

クロスメディア型レコメンデーション の提案と評価

Proposal and Evaluation of Cross Media Recommendation

柳原 正[♡]帆足 啓一郎[♡] 小野 智弘[◇]松本 一則[♡] 滝嶋 康弘[◇]Tadashi YANAGIHARA,
Keiichiro HOASHI, Chihiro ONO,
Kazunori MATSUMOTO, Yasuhiro TAKISHIMA

本論文では、「本」や「CD」などの商品種別（以下、メディア）に関する履歴情報を分析し、異なるメディアの履歴情報を相互に関連付けることで、少ない履歴情報からでも適切な商品を推薦する手法を提案する。具体的には、推薦対象となるメディアが与えられたとき、他メディアの履歴情報を参照して、メディア間で同時に出現する共通ユーザに着目し、共通ユーザの数が最大となる他メディアの履歴情報を推薦対象となる履歴情報と組み合わせた上で商品を推薦するクロスメディア型レコメンデーションを提案する。評価実験では、推薦対象となるメディアを共通ユーザの数が多く他メディアを参照することで、精度が向上することを確認した。また、従来手法との比較においても、本手法の方が推薦結果の精度が高いことを確認した。

In this paper, we propose a recommendation method called Cross Media Recommendation, which analyzes the item's media (such as information on whether the item is a book or CD) and associate an item's media with another media to provide accurate recommendations of items belonging to a media with scarce history information. In detail, when given a specific media to be recommended, Cross Media Recommendation refers to other media and focuses on common users, which are users who co-appear between the two media, and combines the media containing the most number of common users in order to create more accurate recommendation. We have confirmed from experiment results that by selecting media with the most number of common user increase the recommendation accuracy the most, and that our proposal outperforms collaborative filtering.

1. はじめに

近年、インターネットではユーザが過去に購入した商品に関する購入履歴を用いて、ユーザの嗜好に合致する商品を自動的に推薦する機能を提供するウェブサイトが普及してい

[♡] 正会員 株式会社 KDDI 研究所
{td-yanagihara, hoashi, matsu}@kddilabs.jp
[◇] 非正会員 株式会社 KDDI 研究所
{ono, takisima}@kddilabs.jp

る。例として、インターネット上で本を中心とした商品を推薦するウェブサイトが挙げられる。このように、ユーザが利用した商品(以降、「アイテム」と呼ぶ)の購入履歴(以降、「履歴情報」と呼ぶ)をもとに、ユーザの嗜好を推測した上でアイテムを自動的に推薦するシステムをレコメンデーションシステムと呼ぶ。レコメンデーションシステムで扱う商品の種類(以降、「メディア」と呼ぶ)は本やCDのように複数存在することが多い。

レコメンデーションシステムの多くは協調フィルタリング[1]と呼ばれる推薦方式に基づく。複数のメディアに関するアイテムを協調フィルタリングで推薦する方法として、メディアごとに協調フィルタリングで推薦する場合と、全アイテムのメディアを区別せずに協調フィルタリングで推薦する場合が考えられる。しかし、いずれの方法も履歴情報が少ないメディアを推薦した場合には、推薦結果の精度が低下してしまう。前者の場合、ユーザの嗜好を推測するために十分な履歴情報が存在しないため、推薦結果の精度が低下する。後者の場合、履歴情報が少ないメディアの嗜好よりも、履歴情報が多い他メディアの嗜好が反映されやすくなり、推薦したいメディアのユーザの嗜好が推薦結果に反映されにくい。このように、従来手法である協調フィルタリングでは、メディアが複数存在した場合、履歴情報が少ないメディアの推薦に対応できない。

本論文では、この問題を解決するために、推薦対象となるメディアが与えられたとき、他メディアの履歴情報を参照して、メディア間で同時に出現する共通ユーザに着目し、共通ユーザの数が最大となる他メディアの履歴情報も参照して、メディア間で同時に出現する共通ユーザに着目し、共通ユーザの数が最大となる履歴情報と組み合わせた上で推薦するクロスメディア型レコメンデーションを提案する。これにより、履歴情報が少ないCDというメディアを推薦するとき、より多くの履歴情報が存在する本というメディアに含まれる履歴情報と組み合わせることによって、CDの推薦精度が向上する。また、映画の視聴履歴を用いた評価実験を通して、メディアが複数存在したときの推薦結果の精度を計測し、協調フィルタリングの推薦結果の精度と比較する。以降、2章では、関連研究について取り上げ、3章では提案内容について述べる。4章では、提案手法の性能評価と考察を行い、5章でまとめを述べる。

2. 関連研究

関連研究として、推薦結果の精度を向上するためにアイテムに付与された属性情報を活用する手法が挙げられる。具体的には、アイテムとアイテムに付与された属性情報に対し、two way aspect model を適用することで、属性情報の有効性を検証し、有効な場合にはその履歴情報に基づいて推薦する研究事例[2]である。文献[2]では、ユーザが商品を購入したり、サービスを利用したりする際に、その行動を決定する潜在変数が商品やサービス内に存在すると仮定していることを特徴とする。評価実験では映画の視聴履歴を対象に、映画の主演が映画の視聴を決定する要因となった可能性について検証した後、履歴情報がない映画を与えられた際に、映画に付与された主演に反応を示す可能性が高いユーザに対して推薦する手法を取る。これにより、履歴情報が全くない商品やサービスについて、高精度の推薦結果が生成可能となる。しかし、このときに利用する要因は一つ(この場合では主演)のみであると仮定しており、要因が複数存在した場合には、

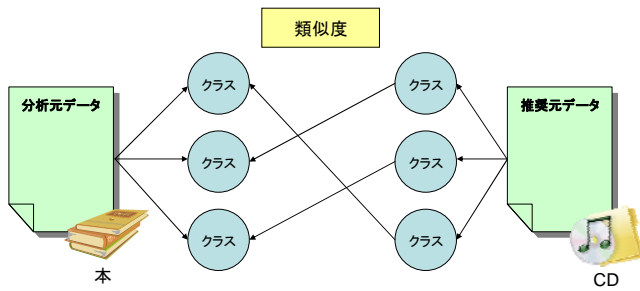


図 1 クロスメディア型レコメンデーションにおけるユーザ群の相関関係

Fig.1 System Configuration of Cross Media Recommendation

どの要因を選択すべきかについての指標が示されていない。他に、文献[2]が参照している手法[3]では、同様に潜在変数が存在すると仮定し、pLSA[4]を用いて潜在変数を推測することにより、疎なデータにおいても高精度な推薦を可能としている。また、協調フィルタリングで使用するユーザ・アイテムの行列に対してクラスターリング手法を適用することで、ユーザやアイテムの特徴がより強調されるため、疎性による精度の低下を防ぎ、推薦結果の精度が高められることも報告されている[5][6]。しかし、これらの手法ではメディアの情報を問わず、履歴情報をすべて同一のメディアであると仮定した上でクラスターリングを行っており、従来手法である協調フィルタリングと同様、メディアが複数存在した場合、履歴情報が少ないメディアの推薦に対応できない。

これらに加え、ユーザ間の関連度とアイテム間の関連度が事前に与えられた上で、ユーザ・ユーザ間の行列及びアイテム・アイテム間の行列に対してクラスターリングを行い、形成されたユーザのクラスとアイテムのクラスの間で推薦する事例[7]も存在する。しかし、推薦する上で有益であるアイテムのクラスが取得できない場合は適用が難しい。本提案ではこの手法と異なり、ユーザのクラスの間で推薦する。

3. 提案手法

ここでは、本論文で提案するクロスメディア型レコメンデーションについて説明する。

3.1. 概要

従来の協調フィルタリングにおいて、推薦対象となるメディアが複数存在したとき、履歴情報が少ないメディアの推薦結果が低下する問題を示した。本論文では、この問題を解決するために、推薦対象となるメディアが与えられたとき、他メディアの履歴情報を参照して、メディア間で同時に出現する共通ユーザに着目し、共通ユーザの数が最大となる他メディアの履歴情報も参照して、メディア間で同時に出現する共通ユーザに着目し、共通ユーザの数が最大となる履歴情報と組み合わせた上で推薦するクロスメディア型レコメンデーションを提案する。

クロスメディア型レコメンデーションでは、メディアごとのユーザが商品を購入する際に、その行動を決定する未知の要因である潜在変数が存在すると仮定し、同一の潜在変数によって商品を購入したユーザは嗜好が似ていると見なし、クラスとして抽出する。クラスが抽出されたあと、推薦対象と

なるメディアに含まれる各クラスに対し、他メディアの一つのクラスを参照することで、選択されたクラス内で人気があるアイテムを推薦する手法である。

以下に、「本」を購入したことがある人に対し、「CD」を推薦する例とし、動作手順を説明する。(図 1 参照)

- 1) 「本」を分析元データと定義し、「CD」を推奨元データと定義する。これらのデータをもとにクラス抽出用データとして利用する。
- 2) 分析元データと推奨元データに対し、クラスターリングなどによって、嗜好が似ていると仮定したユーザの集合であるクラスを抽出する。
- 3) 分析元データの各クラスに対し、推奨元データの各クラスへの関連度を求める。その後、分析元データの各クラスに対し、最も類似度が高い推奨元データのクラスを参照し、推奨元データのクラスにおいて最も利用されているアイテムを推薦する。

3.2. 手法詳細

ここでは本手法の動作手順に関する詳細説明を行う。

クラス抽出用データの生成

本提案では 2 種類のデータを必要とする。推薦対象となるメディアに関する履歴情報である「推奨元データ」と、推薦精度を向上するために用いる他メディアの履歴情報である「分析元データ」である。上記の例では、「本」を分析元データ、「CD」を推奨元データとして定義する。

クラス抽出用のクラスターリング

分析元データおよび推奨元データのそれぞれで嗜好が似たユーザを集合として抽出する。具体的には、分析元データに含まれる履歴情報に対し、クラスターリング手法を適用することで、分析元データに出現するユーザをクラスターリングする。本論文で紹介する実装例では、潜在クラスモデルに基づいた分析手法[8]を使うことで、嗜好が似たユーザを潜在クラスと呼ばれるユーザの集合として分類する。

潜在クラスモデルとは、ユーザの行動において、観測可能な要因である顕在変数以外に、観測できない要因である潜在変数が存在すると仮定し、同じ潜在変数によって行動を起こしたユーザの集合を潜在クラスとして表現するモデルである。潜在変数は観測不可能であるため、直接求めることができないが、以下の手順で示す潜在クラスの抽出方法[10]により、推測が可能となる。

- 1) ユーザ i がアイテム j を購入するとき、ユーザ・アイテム行列内の i 行 j 列の値 X_{ij} を 1 とし、購入していなければ X_{ij} を 0 とする。
- 2) 潜在クラス T のクラス数 t を決定する。これは AIC や BIC に基づくクラス数数のモデル検定や、 t の値を変更し、精度が最もよくなる最適値を発見することで定められる。
- 3) 潜在クラスモデルで利用するパラメータを EM アルゴリズム[9]を使って推定する。
- 4) ユーザ i が各潜在クラスに対する帰属確率が最も高い潜在クラス T_x を求め、ユーザ i は潜在クラス T_x に帰属すると見なす。

なお、EM アルゴリズムは値の最尤推定を行う際に利用する初期値次第で局所解に陥る可能性が高いため、局所解に陥りにくい確定的アンニリングを用いたEM アルゴリズム [11]を用いることで、より精度が高い推薦結果が得られることが予備実験[12]にて示されている。

クラスごとの関連度の計測

本提案では、推奨元データに含まれるアイテムを推薦する際に、推奨元データのそれぞれの潜在クラスに対し、最も関連がある分析元データの潜在クラスに含まれるユーザへ推薦する。このとき、分析元データをメディア A、推奨元データをメディア B としたとき、メディア A に含まれる潜在クラス C_{a1}, C_{a2}, \dots に対し、メディア B に含まれる潜在クラス C_{b1}, C_{b2}, \dots の中から、最も関連度が高い潜在クラス C_r を求めなければならない。そこで、潜在クラス間の類似度は以下のように求める。

- 1) メディア A の潜在クラスを C_{a1}, C_{a2}, \dots とする。メディア B の潜在クラスを C_{b1}, C_{b2}, \dots とする。このとき、メディア A 及びメディア B の両メディア内に履歴情報を持つユーザを共通ユーザ u と定義する。
- 2) 任意の潜在クラス C_{ax} に含まれるユーザに対し、メディア B 内のアイテムを推薦する際に、潜在クラス C_{by} に含まれる共通ユーザ u_x に着目する。
- 3) メディア A とメディア B の履歴情報のどちらにも同時に出現する共通ユーザ u_x はメディア A 内の各潜在クラス C_{ax} に対する帰属確率および、各メディア B 内の各潜在クラス C_{by} に対する帰属確率を共に持つため、潜在クラス C_{ax} における共通ユーザ u_x による t 個のクラスに対する帰属確率を次元の値としたベクトルと潜在クラス C_{by} における共通ユーザ u_x による t 個のクラスに対する帰属確率のベクトルが生成できる。
- 4) 潜在クラス C_{ax} における共通ユーザ u_x による帰属確率のベクトルと潜在クラス C_{by} における共通ユーザ u_x による帰属確率のベクトルの類似度は、ピアソン相関係数やコサイン類似度などの公式で計算する。

以上により、メディア A 内の潜在クラス C_a とメディア B 内の潜在クラス C_b 間の各潜在クラスの類似度が求まる。あとは、潜在クラス C_a 内の各潜在クラスに含まれるユーザに対し、それらのユーザが帰属する潜在クラス C_a と最も関連度が高い特定の潜在クラス C_b を選択し、その潜在クラス C_b 内で履歴情報の件数が最も多いアイテムから順に推薦する。

4. 評価実験

提案手法の有効性を評価するため、協調フィルタリングの実装例である Item based top- n Recommendation [13] と比較実験を行った。

4.1. 実験に用いるデータ

GroupLens [15] プロジェクトにて公開されている映画を鑑賞したユーザの 100,000 件分の履歴情報である movielens [16] の 100,000 data set を用いることで、映画のジャンル情報を「メディア」とし、映画を「アイテム」とし

て見立てる。データに含まれるジャンル情報は 19 種類であるが、「unknown」のジャンルに該当する映画は 1 件のみであったため、メディアとしては無視した。

評価データの生成方法

各手法の入力となる評価データの生成方法は文献 [13] で提示された手順に従う。

- 1) 分析元データである履歴情報を D_A 、推奨元データの履歴情報を D_R とする。このとき、 D_A と D_R で共に履歴情報を持つ共通ユーザを U_{AR} とする。
- 2) D_A および D_R に含まれる全ユーザが視聴した映画に対する評価として、1 - 5 の範囲における評価値が含まれている。(1 が「とても嫌い」、2 が「嫌い」、3 が「どちらでもない」、4 が「好き」、5 が「とても好き」) このとき、それぞれの評価値に対し、推薦すべき映画(4 および 5 の評価された映画)を 1 とし、推薦すべきではない映画(1, 2, および 3 と評価された映画)を 0 に置き換える。
- 3) D_R 内の共通ユーザ U_{AR} が持つ履歴情報に対し、1 の評価値として置き換えられた履歴情報を 1 つ選択し、削除する。このとき、削除した履歴情報を D_E とし、残りの履歴情報を D_R とする。
- 4) 文献 [13] の手順に従ったままでは、 D_A 及び D_R の両方の履歴情報を使って、協調フィルタリングで学習を行う場合、評価対象のデータに同じアイテムが含まれる可能性がある。これを避けるため、 D_A 及び D_R の両方の履歴情報に含まれるアイテムについては評価対象から外す。

推薦結果の生成方法

クロスメディア型レコメンデーションでは、分析元データおよび推奨元データからそれぞれ潜在クラスを抽出する際に、クラス数 t をそれぞれ $t = 10$ として抽出した。履歴情報 D_A と履歴情報 D_R から本提案もしくは協調フィルタリングにより推薦結果を計算した後、出力された推薦結果と D_E に含まれる評価値を比較する。例えば、特定の共通ユーザ U_x に対し、アイテム i を推薦すべきであると評価されたとき、共通ユーザ U_x が実際にアイテム i を利用したことがあり、且つ U_x のアイテム i に対する評価値が D_E に含まれていれば、正解とする。

精度の計算方法

精度の計算方法として、文献 [13] に掲載された精度の計算方法を用いる。具体的には、推薦結果と削除対象の値を比較し、値が一致した正解の件数 (Number of hits) を推薦結果に含まれる全件数 (n) で割った値を精度 (hit-rate) として評価する。

$$hit - rate(HR) = \frac{Number\ of\ hits}{n} \dots (1)$$

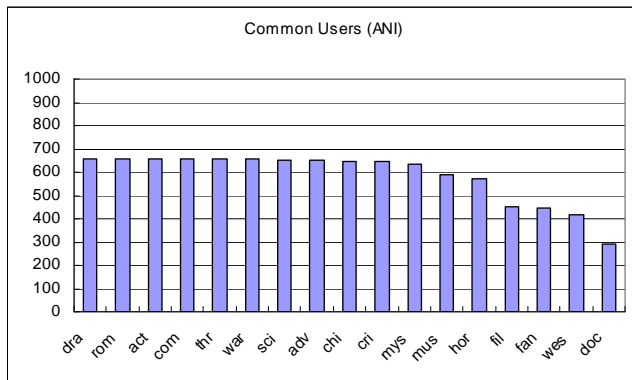


図 2 “Animation”を推奨元データとしたとき、その他のメディアを参照したときの共通ユーザ数

Fig. 2 Number of Common Users of each Media when Associated with “Animation”

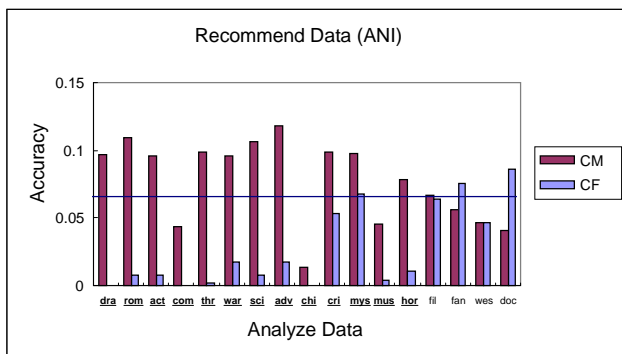


図 3 “Animation”の推薦を行うときの精度

Fig. 3 Recommendation Accuracy of “Animation”

表 1 グラフ内の名称に関する表

Tab 1. Table for Meaning of Abbreviations

本文内の名称	グラフ内の名称	本文内の名称	グラフ内の名称
クロスメディア型レコメンデーション	CM	Drama	DRA
協調フィルタリング	CF	Romance	ROM
共通ユーザ	Common Users	Action	ACT
精度	Accuracy	Comedy	COM
推奨基データ	Recommend Data	Thriller	THR
分析基データ	Analyze Data	War	WAR
		Science	SCI
		Adventure	ADV
		Children	CHI
		Mystery	MYS
		Musical	MUS
		Horror	HOR
		FilmNoir	FIL
		Fantasy	FAN
		Western	WES
		Documentary	DOC

4.2. 評価結果と考察

クロスメディア型レコメンデーション(以下, CM と略す)および協調フィルタリング(以下, CF と略す)において, 履歴情報が少ないメディアで推薦結果を計算する場合, どのメデ

ィアを分析元データとして選択すべきかについて検証する. このため, 履歴情報が少ない分析元データ, 平均的である分析元データ, および履歴情報が多い分析元データを用いた場合の精度を測定した. そこで検証には, 以下の3つの分析元データを用いて検証した.

- ・ 履歴情報の件数が少ないメディアを推薦する場合: Animation (ANI)
- ・ 履歴情報の件数が平均的なメディアを推薦する場合: Children (CHI)
- ・ 履歴情報の件数が多いメディアを推薦する場合: Comedy (COM)

履歴情報の件数が少ないメディアの推薦

分析元データとなりえる18種類のメディアを履歴情報の件数が比較的少ない「Animation (ANI)」と参照したときの共通ユーザの数を図 2 に示す. 「Animation (ANI)」の履歴情報の件数は計3605件であった. また, 図 3 に分析元データとなりえる18種類のメディアを用い, 履歴情報の件数が比較的少ない「Animation (ANI)」の映画を推薦するときの精度を示している. 図 3 内の横軸では, 各分析元データの履歴情報と「Animation (ANI)」の履歴情報の両方を参照した際に出現する共通ユーザの多い順にメディアを左から右へ並べられている. また, 共通ユーザが最も多くなるジャンルは「Drama (DRA)」, 次に「Romance (ROM)」, 次に「Action (ACT)」, といった順に並べられる. 棒グラフ内のうち, 左側はクロスメディア型レコメンデーション(CM), 右側は協調フィルタリング(CF)で計算した推薦結果の精度を表す. グラフ内の横線はCFを使い, 同一のメディア内で推薦したときの精度を表す. (図面の都合上, 図内で略称などを用いた. 略称の意味については, 表 1を参照)

図 3 に着目すると, CM は「Fantasy (FAN)」, 「Western (WES)」, 「Documentary (DOC)」を除くジャンルではCF より精度が高いことが分かる. 共通ユーザが最も多いメディアである「Drama (DRA)」を参照したときの精度(0.096)をCFで推薦したときの精度 (0.065) と比べて, 精度向上率が約 1.47 (= 0.096 / 0.065) であった. また, CFの精度が最大となる「Documentary (DOC)」を分析元データとして選択したときの精度(0.085)と比べて, 精度向上率が1.12 (= 0.096 / 0.085) であった.

なお, 「Comedy (COM)」および「Children (CHI)」の精度がCM及びCF共に精度が低い要因としては, これらのジャンルに所属するアイテムは「Animation (ANI)」というジャンルにも所属するものが多く, 4.1. の4) において, 評価対象とならなかったことが考えられる. これらのことから, 履歴情報が少ないメディアを推薦する場合, CFでは履歴情報が多いメディアの履歴情報を用いると精度が低下してしまうのに対し, CMではメディア別に特徴を取り出しているため, 推薦結果の精度が向上できると言える. これにより, 履歴情報が豊富な分析元データを使い, 履歴情報が少ないデータを推薦するとき, CF よりもCMが適切であるといえる.

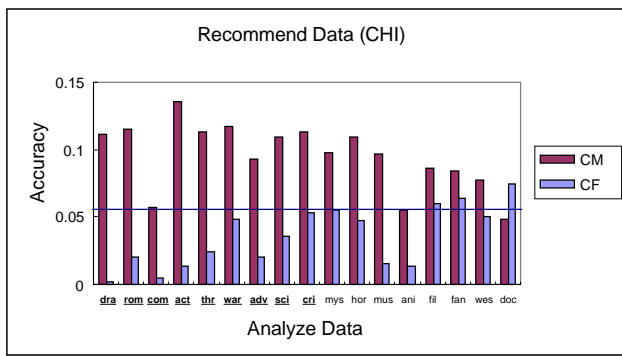


図 4 “Children”の推薦を行うときの精度

Fig. 4 Recommendation Accuracy of “Children”

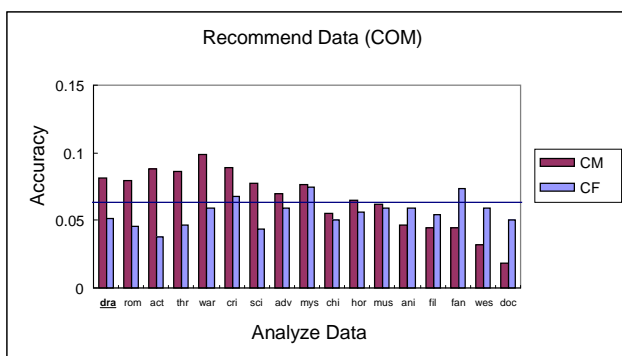


図 5 “Comedy”の推薦を行うときの精度

Fig. 5 Recommendation Accuracy of “Comedy”

履歴情報の件数が平均的なメディアの推薦

図 4 では、18 種類のメディアのうち、履歴数が中間の 9 位である「Children (CHI)」のジャンルに該当する映画を推薦する場合、分析元データとして特定の他メディアの履歴情報を利用したときの精度を表す。横軸では、各分析元データの履歴情報を参照した際に表示する共通ユーザの多い順にメディアを左から右へ並べられている。

ここでは、CM の精度は Documentary を除くジャンルにおいては CF の精度よりも高いことが分かった。また、「Comedy (COM)」を除く、共通ユーザが最も多いメディアの分析元データを参照したときに最も良い精度が得られる。共通ユーザが最も多いメディアである「Drama (DRA)」を参照したときの精度(0.111)を CF で推薦したときの精度(0.055) と比べて、精度向上率が 2.01 (= 0.111 / 0.055) であった。また、CF の精度が最大となる「Documentary (DOC)」を分析元データとして選択したときの精度 (0.074) と比べて、精度向上率が 1.5 (= 0.111 / 0.074)であった。

履歴情報の件数が多いメディアの推薦

図 5 では、18 種類のメディアのうち、履歴情報が多いジャンルのうち、「Comedy (COM)」に該当する映画を推薦する場合、他メディアの履歴情報を分析元データとして利用したときの精度を表す。横軸では、各分析元データの履歴情報を参照した際に表示する共通ユーザの多い順にメディアを左から右へ並べられている。

この場合では、共通ユーザが最も少ない組み合わせである

4 種のジャンルを除いて、CM の精度が CF よりも高いことが分かる。

共通ユーザが最も多いメディアである「Drama (DRA)」を参照したときの精度 (0.081) を CF で推薦したときの精度 (0.062) と比べて、精度向上率が 1.30 (= 0.081 / 0.062) であった。また、CF の精度が最大となる「Mystery (MYS)」を分析元データとして選択したときの精度 (0.074) と比べても、精度向上率が 1.09 (= 0.081 / 0.074)であった。

以上により、履歴情報が少ないメディアを推薦する際に、より履歴情報が多いメディアの履歴情報を利用することで、より高精度な推薦が生成できることが分かった。また、推奨元データに対し、共通ユーザが最も多くなる分析元データを参照することで精度が最も向上することも分かった。そして、履歴情報が多いメディアを用いて、履歴情報が少ないメディアを推薦する場合、従来の協調フィルタリングと比べ、クロスメディア型レコメンデーションはより精度が高いことが分かった。これにより、クロスメディア型レコメンデーションの有効性を示すことができた。

5. 今後の課題

今後の課題として、以下の 3 項目を挙げる。

- 今回の評価実験では、一般に公開されている movielens のデータを利用し、異なるジャンルを異なるメディアと見立てて評価した。ところが、ジャンル間の共通ユーザが数多く存在しており、必ずしも嗜好の差がある異なるメディアを再現できているとは言えない。今後は他のメディアのデータを入手し、有効性の評価を実施したい。
- 本稿では、共通ユーザ数が多いメディアを組み合わせることで推薦精度が向上することを示した。今後はメディア間の共通ユーザの割合と精度向上率との定量的な関係を抽出することにより、さらなる精度向上を目指したい。
- 評価データには時刻情報が付与されているが、今回は利用していない。評価の順序関係などを利用するとさらに精度が向上できる可能性がある。

6. まとめ

本論文では、推薦対象となるメディアが与えられたとき、他メディアの履歴情報を参照して、メディア間で同時に出現する共通ユーザに着目し、共通ユーザの数が最大となる他メディアの履歴情報も参照して、メディア間で同時に出現する共通ユーザに着目し、共通ユーザの数が最大となる履歴情報と組み合わせる上で推薦するクロスメディア型レコメンデーションを提案した。評価実験では、従来のレコメンデーションシステムで複数のメディアから構成される履歴情報を推薦対象としたときに、履歴情報が少ないメディアを推薦する場合では、協調フィルタリングのときと比べ、精度向上率が 1.12 であることを示した。さらに、任意の推奨元データに対し、共通ユーザが最も多くなる分析元データを参照することで精度が向上することを確認した。

[謝辞]

日頃ご指導いただくKDDI研究所の秋葉重幸所長，松本修一副所長，および菅谷史昭執行役員に深く感謝致します。

[文献]

- [1] P. Resnick and H. Varian. "Recommender Systems." *Communications of the ACM*, Vol.40, No.3, pp. 56–58, 1997.
- [2] A. Schein, R. Popescul, L. Ungar and D. Pennock. "Methods and Metrics for Cold-start Recommendations." In *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 253–260. ACM Press, 2002.
- [3] T. Hofmann and J. Puzicha. "Latent Class Models for Collaborative Filtering." In *Proceedings of the 16th International Joint Conference on Artificial Intelligence*, pp. 688–693, 1999.
- [4] T. Hofmann. "Probabilistic Latent Semantic Indexing." In *Proceedings of the 22nd International Conference on Research and Development in Information Retrieval (SIGIR'99)*, pp. 50–57, 1999.
- [5] M. O'Connor and J. Herlocker. "Clustering items for Collaborative Filtering." In *Proceedings of the Workshop on Recommender Systems: Algorithms and Evaluation (Part of SIGIR'99)*, 1999.
- [6] L. Ungar and D. Foster. "Clustering Methods for Collaborative Filtering." In *Proceedings of AAAI Workshop on Recommendation Systems*, 1998.
- [7] H. Chen, K. Furuse, N. Ohbo and S. Nishihara. "Improving Collaborative Filtering by W-Clustering." In *Proceedings of the IASTED Int'l. Conf. on Databases and Applications*, pp 81–86. , 2004.
- [8] P.F. Lazarsfeld. "The Logical and Mathematical Foundations of Latent Structure Analysis." *Measurement and Prediction*, 1950.
- [9] A.P. Dempster, N.M. Laird and D.B. Rubin. "Maximum Likelihood from Incomplete Data via the EM Algorithm." *J. R. Statistical Society*, Vol. Series B, No.39, pp. 1–38, 1977.
- [10] 岡太彬訓, 木島正明, 守口剛. マーケティングの数理モデル. 朝倉書店, 2001.
- [11] 上田修功, 中野良平. "確定的アニーリング EM アルゴリズム." 電子情報通信学会論文誌 D-II, Vol. J80-D-2, No.1, pp. 267–276, 1997.
- [12] 柳原正, 帆足啓一郎, 松本一則, 菅谷史昭. "確定的アニーリング EM アルゴリズムを用いたクロスメディアレコメンデーション方式の提案." 第 5 回情報科学技術フォーラム (FIT2006), No.2, pp. 75–76, 2006.
- [13] M. Deshpande and G. Karypis, "Item-based top-*n* Recommendation Algorithms". *ACM Transactions on Information Systems (TOIS)*, Vol. 22, No. 1, pp. 143-177, 2004.
- [14] J. Breese, D. Heckerman and C. Kadie, "Empirical Analysis of Predictive Algorithms for Collaborative Filtering". In *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence*, pp. 43-52, 1998.
- [15] GroupLens.
<http://www.cs.umn.edu/Research/GroupLens>
- [16] MovieLens. <http://movielens.umn.edu>

柳原 正 Tadashi YANAGIHARA

2002年 慶大・環境情報卒. 2004年 同大大学院修士課程了.
2005年 KDDI(株)入社. 現在, (株)KDDI 研究所にて, レコメンデーションシステム, テキストマイニング等の研究に従事. 2007年度 FIT ヤングリサーチャー賞受賞. 日本データベース学会会員, 情報処理学会会員, 電子情報通信学会会員, ACM 会員.

帆足 啓一郎 Keiichiro HOASHI

1995年 早稲田大学・理工学部・情報科卒. 1997年 同大大学院修士課程了. 同年国際電信電話(株)(現・KDDI 株式会社)入社. 現在, (株)KDDI 研究所にて, マルチメディア情報検索, 情報フィルタリング, レコメンデーションシステム等の研究に従事. 2004年度 FIT ヤングリサーチャー賞受賞. 日本データベース学会会員, 情報処理学会会員, ACM 会員. 工博.

小野 智弘 Chihiro ONO

1992年 慶大・理工学部 電気工学科 卒. 1994年 同大大学院修士課程了. 同年国際電信電話(株)(現・KDDI (株))入社. 1999年~2000年 スタンフォード大学電気工学科客員研究員. 現在, (株)KDDI 研究所にて, データマイニング, レコメンデーションシステムの研究に従事. 1996年度情報処理学会学術奨励賞受賞. 電子情報通信学会会員. 工博.

松本 一則 Kazunori MATSUMOTO

1984年 京大・工・情報工学卒. 1986 同大大学院修士課程了. 同年国際電信電話(株)(現・KDDI (株))入社. 現在, (株)KDDI 研究所にて, マルチメディア検索, コンテンツ配信の研究開発に従事. 人工知能学会 1998年度研究奨励賞, 電子情報通信学会 2000年度論文賞, FIT2003 船井ベストペーパー賞を受賞. 日本データベース学会会員, 電子情報通信学会会員.

滝嶋 康弘 Yasuhiro TAKISHIMA

1986年 東大・工・電気卒. 1988年 同大大学院修士課程了. 同年国際電信電話(株)(現・KDDI (株))入社. 現在, (株)KDDI 研究所 知能メディア グループリーダー. この間, 動画の符号化方式, 動画通信システム, 情報理論の研究・開発に従事. 電子情報通信学会, 映像情報メディア学会, 画像電子学会会員. 工博.