

修士研究の方向性

紹介論文導入

予測の概要

予測手法

まとめと今後の
課題

【論文紹介】

マイクロジオデータと機械学習手法を用いた メッシュ間 OD 交通量の推定

武藤 克弥 (Katsuya Muto)
u255018@st.pu-toyama.ac.jp

富山県立大学大学院 電子・情報工学専攻 情報基盤工学部門

May 24, 2022

OD 交通量

- 出発点 (Origin) から到着点 (Destination) までの人・車の流動を計測したもの
- Origin-Destination Matrix(OD 行列) で表現→図 1
- 交通に限らず、物や情報の流入・流出量を行列にすることもできる

	着地		
発地	アジア	米国	欧州
アジア	100	120	50
米国	80	200	50
欧州	20	50	230

図 1: OD 行列の例

OD 行列推定

- 地域活性化や都市の整備や交通計画において、ヒト・モノの流入流出を把握することが重要→ OD 行列を推定することで把握
- 機械学習を使って推定

研究背景

- 人口減少，中心街活性，環境の観点からコンパクト・プラス・ネットワークの形成が重要視
- 都市部施設の配置方法，それによる交通量の影響などの評価が必要

目的

- OD 交通量推定モデルを構築し，交通整備・施設の誘導が OD 交通量に与える影響を予測する。

建物用途別延べ床面積，滞在人口，移動時間から OD 交通量を推定
→その際に用いるデータ

都市計画基礎調査データ

- 都市計画を行うための人口，産業，土地利用，建築物，交通等の調査データ

→「小分類」の建物ごとの延べ床面積の平均値等を用いる

大分類	小分類
商業	業務施設，集合販売施設，専用店舗施設，宿泊施設，興業施設，風俗営業施設，遊戯施設，スポーツ施設，地方国家施設，自治体施設
住宅	専用住宅，共同住宅，一般店舗併用住宅，飲食店併用住宅，作業所併用住宅，事務所併用住宅
文教 厚生	教育施設，研究施設，文化施設，宗教施設，記念施設，医療施設，運動施設，社会福祉施設，厚生施設
工業	重化学工業施設，軽工業施設，サービス工業施設，家内工業施設，供給処理施設，運輸倉庫施設，通信施設
その他	漁業施設，農業施設

Table 1: 都市計画基礎調査データの建物区分

使用するデータ 2

5/12

モバイル空間統計

- Docomo ユーザの携帯から収集された位置情報を利用
- ある時刻, 特定エリア (メッシュ) における, 性別, 年代の分布などを取得可能

→ 7時から9時の滞在人口・流動人口 (OD 交通量) を使用

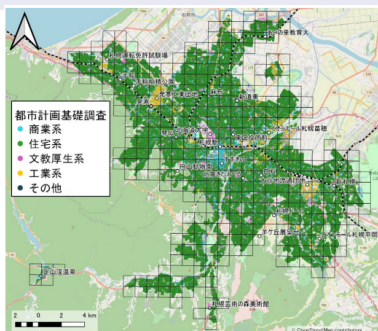


図 2: 各メッシュごとの建物 (表 1 の大分類) の色分け分布

	滞在人口	OD 交通量
日付	2020 年 1月 20~24 日	2019 年 10月 17 日
曜日	平日	木曜日
時刻	7, 8, 9 時	7~9 時台
集計規模	500m メッシュ	1 km メッシュ
集計方法	各時間の平均値	総和
単位	人	trip

図 3: モバイル空間統計データ詳細

推定移動時間

- メッシュの中心点から中心点への移動を仮定
- 8時30分までの到着を条件
- 公共交通機関、自動車移動の2通りの平均移動時間などを計算

OD 交通量は半径 1km 以上、移動先で 1 時間滞在で 1 トリップ (1 人分の移動) とカウント

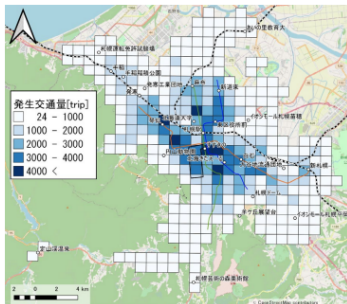


Table 2: 発生交通量の分布

用いるデータセット

- ～施設：369 メッシュに対する施設の延べ床面積の代表値
- 滞在人口～時：369 メッシュ間の代表値
- ～移動時間：4863 通りの OD(出発地-到着地) ペアに対する代表値

	業務施設	集合販売施設	専用品舗施設	宿泊施設	興業施設	風俗営業施設	遊技施設	スポーツ施設	地方国家施設	自治体施設
平均値	20131.7	4731.9	13297.3	5435.5	167.5	606.2	2777.8	513.8	947.0	1627.6
標準偏差	85137.1	26847.7	30181.8	32726.3	2273.0	10942.3	7727.5	2443.2	7395.6	9278.7
中央値	4245.0	0.0	5202.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	専用住宅	共同住宅	一般店舗併用住宅	飲食店併用住宅	事務所併用住宅	作業所併用住宅	教育施設	研究施設	文化施設	宗教施設
平均値	10535.6	128435.1	3301.9	216.8	2806.7	272.7	12811.8	370.3	1657.6	1572.3
標準偏差	70806.1	172564.4	4048.8	347.8	3087.6	404.5	29798.8	2370.7	7428.3	3071.7
中央値	101693.0	52125.0	1958.0	97.0	1754.0	69.0	7757.0	0.0	435.0	365.0
	記念施設	医療施設	運動施設	社会福祉施設	厚生施設	重化学工業施設	軽工業施設	サービス工業施設	家内工業施設	供給処理施設
平均値	19.6	7718.2	563.5	4927.1	153.3	28.9	4453.5	1557.4	406.1	1829.6
標準偏差	267.8	14467.9	2410.0	5873.2	1068.7	182	15257.2	3964.2	1146.9	8247.2
中央値	0.0	2465.0	0.0	2727.0	0.0	0.0	49.0	0.0	0.0	354.0
	運輸倉庫施設	通信施設	漁業施設	農業施設	滞在人口7時	滞在人口8時	滞在人口9時	公共交通移動時間	自動車移動時間	OD交通量
平均値	9722.2	1283.5	1.0	138.4	3793.1	3779.9	3759.7	22.6	11.5	74.3
標準偏差	24902.9	6765.7	19.0	302.7	3328.3	3847.8	4550.9	11.8	8.5	79.5
中央値	1173.0	0.0	0.0	0.0	2980.8	2797.8	2776.8	20.0	8.8	49.0

Table 3: 各説明変数とその代表値

予測の流れ

- 説明変数: 76 種類 (表 4), 目的変数: OD 交通量 (表 4 右下)
- OD ペア 4863 通其中 90%学習データ, 10%検証データとして 10 分割交差検証

① ランダムフォレストで変数選択

→ OD 交通量に対する各延べ床面積の寄与度から評価

② グリッドサーチで決定木数, 最大深さを設定

③ 寄与度の高い変数を用い, NN で OD 交通量を予測

変数	変数名	種類	単位
説明変数	出発地 建物用途別延床面積	34	m ²
	到着地 建物用途別延床面積	34	m ²
	出発地 滞在人口 (7,8,9時)	3	人
	到着地 滞在人口 (7,8,9時)	3	人
	自動車移動時間	1	分
	公共交通移動時間	1	分
目的変数	メッシュ間 OD 交通量	1	tip

Table 4: 説明変数と目的変数の概要

NN の構成

- 入力・中間・出力層と各層のユニット (ニューロン) から構成
- 非線形データへの当てはまり，汎化性能ともにバランスが良い

$$u = \theta + \sum_{i=1}^K w_i x_i \quad (1)$$

u : 次の層への入力, θ : バイアス, K : ある層のユニット数, x_i : ユニット i への入力, w_i : ユニット i に対する重み

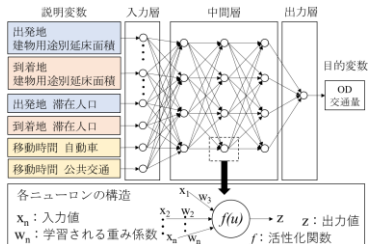


図 6: NN(ニューラルネットワーク) の構成

NN のパラメータ決定

- (入力層→中間層→中間層) 活性化関数を Relu 関数 ($y = x(x > 0)$)
- (中間層→出力層) 恒等関数 $y = x$
- 中間層は 3 層で固定
- 平均二乗誤差が最小となるようバイアス, 重みを決定

$$u = \theta + \sum_{i=1}^K w_i x_i \quad (2)$$

u : 次の層への入力, θ : バイアス, K : ある層のユニット数, x_i : ユニット i への入力, w_i : ユニット i に対する重み

分析結果

- ランダムフォレスト：寄与度が高い上位 17 種の説明変数を採用
- 予測値に対する実測値との関係で決定係数 $R^2 = 0.81$ で強い相関

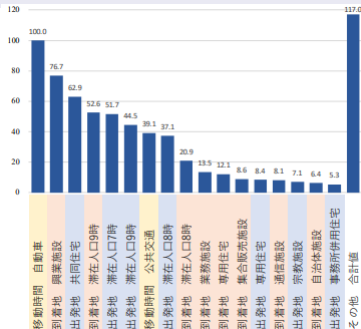


図 7: ランダムフォレストによる変数の寄与度

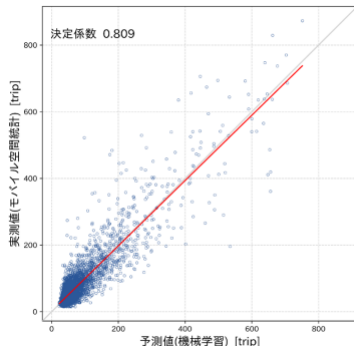


図 8: 予測値と実測値の関係

【紹介論文】まとめと今後の課題

- 滞在人口，メッシュ間時間関係，建物の延べ床面積が OD 交通量に強い影響を与えていることが分かった.
- ランダムフォレストで棄却した変数が大幅に変化すると，正確な OD 交通量予測できない可能性があり，変数選択と起こりうる誤差をどのくらい許容するかをさらに検討していく
- 実際の交通流データに基づく移動データを用いたモデル構築をしていく

修士研究の今後の予定

- 卒業研究や今回の紹介論文のような特徴量作成，モデル作成，予測の流れを OD 交通や経路ネットワークに適用していきたい