

ICカードデータの時空間利用特性による 市電利用者のクラスタリング

森田 琢雅¹・溝上 章志²・中村 嘉明³

¹学生会員 熊本大学大学院自然科学研究科 (〒860-8555 熊本市黒髪2-39-1)

E-mail:166d8833@st.kumamoto-u.ac.jp

²正会員 熊本大学教授 大学院自然科学研究科 (〒860-8555 熊本市黒髪2-39-1)

E-mail:smizo@gpo.kumamoto-u.ac.jp

³正会員 構造計画研究所 (〒869-1235 熊本県菊池郡大津町室1315)

E-mail:nakamuray@kke.co.jp

地方都市におけるICカードデータを活用した利用実態の把握が近年注目を集めている。これまで様々な都市で利用者属性と行動を把握する分析が行われているが、その多くは時間的な利用特性を利用して利用者の分類を行ったものであり、空間的な利用特性を考慮していない。本研究では、まず停留所ごとに集計したICカードデータによって停留所の階層的クラスタリングを行い、公共交通沿線上のゾーン分けを行った。次に利用者ごとに各ゾーンの利用率を含む利用特性種々の変数を算出し、非階層的クラスタリングを行って利用者の分類とその特徴の整理を行った。これらの結果と日々の運行実績から得られる任意の期間・区間における遅延時間との比較を行うことによって、市電の遅延に影響を与える要因の考察を行った。その結果、休日前日の特定路線において一定以上の来訪者が利用すると遅延時間のばらつきが大きくなる性質を発見した。

Key Words : smart card data, public transport in rural city, cluster analysis, actual diagram

1. はじめに

地域公共交通網の再編は利用実態や運行実績に関する詳細なデータを使って検討されるべきである。しかし、網羅的で大量のデータを収集するために従来から行われてきた乗り込み調査などは実施のための費用や精度の面で十分な情報を得るのが難しい。そういった中、筆者らを含め多くの研究者がICカードデータを公共交通の利用実態の把握に活用する研究を行っている。西内・轟ら¹⁾はICカードデータから算出した利用者毎の時間的・空間的なトリップ依存度を用いてクラスター分析を行い、定期利用者と非定期利用者のトリップ特性の違いを示した。嶋本・北脇・宇野・中村ら²⁾は、英国・ロンドンのICカードデータを用いて、料金支払い形態ごとの利用回数の変動を分析し個人の利用回数の変動の4割以上を支払い形態の分類で説明できることを示した。また西内・カ石・兵頭・轟ら³⁾は公共交通利用者の利用間隔の日々のパターンに注目し、高知市都市圏における利用者の利用特性を整理している。

このように、交通系ビッグデータと呼ばれるICカード

データから利用者の特性を明らかにする研究は注目を集めているが、これまでの研究はICカードの券種や利用時間特性、同一停留所の利用率などによって利用者情報の抽出を行っているのが大半であり、都市の空間的な特徴とそこでの利用者の利用形態を考慮した分類を行っている研究はない。本研究ではその足掛かりとして、熊本市電の停留所を空間的に分類し、利用者毎に利用時間帯や分類したゾーンの利用率など、空間的な利用特性も用いて利用実態の詳細な分析を行う。さらにそれらの結果を運行実績と比較することで、利用と運行の両面からの分析を試みる。

2. 利用データ

(1) 分析に用いるデータ

本研究では熊本市交通局に蓄積された2016年2月1日から2週間の熊本市電のICカードデータを用いる。データ項目などを表-1に示す。また、今回はカード1枚のデータを利用者1人のデータとして扱う。また熊本市電で利

用可能な全国相互利用ICカードの「でんでんnimoca」の定期券としてのICカード利用者を「定期券利用者」、そうでない利用者を「一般利用者」とする。

一方、降水時の利用回数を算出するために用いる天候データは、気象庁HP⁹⁾にある熊本市の毎日の9時、15時、21時の気象観測データを、それぞれ朝、昼、夜の降水の有無のデータとして用いた。

分析対象期間の利用者1人あたりの週平均利用回数は一般利用者が1.43回/週、定期利用者が8.67回/週であった。それぞれの利用者の利用回数分布を図-1、図-2に示す。一般利用者は5回以下の利用者が約9割を占め、多くが期間中に数回の利用である。定期利用者はおおむね週8回から10回の利用が多く、通勤や通学で主に平日の往復に利用していると思われる。

(2) 熊本市電沿線について

熊本市電の路線図を図-3に示す。系統は熊本駅（田崎橋）から東区の健軍町を繋ぐA系統と上熊本駅から同じく健軍町を繋ぐB系統の2つがあり、2系統は市内中心部の辛島町電停より東側では同軌道上を走行する。

表-1 ICカードから得られるデータの内容

	内容
期間	2016年2月1日～2016年2月14日
有効乗降処理	168,136回
出現カード枚数	37,688枚
データ項目	カードID, 券種, 利用日時(分単位), 乗車電停, 降車電停, 乗車人数, 車両番号, 系統 等

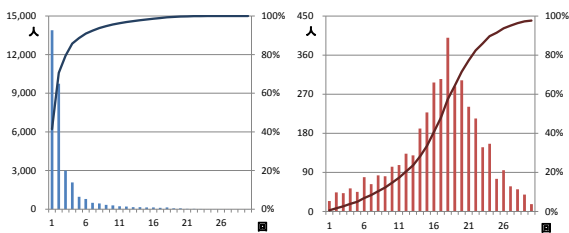


図-1 利用回数分布 (一般) 図-2 利用回数分布 (定期)



図-3 熊本市電路線図

阿蘇地域から流れてくる白川以西の上通や下通を中心とした通町桜町地区が中心市街地であり、通町筋や熊本城・市役所前、花畑町などの電停が位置する。熊本市電は観光客には西部の熊本駅から東部の水前寺公園や動植物園などへのアクセスに、市民には沿線の中学校や高校への通学、中心市街地への通勤や買い物、鉄道駅へのアクセスなどに利用されている。

3. 停留所のゾーン定義

(1) 停留所の定義と説明変数

次章では利用者ごとに各ゾーン（いくつかの連続した電停の集合）の利用率を算出するために、路線沿線上でゾーンの境界を決める。これは電停ごとの位置や利用状況によって直観的に行うこともできるが、分析者の主観で行うことは好ましくない。今回は、ICカードから得られる実際のデータから停留所ごとの利用実態を表す変数を算出し、クラスタリングの結果から沿線上のゾーン境界を見つけることを試みた。

停留所の実態を説明する要因として、今回は表-2に示す変数を算出し、階層的クラスタ分析によって電停を分類した。この方法の特徴は停留所の空間的な位置に関する変数は一切与えていないところにある。時間帯別のOD比を用いることで、トリップが集中する電停と発生する電停を分類し、鉄道駅や学校との隣接ダミーといった周辺環境に関する変数を導入した。そのため他都市でも同様に実施可能な方法になっている。

(2) 停留所の分類結果とゾーン分け

分析結果であるデンドログラムを図-4に示す。今回は同クラスターの電停が空間上に最もまとまって集合するクラスター数が6になる場合で電停を分けた。クラスターごとの特徴を示したのが表-3である。クラスター1は平日休日共に朝の出発地としての利用、夜の目的地として

表-2 停留所クラスタリングの変数

変数	説明
系統	A系統に属す(0,1) B系統に属す(0,1)
鉄道乗換	150m以内に鉄道駅(0,1)
学校	150m以内に学校(0,1)
発着数	平日 1日あたりの発着数 休日 1日あたりの発着数
OD比	平日 朝 5:00-12:00 平日 OD比 平日 昼 12:00-18:00 平日 OD比 平日 夜 18:00-25:00 平日 OD比 休日 朝 5:00-12:00 休日 OD比 休日 昼 12:00-18:00 休日 OD比 休日 夜 18:00-25:00 休日 OD比

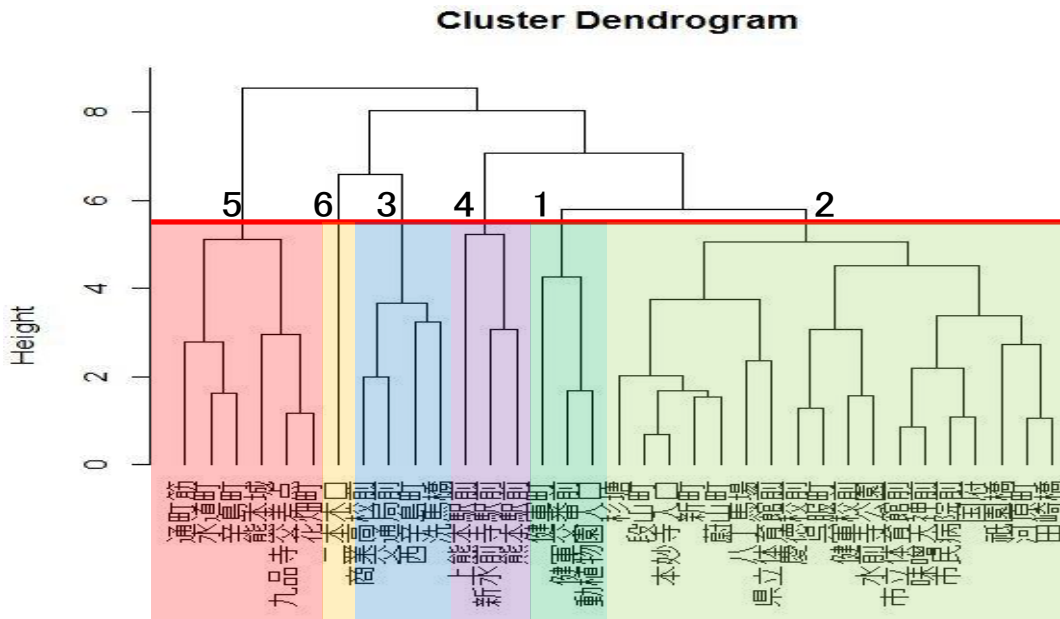


図4 デンドログラム

表-3 停留所クラスタリングの結果

クラスター	個数	鉄道	学校	平日発着	休日発着	平日 発/着 朝→昼→夜	休日 発/着 朝→昼→夜
1	3	0	0	24,786	6,063	—	—
2	18	0	4	6,701	1,794	—	—
3	4	0	3	7,880	1,685	—	—
4	3	3	0	42,318	13,246	—	—
5	6	0	0	33,484	11,121	—	—
6	1	0	0	5,883	996	—	—

の利用が多い電停である。クラスター2は時間帯によるOD比の変化がクラスター1と同じであるが利用者数がクラスター1に比べると1/4程度と少ない。学校隣接の電停も含まれている。クラスター3は朝の目的地としての利用、昼の出発地としての利用が多く、学校隣接の電停が多いことから、主に学校への通学に利用されている電停であると考えられる。クラスター4に分類された3つの電停はすべて鉄道駅に隣接する電停であり、利用者数も最も多い。クラスター5は朝の目的地としての利用、夜の出発地としての利用が多い電停である。利用者数もクラスター4に次いで多い。クラスター6は平日と休日で時間帯別OD比の変化が異なる電停一つのみが分類されている。

電停のクラスタリングの結果を地図上に記述したものが図-5である。電停の上の数字がクラスターの番号を示す。クラスター分析の際には電停の空間的な位置変数は一切用いていないにもかかわらず、クラスター1とクラスター2とクラスター5ではそれぞれ空間的に連続した電停で構成されている。この結果を用いて今回は、学校隣接のため通学利用が多いクラスター3や鉄道乗換利用が多いクラスター4を除いて、クラスター番号が変わるところを境界として、熊本市電沿線を健軍北部、東部、中央部、南西部、北西部の5つのゾーンに分割した。



図-5 電停クラスタリングの結果

4. 利用者の分類

(1) 説明変数

利用者を分類する説明変数については表-4に示す。利用日特徴として平日、休日の利用回数、降水時の利用回数、時間帯特徴として朝ピークと夕方ピークの利用回数、そして3章で定義したゾーン別の利用回数をそれぞれ算出し変数としている。

(2) k-means法とGap統計量

利用者の分類方法にはk-means法の初期値の選択に改良を行った非階層的手法のk-means++法を採用した。k-means法はあらかじめクラスター数kを決め、評価関数Lを最小にする分割を探す。評価関数Lはデータと対応する代表ベクトルの距離の二乗和の最小値であり、次式で表される。

$$L(c_1, \dots, c_k) = \sum_{i=1}^n \min_{l=1, \dots, k} \|x_i - c_l\|^2 \quad (1)$$

一方、クラスタ数はGsp統計量を用いて決定する。対象データをクラスタ k で分類した際の評価値を L 、同範囲に一樣分布するデータをクラスタ k で分類した評価値を L' とすると、Gap統計量は次式で表される。

$$Gap_k = \log \frac{L'_k}{L_k} = \log L'_k - \log L_k \quad (2)$$

図-6はクラスタ数 k が変化した時のGap統計量の変化を示している。今回はこの結果から、クラスタ数が増えても値が変化しなくなる $k=9$ を採択した。

(3) クラスタ分析の結果

クラスタリングの結果を表-5に示す。なおクラスターNo.は人数の多い順に付与している。クラスター1から3までは利用日や時間帯、降水時の利用に特徴があり、クラスター4から9までは利用場所に特徴がある。以下にそれぞれのクラスタについて特徴を述べる。

クラスター1：平均利用回数が1.7回で、主に期間中、数日だけの利用者である。乗換電停と南西部の利用率が高く、低い朝ピーク利用率と高い夕方ピーク利用率から、熊本駅を利用した来訪者である。

クラスター2：休日と降水時の利用が多いこのクラスターは利用回数も3.9回と低い。主に雨の時に市電を利用する人や休日の利用が多い利用者である。

クラスター3：利用回数は中頻度の8.5回で、休日の利用率が低いのが特徴である。降水時の利用率は最も少なく、利用電停に特徴は見られない。熊本に在住する中頻度利用者であると考えられる。

クラスター4：乗換電停の利用が多いこと、南西部と中央部の利用が多いことから、熊本市外から通勤、もしくは市外へ通勤している利用者であると考えられる。

クラスター5：健軍北部と中央部での利用が多い高頻度利用者であり、朝ピーク利用率が最も高く乗換電停と学校電停の利用はほとんどない。熊本市東部に居住する中央部への通勤利用者と考えられる。

クラスター6：唯一、北西部での利用率が高いグループであり、乗換電停として上熊本駅前を利用していると考えられる。

クラスター7：大半が東部と中央部の利用であり、乗換電停として新水前寺駅前の利用が多い利用者と考えられる。

表-4 利用者クラスタリングの変数

変数	説明
定期利用者	定期利用者であるか(0,1)
平日利用回数	平日乗車回数
休日利用回数	休日乗車回数
朝ピーク利用数	平日 7:00~9:00の乗車回数
夕方ピーク利用数	平日 16:00~20:00の乗車回数
降水時利用回数	降水時の利用回数
乗換電停選択数	JR 駅隣接の電停利用回数
学校電停選択数	学校隣接の電停利用回数
ゾーン別利用数	健軍北部の電停利用回数
	東部の電停利用回数
	中央部の電停利用回数
	南西部の電停利用回数
	健軍北区の電停利用回数

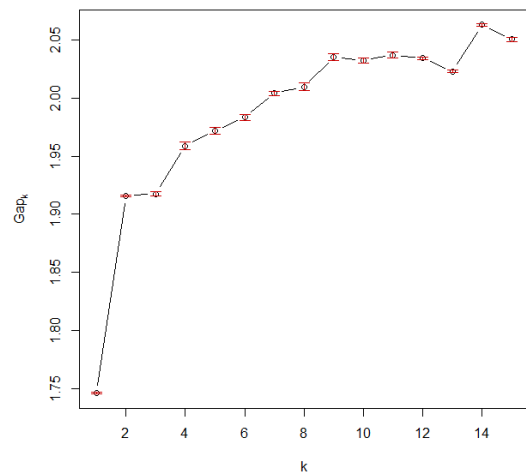


図-6 Gap統計量

表-5 利用者クラスタリングの結果

クラスター	人数	定期券利用者	利用回数(回/2週)	利用率											
				休日	朝ピーク	夕方ピーク	降水時	乗換電停	学校電停	健軍北部	東部	中央部	南西部	北西部	
1	26,116	91	1.7	28%	8%	32%	10%	55%	16%	15%	33%	81%	47%	12%	
2	4,646	43	3.9	64%	8%	27%	44%	46%	17%	19%	39%	79%	34%	16%	
3	2,422	545	8.5	10%	30%	34%	9%	40%	25%	26%	49%	66%	30%	14%	
4	1,004	785	18.1	20%	33%	31%	13%	70%	14%	11%	9%	73%	93%	1%	
5	884	637	18.6	19%	36%	34%	13%	3%	6%	94%	8%	89%	5%	2%	
6	787	593	18.1	20%	30%	32%	14%	41%	17%	13%	25%	52%	3%	88%	
7	746	584	18.4	22%	33%	31%	14%	56%	4%	1%	95%	88%	3%	2%	
8	720	611	20.1	19%	33%	36%	13%	29%	86%	13%	88%	49%	17%	4%	
9	363	300	18.9	18%	29%	32%	14%	59%	6%	89%	87%	14%	5%	1%	

クラスター8：学校電停利用率が高いグループである。利用回数は20.1回，定期券利用者の割合は84.9%と最も高く，通学利用の学生と考えられる。
 クラスター9：健軍北部と東部の利用が多く，乗換電停としては新水前寺駅前の利用が多いグループである。

5. 運行実績ダイヤの利用

(1) 運行実績ダイヤとは

ICカードデータから得られる電停ごとの乗車時刻と降車時刻のうちの最早時刻を到着時刻，最遅時刻を出発時刻と仮定することによって，時空間グラフ上に個々の車両の移動軌跡を描くことができる。筆者らはこの車両軌跡一台毎に対応する時刻表のデータを紐付けしたものを運行実績ダイヤと呼んでいる。車両軌跡についてはGPSデータなどがあれば詳細を明らかにできるが，1)乗客の乗降データが同時に取得できる点，2)高価なロケーションシステムを導入しなくてもよい点で運行実績ダイヤを分析することは有効である。なお，ICカード利用者の乗車・降車がない電停では通過時刻が不明となる。このような場合，図-7を例に，以下のようにして通過時刻を推定した。もし，電停Bにおける通過時刻が不明の場合，AC間実測所要時間（10分）にAB間時刻表所要時間/AC間時刻表所要時間（2/5）を乗じたもの（4分）をAB間所要時間とし，それをA出発時刻に加算したもの（9:04）をBの到着時刻および出発時刻とする。なお今回の実績

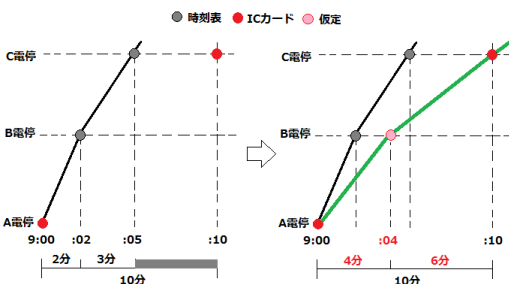


図-7 電停の通過時刻の算出方法

ダイヤの可視化についてはフリーのダイヤ編成支援システム「その筋屋」を使用している。

以降の運行実績ダイヤの筋の配色を図-8，図-9に記す。図-10，図-11は2/1(月)A系統の利用者数と遅延時間を可視化したものである。対象期間の日ごとにこれらを作成し，毎日の利用実態を視覚的に分かり易く表現することができた。図-10からは，7:00頃～8:30頃までの健軍北部～中央部と熊本駅～中央部の朝ピーク，12:00頃と13:00頃に起こった一時的な利用増加，17:00頃～21:00前後までの夕方ピーク，18:00頃に熊本駅から中央部に向かう利用者の増加などが見てとれる。図-11からは，8:30頃と11:30頃の水前寺公園から健軍町で一時的な大幅な遅れが起こったことや，夕方18:30頃～20:00頃に10分前後の遅れが多くなっている事などが視覚的に分かる。両図を比較して，乗客が多い時間に遅延が大きくなるのではなく，特に夕方ピークにおいては多くの乗客が降車した後の運行で遅延が大きくなったままであることが分かる。

(2) クラスタリング結果を加味した運行考察

分類したクラスターごとに運行実績ダイヤを作成し，いつ，どこで，どういった利用者が市電を利用し，いつ，どこで運行上の問題（遅延）が発生しているのか，また両者にどのような関係があるのかを明らかにすることを試みる。図-12，図-13，図-14，図-15はそれぞれ2/1(月)A系統におけるクラスター1（来訪者），クラスター4（熊本駅通勤者），クラスター5（東区通勤者），クラスター8（通学者）の実績ダイヤを示したものである。これらの図から，クラスターごとに利用の時間帯と区間が異なっていることが分かる。この方法を使って運行実績ダ

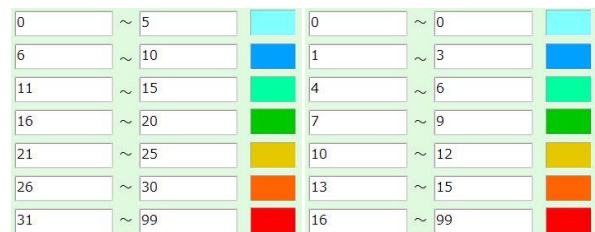


図-8 乗車人数の凡例(人) 図-9 遅れ時間の凡例(分)

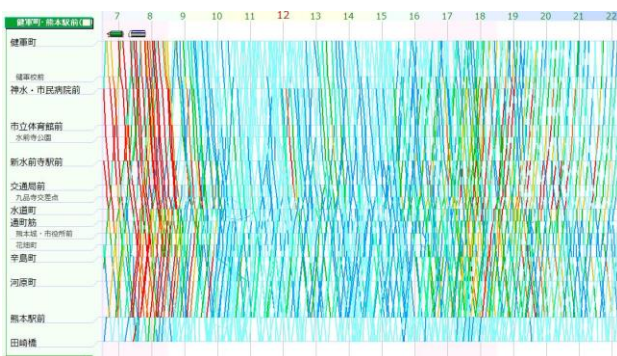


図-10 2/1全乗客数

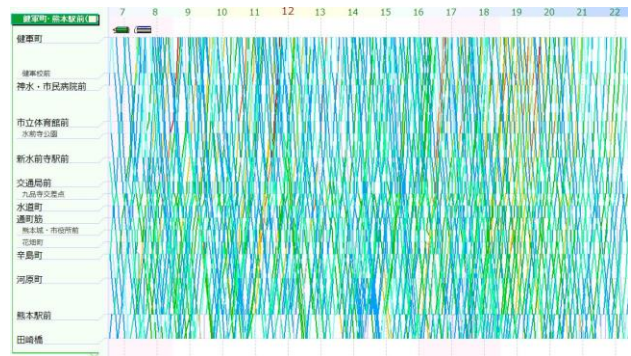


図-11 2/1遅延時間

イヤを様々な形で集計し、その結果から得た発見の一部を紹介する。

表-6は14日間の18:00~25:00におけるA系統健軍方面行の各運行終着電停での遅れ時間の平均と各クラスターの利用者数を示している。このデータの見方から、どのクラスターの利用者が増えた場合に遅延時間やそのばらつきが大きくなるのかを検討した。データの期間が狭いことから断定することはできないが、主に休日前日の夜にクラスター1の利用者が大きく増えて約1,000人を超えたときに、通常日と比較して遅延時間の標準偏差が約1分52秒大きくなる傾向が見られた。週末の夜に熊本駅から中央部に移動する来訪者による遅延時間のばらつきが起き

た後、中央部から東部や健軍北部へ帰宅する定期利用者によってさらなる遅延時間のばらつきの増加が起きていると考えられる。

また、降雨時の利用が多いクラスター2の利用が多いときには遅延時間の平均値、標準偏差ともに大きくなるのが予想されたが、そのような傾向は見られなかった。

今回の手法を繰り返し行うことで、どういった利用者がどれほど増減することでどれくらい運行に影響を与えているのかといったことを発見することができる。その都市の公共交通網再編の際に活かせる発見は、このICカードデータの分析手法から得られると考えている。

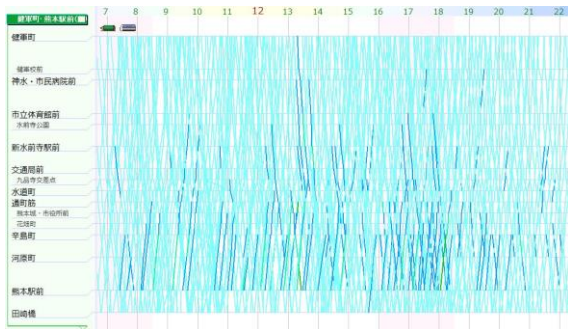


図-12 2/1 クラスター1 (来訪者)

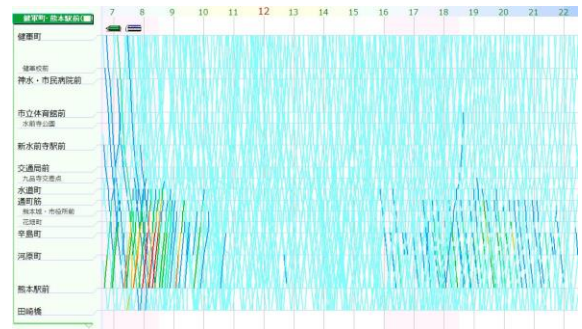


図-13 2/1 クラスター4 (熊本駅通勤者)

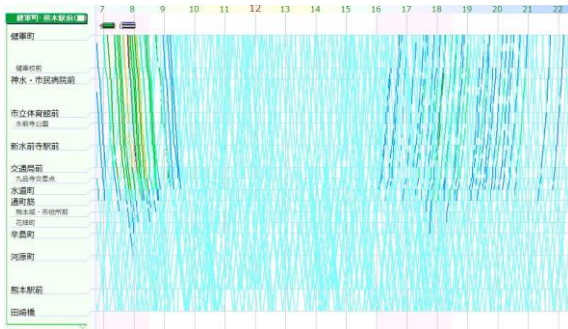


図-14 2/1 クラスター5 (東区通勤者)

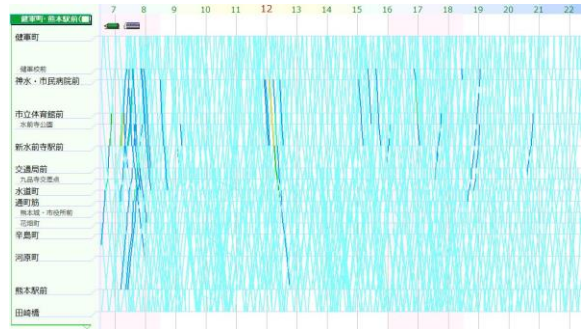


図-15 2/1 クラスター8 (通学者)

表-6 18:00~25:00 A系統健軍方面行の平均遅延時間と各クラスターの乗車数

日付	平均遅延	標準偏差	クラスター									総数
			1	2	3	4	5	6	7	8	9	
2/1	7:38	3:48	335	46	229	101	258	19	205	137	75	1,405
2/2	7:26	3:55	634	81	359	491	272	32	252	241	111	2,473
2/3	8:10	4:05	660	101	359	497	247	25	251	226	135	2,501
2/4	8:28	3:27	708	89	385	460	287	37	242	218	109	2,535
2/5	7:42	3:07	957	145	371	445	251	36	234	239	105	2,783
2/6	7:31	3:31	388	777	74	249	148	22	139	104	49	1,950
2/7	4:33	2:28	428	208	62	220	119	18	130	87	44	1,316
2/8	7:32	3:05	642	92	394	475	290	33	263	255	128	2,572
2/9	7:54	3:26	775	111	343	479	298	38	272	265	139	2,720
2/10	9:19	6:01	1,018	141	410	470	275	48	256	263	114	2,995
2/11	5:11	3:43	542	252	94	242	130	18	132	123	38	1,571
2/12	7:49	4:40	1,057	254	321	452	250	37	250	237	101	2,959
2/13	7:59	3:40	351	849	52	231	139	25	144	97	51	1,939
2/14	4:51	3:29	413	259	50	187	124	19	114	82	28	1,276

6. おわりに

本研究では利用者ごとに利用時間や利用空間の特徴を算出し、クラスタリングによる利用者の分類、また利用実態の可視化を試みた。その中でICカードデータと停留所周辺のデータを使って停留所を空間的な分類を定義し、各利用者の利用時間や利用空間の特徴を算出することによって詳細な利用者の整理ならびに利用実態の分析を可能にした。さらに実績ダイヤとクラスタリングの結果を用いて運行遅延時間に影響を与えている利用者グループを考察した。しかし遅延時間への影響要因の考察はまだ十分に行えていないと感じており、この点は今後の課題としたい。

謝辞：本研究を行うにあたり貴重なデータを提供いただいた熊本市交通局には、この場を借りて深く御礼申し上げます。

参考文献

- 1) 西内裕晶, 轟朝幸: 交通マーケティング手法検討のための IC カードデータを活用した利用者行動特性の把握, 土木学会論文集 F3(土木情報学), Vol.68, No.2, pp.II_8-II_17, 2012.
- 2) 嶋本寛, 北脇徹, 宇野伸宏, 中村俊之: IC カード利用履歴データを用いた公共交通需要変動分析, 土木学会論文集 D3(土木計画学), Vol.70, No.5(土木計画学研究・論文集第 31 巻), pp.I_605-I_610, 2014.
- 3) 西内裕晶, 力石真, 兵頭知, 轟朝幸: IC カードデータを活用した公共交通利用者の利用間隔の変化に関する基礎分析, 第 55 回土木計画学発表会・講演集, 2017.
- 4) 森田琢雅, 溝上章志, 中村嘉明: IC カードデータによる熊本市電利用者の行動特性分析とダイヤ編成への活用, 第 54 回土木計画学研究発表会・講演集, 2016.
- 5) 国土交通省気象庁「過去の気象データ検索」: <http://www.data.jma.go.jp/obd/stats/etrn/>
- 6) 熊本市交通局 HP: <http://www.kotsu-kumamoto.jp/>

(?????.?? 受付)

CLUSTERING OF TRAM USERS USING SPATIO-TEMPORAL CHARACTERISTICS FROM SMART CARD DATA

Takumasa MORITA, Shoshi MIZOKAMI and Yoshiaki NAKAMURA

In recent years it has been attracting attention to analyze usage situation using smart card data in public transportation of rural cities. Cluster analysis to classify users in various cities has been done so far, but most of them do not consider spatial usage characteristics. In this study, we clustered stops using smart card data collected for each stop, and divided zones on public transportation. Next, we calculated each variable usage characteristics including utilization of each zone for each user, and performed non-hierarchical clustering to classify users and organize their features. In addition, we created the daily performance schedule, aggregated the delay time and the number of users per cluster, and considered factors that affect the delay of the tram. As a result, we discovered the possibility that the variance of the delay time becomes large if more than a certain number of passengers use a specific route.