

**科学研究費助成事業 研究成果報告書**

平成 29 年 4 月 24 日現在

機関番号：14301

研究種目：若手研究(B)

研究期間：2014～2016

課題番号：26820211

研究課題名(和文)交通系ICカードデータを活用した公共交通のOD需要推計モデルの構築と精度検証

研究課題名(英文) Estimation of passengers alighting at a bus stop based on smart card data using bayesian networks and neural networks as machine learning.

研究代表者

中村 俊之(Nakamura, Toshiyuki)

京都大学・工学研究科・助教

研究者番号：10419062

交付決定額(研究期間全体)：(直接経費) 3,000,000円

研究成果の概要(和文)：交通系ICカードは料金支払いのシステムとして、バス・鉄道事業者に広く導入され、利用者の利用時間や乗降実態、利用頻度や料金等、人々の行動把握なデータが収集されている。

本研究課題では、こうした交通系ICカードを活用し、ベイジアンネットワーク、ニューラルネットワークを援用した公共交通のOD需要推計モデルの開発並びに精度検証を目的とするものである。両手法ともに、乗車人数規模別に行うことで一定の精度での予測が担保されることが示唆されたが、交通計画の策定時に本研究成果の利用する際には、影響要因の把握が可能なBNNを用いることが有益であることが、交通事業者へのヒアリングを通じて明らかになった。

研究成果の概要(英文)：Smart card systems are widely introduced as payment (electric fare collection) systems to bus and railway operators. Electronic fare collection data such as smart card records offer new opportunities though. Smart cards have now been available in many cities since several years so that records of public transport usage frequency records over prolonged periods of time can be obtained.

The Aim of this research is to utilize such transportation smart cards, to develop a model for estimating OD demand of public transportation using Bayesian network, neural network. Both methods suggested that prediction with a certain precision is guaranteed by doing by number of passengers, but it is possible to grasp the influential factors when using this research result at the time of formulating the traffic plan. It was clarified through interviews with transportation companies that it is beneficial to use BNN.

研究分野：交通計画

キーワード：スマートカード 公共交通 バス ニューラルネットワーク ベイジアンネットワーク 機会学習

### 1. 研究開始当初の背景

交通事業者ならびに国、地方自治体の計画策定者が公共交通の利用促進を図る際には、交通需要を把握することが必要不可欠である。我が国では、交通需要の把握に四段階推定法が用いられてきた。四段階推定法を適用することで、公共交通利用者がどのような起終点（以下、OD）を有しているのか、またどのような移動目的で公共交通を利用しているのかを把握することが可能となる。その結果をもとに、交通事業者、計画策定者は、公共交通のサービスレベルを決めることができる。しかしながら、実際、多くの交通事業者や計画策定者は、利用者の OD や移動目的を十分に把握するには至っていない。

近年普及が進んでいる交通系 IC カードにより、365 日 24 時間利用者のデータ収集がされている。具体的には鉄道、バス利用者が改札、乗降の際に IC カードを専用端末にタッチすることで、自動的にデータが収集される。このデータを活用し、公共利用者の OD を把握することが期待されている。しかしながら、料金均一制のバスでは乗車時、もしくは降車時のみ、専用端末にタッチすることから、O（起点）、もしくは D（終点）に限りデータ収集が行われ、依然として利用者の OD は把握することができない。加えて、交通系 IC カードから収集されるデータでは、移動目的が収集されておらず、交通計画の策定、交通施策の実施にあたっては不十分なデータである。

### 2. 研究の目的

上述のような背景のもと、交通系 IC カードデータを利用し、特に OD 需要の把握が困難なバス交通に着目し、OD 需要推計の手法論を構築することが本研究課題における目的である。

これまでバス交通に関して、その利用特性の分析などはこれまで数多くなされてきており、利用の実態は掴めてきているものの、1 つ 1 つの停留所のデータを把握することの難しさなどから、広域エリアの停留所に着目した OD 推計を行っている研究はあまり見られない。また、その OD 推計を行っている研究に関しても、適合度が低かったり、推計に使用するデータ制約などから特定の地域でしか適用できなかったりと汎用性の高いモデルは存在していない。さらに、本研究で用いる推計手法である、ニューラルネットワーク（以下、NN）とベイジアンネットワーク（以下、BN）といった機械学習による推計は、交通分野に限定すると高速道路を対象としたものが多く、バスの停留所別乗降車人数や OD の推計に使われたことはない。

以上から本研究課題では、利用特性の分析を行った既存研究を参考にして、停留所ごとに推計に有効であると考えられる変数を抽出し、その変数をもとに NN と BN を利用して、OD 需要推計のうち、停留所別降車人数の推計を行う。

### 3. 研究の方法

本研究課題遂行にむけて、しずてつグループが導入している交通系 IC カードにより収集されるデータを利用する。

#### （1）研究の方法と手順

停留所別降車人数の推計には NN や BN といった近年ビッグデータ解析に注目されている機械学習を適用する。具体的には図-1 に示す手順にて実施する。

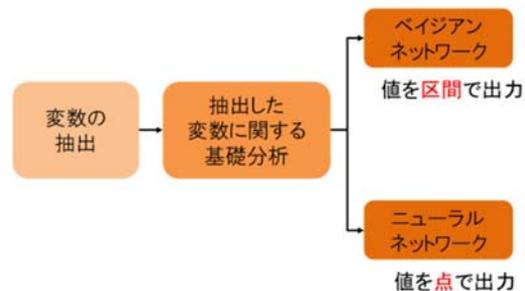


図1 停留所別降車人数推計の手順

#### ①推計に用いる変数の抽出

推計に有効であると考えられる変数を考え、抽出する。つまり、降車人数に影響を与えると考えられる変数をリストアップし、取得したデータより、推計に用いるデータセットを作成する。

#### ②抽出した変数に関する基礎分析

抽出した変数に関する基礎分析を行い、従属変数と各独立変数との関係性を確認する。ここでは、主にクロス集計を行うことでその特性を把握する。

#### ③NN と BN を援用した停留所別降車人数の推計

これまでに作成した各変数を用いて、停留所別降車人数の推計を行う。推計手法として、機械学習である NN と BN を用いる。

#### （2）研究対象地域

なお、本研究課題は、静岡県静岡市を分析対象地域とする。この対象地域内には、新静岡駅の本ターミナルなど、複数のバス路線が存在し利用者数が多い停留所から、バス路線は単線で運行本数が少なく、利用者数も少ない停留所まで他にも様々な特徴をもつ停留所が存在している。また、対象とする停留所の数は 974 個が存在している。

### 4. 研究成果

3 章にて示した手順に基づき、最後に成果をまとめる。

#### （1）推計に用いる変数の抽出

独立変数としては IC カードデータより集計される動的な変数と停留所が個別に有する

静的な変数を定義している。

動的な変数は、利用するデータ期間やエリアにより、その値が異なる変数である。ICカードデータから抽出される動的な変数として、乗車人数、低頻度利用者割合、高頻度利用者割合、朝時間帯利用割合、夕時間帯利用割合、平日利用割合、片道利用者割合の7変数を抽出した。

静的な変数は、交通事業者が路線や時刻表等のサービスレベルの更新作業を行わなければ不変の変数であり、停留所に停車する路線数やバス停留所の位置に関する情報を含むものである。静的な変数として、路線重複数、一方向停車ダミー、始バス7時ダミー、終バス20時ダミー、最近隣停留所間距離、そして位置ダミーとして、駅、商業施設、学校、病院、住宅地、レジャー施設、役所の12変数を抽出した。

### (2) 抽出した変数に関する基礎分析

抽出した変数に対して、実施した基礎分析結果を簡単に示す。

図2は乗降比カテゴリ別の乗車人数カテゴリ構成比であり、比率をとった場合、乗車人数が多い停留所ほど、計算上規模の影響を受け、乗降比は1に近い値をとる傾向を示している。

図3は乗降比カテゴリ別の高頻度利用者割合カテゴリ構成比であり、高頻度利用者割合が小さい停留所ほど乗降比は1から離れた値をとる傾向を示している。

一部のクロス集計の結果のみを示しているが、機械学習における推計にあたって、利用する全ての独立変数と従属変数間でクロス集計は実施し、独立変数ごとに従属変数に対して、特定の影響を確認した。

### (3) NNとBNを援用した停留所別降車人数の推計

#### ①推計時の前提条件

NNでは降車人数を、BNでは乗降比を従属変数として設定する。この際、乗降比は15の

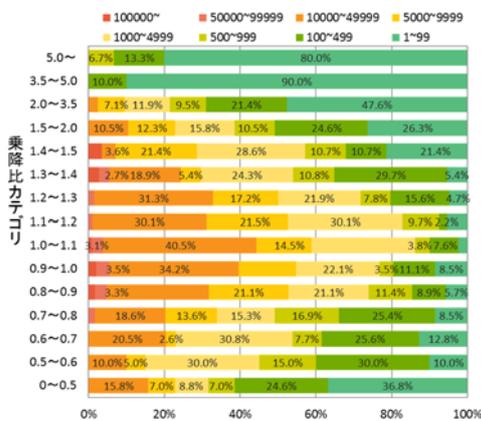


図2 乗降比カテゴリ別の乗車人数構成比割合構成比

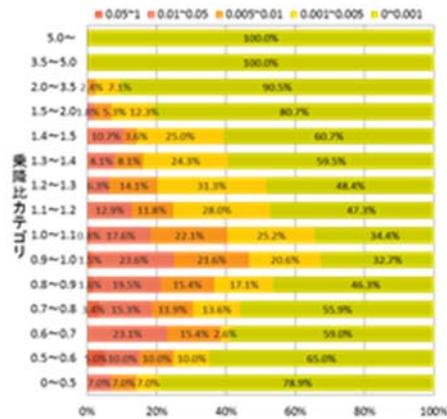


図3 乗降比カテゴリ別の高頻度利用者

カテゴリに分け、各カテゴリの中間値を乗車人数に掛けることで降車人数の推計を行う。

本研究における推計値の比較対象は、全利用が同一停留所間での往復利用を仮定し、降車人数を推計したモデル（以下乗降車数一致モデル）を設定する。全利用者の乗降車が往復利用を仮定したモデルでもデータが大量かつ長期間蓄積することで、利用者の特徴の把握がある程度は可能である。機械学習を用いることで、その状況よりも精度良く推計を行うことができるのかを見極める必要がある。

また、乗車人数の規模によって停留所の特徴が異なると考え、本研究では停留所を乗車人数規模によって4つのカテゴリに分類し、推計を行った。

#### ②推計手法の適用結果

4つの乗車人数規模に対してNNとBNを適用した結果を表1に示す。この結果から、乗車人数規模1~99人での平均誤差率と乗車人数規模10,000~59,999人のニューラルネットワークにおける平均誤差率を除き、乗降車数一致モデルと比較して推計精度は改善されていることが確認できる。

乗車人数規模1~99人での平均誤差率のみ乗降車数一致モデルよりも悪い結果となった理由は、外れ値に関する学習の影響が考えられる。乗車人数95人に対して降車人数12,484人という停留所が存在しており、学習データとしてこのデータを採用することで、推計結果が大きくなる方向に影響を及ぼしたと考えられる。また、BNよりも、NNによる平均誤差率の方が大きく外れた結果となった原因は、乗降比の上限値が5に設定してあることが影響していると考えられる。

次に、推計値と実測値の散布図（誌面の都合上、10,000~59,999人）での結果を図4から図6に示し、考察を行う。

この散布図より、乗降車数一致モデル、NN、BNモデルでの推計値と比べて、ベイジアンネットワークモデルでの推計値はより45度線に近いものが多い結果となる一方で、45度線から大きく外れた推計値も多くことが確認できる。

表 1 推計結果のまとめ

乗車人数規模	比較対象	平均差	平均誤差率	RMSE
1~99人	乗降車数一致	138.9	78.20%	1121.1
	NN	51.3	239.10%	107.9
	BN	127.3	80.30%	1085.8
100~999人	乗降車数一致	190	58.80%	456.5
	NN	157.5	55.00%	339.7
	BN	171.3	37.70%	363.3
1,000~9,999人	乗降車数一致	977.6	28.60%	1810.7
	NN	913.4	27.10%	1467.6
	BN	713.1	17.90%	1633.3
10,000~59,999人	乗降車数一致	3230.3	18.50%	4911.6
	NN	3173.3	18.60%	4623.4
	BN	2627.7	15.20%	4798.9

推計降車人数(人/年)

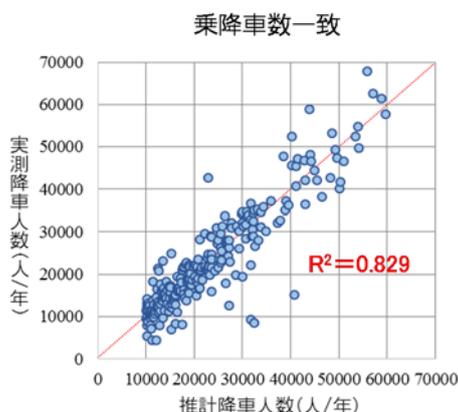


図 4 乗降車数一致時の散布図

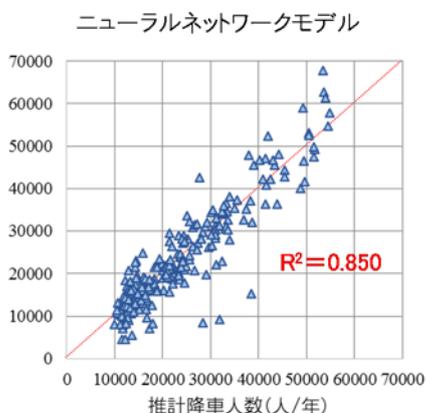


図 5 NN モデルの散布図

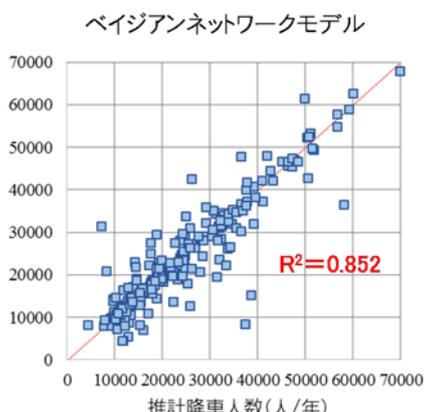


図 6 BN モデルの散布図

③推計手法の適用結果のまとめ

推計手法別の適用結果の特徴を整理する。乗降車数一致モデルは、停留別にデータの集計のみを行えば良く、データ処理も簡単であり、簡潔で分かりやすいが今回の機械学習の結果と比較して、精度の面で劣ることがわかった。

今回適用した機械学習を用いた手法について、NN モデルは、低精度な推計値が少なく、RMSE では良い結果が得られたものの、高精度な推計値も少ない。BN モデルは、高精度な推計値が多く、平均差と平均誤差率では良い結果が得られたものの、低精度な推計値となることも多い結果となった。このように、各モデルには一長一短が存在することから、推計時に用意できる変数や乗車人数規模に応じて、現状ではモデルを使い分けることが求められる。

一方で BN 適用時には利用変数間の関係が把握できる結果として把握できるのに対して、NN 適用時には変数間の関係は分析者でも把握ができない。すなわち得られた結果を用いたその影響を説明するという場合には BN を適用せざるを得ないこと、すなわち、交通計画の策定時に本研究成果の利用する際には、影響要因の把握が可能な BNN を用いることが有益であることが交通事業者へのヒアリングを通じて明らかになった。

④研究を通じた課題

最後に、本研究課題では実際に料金均一制のバスでは乗車時、もしくは降車時で運用されている地域に対しても、本手法を適用することを想定していたが、データ収集の都合上、困難であった。そのため、今回モデル構築時とモデル検証時のサンプルとしてともに同一都市での停留所データを使用しているため、精度が高めにしていることが懸念される。今後料金均一制のバスが運行しており、乗車時もしくは降車時にしかデータ収集されていない東京都バスや京都市バスのエリア等の他地域へ本手法を適用することを考慮すると、諸地域の様々な特徴をもったサンプルを学習させることが望まれる。

5. 主な発表論文等

〔学会発表〕(計 5 件)

- 1) 渡邊悠介・中村俊之・宇野伸宏・Jan-Dirk Schmoecker:バス IC カードデータを用いたベイジアンネットワークによる停留所別降車人数推計に関する研究, 土木学会年次学術講演会講演概要集, CD-ROM, Vol.70, 2015.
- 2) 渡邊悠介・中村俊之・宇野伸宏・Jan-Dirk Schmoecker・山崎浩気: IC カードデータを用いた機械学習による停留所別降車人数推計手法の研究, 土木計画学研究・講演集, CD-ROM, Vol.52, 2015.

- 3) Watanabe, Y., Nakamura, T., Uno, N., Schmoecker, J.D.: Estimation of Passengers Alighting at a Bus Stop Based on Smart Card Data Using Bayesian Networks, The 20th international conference of Hong Kong society for transportation studies, Hong Kong 2015.
- 4) Nakamura, T., Nakamura, N., Schmöecker, J.D., Uno, N., Iwamoto, T.: Urban Public Transport Mileage Cards: Analysis of their potential with smart card data and an SP survey, Transportation Research Board 95 th Annual Meeting, Washington, D.C. 2016.
- 5) Nakamura, T., Schmöcker, J.D., Uno, N., Akihiko Shimizu, Iwamoto, T.: Explaining Passengers' Bus Stop Choice Using Smart Card Data, 2nd International Workshop on Automated Data Collection Systems: Improving Urban Public Transport Planning and Operations

[図書] (計 1 件)

- 1) Kurauchi, F.,Schmöcker, J.D.: CRC Press(Taylor & Francis Group), Public Transport Planning with Smart Card Data, ISBN-13: 978-1498726580 (2016).

[その他]

ホームページ：所属機関

<http://www.um.t.kyoto-u.ac.jp/ja>

ホームページ：所属研究室

<http://trans.kuciv.kyoto-u.ac.jp/its/index.html>

## 6. 研究組織

### (1)研究代表者

中村 俊之 (NAKAMURA, Toshiyuki)

京都大学・大学院工学研究科・助教

研究者番号：10419062