

モビリティデータを用いた集積エリアに関する研究*

野際 大介^{*1}, 藤原 明広^{*2}

A Study on Agglomerated Areas Using Mobility Data

Daisuke NOGIWA^{*1} and Akihiro FUJIHARA^{*2}

^{*1} Department of Industrial Business and Engineering

^{*2} Department of Management and Information Sciences

As we use mobile phones, mobility data is accumulated because of GPS or data transaction. Especially, mobility data by data transaction is huge quantity. Therefore, we can analyze human mobility as detail. However, huge data quantity prevents us from calculating and analyzing some statistics. We have troubles in treating big data by using spread sheet application software. In this research, we suggest that statistical modeling has some solution of its trouble. And the statistical modeling is introduced random effect term in linear regression. We illustrate this model applied for mobility big data and advantage of human mobility analysis.

Key Words : Big Data Analysis, Mobility Data, Statistical Modeling, Random Effect

1. はじめに

近年、携帯電話といった移動媒体の普及が普及しており、Fig.1 によれば 2012 年以降人口普及率は 100%を越える状況となっている。また、携帯電話はスマートフォンやタブレット端末といった高機能なスペックを備える機器にとってかわり、我々の生活には欠くことのできない存在となっている。もちろん、これらの状況は日本だけでなく世界でも同様であり、移動媒体を用いたさまざまなサービスが展開されている。とりわけ、スマートフォンやタブレット端末においては、GPS (Global Positioning System) が標準搭載され、所有者自身の詳しい位置情報が簡単に取得で可能になった。また、このような高機能端末でなくとも、移動媒体が受信する電波の基地局の位置関係からも位置情報が取得可能になった。これら移動媒体から得られた所有者の位置情報が秒単位でデータ化されることで膨大なデータを形成することになる。GPS や基地局から推定される所有者の位置情報を記録したデータは、近年注目を浴びているビッグデータとしてさまざまな場面に活用の方向性が期待されている。例えば、本田技研工業から発売された自動車に搭載されたカーナビゲーションは、GPS データを常に通信によって蓄積し、VICS 情報だけでなく、個々の自動車の動きから渋滞状況を予測し、その結果をタイムリーにフィードバックするシステムが構築されている。これらのビッグデータは、2011 年に発生した東日本大震災で通行不可能になった道路の情報を収集し、その分析結果を公表するといった災害時への活用も行われている。国土交通省においてもこれらの活用が議論されている。

一方、移動媒体といった普及率がさらに高い端末では、通信局からの電波受信の位置関係から犯罪捜査にも活用され、犯人検挙や誘拐被害者の保護などに役立っている。しかしながら、上述の自動車に搭載された GPS といった交通に特化したデータは活用範囲が極めて明確であるが、個人が所有する移動媒体から蓄積されたデータはあまりにも膨大であり、その取扱いが難しく活用の汎用性が見いだせていない現状がある。スマートフォンやタブレット端末といった高機能端末に搭載された GPS データからであれば、詳細な位置情報を知ることができるので、所有者がダウンロードしたアプリケーションから対応が可能であるため、さまざまなサービスの方向性が展

* 原稿受付 2015 年 2 月 27 日

^{*1} 産業ビジネス学科

^{*2} 経営情報学科

E-mail: d-nogiwa@fukui-ut.ac.jp, fujihara@fukui-ut.ac.jp

開されている。いずれにせよ、その活用性は限定的であり、得られた膨大なデータを活用するまでには実務上至っていないのが現状である。

そこで本研究では、移動媒体の通信から得られたビッグデータとりわけ、その所有者の移動状況が把握できるモビリティデータを用いて、個人の移動状況の変化ならびにパターンを統計的に分析することにより、新たな知見を得ること目的としている。試みとしては、第一にモビリティデータが GPS 搭載端末だけではなく、通信から得られた比較的詳細ではないが位置情報を持つ移動媒体すべてを使用したこと。さらに、これらのモビリティデータ解析に統計モデルを導入したこと。これらの二点が大きく挙げられる。前者については、いわゆるビッグデータとして近年注目される領域である。データ量が膨大であるため取扱いが難しく、ランダムサンプリングや対象者を絞り込む形で分析を進めることが一般的である。しかしながら、本研究では、得られたすべてのデータを用い、さらに精度の高い研究に対応している。ビッグデータ分析は、汎用計算機やスプレッドシートタイプのソフトウェアによる分析（単純集計を含む）は困難であり、一定のプログラミング技術が必要となる。それ故に、実務において、ビッグデータ分析の実践は難しいといわれている。さらに学術的な観点からもサンプリングすることなくすべての情報を活用できるメリットがある。本研究では大規模データの取扱いが容易であり、大企業におけるデータベース管理にも活用されている SAS システムを導入し、分析を進めることとした。後者については、統計学や計量経済学を用いた社会科学におけるデータ解析では比較的盛んに行われている一方、交通工学や行動科学では多く応用がされていないモデルを用いることで、外生的な効果を計測できる利点を主張するものとした。

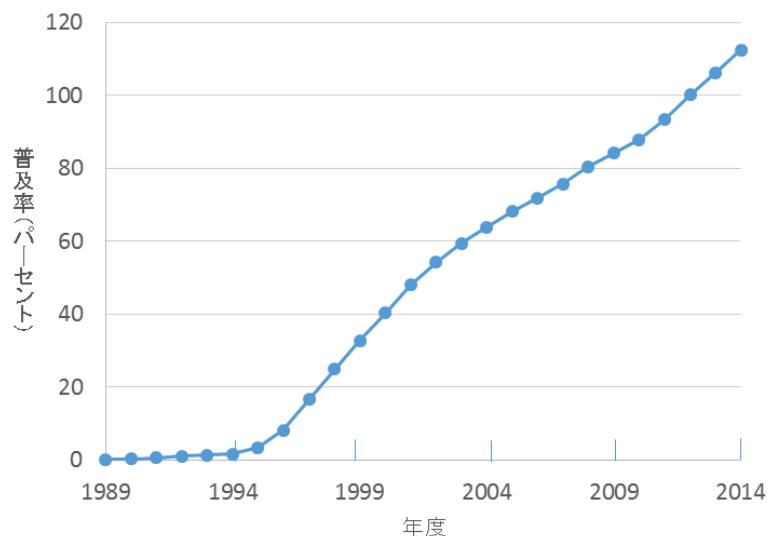


Fig. 1 日本における携帯電話の普及率（総務省 電気通信サービスの加入契約数等の状況調査⁽¹⁾）

2. 先行研究

2.1 モビリティデータを用いた先行研究

人間の移動パターン (Human Mobility Patterns) の統計的な性質に関する研究は、主に複雑系やネットワーク科学を研究する学者達により、携帯電話や GPS センサ等が普及した頃から活発に行われるようになってきた。人間の移動パターンの研究で最初に注目を集めたのは、Brockmann ら⁽²⁾による研究で、人間の旅行による移動距離の統計分布がベキ分布（両対数プロットすると、裾野が直線になる分布）に従うというものである。これは、アメリカの紙幣のシリアル番号がどの銀行にあるかを記録したデータベースを利用して解析した結果から導き出された。この分布は後に行われた GPS センサや携帯電話の通話履歴（Call Detailed Record (CDR)）を利用した解析^(3,4)でも同様の結果が得られている。ただし、ベキ分布に従うのはあくまで沢山の人のデータを一緒に集めてとった分布のみであって、一人ひとりのデータの移動距離分布ではベキ分布に従わないことも知られている⁽⁴⁾。また一般に、人の移動の時間経過（トレース・データ）を抽出して、移動に関するエントロピーを解析した結果、その

値は非常に低い値を取ることが分かっている⁽⁵⁾。これは、人間の移動はランダム・ウォークのように全ての場所をくまなく移動するのではなく、非常に偏った動きをすることを意味する。これは人間の移動パターンが非常に計画的かつ規則的になっていることを示している。この結果を逆に解釈すると、人間の行動パターンは非常に予測しやすいということになる。沢山の人の大量なトレース・データを数点抽出することができるだけで、それがどの人のトレース・データであるかを高い確率で当てることができるという研究も報告されている⁽⁶⁾。また、人間が繰り返し訪問する場所の訪問回数もベキ分布に従うことが知られている⁽⁷⁾。ベキ分布は所得分布がそれに従うことが知られているが、このことから非常に格差のある分布といえる。訪問回数がベキ分布に従うことから、人が移動により訪れる場所については非常に格差が大きくなるということが出来る。これらの高頻度に移動する場所を点にして、その間の移動があった点のペアを線で結んだグラフのことを人の移動モチーフ (Human Mobility Motif) と呼ぶが、このグラフ構造も取り得るグラフ構造全体のうちの非常に少ないパターンしか起こりえないことが報告されている⁽⁸⁾。

また、携帯電話の通話履歴は、地域同士でのコミュニケーションの強さに関する性質についても調べることができる。CDR データの解析により、携帯電話での累積通話時間と通話があった二点間の距離にベキ分布が成立することが分かっている。このベキ指数が2に従うという報告もある。これは、万有引力の法則（力が距離の2乗に反比例する物理法則）と同じになることから、コミュニケーションの重力法則 (gravity law) と呼ばれたりする⁽⁹⁾。また、通話履歴の時間ゆらぎ（移動平均を取り除いた変動値）を見ると、災害等の異常が起こっていることが分かることから、異常検出への応用が期待できることも知られている。

人間の移動パターンから人間同士の遭遇に関する統計的な性質について解析する研究も行われている。以前に出会ってから次に出会うまでの時間間隔のことを遭遇間隔時間 (intercontact time) と呼ぶ。この統計分布はベキ分布に従うことが報告されている。また、Fujihara・Miwa⁽¹⁰⁾は、Bluetooth や Wi-Fi といた近距離無線通信技術を用いた実験を行い、人間の遭遇特性のデータを収集し、人間同士の遭遇頻度分布がベキ分布に従うことを報告している。また、遭遇の大多数が一期一会（実験期間中で一度しか出会うことができない）の遭遇であったことも報告している⁽⁸⁾。人間同士の遭遇はウィルスの伝播の過程から感染経路を特定したり、ウィルス拡散の収束時期を予測するなどの応用が期待されている。また、遅延耐性ネットワークやモバイル日和見型ネットワークといった、人間同士の近接遭遇を利用した情報転送を行うことで、インターネットを使える環境を広げていくという発想の次世代ユビキタス情報通信技術の通信性能評価にも役立つと考えられる⁽¹¹⁾。

以上のように、人間の移動や遭遇に関する統計的な性質にはベキ分布が出てくることが多い。また、ベキ分布は格差がある分布であることから、その時空間パターンはランダムではなく、パターンが存在することがデータによって実証されつつある。このことから人間の行動パターンは予測しやすいため、行動パターンの監視に利用することもできる。しかし、人間の行動パターンは非常にプライバシーに関わる内容であるため、ヒューマン・モビリティの研究はプライバシーに配慮しながら、今後も展開されていくことが予想される。

2.2 モビリティデータと実務に関する先行研究

モビリティデータやこれらに関連して蓄積されたビッグデータは、先述した交通分野だけではなく、さらに実務的な活用のための研究がなされている。とりわけ、福岡大学都市空間情報研究所では、博多駅や天神などの福岡市内に関する個人の回遊ならびに商業施設へのマーケティング施策を提案している。齋藤・梶井・中嶋⁽¹²⁾、斎藤・山城・中嶋⁽¹³⁾では、福岡市内の商業施設に関わる顧客の回遊パターンに対してモビリティデータを用いて推定する研究が行われている。これらの研究は、GPS のモビリティデータから顧客の回遊パターンを推定する試みであり、とりわけ位置状態の遷移に注目した研究である。これらはモビリティデータから都市や地域単位での顧客の回遊を移動媒体から推測し、その結果をマーケティングの実務あるいは地域活性化に役立てるための研究といえよう。実務では、回遊状況を加味したクーポン券の提案や今後の買い回りの予測について対応する試みがなされている。一方、長尾・川村・山本・大内⁽¹²⁾では、GPS でのモビリティデータを用いた観光を目的とした回遊の可能性について議論している。この研究では、北海道における観光客の移動についての研究であり、観光活性化といった目的で当該データの可能性について議論している。実際観光地などでもスマートフォンのアプリケーションを用いて観光客への案内や情報を発信する試みがなされている。

以上のように、実務的な側面に対応した研究は行われているが、移動媒体全データではなく精度の高いGPSデータに基づいた対応がされていることが挙げられる。これらは、詳細なデータが取得できる反面、スマートフォンやタブレット端末の所有者の年齢層やGPSアプリケーション利用に対する態度によって、対象者が限定されてしまう恐れがある。また、これらの端末に利用制限がかけられているビジネスマンなどの行動を除かれてしまう可能性もあると考えられる。その点で詳細なデータは取得できないものの、基地局からの受信でエリアを特定したデータはより多くの可能性をもち、そのデータ全体を網羅し、サンプリングをしないビッグデータで分析することは有益であると考えられる。

3. モビリティデータについて

3.1 データの概要

本研究で用いるデータは、Data for Development (D4D) Challenge Senegal⁽¹⁴⁾において提供された高解像度な人の移動のトレース・データを用いた。これは、2013年の一年間のセネガル全体の人の移動を最寄りの電波塔IDの時系列変化データとして抽出されたものである。このデータはOrangeというフランス系の移動体通信会社の系列で、セネガルでの移動体通信プロバイダであるSonatelが独自のプライバシー保護をしたフォーマットにより提供するものである。トレース・データの数320,000人分に達している。ただし、各トレース・データは二週間ずつに区切られており、二週間以上にわたった人の移動パターンは解析できない。電波塔IDの緯度・経度情報は対応関係が別データにより分かっているために、人間がどの電波塔の近くにいるかという時系列が分かる。セネガルの首都のダカールには、セネガル全体の約1/3の電波塔があるため、誤差が数kmで人間のトレース・データが取れることになる。また、セネガルの日時降雨量に関するデータもD4D Challenge Senegalから提供されたものを用いた。ダカールではYoffという観測所を代表点として、その場所の降雨量が2013年の一年間与えられたデータを用いた。セネガルの気候は主に雨季(7~10月)と乾季(それ以外の月)に分かれる。雨や降雨量の影響を見るために、主に雨季のトレース・データを活用して研究を行った。また、人の移動を観察するために、ダカールにある402個の電波塔の位置情報を用いた。

3.2 基礎集計

当該データは2013年1月1日~2013年12月16日の期間で取得したデータである。前節の説明のとおり、移動媒体からデータならびに音声通信をした時刻と基地局番号がIDごとに記録されたデータである(Table 1)。すべてのレコード数は1,204,451,385であり、データ量もテキストベースで34.4Gになるビッグデータである。個人情報保護の観点から2週間ごとにID番号が振りなおされており、すべての具体的な人数はデータから特定することはできない。日別に個人別集計をしても、延べ102,185,425人分の個人別移動情報が記録されている。

これらの延べ人数を対象に個人の行動に関して基礎集計する。まず、個人の行動パターンを把握するために1日あたりの回遊傾向を分布で確認する。Fig. 2から1日あたり約40%人は回遊エリアが1か所に留まっていることが確認できる。2か所も25%、3か所も15%となり、約8割の人が1日に回遊3か所以下のエリアで活動していることが確認できた。また、連続した複数時点で同じ基地局エリアに存在する際には、その通信時刻の差分から滞在時間を予測することができる。実際に1日1基地局あたりの集計を個人別に行い、その結果をFig. 3ならびにFig. 4に示した。最大滞在時間(秒)ならびに平均滞在時間(秒)の分布は両方ともに0~10,000秒までの頻度が高く、次いで80,000秒以上が多くなっており、滞在時間が短い人と長い人の分布が分かれる結果となった。

Table 1 モビリティデータの構造

ID 番号	日付	基地局番号
1	2013/1/7 13:10:00	231
1	2013/1/7 17:20:00	231
1	2013/1/7 18:40:00	312
1	2013/1/8 10:50:00	231
2	2013/1/7 8:20:00	124
2	2013/1/7 9:10:00	124

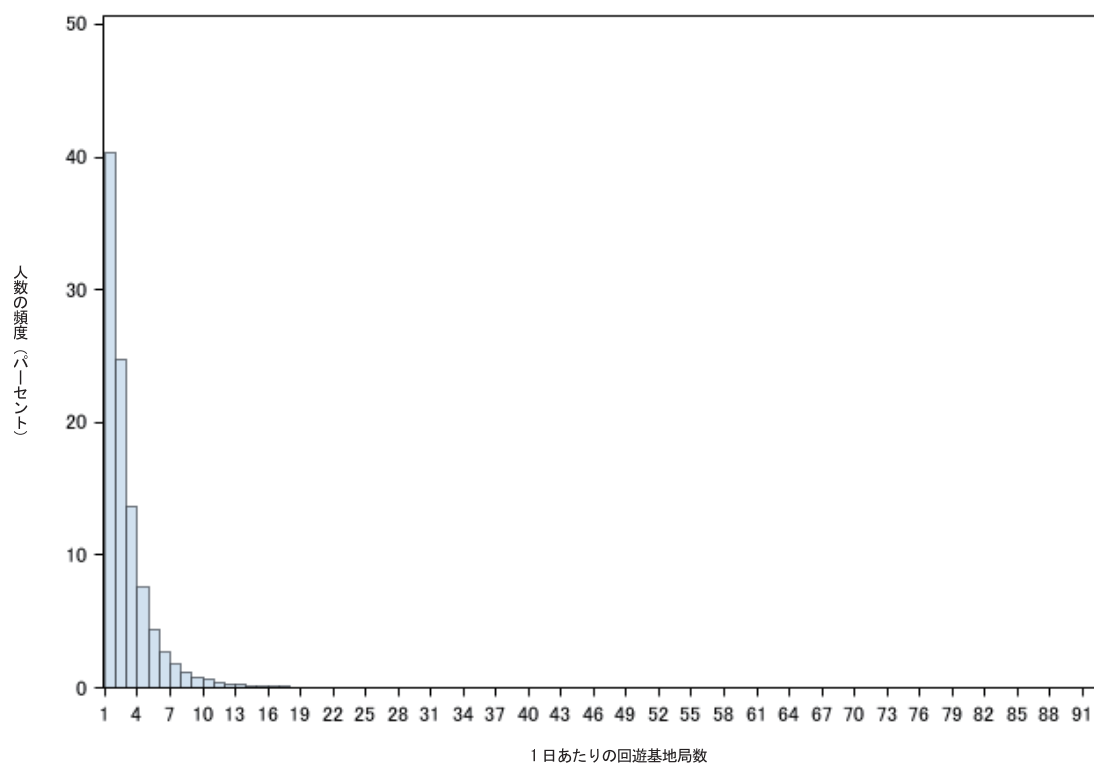


Fig. 2 1日あたりの個人別回遊基地局数の分布

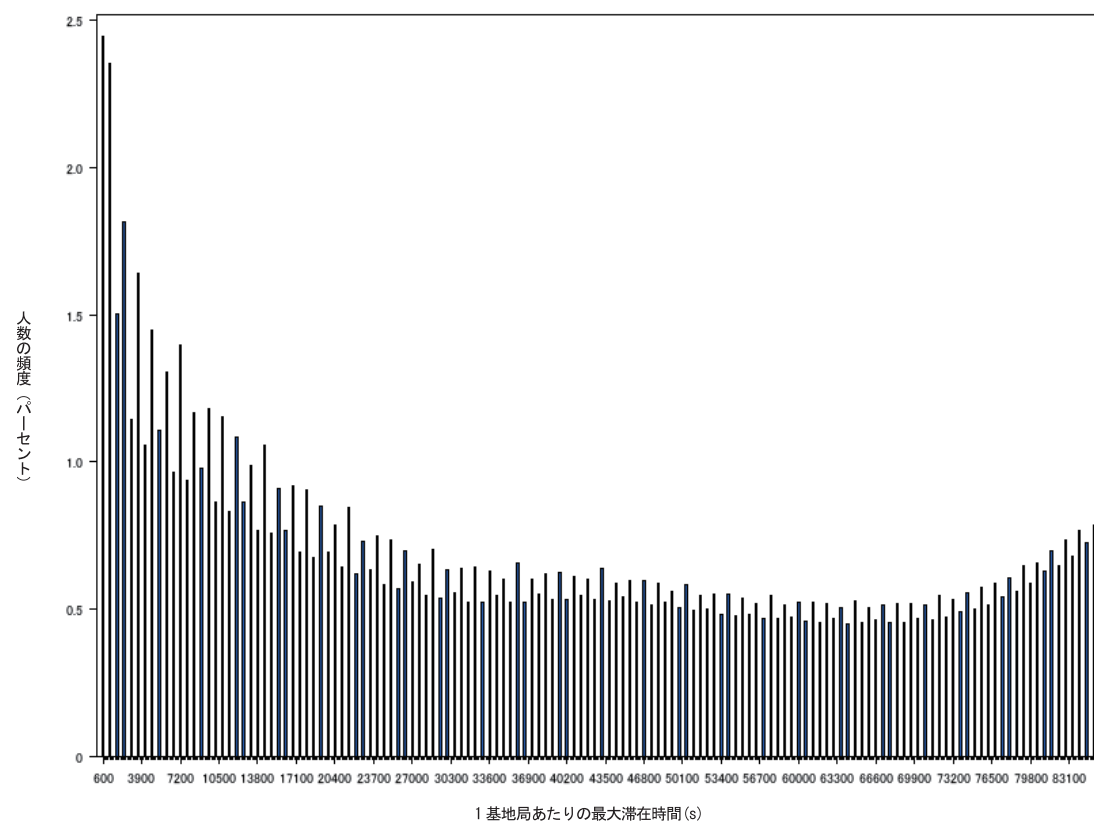


Fig. 3 1日1基地局あたりの個人別最大滞在時間 (秒) の分布

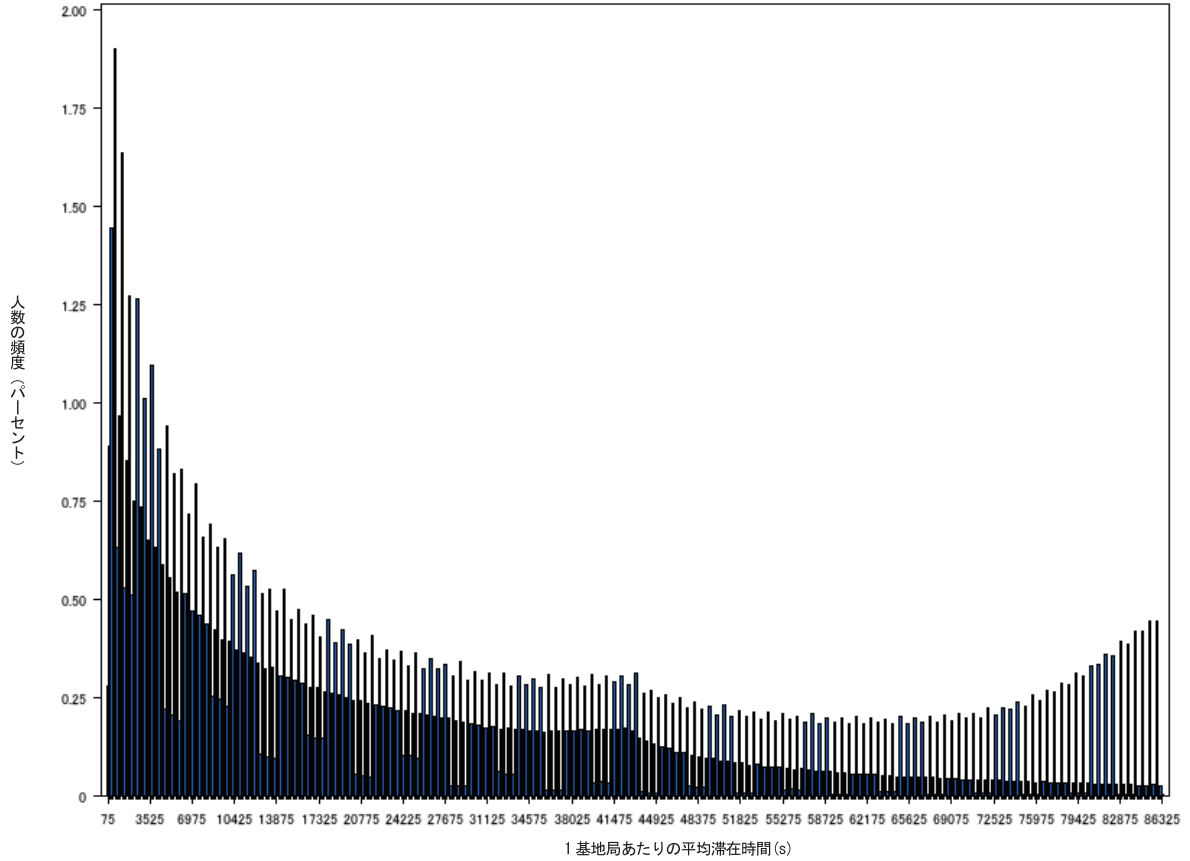


Fig. 4 1 日 1 基地局あたりの個人別平均滞在時間の分布

4. 統計モデルとの推定結果

4.1 統計モデルについて

本研究では、移動媒体から得られた基地局があるエリアへの集積状況を統計モデルで記述する．具体的には、基地局があるエリアの人の集積が外生的な要因によってどのように変化するのかを確認するモデルである．当該分析では、とりわけ人間の行動に影響されるであろう天候要因ならびに季節要因、そして休祝日の3つの要因について分析を行う．これらの要因を選んだ理由としては、先行研究で述べられた実務的な側面を鑑みた結果、最も知るべき情報であり、これらの外生的な要因は行動パターンに影響する要因であると考えられる．これらの影響度を知ることによって交通渋滞の予測や公共施設の拡充だけでなく、商業面からでもさまざまな施策を提案可能になるとと思われる．

ここで t 時点の i エリアに関する 1 日あたりの集積人数を N_{it} とし、そのときの一人あたりの滞在時間を V_{it} とする．このとき、集積人数に関する統計モデルは

$$\log(N_{it}) = \beta_0 + \beta_1 RAIN_{it} + \beta_2 LRAIN_{it} + \beta_3 HOLIDAY_{it} + \beta_4 SEASON_{it} + \varepsilon_{it} \quad (1)$$

また、滞在時間に関するモデルは

$$\log(V_{it}) = \beta_0 + \beta_1 RAIN_{it} + \beta_2 LRAIN_{it} + \beta_3 HOLIDAY_{it} + \beta_4 SEASON_{it} + \varepsilon_{it} \quad (2)$$

として表すこととする．ただし、 $\beta_k (k = 0, \dots, 4)$ は推定パラメータであり、 $\varepsilon_{it} \sim N(0, \sigma^2)$ と仮定する．また、外生変数は以下のとおりである．

$RAIN_{it}$: t 時点の i エリアが雨天のダミー変数

$LRAIN_{it}$: $t-1$ 時点の i エリアが雨天のダミー変数

$HOLIDAY_{it}$: t 時点の i エリアが休祝日のダミー変数

$SEASON_{it}$: t 時点の i エリアが雨季のダミー変数

また、(1)と(2)に対して、エリアは基地局であるゆえ、その位置関係や大きさ、施設配置などもさまざまであることが考えられる。上記外生変数の影響度を得るためには、これらの変動を加味する必要があると考える。しかしながら、基地局の大きさや施設配置などさまざまなノイズが考えられ、この状況をすべて具体的に記述するモデルは複雑であり、実用的でもないと考えられる。そこで、このエリアがもつ特性を確率分布として捉え、以下のモデルを提案モデルとする。

$$\log(N_{it}) = \alpha_i + \beta_0 + \beta_1 RAIN_{it} + \beta_2 LRAIN_{it} + \beta_3 HOLIDAY_{it} + \beta_4 SEASON_{it} + \varepsilon_{it} \quad (3)$$

$$\log(V_{it}) = \alpha_i + \beta_0 + \beta_1 RAIN_{it} + \beta_2 LRAIN_{it} + \beta_3 HOLIDAY_{it} + \beta_4 SEASON_{it} + \varepsilon_{it} \quad (4)$$

ここで、 $\alpha_i \sim N(0, \sigma_\alpha^2)$ として確率分布を仮定する。提案モデルは一般的に統計学ではランダム効果モデルや混合モデルと呼ばれるモデルである。エリアごとを確率的な変動ととらえた所以は、エリアの数が著しく多く、これらすべてに対してエリアダミー変数を用意し、分析すると説明変数の数が膨大になり、統計的実用的な面からも課題があると考えられる。このような集積エリアのノイズを確率的に表現し、簡易的なモデルを構築した先行研究はなく、新規性のあるモデルであることが挙げられる。加えて、ビッグデータ分析としての試みも先行研究にはなく、これだけ膨大なモビリティデータを生かした分析をする事例もない。一方で、都道府県や国などを対象にした都市経済学や地域経済学の分野でこのようなモデルでの研究は盛んに行われており、これらの分析手法として計量経済学におけるパネルデータ分析が有名である。

4.2 推定結果

前節で提案したモデル(1)～(4)を本研究で対象とするモビリティデータを用いて推定をする。これらの推定には SAS システムのプロシジャを活用し、推定方法は最尤推定とした。また、(1)と(3)ならびに(2)と(4)のモデルの比較検討は、AIC (Akaike's Information Criterion : 赤池情報量規準) ならびに BIC (Bayesian Information Criterion : ベイズ情報量規準) の 2 つの情報量規準を用い、その値が小さい方をモデルの適合度が高いと判断し分析を進める。また、変数間の相関を調べるために Table 2 に変数間の相関係数行列を示した。各変数ともに大きな相関は見られず独立性が担保されていると考えられる。

Table 3 では、(1)と(3)ならびに(2)と(4)ごとに、AIC ならびに BIC それぞれの情報量規準でランダム効果を導入したモデルが比較検討結果妥当であると考えられる。また、パラメータの推定結果も符号の逆など両者大きく異なった結果となっておらず、推定上の欠陥はないと思われる。それゆえ、推定結果の解釈は(3)ならびに(4)に基づき進める。

(3)の集積人数を説明するモデルでは、雨天ダミーの推定値以外すべて有意水準 95%で 0 を棄却した。前日が雨天である場合にはエリアの集積人数が増加する傾向にあるということがわかった。また、祝日も同様な結果であった。一方、雨季の状況下では、集積人数を減らす傾向にあり、雨季は回遊が少なくなるということがわかった。(4)の滞在時間を説明するモデルでは、すべて有意水準 95%で 0 を棄却した。雨天時であると 1 エリアに滞在する時間が長くなる傾向にあるが、前日が雨である場合にはその逆で滞在時間が短く活動的になり、(3)のモデルとも整合性がとれているように思われる。また、休祝日は滞在時間が増える傾向にあり、自宅などで過ごす人が多くなると思われる。最後に雨季も同様に滞在時間が増える傾向にあった。とりわけセネガルでは、土地の水はけが悪く、雨季に洪水になることが多く、その要因である可能性が考えられる。

5. 考察とまとめ

本研究では、モビリティデータで個人の移動に関する要因を探る分析手法を提案し、セネガルで収集された人の移動に関する膨大なデータに適応し、別に行った基礎集計結果と比較することで、セネガル国内における住民の移動に関する特徴の抽出を試みた。ビッグデータはさまざまな分野で応用が期待されているが、それゆえハン

ドリリングが難しく、モデル構築以前の集計も困難である状況であった。これらを鑑み、複雑な手法ではなく、比較的簡便で実務に耐えうるランダム効果を導入した統計モデルを提案し、その実証結果を示した。実証結果についても、雨天や休祝日、季節性のダミー変数など比較的容易に準備できる変数を導入し、その影響度を確認した。この条件下での人の行動パターンの分析には当該分析手法の枠組みが有効であると考えられる。

また、本研究をより発展的にすることも可能であると思われる。具体的には、集積エリアへの影響度合いを人の行動パターンの違いによって、天候、休祝日、季節の影響度知ることができるモデルの構築も可能である。例えば、1日に回遊するエリアの数の多い人は雨の影響度得やすいといったモデルである。このモデルは当該ランダム効果モデルをさらに発展させ、階層モデルとして記述することが求められる。外生的な影響と人のパターンを結びつけることで、よりきめ細やかな傾向を洗い出すことができる。しかしながら、ビッグデータの領域である故計算量が膨大であり、これらを解決する課題があると言えよう。近年、統計学ではコンピュータの目覚ましい発展から計算機統計学の分野の研究が盛んに行われている。とりわけ、ベイズ統計学は事前分布と事後分布の評価からパラメータを推定する枠組みである。当該発展モデルを想定するにはベイズ統計学の分野を応用する必要がある。ただし、サンプリングに時間を要したため、本研究では取り上げなかった。今後、計算時間の短縮を加味して更なる実用に耐えうるモデリングの提案を行っていきたいと考える。

Table 2 変数間の相関係数行列

	<i>RAIN</i>	<i>LRAIN</i>	<i>HOLIDAY</i>	<i>SEASON</i>
<i>RAIN</i>	1	0.467	-0.003	0.455
<i>LRAIN</i>	0.467	1	0.029	0.455
<i>HOLIDAY</i>	-0.003	0.029	1	-0.021
<i>SEASON</i>	0.455	0.455	-0.021	1

Table 3 モデルの推定結果（最尤法）

	変数	β	推定値	標準誤差	P値	AIC
(1)	切片	0	5.5923	0.0024	0.0000	1809989
	雨天ダミー	1	0.0004	0.0075	0.9580	
	前日雨天ダミー	2	0.0073	0.0075	0.3296	BIC
	休祝日	3	0.0166	0.0039	0.0000	1810000
	雨季	4	-0.0636	0.0048	0.0000	
(2)	変数	β	推定値	標準誤差	P値	AIC
	切片	0	10.2900	0.0013	0.0000	1155138
	雨天ダミー	1	0.0555	0.0041	0.0000	
	前日雨天ダミー	2	-0.0823	0.0041	0.0000	BIC
	休祝日	3	0.0743	0.0021	0.0000	1155149
	雨季	4	0.1279	0.0026	0.0000	
(3)	変数	β	推定値	標準誤差	P値	AIC
	切片	0	5.5454	0.0326	0.0000	315694
	雨天ダミー	1	0.0038	0.0018	0.0349	
	前日雨天ダミー	2	0.0129	0.0018	0.0000	BIC
	休祝日	3	0.0165	0.0009	0.0000	315705
	雨季	4	-0.0647	0.0012	0.0000	
(4)	変数	β	推定値	標準誤差	P値	AIC
	切片	0	10.2773	0.0152	0.0000	503229
	雨天ダミー	1	0.0558	0.0022	0.0000	
	前日雨天ダミー	2	-0.0820	0.0022	0.0000	BIC
	休祝日	3	0.0742	0.0011	0.0000	503240
	雨季	4	0.1289	0.0014	0.0000	

文 献

- (1) 総務省 通信利用動向調査 平成 26 年総務省 電気通信サービスの加入契約数等の状況調査
- (2) Brockmann, D. et al., “The scaling laws of human travel,” *Nature* 439, 462-465 (2006).
- (3) Rhee, I. et al. “On the Levy-walk Nature of Human Mobility: Do Humans Walk like Monkey?” The 27th IEEE International Conference on Computer Communications (IEEE INFOCOM 2008), 924-932 (2008).
- (4) González, M. C. et al., “Understanding individual human mobility patterns,” *Nature* 453, 779-782 (2008).
- (5) C. Song, et al., “Limits of Predictability in Human Mobility,” *Science*, vol. 327, no. 5968, pp.1018-1021 (2010).
- (6) Y.-A. de Montjoye, et al., “Unique in the Crowd: The privacy bounds of human mobility,” *Scientific Reports*, vol. 3, no. 1376 (2013).
- (7) C. Song, et al., “Modelling the scaling properties of human mobility,” *Nature Physics* (Advanced Online Publications) 7, 713 (2010).
- (8) C. M. Schneider, V. Belik, T. Couronné, Z. Smoreda, and M.C. González, “Unravelling daily human mobility motifs,” *Journal of The Royal Society Interface*, Vol. 10 (2013).
- (9) V. D. Blondel, A. Decuyper, G. Krings, “A survey of results on mobile phone datasets analysis,” <http://arxiv.org/abs/1502.03406>
- (10) A. Fujihara and H. Miwa, “Homesick Levy Walk: A Mobility Model Having Ichi-Go Ichi-e and Scale-Free Properties of Human Encounters,” *IEEE COMPSAC 2014*, pp.576-583 (2014).
- (11) 斎藤参郎, 梶井昌邦, 中嶋貴昭, “都心商業空間における商業施設への消費者来街者数と回遊パターンの同時推定逆問題について” *地域学研究* Vol. 30 No.1 (1999), pp. 213-229
- (12) 斎藤参郎, 山城興介, 中嶋貴昭 “隠れマルコフモデルによる地区間歩行者流動パターンの動的推定-福岡市都心部での大名・国体道路・今泉地区への適用” *地域学研究* Vol. 34 No.1 (2003), pp. 69-90
- (13) 長尾光悦, 川村秀憲, 山本雅人, 大内東, “GPS ログからの周遊型観光行動情報の抽出”, 通信学会技術研究報告, 人工知能と知識処理 Vol. 105 No.224 (2005), pp. 23-28
- (14) D4D Challenge Senegal, <http://www.d4d.orange.com/en/presentation/data> (参照日 2015 年 2 月 27 日)

(平成 27 年 3 月 31 日受理)