

# 項目反応理論を用いた観光地分類手法の提案

咲山 拓哉 山田敬三 高木正則 佐々木淳

岩手県立大学大学院ソフトウェア情報学研究科 〒020-0693 岩手県滝沢市菓子 152-52

Email: g231n011@s.iwate-pu.ac.jp, {k-yamada, takagi-m, jsasaki}@iwate-pu.ac.jp

**あらまし** 現在観光地の分類方法として統一した見解がないことが問題となっている。そこで項目反応理論を観光履歴に対して適用し、観光地ごとの特性値を推定、内部一貫法の考えをもとにジャンルとしての適応度を計算し、進化計算で分類を発見する。

**キーワード** 項目反応理論, 遺伝的アルゴリズム, SNS, 観光, 分類

## 1. はじめに

情報化技術の発展などから誰もが大量の情報を得ることができるようになった情報過多の現在、欲しい情報が何かわからないという状況や探している情報を見つけ出せないといった状況を打破するため、利用者にとって有用な情報を見つけ出す推薦システムが考案されている[1]。今回は、観光分野の観光地を推薦するものについて考える。

## 2. 観光地の分類

### 2.1 現状

観光資源の定義や分類方法・範囲などについては様々な解釈がなされ見解が統一されていない。この理由としてわが国においては「規範性のある観光資源制度が存在しないため、観光資源の定義に見解が統一されないまま論議が行われてしまう」状態にあることが挙げられている。そして、現在観光資源の分類に統一した見解が求められている[2]。

### 2.2 仮説

本研究では、観光者は観光地の決定の際に、観光者の嗜好よりも観光者の観光に対する何らかの能力が関係しており、この能力に近い人は似たような観光地に行くという仮説を立てる。

### 2.3 用いる理論

仮説に基づき、観光地は観光者の何らかの能力によって行く確率の高い観光地や低い観光地を算出でき、確率的に観光地を分類することができるのではないかと考えた。そこで、本研究では学力のテストで用いられる項目反応理論を観光分野に適応させ、推定される観光地ごとの特性値を用いて観光地の分類を明らかにする。また、その分類を今後の推薦に応用することを考えている。

## 3. 研究の目的

本研究では、観光地の推薦を行う前段階として項目反応理論を用いて現在の統一されていないあいまいな分類とは違う観光地の分類を行うことを研究の目的とする。

### 3.1 既存の観光地推薦システム

既存の観光地推薦システムとして樽井[3]は、協調フィルタリングとコンテンツ分析を利用した研究を行っている。これは、観光地の特徴を最もよく表している観光特性を要素に持

つ観光地特徴ベクトル、利用者履歴ベクトル、利用者特徴ベクトルを利用して推薦を実現している。観光地特徴ベクトルの生成では、観光地の持つ観光特性の要素を wikipedia から引用し試験的に 16 の要素で構成している。このシステムの課題の 1 つとして樽井は、観光地の特徴を表現するための観光特性の要素についての検証を挙げている。このように既存の研究でも観光地の分類（観光特性）を明らかにすることが求められている。

### 3.2 項目反応理論

項目反応理論は、過去の受験者のテストに対する応答のパターンから各項目の特性値および新規の受験者の能力を最尤推定を用いて推定する理論で主に学力テストによく用いられる[4]。また、受験者の能力を逐次推定しながら受験者の能力を動的に短時間で測ろうとする適応型テストも注目されている。

推定される項目の特性値としてよく使われるものが、項目識別度と項目難易度である。項目識別度とは学力テストにおいてその項目がどのくらい受験者の能力を識別する力があるかを表す値である。つまり項目識別度が高い項目は受験者の能力をより明確に分けることができる。項目難易度とは項目の難易度を表すものでこの値が高いほど、能力の低い受験者の正答率は下がるようになっている。

観光の分野に項目反応理論を適用させるために、学力テストにおける問題への応答を、観光地へ行ったことがあるか否かという応答に置き換え、受験者（観光者）が観光地 1 つ 1 つにたいして行ったことがあるか否かの 2 値データを用いることで項目反応理論を実現する。

## 4. 提案

### 4.1 手法の概要

本研究は、観光履歴の 2 値データ（ユーザが観光地に行った or 行っていない）に項目反応理論を用いて、観光地ごとの特性を推定し、その特性値を用いて観光地を分類していくという流れになっている。提案手法利用の流れを図に表したものが図 1 である。本研究では進化計算の 1 つである遺伝的アルゴリズムの考えを用いて、評価関数に分類の理想とされる特性値の状態を設定し、進化計算で分類を発見する。後節からは発見までの流れを詳しく説明する。遺伝的アルゴリズム

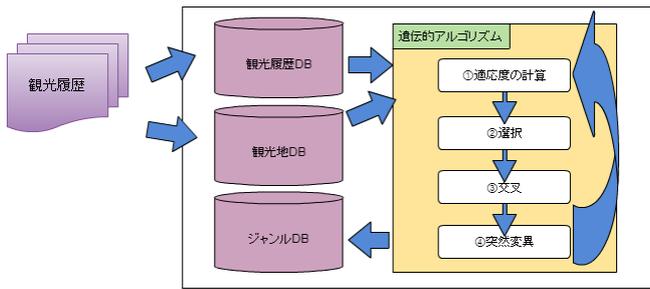


図 1 一点交叉時の配列間の動き

ムの概要は棟朝の著書[5]を参考にする。

### 4.2 遺伝的アルゴリズム

遺伝的アルゴリズムは文字列で表現される集団に対し、遺伝的操作を繰り返して行うことで目的の解に近似な最適解を得ようとするアルゴリズムである。アルゴリズムの手順として主に、

- ① ランダムに初期個体集団を生成
- ② 集団に対して選択を適用し、優れた個体を選ぶ
- ③ 集団内の個体ペアに対して交叉を適用する
- ④ 集団内の個体に対して突然変異を適用する
- ⑤ 停止条件になるまで②-④を繰り返す

となっている。ここでの停止条件は集団の適応度に閾値を決めたり、繰り返す回数（世代数）などを停止条件とする。

### 4.3 事前の準備

観光履歴をもとに計算を行っていくため、事前に観光履歴の2値の行列データを用意することが必要となる。

### 4.4 個体集団の生成、符号化

集団を文字列で表現するため、符号化を行う。今回は観光地の総数と同じ長さの1or0の配列を生成する。配列の中で1は選択された観光地、0は選択されなかった観光地として1つの観光地の集団を表現する。

### 4.5 適応度の計算

本研究では観光リストの信頼性を内部一貫法により測り、適応度とする。すなわち、いくつかの折半についての項目難易度の相関係数の平均を適応度とする。

まず作成された個体の1となっている部分の観光地だけの観光履歴に対して項目反応理論を適用する。これにより、xの集団での観光地ごとの項目識別度と項目難易度が推定される。次に配列内の観光地の集団をxとして、xを

$$x_1 + x_2 = x, \quad ||x_1| - |x_2|| \leq 1 \quad \dots(1)$$

となるような折半を行い、部分集合 $x_1, x_2$ を作成する。Xと同様に、 $x_1, x_2$ に項目反応理論を適用することで $x_1, x_2$ の状態の観光地の特性値が推定される。内部一貫法を用いて部分集合 $x_1, x_2$ での特性値とxの特性値の観光地の個々の相関係数を計算し、平均を求める。内部で一貫の状態（相関係数が高い状態）であれば適応度が高いとし、この適応度をもとに選択を行っていく。特性値は項目難易度や項目識別度が推定されるが、本研究では項目難易度のみを用いる。

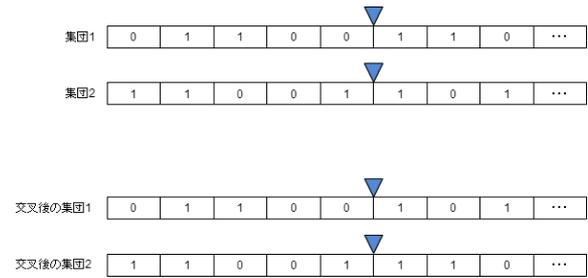


図 2 一点交叉時の配列間の動き

### 4.6 選択

選択は、集団内で適応度が高い個体を次世代へと選択的に残す操作である。選択の手法としてルーレット法やトーナメント法などが提案されている。本研究ではルーレット法を用いて選択を行う。ルーレット法は個体の適応度に応じた確率で次世代の個体を選択するものである。つまりルーレット法を用いると次世代で適応度の高い個体を選択される確率が高くなり、逆に適応度の低い個体は選択される確率が低くなる。

### 4.7 交叉

交叉は、あらかじめ交叉確率を定め、集団内から選ばれた2つの個体の間で部分列を交換する操作である。また、交叉には一点交叉、二点交叉、一樣交叉などが提案されているが本研究では一点交叉を用いる。一点交叉は2つの個体の交叉点を1点だけ決め、交叉点以降の文字列を交換する交叉である。一点交叉を図で表したものを図2に示す。図2では集団1と集団2の交叉点が5文字目と6文字目の間になったと想定し交換を行っている。

### 4.8 突然変異

突然変異は個体を構成する文字の一部を別の文字に変更する操作である。突然変異では、個体を構成するそれぞれの文字について確率を決めそれにより別の文字に変更する。本研究では1or0の文字列のため1となっている部分が突然変異した場合には0に、0となっている部分が突然変異した場合には1に変更する。突然変異を図で表したものを図3に示す。図3では集団1の4文字目が突然変異すると想定し突然変異前と突然変異後の集団1を示している。

## 5. おわりに

本提案では項目反応理論を用いて観光地の特性値を推定し、それを用いて観光地の分類を行うことを提案した。この手法を用いることにより統計的に観光地の分類を行うこと

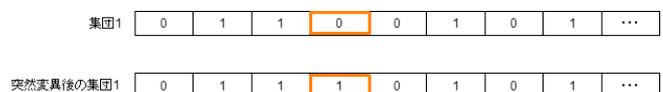


図 3 突然変異時の配列の動き

ができ、観光地推薦に用いることができるのではないかと考える。分類された観光地集団の評価が高いものは内部一貫性が高いものしているが、実際に分類された観光地集団を見てもらい客観的にどう感じるかアンケートを取ることも評価となるのではないかと考える。今後はこの手法で発見された観光地分類を用いることに加えて項目反応理論を用いて観光地を推薦することを試みる。その後、どのようなデータセットならばどのような結果が出るのかなど1つのパターンではなくいろいろなパターンの観光履歴を用いて実験を行いたいと考えている。

## 謝辞

本研究は JSPS 科研費 15K01950 の助成を受けたものです。

## 参考文献

- [1] 神島敏弘, 推薦システム-情報過多をのりきる, 情報の科学と技術 56 巻 10 号, 2006
- [2] 高橋光幸, 観光資源の定義と分類に関する考察, 富山国際大学現代社会学部紀要第 6 巻, 2014
- [3] 樽井勇之, 協調フィルタリングとコンテンツ分析を利用した観光地推薦手法の検討, 上武大学, 上武大学経営情報学部紀要第 36 号, 2011
- [4] 高橋正視, 項目反応理論入門-新しい絶対評価, イデア出版局, 2002
- [5] 棟朝雅晴, 遺伝的アルゴリズム その理論と先端的手法, 森北出版, 2008
- [6] 小松一星, 山田敬三, 高木正則, 佐々木淳, チェックインサービスを利用した観光履歴収集システムの提案, 電子情報通信学会, 2014 年総合大会講演論文集, 2014