

行動履歴を利用したグループレコメンデーションのための

コンテンツフィルタリング手法

A Contents Filtering Method for Group Recommendation using Life Log

瀬古 俊一[†]
Shunichi SEKO茂木 学[†]
Manabu MOTEGI八木 貴史[†]
Takashi YAGI武藤 伸洋[†]
Shin'yo MUTO阿部 匡伸[‡]
Masanobu ABE

1. はじめに

近年、個人の行動履歴から嗜好を推定し、その個人の嗜好に合った情報を提供するレコメンドシステムの研究・サービスが活発化している。また、iPadのような大型のディスプレイを備えた携帯端末の流行やデジタルサイネージの台頭により、同一ディスプレイを複数人で閲覧し、情報を取得するという機会が増えると予想される。このような複数人のグループに情報を提示する場合、そのグループの特性を考慮する必要があり、従来の個人の嗜好推定と同様のアプローチでは適切な情報選択が困難な場合がある。そこで本論文では、グループの特性を活かしたフィルタリング手法を提案する。本手法はグループメンバー間の関係性に着目し、各メンバーの嗜好とそのグループの行動履歴を用いることで、より高精度なフィルタリングを実現するものである。

2. グループに適したコンテンツフィルタリング手法の代表例

グループに適したコンテンツフィルタリング手法を実現するには、あるコンテンツ情報が対象グループにとって適した情報か否かを判定する必要がある。これを行う代表的な手法として、以下の3つがあげられる。

1. 各メンバーの嗜好スコアを用いた推定手法
2. グループの行動履歴を用いた推定手法
3. 正解付きコンテンツ情報を用いた推定手法

上記3手法は、それぞれ入力データが異なる。以下に各手法の特徴を述べる。

2.1 各メンバーの嗜好スコアを用いた推定手法

各メンバーの嗜好スコアを用いた推定手法とは、各メンバーのコンテンツに対する嗜好スコアを加重平均などで統合し、その結果をグループのコンテンツに対する嗜好スコアとする手法である。ここで、コンテンツとは、例えば〇〇中華料理店などの飲食店のことをさし、個人の嗜好スコアとはその飲食店を好む度合いが5点などの定量化された値のことである。伊藤ら[1]は、個人の行動履歴(例えば中華料理店の情報を閲覧したなど)から個人の嗜好スコアを推定する手法を提案している。グループに対する推薦を行う際は、ユーザが任意の加重を入力し、各メンバーのコンテンツに対する嗜好スコアに、この加重を乗算してグループのコンテンツに対する嗜好スコアを推定する。そして、グループの嗜好スコアが高いコンテンツを対象グループに適する情報として提示する手法が述べられている。以上のように、各メンバーの嗜好スコアを用いた推定手法は、個々人の嗜好ス

[†]NTTサイバースリユーション研究所 NTT Cyber Solutions Laboratories

[‡]岡山大学 Okayama University

コアが算出されていれば推定可能である。したがって、比較的収集が容易な個々人のデータを入力とするだけで推定が行える。

2.2 グループの行動履歴を用いた推定手法

グループの行動履歴を用いた推定手法とは、例えばそのグループで中華料理屋に訪れる頻度が高ければ、そのグループは中華料理を好む度合いが高いと推定するものである。このように、実際に起こした行動を基に嗜好を推定する手法である。例えば、上述の伊藤ら[1]の提案する手法における個人の行動部分をグループの行動とすると、同様の手法でグループの嗜好を推定可能となる。以上のように、グループの行動履歴を用いた推定手法は、実際に起こった行動を用いて推定を行う。そのため、一定数の行動履歴が蓄積されなければ十分な推定を行えない可能性がある。

2.3 正解付きコンテンツ情報を用いた推定手法

正解付きコンテンツ情報を用いた推定手法とは、対象グループにとって適しているコンテンツと不適なコンテンツの情報を予め取得し、それに基づいて未知のコンテンツ情報の適・不適を推定する手法である。例えば、対象グループにいくつかのコンテンツ情報(例えば〇〇中華店、△△ラーメン屋など)を提示し、その情報がグループにとって適しているか否かのアンケートを取る。そのアンケート結果を機械学習などで学習を行い、その学習結果に基づいて未知のコンテンツ情報が対象グループにとって適した情報であるか否かの推定を行う。学習時には、コンテンツ情報をいくつかの特徴量(例えば飲食店であれば料理のジャンルなど)に分解し、それぞれの特徴量をパラメータとして学習させる。それにより、どのような特徴量をもつ情報が対象グループにとって適した情報であるかを推定可能となる。以上のように、正解付きコンテンツ情報を用いた推定手法は、対象グループに適する(あるいは不適)と確定された情報を基に推定を行う。そのため、十分な推定を行うには一定量の正解付きコンテンツ情報が必要である。

3. "グループ"であることに着目した手法の提案

2章では、グループに適した情報を選択するための手法として3つの例をあげた。しかしながら、2章で述べた手法はどれも個人に適した推定手法をグループに適用した手法であり、以下に述べるグループの特性を考慮したものではない。したがって、グループの特性に着目することで、推定精度を向上できる可能性がある。

個人とグループの最大の違いは、相手がいるか否かである。相手がいるときは、一人での行動とは違った行動を起こす可能性がある。井上[3]は、一人のときと同伴者がいる時の消費・購買行動は、購買の動機から結果にいた

るまでいくつかの違いが観察されると述べている。また、同判者がいる場合でも、同判者の違いによって消費・購買行動の傾向が変化することも述べられている。このように、一人での行動とグループであるときの行動は変化する。また、グループのメンバ構成によっても変化する。以上のことを踏まえると、対象グループの特性を考慮することは、適した情報推定を行う際の推定精度向上に有用であると考える。

本論文では、グループの特性を考慮する一つの手法として、グループ内の関係性に着目する。グループ内の関係性とは、例えばグループメンバ間の力関係や嗜好の偏り具合などといった傾向のことである。グループ内の関係性は、あるコンテンツ情報に対する各メンバの嗜好スコアを特徴量としたとき、どのような特徴量を有するコンテンツ情報がグループの行動として起こる頻度が高いかを調べることで抽出可能である。例えば、あるメンバの嗜好スコアが高いコンテンツ情報がグループの行動として頻繁に起こっていた場合、そのメンバはグループ内で強い力を持っていると推測可能である。このように、各メンバの嗜好スコアを特徴量とした空間(以下、嗜好空間)を考慮すると、グループ内の関係性を示すことが可能となる。図1は User A と User B から成るグループの嗜好空間上において、グループでの行動が頻繁に起きている空間を図示した例である。図1の①は、行動頻度が高い空間が User A の嗜好スコアが高い空間に固まっているため、User A が強い力を持つ関係であるといえる。図1の②は、①とは反対に User B の嗜好スコアが高い空間の行動頻度が高いため、User B の力が強い関係といえる。図1の③は、お互いの嗜好スコアが似通った空間の行動頻度が高いため、どちらかの嗜好に偏るよりはお互いの嗜好に近いものを選択する傾向にあり、バランスを重視する関係といえる。図1の④は、少なくともどちらか一方の嗜好スコアが高い空間の行動頻度が高いため、どちらかが必ず満足するようなものを選択する傾向にある関係といえる。このように、嗜好空間において行動頻度が高い空間を抽出することにより、グループ内の関係性を示せる。これによりグループ内の関係性を抽出可能となり、グループの特性を考慮したフィルタリング手法が実現可能となる。

4. アプローチ

本章では、3章で提案したグループ内の関係性を考慮した、対象グループに適したコンテンツ情報推定手法の実現方法について述べる。

本手法は、各メンバの嗜好スコアとグループの行動履歴を入力データとして、コンテンツ情報が対象グループに適している情報か否かを推定する。ここで、各データは飲食店を例にすると以下となる。

- 嗜好スコア：○○中華料理店に対する好みを定量化した情報
- グループの行動履歴：○○中華料理店にグループで実際に訪れたことを示す情報
- コンテンツ情報：○○中華料理店や△△ラーメン屋等の飲食店に対する情報

まず、グループの行動履歴を嗜好空間上のコンテンツ情報に反映する。例えば、○○中華料理店に訪れたことを示すグループの行動履歴であれば、○○中華料理店というコ

ンテンツ情報に、実際に訪れたことを示す情報(例えば訪問回数)が付加される。グループの行動履歴を反映したコンテンツ情報を行動履歴コンテンツ情報と呼ぶ。

各コンテンツ情報が対象グループに適しているか判定するために、そのコンテンツ情報と行動履歴コンテンツ情報との嗜好空間上における類似性を用いる。この類似性を定量化したものを類似度と呼び、行動履歴コンテンツ情報すべての類似度を総和した値を基に対象グループに適した情報であるか否かを判別する。この類似度の総和が高いということは、嗜好空間上の行動頻度が高い空間に属していることを表している。そのため、類似度の総和が閾値以上のものを対象グループにとって適した情報であると判別する。これを各コンテンツ情報に対して行うことで、対象グループに適した情報を推定可能となる。

コンテンツ情報と行動履歴コンテンツ情報との類似度の算出法として、3つの手法をあげる。図2は User A と User B から成るグループの嗜好空間上にコンテンツ情報3つ(白で表示)と行動履歴コンテンツ情報1つ(黒で表示)をプロットした場合の例を示している。図2に示した類似度算出法1は、コンテンツ情報と行動履歴コンテンツ情報とのユークリッド距離を用いて類似度を算出する手法である。ユークリッド距離が近いほど類似性が高いとし、例えば一定距離以内にあるコンテンツ情報の類似度を1(範囲外のデータの類似度は0)とする。あるいは、ユークリッド距離の逆数の値を類似度とする。図2に示した類似度算出法2は、コンテンツ情報と行動履歴コンテンツ情報との cosine similarity の値を用いて類似度を算出する手法である。cosine similarity の値が高いほど類似性が高いとし、例えば一定値以内にあるコンテンツ情報の類似度を1(範囲外のデータの類似度は0)とする。あるいは、その値そのものを類似度とする。図2に示した類似度算出法3は、嗜好空間をn個に分割したクラス(図の例では各メンバの嗜好スコアを5等分した25個に分割したクラス)を用いて類似度を算出する手法である。コンテンツ情報と行動履歴コンテンツ情報が同一クラスにある場合を類似しているとし、同一クラスにあるコンテンツ情報の類似度を1、異なるクラスにある情報の類似度を0とする。上述した手法を用いることで、

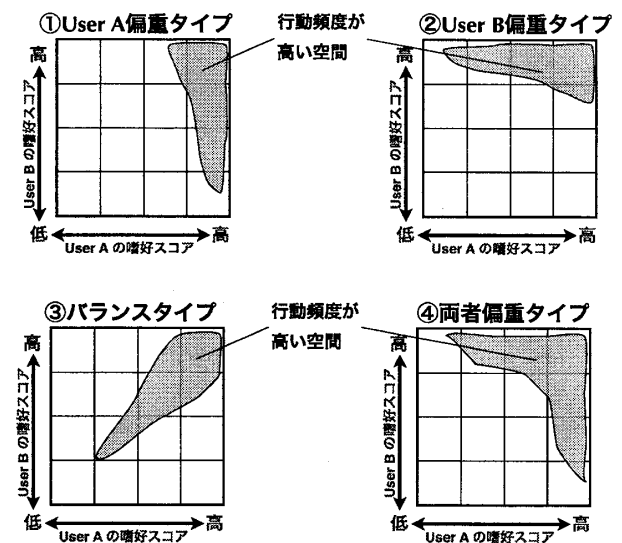


図1 嗜好空間で表せるグループ内の関係性の例

コンテンツ情報と行動履歴コンテンツ情報との類似度を算出可能となる。いま、グループメンバー数を n 、メンバー i のコンテンツ情報 a に対する嗜好スコアを $u_{i,a}$ 、各メンバーの嗜好スコアを特徴量とした行動履歴コンテンツ情報 j を $H_j = (u_{1,j}, u_{2,j}, \dots, u_{n,j})$ 、各メンバーの嗜好スコアを特徴量としたコンテンツ情報 k を $I_k = (u_{1,k}, u_{2,k}, \dots, u_{n,k})$ と表したとき、コンテンツ情報 k の行動履歴コンテンツ情報 j に対する類似度 $sim(k, j)$ の算出式は以下ようになる。

- 類似度算出法 1

$$sim(k, j) = \frac{1}{\sqrt{\sum_{i=1}^n (u_{i,k} - u_{i,j})^2 + 1}}$$

- 類似度算出法 2

$$sim(k, j) = \cos(I_k, H_j) = \frac{I_k \cdot H_j}{\|I_k\|_2 \times \|H_j\|_2}$$

- 類似度算出法 3

$$sim(k, j) = \begin{cases} 1 & \text{if } class(I_k) = class(H_j) \\ 0 & \text{if } class(I_k) \neq class(H_j) \end{cases}$$

$class(x)$ は x が属するクラスを表す

上記の式を用いることでコンテンツ情報 k の行動履歴コンテンツ情報 j に対する類似度を算出可能となる。ただし、上記の類似度算出は行動履歴コンテンツ情報 1 つに対する類似度であるため、コンテンツ情報 k が対象グループに適した情報であるかを判別するには、各行動履歴コンテンツ情報との類似度の総和を求める必要がある。いま、行動履歴コンテンツ情報数を m と表したとき、コンテンツ情報 k が対象グループに適した情報であるかを判別するための値 (以下、適合度) s_k は以下のように算出される。

$$s_k = \sum_{j=1}^m sim(k, j)$$

コンテンツ情報 k の適合度 s_k が閾値以上かを判定し、閾値以上であればコンテンツ情報 k は対象グループに適した情報であるとする。反対に閾値に満たなければ不適な情報であるとする。これをコンテンツ情報の集合すべてに対して行い、不適な情報と判定されたものを除去することで、対象グループに適したコンテンツ情報のフィルタリングが可能となる。

5. 検証実験

4章で提案したアプローチと2章で示した手法との比較を行うことにより、それぞれの手法にどのような特性があるか検証を行った。実験概要は以下の通りである。

- 評価基準
 - 最適情報再現率
 - 非最適情報適合率
- 比較手法
 - 各メンバーの嗜好スコアを用いた推定手法
 - グループの行動履歴を用いた推定手法
 - 正解付きコンテンツ情報を用いた推定手法
 - 提案手法(類似度算出法3を利用)
- データセット
 - 対象コンテンツ
 - ◇ 飲食店
 - 被験者
 - ◇ 一緒に外食に行く頻度が高い男女の2人組1組
 - ◇ 男女ともに20代
 - メンバーの嗜好スコア
 - ◇ 被験者にアンケートを行うことで取得

- ◇ 219項目に対する嗜好度合いを5段階評価
- グループの行動履歴
 - ◇ レシートや手帳に記録された情報から抽出
 - ◇ 約1年半、75店舗分
- 評価判定用コンテンツ情報
 - ◇ コンテンツ情報に被験者の組にとっての適合度合いを付与したもの
 - ◇ 被験者の組にアンケートを行うことで取得
 - ◇ 実在する470店舗に対する適合度合いを5段階評価

5.1 実験内容

最適情報再現率と非最適情報適合率を評価基準として検証を行う。最適情報再現率とは、評価判定用コンテンツ情報の集合に含まれている対象グループにとって適する全コンテンツ情報のうち、フィルタリング後にどれだけ保持できたかを示す値である。最適情報再現率は以下の式を用いて算出し、値が高いほど良い結果であることを示している。

$$\text{最適情報再現率} = \frac{\text{フィルタリング後の最適情報数}}{\text{全最適情報数}}$$

非最適情報適合率は、フィルタリング後の評価判定用コンテンツ情報集合のうち、対象グループにとって不適な情報をどれだけ含んでいないかを示す値である。つまり、フ

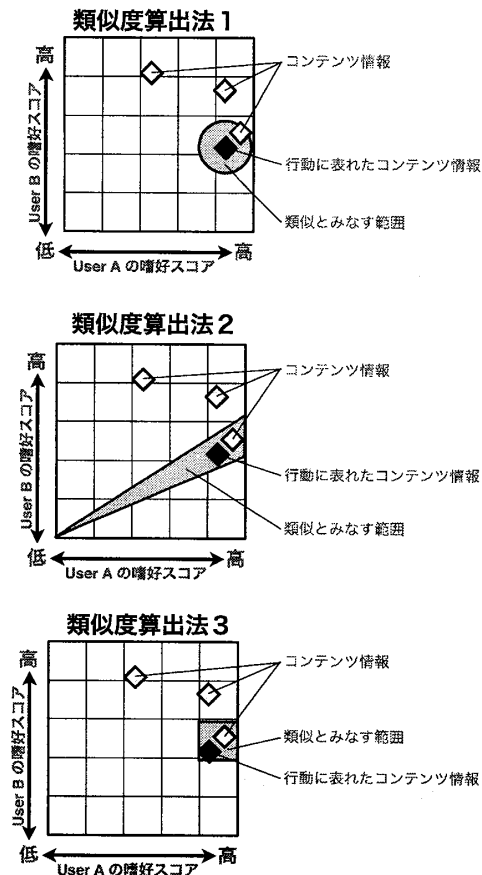


図2 嗜好空間上における類似度算出法の例

フィルタリング後にどれだけ不適なコンテンツ情報を除去できたかを表す。非不適情報適合率は以下の式を用いて算出し、値が高いほど良い結果であることを示している。

$$\text{非不適情報適合率} = 1 - \frac{\text{フィルタリング後の不適情報数}}{\text{フィルタリング後の全情報数}}$$

最適情報再現率と非不適情報適合率が共に高いという結果が得られれば高精度のフィルタリングが行えたと評価する。

比較する手法は、各メンバーの嗜好スコアを用いた推定手法(以下、従来手法 1)、グループの行動履歴を用いた推定手法(以下、従来手法 2)、正解付きコンテンツ情報を用いた推定手法(以下、従来手法 3)、提案手法の 4 つで行う。従来手法 1 は、コンテンツ情報に対するグループの嗜好スコアを、各メンバーのコンテンツ情報に対する嗜好スコアの平均値として推定を行った。嗜好スコアは被験者にアンケートを行うことで算出する。アンケートの詳細は後述する。従来手法 2 は、グループの行動履歴に蓄積されているコンテンツ情報の出現回数を用いて推定を行った。グループの行動履歴の抽出については後述する。従来手法 3 では、被験者にコンテンツ情報に対する適合度合いを回答してもらうアンケートを行い、その一部を学習データとし、残りをテストデータとして推定を行った。学習アルゴリズムは決定木を用いた。提案手法は 4 章で述べた手法であり、類似度の算出は類似度算出法 3 を用いて行う。なお、分割するクラス数は、各メンバーの嗜好スコアの理論的な最小値と最大値の間を 20 等分した数(2 人組みであれば 400 クラス)とした。これら 4 つの手法の最適情報再現率と非不適情報適合率を比較し、どの手法によるコンテンツフィルタリングが優れているかを評価する。

実験に用いたデータセットは、対象コンテンツを飲食店情報とした。被験者は一緒に外食に行く頻度が高い男女の 2 人組 1 組とし、年齢は男女ともに 20 代である。被験者それぞれの嗜好スコアは個々にアンケートを取ることで取得した。アンケート内容は料理のジャンルに対しての嗜好度合いを 5 段階(嫌い、やや嫌い、普通、やや好き、好き)の評価で行った。料理のジャンルは飲食店情報サイトを参考にし、全 219 項目を設定した。グループの行動履歴は、被験者の組が実際に飲食店で食事をした際のレシート及び手帳に記録されていた情報を基に抽出を行った。抽出した情報は日時と店舗名である。また、料理のジャンルを決定する際に利用した飲食店情報サイトに掲載されている飲食店情報と実際に訪れた店舗名をマッチングし、その店舗に該当する料理のジャンルを導出した。つまり、グループの行動履歴の 1 つは、日時と店舗名、その店舗に該当する料理のジャンルとなる。店舗に該当する料理のジャンルは 1 店舗に対して 1~3 つとなり、その数は店舗によって異なる。評価判定用コンテンツ情報は実在する飲食店とし、その情報は料理のジャンルを決定する際に利用した飲食店情報サイトに掲載されている情報 470 店舗分を利用した。この評価判定用コンテンツ情報が被験者の組に適しているか否かを判定し、最適情報再現率と非不適情報適合率を算出する。これらの情報が最適か否かは被験者の組にアンケートを行うことで導出した。アンケート内容はグループにとってその飲食店の適合度合いを 5 段階(低い、やや低い、中、やや高い、高い)の評価で行った。最適情報、不適情報はこの適合度合いを用いて判断する。対象グループにとって適しているか否かという判断は、以下の 2 つのケースで行った。

1. 適合度合いが“高い”飲食店を最適情報、それ以外を不適情報とする
2. 適合度合いが“高い”を最適情報、適合度合いが“やや低い”または“低い”を不適情報とする

2 つのケースを用いる理由としては、応用先のサービスによって求められる結果が異なるためである。ケース 1 は最適情報のみを抽出したい場合の評価になり、ケース 2 は不適情報をできる限り除去したい場合の評価になる。両ケースにおいて良い結果が得られる手法は、どちらの目的でも使用可能な手法といえる。

5.2 実験結果

上記検証実験の結果を以下に示す。図 3 はケース 1 による判断で各手法の推定結果を評価したものである。横軸はフィルタリングを行ったことで元の評価判定用コンテンツ情報の集合から何%の数を除去したか(以下、データ除去率)を示しており、縦軸は最適情報再現率あるいは非不適情報適合率を示している。最適情報再現率も非不適情報適合率も数値が高いほど良い結果であることを示している。図 3 によると、最適情報再現率も非不適情報適合率も各手法で差が顕著にあらわれるのはデータ除去率が 80%以上の場合である。図 4 は図 3 のデータ除去率 80%以上の領域を拡大した図である。図 4 によると、従来手法 3 は最適情報再現率と非不適情報適合率の両者ともに高い結果が得られる。提案手法がその次に高い結果が得られる。従来手法 1 と従来手法 2 はデータ除去率 90%以上では差がみられないが、80~90%の最適情報再現率は従来手法 2 が若干勝っている。また、データ除去率が 90%以上では、従来手法 1 および 2 の最適情報再現率と非不適情報適合率は、他の 2 つの手法に比べて劣る結果となっている。

図 5 はケース 2 による判断で各手法の推定結果を評価したものである。図 3 と同様、横軸はデータ除去率、縦軸は最適情報再現率あるいは非不適情報適合率を示している。図 5 によると、最適情報再現率において各手法の差が顕著に現れるのはデータ除去率が 80%以上となる領域である。図 6 は図 5 のデータ除去率 80%以上の領域を拡大した図である。図 6 によると、ケース 1 と同様に、従来手法 3 は最適情報再現率と非不適情報適合率の両者ともに高い結果が得られる。ついで、非不適情報適合率を優先するのであれば従来手法 2、最適情報再現率を重視するのであれば提案手法となる。従来手法 1 は、ケース 1 と同様、他の手法に比べ劣勢の結果となった。

5.3 考察

以上の検証実験結果から得られた考察を以下に述べる。まず、各手法の効果があるケースに関しての考察を行う。従来手法 1 はどちらのケースにおいても高精度なフィルタリングは行えなかった。しかし、ケース 2 ではデータ除去率が 65%前後であれば最適情報再現率と非不適情報適合率が高い値でバランスよくとれている。そのため、データ除去率が 65%前後であれば、従来手法 1 でも効果が得られる可能性がある。従来手法 2 は、データ除去率 80%前後であれば、ケース 1, 2 共に効果のあるフィルタリングが行えるといえる。ただし、データ除去率 90%以上になるとケース 2 の非不適情報適合率以外のフィルタリングの精度は低い。そのため、従来手法 2 で高精度なフィルタリングを行える可能性が高いのは、データ除去率 80%前後のときである。従来手法 3 は、どちらのケースにおいても最適情報再現率、

非不適情報適合率共に高い結果が得られる。したがって、どちらのケースに応用しても効果のあるフィルタリングを行えるといえる。提案手法も、どちらのケースに対しても効果のあるフィルタリング効果を得られるといえる。ただし、従来手法3よりは劣る結果となった。

次に、各手法に関して必要な入力データ収集の容易性という観点から考察を行う。上述したように最も効果的なフィルタリングを行える従来手法3は、各メンバーの嗜好スコアのほかに学習に用いるデータが必要になる。この学習に用いるデータはグループ毎に必要なため、膨大な量のアンケートをユーザに課すことになる。そのため、ユーザにかかる負担は他の3手法に比べて大きいと考えられる。一方、次に良い結果が得られた提案手法は、各メンバーの嗜好スコアとグループの行動履歴が必要となる。しかし、グループの行動履歴は環境さえ整えば自動的に収集・蓄積できる可能性がある[4][5]。したがって、従来手法3に比べてデータの収集は若干容易になる可能性がある。また、従来手法2は、グループの行動履歴のみが必要となる。これは、提案手法と比較すると、各メンバーの嗜好スコアが必要ないため、データ収集が容易といえる。最後に、従来手法1は、各メンバーの嗜好スコアが必要となる。グループに関する情報は必要なく、各個人の情報を利用可能なため、4つの手法の中で最もデータ収集が容易であるといえる。

以上のことを踏まえると、必要な入力データ収集の容易性とフィルタリング効果の大きさはトレードオフの関係にあるといえる。したがって、サービス等に応用する際には必要精度とユーザの負担を考慮して、どの手法が最も適しているかを上述した考察に基づき判断する必要がある。

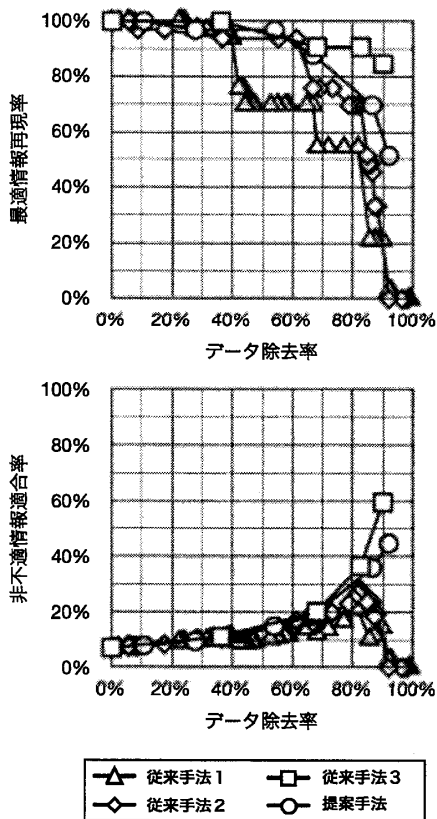


図3 ケース1における各手法の結果

6. まとめと今後

本研究では、グループに適した情報を提供する際のフィルタリング手法について述べた。その中でグループの関係性に着目し、グループに適した情報の判定を嗜好空間上における行動履歴コンテンツ情報との類似性を利用して行う手法を提案した。検証実験の結果、提案手法はグループにとって最適な情報の抽出を重視するケースおよび不適な情報の除去を重視するケースどちらにおいても効果がある可能性を示した。そして、従来手法と比較して各手法の特性を考察した。

今後は、今回の考察結果の信頼性を高めるために、より多くの被験者を用いて検証を行う予定である。また、必要データの取得容易性とフィルタリング効果の有用性はトレードオフの関係にあるという新たな課題が発見されたため、その課題の解決にも取り組んでいく。

参考文献

- [1] 伊藤 浩二, 他, “行動支援サービスのためのユーザ理解モデルの検討”, 信学技報, Vol. 109, No. 272, 2009.
- [2] 瀬古 俊一, 他, “集団へのレコメンドに適用可能な同行者との影響度合い定量化手法”. IPSJ-SLDM10144048, Vol.2010-SLDM-144 No.48, pp. 1-8, 2010.
- [3] 井上淳子, “購買行動における同伴者の影響: 母娘ショッピングの観点から”, 産研アカデミック・フォーラム, vol.13, p29-40, 2005.
- [4] 厩本 純一, 他, “PlaceEngine:実世界集合知に基づく WiFi 位置情報基盤”, インターネットコンファレンス 2006, pp.95-104, 2006.

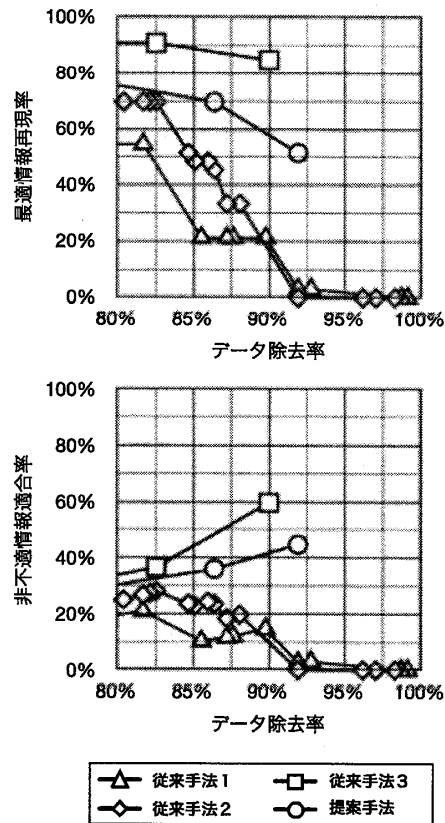


図4 ケース1における各手法の結果 (データ除去率80%以上の領域を拡大)

- [5] 瀬古俊一, 他, “誤差情報を考慮した同行判定手法”, 情報処理学会研究報告, Vol.2008, No.110, p65-72, 2008.
- [6] Gediminas Adomavicius, et al., "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions", IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, Vol.17, No.6, pp.734-749, 2005.

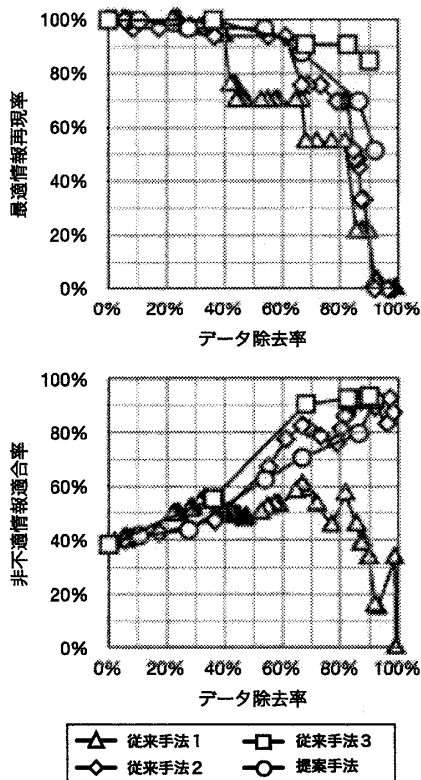


図5 ケース2における各手法の結果

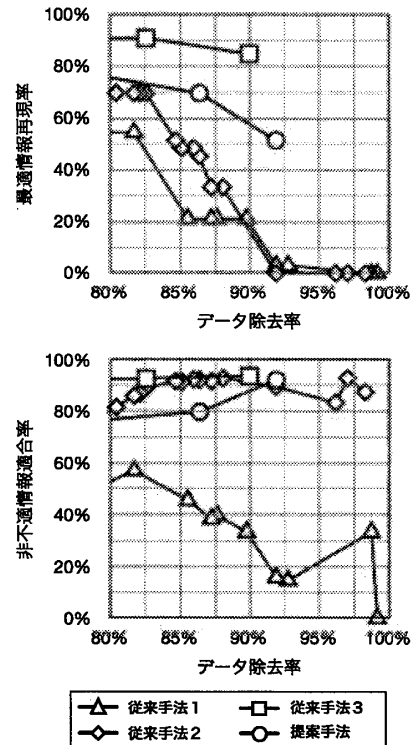


図6 ケース2における各手法の結果 (データ除去率80%以上の領域を拡大)