

回遊行動モデルに基づく重要観光スポット抽出法

鈴木 優伽† 伏見 卓恭†
静岡県立大学経営情報学部

斉藤 和巳† 風間 一洋††
和歌山大学システム工学部

1 はじめに

近年、複雑ネットワークの観点から現実の電力網や道路網の分析が行われている。文献[1]では、ネットワークにおける既存の中心性指標[2]を拡張した指標を用い、道路網などにおける重要な施設を検出することで、ネットワーク上での施設配置問題における数理的な一解法を示している。さらに文献[3]では文献[1]で提案された指標を各スポットの人気度を考慮するように改良することで、より現実の需要にそった施設の検出が可能であるとしている。本研究では、文献[1],[3]で提案された指標をベースに、新たに回遊者の行動モデル[4]を組み込んだ、重要観光スポット抽出法を提案する。具体的には、Levy flight と呼ばれる乱歩モデルを土台にした、回遊者の基本行動モデルを採用し、各観光スポットの人気度だけではなく、スポット間の距離を考慮した重要施設の検出を目指す。さらに、収集した観光情報を用いて回遊者行動モデルを4種類のパラメータの組合せで評価して、パラメータ設定に関する特性や有効性などを検証する。

2 提案手法

2.1 回遊者行動モデル

回遊者の行動をモデル化するために、Levy flight モデルに観光スポットの人気度と距離に依存したパラメータを導入する。

Levy flight モデルとは、ランダムウォークの一種である。進行方向は常にランダムに選択され、稀に非常に長い距離の移動が行われることが特徴であり、ランダムウォークに比べ、効率の良い探索が行われることが示唆されている。観光スポット集合を $S = \{s, t, v, \dots\}$ 、スポット s からスポット t への移動に必要な距離を $d(s, t)$ とする。

この時、回遊者がスポット s の後にスポット t を訪れる確率 $p_\delta(t|s; \delta)$ は、Levy flight モデルにパラメータ δ を設定した次式で定義できる。

$$p_\delta(t|s; \delta) = \frac{d(s, t)^{-\delta}}{\sum_{v \in S} d(s, v)^{-\delta}} \quad (1)$$

ここで δ は指数係数に対応する非負パラメータであり、確率 $p_\delta(t|s; \delta)$ は距離に反比例して小さくなる。また、各スポット t に対し、スポットの人気度 $f(t)$ に比例したスポット選択確率が以下のように定義できる。

$$p_\phi(t; \phi) = \frac{f(t)^\phi}{\sum_{t \in S} f(t)^\phi} \quad (2)$$

本研究では、回遊者行動モデルとして回遊者がスポット s の後にスポット t を訪れる確率 $p(t|s; \delta, \phi)$ を式(1), (2)を用いて次式で定義する。

$$p(t|s; \delta, \phi) = \frac{p_\delta(t|s; \delta)p_\phi(t; \phi)}{\sum_{v \in S} p_\delta(v|s; \delta)p_\phi(v; \phi)} \quad (3)$$

2.2 重要スポット抽出法

文献[1]では、スポット集合 $R \subset S$ に対し、ある観光スポット s からあるスポット $v \in R$ を経由し、スポット t へ行く際のスポット v の立ち寄りやすさを示す回遊中心性を以下のように定義している。

$$\text{smBWC}(R) = \sum_{s \in S} \sum_{t \in S} D(s, t; R)^{-1} d(s, t) \quad (4)$$

ここで $D(s, t; R)$ とは、ある観光スポット s から任意のスポット $v \in R$ を経由し、他の観光スポット t へ行く際の寄り道距離であり、次式で表される。

$$D(s, t; R) = \min_{r \in R} \{d(s, r) + d(r, t)\} \quad (5)$$

$\text{smBWC}(R)$ を最大にするスポット集合を求めることで、回遊者の回遊性や利便性を向上させるスポットが抽出できる。

2.3 行動モデル付き重要スポット抽出法

本研究では式(3)を用い、回遊者の行動モデルを組み込んだ回遊中心性を以下のように定義し、この目的関数を最大にするようなスポット集合 R を求める。

$$\begin{aligned} \text{psmBWC}(R) &= \sum_{s \in S} \sum_{t \in S} p_\delta(s; \phi) p(t|s; \delta, \phi) D(s, t; R)^{-1} d(s, t) \end{aligned} \quad (6)$$

$\text{psmBWC}(R)$ において、パラメータ $\delta = 0, \phi = 0$ とすれば、 $\text{smBWC}(R)$ に帰着されるので、本提案モデルは従来の提案法の自然な拡張である。

3 評価実験

3.1 データセット

レビューサイト「trip advisor」から収集した京都府及び神奈川県の観光スポットのデータを用いて、提案法の有効性を検証する。2節における距離 $d(s, t)$ を各スポットの緯度経度情報と近隣道路網から求め、各スポットのレビュー数を人気度 $f(t)$ とする。京都府の全スポット数は428、レビュー数は22873であり、神奈川県の全スポット数は435、レビュー数は13473である。また、パラメータの組合せからの4つの回遊者モデル、M1:ランダム移動型 ($\delta = 0, \phi = 0$)、M2:距離重視型 ($\delta = 1, \phi = 0$)、M3:人気度重視型 ($\delta = 0, \phi = 1$)、M4:距離・人気度重視型 ($\delta = 1, \phi = 1$)、を用いてスポットを抽出した。

3.2 可視化結果の分析

各モデルで抽出されたスポットの性質や違いについて可視化で定性的に評価する。図1に $K = 10$ の際の京都府における抽出スポットを、M1で抽出されたスポットをピンク▲、M4で抽出されたスポットを黄色★、それ以外のスポットを青●で示す。ここでは紙面の都合上、顕著な違いが見られたM1, M4の結果のみを述べる。M1によって抽出されたスポットは、祇園、宝篋院、河合勘次郎記念館、京都国立現代美術館、旧三上家住宅、寺田屋、武信稲荷神社、西陣織会館、仁左衛門の湯、幸神社である。また、M4で抽出されたスポットは、祇園、金閣寺、嵐山、三十三間堂、銀閣寺、二条城、大

Extracting important sightseeing spots based on rambling activity model
†Yuka SUZUKI †Takayasu FUSHIMI †Kazumi SAITO
††Kazuhiro KAZAMA
†School of Management and Information, University of Shizuoka
††Faculty of Systems Engineering, Wakayama University

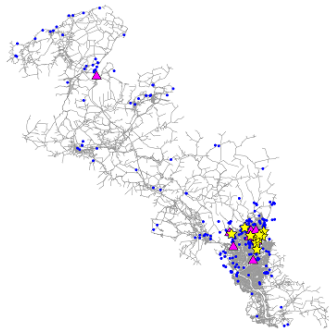


図 1: 京都府抽出スポット

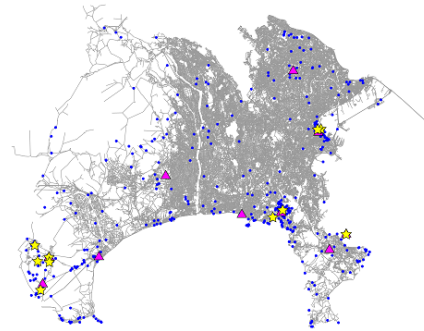


図 2: 神奈川県抽出スポット

表 1: 平均スポット間距離 μ_{dist} , 平均人気度 μ_{pop}

地域	モデル	μ_{dist} (km)	μ_{pop} (review)
京都	全スポット	30.8	53.4
	M1	26.9	125.8
	M2	6.56	36.1
	M3	6.18	1000
	M4	5.76	1011
神奈川	全スポット	35.0	31.0
	M1	37.4	14.2
	M2	42.0	74.7
	M3	40.4	288.4
	M4	44.6	275

雲院, 伏見稲荷大社, 南禅寺, 平安神宮である。どの抽出スポットも近隣に多くのスポットが存在しており, ちょっと立ち寄るのに適している。M1で抽出されたスポットに比べ, M4で抽出されたスポットは, より京都の観光地として著名なスポットであり, 本研究の目的である人気度と距離を考慮した回遊者の要求を満たす重要施設の検出という点に即している。図2に $K=10$ の際の神奈川県における抽出スポットを示す。M1によって抽出されたスポットは, 玉村豊男ライフアートミュージアム, 清閑亭, 箱根美術館, 駒形神社, 大明寺, 猿島, 大巧寺, 鶴沼海浜公園スケートパーク, 馬車道, 横浜歴史博物館である。一方 M4で抽出されたスポットは, 箱根神社, 箱根小涌園ユネッサン, 大涌谷, 強羅公園, 猿島, 長谷寺, 鶴岡八幡宮, カップスードルミュージアム, 横浜ランドマークタワーであり, 京都の時と同様, M4の方が著名な観光スポットを抽出できており, 提案法の有効性が確認できる。

3.3 平均スポット間距離, 平均スポット人気度の分析

抽出されたスポット群が距離の近さと人気度の高さを満たしているか調べるために, 平均スポット間距離 μ_{dist} と平均スポット人気度 μ_{pop} を以下の式で求め, 定量的に評価する。

$$\mu_{dist}(R) = \frac{1}{|K|(|K|-1)} \sum_{s \in R, t \in R} d(s, t)$$

$$\mu_{pop}(R) = \frac{1}{|K|} \sum_{t \in R} f(t)$$

表1に, 京都府, 神奈川県の各観光スポットにおける μ_{dist}, μ_{pop} , 各モデルで抽出されたスポットにおける μ_{dist}, μ_{pop} の値を示す。表1を見ると, 京都府では距離を考慮する M2 と M4 は, μ_{dist}, μ_{pop} が京都府の全体平均と比べ減少傾向にあるが, 全スポット平均に比べ減少傾向にあるが, 神奈川県では逆に増加

傾向にある増加傾向にある。これは, 図1, 2を見ても確認できるように, 京都府は大きな京都盆地があり, その中に観光スポットが集中して存在しているが, 神奈川県は小さな平野が分散し, 観光スポットが互いに離れた横浜, 鎌倉, 箱根などに局所的に集中して存在しているためである。逆に, 本研究のスポット抽出法を用いることで, その地域の重要観光スポットが局所的に存在しているのか, 分散性があるのかを検証することが可能であり, 地域毎の観光スポットの分布特性を明らかにできると考えられる。一般的に, 観光スポットは地域ごとにある種のクラスタを形成しており, 回遊者は各クラスタ内のスポットのみを移動する傾向があるが, 回遊者が特に望む場合に稀に他クラスタに移動すると考えられる。京都と神奈川県の事例から, 提案手法ではこの相反する行動をモデル化できていることがわかる。

4 おわりに

本研究では, 従来のスポット抽出法に回遊者の行動モデルを組み, その有効性を検証した。パラメータ設定による抽出スポットの違いや地域の観光施設の分散性による抽出スポットの違いを明らかにすることができた。今後は, 他の観光地データによる検証を行うとともに, 各スポットごとに「ジャンル」などの属性値情報を与え, より回遊者の行動に適した抽出法への拡張を試みる。

謝辞

本研究は, 総務省 SCOPE(No.142306004), ふじのくに地域・大学コンソーシアム学術研究及び, 科研費 (No.26330345) の補助を受けた。

参考文献

- [1] 伏見卓恭, 齊藤和巳, 池田哲夫, 武藤伸明, 風間一洋. 実距離を考慮した中心性指標の提案と重要観光スポット抽出への応用. 第102回知識ベースシステム研究会 (SIG-KBS102), 2014.
- [2] L. Freeman. Centrality in social networks: Conceptual clarification. *Social Networks*, Vol. 1, No. 3, pp. 215–239, 1979.
- [3] 伏見卓恭, 齊藤和巳, 池田哲夫, 武藤伸明, 風間一洋. 人気度を考慮した重要スポット抽出法. 情報処理学会第11回ネットワーク生態学シンポジウム, 2014.
- [4] Song C, Koren T, and Barabasi A-L Wang P. Modelling the scaling properties of human mobility. *Nature*, pp. 818–823, 2010.