

# パーソナライズ情報提案エージェントの ユーザ嗜好学習に関する一考察

柴田 弘 寺元 光生

NTT 未来ねっと研究所 〒180-8585 東京都武蔵野市緑町 3-9-11

E-mail: {shiba, tera}@ma.onlab.ntt.co.jp

近年の WWW をベースとした放送技術, P2P 技術の発展により, 個人向け情報発信サービス(パーソナライズ情報提案サービス)が活発化するとともに, ネットワークに存在する膨大なコンテンツの中からユーザの嗜好に合致するコンテンツをカスタマイズする技術が注目されている. 本稿では, ユーザの代理人として, ユーザからの明示的な指示がなくても, ユーザ嗜好に合致するコンテンツを自律的に探索して提案し, 当該コンテンツ利用中のユーザの挙動から自動的にユーザ嗜好を学習・推測するユーザエージェント(UA)を実現するための, ユーザ嗜好学習アルゴリズムを提案する. さらに, 本アルゴリズムの有効性を実験により検証するとともに, 学習方法のさらなる高精度化手法を提案する.

## A study on learning user-preference for personalized information recommending agents

Hiroshi SHIBATA and Mitsuo TERAMOTO

Network Innovation Laboratories, NTT Corporation 3-9-11 Midori-cho, Musashino-shi, 180-8585 Japan

E-mail: {shiba, tera}@ma.onlab.ntt.co.jp

These days, personalized information recommending services are attracting world wide attention due to the development of broadcasting technology based on WWW and P2P technology. In this paper, we propose a method of implementing our active autonomous agents that learn user's preference, discover and recommend contents matched with his/her preference without his/her guidance. We make an experiment on our learning algorithm of the agents and propose the way to improve the method of learning user-preference.

### 1. はじめに

近年の WWW をベースとした放送技術, P2P 技術の発展により, 誰もが容易にコンテンツ提供者になることができる. しかし一方で, 膨大なコンテンツの中からユーザが自身の嗜好に合致するコンテンツを選択的に取得することが困難になっている. そのため, ネットワークに存在する膨大なコンテンツの中からエンドユーザ(以後, 単にユーザと呼ぶ)の嗜好に合致するコンテンツを選出し, カスタマイズして提案するサービスが注目されている. このようなパーソナライズ情報提案サービスとして, ユーザの好みに応じてカスタマイズ可能なインターネットラジオ[1]やエージェントテレビ[2], インタラクティブ TV[3]等が実現されている. しかし, これらのサービスにおいては, ユーザから入力された好み情報や選択内容を基にユーザに提供するコンテンツを決定していくため, ユーザが事前知識(音楽や映画等のコンテンツのタイトルやジャンル, 出演俳優 / 演奏

アーティスト等)を持っている必要があり, またコンテンツ鑑賞時にユーザが常にさまざまなデータ入力やメニュー選択等を行うことが求められる. ユーザに対してサービス利用時に負担がかかっている.

このような背景から, 著者らは, ユーザ嗜好を学習して, ユーザ嗜好に合致するコンテンツを探索・カスタマイズ・提案等を行うユーザエージェント(UA), および, コンテンツのメタ情報を学習し, メタ情報に基づいてコンテンツ配信等を行うコンテンツエージェント(CA)を提案した[4][5][6]. アプリケーションは, UA および CA に対して学習の材料となる情報を入力し, その学習結果から UA および CA が提案するコンテンツ等を表示・組み合わせ提供等を行う. 例として, ユーザ嗜好に合致する動画コンテンツ等を組み合わせる番組シナリオを構築し, 提供するプログラマブルムビーサービス(PM サービス)等がある.

本報告では, 上記のようなコンテンツ提案サービス実現のため, ユーザが意識して自身の嗜好情報や, コンテンツ検索のためのキーワードを入力

しなくても、コンテンツ利用中のユーザの自然なアクション(映像等のコンテンツをクリップの最後まで鑑賞,映像クリップのスキップ/巻き戻し,表示サイズ変更等)から,自動的にユーザ嗜好を学習し,ユーザの負担を軽減することが可能な,ユーザ嗜好学習方法を提案し,UAの実現方法を説明する.さらに,本UAおよびサンプルアプリケーションとしてのPMサービスを用いた実験による学習結果の検証と,その結果に対する考察を述べる.

## 2. パーソナライズ情報提案エージェント

本節では,UAおよびCAの概要と,これらのエージェントのパーソナライズコンテンツ提案サービスへの適用について説明する.

UAはユーザの代理人としてユーザ嗜好に基づいてふるまうピアであるため,ユーザおよびアプリケーションからユーザ嗜好学習の材料となるさまざまな情報を収集し,ユーザ嗜好を管理しなければならない.そこでUAは,ユーザ嗜好を,ユーザの興味(好きな俳優,ジャンル),プロフィール(年齢,住所),環境(時刻,所持金)等のマルチ属性(以後これらをユーザ嗜好属性と呼ぶ)に分類することにより,詳細に表現する.すなわち,表1のようなマルチ属性から成るユーザ嗜好属性を扱うものとする.マルチ属性は,表に示した構成で,適用するサービスに応じて新たな他の属性を使用可能である.次に,UAは,ユーザ嗜好に合致するコンテンツを,SIONet[7][8][9]等の探索ネットワークを用いて発見することが求められる.このとき,ユーザ嗜好属性とマッチング可能なコンテンツ属性がコンテンツに付与されている必要があるため,CAは表2のような構成のコンテンツ属性を扱う.その後,UAは探索により発見したコンテンツ群から,ユーザに提案するに相応しいコンテンツのみを抽出し,パーソナライズする.すなわち,UAは,ユーザ嗜好に応じてコンテンツ群をカスタマイズする.さらに,コンテンツの積極的な提案が求められるサービスにおいては,ユーザに相応しいコンテンツを,UAが自動選択・表示する等の機能も求められる.

次に,ユーザ嗜好に応じたコンテンツのカスタマイズおよび提案を実現する方法について述べる.コンテンツ属性およびユーザ嗜好属性をUA,CA,コンテンツ提案アプリケーションが共通に把握できるデータ形式に変換したものを,それぞれ,コンテンツキーワード(CK),ユーザ嗜好キーワード(UPK)と呼ぶ.UAは発見したコンテンツ群をパーソナライズするため,CKとUPKを照合し,コンテンツ属性とユーザ嗜好属性との合致度を算出して,その得点の高いコンテンツの順にランク付けする.このとき,得点の算出にはUPKとCKの属性を構成するベクトル(小項目およびその重み値)の内積等により,属性毎の合致度を求めることができるが,さらにUAは,マルチ属性を総合的に判断する必要がある.例えば,発見したコンテンツ群のうち,1つはユーザの俳優好みによく合致するが他の属性に関しては合致度が低く,別のコンテンツはユーザのジャンル好みに合致

する等の状況になった場合,どちらのコンテンツが,ユーザにとって相応しいのか判断するのは困難である.そこで,本論文で提案するUAは,マルチ属性なユーザ嗜好属性のうち,それぞれがどの程度重要な要素であるかという情報を持つ.すなわち,ユーザ嗜好の各属性(大項目)に重み付けを行う.大項目重み値の学習に関しては強化学習[10]に基づいた学習方式により実現可能であることを文献[4][5][6]において示した.そのため,本報告では,小項目値および小項目重み値に関するUAの学習方法について述べる.

表1 ユーザ嗜好属性  
Table 1 User-preference attributes

属性(大項目)		小項目	
Actor	63	Tom Cruise	81
		Brad Pitt	36
		:	:
Genre	22	Horror	74
		Action	48
		:	:
Price	34	less than 20 yen	/
:	:	:	:

表2 コンテンツ属性  
Table 2 Contents attributes

属性(大項目)		小項目	
Actor		Brad Pitt	70
		Meg Ryan	50
		:	:
Genre		Horror	80
		Comedy	30
		:	:
Price		10 yen	/
:	:	:	:

## 3. 基本学習アルゴリズム

本章では,ユーザ嗜好属性における小項目(詳細なユーザ嗜好情報)毎の重み値の学習方法を示す.映画等のビデオクリップを対象としたサービスであれば,俳優(Actor)好み,およびジャンル(Genre)好み等の属性があり,さらに俳優好み属性の小項目として“Tom Cruise”,“Brad Pitt”,“Meg Ryan”等のように詳細な項目が存在する.この項目が小項目であり,それぞれの好み度を示す付加情報として重み値を付与することにする.(音楽クリップであれば,俳優好みの代わりにアーティスト好み,といった属性が存在することとなる).このような小項目の学習方法として,ユーザの行動履歴に基づいたユーザモデリング手法をベースに,ユーザに対する負担を小さくし,かつ,ユーザからの学習材料となる情報を効果的に収集するため,コンテンツ利用中のユーザの自然なアクション(映像等のコンテンツをクリップの最後まで鑑賞,映像クリップのスキップ/巻き戻し,表示サイズ変更等)を用いることにする.

以下の式では,学習回数すなわちユーザのアクションに基づくフィードバック数を  $n$  と表し,当該アクション( $n$  回目のユーザのアクション)に基づく評価値(フィードバック値)を  $V_n$  と表す.ユーザの

俳優好み属性の小項目“ $Actor_j$ ”に関する重み値を  $Au_n(Actor_j)$  , 値域を  $-100 \leq Au_n(Actor_j) \leq 100$  とし, 正の値は興味が深い(ファンである)ことを意味し, 0 は興味なし, 負の値は嫌悪していることを表すとする. また, ユーザが鑑賞したコンテンツに付与された CK における“ $Actor_j$ ”の重み値を  $Ac(Actor_j)$  , 値域を  $-100 \leq Ac(Actor_j) \leq 100$  とし, 正の値は“ $Actor_j$ ”がファンに好評な役柄で当該コンテンツに出演していることを示し, 0 は出演無し, 負の値はファンに不評な役柄で登場していることを表すとする. UPK における小項目重み値  $Au_n(Actor_j)$  は, コンテンツ鑑賞中のユーザのアクション等に基づいて(1)式を用いて随時更新することができる.

$$Au_n(Actor_j) = Au_{n-1}(Actor_j) + Ac(Actor_j)\mu V_n \quad (1)$$

ここで  $\mu$  は, 現在のユーザの好みを UA の学習データに加算する際の加算率を意味し, ユーザのアクションに基づく  $V_n$  を調整するためのパラメータである. なぜならば, ユーザ毎にスキップ等の操作に対する意味が異なると考えられ, 例えば, 好き嫌いにかかわらず頻繁にスキップをするユーザ(頻繁に TV のチャンネルを変更するユーザと同様)は, スキップ一回あたりの評価フィードバックを弱くするなどの調整が必要となるからである. 一方, 減多にスキップなどの操作を行わず, UA によって提案されるがままのコンテンツを鑑賞するユーザであれば, 最後までビデオクリップを見たとしても本当に当該コンテンツを好んでいるかどうか判断できないため,  $\mu$  を小さくして UA の学習データへのフィードバックを低減するようにする必要がある.

また, 上記小項目重み値更新時に, 最大値あるいは最小値を越えた場合の調整方法として, 重み値の絶対値の最大値  $Au_{max_n}$  を算出する.

$$Au_{max_n} = \max \{ |Au_n(Actor_j)|; \forall Actor_j \} \quad (2)$$

この値が, 100 を超えた場合, 次式で算出される重み値調整係数  $\eta$  を用いて, 全俳優に対する重み値(全ての小項目“ $Actor_j$ ”に対する小項目重み値)に乗算する.

$$Au_{max_n} > 100 \Rightarrow \eta = 100 / Au_{max_n} \quad (3)$$

$$Au_n(Actor_j) \leftarrow \eta Au_n(Actor_j) \quad (4)$$

ここで記号“ $\leftarrow$ ”は, 左辺のパラメータ値を右辺の計算結果に更新することを意味する. 以上の計算式を使用した学習方式により, ユーザが自身のユーザ嗜好属性のさまざまな重み値等を意識しなくても, ユーザのコンテンツ鑑賞や鑑賞時の挙動から, それらの値が UA によって自動的に学習され設定される. 以上を, 本提案 UA の基本学習アルゴリズムとする.

#### 4. 学習アルゴリズムの評価

本章では, UA・CA を適用して実現できるアプ

リケーションとして, プログラブルムービーサービス(PM サービス)を用いて, 3章で示した UA のユーザ嗜好基本学習アルゴリズムの妥当性を検証する. 検証方法としては, UA が, PM サービス利用中のユーザの挙動から, ユーザ嗜好属性小項目(詳細なユーザ嗜好情報)およびその重み値を学習する. この学習結果と, ユーザからアンケートで得た理想的なユーザ嗜好の値とを比較し, 評価する.

##### 4.1. PM サービス

PM サービスとは, ユーザ嗜好に応じて映像番組の構成を組み立て, ユーザ毎に内容をカスタマイズした映像番組を提供するサービスであり, ユーザの嗜好に応じてストーリーの変化するマルチナリオドラマや, 好みの音楽や映像クリップのみを集めたお気に入り音楽/映像集等を提供可能である. このサービスは, ユーザの興味(好きな俳優, ジャンル), プロファイル(年齢, 住所), 環境(時刻, 所持金)等を総合的に判断し, 番組の材料として利用可能なコンテンツ(ビデオクリップ, テキストデータ, 静止画像データ等)の中から, ユーザに提供するに相応しいものを自動的に特定し, ユーザにとって TV 鑑賞をしているかのように, コンテンツを連続的にユーザに提供し, 表示していく.

PM サービスシステムは, SIONet と, UA および CA を用いて実現される. PM サービスアプリケーションは, ユーザ(視聴者)ピアにおいては, UA に学習材料を入力し, UA の提案するコンテンツを利用して, 多種多様なコンテンツの組み合わせで実現される映像番組(リッチコンテンツ)を表示する. 一方, コンテンツ提供者ピアにおいては, 動画やテキスト, 静止画等のさまざまなメディアのコンテンツを管理し, UA からの要求に応じて CA がユーザピアにコンテンツを配信する. 図 1 は, ユーザピアにおける PM サービス表示画面である. ユーザはこのようなアプリケーション画面に表示されるリッチコンテンツを鑑賞し, 好みに応じて“そのまま鑑賞し続ける”, “画面サイズ変更”, “スキップ”等の操作を行う. このとき, PM サービスアプリケーションは, この画面に対するユーザの操作アクションを UA に対して入力する. この操作情報は UA のユーザ嗜好学習材料となる. PM サービスにおいて, 鑑賞中のコンテンツに対するユーザの評価値  $V$  とユーザの挙動に関して, 表 3 のような対応関係を定義した.

UA によって提案された映像コンテンツ(映像番組の 1 シーン, あるいは, 音楽のビデオクリップ 1 曲分)をユーザが最後まで見た場合は  $V=1.0$  が UA にフィードバックされる. PM サービスは, TV 放送のように, ユーザがあまり意識して操作入力を行わなくてもコンテンツが連続的に提供されるサービスであるため, 一般的にこの評価になることが多いと考えられる. ただし, ユーザの興味に合致しないシーンが登場した場合, ユーザによってスキップ等のアクションが行われ, このとき  $V=-1.0$  の評価を与えることにする.



図 1 PM サービス表示画面  
Fig. 1 PM service player

表 3 ユーザの挙動に基づくフィードバック情報  
Table 3 Feedback based on user's action

ユーザの挙動	評価値 V
鑑賞中コンテンツを巻き戻して再鑑賞	1.0
鑑賞中コンテンツを最後まで見終える	1.0
鑑賞中コンテンツの表示画面サイズを拡大	0.5
コンテンツ鑑賞中に他の APL を前面に表示	- 0.2
鑑賞中コンテンツの表示画面サイズを縮小	- 0.5
鑑賞中コンテンツをスキップまたは途中終了	- 1.0

## 4.2. 実験方法と結果

次に、PM サービスを用いたユーザ嗜好学習実験方法を以下に示す。

- ・ 被験者は約 1 時間、PM サービスを利用して音楽ビデオクリップを鑑賞する。鑑賞中の操作(クリップのスキップおよび巻き戻し、最後まで見終える等)は、被験者の意思に任せる。
- ・ PM サービスで提供するコンテンツは、さまざまなジャンルの音楽をメインとしたビデオクリップであり、1クリップあたり 3分~4分に収まるもののみである。また、1クリップが1曲に相当する。本実験で用いたコンテンツは全 40 種類である。
- ・ コンテンツ(音楽クリップ)に付与するメタ情報には、コンテンツ属性小項目を、当該音楽クリップに相応しいと思われる音楽ジャンルとし、その小項目重み値は一律 70 の値を与えた。例えば、ある音楽クリップが、“Rock”と“Rap”の要素が盛り込まれているならば、当該コンテンツのコンテンツ属性には“Rock”と“Rap”の項目名で示される小項目が存在し、各小項目の重み値は一律 70 である。
- ・ PM 利用後、ユーザから理想的なユーザ嗜好属性情報(具体的には、表 1 の“Genre”属性の小項目およびその重み値)をアンケートで収集する。この結果を  $\vec{P}_R$  とおく。
- ・ PM 利用中に学習した UA の学習結果を  $\vec{P}_A$  と

おき、 $\vec{P}_R$  との類似度  $\delta$  を以下の式で求め、学習結果の妥当性を比較・評価する。

$$\delta = \frac{\vec{P}_R \cdot \vec{P}_A}{\|\vec{P}_R\| \|\vec{P}_A\|} \quad (5)$$

このような実験を、17 人の被験者に対して行った。なお、実験用に準備したコンテンツは全 40 種類であるため、実験中に同一コンテンツが複数回ユーザに提供されることもしばしばあった。表 4 に典型的な被験者の挙動として、実験中の操作アクション回数を示す。

表 4 に示した被験者は、“最後まで見る”等のプラス評価をフィードバックする挙動に比較して、“スキップ”といったマイナス評価フィードバック回数が多い。特にスキップに関しては、過去に見たコンテンツは全てスキップしている。この被験者は一例であるが、他の被験者もこのユーザと同様の挙動を示した。

表 4 被験者の挙動  
Table 4 The behavior of a test user

配信されたコンテンツ数	48	初めてのコンテンツ	37
		過去に見たコンテンツ	21
最後まで見た回数	15		
スキップした回数	33	初めてのコンテンツに対するスキップ	12
		過去に見たコンテンツに対するスキップ	21
その他(巻き戻し、表示サイズ変更等)の回数	0		

この実験データを基に UA の学習結果  $\vec{P}_A$  を算出し、アンケートにより得た  $\vec{P}_R$  との類似度  $\delta$  を算出する。ここで、(1)式における  $\mu$  の値を一律 (“スキップ”に対しても、“最後まで見る”に対しても同じ値)  $\mu = 0.5$  として算出したところ、類似度  $\delta = -0.20$  となった。すなわち、学習結果はユーザ嗜好を適切に学習していないこととなる。そこで、(1)式における  $\mu$  の値を“スキップ”に対する調整値  $\mu_{Skip}$ 、および“最後まで見る”に対する  $\mu_{Play}$  を変動させ、その値に基づいた学習値を用いて類似度を計算した。この計算結果を図 2 に示す。

図 2 のグラフにおいて、UA 学習の調整値と類似度の関係を視覚化するため、調整値  $\mu_{Skip}$  と  $\mu_{Play}$  を(6)式で表される一定の関係を持たせて  $\mu_{Skip}$  および  $\mu_{Play}$  の値を変化させた。図 2 の横軸は、(6)式の関係を満たす  $\mu_{Play}$  の値であり、 $\mu_{Play}$  の値を大きくするほど  $\mu_{Skip}$  は小さくなる。すなわち、グラフの右へいくほど、“最後まで見た”に対するプラス評価フィードバックを強くなり、逆に左へ行くほど、“スキップ”に対するマイナス評価フィードバックが強くなる。

$$\mu_{Play} + \mu_{Skip} = 1.0 \quad (6)$$

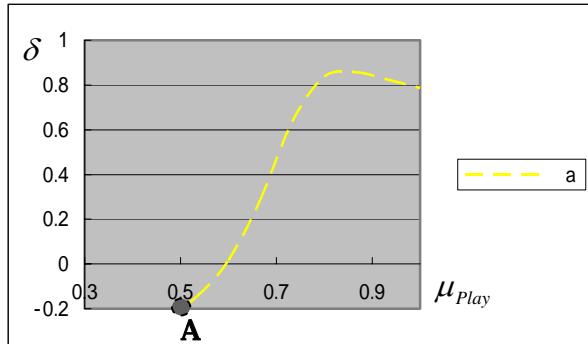


図 2 UA による学習結果の類似度

Fig. 2 Similarity between  $\bar{P}_R$  and  $\bar{P}_A$ .

図 2 において、点 A は、 $\mu_{Play}$  と  $\mu_{Skip}$  の値を等にして UA 学習値  $\bar{P}_A$  を算出したときの、 $\bar{P}_R$  との類似度であり、その値は  $\delta = -0.20$  である。これに対し、 $\mu_{Play}$  の値を大きくすると、学習結果が理想値に近づくことが分かる。この被験者はスキップ回数が 33 回であるのに対し、最後まで見た回数が 15 回であり、マイナス評価フィードバック回数が多く、そのため、学習された小項目の重み値の多くがマイナスの値となるが、アンケートで得たユーザの理想値では、マイナスの項目よりもむしろプラスの項目が多いため、学習値と理想値とは正反対の結果となってしまう。これは、スキップによるマイナス評価フィードバックの影響が強くなり、プラス評価の学習結果が目立たなくなってしまうことを意味する。そこで、スキップ操作に対するマイナスフィードバックの強さを調整するため、スキップ操作の要因と意味を分析すると、以下のように考えられる。

- 一度見たコンテンツをもう一度見るよりも、他のコンテンツを見たいと思うユーザがいる。(表 4 の被験ユーザには、過去に見たコンテンツが表示されたことが 21 回あり、これらに対して当該ユーザは全てスキップしている。)

- 音楽等のコンテンツにはさまざまなキーワード(小項目)で示される(さまざまなジャンルが融合されている)ものも多く、好きなジャンルの要素があっても、好きでは無い要素が含まれているとスキップされることがある。(例えば、純粋な“Rock”が好きなユーザであれば、“Rock”と“Rap”の融合した音楽に対して、スキップすることがあり得る。この場合、“Rock”と“Rap”の小項目に対してマイナス評価がフィードバックされる。)

- 映画やドラマ等の番組に比べて、音楽集のような番組は、映像クリップ(シーン)間の繋がりが弱く、1つ1つのクリップが独立しているため、あるクリップをスキップしても番組全体としてストーリーが分からなくなる等の問題が無い。

このような“スキップ”の要因や頻度はユーザにより異なるため、次節では、ユーザに応じたフィードバック値の調整方法について検証する。

### 4.3. 学習フィードバック値調整とその効果

前節の結果より、学習パラメータ  $\mu_{Skip}$  および  $\mu_{Play}$  をユーザおよびコンテンツに応じて変動させ、ユーザ嗜好を学習する必要があると考えられる。そこで、これらの学習パラメータの調整方法を全被験者の学習結果を参考に導出することにする。調整方法の導出のため、次のような調整を施し、その効果をグラフ化により視覚的に検証した。

(調整 1) 一度表示したコンテンツを再度表示するとスキップするユーザが多いため、二度以上表示されたコンテンツに対する“スキップ”の評価フィードバックを低減する。このような条件の下で、さらに図 2 と同様の方法で、“スキップ”に対する調整値  $\mu_{Skip}$ 、および“最後まで見る”に対する  $\mu_{Play}$  を変動させ、その値に基づいた学習値を用いて類似度を計算したところ、図 3 のようなグラフが得られた。

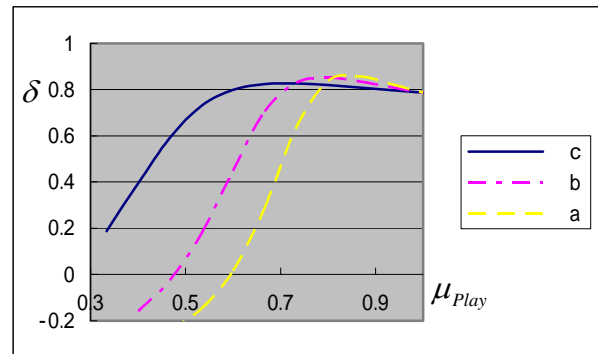


図 3 調整 1 の効果

Fig.3 Effect of the adjustment 1.

図 3 において、 $\delta_a$  は、図 2 のグラフと同様、二度以上同じコンテンツを表示した際のスキップに対する評価フィードバックを低減しないパターンである。これに対して  $\delta_b$  は、同じコンテンツに対するスキップの評価を半分に低減した場合であり、 $\delta_c$  は、前記のスキップ評価を無効化した場合の類似度である。このグラフより、同じコンテンツを表示した際に起こるスキップによるマイナス評価を低減するほど、類似度が高くなる  $\mu_{Play}$  の値域が広がる。例えば、 $\delta_a > 0.8$  を満たす値域は  $0.78 < \mu_{Play} < 0.98$  であるが、 $\delta_c > 0.8$  を満たす値域は  $0.60 < \mu_{Play} < 0.92$  となる。したがって、過去に表示したコンテンツの再表示に対するスキップ評価を低減することにより、UA の学習結果を理想値に近づけるための  $\mu_{Play}$  の調整が容易になり、また、ユーザの操作におけるノイズ(好み以外の原因による操作アクション)の影響から学習結果が著しく悪くなることを避けることができる。

(調整 2) 図 4 のようなグラフにおいて、類似度  $\delta_c$  がピークになるときの  $\mu_{Play}$  および  $\mu_{Skip}$  と、当該ユーザのスキップ率(全操作アクション回数のうち、スキップ回数が占める割合)  $P_{Skip}$  の関係を求める。ただし、過去に表示したコンテンツの再表示に対するスキップは学習の対象外であるため、ここでの  $P_{Skip}$  の算出に用いるスキップ回数には、この回数を除外することとする。被験者 17 人分のこの 2 つのパラメータをプロット(横軸に  $P_{Skip}$ 、縦軸に  $\mu_{Play}$  および  $\mu_{Skip}$ )したところ、図 4 のようなグラフが得られた。

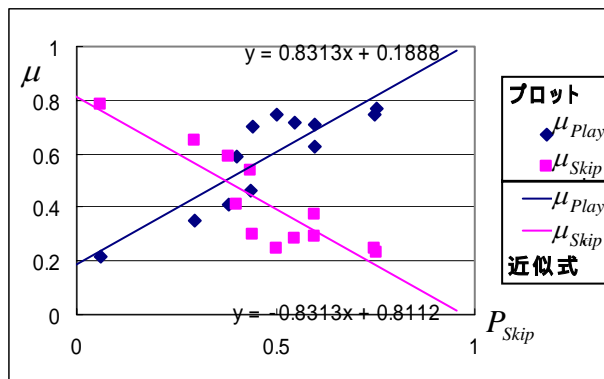


図 4 スキップ率とフィードバック調整値の関係  
Fig. 4 Relation between  $P_{Skip}$  and  $\mu$ .

図 4 に示されるように、被験者のスキップ率  $P_{Skip}$  が高いほど、UA の学習結果を理想値に近づけるためには、 $\mu_{Play}$  を大きく(すなわち  $\mu_{Skip}$  を小さく = 1 回あたりのスキップによるマイナス評価フィードバックを弱く)することが効果的である。図 4 から求められる近似式はおおよそ以下のようになる。

$$\mu_{Play} = 0.8P_{Skip} + 0.2 \quad (7)$$

$$\mu_{Skip} = 1.0 - \mu_{Play} = -0.8P_{Skip} + 0.8 \quad (8)$$

この結果より、(1)式の学習方法における、フィードバック調整値  $\mu$  として、“最後まで見る”に対しては(7)式で求められる  $\mu_{Play}$  の値を、“スキップ”に対しては  $\mu_{Skip}$  の値を用いて学習結果  $\bar{P}_A$  を算出し、理想値  $\bar{P}_R$  との類似度を求めた。なお、“巻き戻し”操作は、“最後まで見る”と同様、プラス評価フィードバックになるため、“巻き戻し”に対するフィードバック調整値  $\mu_{Rewind}$  は  $\mu_{Play}$  と同じ値とした。その他の操作(画面サイズ変更等)は、本実験では対象外とした。

以上の条件の下での全被験者に対する学習結果の類似度は、平均 0.84、標準偏差 0.08 となり、安定して高い値が得られた。一方、フィードバック値の調整を行わずに学習した結果の類似度は、平均 0.41、標準偏差 0.45 であった。

この結果より、フィードバック値の調整を行わ

ない場合、ユーザによっては学習結果が理想値から著しく離れてしまうことがある。(例えば、過去に見たコンテンツに対するスキップが多いユーザに対して顕著に現れる)。これに対して、提案方式のように、ユーザのスキップ率や過去に表示したコンテンツかどうかのチェックを基にした本提案のフィードバック値調整方法を適用すると、UA の学習結果は、全被験者に対して理想値と類似度を高くでき、大きく離れることが無く安定させることができた。

## 5. おわりに

本報告では、コンテンツ鑑賞中のユーザの挙動からユーザ嗜好を学習する UA の基本学習アルゴリズムを提案し、この基本アルゴリズムを実験により検証した。さらに実験結果から、本学習アルゴリズムの学習過程におけるフィードバック値(コンテンツ評価値)をユーザの操作頻度およびコンテンツに応じて調整する方法を提案し、その効果を示した。今後はさらなるユーザの特徴的な挙動を実験等により抽出し、学習アルゴリズムを洗練させる予定である。

## 謝辞

研究を進めるにあたって、有益なご指導・ご意見をくださった NTT 未来ねっと研究所および NTT ネットワークサービスシステム研究所の皆様へ深く感謝いたします。

## 文献

- [1] インターネットラジオ: <http://spinner.com>
- [2] 村崎康博, 金淵培, 柴田正啓, 浦谷則好: “ユーザモデルエージェントによる番組選択システム”, 信学技報 AI2001-50, pp.25-31, Nov. 2001.
- [3] MSN TV: <http://www.webtv.com>
- [4] H. Shibata and T. Hoshiai: “Method for personalized information recommending agents to learn user preference for programmable movie service”, APCC2001, pp.201-204, Sep. 2001.
- [5] 柴田弘, 星合隆成: “パーソナライズ情報提案サービスにおける自律エージェント”, 信学技報 KBSE2001-56, pp.53-60, Jan. 2002.
- [6] H. Shibata, T. Hoshiai, M. Kubota and M. Teramoto: “Agent technology recommending personalized information and its evaluation”, IWADS2002, pp.176-183, Nov. 2002.
- [7] 星合隆成, 柴田弘: “御用聞き社会構築に向けてのコンテンツ情報流通網と自律分散照合環境アーキテクチャ”, 信学会論文誌 D-1, Vol.J83-D-1, No.9, pp.1001-1012, Sep. 2000.
- [8] 星合隆成: “意味情報ネットワーク SIONet におけるエンティティのオンライン増減設機構”, 信学会論文誌 B, Vol.J85-B, No.2, pp.180-199, Feb. 2002.
- [9] 星合隆成: “P2P の理念および実現技術:SIONet の全貌”, P2P Conference in JAPAN 2002 Spring, April 12 2002, <http://www.p2pconf.com>
- [10] R. S. Sutton and A. G. Barto: “Reinforcement Learning - An Introduction”, MIT Press, 1998.