

博士論文

データ同化手法を用いた
人の流れの高精度推定手法の研究

中村 敏和

目次

第1章 はじめに	1
1.1 研究の背景	1
1.1.1 人の流れ推定のニーズ	1
1.1.2 既往の人の流れの計測技術	2
1.1.3 既往の人の流れの推定手法	3
1.1.4 人の流れの計測・推定技術の課題	3
1.2 研究の目的・独自性	4
1.3 本論文の構成	5
第2章 既往の研究事例	6
2.1 計測技術に関する既往研究	6
2.1.1 カメラ・レーザスキャナセンサを用いた広域での非集計観測	7
2.1.2 GPS	8
2.1.3 IC カード	9
2.1.4 PT 調査	10
2.2 シミュレーションモデルに関する研究事例	11
2.2.1 選択行動モデル	12
2.2.2 流体モデル	13
2.2.3 離散モデル	13
2.2.4 セルオートマトンモデル	14
2.2.5 経路選択モデル	15
2.3 観測値を用いた高精度化	16
2.4 現状の課題と本研究の位置づけ	16
第3章 提案する人の流れ推定手法のフレームワーク	23
3.1 提案する人の流れ推定手法の全体フレーム	23
3.2 データ同化手法	24

3.3	提案手法の方法論的詳細	27
3.3.1	パーティクルフィルタ法	29
3.3.2	最適内挿法	31
3.3.3	人の流れ推定への適用	32
3.3.4	大規模計算時の重み更新の近似	35
3.4	本研究での検証・考察対象	37
第4章	駅構内での人の流れ推定への適用	39
4.1	駅構内での人の流れ推定のフレームワーク	39
4.2	歩行者モデル	40
4.2.1	グリッドトラジェクトリーの生成	42
4.2.2	グリッドトラジェクトリーの平滑化	43
4.2.3	まとめ	44
4.3	データ同化手法の適用	45
4.3.1	手法の概要	45
4.3.2	具体例	46
4.4	大崎駅構内での実験	50
4.4.1	研究対象領域	50
4.4.2	使用するデータ	51
4.4.3	実験の概要	54
4.4.4	データ同化手法の基本的な検証	55
4.4.5	歩行者モデルとデータ同化手法の融合	58
第5章	東京都市圏での人の流れ推定への適用	66
5.1	東京都市圏での推定の全体フロー	66
5.2	パーティクルの生成	68
5.2.1	トリップ集合の生成	69
5.2.2	出発時間の分散	69
5.2.3	パーティクル生成	73
5.3	東京都市圏でのシミュレーションモデル	73
5.3.1	データ指向な人の流れ推定手法	73
5.3.2	行動選択モデルを用いた人の流れ推定手法	75
5.4	東京都市圏での実験	77

5.4.1	実験条件	77
5.4.2	データ指向な人の流れ推定手法での実験	79
5.4.3	行動選択モデルを用いた人の流れ推定手法での実験	84
第 6 章	結論	95
6.1	本研究の成果	95
6.2	今後の課題	97
6.2.1	より理論的な考察	97
6.2.2	フィルタリング方法の改善	97
6.2.3	近似計算の手法改善	98
6.2.4	大規模推定への適用へ向けた計算コストの削減	98
6.2.5	経路推定の導入	98
6.2.6	非集計観測値の利用	98
6.2.7	リアルタイム推定	99
6.3	今後の展望	99

表 目 次

2.1 個人属性の例	11
4.1 赤線を通過したかどうか	46
4.2 パーティクルとその重み	49
4.3 重みの更新	50
4.4 SICK LMS200 の仕様	52
4.5 出発出入口の頻度	53
4.6 到着出入口の頻度	53
4.7 入力データ	55
5.1 説明変数	77
5.2 実験条件	79
5.3 2 レベル NL モデル目的地選択パラメータ	85
5.4 2 レベル NL モデル目的選択パラメータ	86
5.5 3 レベル NL モデル交通手段選択パラメータ	86
5.6 3 レベル NL モデル目的地選択パラメータ	87
5.7 3 レベル NL モデル目的選択パラメータ	87

目次

2.1	人の流れの計測方法	6
2.2	カメラを用いた人の流れの計測	7
2.3	レーザスキャナを用いた人の流れの計測	7
2.4	カメラやレーザスキャナによる観測（左図）とそのギャップ（右図）	8
2.5	IC カード利用例	9
2.6	拡大に用いるデータ内属性の項目	11
3.1	研究の目的	23
3.2	全体フレーム	24
3.3	データ同化の気象への適用例	25
3.4	データ同化のイメージ	25
3.5	データ同化の効果	26
3.6	サンプル集合生成の概念図	31
3.7	推定手法のイメージ	33
3.8	パーティクルの生成	33
3.9	パーティクルの重み更新のイメージ	33
4.1	推定手法のフレームワーク	39
4.2	ポテンシャルマップ	40
4.3	グリッド分割	41
4.4	グリッドトラジェクトリーの例	41
4.5	グリッドトラジェクトリーの平滑化	42
4.6	パーティクルの生成	43
4.7	パーティクルの評価	44
4.8	平滑化の結果例	44
4.9	歩行者モデルのまとめ	45
4.10	具体例の状況	46

4.11 生成されたパーティクル	48
4.12 大崎駅構内と 11 の出入口	51
4.13 レーザスキャナによる歩行者計測	51
4.14 レーザスキャナセンサの配置図	52
4.15 レーザスキャナによる歩行者計測結果	53
4.16 7:00-7:05 の歩行者の軌跡	54
4.17 歩行者モデル用データ (8:00-8:10) の 6 番出入口からの軌跡	55
4.18 各出口の通過人数	56
4.19 30 秒ごとの出口の通過人数の RMSE	56
4.20 実際のデータのグリッド通過量	57
4.21 推定値のグリッド通過量 (左図:データ同化なし, 右図:データ同化あり) . . .	57
4.22 0.05 秒刻みのグリッド人数の RMSE	57
4.23 30 秒ごとの出口の通過人数の RMSE	58
4.24 実際のデータのグリッド通過量 (再掲)	59
4.25 推定値のグリッド通過量 (左図:データ同化なし, 右図:データ同化あり)(再掲)	59
4.26 歩行者モデルを用いた推定値のグリッド通過量 (左図:データ同化なし, 右図: データ同化あり)	59
4.27 0.05 秒刻みのグリッド人数の RMSE	60
4.28 シミュレーションモデルのみでの推定値の OD 表の実際との差	60
4.29 データ同化ありでの OD 表の実際との差	61
4.30 時間間隔と OD 表の RMSE の関係	61
4.31 パーティクルフィルタ法による予測変化のイメージ	62
4.32 観測地点となる 30 箇所の 3m × 3m グリッド	62
4.33 パーティクルフィルタ法による予測値の変化	63
4.34 厳密な計算と近似計算の比較	64
4.35 厳密な計算と近似計算の重みの誤差割合の確率分布	64
4.36 厳密な計算での計算時間とパーティクル数の関係	65
4.37 近似計算での計算時間とパーティクル数の関係	65
5.1 東京都市圏での推定の全体フロー	67
5.2 トリップ集合の生成	68
5.3 H10 東京都市圏 PT 調査データの出発時間の分布	72

5.4	推定されたトリップの出発時刻密度関数	72
5.5	行動パターン選択によるトリップ推定	74
5.6	NL モデルのツリー構造	75
5.7	検証方法の概念	78
5.8	時間帯別の鉄道駅乗車数の RMSE	80
5.9	時間帯別の鉄道駅降車数の RMSE	80
5.10	道路交通センサス対象道路における時間帯別 RMSE	81
5.11	全道路を対象にした時間帯別の RMSE	82
5.12	道路交通センサス対象道路における時間帯別 RMSE	83
5.13	道路交通センサス対象道路における時間帯別 RMSE	83
5.14	道路交通センサス対象道路における時間帯別 RMSE	84
5.15	NL モデルのツリー構造（再掲）	85
5.16	時間帯別の鉄道駅乗車数の RMSE	88
5.17	時間帯別の鉄道駅降車数の RMSE	89
5.18	道路交通センサス対象道路における時間帯別 RMSE	89
5.19	NL モデルを適用する利点	90
5.20	鉄道駅における乗車・降車客数の時間帯別 RMSE の 1 時間平均とパーティ クル数の関係	91
5.21	鉄道駅における乗車・降車客数の時間帯別 RMSE の 1 時間平均と観測地点 数の関係	92
5.22	鉄道駅における乗降客数の時間帯別 RMSE の 1 時間平均と観測鉄道駅の情 報量の平均の関係	93
5.23	鉄道駅における乗降客数の時間帯別 RMSE の 1 時間平均と観測鉄道駅の観 測人数の平均の関係	93

第1章 はじめに

1.1 研究の背景

近年の多様なセンサの発展・高性能化に伴い、セキュリティやマーケティング、コンテクストアウェアサービスなどを中心として、センサ情報を利用した人の流れの計測及びその応用に対する期待が高まってきている。駅構内や公園等の歩行者空間では、交通バリアフリー法の制定やユニバーサルデザインの考え方が広まったことで、重点整備がなされてきており、歩行者の動線等を計測・推定することは、歩行者空間の質の向上へ大きく役立つと考えられる。また、都市圏での広範囲な人の流れに目を向けると、地震や火災発生あるいは大規模イベント開催における混乱に伴う二次的災害や、ターミナル駅等の交通結節点における混雑などを解消する必要性から、ダイナミックに時々刻々と変動する多くの人々の流動を日常的に把握する必要性が出てきている。このように、歩行者空間から都市内の広範囲な動きまで、様々な規模で人の流れを推定することへの重要性が高まってきている。

こうした背景のもと、本研究では、観測値とシミュレーションモデルを融合して、人の流れを高精度に推定する手法を提案する。そして、大崎駅構内での人の流れ推定と東京都市圏での人の流れ推定へ適用する。この二つの領域では、計算量やメモリ領域等の観点や、観測値の取得間隔などの違いから、前者ではより理論上正確な手法の提案を行い、後者に対しては、その近似的手法を提案し、適用する。

1.1.1 人の流れ推定のニーズ

歩行者空間の安全性・快適性は、密度や動線のわかりやすさ、適度なにぎわいなど、様々な要因によって決定される。また、マーケティングなどを考えると、人が良く通る場所や進行方向などの情報を推定できることの意義は大きい。

駅構内に絞って考えると、まずは、人の流れを推定し、混雑度の推定を行うことで、より安全な歩行者空間の実現が期待できる。駅構内には朝・夕のラッシュを中心に、非常に多くの人が集まるため、その歩行者空間の安全性・快適性を向上させることは社会的にも大きな利益を生むことができると考えられる。また、マーケティングの効果も高いため、人

の流れが推定できればより効率的な配置が可能となる。さらに、駅構内では構内のみならず周辺状況の案内も期待されているため、それらの配置や表示方法の最適化にも大きく寄与すると考えられる。

また、高度に発展した現代都市の居住者や管理者、商業を営む者にとって、生活レベル向上、サービス向上やマーケティングの観点から、都市内の人の流れを把握することがより重要になってきている。混雑の解消や都市交通計画の策定などのために人の流れの把握が重要になっている他、近年ではインフルエンザなど感染症のパンデミックが問題視されている中、さらなる感染拡大を防ぐため感染経路のシミュレーションに用いられるなど、様々な分野に応用され、その価値はさらに高まっている。

このように、人の流れを推定することで、非常に多岐にわたる応用が望めるが、それぞれの環境によって求められる精度やニーズも様々である。朝のラッシュ時などは人口密度が非常に高いケースが考えられるため、高精度な計測・推定は非常に困難となる。このような公共空間におけるニーズとして、構内の流動調査、人数推定、混雑度推定などから、混雑の緩和、災害時における避難誘導など、基本的には利用者が利用しやすい空間構造を評価することが挙げられる。一方で、セキュリティやマーケティングなど、個人レベルの行動履歴が重要な役割を果たす場合、その軌跡や進行方向などを含めたより高精度な推定が望まれる。また、セキュリティや災害時の救助への応用のためには、リアルタイムに推定し利用できることが望ましい。

本研究では、基本的には駅構内のラッシュ時における人の流れの推定を行い、混雑度や人数推定を推定することを主目的とするが、将来的により高精度な推定への改良が期待できる手法を提案する。

1.1.2 既往の人の流れの計測技術

駅構内などの自由空間での計測方法として、計測対象の人物がセンサを所持するタイプと、外部から非接触で計測するタイプの2つのタイプがある。前者は、GPSやRFIDタグ、加速度センサ等の装置を能動的に携帯することで位置を計測するため、基本的に周囲の混雑状況などに依存されずにデータを取得できる。近年は、スマートフォンの普及により、こうした計測を行いやすくなってきている。ただし、データの取得間隔によっては、その間を補間する必要がある。また、一部の人の計測しかできないため、全歩行者の推定などを行う場合は推定する必要がある。

後者のタイプは、カメラやレーザスキャナ、人感センサ等のセンサを環境側に設置し、被観測者に対して非接触に計測するため、駅構内や店舗などにおける不特定多数の人を計測

したい場合に有効である。ただし、人の ID などの情報は直接的には分からないため、何らかの推定を行う必要がある。

カメラやレーザスキャナを用いて局所的に通行量や人の流れを推定する技術が発展してきている一方で、都市圏などの広範囲においてすべての人の流れを計測することは不可能に近いことである。また、IC カードや ETC の普及により、断面交通量や駅の乗降客数をより容易に取得できるようになるとも考えられるが、同様に局所的な計測に留まってしまう。

今まではこうした都市圏レベルの人の流れの把握のために、パーソントリップ（以下 PT）調査データなどで偏りのないサンプルを調査することにより全体を推定するという作業が行われてきた。実際に PT 調査は国勢調査と整合を取る形で拡大係数を保持しているため、属性ごとにサンプル率の逆数を拡大係数として算出し、これを書くサンプルに与えることで、人一人の移動を拡大係数分の人数の移動としている。

1.1.3 既往の人の流れの推定手法

歩行者空間の混雑度などを、歩行者モデルを構築して推定するという試みが多くなされてきている。新たに設計した歩行者空間に対し、ある歩行者モデルで記述された人々を発生させ、その動線や混雑度を推定する。壁や柱からの距離、他の人からの距離や交錯、すれ違いなどを記述するために、ベイズ推定を用いたモデルやゲーム理論を用いたモデル、流動マップを用いたモデルなどが提案されている。

都市圏レベルでは、前述のように PT 調査を拡大して全体の人の流れとして用いることが多い。他には、人の行動モデルを PT 調査などから構築し、シミュレーションすることで全体を推定する研究や、完全な人の流れの推定ではないが、どの出発地からどの到着地にどれくらいの人数が移動しているかを推定する OD 推定などの研究が良くなされてきている。

1.1.4 人の流れの計測・推定技術の課題

前述のように、自由空間での計測技術には大きく分けて 2 通りの方法があるが、どちらもメリット・デメリットを抱えている。対象人物がセンサを所持する場合は、人の ID などの情報を得られる一方で、多くの人を同時に計測することが困難であり、データの補間が必要な場合もある。非接触で計測する場合は、同時に多人数を計測することが可能であるが、人の ID などの情報を推定する必要がある。また、ある程度の広さの計測は可能であるが、都市圏の駅構内などのある程度の広さがある領域での計測はコスト設置場所などの問題があり困難である。

歩行者モデルを用いた人の流れの推定手法に関する課題としては、出発地・到着地が必要であること、非線形な動きをする人の行動を数値モデルで表すことの限界、などが挙げられる。一般に、歩行者モデルは、出発地と到着地が与えられることを前提として構築されているが、実際に OD を得ることは非常に難しい。また、天候やイベント、地面に落ちているゴミ、など大小様々な要因で人の流れは変化するため、それをすべて数値モデルで記述することは非常に困難である。

都市圏での人の流れの計測については、その規模から全数を計測しきることが不可能と言えるということが大きな問題点となり、全体の把握には何らかの推定が必要になる。

PT 調査を拡大する場合、人一人の移動を拡大係数分の人数の移動とするため、同じ属性を持つ人々が数十人集まり、一日中同じ経路、同じ行動を取ることになる。これはかなり単純化された考え方であるため、道路や鉄道路線単位などの詳細な状況を把握するには、様々な観測データを加味し、動的に推定できることが今後必要であると考えられる。

行動モデルを構築する推定手法に関しては、自由空間での歩行者モデルと同様に、非線形性が強く、様々な要因が複雑に絡み合っているため、すべてを数値モデルで記述することは非常に困難である。また、パラメータの算出を過去データから行うしかなく、過去の状態に依存してしまうという問題点もある。

1.2 研究の目的・独自性

このように、計測、シミュレーションモデルともに課題が残っている。本研究では、計測の課題である観測のギャップを推定することで埋め合わせ、逆に、歩行者モデルの課題である出発地・到着地を事前に設定する必要があるという点を断片的な観測値で補間する、というフレームワークの構築を提案する。つまり、計測とシミュレーションモデルを統合することでより高精度な人の流れの推定を行うことを目的とする。統合には後述するデータ同化手法を用いる。

観測値とシミュレーションモデルを融合して精度向上を目指す研究は他にもいくつかあるが、本研究と他の研究との違いとして、本研究で求める対象が移動軌跡の確率分布であり、これは非連続で離散的で分散も計算できない変数の確率分布である点が挙げられる。こうした対象に対してデータ同化手法を用いた推定手法を提案している研究はまだなく、新規性が高いと言える。交通量などの数値だけでなく、個々人の移動軌跡の確率分布を推定する事で、マーケティングや防犯、イベント開催時の対応など、人の属性が重要になる事柄に対し、より効果の高い情報を提供できる。また、これにより確率的シミュレーション

モデルであればその構造を限定せずに適用できる手法の提案をすることも独自性のひとつである。

1.3 本論文の構成

本論文は以下のように構成されている。

第二章：既往の研究事例

第三章：提案する人の流れ推定手法のフレームワーク

第四章：駅構内での人の流れ推定への適用

第五章：東京都市圏での人の流れ推定への適用

第六章：結論

第二章では、計測技術・手法とシミュレーションモデル、観測値とシミュレーションモデルの融合に関する既往の研究事例について触れる。第三章では、本研究で提案する人の流れ推定手法のフレームワークについて述べ、方法論的位置づけおよび理論的裏付けについて述べる。そして、第四、五章で大崎駅構内と東京都市圏を対象に、それぞれの人の流れ推定への適用について述べる。最後に第六章で、本論文の結論として、研究成果と今後の課題について述べる。

第2章 既往の研究事例

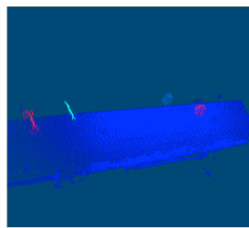
本節では，人の流れの計測技術と推定技術について，既往の研究事例に触れていく．

2.1 計測技術に関する既往研究

人の流れの計測方法として，図 2.1 のようにいくつか発展してきている．例えば，狭域で集計観測値を得る方法として，ETC での通過量や IC カードによる鉄道やバスの利用情報が得られる．近年の携帯電話の普及により，基地局情報などでより広域に集計観測値を得ることもできるようになっている．一方で，一人一人の移動軌跡を計測する手法としては，狭域でカメラを用いたビジョンベースの事例 (Snidaro (2005), Yang (2003)) と，ロボティクス等の一部の分野で発展しつつあるレーザスキャナによる事例 (Shao (2008)) などの研究がなされており，広域では PT 調査や大都市交通センサスなどのアンケートベースの計測が行われてきた．以下，これらの計測技術や得られる情報について簡単に述べる．



カメラ



レーザスキャナ
センサ



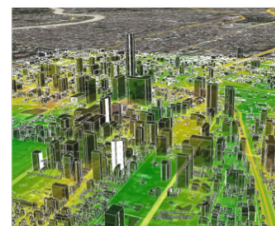
PT調査や大都市交通センサス



ETCにおける通過量情報



ICカードによる鉄道・
バス利用情報



携帯電話の基地局情報

図 2.1: 人の流れの計測方法

2.1.1 カメラ・レーザスキャナセンサを用いた広域での非集計計測

カメラの利点の一つは、色情報やエッジ形状など個々の追跡対象の特徴量を取得できることである。カメラを用いた計測の多くは、背景差分を行い人物領域を抽出し、その位置を追跡して計測を行う。例えば、追跡対象の色分布に基づいて MeanShift アルゴリズムによって計測を行う手法や、パーティクルフィルタによる計測手法などが提案されている。また、高解像度の顔画像から画像マッチングによって個人を特定する手法や、顔の向きを推定する手法なども提案されている。



図 2.2: カメラを用いた人の流れの計測

レーザスキャナを用いた計測としては、水平にレーザセンサを当て、足首の位置を取得する手法や、複数台のレーザスキャナを組み合わせで 3 次元で人の形状を取得する手法などが提案されている。レーザスキャナを用いた計測の利点として、証明条件や設置条件に影響を受けにくい、プライバシーを侵すことなく計測できる、天井が低く俯瞰しにくい場所でも対応できる、などが挙げられる。

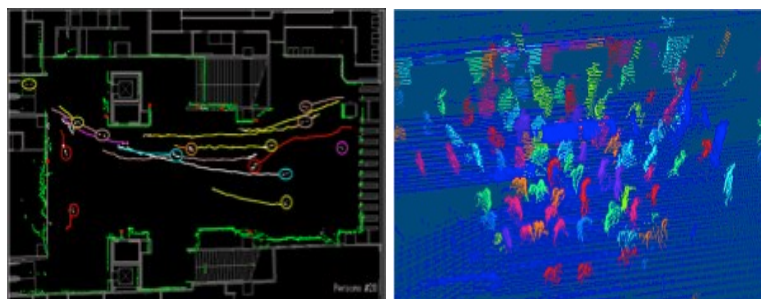


図 2.3: レーザスキャナを用いた人の流れの計測

どちらの事例でも、近年の研究による発展でかなりの精度で計測できるようになっている。しかし、計測可能領域には限りがあり、広範囲に渡って計測するには多くの機器を設置し、統合しなければならないが、コストなどの観点から非常に困難である。そのた

め、例えば東京駅構内で人の流れを計測しようとする、どうしても観測できないギャップが生じてしまう(図2.4)。図左のようにカメラで赤い領域の観測を行ったとしても、図右の紫部分のような観測できないギャップが生じてしまう。

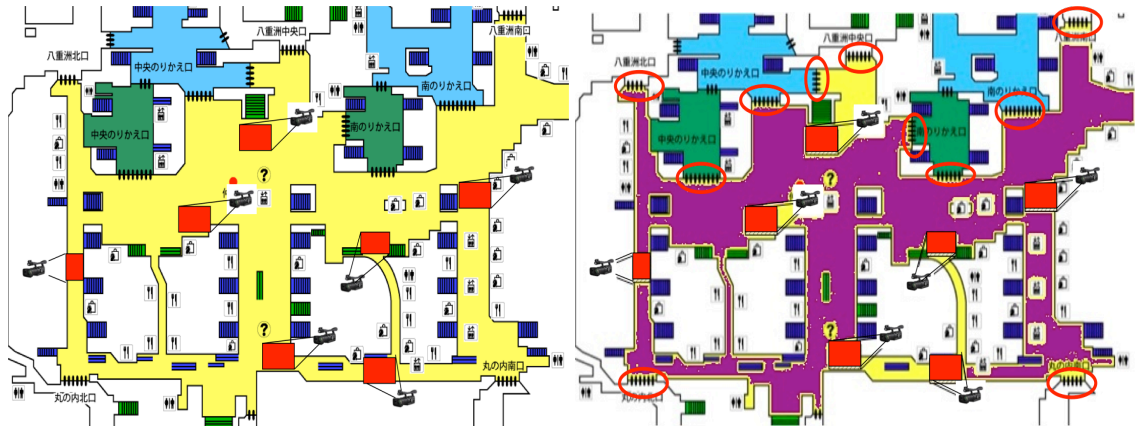


図 2.4: カメラやレーザスキャナによる観測（左図）とそのギャップ（右図）

本研究では、これらのギャップを歩行者モデルで補間することで、高精度な人の流れの推定を行うことを目指す。

2.1.2 GPS

近年、GPSは携帯電話に標準搭載されてきている。2007年4月より第3世代と呼ばれる携帯電話ではGPSの内蔵が義務付けられているため、日本国内で流通する携帯電話の多くは利用可能である。流通している携帯電話の台数は1億1000万台（2010年01月末現在、[電気通信事業者協会]より引用）であり、日本中で広く使われているといえる。GPSを利用する携帯アプリケーションも数多く存在し、現在地から地図や周辺の施設情報を組み合わせて表示するものが例として挙げられる。GPSと携帯電話を利用したシステムを構築する場合、通信履歴が残るため利用者の行動を特定可能であり、個人情報の保護を多分に考慮に入れる必要がある。とはいえ観測頻度の点でも数秒単位で連続的に測位可能であり、移動の様子を計測する手段としては申し分ないシステムである。

ただしGPSはPT調査に利用するにあたり問題となる点がある。移動の様子は克明に記録することが可能だが、移動の目的、交通手段は明らかではないことがあげられる。確かに速度が5km/hであれば歩行、100km/hで高速道路に沿って移動していればおそらく車で移動しているであろうことがわかるが、自動的に目的を判別することは難しい。このため調査を行うには、道路交通センサスのオーナーインタビューOD調査と同様にユーザーに

記録の方法を通知し実践させる必要があり，簡便にサンプル数を多く取ることに難がある．

G P Sを利用した都市交通データの研究として岡部ら (2005) が，速度を計測できる機器を積み込み調査するプローブという手段を利用してO Dごとの所要時間を後述するセンサデータと比較している例として松本 (2007) がある．

2.1.3 IC カード

非接触型 I C カード乗車券は，例として S U I C A，P A S M O，I C O C Aなどが挙げられる．2001 年の S U I C Aサービス提供開始以降，2009 年 4 月，S U I C Aと P A S M Oの発行枚数は 4000 万枚を超え，首都圏での I C カードの乗車券利用は非常に盛んである．また日本各地への I C 乗車券の普及が進んでいる．営業地域を超えて J R 各社や私鉄で I C 乗車券の相互利用が進み，利便性は増してきている．この I C 乗車券には履歴として乗車駅，降車駅が記録されている．

I C 乗車券は乗車から降車の両方を記録しており，O D 調査と親和性が高い．この乗車券自体が保存する履歴のほか，改札を利用するごとに全利用者の履歴が鉄道会社管理により保管されている．多くの乗客が I C カードを利用しているため，この履歴を利用することで鉄道やバス利用状況を正確かつ大量に，遅滞なく収集することが可能である．

ただし I C 乗車券は利用履歴の利用者以外への公開を行わない契約でのサービスであるため，データ利用の形式及び方法が大幅に制限されると考えられる．また東急線と東京メトロ線のように相互乗り入れ路線が増えることで乗車経路の特定が困難になることが予想される．



図 2.5: IC カード利用例

2.1.4 PT 調査

PT 調査とは、人の一日の行動を「目的を有する移動＝トリップ」ごとに分割し、数%のサンプルを抽出しアンケート方式で調査したものである。「パーソントリップ」は、一人の人間がある目的を持って出発地から到着地までに要する移動の単位のことであり、アンケート用紙に記載し郵送回答してもらうことによって、調査対象者の一日の行動移動を知ることができる。アンケートには性別・年齢などの個人情報も記載されることで、個人属性ごとのデータの抽出も可能である。

またパーソントリップ調査において、1トリップは1～複数個の「サブトリップ」で構成されている。トリップは「出勤」「営業」などのように目的を持つトリップであるが、例えば、ある「出勤」トリップを構成する移動が「自転車・徒歩・鉄道・徒歩」という4つに細分化される時、このトリップは4つのサブトリップから構成されていることになる。

PT 調査は対象都市圏で数%のサンプルを抽出して行うため、年代性別嗜好などの偏りなく人の行動を調査できる点で、他の行動調査データに比べて利用価値が高いと言える。

前述のように、PT 調査データは対象都市圏からのサンプリングデータであるため、対象都市圏全体の人の流れを把握するには、適切に拡大する必要がある。層別拡大法に関する研究として山形(1984)などがある。現状では、PT 調査データを拡大する方法として、属性ごとにサンプル率の逆数を拡大係数として算出し、これを各サンプルに与えるという層別拡大方法がよく用いられている。具体的には、表 2.1 のような個人属性で分類した上で、国勢調査データと層別に人数を比較し拡大を行う。

例えば、個人が4変数(a, b, c, d)の属性を持ち、各変数が4種類(1, 2, 3, 4)の値をとるとすれば、この地区には $4^4 = 256$ 種類の個人属性を持つ住民が居住していることになる。256種類各々のサンプル調査人数と、国勢調査より得られる256属性各々の現実の人数の比を拡大係数として用いれば、属性ごとに異なる拡大係数を与えることによって現実の再現性に富んだ拡大を行うことが可能になる。a=0, b=1, c=2, d=1の値(属性)を持つ人が現実には100人、サンプルに4人いるとすれば、抽出率は4%であり拡大係数は25である。その人たちは一人あたり25人分の重みを持っていると考えられ、他の属性を持ち各々の拡大係数を付与される人たちと最終的に一つの「拡大後サンプルデータ」として統合され、推測された現実の様子として扱われる。

東京都市交通計画協議会(2001)によると、上記の拡大係数設定に利用されるカテゴリの中で、制約変数は図 2.6 のように「居住地＝275種類」「性別年齢別＝15種類」「自動車保有の有無＝2種類」「世帯人数＝3種類」の4項目が利用されており、H10年度版では全ての項目数の積である計24750通りの属性に分けられる。

表 2.1: 個人属性の例

個人属性	世帯属性	自動車利用	居住地
<ul style="list-style-type: none"> ・ 性 ・ 年齢 ・ 職業 ・ 運転免許の有無 	<ul style="list-style-type: none"> ・ 世帯人数 ・ 世帯構成 	<ul style="list-style-type: none"> ・ 世帯保有自動車台数 	<ul style="list-style-type: none"> ・ 居住地ゾーン

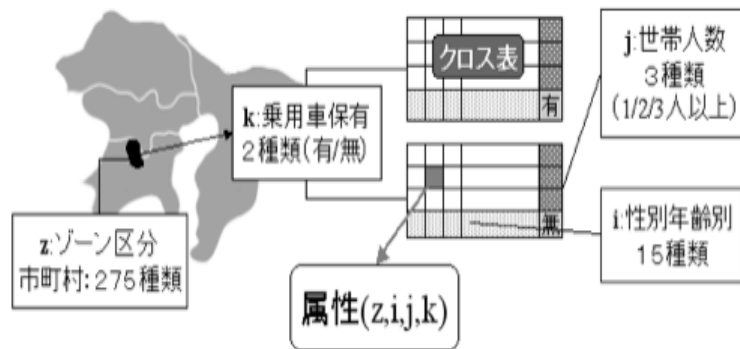


図 2.6: 拡大に用いるデータ内属性の項目

2.2 シミュレーションモデルに関する研究事例

このように、人の流れの計測技術の発展により、様々な情報の取得が可能になってきている。しかし、全ての人を完璧に計測しきることは不可能であり、全体の人の流れの情報を取得するためには何らかの推定手法の適用が必要となる。前述の PT 調査の拡大も推定手法のひとつと言える。そのような推定手法として、数値モデルを用いてシミュレーションモデルを構築する研究が多くなされている。シミュレーションモデルを構築するモデルとして、経路や交通手段などの選択を決定する選択行動モデル、自動車や群衆の移動を流体で捉えて推定する流体モデル、車両や歩行者をエージェントとみなして離散的な物体として移動させる離散モデル、対象領域を個々人や個別の車両が入るセルに区切りセル間を確率に従って動くセルオートマトンモデル、自由空間での経路選択モデルとしてネットワークモデルやポテンシャルモデルなどが提案され、これらが組み合わされて用いられている。本節では、これらのシミュレーションモデルについて研究事例に触れる。

2.2.1 選択行動モデル

ロジットモデルとその派生モデルを用いた非集計行動モデルの構築が多数研究されてきている。非集計行動モデルは確率効用理論のもと、一人一人の選択行動を直接表現できる手法として、モデルの理論的な構築が可能であることから良く用いられている。その利点として、モデル構築に必要なデータが少なく済み、変数を操作することで政策評価が容易にできることなどが挙げられている。多くの派生モデルが開発されているが、ここでは交通の需要予測等で特に良く用いられており、本研究でも用いる、ロジットモデルから MNL (Multinomial Logit) モデル、NL (Nested Logit) モデルに至る研究について簡単に触れておく。

まず、ロジットモデルを用いた非集計モデルの原型は、Mc Fadden (1974) が確率効用理論より、MNL モデルと MNP (Multinomial Probit) モデルを導出したことにより始まった。適用の容易さから MNL モデルが着目され、以降改良が加えられていく。特に必要な改良として、MNL モデルが持つ、選択肢間が独立でないといけないという IIA 特性に対処することが求められた。これは、赤バス-青バス問題として良く知られており、例えば、交通手段選択モデルを構築する際に、鉄道、自動車、徒歩といった選択肢とともに、赤バスと青バスという選択肢が合った場合、赤バスと青バスに強い相関があり、MNL モデルではそれを的確に表現できない。

Ben-Akiva (1974) は IIA 特性に対処するために、目的地と交通手段との選択を段階的に表現し、周辺確率に現れるログサム変数にパラメータを持たせるモデルを示した。それが NL モデルの開発へとつながり、相関のある選択肢を同一レベルにしたツリー構造を用いて、IIA 特性の対処に成功した。NL モデルについては、ネットワーク空間での人の流れ推定をする際に詳しく述べる。

交通需要予測をする際に四段階法が長らく用いられたが、いくつかの問題点が指摘され、非集計行動モデルがその課題解消に用いられ、交通需要統合型ネットワーク均衡モデルとして発展してきている。詳しくは土木学会 (1998) にまとめられている。都市圏を対象に行動モデルを構築し、統合均衡モデルと融合して推定した研究として、円山ら (2002)、金森ら (2009) などがある。

また、歩行者の移動モデルとしても、ロジットモデルを用いた離散選択モデルが提案されている。Antonini (2006) は、進行方向を 33 のブロックに分割し、それぞれのブロックに対し、壁や他の人々の影響を数値化し効用を求めることで、確率的に 1 つのブロックを選択するというモデルを提案している。

ロジットモデルを用いた選択行動モデルの課題として、パラメータ同定の困難さが挙げ

られる。現状では、モデルを熟知している専門家が、経験的にパラメータ探索の初期値を調整しながら算出している。最尤法などで最適値を求めることが一般的だが、極値が多く存在することなどから、調整の仕方によってパラメータ算出結果が変わってきてしまうという問題もある。

2.2.2 流体モデル

自動車の流れや歩行者流を流体とみなして近似する研究も多くなされている。Lighthill and Whitham (1955) は kinematic wave の理論を確立し、交通の流れに対して応用している。kinematic wave の理論はある場所の流量 q と密度 k が関数関係で与えられていれば、どのような一次元の流れの問題にも適用できる理論である。それとは独立に Richards (1956) も同様に交通流の推定に流体モデルを用いるモデルを提案している。これらの理論に、交通量 Q -密度 K 曲線や車間距離 S -速度 V 曲線を組み合わせて交通流を推定していく。流体モデルを用いた交通シミュレーションとして代表的なものに Hall et al. (1980) などがある。これは連続流体として交通量を表現している。

また、歩行者流を流体近似したモデルもいくつか提案されている。一方向の流れに対して障害物周りの群衆の移動について理論解の導出を試みたものもあるが (Hughes (2002))、多くのモデルは流体近似した群集流の状況を数値計算により解くものである。中 (1978) や Algahtahi and Mahmassani (1990) は、移動可能空間をメッシュ分割し、メッシュ内の歩行者密度から移動歩行者数を決定していくモデルを提案している。

流体近似するモデルの課題として、 Q - K 曲線や S - V 曲線を精度よく推定するための方法がまだ確立されていないことが挙げられる。また、これら曲線で表される関係には個人差があり、そのばらつきを表現する必要もある。ネットワーク形状や、道路形状、移動空間と障害物、歩行者同士の交差角度など、様々な条件で関係性が変化すると考えられるため、一般的に全てを考慮しきるのは難しいと言える。

2.2.3 離散モデル

離散モデルは、個々の車両や個々人を単位として、その動きを記述するものである。Pipes (1953) が車の運動を現象論的に考え、車の運動方程式を導入して以降、Chandler (1958) によって運転手の加速時の時間的な遅れを取り入れたモデルが提案され、Kometani and Sasaki (1959) が速度変化に前の車との相対距離と取り入れるなど、改良が繰り返されてきている。ただし、運動方程式を用いるシミュレーションモデルには、個々人や個別の車の意思決定

や違いを表現することが難しいなどの課題があり、運動方程式だけで完全に表現するのは難しいと言える。

歩行者流のシミュレーションに対しても、運動方程式を用いたモデルが研究されている。代表的なモデルは反発力型のもので、障害物と歩行者の間に同極間の磁力のような斥力が働くと仮定するものである。斥力と目的地方向へのベクトルとを合わせて速度ベクトルを算出する。Helbing (1995) による Social Force モデルや Teknomo (2006) によるモデルなどが挙げられる。

自動車交通シミュレーションを離散的に行うモデルとして、S-V 曲線を用いて速度を求めて逐次的に移動させていくものも多く開発されている。Rathi and Santiago (1992) は、ネットワーク上の個別車両の動きを、経路は外生的に固定配分して、シミュレーションするモデルを提案している。これに対し、森津ら (1991) や井上 (1993)、吉井ら (1995) は交通状況に応じて経路選択がなされるようなモデルを構築している。その後、運転手の属性や得られる情報の質や量を考慮したモデルが構築されてきている。

歩行者流のシミュレーションを離散的に行うモデルには、移動コストを定義し、コスト最小化できる速度ベクトルを算出していくものも提案されている。Hoogendoorn (2003) は、加速度を決定する費用として、希望速度、他の人との近接、速度変化をコストとし、最適な速度ベクトルを求めている。また、浅野ら (2007) は、進行方向に存在する人の行動を予測し、目的地方向と障害物の影響を加味してコストを計算し、最適な方向・速度で進むモデルを構築している。

2.2.4 セルオートマトンモデル

セルオートマトンは有限の離散内部状態を持つセルの集合とその状態の離散的な時間発展で成り立っている。セルの状態の更新は、周囲のセルの状態と自信の状態から状態遷移ルールに従い行われる。Nagel and Schreckenberg (1992) はセルを格子点とみなし、格子の状態として車の速度または存在しないことを表す符号を持たせ、進行方向の格子点と速度からセルの状態遷移を行うモデルを提案している。歩行者流のシミュレーションモデルとしては、が歩行可能空間を歩行者のサイズに等しい幅のセルに区切り、セル間を周辺歩行者や障害物の位置を考慮しながら遷移するモデルを提案している他、いくつか研究されている。

セルオートマトンモデルを用いた歩行者流のシミュレーションモデルとして、フロアフィールドモデルと呼ばれるモデルも開発されている。フロアフィールドモデルは各セルに状態とは別にフロアフィールドという状態量を持たせ、近傍セルのフロアフィールドを

比較して移動する。フロアフィールドは、歩行空間の静的状態と目的地からモデル化される静的フロアフィールドと、周辺の歩行者の流れに追従するなどの性質をモデル化する動的フロアフィールドからなり、避難時のシミュレーションなどが行われている (Burstedde et al. (2001))。

2.2.5 経路選択モデル

経路選択モデルは、経路生成と経路選択の2つの段階に分けることができる。Dial (1971) など2つの段階を同時に行う研究もあるが、2つの段階が含まれる必要があることには変わりない。

Dial (1971) はロジットモデルを用いて確率的に経路を生成・推定する手法を提案した。このモデルを拡張した経路選択モデルの研究はいくつかなされている (朝倉ら (2000), 兵藤ら (2009))。しかし、パラメータ算出の必要性があることや計算コストが大きくなってしまふこと、似たような経路が多くできてしまうこと等の問題点がある。

2つの段階を別々に行う手法もいくつか提案されている。経路生成ができれば、その経路集合から選択モデルを構築する等して経路選択をすることは比較的容易であるため、特に経路生成に関する研究が多くなされている。k-最短距離法は最も単純な生成手法の一つで、ネットワーク空間のある始点ノードから終点ノードまでk番目に最短である経路までを生成する手法である。こういった最適経路を生成するアルゴリズムに Dijkstra 法等が適用でき、 $O(V \log E)$ の計算量で生成できる。ここで、 V はネットワーク空間のノード数、 E はエッジ数である。最適経路の生成をベースとした経路生成手法に、リンクペナルティー法やリンク削除法など (De la Barra et al. (1993)) や最短経路、最小費用経路、最短時間経路などを求めて経路集合とするラベル法 (Ben-Akiva (1984)) などがある。

経路生成のもう一つのアプローチとして幾何的なものが挙げられる。Lombard and Church (1993) は始点と終点に対し、経由点を設定し、その経由点を通る最短経路を求めることで経路集合を生成する、GSP (Gateway Shortest Path) 法を提案している。また、木下ら (2003) は GPS を用いつつ重複率による経路の分類・集約を行うことで尤もらしい経路集合の生成手法を提案している。しかし、GSP 法はいくつかの問題点が指摘されている。Akgun et al. (2000) は GSP 法の問題点として、生成された経路がループを含む場合があるということを挙げている。

歩行者流の経路推定モデルとしては、歩行空間上にノードとリンクからなるネットワークを設定し経路選択を行うネットワークモデル (塚口ら (2002)) や目的地までのポテンシャルコストにより経路選択を行うポテンシャルモデル (横山ら (1995)) などが提案されている。

ネットワークモデルは、交通量などに応じたリンクコストを与えることで、車両の交通ネットワークと同様に経路選択・配分を行うものである。そのため、交通ネットワーク理論における知見をそのまま適用できるという利点がある一方で、歩行者の移動をリンクに集約することで詳細な容量評価に影響が出る恐れがある。ポテンシャルモデルは、まず任意の点から目的地までの移動コストを算出し、移動コストのポテンシャル面を描く。次にこのポテンシャル面に対して勾配の急な方向に向かって歩行者が移動しようとする。ポテンシャルモデルでは、最適な方向を常に探索することのできる反面、同じ目的地の歩行者については経路が一意に決まってしまうという問題がある。

2.3 観測値を用いた高精度化

GPS 等、様々な観測機器を用いた交通行動モデルへのフィードバックは 2000 年以降、牧村 (2000)、兵藤 (2002)、山田ら (2010) を始めとして多数見られるようになってきた。さらに最近では、携帯電話の基地局情報などを用いて、よりダイナミックな人の分布の時間変動の推定を行うようなものも出てきている (Ratti et al. (2006) など)。さらにこうしたデータ取得・観測が負担になりすぎないように、Opportunistic な観測を目指したものとして、Lane et al. (2008) がある。

また、観測値を用いた高精度化のアプローチとして、シミュレーションモデルのパラメータを補正するという研究もいくつか行われている。森地ら (1987) はロジットモデルで構築される非集計行動モデルのパラメータに対し、観測値としていくらかの集計データが得られたときにベイズ推定を用いてパラメータの最適化を行っている。

推定方法としてデータ同化手法を用いたものもいくつか研究されている。Nanthawichit et al. (2003) や Herrera and Bayen (2008) は高速道路の自動車容量と速度を推定する際に、センサを搭載する車から得られる個別の位置と速度情報を用いてデータ同化を行い精度を高める手法を提案している。また、Cheng et al. (2006) は携帯電話使用履歴を観測データとしてデータ同化し、同様に自動車容量と速度を推定するモデルを提案しているなど、近年データ同化を用いた推定・予測が多く行われるようになってきている。

2.4 現状の課題と本研究の位置づけ

以上のように、人の流れ情報を取得するために、計測技術と推定技術の両面から研究がなされ、発展してきている。しかし、それぞれいくつか課題がある。計測技術については、広範囲にすべてを計測することが困難であるという点が挙げられる。一人一人の軌跡情報

を取得したい時、カメラやレーザスキャナセンサで局所的にデータ取得したり、GPS データやアンケート調査でサンプリング調査したりすることはできるが、どちらも領域や人数に観測できないギャップができてしまう。

推定技術についても、選択行動モデル、流体モデル、離散モデル、セルオートマトンモデルなど、様々な観点から研究がなされてきている。しかし、人の行動の多様性からモデルで完璧に表現しきるのは難しく、モデル自体に誤差が生じてしまう。また、シミュレーションモデルの構築には、多くの場合過去データからパラメータの算出等が必要であり、モデル自体は過去の状態を最適に表現したものとなり、現状とずれが生じてしまう可能性も大きい。実用的な側面から見たとき、パラメータの同定が難しく、満足する値を算出するのに多大な労力と時間を要するという問題もある。

そうした中、観測値によってシミュレーションモデルの精度を向上させる研究や、データ同化手法を用いて推定値を補正する研究も行われてきている。しかし、それらの研究は、推定する変数が交通流量や速度、シミュレーションモデルの持つパラメータ値などの連続な数値であることがほとんどである。また、シミュレーションモデルの構造を限定し、そのモデルに対して最適な手法を提案することがほとんどである。

このような背景から、本研究では計測と推定という両面を融合し、より高精度に人の流れ全体を推定することを目指す。観測値でシミュレーションモデルの精度向上を目指す他の研究との違いとして、主に以下の2点が上げられる。

- 推定する対象が移動軌跡の確率分布であり、これは離散的な変数に対する確率分布であり、連続でない、微分できない、分散が計算できないものである。
- シミュレーションモデルを限定しない。

移動軌跡の確率分布という、離散的で分散が計算できないような対象に対してデータ同化手法を用いる研究はまだ行われていない。また、その確率分布は変数（移動軌跡）が離散的なため直接求めることはできないので、変数（移動軌跡）をサンプリングして、確率つきサンプリング集合を生成することで代替する。このような形にすることで、2点目のシミュレーションモデルに依存しない手法の提案が行える。これは各シミュレーションモデルに最適なものではない可能性があるが、実用的な面で考えたときに、様々な種類のモデルを適用できることが大きな利点を持つと考えられる。

参考文献

- Snidaro, L., 2005. Video security for ambient intelligence. IEEE Trans. SMC, Part A, vol. 35, no. 1, pp.133-144.
- Yang, D. B., 2003. Counting people in crowds with a real-time network of simple image sensors. IEEE Conference on Computer Vision, vol. 1, pp. 122-129.
- Shao, X., 2008. Tracking a Variable Number of Pedestrians in Crowded Scenes by using Laser Range Scanners. IEEE International Conference on Systems, Man and Cybernetics (SMC 2008), pp. 1545-1551.
- 岡部博志, 福田 敦, 石坂哲宏, 2005. プローブカーシステムを活用した開発途上国における都市交通データ収集の可能性に関する研究 タイ国バンコクを対象として. 土木計画学研究・講演集 Vol: 31, pp.191-195.
- 松本修一, 2007. GPS 携帯のプローブパーソン調査への適応性に関する基礎的研究. 土木計画学研究講演集.
- 塚口博司, 松田浩一郎, 2002. 歩行者の絵色選択行動分析. 土木学会論文集, 2002. 709 IV-56: p.p.117-126.
- 横山秀史, 目黒公郎, 片山恒雄, 1995. 避難行動解析へのポテンシャルモデルの応用. 土木学会論文集, 1995. 507(I-31): p.p.225-232.
- Antonini, G., 2006. Dynamic Cell Transmission-Based Pedestrian Model with Multidirectional Flows and Strategic Route Choices. Transportation Research Part B 40 (2006), pp. 667-687.
- 東京都市交通計画協議会, 2001, 平成 10 年東京都市圏パーソントリップ調査（総合都市交通体系調査）報告書.
- 山形耕一, 1984, パーソントリップ調査における調査不能誤差と層別拡大に関する研究, 土木学会論文集, No. 343, pp.121-129.

- Mc Fadden, D., 1974, Conditional Logit Analysis of Qualitative Choice Behavior, In *Frontiers in Econometrics*, ed. P. Zarembka, Academic Press, pp.105-142.
- Ben-Akiva, M. E., 1974, Structure of Passenger Travel Demand Models, *Transportation Research Record* 526, pp.26-42.
- 土木学会, 1998, 交通ネットワークの均衡分析-最新の理論と解法-, 丸善.
- 円山琢也, 原田昇, 太田勝敏, 2002, 大規模都市圏への交通需要統合型ネットワーク均衡モデルの適用, *土木計画学・論文集*, Vol.19, No.3, pp.551-560.
- 金森亮, 森川高行, 山本俊行, 三輪富生, 2009, 総合交通戦略の策定に向けた統合型交通需要予測モデルの開発, *土木学会論文集 D*, Vol. 65, No. 4, pp. 503-518.
- M. J. Lighthill and G. B. Whitham, 1955, On kinematic waves. I. Flood movement in long rivers, *Proc. R. Soc. Lond. A*, Vol. 229, No. 1178, pp.281-316.
- M. J. Lighthill and G. B. Whitham, 1955, On kinematic waves. II. Traffic Flow on Long Crowded Roads, *Proc. R. Soc. Lond. A*, Vol. 229, No. 1178, pp.317-345.
- Paul I. Richards, 1956, Shock Waves on the Highway, *Operations Research*, Vol. 4, No. 1, pp.42-51.
- Hall, M. D., Vliet, D. Van, and Willumsen, L. G., 1980, A simulation assignment model for the evaluation of traffic management schemes, *Traffic Eng. & Control*, Vol. 21, No. 4, pp.168-176.
- S. A. H. Algadhi and H. S. Mahmassani, 1990, Modelling crowd behavior and movement: application to Makkah pilgrimage, *Transportation and Traffic Theory*, pp. 32-38.
- 中祐一郎, 1978, 鉄道駅における旅客の交錯流動に関する研究, *鉄道技術研究報告*, 1079, pp.1-131.
- Hughes, R. L., 2002, A continuum theory for the flow of pedestrians, *Transportation Research Part B*, 36, pp.507-535.
- L. A. Pipes, 1953, An Operational Analysis of Traffic Dynamics, *Journal of Applied Physics*, Vol.24, No.3, pp.274-281.
- R. E. Chandler, R. Herman, and E. W. Montroll, 1958, Traffic Dynamics: Studies in Car Following, *Operations Research*, Vol.6, pp.165-184.

- E. Kometani and T. Sasaki, 1959, A safety index for traffic with linear spacing, *Operations Research*, Vol.7, pp.704-720.
- Rathi, A. K. and A. J. Santiago, 1992, Urban Network Traffic Simulation: TRAF-NETSIM Program, *Journal of Transportation Engineering*, Vol.116, No.6, pp.734-743.
- 森津秀夫, 大原竜也, 多田典史, 井上琢弥, 1991, 経路誘導による交通ネットワークフローの変化に関する分析, *土木計画学研究・論文集*, No.9, pp.37-44.
- 井上博司, 1993, 道路網における交通流動の動的シミュレーション手法, *土木学会論文集*, No.470/IV-20, pp.87-95.
- 吉井稔雄, 桑原雅夫, 森田緯之, 1995, 都市内高速道路における過飽和ネットワークシミュレーションモデルの開発, *交通工学*, Vol.30, No.1, pp.33-41.
- Helbing, D., Molnar, P., 1995, Social force model for pedestrian dynamics, *Physical Review E*, Vol.51, No.5, pp.4282-4286.
- Teknomo and Kardi, 2006, Application of microscopic pedestrian simulation model, *Transportation Research Part F*, Vol.9, pp.15-27.
- Hoogendoorn, S. P. and P. H. L. Bovy, 2003, Simulation of Pedestrian Flows by Optimal Control and Differential Games, *Optimal control applications and methods*, Vol.24, pp.153-172.
- 浅野美帆, 桑原雅夫, 2007, 先読み行動を考慮した歩行者交通流シミュレーション, *生産研究*, Vol.59, No.3, pp.38-41.
- K. Nagel and M. Schreckenberg, 1992, A cellular automaton model for freeway traffic, *J. Phys. I France*, Vol.2, pp.2221-2229.
- C. Burstedde, K. Klauck, A. Schadschneider, J. Zittartz, 2001, Simulation of pedestrian dynamics using a two-dimensional cellular automaton, *Physica A*, Vol.295, pp.507-525.
- Dial, R.B., 1971. A probabilistic multipath traffic assignment model which obviates path enumeration. *Transportation Research*, 5(2), p. 83-111.
- 朝倉康夫, 羽藤英二, 2000, 交通ネットワーク上の経路選択行動：観測と理論, *土木学会論文集*, No.660, IV-49, pp.3-13.

- 兵藤哲郎, 遠藤弘太郎, 萩野保克, 西隆太, 2009, Path Size Dial Logit モデルの提案とその適用可能性, 交通工学, Vol.44, No.4, pp.66-75.
- De la Barra, T., B. Perez, and J. Anez, 1993. Multidimensional Path Search and Assignment. Proceedings of the 21st PTRC Summer Meeting, pp. 307-319.
- Ben-Akiva, M., M.J. Bergman, A.J. Daly, and R. Ramaswamy, 1984. Modelling Inter Urban Route Choice Behaviour. Proceedings of the 9th International Symposium on Transportation and Traffic Theory, VNU Press, Utrecht, pp. 299-330.
- Lombard, K., Church, R. L., 1993, The gateway shortest path problem: Generating alternative routes for a corridor location problem, Geographical Systems, 1, pp. 25-45.
- Akgun, V., Erkut, E., and Batta, R., 2000, On finding dissimilar paths, European Journal of Operational research 121, pp. 232- 246.
- 木下統英, 片山哲平, 小堀雅嗣, 朝倉康夫, 2003, 道路ネットワーク上での経路選択枝集合の生成手法, 土木計画学研究・講演集, Vol.28, VIII(275).
- 牧村和彦, 2000, 高度情報機器を用いた歩行者行動モニタリングと移動支援, 交通工学, Vol. 35, No. 4, pp. 40-45.
- 兵藤哲郎, 2002, 交通需要予測の課題と展望－新しいデータ, 新しいモデル, そして需要予測不信－, 「交通工学」企画特集「観測と理論」
- 山田孝太郎, 羽藤英二, 2010, 位置データ指向型交通機関選択モデルのパラメータ特性, 第30回交通工学研究発表会論文集, pp. 225-228.
- Ratti, C., Pulselli, R. M., Williams, S. and Frenchman, D., 2006, Mobile Landscapes: using location data from cell phones for urban analysis, Environment and Planning B: Planning and Design, Vol. 33, No. 5, pp. 727-748.
- Lane, N. D., Eisenman, S. B., Musolesi, M., Miluzzo, E. and Campbell, A. T., 2008, Urban Sensing Systems: Opportunistic or Participatory?, Proceedings of 9th Workshop Mobile Computing Systems and Applications (WMCSA 08), ACM Press.
- 森地茂, 屋井鉄雄, 平井節生, 1987, 個人データと集計データとの統合利用によるモデル構築方法, 土木計画学研究・論文集, No.5, pp.51-58.

- C. Nanthawichit, T. Nakatsuji, and H. Suzuki, 2003, Application of probe-vehicle data for real-time traffic-state estimation and short-term travel-time prediction on a freeway, Transportation Research Record, Vol.1855, pp.49-59.
- Herrera, J. C. and Bayen, A. M., 2008, Traffic Flow Reconstruction Using Mobile Sensors and Loop Detector Data, Transportation Research Board 87th Annual Meeting.
- P. Cheng, Z. Qiu, B. Ran, 2006, Particle Filter Based Traffic State Estimation Using Cell Phone Network Data, IEEE ITSC, pp.1047-1052.

第3章 提案する人の流れ推定手法のフレームワーク

3.1 提案する人の流れ推定手法の全体フレーム

以上を踏まえ、本研究の目的は、計測とシミュレーションモデルを統合することでより高精度な人の流れの推定を行うことである。図 3.1 のように、観測値、シミュレーションモデルともに、それだけでは全体の人々の流れ情報を取得するのに不十分である。そこで、本研究では、データ同化手法を用い、シミュレーションモデルの推定値を観測値で補正し、より高精度な推定を行うことを目指す。

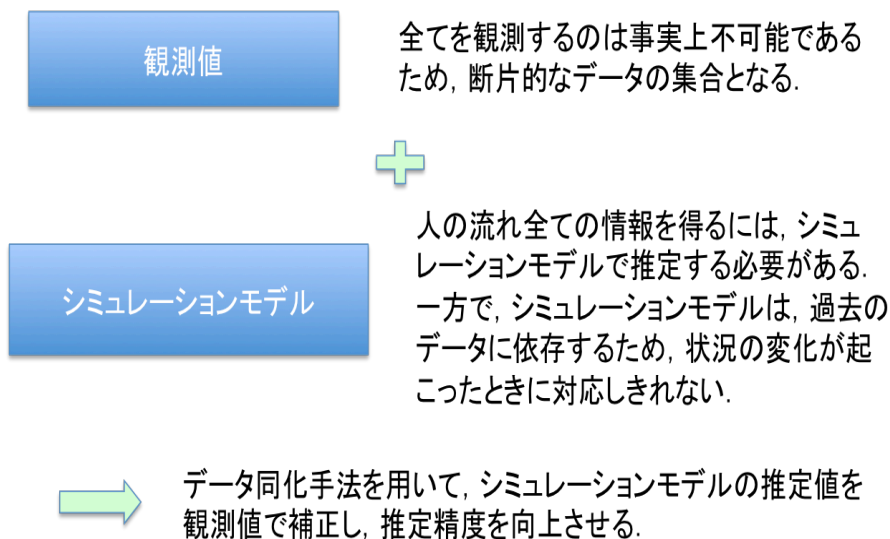


図 3.1: 研究の目的

提案するデータ同化手法を用いた人の流れ推定手法の全体フレームを図 3.2 に示す。まず、対象とする領域での人の流れを推定するモデルを構築する。このとき推定モデルは、最適解のみを推定するのではなく、確率的にシミュレーションできるモデルを構築する。次に、構築したモデルを用い、移動軌跡候補（パーティクル）を確率的に多数生成する。ここでパーティクルを確率的に多数生成できるようにするために、推定モデルが確率的にシミュレーションできるモデルに限定される。そして最後に、データ同化手法のひとつであ

るパーティクルフィルタ法を用い、取得できた観測値によって各パーティクルを評価し、重み付けする。最終的に得られる結果は、この観測値によって評価された各パーティクルの重みで、推定モデルによる各移動軌跡の確率分布を観測値で補正した確率分布を得ることに等しくなる。

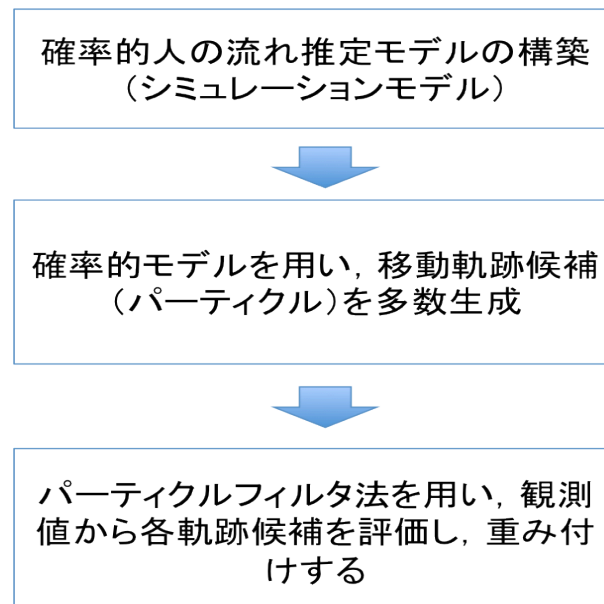


図 3.2: 全体フレーム

3.2 データ同化手法

次に、観測値と歩行者モデルの統合に用いるデータ同化手法について簡単に述べる。

データ同化とは、地球科学の分野において発展してきた手法で (Wunsch (1996), Daley (1991), Barnett (2002), 図 3.3), 観測値を用いて数値モデルの再現性を高める手法である。簡単に言えば、モデルに実際の観測値を入力して、より現実に近い結果が出るようにすることを目指す。

地球科学においては、非線形性の高い自然現象を数値モデルによって再現する手法がある。特に気象学では、これが天気予報に大きく貢献している。データ同化は、例えば数値気象予報モデルに気温・気圧・湿度・風向・風速などのデータを入力し、専門的には「初期場」と呼ばれる、物理的パラメータの空間的分布状態を作り出す作業である。ただ、入力されるデータは空間的に偏りが大きいため、データの少ないところでは精度が低くなる。

図 3.4 はモデルと観測データと同化後の推定値の関係を表現したものである。クロ曲線のようにモデルによって推定値の候補群を生成し、候補群から観測値に最も適する候補を

選び、最終的な推定値とすることで、モデルによる推定値と観測値の同化を行う。図 3.5 はデータ同化によって得られる効果の可能性を示したものである。青線が時点ごとのシミュレーションモデルによる予測値、水色の丸が観測値、赤い丸が補正值、赤線がデータ同化手法による補正值であり、観測値で予測値を補正することで、傾向が同じであっても結果が大きく違ってくる場合があることを示している。

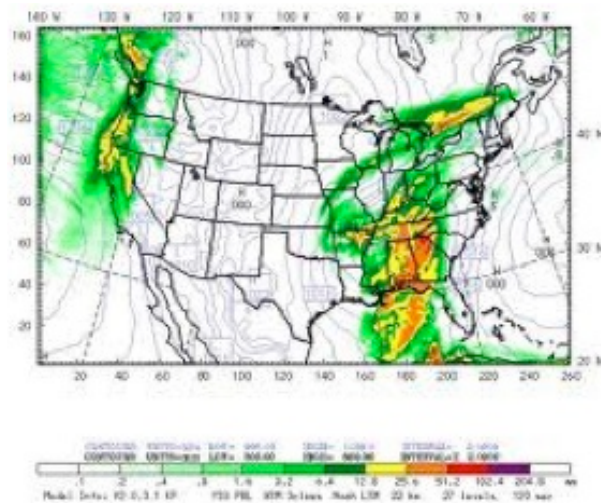


図 3.3: データ同化の気象への適用例

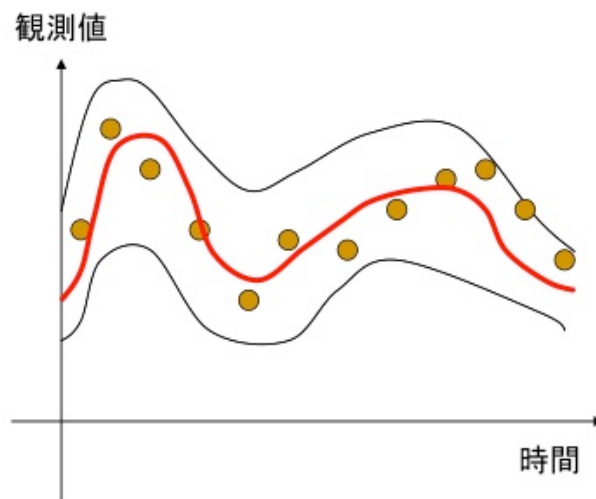


図 3.4: データ同化のイメージ

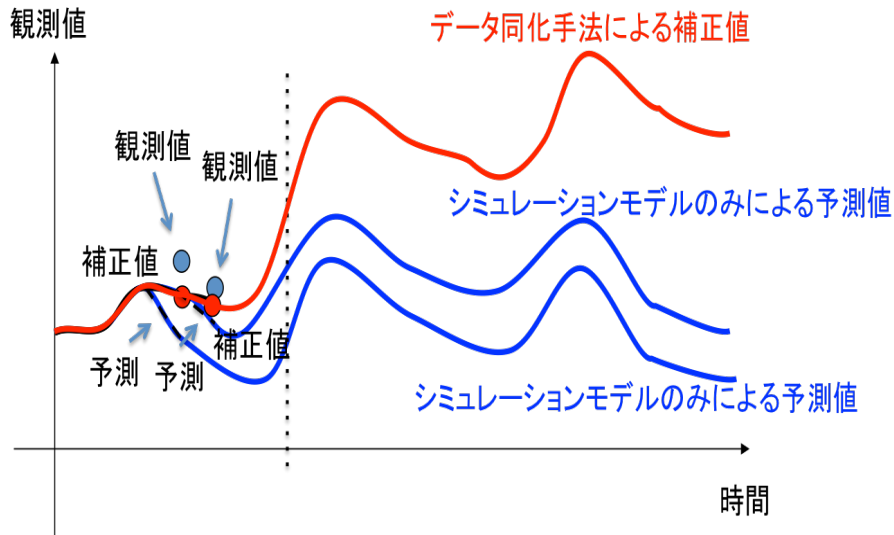


図 3.5: データ同化の効果

データ同化手法には、推定値と観測値を補正する手法として、様々な手法が提案されており、直接挿入やナッジング、3次元データ同化手法、4次元データ同化手法などがある。直接挿入は観測が得られている部分に観測値を直接挿入する手法である。ナッジングとは観測が得られている部分について、推定値を補正するものである。また、3次元データ同化手法は誤差分布や確率分布を用いて、理論的により高精度な推定を行うもので、最適内挿法や3次元変分法などが提案されている。4次元データ同化手法は誤差分布や確率分布を過去の情報から変化させていくもので、予測値に過去の情報を加味することができる。4次元データ同化手法には、カルマンフィルタなどの逐次データ同化手法と、4次元変分法などが提案されている。

直接挿入は前述のように、観測値を直接挿入するだけなので、詳しい説明は省く。ナッジングという手法は、観測値と推定値による観測値の予測値の差にナッジング係数を乗算することで、推定値を補正していく。これらの方法は簡便に利用できるが、観測値に近づけているだけで推定値が正しい値に近づいている保証がない。

そこで、3次元データ同化手法では、推定値と観測値の誤差分布や確率分布を用いて理論的に高精度な推定を行うことを目指している。3次元データ同化では、観測値を条件とした時の求めたい x の確率分布をベイズの定理から求め、その確率分布を最大化する x を算出することで推定値を補正する (Lorenz (1986))。確率分布を最大にする x を求める過程で、最適内挿法では観測モデルが線形の場合に対し、行列計算で推定値の補正を行う。3次元変分法では、観測モデルが非線形の場合にも対応しており、共役勾配法や準ニュートン法などを用いて最大値の探索を行う (Lorenz et al. (2000))。

4次元データ同化の逐次データ同化手法の提案も対象となる関数の性質を拡張する形で多く行われている。まず、Kalman (1960) はモデルが線形でかつ誤差が正規分布で表される際の補正方法としてカルマンフィルタを提案している。カルマンフィルタは、1時点前の状態から線形モデルを用いて予測を行い、その予測と観測値の残差の共分散を求め、残差の期待値が最小二乗法で最適となるように予測を更新する。それによって推定値の補正ができるが、線形モデルと誤差が正規分布という制約がある。

拡張カルマンフィルタ (片山 (1983)) は、カルマンフィルタを非線形なモデルにも適用できるように拡張したものである。ただし、モデルが微分可能であるという制約がつき、誤差については同様に正規分布に従っている必要がある。

アンサンブルカルマンフィルタ (Evensen (1994), Evensen (2003)) は、状態変数の確率分布を実現値集合であるアンサンブルによって保持し、観測を得るごとに、観測モデルとアンサンブルからカルマンフィルタを用いて補正する。状態変数の確率分布をサンプリングで代替することで、モデルが微分可能でない場合でも適用できる。

パーティクルフィルタ (Kitagawa (1996), Gordon et al. (1993)) も、アンサンブルカルマンフィルタと同様にサンプリングを行って非線形なモデルに適用できるようにした手法である。アンサンブルカルマンフィルタとパーティクルフィルタの違いは、アンサンブルカルマンフィルタがカルマンゲインをもとにフィルタリングを行うのに対し、パーティクルフィルタでは重み関数によるリサンプリングでフィルタリングを行うことである。

3.3 提案手法の方法論的詳細

本研究では最終的に、推定したい移動に対して、候補となる移動軌跡を列举し、各移動軌跡の選択確率を求めることが目的である。このとき、人の移動を推定する際、考えられる移動軌跡候補を全て考慮した上で推定することは現実的でなく、確率的にサンプリングして代替する必要がある。サンプリングして、各候補の選択確率を観測値から尤もらしい値を算出して与えることを繰り返す手法にパーティクルフィルタ法があり、本研究では移動軌跡の確率分布推定にパーティクルフィルタ法を用いる。

移動軌跡の確率分布を求めるのが目的となるが、求める確率分布の変数が連続でなく離散的であるため、分散が存在しない。この点が、交通分野にデータ同化手法を適用した他の研究よりも困難な点として挙げられる。そのため、離散的にサンプリングされた移動軌跡候補とシミュレーションモデルから算出されたそれらの選択確率に対して最適化計算を行い、観測値に即した選択確率を求める必要がある。この最適化計算は枠組みとしては3

次元データ同化と同様であり、本研究でも最適内挿法と同様の計算方法を用いて選択確率を計算する。

離散的な変数を求める際にシミュレーションモデルと観測値を融合して最適値を算出する問題に対して、隠れマルコフモデルが良く用いられている。隠れマルコフモデルでは、有限個の離散的な変数に対し、初期状態分布と、それらの変数を遷移する状態遷移確率行列をモデルとして与え、観測値が得られた際に初期状態分布、状態遷移確率行列、観測確率分布の最適値を算出できる。しかし、モデルとして遷移状態と状態遷移確率分布を与えなければならず、シミュレーションモデルの構造が限定されてしまう。例えば、歩行者空間の人の流れの推定に対して、セルオートマトンモデルを用いてセル間の移動を遷移確率行列を用いて表したとすると、比較的容易に適用できる。しかし、離散モデルを用いたい場合には、移動速度や向き等も状態となり、有限個の状態で表すのが非現実的になることも多いと考えられる。また、そもそもシミュレーションモデルを有限個の状態と状態遷移確率行列に変換する必要があることは、実用化を考えるとマイナスとなる。それに対し、パーティクルフィルタ法を用いることで、移動軌跡の列挙および選択確率の付与は、事前に用意したシミュレーションモデルで、乱数を用いる等で行うことができる。これによりシミュレーションモデルをそのまま当てはめて用いることができ、適用への障害が少なく済む。

また、サンプリング近似をアンサンブルカルマンフィルタで行うという方法も考えられる。パーティクルフィルタ法を用いることで起こりうる問題のひとつである縮退が起こりにくいこと、サンプル数がより少なくて済むことなどが利点として挙げられる。しかし、これらの利点はサンプリング近似した集合に対して平均や分散を計算できることからきており、本研究で求める移動軌跡の確率分布はそうした統計値が計算できないため、そのまま適用することは難しい。歩行者空間での人の流れを推定する際などに、その移動軌跡が連続的な変数として扱うことができる状態で、そうした統計値を定義することができる場合に限り適用可能となる。本研究では、状況を限定せずに適用できるフレームワークの提案をするため、そうした状況には触れないが、アンサンブルカルマンフィルタを用いることができる場合はその方が計算コスト、堅牢性などに優れていると言える。

以上から、本研究では、移動軌跡の確率分布自体はパーティクルフィルタ法の枠組みで逐次計算していき、各パーティクルの選択確率は最適内挿法を用いて算出する。つまり、細かく見ると2つのデータ同化手法が組合わさっていると言える。これは求めたい変数が離散的な確率分布であり、かつ、離散的な変数を全列挙して推定することが困難なことからきている。交通分野でこうした問題を解いた例はまだなく、本研究の独自性の一つと言える。

以下、本研究で用いるパーティクルフィルタ法と最適内挿法について述べる。

3.3.1 パーティクルフィルタ法

パーティクルフィルタは、元々制御理論の分野で研究されてきた理論である時系列フィルタリングの一手法である。過去に観測された信号系列から、未来の値を予測したり、ノイズを含む観測値から対象の状態を推定したりする問題などが取り扱われてきた。パーティクルフィルタ法は、その中でも、非線形、非ガウス型の状態空間モデルに対して、効率よく状態を推定する手法である。近年コンピュータビジョンの分野でも、対象追跡や輪郭追跡などにおいて注目されており、パーティクルフィルタを用いた追跡手法が数多く提案されている。

以下、パーティクルフィルタ法について詳しく述べる。まず、パーティクルフィルタ法は求めたい時点 t での状態 X_t を一時点まえの状態 X_{t-1} から求める状態遷移モデルと、時点 t での観測値 Z_t について状態 X_t から求める観測モデルの2つの段階から成っている。状態遷移モデル f_t は、

$$X_t = f_t(X_{t-1}, v_{t-1}) \quad (3.1)$$

と表される。ただし、 v_{t-1} はシステムノイズである。また、観測モデル h_t は、

$$Z_t = h_t(X_t, w_t) \quad (3.2)$$

と表される。ただし、 w_t は観測ノイズである。

パーティクルフィルタの基本的アイデアは、事後分布 $p(x_t|Z_t)$ 、及び事前分布 $p(x_t|Z_{t-1})$ をこれらの条件付き分布に従う多数のサンプル（粒子）を用いて近似するという点である。サンプリングによって、非線形なモデルを扱うことができる。事前分布 $p(x_t|Z_{t-1})$ に従うサンプル集合 $S_{t|t-1} = \{s_{t|t-1}^{(1)}, \dots, s_{t|t-1}^{(N)}\}$ 、事後分布 $p(x_t|Z_t)$ に従うサンプル集合を $S_{t|t} = \{s_{t|t}^{(1)}, \dots, s_{t|t}^{(N)}\}$ とする。このとき事前分布、事後分布はサンプル集合を用いて以下のように近似される。

$$p(x_t|Y_{t-1}) \approx \frac{1}{N} \sum_{i=1}^N \delta(x_t - s_{t|t-1}^{(i)}) \quad (3.3)$$

$$p(x_t|Y_t) \approx \frac{1}{N} \sum_{i=1}^N \delta(x_t - s_{t|t}^{(i)}) \quad (3.4)$$

ただし、 $\delta(\cdot)$ はディラックのデルタ関数である。

パーティクルフィルタのポイントは、二種類のサンプル集合 $S_{t|t-1}, S_{t|t}$ をそれぞれ事前分布 $p(X_t|Z_{t-1})$ 、事後分布 $p(X_t|Z_t)$ に従って生成することである。これらの分布は式3.1、式

3.2に従って推定され、パーティクルフィルタでは、これらの式を、次のような手順でサンプルに適用しながら、逐次的にサンプル集合を生成する。図 3.6 にサンプル集合の生成の概念図を示す。

1. 初期化： $i = 1, \dots, N$ について $s_{0|0}^{(i)} \sim p_0(x)$ を生成する。ただし、 $p_0(x)$ はあらかじめ与えた初期分布である。 $t = 1$ として以下の手順を実行する

2. 予測：各サンプルについて、時刻 t における予測サンプル $s_{t|t-1}^{(i)} \sim p(X_t|X_{t-1} = s_{t-1|t-1}^{(i)})$ を次の手順で生成する。

(a) $i = 1, \dots, N$ について、次元の乱数としてシステムノイズ $v_t^{(i)} \sim q(v)$ を生成する。ただし、 $q(-)$ はあらかじめ設定したシステムノイズ v の分布である。

(b) 式 (2) のシステムモデル f_t に従って時刻の各サンプル $s_{t-1|t-1}^{(i)}$ を遷移させて予測サンプルを生成する。

$$s_{t|t-1}^{(i)} = f_t(s_{t-1|t-1}^{(i)}, v_t^{(i)}) \quad (3.5)$$

3. 尤度推定：各予測サンプルについて、重みを推定する。

$$\pi_t^{(i)} = \frac{p(Z_t|X_t = s_{t|t-1}^{(i)})}{\sum_{i=1}^N p(Z_t|X_t = s_{t|t-1}^{(i)})} \quad (3.6)$$

$p(Z_t|x_t = s_{t|t-1}^{(i)})$ は、状態 x_t が $s_{t|t-1}^{(i)}$ であったときに、観測 Z_t を得る確率（尤度）であり、式 3.2 の観測モデル h_t に従って推定する。

4. フィルタ： $S_{t|t-1} = \{s_{t|t-1}^{(1)}, \dots, s_{t|t-1}^{(N)}\}$ からそれぞれ $s_{t|t-1}^{(i)}$ を重み $\pi_t^{(i)}$ に比例する割合で N 個復元抽出し、 $S_{t|t} = \{s_{t|t}^{(1)}, \dots, s_{t|t}^{(N)}\}$ とする。このとき、 $s_{t|t}^{(i)}$ は、重み $\phi_t^{(i)}$ の確率に従ってランダムに抽出してもよく、また重み $\pi_t^{(i)}$ が大きい順に $s_{t|t}^{(i)}$ を $[\pi_t^{(i)} N]$ 回ずつ、 N 個になるまで順に選択してもよい。 $t = t + 1$ として、2 の予測へ。

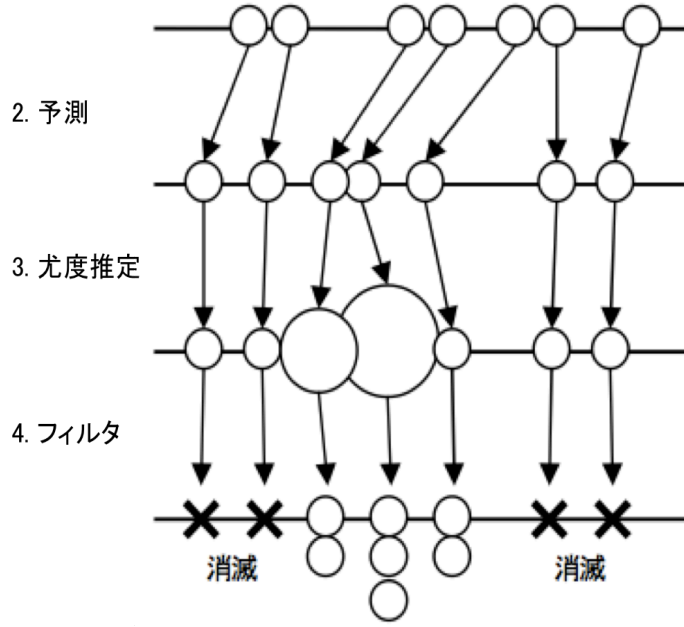


図 3.6: サンプル集合生成の概念図

またある時刻における推定値 \hat{X}_t は次式のように尤度による重み付平均によって期待値として求めることができる。

$$\hat{X}_t = E(X_t) \approx \frac{1}{N} \sum_{n=1}^N s_{t|t}^{(n)} = \sum_{n=1}^N \pi_t^n s_{t|t-1}^{(n)} \quad (3.7)$$

3.3.2 最適内挿法

最適内挿法は観測モデルが線形で、

$$Z_t = h_t(X_t, w_t) = H_t X_t + w_t \quad (3.8)$$

と表される時に X_t を最適化する手法である。

時点 t において、求めたい X_t と観測値 Z_t があるとき、ベイズの定理より、

$$p(X_t|Z_t) = \frac{p(Z_t|X_t)p(X_t)}{\int p(Z_t|X_t)p(X_t)dX_t} \quad (3.9)$$

となる。ここで、 $p(X_t)$ が平均 $X_{t,b}$ ($X_{t,b}$ はシミュレーションモデルによる推定値)、分散 V 、 $p(Z_t|X_t)$ が平均 $H_t X_t$ (観測モデルによる推定値)、分散 R の正規分布に従うと仮定すると、

$$p(X_t) = \frac{1}{\sqrt{(2\pi)^k |V|}} \exp\left(-\frac{1}{2}(X_t - X_{t,b})^T V^{-1}(X_t - X_{t,b})\right) \quad (3.10)$$

$$p(Z_t|X_t) = \frac{1}{\sqrt{(2\pi)^k |R|}} \exp\left(-\frac{1}{2}(Z_t - H_t X_t)^T R^{-1}(Z_t - H_t X_t)\right) \quad (3.11)$$

となり、これと式 3.9 から、 $p(X_t|Z_t)$ を最大化する X_t は、

$$J = (X_t - X_{t,b})^T V^{-1} (X_t - X_{t,b}) + (Z_t - H_t X_t)^T R^{-1} (Z_t - H_t X_t) \quad (3.12)$$

を最小にする X_t となる。この最小値をとる X_t は、

$$\frac{\partial J}{\partial X} = V^{-1} (X_t - X_{t,b}) - H_t^T R^{-1} (Z_t - H_t X_t) = 0 \quad (3.13)$$

の解となるので、これを解くと、

$$X_t = X_{t,b} + V H_t^T (H_t V H_t^T + R)^{-1} (Z_t - H_t X_{t,b}) \quad (3.14)$$

となる。

3.3.3 人の流れ推定への適用

本節では、パーティクルフィルタ法を人の流れ推定にどう適用するかについて述べる。

まず、図 3.7 に推定手法のイメージを示す。図の上部から対象領域に入った人がいたとして、黒点線と赤線の 3 パターンの移動軌跡候補を生成したとする。このとき図の右部分でカメラで観測を得たとすると、赤線の移動軌跡が選択された確率が一番高いと考えられる。対象人数が一人であれば以上のように推定すれば問題ないが、自由空間、ネットワーク空間ともに対象人数は複数人であるため、集計観測値による補正が必要となる。

図 3.8 と図 3.9 は集計観測値に対してどう適用するかを示したイメージである。まず、図 3.8 のように、ある移動を推定する際に、シミュレーションモデルを用いて複数の移動軌跡の候補（パーティクル）を生成する。そして例えば、図 3.9 のように道路利用の観測が少ない、駅利用数の観測が多いなどの観測値を取得できた際に、それに合わせてパーティクルの重みを更新する。図 3.9 の青線、青点線の太さはパーティクルの重みを表しており、観測値が少ないパーティクルは重みが小さく、観測値の多いパーティクルは重みが多くなる。以下、数式を用いて詳しく述べる。

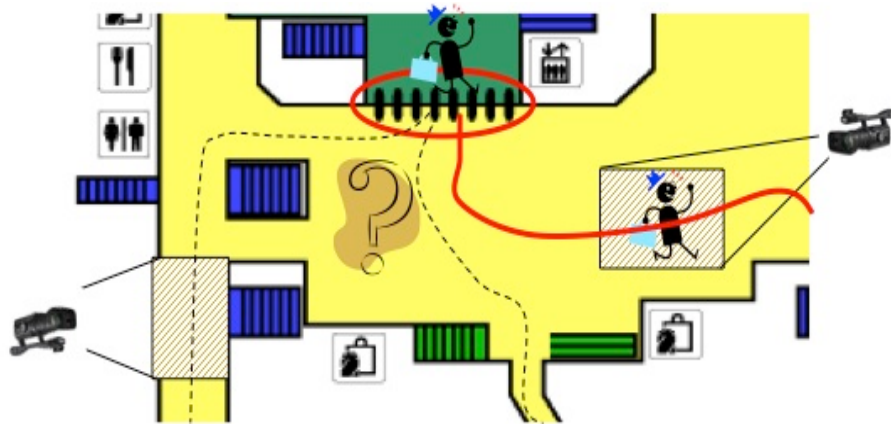


図 3.7: 推定手法のイメージ



図 3.8: パーティクルの生成



図 3.9: パーティクルの重み更新のイメージ

状態遷移モデルとして、構築した確率的なシミュレーションモデルを適用する。ある時点 t での状態 X_t は一時点前 $t-1$ から t までの移動軌跡の確率分布となる。 i 番目の人に対する移動軌跡を $T_{i,t-1 \rightarrow t}$ とすると、

$$X_t = p(T_{i,t-1 \rightarrow t}) \quad (3.15)$$

となる。 $T_{i,t-1 \rightarrow t}$ は離散的なので、 N_t 個サンプリングされた重みつき軌跡集合

$$TS_{i,t-1 \rightarrow t} = \{\{t_{i,t-1 \rightarrow t}^{(1)}, w_{i,t}^{(1)}\}, \dots, \{t_{i,t-1 \rightarrow t}^{(N_t)}, w_{i,t}^{(N_t)}\}\} \quad (3.16)$$

の推定をすることになる。

次に、各移動軌跡候補の重み $w_{i,t}^{(j)}$ を観測値を用いて更新する。この重みを全 N_p 人分並べた行列 $W_t = \{w_{1,t}^{(1)}, \dots, w_{1,t}^{(N_t)}, w_{2,t}^{(1)}, \dots, w_{2,t}^{(N_t)}, \dots, w_{N_p,t}^{(1)}, \dots, w_{N_p,t}^{(N_t)}\}^T$ と、各移動軌跡候補 $t_{i,t-1 \rightarrow t}$ が各観測地点で観測されるかどうかを表す行列 H_t を用いて、観測モデルは、

$$Z_t = H_t W_t + \epsilon \quad (3.17)$$

と表すことができる。ここで、 H_t は行数が観測地点数、列数が重み行列 W_t の行数となり、 δ_{ij} を、

$$\delta_{ij} = \begin{cases} 1(i \text{ 番目の重みに対応する移動軌跡が観測地点 } j \text{ で観測されるとき}) \\ 0(\text{それ以外}) \end{cases} \quad (3.18)$$

とすると、 H_t の i 行 j 列の要素は δ_{ji} となる。そして、最適内挿法の式 3.12 の X を W で置き換えて、この J を最小にする W を求める。しかし、最適内挿法と違って、人ごとに $\sum_j w_{i,t}^{(j)} = 1$ とならなければならないので、その制約のもと、 J を最小にすることになる。これは、ラグランジュの未定乗数法より、

$$J = (W_t - W_{t,b})^T V^{-1} (W_t - W_{t,b}) + (Z_t - H_t W_t)^T R^{-1} (Z_t - H_t W_t) - \sum_i \lambda_i (1 - \sum_j w_{i,t}^{(j)}) \quad (3.19)$$

を最小にすることで W_t を算出することができる。この最小値をとる W_t は、

$$\frac{\partial J}{\partial W} = V^{-1} (W_t - W_{t,b}) - H_t^T R^{-1} (Z_t - H_t W_t) - (H_t^T R^{-1} H_t + V^{-1})^{-1} A \Lambda = 0 \quad (3.20)$$

$$\sum_j w_{i,t}^{(j)} = 1 (i = 1, 2, \dots, N_p) \quad (3.21)$$

の解となる。ここで,

$$A = \begin{pmatrix} 1 & 0 & & 0 \\ \vdots & \vdots & \dots & \vdots \\ 1 & 0 & & 0 \\ 0 & 1 & & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 1 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & & 1 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & & 1 \end{pmatrix} \quad (3.22)$$

$$\Lambda = (\lambda_1, \lambda_2, \dots, \lambda_{N_p})^T \quad (3.23)$$

である。式 3.20 より,

$$W_t = W_{t,b} + V H_t^T (H_t V H_t^T + R)^{-1} (Z - H_t W_{t,b}) + (H_t^T R^{-1} H_t + V^{-1})^{-1} A \Lambda \quad (3.24)$$

となり, これと式 3.21 から Λ を求め, 式 3.24 に代入すれば, W_t を求めることができる。

こうして, 時点 t の移動軌跡のサンプリングによる確率分布を求めることができる。以降, シミュレーションモデルでリサンプリングを行い, 逐次的に推定していく。

最後にリサンプリングの方法について述べる。リサンプリングの方法は, 重みが大きいものを残し, 小さいものを消滅させるという操作が良く行われる。しかし, 本研究の提案手法でそのような操作を行うと, 元のサンプリング集合に同じ移動軌跡が複数含まれた場合, それらが残りにくくなってしまう。実質的には同じ移動軌跡が複数含まれた場合, それらの重みの合計がその移動軌跡を評価する際に使われるべきである。そこで, 本研究では乱数を用いてリサンプリングを行う。元もサンプリング集合からその重みに従ってランダムにサンプリング集合を生成し直し, それらから次時点以降の移動軌跡をシミュレーションモデルで推定する。こうすることで, 元のサンプリング集合に同じ移動軌跡が含まれていても, それを考慮してリサンプリングを行うことができる。

3.3.4 大規模計算時の重み更新の近似

式 3.24 と式 3.21 の制約から, W_t を算出することができるが, この行列計算には, 計算時間, メモリ領域ともに $O(W_t \text{の要素数}^2)$ となる。これは小規模であれば大丈夫だが, 都市

圏などの大規模範囲を対象とした推定に適用するのは難しい問題となってしまう。

そこで、都市圏などの大規模領域を対象とする際は、何らかの仮定をおいて W_t の近似値を求める必要がある。本研究では、以下の仮定をおいた近似計算を提案し、適用する。

- 分散共分散行列 V , R の共分散は 0 とみなす。つまり, V , R は対角行列となる。
- 各移動軌跡候補は 1 時点に複数箇所で観測されない。

2 つ目の仮定はそうでない場合も多いと考えられるが、大規模領域で密に観測するのは難しいこと、それによって、各移動軌跡候補がほとんどの観測地点で観測されないと考えられることから、この仮定をおいた推定で十分であると判断した。

この仮定から、対角行列 V の i 行 i 列の値を v_i , 対角行列 R の i 行 i 列の値を r_i とすると、

$$VH_t^T = (\delta_{ij}v_i) \quad (3.25)$$

$$H_tVH_t^T = (b_{ij}) \quad (3.26)$$

$$b_{ij} = \begin{cases} \sum_k \delta_{kj}v_k (i = j) \\ 0 (i \neq j) \end{cases} \quad (3.27)$$

$$H_t^TR^{-1}H_t = (c_{ij}) \quad (3.28)$$

$$c_{ij} = \begin{cases} \sum_k \frac{\delta_{ik}}{r_k} \\ 0 (i \neq j) \end{cases} \quad (3.29)$$

となり、 $Z - HW_{t,b}$ の j 番目の要素を dz_j とすると、

$$VH_t^T(H_tVH_t^T + R)^{-1} = \left(\sum_k \frac{\delta_{ik}v_i dz_k}{\sum_l \delta_{lk}v_l + r_k} \right) \quad (3.30)$$

$$(H_t^TR^{-1}H_t + V^{-1})^{-1} = \left(\frac{v_i}{v_i \sum_k \frac{\delta_{ik}}{r_k} + 1} \right) \quad (3.31)$$

となる。また、よって、 $W_{t,b}$ の i 番目の要素を $w_{t,b,i}$ とすると、式 3.24 より、 W_t の i 番目の要素 $w_{t,i}$ は、

$$w_{t,i} = w_{t,b,i} + \sum_k \frac{\delta_{ik}v_i dz_k}{\sum_l \delta_{lk}v_l + r_k} + \frac{v_i}{v_i \sum_k \frac{\delta_{ik}}{r_k} + 1} \lambda_{\lfloor \frac{i}{N_p} \rfloor + 1} \quad (3.32)$$

となる。これと式 3.21 から λ を求め、式 3.32 に代入すると、 $w_{t,i}$ の近似値を求めることができる。この計算は計算時間、メモリ領域ともに $O(W_t \text{の要素数})$ で行うことができる。

3.4 本研究での検証・考察対象

以上のように、パーティクルフィルタ法と最適内挿法を用いた人の流れ推定手法の提案するにあたって、本研究で検証・考察していくことについて述べる。

前述のように、本研究では大崎駅構内と東京都市圏を対象として、提案する人の流れ推定手法の適用実験を行い、検証・考察を行う。大崎駅構内では厳密に重みを計算し、正規分布の仮定とベイズの理論を用いて重みを更新する。一方で、東京都市圏では同様の計算だと計算リソースが $O(\text{人の数} * \text{パーティクル数})$ になってしまうので、そのまま適用することが事実上不可能だと考えられる。そこで、本章で提案する近似計算を行い重みを更新する。また、東京都市圏では1時間ごとに観測値が取得していると想定するため、今回の実験ではパーティクルフィルタ法のパーティクルのリサンプリング操作は行わない。そのため、東京都市圏での推定は、実質的には、サンプリング集合で離散的な移動軌跡を近似的に表現し、3次元データ同化手法を用いて各サンプルの確率を観測値で補正して推定するというを行うことになる。

大崎駅構内では、主に以下の点について実験・考察を行う。

- パーティクルフィルタ法と最適内挿法を用いた提案手法により精度向上に成功しているかどうか。
- パーティクルフィルタ法を用いることで予測値の改善ができているか。
- 近似計算で十分な精度を期待できるか。計算速度はどれくらい改善されるか。

また、観測地点の効果や、観測値の取得間隔の違いなども分析する。

東京都市圏では、近似計算を行い、主に以下の点について実験・考察を行う。

- シミュレーションモデルの精度の影響
- 観測値のコストパフォーマンスの評価

これらを分析することで、より実用性の高い提案が行える。

参考文献

Wunsch, C., 1996, The Ocean Circulation Inverse Problem, Cambridge University Press, Cambridge.

Daley, R., 1991, Atmospheric Data Analysis, Cambridge University Press, Cambridge.

Bennett, A. F., 2002, Inverse Modeling of the Ocean and Atmosphere, Cambridge University Press, Cambridge.

A. C. Lorenc, 1986, Analysis methods for numerical weather prediction, Quarterly Journal of the Royal Meteorological Society, Vol.112, pp.1177-1194.

A. C. Lorenc, S. P. Ballard, R. S. Bell, N. B. Ingleby, P. L. F. Andrews, D. M. Barker, J. R. Bray, A. M. Clayton, T. Dalby, D. Li, T. J. Payne, F. W. Saunders, 2000, The Met. Office global three-dimensional variational data assimilation scheme, Quarterly Journal of the Royal Meteorological Society, Vol.126, pp.2991-3012.

Kalman, R. E., 1960, A new approach to linear filtering and prediction problems, Journal of Basic Engineering, 82, 35-45.

片山 徹, 1983, 応用カルマンフィルタ, 朝倉書店.

Evensen, G., 1994, Sequential data assimilation with a non-linear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, Journal of Geophysical Research, 99, 10143-10162.

Evensen, G., 2003, The ensemble Kalman filter: Theoretical formulation and practical implementation, Ocean Dynamics, 53, 343-367.

Kitagawa, G. 1996, Monte Carlo filter and smoother for non-Gaussian nonlinear state space models, Journal of Computational and Graphical Statistics, 5, 1-25.

Gordon, N. J., Salmond, D. J. and Smith, A. F. M., 1993, Novel approach to nonlinear/non-Gaussian Bayesian state estimation, IEE Proceedings F, 140, 107-113.

第4章 駅構内での人の流れ推定への適用

4.1 駅構内での人の流れ推定のフレームワーク

本節では、駅構内での人の流れ推定への適用として、駅構内での歩行者流の推定を目的に、観測値と歩行者モデルを統合する手法の概要について述べる。前述のように、コストなどの面から、広範囲な領域での人の流れには観測値が無い空間が生じてしまう。そこで、そのギャップを歩行者モデルで補うことを目指す。

全体のフレームワークを図 4.1 に示す。前提条件として、対象領域に入る人の時間と出入口情報は得られるものとする。まず、対象領域に入る人を検知した時点で、エージェントを生成する。次に各エージェントについて、歩行者モデルを用いて確率的に軌跡候補を生成する。最後に生成した軌跡候補を、データ同化手法と観測値によって評価し重み付けすることによって、観測値に即した軌跡候補の選択を行う。つまり、歩行者モデルによって一つに軌跡を決定するのではなく、確率的にいくつかの軌跡候補を生成し、観測値に即すように軌跡の組み合わせを得るという流れで推定を行う。

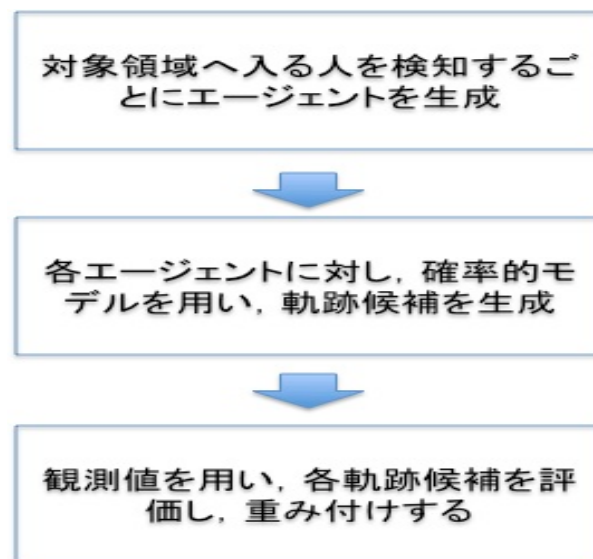


図 4.1: 推定手法のフレームワーク

4.2 歩行者モデル

軌跡候補生成の際に用いる歩行者モデルについて述べる．本研究では，衝突回避モデルについては考慮せず，ある程度大まかな経路選択モデルを構築することを目指す．

経路選択モデルの基本として，ポテンシャルモデルを用いる．前述のように，ポテンシャルモデルには，同じ目的地の歩行者については経路が一意に決まってしまうという問題がある．本研究では確率的に多くの軌跡候補を生成する必要があるため，一意に決まってしまうという欠点を改善したモデルを新たに提案する．

ポテンシャルモデルでは，まず図4.2のように，任意の点から目的地までの移動コストを算出し，移動コストのポテンシャル面を描く．次にこのポテンシャル面に対して勾配の急な方向に向かって歩行者が移動しようとする．本研究では，このように最も急な方向に向かって移動すると，多くの候補を生成できないので，各地点のポテンシャルから確率的に移動方向を決定するモデルを提案する．

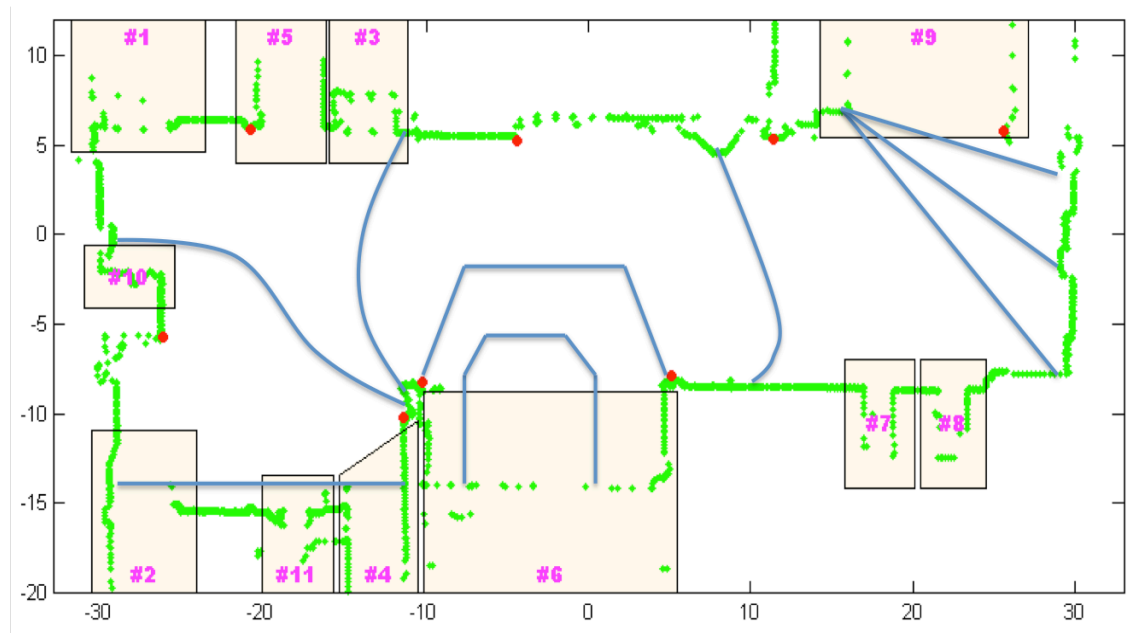


図 4.2: ポテンシャルマップ

まず，対象領域を $1\text{m} \times 1\text{m}$ 四方のグリッドに分割する (図 4.3)．次に出発グリッドから到着グリッドまでグリッドを結んでできる軌跡を生成する (図 4.4)．このグリッドでできた軌跡をグリッドトラジェクトリーと呼ぶことにする．最後に，このグリッドトラジェクトリーを平滑化し，軌跡を生成する (図 4.5)．以下，詳細を述べていく．

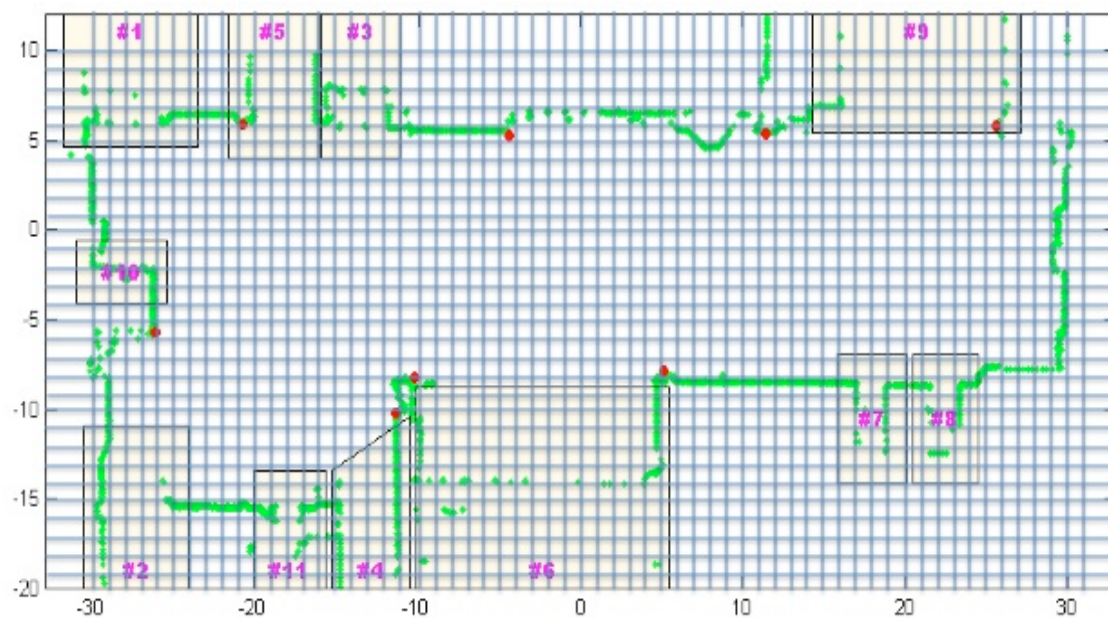


図 4.3: グリッド分割

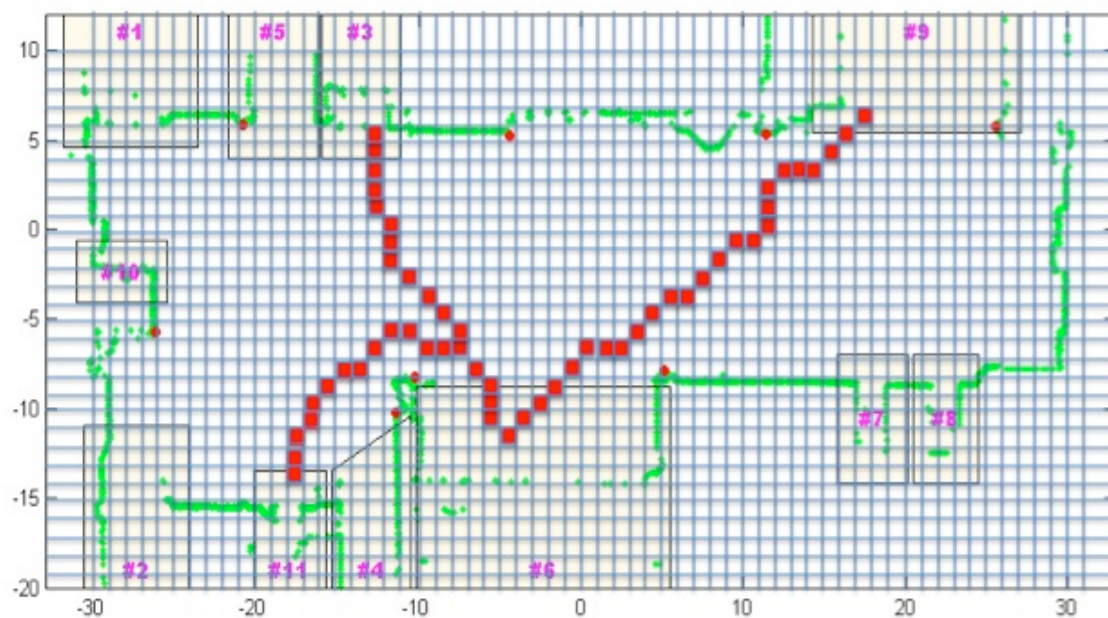


図 4.4: グリッドトラジェクトリーの例

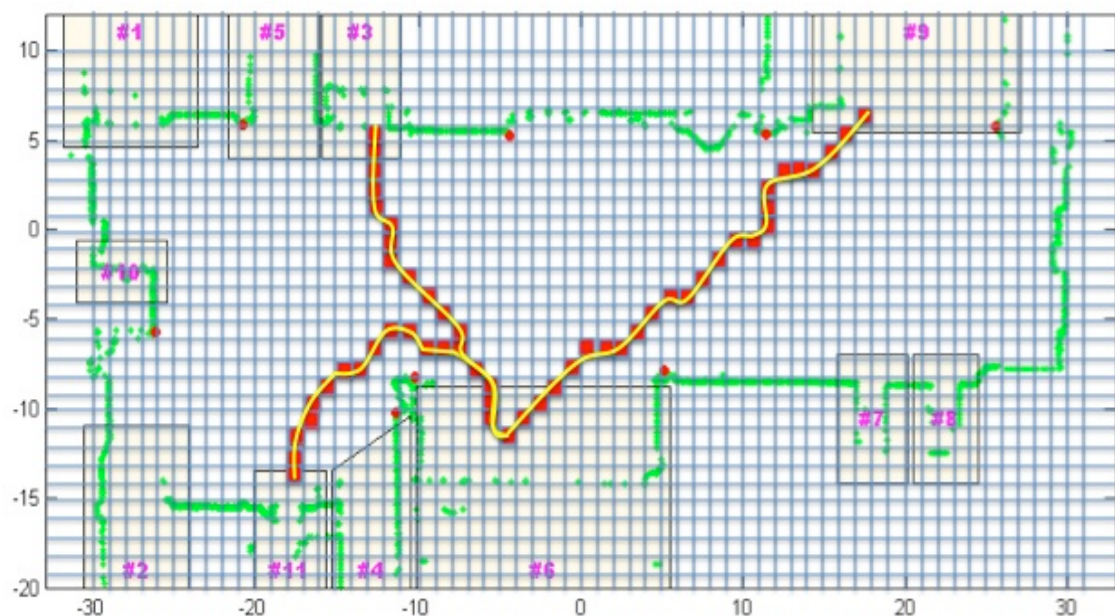


図 4.5: グリッドトラジェクトリーの平滑化

4.2.1 グリッドトラジェクトリーの生成

まず、対象領域をグリッド分割した後、ポテンシャルマップを作成し、グリッドトラジェクトリーの候補を生成する手順について述べる。

ポテンシャルマップの生成は手動で行われることが多かったが、本研究では自動的に生成することを目指す。ポテンシャルの計算手法として、離散最短距離による手法や拡散方程式による手法などが考えられる。本研究では、データ同化手法による観測値と歩行者モデルの融合に力を入れるため、シンプルな離散最短距離による手法を用いる。離散最短距離とは、目的地グリッドから各グリッドまで通過可能なグリッドをつないだときの最短距離である。具体的には、各グリッドの中心点をノードとし、近隣グリッドのうち通過可能なグリッドのノードへリンクをつなぎネットワークを構築し、ダイクストラ法を用いて最短距離を算出した。また、人は壁からある程度以上離れて歩くことが多いと考えられるので、壁からの距離もポテンシャルの説明変数に含める。それらを踏まえ、グリッド i のポテンシャル po_i を、離散最短距離 l_i 、壁からの距離 w_i を用い、以下の式を用いる。

$$po_i = l_i + w_i \quad (4.1)$$

次にポテンシャルを用いてグリッドトラジェクトリーを生成する。まず出発グリッドからその近隣 8 グリッドのうち、通過可能かつポテンシャルが低くなるグリッドを次グリッ

ドの候補とする。そして、以下の式で各候補グリッド i の選択確率 p_i を求める。

$$p_i = \frac{\log(po_{now} - po_i)}{\sum_i \log(po_{now} - po_i)} \quad (4.2)$$

次グリッドの候補から確率的に次グリッドを選択し、同様に次グリッドを選択していく。
以上のように、グリッドトラジェクトリーの候補を確率的に生成する。

4.2.2 グリッドトラジェクトリーの平滑化

次に生成したグリッドトラジェクトリーから滑らかな軌跡を生成する手順について述べる。平滑化にはパーティクルフィルタ法を用いる。

グリッドトラジェクトリーの中心点の点列を目標点とし、1 時点後の目標点（つまり次グリッドの中心点）を目標点とし、以下に示す Step を起点データから順次繰り返し処理を行う。まず、起点観測データ周辺に 10 個のパーティクルを生成・拡散し、初期値とする。目標点は、2 番目のグリッドの中心点とする。

Step1 : パーティクルの生成

直前の段階でフィルタリングされた 10 個のパーティクルに対し、半径 1.0m の円内にランダムにパーティクル 1 つにつき 10 個生成・拡散する (図 4.6)。

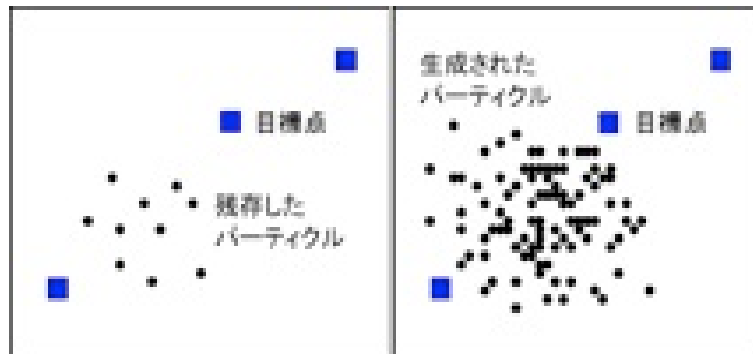


図 4.6: パーティクルの生成

Step2 : パーティクルの評価

生成・拡散された 10×10 個のパーティクルに対して、目標点までの距離の逆数を尤度とし、尤度の高い 10 個の点以外のパーティクルをフィルタリング（消滅）させる (図 4.7)。

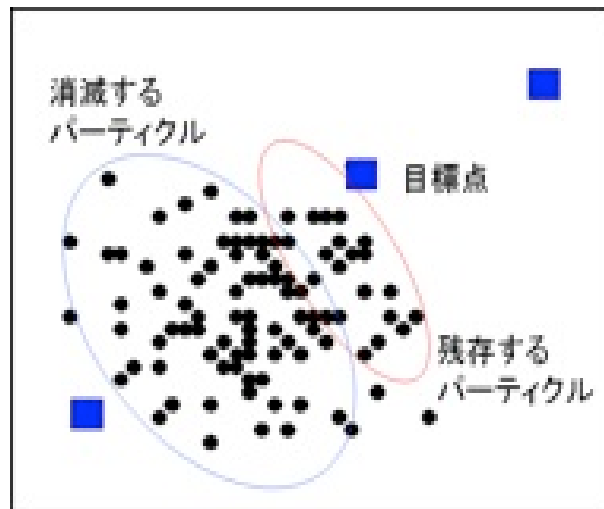


図 4.7: パーティクルの評価

Step3 : 目標点の切り替え

フィルタリングされた 10 個のパーティクルの重心と目標点との距離が、基準距離の 1/2 より小さくなったとき、目標点を次の GPS 観測データに切り替え、基準距離を重心と次の GPS 観測データとの距離とする。

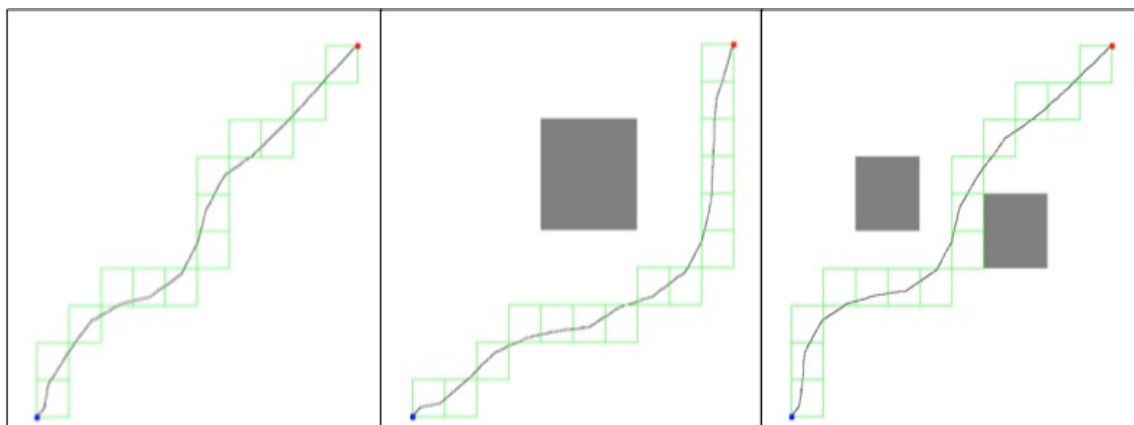


図 4.8: 平滑化の結果例

4.2.3 まとめ

以上のように、軌跡候補を生成する。ポテンシャルマップは目的地ごとに作成する必要があるため、生成する軌跡候補数を各目的地に割り振り、目的地ごとに軌跡候補を生成する。下に歩行者モデルのフレームを示す (図 4.9)。

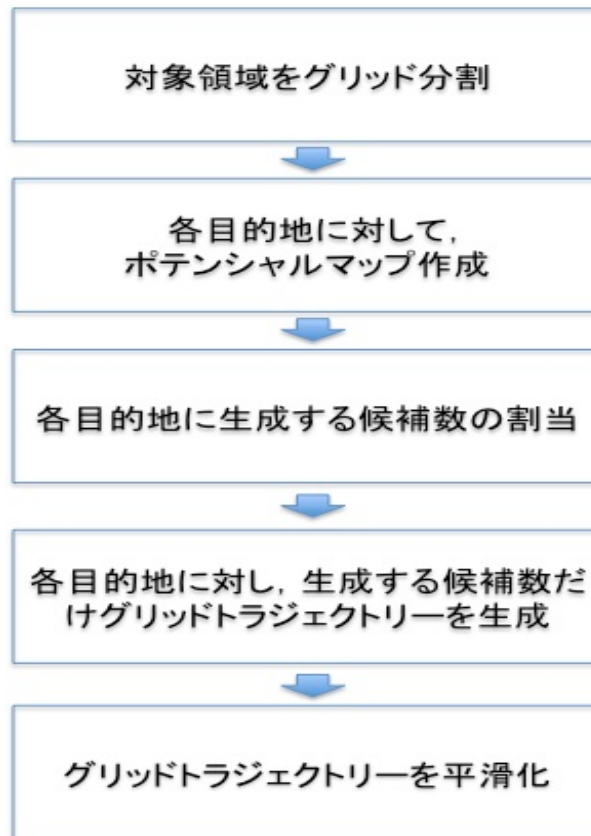


図 4.9: 歩行者モデルのまとめ

4.3 データ同化手法の適用

本節では、データ同化手法による観測値と歩行者モデルの融合について述べる。

4.3.1 手法の概要

観測値と歩行者モデルの統合のために、パーティクルフィルタ法を用いてデータ同化を行う。

軌跡推定モデルによって得られる候補の選択確率を観測値によって更新していく。まず、エリア内にある出入口から入った人に対し、歩行者モデルに従って、パーティクルとして候補となる軌跡を持たせ、各パーティクルの重みとして、 $1/\text{パーティクル数}$ を与える。例えば、計 100 の軌跡候補（パーティクル）を生成する場合、それぞれのパーティクルを重み $1/100$ として生成する。

その後、最適内挿法によって重みを更新し、標準化と繰り返し計算をして、最終的な重みを得る。最後に、最終的に得られる各パーティクルの重みを選択確率とし、得られた確

率から乱数を用いて軌跡選択を行い推定軌跡とする,

4.3.2 具体例

観測値と歩行者モデルの統合手法について, 具体例を示しながら詳述する.

具体的な状況として, 対象領域に4人が入り, それぞれ図4.10の点線のような軌跡で行動する状況を考える. ID1の人は出入口6から出入口9へ, ID2の人は出入口2から出入口9へ, ID3の人は出入口1から出入口6へ, ID4の人は出入口9から出入口6へ移動する. また, 図4.10の赤線上を通過した人数を計測し観測値として取得できると仮定する. このとき, 4人が赤線上を通過したかどうかを見ると, ID2と3は通過しており, ID1と4は通過していない(表4.1). そのため, 赤線上を通過した人数が2人であるという観測値を取得できると仮定する.

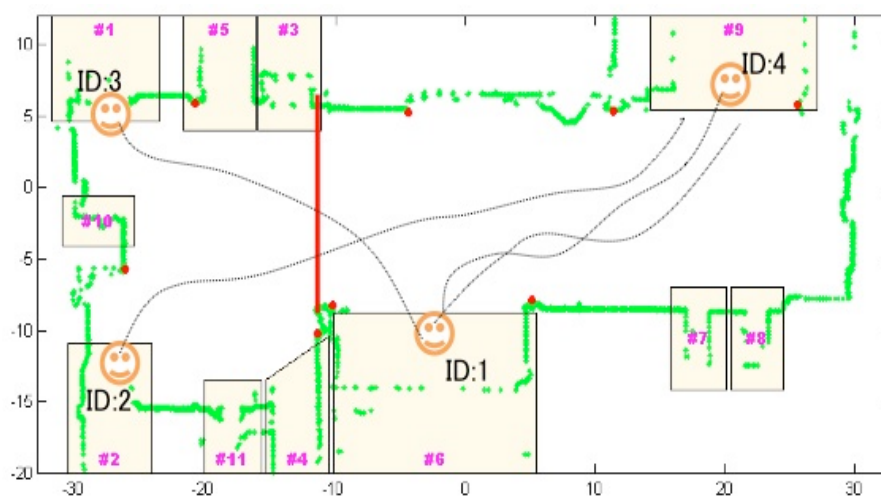


図 4.10: 具体例の状況

表 4.1: 赤線を通過したかどうか

ID	通過
1	×
2	○
3	○
4	×

この具体例では, 各人に対し2つずつのパーティクルを生成するとする. 各人のパーティクル数が2なので, 各パーティクルの重みの初期値は0.5となる. 各パーティクルは図4.11

のように生成する。ID1 には出入口 6 から入り出入口 5 と 9 から出るパーティクル，ID2 には出入口 2 から入り出入口 6 と 9 から出るパーティクル，ID3 には出入口 1 から入り出入口 6 と 9 から出るパーティクル，ID4 には出入口 9 から入り出入口 5 と 6 から出るパーティクル，を生成する。この場合，赤線上を通過したパーティクル数は，ID2 と 3 が 2 つ，ID1 と 4 が 1 つとなる (表 4.2).

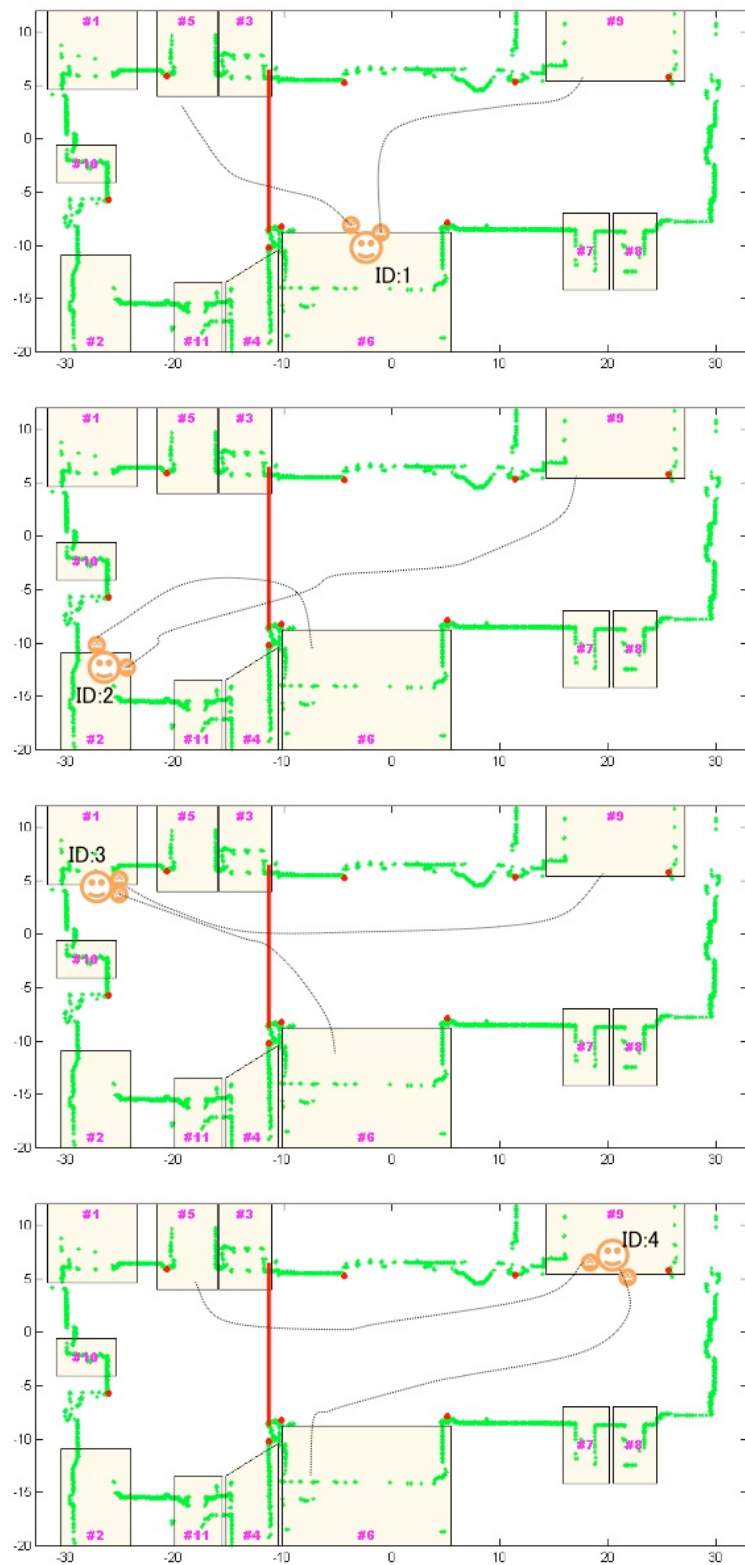


図 4.11: 生成されたパーティクル

表 4.2: パーティクルとその重み

ID	通過	重み
1	○	0.5
1	×	0.5
2	○	0.5
2	○	0.5
3	○	0.5
3	○	0.5
4	○	0.5
4	×	0.5

以下、これらのパーティクルの重みを観測値を用いて更新する。赤線上を通過した実際の人数は2である。一方で、計6つのパーティクルが赤線上を通過しており、その重みの合計は $6 \times 0.5 = 3$ (人) となる。これは実際の観測値である2人より多いため、更新により赤線上を通過するパーティクルの重みを小さくする必要がある。そこで、式 3.24 と式 3.21 により重みを更新する。

まず、式 3.24 において、

$$W_{t,b} = (0.5, 0.5, 0.5, 0.5, 0.5, 0.5)^T \quad (4.3)$$

$$H_t = (1, 0, 1, 1, 1, 1, 1, 0) \quad (4.4)$$

となる。ここで、 W_t と Z_t の分散共分散行列 V と R について、ともに共分散は0で、標準偏差を W_t は $W_{t,b}$ の 10%、 Z_t はその 1% と仮定すると、分散はそれぞれその 2 乗になるので、

$$V = \begin{pmatrix} 0.05^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.05^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.05^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.05^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.05^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.05^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.05^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.05^2 \end{pmatrix} \quad (4.5)$$

$$R = (0.02^2) \quad (4.6)$$

となる。これを式 3.24 に代入し、式 3.21 で連立させ Λ を求め、式 3.24 に代入すると、

$$W_t = (0.069, 0.931, 0.5, 0.5, 0.5, 0.5, 0.069, 0.931)^T \quad (4.7)$$

となる。結果として重みは表 4.3 のように更新される。赤線を通過していない ID:1 と ID:4 について、赤線を通過していないパーティクルの重みが大きくなり、赤線を通過するパーティクルの重みが小さくなっている。

表 4.3: 重みの更新

ID	通過	重みの初期値	重みの更新値
1	○	0.5	0.069
1	×	0.5	0.931
2	○	0.5	0.5
2	○	0.5	0.5
3	○	0.5	0.5
3	○	0.5	0.5
4	○	0.5	0.069
4	×	0.5	0.931

4.4 大崎駅構内での実験

4.4.1 研究対象領域

本研究では、大崎駅構内を対象領域とし、提案する人の流れの推定手法と検証を行う。大崎駅は一日平均乗車人員が約 12 万人であり、JR 東日本で 24 番目の多さである。山手線、りんかい線、湘南新宿線、埼京線が通っており、検証に適度な広さであると考えられる。本研究では、大崎駅のコンコース 60 m × 20 m ほどの広さを対象とする。この中に、改札口や階段、トイレなど 11 の出入口が存在する。

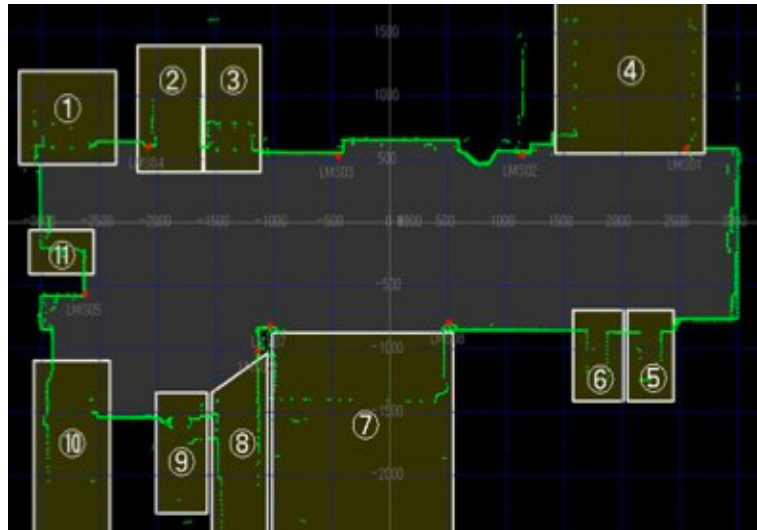


図 4.12: 大崎駅構内と 11 の出入口

4.4.2 使用するデータ

本研究では、大崎駅構内で複数台のレーザスキャナによって取得された人の流れデータを用いる [3]。以下、その取得方法の概要を述べる。

人の流れ計測のために使用されたレーザスキャナは SICK 社の LMS-200 であり (表 4.4), これを床から約 20cm の高さに設置し、水平の計測断面を歩行者の足首を狙って計測された (図 4.13)。また、大崎駅構内に 8 台のレーザスキャナを設置し、それらを統合することで広範囲に渡って高精度な計測がなされた (図 4.14)。



図 4.13: レーザスキャナによる歩行者計測

表 4.4: SICK LMS200 の仕様

計測範囲	最大 180°
角度分解能	0.25-1.0°
距離分解能	10mm
システム誤差	± 1.5mm
測定可能距離	≤ 30m
消費電力	20W
仕様周囲温度	0-50 °C
重量	4.5kg
外形寸法	156 × 155 × 210mm
波長帯	905nm (近赤外)
レーザクラス	1A (アイセーフ)
サンプリングシート	37.5Hz (デフォルト値)

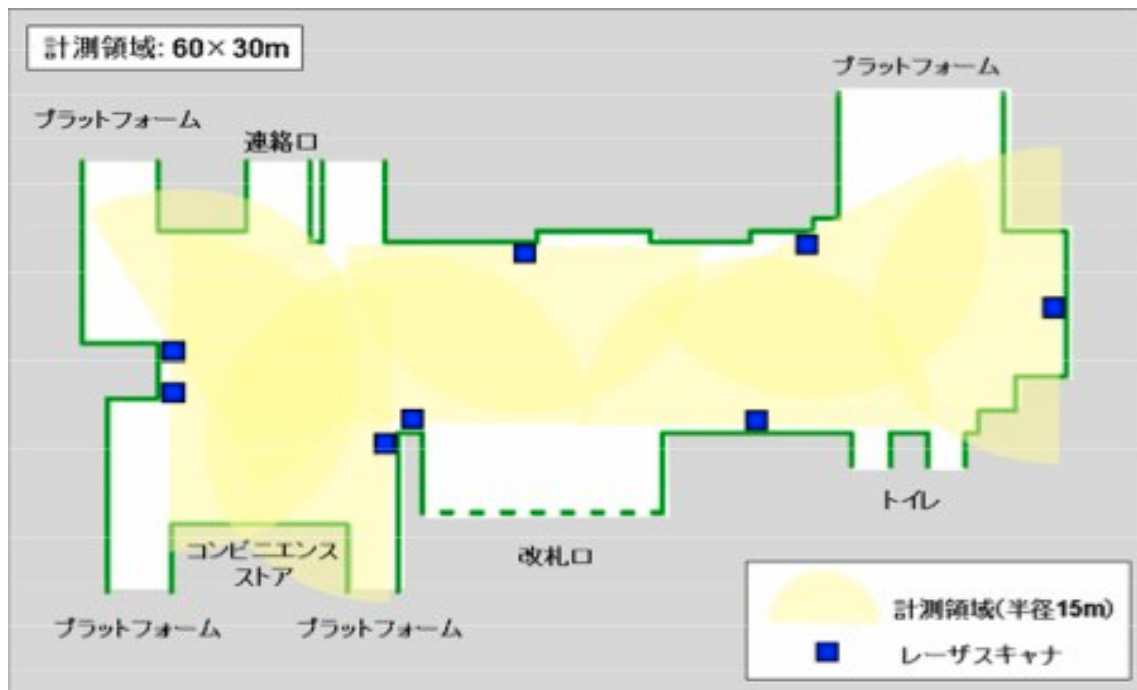


図 4.14: レーザスキャナセンサの配置図

本研究では、2つのデータセットを検証に用いる。一つは 7:00-7:05 の 5 分間のデータ、もう一つは 8:00-8:10 の 10 分間のデータである。前者は 323 人、後者は 2299 人の歩行者が計測され、11 の出入口のうちの一つから対象領域に入り、11 の出入口のうちの一つから出

ている。

出入口の利用頻度を見てみると、出発・到着ともに改札口である出入口6の利用頻度が多くなっている。興味深い点として、出入口9を出発出入口とする人の数が、7:00-7:05では0人であるのに対し、8:00-8:10では640人となり、出入口6よりも多くなっていることが挙げられる。人の流れを推定することの難しさが表れているといえる。

計測結果として、図4.15のように足首の軌跡が取得できた。7:00-7:05の歩行者の軌跡を見てみると、多くの人が通過している空間とほとんど人が通っていない空間があることが分かる(図4.16)。こういった特徴を高精度に推定することを目指す。

表 4.5: 出発出入口の頻度

出発出入口	1	2	3	4	5	6	7	8	9	10	11
7:00-7:05	21	49	79	43	9	64	7	0	0	6	45
8:00-8:10	108	230	383	258	93	361	48	11	640	25	142

表 4.6: 到着出入口の頻度

到着出入口	1	2	3	4	5	6	7	8	9	10	11
7:00-7:05	23	11	10	9	32	128	10	0	68	1	31
8:00-8:10	240	160	125	50	257	997	42	15	254	15	144

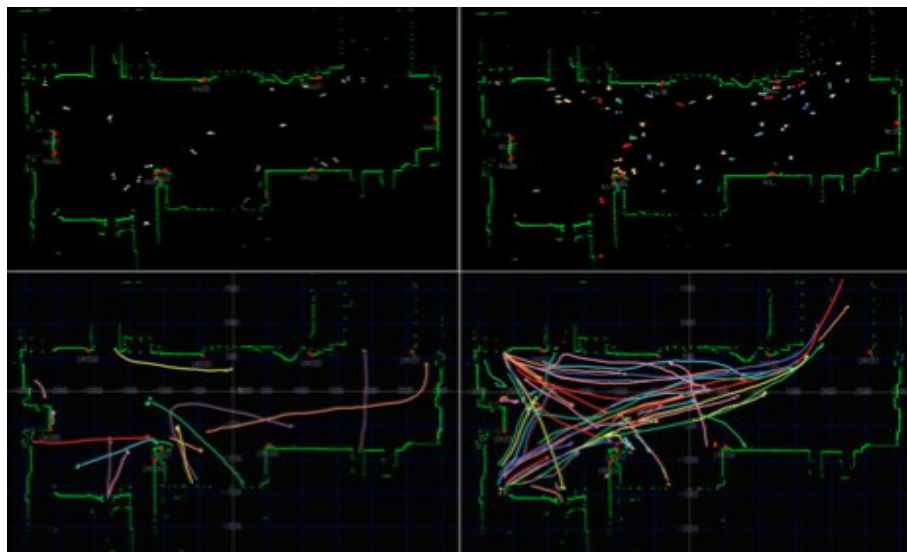


図 4.15: レーザスキャナによる歩行者計測結果

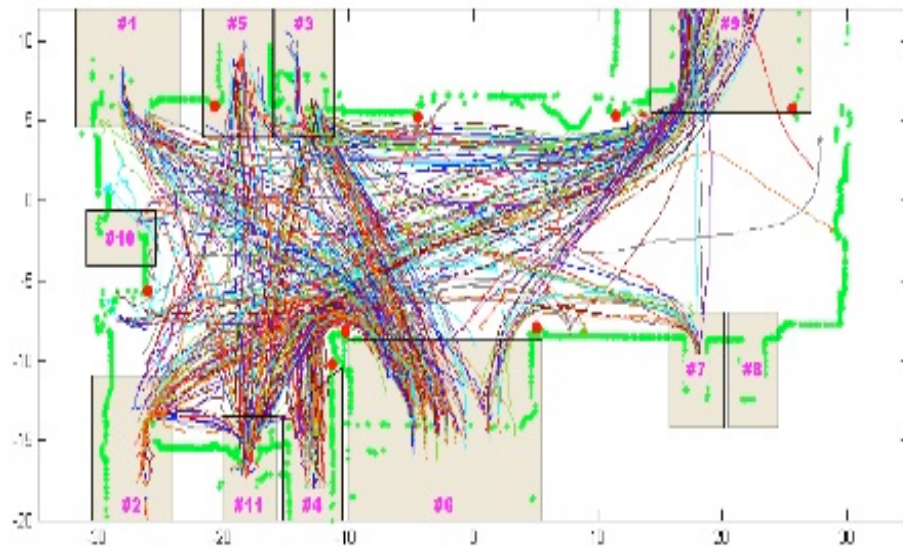


図 4.16: 7:00-7:05 の歩行者の軌跡

4.4.3 実験の概要

大崎駅構内における人の流れデータのうち、7:00-7:05 のときのデータを用いて手法の検証を行う。検証は、データ同化手法の検証と、歩行者モデルとデータ同化手法を合わせた総合的なフレームワークの検証に分けて行う。

データ同化手法の検証では、データ同化手法の効果のみを検証するため、歩行者モデルとして提案するモデルではなく、実際のデータを用いたシンプルなモデルを利用する。具体的には、8:00-8:10 のデータを検証用歩行者モデルとして利用し、データをそのままランダムに与えることで代用する、例えば、軌跡推定モデル用データには、6 番出入口から対象領域に入った人の軌跡が 361 あるので、6 番出入口からエリアに入る人には、各軌跡が $1/361$ の確率で与えられる (図 4.17)。

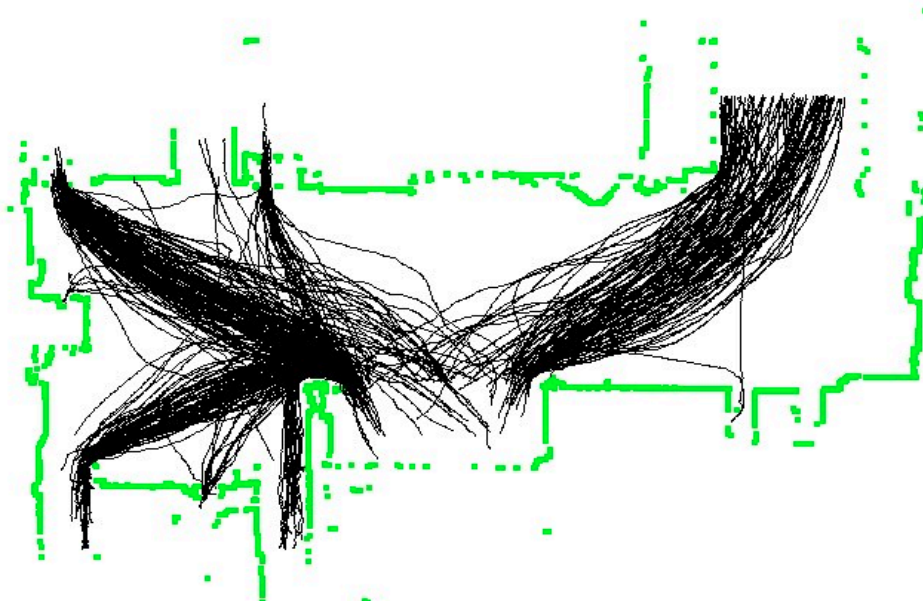


図 4.17: 歩行者モデル用データ (8:00-8:10) の 6 番出入口からの軌跡

また，推定の際の前提条件として，前述したように，対象領域に入る人の時間と入口情報は得られるものとする．その上で，いくつかの出口での通過人数やある断面での通過人数などの観測値が得られたと仮定し，データ同化手法を用いて推定を行う．

4.4.4 データ同化手法の基本的な検証

まず，観測値として，各出入口から対象領域を出る人の数を 30 秒間隔で得られたと仮定して検証を行った．つまり，データ同化手法ありの推定の場合，入力データとして前提条件の対象領域に入る人の時間と出入口情報に加えて，観測値として各出口での利用人数が 30 秒間隔で得られるとし，データ同化手法なしの場合，前提条件のみを入力データとして推定を行う．そして，データ同化を行った場合と行わなかった場合について比較を行う．

表 4.7: 入力データ

	データ同化なし	データ同化あり
対象領域に入る人の時間と入口情報	○	○
各出口の利用人数	×	○

まず，各出口の利用人数がデータ同化手法を用いることで真値に近づくかどうか調べる，図 4.18 は各出口の通過人数を示したものである．左が真値，真ん中がデータ同化ありでの推定，右がデータ同化なしでの推定結果である．図より，全ての出口の通過人数について，

データ同化を行わないときよりも行った時の方が真値に近いことが分かる。図 4.19 は 30 秒ごとの出口の通過人数の RMSE を計算したものである。実線がデータ同化ありによる推定、点線がデータ同化なしでの推定結果である。図から全ての時間帯で大きく誤差を減らすことができていることが分かる。ただし、データ同化手法により各出口の通過人数を同化しているため、出口の利用人数に対する RMSE が改善されるのは自明であるとも言え、この結果から、その期待された効果が得られることが確認されたと言える。

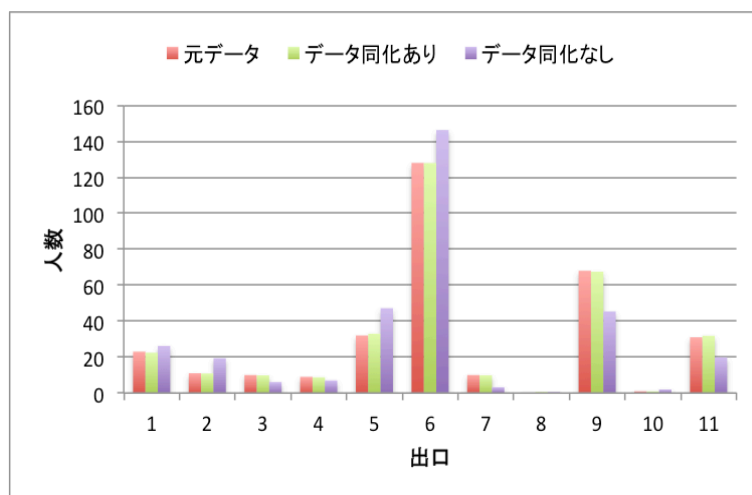


図 4.18: 各出口の通過人数

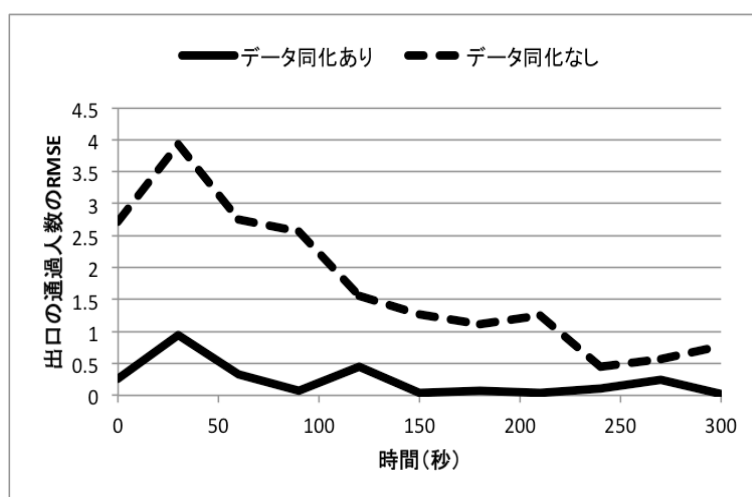


図 4.19: 30 秒ごとの出口の通過人数の RMSE

次に、軌跡について検証する。対象エリアを $1\text{m} \times 1\text{m}$ のグリッドに分割し、各グリッドを通過した人数を計算した。図 4.20 が実際のデータで、図 4.21 が推定結果である。図 4.21 の左図と右図を比較すると、右上の部分などがデータ同化を行うことでより実際のデータに近い値となっていることが分かる。図 4.22 は 0.05 秒刻みのグリッド人数の RMSE を比

較したものである。青がデータ同化ありでの結果、赤がデータ同化なしでの結果で、若干ではあるが、データ同化ありの結果の方が誤差が小さくなっている。ただし、シミュレーションモデルがある程度実際の人の流れを表現できているため、改善度合いは少なくなっている。



図 4.20: 実際のデータのグリッド通過量



図 4.21: 推定値のグリッド通過量 (左図:データ同化なし, 右図:データ同化あり)

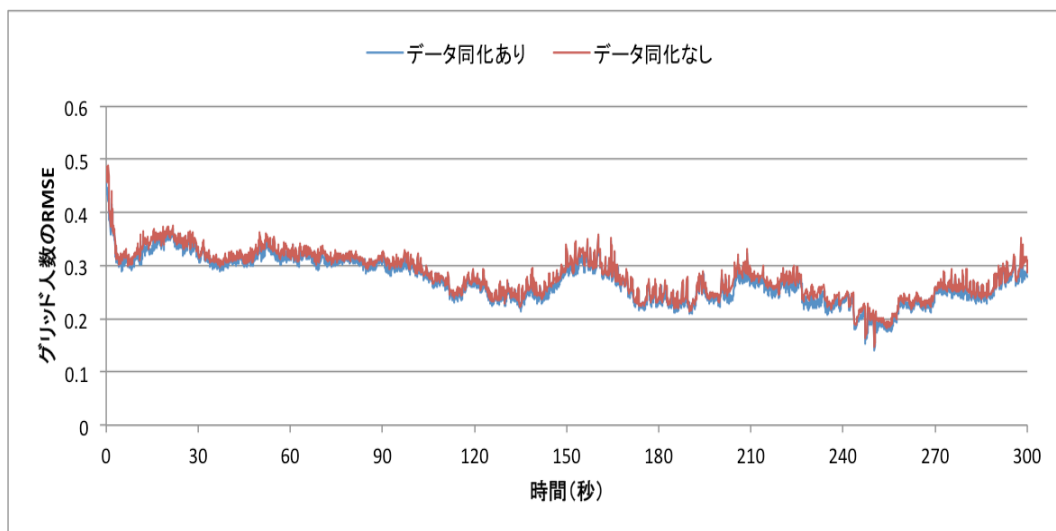


図 4.22: 0.05 秒刻みのグリッド人数の RMSE

4.4.5 歩行者モデルとデータ同化手法の融合

次に、ポテンシャルマップを拡張して構築した歩行者モデルとデータ同化手法を融合した場合について考察を行う。観測値として、全出口で利用人数を30秒ごとに取得できると仮定する。図4.23は、30秒ごとの出口の通過人数のRMSEを計算したものであり、実線がデータ同化あり、点線がデータ同化なしでの推定結果を示す。図から、出口での通過人数はデータ同化により大きく改善しており、期待通りの結果となっている。

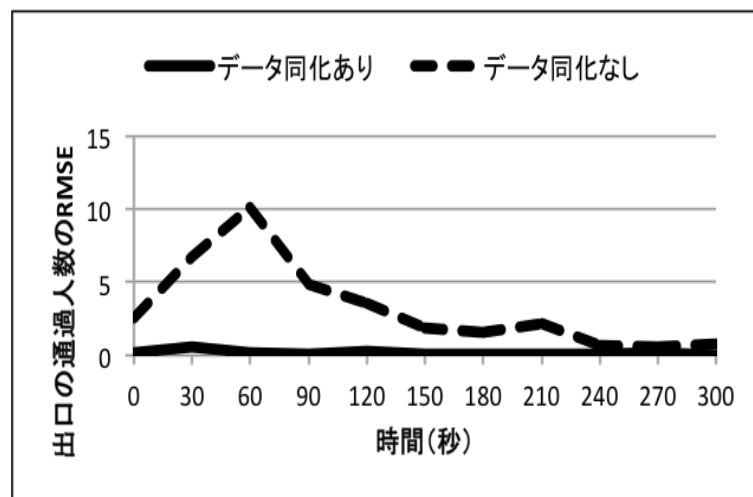


図 4.23: 30 秒ごとの出口の通過人数の RMSE

また、同様に、各グリッドの通過人数を計算し、軌跡の検証を行う。図4.26がその結果である。図4.25と図4.26の左図を比較すると、歩行者モデルを用いた方が平均的になっている。実際のデータを軌跡生成用に用いることに比べ他の予備知識なしに軌跡を生成しているためである。これに対し、図4.26の右図を見ると、データ同化を行うことで真値へかなり近づいていると言える。歩行者モデルだけでは混雑などの特徴がはっきり表れていないが、データ同化を行うことで、ある程度混雑状況を表すことができていると言える。図4.27は5秒ごとのグリッド人数のRMSEであり、青がデータ同化あり、赤がデータ同化なしでの推定結果である。図からデータ同化を行うことで精度を改善できていることが分かる。以上から、歩行者モデルを用いても、データ同化を行うことで、実際のデータを軌跡生成用に用いた場合かなり近づけることができていると言える。



図 4.24: 実際のデータのグリッド通過量 (再掲)

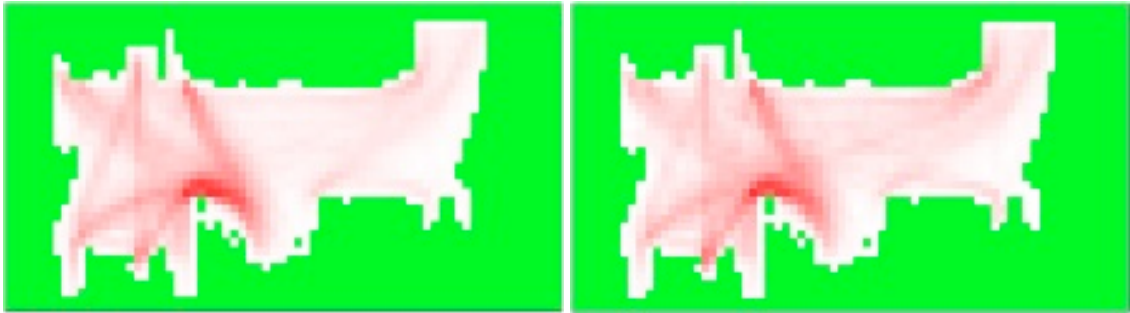


図 4.25: 推定値のグリッド通過量 (左図:データ同化なし, 右図:データ同化あり)(再掲)



図 4.26: 歩行者モデルを用いた推定値のグリッド通過量 (左図:データ同化なし, 右図:データ同化あり)

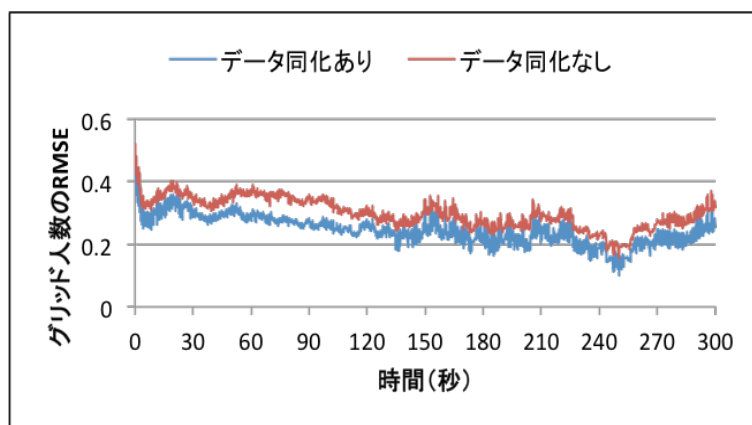


図 4.27: 0.05 秒刻みのグリッド人数の RMSE

最後に、推定結果の OD について精度を検証する。図 4.28 はシミュレーションモデルのみでの推定値と実際の OD 表との差、図 4.29 はデータ同化ありでの実際の OD 表との差である。シミュレーションモデルは実際の OD 表を反映していないため、図 4.28 のようにある程度実際と差ができています。それに対し、提案手法で出口での通過人数で推定値を補正すると、図 4.29 のように実際との差を小さくすることに成功している。RMSE を計算すると、シミュレーションモデルのみの推定値は 5.47、データ同化ありでの推定値は 2.20 となり、50%以上の精度向上に成功している。既存のシミュレーションモデルの中には OD 表を必要とするものが少なくない中、観測値での補正により OD 表の精度も高められるということは、実用化を考えた際に大きな利点となる。

	1	2	3	4	5	6	7	8	9	10	11
1	0.00	3.27	0.00	0.00	3.93	-5.70	-2.00	3.33	-0.40	-1.00	-1.43
2	0.00	6.57	0.00	6.00	-2.47	-13.63	5.93	5.67	-6.47	0.00	-1.60
3	10.80	0.00	10.40	10.47	6.13	-35.40	6.00	0.00	-4.93	0.00	-3.47
4	4.87	5.63	0.00	0.00	-0.07	-20.73	4.63	5.70	-1.80	0.00	1.77
5	0.73	1.03	0.80	1.30	0.67	-2.33	-0.03	0.00	-3.10	1.00	-0.07
6	-15.70	-3.47	-2.10	-2.83	4.50	6.77	5.80	0.00	-5.03	6.60	5.47
7	0.67	0.87	0.70	0.50	0.83	-2.97	0.00	0.00	-2.07	0.80	0.67
8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10	1.37	0.00	0.00	0.00	0.90	-0.97	-1.00	0.00	-1.57	0.00	1.27
11	5.53	5.53	-2.00	6.50	-5.77	-14.70	-1.00	0.00	-0.43	6.33	0.00

図 4.28: シミュレーションモデルのみでの推定値の OD 表の実際との差

	1	2	3	4	5	6	7	8	9	10	11
1	0.00	0.55	0.00	0.00	1.88	-0.39	-2.00	-0.14	3.33	-1.00	-2.23
2	0.00	0.00	0.00	1.28	-4.58	2.75	2.10	0.19	-0.35	0.00	-1.38
3	4.65	0.00	0.00	0.45	3.26	-0.57	-1.67	0.00	-2.76	0.00	-3.36
4	2.16	2.13	0.00	0.00	-1.38	-5.01	0.68	0.07	0.49	0.00	0.87
5	1.23	0.78	0.62	1.00	0.00	-0.65	-0.56	0.00	-2.86	-0.06	0.49
6	-13.21	-4.01	0.65	-4.44	6.77	0.00	3.58	0.00	3.91	1.97	4.80
7	0.90	0.07	0.60	0.37	0.33	-1.64	0.00	0.00	-1.75	-0.10	1.21
8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10	-0.08	0.00	0.00	0.00	0.17	2.90	-1.00	0.00	-2.02	0.00	0.03
11	4.11	0.29	-2.00	0.69	-6.75	3.21	-1.00	0.00	2.30	-0.84	0.00

図 4.29: データ同化ありでの OD 表の実際との差

次に、時間間隔と精度の関係について検討する。図 4.30 は、観測の時間間隔と OD 表の RMSE の関係を示したものである。時間間隔が短いほど精度が高くなっていることが分かる。20 秒あたりからはあまり RMSE に差はない。これは、対象領域での移動が 30 秒ほどのものが多いことから来ていると考えられる。

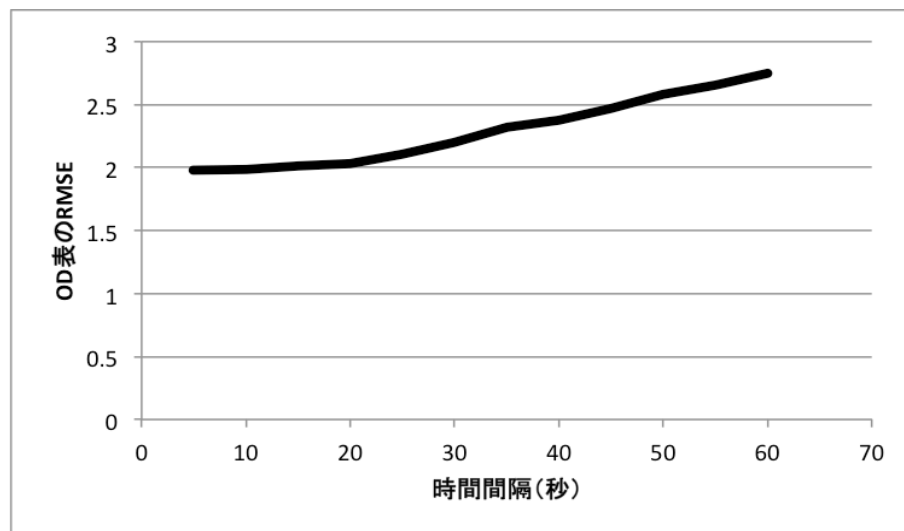


図 4.30: 時間間隔と OD 表の RMSE の関係

次に、パーティクルフィルタ法により予測精度が向上したかどうかを検討する。本研究では、図 4.31 のように、重みの更新後リサンプリングを行い、予測を変化させる。これにより、より尤もらしい移動軌跡に対して次時点までの移動軌跡を予測できるため、予測精

度の向上が期待できる。

そこで、実際に予測精度を向上できているかについて検討するため、図 4.32 のような $3\text{m} \times 3\text{m}$ グリッド 30 箇所て 10 秒ごとに通過人数を観測できると仮定して、リサンプリングを行う場合と行わない場合について予測値の精度を比較する。ここで、図 4.32 の 30 箇所は通過人数が多い上位 30 箇所である。比較的密に、かつ時間間隔が短い観測が行える際に、予測精度の違いが出やすいと考え、このような設定で実験を行う。

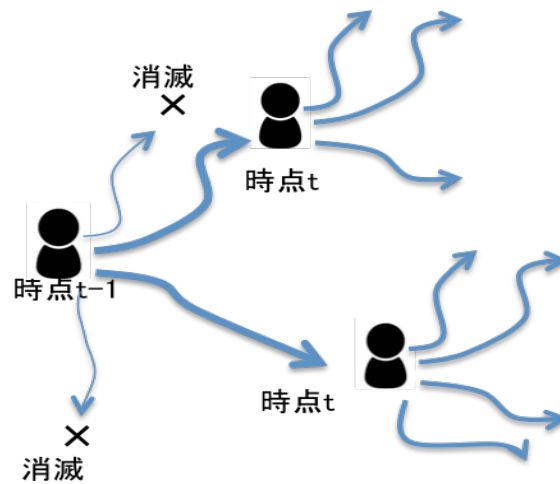


図 4.31: パーティクルフィルタ法による予測変化のイメージ

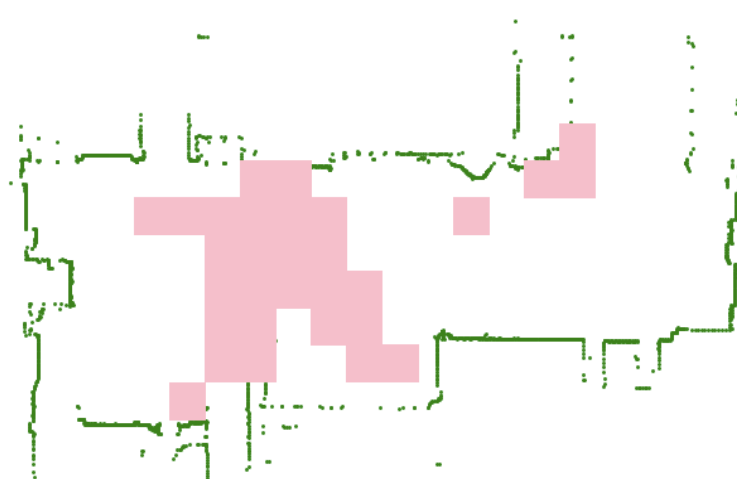


図 4.32: 観測地点となる 30 箇所の $3\text{m} \times 3\text{m}$ グリッド

図 4.33 はリサンプリングする場合としない場合で、予測値に対してグリッド人数の RMSE を計算したものである。青がリサンプリングした場合、赤がリサンプリングしなかった場合であり、若干ではあるが、ほぼ全ての時間帯でリサンプリングした場合の方が精度が良

くなっていることが分かる。ただ、劇的に改善するという結果ではないので、パーティクルフィルタ法によるフィルタリングよりも、最適内挿法による観測値の同化の方が重要であると言える。

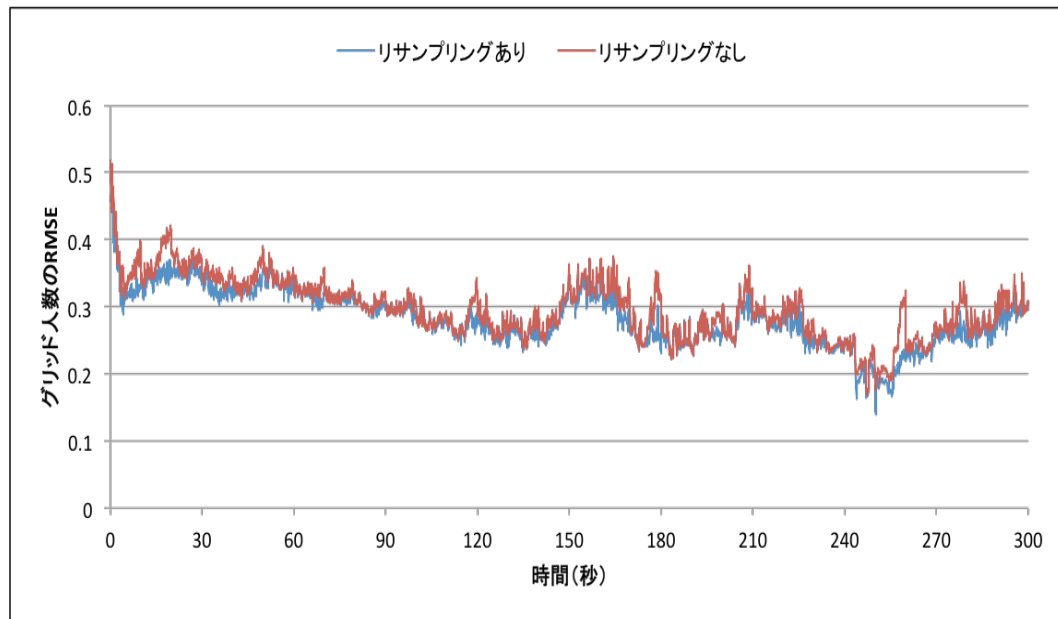


図 4.33: パーティクルフィルタ法による予測値の変化

最後に、重みの更新を近似計算した場合、厳密に計算した場合とどれだけ差が出るのかを検討する。本研究の提案する近似計算手法では、重みと観測値それぞれについて共分散はない、各移動軌跡が1時点間に複数箇所で観測されない、という2つの仮定を置いて近似した。しかし、特に後者についてはそうでない場合も多いと考えられる。そこで、後者が起こる可能性が高い設定で実験し、厳密に計算した場合と近似計算した場合でその差がどの程度であるかを検討する。実験の設定はパーティクルフィルタ法の効果の検討と同じ設定を用いる。この設定では、10秒間隔で3m × 3m グリッド30箇所で通過人数を観測できるというものである。人の歩く速度は秒速1.3mほどと考えられるため、10秒で10数m移動する。10数m移動すれば、複数の3m × 3m グリッドで観測される可能性も高いと考えられるため、求める設定となっていると言える。

図 4.34 は、厳密な計算と近似計算の結果をプロットしたものである。図のように、ほぼ同じ値が計算できており、相関係数は0.99、RMSEは0.003であった。また、図 4.35 は厳密な計算と近似計算の重みの誤差割合の確率分布である。95%有意で誤差割合は2%未満となっている。この誤差範囲であれば、十分に使うことが可能であると言える。

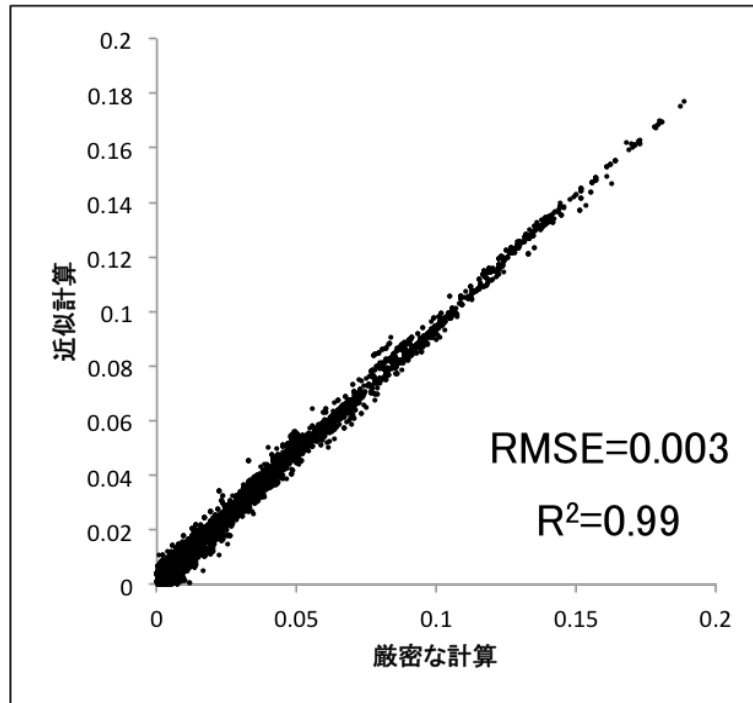


図 4.34: 厳密な計算と近似計算の比較

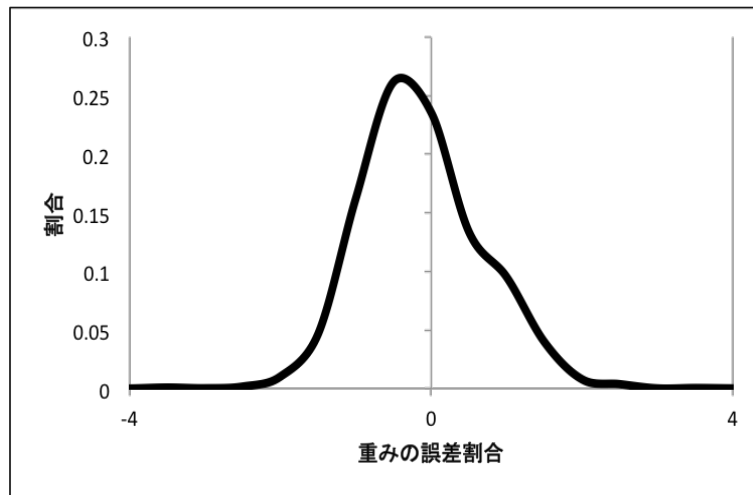


図 4.35: 厳密な計算と近似計算の重みの誤差割合の確率分布

また、厳密な計算と近似計算の計算時間とパーティクル数の関係を示したものが図 4.36 と図 4.37 である。図から近似計算にすることで計算時間が大きく改善されていることが分かる。例えば、パーティクル数が 100 個の場合、計算時間が約 1/1000 になっている。増加の傾向も、厳密な計算ではパーティクル数の 2 乗に比例する形となっているのに対し、近似計算ではパーティクル数にほぼ比例する計算時間となっている。

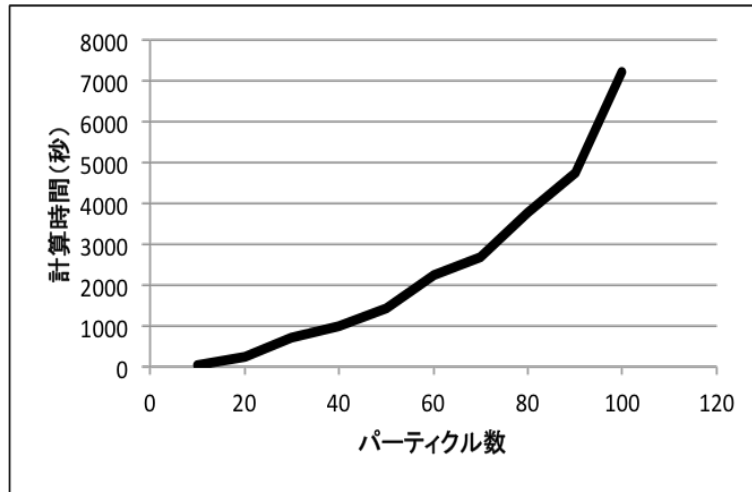


図 4.36: 厳密な計算での計算時間とパーティクル数の関係

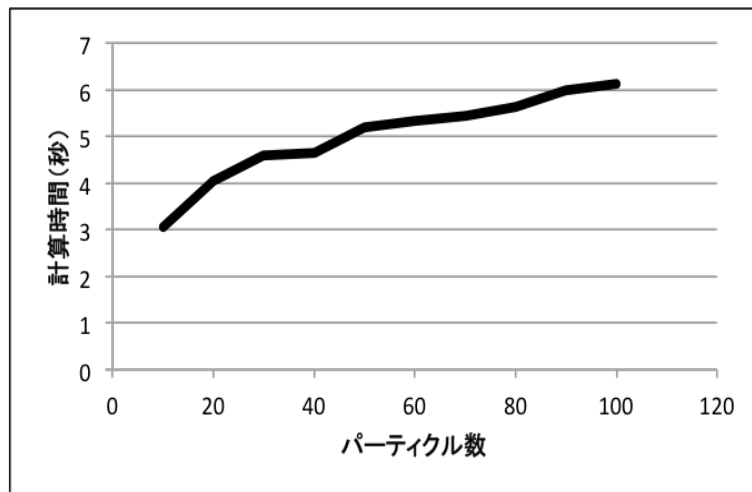


図 4.37: 近似計算での計算時間とパーティクル数の関係

第5章 東京都市圏での人の流れ推定への適用

本章では、都市圏での人の流れ推定として、東京都市圏での人の流れの推定を目的に、提案する人の流れ推定手法の適用方法について述べる。前述したように、東京都市圏では、ある程度広範囲かつ道路ネットワークや鉄道ネットワークに依存した移動を行うと考えられる。大崎駅構内での人の流れ推定への適用と比較して、対象範囲及び人数が多くなることで大きな違いとして挙げられる。そのため、重みの更新手法を近似的な手法にし、計算量を減らした推定を行うことを目指し、推定の計算時間を考慮に入れる必要があることや、観測値を得るためのコストも増大することから、効果の高い観測値を選択する必要があることに着目し、検証・考察していく。

5.1 東京都市圏での推定の全体フロー

本節では、東京都市圏での人の流れ推定手法の全体フローについて述べる。本研究で提案する観測値による推定値の補正手法は、多様な確率的シミュレーションモデルに適用できる事が強みの一つであるが、ネットワークでの推定を行う上で計算量や観測値の効果などを検証・考察するために、トリップベースでの人の流れ推定手法を提案する。図 5.1 にその全体フローを示す。大きく分けると、初期準備、解候補（パーティクル）の生成、パーティクルの重み付け、の3つの段階に分けられる。

まず、初期準備として、データ同化に用いる観測値とシミュレーションモデルの構築に用いるデータを用意する。観測値として、1時間ごとの駅乗降客数や道路交通センサスの調査地点での1時間ごとの交通量など、シミュレーションモデル構築用のデータとして、PT調査データなどを準備する。

パーティクルの生成では、対象領域をPT調査で用いられているゾーンのような小領域に分割した上で、それぞれのゾーンから出発する移動を生成していく。生成手法自体についても、今まで良く用いられてきたPT調査データをもとに構築する。まずPT調査データのトリップを出発ゾーンごとに分け、トリップ集合を作る。そのイメージを図 5.2 に示す。そして、各出発ゾーンから発生するトリップ数を算出し、カーネル推定で得た出発時間の生起確率を用いてその出発時間を与える。そして最後に、発生するトリップに対し、シミュ

レーションモデルを用いてその移動経路を確率的に生成する。パーティクルの生成については、次節でより詳しく述べる。

3章で述べたように、このパーティクルの生成手法を他の手法で置き換えることで、提案する補正手法は様々なシミュレーションモデルに適用できる。例えば、非集計行動モデルを構築し、エージェントシミュレーションでの推定値に対して適用すると、一人一人により焦点を当てた推定ができる。マーケティングでの利用のために推定する際などでより有用な推定となるだろう。本研究では、そうした具体的な事例への適用よりも、より汎用的な検証・考察を行うことに重点を置くため、本節にあるトリップベースの推定のみを行う。

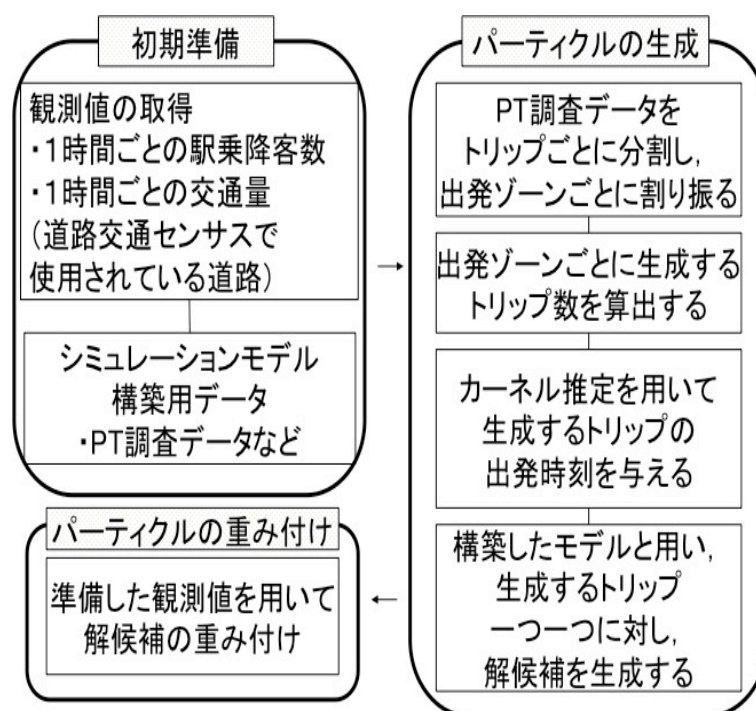


図 5.1: 東京都市圏での推定の全体フロー

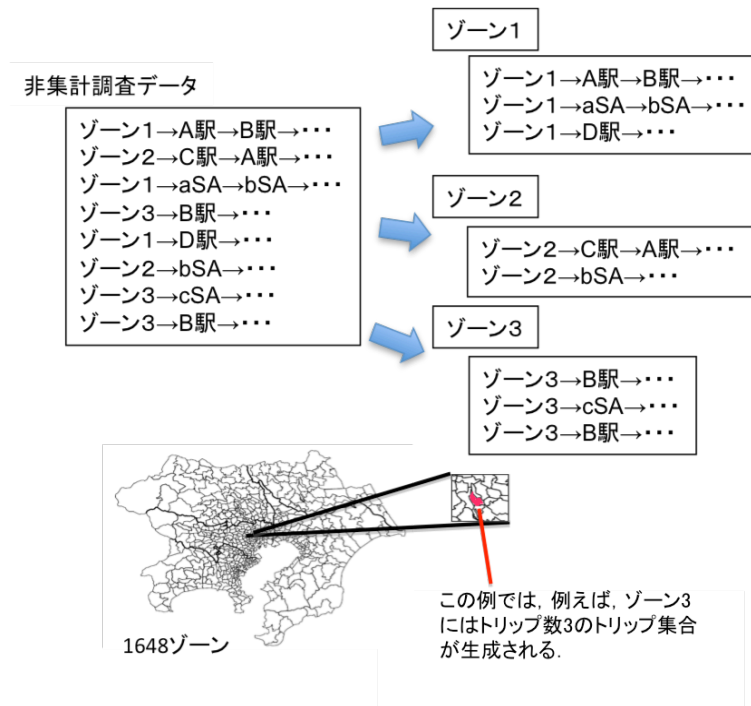


図 5.2: トリップ集合の生成

5.2 パーティクルの生成

本節では、データ同化を用いた推定手法のうち、パーティクルを生成する方法について述べる。人の流れの推定として今まで良く用いられてきた方法として、前述したようなPT調査データの各サンプルに拡大係数を与え、全数を推定するという手法が良く用いられてきている。例えば、東京都市圏では、小ゾーンの性別・年齢・自動車保有の有無別の人口に対するサンプルの抽出率から拡大係数を計算している。しかし、本研究で提案する、データ同化手法を用いて、確率的に得られるパーティクルを観測値で評価する手法を構築するためには、多くのパーティクルを確率的に生成するモデルを構築する必要がある。

そこで、本研究では、PT 調査データからトリップベースで確率的にパーティクルを生成するモデルを構築する。まず PT 調査データをトリップごとに分割し、それを各出発ゾーンごとにパーティクルとし、出発ゾーンごとにトリップを発生させる。出発ゾーンごとのトリップ発生数は、各ゾーンの人の流れデータに占めるトリップ数を抽出率で割ったものとし、出発時間をカーネル密度推定を用いて分散させる。最後に、出発ゾーンごとに発生させるトリップとして、何らかの形で構築したシミュレーションモデルを用いて確率的にトリップ候補を生成することで、パーティクルを生成する。以下に、その手順を詳しく述べる。また、用いるシミュレーションモデルについては次節で述べる。

5.2.1 トリップ集合の生成

まず、PT 調査データをトリップごとに分割し、出発ゾーンごとに集約し、トリップ集合を生成する。図 5.2 のように、各トリップをその出発ゾーンごとに分割し、各ゾーンにひもづけることでトリップ集合が生成される。例えば、図 5.2 でゾーン 3 にはトリップ数 3 のトリップ集合が生成され、また例えば、東京都市圏での PT 調査データ中の全約 200 万トリップのうち 300 トリップがある出発ゾーンから発生している場合、そのゾーンにはトリップ数 300 の集合が生成される。

5.2.2 出発時間の分散

次に、発生させるトリップの出発時間を分散させる。ここで、PT 調査データはアンケート調査であるため、〇時 00 分や〇時 30 分などのラウンドナンバーに時刻が偏っている。図 5.3 は H10 東京都市圏での PT 調査データの出発時刻とその頻度を表した図である。1 分刻みでその時刻を出発時刻とするトリップ数を縦軸に表していて、実際にラウンドナンバーに偏っているため極端に多い時刻とそうでない時刻で滑らかでないことがわかる。そのため、PT 調査データから得られる出発時刻の偏りを平滑化して補正し、出発時刻の偏りをなくす必要がある。

そこで、本研究では、PT 調査データから得られたトリップ集合の出発時間に対して、カーネル密度推定を行い、発生させるトリップ数を分配する。カーネル密度推定は、以下の式で表される。

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right) \quad (5.1)$$

ここで、 n は標本数、 x_i は i 番目の標本値、 K はカーネル関数、 h はバンド幅である。本研究では、このカーネル密度関数を x を出発時間として、 n に PT 調査データのトリップ数、 x_i に各トリップの出発時間を代入することで、出発時間密度関数の算出に用いる。カーネル関数として、Gaussian 関数、rectangular 関数、epanechnikov 関数などが提案されており、バンド幅算出方法として、Silverman (1986) にあるように、標準偏差を用いた方法、さらに四分位範囲を用いた方法、クロスバリデーションを用いた方法などが提案されている。カーネル密度推定の結果は、カーネル関数よりもバンド幅 h に大きく依存するため、カーネル関数としては良く用いられている Gaussian 関数を用いる。Gaussian 関数は以下のように定式化される。

$$K\left(\frac{x - x_i}{h}\right) = \frac{1}{\sqrt{2\pi}} \exp^{-\frac{(x-x_i)^2}{2h^2}} \quad (5.2)$$

バンド幅の算出方法は, Silverman (1986) にあるように, 推定された密度関数と元の密度関数の二乗誤差を最小とするバンド幅を算出して用いられることが多い. 二乗誤差は以下で表される.

$$MISE(\hat{f}) = \int E(\hat{f} - f(x))^2 dx \quad (5.3)$$

$$= \int MSE_x(\hat{f}) dx \quad (5.4)$$

$$= \int (E\hat{f}(x) - f(x))^2 dx + \int var\hat{f}(x) dx \quad (5.5)$$

ここで, $\hat{f}(x)$ の期待値 $E\hat{f}(x)$ は以下のように表される.

$$E\hat{f}(x) = \int \frac{1}{h} K\left(\frac{x-y}{h}\right) f(y) dy \quad (5.6)$$

よって,

$$E\hat{f}(x) - f(x) = \int h^{-1} K\left(\frac{x-y}{h}\right) f(y) dy - f(x) \quad (5.7)$$

となる. $y = x - ht$ とおき, 式 5.7 に代入すると,

$$E\hat{f}(x) - f(x) = \int K(t)(f(x - ht) - f(x)) dt \quad (5.8)$$

となる. テイラー展開 $f(x - ht) = f(x) - ht f'(x) + \frac{1}{2} h^2 t^2 f''(x) + \dots$ を用いると,

$$E\hat{f}(x) - f(x) = -h f'(x) \int t K(t) dt + \frac{1}{2} h^2 f''(x) \int t^2 K(t) dt + \dots \quad (5.9)$$

$$\approx \frac{1}{2} h^2 f''(x) k_2 \quad (5.10)$$

となるため (ただし, $k_2 = \int t^2 K(t) dt$),

$$\int (E\hat{f}(x) - f(x))^2 dx \approx \frac{1}{4} h^4 k_2^2 \int f''(x)^2 dx. \quad (5.11)$$

である. また,

$$var\hat{f}(x) = \frac{1}{n} \int \frac{1}{h^2} K\left(\frac{x-y}{h}\right)^2 f(y) dy - \left(\frac{1}{nh} \int K\left(\frac{x-y}{h}\right) f(y) dy \right)^2 \quad (5.12)$$

$$= \frac{1}{n} \int \frac{1}{h^2} K\left(\frac{x-y}{h}\right)^2 f(y) dy - \frac{1}{n} (f(x) + E\hat{f}(x) - f(x)) \quad (5.13)$$

$$\approx \frac{1}{nh} \int f(x - ht) K(t)^2 dt - \frac{1}{n} (f(x) + O(h^2))^2 \quad (5.14)$$

となるため (ただし, $y = x - ht$), テイラー展開を用いると,

$$var\hat{f}(x) \approx \frac{1}{nh} \int (f(x) - ht f'(x) + \dots) K(t)^2 dt + O(n^{-1}) \quad (5.15)$$

$$= \frac{1}{nh} f(x) \int K(t)^2 dt + O(n^{-1}) \quad (5.16)$$

$$\approx \frac{1}{nh} f(x) \int K(t)^2 dt \quad (5.17)$$

$f(x)$ は確率密度関数であり, $\int f(x)dx = 1$ より,

$$\int \text{var} \hat{f}(x) dx \approx \frac{1}{nh} \int K(t)^2 dt \quad (5.18)$$

となる. よって, 式 5.5 に式 5.11 と式 5.18 を代入すると, 二乗誤差は,

$$MISE(\hat{f}) \approx \frac{1}{4} h^4 k_2^2 \int f''(x)^2 dx + \frac{1}{nh} \int K(t)^2 dt \quad (5.19)$$

であり, 式 5.19 を最小とするバンド幅 h_{opt} は, 以下のように表される.

$$h_{opt} = k_2^{-\frac{2}{5}} \left(\int K(t)^2 dt \right)^{\frac{1}{5}} \left(\int f''(x)^2 dx \right)^{-\frac{1}{5}} n^{-\frac{1}{5}} \quad (5.20)$$

もとの密度関数を単峰性とみなすか, 多峰性とみなすかや, データが密に存在するか, 疎に存在するかなどの状況によって選択する必要がある. 例えば, 標準偏差を用いる手法は元の密度関数に正規分布を想定した上で最小誤差となるバンド幅を算出するものであるため, 多峰性の密度関数の推定には向かない. 多峰性の密度関数を推定するために, Rudemo (1982) や Bowman (1984) がクロスバリデーションを用いる方法を提案しているが, 観測データそのものとのクロスバリデーションを行っているため, データが密に存在しないといけなため, データが密でない今回のケースでは適用しにくい. また, 今回のケースでは, ピークが朝夕にある場合が多く, ピーク数が多くても 5 程度あれば十分である. そこで, 本研究では, 四分位範囲を用いて多峰性の密度関数にもある程度対応できるようにされた, Silverman (1986) が提案したバンド幅の算出方法を用いる. そのバンド幅 h_{opt} は, 式 5.20 に, カーネル関数として Gaussian 関数を代入し, 密度関数として正規分布を仮定すると, 以下のように表される.

$$h_{opt} = 1.06 \times \min \left(\sqrt{\text{var}(X)}, \frac{iqr(X)}{1.34} \right) \times n^{-\frac{1}{5}} \quad (5.21)$$

ここで, var は標本 X の分散, iqr は標本 X の四分位範囲である.

図 5.4 は図 5.3 をカーネル密度推定を用いて平滑化したものである. カーネル密度推定を用いることで, 図 5.3 のようにラウンドナンバーに時刻が偏っていたものを, 図 5.4 のように平滑化することができている. 朝夕の大きなピークと昼の小さなピークを再現しながら時刻による偏りを解消できている. 図 5.4 は H10 東京都市圏 PT 調査データの全てのトリップの出発時刻について平滑化したもので, 本研究では, 出発ゾーンゾーンごとのトリップ集合の出発時刻について同様の操作を行い, 人の流れ推定に推定された出発時刻密度関数を用いる.

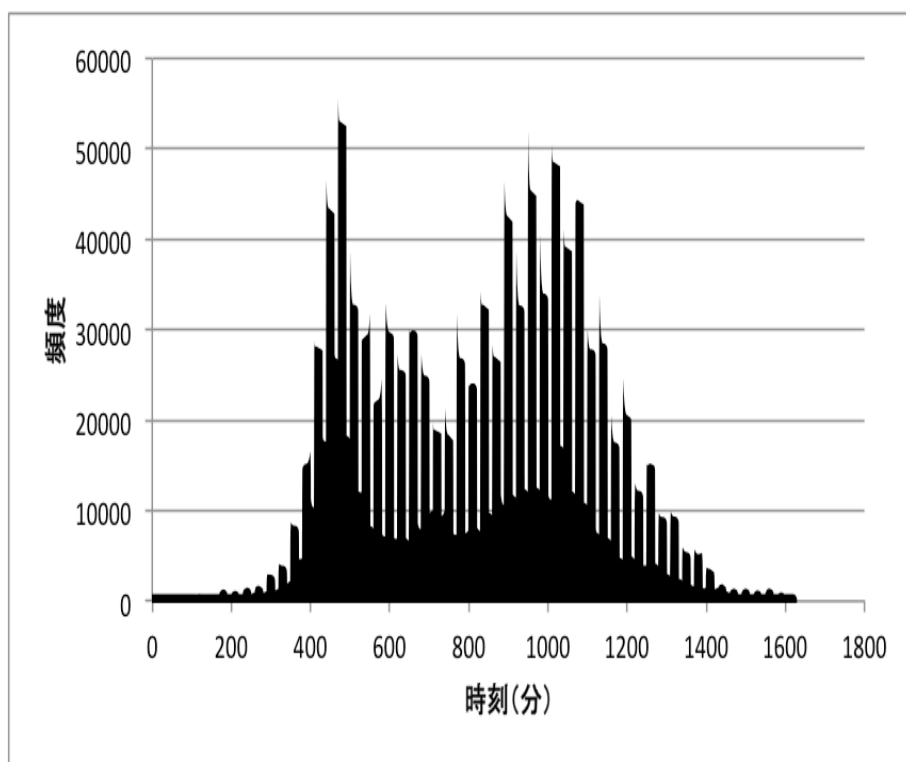


図 5.3: H10 東京都市圏 PT 調査データの出発時間の分布

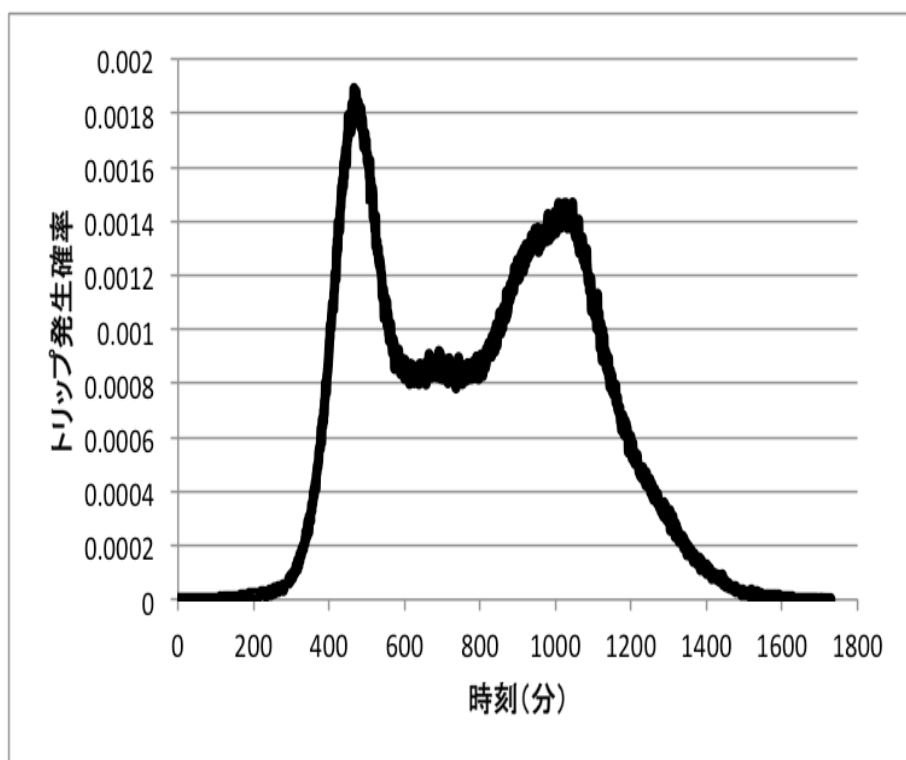


図 5.4: 推定されたトリップの出発時刻密度関数

5.2.3 パーティクル生成

各出発ゾーンごとにカーネル密度推定で求めた出発時刻密度関数を用いて分散された出発時間ごとにトリップを生成する。そして、生成される各トリップについてその解候補（パーティクル）となるトリップをいくつか生成する。パーティクルの生成には確率的にトリップを生成できる何らかのシミュレーションモデルを用い、生成されたパーティクルにはそれぞれ重みを与える。重みの初期値はシミュレーションモデルを反映したものとなるが、本研究では乱数を用いて生成する段階でシミュレーションモデルを反映させるため、結果として N 個のパーティクルを生成した場合、それぞれのパーティクルは $1/N$ の重みを持つことになる。

5.3 東京都市圏でのシミュレーションモデル

2章で述べたように、都市圏での人の流れ推定手法として、いくつかのシミュレーションモデルや全数拡大手法が提案されてきている。本研究は観測値でそうしたシミュレーションモデルの推定値を補正することによる精度改善への効果を研究することに重点を置いているため、パーティクルを生成するシミュレーションモデルとして、実データから行動パターンを生成して確率的に選択するデータ指向な手法と、多くの機会で用いられているロジットモデルを用いた行動選択モデルを構築して推定する手法の2つについて取り上げ、検証・考察を行う。

5.3.1 データ指向な人の流れ推定手法

まず、実データから行動パターンを生成して、そのパターンから確率的にパーティクルを生成する手法について述べる。2章で述べたように、近年データ取得手法や機器の発展により、行動パターンを生成するためのデータとして、様々なデータを用いることができるようになってきている。その中で、本研究では、現在最も偏りが少なく、データ量が多いPT調査データを行動パターンとして用いる。

出発ゾーンごとにPT調査データからトリップ集合を生成するまでは前節で説明した通りである。その後、各出発ゾーンのトリップ集合に含まれるトリップの行動パターンを用いて、パーティクルを生成する。トリップ集合のトリップはPT調査データのトリップであるため、出発時刻や到着時刻などのデータを持っているが、その情報を時刻ではなく経過時間にすることで、行動パターンとして用いる事ができる。例えば、7:00に出発ゾーン内

の自宅から出発し、7:10にA駅で電車に乗り、7:30にB駅で降車し、7:35に会社に到着したというトリップがあった場合、自宅を出発した後、10分後にA駅で電車に乗り、30分後にB駅で降車し、35分後に会社に到着するというような行動パターンとなる。トリップを行動パターンとして用いることで、各出発ゾーンは行動パターン集合を持つことになる。そして、ある出発ゾーンから発生するトリップがある場合、その行動パターン集合からランダムに行動パターンを選択していくことでパーティクルを生成できる。田尾手羽、図5.5はトリップ集合を生成し、トリップ数3のトリップ集合が生成されるゾーン3では、行動パターンが1/3の確率で選択される様子を表している。さらに、出発ゾーンに含まれる行動パターンを全てパーティクルとして生成し、それぞれに1/行動パターン数の重みを持たせ、複数のパーティクル生成を行う。例えば、あるゾーンのトリップ集合に300トリップが存在する場合、それぞれ重み1/300のパーティクルとされて、観測値で評価される。

携帯電話やスマートフォンの普及によりGPSデータの取得が容易になっていることから、GPSデータをよりリアルタイムに行動パターンとして利用することで、精度向上やコスト削減ができる可能性が考えられる。実データをもとにしたデータ指向な手法は単純ではあるが、大量データを取得し保持できるようになった現在の状況を有効に活かすことができるため、発展性の高い手法であると言える。

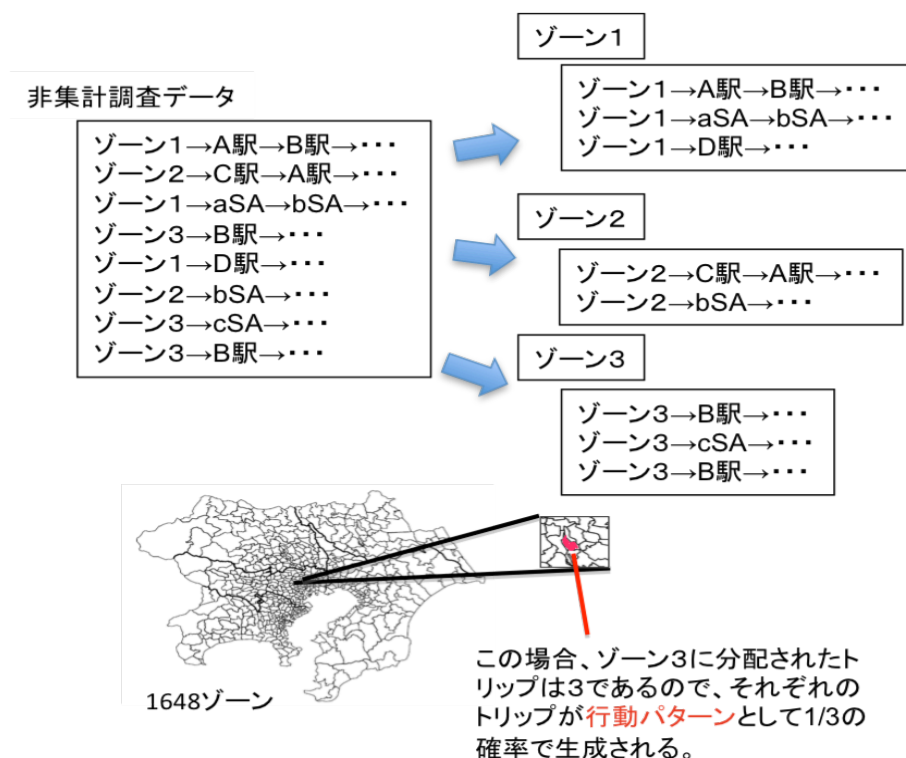


図 5.5: 行動パターン選択によるトリップ推定

5.3.2 行動選択モデルを用いた人の流れ推定手法

次に、行動選択モデルを用いた人の流れ推定手法について述べる。ロジットモデルとその派生モデルを利用した行動選択モデルの構築と適用は、今まで多く行われてきている。本研究では、いくつかある離散選択モデルのうち、Nested Logit モデル（以下、NL モデル）を用いて行動選択モデルの構築を行う。NL モデルは、通常の Logit モデルや Multinomial Logit モデルが持つ IIA 特性を回避することができる離散選択モデルであり、選択肢間の誤差項に相関のある選択問題を扱えることが特徴である。NL モデルは複数レベルからなるツリー構造を持っており、そのレベルごとに相関のない選択肢を比較することで、選択肢間の誤差項に相関のある選択問題を扱えるようになっている。

本研究で用いる NL モデルのツリー構造を図 5.6 に示す。本研究では、交通行動選択が図 5.6 にある 3 レベルの選択ツリー構造で表現できると仮定し、PT 調査データからそれぞれの選択肢の選択確率を求める際に必要となるパラメータを算出する。また、移動経路は最短経路を用いる。以下、図 5.6 のツリー構造を持つ NL モデルの選択確率の算出方法について述べる。

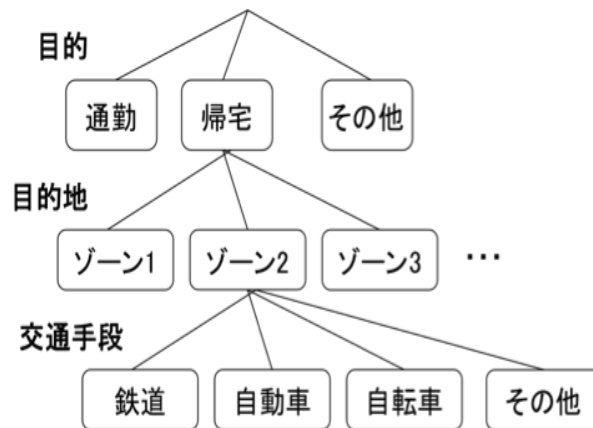


図 5.6: NL モデルのツリー構造

まず、活動目的 i 、目的地 s 、交通手段 m を選択する確率 $P(i, s, m)$ は、

$$P(i, s, m) = P(i)P(s|i)P(m|i, s) \quad (5.22)$$

と表される。ここで、 $P(i)$ は活動目的 i の選択確率、 $P(s|i)$ は活動目的が i のときの目的地 s の選択確率、 $P(m|i, s)$ は活動目的 i 、目的地 s のときの交通手段 m の選択確率である。以下、 $P(m|i, s)$ 、 $P(s|i)$ 、 $P(i)$ の順に詳しく述べる。

活動目的 i 、目的地 s のときの交通手段 m の選択確率 $P(m|i, s)$ は、以下のように表される。

$$P(m|i, s) = \frac{\exp(\mu_{s|i} V_{m|i, s})}{\sum_j \exp(\mu_{s|i} V_{j|i, s})} \quad (5.23)$$

$$V_{m|i, s} = \sum_j \beta_{j|i, s, m} x_{j|i, s, m} \quad (5.24)$$

ここで、 $V_{m|i, s}$ は活動目的 i 、目的地 s のときに交通手段 m を選択する効用の確定項で、 $x_{j|i, s, m}$ は効用を算出するさいに関連する変数、 $\beta_{j|i, s, m}$ はそのパラメータ、 $\mu_{s|i}$ は活動目的 i 、目的地 s における交通手段選択に関するスケールパラメータである。

次に、活動目的が i のときの目的地 s の選択確率 $P(s|i)$ は、以下のように表される。

$$P(s|i) = \frac{\exp(\mu_i V_{s|i} + \frac{\mu_i}{\mu_{s|i}} S_{s|i})}{\sum_j \exp(\mu_i V_{j|i} + \frac{\mu_i}{\mu_{j|i}} S_{j|i})} \quad (5.25)$$

$$V_{s|i} = \sum_j \beta_{j|i, s} x_{j|i, s} \quad (5.26)$$

$$S_{s|i} = \ln \sum_j \exp(\mu_{s|i} V_{j|i, s}) \quad (5.27)$$

ここで、 $V_{s|i}$ は活動目的 i のときに目的地 s を選択する効用の確定項で、 $x_{j|i, s}$ 、 $\beta_{j|i, s}$ は効用を算出するさいに関連する変数、パラメータ、 $S_{s|i}$ は交通手段選択に関するログサム変数、 μ_i は活動目的 i における目的地選択に関するスケールパラメータであるである。

最後に、活動目的 i の選択確率 $P(i)$ は、以下のように表される。

$$P(i) = \frac{\exp(\mu V_i + \frac{\mu}{\mu_i} S_i)}{\sum_j \exp(\mu V_j + \frac{\mu}{\mu_j} S_j)} \quad (5.28)$$

$$V_i = \sum_j \beta_j x_j \quad (5.29)$$

$$S_i = \ln \sum_j \exp(\mu_i V_{j|i} + \frac{\mu_i}{\mu_{j|i}} S_{j|i}) \quad (5.30)$$

ここで、 V_i は活動目的 i を選択する効用の確定項で、 x_j 、 β_j は効用を算出するさいに関連する変数、パラメータ、 S_i は活動目的選択に関するログサム変数、 μ は活動目的選択に関するスケールパラメータであるである。

ここで、ツリー構造の上位レベル（図 5.6 において、交通手段選択がレベル 1、目的地選択がレベル 2、活動目的選択がレベル 3 なので、交通手段選択が最上位レベル）ほど誤差が大きくなる。そのため、スケールパラメータ $\mu_{s|i}$ 、 μ_i 、 μ について以下の式を満たす必要がある (Carrasco et al. (2002), Hensher et al. (2002))。満たせない場合は、ツリー構造が間違っていると考えられる。

$$\mu < \mu_i < \mu_{s|i} \quad (5.31)$$

効用の確定項を算出する際に用いる変数を表 5.1 に示す。これらの説明変数に対応するパラメータとスケールパラメータを PT 調査データから求めることになる。具体的には、最尤法を用い、尤度が最大になるような最適パラメータセットを探索し、モデルを構築する。

表 5.1: 説明変数

選択レベル	説明変数
目的選択	出発時刻ダミー
目的地選択	$\ln(\text{目的地の通勤者数})$
	$\ln(\text{目的地の居住者数})$
	目的地までの距離の逆数
交通手段選択	所要時間

そしてモデルを構築した後、そのシミュレーションモデルを用いてパーティクルを生成する。シミュレーションによる推定では、最大確率となるトリップを選択したり、乱数を用いてトリップをランダムに選択したりすることで解を得ることが多い。本研究では1つのトリップをシミュレーションモデルから選択する代わりに、乱数を用いて複数のパーティクルを生成し、それらを観測値で評価することで補正する。実際には、 N 個のパーティクルを生成する場合、モデルから得られるトリップ候補の確率場から、 N 回乱数を用いてトリップを選択しパーティクルとする。

5.4 東京都市圏での実験

本章では、前章で述べたネットワーク空間での人の流れ推定手法を用いて、提案するデータ同化手法を用いた推定値の補正について検証・考察を行う。観測値による補正によって精度向上が見込めるかという検証を行う他、シミュレーションモデルの違いによる精度の違いや、パーティクル数と精度の関係、観測地点のコストパフォーマンス分析などを行い、提案手法を用いて推定する際に重要となる点について分析していく。

5.4.1 実験条件

実験は東京都市圏を対象として行う。東京都市圏は移動人口が日本の中では桁外れに多く、また、公共交通利用者も自動車利用者も十分いることから、実験対象として十分であると言える。しかし、東京都市圏の人の流れの全体を実際に計測することは難しいため、

PT 調査データから東京都市圏の人の流れを推定しても検証することが難しい。その一方で、完全に仮想の都市空間で仮想的な人の流れや観測データを想定しても現実への展開性が見込みにくい。そこで、本研究では提案する手法の検証を行うために、H10 東京都市圏での PT 調査データ約 80 万人分の人の流れを全体の真値とするような仮想都市圏を想定して実験を行う。

具体的には、推定すべき仮想都市圏約 80 万人分の人の流れに対して、仮想都市圏から 3% のサンプルを抽出し、非集計の観測データ、疑似 PT 調査データとする。そして、非集計の観測データから仮想都市圏を推定し、その推定値と仮想都市圏の人の流れ（PT 調査データ自体）を比較することで検証を行う。図 5.7 はその検証方法の概念を示したもので、検証可能である仮想都市圏と疑似 PT 調査データの関係を利用して検証を行う。また、データ同化に用いる集計の観測データとしては、東京都市圏の鉄道駅における 1 時間ごとの乗降客数と、道路交通センサスの観測地点における 1 時間あたりの交通量を得ることができると仮定して用いる。つまり、本研究の実験で用いる観測値は、PT 調査データ自体からなる仮想都市圏で実験のために作成されたものである。表 5.2 はこれらの実験条件を表したもので、真値として H10 東京都市圏 PT 調査データ、観測データとして、道路通過交通量、駅の乗降客数、疑似 PT 調査データを真値から作成した。また、道路データとして DRM(Digital Road Map)、鉄道駅は H10 東京都市圏 PT 調査当時の鉄道駅データを用い、表にある計算環境で人の流れ推定実験を行った。

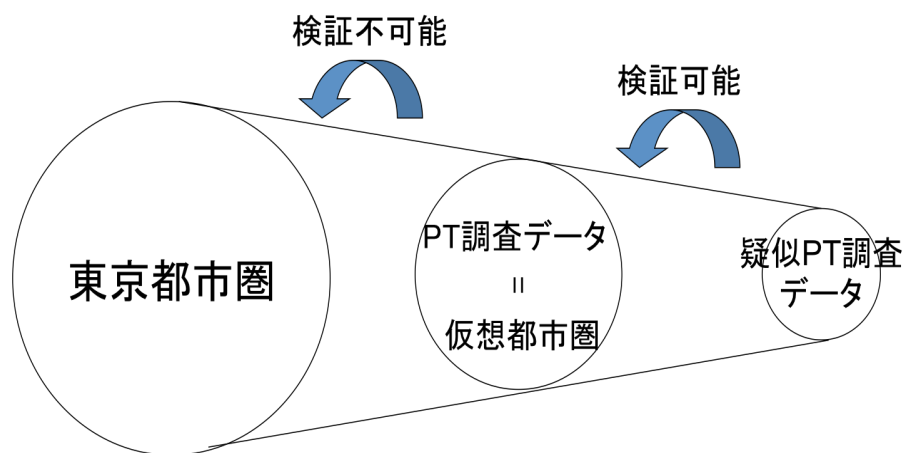


図 5.7: 検証方法の概念

表 5.2: 実験条件

種別		対象データ	説明
データ	真値	PT 調査データ	H10 東京都市圏 PT 調査データ
	観測データ	道路通過交通量	道路交通センサスの観測地点を利用：2105 箇所 (1 時間あたりの通過交通量) 真値とみなす 80 万人分の PT 調査データから作成
		駅の乗降客数	首都圏の全駅 (約 1400 駅) の乗降客数 (1 時間あたり) 真値とみなす 80 万人分の PT 調査データから作成
		非集計調査データ	2.4 万人の疑似 PT 調査データ (全体 80 万人分からランダムに 3%抽出)
	インフラデータ	道路データ	DRM (Digital Road map) を利用
		駅データ	H10 東京都市圏 PT 調査当時の鉄道駅データを利用
計算環境	ハードウェア		CPU : Core 2 Duo 3GHz, メモリ : 3GB
	プログラミング言語		Java

5.4.2 データ指向な人の流れ推定手法での実験

まず最初に、実データから行動パターンを生成し、その行動パターンを用いて人の流れを推定するデータ指向な手法を用いた実験を行い、検証・考察する。

鉄道駅の乗降客数や道路交通量について、提案手法により観測値で補正しつつ推定された人の流れと補正なしで単純に拡大した場合とで比較を行った。ここで、「単純に拡大」とは、複数のパーティクルを生成する代わりにランダムにトリップを一つ選択し、それを推定トリップとした仮想都市圏全体の人の流れ推定を指す。つまり、この「単純に拡大」した場合と比較することで、観測値を同化することによる効果を検証できると考えられる。

図 5.8 は時間帯別の鉄道駅乗車数に対する RMSE、図 5.9 は降車数に対する RMSE を計算したものである。ここで、乗車数と降車数はともに乗り換えは含まれていない。乗車数、降車数ともに、データ同化を用いた提案手法の方が単純な拡大に比べて精度が高いことが分かる。ただし、朝夕のラッシュに当たる時間帯では、その利用数の増加とともに誤差も大きくなっている。これは、朝夕のラッシュ時の移動の多様性に対し、疑似 PT 調査データを行動パターンとして生成した各トリップのパーティクルの多様性が低かったためと考えられる。今回の実験では、約 200 万トリップの推定のために約 2000 万のパーティクルが生成されたため、1 トリップあたり平均で 10 パーティクル生成されたことになる。この多様性を高めるためには、GPS データなどを用いて行動パターンを増やしたり、行動選択モデ

ルを構築したりする必要があり、生成するパーティクルの多様性をを高めることがさらなる精度向上につながると考えられる。その際に、発駅着駅ともに、最寄り駅として複数候補が存在する場合も考えられるため、その組合せを考慮した上でパーティクルを生成することが望ましい。

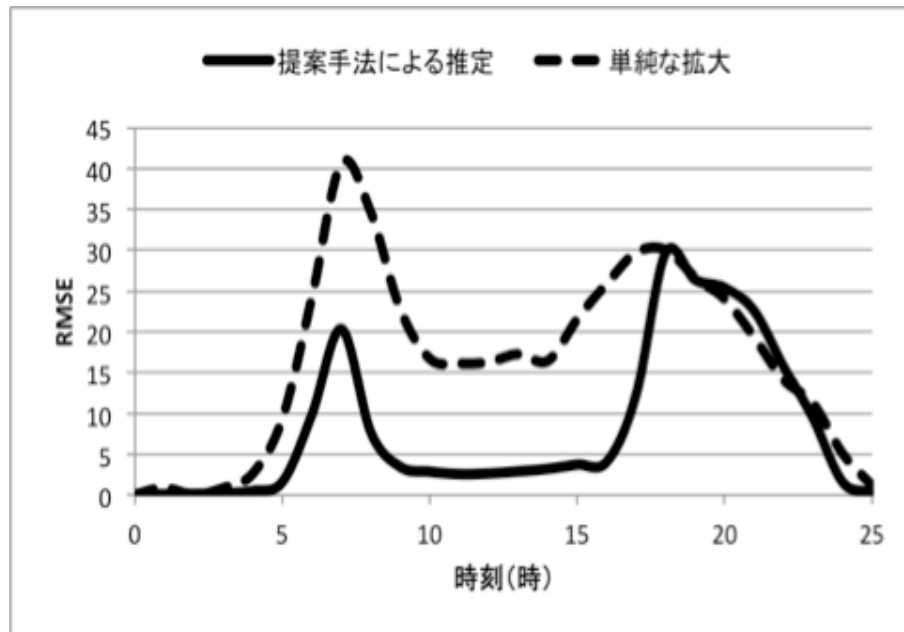


図 5.8: 時間帯別の鉄道駅乗車数の RMSE

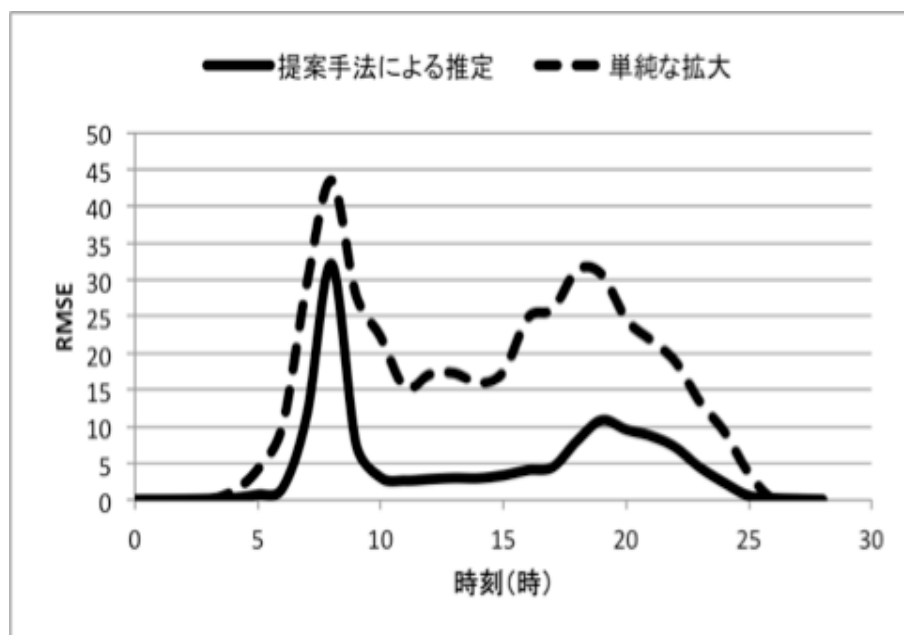


図 5.9: 時間帯別の鉄道駅降車数の RMSE

また、図 5.10 に観測値として用いた道路交通センサスの観測地点における道路交通量の時間帯別 RMSE、図 5.11 にそれ以外の道路も含めた全道路を対象とした時間帯別 RMSE を示す。道路交通センサスの観測地点における誤差は、データ同化を用いた提案手法によって約 50%ほど軽減できていることが分かる。全道路に対しても、単純に拡大した場合と比較すると誤差を小さくできているが、大きく改善するまでには至っていない。これは、道路交通センサスの観測地点が 2105 地点であり、その質、量ともに最適ではないからだと考えられる。そのため、今後の課題として、観測値を得る道路の割合やその選択について検討する必要がある。

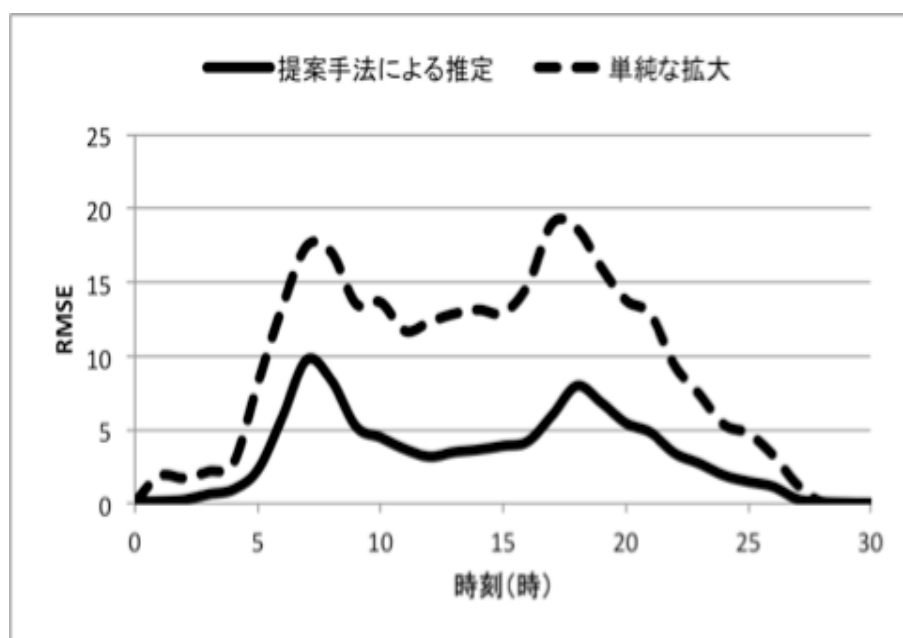


図 5.10: 道路交通センサス対象道路における時間帯別 RMSE

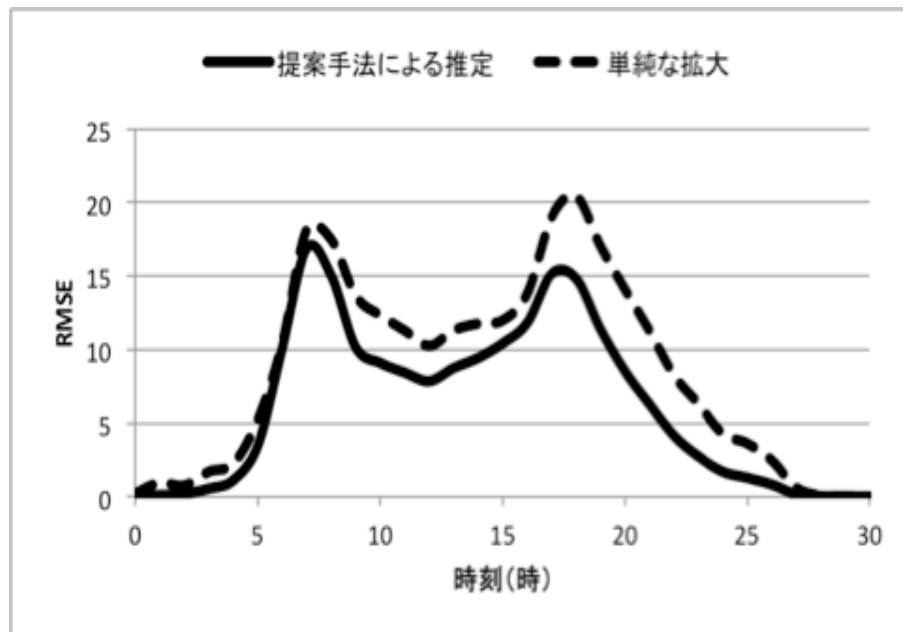


図 5.11: 全道路を対象にした時間帯別の RMSE

次に、データ同化に用いる観測値の違いによる推定精度の差異について検討する。

まず、観測値を得ることができる鉄道駅数を変化させた場合について考察する。ここまでの検証では全駅で観測値を得ることができると仮定していたのに対し、利用数上位 10%の鉄道駅でのみ観測値を得ることができると仮定した場合の推定値について検証を行う。

その結果として、図 5.12 に乗車数に対する時間帯別 RMSE、図 5.13 に降車数に対する時間帯別 RMSE を示す。どちらの場合でも、上位 10%の鉄道駅での観測のみで、単純な拡大に比べ精度を向上させることができている。さらに、全駅で観測した場合に比べると精度は落ちるが、ほぼ同等の精度で推定できていると言える。この結果から、利用数の多い駅で観測値を得ることでコストパフォーマンスの高い推定を行うことができると考えられる。

次に、観測する道路を利用数上位の道路（2105 地点）として推定を行い、道路交通センサスの観測地点（2105 地点）での観測から推定した場合と比較した。その結果として、図 5.14 に全道路における時間帯別 RMSE を示す。道路交通センサスの観測地点を用いた推定では、単純な拡大から少ししか精度の向上が見られなかったが、利用数上位の道路を用いた推定では、約 40%精度が向上している。道路の場合においても、利用数が多い道路で観測を行うことで、効果の高い推定を行えることが分かる。

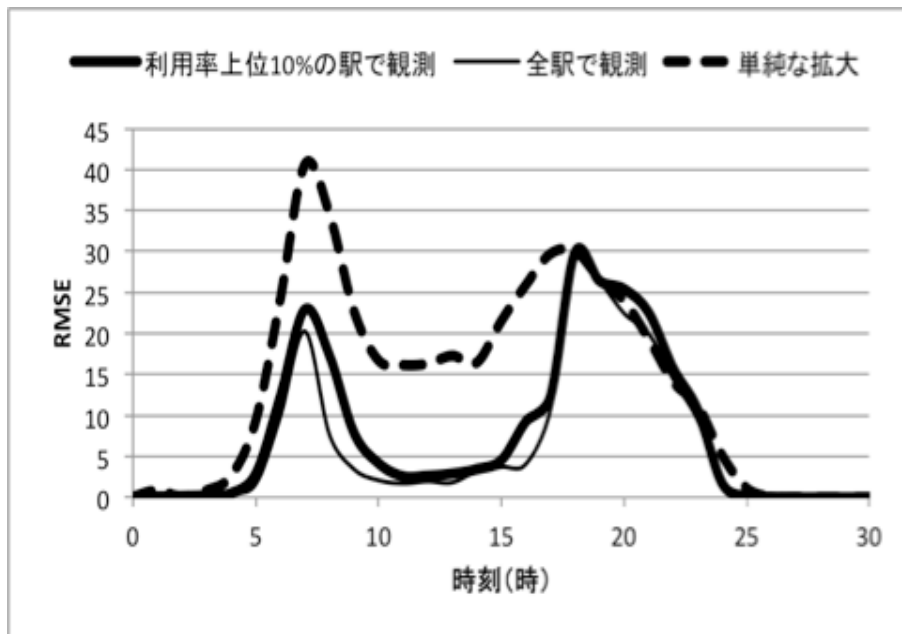


図 5.12: 道路交通センサ対象道路における時間帯別 RMSE

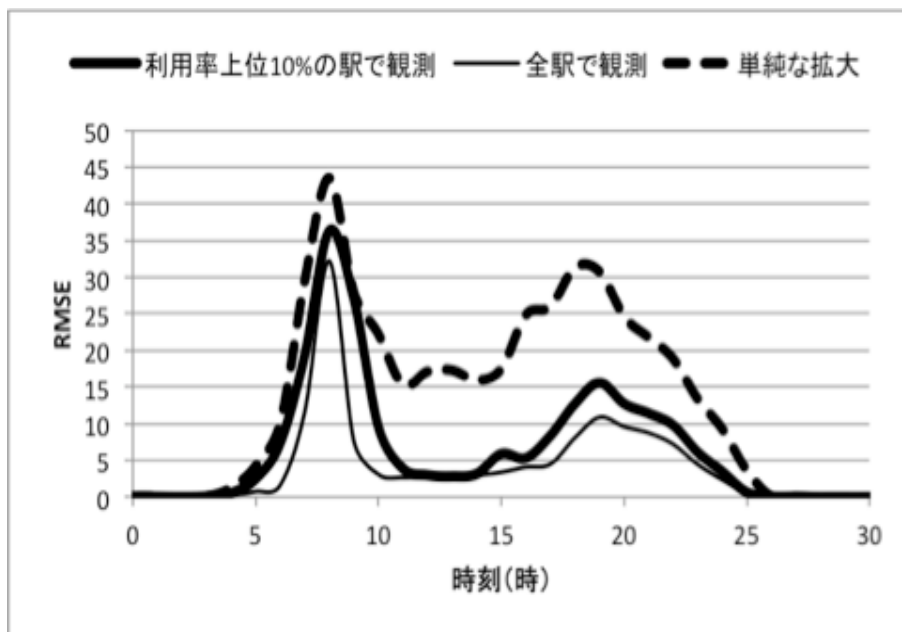


図 5.13: 道路交通センサ対象道路における時間帯別 RMSE

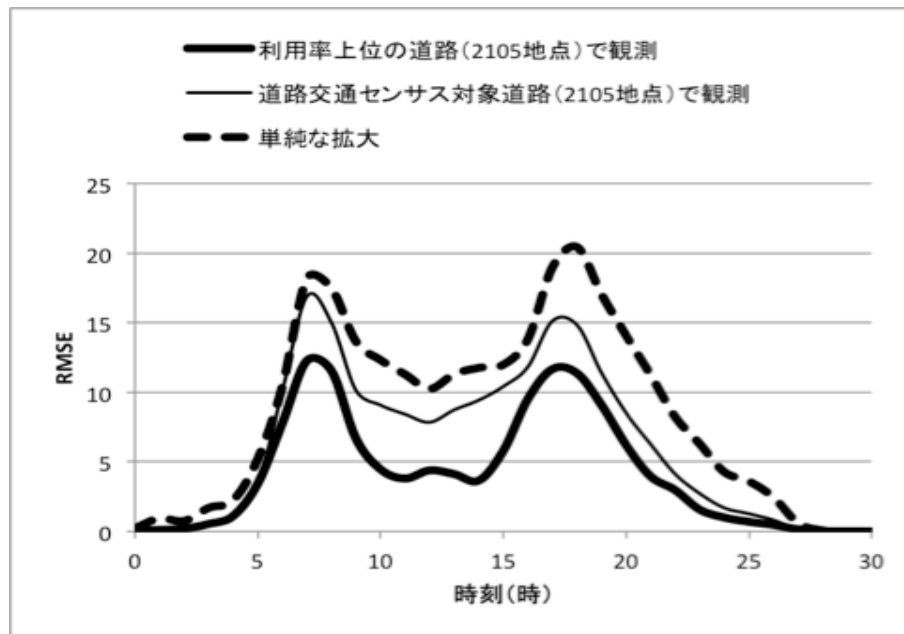


図 5.14: 道路交通センサス対象道路における時間帯別 RMSE

5.4.3 行動選択モデルを用いた人の流れ推定手法での実験

次に、NL モデルを用いて行動選択モデルを構築し、それと提案する観測値での補正手法を融合した人の流れ推定手法について検証・考察を行う。行動パターンからパーティクルを生成する手法では、その多様性の少なさから特に朝夕のラッシュ時に改善の余地が残っていた。

構築する NL モデルのツリー構造を図 5.15 に再掲する。ここで、モデルの説明力と精度の向上についても検討するために、図 5.15 のツリー構造を持つモデルと、そこから交通手段選択を省き、2 レベルの NL モデルかつ交通手段はランダムに選択されるモデルの 2 つを構築し、その比較も行う。

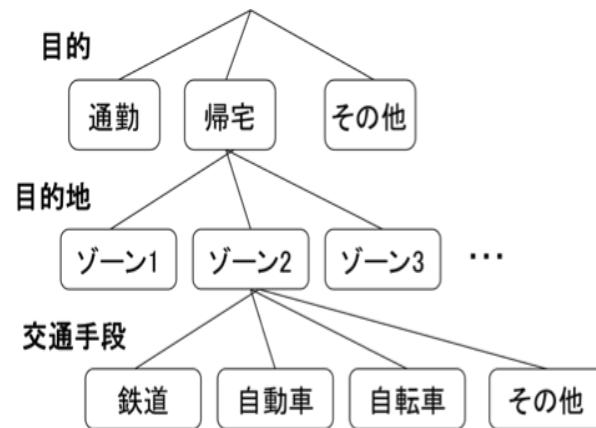


図 5.15: NL モデルのツリー構造（再掲）

まず、2 レベルの NL モデルのパラメータを最尤法を用いて算出した結果について述べる。表 5.3 は目的地選択をする際のパラメータで、表 5.4 は目的選択のパラメータ算出結果である。どのパラメータも 5% 有意で有効であり、数値の正負も全て正の値になっており、想定通りの結果となっている。目的選択パラメータのスケールパラメータは、上位レベルの目的地選択のスケールパラメータより小さい値になっており、式 5.31 を満たしている。

表 5.3: 2 レベル NL モデル目的地選択パラメータ

変数名	推定値（通勤目的）	推定値（帰宅目的）	推定値（その他目的）
スケールパラメータ	1.0 -	0.6548 **	0.9770 **
ln(目的地の通勤者数)	0.5737 **	-	-
ln(目的地の居住者数)	-	0.6120 **	-
距離の逆数	0.4208 **	1.48 **	1.2619 **
サンプル数	10935	25696	23497
初期尤度	-16448.70	-45453.09	-37898.09
最終尤度	-14980.48	-42597.15	-36515.38

** : 5% 有意 * : 10% 有意 - : 設定値

表 5.4: 2 レベル NL モデル目的選択パラメータ

変数名	推定値（通勤）	推定値（帰宅）	推定値（その他）
スケールパラメータ		0.01445 **	
定数項	-3.4748 **	-3.0579 **	-
2-8 時台	3.1565 **	-	-
17-23 時台	-	1.6527 **	-
7-13 時台	-	-	1.1470 **
サンプル数		60128	
初期尤度		-66057.36	
最終尤度		-42294.10	

** : 5%有意 * : 10%有意 - : 設定値

次に、交通手段選択を含めた 3 レベルの NL モデルのパラメータ算出結果について述べる。表 5.5 は交通手段を選択する際のパラメータ、表 5.6 は目的地選択をする際のパラメータで、表 5.7 は目的選択のパラメータ算出結果である。どのパラメータも 5%有意で有効であり、数値の正負も全て想定通りの結果となっている。目的選択、目的地選択、交通手段選択のスケールパラメータは、上位レベルのスケールパラメータが大きい値となっており、式 5.31 を満たしている。

表 5.5: 3 レベル NL モデル交通手段選択パラメータ

変数名	推定値
スケールパラメータ	1.0 -
定数項（徒歩）	4.5564 **
定数項（自動車）	4.3777 **
定数項（鉄道）	4.2695 **
サンプル数	60128
初期尤度	-83355.11
最終尤度	-45831.17

** : 5%有意 * : 10%有意 - : 設定値

表 5.6: 3 レベル NL モデル目的地選択パラメータ

変数名	推定値 (通勤目的)	推定値 (帰宅目的)	推定値 (その他目的)
スケールパラメータ	0.51 **	0.703 **	0.4479 **
ln(目的地の通勤者数)	0.9065 **	-	-
ln(目的地の居住者数)	-	0.5727 **	-
距離の逆数	0.7448**	1.3224 **	2.2373 **
サンプル数	10935	25696	23497
初期尤度	-16448.70	-45453.09	-37898.09
最終尤度	-14985.81	-42576.26	-35903.82

** : 5%有意 * : 10%有意 - : 設定値

表 5.7: 3 レベル NL モデル目的地選択パラメータ

変数名	推定値 (通勤)	推定値 (帰宅)	推定値 (その他)
スケールパラメータ		0.0573 **	
定数項	-4.5205 **	-0.5981 **	-
2-8 時台	4.8215 **	-	-
17-23 時台	-	2.8097 **	-
7-13 時台	-	-	1.6665 **
サンプル数		60128	
初期尤度		-66057.36	
最終尤度		-38288.03	

** : 5%有意 * : 10%有意 - : 設定値

以上のように構築したモデルを用い、まずはパーティクル生成手法の違いによる精度の変化について考察する。比較対象として、行動パターンをもとにパーティクルを生成する手法、活動目的と目的地選択を NL モデルで記述して生成する手法、交通手段選択まで含めた NL モデルを用いて生成する手法の 3 パターンを用意し、それに観測値による補正を行わない単純な拡大を加えた 4 つのパターンについて比較・考察を行う。また、推定する 1 トリップにつきパーティクルを 5 個生成し実験を行った。

結果として、図 5.16 に時間帯別の鉄道駅乗車数の RMSE、図 5.17 に時間帯別の鉄道駅降

車数の RMSE, 図 5.18 に道路交通センサ対象道路における時間帯別 RMSE を示す。全ての図から、NL モデルを用いてパーティクルを生成することで、単純な拡大や疑似 PT 調査データを行動パターンとしてパーティクルを生成する手法よりも観測値に即した推定結果を得ることができていることが分かる。これは図 5.19 のように、NL モデルを適用することで、確率的により多様なパーティクルを生成することが可能となっているからだと考えられる。しかしその一方で、NL モデルを用いた 2 つのパターンでは大きな違いはない。2 つのモデルは交通手段選択モデルがあるかどうかの違いであり、このことから交通手段選択モデルがなくても観測値を使うことで精度を上げることができていると言える。

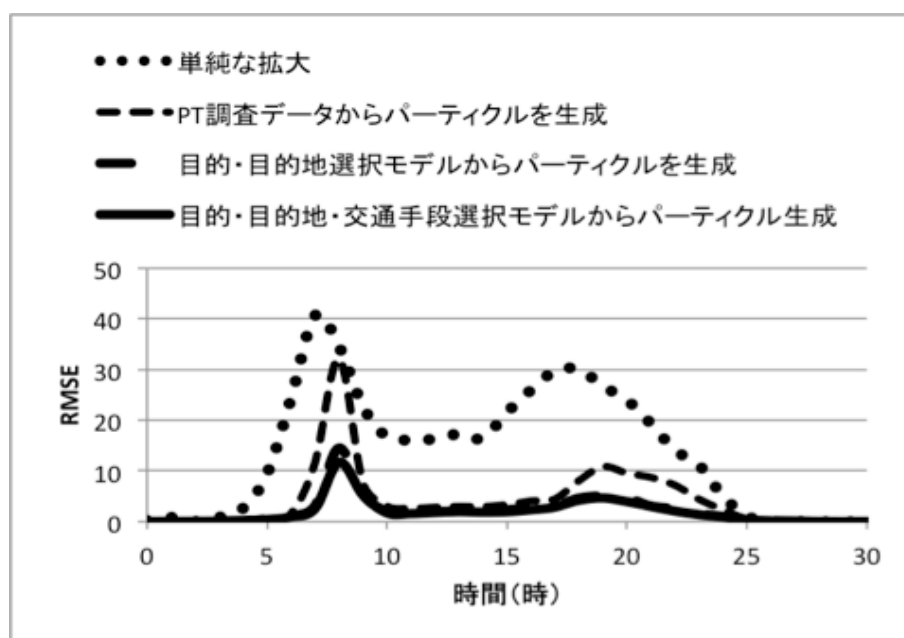


図 5.16: 時間帯別の鉄道駅乗車数の RMSE

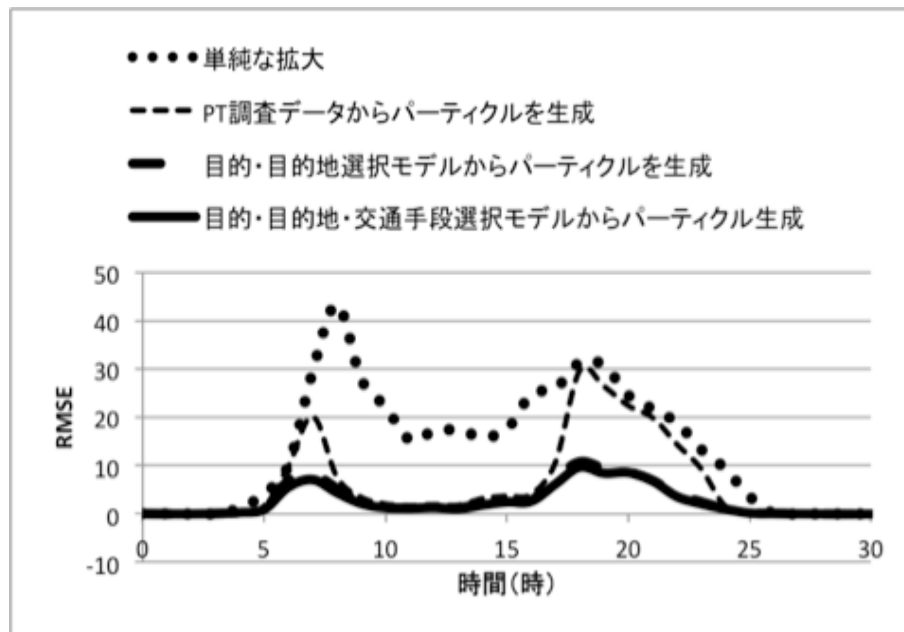


図 5.17: 時間帯別の鉄道駅降車数の RMSE

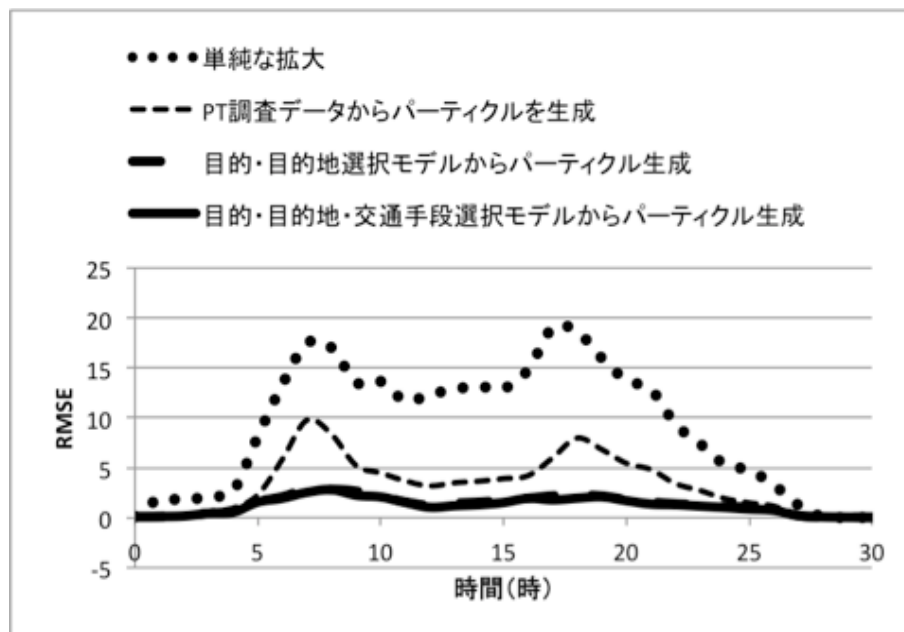


図 5.18: 道路交通センサス対象道路における時間帯別 RMSE

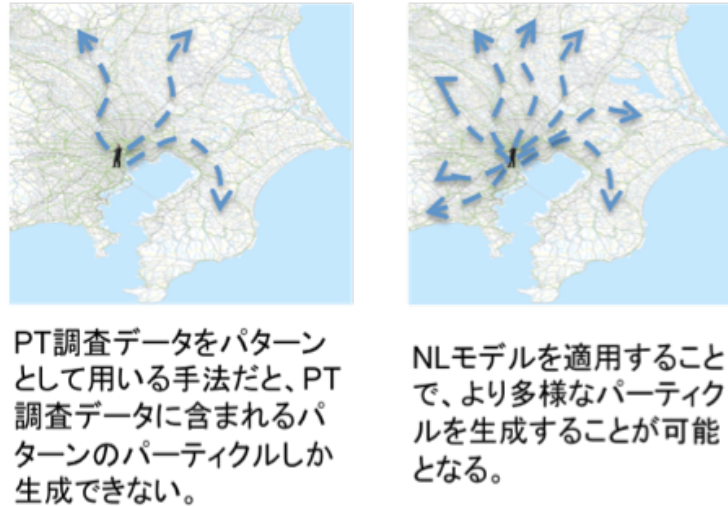


図 5.19: NL モデルを適用する利点

次に、パーティクル生成数と精度の関係を調べてみた。パーティクル生成手法としては、NL モデルを活動目的、目的地、交通手段選択の 3 レベルまで構築したものを使う。パーティクル数を 5 個から 5 個刻みに、10 個、15 個、…、100 個まで増やし、駅乗降客数との RMSE からパーティクル数と精度の関係を考察する。図 5.20 はその結果として、鉄道駅における乗車数、降車数の時間帯別 RMSE の 1 時間あたりの平均値とパーティクル数との関係を示したものである。図 5.20 から 100 個まで増やした段階でほぼ収束していることが分かる。パーティクル数はモデルで記述される確率分布を表現できるだけの数があることが理想であるため、計算リソースとの兼ね合いがあるが、誤差が落ち着く段階まで増やせることが望ましい。計算リソースとの兼ね合いは、東京都市圏のような広範囲な推定を行う場合に大事な課題の一つに挙げられる。

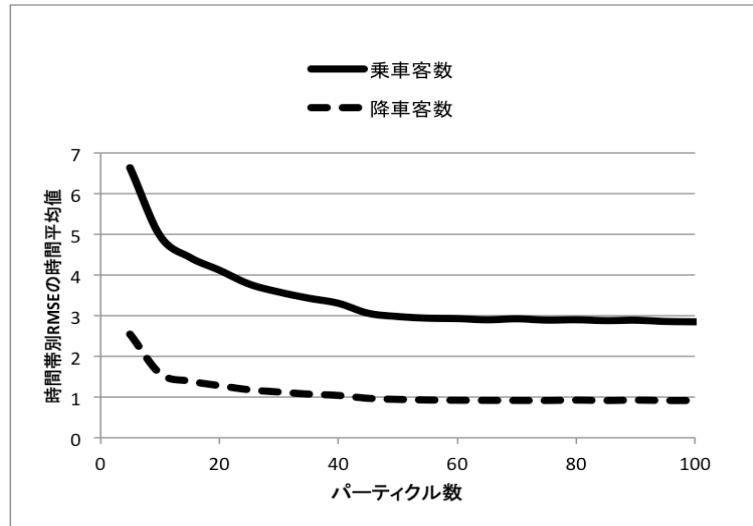


図 5.20: 鉄道駅における乗車・降車客数の時間帯別 RMSE の 1 時間平均とパーティクル数の関係

次に、観測地点数と精度の関係について考察を行う。図 5.21 は全ての駅で観測していたのを全体の 9 割、8 割、…、1 割と減らしていったときの時間帯別 RMSE の平均値を示したものである。ここで、観測地点とする鉄道駅は利用客数が多い順に選択した。図 5.21 から、観測地点数が全鉄道駅の 5 割から 6 割より少なくなると精度が急に悪くなっていくことが分かる。これは自由空間における実験での結果とも合致している。逆に、全鉄道駅の半数程度の観測で、全鉄道駅で観測するのと同程度の効果が得られると言え、コストパフォーマンスを高めることが要求される広範囲での推定で有用な結果だと考えられる。

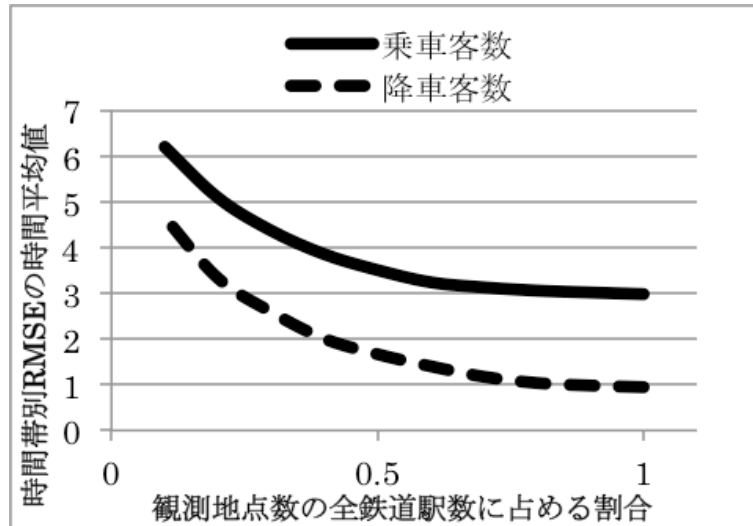


図 5.21: 鉄道駅における乗車・降車客数の時間帯別 RMSE の 1 時間平均と観測地点数の関係

最後に、観測値の質とその効果について検討を行う。ここでは鉄道駅の観測値、乗降客数について考察を行う。鉄道駅 i の観測値の質を表す指標 I_i として、以下の値を用いる。

$$I_i = -N_{ij} \sum_j P_{ij} \log(P_{ij}) \quad (5.32)$$

ここで、 j は観測値が乗車客数の場合はその降車駅、降車客数の場合はその乗車駅を表し、 P_{ij} は観測値全体に対する、鉄道駅 i で乗車（または降車）し鉄道駅 j で降車（または乗車）した人数の割合である。目的駅や出発駅がばらけているほど I_i は大きい値となり、関係する駅が多いほど大きい値を取ることになる。例えば、単線駅は利用のされ方の種類が少ないため小さい値に、乗換駅は利用のされ方が様々なので大きい値になりやすい。

この指標を用いた分析結果として図 5.22 を示す。鉄道駅の観測地点としてランダムに半数を選択肢、それぞれについて式 5.32 の値の平均と時間帯別 RMSE の 1 時間あたりの平均値を比較した。乗車と降車のどちらの場合も情報量の平均が高くなると誤差が小さくなっている傾向が見て取れる。図 5.23 は観測人数の平均と精度の関係を示したものである。この図から、観測人数が多くなると精度が良くなることが分かるが、この時の相関係数は 0.6 で、図 5.22 の相関係数は 0.77 であることから、情報量を組み合わせることで、さらに精度の高い地点の抽出ができていけると言える。このことから、関係する駅が多い乗換駅のような地点での観測値を用いることで精度向上の効果を高めることができると考えられる。

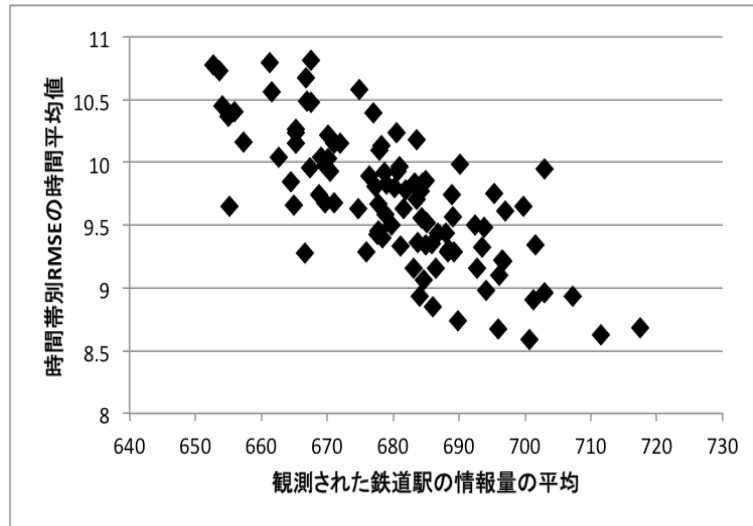


図 5.22: 鉄道駅における乗降客数の時間帯別 RMSE の 1 時間平均と観測鉄道駅の情報量の平均の関係

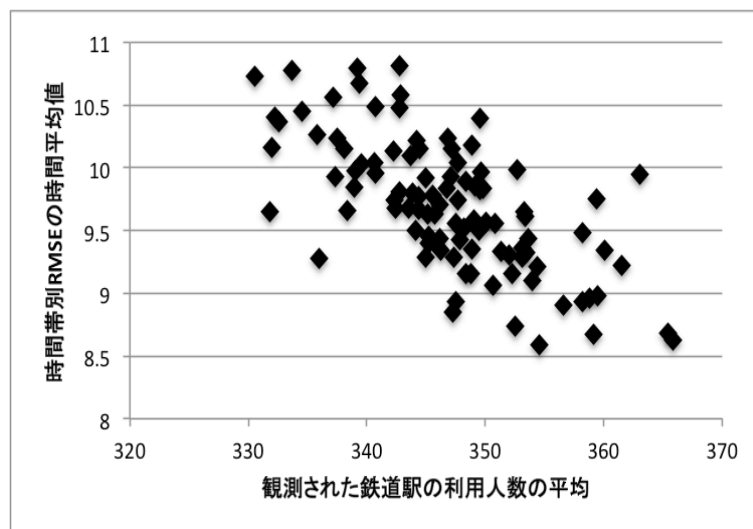


図 5.23: 鉄道駅における乗降客数の時間帯別 RMSE の 1 時間平均と観測鉄道駅の観測人数の平均の関係

参考文献

- B.W. Silverman, 1986, "Density Estimation for Statistics and Data Analysis" Chapman & Hall/CRC
- Rudemo, M. 1982. Empirical choice of histograms and kernel density estimators. Scandinavian journal of Statistics, 9(2), 65-78.
- Bowman, A. W. 1984 An alternative method of cross-validation for the smoothing of density estimates. Bio-metrika, 71(2), 353.
- H. Shimazaki and S. Shinomoto, 2010, "Kernel bandwidth optimization in spike rate estimation", J Comput Neurosci 2010, 29:171-182
- J.L.Bowman, 2000, M.E.Ben-Akiva: Activity-based disaggregate travel demand model system with activity schedules, Transportation Research Part A 35.
- Carrasco JA, de Dios Ortúzar J, 2002, Review and assessment of the nested logit model. Transport Reviews 22(2):197-218
- Hensher DA, Greene WH, 2002, Specification and estimation of the nested logit model: alternative normalisations. Transportation Research Part B 36(1):1-17

第6章 結論

6.1 本研究の成果

本研究では、確率的に人の流れ推定するシミュレーションモデルと、データ同化手法を用いた観測値の融合手法を提案し、大崎駅構内での人の流れ推定への適用と東京都市圏での人の流れ推定への適用を行い、それぞれについて検証した。

全体のフレームワークとして、確率的にシミュレーションができる全てのモデルを対象として、データ同化手法の一つであるパーティクルフィルタ法と最適内挿法を用いてモデルの推定値を観測値で補正する手法を提案した。データ同化手法を融合することの利点として、モデルによる推定をベースに、モデルの説明力を保持したまま、推定値を補正できることが挙げられる。観測値とシミュレーションモデルを融合して精度向上を目指す研究は他にもいくつかあるが、本研究と他の研究との違いとして、求める対象が移動軌跡の確率分布であり、これは非連続で離散的で分散も計算できない変数の確率分布である点が挙げられる。こうした対象に対してデータ同化手法を用いた推定手法を提案している研究はまだなく、新規性が高いと言える。具体的には、パーティクルフィルタ法を用いて重みつきサンプリング集合の生成とフィルタリングを行い、各サンプルの重みを観測値で最適内挿法を用いて補正することで、観測値に即した推定手法を提案した。また、最適内挿法を用いて厳密に推定するには、 $O((\text{推定する人数} * \text{パーティクル数})^2)$ の計算コストがかかってしまうため、 $O(\text{推定する人数} * \text{パーティクル数})$ の計算コストで計算できる近似計算手法も提案した。

本研究では、大崎駅構内と東京都市圏に対して、提案手法の検証を行った。大崎駅構内では、厳密な計算を行い、提案手法の効果を検証し、さらに、パーティクルフィルタ法による予測値の改善と近似計算の精度、計算時間の比較を行った。東京都市圏では、近似計算を用い、シミュレーションモデルと精度の関係や、観測地点数と精度の関係、観測地点の評価を行った。

大崎駅構内における従来の歩行者モデルでは、同じ目的地の歩行者については経路が一意に決まってしまうというなどの課題があったが、ポテンシャルマップ法を拡張することで、確率的に軌跡を決定できるモデルを構築した。経路選択を確率的に行うモデルはこれ

までほとんどなく、既存のポテンシャルマップという研究成果を活かした上で応用できることから、今後多くの分野で活用することが可能だと考えられる。人の流れをシミュレーションしたり、現状推定したりするニーズは非常に高まってきているため、大きく貢献できる。さらに、ロボティックスの分野でも経路選択は非常に重要な要素であるため、応用していく余地があると考えられる。

大崎駅構内での実験結果として、シミュレーションモデルのみの推定値に対し、提案手法により観測値で補正することで精度を向上させることに成功した。また、若干ではあるが、パーティクルフィルタ法により予測値の精度を向上させることにも成功した。ただし、最適内挿法による重みの更新の方がより精度向上に寄与しており、フィルタリングによる精度向上については課題のひとつとして挙げられる。近似計算の精度検証も行い、結果として、近似計算で十分な精度を確保できていた。また、計算速度がパーティクル数 100 個の場合で約 1/1000 になった他、増加率もほぼ比例にすることに成功した。

東京都市圏におけるシミュレーションモデルとして、データ指向なモデルと NL モデルを用いたモデルを挙げ、それぞれについて提案手法を適用した。後者のモデルは都市圏のシミュレーションモデルとして良く用いられているもので、前者のものより高い精度となったが、一方で、GPS データなどの取得が容易になったことから、前者のデータ指向なモデルとデータ同化手法を組み合わせることで簡易に推定するということが今後重要になってくると考えられる。また、6 割程度の観測地点で全数観測した場合と同等の精度を確保できることが分かった。観測値の効果などについても検討し、多様な動線の集まる観測地点での観測値の効果が高いことが分かった。

第二章では、計測技術・手法とシミュレーションモデル、観測値によるシミュレーションモデルの補正に関する既往の研究についてまとめ、その課題を抽出することで、本研究の目的を明確にした。

第三章では、本研究の提案手法のフレームワークについて述べ、データ同化手法に触れるとともにその適用方法を説明した。軌跡の確率分布算出にパーティクルフィルタ法を、サンプリングされた移動軌跡候補の重みを観測値で更新する手法に最適内挿法を用いる手法を提案した。また、大規模推定への適用を可能とするための近似計算手法についても提案した。

第四章では、駅構内での人の流れ推定への適用について述べた。まず推定手法全体のフレームワークについて述べ、次に歩行者モデルについて詳述した。歩行者モデルとして、ポテンシャルモデルを拡張することで、確率的に軌跡を生成することのできるモデルを構築した。また、大崎駅構内で提案する推定手法の検証を行い、提案手法により精度向上でき

ることや、用いる観測値などについて検討した。まず各出口の利用人数を観測値として得られると仮定して推定を行い、提案するデータ同化手法を用いることでより精度の高い推定を行えることを示した。また、パーティクルフィルタ法による予測値の改善や近似計算の精度についても検証した。

第五章では、都市圏での人の流れ推定への適用について、東京都市圏を対象とした推定について述べた。シミュレーションモデルとして、データ指向なモデルとNLモデルを用いたモデルについて述べ、それらを用いて実験を行った。提案手法で推定値を補正することで精度できることを示した。また、パーティクル数や観測地点についての検討も行い、パーティクル数は100個程度あれば十分であり、観測地点数は6割程度あれば十分であることを示した。これらは、計算量が重要となる、都市圏などの広範囲な領域で推定する時に有用な知見となる。また、多様な動線が集まる観測地点での効果が高いことも示した。

6.2 今後の課題

今後の課題として、手法については、観測値の効果などのより理論的な考察、フィルタリング方法の改善、近似計算方法の改善などが挙げられ、実用的な面からは、大規模推定への適用のための計算コスト削減、経路推定の導入、非集計観測値の利用、リアルタイム推定へ向けた計算環境の整備などが挙げられる。以下、それぞれ簡単に述べる。

6.2.1 より理論的な考察

観測値の効果の検討などを適用実験で行ったが、より理論的に考察できた方が望ましい。数式を使って理論的に考察したり、シンプルなシミュレーションを行って、より要因を分けやすい状況で実験したりして考察できると良い。

6.2.2 フィルタリング方法の改善

本研究は、パーティクルフィルタ法を用いて予測値の改善を行える手法を提案したが、フィルタリングについては大きな効果があるものとはなっていない。これはフィルタリングの対象が各人の移動軌跡の確率分布であり、シミュレーションモデルに直接関与しないことも要因の一つであると考えられる。そのため、補正結果を元に、その後のシミュレーションモデルによる推定を改善していく仕組みを加えることができれば、より観測値を活かした予測が行えると考えられる。

6.2.3 近似計算の手法改善

大規模計算への適用のために、本研究では2つの仮定をおいた近似計算手法を提案した。しかし、仮定のうちのひとつである、各サンプルは1時点間に複数箇所で観測されない、という状況とはならない場合も多いと考えられる。そこで、厳密に計算する際に用いる行列計算を他の仮定において、より高精度に近似する手法がある可能性も高いので、それについて検討することは大きな意義があると言える。

6.2.4 大規模推定への適用へ向けた計算コストの削減

近似計算によってある程度現実的な計算時間にはなったが、それでも東京都市圏全数の推定などを想定すると、計算コストを削減する必要がある。人数が増えるとその分同じような条件の人が多くなると考えられるので、同一条件の人をまとめて集団として推定する方法が考えられる。

6.2.5 経路推定の導入

経路推定を導入することも課題の一つとして挙げられる。都市内での移動を考える際、道路単位で推定できると、施策、防犯、マーケティングなど、多くの面でより重要な情報となる。一方で、移動の多様性や候補の多さなどから、経路推定モデルを構築することは簡単ではなく、それだけで大きな研究テーマの一つと言える。また、その推定をする際にも、観測値と融合し精度を向上させることが有益であると考えられる。GPS データなどを活用してよりデータ指向に経路推定を行うことは今後の課題として取り組むのに大きな価値があると言える。

6.2.6 非集計観測値の利用

本研究では集計観測値を同化する手法について提案した。それに加えて、非集計観測値も融合できると幅が広がり、実用性も高まる。非集計観測値を同化するには、基本的には、そこから平均速度など、代表値を計算して用いるなど、うまく加工する必要がある。非集計観測値はその可能性が大きい分だけ利用するための処理が必要となる。そうした利用方法を検討することも課題の一つと言える。

6.2.7 リアルタイム推定

本研究では、将来的にはリアルタイムに推定することを目指して手法を構築した。提案手法は、逐次的に重みの更新を行っているので、どの時点でもその時点での重みを用いて次時点の位置を推定することができるため、リアルタイムで推定することも十分可能であると考えられる。しかし、リアルタイムに推定を行うには、観測データの転送時間や推定の計算時間、ハードウェア構成などを検討していかないといけないため、今後取り組むべきテーマの一つと言える。

6.3 今後の展望

本研究の提案手法が実用化されると、例として以下のような利用が期待できる。

- 建物内や駅、ショッピングモールなどで、シミュレーションモデルと合わせて、より現実に即した推定を行う。
- PT 調査などの都市圏に対する調査を、取得困難な非集計データを減らし、比較的取得しやすい集計観測値で代替できる。
- リアルタイムに近い形で常に推定できる状況になれば、官民、個人を問わず様々なサービスや改善に利用できる。

例えば、駅構内で案内板の設置方法を変えた際に、その変化を分析できるとその後の改善に役立つ。また、ショッピングモールなどで各店舗が集客のための工夫を日々積み重ねているが、日々の人の流れを推定できれば、そういった工夫の効果やさらなる改善へつながる。さらに、快適な空間作りには人の動線を考慮することは必要不可欠であるため、期待した人の流れになっているかなどを確認することで、空間作りやまちづくりにも大きな利点となる。このように日々の人の流れを把握できると、マーケティング、設計、防犯など様々な面で効果が期待できる。

また、現在ではこうした人の流れを把握するために、PT 調査や大都市交通センサスなどのアンケートベースの調査が良く行われている。都市圏規模の調査はサンプリング調査であっても、大規模なものとなり、コストがかかる。例えば、東京都市圏は3%のサンプル率だが、80万人の調査を行っている。こうした個人データを減らし、比較的取得容易な集計観測値で補えるようになると、コストの面で効果が高いだけでなく、調査の頻度を上げるなどの効果も生み出せる。

このように、本研究の適用によりいくつかの効果が期待できるが、そもそもリアルタイムに近い形で人の流れを常に推定できる状況になれば、その効果は計り知れないと考えられる。近年スマートフォンの急速な普及により、こうした情報と実生活を結びつけることが容易になってきている。例えば、普段行かない場所へ行くことになった場合、多くの人が経路検索を行い、乗り換え案内や行き先案内を利用している。そこに人の流れの情報が入れれば、混雑した時間を外して移動したり、似たような人が良く利用する経路を探したりといった付加価値をつけてサービスを提供することができるようになる。Webサービスの活用と提供が容易にできるようになったことから、人の流れを推定するだけで多くのサービスへとつながると考えられる。

謝辞

多くの方々のご指導ご協力があり、ここに一篇の博士論文を仕上げることができました。皆様への感謝の気持ちを込めて、ここにお礼の言葉を述べさせていただきます。

主査である柴崎亮介教授には、研究の指導をはじめ、国内外の学会への参加など、二年に渡る研究生活において充実した環境を提供していただきました。特に、先生の研究への心構えや熱意に大きな刺激を受けることができ、大きな糧となっています。日本の地理空間情報の未来を担う活動で日々多忙である中、多く時間を割いて助言や指導をいただくことができ、非常に価値あるものとなりました。心より感謝いたします。

副査の出口敦教授には、都市計画の観点から貴重なアドバイスを頂き、論文構成、要旨に至るまで丁寧に指導して頂きました。また、曖昧になっていた言葉の使い方にもご指摘を頂くことができ、より伝わりやすい論文になったと思っています。厚くお礼申し上げます。

副査の貞広幸雄教授には、特に結果の分析手法や考察について深く議論して頂き、統計的な観点から多くの示唆を頂きました。全てを活かしきれなかったところが残念ですが、今後の活動に活かしていきたいことを多く教えて頂きました。厚くお礼申し上げます。

副査の布施孝志准教授には、方法論的な観点から貴重なアドバイスを頂くことができ、より理論的裏付けの取れた推定手法の構築ができました。また、多研究との比較についても細かくご指摘くださり、より明確に違いを考察でき、位置づけがはっきりしました。厚くお礼申し上げます。

副査の関本義秀特任准教授には、同じプロジェクトに参加させて頂き、多くの失敗にも粘り強く指導していただくことができ、優しく時に厳しいそのご指導によって、本論文を書き上げることができました。豊富な知見から多くの貴重なアドバイスを頂くことができ、また、研究への真摯な姿勢に学ぶことが多く、大変勉強になりました。厚くお礼申し上げます。

邵肖偉特任講師には、日々の研究で本当に多くの貴重なアドバイスを頂きました。壁にぶつかったときや間違った方向へ進んでしまいそうになったときなど、的確なご指導をしてくださり、心より感謝しています。本論文のデータの提供や、手法へのアドバイスなどを頂くことができたおかげで、本論文を仕上げることができました。厚くお礼申し上げます。

薄井智貴特任助教には、特に、修士課程初期から中期にかけて大変お世話になりました。その際、他のプロジェクトで貴重なご指導をして頂けたことが、本論文の執筆につながりました。論文執筆方法や発表資料作成方法などのご指導をしてくださり、大変勉強になりました。厚くお礼申しあげます。

研究室のメンバーの方々には、研究室での生活において、何かとお世話になりました。いろいろと助けて頂き、実りある研究室生活を送ることができました。本当にありがとうございました。

研究室秘書の秋枝久美子氏、高野かおる氏、本間理恵子氏、川端晴子氏には、海外学会参加や各種申請や登録などの複雑な手続きに対し逐一的確な指示を頂きました。自身の不用意さからお手数をかけることも多く、大変お世話になりました。厚くお礼申しあげます。

ここに挙げた方々以外にも、研究室内外を問わず、また研究以外での活動においても、多くの方々の力添えによって論文を完成させることができました。この場をお借りして、謹んでお礼申し上げます。

最後に、物心両面において支えてくれた両親に最大限の感謝を致します。