

データオリエンテッド交通研究

日下部 貴彦¹

¹正会員 東京工業大学助教 大学院理工学研究科 (〒152-8552 東京都目黒区大岡山2-12-1-M1-20)
E-mail:t.kusakabe@plan.cv.titech.ac.jp

目覚ましい情報通信技術の発展によって交通分野におけるデータに対する関わり方が変わりつつある。近年では、実際に、継続的・長期的に収集された大量の交通に関わるデータの蓄積がみられるようになってきている。一方、これまでの交通分野では、空間的に広い範囲のデータを時間的に高密度でかつ長期的に収集することは困難であり、変動に対する理論的なフレームは必ずしも十分でなかったと推察される。本稿は、これまでの交通分野での研究について、データ取得と蓄積という視点から整理し、新たな枠組みであるデータオリエンテッドな交通研究に焦点をあて、蓄積データを背景とした交通変動の分析に関する展望についてまとめるものである。

Key Words : *data oriented, passive data, data mining, data fusion, transport dynamics*

1. はじめに

交通分野の研究の難しさに系全体の観測の難しさがあげられる。交通行動は、人の何らかし意志や欲求に基づき、それに伴った活動の目的を達成するために人や物が時空間上を移動することによって発生する。すなわち、活動の発生及び、時空間上での移動を観測することが、交通という系の全体を捉えるためには不可欠である。したがって、長期間に及ぶ観測をシステム全体で行うことが系の全体像を捉えるために必要だと考えられるが、交通行動が行われる範囲は広大であり、日周期から年周期に至る変動を都市や国土の規模で観測するような包括的な観測を行うことは非常に難しい。このことから、観測データの使用目的に応じて、空間的要素を限定するか、時系列的要素を限定することが求められ、そのうえで、交通計画や、交通管理での実務や研究で利用される観測及びモデルの体系が構築されてきていると考えられる。

例えば、道路整備など空間的な計画を行う交通計画の分野では、時系列的な変動のデータ取得を限定する一方で、空間的なデータの取得に主眼がおかれ、様々な交通調査の手法が開発・体系化され適用されてきた。一方で、分単位の時系列での特性に応じた交通制御等を行うことが求められる交通工学の分野では、高速道路や幹線道路に観測領域を限定するなど、空間的なデータの密度や範囲を限定する一方で、交通管制システムのような時間的に詳細なデータ取得が可能なシステムが構築されてきた。しかし、以上に示したような時間・空間の各断面からの観測では、交通現象の完全な記述は困難であることが一

般に知られている。例えば、交通管制システムで取得可能な、交通量などの空間的断面データからは、発生・集中や経路交通量を特定することはできない。一方、時間的断面のデータであるパーソントリップ調査のような大規模な調査は、典型的だとされる一日を対象として行われることが一般的であり、変動はおろか、平均的な状態でさえ分からない実情がある。

北村¹⁾は、「観測された交通現象が均衡状態を表すというのは、各行動主体がそれを取り巻く環境、あるいはそれがもつ選好構造に変化が生じたとき、直ちに反応を示すというきわめて強い過程の上にのみ成り立つもの」であり、経時変化ではなく、状態を重視することは、「解析上の、あるいは観測上の便宜性を主たる理由とするのではないかという仮説」を示している。また、断面データから導かれるモデルを将来予測に適用することの問題は繰り返し指摘²⁾されてきた一方で、観測の困難さとモデルの複雑さがこのような研究の方向性とどまらせていることを述べている。つまり、これまでの交通研究は、コスト面での難しさ、技術面での難しさから、断面データという制約のもとに設定された課題に対してアプローチするものであったことが示唆される。

空間的、時間的に連続なデータが究極の顕示的な交通データの集合であることを前提として考えると、現況を示すことですら多大なコストが必要である点が交通学研究を難しくしている一つの原因といえる。時間・空間双方の変動がある系に対して不完全な観測を前提とせざるを得ないなかで、如何に系を分割し、よりよい計画や運用方法を構築していくかが交通学研究に課された大きな

技術的課題であったと考えられる。したがって、観測系の発展により、変動を広範囲にかつ連続的に捉えることができれば、新しい交通学研究の局面が訪れるかも知れない。

近年の情報通信技術の発展はめまぐるしい。数多くの観測技術の発展が見られるだけでなく、収集されるデータの蓄積がとてつもないスピードで進んでいる。世の中に蓄えることが可能なデータ量は、1986年に2.6 exabytesであったが、2007年には29 exabytes (1 exabyte = 10^{18} bytes)にまで増加したという推計³⁾がある。さらに、2013年では4.4 zettabytes (1 zettabyte = 10^{21} bytes)に達しており、2020年には、44 zettabytesに達するという民間調査会社の予測⁴⁾もある。このような、とてつもない情報量の増大 (情報爆発⁵⁾とも) は、多くの分野で研究者がとりあかうデータ量を飛躍的に向上させている。今後5年で10倍に増えるデータの中には、交通系ICカードの料金収受データ、商用車等の管理やナビゲーションを目的としたスマートフォンや自動車から得られる位置情報など交通分野にも関連深いものが含まれるとされている。これらのデータは、料金収受等の本来の目的から副次的かつ自動的に得られるものであり、データ取得のコストはほとんどかからない。また、交通利用者がサービスを利用している限り取得されるという点で継続的かつ長期的なデータ収集が期待される。

このようなめまぐるしい発展を間近にして、羽藤⁶⁾は、交通分野の研究でのデータオリエンテッドアプローチの必要性を述べている。これは、観測系の制約が大きい交通分野の研究において、これまでにない時間的・空間的に詳細なデータセットが大量に入手可能になる可能性とそのようなこれまでにない大量のデータからの知識を得るための方法を整える必要性を示すものである。羽藤⁶⁾は、このような方法を整えることにより、これまでのデータセットではとらえられなかった解像度や時間的スケールでの交通現象への理解を進め、このことを前提として、モデルフレームの見直しにつながることを期待している。北村⁷⁾が述べているように交通分野の多くの研究は断面データを前提として発展を遂げてきていることから、交通現象の経時的変化を前提としたフレームワークは未成熟であると考えられる。とりわけ、これまでの交通分野の研究では、観測に関する制約が大きいことから、いかにその制約を回避するための創意工夫を行うかという指向で進められてきたものが多く、大量のデータがあることを前提とした研究アプローチ、とりわけ、調査以外の副次的に得られたデータから知識獲得を行うことに対しての研究については、必ずしも十分ではない。したがって、このような蓄積データに対する研究を発展させることは、交通現象の変動を理解し、交通分析を高度化させることに寄与すると期待されている。

以上に示した情報通信技術の発展は、交通分野におけるデータに対する関わり方を変える可能性があり、断面データからの脱却とそれにとまなう交通研究のパラダイムシフトの可能性もはらんでいる。このようなことを実現するためには、これまでの交通分野で研究されてきたデータとは異なる性質のデータと向き合うための準備を要するだろう。そこで、本稿は、これまでの交通分野での研究について、データ取得と蓄積という視点から整理し、交通分野でのデータ研究の枠組みとデータオリエンテッドアプローチによる交通変動の分析に向けた展望についてまとめるものである。

2. 交通研究とデータ

前章では、交通データの収集には、制約が多く、目的に応じた断面データの収集に注力されてきたことについて述べた。交通分野のデータ収集は、1990年代の情報通信技術の発展を節目に、高度化、多様化しつつある。また、2000年代からはデータの蓄積も行われる様になってきている。そこで、本章では、交通分野におけるデータ収集の主要な方法について、「交通観測の体系化の時代」、「情報化の時代」、「データ蓄積の時代」に分けてレビューを行うことを通して、これまでのデータ収集の特徴をまとめる。

(1) 交通観測の体系化の時代

本節では、空間的断面での観測を行う交通流観測と時間的断面で行われる交通調査のそれぞれについてレビューを行い、1990年以前のそれぞれのデータ収集方法の発展について述べる。

交通流観測で、定量分析を行うためのデータ取得が始まったのは、1930年代である。交通流の観測は、定点に設置した機器を用いて、時系列的に連続なデータ取得を行い、交通流モデルの同定、交通容量の測定に用いられている。Greenshieldsら⁷⁾は、16mmフィルムの映画用カメラを用いて、車間距離と速度を観測することで、交通流の基本的な性質であるQK関係(Fundamental Diagram)の推定をおこなっている。1950年代、Olcott⁸⁾やGreenberg⁹⁾は、テレグラフから操作して、人がキーを押すことでペンレコーダーを用いて記録紙に印字する方式で交通流の観測をおこなっている。1960年代に入ると、インダクションループ式車両検知器 (感知器、トラカンとも) が登場¹⁰⁾し、アメリカの各地の交通管制システムに導入¹¹⁾され、自動的な観測が行われるようになる。日本でも、大阪万博の時期に合わせて、1969年に阪神高速道路で車両検知器を備えた交通管制システムが導入¹²⁾された。1970年代中盤からは、画像処理を用いた観測方法がアメリカ

Federal Highway Administration (FHWA)などで研究され始める¹⁰⁾。以上のように、都市高速道路等の空間的に限定された範囲では、1960年代から、継続的かつ長期間のデータ取得が可能となっていた。しかし、これらのデータを蓄積することは容易ではなかったと推察される。継続的な観測を活かした研究として、飯田・高山¹³⁾は、車両検知器からえられた1年間の日交通量の変動を分析しているがこうした研究は少ない。

都市規模の交通の空間的な分布を観測することに主眼においた、交通調査では、インタビューや紙面上での質問調査が主流となっている。交通調査が体系化されたのは、1940年代にアメリカのBureau of Public Roadsがガイドラインを発行したのが初めであり¹⁴⁾、Kansas City, Lincoln, Little Rock, Memphis New Orleans, Oklahoma City, Savannah, Tulsaで実際に調査が実施されている¹⁵⁾。日本においても、1950年代には、米谷¹⁶⁾が、路側でのOD調査の問題点を指摘し、標本調査の提案を行っている。1967年には、広島都市圏でパーソントリップ調査が実施され、1968年に首都圏でも実施されている¹⁷⁾。

1970年代には、ダイアリ調査が実施され、複数日の変動に着目した調査が実施され始めている¹⁸⁾¹⁹⁾。さらに、行動の変動を捉えることを意図した調査として、同一のサンプルに対する調査を一定期間及び一定期間毎に行うパネル調査も実施されている²⁰⁾²¹⁾。日本では、大規模な世帯交通調査でのパネル調査は実施はされていないが、交通手段選択行動などを対象としてパネル調査が研究されている²²⁾²³⁾。パネル調査では、消耗バイアス、パネル効果、パネル拘束などの問題があり、こうしたバイアスや統計的性質が研究²⁾がされてきた。またこれらの調査は、インタビューや紙面での質問調査であることから、回答者・調査実施者の負担が大きく、調査票のコーディングなどの費用も大きいことから、調査の範囲や期間、解像度の制約となっていた。また、より詳細な人の流動を観測した研究に、佐佐木・松井²⁴⁾が動物園内の人の移動を観測したものがある。この研究のデータ収集では、調査員が観測対象者を追跡し、記録を行っている。

以上に示したように、交通分野で使用されるデータは、交通管制、交通計画などの目的に応じて、時間断面及び空間断面のそれぞれについてのデータの収集方法が構築され体系化されてきた。一方で、都市規模での空間上の動的な交通データを長期間収集することは、費用や労力の大きさから容易ではなかったといえる。

(2) 情報化の時代

1980年代半ばごろから、データのデジタル化、観測技術・計算技術の汎用化が新たな観測シーズを産みだし、実装がなされてきた。新たに登場した技術的なシーズにより、これまでデータ収集が難しかった詳細な交通行動

や流動データを継続的に収集可能とし、調査・観測手法の体系的改善をはかることを志向する研究が多く成されている。特に、コンピュータを用いた行動実験・調査などの実施、GPS (Global Positioning System) などによる観測技術の登場、画像処理技術の汎用化が大きく目立つ。

1980年代ごろから、仮想的な状況に対する行動調査の方法が数多く発展している。例えば、交通流シミュレータを活用した室内実験による出発時刻や経路の選択行動の観測がなされている²⁵⁾。また、運転挙動に関する室内実験の方法としてドライビングシミュレータの開発も行われた²⁶⁾。経路選択行動等を扱うStated Preference 調査でも、コンピュータを用いた質問法が主流となった²⁷⁾。また、インタビュー調査では、Computer-Assisted Telephone Interviewing (CATI)やComputer-Assisted Self-interviewing (CASI)も利用されている¹⁴⁾。2000年代には、これらのコンピュータ上での質問をWebを用いて行うことも可能となった²⁸⁾。また、パーソントリップ調査等の大規模な調査でも一般的にWebからの入力が入力されるようになり、回答者の負担やコーディングの手間が軽減されている²⁹⁾。

実現象の調査では、GPSやPHS等を活用した観測システムが開発され、車両の軌跡³⁰⁾³¹⁾や人の行動の記録³²⁾³³⁾³⁴⁾³⁵⁾に用いられるようになり、これまで困難だった経路観測や比較的長期間の追跡調査が可能となった。

画像処理技術の高度化、汎用化も観測技術の発展につながっている。観測範囲は、カメラの画角や解像度によって限られるものの、撮影範囲内の車両等の軌跡を詳細に追跡可能であったり、あとから視認して状況を確認できるなどのメリットがあり、調査で用いることが多い。車両を対象としたものでは、交通管制システムでのリアルタイムな交通監視³⁶⁾や、交通流調査での車両挙動観測³⁷⁾に応用されている。また、交通流観測ツールとして、より容易に適用できるように汎用的なハードウェアを用いたシステム化³⁸⁾³⁹⁾もなされている。歩行者の観測についても、多くの研究⁴⁰⁾⁴¹⁾がなされるようになった。以上に示した方法以外にも、車両挙動観測での加速度等の各種のセンサーの活用⁴²⁾や、車両の上空からの撮影⁴³⁾による観測、歩行者観測でのレーザースキャナの活用⁴⁴⁾などもみられる。また、スマートフォン利用者やナビゲーションシステムが発するBluetooth やWi-Fiの信号を利用して、定点を通過する道路上の利用者を識別することで旅行時間等の観測を行う方法もみられる⁴⁵⁾。

以上に示したように、情報化の時代には、コンピュータを用いた行動実験・調査などの実施、GPSなどによる観測技術の登場、画像処理技術の汎用化が大きく目立つ。これらの技術を用いて得られる情報を交通分析のためのデータへ変換する、あるいは、これらの情報技術を交通観測のなかでシステム化するための研究が多く成されている。

(3) データの蓄積の時代

2000年ごろから、交通分野でのデータ蓄積は交通管制データを中心に行われてきた⁴⁶⁾⁴⁷⁾。また、プローブパーソン調査による行動データをOLAP (Online Analytical Processing) 指向のデータベースに実装するための開発が検討され、蓄積データとしての活用も模索された⁴⁸⁾。このようなデータの蓄積により、より長期間の変動の分析が可能となりはじめている。さらに、データの蓄積は交通観測の分野のみならず、さまざまな業務・サービスに伴うデータが蓄積されている。このような業務・サービスに伴うデータには、交通に関連するものも多くあり、既往の交通分野での観測データに対して種類、量ともに上回ろうとしている。プローブカーデータは、道路上の車両から得られる情報から交通状況を観測するものであるが、近年では、運送業者やタクシー、バスなどが運行管理を目的としてGPSを登載することが多くなっており、これに伴うデータを活用しようという試みが実験的に実施されてきた⁴⁹⁾⁵⁰⁾⁵¹⁾。さらに、実際にデジタルタコグラフを登載した商用車等から得られるデータのデータベース化⁵²⁾が進んでいる。プローブカーデータ以外にも、道路交通では、自動車のナンバープレートを認識するAVI (Automatic Vehicle Identification) データ⁵³⁾、ETC (Electronic Toll Collection) データ⁵⁴⁾の活用が検討されている。鉄道に関するデータについても、交通系ICカードデータ⁵⁵⁾、自動改札機データ⁵⁶⁾、車両応加重データ⁵⁷⁾などが着目されている。また、交通事業・運輸業以外の直接的に交通には関わらない業種で収集されたデータも活用が見込まれる状況となっている。例えば、ナビゲーションサービスのデータ⁵⁸⁾、列車乗り換え情報の検索システムのログデータ⁵⁹⁾などの交通観測を直接的な目的としているわけではないデータが大量に蓄積され、交通分野の分析への活用が模索されている。

以上に示したような蓄積データを用いた研究の広がり背景には、ストレージ容量の拡大とデータウェアハウスの構築、情報サービスの多様化によってもたらされた受動的に収集されたデータの存在が大きい。検知器データをはじめとする交通管制データのように交通観測を目的としたデータだけでなく、ETCや交通系ICカードデータのように料金収受を目的として収集されたデータ、運行管理やナビゲーションを目的として得られたプローブデータなど、データ収集の本来の目的が交通分野の分析に比較的近いものから遠いものまでであることが分かる。いずれのデータでも、業務等の本来の目的がある限り、時間的に継続的に自動的に収集され続けることが期待できることが特徴である。

(4) データ研究の類型化

交通分野の観測の歴史を概観すると、1930年代は、計

測自体が困難であり、当時の最新の技術を駆使して定量分析に耐える観測を行っていた時代であるといえる。1940～60年代には、管制システムやパーソントリップ調査の登場があるように、体系的な観測が構築されてきたといえる。1970～80年代には、交通管制の分野では、超音波式車両検知器のような観測装置の開発等により実装のコストを下げるとともに、空間的な観測範囲、密度を上げる努力がなされてきている。交通調査では、パネル調査のように時系列的な観測方法及びその解析方法に関する研究が模索されている。1990年ごろから、情報通信技術を用いた調査・観測手法の開発がはじまり、2000年代にはデータの蓄積がはじまった。2000年代後半には、交通観測以外の目的で収集されたデータの蓄積が多くなり、研究されだしている状況にある。

これまでの交通分野で行われてきた観測を含む研究をデータ研究と見なしたとき、ニーズ先行型、シーズ先行型、データ先行型の3つ類型に分けることができる。ニーズ先行型では、1990年代までの研究にみられるように、交通管制や交通計画のニーズを充足するために観測方法が開発・体系化されてきた。シーズ先行型は、1990年以降にみられる情報技術のシーズに対応した観測体系の拡張に対応する。データ先行型には、2000年代後半からはじまった大量の蓄積データを活用する研究が対応する。

ニーズ先行型の研究では、観測目的やモデルの構築が先行し、モデルに必要な観測値を得ることが最優先される一方で、モデルの枠組みもまた観測の限界に制限されてきたものと推察され、断面データの解析に適したフレームワークへと落ち着いてきたものと推察される。

シーズ先行型の研究では、詳細なデータ取得とより長期間・連続的に観測可能な方法が多く開発されてきたが、データ取得には調査等の実施を前提としており、管制システム等実装されているものを除くと、年単位の長期間の変動を連続的に捉えるのに必要なデータを収集することは容易ではない。

データ先行型の研究では、データの取得が先行していることから、期間・継続的なデータが既に存在しているという前提で研究が展開される点でニーズ先行型、シーズ先行型の研究と異なる。このようなデータがあることで、これまでの断面データでは見えてこなかった、時間軸の変動も捉える可能性もあるといえる。一方で、データ先行型研究の初期の研究では、業務等で得られたデータから交通分野の分析で必要となる変数や指標を推定するものが多い。また、プローブカーの普及が少なかったころは、実際の業務で得られたデータではあるものの、実証実験として車載された機器で得られた比較的短期間のデータに対する研究が中心である。

3. データオリエンティッド研究

前章では、交通分野でのデータ収集の類型を示し、データ先行型の研究環境が整いつつあることを示した。このようなデータは、本来の業務が継続する限り、継続的かつ長期的に収集されるという点で利点がある。また、プローブカーデータのように空間的な広がりをもつデータや、ETCや交通系ICカードデータのように利用者のIDを識別することができるデータなどがあり、これまでの車両検知器などの観測データとは異なるデータ項目が収集されていることもある。また、分析者は、観測という非常に大きなコストがかかる段階を省いた上で、一定の分析が行えることにメリットがある。一方で、これらの観測データは、分析者の意図している目的でデータが収集されているとは限らず、必ずしも既往の分析のフレームワークで必要とされている項目が収集されているとは限らない。例えば、料金収受を目的としている交通系ICカードのデータでは、交通調査で収集されるデータのような、起終点や交通目的などの情報は含んでいない。つまり、交通系ICカードデータをこれまでの分析フレームワークでの交通行動データとみなしたときには、不完全なデータということになり、これまでと異なる分析フレームワークが必要である。

「観測」を分析の目的にしたがってデータを収集する行為であると定義した場合、データ先行型の研究では、データの「収集」と「分析」が必ずしも同一目的によって行われるものではなく、分析者の意図とは異なるところで、収集されたデータがあらかじめ存在しているということになる。したがって、分析者は、巨大なデータの中から「観測」する必要があるといえる。また、これまでの観測と分析が一体であったニーズ先行型及びシーズ先行型の研究では、「観測」を行わなければ、データも収集されないことから、分析者が観測の事前に分析対象となりうる事象に気づくことがなければ、データそのものが存在しないということになる。一方で、データ先行型研究では、蓄積されたデータが既に存在していることから、分析対象となりうる事象をデータから発見し、新たな分析対象を定めることも期待できる。つまり、データから現象を発見するというのも研究の一つのステップとなり得ることから、このような発見を大量のデータから効率的に行う手法も必要となる。

データ先行型の研究を前提に、第1章で述べたような交通現象の変動を理解し、経時的变化を取り込んだフレームワークを構築するための前段階のステップとして、以上に示したように、交通分野のデータとして不完全なデータの取り扱いを検討する研究や、データに含まれている変動についてデータオリエンティッドなアプローチで把握し、発見するとプロセスに対する研究が必要とな

る。さらに、発見された変動をモデルとして記述し、変動の検出方法を構築することで、データ全体の変動を俯瞰し、理解することができると考えられる。

著者の研究では、データに含まれる変動を把握するプロセスとして可視化技術を開発している⁽⁶⁾⁽¹⁾⁽²⁾。また、交通系ICカードデータに含まれる変動を理解するための方法として、乗車列車の推定方法⁽⁶⁾や交通目的の推定手法⁽⁶⁾を構築している。本章では、データ先行型研究での方法と変動の分析の例として、可視化技術、交通系ICカードデータの交通目的の推定手法について述べる。

(1) 可視化技術の構築⁽⁶⁾⁽¹⁾⁽²⁾

都市高速道路網や鉄道網で数年以上の長期間に渡って収集されたデータや交通系ICカードデータを俯瞰し、データに含まれる変動の特徴を見つけ出す、発見的な分析に適用することを目的に、可視化技術の構築を行った。

長期に渡って観測された大量のデータをできるだけ、効率的かつ正確に把握するために、適切な配色を用いた情報の正確な表示、「高速かつシームレスな表示」を要件としてシステムを構築した。構築した方法では、縦軸に日付、横軸に時刻をとり、色で観測値を表現し、地点毎に図化する。このような図を見たい地点について、複数枚の図を並べて表示したり、順にブラウジングすることで、長期間の変動を視覚によって俯瞰する。このような手順を効率的に行うため、例えば、交通系ICカードデータの可視化では、定点集計データをデータベース化し、要件を満たす可視化を行っている。また、色彩工学⁽⁶⁾の知見に基づいた配色により、色の見え方の知覚的な差と観測値の知覚的な変化量が比例するように色を割り当て、交通量や速度などの変動を色の知覚的な変化から把握するための方法を構築し実装した。

可視化技術を用いた分析では、観測地点や観測期間・時刻帯などの条件を指定して可視化を行い、可視化によって得られた画像を視認し、特徴を探索する。例えば、ある高速道路路線の上流から下流までの画像を順次ブラウジングし、気になった現象がある地点では、カラーバーに割り当てる値域を変更したり、特定の時間帯等を拡大したりし、変動の特徴を見て理解するような使い方をする。条件を変化させながら特徴の探索を行い、探索によって特徴が発見された場合には、従来型の研究と同様に、データのサンプリング、現象に対するモデル化及び定量化と検証を行う研究プロセスに移行する。

図-1に可視化を行った例を示す。図は、2003年3月2日～2006年5月31日の平日(土曜・日曜・祝日及び12月30日～1月3日、8月12日～16日を除く日)に、阪神高速松原線の本線料金所の直下部(上り9.0KP(二車線))の5分間断面交通量を可視化したものである。この図は、松原線のデータの特徴を探索していた際に得られた図の一例であ

る。図によると、2003年3月から2005年6月までの期間に、朝の9時付近に特徴的な白色の縦線が現れている。これは交通量が低下していることを表している。この現象は、料金所の運用状況の影響であり、2005年頃からのETCの普及に伴って影響が減少していると推察できる。

図-2に、鉄道会社AのICカードデータを用いた可視化の一例を示す。この図は、2007年7月2日～2008年3月31日の平日に関して、高校や大学が付近にある中規模のC駅での1分毎の入場者数を示している。図では、朝のピーク時に細い縦縞が見られ、15時から17時にかけて間

隔の広い縦縞が見られる。より詳しい分析を行うと、朝の縦縞は、列車の発車時刻間隔に対応しており、午後の縦縞の間隔は、周辺の学校の終了時限にあわせて利用者が駅にやってくる状況を示していることがわかった。また、2007年8月、12月にはこの間隔の広い縦縞がない期間があり、夏期、冬期の長期休暇の状況も現れている。

以上に示したように、可視化技術を用いた特徴の発見分析では、必ずしも事前の予見や知識を前提として現象を探索するのではなく、過去にさかのぼって変動の特徴を見つけ出すことに重点が置かれる。

(2) 交通系ICカードデータの交通目的の推定⁽⁶⁴⁾

交通系ICカードのデータでは、改札を通過したICカード利用者全員の改札通過時刻及びカードを識別できるID情報が記録される。このことから、利用者のIDを識別したうえで、追跡的な分析ができることから、既往の利用者のIDを識別できない定点データとは異なる変動のとらえ方が可能となる。例えば、ID情報が付加された情報では、利用頻度を集計することができ、需要の変動が、単純な利用者の増減により引き起こされているものなのか、個々の利用者の行動が変わったことによる利用頻度の増減により引き起こされているものなのかを知ることができる。一方で、ICカードデータは、交通行動を収集することを目的としたデータではないことから、交通行動データとして完全な形ではない。例えば、公共交通機関の形態によっては、乗車又は降車時のデータ取得いない場合もあり、分析の目的に応じた推定方法が提案されている⁽⁶⁶⁾⁽⁶⁸⁾。また、ICカードデータでは、交通目的やトリップの出発地、目的地など、従来の交通行動を収集することを目的としたデータで収集されている項目を収集できない。したがって、交通機関利用者の変動をみつけた際には、データマイニング等の手法などを用いて、利用者集団を行動パターンで束ねることで、より変動の特性が見えやすくなると考えられる（例えば、Ma et al., 2013⁽⁶⁹⁾）。しかし、そうして抽出されたパターンの解釈は分析者の知見にゆだねられることになる。そこで、本研究では、データマイニングを行う際に、ICカード利用者の行動の変動の要因に関する解釈をより行いやすくすることを目的として、トリップ目的等の項目を推定する

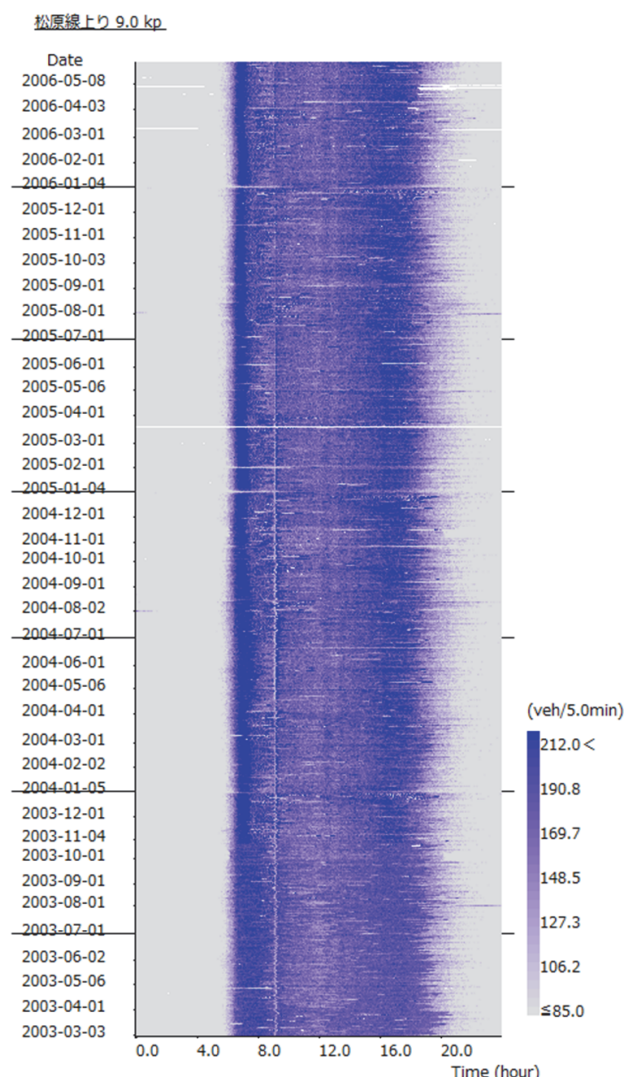


図-1 車両検知器データの可視化の一例

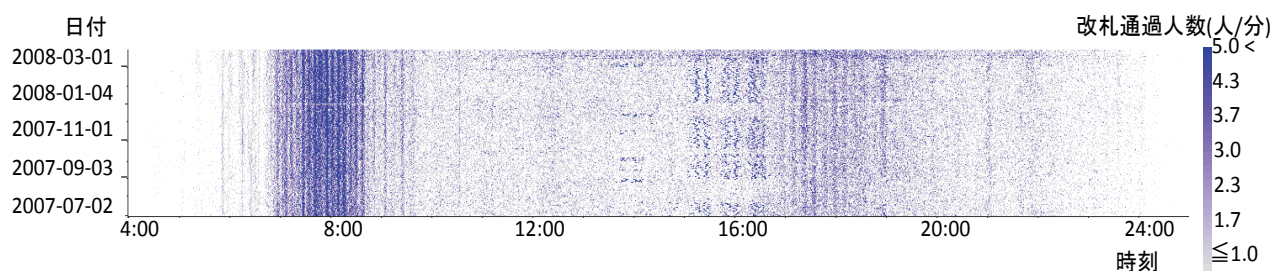


図-2 交通系ICカードデータの可視化の一例

ことを提案した。

本研究の基本的な手順は、ナイーブベイズ分類器⁷⁾を用いた交通目的判別モデルをパーソントリップ調査のデータを用いて作成し、交通目的判別モデルをICカードデータに適用することで、ICカードデータに含まれるトリップの交通目的を判別するものである。本研究では、ICカードデータとパーソントリップ調査データの双方に含まれる変数である「降車時刻」と「次回乗車までの時間間隔」を引数として、トリップ目的を推定している。

以上の方法を用いて、判別を行ったICカード利用者のトリップ目的別利用者数の日変動を示したものが、図-3である。なお、データは、都市部に路線をもつ鉄道会社A社a駅で観測されたICカードデータで、対象の期間は、2007年10月1日から2009年5月31日までの平日である。また、この間に観測されたa駅を乗降するトリップ数は、6,913,979トリップであった。交通目的判別モデルの推定に用いたパーソントリップ調査データは、2002年に実施された第4回京阪神パーソントリップ調査によるデータである。図より、私用以外の目的のトリップでは、お盆や年末年始にトリップ数が減少していることがわかる。帰宅目的のトリップは他の目的のトリップに比べて、4月から9月にかけてのばらつきが大きくなっている。これは、この路線の他の駅近辺にあるイベント施設で行われているイベントの期間に一致しており、その帰宅の乗客が他の社局の路線への乗り換えているものを捉えたものだと考えられる。

図-4は、推定されたトリップ目的ごとの利用日数の月別の分布を示している。通勤目的のトリップでは、利用日数のピークは平日の日数と一致している月が多いことがわかる。つまり、すべての平日で鉄道を利用する規則的な行動パターンを行う人が多いことが把握できる。年末年始がある12月、1月とお盆のある8月では、最頻値が出現する日数が平日の日数よりも小さくなっている。また、最頻値の値が小さくなっていることから利用日数のばらつきも大きくなっていることがうかがえる。これらの月では、鉄道利用のパターンが多様であることがわかる。私用又は業務目的のトリップでは、最頻値は、月1日の利用となっており、どの月も約2割以上を占めている。また、私用・業務目的や帰宅目的のトリップでは、年末年始、お盆による変動はあまり顕著ではないことが分かる。

このように、ICカードデータを用いた変動分析では、従来の調査データでえることが難しい、トリップパターンの経時的変化を分析できる。一方で、交通の文脈も考慮して変動を解釈するためには、必要に応じた推定方法が必要である。例えば、本研究で示した、トリップ目的と利用頻度の関係は、パーソントリップ調査のデータ、ICカードデータのそれぞれ単独のデータセットは推定す

ることはできない。本研究で示したように、データ項目を補完して用いることで、はじめて明らかになる変動についての関係があることがわかる。

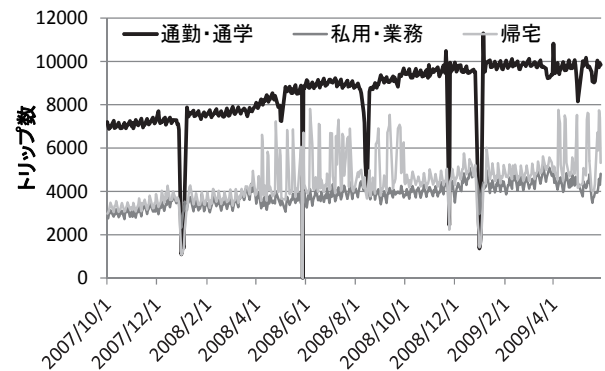


図-3 トリップ目的の推定結果

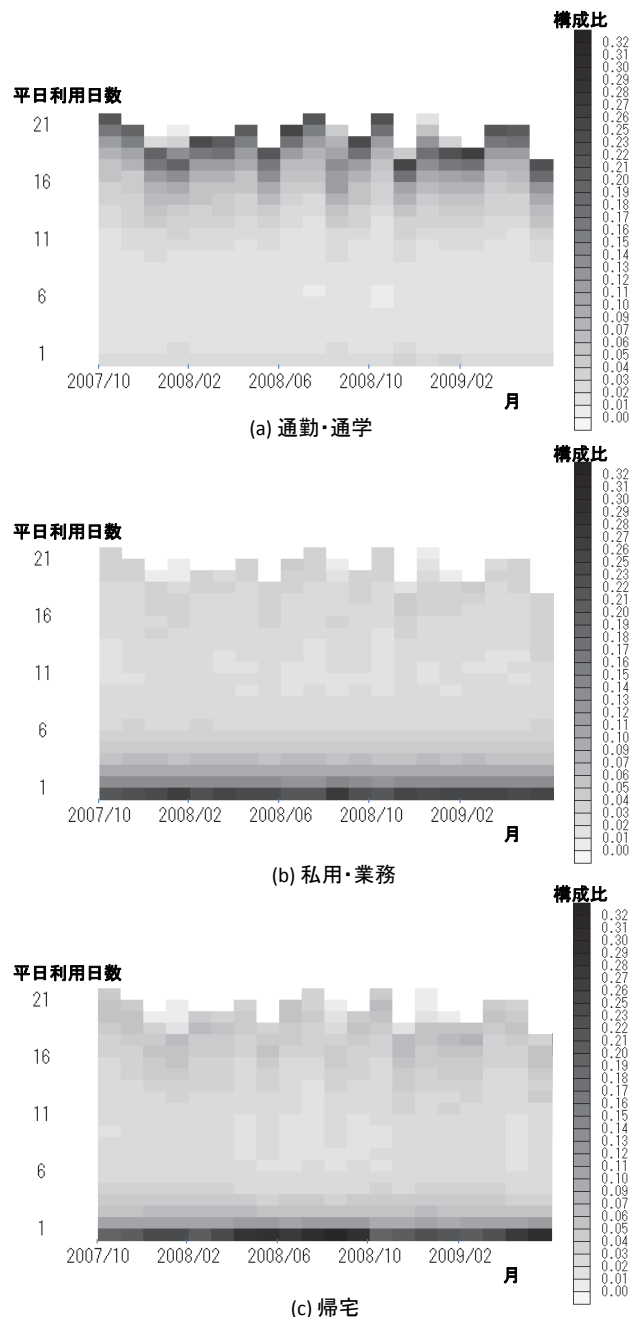


図-4 トリップ目的別利用日数の推定結果

4. おわりに

本稿では、交通の変動を分析することの重要性を示したうえで、長期間・継続的に蓄積されたデータを前提としたデータ先行型の研究環境が整いつつあることを示し、交通の変動を捉えるためのデータオリエンテッドな研究アプローチで必要となる研究の要素について示した。データオリエンテッドな研究では、交通分野のデータとして不完全なデータの取り扱いや、データに含まれている変動の発見、発見された変動をモデルとして記述というプロセスが必要であることを述べた。

第3章では、大量のデータからの現象の発見や解釈のための手法として、可視化技術や交通目的の推定手法を示した。これらの手法は、交通変動の把握・理解を行うものとして位置づけているが、より直接的な応用として、施策実施後に、想定したような感度での影響が現れているかについての継続的なモニタリングを行うことや、既往のパーソントリップ調査等で確認された変化の軌跡の確認等を行うことで変動の原因を推測することに役立つと考えられる。もちろん、施策の重要度に応じてプローブパーソン調査のような能動的調査を行うことも必要であるが、新たな調査が必要な分析には蓄積データの分析に比べて大きな費用がかかることから、調査を行うことの効果を検討したうえで実施する必要があるだろう。

蓄積データの実務への展開では、データ収集者とのデータ共有の問題やプライバシーの問題などもある。また、データ取得者と分析者が異なる場合には、分析に適したデータ構造でのデータ提供がなされるようなフレームワークを構築していくことも重要である。例えば、牛木ら⁷⁾の研究では、プローブカーデータを突発事象検出への適用する際の手法を検討しているが、実装可能な手法は、データ収集の方法のみならず、提供されるデータの形態によって限定されうることを示している。つまり、データの収集や提供の際の規格が決まる前に、交通分野での利用で最低限必要なデータの仕様を示す必要があるといえ、データ収集に先行した手法の開発は重要であると考えられる。このような背景から、例えば、瀬尾ら²⁾の研究では、今後、自動運転等の普及により収集される可能性がある車間距離データに着目した、交通状態推定手法を構築し、推定精度等の検証を行っている。

第1章で示したように、今後5年間で蓄積データの量は飛躍的に伸びることが予測されており、利用できる蓄積データの種類も多くなると考えられる。このことから、今後の研究では、これまでの研究でのデータを単独で用いるだけでなく、複数のデータセットを用いることで、それぞれのデータセットの不完全な部分を補完し、より時間的・空間的に高密度・広範囲ので変動分析を行うことが期待される。

最後に、今後の研究では、蓄積データによる分析を進展させるだけでなく、そこから得られる交通現象の変動について知見を積み重ねることにより、変動の理解・理論化につなげることで、交通分析や交通計画の高度化を通した土木計画学分野への貢献が期待される。そのためには、巨大なデータを処理する能力がある人材はもちろんであるが、このようなデータが映し出す現実に立ち向かえる人材の育成も必要であるだろうと思う。

謝辞：本招待論文を執筆する貴重な機会をくださった土木計画学研究委員会に深く御礼申し上げます。本論文で紹介した著者の研究は、東京工業大学朝倉康夫教授、神戸大学井料隆雅教授との共同研究が含まれます。朝倉康夫教授、井料隆雅教授には、著者の学部生時代からご指導をいただき、現在に至るまで大変多くのことを教授いただきました。また、阪神高速道路株式会社、首都高速道路株式会社、匿名の鉄道会社には、データ先行型の研究を行うに当たって不可欠なデータ提供をしていただきました。ここに深く感謝の意を表します。

参考文献

- 1) 北村隆一：変動についての試行的考察，土木計画学研究・論文集，Vol. 20，pp. 1-15，2003。
- 2) Kitamura, R: Panel analysis in transportation planning: An overview, *Transportation Research Part A*, Vol.24, No.6, pp 401-415, 1990.
- 3) Hilbert, M., and López, P. : The world's technological capacity to store, communicate, and compute information, *Science*, 332(6025), pp.60-65, 2011.
- 4) Turner, V., Gantz, J. F., Reinsel, D., and Minton, S. : The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things, <http://idcdocserv.com/1678> (2014年9月16日アクセス), *International Data Corporation (IDC)*, 2014.
- 5) 喜連川優：情報爆発のこれまでとこれから，電子情報通信学会誌，Vol. 94, No.8, pp.662-666, 2011.
- 6) 羽藤英二：交通データとモデル-プローブパーソントリップ調査の実行可能性と課題-，第27回土木計画学研究発表会・講演集，Vol.27, 259, 2003.
- 7) Greenshields, B. D., Thompson, J. T., Dickinson, H. C. and Swinton, R. S. : The photographic method of studying traffic behavior, *Highway Research Board Proceedings*, Vol. 13, pp. 382-399, 1933.
- 8) Olcott, E. S. : The influence of vehicular speed and spacing on tunnel capacity, *Journal of the Operations Research Society of America*, Vol. 3, pp. 146-167, 1955.
- 9) Greenberg, H. : An Analysis of Traffic Flow, *Operations Research*, Vol. 7: pp.78-85, 1959.
- 10) Klein, L. : Traffic detector handbook third edition—volume I, *Federal Highway Administration Report*, No. FHWA-HRT-06-108, 2006.
- 11) Robert L. Gordon, P. E., Warren Tighe, P.E. : Traffic control systems handbook, *Federal Highway Administration Report*, No.FHWA-HOP-06-006, 2006.

- 12) 溝渕修治：阪神高速道路公団における情報提供技術と効果について，交通工学，Vol.34，No.4，pp.27-30，1999.
- 13) 飯田恭敬，高山純一：高速道路における交通量変動特性の統計分析，高速道路と自動車，Vol.24，No.12，pp.22-32，1981.
- 14) Travel Survey Methods Committee (ABJ40): The On-Line Travel Survey Manual: A Dynamic Document for Transportation Professionals, <http://www.travelsurveymanual.org/> (2014年9月16日アクセス), *Transportation Research Board*, 2014.
- 15) Stopher, P. R. and Metcalf, H. M. A. : NCHRP Synthesis of Highway Practice 236: Methods for Household Travel Surveys, pp.11, Transportation Research Board, National Research Council, Washington DC, 1996.
- 16) 米谷栄二：起終点調査にもとづく街路網計画について，土木学会論文集，No. 34，pp.1-35，1956.
- 17) 黒川 洸：都市圏パーソントリップ調査の歴史，土木学会誌，Vol.98，No. 10，pp.12-13，2013.
- 18) Hanson, S. and Huff, J.O. : Assessing day-to-day variability in complex travel patterns, *Transportation Research Record*, 891, pp.18-24, 1982.
- 19) Pas, E.I. : Weekly travel-activity behavior, *Transportation*, Vol.15, pp.89-109, 1988.
- 20) Kitamura, R. and Bovy, P. H. L. : Analysis of attrition biases and trip reporting errors for panel data. *Transportation Research Part A*, Vol.21, pp.287-302, 1987.
- 21) Murakami, E. and Watterson, W.T. : The Puget sound transportation panel after two waves, *Transportation*, Vol.19, pp.141-158, 1992.
- 22) 内田敬，飯田恭敬：交通行動パネル調査の方法論的検討，土木計画学研究・論文集，Vol.11，pp.319-326，1993.
- 23) 杉恵頼寧，羽藤英二，藤原章正：パネルデータを用いた交通機関選好意識のダイナミック分析，土木計画学研究・論文集，Vol.10，pp.31-38，1992.
- 24) 佐佐木綱，松井寛：会場内の観客流動モデル，土木学会論文集，Vol. 159，pp. 90-95，1968.
- 25) Mahrnassani, H. S. : Dynamic models of commuter behavior: experimental investigation and application to the analysis of planned traffic disruptions, *Transportation Research Part A*, Vol. 24, pp.465-484, 1990.
- 26) 金子哲也，大口敬，飯田克弘，渡邊亨：ドライビング・シミュレータを用いた追従挙動特性の研究，土木計画学研究・講演集，Vol.23，pp.801-804，2000.
- 27) Hensher, D. A. : Stated preference analysis of travel choices: the state of practice, *Transportation*, Vol.21, pp.107-133, 1994.
- 28) 日下部貴彦，辻本洋平，朝倉康夫：旅行時間信頼性情報による高速道路利用者の行動変化の分析，土木学会論文集 D3，Vol.68，No.5，pp.I_781-I_792，2012.
- 29) 遠藤俊宏，田畑浩規，反納敏，辰巳浩，梶田佳孝：パーソントリップ調査におけるWeb回答方式導入の有効性とその特性に関する考察，土木計画学研究論文集，Vol.24，pp.83-90，2007.
- 30) Zito, R., Este, G. D., Taylor, M. A. P. : Global positioning systems in the time domain: How useful a tool for intelligent vehicle-highway systems?, *Transportation Research Part C*, Vol.3, pp.193-209, 1995..
- 31) Sermons, M. W. and Koppelman, S. : Use of vehicle positioning data for arterial incident detection, *Transportation Research Part C*, Vol.4 (2), pp.87-96, 1996.
- 32) 大森宣暁，室町泰徳，原田昇，太田勝敏：交通行動調査へのGPSの適用可能性に関する研究，交通工学研究発表会論文集，Vol.18，pp.5-8，1998.
- 33) Murakami, E. and Wagner, D.P. : Can using global positioning system (GPS) improve trip reporting?, *Transportation Research Part C*, Vol.7, pp.149-165, 1999.
- 34) 朝倉康夫，羽藤英二，大藤武彦，田名部淳：PHSによる位置情報を用いた交通行動調査手法，土木学会論文集，Vol.653，pp.95-104，2000.
- 35) Asakura, Y. and Hato, E. Tracking survey for individual travel behaviour using mobile communication instruments, *Transportation Research Part C*, Vol.12, pp.273-291, 2004.
- 36) 上條俊介，松下康之，池内克史，坂内正夫：時空間Markov Random Filed モデルによる隠れにロバストな車両トラッキング，電子情報通信学会論文誌 D-II，Vol.J83-D-II，No.12，pp.2597-2609，2000.
- 37) 倉内文孝，宇野伸宏：交通現象を理解するための画像観測データの活用 (特集 交通工学における画像処理の適用)，交通工学，Vol.42，pp.12-19，2007.
- 38) 鈴木一史，中村英樹：交通流解析のための簡便な多機能ビデオ画像処理システムの開発，交通工学研究発表会論文報告集，Vol.24，pp.69-72，2004.
- 39) 鈴木一史，中村英樹：簡便な交通流観測ツールとしての画像処理へのニーズと実用課題 (特集 交通工学における画像処理の適用)，交通工学，Vol.42，pp.20-27，2007.
- 40) 日比野 直彦，中山 泰成，内山 久雄，高平 剛：鉄道駅における歩行者データの取得および活用方法に関する一考察，土木計画学研究・論文集，Vol.21，pp.781-787，2004.
- 41) 布施孝志，中西航：歩行者挙動モデルを統合した人物自動追跡手法の構築，土木学会論文集 D3，Vol.68，No.2，pp.92-104，2012.
- 42) 古屋秀樹，牧村和彦，川崎茂信，赤羽弘和：車載型センサーを用いた車両挙動特異地点，抽出に関する基礎的研究，土木計画学研究・論文集，Vol.21，pp.787-795，2003.
- 43) Hoogendoorn, S. P., van Zuylen, H. J., Schreuder, M., Gorte, B. and Vosselman, G. : Microscopic traffic data collection by remote sensing. *Transportation Research Record*, 1855, pp. 121-128, 2003.
- 44) 中村克行，趙 卉菁，柴崎亮介：レーザスキャナと画像センサの融合による歩行者追跡，画像センシングシンポジウム講演論文集，Vol.11，pp.177-180，2005.
- 45) 北澤俊彦，塩見康博，田名部淳，菅芳樹，萩原武司：Bluetooth 通信を用いた旅行時間計測に関する基礎的分析，土木計画学研究・講演集，Vol.47，303，2013.
- 46) Al-Deek, H. and Abd-Elarhman, A. : An evaluation plan for the conceptual design of the Florida transportation data warehouse, Phase-1-, Final Report, Center for Advanced Transportation Systems Simulation, Department of Civil & Environment Engineering, University of Central Florida, 2002.

- 47) 河野浩之, 高田裕之, 石井康裕, 久利良夫: データマイニングによる障害継続時間予測モデルの検討ー阪神高速道路における交通データウェアハウスの活用, 人工知能学会第 20 回全国大会, 1E2-1, 2006.
- 48) 河野浩之, 羽藤英二: 交通計画のための空間データウェアハウス構成技術, 電子情報通信学会技術研究報告. AI, 人工知能と知識処理, Vol.103, No.306, pp.9-14, 2003.
- 49) 田中康仁, 小谷通泰, 中村賢一郎: プローブデータを活用した貨物車による配送活動の実態分析, 土木計画学研究・論文集, Vol.22, pp.715-722, 2005.
- 50) 三輪富生, 森川高行: プローブカーデータを利用した経路選択行動に関するモデル分析, 土木計画学研究・論文集, Vol.21, pp.553-560, 2004.
- 51) 宇野伸宏, 永廣悠介, 飯田恭敬, 田村博司, 中川真治: バスプローブデータを利用した所要時間信頼性評価手法の構築, 土木計画学研究・論文集, Vol.23, pp.1019-1028, 2006.
- 52) 例えば, 富士通株式会社: 車の詳細な走行挙動がわかる商用車プローブデータ・サービスを提供開始, <http://pr.fujitsu.com/jp/news/2014/07/8-2.html> (2014 年 9 月 16 日アクセス), 2014.
- 53) 内田敬, 藤由雅裕: 経路情報とリンクフローを用いたランプ間交通量推定法, 土木計画学研究・論文集 Vol.16, pp.675-680, 1999.
- 54) 上田大樹, 井料隆雅, 朝倉康夫: 長期 ETC 統計データによる異なるランプ間 OD 交通量と旅行時間の相関分析, 交通工学, Vol.49, No.3, pp. 43-52, 2014.
- 55) Pelletier, M., Trépanier, M. and Morency, C.: Smart card data use in public transit: a literature review, *Transportation Research Part C*, Vol.19, pp.557-568, 2011.
- 56) 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.
- 57) 森田泰智, 太田雅文, 窪田崇斗, 家田仁: 車両応荷重データを用いた都市鉄道の時間帯別・車両別混雑率の分析, 土木計画学研究発表会講演論文集, Vol.36, 11, 2007.
- 58) 太田恒平, 大重俊輔, 矢部努, 今井龍一, 井星雄貴: 携帯カーナビのプローブ交通情報を活用した道路交通分析, 土木計画学研究発表会・講演集, Vol.47, 323, 2013.
- 59) 石村 怜美・梶原 康至・太田 恒平: 乗換検索サービスの経路選択データを用いた公共交通の経路選択行動分析, 土木計画学研究発表会・講演集, Vol.49, 7, 2013.
- 60) 日下部貴彦, 井料隆雅, 朝倉康夫: 車両検知器データを用いた交通流可視化技術の開発, 交通工学, Vol.43, No.4, pp.59-68, 2008.
- 61) Kusakabe, T., Iryo, T. and Asakura, Y. : Data mining for traffic flow analysis: Visualization approach, In: Barceló, J. and Kuwahara, M. (eds.) *Traffic Data Collection and its Standardization*, pp.57-72, Springer Series, 2010.
- 62) 日下部貴彦, 中島良樹, 朝倉康夫: 可視化技術をもちいた交通系 IC カードデータの分析, 土木計画学研究・講演集, Vol.39, 244, 2009.
- 63) Kusakabe, T., Iryo, T. and Asakura, Y. : Estimation method for railway passengers' train choice behavior with smart card transaction data, *Transportation*, Vol.37, pp.731-749, 2010.
- 64) Kusakabe, T. and Asakura, Y. : Behavioural data mining of transit smart card data: A data fusion approach, *Transportation Research Part C*, Vol.46, pp.179-191, 2014.
- 65) 日本色彩学会: 新編色彩科学ハンドブック, 東京大学出版会, 1989.
- 66) Trépanier, M., Tranchant, N. and Chapleau, R. : Individual trip destination estimation in a transit smart card automated fare collection system, *Journal of Intelligent Transport Systems*, Vol.11 pp.1-14, 2007.
- 67) Chu, K. K. A. and Chapleau, R. : Enriching archived smart card transaction data for transit demand modeling, *Transportation Research Record*, 2063, pp.63-72, 2008.
- 68) Seaborn, C., Attanucci, J. and Wilson, N.: Analyzing multimodal public transport journeys in London with smart card fare payment data, *Transportation Research Record*, 2121, pp.55-62, 2009.
- 69) Ma, X., Wu, Y.-J., Wang, Y., Chen, F. and Liu, J. : Mining smart card data for transit riders' travel patterns, *Transportation Research Part C*, Vol.36, pp.1-12, 2013.
- 70) Rish, I. : An empirical study of the naïve Bayes classifier, *IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence*, 2001.
- 71) 牛木隆匡, 日下部貴彦, Nguyen Xuan Long, 朝倉康夫: 交通容量低下を考慮したプローブカーデータによる突発事象検出手法, 土木計画学研究発表会・講演集, Vol.48, 277, 2013.
- 72) 瀬尾亨, 日下部貴彦, 朝倉康夫: 車間距離測定プローブカーを用いた都市高速道路の交通流観測実験報告, 交通工学研究発表会論文集, Vol.34, pp.277-283, 2014.

(2014. 9. 19 受付)

DATA-ORIENTED APPROACH IN TRANSPORT STUDIES

Takahiko KUSAKABE

Recent advances of information technologies enable us to collect huge datasets relating to transport systems. Passive data, such as traffic detector data, smart card data and probe vehicle data, can be automatically and continuously collected along with operation of the systems. The long-term continuous records of travellers' use of the transport systems can be stored in such passive datasets. This paper summarises the history of data collection in the transport studies in order to describe differences between conventional datasets and passive ones. Based on the passive datasets, the perspectives of the studies on variabilities in transport network are discussed.