

研究論文

行動履歴と嗜好に基づく グループ向けコンテンツ推薦手法の提案

瀬古 俊一^{1,a)} 八木 貴史^{2,b)} 茂木 学^{1,c)} 武藤 伸洋^{1,d)} 小林 透^{1,e)}

受付日 2012年4月11日, 採録日 2012年8月31日

概要: 本論文は, 家族や友人, 恋人といった互いを知っているグループに対して適したコンテンツを推薦するアルゴリズムを提案する. 推薦精度向上のために, グループメンバー内の誰の嗜好が重視されるかなどのコンテンツ選択傾向に着目した. そこで, 本論文では個人個人のコンテンツに対する嗜好と, グループでの行動履歴を用いてグループメンバー間のパワーバランスを推測し, そのパワーバランスに基づいて推薦スコアを算出する手法について提案する. TV 番組を対象にした推薦精度の検証実験を行った結果, 提案手法は推薦精度を向上させ, 興味があるが意外・知らなかったコンテンツを推薦するのに有益であることを示した.

キーワード: コンテンツレコメンデーション, グループレコメンデーション, レコメンダシステム, 嗜好推定

Content Recommendation for Known Group Based on Behavioral History and Individual Preference

SHUNICHI SEKO^{1,a)} TAKASHI YAGI^{2,b)} MANABU MOTEGI^{1,c)} SHINYO MUTO^{1,d)}
TORU KOBAYASHI^{1,e)}

Received: April 11, 2012, Accepted: August 31, 2012

Abstract: This paper proposes an algorithm to estimate the appropriate content for groups of people who know each other. In order to achieve high recommendation accuracy, we focus on the group characterize, which is the content selecting tendency of the group. Our algorithm estimates the group characterize and calculates recommendation scores based on the Power Balance Map using individual preference for genres and shared history. We verify that the proposed algorithm recommends appropriate content for groups. Evaluation results show that our proposal improves recommendation accuracy and may help to find Novel Content.

Keywords: content recommendation, group recommendation, recommender system, preference estimation

1. はじめに

近年, ビデオホスティングサービスやウェブサービスの台頭により世界中からアクセス可能なコンテンツの数が

増大している. それにより, 利用者は興味のあるコンテンツを見つけ出すことが困難になってきている. このような背景から, 様々な分野で利用者に適したコンテンツを推薦するレコメンダアルゴリズムに関する研究がなされており [1], 利用者の行動履歴を用いたレコメンダ技術の研究もさかんに行われている. たとえば, Amazon.com [2] では利用者の購買履歴をもとに商品を推薦する技術をシステムに組み込んだサービスを行っている. また, Tezuka ら [3] は操作履歴を利用した飲食店のレコメンダサービスに関してフィールド実験を行っており, Nakamura ら [4] は視聴履歴や訪れた場所などの記録から TV 番組を推薦する手法を提

¹ NTT サービスエボリューション研究所
NTT Service Evolution Laboratories, NTT Corporation,
Yokosuka, Kanagawa 239-0847, Japan

² NTT レゾナント株式会社
NTT Resonant Inc., Minato, Tokyo 108-0023, Japan

a) seko.shunichi@lab.ntt.co.jp

b) t-yagi@nttr.co.jp

c) motegi.manabu@lab.ntt.co.jp

d) muto.shinyo@lab.ntt.co.jp

e) kobayashi.toru@lab.ntt.co.jp

案している。このように、レコメンド技術は様々なサービスに適用可能であり、非常に注目されている研究分野である。しかしながら、現在さかんに行われているレコメンド技術は個人を対象にしたものが大多数である。レコメンド技術は個人向けのみだけでなく、グループ向けにも適用可能となるべきである。たとえば映像コンテンツであれば、個人1人だけで視聴するのみでなく、家族や友達といったグループでも視聴する利用シーンが数多く存在する。このようなグループ向けのレコメンデーションは個人向けより複雑であり、個人向けと同じ手法では推薦困難な場合がある。たとえば、個人向けに映像コンテンツを推薦する場合は、その視聴者だけの好みに基づいて行えばよい。しかし、夫婦のような2人組で映像コンテンツを視聴する場合、互いの好みをどのように考慮すれば最適な推薦が可能かは一様に定まらない。この問題を解決するために本論文では、グループに適したコンテンツを推薦可能なレコメンド手法を提案する。

以下に本論文の構成を記す。2章ではグループ向けレコメンド（グループレコメンデーション）分野における既存のアプローチやアルゴリズムについて述べる。3章ではグループに適したコンテンツの推薦手法を提案する。4章ではTV番組を推薦コンテンツの対象として提案手法の評価や考察を行う。最後に5章では本論文の結論を述べる。

2. 先行研究

グループレコメンデーションのアプローチは、以下の2つに大きく分類される。

アプローチ1 仮想個人化法

アプローチ2 統合法

各アプローチの内容および課題について以下に記す。

2.1 仮想個人化法

仮想個人化法は、グループを「仮想的な1人の利用者」と見なすアプローチである。グループを1人の利用者とするにより、個人向けレコメンド技術を利用してグループへの推薦が可能となる。たとえば夫婦の場合、夫婦2人で一緒に体験したコンテンツ（視聴した映像コンテンツなど）の履歴を蓄積していき、その行動履歴に基づいて推薦を行う。仮想的に1人の利用者で見なすため、既存の個人向けレコメンド技術を利用可能という利点がある。しかし、Yuらはグループの行動履歴は個人の行動履歴よりも収集可能な機会が少なく、長期間収集しなければ良い推薦が行えないという欠点を指摘している [5]。

2.2 統合法

統合法は、個々人のプロフィール（映像コンテンツや映像ジャンルに対する嗜好など）や個々人の推薦結果を統合してグループ向けの推薦を行う手法である。本論文では

個々人のプロフィールを統合する手法をプロフィール統合法、個々人の推薦結果を統合する手法を推薦結果統合法と呼ぶ。どちらの統合法においても個人向けのデータのみで推薦可能であるため、グループの行動履歴は不要といった利点がある。しかしながら、Yuらはプロフィール統合法と推薦結果統合法を比較し、プロフィール統合法の方が優れた結果を導き出せることを述べている [5]。また、Berkovskyらによる実験においてもプロフィール統合法が優れていると結論付けられている [11]。したがって、統合法では個々人のプロフィールを統合して推薦を行う手法の方が妥当なアプローチであるといえる。

個々人のプロフィールの最適な統合手法について様々な研究者が提案を行っている。Masthoffはグループの満足度を向上させるには、グループの各メンバが好まないものを除去してから統合することが効果的であると主張している [6]。この主張はO'Connerらの実験で検証されており、その実験では単純にグループメンバの嗜好スコアを平均するより効果的であると述べている [7]。Jamesonらは各メンバの嗜好スコアの平均が高く、分散が小さいコンテンツを推薦するのが良いと主張している [8]。Goren-Barら [9]は視聴時間帯に基づいた重みを用いた加重平均手法を、Yuら [5]やShinら [10]は各メンバの嗜好スコアの分散具合に基づいた重みを用いた加重平均による手法を提案している。

しかしながら、上述したプロフィール統合法では、各手法の特徴とグループの特徴とが合致しなければ効果が薄いという欠点がある。この欠点を視覚的に説明するために、各手法の特徴を表したグラフを図1に示す。各グラフはユーザAとユーザBの2人組であった場合を例としており、横軸はユーザAの嗜好スコアを、縦軸はユーザBの嗜好スコアを示している。また、塗りつぶされたエリアは

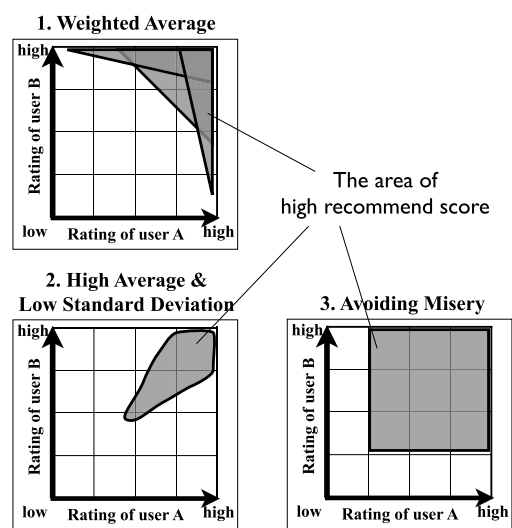


図1 各統合手法の特徴を個々人のプロフィール（嗜好スコア）との関係で表したグラフ

Fig. 1 These graphs show the relation between aggregate methods and individual profiles.

グループに適しているコンテンツがプロットされる領域を示している。図1のグラフ1は、加重値によって推薦順位が高くなるエリアが異なるため複数の領域が描画されており、図の例はユーザAもしくはユーザBに重みをおいた場合と重みが均等の場合という3つの例を表している。図1より、統合法は個々人のプロフィールを統合する手法によって推薦される結果が異なることが分かる。これらのどの手法が最適であるかは、グループの関係性や特徴などによって異なると考えられる。Soteloらはグループの関係性に着目し、グループのメンバ個々人のプロフィールが類似している場合と異なっている場合とで適用する統合アルゴリズムを切り替える手法を提案している [12]。しかし、グループの関係性や特徴はこのような2種類のみで分類可能なものではなく、もっと多様で複雑である。したがって、既存の手法はグループの多種多様な関係性や特徴に対して柔軟に対応していないため、これらを考慮した手法が必要となる。

3. 提案手法

3.1 アプローチ

前章では、仮想個人化手法ではグループの行動履歴が多量に必要であるという欠点があり、統合法ではグループの多種多様な関係性や特徴に対して柔軟な対応ができないという欠点を述べた。そこで本手法ではこれらの欠点を解決するために、Power Balance Map という考えを用いた推薦手法を提案する。Power Balance Map とは、各メンバの嗜好スコア（たとえば映像コンテンツや、その映像コンテンツに紐づいている映像ジャンルに対する好みの度合いを定量化したもの）を軸とした空間上に、グループで一緒に体験したコンテンツ（たとえば視聴した映像コンテンツ）をプロットした分布図と定義する。図2は、ユーザAとユーザBからなるグループにおけるPower Balance Mapの例を示した図である。横軸はユーザAのコンテンツに対する嗜好スコアを、縦軸はユーザBの嗜好スコアを示しており、嗜好スコアがHigh方向にあればあるほど好みが強くなることを表している。菱形の点はこのグラフ上にプロットされた2人での行動履歴（たとえば一緒に視聴した映像コンテンツ）1つ1つを示している。なお、コンテンツをこのグラフ上にプロットするために各メンバの嗜好スコアを成分としたベクトルを、本論文では嗜好ベクトルと呼ぶ。グループの行動履歴をこのような分布図（Power Balance Map）で表すことにより、そのグループの関係性や特徴が推測可能といえる。たとえば、図2中のMap1では、ユーザBよりユーザAが好きなコンテンツがよく体験されているということから、ユーザAの嗜好が優先される関係であると推測される。反対にMap2では、ユーザBの方が優先されていると見なせる。Map3ではユーザAとユーザBともに同程度の嗜好を持つコンテ

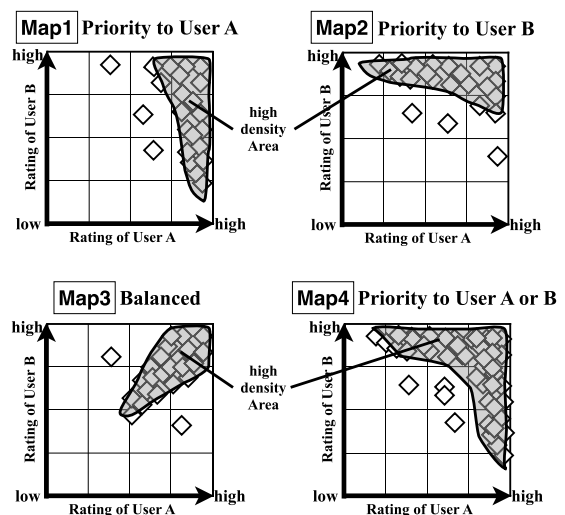


図2 ユーザAとユーザBからなるグループのPower Balance Mapの例

Fig. 2 Power Balance Map examples in case of user A and user B.

ンツをよく体験していることから、どちらかの嗜好に偏らせるよりは、同程度好きなものを選択する傾向を持つグループであると推測される。Map4は、少なくとも片方が好きなコンテンツをよく体験していることから、互いの嗜好が普通程度のコンテンツよりも、必ずどちらかが好きなコンテンツを選択する傾向を持つグループであると見なせる。このように、Power Balance Mapの高密度領域に着目することで、「恋人という関係だから女性の嗜好に偏らせる」といった固定的なルールベースではなく、カップルAは女性寄り、カップルBは男性寄り、カップルCは嫌いなものでなければ相手が好きなものにあわせるなど、同一のデモグラフィック属性であっても個々のグループの多様な関係性や特徴を考慮可能となる。したがって、このPower Balance Map上の高密度領域に属するあるいは近いコンテンツを推薦することで、グループにとって満足度の高い推薦が可能になる。さらに、この高密度領域を利用することで、たとえグループの行動履歴に現れていないコンテンツでもこのPower Balance Map上にマッピングできれば、グループの関係性や特徴に基づいて推薦スコアが算出可能になり、仮想個人化法よりも少ないグループの行動履歴で推薦精度を高めることも可能になる。

3.2 実現方法

図3はPower Balance Mapを用いたコンテンツ推薦システムの流れ図を示したものである。前提として、グループを構成する各メンバのメタデータ（映像ジャンルなど）に対する個々人の嗜好スコア（Rating List of each Member）とグループでの行動履歴（Behavioral History of Group）は既存の手法を用いて取得されているものとする。個々人の嗜好スコアは、たとえば、事前のプロファイル登録の一環

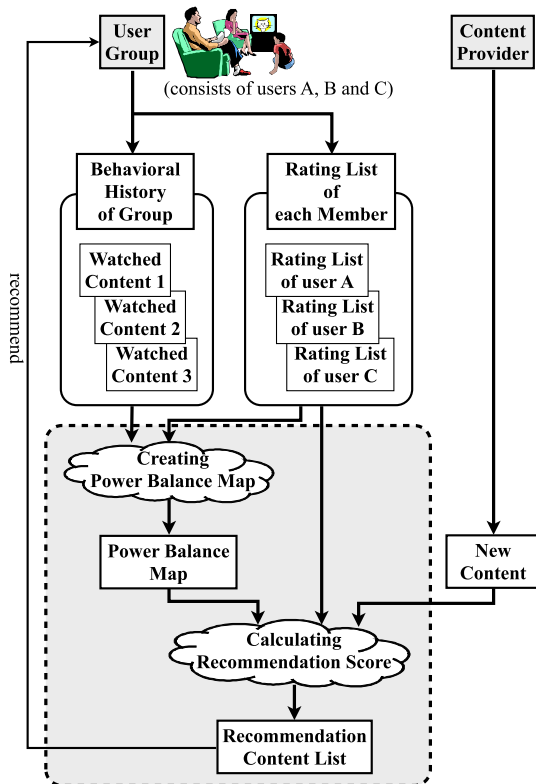


図 3 提案アプローチを用いた推薦システムの流れ図
 Fig. 3 System flow of our recommendation strategy.

として GUI などを用いて (各ユーザに尋ねて) 収集するか、個々人の行動履歴を用いて個人の嗜好スコアを推薦するアルゴリズム [3] を用いて算出する。グループでの行動履歴も同様に、GUI などを用いてユーザが記録することで収集する。提案手法は、コンテンツに対する個々人の嗜好スコアとグループでの行動履歴から Power Balance Map を作成し、それを用いて推薦対象コンテンツ (New Content) が対象グループに適しているか否かの度合いを示すレコメンドスコアを算出する (Calculating Recommendation Score)。このレコメンドスコアに基づいて順序付けられたコンテンツの推薦リスト (Recommendation Content List) を作成し、その順序が高いものから推薦を行う。以下に提案手法の詳細を、映像コンテンツ推薦を例にとって説明する。

まずはじめに、Power Balance Map を作成するために、グループでの行動履歴 (視聴済み映像コンテンツの集合) 内にある各コンテンツの嗜好ベクトルを算出する。今回前提として与えられている個々人の嗜好スコアは映像コンテンツに対するものではなく映像ジャンル (SF, アクションなど) に対するものであるため、まずは映像ジャンルに対する嗜好スコアを用いて視聴済み映像コンテンツに対する個々人の嗜好スコアの算出を行う。もしコンテンツが2つ以上のジャンルを持っている場合、各ジャンルに対する嗜好スコアの平均値がそのコンテンツに対する嗜好スコアとなる。対象とするコンテンツを c 、そのコンテンツに紐付けられているジャンル集合を G 、対象とするユーザを m 、

表 1 ジャンルに対する嗜好スコアの例
 Table 1 A sample of rating table for genre.

Genre	User A Rating	User B Rating
Action	5	2
SF	5	1
Comedy	3	3

そのユーザのジャンル g に対する嗜好スコアを $r_{m,g}$ とすると、あるユーザのあるコンテンツに対する個々人の嗜好スコア $u_{m,c}$ は式 (1) となる。

$$u_{m,c} = \frac{1}{|G|} \sum_{g \in G} r_{m,g} \quad (1)$$

続いて、式 (1) より算出された各メンバ個々人の嗜好スコアから、コンテンツに対する嗜好ベクトルを作成する。グループのメンバ数を n とすると、あるコンテンツに対する嗜好ベクトル V_c は式 (2) と定義される。

$$V_c = \{u_{m_1,c}, u_{m_2,c}, \dots, u_{m_n,c}\} \quad (2)$$

たとえば、ユーザ A (m_A) とユーザ B (m_B) の各ジャンルに対する嗜好スコアが表 1 の値である場合、Action と SF のジャンルを持つ映像コンテンツ α の嗜好ベクトル V_α は以下のように算出される。

$$\begin{aligned} V_\alpha &= \{u_{m_A,\alpha}, u_{m_B,\alpha}\} \\ &= \left\{ \frac{r_{m_A,Action} + r_{m_A,SF}}{2}, \frac{r_{m_B,Action} + r_{m_B,SF}}{2} \right\} \\ &= \left\{ \frac{5+5}{2}, \frac{2+1}{2} \right\} \\ &= \{5, 1.5\} \end{aligned}$$

各メンバの嗜好スコアを軸とした空間に、嗜好ベクトルを求めた視聴済み映像コンテンツをプロットする。この処理をグループでの行動履歴内にある視聴済み映像コンテンツすべてに対して行うことで、Power Balance Map が完成する。

次に、推薦対象コンテンツ (たとえば未視聴の映像コンテンツ) がグループに適しているか否かを判断するためのレコメンドスコアを算出する。レコメンドスコアの算出には、各推薦対象コンテンツに対する各メンバの嗜好スコアと Power Balance Map を利用する。まず、行動履歴と同様の手法で、推薦対象コンテンツの嗜好ベクトルを算出する。続いて、算出した推薦対象コンテンツの嗜好ベクトルと Power Balance Map を構成する各視聴済みコンテンツの嗜好ベクトルとの類似度を求める。3.1 節で述べたように、Power Balance Map 上の高密度領域に属するあるいは近い推薦対象コンテンツがグループに適しているといえる。つまり、推薦対象コンテンツと Power Balance Map を構成する各視聴済みコンテンツとの類似度が高いほどグループに適しているコンテンツといえるため、類似度の総和をレコメンドスコアとする。図 4 は、推薦対象コンテンツを

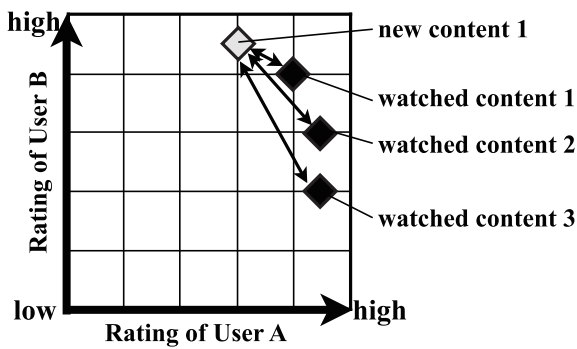


図 4 Power Balance Map 上における推薦対象コンテンツ (new content) と各視聴済みコンテンツ (watched content) の類似度算出法

Fig. 4 The method used to calculate similarity level between new content and watched content on the Power Balance Map.

Power Balance Map 上にプロットし、各視聴済みコンテンツとの類似度を求めるためのユークリッド距離を示した図である。この図を例にすると、推薦対象コンテンツ (new content 1) と各視聴済みコンテンツ (watched content 1~3) との類似度を式 (3) を用いて算出し、算出された類似度の総和 (式 (4)) が new content 1 の recommends コアとなる。

$$sim(\mathbf{V}_c, \mathbf{V}_h) = \frac{1}{\|\mathbf{V}_c - \mathbf{V}_h\| + 1} \quad (3)$$

$$S_c = \sum_{h \in H} sim(\mathbf{V}_c, \mathbf{V}_h) \quad (4)$$

式 (3), (4) において、推薦対象コンテンツ c の嗜好ベクトル \mathbf{V}_c と視聴済みコンテンツ h の嗜好ベクトル \mathbf{V}_h の類似度を $sim(\mathbf{V}_c, \mathbf{V}_h)$ で示しており、推薦対象コンテンツ c に対する recommends コアを S_c で表しており、視聴済みコンテンツの集合を H で示している。上記の手順をすべての推薦対象コンテンツに対して行い、recommends コアの高い順に推薦を行う。このように、Power Balance Map を利用することで、グループに適するコンテンツが推薦可能となる。

4. 検証実験

提案手法の妥当性を検証するために被験者夫婦 3 組の TV 番組視聴履歴を 5 週間収集し、そのデータの分析を行った。なお、被験者夫婦はどの組も同一のブロードキャストエリアに住んでいるため、視聴可能な地上波放送番組は平等である。今回の実験では、TV 番組視聴履歴は夫婦 2 人で視聴した番組のみを対象とし、どちらか片方だけしか視聴していない番組は除外した。以下に検証実験の詳細を述べる。

4.1 検証実験用データセット

検証実験を行うために、個々人の嗜好スコア、夫婦での

視聴 TV 番組履歴データ、推薦対象 TV 番組データという 3 種類のデータを収集した。個々人の嗜好スコア、夫婦での視聴 TV 番組履歴データは Power Balance Map を作成するために使用したデータである。また、それら 2 つのデータは 4.3 節で後述する比較手法での推薦スコア算出においても利用した。推薦対象 TV 番組データは推薦精度を求めるために使用した。以下にそれぞれのデータの詳細を記す。

個々人の嗜好スコア： 個々人の嗜好スコアとは、各 TV ジャンルに対する個人の嗜好を示す値であり、図 3 中の Rating List of each Member にあたるデータである。本実験では、被験者個人に各 TV ジャンルに対する嗜好をアンケートで収集した。嗜好スコアは 5 段階評価で表し、1 が嫌い、3 が普通、5 が好きとなり、2 と 4 はそれぞれの間となる。なお、TV ジャンルは TV 番組表サイト*1 で実際に使用されている 104 種類 (国内ドラマ、野球、トークバラエティ、クイズなど) を対象とした。

夫婦での視聴 TV 番組履歴データ： 夫婦での視聴 TV 番組履歴データとは、被験者夫婦 2 人で一緒に視聴した TV 番組の記録であり、図 3 中の Behavioral History of Group にあたるデータである。被験者夫婦は一緒に視聴した番組名を 5 週間毎日記録し、この番組名と上述の TV 番組表サイトから該当する TV ジャンル (1~3 つ) を合わせたものが夫婦での視聴 TV 番組履歴データとなる。実験期間中に放送された TV 番組データは全 12,234 番組であり、そのうち各被験者夫婦の視聴された番組数はそれぞれ 176, 76, 61 番組であった。

推薦対象 TV 番組データ： 推薦対象 TV 番組データとは、被験者夫婦に推薦する TV 番組の集合であり、図 3 中の New Content にあたるデータである。本実験では、このデータを推薦精度を算出するための評価データとして使用するために、各被験者夫婦に対して未放送の TV 番組データ 1 つ 1 つに “2 人で視聴するのに番組として「適」「不適」のいずれかを回答するアンケートを行った。この「適」か「不適」かの回答結果と、番組名とその番組該当する TV ジャンル (1~3 つ) で構成されたものが 1 つの推薦対象 TV 番組データとなる。なお、このアンケートはお互い相談したうえで「適」「不適」を回答してもらった。番組数は各被験者ごとに 763 番組であり、視聴 TV 番組履歴データの収集期間とは異なる 2 日分の TV 番組データを基にしている。そのため、夫婦での視聴 TV 番組履歴データとはすべて異なる TV 番組データとなっている。

*1 <http://tv.so-net.ne.jp/>

4.2 評価基準と評価手順

提案手法の有効性を示すために2つの評価基準を用いて評価を行った。1つは被験者夫婦2人にとって適していた番組がどの程度含まれていたかの推薦精度を表す適合率(以下, Appropriate Precision)である。もう1つは被験者夫婦2人にとって興味があるが意外・知らなかったTV番組がどの程度含まれていたかの推薦精度を表す適合率(以下, Novelty Precision)である。適合率は情報検索システムの研究分野で使用されている最も典型的な評価基準の1つである [14]。そして、この評価基準はレコメンデーションの研究分野でも数多く利用されており [15], [16], 有効性を示すには妥当な評価基準といえる。評価方法の手順を以下に記す。

Step1 任意の被験者夫婦1組の推薦対象TV番組データ全763件を抜き出す。

Step2 抜き出した推薦対象TV番組データに対して、提案手法と各比較手法(比較手法の詳細は4.3節で後述)それぞれの推薦結果を出力する。

Step3 Step2で出力された各推薦結果を上位から順に参照し、「適」と回答されている推薦対象TV番組データが上位K個にどれだけ含まれていたかを2つの適合率を用いて評価する。

Step4 すべての被験者夫婦に対してStep1からStep3までを繰り返す。

Step5 すべての被験者に対する評価結果のoverallで3手法を比較する。

すなわち、正解データとなる推薦対象TV番組データを先に収集した後に推薦結果を照らし合わせて評価するという順序で行った。なお、上記手順内のStep3で同率順位が存在した場合、ゴルフのストローク順位と同様に、同一スコアのTV番組は同一順位とし、次の順位は同一スコアであったTV番組の数だけ後ろの順位から開始する定義とした。たとえばTV番組A~Eが、A:5点, B:4点, C:4点, D:4点, E:3点であった場合、順位は1位:A, 2位:B, C, D, 5位:Eとし、適合率も上位1位までの適合率, 2位までの適合率, 5位までの適合率という算出となる。そのため、手法によって算出される適合率の数は異なる場合がある。以下に評価基準となる2つの適合率の詳細を述べる。

4.2.1 Appropriate Precision

Appropriate Precisionは推薦したすべての推薦対象TV番組データ集合(図5中のRecommended Programs)のうち、夫婦2人で視聴するのに適していると回答したTV番組(図5中のAppropriate Programs)がどの程度含まれていたかを示す確率である。図5で示した略記号を用いると、Appropriate Precisionは以下の式(5)で求められる。

$$\text{Appropriate Precision} = \frac{R \cap A}{R} \quad (5)$$

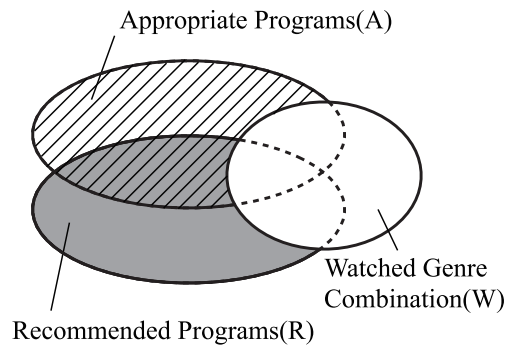


図5 各適合率を説明するための情報集合を示した Venn diagram
Fig. 5 Venn diagram to explain each precision metric.

4.2.2 Novelty Precision

近年、レコメンドシステムはユーザに適しているだけでなく、興味深いコンテンツも見つけ出せるべきだということが叫ばれている。Herlockerらは、レコメンドシステムは高い推薦精度(コンテンツの適合性)を誇るだけでなく、ユーザにとって有益となる推薦(推薦の有用性)も備えるべきだと主張している [17]。文献 [17]では、NoveltyやSerendipityといった「ユーザにとって興味があるが意外・知らなかったコンテンツ」を推薦の有用性を測る基準として紹介している。この考えに基づき、Novelty Precisionという評価基準を定義した。Novelty Precisionとは、推薦対象TV番組データから視聴履歴に含まれるジャンルの組合せとなるTV番組データを除いた場合の適合率と定義する。つまり、「被験者夫婦に適している」かつ「知らない・意外と思われる」TV番組をどれだけ推薦可能かを示す指標となる。今回の実験では、「知らない・意外と思われるTV番組」を「まだ視聴履歴に現れていないTV番組ジャンルの組合せ」かつ「被験者夫婦に適しているTV番組」と定義した。したがって、Novelty Precisionは、図5を例にすると、「推薦したTV番組(R)」かつ「まだ視聴履歴に現れていないTV番組ジャンルの組合せとなるTV番組(W^c)」の集合のうち、「夫婦2人で視聴するのに適していると回答したTV番組(A)」がどの程度含まれていたかを示す確率となり、以下の式(6)で求められる。

$$\text{Novelty Precision} = \frac{R \cap A \cap W^c}{R \cap W^c} \quad (6)$$

4.3 ベースライン

提案手法の優越性を検証するために、以下の2つのグループレコメンデーションアルゴリズムをベースラインとして比較する。

- (1) 加重平均法 (Weighted Average Method)
- (2) 仮想個人化法 (Virtual User Method)

加重平均法とは、個々人の嗜好スコアをある指標に基づいて統合し、その結果をグループでの嗜好スコアとしてコンテンツの推薦を行う手法である。この手法は統合法に基づいたグループレコメンデーションの手法として最も一

一般的な手法である。本実験では、Yu らが提案している手法 [5] に基づいて実装を行った。Yu らが提案するアルゴリズムは、各メンバー間で正規化した嗜好スコアの分布を適用させた手法であり、通常の加重平均よりも推薦精度を向上させるという特徴を持っている。このアルゴリズムに基づいた加重平均法をベースラインの1つとした。本実験で用いた加重平均法の推薦手順を以下に示す。

Step1 被験者夫婦の各々人(夫, 妻)の各 TV ジャンルに対する嗜好スコアを収集する。

Step2 収集した嗜好スコアを、夫・妻それぞれにおいて正規化する。

Step3 各 TV ジャンルに対するグループでの嗜好スコアを、Step2 で正規化した各々人の嗜好スコアの加重平均により求める。

Step4 Step3 で求めた TV ジャンルに対するグループでの嗜好スコアに基づいて検証用 TV 番組データの嗜好スコアを算出する。

Step5 嗜好スコアが高い番組順に推薦を行う。

仮想個人化法とは、グループを「1 人の仮想的なユーザ」と見なし、仮想個人ユーザの行動履歴から嗜好を推定してコンテンツの推薦を行う手法である。本実験では、夫婦での視聴履歴に基づいて各 TV ジャンルに対する嗜好を推定し、その結果を用いて推薦を行う手法をもう 1 つのベースラインとした。本実験で用いた仮想個人化法の推薦手順を以下に示す。

Step1 仮想的な 1 人の視聴者(被験者夫婦)の視聴履歴を収集する。

Step2 視聴済み TV 番組に紐付けられた「TV ジャンル」の頻度を求める。

Step3 Step2 で求めた TV ジャンルの頻度に基づいて検証用 TV 番組データの嗜好スコアを算出する。

Step4 嗜好スコアが高い番組順に推薦を行う。

4.4 検証結果

図 6 は提案手法 (Proposed Method)、加重平均法 (Weighted Average)、仮想個人化法 (Virtual User) の推薦結果上位 38 位 (上位 5%) までの Appropriate Precision (全被験者の overall) を示した図である。横軸は上位 K 位 (上位 K%) を、縦軸は Appropriate Precision の値を示している。この図より、提案手法は上位 1 位こそ加重平均法に劣っているものの、他の 2 手法と比べて安定した推薦精度を出せるという結果が得られた。

図 7 は、各手法の上位 38 位 (上位 5%) までの Novelty Precision (全被験者の overall) を示した図である。この図より、提案手法は既存の 2 手法に比べ「ユーザにとって興味があるが意外・知らなかったコンテンツ」を探す場合においても効果的であるといえる。ここで図 6 の適合率と比較してみると、加重平均法の精度が大幅に低下しているこ

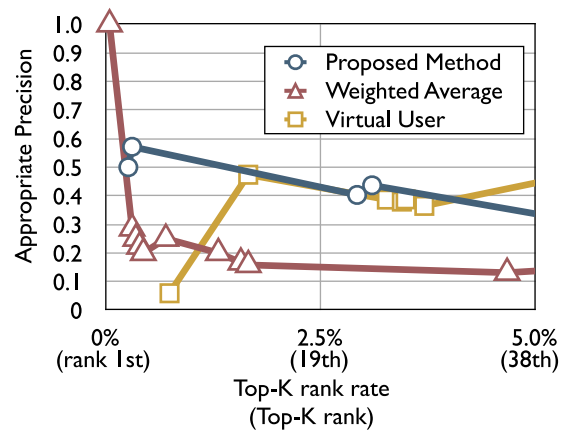


図 6 提案手法 (Proposed)、加重平均法 (Weighted Average)、仮想個人化法 (Virtual User) の Appropriate Precision

Fig. 6 Appropriate Precision of Proposed Method, Weighted Average Method and Virtual User Method.

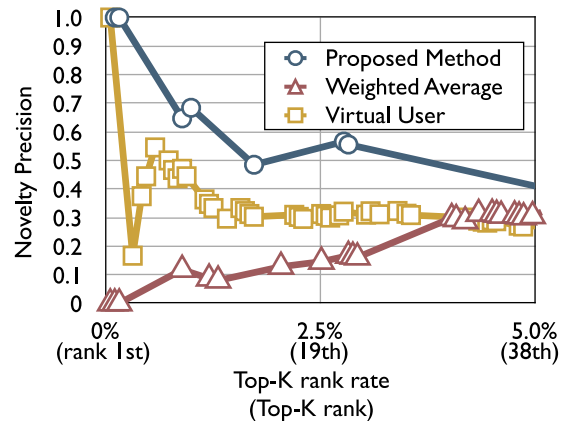


図 7 各手法の Novelty Precision

Fig. 7 Novelty Precision of each method.

とが分かる。このことから、加重平均法はユーザにとって興味があるが意外・知らなかったコンテンツを探す場合には不向きな手法であると推測される。

提案手法と仮想個人化法を比較すると、提案手法の方が高い Novelty Precision を達成している。これは、「夫は好きであるが、妻は普通くらいの TV 番組ジャンルを 2 人でよく見ているため、まだグループの視聴履歴として現れていない TV 番組のジャンルであっても好きであるに違いない」と推定する Power Balance Map の特徴が効果的に働いたと推測される。つまり、グループの視聴履歴と夫婦各々人の各 TV 番組ジャンルに対する嗜好とを組み合わせる Power Balance Map を作成することにより、仮想個人化法より少ない履歴数で適する番組を推薦可能なことを示している。

4.5 考察

図 8 は、各被験者の Power Balance Map、適した番組 (Appropriate Programs)、不適な番組 (Inappropriate Programs)、興味があるが意外・知らなかった番組 (Novel

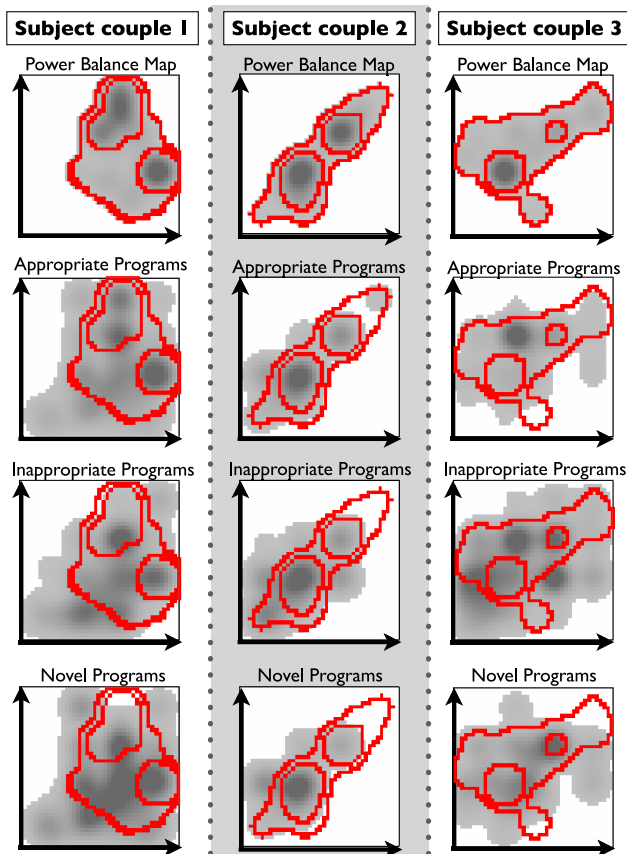


図 8 各被験者の Power Balance Map, 適した番組 (Appropriate Programs), 不適な番組 (Inappropriate Programs), 意外かつ適した番組 (Novel Programs) の分布頻度を示したヒートマップ

Fig. 8 Each subject's Heat map of Power Balance Map, Appropriate Programs, Inappropriate Programs and Novel Programs.

Programs) の分布頻度を表したヒートマップを示した図である。横軸は夫のコンテンツに対する嗜好スコアであり、縦軸はその配偶者のコンテンツに対する嗜好スコアとなっている。各ヒートマップ上で色が濃い領域がコンテンツが密集している高密度領域を示しており、色が薄くなるほど密度も低くなり、白色の領域はコンテンツがプロットされていないことを表している。また、Appropriate Programs, Inappropriate Programs, Novel Programs のヒートマップ上の実線は、Power Balance Map の主な高密度領域とその分布範囲を示している。

各被験者夫婦間で Power Balance Map のヒートマップを比較すると、それぞれ異なった特徴を持った Power Balance Map が作成されていることが分かる。したがって、4.4 節で述べた各適合率の結果とあわせて考えると、提案手法は様々な関係性や特徴を持ったグループに対応可能な手法であるといえる。

各々の被験者内で Power Balance Map と適した番組のヒートマップを比較すると、両者の高密度領域が一致する傾向にあり、適した番組の密度が高めの領域も Power

Balance Map の分布範囲に収まる傾向にあることが分かる。また、適した番組の全体的な分布範囲も Power Balance Map の分布範囲と類似した傾向にあり、不適な番組よりもはみ出している領域が少ない傾向にあることが示されている。しかし、不適な番組の高密度領域が Power Balance Map の高密度領域に含まれてしまう場合も存在している。このことが、Appropriate Precision の上位 1 位が加重平均法より低い結果となっている原因であると考えられる。今回の実験では、嗜好スコアを TV ジャンルのみ使用するという少ない特徴量を用いて行っているため、適切な特徴量を増やすことでこの問題は緩和可能である。

各被験者の興味があるが意外・知らなかった番組のヒートマップに着目すると、夫婦個々人の嗜好スコアがどちらも最高としている領域（ヒートマップの右上の領域）の密度が低いことが分かる。特に被験者夫婦 2 と 3 は、適した番組はこの右上の領域に存在しているにもかかわらず、興味があるが意外・知らなかった番組は存在していない。逆に、Power Balance Map の高密度領域の内部もしくは周辺に、興味があるが意外・知らなかった番組の高密度領域が存在している。このことが加重平均法における Novelty Precision の低さの原因と考えられる。すなわち、夫婦に適した番組は互いの嗜好スコアの高い領域に確かに存在するが、レコメンドの満足度や有用性を考慮した場合、嗜好スコアが高いものどうしという観点だけでは不十分であるといえる。また、“互いが好きどうしのもの”の中に含まれる適した番組は意外性が低く、すでに知っている番組である可能性が高いともいえる。したがって、グループレコメンドにおいてユーザの満足度を向上させるには“互いが好きなものを推薦すれば良い”というアプローチではなく、Power Balance Map などの個々人の嗜好以外の観点を考慮する必要がある。

5. まとめ

本論文では、個々人の嗜好スコアとグループでの行動履歴から Power Balance Map を算出する手法と、その Power Balance Map に基づいてグループに適したコンテンツを推薦する手法の提案を行った。夫婦で視聴した TV 番組の履歴を収集した検証実験を行うことで、Power Balance Map に基づいたグループレコメンドは既存手法よりも安定した推薦精度が出せることを示した。Power Balance Map 上の高密度な領域とグループに適したコンテンツの間には類似した傾向があることを示し、グループの様々な関係性や特徴に対して適応可能なことを確認した。また、個々人の嗜好スコアがどちらも最高となる領域にはグループに適したコンテンツが存在するが、意外性がないなどの推薦の有用性が低いコンテンツである可能性が高いことも示した。そして、Power Balance Map を利用したグループレコメンドは、興味があるが意外・知らなかった

コンテンツを推薦するのに役に立つ可能性を示した。

今後は、さらなる推薦精度向上のため、TV ジャンル以外の特微量（出演者、監督、映像の雰囲気など）も用いた分析と検証を行っていく。さらに、他ドメイン（飲食店、旅行など）への展開した場合の有効性も確認していく予定である。また、本論文では2人で構成されるグループを対象とした検証であったため、3人以上のグループに対しても有効なことを検証していく。

参考文献

- [1] Adomavicius, G. and Tuzhilin, A.: Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions, *IEEE Trans. Knowledge and Data Engineering*, Vol.17, No.6, pp.734-749 (2005).
- [2] Linden, G., Smith, B. and York, J.: Amazon.com recommendations: item-to-item collaborative filtering, *IEEE Internet Computing*, Vol.7, No.1, pp.76-80 (2003).
- [3] Tezuka, H., Ito, K., Murayama, T., Seko, S., Nishino, M., Muto, S. and Abe, M.: Restaurant recommendation service using lifelogs, *NTT Technical Review*, Vol.9, No.1 (2011).
- [4] Nakamura, Y., Itou, T., Tezuka, H., Ishihara, T. and Abe, M.: Personalized TV-program recommendations based on life log, *International Conference on Consumer Electronics (ICCE), 2010 Digest of Technical Papers*, pp.143-144 (2010).
- [5] Yu, Z., Zhou, X., Hao, Y. and Gu, J.: TV program recommendation for multiple viewers based on user profile merging, *User Modeling and User-Adapted Interaction*, Vol.16, No.1, pp.63-82 (2006).
- [6] Masthoff, J.: Group Modeling: Selecting a Sequence of Television Items to Suit a Group of Viewers, *User Model and User-Adapted Interaction*, Vol.14, pp.37-85 (2004).
- [7] O'Connor, M., Cosley, D., Konstan, J. and Riedl, J.: PolyLens: A Recommender System for Groups of Users, *ECSCW 2001*, pp.199-218 (2001).
- [8] Jameson, A. and Smyth, B.: Recommendation to Groups, *The Adaptive Web*, pp.596-627 (2007).
- [9] Goren-Bar, D. and Glinansky, O.: FIT-recommending TV programs to family members, *Computers & Graphics*, Vol.28, No.2, pp.149-156 (2004).
- [10] Shin, C. and Woo, W.: Socially Aware TV Program Recommender for Multiple Viewers, *IEEE Trans. Consumer Electronics*, Vol.55, No.2, pp.927-932 (2009).
- [11] Berkovsky, S. and Freyne, J.: Group-Based Recipe Recommendations: Analysis of Data Aggregation Strategies, *Proc. 4th ACM conference on Recommender systems* (2010).
- [12] Sotelo, R., Blanco-Frenandez, Y., Lopez-Nores, M., Gil-Solla, A. and Pazos-Arias, J.J.: TV Program Recommendation for Groups Based on Multidimensional TV-Anytime Classifications, *IEEE Trans. Consumer Electronics*, Vol.55, No.1, pp.248-256 (2009).
- [13] Rodgers, J.L. and Nicewander, W.A.: Thirteen Ways to Look at the Correlation Coefficient, *The American Statistician*, Vol.42, No.1, pp.50-66 (1988).
- [14] Cleverdon, C.W., Mills, J. and Keen, M.: Factors determining the performance of indexing systems, *ASLIB Cranfield project* (1966).
- [15] Gunawardana, A. and Shani, G.: A survey of accuracy evaluation metrics of recommendation tasks, *The Journal of Machine Learning Research*, Vol.10, pp.2935-2962 (2009).
- [16] Cremonesi, P., Koren, Y. and Turrin, R.: Performance of recommender algorithms on top-n recommendation tasks, *Proc. 4th ACM conference on Recommender systems*, pp.39-46 (2010).
- [17] Herlocker, J.L., Konstan, J.A., Terveen, L.G. and Riedl, J.T.: Evaluating Collaborative Filtering Recommender Systems, *ACM Trans. Information Systems*, Vol.22, No.1, pp.5-33 (2004).
- [18] Sarwar, B., Karypis, G., Konstan, J. and Riedl, J.: Item-based Collaborative Filtering Recommendation Algorithms, *Proc. International World Wide Web Conference* (2001).
- [19] Seko, S., Motegi, M., Yagi, T. and Muto, S.: Video Content Recommendation for Group Based on Viewing History and Viewer Preference, *2011 IEEE International Conference on Consumer Electronics (ICCE)*, pp.359-360 (2011).



瀬古 俊一

1982年生。2008年慶應義塾大学大学院政策・メディア研究科修士課程修了。同年日本電信電話(株)入社。現在、NTT サービスエボリューション研究所研究員。レコメンド技術に関するアルゴリズムやユーザインタフェースの研究に従事。2011年IEEE ICCE Special Merit Awards, FIT 2010 ヤングリサーチアワード受賞。



八木 貴史 (正会員)

1992年慶應義塾大学大学院理工学研究科計算機科学専攻修士課程修了。同年NTT入社。現在NTT レゾナント株式会社サーチ事業部ビジネス推進部門長。検索・ポータル関連サービスの企画・開発に従事。電子情報通信学会、日本バーチャルリアリティ学会、ACM 各会員。



茂木 学 (正会員)

1995年東京工業大学大学院修士課程修了。同年NTT入社。現在、NTT サービスエボリューション研究所研究主任。博士(工学)。1997年第2回ロボティクス・シンポジウム論文賞、1998年日本ロボット学会研究奨励賞、2008年医療の質・安全学会ベストトライアル賞等を受賞。日本ロボット学会会員。



武藤 伸洋

1990年早稲田大学大学院理工学研究科修士課程修了。同年NTT入社。現在、NTTサービスエボリューション研究所主幹研究員。この間2000～2003年国立情報学研究所客員助教授。日本ロボット学会、日本機械学会各会員。



小林 透 (正会員)

1985年東北大学工学部精密機械工学科卒業。1987年同大学大学院工学研究科修士課程修了。同年NTT入社。以来、ソフトウェア生産技術、ユビキタスコンピューティング、情報セキュリティなどの研究開発に従事。現在、NTTサービスエボリューション研究所主幹研究員。電子情報通信学会、IEEE各会員、博士(工学)。