

交通データと消費データの重ね合わせによる 生活行動の試行的分析

神田 佑亮¹・馬本 迪奈²

¹正会員 呉工業高等専門学校教授 環境都市工学分野 (〒737-8506 呉市阿賀南 2-2-11)
E-mail: yusuke-k@kanda-labo.net

²学生非会員 呉工業高等専門学校 環境都市工学科 (〒737-8506 呉市阿賀南 2-2-11)

近年のキャッシュレス化の推進と定着により、日常生活における様々なシーンでのデジタルデバイスによる支払いが浸透してきている。こうしたキャッシュレスによる支払いは、IC カードやユーザーが保有するデバイスやアプリに準じた固有 ID とともに、いつ、どの施設で、いくら購入されたかといった情報(レコード)が記載され、補完される。こうした情報は地域の公共交通マネジメントにも十分に活用可能である。

本論文では、広島県庄原市で試験導入した AI オンデマンド交通の利用者を対象に、交通系 IC カード、地域キャッシュレスカード、デマンド交通利用ログデータベースを統合して分析し、公共交通利用以外も含めた移動の実態を明らかにした。

Key Words: data-fusion, Evidence-Based Policy Making, Transportation Policy

1. はじめに

近年のキャッシュレス化の推進と定着により、日常生活における様々なシーンでのデジタルデバイスによる支払いが浸透してきている。こうしたキャッシュレスによる支払いは、IC カードやユーザーが保有するデバイスやアプリに準じた固有 ID とともに、いつ、どの施設で、いくら購入したかといった情報が記録される。

こうした情報を地域の公共交通マネジメントに活用する動きは、以前から展開されている。バスや鉄道などの公共交通 IC カード利用ログデータの活用により、公共交通利用者の属性、利用 OD、利用時間帯などの情報をもとに、路線網や運行本数等の計画にも反映されている。

一方で、公共交通 IC カードによる利用ログだけでは、把握可能な移動パターンは公共交通のみに限定されるため、移動の目的を把握したり、自動車や自転車など、公共交通以外の手段を用いた際の移動特性について把握することは極めて困難である。

換言すれば、近年急速に普及が進む消費データと公共交通データを統合することができれば、その可能性は大きく広がる。データに基づく議論の範囲も、公共交通利用移動から、移動全体へ議論が広がるのは間違いない。

しかしながら、交通分野のデータと他のデータを統合して分析するために解決すべき課題は多い。まずクリア

すべき課題はデータプライバシーの問題である。「個人情報の保護」は非常にセンシティブな事項であり、個人を特定可能なデータを外部で取り扱わないといったデータハンドリングの面や、データ取得時での合意は当然のことながら、そのデータを取り扱うことによる社会的意義の合意も必要となってくる。

加えて、個人を特定可能な情報を削除して扱うことが一般となっている中、異なる複数のデータレコードから、どのように「一連のデータ」として活用するかといった面も課題となる。さらに、一般的にこうしたデータは容量が大きくハンドリングが容易ではないこと、それらのデータを扱う人材が豊富でないことも大きな課題である。

筆者は、中山間地域での地域モビリティ改善プロジェクトの一環で、「交通」と「消費」の2分野のデータから、地域交通マネジメント政策を議論する実践機会を得た。本論文では、これらのデータベースを統合して分析した際の課題と対応、データフュージョンによる新たな可視化領域の可能性を示す。

2. 分析に用いたデータの特性

本論文で取り扱うデータは、中国地方の中山間地域でデマンド交通の導入による可能性を検証することを目的に得たものである(2021 年度後半に実施)、デマンドバス

の運行は、市街地中心部と郊外の農村地域を接続する形で、主に平日に運行された。デマンドバスは事前予約制であり、予約にあたり、インターネットまたは電話で受け付けた。なお、利用にあたっては事前登録制とはしなかったが、円滑な予約処理のため、氏名、住所、連絡先等の事前の登録を推奨した。

表-1に分析で用いたデータを示す。入手可能なデータは、①公共交通決済ログデータ（以下、公共交通決済ログ）、②デマンド交通予約・利用ログデータ（以下、デマンド利用ログ）、③域内店舗等でのキャッシュレス決済ログデータ（以下、消費ログ）の3種類である。

データのカバー状況は、①公共交通決済ログは、地域での公共交通利用の8割程度の支払いで使用されている。②デマンド利用ログは、全ての利用履歴が把握可能である。③消費ログは、域内の人口に対し約9割の住民が保有しており、店舗カバー率も高く、域内の消費額の約5割がこのカードを通じて決済されている。

ただし、いずれのデータベースも独立したものであり、データベースの管理者・保有者も異なる。そのため、上記3種類のログデータ相互で共通のユーザーIDは付与されておらず、データに共通するキー項目は存在しない。加えて、個人情報保護の観点から、個人名や住所の詳細、電話番号、年齢、生年月日などの情報は除外され、また、利用者固有のIDも変換して提供を受けた。

交通と消費のデータ・フュージョン分析では、複数の

データに含まれる内容を最大限活用するには、これらの異なるデータベースを紐づける必要がある。しかしながら、上記の事情から、3つのデータベースを対に紐づけるデータ項目が存在しない。そのため、①公共交通決済ログと②デマンド予約ログでは、当該交通の乗車時または降車時の記録（時刻等）から両者のデータベースに記録された同一のイベントをマッチングした。また、③消費ログと交通系データベース（①、②）のマッチングは、デマンドバス乗車時にキャッシュレスカードへの乗車ポイント（10円相当）を付与することで、デマンドバス交通利用の記録が、③消費ログに記録され、3つのデータベースの連携を図った。

上記による紐付け分析の結果、②デマンド予約ログに記録のあったID34件のうち、①公共交通決済ログとマッチングできたIDが11件③消費ログとマッチングできたIDが21件、全てのデータベース間でマッチングできたIDが10件であった。なお、全ての利用者が公共交通ICや地域キャッシュレスICを保持しているわけではない。

3. 試行的分析

a) デマンド交通利用者の外出状況

分析事例の一例として、消費ログデータに記録のあった日を外出した日と見なし、うち、その日の公共交通利用の有無別に集計し、決済のあった店舗の種類から移動目的を設定し集計したものである（図-1）。公共交通利用をする人であっても、公共交通の利用記録のない日に買い物や余暇（温浴施設）等の店舗・施設での決済が記録されている日があり、そうした小移動パターンのある人も多い。こうしたデータに表現される背景には、家族や知人等の送迎や相乗りにより日常生活移動が支えられていることが窺える。実際に事業実施地域の住民を対象に別途実施したアンケートでは、回答者(n=158)の約3割が、週に2回以上家族や知人のための移動や送迎を担っていると回答し、そのうち、約4割が「負担に感じる」という回答があった。

表-2は、デマンドバス運行地域において、路線バスも併用されている地域での往復の利用状況を示したものである。この地域では2021年11月から2022年1月までの3ヶ月間、1日あたり3往復のデマンドバスと、1往復の路線バス便が運行された（早朝の中心部行きおよび終発の郊外行き）。デマンドバスの利用状況は予約ログから、路線バスの利用ログは交通系ICカード利用ログに記録されているが、これらの2つの異なるデータベースを、上記の要領で紐づけて分析を行った。

表-2から、デマンドバス1（昼前）に郊外部から中心部方面に乗車し、デマンドバス2（昼前）で中心部から

表-1 分析に利用したデータの特性

データ種別	①公共交通決済ログ	②デマンド交通予約・利用ログ	③域内店舗等でのキャッシュレス決済ログ
管理者	公共交通事業者の連合組織	実証実験事務局(地域内)	実証実験対象地域内の商工団体
属性データ	なし	・年齢 ・性別 ・住所(郵便番号)	・年齢 ・性別 ・住所(郵便番号)
交通データ	・利用日時(分・秒) ・乗車・降車バス停 ・支払額・方法	・利用日時(分・秒) ・乗車・降車バス停 ・予約日時	・乗車ポイント付与の際の日時・端末ID
消費データ	なし	なし	・来訪店舗・端末ID ・決済日時(分・秒) ・決済金額

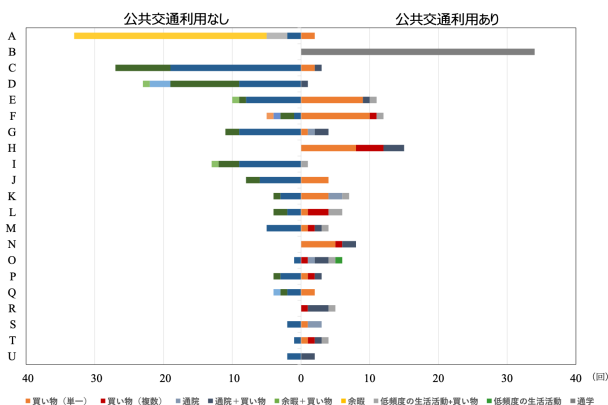


図-1 公共交通およびその他の交通手段での外出回数(個人別)

郊外に帰宅している人が最も多い。次いで、デマンドバス 2 (夕方) に乗車し、デマンドバス 3 (夕方) のデマンドバスで帰宅している人も多いことが読み取れる。3つのデータベースを統合して判断すると、前者は主に通院や買い物、後者は買物のために利用されているという傾向が読み取れる。そして、利用者の多くが、午前中に外出し戻る、または午後に出発し夕方までに戻るといふ、外出の滞在時間が短い傾向にあった。

また、路線バス (朝・夜) の往復や朝の路線バスのみで、帰りの路線バス・デマンドバス利用の記録のない利用も多い。これは、この地域に住む学生が通学時に利用し、朝はバスで出かけるものの、帰りは保護者の送迎等により帰宅していることが窺える。

b) 地域別の外出回数

地域キャッシュレス消費データに記録のあった 60 代以上のユーザーを対象に、地域別(郵便番号別) の決済日数ウを分析した。この地域では、地域の住民の 9 割が地域キャッシュレス IC を保有し、また、生活消費の過半をこのキャッシュレス IC を通じて支払っている。加えて、食料品店やホームセンター等は市街地中心部にあり、周辺部にはほとんど存在しない。即ち、決済が記録された日は、少なくとも消費を伴う外出を行った日とみなすことができる。

平均外出回数と拠点までの距離の回帰分析を図-2 に示す。拠点までの距離が近い地区ほど、平均外出回数が多く、相関係数の値が 0.72 であった。このことから、平均外出回数と拠点までの距離には強い負の相関があると言える。

各地区の年代別の外出日数の分析を行った。一例として、中心市街地に近い川東地区を図-3、中心市街地まで遠い地区として帝積 (帝積宇山・帝積山中・帝積始終・帝積未渡) 地区を図-4 に示す。

分析の結果、拠点までの距離が近い地区ほど年代に偏ることなく外出しており、拠点までの距離が遠い地区ほど 60・70 代に偏って外出していることが明らかとなった。これは、拠点までの距離が近い地区は自転車やバスなど様々な移動手段で外出できる環境が整っているのに対し、中心部の拠点までの距離が遠い地区は移動手段が限られるため、高齢になると距離やモビリティの環境が決済日数、即ち外出日数にも影響することが窺える。

4. まとめと今後の展望

管理者が異なる複数のデータベースを統合して扱う際の課題は多いものの、キーとなるイベントを設定することで、これまでに捉え切れなかった移動パターンを捉え

表-2 デマンド便・路線バス便の往復での利用状況

	郊外部発				
	路線バス (朝)	デマンドバス1 (昼前)	デマンドバス2 (夕方)	デマンドバス3 (夜)	なし
デマンドバス1 (朝)	---	---	---	---	---
デマンドバス2 (昼前)	9.5%	28.5%	---	3.6%	---
デマンドバス3 (夕方)	---	6.6%	12.4%	---	0.7%
路線バス (夜)	10.9%	---	0.7%	2.9%	0.7%
なし	19.0%	2.2%	0.7%	1.5%	---

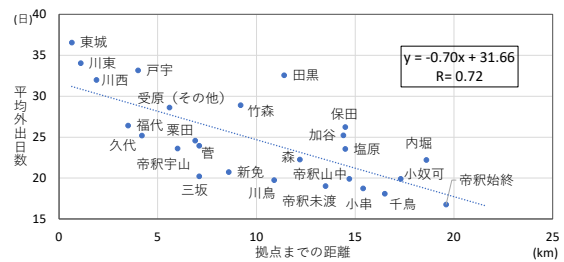


図-2 平均外出回数と拠点までの距離の相関関係

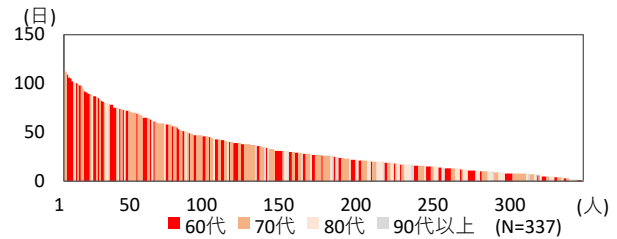


図-3 市街地に近接する地域での年代別外出回数(川東地区)

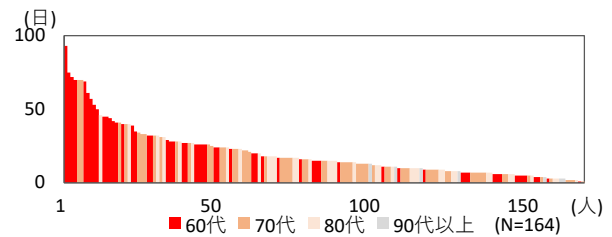


図-4 市街地に近接する地域での年代別外出回数(帝積地区)

ることができ、また、調査時期を問わない、時間横断的なデータの取得も可能になる。今回分析した事例では、データを保有し、提供の意思決定が可能な主体が地域内に存在し、かつ、それぞれ捕捉率の高いデータベースを構築していたため、データ・フュージョン分析が可能となったことや、過疎地域であり人口・経済規模が小さいため、比較的扱いやすいデータ量であった。この点からの示唆として、対象とする範囲が拡大し、データの量が大きくなると、イベント発生ベースでのマッチングの精度を上げなければならなくなる可能性がある。反面、地方レベルではデータ提供のための意思決定が大都市と比較してシンプルで利用しやすい環境が整っている。DX推進の面では一見不利に見える地方部であるが、こうしたポテンシャルを有していることも意識した上で、地域交通を含めた地域の活性化のマネジメントを図っていく必要がある。