

## 人間の基本的欲求を利用した異なるサービスをつなぐユニバーサル・プロファイルの提案と評価

### Proposal and Evaluation of Universal Profile Based on Human's Basic Desires

秋山 高行<sup>†</sup> 寺本 やえみ<sup>†</sup> 小原 清弘<sup>†</sup> 谷崎 正明<sup>†</sup>  
Takayuki Akiyama Yaemi Teramoto Kiyohiro Obara Masaaki Tanizaki

#### 1. 背景と目的

近年、インターネットに代表される情報通信技術の普及とその商業利用の拡大により、ユーザがアクセスすることが可能な情報やコンテンツ（本やテレビ番組、観光のようなサービスも含む）は、多様化・大量化している。こうした状況では、ユーザは、大量の無用なコンテンツの中から自分にとって有用な情報を選択しなければならず、ユーザのコンテンツ選択負荷が増大しているという問題がある。一方、選択可能なコンテンツの幅が広がると共に、ユーザのコンテンツ選択理由や行動形態の多様化が加速している。そのため、コンテンツ提供者においては、多様化するユーザニーズの把握が困難になり、従来のような不特定多数へ向けた広告や新製品開発におけるマーケティングリサーチの効果が大幅に減少しているという問題がある。

このような背景のもとで、個々のユーザの行動をモニタリングしてプロファイル（嗜好などの特性情報）を抽出し、ユーザに最適なコンテンツを提示する推薦システムの必要性が、コンテンツ提供サービスにおいて拡大している。

推薦システムとは、概して、2つのモジュールで構成される（図1）[1]。一つは、ユーザの行動履歴から選好に影響する因子（プロファイル）を抽出するプロファイル抽出するモジュールである。このモジュールによって、ユーザが「何を求めているのか」を判定する。もう一つは、コンテンツとプロファイルの適合度を計算し、ユーザに適したコンテンツを推薦する推薦エンジンである。このモジュールによって、ユーザの要求に適したものを提供する。

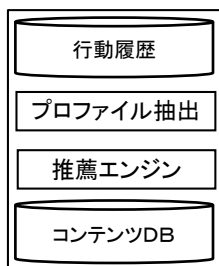


図1. 推薦システム

従来の推薦システムでは、各サービス事業者が独自にシステムを構築し、当該システム上でのユーザの行動履歴を蓄積し、ユーザのプロファイルを抽出する。例えば、テレビ番組の推薦システムにおいては、ユーザの所有するテレビでの視聴履歴を蓄積し、また、ECサイトにお

いては、サイト内での商品選択・購買履歴を蓄積する。このようなシステムでは、ユーザの単一サービスの行動履歴しか蓄積されないため、ユーザの行動の一側面しか捉えられない。しかし、ユーザの生活行動は図2のように幅広く多岐にわたる。ユーザの行動の一側面だけを捉えることでは、ユーザの特性を十分に表現できるプロファイルを抽出することは困難である[2,3]。

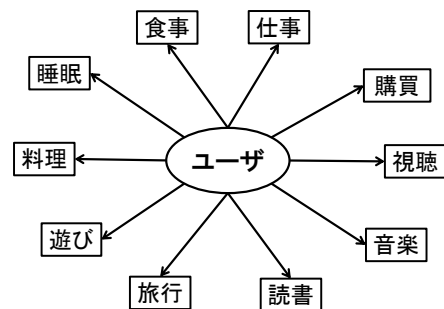


図2. 人間の生活行動の例

その一方で、近年、スマートフォンの普及や電子書籍端末の登場、インターネット対応テレビの普及に代表されるように、ユーザの利用する機器の情報化・ネットワーク化が進んでいる。また、EC（Electronic Commerce）サイトの市場拡大や電子決済の普及に伴い、ユーザの行動履歴の電子化が急速に進んでいる。一つの端末、もしくは、アカウントで、図2のような幅広い行動の履歴を収集可能になっている。

そのため、複数のサービスをまたがるユーザ行動履歴から各サービスで共通利用可能なユーザプロファイルを抽出し、提供するような仕組みが必要である。

そのような仕組みを実現するためには、全てのサービスの行動に影響するユーザの特性を表現するユーザプロファイルが必要である。そのようなユーザプロファイルを、本研究では「ユニバーサル・プロファイル」と呼ぶ。

ユニバーサル・プロファイルを各サービスで共通利用する仕組みとして、筆者らは図3のような構成を提案する。ユーザの全行動履歴を統合して管理し、ユニバーサル・プロファイルを抽出するシステムを、各サービスシステムとは独立に構築する。これを以下では、ユニバーサル・プロファイラ・システムと呼ぶ。各サービスでは、ユニバーサル・プロファイラ・システムとデータ交換する共通インタフェースを実装すればよく、新たなシステム改変が小規模で済むことになる。共通インタフェースでは、各サービスからユーザの行動履歴を収集する際に、ユニバーサル・プロファイルを算出可能なデータへ変換し、また、抽出されたユニバーサル・プロファイルとコンテンツの適合度を算出可能なデータへ変換する機能を持つ。

<sup>†</sup> (株)日立製作所中央研究所, Hitachi, Ltd., Central Research Laboratory

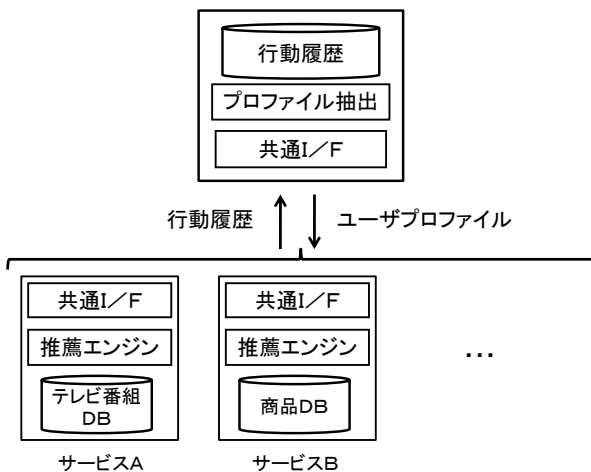


図3. 共通プロフィール提供システム  
(ユニバーサル・プロファイラ・システム)

このようなユニバーサル・プロファイラ・システムを実現するために、本論文では、以下の項目を提案・検証する。

- ユニバーサル・プロフィール
- プロファイル抽出方式
- 共通インタフェースの方式

本論文は以下のように構成される。2章で異なるサービスをつなぐプロフィール抽出に関連研究を紹介する。3章では、異なるサービス間でも利用可能な共通のユーザプロフィール「ユニバーサル・プロフィール」を提案し、それに基づく各方式を提案する。そして、4章では、小規模な評価実験とその評価結果を説明する。5章でまとめ、今後の課題を述べる。

## 2. 関連研究

異なるサービスをまたがる行動履歴からプロフィールを抽出する方式は、二つの方式が考えられる。コンテンツベース方式とユーザベース方式である[1]。

コンテンツベース方式では、コンテンツの特徴をコンテンツ情報によって定義（テレビ番組では、電子番組表に記載されているカテゴリや出演者など）、ユーザのコンテンツ選択履歴から出現頻度などの重みを利用して、ユーザが興味のある特徴を抽出する。そうして抽出された特徴を利用して、興味の高いコンテンツを提示する。そのため、従来では、コンテンツの記述がサービスに依存したものになり、一般性が無いものになる。この方式によって複数のサービスのコンテンツを定義すると、特徴数が膨大になり、さらに、サービス間を横断するような特徴の統一辞書、またはオントロジの設計が必要となる[4, 5]。そのような統一辞書やオントロジを作成するのは困難であり、方法論も明らかにならず、また、特徴数が膨大になり、計算コストも爆発的に増大する。

一方、ユーザベース方式では、コンテンツの特徴は利用せずに、複数ユーザの選択履歴を利用して、当該ユーザと選択履歴が似たユーザを抽出することにより、未選択コンテンツの興味度を推定し、興味度の高いコンテンツを提示する。この方式では、コンテンツの利用履歴によってユーザを定義することになり、ユーザの特徴数はコンテンツ数になる。連携するサービスが増えればコン

テンツ数もますます膨大になり、ユーザのプロファイルの管理のコストが増大する。また、実際のプロファイルは似ていても、コンテンツの利用履歴が異なれば、異なるプロファイルであると認識するという問題もある。

以上のように、従来方式では、ユーザをコンテンツに含まれる情報によって記述し、その内容については考慮しない。そのために、コンテンツの内容が異なるサービス間でのデータ交換は困難だった。そもそもコンテンツを利用することでユーザは何を得ているのか、ユーザは何を得ることを目的としてコンテンツを利用するのか、を考慮する必要がある。

そこで本論文では、まず、ユーザの消費者行動のメカニズムを説明する消費者行動モデルと行動に影響を及ぼすと思われる欲求理論を調査し、それに基づいたユニバーサル・プロフィールを提案する。

## 3. ユニバーサル・プロフィールの検討

### 3.1 人間の消費者行動に関する考察

ユーザの日常生活では、図2のように様々な種類の行動が存在する。しかし、あらゆる行動は、ユーザの特性（プロフィール）が影響していると考えられる。

全ての行動に影響するユーザの特性を理解するためには、人間の行動について深く理解する必要がある。本研究では、特に人間の「消費者行動」を対象とする。「消費者行動」とは、「特定の時点における特定の商品の購買行為と、それに先立ち、またそれに続く、一連の継続的な行為であり、人間の意志決定行動のうちの特定の側面」を指す。従来マーケティングリサーチや社会心理学の世界では、消費者の選好に影響を及ぼすと思われる欲求や状況の依存性、購買を決定するメカニズムに関して様々な仮説が導入されてきた[6]。ここでは、消費者行動へ影響する因子を考察するため、消費者の外的な刺激に対する反応をモデル化した基本的な「EKBモデル」

(Engel Kollat Blackwell Model) を例に考察する(図4)。このモデルでは、消費者の意志決定行動は、「欲求喚起」「情報探索」「選択代案評価」「購買行為」「購買後の評価」という5つの行為から構成されている。これらの行為の進行に影響を与える要因としては、経済的、心理的、社会的、文化的要因があると想定されている。

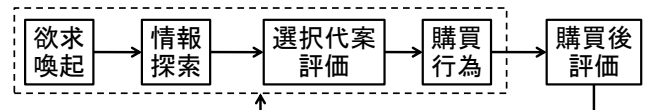


図4. 消費者行動モデル (EKBモデル)

消費者行動理論によれば、消費者の行動の起点には「欲求喚起」があり、行動の目的は「欲求を充足させること」である。消費者行動は、欲求を充足させるか否かが最も重要な条件として行われる。つまり、ユーザプロフィールを欲求の充足度によって記述し、コンテンツを充足する欲求の種類によって記述することが可能であれば、サービスの種類に関わらないユーザプロフィールを抽出することが可能である。

欲求に関しては、これまで多くの心理学者や哲学者たちが独自の理論を生み出している。本研究では、その中

で最も一般的なマレーやマズローの欲求理論、ステイブン・リースによる欲求分類をベースにした消費者心理に関わる研究の中から、十分に科学的な検証がなされている以下の19の欲求軸を採用した(表1)。

表1は、さまざまな地位や年齢の何千人もの被験者に対する心理テストによる分析によって抽出された欲求軸である[7]。本研究では、ユニバーサル・プロファイルを利用者の欲求の充足度によって記述することを提案する。次節では、基本的欲求の充足度をユニバーサル・プロファイルとした提案方式について述べる。

表1. 基本的欲求軸

①	個性・自由	自分の理想を実現したい、思い通りにしたいなど、個性・自由に関する欲求
②	所有	モノを集めたい、特別なものを所有したいなど、獲得・所有に関する欲求
③	摂取・消費	買い物や食事で大量に消費したい、こだわりたいなど、摂取・消費に関する欲求
④	活動・運動	物事を現実的に考えたい、行動を起こしたい、身体を動かしたいなど、活動・運動に関する欲求
⑤	将来展望	貯めたい、見通しを立てたいなど、将来に関する欲求
⑥	秩序	不自然なものは直したい、完璧にしたい、整頓したいなど、秩序に関する欲求
⑦	合理性	物事を進める際に合理的に行ないたいなど、合理性に関する欲求
⑧	持続	継続している事をそのままやめずに続けたいなど、継続・持続に関する欲求
⑨	回復	問題があるものを解消したい、失ったものをとりもどしたいなど、回復に関する欲求
⑩	知的的好奇心	未知のものに興味をもつ、知らないことを学びたいなど、知への好奇心に関する欲求
⑪	感性体験	美や感動、創作や遊びを体験したいなど、感性に関する欲求
⑫	非日常体験	驚きや興奮を味わいたい、挑戦したいなど、非日常性に関する欲求
⑬	愛情	愛しいと感じる、面倒をみたいなど、愛情に関する欲求
⑭	性・恋愛	異性から好かれたい、性的な世界に触れたいなど、性・恋愛に関する欲求

⑮	やすらぎ	癒されたい、ストレスからの開放、恥を予防したいなど、心のやすらぎに関する欲求
⑯	反発	他人とは反対の行動に出たいなど、反発に関する欲求
⑰	向上・達成	人から認められたい、尊敬を勝ち取りたいなど、対人的な地位の向上・達成に関する欲求
⑱	つながり	他人とつながりたい、気持ちを共有したいなど、社会的なつながりに関する欲求
⑲	社会的帰属	地域や社会を守りたい、地球全体規模で正しいことをしたいなど、社会への帰属意識に関する欲求

### 3.2 提案方式

人間の基本的欲求を利用して、図3のような共通インタフェースを構築するには、異なるサービスにまたがる行動履歴からユーザが充足したい基本的欲求を抽出するプロファイル抽出方式と、コンテンツを利用することでユーザが充足する基本的欲求を推定するコンテンツの充足欲求推定方式が必要である。このようなインタフェースが有効であれば、各サービスでは、ユーザの基本的欲求の充足度合いを交換すればよい。本節では、各提案方式について述べる。

#### 3.2.1 プロファイル抽出方式

まず、ユーザが充足したい欲求を抽出する方式について述べる。

消費者行動理論によれば、欲求とは、現実の状態と望ましい状態との間の「不均衡」ないし「ずれ」を意味する。その「ずれ」を望ましい状態に戻すために、ある特定の行動を起こす。本研究では、望ましい状態を各欲求の充足度合いの平衡化された状態(以下、欲求の平衡状態と呼ぶ)と捉え、生活行動の中で安定していくものと考える。1回の行動によって欲求は充足されているものとみなし、各欲求の充足度は充足させた回数全体の行動回数における割合によって計算する。

$$S_i = \frac{C_i}{\sum_{n=1}^{19} C_n} \quad (\text{式1})$$

$S_i$ は欲求軸*i*の充足度、 $C_i$ は欲求軸*i*の充足回数である。しかし、生活の中で充足させる頻度は、欲求によって異なる。例えば、ほぼ毎日定期的に行う食事に関わるような欲求を充足させる頻度は、社会的帰属のような欲求を充足させる頻度よりも、ユーザに関わらず高くなる。そこで、ユーザ個人の欲求の強さを表現するために、平衡状態における欲求の強さを(式2)のように、ユーザ平均との比率で表現する。

$$P_{i,k} = \frac{S_{i,k}}{\frac{1}{U} \sum_{n=1}^U S_{i,n}} \quad (\text{式2})$$



$P_{i,k}$ はユーザ  $k$  の欲求  $i$  の平衡状態での強さ、 $U$  は全ユーザ数、 $S_{i,k}$  は (式1) で計算されるユーザ  $k$  の欲求  $i$  の充足度である。

以上をまとめたプロフィール抽出モジュールの構成を図5に示す。

### 3.2.2 コンテンツの充足欲求推定方式

次に、抽出された欲求をどのコンテンツが充足するかを推定する方式について述べる。また、ユーザの欲求の充足度合いを測定するためには、コンテンツを利用することによってどの欲求が充足されたのかを推定する必要もあるため、コンテンツと欲求の関係性を推定する方式は重要である。

コンテンツを利用することによって充足する欲求を推定するために、コンテンツに含まれる情報によってユーザは欲求を充足するかどうかを判断しているという仮定を設定し、この仮定を利用した簡単な確率モデルを導入する。コンテンツをコンテンツ情報に含まれる単語によって記述し、単語が各欲求において出現する確率を計算することで、コンテンツが欲求を充足する確率を算出する (式3)。

$$\Pr(\text{desire}|\text{content}) = \prod_{\text{term} \in \text{content}} \Pr(\text{desire}|\text{term}) \quad (\text{式3})$$

$\Pr(\text{desire}|\text{content})$  は、コンテンツ  $\text{content}$  が与えられた場合に欲求  $\text{desire}$  を充足する確率を表し、 $\Pr(\text{desire}|\text{term})$  はコンテンツ  $\text{content}$  に含まれる単語  $\text{term}$  が欲求  $\text{desire}$  を充足するコンテンツに出現する確率を表す。本方式の意味するところは、各欲求を充足するコンテンツには、各欲求に特徴的な単語が含まれる、ということである。例えば、欲求3 (摂取・消費) を充足するコンテンツには、食事に関わる単語が特徴的に出現するであろう。また、一般的に高頻度で出現するような単語 (一般語) やコンテンツ情報に意味を成さない不要語は、各欲求において同程度の確率で出現すると考えられるため、欲求推定確率に陽に表れないであろう。

さらに、19の欲求のうちのいずれかを充足するコンテンツに必ず出現すると仮定しているため、次式が成り立つことを仮定している。

$$\sum_{\text{desire}} \Pr(\text{desire}|\text{term}) = 1 \quad (\text{式4})$$

本方式では、欲求を充足させたユーザの実例を収集し、それをもとに  $\Pr(\text{desire}|\text{term})$  を計算し学習する。次式のように、当該単語の全コンテンツにおける出現回数に対する当該欲求を充足するコンテンツにおける出現回数で算出する。

$$\Pr(\text{desire}|\text{term}) = \frac{N_{\text{term} \cap \text{desire}}}{N_{\text{term}}} \quad (\text{式5})$$

$N_{\text{term}}$  は、全コンテンツにおける単語の出現回数、 $N_{\text{term} \cap \text{desire}}$  は欲求を充足するコンテンツにおける単語の出現回数を表す。

以上をまとめたコンテンツ欲求推定モジュールの構成を図6に示す。

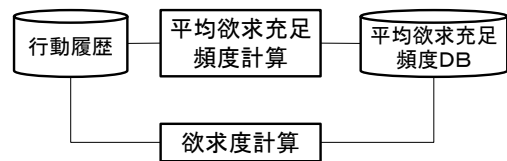


図5. プロフィール抽出モジュール

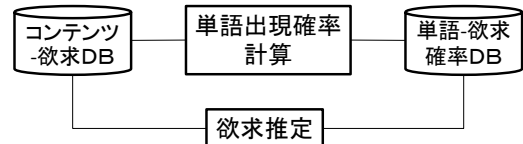


図6. コンテンツ充足欲求推定モジュール

## 4. 評価実験

### 4.1 実験方法

本章では、提案方式の評価実験とその結果について述べる。実際のユーザ行動を観測するために、筆者を含む12名の被験者にアンケート調査を実施し、毎日の視聴履歴と購買履歴を収集した。データセットを表2に示す。

表2. データセット

ユーザ数	12名
性別	男性8名、女性4名
年代	20代後半～30代前半
収集期間	2010/05/24～2010/07/31

被験者は、毎日、昨日視聴したテレビ番組 (録画番組やDVDなども含む) と、昨日購入したもの (ネットショッピング含む) をメールにて回答する。回答内容は、視聴したテレビ番組と購入したもののタイトルとそれによって満たした欲求を19の欲求軸の中から選択する。欲求は複数を満たすことも可とし、複数回答可とした。ただし、被験者は視聴履歴や購買履歴を記憶して入力するため、回答の負荷が高い。また、視聴したテレビ番組、購入した商品名や場所を入力するのは、プライバシーに関わるため、回答する範囲は、被験者の裁量に任せた。

### 4.2 評価結果

#### 4.2.1 ユニバーサル・プロフィール

まず、各ユーザの回答日数、1日当たりの視聴番組数と購買アイテム数について述べる。図7は、各被験者のアンケート回答日数である。被験者平均40.91日 (標準偏差17.13) の回答数で、最も回答数が多かった被験者は66日、最も少なかった被験者は17日であった。前述したように、本調査は回答の負荷が高いため、土日・祝日などの視聴番組数や購買数が多い日の回答は少なくなる傾向が見られた。今後、大規模な検証を行うためには、簡便な記録方法の検討などの継続的な回答の負荷を低減する方法を考える必要がある。

各被験者の1日当たりの視聴番組数は、被験者平均1.14番組 (標準偏差0.51番組)、最も視聴した被験者は1日当たり1.85番組、最も少なかった被験者は1日当

り0.02番組であった。

また、各被験者の1日当たりの購買アイテム数は、被験者平均3.23(標準偏差1.03)で、最も多かった被験者は5.11、最も少なかった被験者は1.00であった。

次に、視聴と購買の各行動における欲求の出現頻度を(式1)によって算出し、被験者の平均の計算結果を示す(図8, 9, 10)。

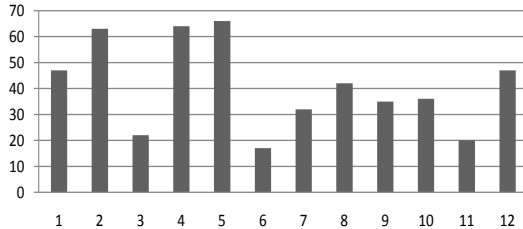


図7. 各被験者の回答日数  
(縦軸: 回答日数、横軸: 被験者番号)

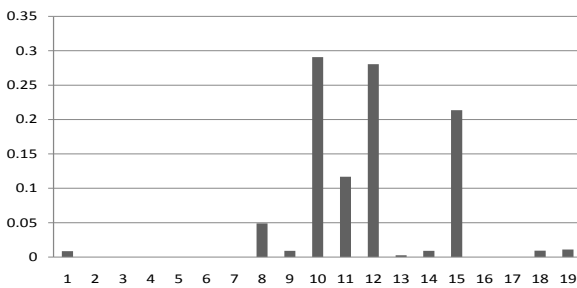


図8. テレビ視聴における各欲求の充足割合 (縦軸: 充足割合の被験者平均、横軸: 欲求軸)

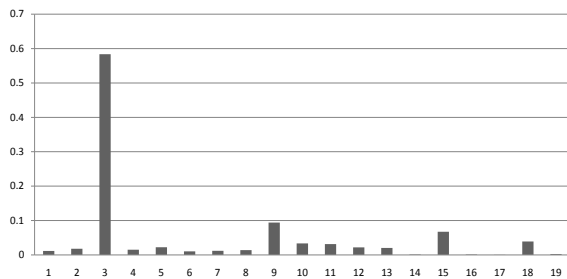


図9. 購買における各欲求の充足割合  
(縦軸: 充足割合の被験者平均、横軸: 欲求軸)

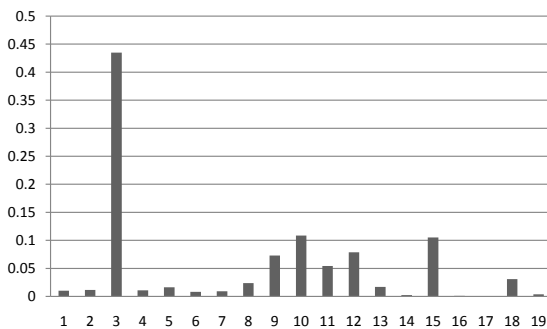


図10. 視聴と購買における各欲求の充足割合  
(縦軸: 充足割合の被験者平均、横軸: 欲求軸)

図8から、視聴行動においては、欲求8(持続)、10(知的好奇心)、11(感性体験)、12(非日常体験)、15(やすらぎ)を充足する割合が高いことが分かる。一方、図9から、購買行動においては、欲求3(摂取・消費)を充足する割合が圧倒的に多く、次いで、欲求9(回復)、15(やすらぎ)となっている。二つの行動を総合して算出した結果(図10)から、日々の生活行動においては、人間の生理的欲求(1次欲求)である食欲に関わる行動「食事」が大半を占めることがわかる。実際に、回答履歴の中では「食品」が占める割合が多かった。

このことから、各欲求を充足する行動には、充足する頻度の偏りが存在するということが分かる。また、テレビ視聴では、知的好奇心や感性体験、商品購買では、摂取・消費、回復、といった、各サービスにおける欲求の充足頻度にも偏りがあることがわかった。そのため、単一の欲求を充足する行動履歴、もしくは、単一のサービスの行動履歴を取得するのみでは、ユーザのすべての欲求の充足度合いを正確に抽出することが困難であるということがわかる。こうした問題を解決するためには、異なるサービスにおける行動履歴から抽出するユニバーサル・プロファイルは有効であるということが言える。

#### 4.2.2 プロファイル抽出方式

次に、プロファイル抽出方式の評価について述べる。

実施した評価実験で得られたデータセットを用いて、筆者のプロファイルを抽出した結果を例として、図11に示す。これによれば、筆者は、欲求5(将来展望)と14(性・恋愛)、17(向上・達成)が強く、欲求6(秩序)と16(反発)、19(社会的帰属)が弱いということになる。

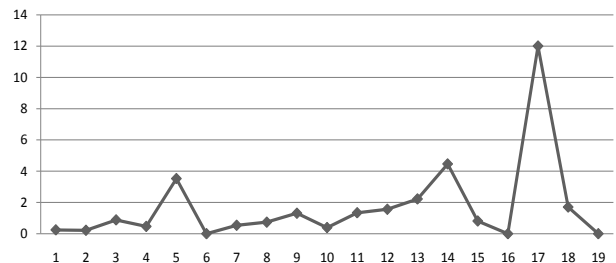


図11. ユーザプロファイルの抽出結果例  
(縦軸: 欲求の強さ、横軸: 欲求軸)

各被験者のユーザプロファイルの抽出結果を図12に示す。各欲求を充足頻度は図10のようになっているが、ユーザ平均を利用することにより、被験者毎に異なるユーザプロファイルを抽出できていることがわかる。

抽出されたユーザプロファイルがユーザの欲求の強さを正確に表現できるのかどうかを検証するため、各被験者のプロファイル抽出結果を直接アンケートによって評価してもらった。具体的には、被験者に対して、抽出結果から強いと思われる欲求の上位3つと、弱いと思われる上位3つを提示し、被験者はそれぞれの結果に対して、「納得できる」を5点、「違和感を覚える」を0点として、5点満点で評価してもらった。8人の被験者に対してアンケートを実施し、その結果を表3に示す。

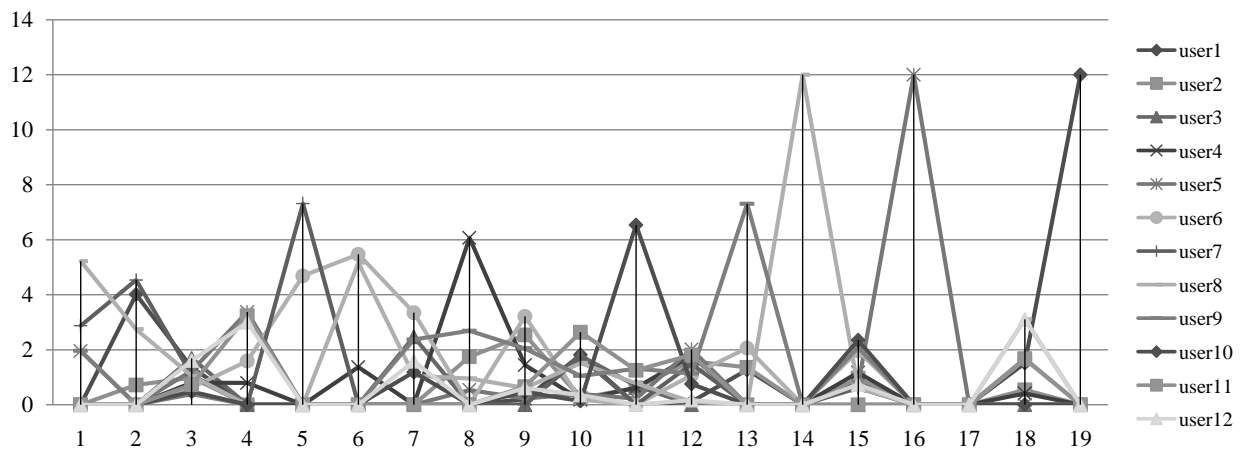


図 12. 各被験者のユーザプロフィールの抽出結果  
(縦軸：欲求の強さ、横軸：欲求軸)

強い欲求に関しては、平均評価値が 4.38、標準偏差が 0.33 となり、被験者の自己評価に概ね即した抽出結果であったということがわかる。その反面、弱い欲求に関しては、平均評価値が 3.21、標準偏差が 0.78 となり、被験者の自己評価とのずれがあることがわかる。この結果から、日々の行動に表れやすい欲求は強い欲求であることは言えるが、日々の行動に表れにくい欲求は必ずしも弱い欲求であるわけではない、ということが示唆される。また、強い欲求に関しては、日々自覚している可能性は高いと考えられるが、弱い欲求に関しては日々自覚していない可能性があると考えられるため、回答結果に有意性が見られなかったとも考えられる。このため、弱い欲求を評価するには、ユーザの主観評価の質問方法に検討の余地がある。

一方、図 10 からわかるように、各欲求を満たす行動には、適した頻度が存在するため、今回の実験期間では十分な頻度で出現しなかった可能性がある。また、今回は、購買行動と視聴行動のみを調査したため、他の行動において充足する欲求が存在することが考えられる。例えば、仕事において充足する欲求もあれば、友人との交流において充足する欲求もある。今後は、被験者数を増やし、実験期間を長くし、取得する行動の範囲を広げる手段を検討する必要がある。

しかし、ユーザの行動の起点には強い欲求があり、コンテンツ推薦において考慮すべきなのは、上位 3 つ程度であると考えられる。つまり、視聴行動と購買行動を取得した結果、本方式を用いてユーザの強い欲求を抽出することは可能であることが示唆される。

表 3. 抽出したプロフィールの評価結果

	強い欲求	弱い欲求
平均値	<b>4.38</b>	<b>3.21</b>
標準偏差	<b>0.33</b>	<b>0.78</b>

#### 4.2.3 コンテンツの充足欲求推定方式

次に、コンテンツの充足欲求推定方式の評価結果について述べる。

コンテンツから単語を抽出するために、テレビ番組情報は EPG (Electric Program Guide) を、商品情報は、一般的な検索エンジンを使って商品名を検索ワードとして検索した結果のトップページに含まれるテキスト情報を、データソースとした。テキスト情報から単語を抽出し、出現頻度を計算するモジュールには、連想検索エンジン「MANTA」を利用した[8]。

EPG と商品データは、実験期間中に得られたすべてについて取得できなかった。しかし、取得できたデータに対して欲求の評価値が付与されているものの中での各欲求の頻度は、図 8、9 の分布とほぼ同一の傾向であった。また、この期間中に、欲求 4 (活動・運動)、7 (合理性)、16 (反発)、17 (向上・達成) を充足する行動は出現しなかった。

詳細データを得られたコンテンツに含まれる単語の出現頻度を計算し、(式 3) によって各コンテンツが充足する欲求を推定した。推定精度の評価は、被験者が実際に利用したコンテンツに関して、欲求の推定確率を計算し、確率上位に実際に被験者がそのコンテンツで充足した欲求が含まれるコンテンツの全コンテンツにおける割合を、欲求推定精度とした。その結果を表 4 に示す。

表 4 より、推定精度は、80% を超えており、また、上位 2 件までに含まれる確率は 90% を超えている。このことから、コンテンツに含まれる単語によって、充足する欲求を推定できる可能性が高い。今回の実験では、12 人の被験者の 2 か月程度の行動履歴を用いたので、被験者数を増やすか、履歴取得の期間を長くすれば、推定精度はより高くなると考えられる。

欲求別に推定精度を評価した結果を図 13 に示す。欲求 3 (摂取・消費) に関わるコンテンツは推定精度が高い。これは、欲求 3 を充足する食事などの行動は高頻度で出現するため、含まれる単語の出現確率を十分に学習していることが要因であると考えられる。このことは、欲求 10 (知的好奇心)、12 (非日常体験)、15 (やすらぎ) を充足するコンテンツの推定精度が高いことにも同様に言える。一方、欲求 14 (性・恋愛) に関わるコンテンツは、行動における出現頻度が低いにも関

ならず、推定精度が高い。これは、欲求14を充足するコンテンツに含まれる単語は、他の欲求を充足するコンテンツに含まれる単語との重複が少ないということが言える。欲求8(持続)にも同様のことが言える。

推定精度が60%以下の欲求は、その欲求を充足するコンテンツに含まれる単語が他の欲求を充足するコンテンツにも含まれるため、推定精度が低い。このような欲求は、被験者によって欲求のとらえ方が異なる、という欲求のユーザ依存性と、状況や時間によってコンテンツ自身の充足する欲求が異なる、などのコンテンツのコンテキスト依存性が起因していると考えられる。

以上より、コンテンツに含まれる単語によってコンテンツが充足する欲求を推定する方式が有効であることが確認できた。

今後はユーザ依存性を考慮するために、被験者数を増加して行動履歴を蓄積し、欲求の捉え方に関するユーザタイプを抽出する、などの改善が必要である。また、コンテンツのコンテキスト依存性を考慮するためには、ユーザのコンテンツ選択時の場所や時間帯などの状況に関するデータを蓄積し、コンテキストを判定する、などのよりユーザ数・データ量を増加して進めていくことが必要である。

表4. 欲求推定精度

Top1に含まれる確率(%)	Top2に含まれる確率(%)	Top3に含まれる確率(%)
80.6	90.4	93.9

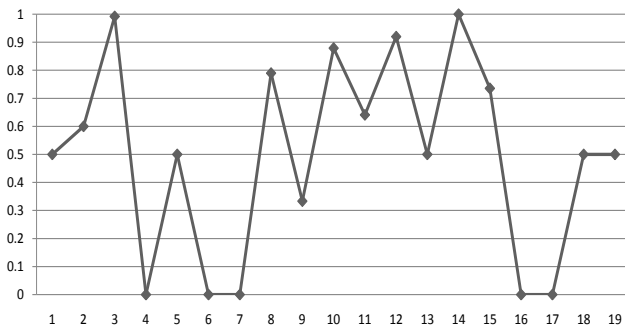


図13. 欲求推定精度の欲求別評価  
(縦軸: Top 1に含まれる確率、横軸: 欲求軸)

#### 4.3 関連方式との比較

異なるサービスにまたがる行動履歴からのユーザプロフィール抽出方式に関しては、多くは協調フィルタリングを利用しており、コンテンツベースで提案されている方式としては、ベイジアンネットワークを利用したものとオントロジベースのものが存在する。

それぞれの方式と提案方式を比較したものが表5である。提案方式は、19の欲求軸によってユーザプロフィールを表現するため、従来の手法よりも軽量である。そのため、様々なシステムや端末に組み込むことが容易であり、ユーザの日常行動のあらゆる場面で利用することが可能である。また、未知の新しいコンテンツサービスも、サービスが提供するコンテンツを本方式で提案する

基本的欲求で表現するだけで、簡単にシステムと接続することができるため、非常に汎用性の高いシステムである。

表5. ユーザプロフィールの方式比較

仕様	次元数	汎用性	多サービスとの接続性	サービス構築コスト
提案方式	19	○	○	小
ベイジアンネットワーク	数百~	×	○	大
オントロジ	数万~	○	△	大
協調フィルタ	数億~	×	○	大

#### 5. まとめと今後の課題

本研究の目的は、ユーザのあらゆる行動履歴からユーザプロフィールを抽出し、ユーザプロフィールに適したコンテンツをコンテンツの属するサービスに関わらず推薦するユーザ満足度の高い推薦システム「ユニバーサル・プロファイラ・システム」を実現することである。本論文では、ユーザの平均的な消費性向と、コンテンツが平均的にどのような消費性向を満たし易いかの双方を、「人間の基本的欲求の組」を用いたプロフィールとして記述することで、異なるサービス間で共通のプロフィールを交換するインタフェースを設計し、ユーザとコンテンツの双方のプロフィールの適合性判定により、ユーザのコンテンツに対する嗜好性を推定する方式を提案した。

プロフィール抽出方式については、ユーザプロフィールを行動における欲求の出現頻度のユーザ平均とのずれによって表現する方式を提案した。コンテンツの欲求推定方式は、コンテンツに含まれる単語によって欲求を推定する方式を提案した。そして、ユーザの視聴・購買履歴取得実験を実施し、プロフィール抽出方式とコンテンツ欲求推定方式の評価を行った。プロフィール抽出方式では、8名のユーザの主観評価において、提案方式によって抽出された上位3件の欲求に対する評価が5点満点中平均4.38点を達成した。これは、推薦に必要であると思われる数の欲求に対してユーザが違和感を覚えないことを意味しており、提案方式が実サービスでも有効であることが確認できた。

コンテンツの充足欲求推定方式では、12人のユーザデータをもとにコンテンツの充足する欲求を80.6%の確率で推定することを確認し、提案方式が有効であることを確認した。

しかし、より高精度にプロフィールを抽出するためには、いくつかの課題がある。まず、各欲求を満たす行動には、適した頻度や適した期間が存在すると推測されるため、今回の実験期間では十分な頻度で出現しなかった可能性がある。また、今回は、購買行動と視聴行動のみを調査したため、他の行動において充足する欲求が存在することが考えられる。例えば、仕事において充足する欲求もあれば、友人との交流において充足する欲求もあ



る。今後は、被験者数を増やし、実験期間を長くし、取得する行動の範囲を広げる手段を検討する必要がある。また、提案方式が、実際にユーザの欲求の強さの表現可能性を確実にするためには、被験者を増やし、統計的な有効性を高めていく必要がある。

また、本論文で提案した方式では、視聴行動も購買行動も一つのコンテンツの利用で欲求が充足するものと仮定した。しかし、実際には、30分の視聴と2時間の視聴、また、100円の商品と10000円の商品では、ユーザの満足度が異なることが想像される。さらに、1時間の視聴と10000円の商品購入というように、単位の比較できない行動における満足度の差異も存在するため、プロフィール抽出の高精度化は、これから研究が必要になると考えられる。

コンテンツの充足欲求推定方式についても、今後より高精度な推定を実現するためには課題が残っている。

まず、欲求の捉え方や充足方法はユーザに依存する部分が多いため、被験者数を増加して行動履歴を蓄積し、欲求の捉え方に関するユーザタイプを抽出する、などの改善が必要である。そのためには、年齢や性別などのユーザに関する情報も得ることによって、ユーザの感性や個性、趣味・嗜好といったものを理解することを検討していく。

また、コンテンツのコンテキスト依存性を考慮するためには、ユーザのコンテンツ選択時の場所や時間帯などの状況に関するデータを蓄積し、コンテキストを判定する、などのよりユーザ数・データ量を増加して進めていくことが必要である。

以上のように実際のユーザ数とデータ量を増加させ、そして、提案方式の統計的な妥当性を検証するためには、ユーザの行動履歴蓄積の負荷を極力低減する必要がある。今後は、実サービスを見据え、サービス運用に必要なプロフィール抽出精度やコンテンツの欲求推定精度を明確にするために、積極的に本システムの有効性を示していく。また、電子マネーやPCの操作ログ、テレビの操作ログの利用を検討し、最適なシステム構成を検討していく。

本研究が目指す「ユニバーサル・プロファイラ・システム」の利用により、今まで個別に収集・解析されてきた人間行動を一元的に管理することができる。また、多様な要素が複雑に絡み合っている人間行動のメカニズムを明らかにできる可能性がある。この結果は、人間はなぜ行動するのか、といった人間行動の原理的理解に基づいた推薦システムへアプローチする重要な足掛かりとなると考えられる。

#### 参考文献

- [1]Herlocker, J., et al. : Evaluating Collaborative Filtering Recommender Systems, ACM Transactions on Information Systems, Vol.22, No.1,pp5-53(2004)
- [2]Adomavicius, G., Tuzhilin, A.: 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*, 17 (6), 734-749. (2005)
- [3]Burke, R.: Hybrid Recommender Systems: Survey and Experiments. *User Modeling and User-Adapted Interaction*, 12 (4), 331-370, (2002)
- [4]玉川奨、他：日本語 Wikipedia からの大規模オントロジー学習、人工知能学会論文誌 25 巻 5 号 SP-C(2010)

[5]Ivaen Cantador, Martin Szomszor, Harith Alani, Miriam Fernandez, Pablo Castells: Enriching Ontological User Profiles with Tagging History for Multi-Domain Recommendations. In: 1st Intl. Workshop on Collective Intelligence and the Semantic Web (CISWeb 2008), pp. 5-19.

[6]Nelson, P. : "Information and Consumer Behavior", Journal of Political Economy, Vol. 78, pp.311-329, (1970)

[7]Steven Reiss : "Who am I? The 16 Basic Desires that Motivate Our Actions and Define Our Personalities", Berkley Trade (2002)

[8]安田知弘：連想検索エンジンのスケーラビリティおよび障害耐性の向上、情報処理学会第69回全国大会, 3D-1, 2007年3月