

論文の内容の要旨

題 目： 人の移動履歴のデータマイニングのための確率モデルの提案とその応用

氏 名： 浅 原 彰 規

近年のモバイルコンピュータ（容易に持ち運び可能な小型コンピュータ）分野の発展は著しいが、このモバイルコンピュータの多くは、測位の機能を有している。さらにこれらに類似するものとして、建物や道路などに設置された機器を用いてそこにいる人や車等の存在を検知する技術がある。これらの測位を利用したサービスは、近年になってさまざまな環境において用いられるようになってきており、人の位置に関する情報が社会に溢れている。

これらの人の位置に関する情報を活用すると、人の移動行動の調査の補完に用いることができると想定される。しかし、現時点で実用になっている事例は多くはない。その原因の一つが、データの入手の困難さの問題である。そこで、調査目的で集められたデータではなく、人が普段の行動をする範囲で収集された位置のデータを副次的に活用することが考えられる。本論文では、このような他の目的で収集された位置の履歴のデータ群から有益な情報を抽出することを移動履歴のデータマイニングと呼ぶ。

移動履歴のデータマイニングは、その特性の差異から、従来の調査とは異なった用途が考えられる。人の移動行動の調査が求められる用途の例として、交通インフラ整備のための交通需要の地点からある地点へ移動したいという需要）の調査を例とする。交通需要調査では、まず四段階推定法などの人の交通行動のモデルを設定し、そのモデルのパラメータを決めるために、パーソントリップ調査のアンケートなどの入手による調査が行われる。これによって人の交通行動のモデルが定まり、新設の道路の需要を求めることができる。このような調査で得られるデータは網羅的、精緻であるが、大きなコストや調査時間がかかる。よって、従来の調査は、道路建設の決定などの十年、二十年という長期的な意思決定には適するが、毎月、あるいは日々の変動に対応せねばならないケース、たとえばバスの臨時便追加などに関するケースには適さない。他方、移動履歴のデータマイニングでは、日々副産物として集まる情報を用いるので、コストはかなり抑えられ、データの鮮度も高い。その反面、データの量そのものは多くとも、網羅性については従来の調査よりも低いと考えられる。

移動履歴のデータマイニングは、従来の調査とは異なった性質をもつので、それに適した分析手法が新たに必要となる。網羅性が低いデータから、短期間で、できるだけ有益な情報を多数抽出しなければならないのである。交通需要調査では各個人に対して交通行動の出発地、目的地、交通手段などが得られることを前提にしているが、移動履歴データではその前提は満たされず、位置情報のみしか得られないことを想定せねばならない。また、目視確認により情報を補完していくようなことも、日々の運用に供することを想定すれば、現実的ではない。つまり、得られた網羅性の低い位置情報だけから、人の位置や移動の全体的な傾向を知るための様々な情報を短時間で自動的に抽出することが必要になるのである。

移動履歴のデータマイニングは、様々な用途が想定されるため、その処理技術も複数想定される。本論文では、この処理技術を要約、推定・予測の2つに分類する。要約とは、人が概況を把握するた

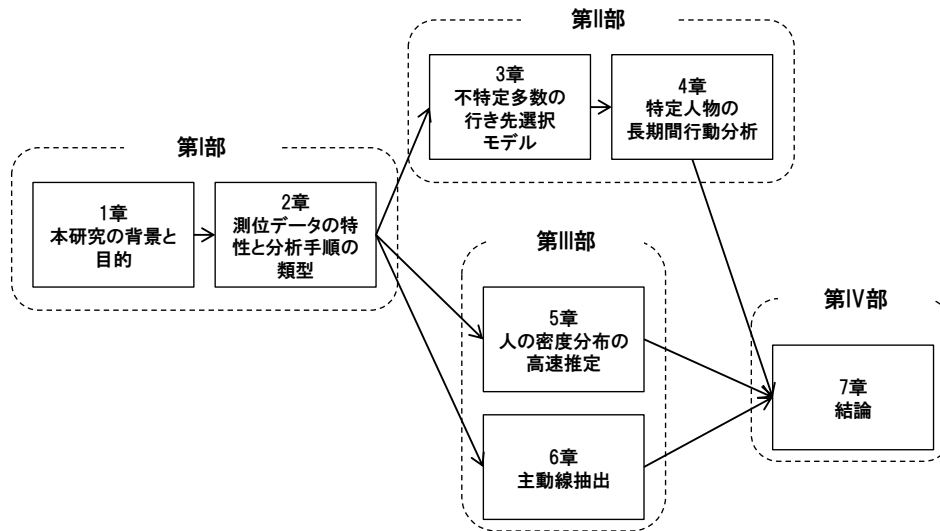


図 1 本論文の構成

めの技術であり、特徴の抽出や異常の検知、可視化などが含まれる。推定・予測とは現状知られていない事実を導き出す技術であり、人の行き先の予測や、移動のパターンにもとづく類似の行動をするグループの検出などが含まれる。意思決定には様々な事情を勘案しなければならないため、抽出される情報は多様なほうがよいが、多すぎると逆に把握が難しくなる。よって抽出情報の観点を絞り、必要な情報を取捨選択できるよう、分析のメニューを揃えていくことが求められる。

ただし、人の行動のように複雑で非決定論的な事象について、それを完全に予測できるような数理モデルをたてることは困難である。そのような場合、経験的な確からしさを定量的に評価する確率モデルが用いられる。本研究でも、確率モデルを用いて人の移動行動をより精緻にモデル化することを検討する。確率モデルは予測に用いることができるが、それだけでなく、確率モデルの中に内部的な状態を含むことにより、推定や要約にも適用できるという利点がある。この観点で先行研究で不足するものの一つが、不特定多数への対応性である。従来研究のうち、指標算出以外の多くの手法は、ある特定個人が十分長期間動線データを計測しつづけることを前提にしており、事前準備されていない不特定多数のデータに対する処理に適した手法はまだ不十分である。

そこで本研究は、蓄積された位置情報だけから情報を抽出するために、確率モデルにもとづく種々の推定・予測と要約の手法を提案して拡充することを目指した。人の移動を扱う確率モデルとして最も基本的なものとしては、マルコフモデルが想定される。これは、人の行動は非決定論的である状態でどう行動するかは一意には定まらないものの、その選択確率だけは定められると仮定したことに相当する。実際には不特定多数の人の行動を一つのマルコフモデルに当てはめられないと想定されるがマルコフモデルは単純ながら拡張性が高いモデルであり、人の行動選択の条件は拡張できる。

以上より、本論文では特にマルコフモデル的な確率モデルに着目し、その拡張によって種々の推定・予測、要約に適用する方法を提案することで、抽出される情報の拡充を目指す。図 1 に本論文の構成を示す。

第 I 部では、この研究の前提条件や目的、定義、問題設定、本研究のアプローチについて述べる。

まず、「第 1 章 本研究の背景と目的」では本研究の背景と位置づけについて概略を示す。「第 2 章 測位データの特性と分析手順の類型」では、測位に関して整理するとともに、分析方式の概略と要素について整頓する。第 I 部では、人の移動にまつわる意思決定に必要な基本的な情報として、対象地域内の移動する人たちに関し、どんな人たちが、いつ、どこにいて、どこへ、どう移動しようとしているかが重要であることを述べている。

次に、第 II 部では、人の行き先を他の人の移動履歴を用いて予測する確率モデルについて第 2 章で規定した手順にもとづき、事例を交えながら議論した。「第 3 章 不特定多数の人の行き先選択モデル」と「第 4 章 特定人物の長期間行動分析」は、予測推定の手法について、主にマルコフ連鎖の確率モデルを用いた方法について述べる。第 II 部では、人の個性をうまく確率モデルとして扱うことが精度向上の鍵になっていたと思われ、時間が経つにつれて個性が顕在化するため、時間変化を扱うモデルが高い性能を発揮していた。つまり、どんな人たちにがいるかに着目した分析であり、移動のパターンや個人ごとの行き先予測という形の情報が得られた。

そこで第 III 部の「第 5 章 人の密度分布の高速推定」と「第 6 章 主動線抽出」では、きの全体像を可視化するため、人の密度分布計算について直前の分布を用い高速化する方法と、マルコフ連鎖的なモデルを適用し主動線を抽出する方法を提案する。ここでは人の個性の情報は集約し、他の情報に焦点を当て可視化している。どこに人がいるかを時間ごとに可視化したのがヒートマップであり、どこへどうやって移動するかを可視化したのが主動線であった。

最後に、「第 7 章 結論」では、全体について俯瞰し、今後の展望について示す。第 II 部および第 III 部の結果を合わせると、マルコフ連鎖的なモデルによって第 I 部で示された人の移動に関する基本的な情報の抽出技法がおおむね網羅されたことになる。したがって、人の移動の平均にマルコフ連鎖的なモデルを適用することによる情報抽出は妥当であり、意思決定のための状況把握がしやすくなったと考えられる。