Wi-Fiパケットセンサを用いた常時データ取得システムの構築と回遊行動特性把握

東北工業大学　学生会員　○菅原直樹
東北工業大学　学生会員　岡部晃之
東北工業大学　正会員　菊池　輝
東北工業大学　正会員　泊　尚志
中央復建コンサルタンツ株式会社　正会員　末　祐介
中央復建コンサルタンツ株式会社　非会員　西村洋紀

1. はじめに

スマートフォン等のWi-Fi通信機器にはそれぞれMACアドレスと呼ばれる固有の識別子が割り当てられている．このMACアドレス(以下，ID)を取得するWi-Fiパケットセンサを用い，観光客の回遊行動を把握する取り組みが全国各地で盛んに行われている[1][2]．

本研究では狭域内において長期的なデータを取得し、安定的な取得を行うシステムの構築と取得データを用いたクラスタリングによるデータの分類を目的とする．

2.調査概要

(1)Wi-Fiパケットセンサ

本研究で用いたWi-FiパケットセンサはRaspberry Pi・無線LANアダプタ・micro SDカード等を用いた簡易に構築可能なデバイスである．これを防水・防塵ケースに内含し、屋根程度の高さに設置した**(図-1)**.

図-1 Wi-Fiパケットセンサ

このセンサはスマートフォン等のWi-Fi通信機能を内蔵した機器の発するプローブリクエスト(以下，PR)を受信し，機器のIDとその収集された時間を記録する．そのPRに各センサに対応した番号を付加し，センサ間の域内移動を観測する．そして，クラウドサービスへ定期的にPRデータとログのアップロードを行った．

(2)調査概要

調査はJR女川駅前を対象としセンサを9台設置した．データ取得期間は2019年12月1日～2020年1月3日の34日間である．なお，12月19日までの一部のセンサは取得されていない日が時折見受けられた．原因の１つとして，センサの長時間稼働が考えられる．そのため，20日以降は1日に1度再起動するよう施した．結果，20日以降のセンサの停止はほとんど見られなくなった．

またセンサの設置場所を選定するため，調査実施前に予備調査を行った．検証用のスマートフォン4台を用い，地域内の駐車場や施設の出入口・レンガ道等の地点でPRデータを取得し，おおよそのセンサ範囲を推定した．予備調査の結果を踏まえて女川駅・シーパルピア女川・駐車場の全域を網羅するようセンサの位置を決定した．各センサの設置箇所を**図-2**に示す．

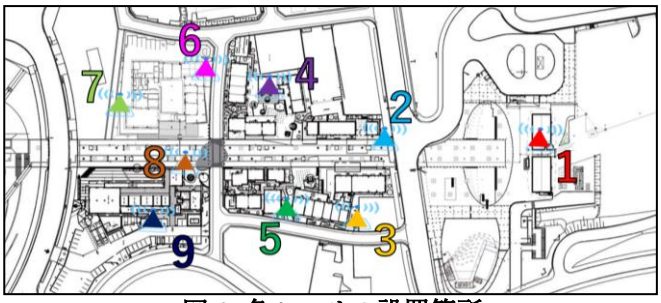

図-2 各センサの設置箇所

3. クラスタリングによる観光客の抽出

(1)データクリーニング

正確なクラスタリングを実施するために，一部のPRデータを除外する必要がある．「パケット数が1つのみのID」は滞在時間の算出が不可能な為除外した．また，「ベンダー”Google”のID」は全PRの中の9割以上を占め，且つパケット数1,2つのIDが大半でクラスタリング結果に影響を及ぼす可能性が高いため除外した．

キーワード　Wi-Fiパケットセンシング，MACアドレス，回遊行動，クラスタリング，長期データ
連絡先：〒982-8577宮城県仙台市太白区八木山香澄町35-1東北工業大学工学部都市マネジメント学科菊池研究室(022-305-3517)

(2)クラスタリング

除外後のID数は34日間合計で約14000件であった．しかしこのIDは観光客だけではなく，従業員やOA機器、地域住民等も含まれている．そこで，これらのカテゴリを分類するためにk-means法を用いたクラスタリングを実施する．使用する変数の一覧を**表-1**に示す．ここで祝日は対象地域のイベント日または年末年始の8日間，休日はイベントを除く土日の7日間，平日はその他19日間をそれぞれ設定した．**表-1**の変数の組み合わせを用い12種のクラスタリングを行った．各クラスタリングに使用した変数を**表-2**に示す．ここでは紙面の都合上No.12の結果を掲載する．結果を**表-3**に示す．

表-3よりgreenやcyan，redは滞在日数が少ないグループであることが分かる．一方blueやblackは滞在日数が多く，滞在時間も長いことからOA機器や従業員の可能性が高い．yellowやmagentaでは際立った特徴は得られなかった．次にクラスタリング結果を裏付けるために，日別・時間帯別観測ID数を集計した．紙面の都合上顕著に特徴が現れたクラスタblueの日別ID数**(図-3)**と時間帯別ID数**(図-4)**について報告する．

表-1 クラスタリングに使用した変数一覧

変数名	説明
①観測日数(日)	期間内に観測された日数
②時間帯別滞在開始時間数(個)	1日の初観測時刻を時間帯別にカウントしたもの 定めた時間帯は8時台，9-11時台，12-13時台，14-15時台，16-22時台，その他の6つ
③時間帯別滞在開始時間割合	時間帯別滞在開始時間数(個)を観測日数で除したもの
④平休祝日数(日)	観測された平日・休日・祝日のそれぞれの日数
⑤平休祝日割合	平休祝日数(日)を観測日数で除したもの
⑥平均滞在時間(分)	滞在時間の和を観測日数で除したもの
⑦平均滞在時間パラメータ	平均滞在時間の値によって1～6のパラメータを付与 1：10分未満　2：10分以上60分未満 3：60分以上90分未満　4：90分以上180分未満 5：180分以上360分未満　6：360分以上

表-2 各クラスタリングの使用変数

No.	使用変数	No.	使用変数
1	②，⑥	7	①，③，⑥
2	④，⑥	8	①，⑤，⑥
3	②，④，⑥	9	①，③，⑤，⑥
4	②，⑦	10	①，③，⑦
5	④，⑦	11	①，⑤，⑦
6	②，④，⑦	12	①，③，⑤，⑦

表-3 No.12のクラスタリング結果

クラスタ	ID数	平均滞在日数	平均滞在開始時間帯割合						平均平休祝日割合			平均滞在時間(分)
			8時台	9-11時台	12-13時台	14-15時台	16時以降	その他	平日	休日	祝日	
green	7609	1.18	3.3%	27.1%	25.1%	16.5%	17.5%	2.2%	44.2%	27.0%	26.6%	20.46
cyan	3547	1.26	3.7%	35.3%	27.3%	9.5%	10.3%	2.4%	47.6%	25.8%	23.8%	99.95
red	1298	1.45	20.1%	33.6%	10.6%	4.0%	5.5%	8.6%	41.3%	25.4%	27.9%	314.07
yellow	842	5.5	15.3%	31.3%	16.6%	10.7%	22.6%	3.5%	61.7%	17.2%	21.1%	164.45
magenta	469	11.61	26.6%	27.6%	12.5%	9.2%	19.2%	4.8%	68.7%	13.1%	18.2%	252.02
blue	442	19.33	44.9%	25.8%	8.8%	5.6%	10.6%	4.3%	73.6%	10.5%	15.9%	383.77
black	203	29.31	34.6%	33.1%	7.7%	5.0%	8.6%	11.0%	58.1%	19.8%	22.1%	510.78

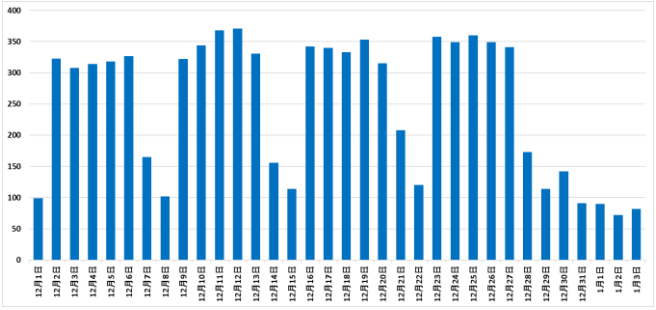

図-3 クラスタ"blue"の日別観測 ID 数

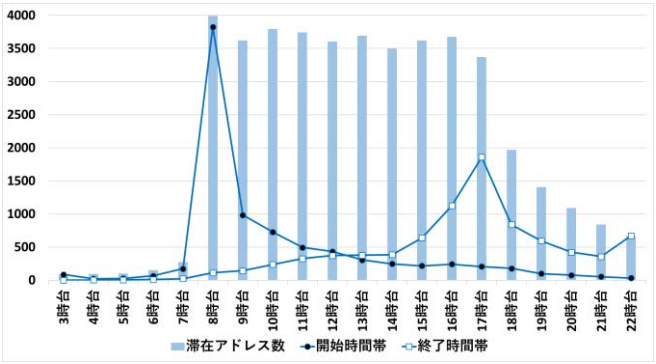

図-4 クラスタ"blue"の時間帯別ID数

図-3では平日のID数が多い一方，休日や年末年始のID数が少ないことが確認できる．また**図-4**では多くのIDが8～17時台に渡り観測されている．これらの特徴からblueには範囲内で働く従業員の携帯端末や，従業員の使用するOA機器のIDが多く該当する可能性が高い．

4. まとめ

本研究では長期的にPRデータを取得してクラスタリングを行い，取得したIDの分類を試みた．クラスタリングではOA機器や従業員など滞在時間帯等が固定化し易いカテゴリのIDに関しては意味づけができたが，観光客や地域住民等の滞在時間が多岐に渡ると思われるIDに関し明確な分類は困難であった．また，データ取得の段階ではセンサ停止の原因特定など課題が残る．今後の展開として，より顕著に分類可能な変数の考案や，安定的なデータ取得に向けての対処を進めていく．

参考文献

1) 一井 啓介，寺部 慎太郎，柳沼 秀樹，康 楠，田中 皓介：Wi-Fiパケットセンサを用いた散策型観光地における観光回遊行動の把握，第57回土木計画学研究発表会・講演集

2) 壇辻 貴生，杉下 佳辰，福田 大輔，浅野 光行：Wi-Fiパケットセンサを用いた観光客の滞在時間特性把握の可能性に関する研究-奈良県長谷寺における試み-，公益社団法人日本都市計画学会 都市計画論文集 Vol.52 No3 2017年 10月