

Aggregate Diversity を考慮した推薦システムに関する研究

情報数理応用研究

5211C026-3 鈴木健史
指導教員 後藤正幸

A Study on Recommender System considering Aggregate Diversity

SUZUKI Takeshi

1 はじめに

近年、EC サイトにおける商品数は膨大な数に上り、それに伴って各ユーザの嗜好に合わせて商品を推薦する機能を持つ推薦システムの重要性が高まってきている。EC サイトにおける推薦システムの購買増への貢献は大きく、より良い推薦システムの導入は EC サイトの経営戦略において重要となっている。推薦システムにはいくつかのアプローチが存在するが、多くの EC サイトでは、導入の簡便性などの理由により、協調フィルタリングに基づくシステムが導入されている。従って、本研究では協調フィルタリングに基づく推薦システムに着目する。

推薦システムは一般に、アクティブユーザの嗜好に適合することに加え、そのユーザにとって未知であるアイテムを推薦することが望まれる [1]。一般に、EC サイトの運営者はユーザがアイテムを既知であるか未知であるかを知ることができないため、推薦システムに関する既存研究の多くは、嗜好の適合するアイテムを予測する手法構築に従事し、その精度を高めてきた。

一方で、既存の推薦手法を用いることで、購買数上位の人気アイテムが集中して推薦されるという問題が存在する。購買数上位のアイテムは、既に広く知られていることが多いため、このようなアイテムのみが推薦されることは望ましくない。そこで、これらの問題を緩和するための手段として、上位アイテムが多く推薦されることを緩和し、代わりにユーザが知っている可能性が低いロングテールアイテム [2] を推薦する方法が考えられる。これは、Adomovicious ら [3] の研究などにおいて、Aggregate Diversity (全体的な多様性) を向上させる分野として注目を集めている。Aggregate Diversity とは、ユーザ全体に推薦したアイテム集合に関する多様性を指す。推薦アイテム集合の Aggregate Diversity が低い状態というのは、ある少数のアイテムのみが集中的に多くのユーザに推薦されている状態を示す。一般に Aggregate Diversity は、システムが推薦したアイテムの集合に対してアイテムユニーク数や Gini 係数などにより測定される。

Aggregate Diversity 向上に関する既存研究は、評価履歴情報をベースにした手法のみが存在する [3]。協調フィルタリングを用いた推薦システムで使用するデータは、主に購買履歴や閲覧履歴などユーザの行動を反映したものと商品やサービスの購入、使用後に与えられる評価履歴などの満足度を反映したものの 2 種類に分けられる。この 2 つのデータの扱いは大きく異なり、それに伴いユーザの嗜好予測モデルも異なる。一般に、多くの EC サイトは評価履歴などの情報を有しておらず情報の量においても不十分な事が多いが、一方で購買履歴や閲覧履歴などのデータは利用できることが多い。しかし、既存研究には購買履歴を扱った研究がなく、本研究において、購買履歴をベースとした手法を提案することは有用である

と考えられる。

以上より本研究では、購買履歴をベースに、嗜好の適合性を考慮することに加え、Aggregate Diversity を向上させる手法の構築を目的とする。構築にあたり、購買履歴をベースとする嗜好予測モデルで代表的な潜在クラスモデルの Aspect Model (以下、AM) [4] を用いる。一般に、協調フィルタリングを用いた推薦手法は主に嗜好推定と推薦リスト作成の 2 段階で構成されているが、他のモデル (LDA やベイズモデルなど) への適用可能性を考え、今回は後者のステップのみに着目し、提案手法を構築する。具体的には、情報検索問題で活用されている語の重み付け手法を援用する。これにより、嗜好の適合度に加え、他ユーザと比較した際の相対的な適合度が高い特徴的なアイテムの推薦を行うことで、本研究が着目した問題を改善することが可能であると考えられる。今回、語の重み付け手法を推薦問題に援用するにあたり、有効性が期待される手法を複数示し、実験的に検討する。本研究では、有効性を検証するための実験を 2 種類用意した。まず、従来の研究に倣い、推薦のベンチマークデータである Movielens data [5] を用いて評価実験を行い、精度と多様性の評価基準のもとで提案手法の有効性を検証する。また、購買数最大化という観点より、市場環境を考慮したシミュレーション実験も併せて行う。これは、未知アイテムの分布状況という市場環境により、未知アイテムの推薦を目的とする手法の有効性が大きく左右されるためであり、EC サイト運営者が手法の採用を検討する際に、より有用な情報を与える。

2 Aspect Model

2.1 Aspect Model

AM はユーザの嗜好とアイテムの特徴を推定するために用いられる確率的潜在クラスモデルである。いま、 I 個のアイテム集合を $\mathcal{X}=\{x_i : 1 \leq i \leq I\}$ とし、 J 人のユーザ集合を $\mathcal{Y}=\{y_j : 1 \leq j \leq J\}$ とする。 (x_i, y_j) はユーザ y_j がアイテム x_i を購入する事象を示し、アイテムとユーザの 2 次元情報のみを扱う。このモデルではユーザとアイテムに対して同一の潜在クラスを仮定し、その集合を $\mathcal{Z}=\{z_k : 1 \leq k \leq K\}$ とする。このモデルは似た嗜好を持つユーザが同じ潜在的なセグメントに所属するという仮定を置いている。また、このモデルは、ユーザとアイテムはただ 1 つの潜在クラスに所属するのではなく、複数のクラスに所属し、嗜好やアイテムの特徴を多面的に表現することが出来る。AM の確率モデルは式 (1) で示される。

$$P(x_i, y_j) = \sum_k P(z_k) P(x_i|z_k) P(y_j|z_k) \quad (1)$$

また、式 (1) の各パラメータ $P(z_k)$, $P(x_i|z_k)$, $P(y_j|z_k)$ は多項分布に従い、それぞれ EM アルゴリズムで推定す

ることができる。

2.2 Aspect Model を用いた推薦方法

ここでは、AM を用いた際の一般的な推薦手順を述べる。まず、同時確率の推定値 $\hat{P}(x_i, y_j)$ を EM アルゴリズムで推定したパラメータ $\hat{P}(z_k), \hat{P}(x_i|z_k), \hat{P}(y_j|z_k)$ を用いて求める。式 (2) によりユーザ y_j のアイテム x_i に対する嗜好適合度の推定値 $\hat{P}(x_i|y_j)$ を求める。

$$\hat{P}(x_i|y_j) = \frac{\hat{P}(x_i, y_j)}{\sum_{x_i} \hat{P}(x_i, y_j)} \quad (2)$$

次に嗜好適合度の推定値 $\hat{P}(x_i|y_j)$ を降順にランキングし、指定数のアイテムを上位から順に推薦する。

3 提案手法

3.1 本研究のアプローチ

本研究では、ユーザのアイテムに対する嗜好の適合性を加味しつつ、Aggregate Diversity を向上させる手法の提案を目的とする。また、前述の通り、嗜好推定後の推薦リスト作成段階にのみ着目し、式 (2) を改善した手法の提案を行う。

従来手法においては、式 (2) の嗜好の適合度合いのみを基準に推薦を行うと、購買数上位のアイテムが多くのユーザに推薦される傾向が強い。これは、多くのユーザの購買数上位アイテムに対する嗜好適合度が、ロングテールアイテムに対するものより高いことを意味する。そのため、この問題を解決する 1 つの考え方として、従来における嗜好の適合度の値に加え、他のユーザの適合度と比較した際の相対的な適合度の高低を加味することが考えられる。これにより、アクティブユーザにとって、他ユーザと比べて相対的に嗜好適合度が低い購買数上位アイテムは推薦がされにくくなる。その一方で、そのユーザに対して、相対的に嗜好適合度が高いロングテールアイテムの推薦が促進される。

他のユーザと比較して相対的に高い適合度を持つアイテムに重み付けをする手法は、情報検索分野で一般に用いられる単語の重み付け手法の考え方と類似しており、これらのアイデアを本研究に援用できると考えられる。これらの文書解析や情報検索を対象とした手法では、Probability Weighted Amount of Information (PWI) [6] などの様々な手法が提案されている。これらの語の重み付け手法は、基本的には「単語の文書内での重要度」に加え、「他の文書と比較した際の、単語の相対的な重要度」を考慮することで構成される。

本研究では、これらの 2 要素を推薦システムへ置き換えたものを、それぞれ絶対適合度と相対適合度とする。前者はあるユーザのアイテムに対する絶対的な嗜好適合度の高さとし、後者を他の全ユーザと比較したときの相対的な嗜好適合度の高さとして定義し、この 2 要素により提案手法を構成し、推薦問題に適用する。

3.2 提案手法

本研究では、文書検索問題に用いられる語の重み付け手法を推薦システムに援用する。語の重み付け手法は文書問題に対してヒューリスティックに提案されていることが多く、システムへの適用にはさらに適した方法がある

と考えられる。本研究では、アクティブユーザ y_j に対するアイテム x_i の評価の際に、絶対適合度と相対適合度の双方を組みこむ評価式を一般的な形で提案する。

絶対適合度を $A(x_i, y_j)$ と相対適合度を $R(x_i, y_j)$ とする。本研究では、推薦すべきアイテムを評価するための新たな尺度として、以下の積モデルと和モデルを提案する。

・積モデル

$$A(x_i, y_j) \times R(x_i, y_j)$$

・和モデル

$$A(x_i, y_j) + R(x_i, y_j)$$

$A(x_i, y_j)$ としては、嗜好の適合度である $\hat{P}(x_i|y_j)$ を用いることが自然である。一方、 $R(x_i, y_j)$ に関しては、アイテム x_i の推定生起確率 $\hat{P}(x_i)$ に対する $\hat{P}(x_i|y_j)$ の相対的な高低を相対適合度とする。 $\hat{P}(x_i)$ は全体的な人気度を表しており、これと比較して $\hat{P}(x_i|y_j)$ が高い場合、ユーザ y_j は他のユーザよりも特にアイテム x_i を好んでいることを意味する。このような相対適合度を表す評価式 $R(x_i, y_j)$ としては、 $\frac{\hat{P}(x_i|y_j)}{\hat{P}(x_i)}$ と $\hat{P}(y_j|x_i) - \hat{P}(x_i)$ といった候補が考えられる。さらに、語の重み付け手法では対数を用いた $\log \frac{\hat{P}(x_i|y_j)}{\hat{P}(x_i)}$ も有用とされていることから、これを相対適合度を表す式として採用可能である。

以上より、提案手法は以下の手法を考えることとする。

$$PWI(x_i, y_j) = \hat{P}(x_i|y_j) \times \log \frac{\hat{P}(x_i|y_j)}{\hat{P}(x_i)} \quad (3)$$

$$WL(x_i, y_j) = \hat{P}(x_i|y_j) \times \frac{\hat{P}(x_i|y_j)}{\hat{P}(x_i)} \quad (4)$$

$$MS(x_i, y_j) = \hat{P}(x_i|y_j) \times (\hat{P}(y_j|x_i) - \hat{P}(x_i)) \quad (5)$$

$$AL(x_i, y_j) = \hat{P}(y_j|x_i) + \log \frac{\hat{P}(x_i|y_j)}{\hat{P}(x_i)} \quad (6)$$

$$AD(x_i, y_j) = \hat{P}(x_i, y_j) + \frac{\hat{P}(x_i, y_j)}{\hat{P}(x_i)} \quad (7)$$

$$AS(x_i, y_j) = \hat{P}(x_i, y_j) + (\hat{P}(x_i|y_j) - \hat{P}(x_i)) \quad (8)$$

式 (3) は、確率で表現された語の重み付け手法である PWI 手法 [6] に相当する。その他の手法は、式 (4) を PWI without Logarithm (WL)、式 (5) を Multiplied Subtraction (MS)、式 (6) を Additional Logarithm (AL)、式 (7) を Additional Division (AD)、式 (8) を Additional Subtraction (AS) とする。

4 実験

以下では、ベンチマークデータを使用した評価実験と市場環境を再現したシミュレーション実験を行い、提案手法の有効性を検証する。前者は、ベンチマークデータにおいて、従来手法と提案手法を精度と多様性の観点より有効性を検証する。後者は、購買数という観点より、市場環境を変化させた場合の提案手法の有効性を検証する。

4.1 ベンチマークデータを用いた評価実験

4.1.1 実験条件

提案手法の有効性を示すため、Movielens データセット [5] を用いて評価実験を行う。ユーザ数 $J = 943$ 、映画数 $I = 1682$ 、総データ数 10 万件であり、訓練データ 8 万件とテストデータ 2 万件に分けられている。ユーザはすべ

てのアイテムのうち最低 20 件を購買している．潜在変数のクラス数 K は，従来手法 [4] と同様に 10 とした．このモデルでは，各パラメータ $P(z_k)$, $P(x_i|z_k)$, $P(y_j|z_k)$ は EM アルゴリズムによって推定されるが，初期値に結果が依存するため，これらをランダムに与えた実験を 10 回繰り返した平均を結果とする．推薦するアイテム数 N を 10, 20, 30 件と変化させた結果を示す．

4.1.2 評価指標

本研究は，Aggregate Diversity を向上させる一方で嗜好に合致するアイテムを推薦する手法の構築を目的としており，手法を評価するため精度と多様性 [3] という 2 つの評価指標を導入する．精度はユーザの嗜好に適合したアイテムをどれだけ推薦できたかどうかの指標であり，式 (9) により算出される．

$$Precision = \frac{G}{NJ} \quad (9)$$

ここで， N は各ユーザに対する推薦アイテム数であり， G はテストデータに存在した推薦アイテムの数である．また，多様性に関する尺度は，全アイテム種類数のうちどれだけの種類のアイテムを推薦できたかを示す多様性指標を使用し，以下の式 (10) で定義される．

$$Diversity\ in\ Top\ N = \frac{\#\{\cup_{y_j \in Y} L_N(y_j)\}}{J} \quad (10)$$

$L_N(y_j)$ はユーザ y_j に推薦したアイテムの集合であり， $\#$ は集合の要素数を示す．この指標により，推薦システムの Aggregate Diversity がどの程度を評価することができる．高い値を示すほど，Aggregate Diversity が向上したことを示す．

4.1.3 実験結果

表 1. AM と提案手法の実験結果

	Top10		Top20		Top30	
	Pre	Div	Pre	Div	Pre	div
AM	0.310	0.259	0.252	0.342	0.217	0.392
PWI	0.232	0.376	0.195	0.488	0.170	0.564
WL	0.227	0.381	0.198	0.508	0.175	0.584
MS	0.275	0.316	0.226	0.404	0.196	0.477
AL	0.016	0.196	0.025	0.376	0.031	0.505
AD	0.015	0.197	0.024	0.376	0.030	0.505
AS	0.287	0.301	0.236	0.390	0.205	0.459

表 1 は，推薦アイテム数を変化させたときの従来手法と提案手法の精度と多様性についての結果である．式 (2) の従来手法を AM と表記する．すべての Top N に対し，精度では AM が最も高い結果となり，提案手法 AS と MS は 0.02 から 0.03 ポイントほど AM より劣る結果となった．一方，多様性に関しては，ほとんどの提案手法は AM より優れた結果となった．提案手法の中では，特に WL が最も多様性が向上し， $N = 30$ において AM より 0.28 ポイントの上昇が見られた．結果より， $N = 10$ における AD と AL 以外は，精度と引き換えに多様性を向上させることが出来ている．

4.1.4 考察

AD 手法と AL 手法の精度が極端に低い結果となった．これは相対適合度がアイテムの推定生起確率に対しての倍数であるのに対して，絶対適合度が 0 から 1 の値を取り，双方を和で取ることで，相対適合度の値に極端に影響を受ける式となるためだと考えられる．そのため，嗜好適合性がほとんど加味されず，精度が極端に低下したと考えられる．モデルでは，加算する項のバランスが取れるように，加重和を取り，その重みパラメータをうまく調節するなどの改良が必要であると考えられる．

また，実験結果より，嗜好適合度の高いアイテムを推薦するということがと既知アイテムの推薦を避けることの間にはトレードオフの関係が存在することが分かる．すなわち，既知アイテムを避けるということは，同時に嗜好の適合度が高いアイテムも避けて推薦を行うことになる．映画作品を例とすると，映画「タイタニック」は，非常に多くのユーザに好まれている作品（多くのユーザに対して嗜好の適合度が高い作品）であるが，同時に非常に多くのユーザはこのアイテムを知っている．

4.2 市場環境を考慮したシミュレーション実験

提案手法は，ユーザの嗜好に適合し，かつ未知であるアイテムをより多く推薦する手法になっているが，手法の有用性は EC サイトの市場環境により大きく変化すると考えられる．市場環境とは，この場合，主にユーザが既知であるアイテムの多寡，及び分布状況を示すが，EC サイトが対象とするアイテムの特性（書籍，映画，音楽など）により変化する．例えば，大手メディア媒体でのプロモーション活動が行われる映画市場においては，既知アイテムが購買数上位アイテムに偏って多いことが想定される一方で，書籍市場は映画市場より既知アイテムが比較的少なく，より分散された分布であることが想定される．

そのため，EC サイトの運営者は，市場環境を見極め，嗜好の適合度と多様性のバランスをとることにより，購買数を最大化する手法を採用する必要がある．そこで，購買数最大化という観点により，未知・既知情報を考慮したシミュレーション実験を行うことで，どのような市場環境において，どの手法が購買数を最も向上させることができるかについて検討する．本実験により，EC サイト運営者は手法を採用する際に，より有用な手掛かりを得ることができる．

4.2.1 シミュレーションモデル概要

シミュレーションのモデルを以下に示す．
STEP1) データ生成フェイズ：ユーザの購買確率を AM に基づき決定し，その確率に従い購買履歴データを作成する．具体的には $P(z_k)$, $P(x_i|z_k)$, $P(y_j|z_k)$ をランダムに設定し，その上で式 (1) により $P(x_i, y_j)$ を算出しこれを真の購買確率とする．Movielens に倣い，全アイテム数 I のうち購買済みアイテムの割合が 0.05 となるように，各ユーザは真の購買確率に従い購買し，これを購買履歴データとする．購買データ作成後，ユーザの未購買アイテムに対する既知未知情報データの作成を行う．また，作成した購買履歴データを下に，以下の式で，アイテム x_i の全ユーザに対する被購買数 B_i ，ユーザ数 I として既知確率 $P_{know}(x_i)$ を以下で算出する．

$$P_{know}(x_i) = \frac{(\frac{B_i}{T})^\alpha}{\sum_i (\frac{B_i}{T})^\alpha}. \quad (11)$$

確率 $P_{know}(x_i)$ に従い既知割合 D になるように既知未知情報データを作成する．その際にパラメータ α を変動させることで，既知アイテムの偏りを変えることができる． α の値を 1 より大きくするほど，購買履歴における購買数分布と比較して，購買数上位アイテムに既知アイテムが偏った市場を想定できる．

STEP2) 推薦フェイズ: 作成した購買履歴データのみを入力とし，従来手法と各提案手法により任意の N 件のアイテムをユーザに推薦する．ユーザがそのアイテムを未知だった場合，推薦されることで既知の状態に変化する．すでに既知状態のアイテムはそのままとなる．

STEP3) 購買フェイズ: ユーザの既知未知情報データと真購買確率 $P(x_i, y_j)$ に基づき，各ユーザの購買シミュレーションを行う．具体的には，ユーザは既知アイテムに対してのみ STEP1 で定めた真の購買確率に従って，購買する．未知アイテムの購買確率は 0 である．

4.2.2 パラメータ条件

シミュレーション条件は以下のものに行う．

表 2:シミュレーション条件

変数	意味	値
J	ユーザ数	500
I	アイテム数	1000
N	推薦アイテム数	10, 20, 30
K	潜在クラス数	10
D	既知割合	0.05, 0.1
α	パラメータ	2

今回， $\alpha = 2$ とし，購買数上位に既知アイテムが偏る市場を想定した．また，実験の対象とする手法は，従来手法に加え，先の実験において精度が極端に低かった AL 手法と AD 手法を除く 4 つの提案手法とし，各手法を比較することで有効性の検証を行う．さらに，STEP3 の購買フェイズにおける購買数の観点より，各手法の評価を行う．

4.2.3 実験結果

既知割合が $D = 0.05, 0.1$ の場合の結果を以下に示す．

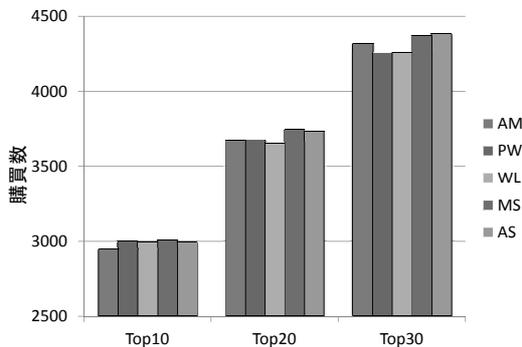


図 1. $D = 0.05$ における AM と提案手法の実験結果

図 1 より，多様性を考慮した各提案手法が従来手法を上回る結果となった．図 1 において，Top10 と 20 においては MS 手法，Top30 においては AS が，それぞれ最も購買数を増加させる手法になっている．また，Top20 における WL，Top20, 30 における WL と PWI が従来手法より劣る結果となった．

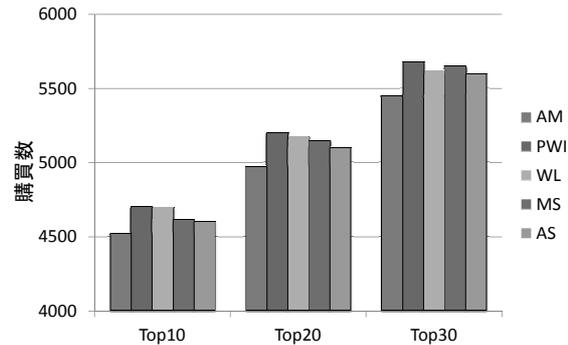


図 2. $D = 0.1$ における AM と提案手法の実験結果

図 2 より，多様性を考慮した各提案手法が従来手法を上回る結果となった．また，図 2 においては Top10, 20, 30 においては PWI が，最も購買数を増加させる手法になっている．

4.2.4 考察

既知割合が 0.05, 0.1 の双方において，従来手法よりも，提案手法が優れている結果となった．これは，人気アイテムに偏っているアイテムを推薦する手法より，各ユーザの嗜好に適合した未知のアイテムを推薦することで総購買数が向上したためと考えられる．また，アイテムに対する既知の割合が高くなるほど，従来手法より提案手法が大幅に上回ることが，図 1 と図 2 より分かる．これは，購買数上位アイテムに対して既知のユーザがより多くなることで，それらのアイテムをより多く推薦する従来手法が購買数増加に貢献できなくなった結果と考えられる．一方で，既知割合が低い市場においては，多様性の比較的高い手法は購買数上位アイテムの推薦を控えていしまい購買数が従来より劣ることが分かる．

5 まとめと今後の課題

本研究では，嗜好に適合するアイテムの推薦に加え，Aggregate Diversity を向上させる手法をいくつか提案した．実データによる実験により，精度と多様性におけるトレードオフが見られた．また，市場環境を考慮した購買シミュレーション実験により，様々な市場条件において提案手法の有効性を確認した．今後の課題として，精度と多様性を双方向向上させる手法の追求とより様々な市場環境を設定し手法の有効性を示していく．

参考文献

- [1] T. Kamishima, "Algorithms for Recommender System(1)," *Journal of Japanese Society for Artificial Intelligence*, vol. 22, No. 6, pp.826–837, 2007.
- [2] C. Anderson, "The Long Tail," *New York: Hyperion*, 2006.
- [3] G. Adomavicius, Y. Kwon, "Improving Aggregate Recommendation Diversity Using Ranking-Based Techniques," *IEEE Transactions on Knowledge and Data Engineering. Forthcoming*, 2011.
- [4] T. Hofmann, and J. Puzicha, "Latent Class Models for Collaborative Filtering," *Proc. 16th International Joint Conference on Artificial Intelligence*, pp. 688–693, 1999.
- [5] MovieLens. "http://www.movielens.org/".
- [6] A. Aizawa, "An information-theoretic perspective of tf-idf measures," *Information Processing and Management*, Vol. 39, pp. 45–65, 2003.