

# Wi-Fi パケットセンサを用いた観測間隔を考慮した代表的観光周遊パターン分析

佐竹 克仁<sup>1</sup>・宇野 伸宏<sup>2</sup>・中村 俊之<sup>3</sup>

<sup>1</sup> 学生会員 京都大学大学院工学研究科 (〒615-8540 京都府京都市京都大学桂 C1-1-209)

E-mail:satake.katsuhito.46x@st.kyoto-u.ac.jp

<sup>2</sup> 正会員 京都大学大学院工学研究科 (〒615-8540 京都府京都市京都大学桂 C1-1-205)

E-mail:uno.nobuhiro.2v@kyoto-u.ac.jp

<sup>3</sup> 正会員 名古屋大学未来社会創造機構 (〒464-8601 愛知県名古屋市千種区不老町)

E-mail:tnakamura@mirai.nagoya-u.ac.jp

京都市では観光混雑の深刻化を受け、観光客の時空間的な分散施策が推進されている。より効果的な分散施策を提案するためには、観光客の周遊行動を把握することが有効であると考えられる。本研究では Wi-Fi パケットセンサを用いた調査により得られた流動データから観光客の代表的な周遊パターンを明らかにし、訪問順の変更等による混雑緩和の可能性を考察することを目的としている。

Wi-Fi パケットセンサによる流動データから、捕捉されたセンサの数や種類、捕捉された時間帯に着目したクラスタリングにより観光客と推定されるデータを抽出した。また、各トリップにおける観測間隔の分布から、同一センサ間の移動であっても所要時間は幅広く分布しており、観光客の周遊パターン分析において観測間隔を考慮する必要性があることを示した。

**Key Words:** *tourist' excursion behavior, Wi-Fi packet sensor, nonhierarchical clustering, tourism demand management*

## 1. はじめに

観光は日本の成長戦略の一つの柱であり、国内有数の観光地である京都市においても 2014 年には「京都観光振興計画 2020」が策定され、様々な観光振興施策が実施されている。しかし、観光客の増加に伴いバスなどの公共交通の混雑、道路や歩行空間の混雑を始めとする種々の問題が深刻化しており、観光の満足度の低下だけでなく住民の市民生活の質の低下を引き起こしている可能性が高い。政府は 2020 年に訪日外国人旅行者数 4,000 万人などの目標に掲げており、今後ますますインバウンドを含めた観光客の増加が予想され、これらの問題が深刻化する懸念が生じている。

観光の満足度をできるだけ損なうことなくこの混雑問題に対処する方策として、観光客の時空間的な分散が考えられる。京都市においても、朝観光、夜観光など観光需要分散による混雑緩和が推進されているが、実際の観光客の周遊行動は把握できていないのが現状である。観光客の周遊行動をより正確に把握することにより、訪問順の変更を推進するなどさらに効果的な混雑緩和施策の提

案やその効果の検証が可能になると考えられる。

観光客の実流動を把握する方法として、本研究では Wi-Fi パケットセンサに着目する。Wi-Fi パケットセンサを利用した人口流動把握等についてはすでに複数の例がある。中西ら<sup>1)</sup>は沖縄・本部半島周辺の 13 箇所に Wi-Fi パケットセンサを設置し、センサ間の OD やトリップチェーンの推定を行っている。また伊藤ら<sup>2)</sup>は京都市の東山エリアに 20 個のセンサを設置して得られたデータから、系列パターンマイニングによって代表的な流動パターンを抽出している。しかし、これらの研究では観測間隔はほとんど考慮されておらず、捕捉されたセンサの順序のみでパターンを決定している。市内の観光周遊行動を念頭におくと、観光客が訪れるスポットすべてにセンサを設置することは困難であるから、センサに捕捉されない移動や滞留も多数存在するといえる。すなわち、捕捉された順序が同じであっても直接その移動を行った場合と観測間に別の観光行動を行った場合とが考えられ、データから得られるトリップチェーンと実態が一致しない可能性がある。観測の時間間隔に着目することで、観測間にどのような観光行動を行ったのかまでは把握でき

ないとしても、別の行動が含まれるか否かを考慮することができる可能性がある。

そこで、本研究では Wi-Fi パケットセンサにより取得した人口流動データを用いて京都市における観光客の代表的な周遊パターンを明らかにするとともに、特に東山エリアを訪れる観光客を対象として時空間的な分散による混雑緩和の可能性について考察することを目的としている。代表的な周遊パターンの抽出については、センサの設置箇所数に限界がある中でより実態に即した結果を得るために、観測の順序だけでなく観測の時間間隔にも考慮した分析を行う点に特徴がある。

図-1 に本研究の構成を示す。まず Wi-Fi パケットセンサから得られたデータから観光客と思われるデータをクラスタリング等の手法を用いて抽出する。次に、系列パターンマイニングを用いて観光客の代表的な周遊パターンを明らかにする。また、センサが比較的密に設置されている東山エリアに着目し、人口集中と混雑度合いとの関連付けを行う。これらの結果を用いて、特に東山エリアを訪問する観光客について訪問順の変更等による混雑緩和の可能性について考察する。

本論文ではこのうちクラスタリングによる観光客データの抽出および、代表的周遊パターン分析のための基礎分析として行ったセンサの観測間隔分析の結果について報告する。

なお、本研究は、国土交通省「道路政策の質の向上に資する技術研究開発」に関する技術研究助成（「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：宇野伸宏）」）の一部として実施している。

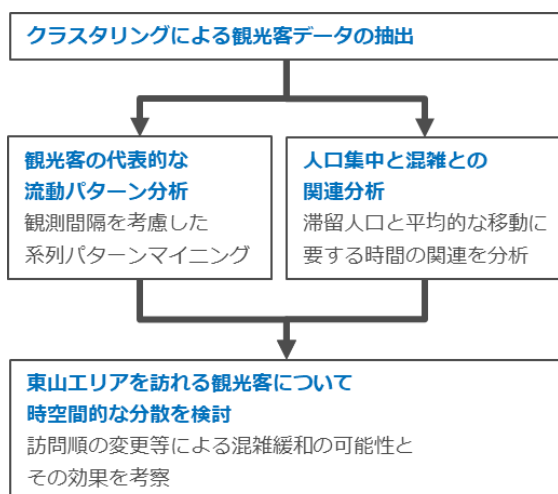


図-1 本研究の構成

## 2. Wi-Fi パケットセンサによる調査および利用データ

### (1) Wi-Fi パケットセンサの概要

近年の急速な通信機器の発達により、多くの人スマートフォンやタブレットなど Wi-Fi 機能を有する機器を持ち運ぶようになった。これらの機器は Wi-Fi 機能をオンにしているとアクセスポイントを探索するための Probe Request と呼ばれる管理パケットを送出しており、このパケットを受信・解析することによって交通流動を計測することが可能である。また、Probe Request には、原則として機器固有のアドレスである MAC アドレスが含まれており、複数箇所でのデータを統合し同一 MAC アドレスのデータを追跡することで滞留時間やトリップパターンの推定が可能である。Wi-Fi パケットセンサの仕組みや性能に関する研究としては望月<sup>3</sup>や、大田<sup>4</sup>などが挙げられる。

Probe Request 単体では個人を特定することはできないが、個人を特定できる情報と紐付けられた場合には個人の行動を追跡される恐れがある。本研究で利用するデータでは MAC アドレスを一方向ハッシュ関数により変換し匿名化した MAC アドレス（Anonymous MAC アドレス；以下 AMAC アドレスと記す）が記録されている。ハッシュ関数のシード値は週に一度変更されるため、これを超えて同一アドレスのデータを追跡することはできない点に注意が必要である。

### (2) 調査の概要

本研究で使用するデータは、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チーム、国土交通省近畿地方整備局による「京都エリア観光渋滞対策実験協議会」、西日本旅客鉄道株式会社の3主体が京都市内の主要観光施設および鉄道駅等に計 39 台のセンサを設置し京都市全域における観光流動の計測を行って得たものであり、観測期間は 2018 年 11 月下旬ごろから 2019 年 2 月末までである。図-2 および表-1 にセンサの設置箇所を記す。センサは京都市の主要観光地および鉄道駅を中心として配置されている。

### (3) 利用データの概要

本研究ではデータ取得期間のうち、以下の 3 週間（21 日間）のデータを利用した。

- ・2018 年 11 月 26 日－12 月 2 日
- ・2018 年 12 月 24 日－12 月 30 日
- ・2018 年 2 月 4 日－2 月 10 日

上記の期間に観測されたデータのうち、(1) 観測回数が 2 回以上、(2) 捕捉されたセンサが 2 箇所以上、(3) 最初の観測と最後の観測の間隔が 5 分以上、という 3 つ

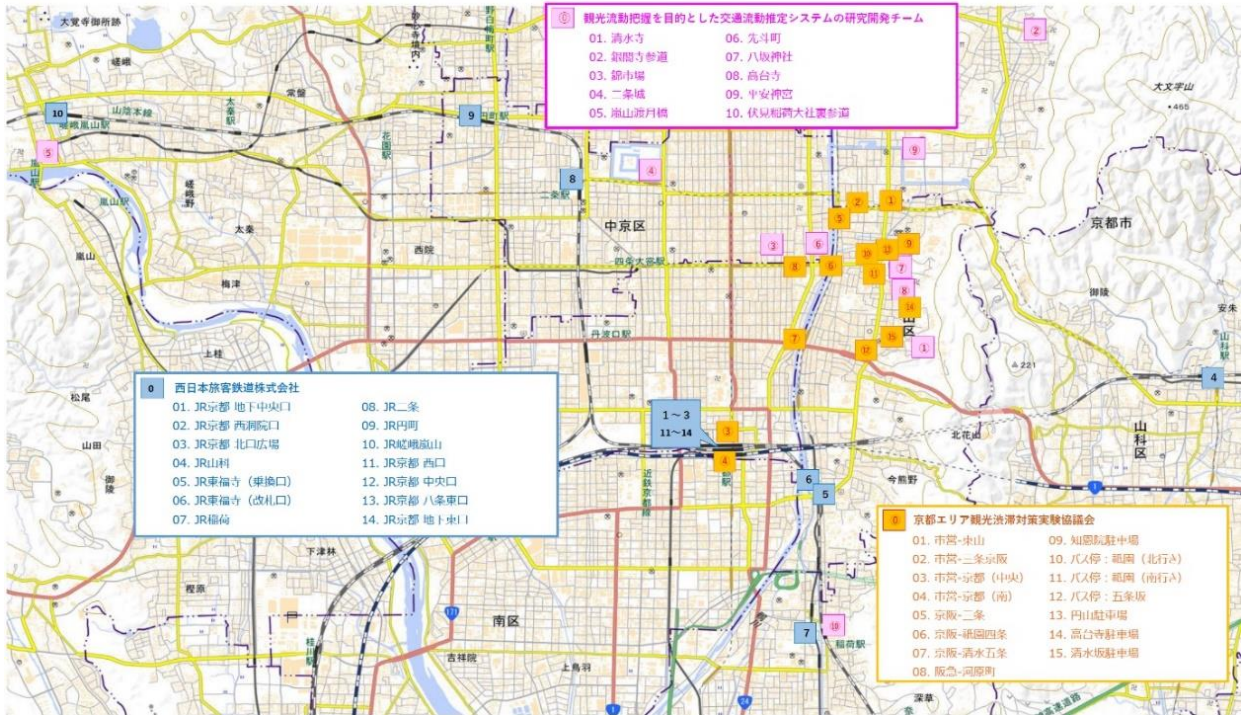


図-2 センサ設置箇所

表-1 センサ設置箇所一覧

	観光流動把握を目的とした交通流動システムの研究開発チーム	京都エリア観光渋滞対策実験協議会	西日本旅客鉄道株式会社
01.	清水寺	01. 市営-東山	01. JR 京都 地下中央口
02.	銀閣寺道参道	02. 市営-三条京阪	02. JR 京都 西洞院口
03.	錦市場	03. 市営-京都 (中央)	03. JR 京都 北口広場
04.	二条城	04. 市営-京都 (南)	04. JR 山科
05.	嵐山渡月橋	05. 京阪-三条	05. JR 東福寺 (乗換口)
06.	先斗町	06. 京阪-祇園四条	06. JR 東福寺 (改札口)
07.	八坂神社	07. 京阪-清水五条	07. JR 稲荷
08.	高台寺	08. 阪急-京都河原町	08. JR 二条
09.	平安神宮	09. 知恩院駐車場	09. JR 円町
10.	伏見稲荷大社裏参道	10. バス停: 祇園 (北行)	10. JR 嵯峨嵐山
		11. バス停: 祇園 (南行)	11. JR 京都 西口
		12. バス停: 五条坂	12. JR 京都 中央口
		13. 円山駐車場	13. JR 京都 八条東口
		14. 高台寺駐車場	14. JR 京都 地下東口
		15. 清水坂駐車場	



の条件を満たす AMAC アドレスのデータのみを分析対象データとした。なお、(2) の捕捉センサの条件では、京都駅に設置された 9 つのセンサ、三条京阪に設置された 2 つのセンサ、祇園バス停に設置された 2 つのセンサはそれぞれ同一のものとして処理した。

### 3. クラスタリングによる観光客の抽出

Wi-Fi パケットセンサから得られるデータには観光客だけでなく、地域住民や通勤通学者、Wi-Fi 機能を有する固定機器などが含まれている。本研究では観光客の流動を対象としているため、これらのデータの中から観光客と想定されるデータを抽出する必要がある。本章ではクラスタリングを用いた観光客データの抽出について述べる。

#### (1) クラスタリングの方針

京都観光総合調査<sup>5)</sup>によると、京都市への観光客のうち 7 割が日帰り、4 泊以上する観光客は全体の 1% 程度である。そこで、本研究では捕捉日数が 5 日以上、データは観光客である可能性が低いとして分析対象外とした。1 週間のうちセンサに捕捉された日数が 1 日、2 日であるデータは観光客である可能性が高いと考えられるが、取得データが限られているためデータから判別することは相対的に困難である。一方、捕捉日数が 3 日から 4 日のデータは観光客とそうでない人のどちらも一定数含まれていると考えられるが、3 日あるいは 4 日分の流動データが得られているため観測状況から観光客らしさを判別できる可能性がある。

以上から、表-2 のように捕捉日数が 3 日、4 日のデータについてはクラスタリングによる観光客の抽出を行い、1 日、2 日のデータについてはその結果を参考にして各指標の閾値を設定し、フィルタリングによって観光客と想定されるデータを抽出するという方針で分析を行う。

表-2 捕捉日数による分類

捕捉日数	観光客らしさ	取得データ量	分析方針
1日-2日	高い	少ない	フィルタリング
3日-4日	中程度	多い	クラスタリング
5日以上	低い	—	除去

#### (2) クラスタリングに使用する指標

まず、今回の調査で設置された 39 個のセンサのうち、鉄道駅に設置されたものを駅センサ、それ以外のセンサを観光地センサと分類した。ただし、「JR 稲荷」、「JR 嵯峨嵐山」は主要観光地に近接している鉄道駅で

あるため観光地センサに含めた。今回の調査では観光地および公共交通機関に多くセンサが設置されているため、より多くのセンサに捕捉されたアドレスほどより観光客らしいと考えられる。特に観光地センサに多く捕捉されている場合は、京都市の主要観光地を周遊している観光客の可能性が非常に高いと考えられる。逆に朝夕を中心に駅センサのみに観測されているアドレスは昼間に京都の観光地を周遊しているとは考えにくく、通勤通学等の目的で駅を利用している人の可能性が高い。以上の方針に基づき、表-3 に示す各変数を用いてクラスタリングを行った。

表-3 クラスタリングに使用する変数

変数名	説明
観測回数	1 日間に観測された回数の最大値および最小値
捕捉センサ数 (個)	1 日間に捕捉されたセンサの数の最大値および最小値
捕捉観光地センサ数 (個)	1 日で捕捉された観光地センサ数の最大値および最小値
観光地センサ割合	3 日間の全観測に占める観光地センサによる観測の割合
昼間観測割合	全観測のうち、10 時から 16 時までの観測割合
初観測時間 (日)	捕捉日数のうち、1 日の最初の観測が 9 時以前であった日数
土休日日数 (日)	捕捉日数のうち土曜・日曜・祝日の日数

#### (3) 連続観測および交互観測データの修正

Probe Request は機器にもよるが数秒から数分程度に 1 回発信されるため、同じ場所に滞留した場合には同じセンサに連続して観測されることになり、観測回数が非常に大きくなるケースが考えられる。そこで、本研究では同一センサにより 3 回以上連続して観測された場合はその最初と最後のみを残す処理を行った。

また、「高台寺」と「高台寺駐車場」、および「円山駐車場」と「バス停：祇園」の 2 組のセンサにおいては双方のセンサで同時あるいは交互に連続して観測されるケースが見られたため、同様に連続 3 回以上の観測が見られた場合にはその最初と最後のみを残し、センサ番号を統合し固有のものに変更する処理を行った。

(4) クラスタリング手法

前述の各変数を利用し、k-means 法による非階層クラスタリングを実施した。クラスタリングは以下の手順により行った。

- 1) 主成分分析を実施，以降は主成分得点を利用。
- 2) エルボー法によるクラスタ数の決定
- 3) k-means++法による非階層クラスタリング

エルボー法とは式(1)で表されるクラスタ内誤差平方和 (SSE 値; Sum of Squared Error) のクラスタ数の増加による SSE 値の減少幅が小さくなる点を最適なクラスタ数とする方法である。ここに、 $C_k (k = 1, 2, \dots, K)$  はクラスタ、 $\mu_i$  はクラスタ  $C_i$  のセントロイドである。

$$SSE_K = \sum_{k \in K} \sum_{x_i \in C_k} (x_i - \mu_i)^2 \quad (1)$$

また、k-means++法は k-means 法の初期値依存性の対策として、2007 年に David Arthur と Sergei Vassilvitskii によって提案された方法<sup>6</sup> であり、初期値選択を以下の手順で行うものである。これにより、より離れたデータが初期値として選ばれやすくなる。

- 1) 1 つ目のクラスタ中心  $c_1$  を全データ  $X$  からランダムに選択する
- 2) 各データ  $x (\in X)$  と最も近いクラスタ中心  $c_i$  の距離  $D(x)$  を計算する。
- 3) 新しいクラスタ中心  $c_i$  を  $x (\in X)$  から  $\frac{D(x)^2}{\sum_{x \in X} D(x)^2}$  の確率で選択する。
- 4) 2), 3) をクラスタ中心  $c_i$  が  $k$  個選ぶまで繰り返す。

(5) クラスタリングによる観光客の抽出結果と考察

捕捉日数 3 日のデータに対して上記のクラスタリング

を行った結果について述べる。図-3 は SSE 値をプロットしたものである。この結果から今回はクラスタ数を 5 とした。表-4 はクラスタごとに各指標の平均値をまとめたものである。クラスタ 1, 2 は観測回数が多く捕捉されたセンサに占める観光地センサの割合も高いことから、今回の調査でセンサを設置した箇所を中心として観光周遊行動を行っている人と推定できる。クラスタ 3 は観測回数は少ないものの、観光地センサ割合は高く、今回のセンサ配置では捉えられない観光客あるいは鉄道駅を利用していない観光客と考えられる。クラスタ 4 は週末型であり、週末に買い物目的で鉄道を利用し市内中心部を訪れている人などが想定される。クラスタ 5 は観光地センサ割合および昼間観測割合が低く、9 時以前に観測されている日数が多いことから、通勤や通学などの目的で駅を利用している人々であると考えられる。

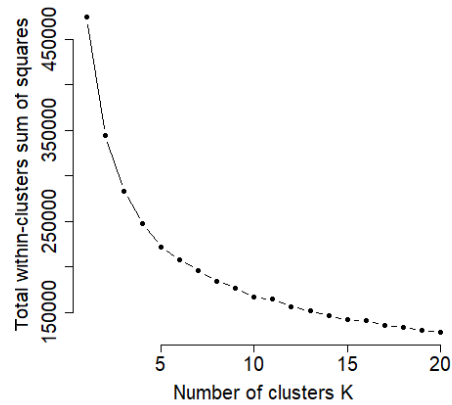


図-3 クラスタ数と SSE 値の関係

表-4 クラスタごとの各指標の平均値

クラスタ	1	2	3	4	5
観測回数の最大値 [回]	9.11	9.72	3.79	3.29	3.44
観測回数の最小値 [回]	4.63	2.33	2.10	2.07	2.14
捕捉センサ数の最大値 [個]	5.25	5.76	3.00	2.67	2.49
捕捉センサ数の最小値 [個]	3.20	2.01	2.02	2.01	2.02
観光地センサ数の最大値 [個]	3.76	4.20	2.09	1.22	0.62
観光地センサ数の最小値 [個]	1.60	0.73	1.11	0.02	0.20
観光地センサ割合	0.69	0.69	0.68	0.33	0.19
昼間観測割合	0.51	0.54	0.47	0.44	0.11
初観測時間 [日]	0.63	0.49	0.59	0.49	2.05
土休日日数 [日]	0.77	0.73	0.97	1.06	0.46
サイズ	3773	5528	14763	15525	9872

表-5はクラスタごとに主なセンサにおける捕捉割合を算出したものである。観光客と推定されるクラスタ 1, 2 では平均的に観光地センサの捕捉割合が高くなっているのは今回のクラスタリング手法からして当然のことであるが、センサごとの捕捉割合では差はあまり見られないことがわかる。クラスタ 3 では観光地での捕捉割合が場所によってはクラスタ 1, 2 に次ぐ高さとなっているが、クラスタ 4, 5 と比較しても鉄道駅の利用割合が低くなっている。クラスタ 4, 5 は山科駅での捕捉割合が高いことが特徴的であり、観光客でない鉄道利用者の可能性が高いと考えられる。

この結果から、本研究ではクラスタ 1, 2 を今回の調査で周遊を捉えられる可能性の高い観光客データとして分析対象とする。

表-5 クラスタ別センサ捕捉割合[%]

	1	2	3	4	5
清水寺	62.6	63.6	31.1	12.7	3.3
高台寺	14.9	15.7	8.8	4.1	0.9
二条城	36.2	36.4	26.3	11.3	6.3
銀閣寺参道	57.8	47.2	32.5	15.7	4.6
伏見稲荷大社 裏参道	50.3	44.9	20.9	8.1	2.4
祇園バス停	78.2	75.1	57.9	27.0	12.7
JR 山科	5.4	5.3	8.2	22.5	22.4

#### 4. 流動および観測間隔に関する基礎分析

クラスタリングにより観光客と想定されると判定したクラスタ 1, 2 のデータを対象として、同一 AMAC アドレスの連続 2 データを 1 つのトリップとして各センサ間の OD 量を算出した。表-6 は OD 量上位 15 位までをまとめたものである。ただし、発着が同一センサであるデータは除いている。この結果から、観光地とそれに近接した公共交通機関に設置されたセンサ間のトリップが多く計測されていることがわかる。また、京都駅と観光地に隣接した JR 各駅間のトリップも多く観測されている。しかし、この OD 量の算出にあたっては観測間隔を無視しており、必ずしも発着の地点間を直行しているとは言えない。

図-4 に「京都駅から JR 嵯峨嵐山」、「銀閣寺参道から JR 嵯峨嵐山」の各トリップについて、発着地間の観測時間の間隔の分布を示す。ただし観測間隔が 3 時間未満のデータのみ集計した。

京都駅から JR 嵯峨嵐山駅の移動については、30 分から 40 分に明確なピークが見られる。京都駅から嵯峨嵐山駅までの所要時間は鉄道を利用するとおよそ 20 分であり、鉄道の運行間隔やセンサの設置位置を考慮するとこの結果は妥当であるといえる。すなわち、多くの人が JR を利用し京都駅から嵐山へ移動していると推定できる。しかし、観測間隔が 1 時間以上のデータも存在しており、これらは京都駅から嵐山へ直行したとは考えにくい。一方、銀閣寺参道から JR 嵯峨嵐山の移動については、観測間隔が相対的に幅広く分布している。これは 2 地点が離れており、鉄道のような直行可能な公共交通機関が存在しないため、連続した観測であっても人により移動経路の差異、別の地点への立ち寄りがあるためと考えられる。

以上より、観測間隔の分布形状にはセンサの組により違いがあり、鉄道駅間などでは比較的時間にピークを持つが、離れたセンサ間では他地点への立ち寄りの可能性も高まり幅広い分布となる可能性があると言える。また、鉄道駅間の移動であっても直行したとは考えにくいデータも一定数存在することから、代表的な周遊パターンの分析にあたっては捕捉されたセンサの順序のみを用いるのではなく、観測間隔も考慮する必要がある。

表-6 センサ間 OD 量 (上位 15 位まで)

発	着	count
八坂神社	祇園バス停	2869
JR 東福寺 (乗換口)	京都駅	2382
祇園バス停	八坂神社	2344
伏見稲荷大社裏参道	JR 稲荷	2274
銀閣寺参道	JR 嵯峨嵐山	2045
京都駅	JR 東福寺 (乗換口)	2031
JR 稲荷	伏見稲荷大社裏参道	1994
JR 嵯峨嵐山	銀閣寺参道	1952
清水寺	清水坂駐車場	1939
JR 稲荷	京都駅	1863
京都駅	JR 稲荷	1858
JR 嵯峨嵐山	京都駅	1684
京都駅	JR 嵯峨嵐山	1602
バス停：五条坂	祇園バス停	1529
祇園バス停	バス停：五条坂	1496

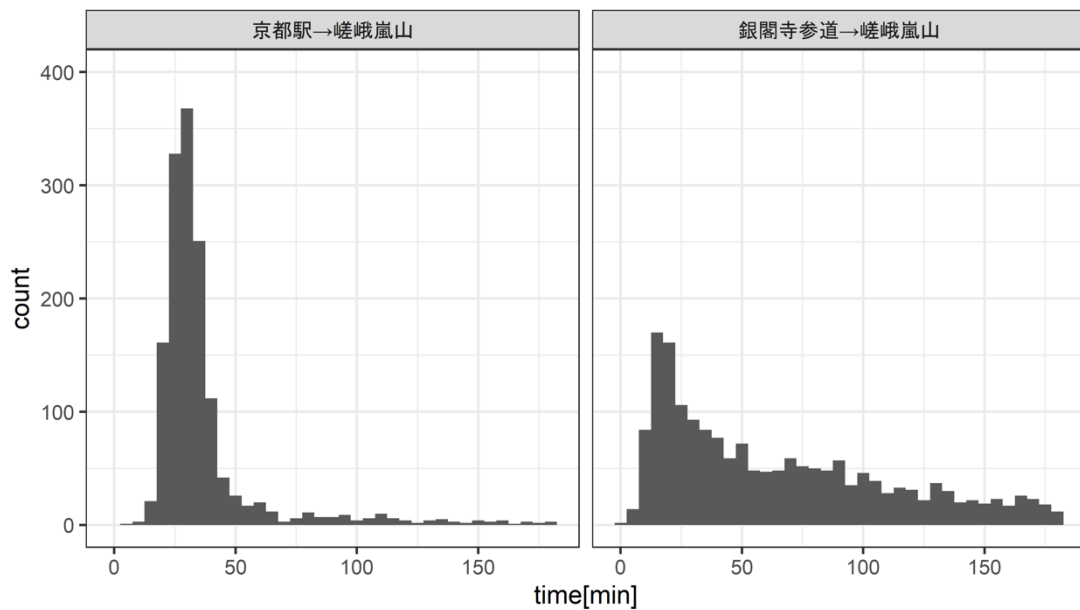


図4 観測間隔の分布

## 5. まとめと今後の展望

### (1) 現在までの成果

Wi-Fi パケットセンサにより得られた流動データから、捕捉センサ数や捕捉された時間帯に着目したクラスタリングにより、観光客と推定されるデータを抽出することができた。また、観測間隔の分布の違いから、周遊パターン分析において観測間隔を考慮する意義を示した。

### (2) 今後の展望

下記の2つの分析を通して、特に東山エリアを訪れる観光客に対して訪問順の変更等による混雑緩和の可能性およびその効果について考察を進めていく予定である。

#### a) 系列パターンマイニングによる代表的周遊パターン分析

分析対象データのうち、捕捉センサ数が一定数以上あり周遊行動が捉えられるデータを対象として系列パターンマイニングによる周遊パターンの解析を行い、代表的な周遊ルートの分析を行う。一日の周遊行動を捉えるため、観測間隔も考慮したパターンマイニングが必要である。各センサ間における観測間隔の分布の分析から着目する時間間隔を決定し、センサの順序と観測間隔が組となったパターンの抽出を行っていく。

#### b) 人口集中と混雑度合いの関連付け

観光地により観光客を含む人口のキャパシティが異なることから、滞留人口が同じであっても混雑度合いは場所により異なると考えられ、混雑度の評価のためには滞留人口と混雑度を結びつける必要がある。本研究では比較的密にセンサが設置されている清水寺や高台寺周辺を

対象として、携帯電話の運用データなどから推定されるエリア内滞留人口と Wi-Fi パケットセンサデータから算出されるセンサ間の移動に要する平均的な時間の関係から、滞留人口と混雑度の関連付けを行う。

### 参考文献

- 1) 中西航, 小林巴奈, 都留崇弘, 松本拓朗, 田中謙大, 菅芳樹, 神谷大介, 福田大輔: Wi-Fi パケットセンサによる観光周遊パターンの把握可能性: 沖縄・本部半島における検討, 土木計画学研究発表会・講演集, Vol. 55, 2017
- 2) 伊藤伸, 倉内文孝, 西田純二: Wi-Fi パケットセンサデータを用いた観光流動把握のための系列パターンマイニング, 土木学会中部支部研究発表会講演概要集, pp.431-432, 2019
- 3) 望月祐洋, 上善恒雄, 西田純二, 中野秀男, 西尾信彦: Wi-Fi パケットセンサを利用した匿名人流解析システムの構築, 情報処理学会研究報告. MBL, [モバイルコンピューティングとユビキタス通信研究会研究報告], Vol.2014, No.45, pp. 1-8, 2014
- 4) 大田香織, 大村真輝, 辻堂史子, 浅尾啓明, 西田純二: Wi-Fi 歩行者流動センサによる計測値からの実数推定手法, 土木計画学研究発表会・講演集, Vol. 57, 2018
- 5) 平成 29 年京都観光総合調査, <https://www.city.kyoto.lg.jp/sankan/cmsfiles/contents/0000240/240130/kyosa29saishu.pdf> (2019.9 入手)
- 6) David Arthur, Sergei Vassilvitskii : k-means++: The Advantages of Careful Seeding, Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA '07, pp. 1027-1035, 2007

(2019.10.4 受付)

ANALYSIS OF REPRESENTATIVE FLOW OF TOURISTS CONSIDERING  
OBSERVATION INTERVALS USING WI-FI PACKET SENSOR DATA

Katsuhito SATAKE, Nobuhiro UNO, Toshiyuki NAKAMURA