

ネットワークシミュレーション用のベンチマークデータセットの構築～車両経路の抽出

(株)熊谷組 エンジニアリング部 正会員 堀口 良太  
 千葉工業大学 正会員 赤羽 弘和  
 東京大学 生産技術研究所 正会員 吉井 稔雄  
 千葉工業大学 大学院 学生会員 花房比佐友  
 大日本コンサルタント 山口 智浩

1. はじめに

近年、交通シミュレーションモデルに関する研究開発が盛んに行われているが、その問題点として実際のデータを使ったモデル検証が十分に行われていないことが指摘されている[1][2]。このため、それらのモデルの精度や適用可能性などについて共通の認識を形成することができず、性能比較や評価の妨げとなっている。一方では、シミュレーションの共通プラットフォームに関する議論が高まりつつあり、目的や問題意識が異なる研究者が1つの共通モデルを使うというよりはむしろ、再現精度や演算速度などの具体的な基準を設け、モデルを評定すべきであるという声もあがっている。このため、モデル開発者が共有することができる検証用データセットの整備が急務とされる。

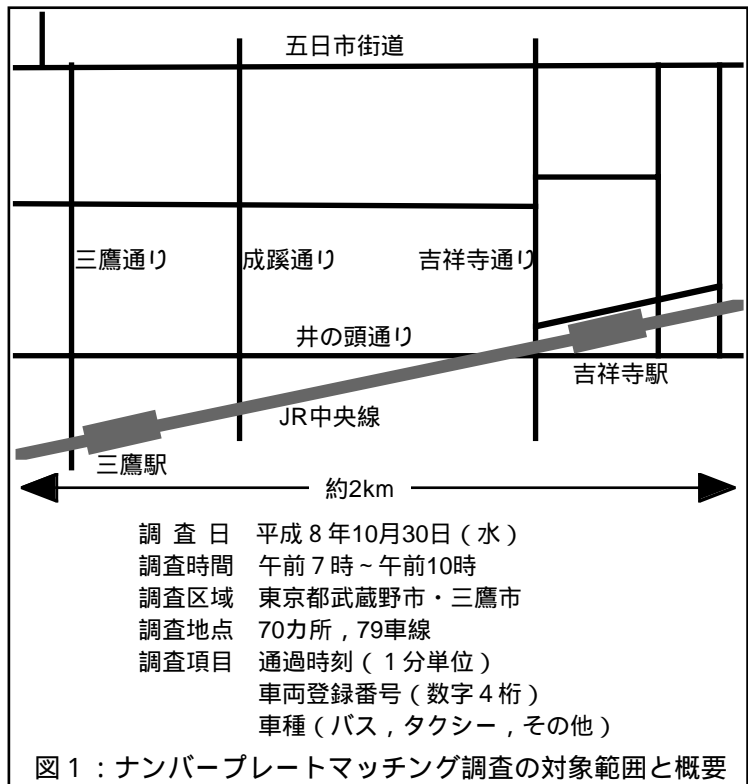


図1：ナンバープレートマッチング調査の対象範囲と概要

このような観点から、本研究ではネットワークシミュレーション用のベンチマークとなるデータセットを構築することを目的としている。本稿ではとくに、図1に示すエリアにおいて行ったナンバープレートマッチング調査[3]の結果を対象として、車両の経路を抽出する過程について述べる。

2. ロバストな経路抽出アルゴリズム

ナンバープレートマッチング調査の結果から車両経路の軌跡を抽出するには、隣接する観測地点間のデータのあいだで車種と車両番号が一致するものをつなげていけばよいが、観測ミスによる軌跡の不連続や、駐停車によるトリップの中断、同一ナンバーの車両軌跡の交錯などを含むため、これらを排除できるロバストなアルゴリズムを構築する必要がある。以下に、本研究で採用したアルゴリズムの概要を示す。

1) データ要素をノードとする時空間グラフの定義

経路抽出の探索空間として、通過時刻、車種、車両番号で属性づけられるデータ要素を頂点集合とするグラフを考える。各データ要素は観測地点と対応する。グラフの枝の初期集合として、隣接する観測地点それぞれの各データ要素について、車種および車両番号が一致し、かつ通過時刻の差が閾値以下のものを接続する枝(マッチング枝)を用意する。閾値は観測地点間の一般的と思われる旅行時間よりも十分長い値(今回は10分)にする。

2) 観測ミスによる不連続な経路軌跡の接続

キーワード：シミュレーション, ナンバープレートマッチング, OD交通量, 経路推定, ベンチマークデータ, AVENUE  
 連絡先：東京大学生産技術研究所, 〒106 東京都港区六本木7-22-1, 電話 03-3402-6231, Fax 03-3401-6286

観測時の車両番号の読み落としや間違いによって、隣接観測地点間でマッチングがとれず、その結果、本来は1本である経路の軌跡が複数の経路軌跡に分断されることが考えられる。これについては、1台の車両が連続した2地点で同時に誤って読みとられる確率は小さいと考え、隣接地点でマッチングされなかったデータ要素については、1つ先で隣接する2地点間で、同様のマッチングを行う。

### 3) ループ・ショートカットを形成するマッチング枝の除去

観測地点の接続関係によっては、本来1本のチェーンとして抽出されるべき車両経路軌跡が途中で分岐・合流して、ループやショートカットを形成する場合がある。グラフの構造からこれらを形成するマッチング枝を同定し、除去する。

### 4) 競合するマッチング枝と無効なマッチング枝の除去

本来入出力のマッチング枝がそれぞれ1以下であるデータ要素に対して、実際はグラフが単純なチェーンにならず、複数のマッチング枝が1つのデータ要素にたいして入出力し、競合する場合がある。また競合はしていないが、途中で駐停車をしたため、旅行時間が極端に大きいマッチング枝も発生する。これらの不正なマッチング枝に対しては、車両番号の読み違いによる偶然のマッチングではないかを調べ、そうでないものについては、観測地点間の所要時間分布に指数正規分布を仮定し、推定パラメータを用いてえられる各マッチング枝の出現確率が小さいものを除去する。その際、誤ったデータを含むものに対して分布パラメータを推定するので、トリム法などのロバストな手法[4]を用いている。

## 3. アルゴリズムの有効性

上記のアルゴリズムを適用した結果を各段階ごとに表1に示す。抽出された経路軌跡数は1)から2)で一旦減少するが、これは1)では分断されていた経路軌跡がつながったためである。その後は段階を追って経路数が増加する。同時に経路の平均旅行時間、平均通過地点数も増加しており、このアルゴリズムが単純に隣接地点でマッチングをとるよりも効率的に車両の経路軌跡を抽出できることがわかる。

表1：各段階での経路抽出の結果

	1)	2)	3)	4)
観測データ要素数	98930	-	-	-
ナンバー読みとり不可	757	-	-	-
経路を構成するデータ要素数	79734	81487	82314	89703
”の割合	81.2%	83.0%	83.8%	91.4%
車両経路軌跡総数	20231	19395	19537	21107
平均旅行時間[分]	2.93	3.39	3.41	3.48
平均通過地点数	3.92	4.2	4.21	4.25

## 4. 今後の展開

今後は抽出された経路と旅行時間などのデータをもとに、この地域でのOD交通量および経路選択行動モデルを同定していく。さらにこれらのデータに、筆者らが開発している交通シミュレーションモデル AVENUE [5] を適用し、モデルの再現能力を検証する。一連の調査およびデータ処理でえられた結果は、シミュレーション用ベンチマークデータとしてインターネット上で公開される予定であるが、AVENUE の検証結果もあわせて公開する。

### 【参考文献】

- [1] 吉井稔雄, 桑原雅夫: ネットワークシミュレーションモデルの適用と評価, 土木計画学研究講演集19(1), p.673, 1996
- [2] 松本 幸正: 交通ネットワークシミュレーションモデルの検証について, 土木計画学研究講演集19(1), p.674, 1996
- [3] 花房比佐友, 山口智浩, 赤羽弘和, 吉井稔雄: 交通シミュレーションシステムの再現性検証用データセットの構築, 第52回年次学術講演会論文集掲載予定, 1997
- [4] UP 応用数学選書「最小二乗法による実験データ解析」中川徹, 小柳義夫著, 東京大学出版会, 1983
- [5] R. Horiguchi, et al.: A Network Simulation Model for Impact Studies of Traffic Management 'AVENUE Ver. 2', The 3rd World Congress on Intelligent Transport Systems, Orlando, 1996