

マイクロシミュレーション型土地利用モデルに対する並列処理の試み*

A Trial of Parallel Computing for a Micro-Simulation Land Use Model*

佐藤勝美**・井ノ口弘昭***・吉村一馬****・北詰恵一*****

By Katsumi SATO**・Hiroaki INOKUCHI***・Kazuma YOSHIMURA****・Keiichi KITAZUME*****

1. はじめに

マイクロシミュレーションの手法を用いた土地利用モデルは、その最小分析単位である個人・世帯などの個々の主体の行動を表現できるとともに、それらの動的な行動の表現も期待できる有効な手法である。しかし、一方で、取り扱うデータ量や計算量が圧倒的に多くなり、それと同時に計算時間も長くなってしまふ。スーパーコンピュータのようなハイエンド計算機を用いれば計算時間を多くかけずに処理を行うことができるが、多額のコストがかかってしまうであろう。ここで想定している土地利用モデルの使い方は、例えば、住民参加や多くの関係者による合意形成などのプロセスにおいてさまざまな都市政策の代替案を考え、それらが将来の土地利用に及ぼす影響を政策指標や GIS などを用いて定量的、かつ視覚的に繰り返し表現しながら、官民一体となった政策検討を行っていくものである。この場合、スーパーコンピュータのようなハイエンドクラスの計算機が利用できるケースの方が少なく、膨大なデータ量と計算量を短時間で処理できる簡易な方法が求められる。

スーパーコンピュータを利用できる環境でなくても長い時間をかけずにシミュレーションをするには PC クラスタ型並列コンピュータを用いた並列処理が有効である。そこで本研究は、PC クラスタ型並列コンピュータを用いたマイクロシミュレーション型土地利用モデルの並列処理を行い、計算時間の短縮を行うことと並列処理の効率分析を行うことを目的とする。

* キーワーズ：並列計算、マイクロシミュレーション、土地利用モデル

** 学生員、関西大学大学院理工学研究科（大阪府吹田市山手町3-3-35、TEL:06-6368-1111、E-mail:nxffd356@ybb.ne.jp）

*** 正員、博士(工学)、関西大学環境都市工学部都市システム工学科（大阪府吹田市山手町3-3-35、TEL:06-6368-0964、E-mail:hiroaki@inokuchi.jp）

**** 正員、修士(工学)、京都市

***** 正員、博士(工学)、関西大学環境都市工学部都市システム工学科（大阪府吹田市山手町3-3-35、TEL:06-6368-0892、E-mail:kitazume@ipcku.kansai-u.ac.jp）

2. モデルの概要

並列処理を検討するためのマイクロシミュレーション型土地利用モデルを作成する。そして、並列処理を意識したプログラムを作成し、PC クラスタ型並列コンピュータによって、計算時間を計測する。計算時間の測定にあたり最大 4 台のコンピュータを使用し台数を変化させながら比較して分析をする。

既存研究も参考にし¹⁾、本研究では、図 1 のようなモデルのフローチャートとした。

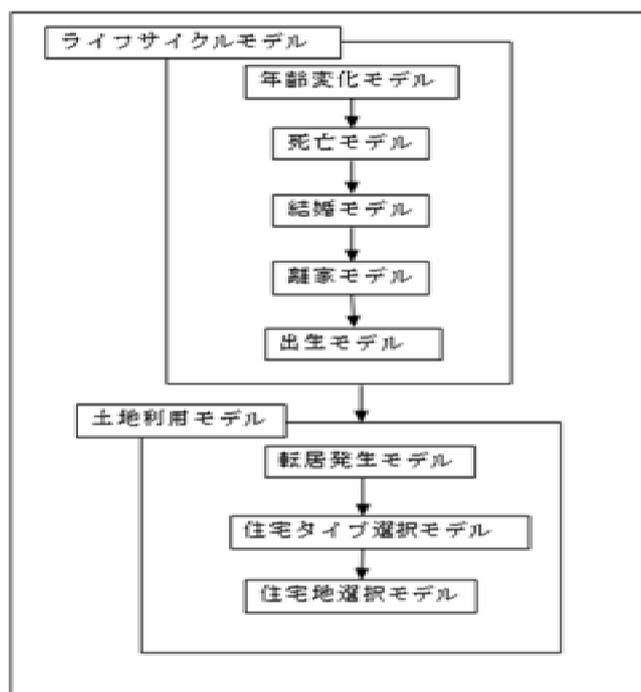


図 1 モデルのフローチャート

ライフサイクルモデルでは個人を主体において人のライフスタイルを年齢変化モデル・死亡モデル・結婚モデル・離家モデル・出生モデルで表現している。計算に用いた人口学的な指標は、国立社会保障・人口問題研究所のデータを用いている。土地利用モデルでは世帯が転居するか、どの住宅タイプを選択するか、住む場所にどこを選択するかを表現している。選択モデルは、標準的な線形関数を用いたロジットモデルタイプのものであり、

パラメータは、平成 18 年度に実施された第 4 回道央都市圏パーソントリップ調査に付随した住居や住み替え意向に関して調査したライフスタイル調査をもとに推計したものを用了。

図 2～4 は、各ステップのモデルの詳細例である。ここでは、年齢変化モデルと離家モデルと土地利用モデルについて示す。

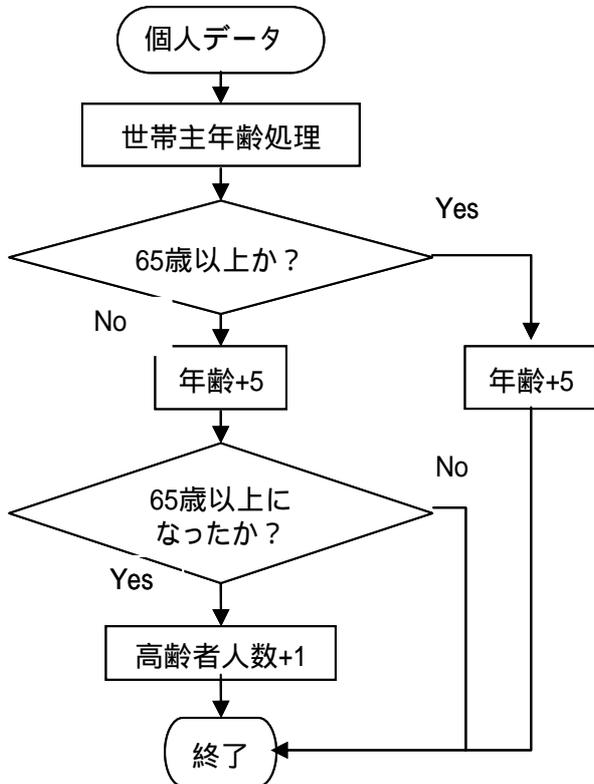


図 2 年齢変化モデル

モデルは、世帯間でのやりとりが存在し、土地利用モデルではゾーン間の相互干渉がある。例えば、PC 間のデータ分担を世帯別に行った場合、離家モデルは PC 間のデータのやりとりを頻繁に行う必要がでてくる。分担をゾーン別に行った場合は、土地利用モデルにおいて同様のことが起こる。ここでのモデルフローは、このような問題を有している。

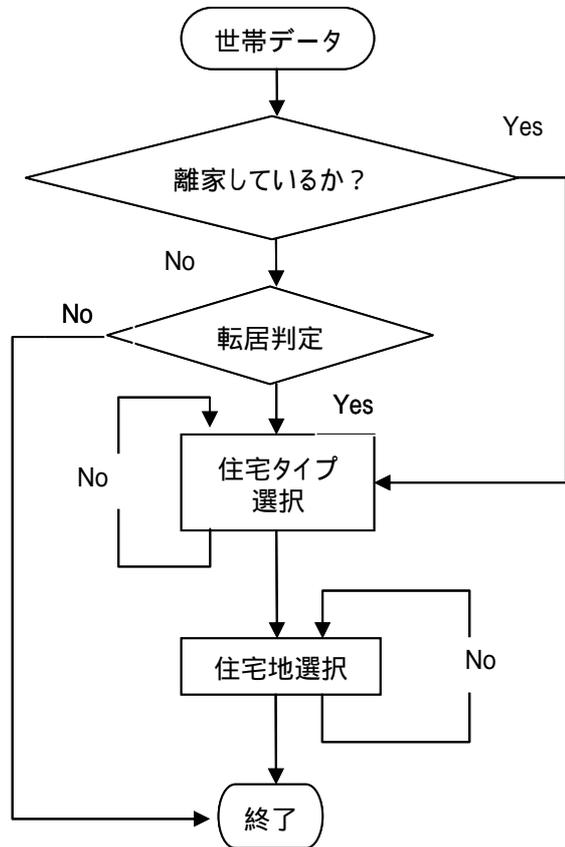


図 4 土地利用モデル

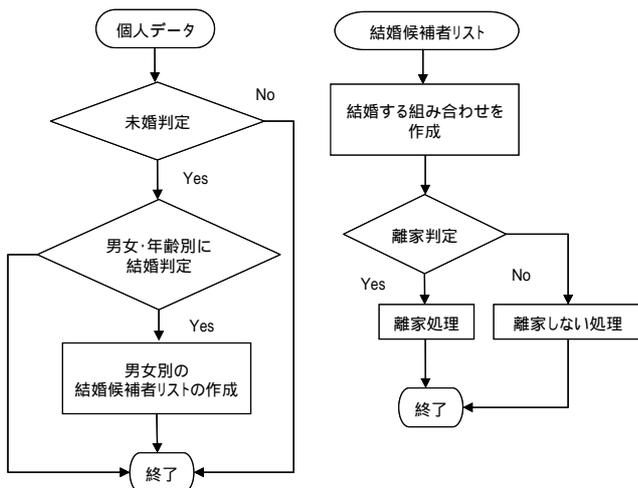


図 3 離家モデル

これらのフローで特徴的なことは、年齢変化モデルは、世帯単位であっても、ゾーン単位であっても、それぞれの計算主体は相互に関連せず独立に計算できるが、離家

3. 計算結果

今回計測にあたり、コンピュータ関連の環境は、サーバー用コンピュータ OS : Linux CentOS、CPU : Xeon 3.2GHz、メモリ : 8.0GB、言語は C、クラスタシステムとして SCore を使用した。²⁾

使用したデータは道央都市圏パーソントリップ調査のデータを元に想定したもので、並列処理の十分な効果が検証できるよう、人口 120 万人、世帯数約 55 万世帯、ゾーン数は 1 辺 10km の 5×5 の仮想都市でシミュレーションを行った。

PC 1 台で計測した結果、最も計算時間が大きかったのは離家モデルの 48 秒で全体の 50.0%であった。その他に計算時間が発生したのは出生モデルの 42 秒で全体の 43.8%、土地利用モデルの 2 秒で全体の 2.1%、その他の処理の 4 秒で全体の 4.1%だった。そして、年齢変

化モデルと死亡モデルと結婚モデルの計算時間はほぼ 0 秒であった。

この結果から、計算時間を要するステップのうちプログラムの関係上並列化できない離家モデルを除いた出生モデルと土地利用モデルの部分の並列化を行い最大 4 台の PC を使用して計算時間の測定を行った。PC 間の分担の仕方は、本研究では、初期の世帯単位での分担とし、図 5 に示すような考え方でおこなった。仮に、 $N_1 \sim N_4$ の世帯集団があったとき、最初の部分の並列化しないステップについては、すべての PC に同じデータをすべて持たせ計算させる。こうすれば、すべての PC が同じ性能を持っているかぎり、同時に計算が終了するはずである。その段階が終わったときに、各 PC は同じデータを持っているので、並列処理において最も時間のかかる PC 間のデータのやりとりを必要としない。

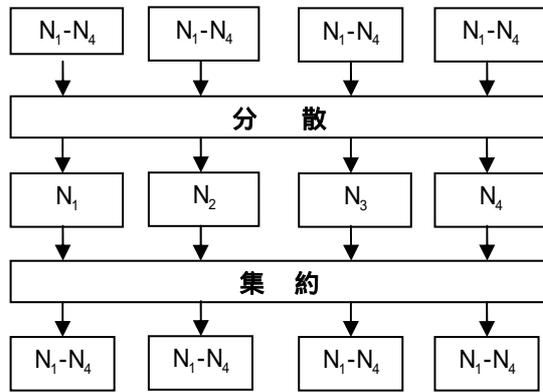


図 5 並列処理分担の考え方

各 PC は、この中から与えられた役割の世帯に対してだけ計算処理を行い、すべての PC が計算終了するまで待つ。その後、互いにデータのやりとりをし、すべての PC が、また同じデータを持つ状態とする。そして、最終的なステップに向けて計算を進める。

表 1 および図 6 は、PC の台数を变化させたときの各ステップでの計算時間である²⁾。この表から、PC の台数が 4 台のとき全体の計算時間の約 1/3 短縮されているのがわかる。これは並列処理を行った部分が全体の約半分しかなかったので短縮される割合が少なくなってしまうためだと考えられる。しかし、並列処理をした部分だけに注目してみると PC の台数が 4 台のときでは計算時間が 1/4 以下になっており、効率的な並列処理が行われていると考えられる²⁾。

表 1 PC を变化させた時の計算時間

計測項目	1 台		2 台		3 台		4 台	
	時間 (秒)	%						
年齢変化モデル	<<1	0.0	<<1	0	<<1	0.0	<<1	0.0
死亡モデル	<<1	0.0	<<1	0	<<1	0.0	<<1	0.0
結婚モデル	<<1	0.0	<<1	0	<<1	0.0	<<1	0.0
離家モデル	48	50.0	48	67.6	48	73.9	48	77.4
出生モデル	42	43.8	17	23.9	11	16.9	8	12.9
土地利用モデル	2	2.1	1	1.4	<<1	0.0	<<1	0.0
その他	4	4.1	5	7.1	6	9.2	6	9.8
合計	96	100.0	71	100.0	65	100.0	62	100.0

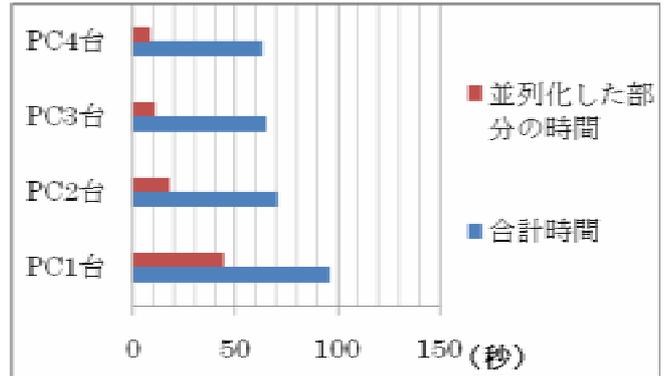


図 6 PC の台数ごとの計算時間(並列・合計時間)

下記の式(1)は Amdahl の法則である³⁾。これで並列化した部分の割合と PC の台数で速度倍率を求めることができる。

$$R(f) = \frac{1}{(1-f) + f/R_p} \quad (1)$$

ここで

f : 全体の処理における並列化可能な処理の割合、

$R(f)$: 総合的な速度向上率

R_p : 並列処理の速度向上率

である。

今回の結果を Amdahl の法則に当てはめたグラフが図 7 となる。

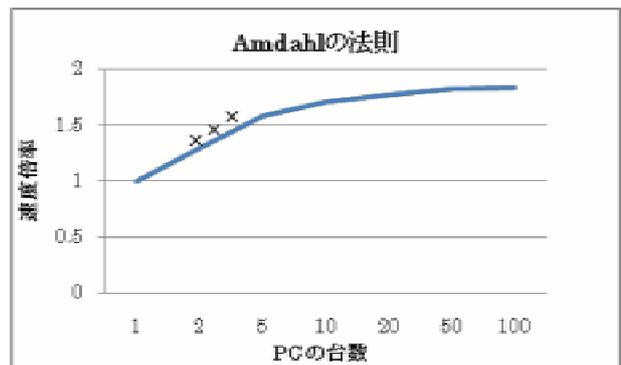


図 7 Amdahl の法則への適用

今回のモデルでどんなに PC の台数を増やしていても速度倍率は約 1.8 倍にしかならないのがわかる。これは並列化した部分の割合が半分以下だったためだと考えられる。そして、実際に計測した計算時間から算出した速度倍率は Amdahl の法則から求めた値より高い結果になった。

また、表 - 2 は、エージェント数を変化させた場合の計算時間の変化をまとめたものである⁴⁾。CPU が 1 の場合と CPU が 4 の場合を行った。この表でわかるように、CPU の数によってエージェント数の減少による計算時間の短縮度合いが若干異なるとともに、そのパターンは出生モデルと土地利用モデルによって異なる。出生モデルでは、1 CPU の場合はエージェント数が半分になれば計算時間は 25.0% になっているのに対し、4 CPU の場合は 26.3% にとどまる。一方で、土地利用モデルでは、ともに 50.0% と減少効果が小さく、CPU の差による変化はあまりみられない。

表 - 2 エージェント数の変化による計算時間

	エージェント数 (千)	出生モデル		土地利用モデル	
		計算時間 (秒)	増加率	計算時間 (秒)	増加率
1 CPU	200	1.17	1.00	0.30	1.00
	400	4.66	3.98	0.59	1.97
	600	10.42	8.91	0.89	2.97
	800	18.62	15.91	1.19	3.97
	1000	29.03	24.81	1.48	4.93
	1200	41.76	35.69	1.78	5.93
4 CPU	200	0.20	1.00	0.17	1.00
	400	0.95	4.75	0.17	1.00
	600	2.07	10.35	0.26	1.53
	800	3.64	18.20	0.34	2.00
	1000	5.32	26.60	0.43	2.53
	1200	7.87	39.35	0.51	3.00

4. 結論と今後の課題

マイクロシミュレーション型土地利用モデルに対して、並列処理による計算時間の短縮は有効であることを確認した。また、その短縮効果は、CPU の数やモデルのフロー、および CPU 間でどのような分担とすることによって変化することも確認した。もちろん、その有効性を高めるためには、並列化部分を多くすることが重要である。

なお、世界の先進的なマイクロシミュレーションを用いた実用土地利用モデルにおける並列処理の現状は、例えば、PECAS が 32 個の CPU で実施しているほかは、UrbanSim ではその実施をテストした段階、ILUTE でも将来的に実行可能としている⁵⁾。このような状況からみても、まだマイクロシミュレーションモデルに対する並列処理の研究は、進めていくべき課題が多いことがわかる。

今後の課題は、今回考慮しなかったコスト面をふまえた最適な PC の台数の検討である。OS に Linux を使

用したが、幅広く利用されている Windows でも利用できるほうが望ましい。

本研究は、平成 20~21 年度科学研究費補助金(基盤研究(B))、(課題番号: 20360232)、研究課題名「詳細属性情報を含む世帯の空間分布予測のためのマイクロシミュレーションシステム」の研究成果の一部をとりまとめたものである。ここに記して、謝意を表したい。

参考文献

- 1) 川島秀樹: 「マイクロシミュレーションにおける個人・世帯のライフサイクル」西日本理論経済学会編「現代経済学研究」第 12 号, 2005
- 2) 佐藤勝美・北詰恵一・井ノ口弘昭: マイクロシミュレーション型土地利用モデルに対する並列処理の効率分析、平成 20 年度土木学会関西支部年次学術講演会講演概要集、CD-ROM、2009.
- 3) 超並列計算研究会ホームページ: <http://www.is.doshisha.ac.jp/SMPP/>
- 4) Kitazume, K., Inokuchi, H., Sato, K. and Yoshimura, K.: Performance of Parallel Computing for a Hypothetical Land-use Microsimulation Model, Computers in Urban Planning and Urban Management'09, CD-ROM, 2009.
- 5) 宮本和明, 北詰恵一, 鈴木温, 杉木直, Varameth Vichiensan :世界における実用都市モデルの実態調査とその理論・機能と適用対象の体系化, 2008.