

人の流動と時空間データセット最前線

関本 義秀

実世界のさまざまな課題・事象と OR が扱うようなさまざまな理論をつなぐものとして、空間データ/時空間データは進んできたと言えるのではないだろうか。つまり、実世界のさまざまな課題・事象が何かしらの概念モデルに基づき表現され、データ化されるが、そうしたデータが単純すぎず複雑すぎず適切で、かつ多くの人が使える状況にあって初めて意味を成す。とくに最近では、防災、交通、マーケティングなど、時々刻々と変化する社会において、人を中心とした、より動的かつ大規模を対象にしたデータのニーズが増しており、今回は「人の流動に関わる大規模な時空間データセット」について、筆者らの経験を踏まえ述べていきたい。

キーワード：時空間データ、人の流動、大規模、携帯端末、パーソントリップ調査

1. はじめに

今回のテーマは「使ってみよう空間データ/時空間データ」であるが、実世界のさまざまな課題・事象と OR が扱うようなさまざまな理論をつなぐものとして、空間データ/時空間データは進んできたと言えるのではないだろうか。つまり、実世界のさまざまな課題・事象が何かしらの概念モデルに基づき表現され、データ化されるが、そうしたデータが単純すぎず複雑すぎず適切で、かつ多くの人が使える状況にあって初めて意味を成す。

筆者が関わっている空間情報の分野は、こうしたデータセットの価値に焦点が当てられ進歩してきたと言えるのではないだろうか。実際に地理情報システム学会 (GIS 学会) では、論文のなかでも「データ論文」というカテゴリーが存在して、データセットそのものの価値について査読を受けることができる。また、後ほど多少触れるが、国際的にもオープンデータやデータチャレンジといった取り組みは増えており、まさにさまざまな分野の人をつなぐわかりやすいものの一つとしてデータセットが存在している格好である。

特に最近では、防災、交通、マーケティングなど、時々刻々と変化する社会において、人を中心とした、より動的かつ大規模を対象にしたデータのニーズが増しており、今回は時空間データという切り口で、筆者らが 2008 年から進めてきた「人の流れプロジェクト (<http://pflow.csis.u-tokyo.ac.jp>)」などの経験を踏まえ、「人の流動に関わる大規模な時空間データセット」

について述べていきたい。

2. 携帯端末の普及前～調査データの活用

まずは、携帯端末が広く普及する前は、こうした移動体に関する時空間データセットは多くなかった。例えばカーナビゲーションシステムは 1990 年代頃からかなり見られるようになっていたが、個人利用が主体で多くの車両データを集めるということにはなかった。

その一方で、マスを対象にするという意味では、交通計画分野では、昔から交通調査は存在し、国内ではパーソントリップ (以下、PT と呼ぶ) 調査や道路交通センサス調査という形で (海外では Household Survey と呼ばれることが多い)、数万人～数十万人を対象に、紙の原票を用いて通過交通量や、各トリップの起終点 (Origin-Destination, 以降 OD と呼ぶ) とその通過時刻、交通手段などを記録していた。これらは、もともと地域間とくに市区町村レベルの人の出入りの数を集計的に調査し、長期の交通計画・地域計画を策定することを主な目的にしていたため、他の用途に利用することはあまり想定していなかった。

しかし、筆者らは、その原票データを、断片的ではあるものの、マスを対象にしている貴重な時空間データとして位置づけ、再利用を試みる研究を 2004 年頃から始め、上記の人の流れプロジェクトにつながっていった。具体的には、トリップ、正確に言うと、交通手段別に記録されているサブトリップ単位の OD の時空間データを、最短経路ベースで 1 分ごとの位置を求め直した、いわば resampling データを作成することである。ここでもとの時空間情報は調査データなので、空間は地名や施設名で表現されるのでジオコーディングで緯度経度に変換する必要があるし、1 分ごとの位置を

せきもと よしひで
東京大学 空間情報科学研究センター
〒 277-8568 千葉県柏市柏の葉 5-1-5

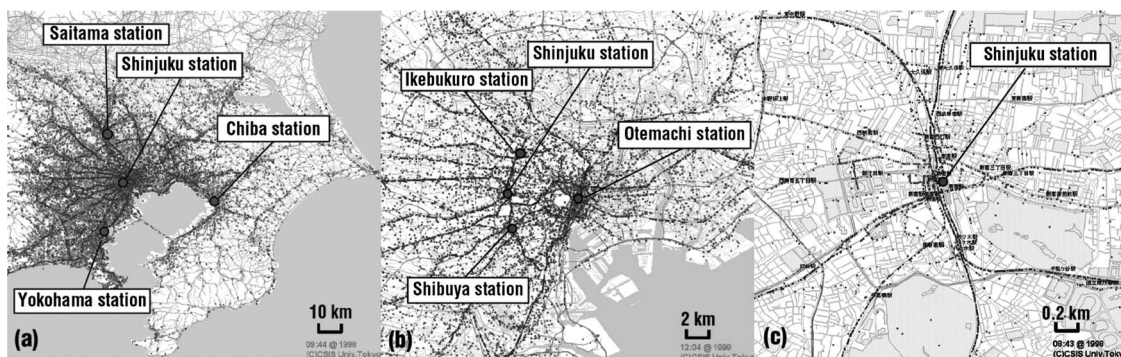


図1 パーソントリップデータを用いた流動再現 (Sekimoto et al.(2011) [1]) をもとに作成. (a), (b), (c) は同じデータについて、スケールを変えて視覚化している. 濃い色は鉄道や徒歩、薄い色は自家用車や自転車による移動を表す)

求める過程では道路や鉄道ネットワークに沿うことによりリアリティが増すため、空間データをふんだんに使うことになる。これらの手法詳細は文献 [1, 2] を見ていただきたいが、結果は図1のとおりであり、1998年の東京都市圏のPT調査のマスターデータをもとに再現したものである。最短経路ベースでサブトリップを内挿するという単純化を行っているものの、各人の1分ごとの位置を求めることにより、ある時間での多くの人の分布状況がわかることとなる。

この時空間データセットはPT調査のマスターデータをもとに作成しているものの、上記のようにある程度の「推定・加工」を行っているので、パーソントリップ調査の主体である国（国土交通省）や地方自治体の承認のもとで、「人の流れデータ」という名前で筆者の現在の所属である東京大学空間情報科学研究センター（CSIS）から産官学の公益に資する範囲で2008年から共同研究の枠組みを通じて無償提供している。2012年10月現在、14都市圏16調査、合計約350万人（延べ人数）分になるが、共同研究の数は数十件にのぼり、感染症の伝搬シミュレーション、環境影響評価、ダイナミックなマーケティングなど、多岐の利用にわたっている。

3. 携帯電話の普及後

3.1 グローバルトレンドと日本の状況

2000年代後半に入ると、携帯電話にGPSの搭載が標準的になるとともに、基地局の情報が研究レベルで使えるようになることにより、位置情報が身近かつ大規模になり、ユビキタスコンピューティングなどの研究コミュニティが一気に広がった。特に、Rattiら [3] やBarabasiら [4] の研究の影響が大きいのではないだろうか。

この辺の研究は裾野が拡大し、現在は個人情報の取り扱いを含め、どのようにしたら社会に根づくものになるかや、社会における人間の移動特性をどのようにとらえるか、などの研究が増えてきているように思える。実際に2012年初頭のダボス会議（世界経済フォーラム）でも、世界の課題解決のために、個人、公的セクター、民間の情報を組合せ、共有財産として扱っていく、というような内容の提言もあった。実際に、携帯キャリアをはじめとしてさまざまなところで、自らのデータを貸与し、新しいアイデア、社会の課題解決などを募る「データチャレンジ」の取り組みなども増えてきた。

ひるがえって日本はどうだろうか？ 日本でも朝倉ら (2000) [5] のように、かなり前に人の交通行動について、PHSを用いて面的に把握しようというものはあったが、通信事業者法による規制もあり、民間企業を含め、社会的に出始めてきたのはかなり最近のこと

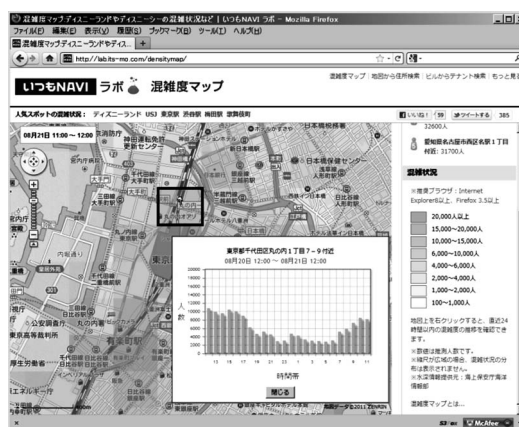


図2 混雑度マップ ((株)ゼンリンデータコム HP より引用 <http://lab.its-mo.com/densitymap/>)

である。具体的には、NTT ドコモ社による基地局情報を集計した「モバイル空間統計」や、NTT ドコモ社が提供するオート GPS 機能を用いたゼンリンデータコム社の「混雑統計データ」(図 2) などがある。法律の運用などの詳細は文献 [6] に譲るが、個人情報の取り扱いもあり、今後は特に、技術と制度、ビジネスの広がりなどバランスを取りながら進められることを期待したい。

3.2 GPS からわかること

われわれ自身も携帯端末のデータを用いた研究を行っている。例えば、ゼンリンデータコム社の協力のもと、混雑統計データ[®]のもととなる、数十万人分の 1 年にわたる非集計のオート GPS データ (個人が特定されないよう秘匿処理を行っている) を用いた分析をここではいくつか紹介したい。

長期にわたるデータがあるため、例えば東日本大震災時の状況などの概観を見ることができる。図 3 は震災直前 14:45 と震災直後 14:57 の人の流動状況の比較を行っており、さまざまな方向へ向かう動きが減り、動きを示す点の数自体も減っていることがわかり、視覚化するだけでもそのインパクトは十分に伝わってくる。

次に一人一人の動きも断片的には見えてくる。図 4

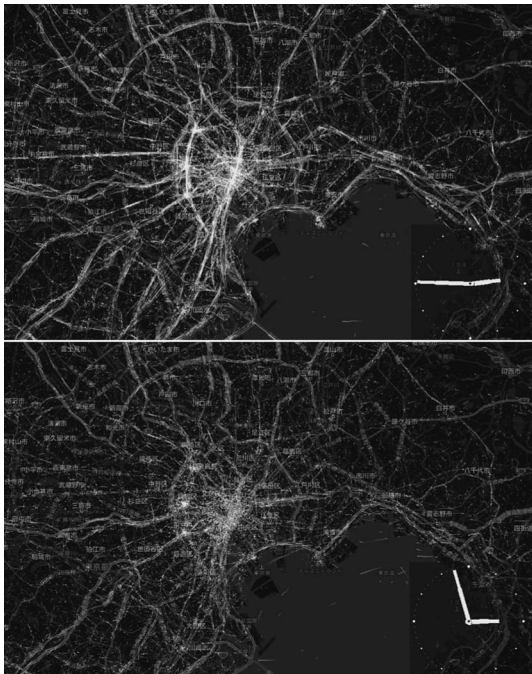


図 3 震災前後の人の流動状況の比較 (東京大学上山研究員提供、上: 震災直前 14:45 の状況、下: 震災直後 14:57 の状況、さまざまな方向へ向かう動きが減り、動きを示す点の数自体も減った状況)



図 4 GPS データによる震災当日の帰宅例 (1 年間のデータを通じ home/office 抽出をあらかじめ行い、震災直後～翌朝までに GPS データが 20 点以上あった人をピックアップ。 (a) は幹線道路を歩いて帰宅したと思われる人、(b) は途中の動きがよくわからなかった人)

は GPS データが多く観測された人 (震災直後から 20 点以上あった人) から 2 人をピックアップしてプロットしたものである。それぞれの人の home/office は 1 年分の実績からそれぞれ夜間、昼間に頻度が高く存在しているグリッドとしている。

その結果、図 4(a) は夕方から office (大手町近辺) を出発し国道 6 号線沿いをひたすら歩いて夜 24 時頃に home (松戸近辺) 到着したことが読み取れる。また、図 4(b) は入間近辺の office を出発し、わざわざ都心を経由してから朝方 6 時頃に東京郊外 (西東京)

に到着している。ただしその途中経路はよくわからず、経路の推定などは今後の課題でもある（詳細は関本ら(2012) [7]を参照のこと）。

震災の日に限らず、その後、人々の活動がどう変わっていったかなどもわかる。図5は福島で原発事故後の1週間の避難シナリオがどのように実行されたかを示している。図5(a)は事故前の3/11の通常の状態である。そして図5(b)は3/12の東京電力から津波によるバックアップ電源の消失などが発表された後の状況で、かなり人が減っている。そして、図5(c)は第1号機や第3号機の爆発後の3/14の状況である。

さらに長期的な復興状況なども知ることができる。例えば震災後に国土交通省観光庁では東北観光博(<http://www.visitjapan-tohoku.org/>)という政策を実施し、各観光ゾーンの活性化を図っている。図6は2010, 2011, 2012年のゴールデンウィークの東北地域外から来た人のトリップ数や各ゾーン間の流動状況を示している。震災直後の2011年が最も少なかったことはもちろんのこと、震災前の2010年より2012年のほうがかなり増えていることが見て取れるため、観光流動が活性化していることがわかる。ただ、2012年は復興需要などで外部から仕事で来ている可能性もあるため、そこは割り引いて考える必要がある。

ここまでの分析はGPSデータそのものや、せいぜいhome/officeなどを使い、広域的な分析を行ってきたが、長期のデータがあるため、より詳細な分析ができると考えている。そのためには単なる時空間データであるGPSデータから個々人の行動の意味づけ、ラベリングを考えることが重要である。こうした研究はAshbrook (2003) [8]以降、情報関係の分野ではかなり見られるようになってきているが大規模なものに適用した事例はなかなかなかった。大野ら(2012) [9]が同じオートGPSデータについて適用したものであるが(図7)、ある人のGPSログをプロットしたものが薄い小さい点であり、特にその人が長くいた所、滞在日数が多いものなどをステイポイント(濃く大きい点)とすると、活動パターンがかなり見えてくるため、今後の行動予測などにもつなげることができる。

3.3 CDRからわかること

一方で最近では、GPSはユーザーの個別端末からの位置情報のアップロードが必要なため、3.2節で紹介した事例はあるものの一般的には難しいことも多く、基地局側のデータを用いたCDR(Call Detail Record)の研究も増えている。例えば図8は、ある1人の20日間の期間に同時に取ったGPSとCDRである。図8(b)

のように、かなり詳細に見ていくとGPSのほうが位置精度が高いことがはっきりわかるものの、図8(a)のように引いた状態で見るとそれほど差がないようにも見える。こうした点から今のところCDRはラフな傾向分析程度にしか用いられていないことが多いようであるが、時間的にはかなり密にデータ取得がなされていることから、筆者らで、人の移動特性(通常、移動は道路や鉄道等のネットワーク上を行う)を考慮して時空間経路の候補などを作成しつつ、経路候補が観測データとしてCDRの基地局範囲内を通過している場合は尤度が高い設定をすることにより尤度最大なもの

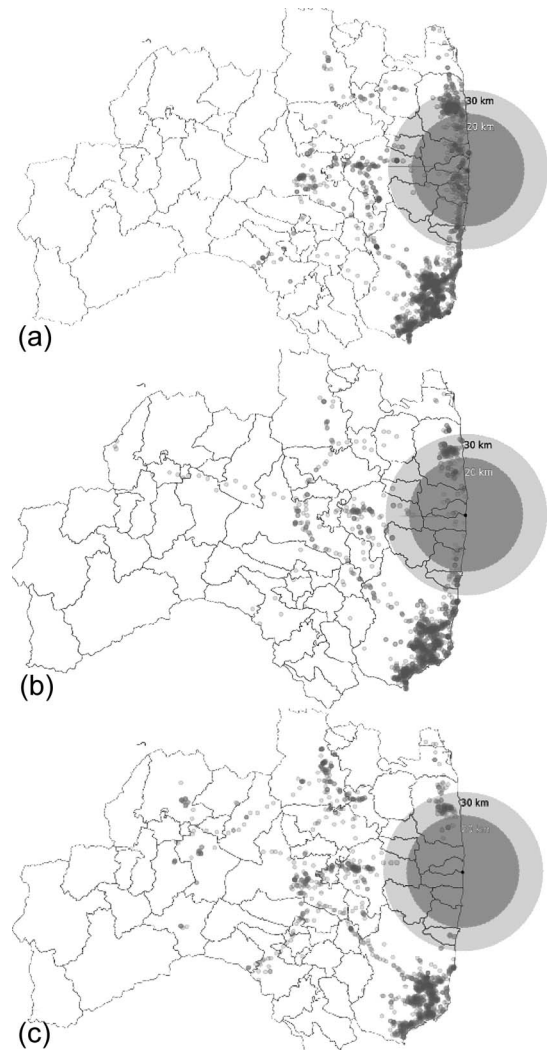


図5 原発事故から一週間の避難状況(東京大学ティーラユットホラノント研究員提供) : (a) 震災前の3/11の状況 (b) 東京電力が福島原発のバックアップ電源装置などの喪失を発表した後の3/12の状態 (c) 三号機の2回目の爆発直後の3/14の状態

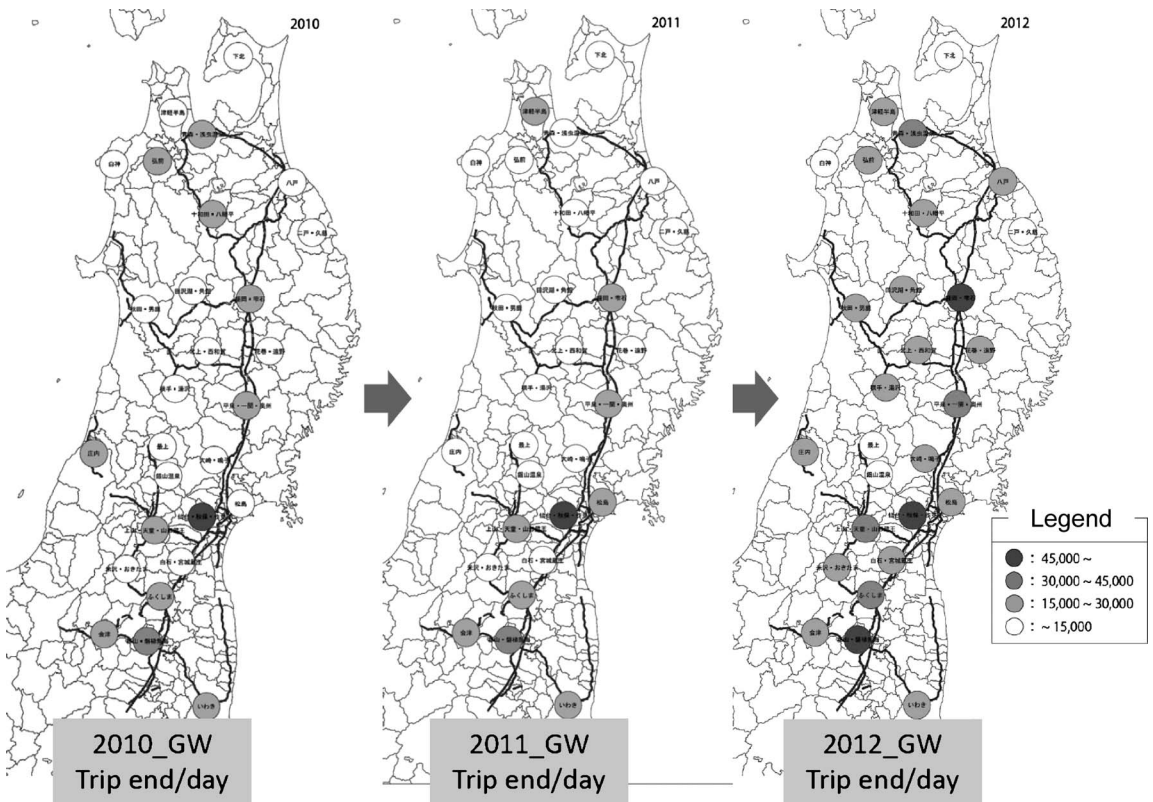


図 6 2010～2012 年のゴールデンウィークにおける東北地方における観光流動の比較
 (東京大学生形協力研究員提供. 円が各観光圏の 1 日あたりの平均トリップ数を表す. 2011 年が震災直後で最も少なく, 2012 年は震災前に 2010 年よりかなり上回っている)



図 7 ある人の全ログおよびステイポイント

に絞り込んでいく手法なども試みている [10].

3.4 今後の流れ

今後の人の流動と時空間データはどういう流れが続

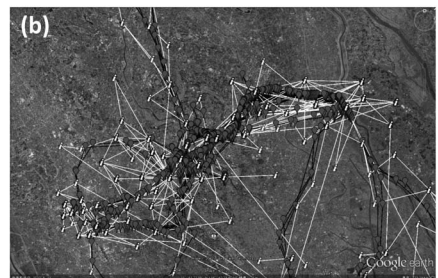
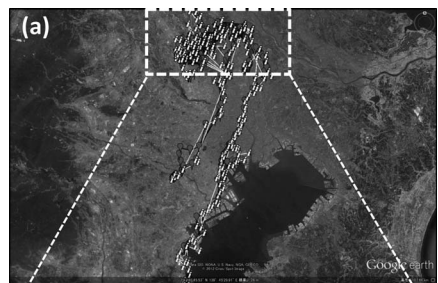


図 8 GPS と CDR の比較
 (黒が GPS, 白が CDR. (a) は全体を表現したもの, (b) はよく訪れる場所を拡大したもの)

くであろうか？国内ではまだまだ、携帯端末のようなリアルタイム性の高い大規模な実データセットが利用できる機会はあまり多くないかもしれないが、携帯端末は海外、特に途上国では社会課題解決のための唯一の情報インフラになっているケースも多い。そうした意味では唯一のソリューションとして期待される部分も多く、今まではラフな分析が多かったが、今後、行動モデルを含めたさまざまな研究が進むであろう。また、その一方で、携帯端末のみのデータであると、事業者依存になる可能性もあるため、パブリックなデータセットの確立も必要になるかもしれない。この辺は人の流れデータセットのように公的調査をベースにしたもののさらなる進化形もあるかもしれない。

4. まとめ

本稿では、空間データ／時空間データの中でも、人の流動と時空間データセットに焦点を当て、その最新動向や研究事例を紹介した。人に焦点を当てた研究は今後ますます増えていくであることが予想される一方で、こうした利用が社会の役に立つような方向での研究をより多く進めていきたい。

参考文献

[1] Y. Sekimoto, R. Shibasaki, H. Kanasugi, T. Usui and Y. Shimazaki, PFLOW: Reconstruction of peo-

ple flow recycling large-scale social survey data, *IEEE Pervasive Computing*, **10**(4), 27–35, Oct.–Dec. 2011.

[2] 薄井智貴, 関本義秀, 金杉洋, 南佳孝, 柴崎亮介, 5都市圏パーソントリップデータの比較と時空間内挿処理の実現, *土木学会土木計画学研究・論文集*, **27**(3), 569–577, 2010.

[3] C. Ratti, R. M. Pulselli, S. Williams and D. Frenchman, Mobile landscapes: using location data from cell phones for urban analysis, *Environment and Planning B; Planning and Design*, **33**(5), 727–748, 2006.

[4] M. Gonzalez, C. Hidalgo and A. Barabasi, Understanding individual human mobility patterns, *Nature*, **453**, 779–782, 2008.

[5] 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*, **653**(IV-48), 95–104, 2000.

[6] 関本義秀, T. Horanont, 柴崎亮介, 解説: 携帯電話を活用した人々の流動解析技術の潮流, *情報処理*, **52**(12), 1522–1530, 2011.

[7] 関本義秀, 中村敏和, 増田祐介, 金杉洋, 大規模なGPS情報をもとにした東京都市圏における震災時の行動分析, *土木計画学研究・講演集*, **45**, CD-ROM, 2012.

[8] D. Ashbrook and T. Starner, Using GPS to learn significant locations and predict movement across multiple users, *Journal of Personal and Ubiquitous Computing*, **7**(5), 275–286, 2003.

[9] 大野夏海, 関本義秀, 中村敏和, T. Horanont, 柴崎亮介, 東京都市圏における長期のGPSデータを用いた移動経路の推定に関する研究, *地理情報システム学会講演論文集*, **21**, CD-ROM, 2012.

[10] 金杉洋, 関本義秀, 黒川茂莉, 渡邊孝文, 村松茂樹, 柴崎亮介, 携帯電話基地局通信履歴に基づく人の移動行動の推定可能性に関する研究, *地理情報システム学会講演論文集*, **21**, CD-ROM, 2012.