

## ID3 を用いた訪問者プロフィールからの観光特性分析手法の提案

手塚 祐樹<sup>†</sup> 小松 一星<sup>‡</sup> 咲山 拓哉<sup>†</sup> 山田 敬三<sup>†</sup> 高木 正則<sup>†</sup> 佐々木 淳<sup>†</sup>  
岩手県立大学ソフトウェア情報学部<sup>†</sup> 岩手県立大学大学院ソフトウェア情報学研究科<sup>‡</sup>

## 1. はじめに

現在、地域活性化施策の一つとして観光が注目されている[1]。一方、旅行者にとって、著名ではない地域の観光地も含む膨大な観光情報の中からその個性に合致した観光先を決定するのは容易ではない。

本稿では、SNS 利用者のプロフィールと訪問先との関係に着目し、ID3(決定木学習アルゴリズム)を用いて観光地毎の訪問者のプロフィールを学習し、プロフィール項目や決定木のエントロピーから観光特性(観光地および旅行者の特性)を分析する手法を提案する。

## 2. 既存研究

樽井[2]は協調フィルタリングを用いた観光地推薦手法の提案をしている。しかし、この手法では、すべてのユーザが行ったことがある観光地を手作業で入力する必要がある。

共著者の小松らは SNS のチェックイン情報を収集し、ユーザが手作業で入力しなくてもチェックイン情報を蓄積し、そのデータを基に観光地を推薦するシステム CheekiTrip[3]を開発している。しかし、このサービスはチェックイン情報を基に協調フィルタリングを用いて推薦を行っており、パーソナライズされた推薦ができないという問題がある。そこで、本研究ではパーソナライズされた推薦を行う方法について提案する。

## 3. 提案

本研究は、観光特性という概念を、観光地の特性(観光地特性)と、観光旅行者の特性(観光旅行者特性)の二つの視点に分け、観光地ごとにどのような観光旅行者特性を持つ人が訪問したかを分析することで、より精度の高い観光推薦の実現を目的としている。

この場合、観光地特性の可視化・数値化が必要となる。本研究では観光地特性を分析する方法として、決定木およびエントロピーを用いる手法を提案し、その手法の有効性を検証する。

決定木を用いる理由としては、観光旅行者にはいくつかの項目(属性)があり、訪問したか否かの特定の項目(クラス属性)の値を決定できるからである。また、その観光地の決定木を構築するアルゴリズムとしては、ID3(決定木学習アルゴリズム)が利用できる。ID3 を用いた場合、各属性の属性値が決定した際の平均情報量の期待値(利得)を求め、その中で利得が最大のものを選び、木のノードにする操作を再帰的にを行い最小の決定木を構築することができる。

## 4. ケーススタディ

### 4.1 概要

本提案の有効性を検証するため、先行研究[4]で収集した CheekiTrip のデータを用いたケーススタディを行った。CheekiTrip は Facebook と Swarm でのチェックイン情報を使用し、協調フィルタリングを用いて観光地推薦をするサービスである。CheekiTrip から得られたデータセットのプロフィール項目を属性にし、スポットに行ったか否かをクラス属性に持つ決定木を観光スポット毎に構築する。そして、構築された決定木を分析し、観光推薦に有意なプロフィールを発見する。

### 4.2 決定木構築システムの開発

観光スポットごとに決定木を迅速に構築するため、Ruby の Gem パッケージ形式で配布されている decisiontree[5]を用い、ID3 アルゴリズムによる決定木構築システムを開発した。

### 4.3 分析方法

データセットのユーザプロフィール情報を属性にし、チェックインの有無をクラス属性として与え、決定木構築システムを用いて決定木を構築する。構築した決定木とその決定木のエントロピーを基に分析を行う。

### 4.4 分析結果

一例として、Motomachi Station(Hyogo)の決定木の一部を図1に示す。この図1のような決

A proposal of analysis method of tourist characteristics from profile information of visitors by using ID3

<sup>†</sup> Yuki TETSUKA, Takuya SAKUYAMA, Keizo YAMADA, Masanori TAKAGI, Jun SASAKI, Faculty of Software & Information Science, Iwate Prefectural University,

<sup>‡</sup> Issei KOMATSU, Graduate School of Software & Information Science, Iwate Prefectural University

定木とその決定木のエントロピーは観光スポットの数だけ作られる。クラス属性値は Motomachi Station (Hyogo) に行ったことがある場合は true, 行ったことがない場合は false となる。

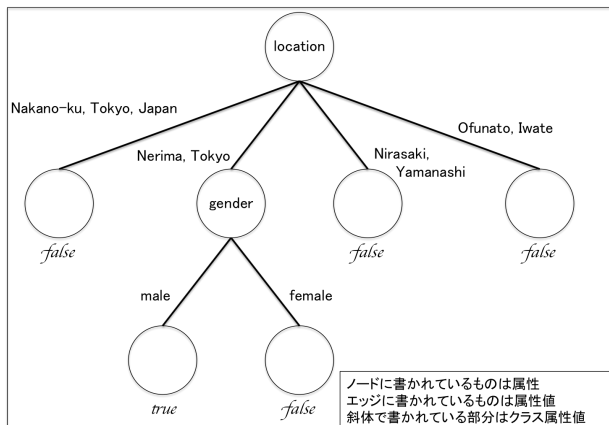


図 1 Motomachi Station (Hyogo) の決定木 (一部)

図 2 に、構築された決定木のエントロピーのヒストグラムを示す。図 2 より、今回のデータでは、エントロピーが低いスポットが多いことが分かった。

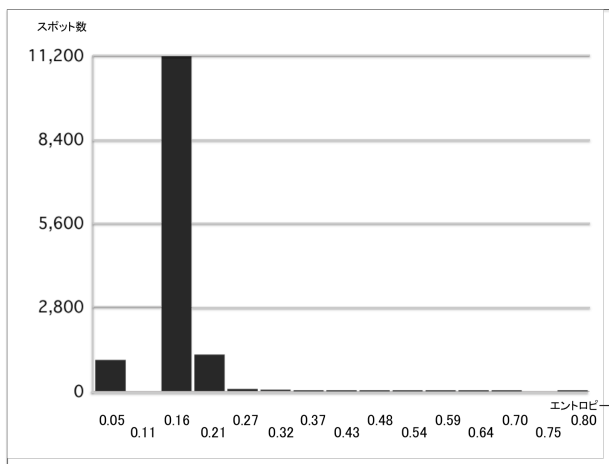


図 1 エントロピーのヒストグラム

表 1 にエントロピーが高い 10 件と低い 10 件のスポットを示す。表 1 より、エントロピーが高いスポットは駅や空港など誰もが行きそうなスポットが多く、エントロピーが低いスポットはあまり知られていないスポットが多い。今回用いたデータで、エントロピーの低いスポットが多いという結果(図 2)から、大衆向けのスポットよりも少数派向け(マニアック)のスポットの方が多いということが分かった。

#### 4.5 考察

本ケーススタディによると、エントロピーが高いスポットを勧める「非個人化推薦」[6](大

衆向けの推薦)と、エントロピーが低いスポットを勧める「個人化推薦」[6](マニアックな推薦)ができる想定される。

表 1 エントロピーの高い 10 件, 低い 10 件

	上位 10 件	下位 10 件
1	東京駅	らめん花月嵐多摩センター店
2	渋谷駅	久屋広場
3	東京国際展示場	赤坂サカス
4	六本木ヒルズ	ラパーク千城台
5	羽田空港	東川町
6	東京ミッドタウン	元風
7	東京スカイツリー	藤沢名物 里のうどん
8	東京国際空港	BISTRO jeujeu
9	岩手県立大学	いとう整形外科
10	成田国際空港	Kuala Lumpur LCCT Departure Lounge

#### 5. おわりに

本稿では、決定木と ID3(決定木学習アルゴリズム)を用いて観光地毎の観光特性(観光地および旅行者の特性の関係)を分析する手法を提案し、ケーススタディを行うことにより観光特性の可視化・数値化が可能であることを示した。このことにより、例えば推薦スポットをエントロピーで指定して提示することでよりユーザーの嗜好に応じた推薦が可能になると考えられる。

今後は、本提案に基づく実用的な推薦システムの開発を進める。

#### 参考文献

- [1] 国土交通省, 観光立国推進基本計画, (2012).
- [2] 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, 上武大学経営情報学部紀要, 第 36 号, pp.1-14, (2011).
- [3] CheekiTrip, <http://cheekitrip.com/>
- [4] 小松 一星ら, 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, 電子情報通信学会総合大会, D-23-4, (2014).
- [5] decisiontree, <https://github.com/igrigorik/decisiontree>
- [6] 神島 敏弘, 推薦システムのアルゴリズム, 人工知能学会誌, 2007 年 11 月号, pp.826-837, (2007).