

人流シミュレーションにおける複数の評価基準を考慮した時間別 OD 表の推定

Inference of Origin Destination Table Differentiated by Time Considered Multiple estimation of Crowd Simulation

川口 英俊 ^{*1*2}
Hidetoshi KAWAGUCHI

野田 五十樹 ^{*2*1*3}
Itsuki NODA

^{*1}東京工業大学 大学院総合理工学研究科 知能システム科学専攻

Tokyo Institute of Technology Interdisciplinary Graduate School of Science Engineering Computational Intelligence and System Science

^{*2}(独) 産業技術総合研究所 サービス工学研究センター

National Institute of Advanced Industrial Science and Technology(AIST) Center for Service Research

^{*3}科学技術振興機構

Japan Science and Technology Agency (JST)

In this paper, we examined the influence by which observation number and time give it to estimate accuracy of Origin Destination Table in crowd simulation. A final purpose of this research is establish data assimilation for multi-agent simulation. we chose from several multi-agent simulation as the subject of an experiment. parameter estimated is a part of Origin Destination Table. The valuation basis are passing number. As a result, we proved that observation number and time have a significant influence.

1. はじめに

マルチエージェント社会シミュレーションにおける大規模データ同化の可能性を探るため、シミュレーション結果(観測値)からのシミュレーション設定値の逆推定の精度について、観測値の点数と精度の関係を調査した。具体的には、エージェントベースの人流シミュレーションにおける OD 表(Origin Destination Table)の推定精度と観測地点、観測時間の設定の関係を実験により検証した。OD 表とは、ある特定の地点間の移動量を表した行列形式の表である。簡単な設定のシミュレーションでも、観測地点と観測時間が推定精度へ与える影響が大きいことが分かった。

本研究の最終的な目的は、マルチエージェント社会シミュレーションのデータ同化手法の確立である。本研究では、データ同化とは現実の観測データとシミュレーションの出力を比較し、観測データに近似した出力となる入力パラメータを推定することと定義した。これは、社会シミュレーションの応用を考える上では、避けることが出来ない問題であり、その手法の確立は重要である。例えば、交通状況やイベント時の群衆移動の初期状態を探索することができれば、災害時などの別ケースを想定したシミュレーションを行うことで、未知の状況での現象の予測・分析の支援を実現することができる。

しかし、これらの社会現象は多様な側面があり、シミュレーション毎のデータ同化の評価基準は無数に考えられ、未だにその理論は未確立である。

本稿では、社会シミュレーションの中から人流シミュレーションをデータ同化の実験対象として議論する。推定するパラメータはエージェントの OD 表の一部分であり、評価基準は地図上の特定地点の通過人数とする。観測地点数と観測時間がデータ同化精度へ与える影響について分析した。

1.1 関連研究

観測データを基に社会シミュレータの精度を向上させようという研究は数多く行われている。しかし、画一的な方法は未だに確立しておらず、各々が独自の方法でデータ同化を行っており、観測者による主観的判断も多く含まれている。

竹内らは Social Force モデルを用いた歩行者・自転車のマルチエージェントシミュレーションにおける現況再現性の検証を行っている [Takeuchi 10]。Social Force モデルの諸パラメータを手作業で調整し、検証指標を観測映像から目視で計測していた。ある程度の精度を確認できたが、計測の労力が大きく、観測者によって誤差がある可能性がある。

上野らは、ミクロ交通流シミュレータのエージェントモデルのパラメータを、観測データから調整している [Ueno 09]。具体的には、センサーから 1 時間あたりの交通量を測り、速度と車頭間距離や加速度の関係を改善していた。結果として、精度の良いシミュレーションを行えたとしている。

上述した研究では、いずれもエージェントに内包される意図や思考に関わるパラメータを調整している。本研究では、そのようなパラメータの画一的な推定方法について検討していく。

2. 問題の定式化

2.1 マルチエージェント社会シミュレーションのデータ同化の前提

本研究で取り上げているデータ同化は、不可観測、または観測に莫大なコストのかかる心理的なパラメータを推定していく。データ同化手法は、大気や津波などの物理シミュレーションにおいては理論が確立しつつある。物理シミュレーションの推定パラメータは、観測が比較的容易であり、空間的、時間的に観測できなかった部分を補完する意味合いが強い。

2.1.1 心理的なパラメータ

マルチエージェント社会シミュレーションの特徴として、人間の心理的な側面をモデル化している点が挙げられる。この心理的な側面とは、行動や意思の決定など論理式で表記できるもの、あるいは、「選好度」のような定量化出来るものを

連絡先: 川口英俊, 東京工業大学 大学院総合理工学研究科 知能システム科学専攻, 〒 226-8503 神奈川県横浜市緑区長津田町 4259, kawaguchi.h.ac@m.titech.ac.jp

指している。例えば、人の移動をモデリングしたエージェントがあったとすると、その出発地と目的地を設定しなければならない。観測するとすれば、アンケート作業が必要である。しかし、通過人数をただ数え上げることと比べればはるかにコストがかかる上、不確実性も高い。出発地と目的地は明確に設定することができるが、そもそも直接の定量化ができないパラメータも存在する。

2.1.2 社会シミュレーションの複雑性

マルチエージェント社会シミュレーションのもうひとつの特徴として、初期設定の不確実性がある。精緻で厳密なモデルがすでに知られている物理現象シミュレーションに比べ、モデルそのものに不明な部分が多分に残る人間の行動を対象とする社会シミュレーションでは、精密なデータ同化が必ずしも正しい初期状態を導くとは限らない。また、扱う量そのものが多様であるため、単純な全体誤差最小化という基準でのデータ同化を定義しにくい。実際には、注目する現象によって全く異なる初期状態がデータ同化により得られることが十分に考えられる。

例えば、人流シミュレーションをデータ同化する際に、人の位置という評価指標と、速度という評価指標があったとする。その際、人の動きの方程式が十分に精密でなければ、位置を合わせる事が出来たととしても、速度を合わせる事が出来なくなる可能性があり、逆の場合も考えられる。

そこで本研究では、多数の評価指標によるデータ同化の解の不確実性を、ゲーム理論におけるパレート解として扱い、パレート解同士の優劣を議論しない立場を取ることにする。

2.2 データ同化の構想

マルチエージェント社会シミュレーションのデータ同化の形式は複数考えられる。

2.2.1 教師あり学習による推定

シミュレーションを条件付き確率分布 $P(y|x)$ として捉える。 y はシミュレーションの出力、 x は入力とする。

このシミュレーションの出力から入力を導出する条件付き確率分布 $Q(x|y)$ をニューラルネットワークの誤差逆伝搬法などを使って関数 $x = F(y)$ として近似する方法である。近似した関数 $F(\cdot)$ に観測データを適用することで入力データ x を得る。

問題となるのが、学習データの収集である。学習するためには、推定したいパラメータの x の収集と y を観測しなければならない。いずれもコストがかかるが、特に x はアンケートなどでしか収集できない場合が多く、更にコストがかかる。

一度 $Q(x|y)$ を近似した場合は、 y を観測さえすれば x を推定することができる。それにより、過去に観測したデータ、将来的に観測されるデータを適用し、その時のパラメータ x を推定することができる。

そもそも、 x の定量化が困難な場合もあり得る。関連研究の [Takeuchi 10] がその例である。その場合は、シミュレーションが社会現象を十二分に再現できていると仮定する。そしてランダムに生成した x をシミュレータへ入力し、 y を得て、それらを学習に用いる。

2.2.2 探索問題としての推定

入力パラメータの推定を探索問題として捉え、進化的計算などの知的な探索手法を適用する方法である。その場合、観測データとシミュレータの出力データの距離を最小化する入力パラメータを探索する。その際、観測データと出力データはそれぞれ同じ構造の多次元ベクトルとなる。それぞれの次元の距離を評価関数とするため、評価関数は複数考えられる。前述した前提に従えば、すべてを最適化する完全最適解は存在しない。そのため、進化的計算の中でも多目的遺伝的アルゴリズムが有

効であると考えられる。

3. 問題設定

本研究ではまず、同化を行うパラメータの推定に、観測がどの程度寄与しうるかを、定性的に分析することを試みる。データ同化での最大の困難は、同化に必要な観測が十分に行えないことである。そこで、そもそもどの程度の観測が必要なのかを知ることは重要な第一歩となる。また、上記のパレート解としてのデータ同化を扱う場合でも、観測の仕方とパレート解全体の形を知ることは重要になってくる。そこで、以下では、与えられた観測によりどの程度、望みの入力データ推定(データ同化)ができるかに焦点を当てて分析・検討をすすめる。

今回は擬似的に簡易的な設定のシミュレータのデータ同化をしたい場合を考える。

3.1 データ同化精度の測定方法

社会的現象の不可観測なパラメータを対象とするため、そもそもデータ同化精度の完全な測定は不可能である。そこで、データ同化精度を測るために、社会的現象をシミュレータに置き換える。そして、データ同化をシミュレータの逆関数の近似とみなす。近似にはニューラルネットワークの誤差逆伝搬法を用いる。その学習の誤差をデータ同化精度とする。

3.2 データ同化の対象シミュレーション

本研究では、人流シミュレータである山下ら [Yamashita 12] の CrowdWalk を使った簡易的な設定のシミュレーションを行う場合を想定する。CrowdWalk は歩行者をエージェントとしてモデル化したマルチエージェントシミュレータである。CrowdWalk の特徴は以下のとおりである。

- Social Force モデルに基づき速度を決定するため、渋滞や対向流の影響を考慮したシミュレーションが可能。
- ネットワークモデルを採用しているため、2次元自由空間モデルに対して軽量であり、大規模シミュレーションや網羅的シミュレーションを実現しやすい。
- エージェントごとに OD や経路、歩行パラメータなどを設定でき、多様な状況をシミュレーションできる。

本研究ではこれらの特徴を活用し、多数のシミュレーションにより、データ同化の定性的な分析を試みる。

3.3 シミュレーション設定

地図データと OD 表を入力パラメータとしてシミュレータへ入力する。OD 表に記述された設定のエージェントをすべてシミュレーション開始時に生成する。エージェントたちは自律的に各々設定された目的地へ向かっていき、到着したエージェントは消滅する。全エージェントが目的地へ到着した時点でシミュレーションは終了し、各エリアのエージェント通過人数を出力する。

3.3.1 地図データ

地図データは 2×5 に区分けされたエリア 2 つを 1 本の橋を繋ぐイメージで設定した。図 1 に概略図を示す。数字の書いてあるマス一つをエリアと定義する。その数字はインデックスである。左右にあるアルファベット A-J はエージェントの出発点および目的地である。エリアをを一つの部屋とみなすならば、外との出入口と捉えることもできる。^{*1}

*1 実際のシミュレーションでは、タイル状の地図ではなくタイルのつながりを表すネットワークとして地図を与えている。

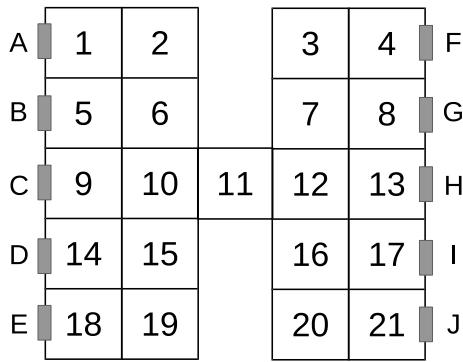


図 1: シミュレーションの地図の概略図

表 1: 実験の OD 表

	A	B	C	D	E	F	G	H	I	J
A	0	9	23	30	44	23	37	65	69	X
B	3	0	19	24	28	14	25	41	61	100
C	2	3	0	13	19	13	17	33	35	70
D	4	5	7	0	6	3	6	12	13	29
E	0	2	2	2	0	7	1	14	13	24
F	0	9	5	12	500	0	6	17	28	41
G	3	3	5	6	14	3	0	21	24	35
H	0	4	6	4	8	6	6	0	17	24
I	0	1	1	0	3	3	3	6	0	13
J	1	1	2	1	2	0	0	5	0	0

3.3.2 エージェントの挙動

それぞれのエージェントは、最短経路で設定された目的地へ移動する。全く距離が同じ経路がある場合は、その中から一様にランダムに選択する。エージェントの進行方向は上下左右のいずれかで、斜めのエリアへ移動することはできない。例えば、C → I へ移動する場合は、9 → 10 → 11 → 12 までは一意に決定するが、そこから 12 → 13 → 17 → I と 12 → 16 → 17 → I という 2 通りの経路が考えられる。

3.3.3 OD 表

エージェントの出発地と目的地は左右に 5 つずつの計 10 箇所設定している。OD 表の設定を表 1 に示す。行が該当するインデックスの出発地からの列インデックスの場所への移動人数である。例えば、A から C へ移動するエージェントの数は 23 体である。本実験では、A から J へ移動するエージェント数 (表 1 右上の X) に絞って推定を行う。なお、マルチエージェント社会シミュレーションの特徴である非線形性を出すために、F から E へ移動するエージェントの数を 500 と大きくし、混雑が起こりやすくしている。X は [0, 999] の間の整数値を一様乱数で決定する。この数値をデータ同化として推定して精度を検証する。

4. 実験

今回は簡易的な設定の人流シミュレータを用いて擬似的にデータ同化を行い、観測地点数と観測時間のデータ同化へ与える影響を確かめる。OD 表の特定地点から他の特定地点への人

数を推定パラメータとする。複数の評価基準としては、21 地点分の観測地点の通過人数を採用した。

シミュレーション結果からの観測データの抽出方法、および分析するパラメータについて述べる。まず、シミュレーションで実際に観測するのは、全シミュレーション時間のうち、最初の 20% と最後の 20% を削除した全 60% の時間で観測する。

データ同化精度への影響を調べたい観測パラメータは、観測地点数と観測時間である。

観測地点数

観測地点数は 1 から 21 の値をとる。観測地点の選択方法はランダムである。観測地点数が多いほどデータ同化精度も良くなると予想できる。

観測時間

観測時間は、上述した 60% の観測時間を基準にして、さらに 20%, 40%, 60%, 80%, 100% の値で観測をする。それぞれの観測時間について序盤、中盤、終盤の観測時期を設定する。例えば、40% の中盤なら、全観測時間のうち最初の 0% から 40% の時間だけ通過人数をカウントする。

パラメータの組み合わせ数は 21 × 13 の全 273 パターンである。それぞれのパターンでデータ同化精度を算出する。

4.1 ニューラルネットワークの学習方法

ここでは、X のデータ同化の問題を、ここでは単純な関数近似とみなす。すなわち、観測データを入力とし、X を出力とする関数を構成するという問題として定義する。そして、3 層ニューラルネットワークのバックプロパゲーション学習により、この関数近似を試みる。そして、その学習精度により、データ同化と観測の間の関係についての分析を行う。

上記の定式化により、ニューラルネットワークの構成は以下のようになる。入力層のノードの数は観測地点の数と等しくなる。この観測地点の数を 1 から 21 までで推定を行い、その際の精度を確認していく。隠れ層のノード数は入力ノードの数の半分 (小数点以下は切り上げ)。

シミュレーションを 2000 回実行し、教師データとテストデータはそれぞれ 1000 個ずつ用意した。それぞれの出力は OD 表の A → J への人数となり、入力観測データの個数と同次元のベクトルである。ニューラルネットワークの学習と評価を観測地点数 1 から 21 個それぞれにおいて 300 回ずつ行い、テストデータによる推定値と実際の値の二乗平均平方根を計算し、それを精度の値とした。観測地点のインデックスのリストは重複なしで毎回ランダムに選択した。

4.2 実験結果と考察

実験結果を図 2、図 3、図 4、図 5 に示す。縦軸が二乗平均平方根で、低いほど精度が高くなる。横軸は観測地点数である。それぞれのプロットが一回のニューラルネットワークの学習の誤差、つまりデータ同化精度にあたる。

4.2.1 観測地点数のみ

図 2 は観測時間が 100% のデータ同化精度、つまりは観測時間を実験の考慮にいれず、純粋に観測地点数の与える影響を分析することができる。そこから、以下のことが分かる。

- 観測地点の数が多くほど、精度も上がり、安定した結果を得ることができる。
- 観測地点が少ない場合でも、選択された場所によってはかなりの精度を維持することができる。

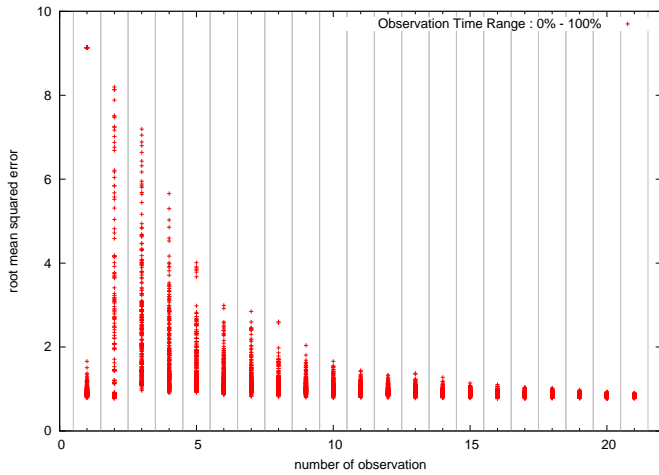


図 2: 観測地点数とデータ同化の精度の関係

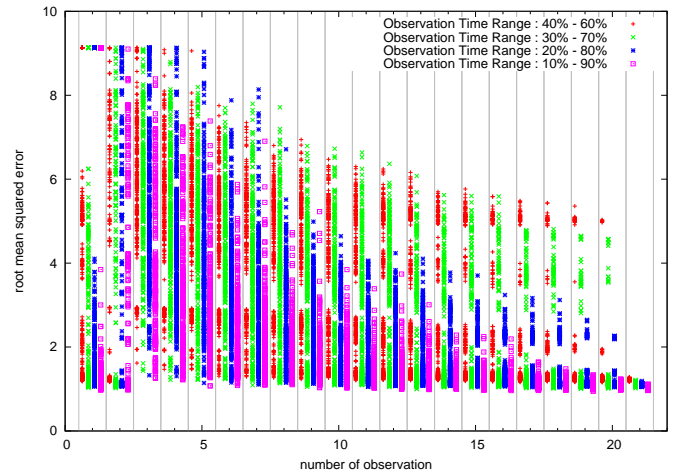


図 4: シミュレーション中盤のデータ同化の精度

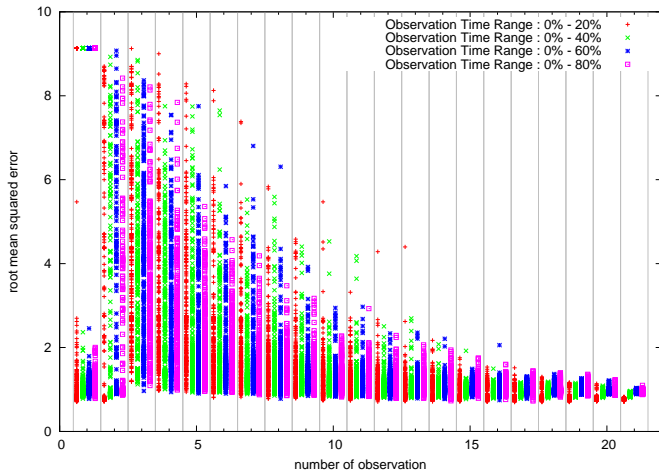


図 3: シミュレーション序盤のデータ同化の精度

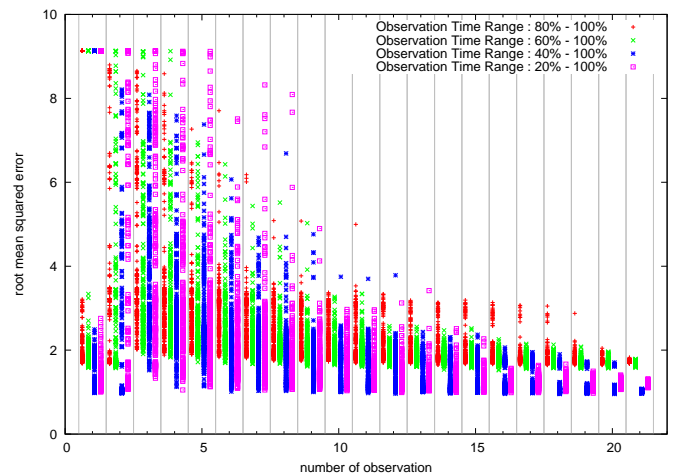


図 5: シミュレーション終盤のデータ同化の精度

- 観測地点数 1 のデータ同化も安定して高精度を出しているが、突出して精度が悪いものもある

観測地点数が多いほど高精度かつ安定している事は自明である。ここで着目すべきは、観測値点数が少ない場合、不安定になっているが観測地点の選択パターンによっては高い精度を維持できている点である。

4.2.2 観測時間の違いによる比較

観測時間が序盤、中盤、終盤の実験結果をそれぞれ 3、図 4、図 5 に示した。凡例は、観測時間から切り出した範囲を示している。観測時間が長いほうがやはり全体的には高精度で安定している。注目すべきは、中盤が序盤、終盤と比較して精度が不安定になっていることである。これはマップ中心での渋滞時間が関係しているのではないかと考えられる。このことから、観測時間の範囲も重要な要素であると考えられる。

つまり、このシミュレーションをデータ同化する場合の観測は、渋滞の起きづらい序盤に行うのが良いと考えられる。

5. おわりに

本稿では、マルチエージェント社会シミュレーションのデータ同化手法の基本構想を述べ、複数の評価基準の基での 1 つの入力パラメータの推定精度についての実験・検証を行った。

今後は 2 つ以上の入力パラメータの推定と複数種類の観測データとの関連を調べ、データ同化手法の洗練化を行っていく。また、実データへの適用も行い検証していく。

参考文献

[Takeuchi 10] 竹内内篤, 河村成人, 大脇鉄也, 濱本敬治, 上坂克巳: 歩行者・自転車マイクロシミュレーションの現況再現性の検証に関する一考察, 土木計画学研究・講演集 41,367(2010)

[Ueno 09] 上野秀樹, 平田洋介, 大場義和: 交通現象を高精度で再現できるマイクロ交通流シミュレータ, 東芝レビュー 64,4,pp.23-26(2009)

[Yamashita 12] 山下倫央, 副田俊介, 大西正輝, 依田育士, 野田 五十樹: 一次元歩行者モデルを用いた高速避難シミュレータの開発とその応用, 情報処理学会論文 53,pp.1732-1744(2012).