

ユーザ行動データによる回遊中心性の拡張 Extending rambling centrality by user activity data

岩崎 一輝[†]
Kazuki Iwazaki

鈴木 優伽[†]
Yuka Suzuki

斉藤 和巳[†]
Kazumi Saito

風間 一洋[‡]
Kazuhiro Kazama

1. はじめに

近年、外国人観光客の増加や2020年の東京オリンピック開催決定などを背景に、観光産業の活性化が重要となっている。観光産業の活性化には、観光地等における回遊者の行動パターンや、主要な観光スポットの特徴などを明らかにしたうえで、新たな観光戦略の構築や観光情報推奨を行う必要がある。

これまでに、寄り道する際の容易さを示す尺度である回遊中心性 [1] を定義することで、道の駅、観光案内所、Wi-Fi スポット等の観光リソースの適切な施設配置候補地集合を提示する技術が提案されている。しかし、従来法では、実際の回遊者の行動や観光スポットの特徴を考慮していないため、必ずしも最適な施設配置が期待できない。本研究では、回遊中心性に実際の回遊者の行動データを組み込んで拡張することで、より回遊者の実行動に沿った施設配置候補を抽出し、観光情報技術への貢献を目的とする。

2. 提案手法

2.1. 回遊中心性

文献 [1] では、あるスポットから別のスポットへ移動する際に、途中で立ち寄る容易さの度を「回遊中心性」と定義している。ここで、観測されるスポットの集合を $S = \{s, t, v, \dots\}$ 、スポット s からスポット t までの実距離を $d(s, t)$ とする。今、スポット s からスポット v を経由してスポット t へ移動する際の、寄り道距離を $D(s, t; v) = d(s, v) + d(v, t)$ とすると、スポット v の回遊中心性は以下で定義できる。

$$\text{mBWC}(v) = \sum_{s \in S} \sum_{t \in S} \frac{d(s, t)}{D(s, t; v)} \quad (1)$$

回遊中心性の高い上位のスポットを施設配置候補とすると、式の性質上、互いに隣接して存在しているスポットなど、一部のスポットが抽出されやすく、特定の回遊者のみが利益を受ける傾向がある。そのため、回遊中心性を集合に拡張した集合回遊中心性を用いることで、一部のスポットが集中して抽出されることを抑え、全体としての効用を高める。スポット集合 S の部分集合を施設配置候補集合 $R \subset S$ とすると、スポット s からスポット t への移動の際に、 R のうちの最寄りのスポット r を経由した最短寄り道距離を $D(s, t; R) = \min_{r \in R} \{d(s, r) + d(r, t)\}$ と表せる。このとき、 R における集合回遊中心性を以下の式で定義する。

$$\text{smBWC}(R) = \sum_{s \in S} \sum_{t \in S} \frac{d(s, t)}{D(s, t; R)} \quad (2)$$

$\text{smBWC}(R)$ において、 $R = \{v\}$ とすれば、 $\text{mBWC}(v)$ に帰着されるので、集合回遊中心性は回遊中心性の自然な拡張であることがわかる。

2.2. 回遊行動中心性

提案法では、集合回遊中心性に回遊者の行動データを組み込むことで、より回遊者の実行動にそった指標へと拡張する。回遊者の集合を M 、回遊者 $m \in M$ がステップ h で訪れたスポットを $r(m, h)$ とし、回遊者 m が訪れた総スポット数を H_m とすると、回遊者 m が滞在したスポット系列は $r(m) = \{r(m, 1), \dots, r(m, h), \dots, r(m, H_m)\}$ と書き表せる。このとき、拡張した回遊中心性を回遊行動中心性とし、以下

で定義する。

$$\text{mbBWC}(v) = \sum_{m=1}^{|M|} \max_{1 \leq h < H_m} \left\{ \frac{d(r(m, h), r(m, h+1))}{D(r(m, h), r(m, h+1); v)} \right\} \quad (3)$$

すなわち、従来法では任意の2スポットを移動する際の寄り道の容易さを算出していたが、提案法では、実際に回遊者が移動した2スポット間での寄り道の容易さを算出する。そのため、距離が比較的近くても、実際にはその組み合わせで訪れることはないような二つのスポット寄り道スポットの抽出が抑えられる。また、提案法は従来法と同様に集合に拡張した集合回遊行動中心性を以下の式で定義する。

$$\text{smbBWC}(R) = \sum_{m=1}^{|M|} \max_{1 \leq h < H_m} \left\{ \frac{d(r(m, h), r(m, h+1))}{D(r(m, h), r(m, h+1); R)} \right\} \quad (4)$$

この目的関数を最大にするような R を貪欲法と局所改善法を用いて算出する。

3. 評価実験

提案手法を、抽出された施設(スポット)の妥当性、及びスポット抽出に用いた実行時間にて評価する。また、妥当性には抽出したスポットの平均人気度を用いる。

3.1. 比較手法

比較手法として、従来の回遊中心性を用いた手法(従来法1)と、回遊者の行動モデル [2] を用いた行動モデルつき回遊中心性を用いた手法(従来法2)を採用する。従来法2で提案されている行動モデルは以下の通りである。

回遊者がスポット s からスポット t への推移確率 $p_1(t|s)$ は、以下で定義できるとする。

$$p_1(t|s; \theta_1) = \frac{d(s, t)^{-\theta_1}}{\sum_{v \in S} d(s, v)^{-\theta_1}} \quad (5)$$

θ_1 は距離に対するパラメータである。また、スポット t の人気度を $f(t)$ 、人気度に対するパラメータを θ_2 とすると、スポット t の選択確率 $p_2(t)$ は次式で表せる。

$$p_2(t; \theta_2) = \frac{f(t)^{\theta_2}}{\sum_{v \in S} f(v)^{\theta_2}} \quad (6)$$

このとき、回遊者はスポット s から別のスポットへ移動する際、以下の確率分布に従うとする。

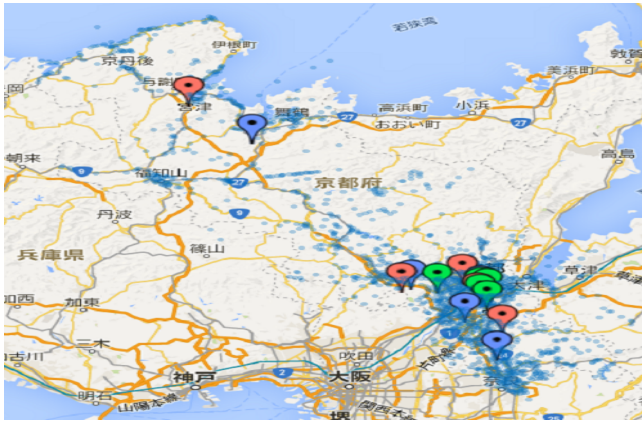
$$p(t|s; \theta_1, \theta_2) = \frac{p_1(t|s; \theta_1)p_2(t; \theta_2)}{\sum_{v \in S} p_1(v|s; \theta_1)p_2(v; \theta_2)} \quad (7)$$

すなわち、行動モデルつき回遊中心性は以下で定義される。

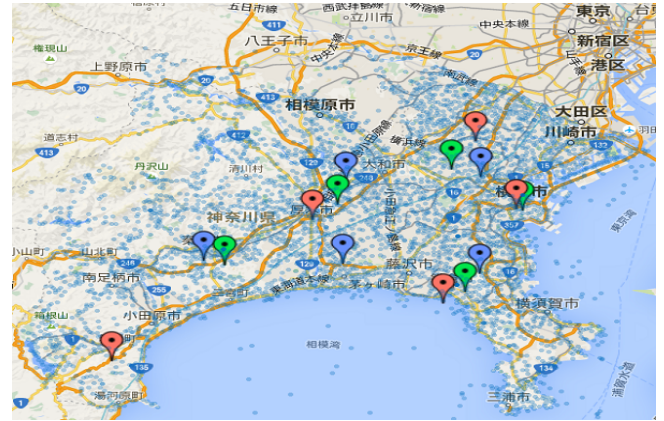
$$\text{psmBWC}(R) = \sum_{s \in S} \sum_{t \in S} p_2(s; \theta_2)p(t|s; \theta_1, \theta_2) \frac{d(s, t)}{D(s, t; R)} \quad (8)$$

[†]静岡県立大学 経営情報学部

[‡]和歌山大学 システム工学部



(a) 京都の抽出スポット



(b) 神奈川の抽出スポット

図1: 各地域のスポット抽出結果

3.2. データセット

オンライン写真共有サイト Flickr に投稿された位置情報付き写真群を用いる。多くの撮影行動がみられる地点は、重要な観光スポットであることが期待できるため、投稿写真群を Mean Shift クラスタリング [3] し、各クラスターの極大点をスポットとする。このとき、各クラスター内での総撮影数を各スポットの人気度と定義する。また本研究では、京都及び神奈川で投稿された写真群を用いた。スポット数・ユーザ数は、京都で 2660・1601、神奈川で 4968・2603 であった。

4. 実験結果

各手法での抽出スポットの平均人気度を従来法1・従来法2・提案法の順で示す。京都では 96・1620・2013 であり、神奈川では 17・23・648 であった。このことから、提案法が従来法1や従来法2にくらべ、人気度の高いスポットを抽出できていることが確認できる。すなわち、提案法を用いることで、より回遊者の効用が高まるようなスポットを提示できると考えられる。

次に、各手法での抽出スポットを実際の地図上にプロットした結果を用いて、提案法を定性的にも評価する。図1に各手法で抽出されたスポットを示す。ただし、本研究では抽出スポット数 $|R| = 5$ とする。従来法1のスポットを青マークで、従来法2のスポットを緑マークで、提案法のスポットを赤マークでプロットした。京都では、提案法で京都駅・亀岡・天野橋立・金閣寺・平等が、従来法1で四条通・愛宕山・桂川・トロッコ亀岡駅・木津川が、従来法2で京都駅前・烏丸通・天龍寺・四条通・伏見稲荷大社が抽出された。神奈川では、提案法で横浜ランドマークタワー・本厚木駅・鎌倉大仏殿高徳院・センター北駅・箱根湯本駅が、従来法1で菅田町・かにざわ公園・瀬上市民の森・秦野駅・文教大湘南キャンパス周辺が、従来法2で海老名駅・カップヌードルミュージアム・鶴岡八幡宮・めん羊の里木里館・中山駅が抽出された。いずれの抽出スポットも、いずれの抽出スポットの近傍にも多くのスポットが見られることから、寄り道に適したスポット抽出ができていることが確認できる。

また、提案法では、従来法1,2で抽出されなかった天橋立や、鎌倉大仏・箱根湯本駅などが抽出できている。この理由として、従来法1,2は、スポット間の距離が重視されるため、人気はあるが、他の観光スポットと少し離れた場所に存在するようなスポットは、寄り道スポットとして抽出されづらい傾向にあるのに対し、提案法は回遊者の行動履歴が重視されるため、少し離れた場所に存在していても、人気のあるスポットは抽出できる傾向があるからだと考えられる。実際に、人気の高いスポットが、鎌倉・横浜・箱根等のエリアに分散して存在している神奈川に着目すると、提案法を用いることで、他の手法と異なり、各エリアからバランスよくスポットを抽出できていることが確認できる。

図2にスポット抽出に用いた実行時間を示す。ただし、京都・神奈川で同様の傾向がみられたため、京都の結果のみと

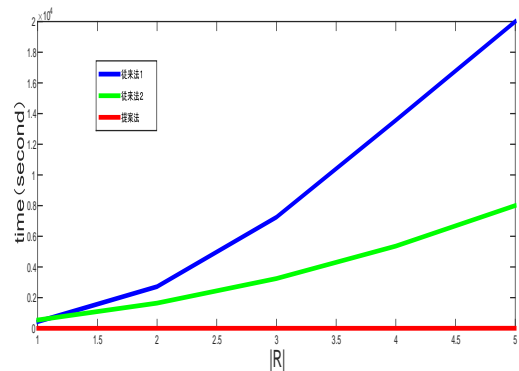


図2: 京都でのスポット抽出時間

する。縦軸に実行時間(秒)を、横軸に抽出スポット数 $|R|$ をとり、青線で従来法1、緑線で従来法2、赤線で提案法の結果をプロットした。図2より、提案法は他の手法に比べ、圧倒的高速にスポットを抽出できることが確認できる。すなわち、妥当な抽出結果を半分近い時間で得ることが可能であり、提案法の有効性が示唆される。

5. おわりに

本研究では、回遊者の行動履歴考慮した回遊行動中心性で、最適な寄り道スポットを抽出する手法を提案した。その結果、提案法はスポットが分散して存在している地域で特に有効的であること、妥当な結果を高速に抽出できることが確認できた。今後は他の地域での検証、及び、より有効性の高いスポット抽出法の提案を試みる。

謝辞 本研究は、総務省 SCOPE(No.142306004)、科学研究費補助金(No.15K00311)の補助を受けた。

参考文献

- [1] 伏見 卓恭, 斉藤 和巳, 武藤 伸明, 池田 哲夫, 風間 一洋, “道路ネットワークに対する実距離を用いた中心性指標の提案と応用”, ネットワークが創発する知能研究会 (JWEIN), 2014.
- [2] 鈴木優加, 伏見卓恭, 斉藤和巳, 風間一洋, “回遊行動モデルに基づく重要観光スポット抽出法”, 情報処理学会全国大会 (IPSJ), 2014.
- [3] D.Comanicu, P.Meer, “Mean Shift: A Robust Approach Toward Feature Space Analysis”, IEEE Transactions on PAMI 24:603-619, 2002.