

2017年度 CSIS 共同研究

No. 748

「もし都市Aが都市Bだったらマップ」から類似地域 の特徴抽出を行うアルゴリズムの開発

報告書

2017年04月

研究代表者

大阪市立大学大学院生活科学研究科居住環境学講座/教授/瀧澤 重志

CSIS教員

貞広 幸雄

2 都市間で類似する地域の特徴抽出法と他都市への適用に関する研究 — 東京・京都・大阪を事例に

正会員 ○土井 亜香里*1
同 瀧澤 重志*2

都市データ解析 特徴抽出法 アナロジー
ZDD 星状領域列挙 線形計画法

1. はじめに

都市のある地域を知ろうとすると、その地域の都市内での位置づけや雰囲気は、都市全体をよく知らねばわからない。その際、知っている都市を基準として、知らない都市をアナロジー（類推）する手法があり、岸本は京都の特定の地域群を東京のそれで喩えることで、アナロジー的に都市を理解する提案をしている^[1]。アナロジーは直感的なわかりやすさやイメージ想起の手助けになるが、多くの場合、対象への深い理解と経験が必要であるため、多くの都市に展開することは難しい。一方、地理的データに基づく客観的な特徴量を用いた地域の地域抽出法に関する研究がある^[2]。

本研究では、人口や業種、用途地域といった地理的な数値情報から、2つの都市の類似地域の特徴抽出を行う方法を新たに開発し、さらに、得られた特徴関数を他都市に適用することで、その有用性を検証する。

2. 対象地の概要

東京 23 区、京都 11 区の各地域（表-1）を特徴関数作成の元となる類似地域とし、区内全域を特徴関数の重み最適化の対象とする。また、大阪 24 区をその検証の対象とし、地域の抽出を行う。

表-1 東京と京都の類似地域と特徴^[1]

地域	東京	京都	特徴
1	赤羽	四条大宮	住む人と飲む人の混在
2	丸の内	烏丸	高層ビル街に煉瓦建築
3~6	JR中央線沿線	叡山電鉄沿線	
3	中野	出町柳	カオスな商店街、小規模
4	高円寺	元田中	カフェ、古着、雑貨、焼き鳥
5	阿佐ヶ谷	茶山	緑地の豊富な団地
6	荻窪	一乗寺	閑静な住宅街、別荘系
7	目白	御所南	富裕層と文教エリア
8	上野	岡崎	緑、自然、広場、街の歴史
9	代官山（青山）	北山	アパレルショップとカフェ
10	品川駅	京都駅	新幹線の発着駅、街自体は静か
11	代々木	御所東	市街地からは離れる。自然派
12	清澄白河	壬生	美術館とコーヒー、ものづくり
13	蔵前	河原町四条～五条	新旧店舗の共存、欧米系外国人
14	松陰神社	北野天満宮	学問とカフェ、焼き菓子の街
15	二子玉川	北大路	家族向け、駅前の賑わい
16	谷中	西陣	古い家々と寺、銭湯、町屋
17	根津	紫野	観光客の街へと変化中
18	竹下通り	新京極通	大人向けグラデーション
19	渋谷～裏渋谷	河原町～烏丸	ティーン向けグラデーション

3. 提案手法

3.1 概要

まず任意の二つの都市の類似地域を示したマップを用意し、複数の数値属性の重み付き和として、各地域の特徴関数を定義する。そして、各都市で類似していると指摘される地域の特徴関数の値が、他の地域よりも高くなるように、特徴関数の重みを線形計画問題として定式化して最適化する。

本研究では各地域を星状領域で与え、小地域ポリゴンの重心を結んでできる形状が星状領域のクラスで記述できるものに限定する。そしてそれらを ZDD (Zero-suppressed Binary Decision Diagram)^[3]と呼ばれる圧縮データ構造を用いて、小地域数が 1~6 までの領域を全列挙して抽出し、重みの最適化に用いる。最後に、最適化された特徴関数を他の都市に適用し、ZDD で列挙された無数の星状地域集合から、特徴関数の値が最大になる地域を高速に抽出するクエリを開発し、検証を行う。

3.2 特徴関数と重みの最適化

考慮する都市の集合を $C = \{A, B\}$ 、都市 $c \in C$ に含まれる小地域の集合を R_c 、類似関係にある地域を $t_c \subset R_c$ 、都市 c の列挙された地域の集合を $ER_c \subset R_c$ とする。さらに、 $x_r = \{x_{r1}, \dots, x_{rn}\}$ を、各小地域 $r \in R_c$ の特徴を表すパラメータのベクトル、 $w_i \in w, i \in \{1, \dots, n\} = N$ を特徴関数のパラメータの重み、 $n = |x_r|$ をパラメータの数とする。

このとき、特徴関数 $f(s)$ は式(1)のように表すことができる。

$$f(s) = \frac{1}{|s|} \sum_{r \in s} \sum_{i \in N} w_i x_{ri} \quad (1)$$

式(1)は重み×特徴の属性（特徴量）の和のかたちをしている。この式の重みを、以下の線計画問題で最適化することで、特徴関数による特徴の表現を目指す。この線形計画問題は、両都市の対象地域の特徴関数の値と、それ以外の列挙された地域の特徴関数の値の差をできるだけ大きくするように重みを最適化する。

$$\text{Maximize } P(w) = d \quad (2)$$

$$\text{s.t. } f(t_A) - f(s) \geq d \quad \forall s \in ER_A \quad (3)$$

$$f(t_B) - f(s) \geq d \quad \forall s \in ER_B \quad (4)$$

$$d > 0 \quad (5)$$

$$-1 \leq f(s) \leq 1 \quad (6)$$

$$\forall s \in \{t_A\} \cup \{t_B\} \cup ER_A \cup ER_B$$

3.3 用いた特徴量と最適化した重み

地域の特徴関数の属性として、人口、事業所数、用途地域を利用した。人口は、平成27年国勢調査^[4]から年齢（5歳階級）別人口、事業所数はH26年経済センサス基礎調査（小地域）^[4]から産業（大分類）別全事業所数、用途地域は国土数値情報ダウンロードサービス^[5]の用途地域データを利用した。これらは各都市内で、平均0、分散1に標準化したものを用いる。

表-1の19地域の内、今回は地域2, 13, 18, 19の4地域の解が最適化できたため、その解を重みとした特徴関数を大阪に適用する。

4. 特徴関数の適用による大阪の類似地域抽出

作成した特徴関数を大阪に適用し、星状領域集合から特徴関数の値が最大となる領域を抽出した。星状領域を構成する小地域数が少ないほど特徴関数の値は高くなる傾向があるため、それらの小地域数毎に値が最大となる領域を求める必要がある（例えば図-1）。

小地域数1~30について、それぞれ最大の値をとる領域を抽出した結果を表-2に示す。



図-1 大阪での領域抽出の例

地域18, 小地域数10, 船場中央（本町周辺～堺筋本町）

表-2 特徴関数を大阪に適用した結果の一覧

地域	東京	京都	大阪
2	丸の内	烏丸	中央区糸屋町1丁目（天満橋～谷四間） 中央区上汐，東平（谷六～谷九間） 谷町六丁目周辺
13	藏前	河原町四条～五条	中央区糸屋町1丁目（天満橋～谷四間） 中央区上汐，東平（谷六～谷九間） 谷町六丁目周辺
18	竹下通り	新京極通	西淀川区西島1丁目（大阪港北） 中央区城見（京橋，大阪城公園） 中央区船場中央（本町周辺～堺筋本町）
19	渋谷～裏渋谷	河原町～烏丸	中央区糸屋町1丁目（天満橋～谷四間） 中央区上汐，東平（谷六～谷九間） 谷町六丁目周辺

今回の結果について、まず、最適解が得られたこれらの4地域は、東京と京都における類似地域が比較的簡便な特徴を有しているように思われる。例えば地域2の丸の内と烏丸はビジネス街であり、地域18の竹下通り（原宿）と新京極通、渋谷～裏渋谷と河原町～烏丸も、その街の名前が既によく知られている。現状の特徴量では、説明に複数の属性が必要な地域よりも、端的に表現できるものの方が特徴関数の作成がしやすいようである。

結果として、地域2, 13, 19についてはそれぞれ似通った地域が抽出され、解釈が難しい。恐らく、あまりいい特徴関数が作成できていないのだと思われるが、この原因としては用いた特徴量の属性数が少ないこと、元となる類似地域の選定などが挙げられる。

特徴量の属性に関しては、人口、事業所数、用途地域を用いているが、実際の街の雰囲気や位置づけにはより多くの特徴がかかわっている。そのため、昼夜間人口比率や地価、商業集積統計等の特徴量を追加することで、より実情に即した特徴関数の作成が行えるのではないかと、これらの属性追加は重みの最適化にも影響し、今回の4地域以外の地域でも特徴関数が得られる可能性がある。

元となる類似地域の選定に関しては、小地域の部分集合として星状領域を作成して行ったが、それぞれの都市内での規模を合わせることで、特徴が顕著に表現できるように星状領域を作成する必要があるのではないかと。

5. まとめ

東京と京都から作成した特徴関数を大阪に適用した結果、いくつかの地域に対し、ある特定の地域の抽出が可能であった。今後は特徴関数の属性追加や、元となる小地域の研鑽を行い、より良い特徴関数を作成できるようにする必要がある。

謝辞

本研究は東京大学 CSIS の共同研究制度を利用しています。

参考文献

[1] 岸本千佳, もし京都が東京だったらマップ, イースト新書Q, 2016
 [2] 前田翠, 関本義秀, 東京23区における地域の特徴量抽出ならびに定量化と地域の類似度算出, 地理情報システム学会第26回研究発表大会, 宮城大学, D45, 2017
 [3] S. Minato, Zero-suppressed BDDs for set manipulation in combinatorial problems, Proceedings of the 30th international conference on Design automation, 1993
 [4] 政府統計の総合窓口 <http://www.e-stat.go.jp/> (平成30年2月6日閲覧)
 [5] 国土数値情報ダウンロードサービス <http://nlftp.mlit.go.jp/ksj/> (平成30年2月6日閲覧)

*1 大阪市立大学大学院工学研究科 前期博士課程

*2 大阪市立大学大学院生活科学研究科 教授, 博士(工)

Graduate Student, Grad. Sch. of Eng., Osaka City Univ.

Osaka City Univ. Grad. Sch. of Human Life Sci., Prof., Dr. Eng.