

# ビッグデータを活用した車両行動推定方式

染谷 一輝†1 清原良三†2 齋藤 正史†3

## 1. はじめに

地方中核都市の一つである石川県金沢市では様々な問題がある。特に問題となっているのが自動車渋滞である。発展した都市であれば公共交通機関を利用することで交通渋滞を避けることができる。しかし地方では公共交通機関が発達している訳ではなく、財政的に莫大な費用をかける施作はむやみに実行できない。道路整備の効果を明確にするためにはシミュレータを用いるが現実と同様の環境をシミュレータ内に実装することは難しい。この問題を解決するため本論文では OD(Origin-Destination)交通量に着目した。

OD交通量を把握する従来手法は存在するが、大規模なアンケートの実施や精度は良いがとても高価なデータであったりと高コストである。本論文ではこの高コストな OD 交通量のデータをモバイル空間統計と道路交通センサスの2つのビッグデータを組み合わせることで、より楽に、より廉価に、低コストで OD 交通量の推定を目指した。

## 2. 提案手法

本論文では実データを用いるのではなく、シミュレータにビッグデータと正解データとなる OD 交通量を実装し、そのログデータで実験する。

ログデータを用いる理由として、OD交通量のデータ取得は難しく、廉価で正確なデータがないため評価ができないこと。現状取得できるビッグデータを用いて分析してもその時の状況しか分からないこと。これらの理由からログデータを利用する。

評価方法はログデータをそれぞれ比較するために一番利用しやすい形であるモバイル空間統計のメッシュ表現を用いる。ビッグデータが1時間ごとにデータを集計していることから、ログデータも同様に1時間ごとにデータを分割することで二次元表現に変形する。二次元表現を  $x$  と  $y$  軸がエリア、 $z$  軸が人数となる三次元表現に変形することで比較し、正当性を確認する。

OD交通量、モバイル空間統計、道路交通センサスをとあるエリアから次どのエリアに移動するかのデータ(以降、次移動エリア表現と呼称する)に変形することで比較が可能となり、モバイル空間統計と道路交通センサスから OD 交通量の作成ができるのではないかと提案する。

## 3. シミュレーション実験

シミュレータは Space-Time Engineering 社のネットワークシミュレータ Scenarigel[1]のマルチエージェントモジュールを用いた。実験対象である石川県金沢市の地図データは OpenStreetMap より利用し、道路交通センサスの計測箇所を適応するために Java Open Street Map Editor を使用して地図データ編集した。モバイル空間統計のデータを調べると、金沢駅周辺のメッシュは 2km メッシュの

ため、全域を把握するために一辺 2km となる 9×9 のメッシュを作成した。

シミュレータの設定は合計人数は 10000 人、20000 人、30000 人の 3 パターン。乗車率は 40%、50%、60% の 3 パターン。最大乗車率は 1 人。自動車速度は 60km . 人の出現間隔は指数分布。実験時間は 21,600 秒とした。

合計人数とはシミュレーションを行なった時に出現する自動車の合計数を表す。乗車率とは自動車に乗る人の割合を表す。乗車しない場合は歩行者として扱う。そのため実験では 3 パターンの人数と 3 パターンの乗車率、計 9 パターンに分けて実験をした。乗車率とは自動車に乗る割合を表す。最大乗車人数とは 1 台の自動車に乗れる人数を表す。自動車の出現間隔とはどの程度の頻度で自動車が出現するかを表す。実験時間はシミュレーション内で経過する時間を表す。6 時間=21,600 秒を実験時間とした。

自動車は出現地点と目的地点、共にランダムとしており、シミュレータ内に存在する建物から建物までを交通量の多い道路を主要として最短距離で移動する。自動車の出現間隔は指数分布であり、目的地に到着した場合は消えるため、総数は常に一定ではない。

## 4. 評価

### 4.1 データ変形

シミュレータを実行して得たログデータを評価するために 2 章提案手法の通りにデータを変形した。二次元表現を表 1、三次元表現を図 1、次移動エリア表現を図 2 に示す。図 1 は人数 30000 人、乗車率 60%、時間帯 3600 秒の時の道路交通センサス。図 2 は人数 30000 人、乗車率 60%、エリア 440 の時の OD 交通量である。移動元となるエリアから他のエリアに移動した人数を表現しているため図 2 のエリア 440 は 0 人となっている。

### 4.2 ビッグデータの信頼性

本論文のデータはログデータのため、そのデータの信頼性がない。そのためモバイル空間統計と道路交通センサスを実データとログデータ、それぞれで回帰式を導出し、比較を行った。実データは研究グループ[2]より使用した。決定係数  $R^2$  は 0.58、回帰式は以下となっている。

$$y = -1.0 \times 10^{-6}x^2 + 0.69x + 1029.7$$

ログデータは表 1 の二次元表現を用いて回帰式を導出した。決定係数  $R^2$  は表 2、回帰式は表 3 に示す。決定係数  $R^2$  の平均値は 0.83 となった。この結果よりログデータは実データと比較した場合、同じとは言えないもののそれなりに似ていると言える。

ログデータは人数と乗車率を変動した 9 パターン実験したが、どのパターンでも似たような結果となり、大きな変化はなかった。

### 4.3 組み合わせ

†1 神奈川工科大学大学院  
Graduate School of Kanagawa Institute of Technology

†2 神奈川工科大学  
Kanagawa Institute of Technology

†3 金沢工業大学  
Kanazawa Institute of Technology

表 1 二次元表現

時間	エリア	人数
3600	000	12
3600	010	34
}		
21600	870	33
21600	880	20

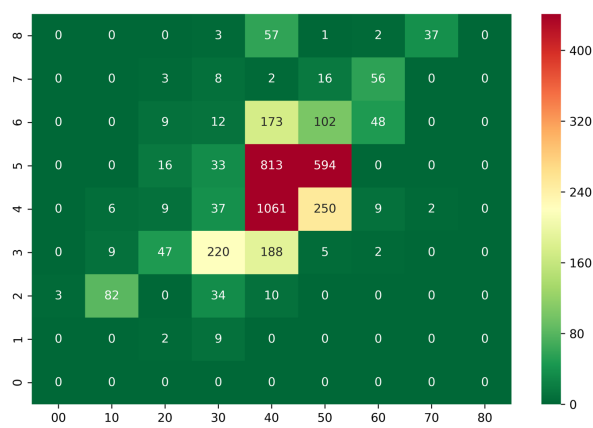


図 1 三次元表現

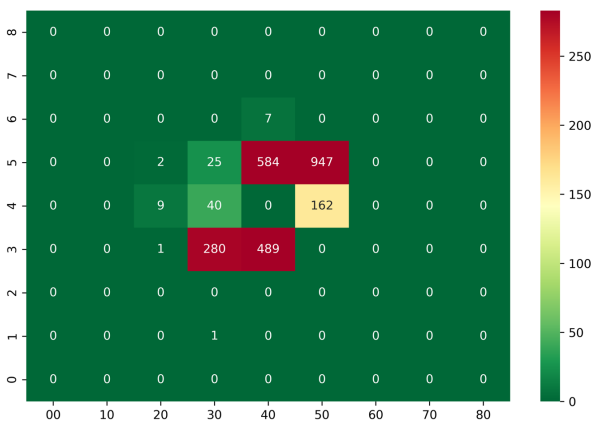


図 2 次移動エリア表現

表 2 ログデータの決定係数 R<sup>2</sup>

乗車率 \ 人数	10000	20000	30000
40	0.82	0.83	0.82
50	0.82	0.82	0.83
60	0.83	0.82	0.83

表 3 ログデータの回帰式

乗車率 \ 人数	10000	20000	30000
40	$y = -9.60 \times 10^{-3}x^2 + 3.25x + 0.87$	$y = -5.37 \times 10^{-3}x^2 + 3.44x + 0.12$	$y = -2.98 \times 10^{-3}x^2 + 3.26x + 2.10$
50	$y = -8.38 \times 10^{-3}x^2 + 3.36x + 0.77$	$y = -4.51 \times 10^{-3}x^2 + 3.48x + 1.16$	$y = -2.73 \times 10^{-3}x^2 + 3.42x + 1.69$
60	$y = -6.24 \times 10^{-3}x^2 + 3.23x + 0.12$	$y = -3.06 \times 10^{-3}x^2 + 3.22x + 2.25$	$y = -1.85 \times 10^{-3}x^2 + 3.25x + 3.04$

OD 交通量を作成するために図2の次移動エリア表現を用いた。結果としてビッグデータに id のような個々を判別する情報がなため次移動エリア表現に変形することができなかった。しかし研究グループの最新成果によると道路交通センサに擬似 id を付与することができた。そのため道路交通センサに id があつたと仮定して比較をした。

その結果、各エリアの決定係数の平均値は 0.48、中央値は 0.49 となり道路交通センサだけでは OD 交通量の予測が困難であることが分かった。平均値と中央値が低い理由として道路交通センサは元となるエリアでも自動車台数が計測していることから外れ値が大きく作用していると考えられる。

## 5. まとめ

本論文ではログデータを変形することでモバイル空間統計、道路交通センサ、OD 交通量の比較をした。ログ

データの信頼性は実データと比べ同じではないがそれなりに似ている結果となった。

しかし目的であるビッグデータから OD 交通量を推定することは達成できなかった。仮に道路交通センサから id のような関連性分かるならば、どのような経路で移動するのか傾向をつかむことができた。

## 参考文献

- [1] Mineo Takai, Jay Martin, Shigeru Kaneda, Taka Maeno: 「Scenargie as a Network Simulator and Beyond」 Journal of Information Processing Vol.27, No.1, pp.2-9
- [2] 齋藤 正史, 波多野 太樹, 清原 良三: 「空間統計データを用いた車両走行推定方式の一提案」 歩歩処理学会 ITS 研究会, 2018-ITS-073