

# 学習コンテンツ推薦のための ユーザ習熟度とアイテム難易度の推定手法

江渕 弘康 三好 康夫  
Hiroyasu EBUCHI Yasuo MIYOSHI  
高知大学理学部応用理学科情報科学コース

## 1. はじめに

探求学習をするにあたり、知識を身につけるためにどのような情報を読んでいくべきかわからなくなり、学習が行き詰まるときがある。そういったユーザに対して、次に読むべき情報を推薦する情報推薦手法の研究が盛んに行われている。

学術的、専門的、技術的な内容のアイテム(学習コンテンツ)には難易度(アイテム難易度)が存在し、また、ユーザには習熟度(ユーザ習熟度)が存在する。アイテム難易度とは、それを読むにあたりどれほどの前提知識を要するかという指標であり、ユーザ習熟度とは、ある分野についてどれだけ知識を持っているかという指標である。しかし、既存の情報推薦技術はそれらを考慮していない。そこで本研究ではユーザ習熟度とアイテム難易度に基づいた推薦システムの構築を目指す。

## 2. 学習コンテンツを対象とした推薦

### 2.1. 習熟度と難易度に基づく推薦手法

既存の情報推薦技術として、ユーザのアイテムに対する過去の評価に基づく協調フィルタリングがある。本研究では、習熟度と難易度に基づくフィルタリングと、協調フィルタリングを組み合わせることで推薦を行う。

方法としては、協調フィルタリングでユーザにとっておすすめのアイテムを選び出し、それらの難易度と自身の習熟度を比較する。習熟度と難易度の差によって、アイテムの推薦対象からの除外や、おすすめ順位の変動が起こる。

### 2.2. 先行研究における習熟度と難易度の推定手法

入野による先行研究[1]で、ユーザ習熟度とアイテム難易度の推定アルゴリズムが提案された。このアルゴリズムの適用対象は、ユーザとアイテムのつながりを表すネットワークである。アルゴリズムには以下の問題点があることがわかった。

- (1) 習熟度はユーザのアイテムを読んだ数、および難易度はアイテムがユーザに読まれた数の影響が大きい
- (2) 習熟度と難易度の値が等しくない

(1)について、あるユーザが他のユーザと比較してアイテムを多く読んだ場合、その分だけ習熟度が高くなる傾向があった。これは難易度についても同様である。(2)について、例えば習熟度が 0.4 のユーザがいたとすると、難易度が 0.4 のアイテムをそのまま推薦できなかった。

### 2.3. 本研究における習熟度と難易度の推定手法

我々は、2.2 節のアルゴリズムで求めていた値をユーザの中心性、およびアイテムの中心性であると考えた。中心性とは、ネットワークにおける各頂点の重要性の評価のための指標である。

本研究では、ユーザ主流度とアイテム主要度、およびユーザ選別力とアイテム品質の 4 つの新たな尺度を導入する。ユーザ主流度とアイテム主要度は、ユーザの中心性およびアイテムの中心性そのもののことである。ユーザ選別力は、ユーザのアイテムに対する目利き力を表す指標で、アイテム品質は、どれほどアイテムが良質であるかを表す指標である。選別力および品質は、中心性の計算時にパラメータによる制限を導入することで得られる。

習熟度は、選別力から主流度を減算した値を初期値とし、その初期値とユーザが読んだアイテムの難易度の平均から求め

る。難易度は、品質から主要度を減算した値を初期値とし、その初期値とアイテムを読んだユーザの習熟度の平均から求める。

## 3. 動作検証

本研究のアルゴリズムの検証は、2.2 節で挙げた先行研究のアルゴリズムの問題点が改善されたことの確認と、習熟度の精度の考察を目的として行う。データは「はてなブックマーク」で、recommendation タグの付与されたブックマークから形成されるネットワークを使用した。recommendation のネットワークにおけるユーザ数は 370、アイテム数は 944、それらの間をつなげるブックマーク数は 2022 であった。

2.2 節の(1)における、ユーザがアイテムを読んだ数およびアイテムがユーザに読まれた数の影響について、どちらに関してもその影響が小さくなった。(2)の習熟度と難易度の値についても改善が確認できた。

また、習熟度の妥当性を評価するため、理想の推薦とシステムの推薦の比較手法である NDPM 法[2]を用いる。NDPM は [0,1] の範囲の実数値を取り、値が小さいほど推薦精度が高い。被験者として、recommendation のネットワークのユーザからブックマーク数が 5~10 のユーザをランダムに 10 名選び、評価者 3 名(a, b, 筆者)が主観的に習熟度の順位を付けた。表 1 は、それらの順位とその平均順位を、アルゴリズムの主流度、選別力、習熟度それぞれの順位と比較したときの NDPM である。どの評価者の順位と比較しても習熟度の結果が最も妥当であった。

表 1 評価者の順位とアルゴリズムの順位を比較したときの NDPM

比較対象	アルゴリズムの順位		
	主流度	選別力	習熟度
a の順位	0.744	0.633	0.433
b の順位	0.670	0.489	0.352
筆者の順位	0.644	0.690	0.598
平均順位	0.747	0.615	0.374

## 4. おわりに

本稿では、学習コンテンツの推薦におけるユーザ習熟度とアイテム難易度の必要性と、それらの推定手法について述べ、推定アルゴリズムの動作検証を行った。

入野の先行研究で提案されたユーザ習熟度とアイテム難易度の推定アルゴリズムには問題点があることがわかり、その解決のためアルゴリズムを改善した。

今後の課題として、ユーザがアイテムを読んだ日付などの時間情報を利用した、さらなるアルゴリズムの改善などが考えられる。

## 参考文献

- [1] 入野 美弥: "ソーシャルデータに基づく学術書籍の難易度推定アルゴリズム", 平成 21 年度高知大学理学部数理情報科学卒業論文, 2010.
- [2] Y.Y.Yao: "Measuring Retrieval Effectiveness Based on User Preference of Documents", Journal of the American Society for Information Science, Vol.46, No.2, pp.133-145, 1995.