

Wi-Fi パケットセンサによる行動把握のためのデータ分類に関する研究

岐阜大学 学生会員 ○伊藤 伸
 岐阜大学 正会員 倉内文孝
 宮津市 正会員 安東直紀
 (株) 社会システム総合研究所 正会員 西田純二

1. 研究の背景と目的

人の移動データは交通施策の立案や評価、災害発生時の避難計画の策定などにおいて重要な基礎情報となる。そのため、従来から交通流動調査が実施されているが、調査精度および頻度、費用面での課題が多い。近年では情報通信機器の発達により、従来調査の補完・代替となりうる詳細な移動データの入手が可能となる手法が開発されつつありその1つが「Wi-Fi パケットセンサ」である。本稿では、Wi-Fi パケットセンサにより得られたデータ（以下パケットデータ）を交通流動把握に利用するために分類を行った結果を報告する。

2. 利用データの概要

2.1 Wi-Fi パケットセンサ

携帯電話やスマートフォンなどの Wi-Fi 機器は基地局との通信のため定期的に管理パケットを発信しているが、これには Mac アドレスと呼ばれる Wi-Fi 機器固有の識別子が含まれる。Wi-Fi パケットセンサは Mac アドレスを収集、時刻、センサの地理情報、受信時の信号強度などと関連付けたデータを作成することで位置情報として扱うことを可能とした装置である。位置情報が収集できることから Wi-Fi パケットセンサを複数地点に設置することで Mac アドレスの位置情報の推移が収集可能となる。なお、Mac アドレスは個人情報保護の観点から匿名化処理が施され、A-Mac アドレス (ID) と呼ばれる匿名の状態に収集されるほか、今回使用したシステムでは週 1 回更新される仕様となっている。

2.2 宮津市における実証実験

本研究では、「観光客の域内動向、交通流動等について調査を行い、今後の地域振興につなげることを目的とし宮津市を代表する観光スポットである天橋立などを含めた 28 か所 (図 1 参照) に設置されたセンサから得られたデータを活用する。なお、周知のために観測地点には調査の趣旨説明を行っているほか、要請があった場合には速やかにデータ削除している。



図 1 Wi-Fi パケットセンサの位置

表 1 分析データの概要

	1月	2月	3月
稼働センサ数	13	28	28
ID 数 (個/日)	14,193	23,495	31,499
観測個数 (個/日)	48,588	72,381	99,695
ID あたり観測数	3.42	3.08	3.17

2.3 分析データの概要

本研究では、前述の 28 のパケットセンサにより、平成 28 年 1 月～3 月までの 3 ヶ月間にわたり記録されたデータを活用して分析を進める。表 1 にデータの概要を示す。おおよそ IID あたり 3～4 程度の観測数が確保されている。また、図 2 に各日の観測パケット数を示すが、1/9～11 の 3 連休、3/5、6 や 3/20、21 の土日などで観測数が多いことから、休日の移動について多くとらえていると期待できる。

2.4 パケットデータ利用上の課題

パケットデータは Wi-Fi 通信機器の位置情報であり、発信者の属性、移動目的、移動手段等の情報は収集不可能であるほか、Wi-Fi 通信を行う固定機器のアドレスまでも収集するためデータの利用には工夫が必要である。

3. クラスタ分析によるデータ分類の可能性

3.1 分析目的

パケットデータには人の流動とは関係のないものも含まれ、さらに「観光客の流動把握」を主目的に位置づけた場合できる限り地元住民と観光客を分類し移動を考察する必要があるが個人属性は入手できない。そのため、クラスタ分析によりデータを分類し、その結果より目的別交通流動が把握可能か検討する。

3.2 分析手法

それぞれの A-Mac アドレス (ID) を 1 データセットと見立て、それを基準とし分類を試みることにし、K-Means 法による非階層クラスタ分析を適用した。ここでは、2月の第1週のデータを活用して分析を試みた。分類要因は、ID あたりの観測回数、時間、地点などから表2のとおり設定した。クラスタ数3~5の分類を試みたが、ここではクラスタ数5の結果を示す。

3.3 分類結果の考察

表3にクラスタ数5の場合の分類結果を示す。暖色系で示したのが他クラスタと比較して値が大きいもの、寒色系が小さいものである。全観測時間がデータを分類する上で大きな要因となっていることがわかる。最もID数が多いのがクラスタ1であり平均観測数が極端に少ないなどの特徴を有しており、このクラスタは移動の分析には適さない。クラスタ2, 5は3日間以上観測されていることもあり住民の移動と推察される。従って、クラスタ3, 4が観光に関連した移動であるのではないかと推察される。クラスタ3の観測時間が60時間程度であるため宿泊を伴うもの、クラスタ4が日帰り型の観光行動の可能性がある。図4に、各クラスタのセンサごとの観測比率を示すが、クラスタ4は城崎温泉観光案内所での観測が突出しているなど、観光行動に近いものが観測されているといえる。

4. おわりに

本研究では、Wi-Fi パケットセンサデータを用い、観光流動把握を目的としたデータ分類について、クラスタ分析を適用した。今後この分類の妥当性を確認するために、これらの移動パターンの詳細な分析を進める予定であり、その結果は講演時に示す。

参考文献

- 1) 森本哲郎ら：Wi-Fi パケットセンサを用いた人流解析と可視化, DEIM Forum 2015 F8-3, 2015

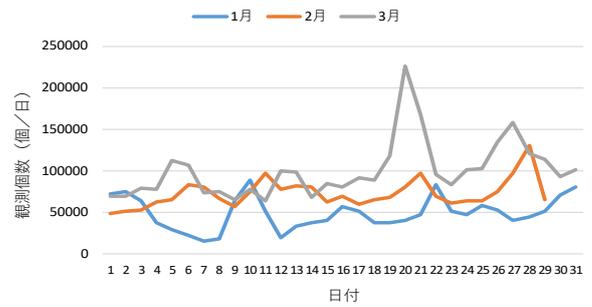


図2 日あたり観測パケット数

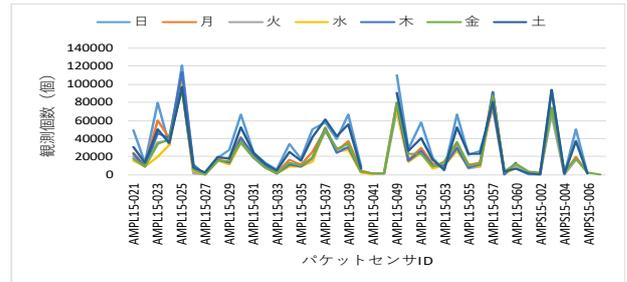


図3 センサごとの観測パケット数

表2 データの分類要因

要因	概要
全観測時間	同一IDで最後に観測された時間と最初に観測された時間の差
訪問センサ数	観測されたセンサの総数
平均観測数	同一IDで観測されたデータの個数
トリップ数	各センサ間の移動数
天橋立滞在	天橋立を経由した割合
高速利用	IC付近での観測の有無
駅利用	駅付近での観測の有無
観光地割合	観光地付近での観測の有無
宿泊割合	宿泊施設での観測の有無

表3 クラスタ分析実施結果

要因	1	2	3	4	5
ID 数	94211	1173	1435	2260	1428
全観測時間 (h)	0.18	124.58	59.38	23.50	96.05
訪問センサ数	1.26	3.44	2.69	2.69	3.08
平均観測数	1.92	77.37	17.58	11.24	26.23
トリップ数	0.36	13.28	8.81	3.20	7.09
天橋立滞在	0.00	0.00	0.01	0.03	0.00
高速利用	0.03	0.22	0.17	0.13	0.25
駅利用	0.08	0.53	0.36	0.35	0.46
観光地割合	0.08	0.02	0.02	0.05	0.02
宿泊割合	0.02	0.02	0.01	0.02	0.01

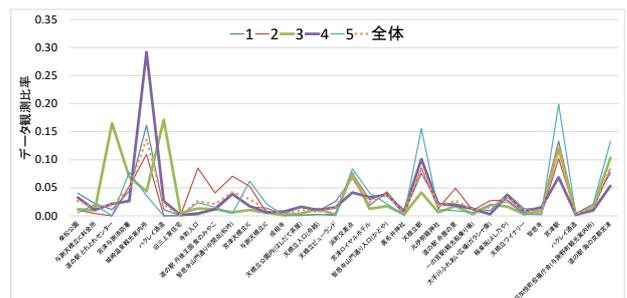


図4 クラスタごとのデータ観測比率