

# 複数スマートフォンのMACアドレスランダム化の非同期性を用いたODデータ自動取得手法の提案と評価

川島 将渡<sup>1,a)</sup> 遠藤 新<sup>1,b)</sup> 垣内 正年<sup>1,c)</sup> 新井 イスマイル<sup>1,d)</sup> 藤川 和利<sup>1,e)</sup>

**概要:** 路線バスにおいて、効率的な運行のために利用者の動向把握は必要不可欠である。利用動向の把握には、OD (Origin-Destination) データの使用が最も効果的であるが、現状、OD データの取得は人手に頼った手法が主流となっている。そのため、本研究では、OD データの自動取得を目的とし、プライバシーや金銭コストの観点から、スマートフォンが発するパケットを用いた OD データの自動取得手法について着目した。しかし、プライバシー保護の観点から、ランダム MAC アドレスが使用されるようになって以降、ランダム MAC アドレスを利用するスマートフォンに対する OD データの自動取得手法は確立されていない。本研究では、Bluetooth の広告パケットから handoff や nearby といった特徴量を抽出し、特徴量の変更と MAC アドレス変更の非同期性を利用したアドレス・キャリアオーバ・アルゴリズムをベースとした、MAC アドレスランダム化の非同期性を用いた OD データ自動取得手法を提案する。京都府京都市の路線を対象に評価実験を行った結果、本提案により、ランダム MAC アドレスを使用するスマートフォンが発するパケットを用いて、OD データを自動取得できることを確認した。

## 1. はじめに

国土交通省の資料において、路線バス事業者が抱える問題が以下のように挙げられている [1].

- 路線バス利用者数の減少に伴う採算悪化により、路線バス事業者の 69%が赤字となっている
- 路線バスを運転する運転手の確保が困難となっている

こうした現状から、路線バス事業者は少ない労働力と車両数で路線バスを運行するため、各バス停間での路線バス利用者数といった利用者の動向把握に努めている。現在、路線バス事業者は利用動向の把握のために、各利用者がどの出発地 (Origin) のバス停で乗車しどの目的地 (Destination) のバス停で降車したかを示す OD (Origin Destination) を集計した OD データを活用している [2]。OD データは、表 1 に示すような形式となっている。この表では、行が乗車バス停、列が降車バス停を意味しており、乗車バス停と降車バス停の交差部分はその区間の利用者数を表している。

従来、OD データの取得は、複数の添乗調査員の目視に

表 1: OD データの形式  
Table 1 Format of OD data

	降車バス停	バス停 1	バス停 2
乗車バス停			
バス停 1		0	7
バス停 2		4	0

よる記録や、利用者へのカード配布・回収といった人手に頼った手法 [3] が主流であった。しかし、複数の添乗調査員が目視によって OD データを記録する手法では、調査員の人件費をかけないために、数か月や 1 年に 1 度しか実施できず、限定的な期間の OD データしか得られない。限定的な期間の OD データをもとにしたダイヤの調整では、季節や大規模イベントによる需要増加に対応できないため、満員のために利用者がバスに乗車できない積み残しが発生する可能性がある。他方で、利用者へのカード配布・回収によって OD データを取得する手法は、添乗調査員が運行中にカードの配布・回収を行うため、カードの配布・回収に時間がかかった場合に運行に支障をきたす。そのため、利用者数の多い混雑した路線バスでは、この手法は OD データの取得に適さない。こうした状況から、人件費等の金銭コストがかからずに OD データを自動取得する手法が必要となる。OD データを自動取得する手法は、IC カード乗車券を用いた手法、カメラを用いた手法、スマートフォンが発するパケットを用いた手法が考えられ、プライバシーや

<sup>1</sup> 奈良先端科学技術大学院大学  
Nara Institute of Science and Technology, Japan  
a) kawashima.masato.kk5@is.naist.jp  
b) endo.arata@itc.naist.jp  
c) masato@itc.naist.jp  
d) ismail@itc.naist.jp  
e) fujikawa@itc.naist.jp

金銭コストの観点から、スマートフォンが発するパケットを用いた手法が OD データを自動取得する手法として適していると考えられる。

## 2. 関連研究

### 2.1 OD データの自動取得手法

スマートフォンが発するパケットを用いた OD データの取得に関する様々な研究がなされてきた [4], [5], [6], [7], [8]. これらの研究では、各利用者はスマートフォンを 1 台のみ携帯しているという仮定のもと、各スマートフォン固有の値である MAC アドレスの有無を観測することで、その MAC アドレスに対応するスマートフォンを携帯する利用者が乗車中か否かを判断する。具体的には、まず、バス車内に存在する複数の利用者のスマートフォンが発するパケットを常に収集し、それらのパケットに記載される MAC アドレスと、パケットを収集した日時を示すタイムスタンプを記録する。そして、各 MAC アドレスに対応するタイムスタンプから、その MAC アドレスが最初に観測された時間と最後に観測された時間を特定し、これらの時間と合致する出発時刻と到着時刻を持つバス停を推定することで OD データを生成する。例えば、Dunlap らは、利用者のスマートフォンが発する BLE (Bluetooth Low Energy) の広告パケットから MAC アドレスを取得し OD データを取得する手法と、利用者のスマートフォンが発する Wi-Fi のプローブ要求パケットから MAC アドレスを取得し OD データを取得する手法 [5] を提案した。しかし、従来のスマートフォンが発するパケットを用いた OD データの取得手法を、現在の多くのスマートフォンに対して適用することはできない。プライバシー保護の観点から、iOS 8.0 もしくは Android 8.0 以降の OS を搭載したスマートフォンでは、アクセスポイントに接続されるまで、乱数生成したアドレスであるランダム MAC アドレスが使用され、この MAC アドレスは定期的に値が変更されるためである [9], [10]. 従来手法では、MAC アドレスが固定されていることを前提として、同じ MAC アドレスが継続して観測されている場合、利用者が乗車中であると推定している。したがって、ランダム MAC アドレスを用いるスマートフォンの場合、従来手法では、定期的な MAC アドレスの変更による元の MAC アドレスの消失を利用者の降車と判断するため、正しい OD データを生成することが困難になる。

### 2.2 ランダム MAC アドレスに対する手法

ランダム MAC アドレスを用いるスマートフォンが登場して以降、そうしたスマートフォンを追跡する様々な研究がなされてきた。Becker らは、BLE の広告パケットに記載される MAC アドレスと一定期間にわたってデバイス固有となるビット列である識別子をペアとして保存しておき、MAC アドレスが変更されたとき、ペアの識別子を基に新し

い MAC アドレスに更新するアドレス・キャリーオーバー・アルゴリズム [11] を提案した。Becker らは、このアルゴリズムを用いることで、最長で 53 分間にわたり iPhone を追跡できたと報告している。しかし、Becker らはこの手法を適用できない場合について、以下を挙げている。

- Android を搭載したスマートフォンの場合、BLE の広告パケットを定期的には送信しないため、アドレス・キャリーオーバー・アルゴリズムの適用が非常に困難である。
- Android もしくは iOS を搭載したスマートフォンの場合、識別子が MAC アドレスと同期して変化することが多く、その場合、アドレス・キャリーオーバー・アルゴリズムの適用が困難である。

### 2.3 ランダム MAC アドレスに対する手法の考察

2.2 節で述べた問題に対するアプローチについて述べる。1 つ目の問題に対して、新型コロナウイルス接触確認アプリである COCOA (Covid-19 Contact-Confirming Application) を用いることで解決できると考えた。COCOA の技術的仕様を検証した結果、COCOA は、他者に接触を通知するために BLE の広告パケットを定期的には送信することが明らかとなったためである。2 つ目の問題に対して、COCOA が送信する BLE の広告パケットに含まれる RSSI (Radio Signal Strength Indicator) を識別子として使用できると考えた。RSSI はゆらぐが、一定期間にわたってある範囲内に値が固定されると考えられるためである。しかし、路線バスの車内を想定した場合、RSSI を用いたアドレス・キャリーオーバー・アルゴリズムは、以下の要因によりスマートフォンの追跡精度が低下する可能性がある。

- 隣接した座席に利用者が座っている場合、RSSI が同一範囲となる可能性が高くなり、その場合、アドレスが変更された際に RSSI を用いて利用者を区別できない。
- 識別子となる RSSI のビット長が短く、識別子間の衝突が起こる可能性が高い。

上記の要因に対し、複数のスマートフォンにおける MAC アドレスの変更タイミングの違い、すなわち、MAC アドレスランダム化の非同期性に注目することで、スマートフォンの追跡精度の低下を抑制できると考えた。路線バスの車内において、車内に存在する全てのスマートフォンのランダム MAC アドレスの更新が同期する可能性は極めて低いためである。この考えを基に、複数のスマートフォンのうち、同期して MAC アドレスが変更されたスマートフォンのみをアドレス・キャリーオーバーの対象とすることで、スマートフォンの追跡精度の低下を抑制する。

本研究では、路線バスの車内環境を考慮した上記のアプ

ローチを基に、RSSI を識別子としたアドレス・キャリアオーバー・アルゴリズムによる OD データの自動取得を目的とする。

### 3. 複数スマートフォンによる MAC アドレスランダム化の非同期性を用いた OD データ自動取得手法

#### 3.1 提案手法の概要

本研究では、路線バスの環境を考慮した、複数スマートフォンによるランダム MAC アドレスの更新の非同期性を用いた OD データ自動取得手法を提案する。この提案手法では、RSSI を用いたアドレス・キャリアオーバー・アルゴリズムにより、2.2 節で述べた MAC アドレスと識別子が同期的に変わる問題を解決する。また、2.3 節で述べたように、RSSI を識別子として用いるだけではスマートフォンの追跡精度が低下するため、アドレス・キャリアオーバーの対象となる MAC アドレスを制限することで、スマートフォンの追跡精度を向上させる。さらに、本提案手法は、バスの利用者と非利用者の判別や利用者の乗降の判別を行う工夫を取り入れている。このような手法により、収集した BLE の広告パケットから、利用者とその利用者の観測開始時刻と観測終了時刻が 1 セットとなった利用者情報データを得る。路線バス用のアドレス・キャリアオーバー・アルゴリズムを用いる上で設定が必要となるパラメータを以下に述べる。

*rssi\_th* アドレスの変更により出現もしくは消失した MAC アドレスと、利用者の乗降により出現もしくは消失した MAC アドレスを判別するために必要となる RSSI の閾値であり、単位は dB である。

*passenger\_th* バスの利用者と非利用者を区別するための閾値であり、単位は秒である。つまり、*passenger\_th* 秒以上観測された MAC アドレスをバスの利用者とする。

*disappeared\_th* 消失した MAC アドレスを判別するための閾値であり、単位は秒である。つまり、*disappeared\_th* 秒以上観測されなかった MAC アドレスを消失した MAC アドレスとする。

*extract\_int* MAC アドレスが変更されて一定時間内に観測された新しい MAC アドレスを抽出するための時間間隔であり、単位は秒である。

路線バス用のアドレス・キャリアオーバー・アルゴリズムのフローチャートを図 1 に示す。このアルゴリズムは、事前に収集した BLE の広告パケットを逐次的に処理する。まず初めに、COCOA のパケットであるか確認する。COCOA のパケットの場合、COCOA のパケットか

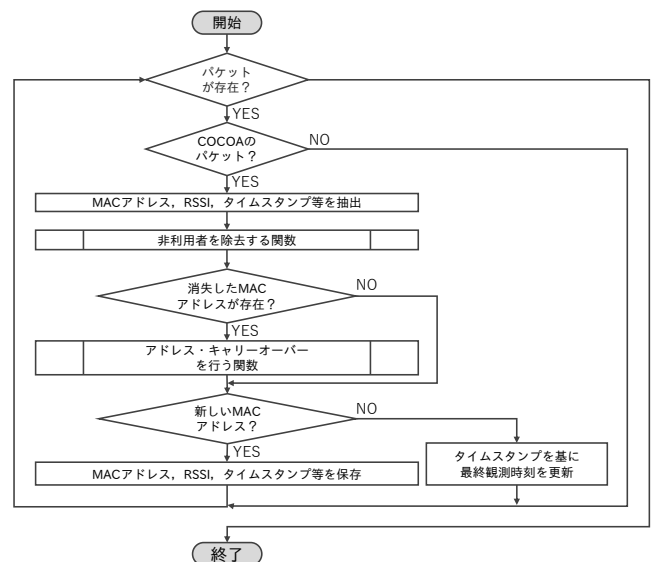


図 1: 路線バス用のアドレス・キャリアオーバー・アルゴリズムのフローチャート

Fig. 1 Flowchart of address carryover algorithm for route-bus.

ら、MAC アドレス、RSSI、タイムスタンプを抽出する。次に、非利用者を除去する関数によりバスの利用者を非利用者を判別する。この関数では、保存済み MAC アドレスに対し、MAC アドレスの観測開始・終了時刻の差が閾値以上の MAC アドレスを利用者の MAC アドレスとする。その次に、消失した MAC アドレスが存在するか確認する。消失した MAC アドレスが存在した場合、アドレス・キャリアオーバーを行う関数により処理を行う。この関数では、まず、複数のスマートフォンにおける MAC アドレスの変更タイミングの違いに着目し、アドレス・キャリアオーバーの対象となる MAC アドレスを制限する。次に、対象の MAC アドレスのうち、消失した MAC アドレスの RSSI と対象の MAC アドレスの RSSI の差が最小かつ差が閾値未満のアドレスを変更後の MAC アドレスとする。このとき、上記の条件に合致する MAC アドレスが存在しない場合、利用者の降車により消失した MAC アドレスとする。最後に、現在の MAC アドレスが新しい MAC アドレスである場合、MAC アドレス、RSSI、タイムスタンプを保存する。そうでない場合、タイムスタンプを基に MAC アドレスの最終観測時刻を更新する。

#### 3.2 提案手法を組み込んだ OD データ自動取得システム

提案手法を組み込んだ OD データ自動取得システムのワークフローを以下に示す。

- (1) パケット収集  
事前に、路線バスの 1 運行における BLE の広告パケットを収集する。
- (2) アドレス・キャリアオーバー  
前フェーズで収集した広告パケットを入力とし、3.1 節

で述べた路線バス用のアドレス・キャリアオーバー・アルゴリズムを用いてアドレスのキャリアオーバーを行い、利用者とその利用者の観測開始時刻と観測終了時刻が1セットとなった利用者情報データを得る。

### (3) OD データ生成

前フェーズで取得した利用者情報データを入力とし、ある1つの運行で停車するバス停の発車時刻と観測開始時刻と観測終了時刻が合致する利用者情報データをカウントすることで、OD データを生成する。

各フェーズについて以下に述べる。

#### 3.2.1 パケット収集フェーズ

このフェーズでは路線バスの1運行分のBLE 広告パケットを収集する。BLE 広告パケットの収集は、路線バスが搭載するCM3 (Compute Module) である Raspberry Pi をベースとしたデジタルタコグラフ車載器 (VR-1000) に Bluetooth 通信モジュールを取り付け、hcidump を用いて行う。収集した BLE 広告パケットは pcap 形式で保存する。デジタルタコグラフ車載器は運転手席のほぼ真上に取り付けられている。

COCOA を使用しているバス利用者が1名の場合、スキャンにより周囲に BLE デバイスが存在しないと判断されるため、COCOA は Rolling Proximity Identifier が変更されたタイミングでのみ広告パケットを送信ようになる。上記のタイミングでのみ広告パケットが送信される場合、収集される MAC アドレスは毎度異なると考えられるため、変更された MAC アドレスと利用者の乗車により新しく出現した MAC アドレスを判別できない。この問題を解決するため、デジタルタコグラフ車載器は一定時間ごとに Bluetooth の ON/OFF を切り替えるように設定されている。こうすることで、COCOA は新たに BLE デバイスが出現したと錯覚するため、Bluetooth の ON/OFF を切り替えるタイミングで、COCOA が広告パケットを送信ようになる。予備実験では、Bluetooth の ON/OFF を切り替えるタイミング (スキャン間隔) を5秒に設定した場合、実験対象の全てのスマートフォンが送信する BLE の広告パケットを収集することができたため、デジタルタコグラフ車載器は5秒ごとに Bluetooth の ON/OFF を切り替えるように設定した。

#### 3.2.2 アドレス・キャリアオーバーフェーズ

このフェーズでは、前フェーズで収集した路線バスの1運行分の BLE 広告パケットを入力とし、これらの広告パケットから、MAC アドレス、RSSI、ペイロード長、観測開始時刻と観測終了時刻を決定するために用いるタイムスタンプを抽出する。次に、抽出した MAC アドレス、RSSI、ペイロード長、タイムスタンプを基にアドレス・キャリアオーバーを行う。最終的な出力として、表 2 に示すように、利用者 ID とその利用者 ID に対応する観測開始時刻、

表 2: 利用者情報データの例  
Table 2 Example of bus-user information data

利用者 ID	観測開始時刻	観測終了時刻
0001	12:00:00	12:15:00
0002	12:00:00	12:15:00
0003	12:10:00	12:20:00
⋮	⋮	⋮

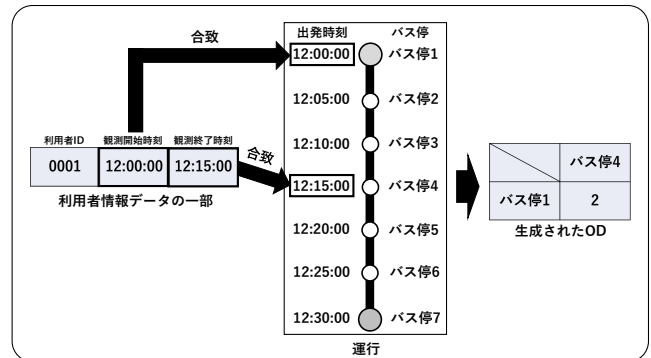


図 2: OD データ生成の例  
Fig. 2 Example of od data generation

観測終了時刻が1セットとなっている利用者情報データが得られる。

#### 3.2.3 OD データ生成

このフェーズでは、前フェーズで取得した利用者情報データを入力とし、ある1つの運行で停車するバス停の発車時刻と観測開始時刻と観測終了時刻が合致する利用者情報データをカウントすることで、OD データを生成する。具体例として、表 2 の利用者情報データを入力した場合の OD データの生成を考える。OD データ生成の例を図 2 に示す。左のデータは、利用者情報データから観測開始時刻が 12:00:00、観測終了時刻が 12:15:00 のデータを抜粋したものである。また、右の図はあるバスの運行を示しており、この運行はバス停 1 を起点、バス停 7 を終点としている。図 2 に示すように、観測開始時刻の 12:00:00 はバス停 1 の出発時刻と合致しており、観測終了時刻の 12:15:00 はバス停 4 の出発時刻と合致しているため、バス停 1-バス停 4 の利用者としてカウントする。最終的な出力として、表 1 に示すような OD データが得られる。

## 4. 評価実験

本章では、京都府京都市の路線バスの実データを用いて OD データの取得精度を評価する。

### 4.1 評価で用いるデータセットを収集する路線

本評価で用いるデータセットを収集する路線として、図 3 に示すような京都府京都市内のバス路線を選定した。本路線は、始点が苔寺・すず虫寺、終点が京都駅前の約 9.3km の路線である。時刻表では、バスが始点から出発して終点



図 3: 評価対象路線の経路図 (地図データ: Google マップ)

Fig. 3 Route map of the route to be evaluated. (Map data: Google map)

へ到着するまで 55 分間かかることが想定されている。また、バス停の数が 33 であることから、最大で計 32 区間の走行時間と 31 回の各バス停での停車時間が存在する。

本路線は、下記の 2 つの要件を満たす点で、提案手法の評価に用いるデータセットを収集する対象として適している。

十分な運行時間を有すること

提案手法の妥当性を検証するためには、MAC アドレスの変更が行われる程度に利用者の乗車時間が長い必要がある。十分な運行時間を有する路線は、乗車時間が長い利用者が多くなると考えるため、提案手法の妥当性の検証での利用に適する。

停留するバス停数が多いこと

停留するバス停数が多い場合、各利用者が異なるバス停間を乗車すると考えられる。つまり、利用者の OD パターンの組み合わせが多数存在すると考えられるため、評価対象としている。

4.1.1 データセットの収集

本節では、4.1 節で記述した路線で収集したデータセットについて説明する。データセットとなる BLE の広告パケットは、路線バスが搭載するデジタルタコグラフ車載器に Bluetooth 通信モジュールを取り付け、hcidump を用いることで収集した。また、図 ?? で示すように、デジタルタコグラフ車載器の後方は透明のボードで仕切られていることや、Bluetooth 通信モジュールの障害により、データセットを収集できない場合を考慮に備えて、Raspberry Pi 3 Model B+ も使用し、同日の同時時間帯でデータセットの収集を行った。このようにすることで、冗長性を確保している。

各データセット収集デバイスの配置を図 4 に示す。図 4 に示すように、デジタルタコグラフ車載器は運転席のほぼ真上に設置しており、Raspberry Pi 3 Model B+ は最後部の左座席に設置している。

4.1.2 比較対象となる正解 OD データ

今回の評価では、目視によってバス利用者がどのバス停

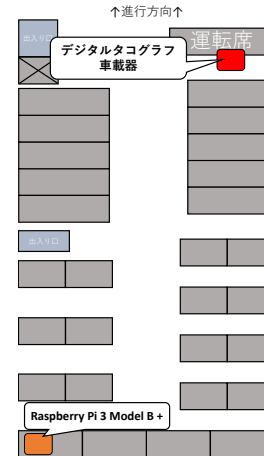


図 4: データセット収集デバイスの配置図

Fig. 4 Layout of data set collection devices.

降車バス停 \ 乗車バス停	庚申前	西大路四條	四條大宮	四條烏丸	烏丸松原	烏丸五條	烏丸七條	京都駅前
苔寺・すず虫寺	0	0	0	0	0	0	0	3
嵐山公園	0	0	0	2	0	0	0	0
嵐山 [京都]	0	0	0	0	1	0	0	2
有栖川 [京都]	0	0	0	0	0	0	1	0
生田口	1	2	0	0	0	0	0	1
帷子ノ辻	1	0	0	0	0	0	0	1
太秦広隆寺前	0	0	0	0	0	1	0	0
庚申前	0	0	0	1	0	0	0	0
西大路四條	0	0	1	0	0	0	0	0
四條烏丸	0	0	0	0	0	0	0	3
烏丸松原	0	0	0	0	0	0	0	1
烏丸五條	0	0	0	0	0	0	0	1

図 5: 評価対象路線における正解 OD データ

Fig. 5 Correct OD data on the routes to be evaluated

で乗車しどのバス停で降車したかを記録した。この記録を基に作成した全ての正解 OD データを図 5 に示す (利用者が存在したバス停区間を橙色に色付けしており、利用者の乗降がなかったバス停は省略している)。今回、利用者が存在したバス停区間は 16 区間であり、利用者の総数は 23 名となっている。正解 OD データの傾向として、乗車するバス停にほぼ偏りはみられないが、京都駅前のバス停で降車する利用者が明らかに多いことがわかる。これは、京都駅から電車やタクシーといった他の公共交通機関に乗り換えを行う利用者が多数存在するためであると考えられる。なお、この OD データは一般の利用者の OD を記録したものであるため、利用者が持つスマートフォンの Bluetooth が ON になっているか、利用者が COCOA を使用しているかは不明である。

4.2 評価方法

本評価では、提案システムが生成する OD データの精度を検証する。そのために、事前に収集したデータセットを入力とし、3.2 節で述べた提案システムを用いて OD データ

	降車バス停	西大路四 条	壬生寺 道	京都 駅前
乗車バス停				
苔寺・すず虫寺	0	0	2	
生田口	0	0	1	
帷子ノ辻	0	0	1	
西大路三条	1	0	0	
四条中新道	0	1	0	
四条大宮	0	0	1	

図 6: デジタルタコグラフ車載器を用いた場合の OD データ  
Fig. 6 OD data when using an onboard digital tachograph

を取得する実験を行う。上述する実験で取得した OD データと正解 OD データを比較することで、OD データの精度を示す。精度の比較では、以下に示す適合率、再現率を用いる。適合率は、COCOA の利用者のみを対象としたときの精度 (式 (1)) であり、単位は%である。また、再現率は、目視で確認した利用者を対象とした時の精度 (式 (2)) であり、単位は%である。ここで  $i$  は乗車バス停、 $j$  は降車バス停、 $N$  はバス停の集合、 $x_{i,j}$  正解 OD データにおけるバス停間の利用者数、 $y_{i,j}$  は提案システムが推定した OD データにおけるバス停間の利用者数、 $c_{i,j}$  はバス停間で利用者が存在するか判別する閾値を示す。

$$\text{適合率} = \frac{\sum_{i,j \in N} c_{i,j} \cdot y_{i,j}}{\sum_{i,j \in N} y_{i,j}} \times 100 \quad (1)$$

$$\text{再現率} = \frac{\sum_{i,j \in N} c_{i,j} \cdot y_{i,j}}{\sum_{i,j \in N} x_{i,j}} \times 100 \quad (2)$$

$$\begin{cases} c_{i,j} = 1 & (x_{i,j} > 0) \\ c_{i,j} = 0 & (\text{otherwise}) \end{cases}$$

### 4.3 評価結果

#### 4.3.1 デジタルタコグラフ車載器を用いた場合の OD データ

データセットとして、デジタルタコグラフ車載器により収集した BLE 広告パケットを使用した場合の OD データを図 6 に示す (利用者が存在したバス停区間を橙色に色付けしており、利用者の乗降がなかったバス停は省略している)。本提案手法により、利用者が存在したと推定されたバス停区間は 6 区間であり、利用者の総数は 7 名となっている。その中で、正解 OD データと比較した際、正しく推定されていた区間は苔寺・すず虫寺-京都駅前、生田口-京都駅前、帷子ノ辻-京都駅前である。また、正しく推定されていた各区間の人数は 2 名、1 名、1 名であり、提案手法で取得された OD データの正解数は 4 名となっている。

#### 4.3.2 Raspberry Pi 3 Model B+を用いた場合の OD データ

データセットとして、Raspberry Pi 3 Model B+ によ

	降車バス停	西大路四 条	壬生寺 道	烏丸 松原	京 都 駅 前
乗車バス停					
苔寺・すず虫寺	0	0	0	2	
有栖川 [京都]	0	0	0	1	
帷子ノ辻	0	0	1	1	
西大路三条	2	0	0	0	
四条中新道	0	2	0	0	

図 7: Raspberry Pi 3 Model B+を用いた場合の OD データ  
Fig. 7 OD data when using an Raspberry Pi 3 Model B+

表 3: 評価指標に基づいた OD データ取得精度  
Table 3 Accuracy of OD data acquisition based on evaluation indices.

	適合率	再現率
デジタルタコグラフ車載器	57%	17%
Raspberry Pi 3 Model B+	33%	13%

り収集した Bluetooth 広告パケットを使用した場合の OD データを図 7 に示す (利用者が存在したバス停区間を橙色に色付けしており、利用者の乗降がなかったバス停は省略している)。本提案手法により、利用者が存在したと推定されたバス停区間は 6 区間であり、利用者の総数は 9 名となっている。その中で、正解 OD データと比較した際、正しく推定されていた区間は苔寺・すず虫寺-京都駅前、帷子ノ辻-京都駅前である。また、正しく推定されていた各区間の人数は 2 名、1 名であり、提案手法で取得された OD データの正解数は 3 名となっている。

#### 4.3.3 評価指標に基づいた OD データ取得精度

各収集デバイスにおける評価指標に基づいた OD データ取得精度を表 3 に示す。この表は、行が収集デバイス、列が評価指標に基づいた OD データ取得精度を示している。デジタルタコグラフ車載器を収集デバイスとして用いた場合、適合率が 57%、再現率が 17%となった。また、Raspberry Pi 3 Model B+を収集デバイスとして用いた場合、適合率が 33%、再現率が 13%となった。

各収集デバイスを比較した際、適合率、再現率ともに、デジタルタコグラフ車載器を収集デバイスとして用いた場合の方が高い。

## 5. 考察

4 章の評価実験では、京都市の路線バスにおける 2021 年 12 月 23 日の 16 時 25 分苔寺・すず虫寺発、17 時 20 分京都駅前着の路線を対象に精度の比較評価を行った。本章では、評価実験で得られた結果から考察を述べ、その後、今後の展望について述べる。

### 5.1 提案手法の有効性について

提案システムの適合率と再現率を用いて、提案手法の有効性を考察する。

適合率に関して、デジタルタコグラフ車載器の適合率が57%、Raspberry Pi 3 Model B+の適合率が33%となっている。4.3節で述べたように、正しく推定された区間における最短区間は帷子ノ辻-京都駅前であり、この区間の所要時間は35分間である。COCOAはランダムMACアドレスを使用し、MACアドレスの変更間隔は最長で約20分と推測されるため、少なくとも1度はアドレスが変更されていると考えられる。したがって、既存手法では、ランダムMACアドレスを用いるスマートフォンを追跡できなかったのに対し、本提案は追跡できることを示した。

しかし、現時点における適合率は実運用に耐える精度ではないと考えられる。適合率の改善点として、路線バスの運行状況を考慮していない点が挙げられる。本提案では、データセットを収集する路線における全てのバス停を対象としているため、通過したバス停が含まれている。そのため、路線バス周辺に存在する非利用者のMACアドレスが通過バス停の通過時刻付近で観測された場合、この非利用者を利用者として誤判別する可能性がある。

実例として、西大路三条と四条中新道は通過バス停である(図5参照)にもかかわらず、デジタルタコグラフ車載器を用いた場合のODデータでは、西大路三条および四条中新道の通過時刻付近で非利用者のMACアドレスが観測されたため、西大路三条-西大路四条、四条中新道-壬生寺道の2つの区間で利用者が存在したと誤って推定されている。

この問題は、路線バスのGPS情報や加速度等から路線バスの停車バス停を推定し、推定したバス停で観測されたMACアドレスのみを利用者と判別することで、解決できると考えられる。したがって、上記の手法によって適合率を改善できることから、本提案は有望である。

再現率に関して、デジタルタコグラフ車載器の再現率が17%、Raspberry Pi 3 Model B+の再現率が13%となっている。再現率低下の原因として、バスの利用者に対するCOCOAの利用者が少ないことが考えられる。総務省の資料報告(接触確認アプリ(COCOA)の活用促進について)[12]によると、2021年12月9日時点でのダウンロード数が2137万件となっている。つまり、日本の総人口を1億2000万人と仮定した場合にCOCOAの利用者は総人口の17%となる。日本の総人口に対するCOCOAの利用者数と再現率が概ね合致することから、バスの利用者に対するCOCOAの利用者が少ないことが、再現率低下の原因であると言える。

## 5.2 適合率低下の原因について

適合率低下の原因として、路線バス周辺に存在する非利用者のMACアドレスを完全に除去できなかったことが考えられる。

デジタルタコグラフ車載器を用いた場合、誤って推定さ

れたバス停区間は西大路三条-西大路四条、四条中新道-壬生寺道、四条大宮-京都駅前の3区間であり、各所要時間は2分、1分、14分となっている。これらの区間において、四条大宮-京都駅前は所要時間が長いこと、アドレス・キャリアオーバーに失敗していると考えられる。しかし、西大路三条-西大路四条、四条中新道-壬生寺道に関しては、西大路四条から京都駅前にかけて道路が渋滞していたため、各区間において路線バス前方に存在する自動車の搭乗者のMACアドレスを観測したと考えられる。つまり、自動車の搭乗者である非利用者のMACアドレスを除去できなかったことが、適合率低下の要因となっている。

また、Raspberry Pi 3 Model B+を用いた場合、誤って推定されたバス停区間は、西大路三条-西大路四条、四条中新道-壬生寺道、帷子ノ辻-烏丸松原の3区間であり、各所要時間は2分、1分、28分となっている。これらの区間において、帷子ノ辻-烏丸松原は所要時間が長いこと、アドレス・キャリアオーバーに失敗していると考えられる。しかし、西大路三条-西大路四条、四条中新道-壬生寺道に関しては、交通渋滞により、各区間において路線バス後方に存在する自動車の搭乗者のMACアドレスを観測したと考えられる。つまり、自動車の搭乗者である非利用者のMACアドレスを除去できなかったことが、適合率低下の要因となっている。

## 5.3 収集位置による精度の違いについて

4.3.3節の評価結果から、適合率と再現率の両方において、デジタルタコグラフ車載器を用いた場合のODデータより、Raspberry Pi 3 Model B+を用いた場合のODデータの方が性能が低いことがわかる。原因として、Raspberry Pi 3 Model B+を用いてBLEの広告パケットを収集する際、Raspberry Pi 3 Model B+を運転席から見て最後部の右側の座席に設置していたため、路線バス車外に存在する非利用者のスマートフォンが送信する広告パケットが、収集した広告パケットに多数含まれていたと考えられる。

## 5.4 今後の展望

提案手法の評価により、既存手法では、ランダムMACアドレスを用いるスマートフォンを追跡できなかったのに対し、提案手法は追跡できることを示した。しかし、提案手法を組み込んだODデータ自動取得システムのODデータ取得精度は低い結果となった。

ODデータの取得精度が低い要因は、5.1節で述べたように、通過バス停の通過時刻付近で観測された非利用者を利用者として誤判別することにある。したがって、路線バスが停車したバス停を推定する必要がある。そこで、今後の課題として、路線バスのGPS情報や加速度等から推定した停車バス停を推定し、推定したバス停で観測された人物を利用者と判別することで、適合率の改善を図る。

## 6. おわりに

路線バスの運行効率を高めるには、バス利用者の利用動向を示す OD データの活用が必要不可欠であり、人手に頼らず取得する手法が望まれる。そのため、本研究では OD データの自動取得を目的とし、バスの利用者数、金銭コスト、決済方法の観点から、スマートフォンが発するパケットを用いた OD データの自動取得手法について着目した。しかし、プライバシー保護の観点から、ランダム MAC アドレスが使用されるようになって以降、ランダム MAC アドレスを利用するスマートフォンに対する OD データの自動取得手法は確立されていない。

本研究では、アドレス・キャリアオーバー・アルゴリズムをベースとした、複数のスマートフォンによる MAC アドレスランダム化の非同期性を用いた OD データ自動取得手法を提案した。この手法は、あるスマートフォンの MAC アドレスの変更後、一定時間内に観測された新しい MAC アドレスを変更後の MAC アドレスの候補とすることで OD データを自動取得する。評価では、京都府京都市の路線を対象に評価実験を行い、正解 OD データと比較した結果、OD データの取得精度は 57.1% となった。この結果は、既存手法がランダム MAC アドレスを用いるスマートフォンを追跡できなかったのに対して本提案は追跡できる一方、本提案の OD データ取得精度が低いことを示した。今後の課題として、路線バスの GPS 情報や加速度から推定した停車バス停を用いてバス利用者と非利用者を判別する手法を取り入れることで、OD データの取得性能が向上する可能性がある。

**謝辞** データ収集にご協力頂いたみなと観光バス株式会社、京都バス株式会社、及び京都市都市計画局歩くまち京都推進室の皆様へ感謝いたします。

## 参考文献

- [1] 国土交通省：地域交通をめぐる現状と課題，<https://www.mlit.go.jp/policy/shingikai/content/001311082.pdf> (2019).
- [2] 滝沢市 都市整備部 交通政策課：路線バス乗り込み OD 調査調査結果，[https://www.city.takizawa.iwate.jp/var/rev0/0102/7936/10\\_hosokusiryou2-0Dtyousa.pdf](https://www.city.takizawa.iwate.jp/var/rev0/0102/7936/10_hosokusiryou2-0Dtyousa.pdf) (2016).
- [3] 国土交通省中部運輸局：バスデータ活用大百科，<https://www.tb.mlit.go.jp/kanto/content/000166077.pdf> (2020).
- [4] Kostakos, V., Camacho, T. and Mantero, C.: Towards proximity-based passenger sensing on public transport buses, *Personal and Ubiquitous Computing*, Vol. 17, pp. 1807–1816 (online), DOI: 10.1007/s00779-013-0652-4 (2013).
- [5] Dunlap, M., Li, Z., Henrickson, K. and Wang, Y.: Estimation of Origin and Destination Information from Bluetooth and Wi-Fi Sensing for Transit, *Transportation Research Record*, Vol. 2595, pp. 11–17 (online), DOI: 10.3141/2595-02 (2016).
- [6] Paradedá, D. B., Junior, W. K. and Carlson, R. C.: Bus passenger counts using Wi-Fi signals: some cautionary findings, *TRANSPORTES*, Vol. 27, pp. 115–130 (online), DOI: 10.14295/transportes.v27i3.2039 (2019).
- [7] Pu, Z., Zhu, M., Cui, Z. and Wang, Y.: Mining Public Transit Ridership Flow and Origin-Destination Information from Wi-Fi and Bluetooth Sensing Data, *IEEE Internet of Things Journal*, Vol. 8, pp. 474–486 (online), DOI: 10.1109/JIOT.2020.3007373 (2021).
- [8] Barceló, J., Montero, L., Marqués, L. and Carmona, C.: Travel time forecasting and dynamic origin-destination estimation for freeways based on bluetooth traffic monitoring, *Transportation research record*, Vol. 2175, pp. 19–27 (online), DOI: 10.3141/2175-03 (2010).
- [9] Apple: Wi-Fi のプライバシー，<https://support.apple.com/ja-jp/guide/security/secb9cb3140c/web> (2022).
- [10] Android オープンソースプロジェクト：MAC アドレスランダム化の実装，<https://source.android.com/devices/tech/connect/wifi-mac-randomization?hl=ja> (2021).
- [11] Becker, J. K., Li, D. and Starobinski, D.: Tracking Anonymized Bluetooth Devices (2019).
- [12] 厚生労働省：接触確認アプリ (COCOA) の活用促進について，<https://www.mhlw.go.jp/content/10900000/000704457.pdf> (2020).