

交通系ICカードデータを用いた長期間の鉄道利用履歴の分析*

Longitudinal Analysis for Railway Users Behavior Based on Smart Card Data*

北野誠一**・中島良樹***・井料隆雅****・朝倉康夫*****

By Seiichi KITANO**・Yoshiki NAKAJIMA***・Takamasa IRYO****・Yasuo ASAKURA*****

1. 背景と目的

従来、鉄道旅客流動の分析には磁気券による改札通過データ^{1) 2) 3)}や、パーソントリップ調査・大都市交通センサスなどのアンケートを用いた集計結果が用いられてきた。前者では正確な鉄道利用者数は取得できるが、利用者毎の起終点(OD)や個人属性は把握できない。一方後者では、ODデータや利用者の属性は得られるが、アンケートに基づいているため正確な乗降時刻や総利用者数を直接計測できない上、一般的にはある1日の行動しか扱うことができない。

近年では特に都市圏を中心にICカード乗車券の導入が進み、その保有者も年々増加している。この交通系ICカードの特長として①長期にわたる多量のODデータが収集できる、1分単位の正確な改札通過時刻が記録されている、カードID(ユニークナンバー)によって個別の利用履歴が識別できることなどがあげられる。これらの特長に加え、本研究で扱うA社系ICカードには④利用者の属性情報が付加されている。ICカードを用いた旅客流動の分析例として、高木⁴⁾らはICカードの乗車記録データを用い、正確な入出場時刻記録を活用した利用者行動の分析方法を提案している。また道路交通の分野ではICカードデータと同様の情報を持つETCデータを用いた交通行動分析の研究事例として、秋元⁵⁾はETCデータの長期にわたる利用履歴を活用し、利用者個人の行動について分析している。

以上をふまえて、本研究では長期間にわたる交通系ICカードデータの特性を活かした分析方法を提案する。詳細な情報が含まれる多量のICカードデータから、今までにない詳細な旅客流動や鉄道利用者属性の特徴を把握することができることに研究の特長がある。

*キーワード: ICカード、長期データ、旅客流動

**学生員、学士、神戸大学大学院工学研究科

(神戸市灘区六甲台町1-1、

TEL078-803-6360、FAX078-803-6360)

***学生員、工修、神戸大学大学院工学研究科

****正会員、工博、神戸大学大学院工学研究科

*****正会員、工博、神戸大学大学院工学研究科

本研究では特に、ICカードデータの集計分析から得られた結果を元に乗降駅間の流動や乗降車駅・居住地の郵便番号情報による空間的分析を行う。

なお長期データを扱うが、本研究では日々の需要の変動に関する時系列分析は対象としていない。

2. 使用データの構造と分析方法

(1) ICカードデータの構造

本研究に用いるICカードデータには、図-1のような項目のデータが含まれている。これらのICカード改札通過データは、分析用に匿名加工された個人番号(分析用ID)によって行動履歴(トリップ情報)と属性情報(フェイス)に関係付けられる。これに加えて別途に作成した郵便番号と市区町村の対応情報や、駅情報などを組み合わせる。

(2) 分析方法

前節に示したICカードデータから、分析の目的に沿った条件を指定してデータを出力する。日付や時間帯などの時間条件や、利用路線・区間・駅などの空間条件、そして利用者の属性条件などを指定する。こうして抽出されたトリップデータを集計し、分析を行う。なおデータ量が膨大であるため、本研究ではSQLを用いてデータの操作を行った。



図-1 ICカードデータ ER図

このようなトリップデータの数え上げ、つまり改札通過枚数の集計結果が主要な分析対象ではあるが、分析用IDによって個人を区別できるという特性を活用して、同一利用者の重複利用を省き、のべ利用回数を構成する“出現人数”の集計も可能である。本研究のように長期データを利用する場合、一定期間内の“出現人数”と“のべ利用回数”の比較分析は定期的利用と臨時的利用の傾向把握などに有効である。また、個人を区別できるという特性は、一定期間における利用回数の集計も可能である。

分析結果の評価の段階では、集計結果の相互比較が特に重要である。この比較に際し、実数に差がある場合は割合を用いた方が比較しやすい。例えば各駅の利用者属性データを比較する場合、駅ごとの利用者数に大きな差があるので利用者属性の割合を用いた方が傾向の解釈に役立つ。

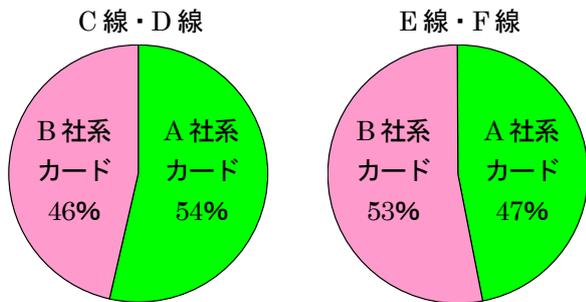


図-2 A社線利用者“出現人数”券種属性

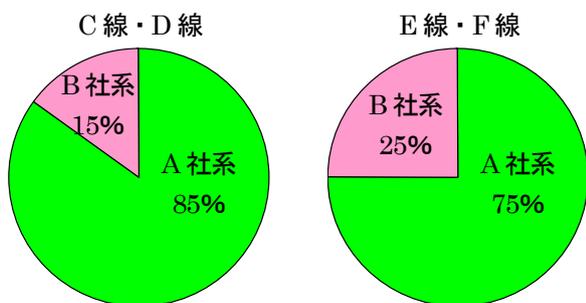


図-3 A社線利用者“のべ利用回数”券種属性

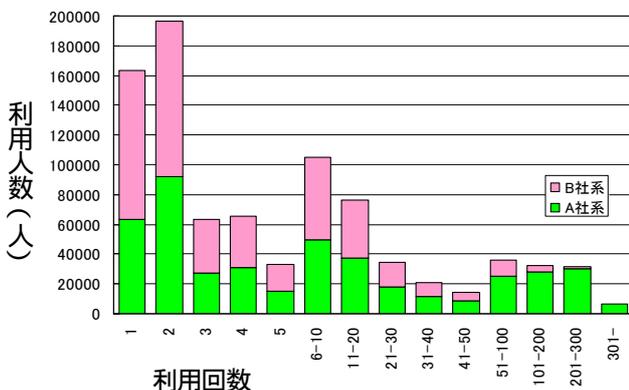


図-4 A社線利用者ヒストグラム

3. カード券種別の利用者属性

(1) のべ利用回数と出現人数

図-2は2007年上半期ICカードデータの全利用回数から、A社各線を対象として、ICカード利用者数の割合である“出現人数割合”を図示したもので、図-3は同データの“のべ利用回数割合”を示したものである。図-3より、A社線内の改札ではA社系カードの使用回数が多いと考えられるが、分析用IDによって識別された出現人数を見るとA社系カード所有者とB社系カード所有者の割合がほぼ同じであることがわかる。

図-4からもA社系ICカード所有者は定期利用などによる自社線内の複数回利用が多いことがわかる。その一方で、低頻度利用に限定すれば、他社カード所有者の利用も半数を超えることがわかる。

図-2から、C線系路線でのA社系ICカード利用者は54%、E線系路線では47%という相違点がみられる。この7%の差が、利用駅によってどのような特徴となって現れるかについて、以下で詳しい分析を行う。

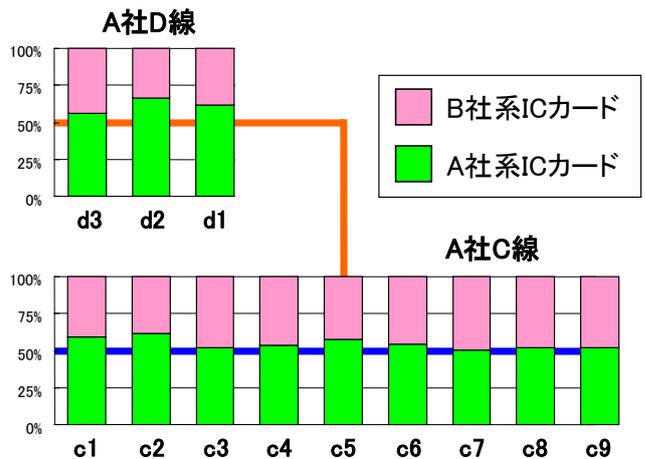


図-5 A社C線系路線 駅別利用者 券種属性

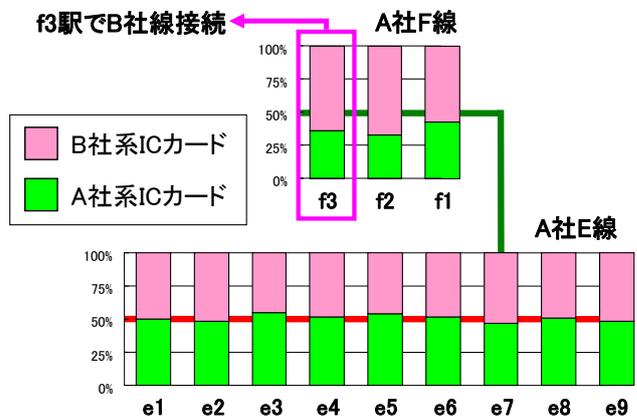


図-6 A社E線系路線 駅別利用者 券種属性

(2) 路線ごとのカード種別割合の比較

各線利用駅ごとに2007年上半期でのA社・B社系ICカードの利用者数を抽出し、グラフ出力して比較する。各駅の両社の利用者数割合を路線図と対応させて配置した。

図-5、図-6から、C・E線の各駅ともA・B両社カードの利用率はほぼ50%で大きな差はみられない。

図-2で示したA社C線系路線とE線系路線におけるA社系ICカード利用回数割合の差異(7%)は、D線・F線におけるA・B社カードの利用率の差によるものと思われる。D線ではA社の割合が6割であるのに対し、F線では逆に4割程度である。F線各駅のB社系ICカード利用者が多い理由は、F線の起点駅でありB社線との乗換駅でもあるf3駅の影響が大きいと考えられる。つまり①F線住民でB社線沿線に勤務するB社系ICカード利用者の影響や、またF線沿線に多くの工場が存在することから、②沿線企業従業員の中にB社線沿線居住のB社系ICカード利用者が存在していることが考えられる。②に関して、次章でさらに詳しく分析を行う。

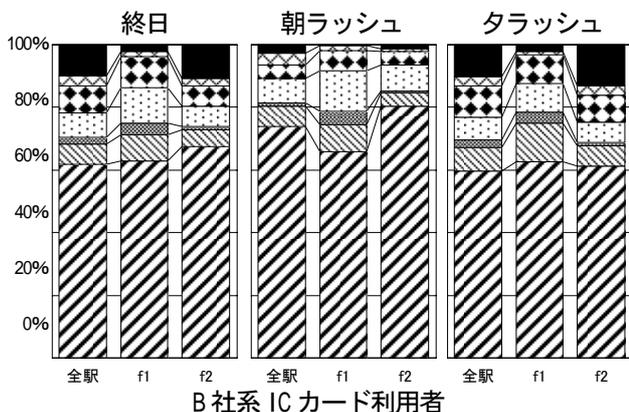
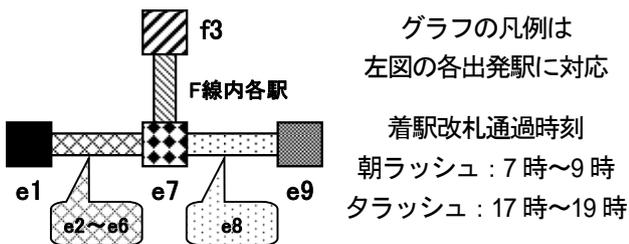
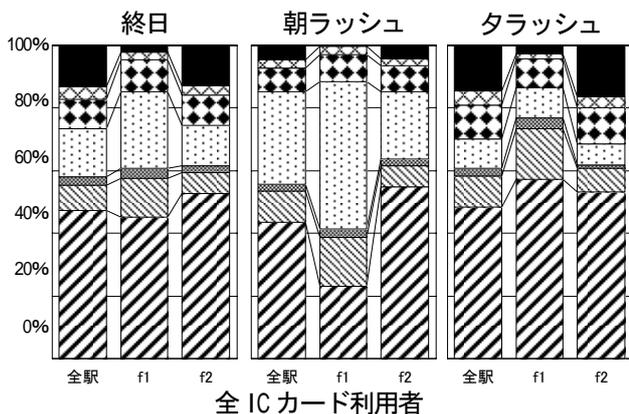


図-7 F線各駅降車利用者の乗車駅分布

4. ICカード利用者の居住地分布

図-7に、F線各駅降車利用者の乗車駅割合図を示す。全利用者に比べて、B社系ICカード利用者の“F線f3駅→f1・f2駅”の利用率(太線斜めハッチ部分)が10~15%ほど高くなっている。このことから、図-6におけるF線各駅でのB社系ICカード利用者率の高さは、f3駅の影響であることがわかった。

ICカードデータには1分単位の正確な改札通過時刻が記録されており、着駅(f1・f2)の改札通過時刻を朝夕の通勤ラッシュ2時間ずつ抽出して分析を行った。朝ラッシュ(7時00分~8時59分)とタラッシュ(17時00分~18時59分)の間にも分布の差異が見られる。f3駅からの利用者割合に着目すると、全利用者では朝ラッシュ時のほうが割合は低く、B社系ICカード利用者ではタラッシュ時のほうが低い。このことは、B社系ICカード所有者によるF線沿線企業への通勤利用の多さを反映しているものと思われる。

図-8に、2007年上半期におけるF線区間の利用頻度分布を示す。この図から、F線区間ではB社系カード利用者がA社系カード利用者より多く存在し、高頻度利用者つまり通勤で利用していると考えられる利用者の中にもB社系ICカードを利用している方が多く存在することがわかる。

そこで、F線のf3駅→f1・f2駅を利用するB社系ICカード所有者がf3駅までB社線を利用しているのか、居住地分布を分析することを考える。しかし、B社系ICカードには郵便番号などの属性データが含まれておらず、またB社線の利用履歴は入手不可能なため、B社系ICカード利用者の行動分析を行うことはできない。

そこで代替的に、“A社F線f3駅→f1・f2駅”を利用したA社系ICカード利用者の郵便番号データを用いて、f3駅でB社線からA社F線に乗り換え、沿線企業に勤務していると考えられるA社系ICカード利用者の分布を分析することにした。

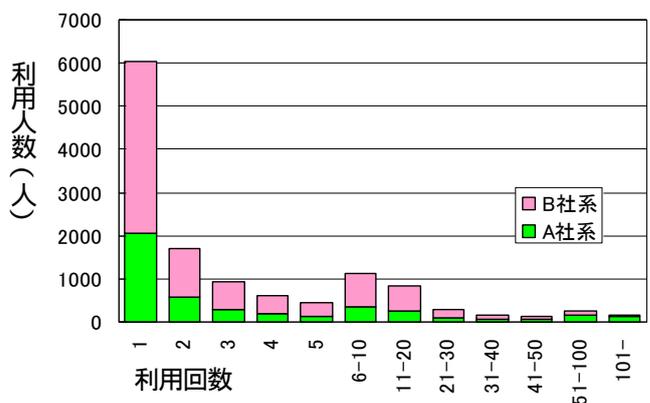


図-8 f3駅→f1・f2駅間利用頻度分布

図-9、10のように郵便番号データを集計して地図上に表現することによって、F線利用者の分布区域を把握することができる。図-10は、2007年上半期内に50回以上“F線f3駅→f1・f2駅”を利用した方の分布である。ここで50回とは上半期に週1回往復利用した場合の回数（6ヶ月×4週×週1日×1日2回≒50）であり、通勤利用・定期的利用の目安とした。

図-9から、三重・滋賀・和歌山県や淡路島在住の“F線f3駅→f1・f2駅”利用者が確認され、F線利用者の居住地範囲の広さがわかる。さらに図-10から、F線沿線企業への通勤と考えられる高頻度利用者の居住地の広さも把握できる。彼らのほとんどがf3駅までB社線を利用していると考えられるが、A社系ICカード所有者なので、居住地が同カードを導入している鉄道・バス会社の沿線であるということも推測できる。

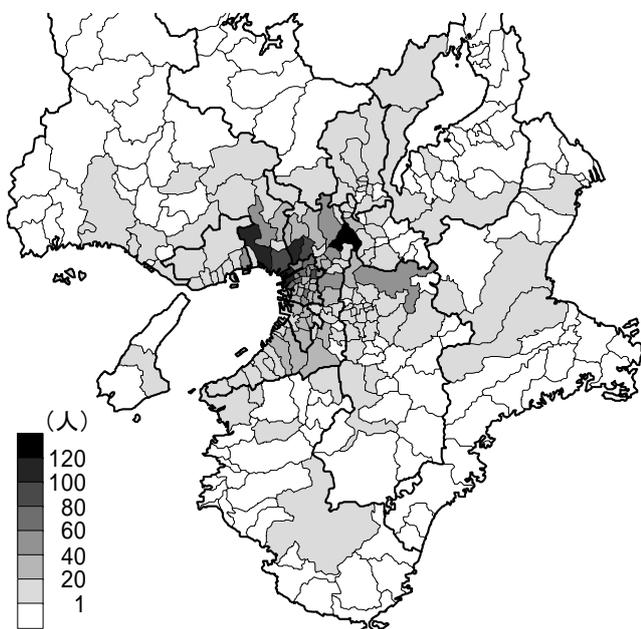


図-9 F線 ICカード利用者 居住地分布

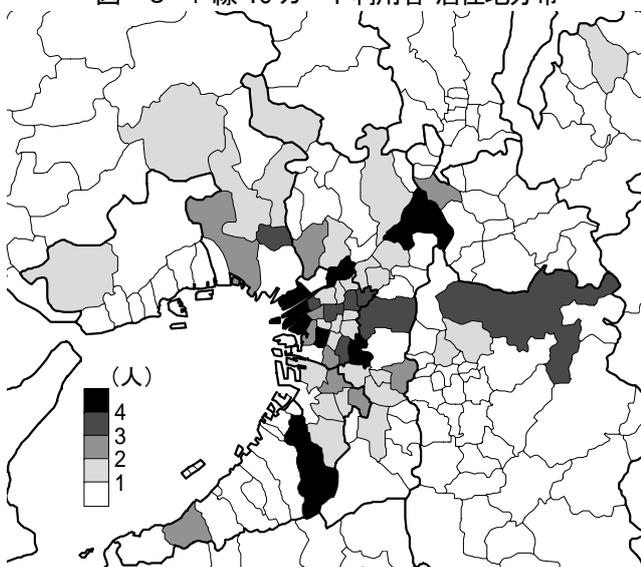


図-10 F線 ICカード 高頻度利用者 居住地分布

この居住地分布分析結果と、図-7におけるB社系ICカード利用者の“F線f3駅→f1・f2駅”利用率、および図-8より該当区間高頻度利用者におけるB社系ICカード利用者の存在から、F線を高頻度利用するB社系ICカード利用者の分布も図-10のように広範囲である可能性があると推測できる。

5. まとめ

交通系ICカードデータの特性を活用し、本研究ではA社路線別のカード券種利用者比較分析によって路線ごとの特徴を発見し、A社F線とB社系ICカード利用者の関連性を示すことができた。今まで経験的に認知はされていたが、ICカードデータの分析によって数量的な把握が可能となった。

本研究で得られた結果は次のように要約できる。

- ①個人履歴別の集計が可能であることを示したこと。
- ②正確な時間帯での比較分析ができることを示したこと。
- ③利用者の空間分布を示したこと。

今後は、7桁の郵便番号情報を活用した市区町村区分よりも詳細な居住地分布表現方法の開発や、その他の属性情報を用いた分析を進める一方で、現段階では十分に活用できなかったクラスタリングや頻出パターン抽出など、データマイニングの解析手法を用いたデータ分析を検討していく。

参考文献

- 1) 明星秀一：自動改札機データを活用した旅客流動推定手法，鉄道総研報告，Vol.20，No.2，pp.23-28，2006。
- 2) 岡村敏之，藤原章正，神野優，杉恵頼寧：共通プリペイドカードによる都市圏内公共交通乗車記録の特性分析，土木計画学研究・論文集，Vol.19，No.1，pp.29-36，2002。
- 3) 窪田崇斗，古谷聡，家田仁：自動改札機・車両応荷重データを用いた都市鉄道の混雑率推定手法に関する研究，土木計画学研究・講演集，Vol.36，CD-ROM，2007。
- 4) 高木勇弥，日下部貴彦，井料隆雅，中島良樹，朝倉康夫：ICカード乗車券システムから得られる改札通過データを用いた鉄道旅客流動の分析，土木計画学研究・講演集，Vol.36，CD-ROM，2007。
- 5) 秋元健吾，小根山裕之，西内裕晶，割田博：ETCデータを用いた首都高速道路における入口選択行動の分析，第27回交通工学研究発表会論文報告集，Vol.19，pp.193-196，2007。