

# Multi-Period 観光にむけた POI クラスタリング手法

森竹 壮汰郎<sup>†</sup> 笠原 秀一<sup>††</sup> 馬 強<sup>†</sup>

<sup>†</sup> 京都大学大学院 情報学研究科 〒 606-8501 京都市左京区吉田本町

<sup>††</sup> 京都大学 学術情報メディアセンター 〒 606-8501 京都市左京区吉田本町

E-mail: <sup>†</sup>moritake@db.soc.i.kyoto-u.ac.jp, <sup>††</sup>kasahara.hidekazu.3m@kyoto-u.ac.jp, <sup>†††</sup>qiang@i.kyoto-u.ac.jp

**あらまし** 1日や午前午後の単位で期間 (Period) を区切って大まかに訪問地域を決め、広範囲・長期間の観光を計画することがよくある。それをサポートするため、POI(Point Of Interests) をクラスタリングして観光計画をする手法が提案されている。しかしながら、既存のクラスタリング手法は地理的距離に基づくものが多く、POI の特徴や実際の観光客の移動パターンへの考慮は不十分である。そこで本研究では、観光客の軌跡データを対象に、POI の特徴を考慮した k-means ベースのクラスタリング手法を提案する。提案手法では、まず軌跡データを用いて POI のハブとオーソリティの値を算出し、交通の軸となる中枢 POI と魅力の高い必見 POI に分類を行う。POI を k-means でクラスタリングする際、クラスタの中心を、必見 POI への移動コストが最小となる中枢 POI になるようミニサム法を使って求める。

**キーワード** 観光, POI(Point Of Interests), クラスタリング, 軌跡データ, Hub/Authority

## 1 はじめに

コロナ禍であっても、観光庁 [1] は 2030 年の訪日外国人数の目標値を 6000 万人に設定しており、観光産業は依然として最も重要な産業とされている。またコロナ禍以前より、分散化・個人化した観光に向けた潮流がある [2] [3]。それらの観光スタイルは、コロナ禍をきっかけにより重視され、ウィズコロナ・アフターコロナの時代においてもこの流れは継続されると考えられる。

広範囲または長期間の観光計画を立てる際、1日や午前午後の単位に期間 (Period) を区切って、その期間に対応させるよう大まかに訪問地域を設定して観光計画を立てることがよくある。このような観光計画をサポートするため、既存の研究では、POI (Point Of Interests) をクラスタリングして地域を設定することで、観光計画をする手法が提案されている [4]。しかしながら、これらのクラスタリング手法は地理的距離に基づいており、POI の特徴や実際の観光客の移動パターンを十分に表現できていない。

我々は、長期間・広範囲の旅程を複数の期間 (Period) に分け、期間ごとに地域を対応させる観光計画を「Multi-Period 観光」と定義し、これを支援するツアープランの作成について研究を行っている。本論文では、実際の観光客の移動パターンを重視しながら、POI の特徴を考慮した k-means [5] ベースのクラスタリング手法を提案する。

具体的な提案手法として、観光客の軌跡データから観光客の移動を表現する遷移ネットワークを構築し、交通の軸となる中枢 POI と魅力の高い必見 POI に分類を行う。例えば京都市内では、中枢 POI は「嵐山駅」、必見 POI が「竹林の小径」や「天龍寺」に該当する。これは、嵐山駅が嵐山エリアへのアクセスのための「交通の軸」として利用され、嵐山駅そのものが

観光目的ではなく、「その周辺の観光スポット (必見 POI)」が観光目的となるケースが適しているためである。

そして、POI を k-means でクラスタリングし、クラスタの中心点を必見 POI への移動コストが最小となる中枢 POI になるようミニサム法を用いて求める。

本研究における貢献は以下の通りである。

- 長期間・広範囲の観光計画を支援するため、観光客の移動特性や POI の特性を踏まえた、k-means とミニサム法 [6] [7] を組み合わせたクラスタリング手法を提案する。
- 京都市内を対象にした軌跡データを使って検証を行なった結果、クラスタリングが適切な方法であることがわかった。

本論文の構成は次の通り：2 節では関連研究、3 節では本研究の具体的な提案手法、4 節では提案手法の実験と評価、5 節では考察、6 節ではまとめと今後の展望について論じる。

## 2 関連研究

観光情報学の分野は「どのようなスポットなのか」「どうやって行くのか」などの様々な切り口から取り組まれている。特に、SNS(Social Network Services: ソーシャルネットワークサービス) やブログなどで生成される UGC(User Generated Content: ユーザ生成コンテンツ) を対象に、利用者視点から観光地に関する観光集合知をマイニングし、観光推薦に利活用する試みが多数行われている [2] [3]。

### 2.1 遷移ネットワークを用いた POI 評価について

J.Sun ら [8] は、実際の軌跡データから POI を抽出し、重み付き HITS ベースのアルゴリズム [9] を適用させることで、POI の人気度を新たな観点から表現する試みを行った。具体的には、POI をその場所の魅力や体験の価値を測るオーソリティと、他の POI へのアクセスのしやすさを測るハブの 2 種類の基準を

導入した。J.Sun らの手法は、POI を観光の要所と交通の要所という新しい評価基準で分類することで、観光客の移動特性を十分に踏まえたスポット推薦を可能にする点が優れている。本研究では、期間 (Period) 単位のツアープランを生成するために、J.Sun らの定義を活かしたクラスタリング手法について提案を行う。

## 2.2 POI 分析について

POI の分析について、UGC などのデータを用いて評判、ユーザーの行動、景観など様々な尺度から分析が行われている。例えば、観光は「光」を「観る」のであることに着目し、景色や景観物に着目した POI 評価・分析手法が存在する。Y. Shen らは [10] は、観光の価値を自然の価値と文化の価値の側面から推定し、ジオソーシャル画像の分析を行なった。具体的には、自然の価値では、観光客の心理的要求を満たすかどうか検証するために、人間の好感度に関連する画像特徴を抽出し、文化の価値では、POI に含まれる文化的要素を認識することで、POI の品質推定を行った。

## 2.3 POI スポット推薦について

宜・山岸ら [11][12] は、ユーザー体験の 4 つの要素 (Who, When, Where, What) を UEM モデルとしてモデリングすることで、旅行者がどこで何をしたかを明らかにしながら、それに基づいてユーザーがやりたいことを体験できるスポットを推薦する試みを行った。この UEM モデルはユーザーごとに異なるスポットの推薦結果を出力する。そのため、本論文で提案するクラスタリング手法を組み合わせることで、出力されたスポットを具体的なツアー計画にする際に役立てることが可能だと考えられる。

## 2.4 POI クラスタリングについて

D.Gavalas ら [4] は、観光客が公共交通機関を移動することを支援するツアープランアプリケーション「eCOMPASS」を開発した。アプリケーション内では、観光に使える時間を最大化、交通機関での待ち時間を最小限にするために POI の訪問順 (旅程) をクラスタリング等を用いて計算している。クラスタリングでは、経度や緯度の地理的距離に基づいて k-means を行い、旅程を作る際には同じクラスタにグループ化された POI を連続して訪問させる方針をとっている。この手法は複数日への観光にも適用が可能であるが、地理的基準のみに基づいているため、観光地の特性や実際の観光客の移動パターンを十分に表現したクラスタリングとはいえない。そこで本研究では、実際の軌跡データから POI の中枢 POI や必見 POI を分析して、それを用いたクラスタリング手法を提案する。

さらに、Mean-shift [13] や DBSCAN [14] などのクラスタリング手法を適用するには、適切なパラメータの設定が必要であるが、それは容易ではない。そこで、J.Sun ら [7] は、既存の観光名所の面積や規模の大きさに合わせて、異なる粒度の POI を発見するための都市適応型クラスタリングフレームワークを開発している。手法では、2 つのクラスタリングアルゴリズムを利用し、異なるデータセットに適したパラメータを自動的に

同定することにより、異なる都市に適応的に対応している。

## 2.5 パッケージ推薦について

C.Li ら [15] は、単一のアイテム・リストだけでなく異質で多様なオブジェクトの組み合わせをパッケージとして推薦する試みを行った。C.Li らは記事・友人・メディアをパッケージとして、ユーザーが複数のパッケージの中からどのパッケージに属するのか推薦を行っている。手法では深層学習のアプローチをとることで、パッケージ内・パッケージ間の特性とユーザーの趣味嗜好を掴んでいる。これに対し、我々は既に存在するパッケージを推薦するのではなく、新たにパッケージを生成することを試みている。

## 2.6 ツアープランニングについて

複数の観光スポットを効率良く周遊するルートの発見は組合せ問題の 1 つである TTDP (Tourist Trip Design Problem) としてモデリングされている。前項で取り上げた D.Gavalas ら [4] の取り組みでも、観光に使える時間を最大化、交通機関での待ち時間を最小限にするプランニングを行うため、POI クラスタリングの後、複数のクラスタから順次 POI を選択することでルート生成を行っている。また、POI 選択の際には時間的余裕を持たせることで、全体としても余裕のあるツアープランニングを可能にしている。さらに、スポット間の移動方法は時間帯によって変化し、各時間帯で最も早く到着できる移動方法が選ばれるようになっている。

また、J.Sun ら [16] は、軌跡データから観光客の移動パターンを分析し、場所と遷移の知識を組み合わせ、場所と遷移の両方に対して報酬を与える移動経路推薦手法を提案している。

画像データを使ったツアープランニングも行われている。X.Lu ら [17] の Photo2Trip では、ジオタグ付き写真から回収した既存の旅行ルートを活用し、ユーザーの好みに応じてカスタマイズした旅行ルート計画を提案している。ジオタグ付き写真から、訪問スポットとスポット内の移動経路、スポットの滞在時間、スポットを訪問する順序、スポット間の移動時間などの軌跡情報を利用できる。そのため、訪問場所、訪問時期/季節、旅行期間、目的地などの個人的な好みをインタラクティブに指定し、システムを利用できる点で優れている。

さらに、直近では強化学習の手法を取り入れたツアープランニングも行われている。OPTW (Orienteering Problem with Time Windows) と呼ばれる、異なる訪問先から集めたスコアの合計を最大化することを目的とした組合せ最適化問題がある。R.Gama [18] らは OPTW 問題を解決するために、強化学習を導入してポインタネットワークモデルを利用した。特定のスポットを訪れる観光客間で変化し得る変数 (開始位置、開始時間、利用可能時間、各 POI に与えられたスコア) をサンプリングしてポインタネットワークを学習させることで、ツアープランニングを行っている。観光の分野ではデータがスパースになる場面が少なくないため、強化学習の手法は今後より効果を発揮すると考えられる。

### 3 提案手法

この節では具体的な提案手法について説明する。概要を図1に示す。大きな流れとして、軌跡データを入力とし、クラスタリングの結果である各期間 (Period) で訪問する POI のクラスタを出力としている。まず、軌跡データをもとに POI の遷移ネットワークを構築し、重み付き HITS アルゴリズムを用いて各 POI のハブ値およびオーソリティ値を求める。次に、ハブ値の高い POI についてクラスタリングを行うが、クラスタの中心をミニサム法で求める。

以下、3.1 節で軌跡データから遷移ネットワークを構築し、3.2 節で POI をクラスタリングする手法について述べる。クラスタリングの結果から、3.3 節でクラスタの中心を調整する手法について述べる。

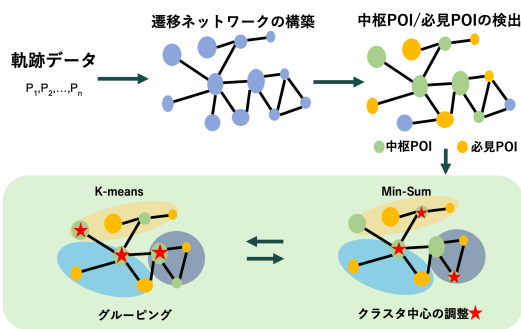


図 1: 提案手法概要

#### 3.1 GPS 遷移ネットワークの構築

J.Sun らの手法 [5] を用いて、GPS ログデータ [19] より抽出された POI データと移動データを用いて遷移ネットワークを作成する。ログデータは、観光客から取得され、各観光客がある時刻にどの地点に存在していたかが記録されている。ユーザー  $u$  の GPS デバイスから、ユーザー  $u$  が訪れた場所を記録したログデータから空間点  $P_u = (p_{u,1}, \dots, p_{u,n})$  を得る。各空間点  $p_{u,i} \in P_u$  はタプル  $(lat_{u,i}, lng_{u,i}, t_{u,i})$  であり、 $lat_{u,i}$  および  $lng_{u,i}$  は場所の座標を、 $t_{u,i}$  は訪問時刻を記録する。

この各地点を Mean-shift や DBSCAN などのクラスタリング手法を用いて処理し、滞在点  $S = (s_1, s_2, \dots, s_m)$  を得る。各滞在点  $s_i$  には、名称、緯度、経度と滞在時間  $Dur_{s_i}$  も関連付けられている。 $Dur_{s_i}$  は  $s_i$  を訪問したユーザーの時間間隔の平均値である。このようにして、全ユーザーの滞在地点の軌跡を  $Tra_j$  として抽出する [20]。

グラフ生成手順としては以下の通りである。

- POI データからグラフを作成する。POI をノードとする。ノード集合を  $V$  とする。
- 移動軌跡を参照し、POI  $P_o$  から  $P_j$  に移動軌跡がある時、有向辺  $e_{o,j}$  を作成する。
- エッジ集合を  $E$  とし、構成されたグラフを  $G$  とする。
- グラフ  $G$  に対して重み付き HITS アルゴリズムを適用させることで、ハブ値ならびにオーソリティ値を求める。各値

から、中枢 POI 並びに必見 POI を検出する。

本研究では、ハブを「多くの観光名所とリンクしている POI」、オーソリティを「多数の移動の要所とリンクされている POI」とする。そのため、ここではハブ値とオーソリティ値を以下のように定義している。

- ハブ値：他のスポットへのアクセスしやすさ
- オーソリティ値：その場所の魅力や経験の価値

この定義から、本研究では POI を 3 つに分類する。

- 中枢 POI：ハブ値が高い POI であり、その場所を中心に観光を進めることができる移動の中枢となる POI
- 必見 POI：オーソリティ値が高い POI であり、魅力溢れる必見のスポットとなる POI
- その他の POI：中枢 POI・必見 POI どちらにも該当しない POI

ここで注意しておきたいのが、中枢 POI と必見 POI は相反するものではなく、良いオーソリティ値をもつ POI は良いハブ値を持つ POI にもなる可能性があることである。

#### 3.2 クラスタリング

1 節でも述べたように、本研究では、長期間・広範囲の旅を複数の期間 (Period) に分け、期間ごとに地域を対応させる観光計画を Multi-Period 観光と定義し、それを支援する Multi-Period プランニングを支援する研究を進めている。Multi-Period プランニングでは、長期間・広範囲の旅を複数の期間 (Period) に分け、期間 (Period) とクラスタを対応させることで、中枢 POI を起点に必見 POI を訪問する観光ルートを構築する。具体的にはまず、観光に費やすことのできる時間予算を午前・午後など数時間単位の期間 (Period) に区切る。区切った期間 (Period) に対し、クラスタリングで生成したクラスタを割り当て、1 期間 (Period) でひとつのクラスタを訪問するよう動機づける。

ここでは、半日を 1 期間 (Period) と定義してユーザーが観光する MultiPeriod を作成し、期間 (Period) 数に応じたクラスタを作成する。例えば、二日間の旅行をする場合、4 つの期間 (Period) が作成される。これらの期間 (Period) の数をクラスタ数として、POI のクラスタリングを行う。

具体的な手続きを以下に示す。

- (1) 任意の中枢 POI を  $K$  個 (=期間数) 選択し、それぞれを期間クラスタの中心として設定する。また、クラスタの中心ベクトルを  $\mu$  とする。また、POI のベクトルには緯度、経度、ハブ値、オーソリティ値から構成される。
- (2) 残りの中枢 POI、必見 POI およびその他の POI と、各クラスタ中心  $\mu_k$  までの距離を計算する。ユークリッド距離が一番近いクラスをその POI の所属クラスタとする。
- (3) ミニサム法 [6] [7] を用いて、クラスタの中心となる中枢 POI を選定する。詳細は 3.3 節で述べる。
- (4) 終了条件を「クラスタの中心に変更がなかった場合」とし、それを満たした場合は終了する。満たしていない場合は、(2) に戻り以降の手続きを再度実行する。

### 3.3 クラスタの中心の決定

ここでは、3.2節の手順(4)で構築されたクラスターについて、ミニサム型施設配置問題(Weber問題)[6][7]になぞらえてクラスタの中心を求め、更新を行う。具体的な手続きを以下に示す。

(1) 各クラスタの中にあるPOIに対して、必見POIを $u$ 、その他のPOIを $v$ 、中枢POIを $x$ と置いて分割する。

(2) 各 $u, v$ から各 $x$ までの距離、及び計算対象の $x_i$ と他の $x$ 間の距離の総和 $T(x_i)$ を計算する。

$$T(x_i) = \sum_{k=1}^K |x_i - u_k| + \sum_{l=1}^L |x_i - v_l| + \sum_{m=1}^M |x_i - x_m|$$

(3) ここで、中枢-中枢、中枢-必見POI間の距離に対して、以下のような2種類の割引率を導入する。割引率を導入する意義として、オーソリティ値の高いスポットと密接なつながりをもつ、ハブ値の高いPOIを選ぶことや、Hub間の移動を容易にすることで、Period間のスムーズな移動を誘導することが挙げられる。なお、 $P_i(\text{authority})$ はそのPOIのオーソリティ値、 $P_i(\text{hub})$ はそのPOIのハブ値、 $\text{Max}(\text{authority})$ は全POI中のオーソリティの最大値、 $\text{Max}(\text{hub})$ は全POI中のハブ値の最大値をそれぞれ表す。また、 $\varepsilon$ は割引率を導入したことで値が0になることを防ぐ定数である( $0 < \varepsilon < 0.001$ )。

- 中枢割引率

$$1 - \frac{P_i(\text{authority}) + P_i(\text{hub})}{\text{Max}(\text{authority}) + \text{Max}(\text{hub})} + \varepsilon$$

- 必見割引率

$$1 - \frac{P_i(\text{authority})}{\text{Max}(\text{authority}) + \text{Max}(\text{hub})} + \varepsilon$$

これらの割引率を各 $u, v$ から各 $x$ までの距離、及び計算対象の $x_i$ と他の $x$ 間の距離に乘算することで、 $T(x_i)$ を定義する。

$$T(x_i) = \text{必見割引率} \times \sum_{k=1}^K |x_i - u_k| + \text{中枢割引率} \times \sum_{m=1}^M |x_i - x_m| + \sum_{l=1}^L |x_i - v_l|$$

(4) この $T(x_i)$ を最小にするPOIである $x$ を求め、それをクラスタの中心とする。

新しく設定した中心に対して、前節のクラスタリングを再度実施し、ミニサム型施設配置問題で求めた中心の位置とクラスタリングで求めたクラスタの中心の結果に変化が生じなくなるまで反復を行う。

## 4 実験

この節では、提案手法の有用性の検証を行う。まず、実験に用いるデータセットについて説明し、次にそれぞれの実験について説明を行う。

### 4.1 データセット

J.SunやXuらの研究[5]で用いられたスポットとその軌跡が含まれたデータセットを用いた。データセットには京都府内の91個のPOIに対して、名称、スポットの説明文、経度、緯

度の情報が含まれている。またここから求めたPOIのハブ値、オーソリティ値の分布は図2,3の通りである。

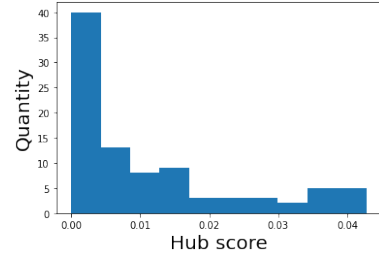


図2: ハブ値の分布

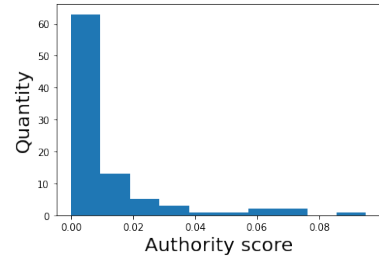


図3: オーソリティ値の分布

### 4.2 実験設定

中枢POI、必見POIの定義として、中枢POIはハブ値が0.035以上のPOI、必見POIはオーソリティ値が0.02以上で、中枢POIに含まれていないPOIとした。

期間(Period)の数として4~8の5通りを設定した。ミニサム手法で用いる距離計算については、割引率区別あり(中枢POI-中枢POI間には中枢割引率を、中枢POI-必見POI間には必見割引率を適用)、割引率区別なし(中枢POI-中枢POI間、中枢POI-必見POI間どちらも必見割引率を適用)、割引率なしの3通りを設定した。期間(Period)5通りと割引率の適用方法3通りを組み合わせ、合計15通りの試行を行った。

評価では、本提案手法で得られたクラスタリング(クラスタ)と人手によるクラスタリング結果(カテゴリ)を比較した。人手によるクラスタリングは、観光客を想定し、データセットにある91個のPOI候補の中から、与えられたPeriod数に応じてPeriod(半日)ごとに訪問したいPOIを分類した。そのため、1Period(半日)で訪問可能なスポットの集合を1つのクラスタとして構成する。

評価指標として、Purity, Inversepurity, Fの3つを用いた。Purityは生成されたクラスタにおいて最もよく現れる人手で定めた正解の頻度を示し、ノイズの少ないクラスタを高く評価する。また、以下のように表す。

$$\text{purity} = \frac{|C_i|}{N} \sum_{i=1}^L \max \frac{|C_i \cap L_j|}{|C_i|}$$

Inversepurityは各カテゴリで定められた要素を多く含むクラスタを高く評価する。Inversepurityは以下のように表す。

$$Inversepurity = \frac{|L_j|}{N} \sum_{i=1}^L \max \frac{|C_i \cap L_j|}{|L_i|}$$

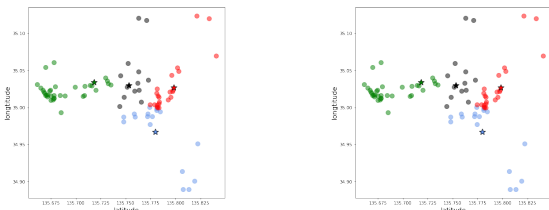
F は上記 2 つの調和平均である。

$$F = \frac{2 \times purity \times Inversepurity}{purity + Inversepurity}$$

ただし、 $N$  をデータ数、本提案手法で作成されたクラスタ集合を  $C$ 、人手で作成したカテゴリの集合を  $L$  とする。

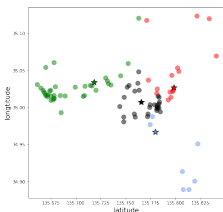
### 4.3 実験結果

ここでは、15 通りの施行の結果得られたクラスタリング結果と Purity について示す。クラスタリングの結果を図 4 から図 8 に、Purity, Inversepurity, F の結果を図 9 から 11, Purity, InversePurity, F の平均値を表 1 から 3 に示す。同じ色で塗られているポイントはそれぞれ同じクラスタに属していることを表し、☆で示されているポイントはクラスタ中心を表す。



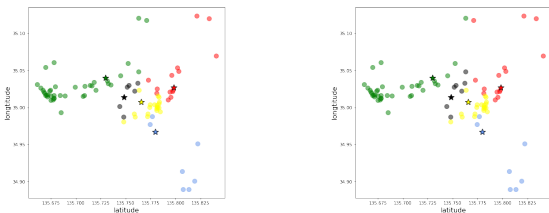
(a) 割引率区別あり

(b) 割引率区別なし



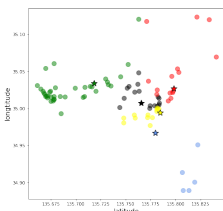
(c) 割引率なし

図 4: Period=4 のクラスタリング結果



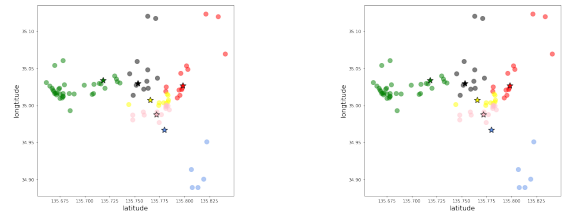
(a) 割引率区別あり

(b) 割引率区別なし



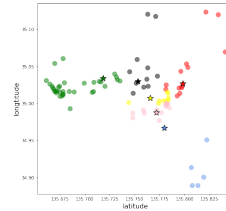
(c) 割引率なし

図 5: Period=5 のクラスタリング結果



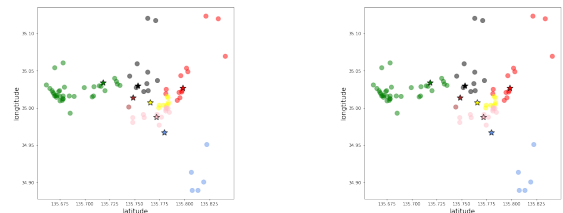
(a) 割引率区別あり

(b) 割引率区別なし



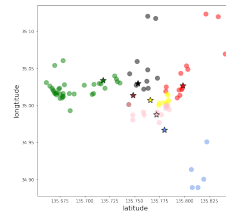
(c) 割引率なし

図 6: Period=6 のクラスタリング結果



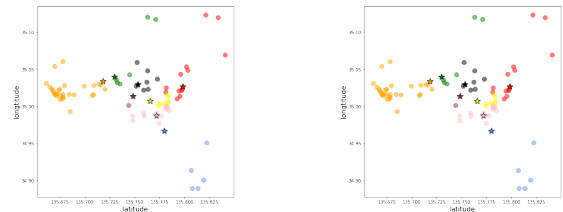
(a) 割引率区別あり

(b) 割引率区別なし



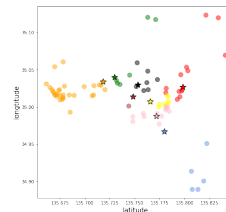
(c) 割引率なし

図 7: Period=7 のクラスタリング結果



(a) 割引率区別あり

(b) 割引率区別なし



(c) 割引率なし

図 8: Period=8 のクラスタリング結果

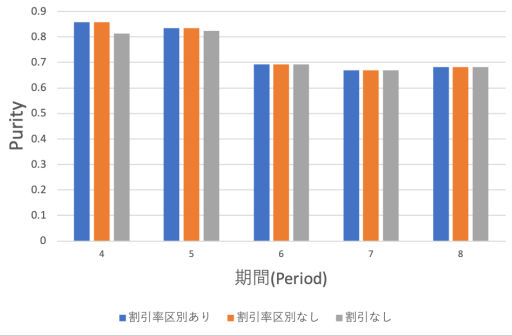


図 9: purity の結果

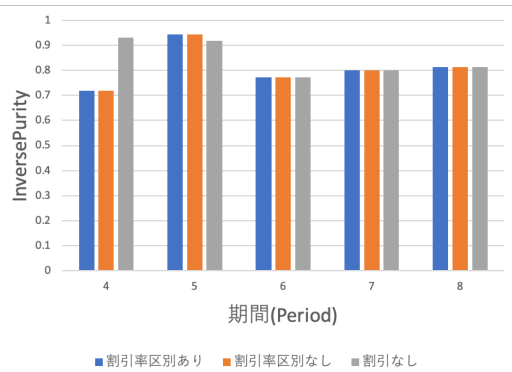


図 10: Inversepurity の結果

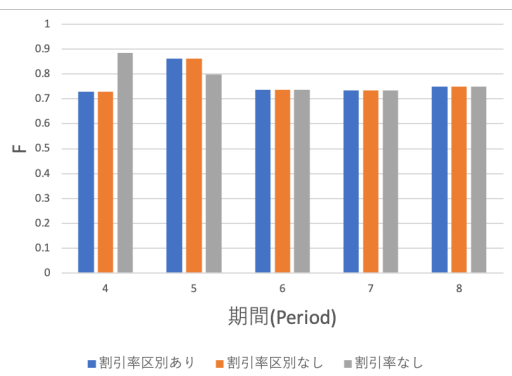


図 11: F の結果

表 1: purity の平均値

期間 (Period)	平均値
4	0.8424
5	0.8314
6	0.6923
7	0.6700
8	0.6810

表 2: Inversepurity の平均値

期間 (Period)	平均値
4	0.7895
5	0.9349
6	0.7716
7	0.8011
8	0.8131

表 3: F の平均値

期間 (Period)	平均値
4	0.7799
5	0.8405
6	0.7361
7	0.7341
8	0.7506

## 5 考 察

Period 数を増やすにつれて、中心調整の割引率の適用法に依らず、中心の変更はあったもののクラスタリングの結果に変更が見られなかった。そのため、Period 数の増加に伴い Purity が低下する現象が見られたが、それぞれの Period においては割引率を導入した際の評価に差が見られなかった。考えられる原因として 2 点考察を行った。

- 標高の考慮について：提案手法では、貴船神社、鞍馬寺と延暦寺など実際にこれら全てを訪問するには山の登り降りが必要な POI が同じクラスタに分類されるなどの事例があった。対して人手による結果では、上記のような POI は山の登り降りを加味するとそのスポットだけで 1Period(半日) 費やすため、各 POI 単体のクラスタが生成される結果となっていた。
- 交通の便の考慮について：提案手法では金閣寺・嵐山エリアを同じクラスタに分類をしていたが、人手による結果ではそれらは別クラスタに分類されていた。この考察として、人間は交通の便による体感距離の違いまで加味していることが考えられるため、人手でのクラスタリング結果と差が生じたと考えられる。

また、今回はユークリッド距離を用いてクラスタリングの距離計算を行なった。しかし、本手法を Multi-Period 観光にするためにはユークリッド距離ではなく実際の「訪問しやすさ」や「訪問にかかる時間」などより実際のシチュエーションを意識した距離計算を行う必要があると考えられる。

さらに、中枢 POI・必見 POI の分類方法についてさらに工夫する必要があると考えられる。これは、高いハブ値をもつ POI は高いオーソリティ値をもつ POI の例が多いため、閾値を設けて分類を行う現在の方法から再検討する必要がある。

## 6 まとめ及び今後の展望

本論文では、長期間・広範囲の旅行を計画する際、旅程を複

数の期間 (Period) に分け、そこに地域を対応させる「Multi-Period 観光」に向けて、実際の観光客の移動特性や POI の特性を踏まえたクラスタリング手法について提案を行った。実験の結果を受けて、クラスタリング方法についてさらなる検討を行う予定である。具体的には、今回行ったユークリッド距離によるクラスタリングだけでなく、移動時間や移動のしやすさに焦点を置いたクラスタリングを試行する予定である。

評価についても、今回行った Purity, Inversepurity, F だけでなく、実際に Multi-Period 観光を適用させることを念頭に定量的・定性的な面から評価を行う予定である。

さらに、本手法においては、

- ホテルの推薦：クラスタの中心にミニサム法を用いてホテルの場所を決める

- 個人化推薦：個人化した広範囲・長期間の観光の支援

- テーマ性を持った観光プランの生成：「坂本龍馬の歴史」等テーマ性を持つ観光の支援

など様々な旅のシチュエーションに対して応用を目指す予定である。

また、本手法は既に行われている個人化にむけた観光スタイルの試みとも親和性が高い。既に、宜や山岸ら [2] は、UEM (User Experience Model) モデルを構築した。このモデルは観光の体験を分析することで、ユーザの好みにあったスポットを推薦することを可能とする。このモデルで複数のユーザが得た複数のスポットに対して、本論文で提案するクラスタリング手法を適用することによって、個人の好みのスポットを十分に訪問することができる観光プランの立案にも繋がると考えられる。

今後、Multi-Period 観光の支援に向けて様々な観点から取り組みを進めていきたい。

## 7 謝 辞

本研究の一部は総務省 SCOPE (201607008) による。

## 文 献

- [1] 観光庁. 観光庁 観光ビジョン実現プログラム 2020. [https://www.mlit.go.jp/kankocho/news02\\_000419.html](https://www.mlit.go.jp/kankocho/news02_000419.html).
- [2] 馬強. UGC を利用した観光資源の発見と推薦. 情報処理, Vol. 62, No. 11, pp. e12-e17, 2021.
- [3] 馬強. 観光情報学の最前線-観光の分散化と個人化を促進する集合知活用情報技術. 情報処理, Vol. 58, No. 3, pp. 220-226, 2017.
- [4] Damianos Gavalas, Vlasios Kasapakis, Charalampos Konstantopoulos, Grammati Pantziou, Nikolaos Vathis, and Christos Zaroliagis. The ecompass multimodal tourist tour planner. *Expert systems with Applications*, Vol. 42-21, pp. 7303-7316, 2015.
- [5] James MacQueen, et al. Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, Vol. 1-14, pp. 281-297, 1967.
- [6] 本間裕大. 施設配置の数理: 種々の最適化視点から見つめる都市 (特集 高校生に伝える or). オペレーションズ・リサーチ, Vol. 60(9), pp. 517-520, 2015.
- [7] 栗田治. 施設配置モデル-社会のための数学の例. オペレーションズ・リサーチ: 経営の科学=Operations research as a management science research, Vol. 41(3), pp. 174-177, 1996.
- [8] Junjie Sun, Yuta Matsushima, and Qiang Ma. Property analysis of stay points for POI recommendation. *DEXA (I)2021*, pp. 374-379, 2021.
- [9] Longzhuang Li, Yi Shang, and Wei Zhang. Improvement of hits-based algorithms on web documents. *WWW2002*, pp. 527-535, 2002.
- [10] Yizhu Shen, Min Ge, Chenyi Zhuang, and Qiang Ma. Sightseeing value estimation by analysing geosocial images. *Int. J. Big Data Intell.*, Vol. 5, No. 1/2, pp. 31-48, 2018.
- [11] Kun Yi, Ryu Yamagishi, Taishan Li, and Qiang Bai, Zhengyang Ma. Recommending pois for tourists by user behavior modeling and pseudo-rating. *arXiv preprint arXiv:2110.06523*, 2021.
- [12] 山岸立, 馬強. 個人適応型観光のためのユーザ体験推薦. 日本データベース学会和文論文誌 Vol.20-J, Article No.15, 2022.
- [13] 岡田和典. ミーンシフトの原理と応用. 情報処理学会研究報告 CVIM, Vol. 162, pp. 401-414, 2008.
- [14] Martin Ester, Hans-Peter Kriegel, Jorg Sander, and Xiaowei Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. *KDD1996*, pp. 226-231, 1996.
- [15] Chen Li, Yuanfu Lu, Wei Wang, Chuan Shi, Ruobing Xie, Haili Yang and Cheng Yang, and Leyu Lin Xu Zhang. Package recommendation with intra-and inter-package attention networks. *SIGIR 2021*, pp. 595-604, 2021.
- [16] Junjie Sun, Chenyi Zhuang, and Qiang Ma. User transition pattern analysis for travel route recommendation. *IEICE Trans. Inf. Syst.*, Vol. 102-D, No. 12, pp. 2472-2484, 2019.
- [17] Xin Lu, Changhu Wang, Jiang-Ming Yang, Yanwei Pang, and Lei Zhang. Photo2trip: generating travel routes from geo-tagged photos for trip planning. *ACMMM2010*, pp. 143-152, 2010.
- [18] Ricardo Gama and Hugo L. Fernandes. A reinforcement learning approach to the orienteering problem with time windows. *Computers and Operations Research 133(C)*, 2021.
- [19] Hidekazu Kasahara, Mikihiko Mori, Masayuki Mukunoki, and Michihiko Minoh. Business model of mobile service for ensuring students' safety both in disaster and non-disaster situations during school trips. *ENTER 2014*, pp. 101-114, 2014.
- [20] Jing Xu, Junjie Sun, Taishan Li, and Qiang Ma. Kyoto sightseeing map 2.0 for user-experience oriented tourism. *MIPR 2021*, pp. 239-242, 2021.