

観光行動分析のためのプローブパーソンデータを用いた中心性指標の評価

大石 真生[†] 齊藤和巳[†] 渡邊 貴之[†]

静岡県立大学大学院経営情報イノベーション研究科

1. はじめに

自治体やDMOにおける観光施策においては、観光客への確実な情報提供とその利便性を向上させるために、Wi-Fiスポットやデジタルサイネージなどの観光資源を最小のコストで適切に配置することが重要である。また、防災拠点や防犯のための監視カメラなどの配置についても同様である。これらの問題を解くために、道路網をネットワークとして捉えることで大規模ネットワークの分析アプローチが適用できる。具体的には、ネットワーク内の各ノードの重要性を指標化する様々な中心性が提案されており、ノードを順位付けすることで重要ノードを抽出することができる[1]。観光分野においては、道路網の交差点や観光スポットをノードとみなし、中心性を計算することができる。また、より多くの観光客に情報提供を行うためのノード抽出の指標として集合媒介中心性が[2]、観光客の回遊性を向上させるための指標として集合回遊中心性が提案されている[3]。本報告では、この2つの中心性によって抽出されたノードとプローブパーソン調査によって収集したデータから得られた中心性を比較し評価を行う。

文献[4]では、ネットワーク分析による都市空間のポテンシャル分析と実行動データを用いた中心性の分析が報告されている。その際に、ネットワーク分析における空間ポテンシャルの指標として媒介中心性が用いられている。また、実行動データとしてプローブパーソンデータを用い、実行動における中心性を求めることで両者の関係を考察している。結果として、ネットワーク上の媒介中心性と実行動の中心性には差があり、実際の移動経路は空間特性に大きく影響されることが報告されている。

本報告では、文献[4]とは異なり、比較する中心性として、媒介中心性を拡張した集合媒介中心性と集合回遊中心性を用いる。これらの集合としての中心性を用いることで、ノード同士の協調的な振る舞いが考慮できるため、実行動との差を縮小できることを示す。さらに、文献[2][4]ではリンクの重みを一切考慮していないため、本報告では重みとして交差点間の距離や所要時間を考慮し、その効果について示す。

2. データセット

本報告では、2015年9月から翌年3月まで実施された静岡県内を旅行する日本人観光客に対して行われたプローブパーソン調査で収集したデータを使用する[5]。調査では、スマートフォンGPSアプリを対象者にダウンロードしてもらいGPSログを収集した。期間中に収集した有効なGPSログデータは128件である。ただし、本報告においては、伊豆半島に限定したGPSログのみ扱う。さらに、移動者の交通手段を自動車だけに限定する。道路網ネットワークは、OpenStreetMap(OSM)をもとに作成した伊豆半島のネットワークデータを用いた。データセットの概要を表1に示す。

表1: 伊豆半島データセット概要

対象者数	測位地点数	交差点数	リンク数
82	10,679	312,267	651,184

Evaluation of Centrality Indicators Using Probe Person Data for Tourist's Behavior Analysis

[†]Mao Oishi [†]Kazumi Saito [†]Takayuki Watanabe

[†]Graduate School of Management and Information of innovation, University of Shizuoka

3. ネットワークの中心性

(1) 集合媒介中心性

まず、ノード集合 V 、リンク集合 E からなる無向ネットワーク $G = (V, E)$ が与えられたとき、ノード $v \in V$ の媒介中心性 $BC(v)$ は以下のように定義される。

$$BC(v) = \sum_{s \in V} \sum_{t \in V} \frac{\sigma_{s,t}(v)}{\sigma_{s,t}} \quad (1)$$

このとき、 $\sigma_{s,t}$ はノード s, t 間の最短パス数を表しており、 $\sigma_{s,t}(v)$ はノード s, t 間の最短パスの中でノード v を通るパス数である。次に、ノード集合 R に対する集合媒介中心性 $GBC(R)$ は以下のように定義される。

$$GBC(R) = \sum_{s \in V} \sum_{t \in V} \frac{\sigma_{s,t}(R)}{\sigma_{s,t}} \quad (2)$$

集合媒介中心性を最大化する K 個のノード集合 R の近似解を求める方法として貪欲法による方法が提案されている[2]。

(2) 集合回遊中心性

道路網の交差点をノードとみなし、各リンクにノード間の実距離を導入する。ノード集合を $S = \{s, t, v, \dots\}$ とし、ノード s とノード t の最短距離を $d(s, t)$ 、 $R \subseteq S$ として考える。ノード s から t への移動において、 $r \in R$ を寄り道ノードとして経由するときの最短距離は次式で定義される。

$$D(s, t; R) = \min_{r \in R} \{d(s, r) + d(r, t)\} \quad (3)$$

集合 R に対する集合回遊中心性は以下のように定義される。

$$RDC(R) = \sum_{s \in S \setminus R} \sum_{t \in S \setminus R \cup \{s\}} \frac{d(s, t)}{D(s, t; R)} \quad (4)$$

集合回遊中心性を最大化する K 個のノード集合 R の近似解を求める方法として貪欲法による方法が提案されている[3]。

(3) 媒介中心性と回遊中心性の比較

媒介中心性と回遊中心性(集合回遊中心性の $K = 1$ の場合)の違いについて考える。ネットワーク A 内のノード s, t 間における、ノード v の寄与度 δ を図1に示す。媒介中心性

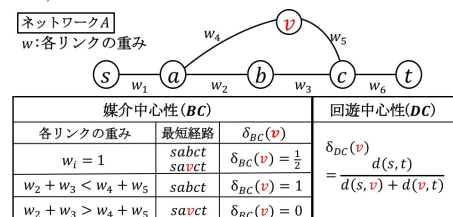


図1: ノード s, t 間におけるノード v の寄与度 δ

では、各リンクの重みによって $\delta_{BC}(v)$ の値は変化する。媒介中心性は、最短経路上に存在し、多くの道路への橋渡しをしているノードほど寄与度の値が高くなる。また、最短経路上にノード v が存在しなければ寄与度は0となる。一方、回遊中心性は、 $\delta_{DC}(v)$ で求められるように、最短経路上にノード v が存在しなくても0とはならず、寄り道距離が短ければ短いほど寄与度の値が高くなる。

(4) 実行動における中心性

実行動における中心性は、文献[4]と同様に定義する。実際の移動データから各ノードの総通過回数を計算し、それを全てのODペア数で割って基準化した値を中心性と定義する。

4. 各種中心性の比較とその評価

伊豆半島のネットワークデータを用いて各種中心性の比較を行う。伊豆半島における各種中心性の分布を図2に示す。色が赤く濃い交差点ほど中心性の値が大きいことを表してい

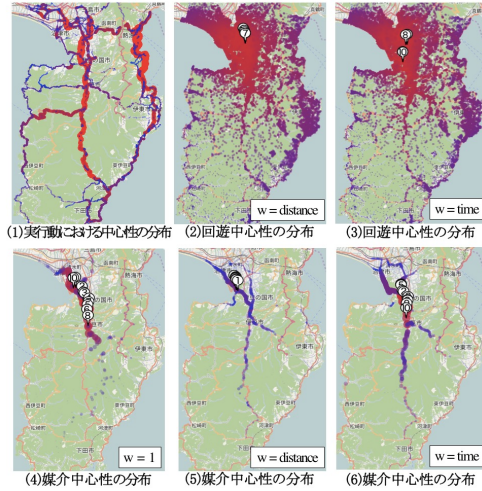


図2: 各種中心性の分布の比較

る。図2(2)-(6)は、中心性の値が大きい上位10位までの交差点にピンを立てている。また、図の右下に考慮したリンクの重み w を示す。 $w = 1$ は重みを1, $w = distance$ は重みとして交差点間の実距離, $w = time$ は重みとして交差点間の所要時間を考慮している。このとき、所要時間はOSMの道路種別をもとに定めた時速と、交差点間の実距離から求めた。例えば、道路種別が motorway なら時速を 80km/h に設定し, living street なら時速を 20km/h に設定する。

媒介中心性では、伊豆半島の中央部を南北に走る国道上の交差点に、回遊中心性ではその国道と周辺に、それぞれ中心性の値が大きい交差点が特に集中して分布している。しかし、両中心性の分布(図2(2)-(6))と実行動における中心性の分布(図2(1))には差がある。具体例としては、図2(6)の媒介中心性の高い交差点だけでは、図2(1)の伊豆半島の東海岸側の実行動における中心性の高い交差点をカバーできない。そこで、ノード同士の協調的な振る舞いが考慮できる集合媒介中心性と集合回遊中心性について評価を行う。両中心性では、 K 個の交差点を抽出するとき、すでに選択された最短経路上にある交差点の影響は除外されるため、特定のエリアに集中して交差点が選択される傾向を改善することができるので、実行動における中心性との差が縮小できると考えられる。

図3(1)-(5)に、重みの異なる両中心性により抽出された上位10位までの交差点を示す。さらに、距離と所要時間の両者の重みを組み合わせて、両者を加味した交差点を抽出することを考える。具体的には、交差点を抽出する際に1個目の交差点の抽出では重みとして実距離を、2個目の交差点の抽出では重みとして所要時間を、と交互に重みを切り替えて集合媒介中心性の増分を計算する。このようにして抽出した上位10位までの交差点を図3(6)に示す。また、図4には図3で抽出された上位交差点を通過したユーザーのカバー率を示す。先ほどの図2(2)-(6)で抽出された交差点と比較して、図3(1)-(6)で抽出された上位交差点は、伊豆半島全体に広がっている。しかし、重みが1の集合媒介中心性で抽出された上位交差点の中には、ユーザーが1人も通過していない交差点が複数存在した(図3(1))。一方で、重みを考慮した図3(2)-(6)の中心性では、上位に抽出された殆どの交差点で通過したユーザーが少からず存在した。そのため、図4では、重みを全て1とした時以外は抽出するノード数を10まで増やすことで、交差点を

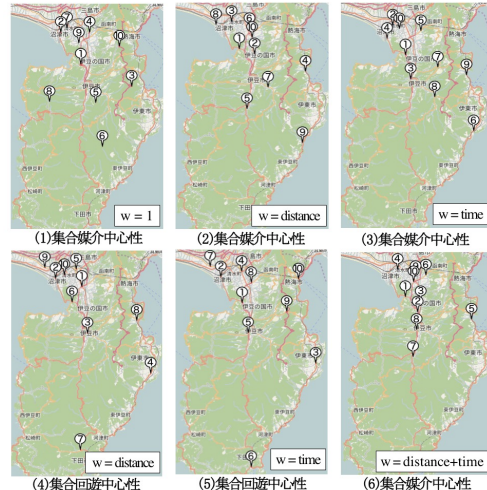


図3: GPSログと各種中心性による抽出上位交差点

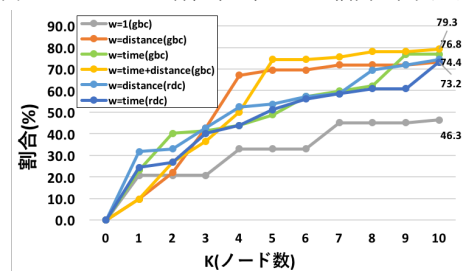


図4: 抽出された上位交差点を通過したユーザーのカバー率

通過するユーザーのカバー率はどれも7割を越えている。つまり、抽出するノード数 K を増やせば増やすほど、ユーザー全体の移動をカバーできるようになり、より現実の移動に近づくと考えられる。

5. 考察とまとめ

本報告では、各種中心性によって抽出されたノードとプローブパーソン調査によって収集したデータから得られた中心性を比較し評価を行った。文献[2][4]では媒介中心性のリンクの重みを一切考慮していないため、本報告では重みとして交差点間の距離や所要時間を考慮した。そして、集合媒介中心性と集合回遊中心性を用いて評価を行った。これらの集合としての中心性を用いることで、ノード同士の協調的な振る舞いが考慮できるため、実行動との差を縮小できることを示した。集合に拡張した両中心性で抽出された上位10位までの交差点を通過するユーザーのカバー率は、重みを全て1とした集合媒介中心性を除いてどれも7割を越えた(図4)。つまり、重みを考慮した集合媒介中心性と集合回遊中心性は現実のユーザーの移動に即した中心性であるといえる。最も割合が高くなったのは、距離と所要時間を考慮した集合媒介中心性となった。

謝辞 本研究は、科学研究費補助基金基盤研究(C)(No.15K00311)の支援を受けて行ったものである。

参考文献

- [1] L.C.Freeman, "Centrality in social networks: Conceptual clarification", Social Networks, vol.1, no.3, pp.215 - 239, 1979年.
- [2] 伏見, 斉藤, 池田, 武藤, "ノード群の協調的振る舞いに着目した集合媒介中心性の提案と応用", 電子情報通信学会論文誌 D. Vol. J96-D, No.5, pp.1158-1165, 2013年5月.
- [3] 鈴木, 斉藤, 風間, "最尤推定にもとづく回遊行動統計モデリング", ネットワークが創発する知能研究会 2015(JWEIN2015), 2015年8月.
- [4] 斉藤, 福山, 羽藤, "ネットワーク分析とプローブパーソンデータを用いた都市における歩行空間分析-横浜中心部を対象として-", 景観・デザイン研究講演集 No.6, 2010年12月.
- [5] 長島, 渡邊, 大久保, 鈴木, 木村 "Web アンケートとGPSアプリによる観光行動調査システムの構築とその応用", 観光情報学会第13回全国大会, 2016年7月.