

ICカードシステムによる改札通過データを活用した乗車列車推定方法の開発*

Estimation of Passenger's Train Choice in Railway Network with Smart Card Ticket System *

日下部貴彦**・高木勇弥***・井料隆雅****・朝倉康夫****

By Takahiko KUSAKABE**・Yuya TAKAGI***・Takamasa IRYO****・Yasuo ASAKURA****

1. はじめに

近年、鉄道などの公共交通機関でICカード乗車券(交通系ICカード)の導入が進んでいる。交通系ICカードの乗車履歴データ(以下ICデータと呼ぶ)は、各鉄道会社で管理されており、鉄道利用者に対するマーケティング等への活用が期待されている。

ICデータの特徴には、①ICカード(利用者)毎の利用履歴が記録されていること、②改札通過時刻を1分単位という詳細な時間解像度で記録していること、③長期にわたる改札通過の観測データが収集できること、④入出場記録が完結されなければ利用者は継続利用できないためデータの欠損が少ないことがあげられる。これらの特徴よりICデータは、時刻毎の利用者の鉄道利用の状況の詳細な分析や、利用者行動の長期にわたる変動を分析するのに適したデータであるといえるだろう。

これまでの鉄道利用者の行動を推定する研究では、大都市交通センサスによるデータを用いる方法が主流であった¹⁾²⁾。しかし、センサスを用いた方法では、都市全体の鉄道旅客流動を把握することが可能である半面、日々の変動をとらえることはできない。一方、各車両の応荷重の履歴データや自動改札機の通過履歴によるデータを用いることで日々の旅客の変動を推定することも可能になってきている³⁾⁴⁾⁵⁾。これらの手法では、個々の車両の乗客数や駅間を移動する乗客数の日々の変動を捉えることが可能である。しかし、これらのデータを用いた方法でも、各利用者を識別した上で変動を捉えることはできない。これに対し、ICデータでは、利用者毎の乗降駅と乗降時刻を知ることができる。その情報から、利用者毎に乗車列車を推定できれば、様々な利用者層での列車選択の特性を明らかにすることが期待できる。

本研究の目的は、ICデータと列車ダイヤを用いて、

*キーワード：ICカード、列車選択、交通量推定

**学生員、工修、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、学士、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

各ICカード利用者が乗車した列車を推定する方法を開発することにある。本研究の方法では、利用者ごとのDay-to-dayの列車選択を推定結果として得ることができる。その結果は、ダイヤ改正に代表される時間的な鉄道サービスの変更に対する個々の利用者の利便性の評価に適用できる。たとえば、ダイヤの改正前に比べ改正後には遅い時刻に出発する列車に乗車し、より早く目的地に到着している利用者が観測されれば、その利用者の利便性は向上していると考えられる。逆に早い時刻に出発する列車に乗車し遅く到着している利用者は利便性が低下していると考えられる。利用者ごとの利便性を評価することで、ダイヤ改正等の施策がどのような利用者にとどの程度の利便性の向上をもたらしたかなどを詳細に調べることができ、鉄道利用客に対するCRM(Customer Relationship Management)に有用な情報が得られると期待される。なお、本研究で対象としている鉄道ネットワークの形状は、駅間ODが決まれば一意に空間的な経路が決定する形状の複線区間に限定される。また、優等列車の設定がある列車ダイヤを想定している。

2章では、本研究の手法を適用することができるICデータの仕様をまとめる。3章で推定方法を構築し、4章でICデータを用いた分析を行い、推定精度等に関して考察を行う。

2. 対象データ

本研究で想定しているICデータの形式は、レコード

| ICデータ | | | | | |
|-------------|-----|-------|-----|-------|----------|
| 使用年月日 | 入場駅 | 入場時刻 | 出場駅 | 出場時刻 | カードID |
| 2007年10月13日 | A駅 | 7:10 | C駅 | 7:23 | A25687DK |
| 2007年10月13日 | A駅 | 7:11 | C駅 | 7:23 | B68677DS |
| 2007年10月13日 | A駅 | 7:11 | B駅 | 7:18 | B67732RR |
| ⋮ | | | | | |
| 2007年10月13日 | C駅 | 18:00 | A駅 | 18:17 | B89751RR |
| ⋮ | | | | | |

| 列車ダイヤデータ | | | | | |
|----------|-----|------|-----|------|--|
| 列車識別番号 | 出発駅 | 出発時刻 | 到着駅 | 到着時刻 | |
| 1 | A駅 | 7:12 | B駅 | 7:17 | |
| 1 | B駅 | 7:19 | C駅 | 7:22 | |
| 2 | A駅 | 7:22 | B駅 | 7:27 | |
| ⋮ | | | | | |

図-1 ICデータ・列車ダイヤデータの例

毎に一回の鉄道利用が記録されているものである。各レコードは「使用年月日、入場駅、入場時刻、出場駅、出場時刻、利用者のカードID」から構成される。入場時刻及び出場時刻は、1分単位で記録されているとする。

列車ダイヤデータは各列車の駅到着時刻と出発時刻が1分単位で記録されているものを想定する。列車ダイヤデータのレコードは、「列車識別番号、出発駅、出発時刻、到着駅、到着時刻」が記録されているものとする。列車識別番号とは、列車を識別するために列車毎に割り当てられる番号である。各データの例を図-1に示す。

3. 乗車列車推定方法の構築

本研究での手法は、高木らによる先行研究⁶⁾でのICデータ分析の結果に基づいて、ICカード利用者の行動を仮定し、この仮定のもとで列車ダイヤから乗車可能な列車の組み合わせを列挙するものである。第一節で乗客の行動に関する仮定を説明する。第二節では、乗客の行動を列車ダイヤ上で記述するために、列車ダイヤを有向グラフとして記述する(以下、これを乗換ネットワークと呼ぶ)。第三節で乗換ネットワークを用いて、利用される可能性のある列車選択の組み合わせについて説明する。第四節では、列車選択の組み合わせを推定するためのアルゴリズムについて述べる。

(1) 乗客行動の仮定

高木らの研究⁶⁾では、図-2に示すように1分単位で観測されたICデータによる利用者の入出場時刻の組み合わせをプロットし、列車ダイヤを重ね合わせることによって分析を行っている。この図では、縦軸に出発駅であるA駅での時刻、横軸に到着駅であるB駅の時刻をとり、A駅で入場しB駅で出場した利用者の入出場時刻に対応する点をプロットしている。また、A駅での出発列車をその出発時刻に応じた横線で示し、B駅の到着列車を到着時刻に応じた縦線で示している。図-3は鉄道会社A社のICデータをこの方法でプロットしたものである。この図の列車カテゴリI, J, K, Lはそれぞれ「特急列車A」、「特急列車B(特急列車Aと停車駅が異なる)」、「急行相当列車」、「各駅停車」を表している。この図からわかるように、乗客は列車の出発時刻にあわせて駅に入場する傾向があり、列車到着後すぐに出場する傾向があることがわかる。高木らの研究では、この傾向を利用することで各利用者の乗車列車が判別可能であることを示唆している。しかし、具体的な判別方法を提案するにはいたっていない。

本研究では、高木らのICデータの分析から得られた「乗客は列車の出発時刻にあわせて駅に入場する傾向があり、列車到着後すぐに出場する傾向がある」という知

見を踏まえて、鉄道利用者の列車選択行動に対して次のような仮定を行う。

乗客は、出発駅への入場から列車乗車までの「待ち時間」と、到着駅での降車から出場までの「時間ロス」の和が最小となるように列車を選択する。

①を満たす列車選択の組み合わせが複数ある場合には、乗り換え回数を最小にするような行動をする。

①と②の双方を満たす複数の列車選択の組み合わせがある場合には、それらを等確率で選択する。

仮定①は、高木らの研究結果での、乗客が列車の出発時刻にあわせて駅に入場する傾向と、列車到着後すぐに出場する傾向を反映したものである。仮定①を満たす列車選択の組み合わせの中で最も合理的なものは、乗り換え回数が最も少ない組み合わせの中であると考えられ

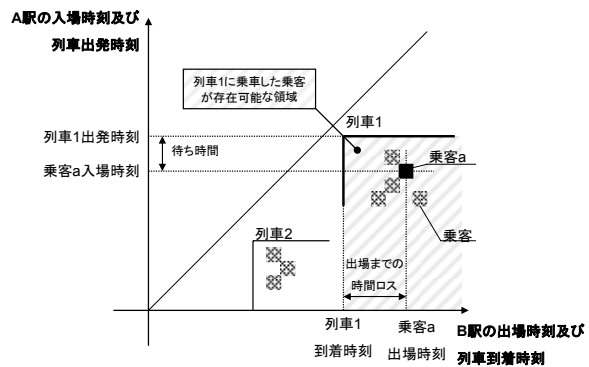


図-2 利用者の入出場時刻と列車ダイヤ

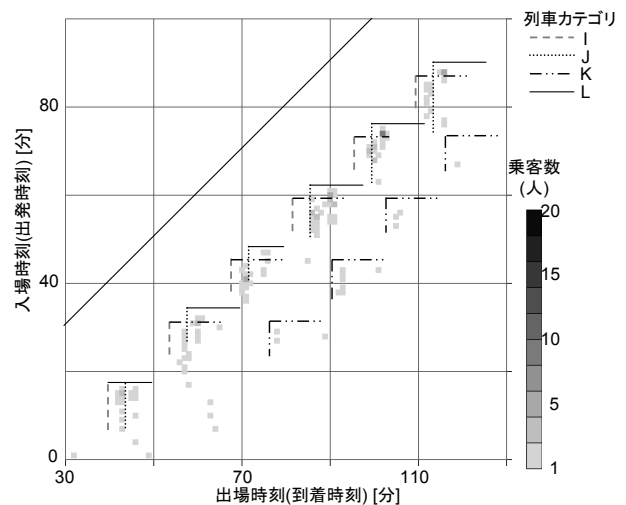


図-3 鉄道会社A社の利用者の入出場時刻と列車ダイヤ

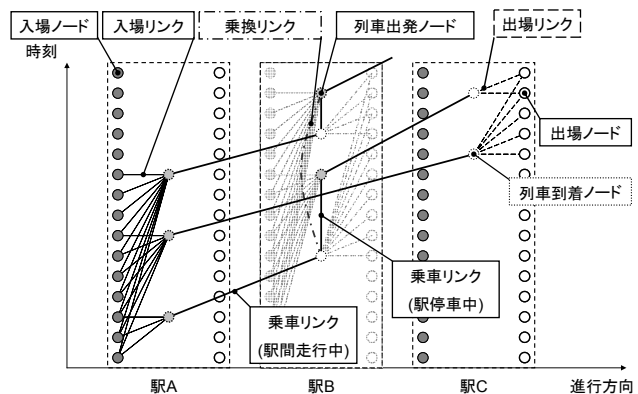


図-4 乗換ネットワークの例

る。仮定②は、選択される可能性が低い（乗り換え回数が多い）組み合わせが選択肢として抽出されることを避けるために設定した。仮定③は、仮定①②を満たす列車選択の組み合わせが複数ある場合には、その中からランダムにひとつの組み合わせを取り出すものである。一見すると仮定①と②を同時に満たす組み合わせは出現しにくいと思われるかもしれない。しかし、乗降駅間に複数の乗り換え可能な駅がある場合には、入場時に乗車した列車（列車A）と出場時に乗車した列車（列車B）の組み合わせを特定できたとしても、どの駅で列車Aから列車Bに乗り換えたかを一意に決めることはできない。例えば、ある緩行列車から複数の駅で乗り換え可能な優等列車に乗り換えを行った場合に、どの駅で乗り換えたかを特定できないことから複数の列車選択の組み合わせが残る。したがって、ICデータと列車ダイヤデータだけでは、複数の組み合わせが残った場合には等確率で選択という仮定を設けざるを得ないのである。このような場合の列車選択に関する行動ルールのモデル化は今後の課題としたい。

仮定①～③のすべてを満たす列車選択行動を記述するためには、仮定①②を満たす列車選択の組み合わせをすべて列挙し、複数の組み合わせが残った場合には、その中からランダムにひとつの組み合わせを取り出せばよい。そこで以降では、1回の鉄道利用毎に仮定①②を満たす乗降駅間の列車選択の組み合わせをすべて列挙するための効率的な手法を構築する。

(2) 乗換ネットワークの構築

入出場駅及び入出場時刻に応じて利用者が利用し得る列車の組合せを列挙する。このために、まず列車ダイヤを時刻と駅の属性を持った図-4のような有向グラフとして記述し、これを乗換ネットワークと呼ぶこととする。乗換ネットワークを用いることで、利用者が乗車した列車及び乗換駅の組合せはネットワーク上での「経路」として記述することができる。以降では、乗換ネットワークのノード及びリンクの定義を行う。

乗換ネットワークでのノード(v)は、入場ノード、出場ノード、列車到着ノード、列車出発ノードの4種類で構成される。各ノード集合をそれぞれ、 V_s 、 V_t 、 V_a 、 V_d とあらわす。入場ノードと出場ノードは、各駅の時刻毎に定義する。各ノードが示す駅は $Station(v)$ で表し、時刻は $Time(v)$ であらわすこととする。また、それぞれの入場ノードと出場ノードは、それぞれ $Station(v)$ と $Time(v)$ の組合せで一意に特定できる。

列車到着ノードと列車出発ノードは列車ダイヤデータから作成するものである。列車到着ノードは各列車の各駅への到着時刻に対応する。また、列車出発ノードは各駅からの出発時刻に対応する。それぞれのノードは、入場ノードなどと同様に時刻及び駅を属性値として持つ。

また、どの列車の到着及び出発を表すのかを示すために列車毎に割り当てられた列車識別番号を属性値として持つ。各ノードの列車識別番号は、 $Train(v)$ とする。

乗換ネットワークでのリンク(a)は、入場リンク、出場リンク、乗車リンク、乗換リンクの4種類のリンクで構成する。それぞれのリンク集合は、 A_s 、 A_t 、 A_r 、 A_c と記述し、リンク集合全体を A とする。また、各リンクの起点ノード及び終点ノードは、それぞれ $V_{st}(a)$ 、 $V_{ed}(a)$ で表す。すべてのリンクに共通する条件は、

$$Time(V_{ed}(a)) - Time(V_{st}(a)) \geq 0 \quad a \in A \quad \dots\text{式1}$$

である。この条件により乗換ネットワークは自己ループのないグラフになる。また、単線又は複線の路線の列車ダイヤを用いる場合には、単純グラフとなる。

入場リンクは、入場ノードから列車出発ノードへの有向リンクであり、出場リンクは、列車到着ノードから出場ノードへの有向リンクである。入場リンク及び出場リンクは、式1の条件と

$$Station(V_{st}(a)) = Station(V_{ed}(a)) \quad a \in A_s \cup A_t \quad \dots\text{式2}$$

を満たす。

乗車リンクは、列車を表すリンクであり、列車ダイヤデータから作成する。駅間を走行中の列車と駅に停車中の列車の双方をリンクとして表す。駅間を走行中の列車を表す乗車リンク a は、起点ノードは列車出発ノード、終点ノードは列車到着ノードであり、式1と

$$\begin{aligned} Station(V_{st}(a)) &\neq Station(V_{ed}(a)) \text{かつ} \\ Train(V_{st}(a)) &= Train(V_{ed}(a)) \end{aligned} \quad \dots\text{式3}$$

を満たす。停車中の列車を表す乗車リンク a は、起点ノードは列車到着ノード、終点ノードは列車出発ノードであり、式1と

$$\begin{aligned} Station(V_{st}(a)) &= Station(V_{ed}(a)) \text{かつ} \\ Train(V_{st}(a)) &\neq Train(V_{ed}(a)) \end{aligned} \quad \dots\text{式4}$$

を満たす。

乗換リンクは、各駅で異なる列車への乗換を示すリンクである。起点ノードは列車到着ノード、終点ノードは列車出発ノードであり、式1と、

$$\begin{aligned} Station(V_{st}(a)) &= Station(V_{ed}(a)) \text{かつ} \\ Train(V_{st}(a)) &\neq Train(V_{ed}(a)) \quad a \in A_c \end{aligned} \quad \dots\text{式5}$$

を満たす。

(1)に示した仮定①②を満たす列車選択の組み合わせを推定するために各リンクのコストを2次元ベクトルで定義する。リンク a のコストは、

$$C(a) = (\tau_a, c_a) \quad \dots\text{式6}$$

ただし,

τ_a : 入場から乗車または, 降車から出場までにかかる時間

c_a : 乗換を示す離散変数. $c_a = \begin{cases} 1 & a \in A_c \\ 0 & \text{otherwise} \end{cases}$

とする. これより, それぞれのリンクのコストは,

$$C(a) = \begin{cases} (Time(V_{ed}(a)) - Time(V_{st}(a)), 0) & \text{if } a \in A_s \cup A_t \\ (0, 0) & \text{if } a \in A_r \\ (0, 1) & \text{if } a \in A_c \end{cases}$$

…式7

となる.

(3) 乗客行動の仮定を満たす経路

ICデータのレコードから, 各鉄道利用に関する入場駅 s_o , 入場時刻 t_o , 出場駅 s_d , 出場時刻 t_d が得られたとき, 経路の起点ノードは,

$$(s_o, t_o) = (Station(v_o), Time(v_o)) \quad v_o \in V_s \quad \dots \text{式8}$$

を満たす出発ノード v_o あり, 終点ノードは

$$(s_d, t_d) = (Station(v_d), Time(v_d)) \quad v_d \in V_t \quad \dots \text{式9}$$

である. v_o, v_d を起終点とする経路の集合を R_{od} とする. R_{od} のうち, (1)で示した仮定をみたすものは,

$$R_c = \left\{ r_c \mid r_c \in R_0, \sum_{a \in A(r_c)} c_a \leq \sum_{a \in A(r_0)} c_a \text{ for } \forall r_0 \in R_0 \right\}$$

…式10

ただし,

$A(r)$: 経路 r に含まれるリンクの集合

$$R_0 = \left\{ r_0 \mid r_0 \in R_{od}, \sum_{a \in A(r_0)} \tau_a \leq \sum_{a \in A(r)} \tau_a \text{ for } \forall r \in R_{od} \right\}$$

と表すことができる. R_0 は仮定①を満たす列車選択の組み合わせの集合であり, 入場時の列車待ち時間と出

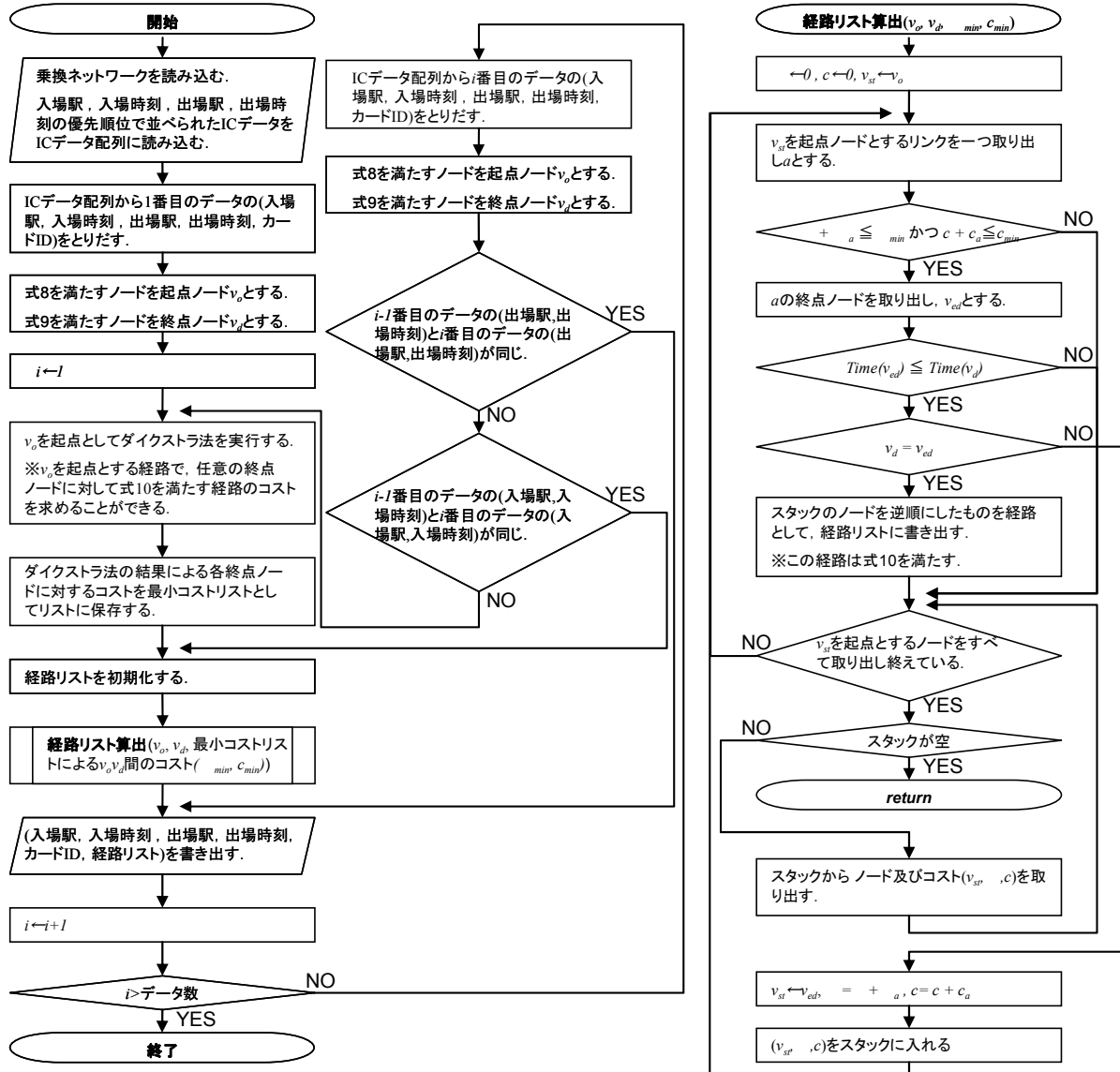


図-5 乗車列車推定のためのアルゴリズム

場時の時間ロスの和 $\sum_{a \in A(r_0)} \tau_a$ が最小となる列車選択の組み合わせを表している。 R_c は、仮定①を満たす列車選択の組み合わせのうち、仮定②を満たすものを表しており、列車選択の組み合わせに含まれる乗換回数 $\sum_{a \in A(r_c)} c_a$ が最小となる組み合わせを表している。

(4) アルゴリズム

ICデータの量は膨大であり、すべての駅間で時々刻々の列車選択の組み合わせの抽出を行うための計算量は極めて大きい。そこで、計算負荷を削減するための効率的なアルゴリズムが必要となる。

本節では、ICデータに含まれるすべての鉄道利用について、式10を満たす乗車列車を推定するためのアルゴリズムについて述べる。図-5に推定アルゴリズムの流れ図を示す。図のアルゴリズムでは、まず、経路探索の回数を減らすために、ICデータのレコードを入場駅、

入場時刻、出場駅、出場時刻の順に並べ替える。次に、レコードを一つ取り出し、ダイクストラ法を用いて、式10を満たす乗換ネットワーク上の経路のコストを求める。「経路リスト算出サブルーチン」で、式10を満たす経路すべてを列挙する。経路リスト算出サブルーチンでは、計算時間を短縮するために、時刻が入場時刻と出場時刻の間で、かつ入場ノードを起点とした経路のコストがダイクストラ法で求めたコスト以下であるサブネットワークのみを探索している。最後に、サブルーチンの算出結果を出力し、次のレコードの乗車記録について推定を行う。この際、前のレコードと入場駅、入場時刻、出場駅、出場時刻のすべてが同じ場合には、探索を省略して、前のレコードの推定結果をそのまま出力する。入場駅、入場時刻のみが前のレコードと同じ場合には、ダイクストラ法による計算を省略して、経路リスト算出サブルーチンを実行し、推定結果を出力する。

4. 推定結果を用いた分析例

本章では、鉄道会社A社のICカード乗車券システムにより取得されたICデータを用いて乗車列車の推定を行う。その狙いは、提案した方法論が実装可能であるかどうかを確認し、列車特定により分析できる内容を示すと同時に、方法論の持つ問題点を把握することにある。第一節では、用いたICデータの仕様を説明する。第二節では、推定結果を分析することによって推定手法の問題点等を明らかにする。第三節では、特定のICカード利用者に着目して分析を行う。

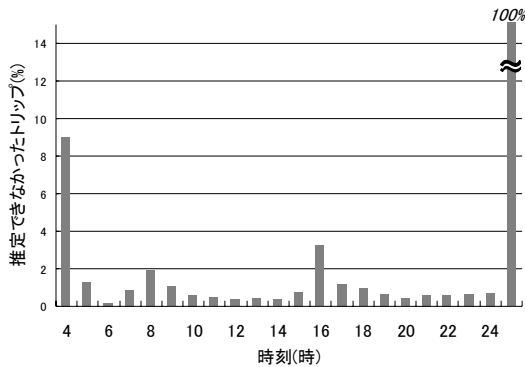


図-6 推定できなかったトリップの割合(時間帯別)

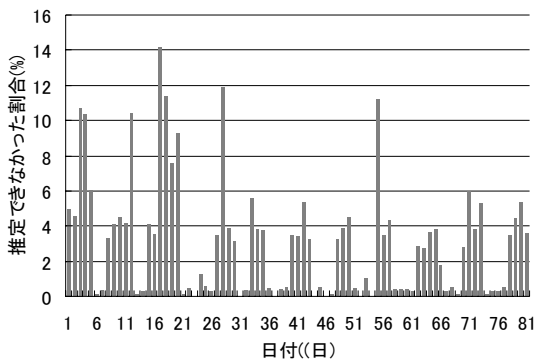


図-7 推定できなかったトリップの割合(日別・16時台)

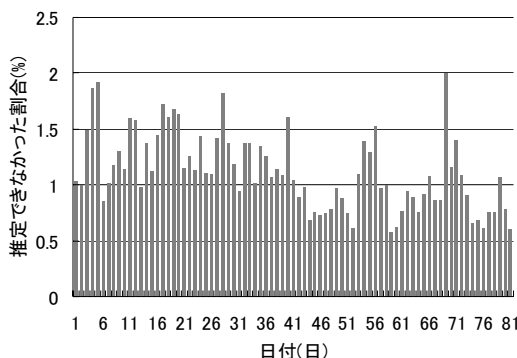


図-8 推定できなかったトリップの割合(日別)

(1) データ

ICデータと列車ダイヤデータの項目は2章で示したものである。推定の対象期間は平日81日間であり、推定期間の第40日目にはダイヤの改正が実施されている。推定対象とした路線は、A社線の一部区間の上り線であり、全区間で複線である。また、駅数は42駅である。なお、列車ダイヤデータでは、ダイヤ乱れや臨時列車の運行を考慮していない。ICデータの対象とした改札間トリップ(以下トリップと呼ぶ)は、対象区間の上り線で入出場した1,425,345トリップである。

なお、列車の運行間隔が概ね20分以内であることを利用して、計算速度を向上させるために、乗換ネットワークの入場リンクは列車出発前20分以内にある入場ノ

表-1 推定結果

| | 平均 | 標準偏差 | 最大 |
|--------------------------|------|------|-------|
| トリップあたりの列車選択の組み合わせ数(個) | 1.31 | 0.96 | 68 |
| 乗換え回数(回) | 0.46 | 0.56 | 3 |
| 入場直後の列車待ち時間と出場の時間ロスの和(分) | 5.47 | 2.69 | 36.25 |

ドとの間に設定した。出場リンクについても列車到着後20分以内にある出場ノードとの間のみに設定した。

(2) 推定結果

期間中のすべてのトリップについて推定にかかった時間は195分であった。1日分あたりの計算時間は平均で2.4分であり、1トリップあたりの推定時間は平均で8.2ミリ秒である。なお、推定で用いたコンピュータは、CPUがIntel Xeon 5110 (1.60GHz) ×2でメモリが4GBである。

すべてのトリップのうち、行動仮説を満たす乗車可能な列車が1つも残らず、乗車列車を推定できなかったものは15,923トリップあった。これは全体トリップの1.1%に相当する。推定できなかったトリップが全トリップに占める割合を時間帯別に見たものを図-6に示す。乗車列車を推定できなかったトリップは、4時台と16時台に集中していることがわかる。なお、25時台は発生率100%となっているが観測数は1である。16時台の推定できなかったトリップの割合を日別に示したものが図-7である。この図によると、乗車列車が推定できないトリップはいくつかの日に集中して発生していることが分かる。また、乗車列車が推定できなかったトリップの割合を全時間帯を対象として示した図-8からも同様の傾向がある

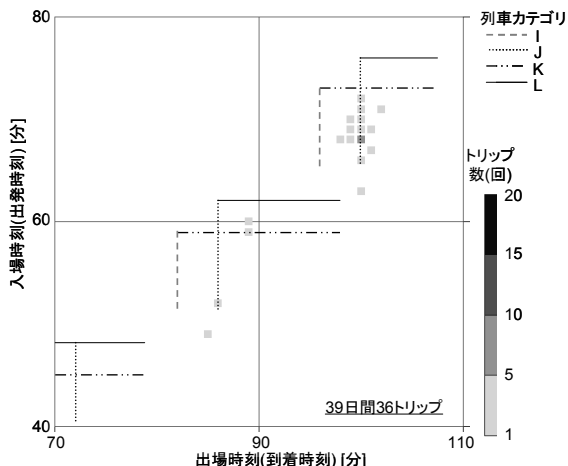


図-9 乗客Qの入出場時刻 (ダイヤ改正前)

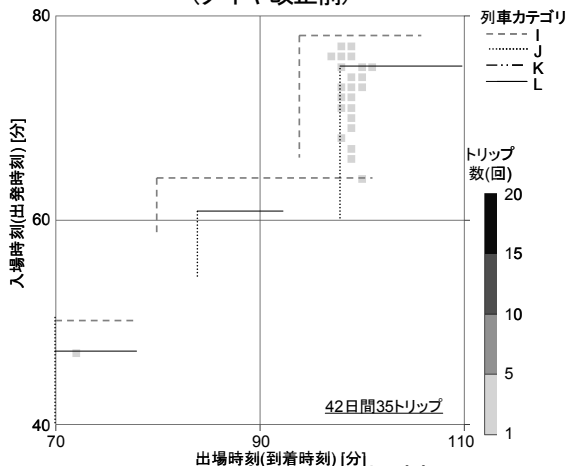


図-10 乗客Qの入出場時刻 (ダイヤ改正後)

ることがわかる。これらの図より、推定できないトリップは、特定の日の特定の時刻に集中して発生していると考えられ、ダイヤ乱れや臨時列車運行等による影響であると推察される。したがって、実際の列車の運行記録が使用可能な場合には、運行記録をダイヤデータに用いることが望ましいといえるだろう。

列車が推定されたトリップについての、トリップあたりの列車選択の組み合わせ数、乗換え回数、入場直後の列車待ち時間と出場の時間ロスの和を表-1にまとめる。トリップあたりの乗車可能な列車選択の組み合わせ数は平均1.31個、標準偏差0.96個であり、多くのトリップでは、1~数個程度の列車選択の組み合わせが選択肢として抽出されていることがわかる。乗換え回数は、ほとんどのトリップで乗換えなしか、たかだか1回程度の乗換えであると推定されていることがわかる。また、乗換え回数の最大値は3回であり、著しく非現実的な乗換を行う列車選択の組み合わせは抽出されていないと考えられる。入場直後の列車待ち時間と出場の時間ロスは、平均5.47分であった。また、最大値は36.25分である。この値には、第二節での乗換ネットワーク設定時に入出場リンクをそれぞれ列車出発前後20分の入出場ノードとの間に設定したことが影響している可能性がある。

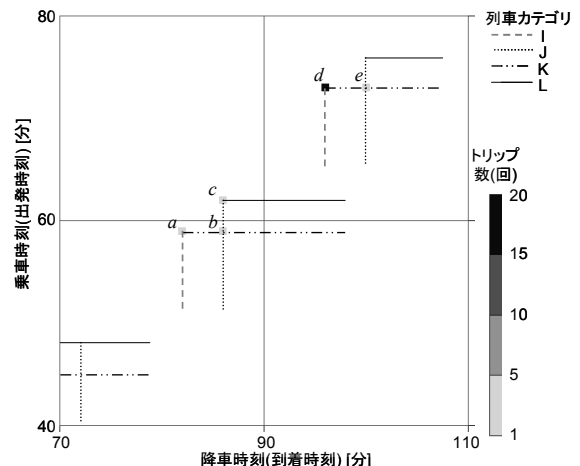


図-11 推定結果による乗客Qの乗降時刻 (ダイヤ改正前)

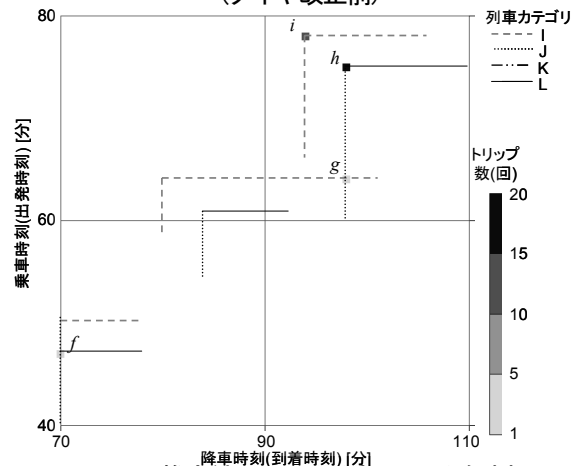


図-12 推定結果による乗客Qの乗降時刻 (ダイヤ改正後)

表-2 乗客 Q の乗降列車の推定結果 (ダイヤ改正前 39 日間 36 トリップ)

| 列車選択の 組み合わせ | S 駅出発 時刻(分) | T 駅到着 時刻(分) | 観測回数 (回) | S 駅での待ち時間(分) | | T 駅での出場までのロス時間(分) | | 同時刻の乗換 パターン数(個) | 乗換回数 (回) |
|----------------|----------------|----------------|-------------|--------------|-----|-------------------|-----|--------------------|-------------|
| | | | | 最小値 | 平均 | 最小値 | 平均 | | |
| a | 59 | 82 | 2 | 7.2 | 8.7 | 2.5 | 3.0 | 1 | 1 |
| b | 59 | 86 | 1 | 0.2 | 0.2 | 2.0 | 2.0 | 1 | 1 |
| c | 62 | 86 | 1 | 2.0 | 2.0 | 1.0 | 1.0 | 1 | 1 |
| d | 73 | 96 | 27 | 1.2 | 4.2 | 1.5 | 2.2 | 1 | 1 |
| e | 73 | 100 | 5 | 2.2 | 4.2 | 1.0 | 1.0 | 1 | 1 |

表-3 乗客 Q の乗降列車の推定結果 (ダイヤ改正後 42 日間 35 トリップ)

| 列車選択の 組み合わせ | S 駅出発 時刻(分) | T 駅到着 時刻(分) | 観測回数 (回) | S 駅での待ち時間(分) | | T 駅での出場までのロス時間(分) | | 同時刻の乗換 パターン数(個) | 乗換回数 (回) |
|----------------|----------------|----------------|-------------|--------------|-----|-------------------|-----|--------------------|-------------|
| | | | | 最小値 | 平均 | 最小値 | 平均 | | |
| f | 47 | 70 | 1 | 0.8 | 0.8 | 1.5 | 1.5 | 1 | 1 |
| g | 64 | 98 | 1 | 0.2 | 0.2 | 1.5 | 1.5 | 3 | 2 |
| h | 75 | 98 | 20 | 0.8 | 3.4 | 0.5 | 0.6 | 1 | 1 |
| i | 78 | 94 | 13 | 1.2 | 3.9 | 1.3 | 2.3 | 1 | 0 |

(3) 特定の乗客に着目した分析

本節では特定のカードIDを持つICカード利用者のトリップに着目する。高木らの研究⁶⁾で提案された図を用いた方法と、本研究で開発した方法による乗車列車の組み合わせの推定結果を比較する。

図-9及び図-10は、ある一人のICカード利用者Qが期間中の通勤時間帯にS駅で入場しT駅で出場したすべてのトリップを図示したものである。図-9はダイヤ改正前の39日間のトリップであり、図-10はダイヤ改正後42日間のトリップである。これらの図の縦軸は、S駅での入場時刻を示しており、横軸はT駅での出場時刻を示している。なお、縦軸、横軸の数値は、ある基準時刻からの相対時刻である。図中の点は、点の位置に対応する時刻にQがICカードを利用したことを示しており、点の濃さはその時刻に入出場した回数を示している。図中の横線はQが入場した駅での列車の出発時刻を示しており、縦線は出場した駅での列車の到着時刻を示している。Qが乗車可能な列車を、優等列車から順に列車カテゴリI, J, K, Lで区別する。なお、列車カテゴリI, Jは、それぞれ停車駅の異なる特急列車であり、Kは急行相当の列車、Lは各駅停車である。ダイヤ改正前にS駅ではカテゴリK, Lが停車し、ダイヤ改正後はカテゴリI, Lが停車するようになった。T駅にはI~Lのすべてのカテゴリの列車が停車する。

まず、高木らの方法⁶⁾でこれらの図からQが乗降した列車の判別を試みる。図-9に着目すると、利用者Qは多くのトリップで時刻70分前後にS駅に入場し、時刻100分前後にT駅を出場している。このことから、多くのトリップでS駅から乗車可能な列車は、73分に出発するカテゴリKの列車か、76分に出発するカテゴリLの列車であるとみなせる。また、T駅での出場時刻より96分にT駅に到着するカテゴリIの列車か、100分に到着するカテゴリJの列車に乘車していると考えられる。したがって、利用者Qが駅入場後直ぐに出発する列車に乘車していると仮定すると、QはカテゴリKの列車に乘車し途中の駅でカテゴリI又はカテゴリJに乗り換えてT駅に到着していると推測できる。ダイヤ改正後の図-10でも同様に考えると、

Qは75分に出発するカテゴリIの列車に乘車してT駅に到着するか、カテゴリLの列車に乘車して途中の駅でカテゴリJの列車に乗り換えてT駅に到着していると推測できる。このように、高木らの方法では、可能な乗降列車の組み合わせを視覚的に判別・推測することはできないもの、一定の基準を設けて乗降列車の組み合わせを推定することはできない。

表-2, 表-3は Q が乗車した列車を本研究の方法で推定した結果である。乗降駅での乗降列車の組み合わせはダイヤ改正前には5通り、ダイヤ改正後には3通り抽出された。なお、表中の「同時刻の乗換パターン数」は、乗降駅での列車は同じで乗り換える駅や途中で乗車した列車が異なるものの数を示している。表-2から、ダイヤ改正前は列車組み合わせ d(S 駅を 73 分に出発する列車Kに乘車し、途中で列車Iに乗り換えてT 駅に至るパターン) が最も多く選ばれていたと推定される。表-3から、ダイヤ改正後は、組み合わせ g(列車Lから列車Jに乗り換え)と組み合わせ h(乗り換えなしで列車Iで移動)が利用されていると推定された。

表-3の組み合わせ g の「T での出場までのロス時間」に着目すると、他の組み合わせに比べ最小値、平均値が小さくなっていることが確認される。これは利用者Qが実際には組み合わせ h を利用したにもかかわらず、推定を誤った可能性もあるといえる。IC データに出場した改札コーナを記録しておくとともに、改札からホームまでの最短時間を乗換ネットワークに組み込むことでこのような誤推定を軽減できると考えられる。

図-11及び図-12は、表-2, 表-3の結果を図示したものである。これらの図は、それぞれ上述の図-9, 図-10と整合している。図-9, 図-10では入出場時刻の組で打たれた点の分布が、図-11では5つの点(aからe)に、図-12では3つの点(fからh)にそれぞれ集約できたことになる。

5. おわりに

本研究では、ICデータに記録された利用者の乗車した列車を推定する方法を開発した。3章では、高木らによ

る先行研究⁶⁾でのICデータ分析の結果に基づいてICカード利用者の行動を仮定し、列車ダイヤデータから構築した乗換ネットワークを用いて、仮定を満たす乗車可能な列車選択の組み合わせを列挙する方法を構築した。4章では、鉄道会社A社のICデータを用いてトリップ毎にICカード利用者が乗車した列車を推定した。計算時間の観点からみると、数年程度の長期間に観測されたICデータであれば数時間から数日程度で推定が可能であることが示された。4章3節では、特定のICカード利用者に着目した分析を行った。この分析により、本研究で提案した乗車列車の推定結果と、従来の図を用いた視覚的推測との整合性を確認するとともに、ダイヤ改正前後の列車利用パターンの変化を定量的に把握できることがわかった。

推定結果の検証には、2つのアプローチが考えられる。ひとつは、個別の利用者の行動が正しく再現されているかを検証するものであり、他のひとつは列車毎に集計して得られた乗車人数を用いて集計的に検証するものである。本研究では、乗車可能な列車選択の組み合わせ数や乗換回数から、著しく非現実的な列車選択は推定されていないことを確認しており、個別利用者の行動再現については一定の検証を行っている。今後、さらに、ICカード利用者に対するプローブパーソン調査を実施すれば、その観測値と推定値を比較することで、個別の行動に関する推定精度をより詳しく検証することが可能となる。集計量による検証には、推定結果から列車毎の輸送人数を算出し、列車毎の応荷重データによる乗車人数の推定値と比較する方法が考えられる。ただし、本研究のように鉄道の全利用者がICカードを利用していない場合には、ICデータによる推定値に対して拡大処理を行って全利用者の乗車列車を推定した上で、応荷重データ等による推定値と比較する必要がある。拡大処理の方法によっては

大きな誤差を伴うことも想定されるため、この検証方法の構築は今後の課題としたい。

今後、長期間に渡って観測されたICデータを用いることで、ICカード利用者ごとの列車選択の特性を把握し、乗客の列車選択に対する選好を明らかにすることも可能になるだろう。このことは、鉄道のサービスレベルの評価につながることを期待できる。また、利用者の日々の列車選択行動を分析することで、ダイヤ改正や沿線でのイベントの影響分析などが可能になるだろう。

参考文献

- 1) 家田仁, 赤松隆, 高木淳, 畠中秀人: 利用者均衡配分による通勤列車運行計画の利用者便益評価, 土木計画学研究・論文集, No.6, pp.177-184, 1988.
- 2) 田口東: 首都圏電車ネットワークに対する時間依存通勤交通配分モデル, 日本オペレーションズ・リサーチ学会和文論文誌, Vol.48, pp. 85-108, 2005.
- 3) 森田泰智, 太田雅文, 窪田崇斗, 家田仁: 車両荷重データを用いた都市鉄道の時間帯別・車両交雑率の分析 - 東急田園都市線を対象として -, 土木計画学研究・講演集, Vol.36, 2007.
- 4) 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.
- 5) 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, 2007.
- 6) 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, 2007.

ICカードシステムによる改札通過データを活用した乗車列車推定方法の開発*

日下部貴彦**・高木勇弥***・井料隆雅****・朝倉康夫****

本研究では、交通系ICカードの乗車履歴データ(ICデータ)を用いて、ICカードを利用した鉄道利用者の乗車列車を推定する方法を構築した。本研究の手法は、列車ダイヤデータから構築した有向グラフを用いて、ICデータに記録された鉄道利用毎に選択したと推定される列車を列挙するものである。鉄道会社による実際のICデータを用いて推定を行い、推定にかかる時間及び推定精度に関する考察を行った。その結果、数年程度の長期間に観測されたICデータについて推定可能であることが確かめられた。

Estimation of Passenger's Train Choice in Railway Network with Smart Card Ticket System *

By Takahiko KUSAKABE**・Yuya TAKAGI***・Takamasa IRYO****・Yasuo ASAKURA****

The aim of this paper is to develop a methodology to estimate trains boarded by each passenger who is recorded on the smart card system. This paper shows the methodology for the estimation and algorithm for its computation. In order to validate computation time and accuracy of the estimation, the empirical analysis is carried out with the smart card data provided by a railway company in Japan. The results show that the proposed method could be used for the estimation with the smart card data observed over a long time period.
