

都市計画用住民データ推定における適合度算出手法の改良

大谷 紀子研究室

0732134 立岩 美紗子

1. 研究の背景と目的

都市の将来にわたる開発計画においてマイクロシミュレーションモデルの活用が進められている。マイクロシミュレーションモデルでは各世帯に関する詳細なデータが必要になるが、すべての世帯に対し聞き取り調査することは困難であり、また住民基本台帳などからデータを取得することは個人情報保護の観点から望ましくない。したがって、IPF 法などにより推定したデータ（以降、推定データと呼ぶ）を利用する。しかしデータの推定手法は複数あり、各々の推定手法の妥当性を問う必要がある。大谷らは、一部の地域に関する実際のデータ（以降、観測データと呼ぶ）が得られている場合の推定データの適合度を定義し、実時間内での適合度算出を可能にするため、共生進化に基づく算出方法を提案している[1]。適合度の平均と分散で適合度算出の有効性を示しているが、分散が十分に小さいとはいえず、より確実に収束する手法が求められる。

本研究では、より正確で処理時間の短い適合度算出手法の提案を目的とし、処理内容を変更して精度向上を目指す。また、別研究により提案される単純な遺伝的アルゴリズム（GA）を利用する算出方法と比較し、共生進化を利用した場合の有効性を検証することで利用価値について言及する。

2. 共生進化を用いた推定データの評価

本研究で扱うデータの構成要素は4人の家族構成員の年齢である。観測・推定データは父・母・子1・子2の年齢を成分としたベクトルで表される。データ間の適合度の指標として、大谷らが定義した推定データと観測データの集合間の最小距離和を利用する。適合度が小さいほど評価が高いことを示すが、すべてのデータの組合せに関して距離を算出すると、計算量はデータ数の階乗のオーダーで増加するため組合せ爆発を伴う。組合せ爆発を回避するため、大谷らが提案したGAの一種である共生進化を用いた適合度算出手法を使用する。GAは、生物の進化過程を模倣した最適解探索アルゴリズムであり、共生進化はGAに比べ局所解への収束を回避しながら効率よく最適解を見つけることができる手法である。共生進化を用いた適合度算出の処理手順を以下に示す。

- 1) 初期世代の部分解集団の個体をランダムに生成後、全体解集団の個体をランダムに生成
- 2) 全体解集団を評価後、部分解集団を評価
- 3) 次世代の部分解集団の生成後、次世代の全体解集団の生成
- 4) 1)~3)をG回(回数は任意)繰り返す
- 5) 全体解の適合度から最良のものを出力

なお、先行研究では部分解集団の個体の染色体を長さ32bit×2の64bitとしている。64bitのうち16bitで一つの世帯番号を表し、推定データと観測データ一世代ずつをペアとして2ペアを表現している。全体解集団の個体の遺伝子は部分解集団の個体を参照するポインタである。

3. 適合度算出手法の精度向上

適合度算出手法の精度向上のため以下の5種類の処理について検討した。なお、データはランダムに作成した100世帯の家族4人分の年齢である。

① 次世代生成の検討

部分解集団の交叉やコピーする個体を変更した。

② 交叉方法の検討

一点交叉だったものを二点交叉に変更し、①に組み合わせた。

③ 個体数の検討

先行研究にて全体解集団・部分解集団の個体数が1000個体だったものを各々a. 900, b. 1100, c. 1200, d. 1500に変更した。

④ 部分解における染色体の表現の検討

A. 8bit を一世帯とした表現型, B. 4bit を一世帯とした表現型に変更した。

⑤ 全体解における表現型の検討

表現型生成でポインタの示す部分解の世帯番号に重複が生じた際の処理方法を変更した。

精度向上実験の結果より、先行研究よりも適合度の分散が小さい手法は③のbと、④のAであり、処理時間が短く適合度に大きなずれが生じないものは③のaで適合度を算出した場合であった。

4. 比較実験

精度向上実験で適合度が向上した③のa, bでの適合度算出手法と、④のAでの適合度算出手法において、データ数を5000世帯分に増やし適合度算出を20回繰り返した。図1に比較実験の結果を示す。精度向上実験では④のAは先行研究の結果よりも適合度・時間ともに良かったが、比較実験では圧倒的に適合度が低くなった。しかし分散は4つの中で一番小さくなった。③のaは精度向上実験と同じく比較的適合度が高く、時間が短くなった。③のbは比較的適合度が高いが、時間が長くなった。

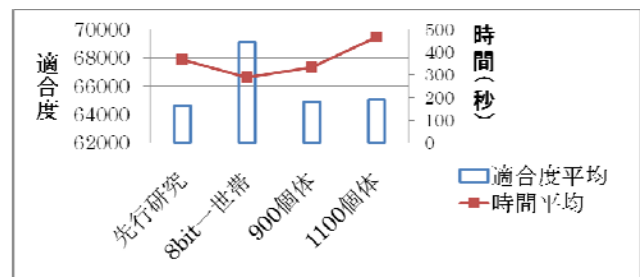


図1：比較実験での適合度の平均

5. 考察

比較実験の結果、適合度の平均が先行研究に達したものはなく、適合度の分散がより小さく処理時間の短い算出手法を発見できなかった。しかし、③のaは10回の実験結果を照らし合わせると先行研究に勝ることもあり、処理時間は先行研究を下回っているため有用であるといえる。また、④のAは適合度の平均が高いものの評価に安定性があり、今後他の個所に変更を加えることにより十分に利用価値は高まると考えられる。

GAによる適合度算出手法と本研究の手法を10回繰り返して比較した結果、GAでは平均適合度493211、平均処理時間64891.64秒となり適合度・処理時間ともに共生進化の利用が妥当という結果を得た。

参考文献

[1]大谷紀子, 杉木直, 宮本和明, “土地利用マイクロシミュレーションにおける観測マイクロデータ集合と推定集合の適合度評価” 第39回土木計画学研究発表会論文集, 136, CD-ROM, 2009.