

都市内交通の最適構成の把握に資する地域分類手法の検討

長谷川大輔・巖先鏞

A Study of Area Segmentation Methods for Understanding Optimal Composition of Urban Traffic Patterns

Daisuke HASEGAWA and Sunyong EOM

Abstract: This paper aims to understand the characteristics of short-distance trips covered by short-distance transportation. First, the amount of short trips in grid units and the cosine similarity based on the azimuth are summarized. Then, area segmentation is performed by several methods using static data, and a method with high accuracy of regression results is examined for each segment.

As a result, clustering by VBGMM using Point of Interest data and population data obtain a more precise the result of regression.

Keywords: 都市内交通パターン(urban traffic patterns), 公共交通(transportation), 地域分類(area segmentation)

1. はじめに

近年 MaaS(Mobility as a Service)の概念が広まり、多様な交通手段を IT サービスによってシームレスに利用する動きがフィンランドの Whim を先駆けに世界中で検討が進んでいる。(MaaS_Global, 2018) 特に短距離交通においては既存の公共交通手段に囚われない相乗り交通やシェアサイクル、キックスケーターといった交通手段の多様化が進んでいる一方で、企業による過剰な投資により無秩序なシェアバイク利用等の弊害も生まれている(駒形, 2017)。こうした交通手段にはそれぞれ道路・地形などの地理的特性や交通需要の密度・分布といった交通需要の特性によって相対的な優位性が異なり、地域によって最

適な交通手段の構成が存在する(長谷川・鈴木, 2017)。そのため、その都市における都市内交通手段の最適構成を把握することは、投資の過不足を把握し、持続可能かつ都市のモビリティを改善する交通体系を構築するために重要な観点と言える。そこで本研究では都市内交通の最適構成に影響を与える交通需要の特性を把握に着目し、起終点・時間属性のあるトリップの密度・類似度といった特徴量を評価する手法を提案する。また、交通データがない場合を仮定し、その特徴量を建物・人口データ等の静的データを用いた分類・回帰分析によるトリップ特徴量の推計を行い、静的データによる潜在的なトリップ量・類似度を把握することを目的とする。

長谷川大輔 〒150-8510 東京都渋谷区渋谷 2-21-1

所属 株式会社ディー・エヌ・エー オートモーティブ事業部 モビリティインテリジェンス開発部

E-mail: daisuke.a.hasegawa@dena.jp

2. 分析手法

図 1 に本研究の分析フローを示す。1/4 地域メッシュ (250m メッシュ) 単位で建物や建物内の

テナントデータから構成される POI (Point of Interest) データ, 人口・年間鉄道駅利用者からなる統計データ, 起終点・時間帯別トリップデータの集計を行い, メッシュ別統計データ, トリップ特徴データを作成する. 次にメッシュ別統計データを用いた地域のクラスタリングを行う. そして, クラスタごとに時間帯別のトリップ量・類似度を推計する回帰モデルを構築し, 時空間別トリップ特徴量を推計する.

本研究では推定精度の向上が見られるクラスタリング手法, トリップ特徴量に影響を与える要素を把握する地域分類方法の検討に着目し分析を行った.

2.1. メッシュ単位データ集計

図2にデータ集計のイメージを示す. 集計対象のメッシュを起点とする距離範囲 r_1, r_2 の範囲内のデータを抽出する. 起点・終点位置, 時間属性のあるトリップデータを対象に, 集計対象とするメッシュより一定の距離範囲・時間範囲でトリップデータをフィルタリングした後, 起点が含まれる地域メッシュと, トリップの発生時間を6時間帯単位で区切った上でトリップ総数と類似度を集計する. 類似度については, Moran' s I等の点群間の距離を用いた空間集積を評価する手法ではバスや LRT といった路線型の交通手段において重要となる線集積の評価が難しい. そこで本

研究ではトリップの方位角の \cos 値の標準偏差を1から引いた値とし, 以下の式で求める.

$$sim_{ab} = 1 - \sqrt{\frac{1}{n_{ab}} \sum_{i=0}^{n_{ab}} (\cos \theta_{t_{ab_i}} - \overline{\cos \theta_{t_{ab}}})^2} \quad (1)$$

ただし,

sim_{ab} : メッシュ a の時間帯 b におけるトリップ類似度

n_{ab} : メッシュ a の時間帯 b におけるトリップ数

$\cos \theta_{t_{ab_i}}$: メッシュ a の時間帯 b のトリップ t_{ab_i} の方位角 θ の \cos 値

$\overline{\cos \theta_{t_{ab}}}$: メッシュ a の時間帯 b のトリップの \cos 平均値

である. トリップが 360° 一様に存在する時は0をとり, 1に近づくほどトリップの類似度が高くなり, 集約されたトリップが発生している.

POI データや統計データについても 250m メッシュ単位での集計を行う. 起点となるメッシュから一定の距離範囲別にその建物数, 建物テナント数, 人口, 世帯数を合計する. また, トリップの類似度に影響を与えと考え, 含まれるメッシュ間での標準偏差を求める.

本研究では $r_1 = 2500m, r_2 = 5000m$ として, 起点メッシュ別に距離範囲別トリップ量, 類似度, POI, 統計データを集計する.

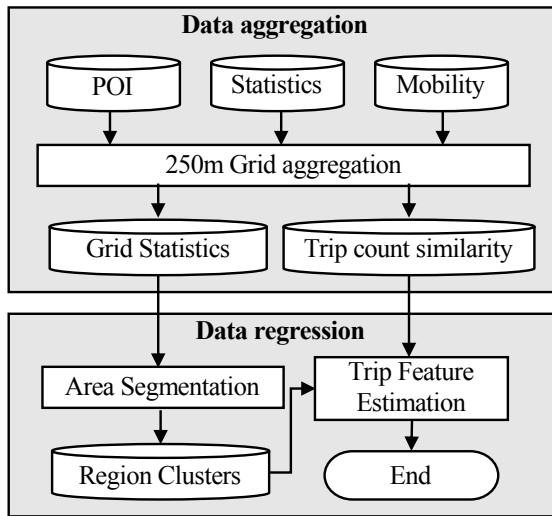


図1 分析フロー

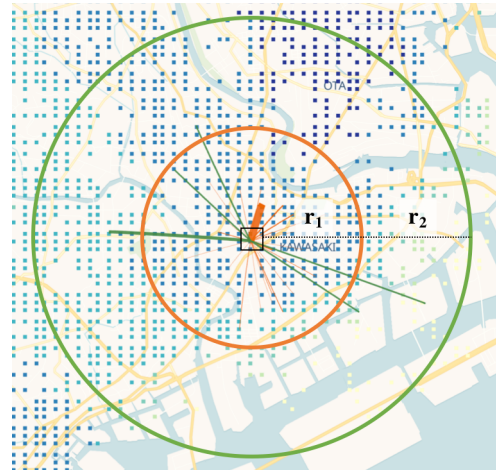


図2 データ集計イメージ

2.2. 地域分類およびトリップ特徴量推定

時空間別トリップ量および類似度をメッシュ別統計データにより推計する．その際の推計精度向上および地域別の交通パターンを把握するために統計データを用いた地域の分類を行う．

地域分類の手法として，クラスタリング手法として X-means クラスタリング (Pelleg and Dan, 2000), 変分混合ガウス分布 (Variational Bayesian Gaussian Mixture Model, VBGMM) クラスタリング (Nasios and Adrian, 2006)を用いる．

どちらもクラスター数を自動決定する手法であるが，X-means クラスタリングは一般的な K-means 法を逐次繰り返していき，クラスター内の尤度関数より求めた BIC (ベイズ情報量基準) の数値が最大値をとるクラスター数が最適とする方法である．

VBGMM クラスタリングは，クラスター毎の正規分布に従う確率分布によって割当を行うガウス分布クラスタリング (Gaussian Mixture, GMM) に変分ベイズ法を組合せた方式で，確率分布自体をベイズ推定によって事前分布から事後分布へと更新するアルゴリズムである．

そして，上記の手法で分類したクラスター別に独立した，距離範囲・時間帯別トリップ量，類似度を非説明変数とした重回帰分析モデルを構築

表 1 分析に用いるデータ

種類	名称	年度	取得元
交通データ	トリップデータ	2008	H20年東京都市圏人の流れデータ
POIデータ	建物数	2018	ゼンリン汎用
	テナント数		デジタルフォーマットデータ
	駅出入口		
統計データ	1/4地域メッシュ人口	2015	国勢調査
	1/4地域メッシュ世帯		
	年間駅利用者数	2017	国土数値情報 駅別乗降客数

表 2 距離範囲別データ集計結果

		$r_1:0m\sim2500m$		$r_2:2500m\sim5000m$	
		Avg	Std	Avg	Std
交通	トリップ数(trip)	359.89	449.28	71.33	78.74
	トリップ類似度	0.38	0.16	0.29	0.16
POI	建物数(棟)	14769.8	10065.8	14233.5	9660.3
	テナント数(件)	29196.1	31105.4	26529.3	24866.7
統計	人口(人)	245100.1	118726.0	238770.6	123213.5
	世帯(世帯)	120235.5	69658.1	116143.1	69966.9
	駅別乗降客数(人)	61497.0	88395.1		

する．説明変数は集計した距離範囲別のメッシュ集計処理を行った POI, 統計データの合計，およびメッシュ間の標準偏差を用いる

3. 分析結果

3.1. メッシュ単位データ集計結果

表 1 に集計対象とするデータとその取得元を，距離範囲別のデータ集計結果を表 2 に示す．本研究では交通データとして東京大学空間情報科学研究センター(CSIS)が提供している H20 年東京都市圏人の流れデータのサブトリップ（出発地から乗換地点までの移動を 1 トリップと計上しているデータ）を用いる．このうち，集計の効率化のために 1 次メッシュ番号 5339 のエリアを起点とし，起終点間の直線距離が距離範囲 r_1, r_2 に含まれ，同一メッシュ内で完結しない 989,237 トリップのみ抽出した．また，集計対象とするメッシュも出発トリップが存在する 2,294 メッシュのみを対象とした．図 3 は距離範囲 r_1 ，6 時～12 時台の起点メッシュ別トリップ数と類似度の集計結果である．これを見ると，駅のあるメッシュはトリップ量が多く，特に新宿や渋谷，東京といったターミナル駅では乗換トリップの空間的集中が見られる．また，最寄り駅へ向かうトリップが多い地点等ではトリップ類似度が高く，バスや相乗りタクシー等の相対的優位性が向上することが考えられる．

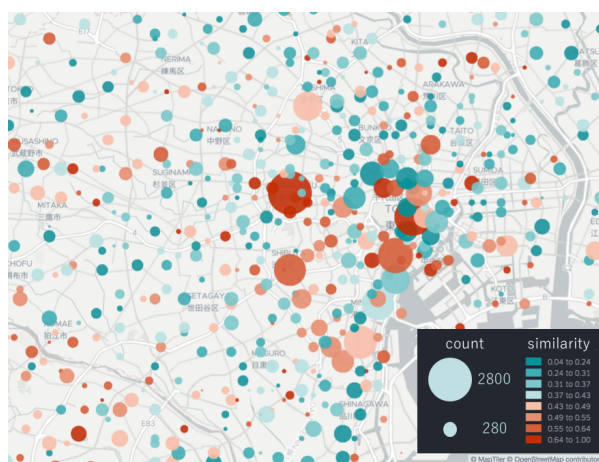


図 3 起点メッシュ別トリップ数と類似度 ($r_1:0m\sim2500m$, 6 時～12 時)

また、POI データについては（株）ゼンリン汎用デジタルフォーマットデータに含まれる建物代表点データ、および各建物内のテナントデータを用いる。また、図3の可視化によって鉄道駅とそれ以外でトリップの特性が異なることが見られたため、同データに含まれる駅の出入口があるメッシュ (mesh_station_entrance) とそれ以外 (mesh_no_station) の二群に分けた上で地域分類と回帰分析を行う。

3.2. 地域分類およびトリップ特徴量推定

表3にクラスタリングなし、X-means、VBGMM による距離範囲別・駅の出入口で分けたメッシュ群別にトリップ量と類似度を推計した際の各時間帯の決定係数の調和平均したもの示す。クラスタリングによる事前処理によって決定係数が2倍弱に増加し、VBGMM がより精度の高い結果を示した。その結果によって分類した結果とその誤差率を図4に示す。都心部から放射環状に広がるクラスタに分類され、外縁部では誤差率の低い地点が多く見られているものの、トリップ数が極端に多い都心部においては、非線形の回帰手法や、建物数だけでなくテナントの種類などによる建物

の分類を行うことで精度の向上が期待できる。

4. おわりに

本研究ではトリップデータの距離範囲別の特徴量を提案し、POI・統計データを用いた特徴量の推計を行った。その結果、都市の外縁部において一定の精度を確認できる分類方法を提案した。道路指標や POI の種別の考慮、回帰手法の検討が今後の課題となる。なお、本研究は東京大学 CSIS 共同研究 No.915 の成果の一部である

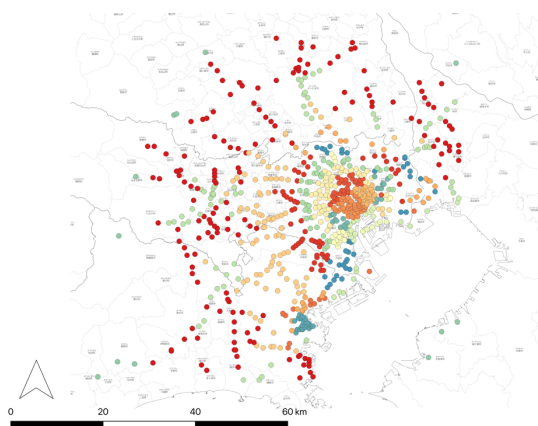
参考文献

- MaaS_Global. (2018): MaaS as a Concept.
<https://maas.global/maas-as-a-concept/>, (accessed: 12/30/2018).
 駒形哲哉(2017)：中国におけるシェアバイクの爆発的普及 -- 「大量生産・大量消費」的シェアリング・エコノミー、アジ研ワールド・トレンド, 9-13, 日本貿易振興機構.
 長谷川大輔, 鈴木勉(2017)：需要密度・移動距離に着目した多様な公共交通システムの優位性に関する理論的考察, 都市計画論文集, 52(3), 1284-1289.
 Pelleg, Dan, and Andrew W. Moore.(2000) X-means: Extending k-means with efficient estimation of the number of clusters, Icm1, 1, 727-734.
 Nasios, Nikolaos, and Adrian G. Bors.(2006) Variational learning for Gaussian mixture models, IEEE Transactions on Systems -Man, and Cybernetics Part-B, 36(4), 849-862.

表3 地域分類手法別決定係数

	distance range	mesh_station_entrance (N=781)			mesh_no_station (N=1513)		
		no_cluster	xmeans	VBGMM	no_cluster	xmeans	VBGMM
trip_count	r ₁	0.3915	0.6502	0.6611	0.2110	0.2166	0.2517
	r ₂	0.3355	0.6345	0.6399	0.0750	0.2284	0.2692
cos_simirarity	r ₁	0.0890	0.3801	0.4098	0.0736	0.2196	0.2501
	r ₂	0.0704	0.4013	0.4646	0.0681	0.2010	0.2378

分類結果



誤差率

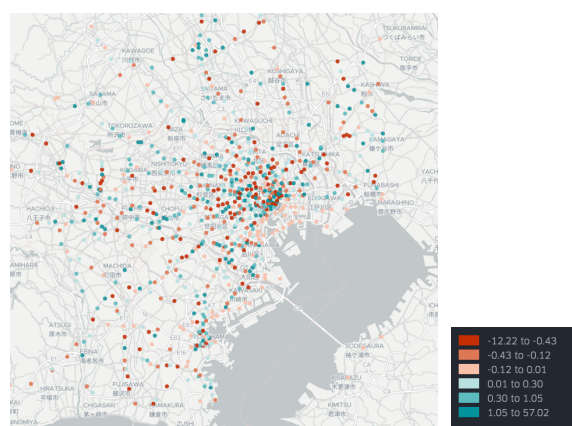


図4 mesh_station_entrance メッシュ分類結果 (VBGMM) と誤差率