

卒業論文 2006年度(平成18年度)

人の物履歴に基づくユーザプロファイリング機構の構築

指導教員

慶應義塾大学環境情報学部

徳田 英幸

村井 純

楠本 博之

中村 修

高汐 一紀

湧川 隆次

慶應義塾大学 環境情報学部

鈴木 慧

suzuk@sfc.wide.ad.jp

人の物履歴に基づくユーザプロファイリング機構の構築

ユーザの嗜好は、特にマーケティングの分野においてサービス提供者に広く利用されている。例えば顧客の購買履歴を分析することで、顧客毎の嗜好や顧客全体の予算傾向を推定できる。近年、PC 上の行動履歴からユーザの嗜好を抽出し、ユーザに合ったサービスのカスタマイズが普及しつつあり、ユーザの利便性を高めている。一方、実世界の自然な動作からユーザの嗜好を抽出し、ユーザの嗜好に合ったサービスを提供することは難しい。

本研究の目的は、実世界におけるユーザの日常動作から嗜好情報を抽出し、ユーザに適したサービスを提供することである。本論文ではユーザがどのような物を机の上や部屋の中に置き、または携帯しているのかといった物利用状況を取得し、取得した物利用状況からユーザプロファイルを生成・更新する。また、抽出したユーザプロファイルを利用して協調フィルタリングを行い、ユーザの嗜好に合った情報の推薦を行うアルゴリズムを提案する。物利用状況の取得手法として、布型 RFID リーダであるスマートふるしきを利用した。

慶應義塾大学 環境情報学部
鈴木 慧

Abstract of Bachelor's Thesis

An Algorithm for Extracting User's Preference based on Historical Usage of User's Objects

The user's preference is being widely used by the service provider especially in the field of marketing to analyze customers' purchasing histories. They can presume an each customers' preference and a trend of the budget of whole customers. In recent years, the user's preference is extracted from the action history on PC. It is being spread a customized service for using that user's preference and it improves the user's convenience. On the other hand, it is difficult for the user to extract the user's preference from behavior in daily life of the real world.

We are focusing on extracting the user's preference from the user's daily behavior in real world, and providing service being match for the user's preference. Therefore, we make a proposal an algorithm for the recommendation based on user's preference which is used of collaboration filtering.

The collaboration filtering use the model of the user's interest made from the user's action histories which are that how kind objects were put on the desk, user's room, and were brought with user.

We used the Smart-Furoshiki of fabric RFID-Reader for historical usage of user's objects.

Kei Suzuki

Faculty of Environmental Information Keio University

目次

第1章	序論	1
1.1	背景	1
1.1.1	PC上の動作履歴を利用した情報推薦	1
1.2	パーソナライゼーション, ユーザプロファイリング, ユーザプロファイル	2
1.3	問題意識	3
1.4	本論文の目的	3
1.5	本論文の構成	5
第2章	ユーザプロファイリングに基づく情報推薦	6
2.1	ユーザプロファイリングに基づく情報推薦の流れ	7
2.2	ユーザプロファイリングフェーズ	8
2.2.1	明示的手法	9
2.2.2	暗黙的手法	9
2.2.3	ユーザプロファイリング手法の比較	10
2.3	情報推薦フェーズ	11
2.3.1	コンテンツに基づく情報推薦	12
2.3.2	協調フィルタリングによる情報推薦	13
2.4	本章のまとめ	14
第3章	人の物履歴に基づく情報推薦の提案	15
3.1	採択するユーザプロファイリング手法と情報推薦手法	16
3.2	機能要件	16
3.3	アプローチ	17
3.3.1	物履歴の取得対象と取得する動作	17
3.3.2	物履歴におけるユーザの嗜好や興味の仮定	18
3.4	物履歴の定義	19
3.5	ユーザの嗜好の分類: 短期的嗜好と長期的嗜好	20
3.6	ユーザプロファイルの生成	20
3.7	ユーザプロファイルの更新	21
3.8	協調フィルタリングによる情報推薦	21
3.9	本章のまとめ	23

第4章	設計	24
4.1	全体概要	25
4.1.1	UPOHのモジュール分け	25
4.2	ハードウェア構成	26
4.3	ソフトウェア構成	27
4.4	物履歴取得部	28
4.5	物履歴管理部	28
4.6	ユーザプロファイリング部	29
4.7	本章のまとめ	29
第5章	実装	30
5.1	概要	31
5.2	実装環境	31
5.3	物履歴取得部	32
5.3.1	バックの物履歴取得	33
5.3.2	本棚の物履歴取得	34
5.3.3	物履歴取得部のまとめ	34
5.4	物履歴管理部	35
5.5	ユーザプロファイリング部	35
5.5.1	OHModelを用いたユーザプロファイル生成アルゴリズム	35
5.6	サンプルアプリケーション	36
5.7	本章のまとめ	37
第6章	UPOHの評価	38
6.1	定量評価	39
6.1.1	OHModelによるレーティングの評価	39
6.1.2	OHModelが利用したユーザインタラクションの評価	40
6.1.3	リコメンデーションの評価	40
6.1.4	UPOHのパフォーマンス評価	41
6.2	定性評価	42
6.3	本章のまとめ	44
第7章	結論	45
7.1	今後の展望	46
7.2	まとめ	46

目 次

1.1	パーソナライゼーションの概要	2
2.1	ユーザプロファイリングに基づく情報推薦の流れ	7
2.2	ユーザプロファイリングモジュール図	8
2.3	3つの本に対するユーザの評価表	13
2.4	協調フィルタリングによる評価後	13
3.1	ユーザの自然な動作を利用した情報推薦の流れ	18
3.2	ユーザ A の評価ベクタ生成	22
3.3	各ユーザに対する情報の推薦	23
4.1	モジュール分け	25
4.2	ハードウェア構成図	26
4.3	ソフトウェア構成図	27
5.1	物利用状況の取得	33
5.2	スマートバッグ	34
5.3	物履歴取得部からの XML	35
5.4	ユーザプロファイリング部からの XML	36
6.1	物履歴数に対するプロフィール生成時間の変化	42
6.2	SmartShelf を使った取得した物履歴数に対するストレージ容量の変化	43

表 目 次

2.1 ユーザプロファイリング手法の比較	11
5.1 実装環境	32
5.2 開発環境	32
6.1 本 10 冊の評価 評価尺度 (1~5)	39
6.2 本 10 冊の評価 評価尺度 (1~5)	40
6.3 本 10 冊の評価 評価尺度 (1~5)	40
6.4 実験環境	41
6.5 人間の自然な行動からユーザプロファイルを生成する手法の比較	43

第1章 序論

1.1 背景

インターネット上のアクセス可能な情報やサービスは、急な増加傾向にある。世界中でアクセス可能な Web ページ数は、1999 年時点で 8 億ページであったが [1]，2005 年時点では 115 億ページに増加している [2]。それに伴い、e-commerce サイトなど既存のサービスだけでなく、SNS など新しいサービスも増加している。ユーザのニーズを満たす可能性が高まる一方、Web 全体の容量は一個人で把握するのがほぼ不可能なほど増大した。

ユーザにとって、このような情報の増加傾向は自分が必要としない情報に出会う機会が増加するため負担になる。また情報の増加に伴いその内容も多様化しており、ユーザのニーズも多様化している。そのため、膨大な情報の中からユーザ個人に適した情報を選択し、個人に合わせてカスタマイズし推薦することが重要となる。

このように、ユーザ個人にカスタマイズするサービスを実現するためには、ユーザの嗜好や興味に関する情報をシステムが取得し、その情報に適した推薦を行う必要がある。

1.1.1 PC 上の動作履歴を利用した情報推薦

近年、PC 上の動作履歴からユーザの嗜好や興味を抽出し、ユーザ毎にカスタマイズしたサービスが普及しつつある。米 Amazon 社 [3] は、過去の購買履歴やサイト訪問履歴などの動作履歴から、ユーザごとにカスタマイズした情報推薦を行っており、ユーザは自分の嗜好や興味に適した商品の中から選択できる。協調フィルタリングを使い、動作履歴に関連するユーザの嗜好を反映した商品を推薦する。顧客の嗜好や興味を抽出するアルゴリズムが優れているため、推薦された商品は顧客の興味を強く引きつける。また、米 Google 社のサービスの 1 つ GoogleNews[4] は、記事の閲覧履歴からユーザごとにカスタマイズした記事を掲載する。ユーザと同様の記事を閲覧した他ユーザの閲覧履歴から、より多くの人数が閲覧したニュースを掲載する。

このように、ユーザの動作履歴を分析し、ユーザの嗜好や興味を抽出することで、ユーザ毎にカスタマイズしたサービスを提供することができる。

1.2 パーソナライゼーション, ユーザプロファイリング, ユーザプロファイル

本論文では, ユーザ毎に合わせたカスタマイズを行うサービスをパーソナライゼーションと呼ぶ. パーソナライゼーションを行うためには, ユーザの嗜好や興味, 目的, 状況を取得し, これらの要素に適した情報推薦を行う必要がある. 本論分では, 上記要素の中からユーザの嗜好や興味を取得する手法についてフォーカスする. また, ユーザの嗜好や興味を取得する技術をユーザプロファイリングと呼ぶ.

ユーザプロファイリングは, ユーザの嗜好や興味の特徴を表すユーザモデルを作成する. このユーザモデルをシステムが利用できるように記述した情報をユーザプロファイルと呼ぶ. ユーザプロファイルとは, ユーザとのインタラクションによって取得したユーザの嗜好や興味に関する情報である. 具体的には前節の例を用いると, ユーザの購買履歴やサイト訪問履歴, ニュース閲覧履歴から取得した, ユーザはどんな商品を好むのか, どんな記事に興味があるのかといった情報である.

システムは, ユーザプロファイルからユーザ個人の嗜好や興味を解釈するため, パーソナライゼーションにおいて, どのような手法でユーザプロファイリングを行うのが重要になる.

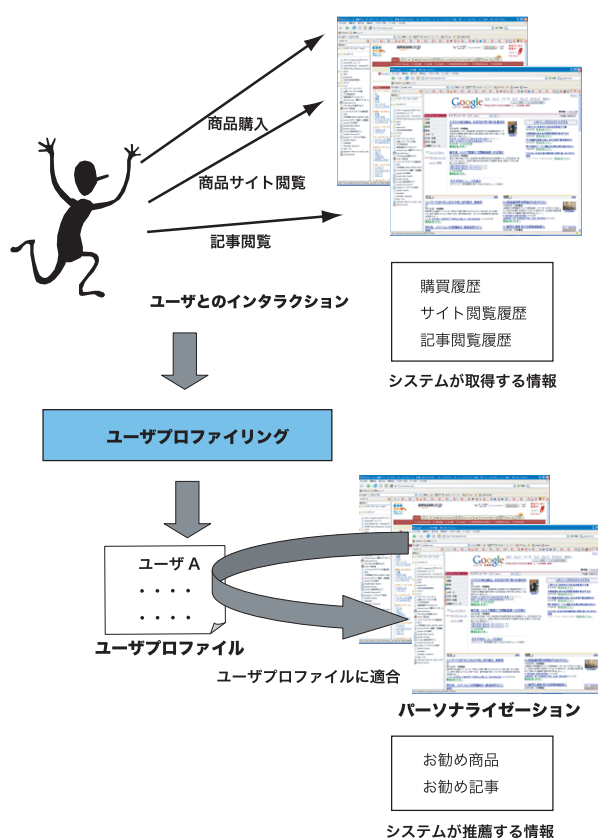


図 1.1: パーソナライゼーションの概要

1.3 問題意識

前節で述べたとおり，ユーザプロフィールは購買履歴や Web 参照履歴など，PC 上でのユーザの動作履歴から抽出できる．一方，PC 上の動作履歴として取得できないユーザの動作から，ユーザプロフィールを抽出し情報推薦に利用することは難しい．例えば Amazon の商品情報推薦システムで，購入後使っていない物は推薦する商品情報に対する影響を低減させるといったことはできない．また，人から貰い受けた物は推薦する商品情報に反映できない．同様に GoogleNews などの記事情報推薦システムにおいて，実世界で読む記事は記事推薦情報に反映できない．

来るべきユビキタスコンピューティング環境では，生活空間にコンピュータやネットワークが溶け込むようになる．現在 PC 上で提供される情報やサービスは，PC の前にいなくても利用できるようになり，普段の生活の中でより自然に利用できるようになると考えられる．このような環境では，実世界のユーザの動作に応じてユーザの嗜好や興味を抽出し，情報やサービスをカスタマイズする必要がある．また嗜好や興味に利用する実世界のユーザの動作は，ユーザにコンピュータを意識させないために，ユーザが日常良く行っているような自然な動作が望ましい．

そのため，実世界の自然な動作からユーザプロフィールを抽出するユーザプロファイリングの手法が必要となる．しかし，まだそのような手法は確立されていない．

1.4 本論文の目的

本論文の目的は，実世界のユーザの自然な動作からユーザの嗜好や興味を抽出し，ユーザに適したサービスを提供する手法の提案・構築である．

人間は自分の好みや関心事に合った本を読んだり，バックに入れて持ち歩いたりする．そのため，ユーザの嗜好や興味は PC 上の動作だけでなく，実世界の自然な動作にも良く反映されていると考えられる．本論文では，ユーザの自然な動作から動作履歴を取得し，この動作履歴からユーザプロフィールを抽出したい．

ユーザが机の上や部屋に置いている物や携帯する物は，ユーザの個性だけでなく，取り組んでいる仕事や興味のある物と関連性が高いと考えられる．そこで実世界のユーザの自然な動作として，ユーザの物履歴に着目する．

本論文では，普段ユーザがどのような物を机の上や部屋の中に置き，または携帯しているのかといった物履歴からユーザプロフィールを抽出する手法を提案する．また，物履歴を取得する対象の範囲を，ユーザが個人で使う机の上と本棚，それにバックの中に絞る．これらの対象には，ユーザのプライベートな物が置かれるため，ユーザの嗜好や興味の特徴を良く現していると考えたためである．本論文では，取得した物履歴の動作履歴から抽出するユーザプロフィールを物利用モデルと呼ぶ．物利用モデルは，ユーザの実世界における物履歴の特徴量を表す．

なお，本研究では市販品に標準で RFID タグがついている環境 [5] を想定環境としている．

本研究では、物履歴を利用してユーザプロフィールを生成するシステム UPOH を提案し、システムを実現するミドルウェアを構築する。また、構築したミドルウェアの有用性を検証する。

1.5 本論文の構成

本論文では、第2章において関連研究を紹介し、第3章では人の物履歴に基づくユーザ嗜好抽出について述べる。第4章では物履歴に基づくユーザプロファイリングシステムであるUPOHの設計について、第5章ではUPOHの実装について述べる。第6章ではUPOHの評価を行い、第7章で今後の第8章ではまとめと今後の課題について述べる。

第2章 ユーザプロファイリングに基づく 情報推薦

本章では，ユーザプロファイリングに基づく情報推薦の流れについて述べ，全体を2つに分け整理する．次に，本論文が対象とするユーザプロファイリングの暗黙的な手法について課題を整理する．最後に，本論文の機能要件を整理する．

2.1 ユーザプロファイリングに基づく情報推薦の流れ

本章では、ユーザプロファイリングに基づく情報推薦の流れを2つに分けて整理する。ユーザプロファイリングを行う段階をユーザプロファイリングフェーズ、ユーザプロファイルに基づいた情報推薦をする段階を情報推薦フェーズと呼ぶ。システム全体の流れを、図2.1に示した。

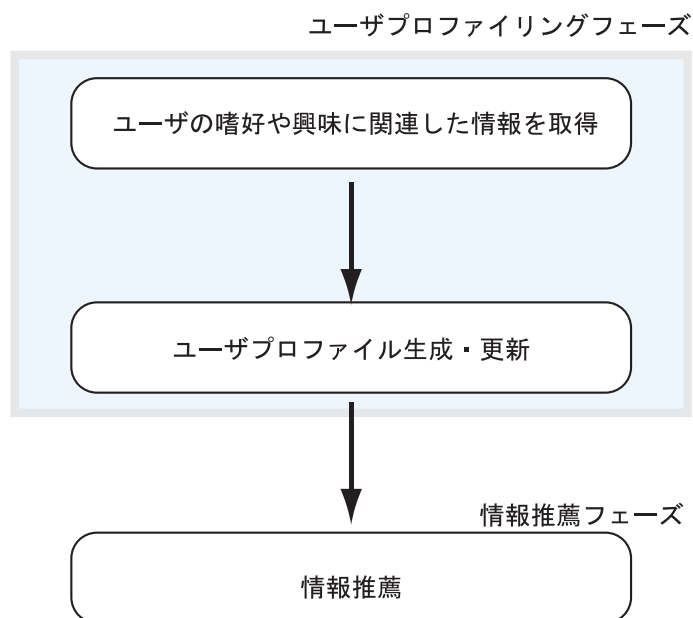


図 2.1: ユーザプロファイリングに基づく情報推薦の流れ

ユーザプロファイリングに基づく情報推薦の目的は、個人に適した情報推薦を行うことである。ユーザが同じ要求を出した場合でも、ユーザの嗜好や興味はユーザ毎に違うため、求める情報の種類や専門性はユーザ毎に異なると考えられる。そのため、情報推薦システムはユーザの嗜好や興味を考慮することが望ましい。

このような情報推薦システムを実現するためには、システムは第1にユーザの嗜好や興味を獲得する必要がある。ユーザの嗜好や興味を獲得するために、システムはユーザの嗜好や興味に関連した情報を取得し、取得した情報から嗜好や興味に関するユーザの特徴を抽出する。例えば、ユーザが参照したWebページからユーザの嗜好や興味を獲得する場合、ページ中の頻出単語やユーザがクリックしたリンクのタイトルなどのデータを取得し、それらデータからユーザの特徴を抽出することが考えられる。また、抽出したユーザの嗜好や興味は、システムが扱えるようにユーザプロファイルとして記述する必要がある。ユーザの嗜好や興味は時間に応じて変化するため、ユーザプロファイルは一度作られるだけでなく更新される必要がある。ユーザプロファイリングフェーズは、ユーザの嗜好や興味に関連した情報を取得し、それら情報からユーザプロファイルを生成・更新する作業を担当する。

システムは第2に、ユーザプロファイルからユーザの嗜好や興味を解釈し、個人に適し

た情報推薦を行う。情報推薦フェーズは、ユーザプロフィールに適した情報推薦を行う作業を担当する。

以上より、ユーザプロファイリングに基づく情報推薦の流れの概要を述べた。ユーザプロファイリングフェーズについて、第 2.2 節で述べ、情報推薦フェーズについて、第 2.3 節で述べる。

2.2 ユーザプロファイリングフェーズ

ユーザプロファイリングフェーズでは、ユーザの嗜好や興味に関連した情報の取得とユーザプロフィール作成・更新を行う。本節では、ユーザプロファイリングフェーズをイメージし易くするために、最初にシステム構成について述べる。本論文ではシステムの説明上省略のために、ユーザの嗜好や興味に関連した情報をユーザ関連情報と呼ぶ。システムは、ユーザ関連情報取得モジュールと、ユーザモデルを作成する User Modeler モジュールと、ユーザプロフィール管理モジュールで構成される。各モジュールの関係を図 2.2 に示す。

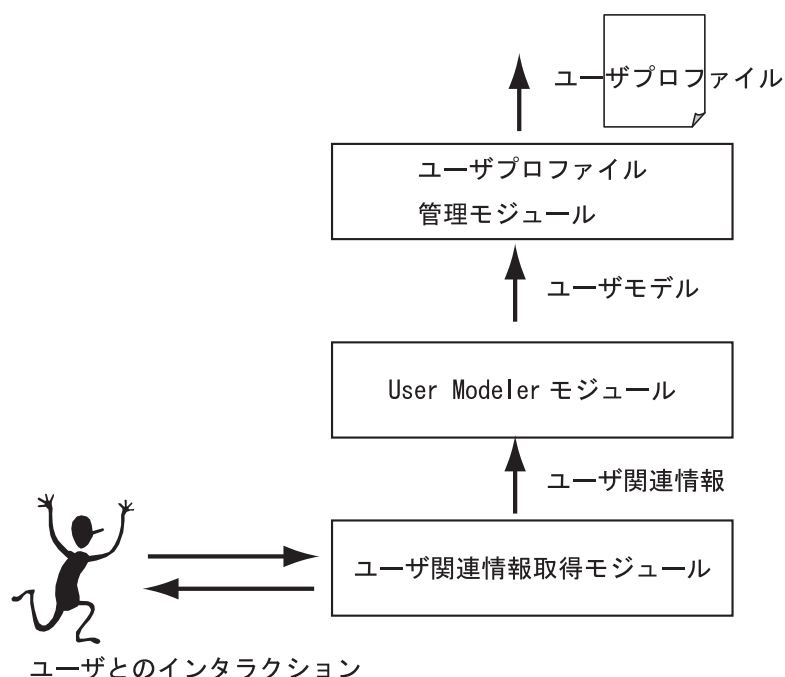


図 2.2: ユーザプロファイリングモジュール図

一般的に、ユーザ関連情報取得モジュールはユーザとのインタラクションから情報を取得する。User Modeler モジュールでは、取得したユーザ関連情報からユーザの嗜好や興味の特徴を表すユーザモデルを作成する。最後に、ユーザプロフィール管理モジュールはユーザプロフィールの生成や更新を行う。

ユーザプロファイルは、ユーザの嗜好や興味の特徴を表すユーザモデルが記述されたファイルである。具体的には、ユーザプロファイルにはユーザの嗜好や興味を表すキーワードや特定のデータ項目などが記述される。例えば、ユーザの Web 参照からユーザプロファイルを作成すると参照ページ中に含まれる頻出単語やページの参照時間が考えられる。この例のようにユーザプロファイルに記述されるユーザモデルは、取得するユーザ関連情報に依存する。

そこで、本節ではユーザプロファイリングの手法をユーザ関連情報の取得において、ユーザの負担という観点から 2 種類に分けて整理する。第 2.2.1 節で明示的手法について述べ、第 2.2.2 節で暗黙的手法について述べる。

2.2.1 明示的手法

明示的手法とは、ユーザ嗜好や興味に関する情報をユーザに直接入力してもらい、ユーザ関連情報を取得する手法である。明示的手法は、大きくアンケート調査手法と、評価付け手法に分けられる。アンケート調査手法は、ユーザの嗜好や興味に関するキーワードを、ユーザに答えてもらう手法である。T. Yan らによる SHIFT[6] は、ユーザに興味のあるトピックのキーワードをメールで送ってもらう。システムは、メールに書かれているキーワードが良く含まれる情報を推薦候補から選択し、ユーザに推薦する。ユーザプロファイルはメールで送られたキーワードのベクトル空間モデル、またはブール代数モデルが記述される。

評価付け手法は、特定の情報に関するユーザの嗜好や興味尺度を答えてもらう手法である。例えば、Semantic Differential 法 [7] を用いた場合、次のようにユーザのジャズに対する好き嫌いの尺度を 1-5 段階で答えてもらう。

ジャズ： 嫌い 1 2 3 4 5 好き

また、特定の Web ページに対しての参考になったかどうかユーザに評価付けしてもらうことで、その Web ページに含まれる情報のユーザの嗜好や興味尺度を取得できる。

システムは評価付けされた対象の情報から、ユーザの嗜好や興味を推定する。Ken Lang による NewsWeeder[8] は、ユーザに特定の Web ページを 1 から 5 の 5 段階で評価してもらい、評価付けされた Web ページに含まれる頻出単語を tf-idf で取得し、評価値に基づいて頻出単語を重み付けし、ユーザプロファイルを作成する。

2.2.2 暗黙的手法

暗黙的手法とは、ユーザに特別な操作を要求せずに、システムが自動的にユーザの嗜好や興味を獲得する手法である。特別な操作とは、ユーザの嗜好や興味を獲得するために要求する操作である。具体的には、ユーザの Web 参照履歴や GPS を用いた行動履歴などのユーザ関連情報を、システムが自動的に取得しユーザの嗜好や興味を抽出する。

暗黙的手法では，取得したユーザ関連情報がユーザの嗜好や興味と関係のない情報を含む可能性を考慮する必要がある．そのため，取得したユーザ関連情報からユーザの嗜好や趣味の特徴を発見するために，データマイニングを行う必要がある．

データマイニング データマイニングとは，統計学，パターン認識，人工知能等のデータ解析の技法を大量のデータに網羅的に適用することで知識を取り出す技術である．データマイニングは特にマーケティングの分野においてサービス提供者に広く利用されてきた．例えば，米国の大手スーパーマーケットチェーンでは，販売データをバスケット分析した結果「顧客はおむつとビールを一緒に買う傾向がある」ことが分かった．そこでこの2つを並べて陳列したところ，売り上げが上昇した．

暗黙的手法は，ユーザの自然な動作を取得し，取得した動作履歴からユーザプロフィール作成を行う．次に暗黙的手法の研究例として，Web ページ閲覧中のマウス操作やユーザの位置情報から得られる行動履歴，視線の注目時間に着目した研究を紹介する．

土方らの `textextractor` は [9]，ユーザの Web 閲覧中のマウス操作を利用して，ユーザが興味を持ったと思われるテキスト部分を全体のテキストから自動的に抽出する．具体的には，ユーザのなぞり読みやリンクポインティング，リンククリック，テキストの選択といったマウス操作を利用している．ユーザが興味を持ったと思われる部分からユーザプロフィールを作成するため，ユーザプロフィールに記述されるキーワードの精度を高めている．

大阪大学の中里祐介らは [10]，GPS を使ったユーザの位置情報で得られる行動履歴からユーザプロフィールを作成する．具体的には，ユーザの訪れる頻度が高い場所はユーザの興味がある場所であるという仮定に基づき，特定の場所への訪問頻度や滞在時間から店舗の重みを増加させる．ユーザプロフィールは，取得した行動履歴から重み付けされたクラス関係図で記述される．

広島大学大学院の脇山らは [11] ユーザが注目していた部分を検出し，注目していた対象からユーザプロフィールを作成する．具体的には，注目している時間が長い領域ほど興味の度合いが大きいという仮定に基づき，注目している領域と注目時間よりユーザの嗜好や興味対象を取得する．ユーザプロフィールは，注目時間を基に重み付けされたベクトル空間で記述される．

2.2.3 ユーザプロフィールリング手法の比較

本節では，先述した明示的手法と暗黙的手法について利点や欠点を整理し，比較する．

明示的手法の利点 取得したユーザ関連情報は，ユーザが直接答えたものであるために信頼性が高い利点がある．そのため，暗黙的な手法に比べて作成されるユーザプロフィールにはノイズが少ない．また，他のユーザの Web ページに対する評価を参考に情報推薦を行うこともできるため，全てのユーザに嗜好や興味に関して答えを要求せずに情報推薦を行うこともできる．

明示的手法の欠点 ユーザの嗜好や興味を抽出するために特別な作業を要求するため、ユーザに負担がかかる欠点がある。アンケートに答える、評価付けするという作業は、本来は必用のない面倒くさい作業であるため、ユーザが協力してくれないことが考えられる。ユーザの嗜好や興味は時間の経過に応じて変化するものであり、システムはその変化を取得できない可能性があるため、個人ごとに適した情報推薦システムを長期で運用するのは難しい。

暗黙的手法の利点 ユーザの嗜好や興味を抽出するために、ユーザに特別な作業を強いな利点がある。そのため、ユーザはパーソナライゼーションのシステムを意識することなく、個人に適した情報推薦を利用できる。また、ユーザに意識的に嗜好や興味について答えさせる必要がないため、ユーザ本人も気づいていない潜在的な嗜好や興味を発見できる可能性がある。

暗黙的手法の欠点 ユーザの自然な動作からユーザ関連情報を取得するため、取得した情報にノイズが含まれる可能性が高い欠点がある。ユーザの自然な動作から得られるユーザ関連情報には、ユーザの嗜好や興味の特徴を現していない情報も含まれる。そのため、システムはユーザ関連情報からうまく抽出しなければならない。また、ユーザ関連情報を取得するために先述した [11] のように、ユーザに特別なデバイスを持たせた場合、デバイスを持つこと自体がユーザにとって負担になるという問題がある。

以上より、明示的手法と暗黙的手法それぞれに利点と欠点があり、それらは双方で補完する関係にある。明示的手法と暗黙的手法の利点と欠点を表 2.1 に示す。

手法	信頼性	ユーザ負担
明示的手法		×
暗黙的手法		

表 2.1: ユーザプロファイリング手法の比較

ユーザの負担という観点からすると、暗黙的手法のみからユーザの嗜好や興味を推定するのが良いが、高い精度を求めるのはなかなか難しい。そこで、明示的手法と暗黙的手法の両方を用いてユーザの嗜好や興味を獲得する手法も考えられる。また、暗黙的手法において、ユーザに特別なデバイスやソフトウェアを利用してもらう必要がある場合、それ自体がユーザの負担にならないようにしなければならない。

2.3 情報推薦フェーズ

情報推薦フェーズでは、ユーザプロファイルを基に情報推薦を行う。ユーザプロファイルに基づいた情報推薦は、コンテンツに基づく情報推薦と協調フィルタリングの2種類に

分けることができる。前者は、推薦する情報の内容に基づき、情報の取捨選択を行う。後者は、ネットワーク上に存在する同じ好みを持ったコミュニティを発見し、そのコミュニティが共通して好む情報を選択する。それぞれの手法によって必要とするユーザプロフィールが大きく違う。コンテンツに基づいた情報推薦の方が、より細かい単位で嗜好や興味情報を必要とする。

それぞれ情報選択の基本的な考え方は異なるが、いずれにしてもユーザはどの情報が好きなのか、あるいはどの情報に興味があるのかという情報が必要となる。ただし、このような嗜好・興味に関する情報をどの程度の粒度で必要かは、それぞれの手法で異なる。コンテンツに基づくフィルタリングでは、テキスト情報の解析をキーワード単位で行うことが多いため、キーワード単位で興味の有無が分かる方が良いと言える。協調フィルタリングでは、推薦する情報単位、つまり Web であればページ単位で、複数のユーザの興味に関する情報を解析するため、ページレベルで興味の有無が分かれば十分なケースが多い。コンテンツに基づくフィルタリングの方が、より細かい単位で興味を推定するため、ユーザプロフィール技術の開発も困難なものになると言える。

第 2.3.1 節でコンテンツに基づく情報推薦について述べ、第 2.3.2 節で協調フィルタリングについて述べる。

2.3.1 コンテンツに基づく情報推薦

コンテンツに基づく情報推薦とは、推薦する情報の内容に基づき情報の取捨選択を行う手法である。システムは、推薦候補の中身を見て、推薦情報は、ユーザプロフィールと推薦候補の情報との類似度から決定される。例えば Web ページを対象とする場合、ページ内のテキスト情報をキーワード単位で解析し、ユーザプロフィールに記述されているキーワードのベクトル空間との類似度から推薦するか否かを決定する。

キーワード間の類似度の算出方法として、完全照合方式と部分照合方式の 2 種類がある。完全照合方式の代表的な手法として、タウベらが開発したブール代数 (Boolean algebra) に基づく検索理論 [Taube 55] が上げられる。数個のキーワードを取り出し、それらを AND や OR, NOT でつないで、適合する文書を探し出す。部分照合方式の代表的な手法として、サルトンらによるベクトル空間モデル (vector space model) [Salton 83] が上げられる。ベクトル空間モデルは、その名の通りキーワードによるベクトルを生成して、これを文書の近さの計算に利用するものである。

次にコンテンツに基づく研究例として、tf-idf を利用したコンテンツに基づく情報推薦の研究を紹介する。

Tomonari Kamba らは [12]、Web ニュースの記事とユーザプロフィールとの部分照合方式による類似度からユーザに推薦記事を提示する。具体的には、Web ニュースの記事を tf-idf を使って解析し、記事から抽出されたキーワードごとに重み付けする。次にユーザプロフィールに記述された重み付けされたキーワードのベクトル空間と、記事ごとに抽出されたキーワードのベクトル空間との類似度を計算し、類似度が高い記事を推薦記事とする。ユーザプロフィールは、ユーザの嗜好や興味に関するキーワードごとに重み付けしたものが必要となる。

2.3.2 協調フィルタリングによる情報推薦

協調フィルタリングとは、ユーザの評価データだけでなく他人の評価データも参考にして情報推薦を行う手法である。具体的には、被情報推薦者のユーザプロフィールと他のユーザのユーザプロフィールとの間の類似度を算出し、類似度の高いユーザの推薦情報を優先して情報推薦に反映する。例えば、複数の映画に対する評価を多数のユーザから集めておき、各ユーザがどの映画を好むのかを把握できるようにする。次にあるユーザに対して情報推薦する場合、同じような映画を好む他ユーザが高く評価している映画を推薦する。

一般的に協調フィルタリングを行う Web サイトにて入手しているユーザ情報には、評価データとデモグラフィックデータの 2 つがある。評価データとは、ユーザから 5 段階のアンケート形式などで得られるアイテムに対する評価値である。デモグラフィックデータとは、性別、年齢などその人の持つ特質を表すデータである。協調フィルタリングはユーザのアイテムに対する嗜好を記録し、そのユーザと似た嗜好を持つ他のユーザグループの情報を基に、未知のアイテムに対するユーザの嗜好を推測する手法である。

ここで、協調フィルタリングをイメージし易くするために、協調フィルタリングの一つである J.Riedl らの GroupLens の方式 [13] を用いて例を紹介する。次に説明される例について、図 2.3 と図 2.4 に示す。

3 つの本の中からお勧めする例 図 2.3 は、3 つの本に対する 4 人のユーザの評価を 1-5 段階で評価付けしたものである (数値が大きいほど良い評価)。

	Effective C++	PHP5 徹底攻略	プログラミング言語 C
A	5	2	?
B	1	4	1
C	4	2	4
D	5	5	5

	Effective C++	PHP5 徹底攻略	プログラミング言語 C
A	5	2	4
B	1	4	1
C	4	2	4
D	5	5	5

反映

図 2.3: 3 つの本に対するユーザの評価表 図 2.4: 協調フィルタリングによる評価後

この例では、ユーザ A が情報推薦システムを利用してプログラミング言語 C を見るべきかどうか教えてもらいたいとする。図 2.3 中では、ユーザ A はまだプログラミング言語 C を読んでいないため、評価項目は ? になっている。ユーザが既に評価付けした他の 2 冊の本の評価を見ると、ユーザ A とユーザ B は他の 2 冊の本に対する評価が、他のユーザ間よりも類似していることがわかる。協調フィルタリングでは、図中の A と C のように、嗜好や興味が類似したユーザの意見を参考にする。この例

では、ユーザ A に対する情報推薦をするために、最も類似したユーザ C の評価を参考にした結果、図 2.4 のようにシステムはユーザ A さんに対してプログラミング言語 C を評価 4 として推薦する。

協調フィルタリングの問題点として、多数のユーザ評価データが必用となるため、評価対象アイテム数や参加ユーザ数が少ない場合、情報推薦の精度は信用できない。また、協調フィルタリングには、ユーザがこれまでに付けてきた評価値が大きく変わらない限り、毎回同じようなグループが特定され、毎回似たようなアイテムが推薦されてしまうという問題がある。

次に協調フィルタリングによる情報推薦の研究例として、ソフトウェアの未使用機能を推薦するものと、商品を推薦する研究を紹介する。

奈良先端科学技術大学院大学の杉村 [14] はソフトウェアを利用する際に、ユーザが実行する回数が多い機能ほど評価を高くし、協調フィルタリングを使いユーザにとって有用な未使用の機能を推薦するソフトウェア推薦システムを構築した。協調フィルタリングを行うために、機能の利用頻度によってユーザのソフトウェアの特定機能に対する嗜好情報を抽出している。その結果、同じような利用状況にあるユーザが高い評価を付けた機能が推薦される。

Greg Linden らの Amazon.com Recommendations [15] は、アイテムとアイテムの間 (item-to-item) で協調フィルタリングを行い商品情報を推薦している。item-to-item 方式を利用しているため、膨大な顧客数と膨大な商品数を抱えるシステムだが、商品数に対する計算量のみを考えれば良くスケーラビリティに優れる。ユーザプロファイル (アイテムプロファイルと言ったほうが適切だが) は、2 つの商品を一緒に買った人がいるかによって商品ごとに重み付けを行うベクトル空間となる。

2.4 本章のまとめ

本章では、まずユーザプロファイリングに基づく情報推薦について述べた。次にユーザプロファイリングに基づく情報推薦の流れについて述べ、全体を 2 つに分けて整理した。2 つに分けたユーザプロファイリングフェーズと情報推薦フェーズについて、それぞれを詳しく説明した。最後に、本論文における機能要件を、ユーザプロファイリングに基づく情報推薦という観点から考察した。次章では、本研究の目的を実現するためのアプローチとそのアプローチの具体的な実現手法である物履歴を基にしたユーザの嗜好情報抽出について述べる。

第3章 人の物履歴に基づく情報推薦の提案

本章ではまず，前章を踏まえて本論文の目的を述べ，その機能要件について述べる．次に目的を実現するためのアプローチについて述べる．本論文で採択するユーザプロファイリング手法が利用する物履歴についての定義を行う．物履歴から抽出するユーザの嗜好や興味について整理し，2種類に分けて述べる．その後，ユーザの嗜好情報の抽出について述べる．最後に，物履歴に基づいた情報推薦手法について述べる．

3.1 採択するユーザプロファイリング手法と情報推薦手法

現在，Web アプリケーションを中心に多くのアプリケーションが，パーソナライゼーションを行うために PC 上の動作履歴を利用している．しかし，PC 上の動作履歴として取得できないユーザの動作から，ユーザプロファイルを抽出し情報推薦に利用することは難しい．

本論文の目的は，ユビキタス環境においてユーザの実世界の自然な動作からユーザプロファイルを抽出して情報推薦を行うことである．想定環境として，現在あらゆる商品につけられているバーコードが Passive RFID タグに置き換わった世界を想定している．

第 2.2.3 節より，ユーザプロファイリング手法として，明示的手法と暗黙的手法がありそれぞれに利点と欠点があり，それらは双方で補完する関係にあることが分かった．本論文の目的として，ユーザの実世界の自然な動作を利用したいため，暗黙的手法を採択する．また，暗黙的手法を補完する役割として明示的手法も取り入れ統合する．

本論文では，商品につけられた Passive RFID からその商品の属性情報を取得して利用することは考えていない．そのため，本論文では第 2.3 節より，情報推薦手法として細かい単位で嗜好や興味情報を必要としない協調フィルタリングを採択する．

本論文では，実世界の自然な動作からユーザの嗜好や興味を抽出するユーザプロファイリングにフォーカスし，ユーザに対する負担の低減と精度の向上を目指す．

3.2 機能要件

本節では，ユーザの実世界の自然な動作に基づく情報推薦システムの機能要件を整理する．ユーザの実世界における自然な動作を取得したいため，ユーザの動作を取得する際に，普段の動作に支障を来さないようにする必要がある．また，生成するユーザプロファイルがユーザの嗜好や興味の特徴を良く反映する必要がある．最後に，推薦される情報がユーザにとって有用である必要がある．機能要件として，ユーザに対する負担の低減と生成するユーザプロファイルの精度向上，推薦情報の有用度の向上が考えられる．次に，それぞれの機能要件についてまとめる．

ユーザに対する負担の低減 ユーザに対する負担を低減するには，ユーザに負担にならない暗黙的手法を利用する必要がある．具体的には，ユーザの嗜好を獲得するために特別なデバイスを意識させないことが必要となる．

生成するユーザプロファイルの精度向上 ユーザプロファイルがユーザの嗜好や興味を表現する精度を向上させるには，取得するユーザの自然な動作履歴に含まれるノイズを低減しなければならない．また，取得した動作からユーザモデルを構築するアルゴリズムの精度向上を行わなくてはならない．

推薦情報の有用度の向上 協調フィルタリングによる情報推薦には，誰もが知っているベスト 10 などの情報が提供される事が考えられる．場合によっては，多勢に推薦され

たベスト 10 よりもセンスのある人のトップ 10の方が有用と感じられる場合がある。そのため、ユーザ間の類似度から多勢の意見を反映させた手法だけでなく、有用な推薦情報を持っている特定のユーザの意見について考えなくてはならない。

3.3 アプローチ

本研究では日常生活の中でユーザの嗜好を抽出するために、ユーザの所有物に着目した。人が物を所有することについて、社会心理学の Russell Belk は、所有物は広い意味での自分を表し、物を所持することは自分が他者からこのように見られたいという欲求を示していると考えている [16]。例えば、[16] の Olson によるとシニア世代では過去・ヒストリーに関する個性が強くなり、Cameron によると子供ができた夫婦では、自己に関する個性ではなく子供へと焦点が移る傾向がある。いずれにしても、人が物を所有することは、何に興味がありどのようなイメージを好むのかを示しており、人の所有物から人の嗜好情報を抽出できる。

三菱電機の Shimizu ら [8] はユーザが何を携帯しているかによってユーザが必要としている情報を判断する推薦システムを提案している。システムは、携帯している物とユーザのいる場所、時間、温度などのユーザプロフィールを組み合わせることでユーザの状況を解釈する。ユーザが携帯している物を取得するために、passive RFID タグと RFID リーダを用いる。

ユーザの所有物の中でユーザがほとんど利用せず忘れてしまっている物は、ユーザの嗜好との関連性は薄い。そこでユーザにとって興味や価値のある物をユーザの所有物から解釈するために、ユーザが普段良く利用したり、身近に置いている所有物の利用状況に着目する。本研究ではユーザの身近にある所有物の利用状況からユーザの嗜好を解釈する。以後、ユーザの所有物の利用状況のことを物履歴と呼ぶ。

3.3.1 物履歴の取得対象と取得する動作

ユーザの日常生活には個人的空間と公共的空間がある。個人的空間とは、自分の部屋やオフィスにある自分の机といった場所のことで、個人的な物や情報を置いたりやり取りしたりする。一方公共的空間とは、駅、お店、バスといった場所のことで、自分だけでなく他人の存在を考慮するため、個人的な物や情報を置いたり、やり取りするのに適さない場所でのことである。

本研究ではユーザの所有物に着目するため、システムはユーザの所有物が含まれる個人的な空間から物履歴を取得する。具体的には、机の上に置いてある物や携帯している物、ユーザ個人の部屋に置いてある物などが挙げられる。これらの個人的空間の物は、ユーザの個性を表すだけでなく、他者からどのように見られたいかという欲求 [16] が反映されているため、ユーザの嗜好情報を取得できると考えたためである。本研究では個人的空間にある物を、ユーザが所有している物と考え、以後単に所有物と呼ぶ。

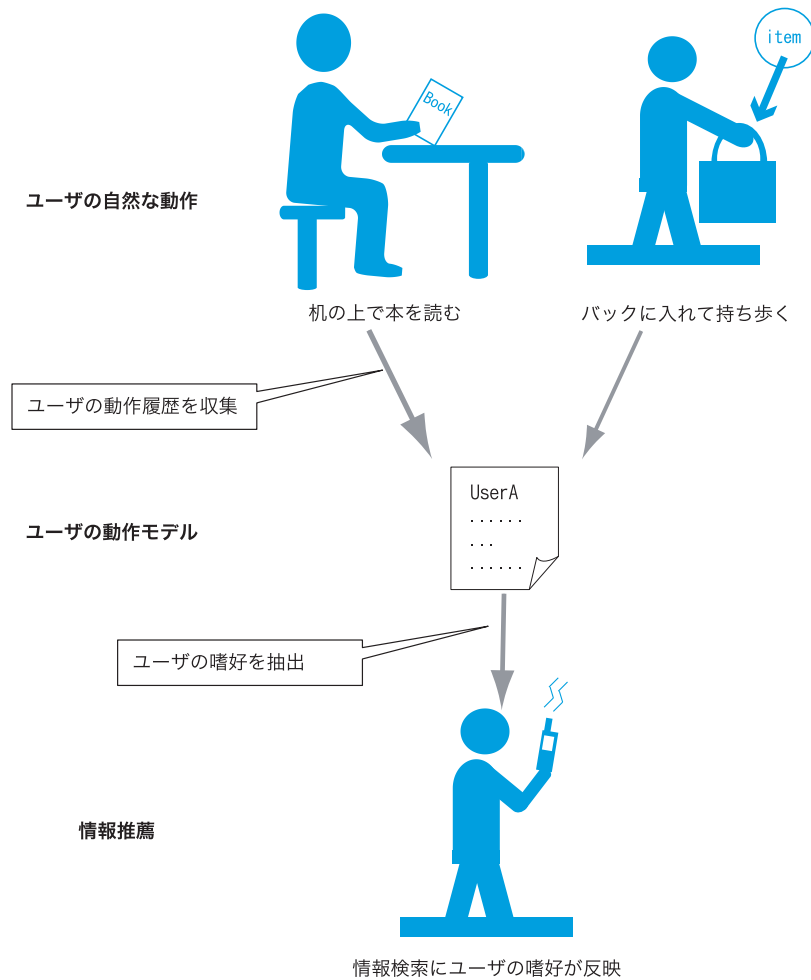


図 3.1: ユーザの自然な動作を利用した情報推薦の流れ

Intel Research の Fishkin ら [6] は物に passive RFID タグを付け、環境側に固定された RFID リーダを設置し、リーダに対するタグの応答率の変化を利用することで、ユーザの動作を検知するアルゴリズムを提案した。ユーザがタグ付けされた物が動かしたことや、物の正面で手を振っている動作や歩いている動作を検出できる。ユーザが物を動かしたことを検知することで、ユーザがどのような物を利用しているのが取得できる。

3.3.2 物履歴におけるユーザの嗜好や興味の仮定

物履歴におけるユーザの嗜好や興味について、本論文では次のように仮定し、第 3.3.1 節で取得した情報に対して重み付けされる。

- ユーザが頻繁に使っている物は、ユーザの嗜好や興味を反映している
- ユーザが近くに置いている物は、ユーザの嗜好や興味を反映している

- ユーザが持ち歩いている物は，ユーザの嗜好や興味を反映している

3.4 物履歴の定義

物履歴を表す指標は，物の特定期間における所有率と利用率とする．システムは物履歴を取得するために，個人的空間にユーザが物を置いた時間を所有開始時間 a_a とし，そこから物を取り出した時間を所有終了時間 a_b として記録する．同様に，所有物をユーザが利用し始めた時間を利用開始時間 b_a とし，ユーザが利用を終了した時間を利用終了時間 b_b として記録する．以後，所有開始時間 a_a と所有終了時間 a_b の間の期間 a_{ab} を生存期間と呼び，利用開始時間 b_a と利用終了時間 b_b の間の期間 b_{ab} を利用期間と呼ぶ．また，システムは物履歴を取得したユーザ空間がどこであるかも記録する．

ここで，ある時間 a から b までの期間を T_{ab} とし， a から b の間の生存期間の集合を $A = \{a_1, a_2, \dots, a_n\}$ ，利用期間の集合を $U = \{u_1, u_2, \dots, u_m\}$ とすると， T_{ab} における所有率 o_{ab} は次の式 3.1 で決める．

$$o_{ab} = \frac{\sum_{k=1}^n a_k}{T_{ab}} \quad (3.1)$$

式 3.1 は，ある期間における生存期間が長いほど大きくなるため，所有率はユーザの所有している期間の足し合わせに比例する．

次に T_{ab} おける利用率 u_{ab} は次の式 3.2 で決める．

$$u_{ab} = \frac{\sum_{j=1}^m u_j}{\sum_{k=1}^n a_k} \quad (3.2)$$

式 3.2 は，生存期間と利用期間の比率を表すため，利用率はユーザの所有物に対する利用頻度を意味する．

ここで物履歴を取得する個人的空間ごとに，ユーザの物履歴から推定できる嗜好情報の違いについて考える必要がある．例えば，机の上に置いたものと携帯している物とでは，同じように全く利用していない物であっても，前者では観賞用に置いてある趣味性の高い物である可能性があり，後者ではバックの中に入れっぱなしになっていて忘れていただけの物である可能性が高いなど，物履歴を取得する場所によって利用状況の示す意味が異なってくる．そのため，物履歴を取得する個人的空間によって，嗜好情報の抽出アルゴリズムを変更する必要があると考えられる．

さらに，利用率の高い所有物ほど重みを上げた場合，辞書や日用品など利用頻度は高いがユーザの興味とあまり関係のない物が，ユーザの嗜好情報に対する影響が高くなる可能性がある．逆に利用頻度は低いがユーザの興味と関連性が高い物はユーザの嗜好情報に余り考慮されない問題がある．前者の辞書や日用品などのユーザの嗜好情報に対する影響に

ついて、ある一定期間以上頻繁に使用している物については、ユーザの嗜好情報に対するノイズとして除去する手法が考えられる。後者について、本研究では物の内容を利用することを考えていないため、物が趣味性の高い物かどうかで利用頻度の低い物の影響を調整することはできない。本稿では、利用頻度の低い物と高い物の間の差を縮めるパラメータを導入し、それを調整することで利用頻度の低い物の影響を反映させる手法を取る。

最後に先述した、物履歴を取得する場所によって利用状況の示す意味が異なってくるものとユーザの嗜好情報に対するノイズや例外に対処するために、ノイズ除去に関するパラメータや利用頻度の低い物の影響を反映させるパラメータを利用状況を取得できるユーザ空間ごとに設定する。

3.5 ユーザの嗜好の分類：短期的嗜好と長期的嗜好

毎日変化するニュースやユーザの日常生活における経験によって、ユーザの興味の対象も移り変わる。そのため、ユーザが良く利用したり身近に置く物も変わり、時間によって物履歴は変化する。変化する物履歴から抽出できるユーザの嗜好情報には、時間ごとの短期的な嗜好と、もっと広い範囲でユーザが元々どのようなものを好むのかといった長期的な嗜好が考えられる。前者は、ユーザがその時に取り組んでいる作業や社会的ニュースの影響が強く、後者はユーザの個性が強く影響する。本研究では、それらの嗜好をそれぞれ、短期的嗜好と長期的嗜好と呼ぶ。本システムを利用するユーザは、ユーザの作業に基づいた情報推薦を望んでいる場合、短期的嗜好に基づいて推薦を受ける。また、自分の個性を強く反映した情報推薦を望む場合は、長期的嗜好に基づいて推薦を受ける。

短期的嗜好はユーザの作業に基づいた嗜好であるため、ユーザが頻繁に利用している物に着目すると良い。短期的嗜好は利用率に基づいて抽出される。長期的嗜好は、ユーザの個性に基づいた嗜好であるため、ユーザが頻繁に利用している物だけでなく、所有しているだけの物にも着目する。長期的嗜好は、所有率と利用率双方を考慮して抽出される。

3.6 ユーザプロファイルの生成

ユーザの嗜好情報に基づいた情報推薦を行うために、膨大な情報の中からユーザの嗜好に合った情報を取り出して推薦を行う情報フィルタリングが必要になる。情報フィルタリングには、推薦候補の内容とユーザの嗜好情報との類似度を比較する手法と、ユーザの嗜好情報と類似した他のユーザの嗜好情報から推薦する手法がある。本研究は、物の内容を考えないため、後者の手法である協調フィルタリングを使った情報推薦を行う。

協調フィルタリングは、ユーザがまだ未知の情報の獲得をする際、嗜好が似ているユーザからのお勧め情報に基づき推薦を行う。例えばC++プログラミング本を所有していてよく利用しているユーザは、C++に興味がある可能性が高く、ユーザは大抵同じようなジャンルの本、CやJavaなどのプログラミング本を所有している可能性が高い。協調フィルタリングを行うことで、同じようにプログラミングに興味のあるユーザが興味を持つ情

報を推薦できる．また，協調フィルタリングでは多くのユーザの意見を反映した情報の推薦を行うため，ユーザの匿名性が保障できる．

協調フィルタリングで必要になるユーザの興味モデルとして，3.1のユーザの所有率と利用率を利用し，物ごとに重み付けした評価ベクタをユーザの興味モデルとして利用して，他のユーザとの類似度を比較する．

ここで，3.1の最後で述べたユーザ空間ごとに設定するパラメータについて決める．ユーザ空間 a について，ノイズ除去に関するパラメータを y_n ，利用頻度の低い物の影響を反映させるパラメータを z_n とする．上記パラメータはシステムを利用するユーザが指定する．

短期的嗜好情報は，1回以上利用された物の利用率に基づいた評価ベクタとなる．ユーザ A がある時間 a から b までの期間 T_{ab} において，短期的嗜好に基づいた推薦を受ける場合，ユーザ A の利用率の集合を $I_A = \{u_{abA,1}, u_{abA,2}, \dots, u_{abA,n}\}$ とし， I_A に1対1に対応したユーザ空間の集合における，ノイズ除去に関するパラメータを $R_A = \{y_{A,1}, y_{A,2}, \dots, y_{A,n}\}$ とし，利用頻度の低い物の影響を反映させるパラメータを $Q_A = \{z_{A,1}, z_{A,2}, \dots, z_{A,n}\}$ とすると，

期間 T_{ab} における評価ベクタ $q_{abA,i}$ は3.1の3.3より次の行列になる．

$$q_{abA,i} = \log_{z_{A,i}} u_{abA,i} \quad (3.3)$$

但し，

$$u_{abA,i} \leq y_{A,i} \quad (3.4)$$

同様に，長期的嗜好に基づいた推薦を受ける場合，所有率の集合を $J_A = \{o_{abA,1}, o_{abA,2}, \dots, o_{abA,n}\}$ とすると，期間 T_{ab} における評価ベクタ $q_{abA,i}$ は3.5より次の行列になる．

$$q_{abA,i} = \log_{z_{A,i}} u_{abA,i} + o_{abA,i} \quad (3.5)$$

但し，

$$u_{abA,i} \leq y_{A,i} \quad (3.6)$$

例えば，第??節で示したように，物履歴には様々な種類があり，アプリケーション毎に必要な物履歴は異なるため，複数物履歴を統合して生成したユーザプロフィールは他のアプリケーションの要求に合わない可能性がある．

3.7 ユーザプロフィールの更新

3.8 協調フィルタリングによる情報推薦

嗜好の似た他のユーザを見つけるために，他のユーザとの嗜好の相関係数を算出する必要がある．相関係数の算出には，先ほど述べたように物ごとの各ユーザの評価ベクタが必要である．ユーザが情報の推薦を受ける時に，各ユーザごとの評価ベクタを生成するの

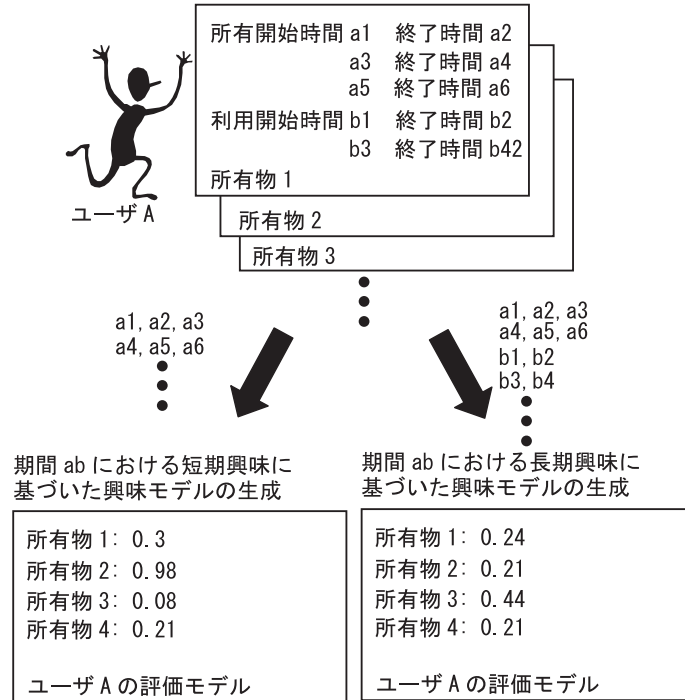


図 3.2: ユーザ A の評価ベクタ生成

は、時間と計算能力のコストが問題になるため、事前に各ユーザ毎ごと評価ベクタを生成しておく必要がある。推薦を受けるユーザは、短期嗜好と長期嗜好に基づいた評価ベクタを生成できるが、他ユーザの推薦の際に利用する評価ベクタについては、事前に生成しておかなくてはならない。過去の物履歴から情報の推薦を受けるといった、ユーザは嗜好情報を抽出する物履歴の期間を指定できるため、他ユーザの推薦に利用される評価ベクタは過去の物に対する評価が含まれているのが望ましい。以上よりシステムは、他ユーザの情報推薦に利用するために、各ユーザの全物履歴から、長期的嗜好に基づいた評価ベクタをユーザごとに生成した評価ベクタを保持する。物履歴に更新のあったユーザのみ、評価ベクタを再生成する。

ユーザの嗜好情報に基づいて情報の推薦を受けるユーザが、短期的嗜好、長期的嗜好を指定した後に、

ある物 i の物履歴におけるユーザ A とユーザ B の間の相関係数を $r_{ab,i}$ とし、先述の興味モデルとピアソンの相関関数から相関係数 r_{ab} は次の 3.7 で決定する。

$$r_{ab,i} =$$

$$\frac{\sum_{i,j} (q_{ab,A,ij} - \overline{q_{ab,A}})(q_{ab,B,ij} - \overline{q_{ab,B}})}{\sqrt{\sum_{i,j} (q_{ab,A,ij} - \overline{q_{ab,A}})^2} \sqrt{\sum_{i,j} (q_{ab,B,ij} - \overline{q_{ab,B}})^2}} \quad (3.7)$$

但し、 $\overline{q_{ab,A}}$ は、ユーザ A の時間 a, b 間の期間における利用状況ベクタの平均を表す。

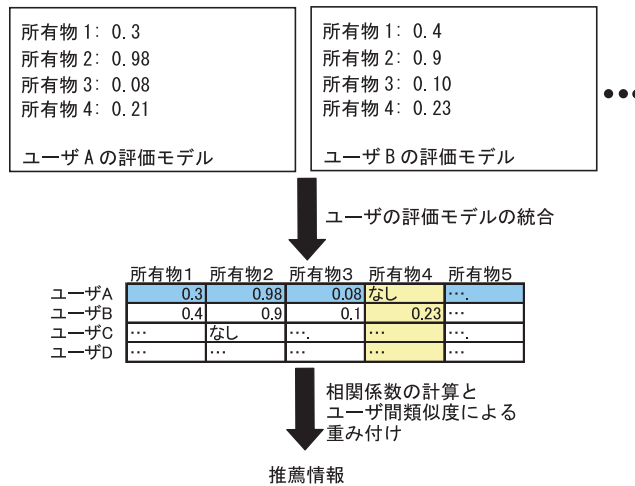


図 3.3: 各ユーザに対する情報の推薦

協調フィルタリングによる情報推薦では、相関係数の絶対値の大きいユーザの意見が反映される。またこのとき、ユーザにとって未知の物の評価は、未知の物を評価しているユーザの評価に相関係数の重みをつけて足し合わせたものである。

3.9 本章のまとめ

本章ではまず、本研究の目的を実現するためのアプローチについて述べた。次に、アプローチで利用する手法についての定義を行い、ユーザの嗜好の分類について述べ、ユーザの嗜好情報の抽出について述べた。次章では、物履歴の具体的な取得方法について述べる。

第4章 設計

本章では，第3章で述べた物履歴に基づいて構築される OHModel を用いてユーザプロフィールを生成するシステムである UPOH の詳細について述べる．最初に，UPOH の全体概要について述べ，システム構成について述べる．次に UPOH を物履歴取得部，物履歴管理部，ユーザプロファイリング部の3つに分けてそれぞれの設計について述べる．

4.1 全体概要

第3章で述べた OHModel を用いたユーザプロファイリングシステムを UPOH と名づける．本節では，最初に UPOH の利用方法について述べ，次に UPOH を3つのモジュールに分けて述べる．

UPOH の利用方法

UPOH は第3章で述べた OHModel を用いたユーザプロファイルを生成，提供するシステムである．そのため，UPOH の利用者は，ユーザプロファイルを利用してサービスを行うアプリケーションとなる．例えば，商品のリコメンデーションを行うアプリケーションが考えられる．

UPOH 利用者は，UPOH サーバにユーザプロファイル送信要求を送ることで OHModel を用いたユーザプロファイルを取得することができる．ユーザの物履歴はユーザが物履歴取得を意識せずに自動的に取得，収集される．

4.1.1 UPOH のモジュール分け

UPOH は物履歴取得部，物履歴管理部，ユーザプロファイリング部に分けられる．UPOH のモジュール分けを図??に示す．

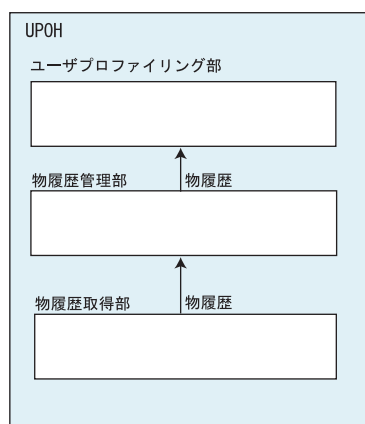


図 4.1: モジュール分け

第??節で述べたように，ユーザプロファイルはアプリケーション依存であるため，UPOH をアプリケーション非依存に構築するためには，アプリケーションにユーザプロファイルを提供するのではなく，物履歴を提供する必要がある．個々のアプリケーションは，必要とする物履歴のみを取得してユーザプロファイルを生成し利用する．そのため，物履歴を扱う部分とユーザプロファイリングはモジュール分けする必要がある．

また，第??節より，物履歴の取得は物履歴取得デバイスによって行われるが，複数の物履歴取得デバイスから取得した物履歴を管理し，物履歴をアプリケーションに提供するためには，物履歴取得部と物履歴管理部はモジュール分けする必要がある．

以上より，UPOHは物履歴取得部，物履歴管理部，ユーザプロファイリング部の3つに分けて構成される．

4.2 ハードウェア構成

UPOHが動作するハードウェア環境を図??に示す．UPOHは物履歴取得デバイス，物履歴管理サーバ，ユーザプロファイル提供サーバで構成される．各デバイス間はネットワークで接続され，物履歴取得デバイスが取得した物履歴は物履歴管理サーバに送信される．また，物履歴管理サーバに保存されている物履歴はユーザプロファイル提供サーバからの要求に従って送信される．次に，各デバイスの機能を述べる．



図 4.2: ハードウェア構成図

- 物履歴取得デバイス

物履歴を取得するデバイス．物の存在を検知し，第??節で述べた物の状態を物履歴として取得する．本研究ではRFID-Readerを用いて物のタグを認識し，物の存在を検知している．また，第??節で述べたように，物取得デバイスではユーザと物とのインタラクションを検知する機能を持つ．例えば，センサやカメラを用いることが考えられる．

- 物履歴管理サーバ
物履歴取得デバイスから送信された物履歴を RDBMS に保存するサーバ。また，ユーザプロフィール提供サーバからの物履歴要求に対し物履歴を送信する。
- ユーザプロフィール提供サーバ
OHModel を用いたユーザプロフィールを利用するアプリケーションからのユーザプロフィール取得要求に回答するサーバ。物履歴管理サーバへ物履歴を要求し，ユーザプロフィールを生成し回答する。

4.3 ソフトウェア構成

本システムのソフトウェア構成図を図 4.3 に示す。第 4.1.1 節より，ソフトウェア構成を物履歴取得部，物履歴管理部，ユーザプロフィール部に分けて設計する。以下は，図 4.1 に示されている各モジュール間のメッセージ番号を元に動作概要を示したものである。

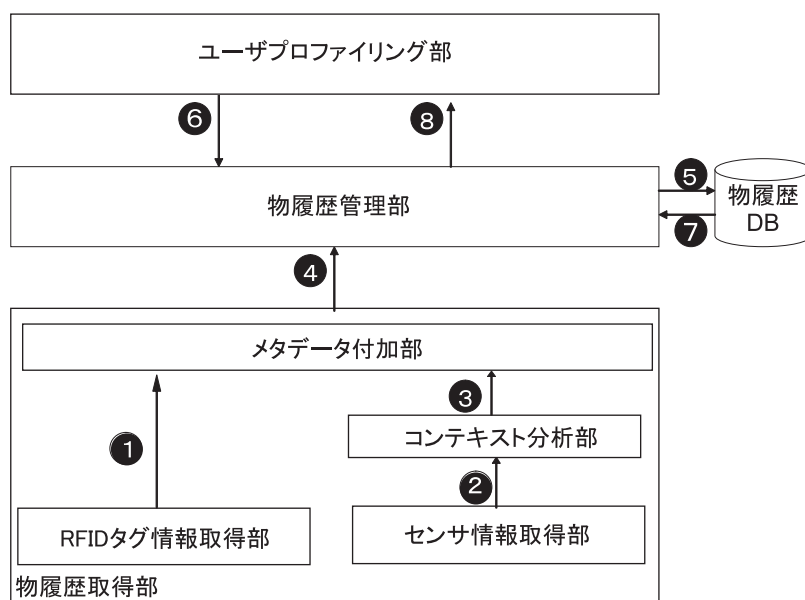


図 4.3: ソフトウェア構成図

1. 物履歴取得部において RFID タグ取得部で取得されたタグ ID をメタデータ付加部へ渡す。
2. 物履歴取得部においてセンサ情報取得部で取得されたセンサ情報をコンテキスト分析部へ渡す

3. 物履歴取得部においてコンテキスト分析部で分析されたコンテキストをメタデータ付加部へ渡す。
4. 物履歴取得部において下位モジュールから渡されたコンテキストを元に、タグIDにコンテキスト情報を付加する。
5. 物履歴取得部はIDタグとセンサから得られたコンテキストメタデータを付加した情報を物履歴管理部に渡す
6. 物履歴管理部は物履歴DBに渡された物履歴を保存する。
7. ユーザプロファイリング部はユーザプロファイルを生成するため、物履歴取得要求を物履歴管理部へ渡す
8. 物履歴管理部は受け取った物履歴取得要求に従って物履歴DBから物履歴を取得しユーザプロファイリング部へ渡す
9. ユーザプロファイリング部は取得した物履歴に基づいてユーザプロファイルを生成

4.4 物履歴取得部

物履歴取得部は、RFIDタグIDを取得し、取得時間から、物がRFIDリーダ周囲に置かれている、または取り出されるイベントを取得する。また、センサが搭載されている場合、センサ情報と統合することで、物やユーザのコンテキストを分析し、物とユーザとのインタラクションを取得する。例えば、加速度センサを利用することで、物が置かれているコンテキストとユーザに携帯されているコンテキストを取得できる。物がユーザに携帯されている場合、物履歴にそのメタ情報を記述することで、ユーザプロファイリング部ではその物とユーザとの関連性が深いと推定する。

4.5 物履歴管理部

物履歴管理部では、物履歴取得部ごとに別のフォーマット、別の保存領域に物履歴を保存する。ユーザプロファイリング部からの物履歴取得の要求に、ユーザプロファイリングが必要とする物履歴だけを簡単に送信することができる。また、ユーザプロファイリング部に送信する物履歴は、その物に関連したIDを付加して送信する。例えば、本の物履歴

を送信する場合，タグ ID に関連した ISBN を付加してユーザプロファイリング部へ渡す．なお，本論文では，ユーザプロファイリング部に物履歴を送信する際にセキュリティを考慮していない．

4.6 ユーザプロファイリング部

ユーザプロファイリングでは，アプリケーションからの要求に応じて取得したい物履歴を指定し，物履歴管理部へ物履歴取得要求を送信する．物履歴の指定は，時間範囲と，物履歴を取得した物履歴取得デバイスの種類を制限することができる．次に，物履歴管理サーバから受信した物履歴に基づいて，OHModel に基づいたユーザプロファイルを作成する．物履歴のフォーマットや含まれる情報は，物履歴取得デバイスごとに違う．そのため，ユーザプロファイルの生成は物履歴取得デバイスごとに違う手法で物履歴を分析し，ユーザプロファイルを生成する．

4.7 本章のまとめ

本章ではUPOHの設計について詳しく述べた．まず，システムの概要とシステム構成について述べ，次にUPOHを構成する個々のモジュールについて述べた．次章では，UPOHの実装について詳しく述べる．

第5章 実装

本章では，UPOHの実装について述べる．最初に，概要と実装環境について述べる．次に，前章と同様に全体を3つに分けてそれぞれの実装について述べる．最後に，サンプルアプリケーションとしてUPOHが生成したユーザプロファイルを用いたお勧め本提案システムについて述べる．

5.1 概要

本節では、UPOHの実装概要を述べる。

最初に、物履歴取得部について述べる。物履歴を取得する物の種類を本に絞る、本にRFIDタグを付けた。本に絞った理由は、ユーザが所持する本はそのユーザの嗜好や興味が反映されるため??、物履歴は本とユーザとの関連性を他のものに比べて良く示していると考えたためである。物履歴取得デバイスが取得する対象として、机の上の本、バックに入っている本、棚に置かれた本の3つに絞った。上記3つの対象を取得する物履歴取得デバイスを、それぞれスマートデスク、スマートバック、スマートシェルフと呼ぶ。詳細については、第??節で述べる。

次に、物履歴管理部について述べる。物履歴管理モジュールでは、物履歴取得部から送られてくる物履歴をMySQL5.0を用いて全てデータベースに保存する。物履歴取得部と物履歴管理部との間で通信する物履歴のデータ形式は、設計で物履歴取得デバイスを限定していないため、データ記述に汎用性、拡張性のあるXMLで通信を行う。同様にユーザプロファイリング部と物履歴部間の通信もXMLで行う。詳細については、第??節で述べる。

次に、ユーザプロファイリング部について述べる。ユーザプロファイリング部では、ユーザプロファイリングクライアントのXMLで送られてくる要求に基づいて、物履歴を物履歴管理サーバに要求して、OHModelを用いてユーザプロファイルを生成する。物履歴を取得した物履歴取得デバイスごとに、物履歴に含まれる情報が違うため、物履歴取得デバイスごとにユーザプロファイル生成アルゴリズムを切り替える。詳細については、第??節で述べる。

最後に、UPOHが提供するOHModelに基づいたUser Profileの利用例として、協調フィルタリングを利用したRecommenderを実装する。詳細については、第??節で述べる。

5.2 実装環境

本節では、UPOHの実装環境について述べる。第??節より、UPOHは物履歴管理サーバ、ユーザプロファイリングサーバと物履歴取得サービスで構成される。実装環境として、1台のWindowsPCで2つのサーバと1つの物履歴取得サービスを実装し、その他にもLinuxPC、WindowsPCそれぞれで物履歴取得モジュールを実装した。そのため、UPOHを動作させるために、計3台のPCを利用した。物履歴管理サーバが動作するPCに、MySQL5.0をインストールした。実装環境を表??に示す。

開発は主にJ2SE(TM) Development Kit 5.0で行った。J2SEで行った理由としては、データベースとの連携を行うために、JDBCライブラリを利用したかったためである。Windows上で動作する物履歴取得部は、Windowsのサービスとして動作させるため、VisualC++で開発した。開発環境について表5.2に示す。

項目	サーバ動作 PC	物履歴取得 PC1	物履歴取得 PC2
CPU	PenM 2.26GHz	PenM 733 1.10GHz	Pen4 3.20GHz
RAM	2GByte	512MByte	2GByte
OS	WinXP Pro S2	WinXP Pro S2	FedoraCore2(Linux2.6.1)
RDBMS ¹	MySQL5.0		

表 5.1: 実装環境

名称	バージョン
J2SE(TM) Development Kit	5.0
VisualC++	7.0

表 5.2: 開発環境

5.3 物履歴取得部

本節では物履歴取得部について詳細に述べる。まず、利用する物履歴取得デバイスについて概要を述べ、次にそれぞれの物履歴取得デバイスごとに物履歴の取得手法を述べる。

第 3.3 節??で述べたように、物履歴の取得範囲はユーザの所有物に関連性が高い物が望ましい。そこで、本研究では、第??節で述べた理由から、物履歴の取得対象をユーザ個人の机上、バック、本棚に絞った。

机上とバックの物履歴を取得するために、スマートふるしき [17] を用いる。スマートふるしきは、8 つの RFID アンテナと 8 つの RFID タグを持つデバイスであり、システムは 8 つの RFID アンテナのどこに RFID タグを置いたのかを検知できる。スマートふるしきは 13.56MHz の RFID タグを読み取る。各アンテナは伝導性の布で出来ており、折り曲げることができるため、布の用に折りたためるだけでなく、折りたたまれた中に挟まれた RFID タグも検知することが出来る。

UPOH では、パッシブ型 RFID の無線通信技術 [18] を使いユーザの物を検出する。パッシブ型 RFID を利用した理由は、安価でほぼ恒久的に作動するためである。ユーザの所有物に RFID タグを付け、RFID リーダを使って読みとることで、タグ付けされた物の物履歴を取得する。

第 5.3 節で机上、第??節で本段、そして第??節でバックを対象とした物履歴の取得手法について述べる。

机上の物履歴取得

机上の物履歴を取得するシステムをスマートデスクと呼ぶ。スマートデスクのデバイスは、スマートふるしきと WebCamera から構成される。スマートふるしきを用いることで、物が机のどこに置かれたかを検知できる。また、

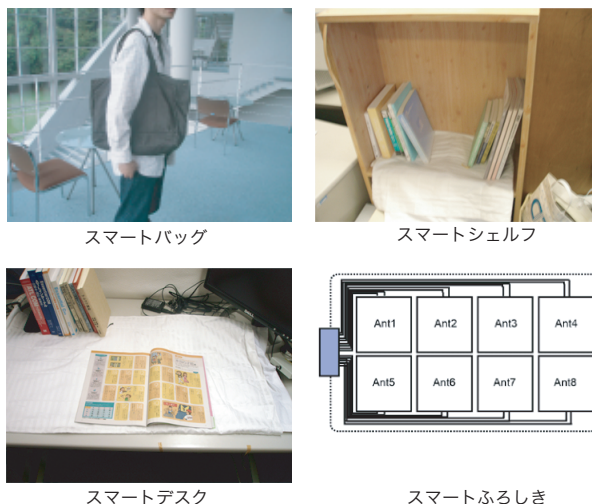


図 5.1: 物利用状況の取得

WebCamera を使い机の上の手の位置を取得し，スマートふるしきの物がどこに置かれているかを検知する機能を統合することで，ユーザが手元においている物を取得することができる．机の上の手の動きを取得するために，Catch Me ミドルウェアを利用する．[19] 本ミドルウェアは，手の肌色領域を認識することで，カメラが捕らえた机の上の手の位置を取得できる．以上より，スマートデスクが取得する物履歴は次の3つ情報を含む

- 物が机の上に置かれた時間/物が机上から取り出された時間
- 物が置かれている机上の位置
- 物が手元に置かれているか否か

5.3.1 バックの物履歴取得

バックの物履歴を取得するシステムをスマートバック??と呼ぶ．スマートバックのデバイスは，スマートふるしきと3軸加速度センサから構成される．スマートふるしきを用いることで，バックの内側にスマートふるしき張り合わせることで，RFIDリーダを持ち歩くことを意識させずに，ユーザのバックの物履歴を取得できる．また，加速度センサをバックに取り付けて利用することで，ユーザがバックを携帯して，移動中であるか否かコンテキストを取得できる．

以上より，スマートバックが取得する物履歴は次の2つの情報を含む

- 物がバックに入れられた時間/物がバックから取り出された時間
- ユーザが移動中か否か



図 5.2: スマートバック

5.3.2 本棚の物履歴取得

本棚の物履歴を取得するシステムをスマートシェルフと呼ぶ。スマートシェルフのデバイスは、RFID-Readerで構成される。本棚の後ろにアンテナを設置することで、棚の中に入った物履歴を取得できる。

以上より，スマートバックが取得する物履歴は次の1つ情報を含む

- 物が本棚に置かれた時間/物が本棚から取り出された時間

5.3.3 物履歴取得部のまとめ

本節では物履歴の取得手法について詳細に述べた。まず，利用するデバイスについて概要を述べ，次にそれぞれのデバイスごとに物履歴の取得方法を述べた。

5.4 物履歴管理部

物履歴管理部は物履歴取得部，ユーザプロファイリング部からのXML形式の要求に答える．物履歴管理部は，物履歴取得部から送られてきたソースデバイスごとに1ヶ月間隔でテーブルを作成し，物履歴を保存していく．XML形式のフォーマット例を次に示す．

物履歴取得部からのXML

```
<xml>
  <objecthistory>
    <msg-num>1</msg-num>
    <sender-typeid> smart-shelf</sender-typeid>
    <msg-type>insert</msg-type>
    <timestamp>2007-01-15 22:12:00.11</timestamp>
    <tag-id>e007000023efa549</tag-id>
    <option>
      <in-out>in</in-out>
    </option>
  </ objecthistory >
</xml>
```

図 5.3: 物履歴取得部からのXML

ユーザプロファイリング部からのXML

5.5 ユーザプロファイリング部

5.5.1 OHModelを用いたユーザプロファイル生成アルゴリズム

第??節で述べたように，物履歴の取得元となる物履歴取得デバイスごとに，ユーザプロファイル生成アルゴリズムが利用できる情報は異なるため，

```
<xml>
  <recommend>
    <method>GET</method>
    <option>
      <duration>
        <hour>1</hour>
      </duration>
    </option>
  </recommend>
</xml>
```

図 5.4: ユーザプロファイリング部からの XML

スマートデスク，スマートバック，スマートシェルフそれぞれについて個別にアルゴリズムを定義する必要がある．

スマートシェルフ

本棚に置かれている時間長をレーティングの基準とする．本棚から取り出した回数を重みとして利用する．ユーザプロファイルを生成は，本棚に置かれている時間長を基準化して，それに本棚から取り出した回数を重み付けした値を，その本のユーザとの関連性を表す数値とする．

5.6 サンプルアプリケーション

本節では，ユーザプロファイルの利用アプリケーション例として，協調フィルタリングを用いた Recommendation Service を実装した．協調フィルタリングエンジンとして，Taste 協調フィルタリングエンジン [20] を利用した．

5.7 本章のまとめ

本章ではUPOHの実装について述べた。まず、概要と実装環境について述べ、次にUPOHを構成する個々のモジュールの実装について述べた。最後に、UPOHが生成するユーザプロファイルの利用アプリケーション例として、協調フィルタリングを用いたお勧め本提案システムについて述べた。次章では、UPOHの評価を行う。

第6章 UPOHの評価

本章では、ユーザと物との関連性を表す、OHModelと、OHModelを用いたユーザプロファイルを生成するシステム、UPOHを評価する。最初に双方の定量的評価を行い、次にUPOHの定性的評価を行う。

6.1 定量評価

本節では、物履歴に基づいたユーザと物との関連性モデルである OHModel と、ユーザプロファイル生成システム、UPOH の定量評価を行う。

まず、物履歴を取得するデバイスとして SmartShelf を選択し、SmartShelf で取得した物履歴に基づいた OHModel の評価を行う。次に、UPOH が生成した User Profile の利用例として、UPOH が提供する協調フィルタリングによるリコメンデーションについて評価する。さらに、OHModel のパフォーマンス評価を行う。

6.1.1 OHModel によるレーティングの評価

OHModel では、ユーザと物との関連性が高いと思われる物ほど、高い値にレーティングされる。本節では、SmartShelf によって構築された OHModel とユーザの直接評価付けとの比較を行い、OHModel の妥当性を検証する。ユーザの直接評価付けは、OHModel がレーティングした物全てに対して 1 から 5 段階で評価付けを行う。実験結果を表 6.1 に示す。

本 ID	A	B	C	D	E	F	G	H	I	J
OHModel のレーティング	2.24	2.24	2.52	2.44	3.87	2.34	2.18	2.49	2.67	2.18
ユーザの直接評価付け	3	4	4	2	3	2	3	1	2	2
本棚から取り出した回数	7	7	8	2	2	1	3	3	3	2

表 6.1: 本 10 冊の評価 評価尺度 (1~5)

表 6.1 より、OHModel のレーティングを見ると各アイテム間のレーティング値に、本 E 以外明確な差がない。本 E の値について、本 E は実験のため 1 冊だけ先に置かれていたため、他の本よりも設置時間が長いことが高いレーティングになったと考えられる。OHModel ではより長い期間置かれていた物に高いレーティング値をつける。次に B、C のレーティング値について、直接評価付けでは 4 という高いレーティング値をつけたが、OHModel のレーティングにはあまり反映されていない。B、C の本棚から取り出した回数は、他の本に比べて明らかに大きい値となっており、この値が大きいほど重みをつけているが、うまく結果が反映されていない。本棚はそれ

ほど頻繁に本を取り出すことがないため，本棚から取り出すことによる重みをより反映しやすくさせる必要があると考えられる．

6.1.2 OHModel が利用したユーザインタラクションの評価

OHModel では物履歴の取得デバイス毎に，ユーザと物とのインタラクションを利用して物にレーティングを行う．

SmartShelf を使って取得した物履歴の場合，物のレーティングは本棚から本を取り出す回数を利用して行われる．利用するインタラクションの妥当性を検証するために，インタラクションを利用せずに OHModel を構築した場合と，前節 6.1.1 で示した，インタラクションを利用した場合の間で比較する．比較方法は，前節 6.1.1 と同じ手法で行う．実験結果を表 6.2 に示す．

本 ID	A	B	C	D	E	F	G	H	I	J
OHModel	2.24	2.24	2.09	3.86	4.48	2.62	2.09	2.80	2.09	2.26
ユーザ	3	4	4	2	3	2	3	1	2	2

表 6.2: 本 10 冊の評価 評価尺度 (1~5)

表 6.2 の結果は，前節の表 6.1 よりも明らかにユーザの意図を反映していない．特にユーザが高いレーティングをつけた B, C の値が低い結果になってしまっている．これより本棚から取り出す回数で重み付けすることは有効に機能することがわかる．

6.1.3 リコメンデーションの評価

システムの有用性を検証するために実装した Recommender サービス、UPOH Recommender が提供する推薦情報について，前節 6.1.1 と同様に直接評価付けして比較する．

本 ID	A	B	C	D	E	F	G	H	I	J
UPOH Recommender	5	5	5	5	5	4	4	3	1	0
ユーザ	4	3	4	4	3	3	3	2	2	2

表 6.3: 本 10 冊の評価 評価尺度 (1~5)

表??の結果は，第6-evaluation-non-interaction節や第6.1節の結果と比較すると，ユーザの評価が良い本が推薦された．これは，第6.1節から分より，アイテム間のレーティングに差がなく，最も本数の多いジャンルが自分の好きなジャンルであった結果である．このことから，アイテム間のレーティングに差が出ない場合，最も冊数の多いタイプの本が推薦されることが分かった．

6.1.4 UPOHのパフォーマンス評価

最初に，User Profileを生成するまでのシステムが要する時間について評価し，次に，物履歴が要求するストレージの容量について評価する．実験環境を表6.4に示す．

項目	仕様
CPU	Intel Pentium M processor 2.26GHz
メモリ	2GB DDR2 SDRAM
ハードディスク回転速度	100GB (7,200rpm / 9.5mm)
OS	WindowsXP Professional Service Pack 2
ネットワーク	1000BASE-T

表 6.4: 実験環境

User Profileの生成

User Profileの生成は，ユーザがユーザプロファイル生成要求を行う度に，物履歴を物履歴データベースから取得し，OHModelを構築する必要がある．そのため，物履歴の増加に従って，データベース検索と物履歴をネットワーク転送時間が増加する．そこで，User Profileの生成時間と生成に利用する物履歴数との関係性を評価する．実験結果をグラフ6.1に示す．

グラフ6.1から，UPOHは物履歴が増加するほど，ユーザプロファイル生成時間が増加することが分かった．今回のUPOH実装では，物履歴のキャッシュを行っていないが，ユーザプロファイルを更新する場合，物履歴をキャッシュしておき，前回のユーザプロファイル生成から更新された物履歴のみを物履歴データベースに要求する必要がある．

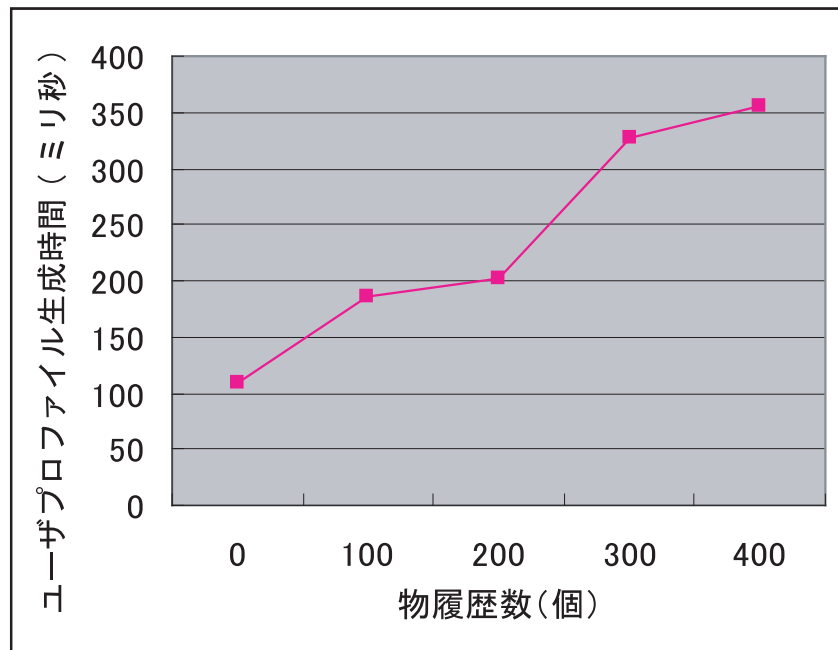


図 6.1: 物履歴数に対するプロフィール生成時間の変化

ストレージ容量

UPOHでは、利用する物履歴取得デバイスごとにデータベースに物履歴を保存する。そのため、システムの利用時間と利用する物履歴取得デバイスが増えるごとに、データベースに保存する容量が増加する。本システムは、ユーザの長期間に渡ってユーザの物履歴を取得することを想定している。そのため、利用時間とストレージ容量との関係性を評価する。結果を表 6.2 に表す。

1つの物履歴辺り最もデータ容量の小さいSmartShelfのみを使った場合、グラフ??より、1日に500byteほど平均で増えることがわかった。1年運用して、5Mなので現在のGByte単位のハードディスク容量では、物履歴ストレージ量が問題になることはない。

6.2 定性評価

本節では、UPOHと既存の関連研究を比較する。既存の人間の自然な行動からユーザプロフィールを生成する手法である、textextractor[9]と、Infor-

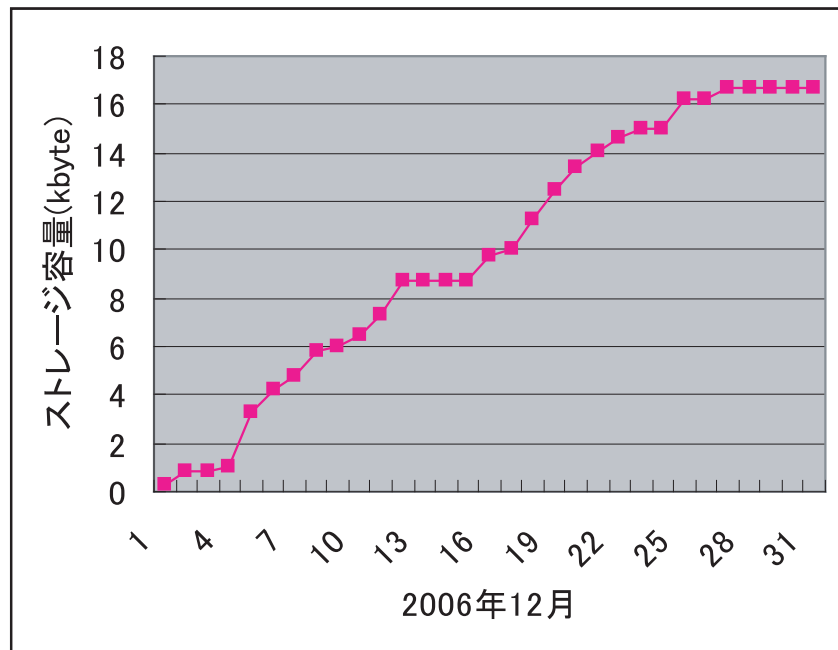


図 6.2: SmartShelf を使った取得した物履歴数に対するストレージ容量の変化

mation Recommendation by collaborative filtering incorporated with gaze detection (本節では, Gaze Detection と呼ぶ) [11] とを比較する. 比較結果を 6.5 に示す.

項目	実世界への対応	嗜好獲得の範囲	ユーザ負担	実現性
textextractor	×			
Gaze Detection				
UPOH				

表 6.5: 人間の自然な行動からユーザプロフィールを生成する手法の比較

表 6.5 の評価項目について説明する. 嗜好獲得の範囲とは, ユーザプロフィールを生成する元となる, ユーザの嗜好に関連した情報を獲得できる範囲のことである. UPOH は, 実世界上の物に RFID tag が付いた全ての物が嗜好獲得の対象となるため, 広い範囲で嗜好獲得ができる. また, 評価項目のユーザ負担は, ユーザに負担をかけない方が となる. textextractor[11] は, マウス操作に基づくため PC 上で閲覧できる情報に限られるため実世界への対応が × になっている. Gaze Detection は人間は興味, 関心のある場所に注目を行うことから, 絵画の部分領域へ注目時間が長いほど, 興味

の度合いが高いと考える．興味度合いを測る対象を評価する前に明示的に，評価を開始するトリガをシステムに与える必要がある．絵画以外にも実世界でユーザがするには，ユーザの視界に移った物を自動で判別し，注目時間から興味度合いを算出する必要があるが，このような手法は難しい．そのため，実世界への対応，嗜好獲得の範囲どちらも となっている．本研究は，実世界への対応が可能であり，嗜好獲得の範囲の広いシステムとして既存研究より優位である．

6.3 本章のまとめ

本章では，UPOH の評価を行った．定量評価として，UPOH が用いる OHModel の妥当性について評価を行った．次に，User Profile の利用例として，UPOH が提供する協調フィルタリングによるリコメンデーションについて評価を行った．また，UPOH の実行速度とストレージ容量についてパフォーマンス評価を行った．最後に，定性評価として既存の関連研究と性能面で比較した．

次章では，今後の展望を述べ，最後に本論文のまとめを述べる．

第7章 結論

本研究では，ユーザの物利用状況を基に，ユーザの嗜好情報を抽出するミドルウェアを構築し，有効に働くかを検証した．本章では今後の展望を挙げ，最後に本論文をまとめる．

7.1 今後の展望

本節では、本研究のこれからの展望について述べる。

本アルゴリズムは、個人的空間におけるユーザの物利用状況を単位時間当たりにおける物の所有率と利用率として表し、物利用状況をユーザの興味モデルとした。また、システムは物利用状態を取得するために、ユーザが物を利用する動作や所有する動作の履歴を保存し、短期嗜好と長期嗜好に分類し、目的によって興味モデルを自分の目的によって変えた。前者は、ユーザの作業に近いものが推薦され、後者はユーザの個性を重視した。これらの値が本当にユーザの嗜好に合うのかユーザ評価実験を基に統計的に有効性を示す。また、今回の実装では本のみを対象にした。ユーザの興味モデルを利用した協調フィルタリングは、どのような種類の物も透過的に扱うことができる。しかし、どのように物を扱うかは物の種類に大きく依存するため、それぞれの物ごとにその物の扱い方を定義する必要がある。今後はユーザの典型的な日常品に対して、その種類ごとに適切な利用状況算出手法を定義していく予定である。

7.2 まとめ

本論文では、個人的空間にある所有物に対するユーザの利用状態を取得することで、実世界の日常動作からユーザの嗜好情報を抽出するアルゴリズムについて述べた。また、物の利用状態を実際に取得するために、布型RFIDであるスマートふるしきを利用し、机の上、バッグ、本棚についてパッシブ型RFIDタグを利用して、所有物の利用状態の取得を行った。プロトタイプ実装としては本に絞って物利用状態を取得するシステムを実装した。

謝辞

本研究の機会を与えてくださり，ご指導を賜りました慶応義塾大学環境情報学部教授徳田英幸博士に深く感謝いたします．また，重要な御助言を頂きました，慶應義塾大学政策・メディア研究科助教授高汐一紀博士の御指導に深く感謝致します慶応義塾大学徳田・村井・楠本・中村・南合同研究会の先輩方には折りにふれ貴重な指導と助言を頂きました．特に，徳田研究室の先生方や先輩方，ACE(Active Computing Environmets) 研究グループの方々に深く感謝いたします．また，駒木亮伯氏，大澤亮氏，青木崇行氏には絶えざる励ましや丁寧なご指導を賜りました．最後に，研究の日々を共に過ごした，今枝卓也氏，その他多くの友人に深く感謝し，謝辞と致します．

平成 19 年 1 月 24 日

鈴木 慧

参考文献

- [1] Steve Lawrence and C. Lee Giles. Accessibility of information on the web. *Intelligence*, Vol. 11, No. 1, pp. 32–39, 2000.
- [2] A. Gulli and A. Signorini. The indexable web is more than 11.5 billion pages. *Proceedings of the 14th International WWW Conference*, 2005.
- [3] Inc Amazon.com. amazon.com. <http://www.amazon.com/>.
- [4] Google. Google news. <http://news.google.co.jp/>.
- [5] Auto-id lab. japan. <http://www.autoidlab.jp/>.
- [6] T. Yan and H. Garcia-Molina. Sift a tool for wide-area information dissemination. *USENIX Technical Conference*, pp. 177–186, 1995.
- [7] C.E. Osgood, G.J. Suci, and P.H. Tannenbaum. The measurement of meaning. *University of Illinois Press*, 1957.
- [8] Ken Lang. News-weeder: learning to filter netnews. *Proceedings of the 12th International Conference on Machine Learning*, pp. 331–339, 1995.
- [9] 土方嘉徳, 青木義則, 古井陽之助, 中島周. Textextractor : web ブラウザ上での操作履歴を用いたテキスト部分抽出. 第8回インタラクティブシステムとソフトウェアに関するワークショップ (WISS2000), pp. 201–206, 2000.
- [10] 中里祐介, 柴田史久, 馬場口登. 歩行者ナビゲーションにおけるユーザの行動履歴からのプロフィール生成手法. 情報処理学会第65回全国大会講演論文集, Vol. 3, pp. 407–408, 2003.

- [11] 脇山孝貴, 吉高淳夫, 平嶋宗. 注視を利用した協調フィルタリングによる興味のある情報の推薦. *WISS2005*, pp. 81–86, 2005.
- [12] T. Kamba, K. Bharat, and M. Albers. An interactive, personalized, newspaper on the www. *4th Intl. World Wide Web Conference*, 1995.
- [13] P. Resnick, N. Iacovou, M. Suchak, P. Bergstorm, and J. Riedl. GroupLens: An open architecture for collaborative filtering of netnews. *Proceedings of ACM 1994 Conference on Computer Supported Cooperative Work*, Vol. 07, No. 1, pp. 175–186, 1994.
- [14] 大杉直樹, 門田暁人, 森崎修司, 松本健一. 協調フィルタリングに基づくソフトウェア機能推薦システム. *情報処理学会論文誌*, Vol. 45, No. 1, pp. 267–278, 2004.
- [15] Greg Linden, Brent Smith, and Jeremy York. Amazon.com recommendations: Item-to-item collaborative filtering. *IEEE Internet Computing*, Vol. 07, No. 1, pp. 76–80, 2003.
- [16] Russell W Belk. Possessions and the extended self. *Journal of Consumer Research*, Vol. 15, pp. 68–139, 1988.
- [17] Ryo Ohsawa, Masayuki Iwai, Takuya Imaeda, Kei Suzuki, Takuro Yonezawa, Kazunori Takashio, and Hideyuki Tokuda. Smartfuroshiki: A sensorized fabrics supporting office activities. *Ubicomp*, 2006.
- [18] Roy Want. An introduction to rfid technology. *Pervasive Computing, IEEE*, Vol. 5, No. 1, 2006.
- [19] 駒木亮伯, 岩井将行, 神武直彦, 高汐一紀, 徳田英幸. Face-connect:顔の動きに基づいたサービス推薦・選択手法. *WISS*, 2005.
- [20] Sean Owen. Taste collaborative filtering for java. <http://taste.sourceforge.net/>.