抽出する。

5. おわりに

本研究では Wi-Fi パケットセンサを用いた代表的観光パターン抽出について検討した。匿名である Wi-Fi パケットセンサデータの分類にクラスタ分析が有効であること、Community Detection 法を使うことで内々の移動量が多いノード群を抽出することができ、移動量に基づいたセンサのグルーピングができることを示した。また、トリップパターンの抽出手法として系列パターンマイニングを採用し観光の出発地と回遊場所の関係について分析を行っている。こちらの結果については発表時に報告する。今後の課題として、複数時点で試算を行うことで、交通施策や観光施策の有無、時期の違いについて比較検討していきたいと考えている。また、訪日外国人の移動の特性を把握するために、外国籍の方に限った分析等を進める必要があると考えている。

謝辞：本研究は、新道路技術会議「観光流動把握を目的とした交通流動推定システムの研究開発」(研究代表者 京都大学 宇野伸宏) の研究成果の一部である。記してここに謝意を表す。

参考文献

- 1) Joerg Reichardt · Stefan Bornholdt: Statistical Mechanics of Community Detection, Physical Review E Vol. 74, 016110, 2006
- 2) 篠原透・沼尾雅之: シーケンシャルパターンマイニング拡張による特徴的なコード進行の抽出手法, DEIM Forum, 2015