

論文 / 著書情報
Article / Book Information

論題	特別講演：交通ビッグデータの活用
著者	福田大輔
出典	JICEレポート, , No. 28, pp. 31-39
発行日	2015, 10
Note	引用・転載不可

「交通ビッグデータの活用」



東京工業大学大学院
理工学研究科
土木工学専攻
准教授

福田 大輔 氏

プロフィール

1999年、東京大学大学院 工学系研究科 社会基盤工学専攻修士課程修了。2001年、東京工業大学大学院 理工学研究科 土木工学専攻助手。2005年に助教授、2007年から准教授として現在に至る。2007～2008年にデンマーク工科大学交通研究所に客員研究員として滞在。

はじめに

東京工業大学の福田です。本日は予定されていた広島大学大学院国際研究科の藤原章正先生の代理となりますが、「交通ビッグデータの活用」というテーマで講演いたします。

今回の「日・韓建設技術セミナー」は26回目ということで、四半世紀以上に及ぶ歴史があるとお聞きしました。このような由緒ある研究セミナーで発表することを大変光栄に思います。力不足で、皆さんの関心に応えられる講演内容になっているかどうか定かではありませんが、できうる限り努めさせて頂きたいと思います。

今回、私が代理の講演者にご指名いただいた理由のひとつに、この講演の後半で紹介しますが、日本の土木学会の中でこの2年ほど活動している「交通関連ビッグデータの社会への実装研究小委員会」で、藤原先生が委員長を、私が幹事長を務めている関係があるかと思っています。

私は東京工業大学大学院の土木工学専攻で、主に交通計画や交通工学の研究や教育に従事しています。ここ数年は、交通ビッグデータも研究活動のひとつとして少しずつ取り組んでいます。

このような経緯もあり、本日の発表のスライドのベースは藤原先生が作成され、それに私が若干の内容を加えた構成となっています。

ビッグデータとは何か その狭義と広義

本日の講演タイトルは「交通ビッグデータの活用」です。

ビッグデータという言葉は、今は交通に限らずさまざまな場面で使われていますが、定義自体も曖昧で、何でもビッグデータと呼ぶ風潮が世の中にあるように思われます。

ビッグデータの統一的な定義はまだありませんが、よく言われているのは、まず狭義には、管理するのが困難な大量のデータ群を指します。ただそうしますと、従来からの交通や社会の動向などに関するさまざまな統計調査もビッグデータということになります。そこで、最近のビッグデータの特徴づける要素として、Volume、Variety、Velocity の3つの頭文字をとって、「3V」と言われる場合が多くなっています。

ビッグデータとして皆さんが一番思い描きやすいのが、最初のVolumeです。非常に大量の、データサイズではギガ、テラ、ペタといったデータが蓄積されている状況を指し、ビッグデータという言葉そのものを示すと思います。

次のVariety。これは交通においてはさまざまな移動体の軌跡や、最近では交通事業者等が運営するICカード等による料金収集の際に得られたデータが挙げられます。あるいはもっと広くは、例えばインターネットのブログやツイッター、SNSなどもビッグデータのひとつの形態です。あるいは動画や画像的なデータなども含め、いろいろなデータのタイプがあるという意味でVarietyという言葉が使われています。

3つ目のVelocity（速度）。これは少し分かりにくいのですが、例えば公共交通のICカードなどにおいて絶えずリアルタイムにデータが蓄積されていく、その蓄積頻度が非常に速いことをVelocityと称しています。

このようにして収集されたビッグデータは、継続的かつ半永久的に収集できるデータ群と考えられます。これが、データそのものを指すときのビッグデータの基本的考え方です。

しかし、ただデータがあって終わりという訳ではありません。データを処理するのは我々人間です。このデータ処理系までも含めて考えるべきであろうと、広義の定義もなされています。データ群を蓄積、処理、そして分析するための技術体系、つまりさまざまなデータ処理の技術もビッグデータと呼びます。より広義には、それらのデータを分析して有用な意味や洞察を引き出す人材や組織、あるいはそれらの教育のための仕組みまで含めてビッグデータと考えるのが一番適切なのではないかと思います。

ビッグデータが急速に広まっている社会的背景

ビッグデータの話が近年、急速に世の中に広まっている背景について説明します。

世界では、ICTの目覚ましい進歩と普及が進んでいます。その傾向は、住民100人あたりの携帯電話契約数のグラフ（図1）に如実に表れています。先進国ではおおむね100人に対して120の携帯電話契約数がある状況に至っています。全世界的な平均のトレンドでもほぼ100%、1人1台という時代に近づいています。開発途上国だけに着目しても、100%にどんどん近づいている状況です。

一方交通系に着目すると、スマートフォンや、さまざまな位置情報のデータ収集技術と、それらの蓄積。さらに、位置情報などの取得が目的でつくられたのではない公共交通のICカードや高速道路の料金収受のためのさまざまなICカードからも、交通の特徴を捉えることができます。このようなパッシブなデータも交通ビッグデータの一つです。

さらに広く言うと、構造物のモニタリングセンサー、例えば高速道路などに設置されている重量計なども含められます。このような、構造物の変異を測るスマートセンサー系も、これは人の動きを直接測るわけではありませんが、交通量の計測にも応用ができ、技術的にも蓄積量が進展している状況と思います。

このような中で、日本では2013年6月14日に「世界最先端IT国家創造宣言」をしています。

こうした社会的背景の中でさまざまなビッグデータが蓄積されています。交通ビッグデータに関しても、人口減少と高齢化といった縮退社会における交通インフラの計画、あるいは、これまでに無い24時間365日の常時モニタリングのツールとしての役割が大いに期待されています。

住民100人あたりの携帯電話契約数

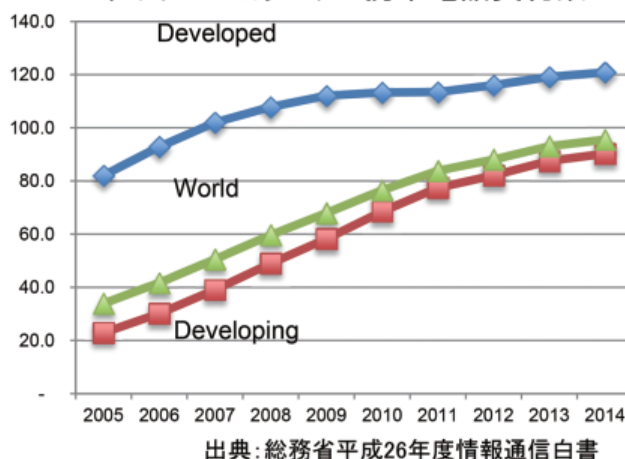


図1 住民100人あたりの携帯電話契約数

それに加え、私が個人的に重要と感じているのは、ビッグデータと既存データの融合による新たな展開の可能性です。

ビッグデータのみで、さまざまな交通インフラの計画やモニタリング等ができるわけでは決してありません。これまでのレガシーのようなもの、具体的にはこれまで行われてきたさまざまな調査のデータや既存の技術と、新しく台頭してきたビッグデータがうまく融合し、互いに代替の関係としてではなく、相互補完する役割としての新たな展開が期待されていると思いますし、研究者としてもそのような補完的な活用方法を検討する必要があります。

「IT と統計学の素晴らしい結婚」

近年、なぜ急にビッグデータについて言われるようになったのか。それは統計解析の技術が、近年大いに向上したことが理由のひとつです。コンピュータの処理能力の向上や、それに伴う機械学習を中心とする統計学的方法論についても学術的にさまざまな進展が見られたのです。

「ITと統計学の素晴らしい結婚」という言葉があります。統計学自体は、1900年代初め頃から提唱された古くからある学問ですが、それが最近の方法論のさらなる進展やデータの処理能力の進展などにより、実践的な社会技術としてさまざまな形でうまく実装されるようになってきました。

日本では、2013年に『統計学が最強の学問である』[ダイヤモンド社]という本が出版され、当時ベストセラーになりました。そのような時代背景の中で、さまざまなビジネス分野におけるマーケティング分析等においても、統計分析に注力するようになってきました。

例えばマイクロソフトやグーグルなどの研究所では、最新

のさまざまな統計解析の技術の蓄積が進んでいます。具体的にはデータマイニング、機会学習、人工知能、自然言語処理といった、実際のデータに基づいた実装可能なさまざまな手法を駆使して、それらを彼ら自身のビジネスインテリジェンスや、競合分析、いわゆるマーケティング的なものに使う戦略について検討をしています。その中では、2つの選択肢のどちらが選ばれそうかをデータ的に把握するA/Bテストなど、様々な最新の方法が検討・適用されています。このように、どんな大量データでも、どんな計算でもできる技術は、統計解析以外にはありません。

この「ITと統計学の素晴らしい結婚」というキーワードに加え、例えばアメリカではヴァリアン [Hal Ronald Varian] が2010年に、「I keep saying the sexy job in the next ten years will be statistics.」と述べています。「sexy」と記されていますが、「今後の10年で最もエキサイティングな職業は、statistics、統計学にかかわるものだ」と述べているわけです。

ヴァリアン氏は、経済学の基礎的な教科書などで有名で、私も学生の頃には、彼の著書を読んで勉強したことがあります。

経済学もさまざまな社会経済データを使い、計量経済学と呼ばれる学問名があるように、いろいろなデータ処理をします。疫学のような、どちらかと言うと自然科学系だけではなく、経済学のような社会科学系からも、同様なことが言われるようになってきた時代背景があると思います。

コッホは細菌学的な観点から スノウは統計分析的な観点から

統計学は古くは、疫学や公衆衛生学と呼ばれる分野で、1800年代の半ば頃に始まった学問と言われています。

ここで、ロバート・コッホ [Robert Koch 1843–1910] と、ジョン・スノウ [John Snow 1813–1858] の2人の名前を挙げます。

コッホは、1883年に病原体のコレラ菌を発見し、「細菌学の開祖」と呼ばれています。コレラ発症のメカニズムを解明して1905年にノーベル生理学・医学賞を受賞しています。しかし、コッホがコレラ菌を細菌学的な観点から見出すよりも前に、スノウは統計分析的な観点から、コレラ菌がどのような社会要因によって生じるのかをエビデンスに基づいて検証し、それに基づいて具体的な処方箋を見出す試みをしています。

スノウは、ロンドンでコレラが流行した当時、統計分析的な観点からコレラの原因が水源の違いによることを見出しました。それによって、例えばコレラに感染する可能性が高い水源の水利用を止める対策を行い、コレラの流行を防いだ実

績があります。

スノウは、このような観点から「公衆衛生学の開祖」と呼ばれ、都市計画や地域計画とも密接に関係があります。当時、このスノウの方法は、コッホのような細菌学的な観点ではなく、あくまでもデータだけ、結果だけに基づいた検証と処方箋の提唱であって、科学的ではないと批判されました。

しかしながら、コッホ以前の細菌学者たちは、コレラの流行を止める具体的な方法を見つけることができなかったわけです。このことから、頭やセンスや行動力に優れた人たちを集めただけでは、つまり（細菌学の）理論的な知識に秀でた人たちを集めただけでも、なかなかシンプルかつ強力な解決策は出てこない。理屈としては一見正しくても、無益もしくは有害であることがしばしばである、とも言われています。

現代はスノウの時代とは様相はかなり異なりますが、もともとのビッグデータ分析の原点は、このジョン・スノウの疫学的な話にあるのではないかと思います。

交通はミクロとマクロの関係性を 考えなければいけない

図2は、交通計画、あるいは交通行動分析と呼ばれる、私たちの研究分野の教科書によく出てくる図です。

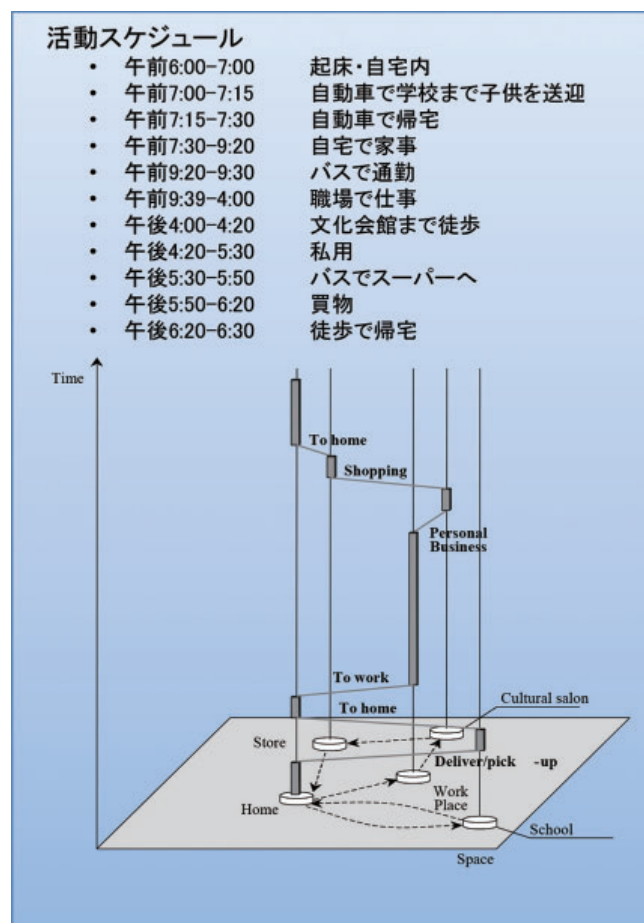


図2 1日の生活行動の時空間パス

図中の平面は地理的な空間で、家があり、店があり、職場があり、それから中心地、子どもをピックアップする場所があります。縦軸方向は時間で、下が朝、上が夜です。このように、1日の生活行動を時空間パスで表現するのが、1970年代以来の交通行動分析の基本的な考え方です。

この時空間パスを、具体的に何時から何時まで、どのような活動をどこで行っているかを記述したのが上のリストで、「活動スケジュール」と言います。朝6時に起き、子どもを送迎して車で移動。仕事をして、買い物をして、帰宅する。このような形で1日の生活行動の時空間パスを記述するわけです。

このように、時空間パスでミクロ的に一人ひとりの交通行動を見たときには、ある地点から別の地点への移動、例えば自宅から職場に移動する軌跡の部分が交通にあたります。

例えば仕事をする、子どもを迎えに行く、買い物をするなどが、人間の1日の生活行動の本来の目的であり、交通とは、そういったさまざまな根源的な活動のために、仕方なく行わなければならない行為です。こういった行為を経済学では、本来の需要から派生して行わざるを得ない「派生需要」と呼びます。それは、「生活活動のついで需要」と言えます。

このような派生需要は、日本全国の時間利用調査などによりますと、人間が1日に持っている時間資源24時間の中で平均約1.5時間ほどの時間量しか現れません。その意味では、24分の1.5しか現れない、稀な事象と考えることもできます。

さらには、このような1日の生活行動の時空間パスが365日まったく同様に繰り返されるかと言うと、決してそうではありません。周辺の環境、道路条件や交通条件がまったく同じであっても、活動には揺らぎが出てきます。まったく同じ条件でも、人間の行動は変動します。また最近では、例えば列車の中でパソコンをインターネットにつないで仕事をするブリーケース・トラベラーといった場合のように、業務と移動の同時実行が可能になるなど、移動と活動の区別ができない状況も生じています。

一方、マクロ的には、一人ひとりの活動が蓄積したとき、例えば道路の渋滞や公共交通の混雑といった社会的な現象が起きます。このような現象の多くは、経済学的には外部不経済と呼ばれ、社会的にはネガティブな影響として現れます。

一人ひとりにとって、自分はこう行動するのが生活のスケジュールの観点から一番望ましいと思い、皆がそれぞれ独自に最適な行動をとった場合、渋滞がより少ない社会的に望ましい結果になるとは限らない、ということです。こうした現象を社会科学全般では「社会的ジレンマ」と呼ぶこともあります。このミクロとマクロの関係性をきちんと考えないといけない点が、交通の難しいところです。

このような状況の中、過去何十年、さまざまなインフラの

投資や政策等を行っても、交通事故や渋滞の発生回数は実はそれほど変わっていません。さらには、交通のための時間が1日に平均1.5時間費やされると述べましたが、これが過去と比べて短くなってきているかと言うと、決してそうではありません。交通体系が整備されると、その分、活動範囲が広くなり、交通に消費する時間配分の量は結局それほど変わらないという指摘もされています。

このような交通行動を的確に捉えるために、これまでのような古くからの交通調査の方法では、細かい側面を追うことは不可能です。この点に関して、さまざまな交通関連のビッグデータが補完的に活用できるのではないかと考えています。

交通にかかわる現象の正しい理解にビッグデータを役立てる

ビッグデータで何ができるか、大きく3つに分けて紹介します。1つ目が「交通にかかわる現象の正しい理解」、2つ目が「多様で複雑な政策の評価」、そして最後が藤原先生と一緒にやってきた「学会としての取り組み」です。

まず、「交通にかかわる現象の正しい理解」です。さまざまな交通政策や交通事業が、本当の意味で社会的に望ましい帰結をもたらすかどうか。それを事前に把握するためには、先ほど示した、人々のさまざまな派生需要としての交通と、その後に起こるさまざまな生活行動のパターンを、より正しく理解することが第一に重要になってきます。

こうした交通行動、交通現象の正しい理解のために、日本を含め多くの国では、ペーパーベースの質問紙によるさまざまな交通調査が従来から行われてきました。代表的なものとして日常の通勤・通学・私事目的等の移動を調査するパーソントリップ調査があり、さまざまな都市圏でこれまで数多くの実例があります。

東京圏でも過去に5回行われ、現在第6回調査の検討がされています。詳細な調査票を世帯構成員一人ひとりに記入してもらい、それに基づいてさまざまな交通現象の理解に役立てようとしています。最初に向かった場所はどこか、何時に家を出たか、目的は何か、手段は何か、どれくらい時間がかかったか、どこで乗り換えたか。その当日あるいは過去の交通の行動を、帰宅後に調査表に記入してもらう形で行われています。

しかし近年、このパーソントリップ調査については、かなりの時間とコストをかけて収集された調査データであるにもかかわらず、例えば回答者が十分に情報を記入していない、調査自体に市民が協力して答えない、いわゆる社会調査の協力率の低下などが、問題として挙げられています。

こうした従来型の調査は、レガシーとして様々な知見を現

在に与えることから重要であり、何らかの形で今後も継続すべきであると思います。それと並行して、今台頭してきているさまざまな交通関連のビッグデータが、このパーソントリップ調査で十分に拾いきれなかった交通の重要な側面を捉えきれる可能性があると考えられます。

GPS 等による移動軌跡データが細かい時空間の分解能を可能に

交通関連のビッグデータとして、日本でも1990年代半ばから現在まで研究や実例として一番多く取り組まれているのが、GPS等による移動軌跡のデータです。

PHSが多く利用されていた時代から、その基地局のデータなどから移動軌跡のデータを得て、それを交通行動の調査に利用していました。現在では、GPS等によって移動軌跡のデータを収集する取り組みが、日本だけではなく全世界的な傾向として増えています。

3年ほど前より、従来の紙ベースの調査とスマートフォンから得られた詳細な移動軌跡の情報をうまく融合して、熊本都市圏の交通計画の策定に役立てようという取り組みが、熊本大学の円山先生のグループを中心に行われました。熊本都市圏を対象に、回答者にスマートフォンのアプリをインストールしてもらい、詳細な移動軌跡のデータを集められています。

移動軌跡のデータは、正確な経路選択情報はもちろん、自動車であればもっとミクロに、道路の利用車線まで、細かい時間的・空間的な分解能で分かるようになってきています。

これらのデータは、従来のさまざまな交通政策や計画の評価、交通の需要予測など以外にも、交差点の右左折や車線変更など、いわゆる交通工学的な側面の政策の検討にも十分に適用が可能です。また、従来の紙ベースの調査ではサンプル数が少なく拾いきれなかった、抜け道を選択する交通や、うろついている回遊交通などのデータも捉えられるようになってきています。

GPSデータは一人ひとりの移動軌跡です。しかし、それらが複数人分集まると、例えば車線変更時に複数のドライバーが互いに駆け引きをして、タイミングをみて追い越し車線に移るといった「複数ドライバー間の繰り返しゲーム」の状況において、個々人の行動から複数のドライバーがうまく協調するような、いわゆるシステム的な考え方への拡張が可能になることも考えられます。

同様に、例えば歩行者や自転車などで、互いに影響を及ぼし合う混み合いが生じる場面においても、この移動軌跡データ群がうまく活用されるようになりつつあります。

移動軌跡データの一番の特徴と思われるのは、個人や車両

のさまざまな行動の、長い時間軸を通しての揺らぎをとることができる点です。特に強みであるのは、例えば大きなイベント、具体的には自然災害や大規模なコンサートなどの社会的なイベントが起きたときに、それが起きていないときのデータと比べてビフォア・アフターの分析などを行うことで、そのイベントの効果でどれほどの交通や需要の変化があったかを的確に捉えられるようになることです。このあたりが、従来型の紙ベースの調査では追えなかった一番の特徴であると思います。

渋滞が延伸していくときや、逆に解消していくときの交通流、自然災害発生後の避難行動、新しい道路区間が供用したときに需要のパターンが空間的に安定してくるまでプロセスなどを、よりつづさに追うことができる。これは特にマーケティング的な観点からも、需要がどう定着していくのかを考えるうえで重要な側面です。このように、変動を的確に、細かい時空間の分解能で追えるのが、移動軌跡データのメリットです。

また、これまでの交通行動分析は、例えば費用が一番安いルート、時間が一番安いルートといった、いわゆる経済原則に則って人々は交通行動や移動をすると捉えて行われてきました。しかし、例えば高齢者あるいは介助者に付随して移動する行動などは、経済原則に必ずしも則りません。そのようなデータも、このGPSによる移動軌跡データを用いて分析できると考えられます。

交通 IC カードからは乗り継ぎ・乗り換え行動のデータ取得も可能

交通ビッグデータと近年言われるようになった一番のきっかけは、今説明したGPSによる移動軌跡データではなく、これから述べる交通ICカード等によるTransaction（料金の支払い等の商取引）に関するデータの蓄積にあると思います。

具体的には、公共交通系のさまざまなICカードや、高速道路における料金収受のETCの普及がきっかけです。これら交通ICカードは、あくまでも利用者からお金を徴収するのが目的です。しかし、例えば公共交通のICカードでは、どの駅のどの改札を何時に通過して、どの駅で何時に降車したかといった形で、その人の駅間の起点と終点の関係を追うことができます。

特に公共交通に関しては、交通モードにかかわらずICカードが相互利用されていく傾向にあります。例えば異なる交通手段間の乗り継ぎや乗り換えの行動も、厳密にそのデータを追えば捉えることができます。

例えば、同じ出発地から目的地に行くにあたって、例えばバスや列車の運行本数が比較的多い朝の時間帯、それらの

本数が相対的に少ない昼間のオフピークの時間帯では、時間や費用が一番少なくなる経路が異なります。（図3）

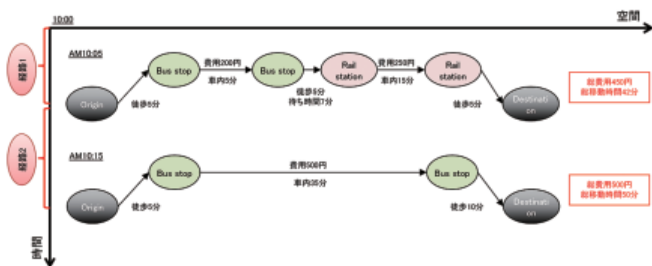


図3 出発時刻によって最短経路とLOS (Level-of-service)が異なる例

図3の経路1（上段）はピーク時の最短経路、経路2（下段）はオフピークの最短経路です。同じ起点と終点であっても、オフピーク時に最短の経路を探すと、起点から終点までバス1本で移動するのが所要時間が一番短くなり、同じ1日の中でも行動の違いが現れることが見えてきます。

このように、その利用者が何時にその公共交通を利用するかによって移動の利便性がどう違って来るかを定量化でき、それを政策にうまく役立てることもできると思います。

藤原先生は広島市で使われている交通ICカードのパスピー（PASPY）を用いて、路面電車とバスの乗り継ぎ行動を分析されています。（図4）

日本の大都市の場合、バスと鉄道の相互連携、同じ人がバスと鉄道をどう乗り継いだかを一気通貫で追える状況には今のところなっていません。バス部分と鉄道部分のICカード利用が完全に切り離されてつながっていないからです。しかし地方では、例えば広島市の場合、複数の事業者の協調により、バスと路面電車の乗り継ぎを一体的に追うことができます。

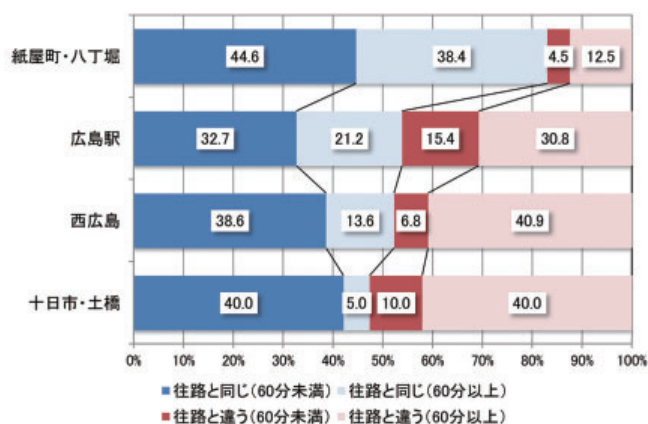


図4 公共交通ICカードを活用した乗り継ぎ行動分析：復路広島市

この図4に示されているのは、各乗車駅の例です。青系の色は往路と復路の経路が同じ人の割合で、そのうち濃い青が乗り継ぎ時間60分未満、薄い青が60分以上の場合。赤系の

色は往路と復路が違う人で、同様に濃い赤が60分未満、薄い赤が60分以上の場合です。

この結果から分かるのは、公共交通の乗り継ぎ客のうち、乗り継ぎ時間60分以上、薄い青と薄い赤を合わせた割合が45%～54%と、駅によって異なることです。もっと重要なことは、往路と復路で乗り継ぎ地点が異なる乗客（赤系）が、少ない駅では17%ぐらいですが、多い駅では半分になる場合もあることです。

こうした事実は、これまでの紙ベースの調査ではなかなか追うことができなかった情報です。交通ICカードによって、これまで把握できなかったこうしたさまざまな交通の側面を正しく追うことができるようになってきます。

ETC 技術で特異日の渋滞予測やきめ細かな料金体系などが可能に

2つ目の「多様で複雑な政策の評価」について、ETC（Electronic Toll Collection：料金収受システム）のビッグデータの例を挙げます。

例えば大型連休のような、1年で日数的に少ないけれどもなり特徴がある特異日の渋滞予想は、ビッグデータに基づくほうが、分析者の経験値や何らかの数学的なモデルに基づいて予測するより、精度が高くなると言われています。

加えて政策的な観点から言うと、高速道路利用の場合、時刻別、距離帯別、車種別、さらに海外事例にある車線別、将来的には高速道路利用の予約別に加えて、これまでの100円や10円単位の料金設定よりもさらにきめの細かい1円単位での料金収受も徐々に対応可能になると言われています。

現在、首都圏の高速道路の料金体系などが細かに議論されています。さまざまなニーズに応じた、きめ細かな料金体系に移行できないかといった検討がなされているわけですが、それができるようになった一番の理由はETCの普及です。

特に近年、ETC2.0という最新の情報サービス、料金収集関係の技術とその普及の大きな傾向が見えてきたことも議論の背景にあると思います。

ETC2.0は、料金収受のみならず加減速情報なども取得できます。例えば高速道路区間における、加減速が著しく起こりやすい、事故リスクが高い場所をピンポイントで特定できるなど、交通安全分析への貢献も期待されています。よりきめの細かい事故データを得ることができるわけです。

それに加えて長期的観点からは、データの長期間の蓄積により、アセットの管理データ、つまり道路インフラなどの維持管理のためのデータへの活用なども考えられます。

事故などの希少事象の評価に 交通ビッグデータは活用できる

交通事故統計とGPS等のプローブデータの例として、高速道路における交通事故のような希少現象の分析に、プローブデータが果たす役割を説明したいと思います。

従来、高速道路に求められている機能とは、渋滞の中を移動しなくていい「快適性」と、より短い時間で目的に到達できるという「速達性」です。こうした機能が、サービスとして第一に求められてきました。加えて近年では、移動の「定時性」、すなわち、予定時間に遅刻せず目的地に着けるといった信頼性や、従来型の調査ではうまく捉えることが難しかった「安全性」に関する側面も、きちんと評価をしないとイケない時代になっていると思います。

統計学の分析の方法論としては、快適性と速達性と言った日常的な状況の評価に関しては、いわゆる古典的な従来型の統計学、具体的にはガウス統計と呼ばれる正規分布を基軸として大数の法則が成り立つような、平均値に基づいて議論すればよいという考え方に基づく統計学の枠組みの中で分析できたわけです。

しかし今後、信頼性や安全性のような、低頻度でしか起こり得ない事故などの希少事象の評価を考えたときには、方法論についても新たな拡張が必要となります。例えば重大なインシデント〔incident:出来事〕が起こる可能性を評価する極値統計学と呼ばれる分布帯や、外れ値などをどのように扱うかに関する知見が必要になってくると考えられます。

安全性に関しては、これまでなかなかデータに基づく検討ができませんでした。どちらかと言うと、そのプロフェッショナルの専門知に委ねる例が多かったわけです。

ニューヨークで2009年1月、飛行機がバードストライク〔bird strike:鳥が衝突する事故〕を受けてハドソン川に不時着した事例があります。このとき、経験のあるパイロットが適切に川に軟着陸して、結果として1人の犠牲者も出ず、「ハドソン川の奇跡」と呼ばれています。これは、航空機発着の数億回のさまざまな成功事例から1回の失敗事例（不時着）を抜き出すという考え方とも捉えることができ、どちらかと言うと、プロフェッショナルの専門知が重要であることを示唆するものです。しかしビッグデータには、このような、ある意味経験値的なものをうまくデータから抜き出す力もあると思います。それによって、外れ値や異常値の分析や、さらに安全性や事故リスクの分析につながると考えられます。

交通ビッグデータの活用により、今述べたようにきめの細かい政策やその分析ができるようになることが期待されます。平常時に限らず非平常時、具体的には大きなインシデントが起こるような場合の評価などにうまく活用できる可能性

があると思います。

「交通関連ビッグデータの社会への 実装研究小委員会」の活動

このような検討を、藤原先生と私と30人ぐらいのメンバーで、2012年11月～2014年10月の2年間、土木学会の中に「交通関連ビッグデータの社会への実装研究小委員会」を立ち上げ、活動してきました。最後に、この我々の取り組みの例を、簡単に紹介します。

ビッグデータが、世の中にどんどん蓄積されています。特に交通系のビッグデータは、さまざまな移動軌跡の情報などに限らず、ICカードやインフラのモニタリングのデータなどが蓄積されています。

こうした状況のもと、我々は、従来型の交通計画、あるいは交通行動分析のパラダイムを超えて、何ができるか、それをうまく社会の中に実装していくにはどうすればいいかという観点から検討を行ってきました。交通の研究者に限らず、維持管理や経済評価、まちづくりの研究者、さらには各分野の実務の方々にも多く加わっていただいています。

活動の総括として、2015年3月にワンデーセミナーを行いました。ここで発表された、交通関連ビッグデータのいくつかの適用事例を3件ほど紹介します。

事例① スマホ Probe Person データを用いた都市内回遊行動の分析

1つ目は、先述の熊本市圏におけるGPS、「スマホ Probe Personデータを用いた都市内回遊行動の分析」です。スマートフォンの中にアプリケーションをインストールしてもらい、回遊行動をはじめとした1日の移動及び活動の情報を得るという、熊本大学の円山琢也先生の研究成果です。（図5）

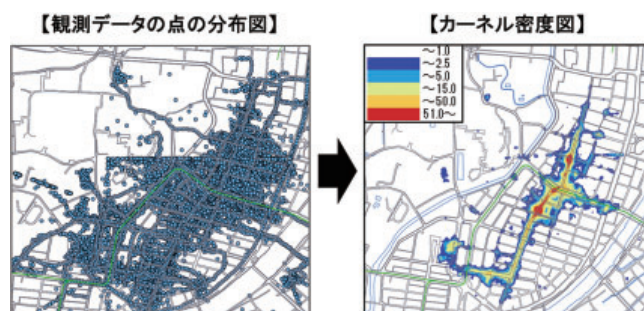


図5 スマホProbe Personデータを用いた都市内回遊行動の分析

円山琢也（熊本大学）
「熊本PT調査と連携したスマホPP調査のデータ分析」
土木計画学研究委員会ワンデーセミナーシリーズ No.74 講演資料, 2015.

これはどちらも熊本市の中心市街地の地図で、左側是个々のスマホProbe Personデータの加入者から収集された立ち

寄りポイントと言いますか、通過ポイントです。この通過ポイントを何10秒間隔に計測して、全員分の移動軌跡を単純にプロットしています。この大量な元データに対し、統計解析技術、機械学習の方法のひとつであるカーネル密度図を用いたのが右側の図です。これは人口密度のように、うまく空間上で視覚化したものです。

都市内の回遊などで、特にどの地点をより多くの人が来訪しているのか、さらにそれが時間帯によってどのように変遷しているのかを直感的に把握できます。ビッグデータに最新の統計解析技術を組み合わせることによって、こうした側面を見ることができます。

事例② モバイル空間統計の都市間交通分析への適用

2つ目のモバイル空間統計の事例は、事例①と同じく携帯電話のデータです。しかし、GPSではなく、各地に設置の電波の受信基地局のどこで今コンタクトがとられたかの情報をもとに、日本全国24時間、そのエリアにいる滞在人口を1時間おきに把握しています。（図6）

こうしたデータは、我が国では1年ほど前から、交通政策、計画の検討の場面でも活用されるようになっていきます。

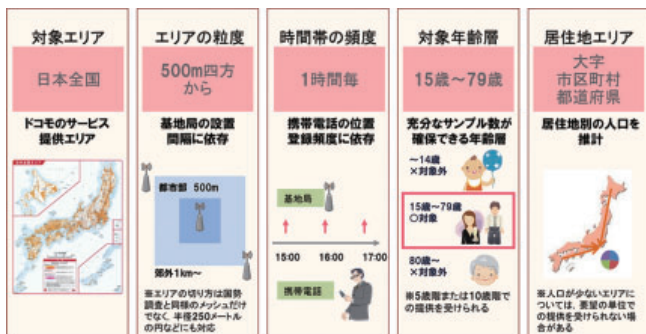


図6 モバイル空間統計

室井寿明（運輸政策研究機構）
「モバイル空間統計による都市間交通分析の可能性」
土木計画学研究委員会フンディセミナーシリーズ No.74 講演資料, 2015.

このモバイル空間統計は、NTTドコモが提供を始めたルートで、500m×500mメッシュで1時間毎の滞在人口が得られる、かなり膨大な量です。NTTドコモはキャリアとしては一番大きく、ほぼ全数かつ全国的な移動が把握可能になります。

これを活用したのが、運輸政策研究機構の室井寿明氏の研究です。（図7）

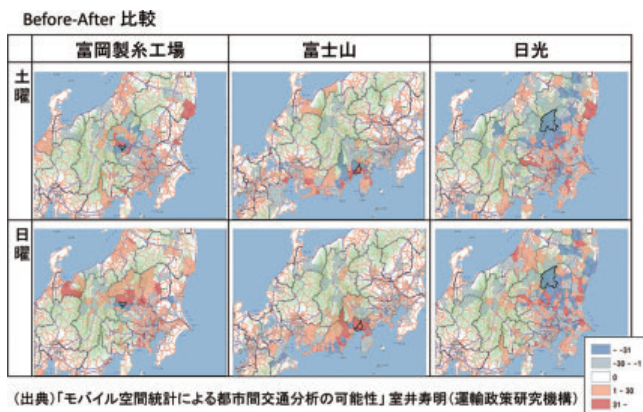


図7 モバイル空間統計の都市間交通分析への適用

室井寿明（運輸政策研究機構）
「モバイル空間統計による都市間交通分析の可能性」
土木計画学研究委員会フンディセミナーシリーズ No.74 講演資料, 2015.

日本の代表的な観光地、左から富岡製紙工場、富士山、日光と3つの例を挙げています。それぞれの地図で太い線が囲まれているところが観光エリアです。

これらの図は、例えば富岡製紙工場の場合、世界遺産に指定される1年前と、指定された1年後に、各地からの来訪者がどれくらい増えているか、ビフォア・アフター比較を示しています。1年前と1年後のデータ、その間に世界遺産に指定されるというイベントが起き、両方の滞在人口の差分をとることで、世界遺産に指定されることによる来客増の効果を推計する試みです。

赤系の色が、来訪客が増えているエリア。例えば富岡製紙工場へは、世界遺産指定によって日曜は金沢方面から訪れる人が増えていることを意味しています。同様のことが、富士山や日光などの観光地でも行われています。

ビッグデータの一番の強みは、政策の前後でデータがとられていれば、単純な比較によって政策の効果を一定程度定量的に把握できるという点にあると思います。

事例③ 交通ビッグデータと既存交通調査の融合的活用

最後の事例は私共が行った都市内交通調査の研究です。先述のモバイル空間統計と似た混雑統計というデータがあります。混雑統計は携帯電話基地局ではなく、GPSのデータが基準になっています。

図8は、東京都23区の各500mメッシュの1時間毎の人口分布図です。通勤時間に鉄道の延伸方向に滞在人口が増えているのが、鉄道のネットワークの概形として見えていると思います。

東京都23区
2012年10月30日
(水曜日)

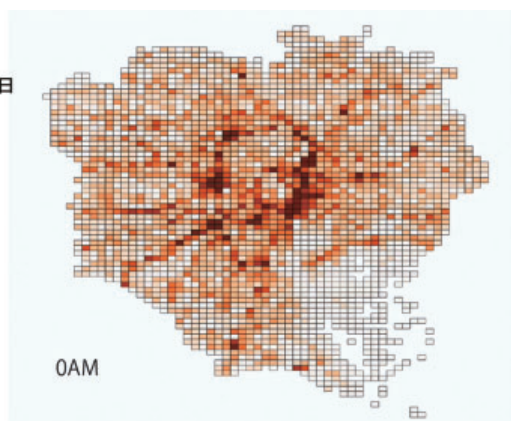


図8 東京都心部の混雑統計表示

このような、ビッグデータとして自動的に収集されるデータと、従来から日本で行われてきた交通調査を補完的に活用できないかと考え、私の研究室では、2012年の混雑統計と、2008年に行われた従来型の東京圏のパーソントリップ調査のデータをうまく融合して、どの区からどの区へ何人ぐらいの人が流動しているかを示すOD表（Origin-Destination Table）を、2008年のものから2012年のものに更新する試みを行っています。

図9は、東京23区で何区から何区へ移動したかのOD表を2008年と2012年で比較し、どれくらいの変化があったのかを示しています。45度の直線上にあれば変化がなかったことになり、直線より上側であればOD交通量が増え、下側であれば減っています。

本来はパーソントリップ調査などが行われる10年に1度ぐらいの間隔でしか構築することができなかった都市圏のOD表を、モバイルビッグデータを補完的に活用することによりパーソントリップ調査が実施されていない年のOD表も獲得することができるというひとつの例です。

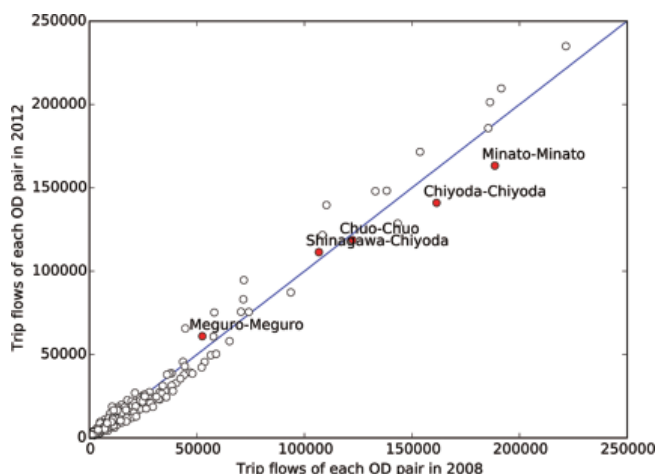


図9 23区内の移動の変化

ビッグデータは公共財としての価値が非常に大きい

交通のデータに限らず多くのビッグデータは、社会に共有されるべき、公共財としての価値が非常に大きいと思います。特に、公共性の側面が強い交通や都市計画、まちづくりなどにおいては、より価値が高いと思います。

一方で、ビッグデータは万能のように思われるかもしれませんが、課題はもちろんあります。GPSなどの測定技術について、それからあくまでもGPSを持っている人たちだけのデータ収集なので、それを持っていない人たちの交通行動を予測するときはどう拡大するか、交通シミュレーション等におけるクローンをどのように生成するのかといった技術的な課題なども残されています。

それにも増して最近よく言われているのは、移動の軌跡や支払いのような細かいプライバシー情報の問題です。この問題に対する適切な対応の検討については、匿名化などの技術も進展していますが、引き続き取り組まなければいけません。

最後が政策への活用です。現在、さまざまな形のオープンデータも徐々に普及してきて、ビッグデータが適用されています。しかし、例えば先述のモバイル空間統計や混雑統計は各500mメッシュで、それ以上の細かい位置情報を我々が分析に使うことはできません。そうすると、より細かい情報が必要な政策（例えば、自転車ネットワークのルート選定など）の分析などには適用できません。そのようなデータ活用や共有に対する公的指針もきちんと考えていく必要があると思います。

本内容は、2015年9月1日に開催された「第26回日・韓建設技術SEMINAR PROGRAM」の特別講演によるものです。