

論文審査の結果の要旨

氏名 新井 亜弓

人口やその動態、日常的な移動、個人属性や世帯属性などに関する情報は様々な分野で不可欠である。人口政策や都市政策に加え、交通政策、健康政策、伝染病への対応、災害対応からマーケティング、イベントの管理、セキュリティなど多くの利用分野がある。

居住地ベースの人口や世帯構成などについては伝統的に国勢調査（センサス）等により収集・更新され、一方交通や健康、マーケティング等については、目的に特化した訪問調査・アンケート調査が定期的に行われてきている。災害対応やイベント対応などについて、迅速なデータ収集が必要なため、現場からの報告やセンサデータ等を集約することが多い。センサスや訪問調査・アンケート調査等については先進国では実施の体制が確立され、予算・人的資源等も継続的に投入される結果、3年や5年に1回という頻度ではあるが、ほぼ着実に実施されている。また災害時等における人々の分布や避難状況を把握する対応体制も整備されている。しかし開発途上国においては、予算や人的資源等の制約により定期的なセンサスでさえ十分実施できず、様々な政策や計画策定に必要な情報が得られないことも少なくない。また、国によってはスラムなどに居住する多数の「不法住民」が正式な政府統計ではカバーされていないこともあり、情報の歪みが生じている。

一方、携帯電話は全世界に普及し、利用台数は70億台に達している。ほとんど全ての国において大人はほぼ1台以上を所有しており、そこから得られるデータを人口推定や人々の移動状況の把握などに使おうとする試みが始まっている。技術的には人口やその動態などの比較的長期間の変化情報に加え、毎日の人々の流動、災害時の避難状況など短期的に大きく変動する情報をも得ることが可能であると期待されている。しかしながら、利用者あるいは携帯電話端末そのものの数や分布、あるいは移動の推定を行っている研究はあるものの、地域の全体人口やその属性についての推定を試みた研究はない。そのため、携帯電話データから得られる情報が、センサス等から得られる人口統計情報をどの程度補完できるのか明らかになっておらず、その利用はまだ政府等の統計調査システムにきちんと位置づけられたものとはなっていない。

本研究は、携帯電話から得られるデータのうち、CDR（Call Detail Record：通話詳細記録）に着目し、そのデータを用いて総人口や居住地・就業地推定、世帯構成、属性分布、さらに時間単位の人々の移動状況等を推定する手法を開発するものである。本研究では移動等も含んだ人口・世帯情報をダイナミックセンサスと呼び、ダイナミックセンサスをCDRデータと現地調査データの組み合わせにより推定する手法を開発した。さらにバングラデシュ・ダッカ市周辺のCDRデータを利用してダイナミックセンサスが作成可能であることを示した。本論文は全8章からなっている。

第1章は研究の背景と目的であり、世界規模での携帯電話の普及が開発途上国を中心に急激に進んでいること、そこから得られるデータの利用可能性が広がっているものの人口関連情報の推定や抽出などに適用した事例は少なく、まだまだ萌芽的な段階にあることなどが整理されている。こうした背景の下に研究の目的が定義されている。

第2章は研究の構成であり、CDRデータと現地調査を組み合わせることで新たな人口統計である「ダイナミックセンサス」を構築する手順が提案されている。以下の各章はこの流れに従って、技術的な課題について解決策が述べられている。

第3章はCDRデータの時空間補間である。CDRデータは通話やデータ通信等が行われた際の基地局位置や通話等の時刻・時間を記録しているが、バングラデシュのような2Gを主体とした携帯電話サービスにおいては通話等の回数

は一日数回と少なく、時間間隔が疎なデータである。また基地局は市街地中心部でも数百メートル間隔、郊外では数キロから10キロ程度の間隔で設置されているため、CDRデータは空間的にも疎なデータである。そこで、まず1ヶ月のデータを時間方向に重ね合わせ、居住地や就業地、その他の日常的な立ち寄り先を空間的に推定した。その上で1ヶ月の間に共通に現れるいくつかの活動パターン（居住地や就業地等で何時から何時までを過ごすのかを表すパターン）をトピックモデルの考え方を適用して抽出し、それらのパターンを毎日のCDRデータに当てはめて時間的な内挿を行った。これにより各日にちに対して居住地、就業地、その他立ち寄り先、移動などの活動ラベルが1時間ごとに内挿される。この内挿された時間別活動パターンを、さらに建物のタイプ別分布と道路ネットワークが表された地図データのうえにマッピングすることで空間補間を実現した。この際、基地局の受信範囲内にある建物を抜き出し、後述する世帯属性（推定値）と関連づけながら、居住建物を推定している。また移動に際しては受信状況を加味した経路内挿を実現し、より実現可能性の高い経路を推定している。

第4章はCDRデータからの個人属性・世帯属性の推定である。属性推定のためにダッカ市内を網羅したサンプル世帯を選定し、現地聞き取り調査を行った。これにより通話や移動の状況と個人属性・世帯属性を個人・世帯別に収集した。現地調査データの分析を通じて、Working male、Housewife、Students、Othersの4つの個人属性がCDRデータから推定しやすいこと、また世帯に小さな子供がいるかないかによって通話や移動の時空間パターンに差があることを明らかにした。その上で機械学習の手法（ランダム・ツリー）をCDRデータから抽出された特徴量に適用し、属性推定を実現した。

第5章は、CDRデータから得られた人口・世帯等の情報から全人口などを推定する手法の開発である。推定に際して、代表性ができるだけ損なわれないようにダッカの1地区を選び、人口・世帯構成、携帯電話の保有状況、所得水準、居住建物の属性などを網羅的に調査した。CDRデータは1携帯電話事業者（事業者Aと呼ぶ）の利用者から得られたものであるため、まず事業者Aの携帯電話の利用者ごとのCDRデータから、その利用者の属する世帯タイプを推定し、それを集計することで「事業者Aを含む世帯と総人口」を推定した。次に、事業者Aによる携帯電話利用者を含む世帯・人口と、ほかの事業者の携帯電話利用者を含む世帯、さらに携帯電話利用者のいない世帯の比を上記の現地調査から求め、それをういて拡大した。これにより携帯電話事業者AのCDRデータから全体の人口・世帯構成等を推定することができる。また、居住地が同じ人口・世帯は携帯電話事業者や携帯電話の保有によらず同じ移動パターンであると仮定して、全体人口・世帯の移動パターン等を推定した。これによりダイナミックセンサスが作成されたことになる。

第6章は、上記の方法によりダッカ市を含む地域でダイナミックセンサスを実際に作成し、視覚化等を行っている。

第7章は、各章の成果をまとめ、ダイナミックセンサスの作成手法全体のフローを整理している。

第8章は結論であり、本研究の結論と今後の展望を述べている。

以上をまとめると、開発途上国での適用を念頭に、携帯電話システムから得られる標準的なデータであるCDRデータを用いて人口や世帯構成から日常的な移動状況等に関する網羅的な情報（ダイナミックセンサス）を推定する方法を開発し、それが適用可能であることをバングラデシュにおいて実証した点が本研究の主要な成果であり、空間情報科学の発展に大きな貢献をしている。さらに論文の成果は柴崎亮介らと共著で公表されているが、論文提出者が主体となって研究を実施しており、論文提出者の寄与は十分である。したがって、博士（環境学）の学位を授与できると認める。