

## オントロジーを用いたユーザ嗜好の多面的特徴抽出方法の提案

檜原渉<sup>1</sup> 撫中達司<sup>1</sup> 森郁海<sup>2</sup> 板垣弦矢<sup>2</sup>

**概要:** 家電製品はそれぞれの役割としての活用に留まり、購入後にその他のサービスの活用に用いられることは少ない。そこで、家電製品の新たな付加価値としてソリューションを新たに付与することが考えられる。本研究では、対象ユーザに適したソリューションをRecommendすることを目的として、オントロジーを用いてユーザの嗜好を説明可能な形でかつ多面的な特徴抽出をするための提案とその検証を行う。結果、ユーザが注目している要素が特徴として抽出されたこと、注目していない要素・情報が無意識的な部分として特徴抽出されたことから、多視点からのRecommendationが可能であることが確認できた。

## Proposal of a multifaceted feature extraction method for user preferences using ontology

WATARU KASHIHARA<sup>1</sup> TATSUJI MUNAKA<sup>1</sup>  
IKUMI MORI<sup>2</sup> GENYA ITAGAKI<sup>2</sup>

### 1. 背景

センサの低価格化、機械学習の高度化により、様々なIoTシステムが開発されている。しかし、多くの提案では新たなセンサ、機器の導入、さらには収集したデータの分析を実行するためのクラウドサービスの導入を前提としており、運用コストなどの面からも利用者にとっては新たなサービス導入の障壁となっている。一方で、家庭内にすでに存在する炊飯器、エアコン、掃除機といった家電製品はそれぞれの役割としての活用に留まり、購入後にその他のサービスで活用されることは少ない。購入後の家電などをその他のサービスとして付加価値を付ける「ソリューション」として提供できれば、新たな機器の導入をすることなく、見守りや生活習慣管理などが可能となる。ソリューションのRecommendにおいては、価格、サービスの人気度、利用者の居住環境などにとどまらず、気象状況などの地球環境の変化、法規制などによる材料等の利用制限などの社会的な変化など、多面的な視点でのRecommendが必要となる。

本研究では、ユーザ嗜好を多面的に捉え、ユーザの現状に応じて新たなサービスの提示を行うことを目的として、オントロジーという表現形式を用いてユーザの嗜好を説明可能な形でかつ多面的な特徴抽出を行う方法の提案とその検証を行う。

### 2. 先行研究

本研究と関係があるものとして、オントロジー構造関連の研究と特徴抽出方法関連の研究について挙げる。

#### 2.1 提案内容

##### 2.1.1 スマートヘルスケアのための身体活動Recommendシステム[1]

ユーザの健康管理に有効なアドバイスを提供することを目的として、オントロジーで管理された基準と身体活動時に取得される生体情報を比較することで各身体活動の状態を評価している。評価は人の身体的な特徴によるもの、身体活動の強度によるもの、生体情報から得られた心拍数、活動強度、心電図の波形を周波数解析することで、自律神経のバランスを示すことができる LF/HF (Low Frequency/High Frequency) という値によるもの、暑さ指数によるものに分類でき、それらの状況の組み合わせから導出している。LF、HF では、それぞれ「交感神経、副交感神経両方の活動性」「副交感神経の影響度」を取得でき、LF と HF 比で表した LF/HF 値を用いることで自律神経のバランスを示すことができる。熱中症に対する評価基準の指針は日本生気象学会によって定義されている「日常生活における熱中症予防指針[2]」を使用している。オントロジーを用いて身体活動の状態を評価する提案であり、実験結果によりオントロジーや指標の定義に従って正しいRecommendが提供できていることが示されているが、Recommendする内容の決定などにおいて、特徴抽出のような手法は抜かれておらず、Recommendの結果の説明性には課題が残る。

<sup>1</sup> 東海大学大学院 情報通信学研究科

<sup>2</sup> 三菱電機株式会社 情報技術総合研究所

### 2.1.2 オントロジを活用した多様な利用者環境に適用可能な IoT サービス構成手法[3]

IoT に関する知識が少ない利用者が IoT デバイスを組み合わせ合わせたサービスを簡単に利用できるようにすることを目的として、「デバイス情報」「ソフトウェア情報」「デバイスの動作」「動作によって影響が出るモノ」という4つのクラスの定義によってデバイスとソフトウェアの関係性を表現するオントロジー構造を提案している。結果として、具体的な IoT デバイスを選択でき、レシピが適用可能か判定できること、異種の IoT デバイスを同一のレシピで利用できたことが確認されている。今後の課題の一つとして、IoT サービスの構成時に利用者の特性を考慮した、適切な IoT デバイスの選択が挙げられている。

### 2.1.3 嗜好に基づくレコメンドソリューションの実現 [4]

総合電子書籍ストアでのレコメンドサービスにおいて、協調フィルタリングを用いたレコメンドにおける3つの問題点を改善する手法とクロスドメインレコメンドの検証を行っている。協調フィルタリングの問題点として、以下の3点が挙げられている。

1. どれだけの類似ユーザを抽出すれば精度の高い推薦ができるか分からない。
2. 購入履歴を嗜好データとしているため、購入数が少ないユーザは、嗜好や推薦の変化が少なく、飽きられてしまう。
3. 類似ユーザが好む書籍が推薦するが、人気書籍の購入数が多く、認知されている可能性が高いことから、意外性のある推薦ができない。

それらを改善する策として、以下の方法が提案されている。

1. 実験的に類似ユーザ数別で推薦精度を確かめることで、どの人数が最も高精度かを検証できる。
2. 閲覧行為は購入よりも気軽にできるため、閲覧履歴も嗜好として捉えることで、嗜好データが蓄積しやすく、変化しやすい。
3. 「人気書籍以外で類似ユーザを比較」というようにユーザ嗜好をフィルタリングすることで意外性のある推薦が行える。

また、協調フィルタリングでは、書籍と関係のない、番組の視聴履歴などを持つユーザに対しての推薦が行えないという問題を解決するために、クロスドメイン環境での TF-IDF(Term Frequency - Inverse Document Frequency)手法を用いた推薦を行っている。TF-IDF は、「TF: 対象文書に含まれる対象単語割合」「IDF: 対象単語が含まれる文書割合の逆数」という意味を示すものであり、結果として、全ての文書を介して対象文書の特徴となる単語を抽出するという手法であるため、「ジャンル、概要の説明文」といった書籍と番組の両方が持つ要素から特徴抽出することで互いの比較ができ、推薦が可能となる。

残された課題として、TF-IDF は、同形異義語が存在して

いたり、単語への重みづけが適切に行われていないと、推薦の精度が下がる可能性があるため、それを補う手法が別に必要になる。

### 2.1.4 コンテンツの多様性を考慮したクロスドメイン推奨 [5]

テレビ番組ごとに適した書籍の推奨を行う目的で、TF-IDF と LDA(Latent Dirichlet Allocation)という特徴抽出手法を併用している。LDA は出現頻度が高い単語からトピックを生成し、対象の文書がどのトピックに関係しているかを数値で表現する手法であり、TF-IDF よりも広い範囲で特徴抽出を行える。LDA の問題点として、説明が不十分であったり、書籍や番組から推奨に有効なトピックを得られなかったりした場合に、推薦の精度が下がる可能性がある。そこで両手法を用いることで、互いの欠点を補う形での推薦が可能となることから、単体よりも精度が高い結果が得られている。

今後の課題として、今回のターゲットドメインが書籍のみであったが、他にも、音楽、映画、家電など、ターゲットドメインの拡大をする際に、商品のカテゴリごとに説明文ごとにまとめて利用することなどから、オントロジーといった類似度計算のための新たな表現形式の利用が必要になることが挙げられている。

## 2.2 先行研究の問題点

先行研究[4]、[5]において、「TF-IDF」「フィルタリング」「LDA」というように、個別に手法を組み込むと、それぞれの手法における視点でしか物事を捉えるができない。つまり、手法と手法の間に内在する暗黙的な視点(関係性)を無視した形となってしまうことが危惧される。例えば、「協調フィルタリング」は、「ユーザが購入した商品と、類似ユーザが購入した商品の関係性」から見た視点で、「TF-IDF」は「商品の説明文と、出現単語との関係性」から見た視点だが、これを踏まえ、新たに「ユーザと、商品の説明文の関係性」という視点が考えられる。この視点は協調フィルタリングでも TF-IDF でも表現されていないが、ユーザは商品を購入する際説明文を読む可能性が高いため、無関係とは考え難く、この二つの手法のみでは説明文を利用して商品同士をレコメンドすることができない(図 1)。

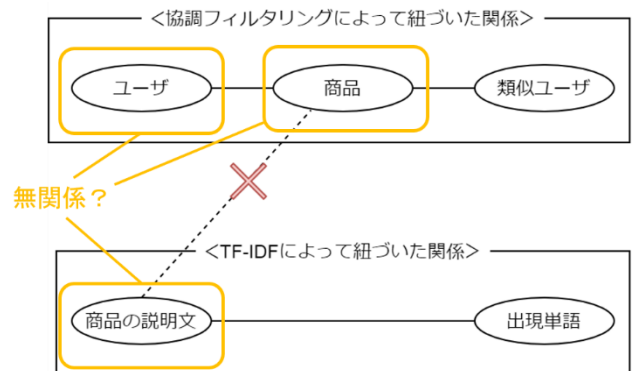


図 1 二つの手法を併用した場合の暗黙的な視点

このようにレコメンドの精度を高めるため併用する手法を増やす度に、関係性が複雑となり、結果として暗黙的な視点が増えてくると、「なぜこの商品が人気なのか、不人気なのか」という理由を説明することが難しくなり、また不人気であってもそのユーザには価値がある商品をレコメンドできる具体的な要素の抽出、ユーザ嗜好に合わせた新たなソリューションの提示が行えなくなる。そこで、手法を併用して視点を増やす際に、暗黙的な視点も明確化される、多面的な視点を持つレコメンド手法が必要となると考える。

### 3. 提案内容

#### 3.1 概要

ユーザ嗜好に適した新たなサービスを提示することを目的とし、要素同士の関連性や、要素を構成するための情報との関係を定義することで、オントロジーを用いた説明性、多視点性、拡張性のある特徴抽出を行う方法を提案する。なお、実装の際はオントロジーモデルを記述するためのツールとして「法造[6]」を用いている。

3.2ではその準備としてオントロジー構造について説明し、3.3では提案する特徴抽出方法について記述する。

#### 3.2 オントロジー構造

オントロジー構造において、ユーザの嗜好を表現する際に、ノードを「概念、要素、情報」という3つに分類して表現するオントロジーモデルを定義した(図2)。ここで、提案するオントロジー構造における「概念」とは、ソリューション、センサ、(ソリューションが属する)カテゴリなどの要素を分類したノードである(図2中の菱形ノード)。次に「要素」とは、「概念」を構成する具体的なセンサ、ソリューションのカテゴリとなる(図2中の丸ノード)。最後に、「情報」は要素に関する具体的な評価値であり、価格、人気度、センサの型番などを保持する(図2中の四角ノード)。また、矢印(プロパティ)によって互いの階層関係を表現しており、「要素」に関してはそれぞれ直接的に関係性を表現している。また、オントロジーモデルは、開発者側が「概念、要素、情報」を定義した内容を管理するための「プラットフォーム知識」とユーザ個人の購入履歴やユーザ嗜好を管理するための「ユーザ知識モデル」という二つのモデルを使用する[7]。すなわち、ユーザ知識モデルはユーザごとに所有していることになる。プラットフォーム知識モデルの構成は図2のような形となるが、ユーザ知識モデルでは、プラットフォーム知識モデルからユーザが購入したソリューションとそれに関係する概念、要素、情報を取り出し、新たに「特徴値」を要素・情報ごとに算出し、それぞれ格納する(図3)。特徴値の詳細について、3.3に記述する。

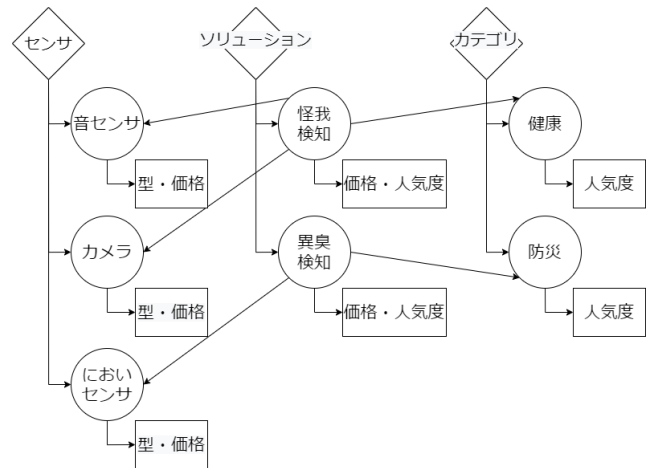
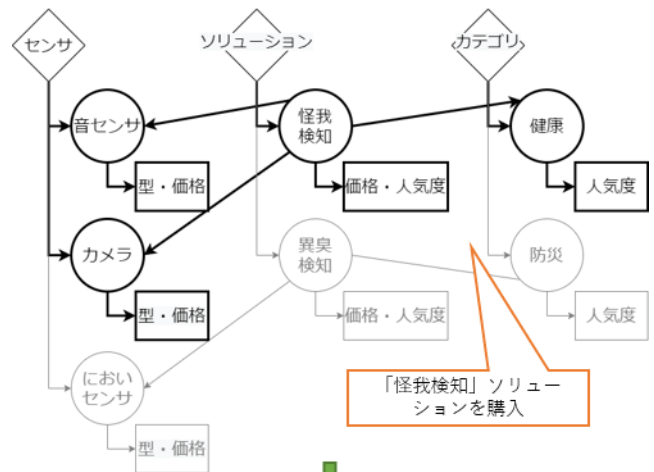


図2 提案するオントロジー構造の例

#### <プラットフォーム知識>



ユーザ知識  
モデルに反映

#### <ユーザ知識>

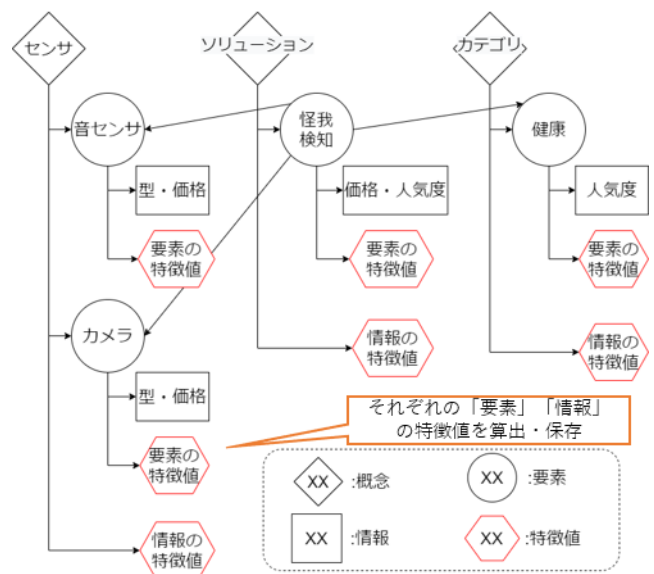


図3 「怪我検知」ソリューションを購入した場合のユーザ知識モデル例

$$c_x \text{ の特徴値} = c_x \text{ のレートの正規化値} = \frac{c_x \text{ のレート} - \min(c_x \text{ のレート})}{\max(c_x \text{ のレート}) - \min(c_x \text{ のレート})}$$

$$c_x \text{ のレート} = \frac{c_x \text{ へのリンク数}}{\text{ソリューションとカテゴリ間のリンクの総数}} = \frac{\text{赤・青・緑線のいずれかの数}}{\text{赤・青・緑線の合計}}$$

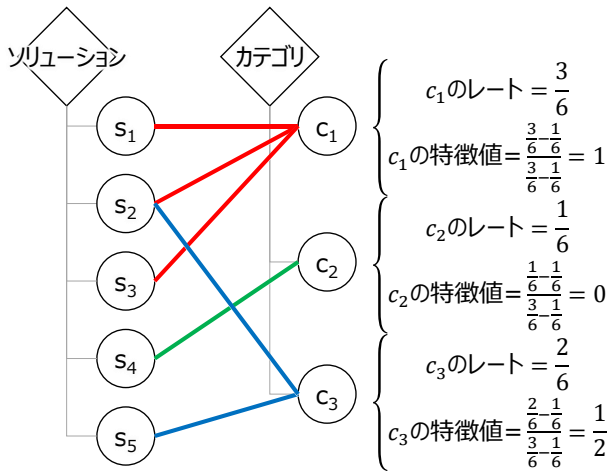


図4 「カテゴリ」要素の特徴値の算出例

### 3.3 特徴抽出方法

ここで述べる特徴抽出方法とは、「ユーザが概念・要素・情報といった対象にどれだけ注目しているか」を定量的に表現する「特徴値」を算出するための方法である。特徴値は、1に近いほど特徴として明確に抽出されていることを意味する。なお、ユーザの購入履歴であるユーザ知識モデルを用いて、「要素」と「情報」それぞれの特徴値を導き出す。

「要素」の特徴値算出方法として、対象「概念」下の「要素」集合を基準として、「要素」ごとに、他の「要素」と紐づく要素数の集合全体に対する割合を求め、正規化を行う(図4)。

「情報」の特徴値算出方法として、「個人基準値」という数値を定義する。個人基準値とは、「ユーザが購入するための目安」を同じ概念に分類される同じ情報(ソリューション1の価格、ソリューション2の価格など)で比較し、0~1で表したものである。これにより異なる情報の数値幅を統一し、異なる情報同士での特徴値の比較が可能となる。対象概念下の同じ情報の個人基準値集合から標準偏差を求め、0~1に正規化を行う。

### 3.4 提案内容の特徴

本研究における提案である特徴抽出手法は、主に「多視点性、説明性、拡張性」という3つの特徴を持つ。

#### 3.4.1 多視点性

オントロジーの構造上、抽象的な視点から具体的な視点まで変更・追加が可能(図5、図6)であるため、「抽象的過ぎるために具体的内容が見えづらい」「具体的過ぎるために全体が見えづらい」という問題を回避できる。

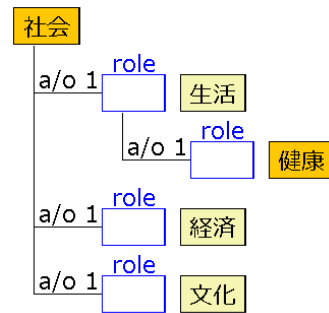


図5 「健康」カテゴリの具体化例

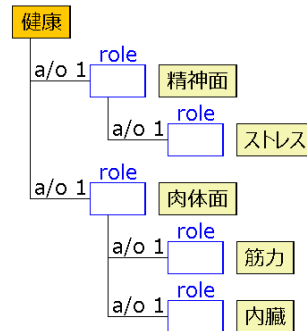


図6 「健康」カテゴリの抽象化例

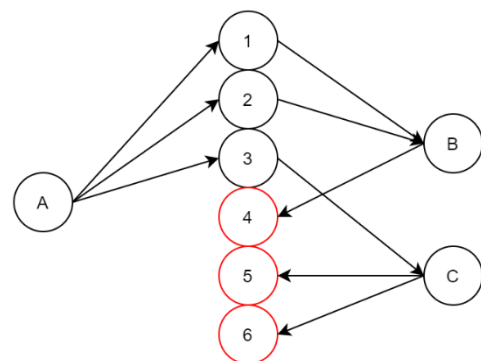
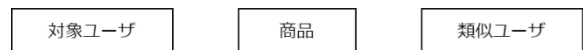


図7 オントロジーで表現された協調フィルタリング例

また、2.2で説明したような「協調フィルタリング」と「TF-IDF」においての問題を解決できる。方法としてそれぞれをオントロジーモデルで表現し(図7、図8)、互いを紐づける(図9)ことで、間接的にユーザと商品の説明文との関係性があることを明確に表現でき、それを用いた特徴抽出も行うことが可能なため、多視点的にユーザ嗜好を捉えられる。以下、図7、図8、図9の詳細を記述する。なお、例として挙げる図は、「情報」や「概念」と「要素」の関係性を示す矢印を略しており、四角が「概念」、その下に配置されている丸が、概念下の「要素」を表現しているとする。

図7では、それぞれユーザが購入した商品が「A:1, 2, 3」「B:1, 2, 4」「C:3, 5, 6」という場合、対象ユーザAには商品4, 5, 6が推奨されることを表している。初めにAから購入商品のノードに向かって矢印で関係が示される。また、購入商品から類似ユーザを算出するために、Aが購入した商品から、その商品を購入していたユーザに向かっ

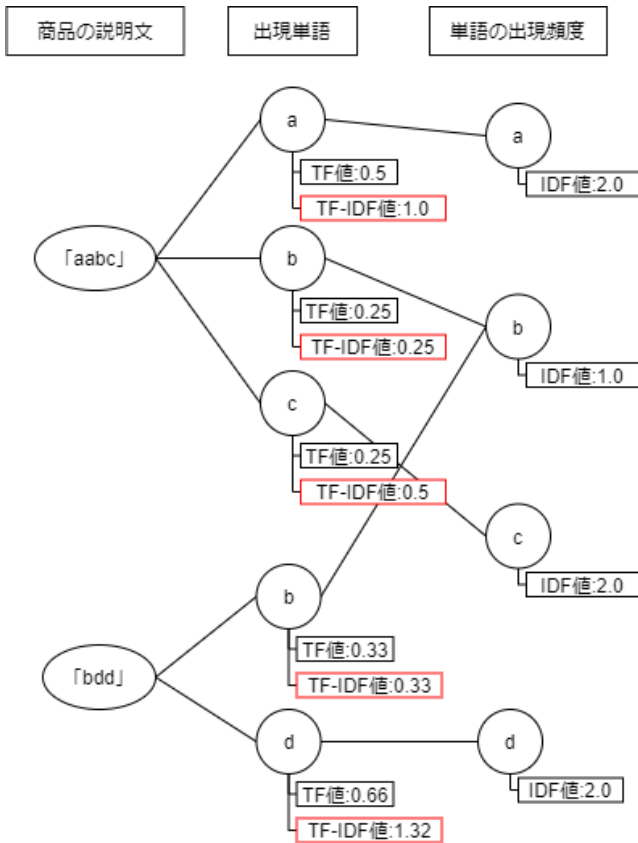


図 8 TF-IDF をオントロジーで表現したもの

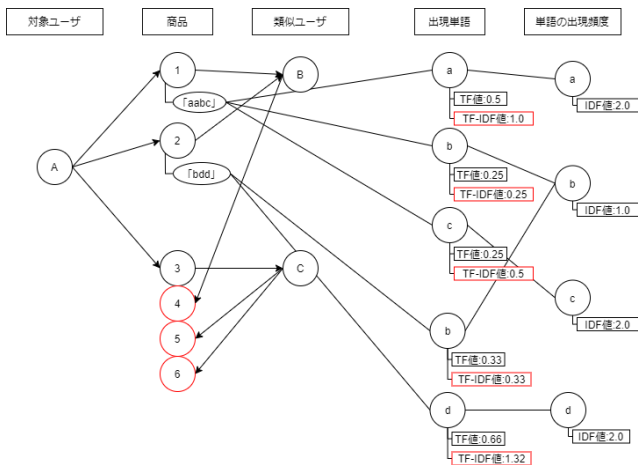


図 9 両手法のオントロジーモデルを紐づけたもの

て矢印で関係が示される。これにより類似ユーザとして B, C が挙がり, B, C の購入商品のうち, A が購入していない商品に対して矢印が引かれ関係が示される。その結果, 推奨される商品は 4, 5, 6 となる。

図 8 は, 「aabc」「bdd」という説明文書に対して TF-IDF を行い, 文書ごとの単語要素に算出結果を記録していることを表している。初めに文書それぞれの単語出現割合, すなわち TF 値を算出する。TF 値は「出現単語」概念内に

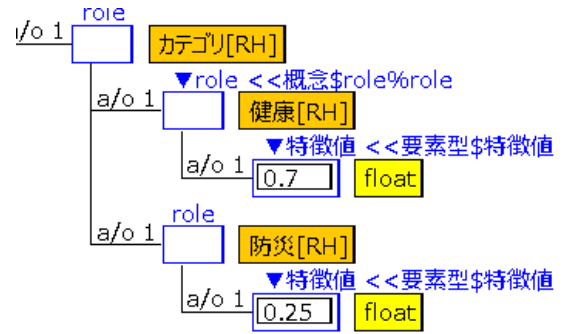


図 10 オントロジーモデルの一部例

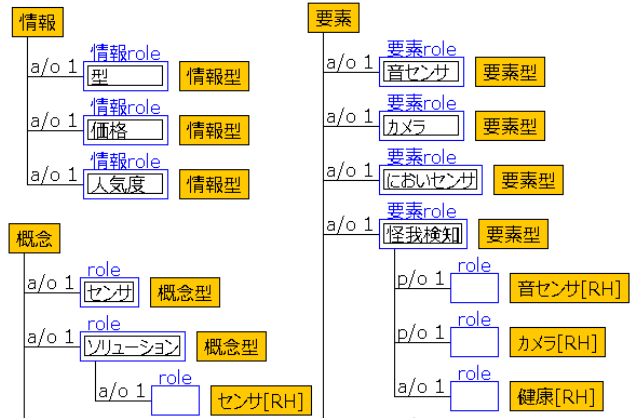


図 11 オントロジーモデル例

ある単語要素に記録され, この例の場合, 説明文 1 の単語では「a:0.5, b:0.25, c:0.25」が, 説明文 2 の単語では「b:0.33, d:0.66」が記録される。次に, 全ての文書のうち対象単語が出現した文書数割合の逆数, すなわち IDF 値を算出する。IDF 値は「単語の出現頻度」概念内にある単語要素に記録される。この例の場合, 総合文書数に対してそれぞれの単語が出現した文書の割合を算出し, その逆数の「a:2, b:1, c:2, d:2」が IDF 値として記録される。最後に, TF 値, IDF 値を掛け合わせ, 文書の出現単語ごとに TF-IDF 値を算出する。

図 9 は, 商品とその説明文で関係性があった場合, その二つを通して, 協調フィルタリングと, TF-IDF の手法の紐づけが行えることを示している。これにより, 2 つの手法を併用したのみでは得られなかった, 「ユーザと商品の説明文の関係性」という暗黙的な視点が内在していたことが示せることになる。

### 3.4.2 説明性

特徴値の全てにラベルが付くため, その大小に対応して意味を表す言葉に置き換えられることから, 数値と要素の関係性において説明性が高い。図 10 のようなオントロジーモデルが存在していたとして, 言葉で表現すると「カテゴリ概念における健康要素の特徴値が 0.7 であるため, 健康カテゴリに注目する傾向にある」と説明ができる。

表 1 検証モデルでのカテゴリ項目

カテゴリ項目	説明
健康	転倒防止などの見守りに関するソリューション
防災	電気やガスによる火災などを防止するソリューション
防犯	戸締りなどを注意喚起し空き巣の侵入を防ぐなどのソリューション
電気	電気代の管理など行うソリューション
外出	外出する上で、家電操作や有用な情報を提供するソリューション

表 2 検証モデルでの情報定義

概念	情報
ソリューション	価格, 人気度
センサ	価格
カテゴリ	人気度, 季節性

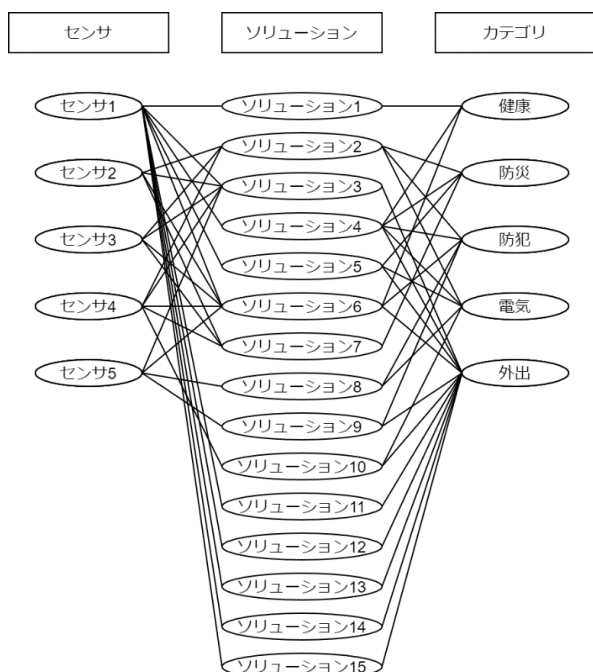


図 12 検証用モデルの全体図

### 3.4.3 拡張性

オントロジーの構造を予め定義することで、概念、要素、情報の関係を明確に表現することが可能となり、同じ形式ですべての要素と関係を表現できる(図 11)。これにより、追加、変更が容易に行うことが可能となり、拡張も行いやすい。

## 4. 検証

### 4.1 検証で用いるモデル

「健康」、「防災」、「防犯」、「電気」、「外出」の5つ(表 1)をソリューションのカテゴリとし、「センサ1」～「セン

i) 紐づきの偏り例を図 17 に記す。

表 3 検証項目

検証番号	注目の仕方
1	「健康」カテゴリを含んでいるもの
2	「外出」カテゴリを含んでいるもの
3	「健康」もしくは「外出」カテゴリを含んでいるもの
4	「防犯」かつ「外出」カテゴリを含んでいるもの
5	安価なもの
6	安価, 時期的に推奨されるもの

サ5」の5つのセンサがすでに設置されているという仮定の下、価格、人気度といった情報値が設定された15のソリューションを用いて検証を行った(図 12)。なお、ソリューションとそれ以外の要素(カテゴリ、センサ)の関係性において、紐づく要素数が大きく異なった場合[i]や、センサやカテゴリごとの供給量が大きく異なっていた場合[ii]といった条件を再現するために、ソリューション数を比較的多くしている。

### 4.2 情報定義とそれに伴った個人基準値の定義

概念ごとの情報定義を表 2 に示す。情報値の設定基準について説明する。

#### (1) ソリューション, センサの情報値

実際に設定される数値は広い範囲に及ぶことが想定されるため、検証では数値を幅広く設定する。しかし、単純に乱数を用いた場合には数値分布に偏りが生じる可能性があるため、予め決めた等間隔の数値配列からランダムに抽出することで、想定外の数値分布を無くしている。なお、ソリューションの価格に関してはセンサの価格を足したものとなっている。

#### (2) カテゴリの情報値

「季節性」という数値は「対象のソリューションが現在の季節に適しているかどうか」を意味している。このため、例えば「1月」では「寒い季節、火災が多い」等の観点から数値を設定している。「人気度」は季節性と正の相関関係になるという仮定のもと、季節性の数値に乱数を加えたものを設定している。

### 4.3 検証項目

検証は6つの項目を設定した(表 3)。検証方法として、それぞれの項目ごとに購入後のユーザ知識モデルを参照し、特徴抽出を行う。検証1を例とすると、ユーザが「健康」カテゴリに注目した上で、ソリューションを購入し、そのユーザ知識モデルから特徴抽出を行う。

ii) 供給の偏り例を図 18 に記す。

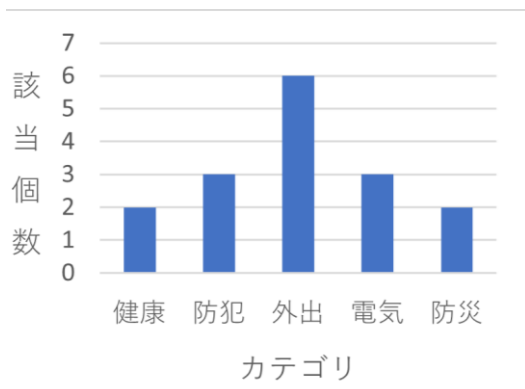


図 13 外出カテゴリに注目して購入した場合のカテゴリ分布

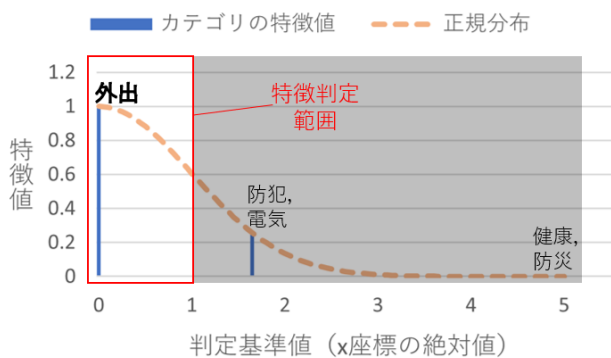


図 14 特徴としての判定例

#### 4.4 判定基準

「特徴として捉えた」「正しく特徴抽出できた」「多面的に特徴抽出できた」という判定基準について説明する。

##### (1) 「特徴として捉えた」という判定方法

図 13 は検証用のモデルで外出カテゴリに注目してソリューションを購入した場合のカテゴリ分布を表しているが、左右対称の山型が見られる。この割合を 0~1 の範囲に正規化した場合、分布の形状が維持された状態でカテゴリの該当数が特徴値に反映されるため、特徴値は正規分布に従うと考えられる。また、カテゴリの要素数などは任意で決まるため、その数に影響されない判定方法としてガウス関数を用いる。特徴値をガウス関数の計算式の y 座標に当てはめ、出力された x 座標の絶対値を特徴値の判定基準値とする。なお、特徴値が 0 の場合、判定基準値が無限になることを考慮して、最大値を 5.0 に制限している。今回は判定基準値が 1 以下のものを「対象が特徴として捉えられる」と判定することとし、検証を行う (図 14)。また、特徴として捉えられたものが複数存在した場合、特徴値を比較してレコメンド時の優先順位を決める。比較する範囲は、特徴値を算出する上での範囲に合わせ、要素の特徴値は概念ごと、情報の特徴値は全てとする。

##### (2) 「正しく特徴抽出できた」という判定方法

表 3 に記述した検証項目ごとで注目した要素・情報値が全て特徴として捉えられた場合、正しく特徴抽出が行えたと判断する。

### (3) 「多面的に特徴抽出できた」という判定方法

ユーザが直接注目している内容とは別に、新たな特徴が抽出された場合、多面的に特徴抽出が行えたと判断する。

#### 4.5 検証結果

評価の結果、表 3 に記述した検証項目ごとで注目した要素・情報値が全て特徴として捉えられた (表 4)。例として、検証 1 (図 15) では、注目していた「健康」カテゴリは判定基準値が 0 で、「1 以下」という条件を満たしているため特徴として捉えられた。また、検証 3 (図 16) では、「健康」もしくは「外出」のカテゴリに注目しており、判定基準値を見るとどちらも 1 以下であるため特徴として捉えられているが、「外出」カテゴリは「健康」カテゴリよりも特徴値が高いため、レコメンド時に優先されることになる。同様に、検証項目すべてにおいて正しく特徴を捉えられており、一部の検証項目では注目していなかった情報が最も特徴として捉えられ、特徴抽出方法の多面性を示せた。4.6 でその具体的な理由と考察を記す。

#### 4.6 設定条件のデータにおける結果の妥当性

##### 4.6.1 ソリューションと紐づく要素数への依存度

センサやカテゴリなどに紐づく要素数が多いソリューションを購入した場合、カテゴリの分布が広がり、特徴抽出の精度が顕著に低下することが懸念された。しかし、表 4 に示す様に、要素・情報の特徴が正しく抽出できていることから、ユーザの注目対象が「要素」「情報」など明確に表現される場合には、正しく特徴が捉えられやすいと考えられる。

##### 4.6.2 ソリューション以外の要素の供給偏り

センサ、カテゴリといった概念下の要素において、紐づくソリューションの数に要素同士で大きく差が出ていた場合、ソリューションが複数購入された際に紐づきが少ないセンサやカテゴリが特徴として上手く抽出されない可能性が考えられた。結果として、検証 1 のように注目するカテゴリが一つの場合、購入したソリューションが「外出」カテゴリなど供給が多いカテゴリに属していたとしても、検証 1 の出力結果 (図 15) が示すように「健康」カテゴリの方が高い数値が出力されることが確認できることから、レコメンドするソリューションには、供給が多い他のカテゴリの影響は小さいと考えられる。

表 4 検証項目と判定結果

検証番号	注目したカテゴリの判定結果
1	健康：○
2	外出：○
3	健康：○, 外出：○ (or)
4	防犯：○, 外出：○ (and)
5	価格：○
6	価格：○, 季節性：○ (and)

or：注目対象のいずれかからソリューションを購入  
and：全ての注目対象に適しているソリューションを購入

[購入ソリューション]:1, 4, 6	
要素の特徴値:	
センサ:	
センサ 1:	1.000(0.0)
センサ 2:	0.000(5.0)
センサ 3:	0.000(5.0)
センサ 4:	0.000(5.0)
センサ 5:	0.000(5.0)
カテゴリ:	
健康:	1.000(0.0)
防災:	0.000(5.0)
防犯:	0.500(1.2)
電気:	0.000(5.0)
外出:	0.500(1.2)

図 15 検証 1 の要素の特徴値 (括弧内は判定基準値)

[購入ソリューション]:1, 3, 4, 6, 12	
要素の特徴値:	
センサ:	
センサ 1:	1.000(0.0)
センサ 2:	0.000(5.0)
センサ 3:	0.000(5.0)
センサ 4:	0.000(5.0)
センサ 5:	0.000(5.0)
カテゴリ:	
健康:	0.667(0.9)
防災:	0.000(5.0)
防犯:	0.333(1.5)
電気:	0.000(5.0)
外出:	1.000(0.0)

図 16 検証 3 の要素の特徴値 (括弧内は判定基準値)

表 5 検証 6 での特徴的だった情報

情報	個人基準値	特徴値
価格	0.299	0.631
季節性	0.680	0.877
カテゴリの 人気度	0.574	0.884

また、検証 3 のように供給の少ない「健康」カテゴリと供給の多い「外出」カテゴリを同時に注目対象とされた場合、供給の偏りによる影響によって「健康」カテゴリが特徴として捉えられない可能性が懸念される。検証 3 の出力結果 (図 16) において、供給の多い「外出」カテゴリが 1.000 を示しており、最も特徴として捉えられたが、「防災」「防犯」「電気」カテゴリの値がそれぞれ 0.000, 0.333, 0.000 であるのに対して、「健康」カテゴリは 0.667 と比較的高い数値が出力されたため、特徴として捉えられると考える。

#### 4.6.3 情報値の特徴

「価格」「人気度」「季節性」という情報に注目した検証 5, 6 については、「安価」「時期的なもの」という注目対象を特徴として抽出できていることが確認できた。しかし、

検証 5 では注目していた「安価」が 0.881 対し、カテゴリにおける「人気度:0.748, 季節性:0.700」も高い値で特徴として捉えられており、検証 6 では「安価, 時期的なもの」に注目していたのにも関わらず、カテゴリにおける「人気度」の特徴値が最大値になっている。表 5 は、検証 6 における価格, 季節性, カテゴリの人気度の個人基準値と特徴値をまとめたものである。価格と季節性はそれぞれ 0.299, 0.680 となっており、「安価, 時期的なもの」と、実際に検証 6 で注目していたような傾向が出ている。カテゴリの人気度は 0.574 となっており、「人気傾向にある」ものに注目していたことになる。季節性と人気度は相関関係を持たせるように定義をしていたため、人気度が特徴として捉えられたことは正しい結果である。このことは、ユーザが注目していた情報とは別に「人気度」も活用することで、ユーザに適した、かつ意外性のあるレコメンドができることを示したものと考えられる。

## 5. おわりに

本研究の目的としていた、「新たなサービスの提示を行う」という点は今回検証できなかったが、暗黙的な関係性を示せたことから、従来の手法では困難だった「ユーザ嗜好を多面的に捉える」という部分は達成できた。この手法を利用して、レコメンドにおいて意外性・柔軟性といった精度向上に繋がれることが考えられる。なお、今回の評価では「判定基準値の閾値を 1」に設定することで、特徴抽出が行えたがどうかを判定したが、閾値の妥当性と設定方法については今後実際のデータを用いて検証を行う。

今後の課題として、オントロジー構造面において、多面性は要素に依存し、その要素を手で追加していることから、視点が欠けている可能性がある。そこで口コミなどを利用した要素の補完が必要と考えられる。また、特徴抽出手法の面において、ユーザ嗜好を数値で示すことが可能となったが、注目する情報値が動的に変化するなど複雑になった場合にも対応できる柔軟な特徴抽出を行えるようにする必要はある。そのため、ニューラルネットワークなどを利用した手法の提案や、購入履歴のほかに、検索履歴や選択した項目などの操作もユーザ嗜好として解析を行いたいと考えている。

## 参考文献

- [1] 洞淵, 高田, 梅田, 城, “スマートヘルスケアのための身体活動レコメンドシステム”, 人工知能学会, セマンティックウェブとオントロジー研究会第 32 回 SWO 研究会, 2014 年 3 月
- [2] 日本生気象学会, “日常生活における熱中症予防指針 (Ver.3),” 2013.
- [3] 和室, 北形, 長谷川, “オントロジーを活用した多様な利用者環境に適用可能な IoT サービス構成手法”, 情報処理学会, 2021-DPS-186, 2021 年 3 月



- [4] 堀井, 富士谷, 堀, 乾, 山村, 芥子, ”嗜好に基づくレコメンドソリューションの実現”, SHARP technical journal, 41, 2015年7月
- [5] 富士谷, 村尾, 望月, 西尾, ”コンテンツの多様性を考慮したクロスドメイン推奨”, 情報処理学会論文誌, 57, 2016年7月
- [6] 法造-オントロジーエディタ 操作マニュアル [http://www.hozo.jp/oe5\\_manual\\_jp.pdf](http://www.hozo.jp/oe5_manual_jp.pdf), 2021/1/18 閲覧
- [7] 板垣, 森, 撫中, ”ソリューションと機器推薦のためのオントロジースキーマ及び推薦方法の検討”, 人工知能学会, セマンティックウェブとオントロジー研究会第53回SWO研究会, 2021年3月

付録

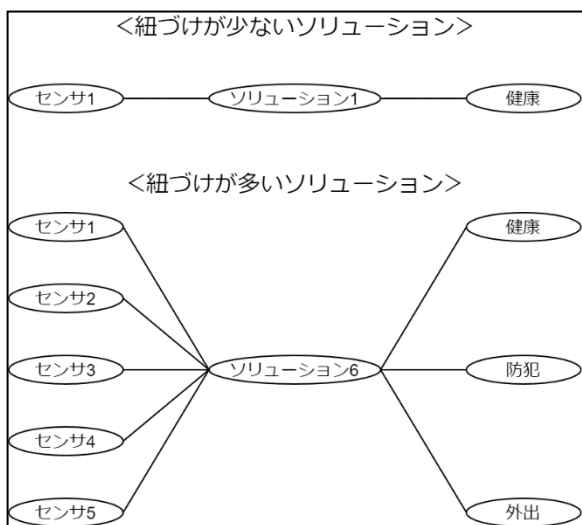


図 17 紐づけの偏り例

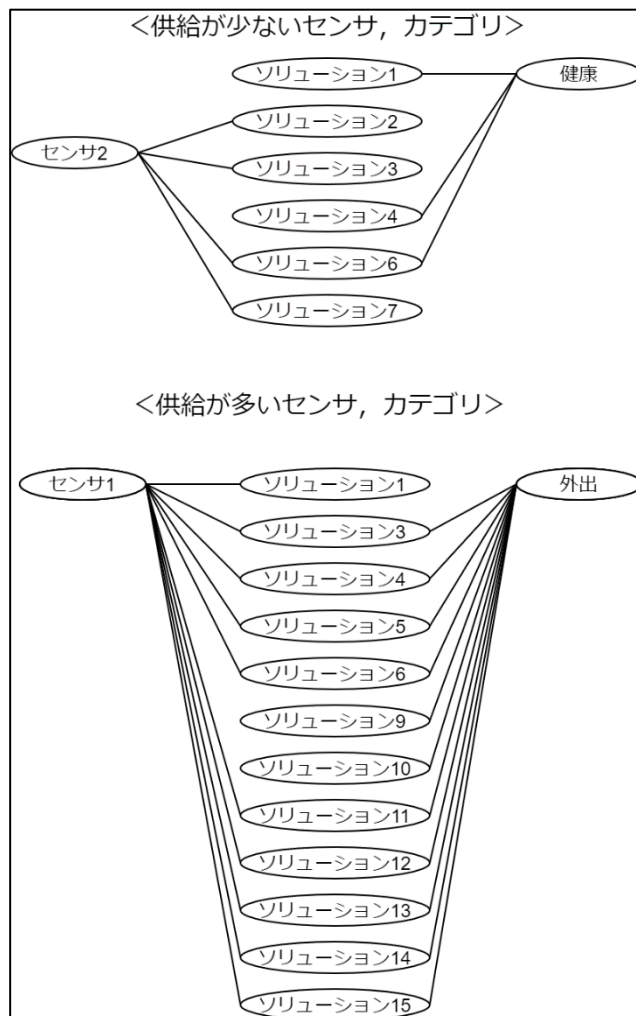


図 18 供給の偏り例