

状況依存型ユーザ嗜好モデリングに基づく Context-Aware 情報推薦システム

奥 健 太[†] 中 島 伸 介[†]
宮 崎 純[†] 植 村 俊 亮^{††}

本論文では、ユーザの状況に応じて適切な情報を提供する状況依存型情報推薦システムのプロトタイプを提案する。膨大な情報からユーザの嗜好に合致する情報を提供する手法として、情報推薦システムに関する研究が行われているが、ユーザのそのときの状況（時間帯や天気、同伴者、予算など）に応じて変化するユーザの嗜好に対し、柔軟に対応することは容易ではない。そこで我々は、状況に応じて変化するユーザの嗜好を適切にモデル化する手法を提案した。本論文では、このモデル化手法を適用した状況依存型情報推薦システムのプロトタイプを提案し、検証実験に基づいて、提案手法の評価を行った。この中で、提案手法であるコンテキスト依存型情報フィルタリングとコンテキスト依存型協調フィルタリングの有効性や特長の違いを明らかにするとともに、対象コンテンツの特徴パラメータの最適化に関して考察した。

Context-Aware Recommendation System Based on Context-Dependent User Preference Modeling

KENTA OKU,[†] SHINSUKE NAKAJIMA,[†] JUN MIYAZAKI[†]
and SHUNSUKE UEMURA^{††}

This paper proposes a prototype of a context-aware recommendation system which provides users with contents appropriate to their contexts. There have been many studies of recommendation systems which provide users with contents suited to their preferences. However, it is not easy to adapt these recommendations to changing contexts (e.g., time of day, weather, companions and budget). Thus, we propose a method which takes into account users' preferences as well as their current context. This paper also introduces a prototype of a context-aware recommendation system using the method, and evaluates our proposed methods. In our experiments, we explore the strengths and weaknesses both Context-Aware Information Filtering and Context-Aware Collaborative Filtering. Lastly, we discuss optimization of the feature parameters of contents.

1. はじめに

Webの発達とともに、ユーザがアクセスできる情報が膨大になり、この膨大な情報の中から適切な情報を獲得することを目的とした情報推薦手法に関する研究が行われてきた。従来の情報推薦手法のうち、最も基本的なものの1つは、ユーザの嗜好に基づいたユーザプロファイルを作成し、これに適合する情報をユーザに対して推薦するものである。

しかしながら、ユーザの嗜好は状況によって大きく変化するものであり、これを適切に学習し、ユーザプロファイルとして表現することは容易ではない。たとえば、焼肉が好きなユーザがいたとしても、朝から焼肉店を推薦することは、必ずしも適切ではない。また、オープンカフェに行くことが多いユーザがいたとしても、雨の日にオープンカフェを推薦するのは適切だとはいい難い。したがって、ユーザの状況（以降、コンテキストとよぶ）を考慮した情報推薦を行うことは重要であるといえるが、コンテキストに応じて変化するユーザの嗜好をどのように学習するのかということが大きな問題である。

最も単純な解としては、考えられるコンテキスト別に学習データを用意し、これに基づいたユーザプロファイルをコンテキスト別に作成することで、各コン

[†] 奈良先端科学技術大学院大学情報科学研究科
Graduate School of Information Science, Nara Institute
of Science and Technology

^{††} 奈良産業大学情報学部情報学科
Department of Informatics, Faculty of Informatics,
Nara Sangyo University

テキストに応じた推薦を行う方法が考えられる。しかしながら、ユーザの各コンテキストに応じた細やかな情報推薦を行うためには、想定するコンテキストを細かく設定する必要があるが、各コンテキストに対応する十分な学習データを確保することは容易ではない。また、想定していないコンテキストにおいては、適切な情報推薦を行うことはできない。

このような問題に対して、我々はこれまでに、コンテキストに応じて変化するユーザの嗜好に対し、これをモデル化する手法を提案してきた^{1),2)}。この手法は、学習データ作成の際に、“対象コンテンツそのものに対する満足度として“満足”もしくは“不満足”という評価を行う”のではなく、あるコンテキスト s における対象コンテンツ A (A_s と表現)と、コンテキスト t における対象コンテンツ A (A_t と表現)はまったく別物として扱う。さらに、これらに対する満足度として“満足”もしくは“不満足”という評価を個別に行う。そして、対象コンテンツの特徴だけでなく、コンテキストに関する特徴次元を追加することで拡張した、多次元特徴空間を想定し、学習済みコンテキスト付きコンテンツ (A_s や A_t など)を配置する。さらに各コンテキスト付きコンテンツに対する満足度 (“満足”・“不満足”)に基づいて、学習するものである。

このコンテキスト付きコンテンツの満足度に基づいて、ユーザの嗜好を学習する手法としては、Support Vector Machine (SVM)³⁾を採用した。SVMは多次元空間に配置された正例および負例を、超平面と呼ばれる平面によって分離することで学習する手法である。このSVMの超平面は、データが決まれば一意に決定されるため、学習データが比較的少ない場合に対しても、その汎化性能が高いとされている。したがって、コンテキストに関する特徴次元を含んだ多次元特徴空間に対し、SVMによって学習することにより、1人のユーザのコンテキスト依存の嗜好を1つの学習モデルで表現することが可能となる。

我々は、コンテキスト依存型情報推薦システムの開発を目標としており、本研究では、この提案手法を適用した情報推薦システムのプロトタイプを実装する。この情報推薦システムは以下の2つの情報推薦方式からなる。

- (a) コンテキスト依存型情報フィルタリング (Context-Aware Information Filtering; C-IF)
- (b) コンテキスト依存型協調フィルタリング (Context-Aware Collaborative Filtering; C-CF)

本研究では、検証実験に基づき、この情報推薦システムの評価もあわせて行う。本論文の成果の概要は以

下のとおりである。

- 主なクラス分類手法との比較に基づく、SVMの性能評価を行った。
- より効果的な情報推薦を実現するために、2つの情報推薦方式(上記の(a)および(b))の有効性や利用場面に応じた特長の違いを分析した。
- コンテンツおよびコンテキストの特徴次元の最適化手法について検討した。

以下、本論文の構成を示す。2章では、情報推薦手法およびコンテキスト依存型アプリケーションの関連研究を取り上げる。3章では、本研究において扱うコンテキストを定義したうえで、その定義に則ったコンテキストの例を示す。4章では、提案手法である、コンテキストに応じて変化するユーザの嗜好のモデル化手法について説明し、5章では、この手法を適用したコンテキスト依存型情報推薦システムのプロトタイプについて説明する。6章、7章および8章では、検証実験に基づき、この情報推薦システムの評価を行う。最後に、9章で本論文をまとめる。

2. 関連研究

2.1 情報推薦手法

情報推薦手法を実現するものとして、

- (a) コンテンツに基づくフィルタリング
 - (b) 協調フィルタリング
- がある⁴⁾。

コンテンツに基づくフィルタリングは、対象コンテンツの内容に基づき、ユーザに推薦する情報を選択する手法である。

岩濱ら⁵⁾は決定木を用いた音楽情報フィルタリングシステムを提案している。このシステムでは、提示された音楽データに対して、“好き”・“嫌い”・“どちらでもない”の3段階でユーザが評価を行い、決定木を作成する。決定木のノードは曲の特徴量についての条件を持ち、葉ノードは“好き”・“嫌い”・“どちらでもない”というクラスを持つ。ここで、システムは“好き”というクラスの葉ノードに達した音楽データを推薦する。

杉山ら⁶⁾は、ユーザの嗜好の経時的変化を考慮しながら、構築したユーザプロファイルに基づいたWeb情報検索手法を提案している。長期間のユーザの嗜好と1日限りのユーザの嗜好との両方を考慮しつつ、閲覧ページにおける単語の出現頻度およびページに対する閲覧時間に基づいてユーザプロファイルを作成する。協調フィルタリングは、ネットワーク上において、類似する嗜好を持つユーザを発見し、その類似ユーザ

が興味を持つ情報を推薦する手法である。たとえば、「ユーザ A がコンテンツ X を好むとすると、コンテンツ X を好む別のユーザ B が好むコンテンツ Y を探し出し、ユーザ A もコンテンツ Y を好むだろう」という推論に基づく推薦手法である。協調フィルタリングは、1 度、人の価値判断に基づいて“良い”と評価されたコンテンツを推薦するため、推薦情報に対する信頼度が高く、Web 上で広く成功を収めているシステムであるといえる。

GroupLens^{7),8)} は、協調フィルタリングによりネットニュースのフィルタリングを行うシステムである。このシステムは、嗜好が類似するユーザの評価値から、対象ユーザのニュース記事に対する評価値を推測し、ニュース記事を推薦するというものである。このシステムは、映画の推薦システムである MovieLens⁹⁾ にも応用されている。

Amazon.com¹⁰⁾ では、ブラウザ上で対象ユーザが商品情報を閲覧すると、「この商品を買った人はこんな商品も買っています」というように、他のユーザの購入履歴に基づいて対象ユーザに商品を推薦する。

このように、ユーザの嗜好に合った情報を提供する情報推薦手法に関する研究は多く行われているが、従来手法では、ユーザのその時々コンテキストに応じた嗜好の変化には対応できない。

2.2 Context-Aware アプリケーション

ユーザのコンテキストに依存したアプリケーションに関して、Context-Aware アプリケーションに関連した研究¹¹⁾ が多く行われている。

COMPASS¹²⁾ では、旅行者に周辺の施設情報やサービスを提供する。Shopping Assistant¹³⁾ では、店内の買物客の位置情報に基づいて、商品情報を提示する。Cyberguide¹⁴⁾ では、現在時刻や旅行者の位置情報に基づいて、観光ガイドなどを提示する。

ただし、ここにあげたように、先行研究では時間帯やユーザの位置情報に基づくものが多く、天気や同伴者、予算、ユーザの目的、気分など、ユーザの嗜好に直接影響を及ぼすようなコンテキストを考慮したアプリケーションはあまり見当たらない。

3. 人の意思決定に影響を及ぼすコンテキスト

コンテキストとは、一般にユーザを取り巻くさまざまな状況を意味するが、我々は独自に、コンテキストを「人の情報選択の判断に対する意思決定に影響を及ぼす状況」と定義する。この定義に則ったコンテキストとして、たとえば以下に示すものがあげられる。

- 時間情報：季節、曜日、時刻、...

- 気象情報：天気、気温、湿度、...
- ユーザ情報：休日であるか、予算、余裕時間、目的、気分、体調、...
- 同伴者情報：関係（家族や恋人、友人、上司・部下など）、人数、...
- 周辺環境：現在地のエリアタイプ（歓楽街や郊外、観光地など）、周辺店舗の分布状況、混雑状況、...
- 時系列的要因：過去の行動履歴、将来の行動予定、...

ここにあげたコンテキストのうち、本研究では、時間情報、気象情報、ユーザ情報、同伴者情報を考慮した情報推薦システムについて検討する。これに加え、周辺環境や時系列的要因もユーザの意思決定に影響を及ぼすものであるが、これらを考慮した情報推薦システムについては別途検討課題とする。

4. コンテキスト依存型ユーザ嗜好モデリング (CP モデリング)

本章では、これまでに我々が提案してきたコンテキスト依存型ユーザ嗜好モデリング手法^{1),2)} について説明する。4.1 節では、単純なモデリング方式による問題点を指摘したうえで、それを解決する本提案手法のアイデアを説明する。4.2 節では、次章で提案するコンテキスト依存型協調フィルタリングを実装するための事前準備として、ユーザのコンテキストを考慮したユーザ類似度の算出方法について説明する。

4.1 ユーザ嗜好モデリングの手順

まず、SVM を用いた場合の、単純なユーザ嗜好モデリングの手順を以下に示す（図 1 (i) 参照）。

- (1) コンテンツデータを p 次元の特徴パラメータ（ジャンルや雰囲気、平均予算など）の形式で表現する。
- (2) ユーザはコンテンツデータ 1 つ 1 つに対して好みの評価（“満足”・“不満足”）を与える。
- (3) 評価付けされたコンテンツデータ集合を学習データ集合として p 次元特徴空間上に配置する。
- (4) 学習データ集合を満足クラスと不満足クラスとに分類する識別平面を構成する。

上記手順により、コンテンツデータに対するユーザの嗜好をモデル化することができるように見える。しかし、人の嗜好は、その人のおかれているコンテキストにより変化するものである。たとえば、「同じ焼肉屋であっても、夜に友だちと行くときには満足だが、朝に 1 人で行くときには不満足」といったように、同じコンテンツであるにもかかわらず、コンテキストごとに評価が異なるということが実際にはありうる（図 1 (ii)）。

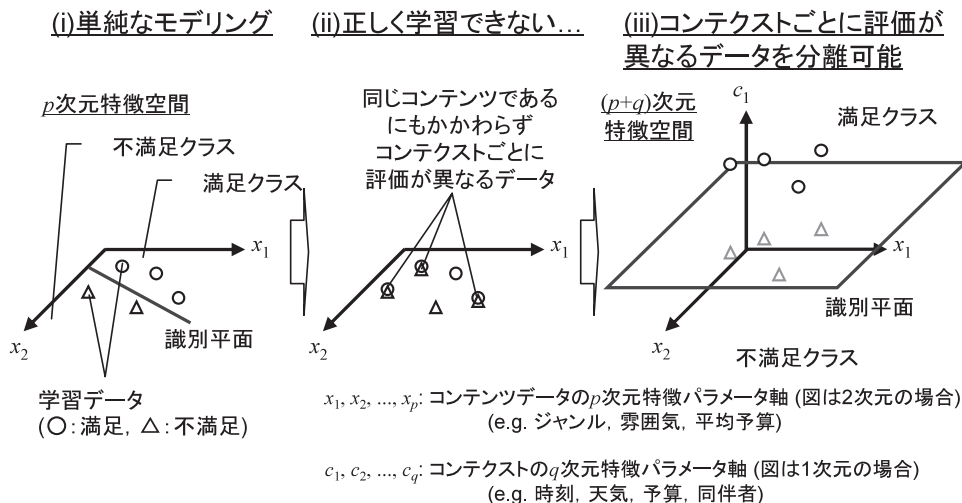


図 1 コンテキスト依存型ユーザ嗜好モデリング (CP モデリング)
 Fig. 1 Context-Dependent User Preference Modeling (CP-Modeling).

このように、その時々コンテキストに応じて多様に
 変化するような人の嗜好をモデル化することは容易で
 はなく、上記手順ではこれを適切にモデル化すること
 はできない。

この問題に対して、我々はデータの表現方法に着目
 する。従来方法では、データの特徴量を表現するも
 のとしてコンテンツの特徴を表すパラメータだけしか
 扱わなかった。これに対し、本研究では、このデータ
 の特徴量表現に対してコンテキストの特徴を表すパラ
 メータも含めて統合する。このようにデータの特徴量
 としてコンテキストの特徴パラメータも含めること
 により、たとえば「夜に友だちと行くときの焼肉屋」と
 「朝に1人で行くときの焼肉屋」とを異なるデータと
 して扱うことができる。これはつまり「同じコンテン
 ツデータであっても、評価時のコンテキストが変われ
 ば異なるデータである」という考えに基づくものであ
 る。これに合わせて、用いる SVM の特徴空間も、コ
 ンテキストの特徴パラメータの次元数だけ拡張する。

その具体的なモデリングの手順を以下に示す
 (図 1(iii) 参照)。

- (1) コンテンツデータを p 次元の特徴パラメータ、
 コンテキストを q 次元の特徴パラメータ (時刻
 や天気、予算、同伴者など) の形式で、それぞ
 れ表現する。
- (2) ユーザはコンテンツデータに対して1つ1つ、
 そのときのコンテキストにおける好みの評価 (“
 満足”・“不満足”) を与える。
- (3) 評価付けされたコンテンツデータの特徴パラ
 メータとその評価時のコンテキストの特徴パラ

メータとを統合したデータ集合を、学習データ
 集合として $(p+q)$ 次元特徴空間上に配置する。

- (4) 学習データ集合を満足クラスと不満足クラスと
 に分類する識別平面を構成する。

我々はこのモデリング手法を、コンテキスト依存型
 ユーザ嗜好モデリング手法 (Context-Dependent User
 Preference Modeling, 以降 CP モデリングと称する)
 とよぶ。CP モデリングにより、コンテキストに応じ
 て変化するユーザ嗜好が適切に学習されたユーザ嗜好
 モデル (以降、CP モデルと称する) を構築することが
 できる。CP モデルの特長は、さまざまなコンテンツ
 の特徴およびコンテキストの特徴に対する複雑なユー
 ザ嗜好をただ1つのシンプルな特徴空間で表現してい
 るところにある。

4.2 ユーザ嗜好モデルに基づくユーザ類似度の算 出手順

従来の協調フィルタリングでは、ユーザ同士の嗜好
 の類似度を算出するための主な方法として、同じコン
 テンツに対して与えた評価値の相関係数に基づく方
 法がある。たとえばユーザ A がカフェ X を好むとし、
 ある別のユーザ B もそのカフェ X を好むとしたとき、
 ユーザ A とユーザ B が類似ユーザとして扱われる。
 しかし、前節で述べたことと同様、同じコンテンツに
 対しても、そのときのコンテキストにより評価が異な
 る。たとえば、ユーザ A が「朝に1人で行くときは、
 カフェ X は満足」としているのに対し、ユーザ B
 が「夕方に恋人と行くときには、カフェ X は満足」と
 していた場合、ユーザ B は「朝に1人で行くときは、
 カフェ X は不満足」と評価する可能性もある。こ

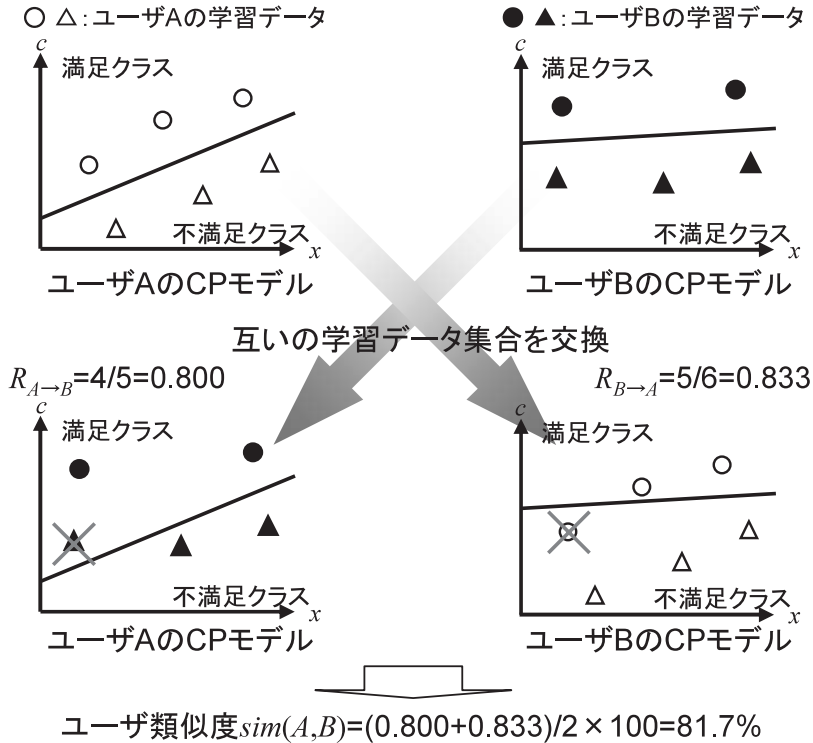


図 2 CP モデルに基づくユーザ類似度の算出
Fig. 2 Calculating user similarity based on their CP-Models.

の場合、ユーザ A とユーザ B は嗜好が類似しているとはいえない。つまり、ユーザの評価時のコンテキストを考慮しないと、ユーザの嗜好の類似度を適切に算出することはできない。

この問題に対し、我々は、互いのユーザの CP モデルの類似度に基づいてユーザ類似度を算出するという方法をとる。前節で説明したように、CP モデルには、そもそもユーザの評価時のコンテキストの情報も含まれているため、このモデルの類似度に基づいてユーザ類似度を算出することにより、評価時のコンテキストも含めて適切にユーザの嗜好の類似度を算出することができる。

その具体的なユーザ類似度の算出手順を以下に示す。図 2 のように、あらかじめ CP モデルが構築された 2 人のユーザ、ユーザ A とユーザ B を対象に、ユーザ類似度の算出方法を説明する。

- (1) 各ユーザの学習データ集合を互いの CP モデル上に交換する。
- (2) ユーザ A の CP モデルにより、ユーザ B の学習データ 1 つ 1 つに対して属するクラス (“満足”・“不満足”) を判定する。
- (3) そのユーザ A のモデルにより判定されたクラ

スとユーザ B の評価値とが一致したデータ数 ($M_{A \rightarrow B}$) を数える (図 2 の例では $M_{A \rightarrow B} = 4$)。

- (4) ユーザ B の全学習データ数 (N_B) に対する一致データ数の割合 ($R_{A \rightarrow B}$) を求める (図 2 の例では $R_{A \rightarrow B} = 4/5 = 0.800$)。
- (5) ユーザ A の学習データ集合に対しても同様にを行い、 $R_{B \rightarrow A}$ を求める (図 2 の例では $R_{B \rightarrow A} = 5/6 = 0.833$)。
- (6) $R_{A \rightarrow B}$ と $R_{B \rightarrow A}$ との平均値をユーザ類似度 ($sim(A, B)$) とする (図 2 の例では $sim(A, B) = 81.7\%$)。

以上の算出方法を一般化すると、ユーザ u とユーザ v とのユーザ類似度 $sim(u, v)$ は、式 (1) により求められる。

$$sim(u, v) = \frac{1}{2} \left(\frac{M_{v \rightarrow u}}{N_u} + \frac{M_{u \rightarrow v}}{N_v} \right) \times 100(\%) \quad (1)$$

このユーザ類似度が高いユーザ同士を類似ユーザとして扱う。

互いのユーザのコンテンツに対する評価値の相関関係に基づく従来の算出方法では、両ユーザがいくつかの同じコンテンツに対して評価を与えていることが前

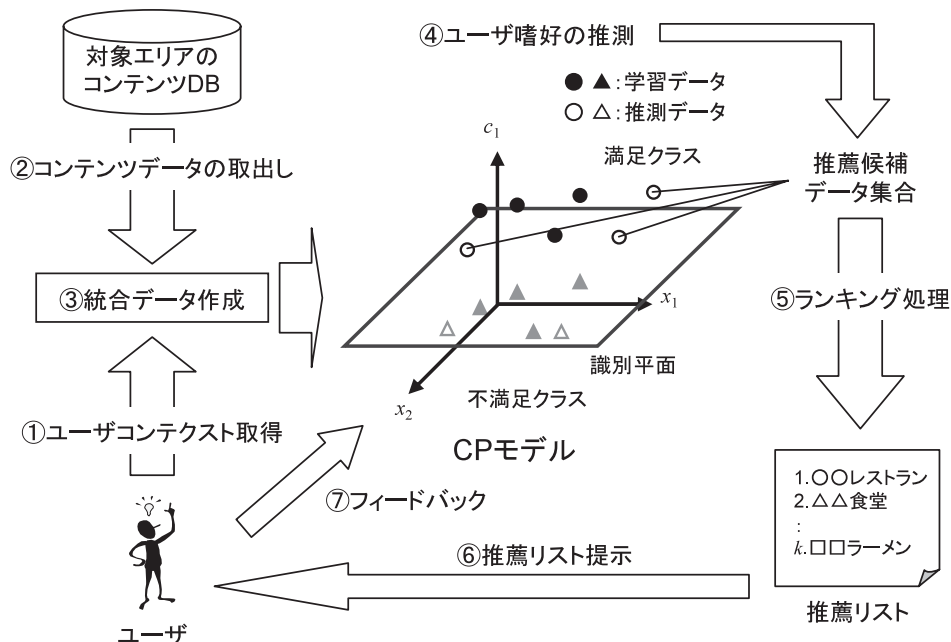


図 3 コンテキスト依存型情報フィルタリング (C-IF) によるシステムの動作手順
 Fig. 3 Procedure of system based on Context-Aware Information Filtering (C-IF).

提となる。これに対して、CP モデルに基づくこの方法では、必ずしもユーザ間で同じコンテンツに対して評価付けされている必要はない。

5. コンテキスト依存型情報推薦システムのプロトタイプ

本研究では、前章で説明した CP モデルを適用したコンテキスト依存型情報推薦システム (Context-Aware Recommendation System) のプロトタイプを実装した。この情報推薦システムは以下の 2 つの情報推薦方式からなる。

- (a) コンテキスト依存型情報フィルタリング (Context-Aware Information Filtering; C-IF)
- (b) コンテキスト依存型協調フィルタリング (Context-Aware Collaborative Filtering; C-CF)

本章では、飲食店情報推薦を例として、これらの情報推薦方式について説明する。

5.1 コンテキスト依存型情報フィルタリング (C-IF)

コンテキスト依存型情報フィルタリング (以降、C-IF と称する) は、対象ユーザの CP モデルに基づいて、データベースから、ユーザのおかれているコンテキストにおけるユーザの嗜好に合致するコンテンツを検索し、そのコンテンツをユーザに推薦しようというものである。

以下、システムの動作手順を示す (図 3 参照)。

- (1) ユーザのコンテキストを自動的に、もしくはユーザによる明示的な入力により取得する。
- (2) 対象エリア (ユーザが現在位置しているエリア) のコンテンツデータベースから、ユーザの現在位置に最も近い n 件のコンテンツデータを取り出す。
- (3) 取り出したコンテンツデータ 1 つ 1 つの特徴パラメータとユーザコンテキストの特徴パラメータとを統合したデータ集合を作成する。
- (4) 統合したデータ集合を CP モデル上に配置し、この CP モデルに基づいて、各データに対するユーザの嗜好 (“満足”・“不満足”) を推測する。
- (5) ユーザの現在おかれているコンテキストにおいて満足されると推測されたデータ集合を推薦候補のコンテンツデータ集合とし、このデータ集合に対してランク付けを行う。
- (6) そのランキングに従って、上位 k 件のコンテンツデータを推薦リストとして、ユーザに提示する。
- (7) 推薦データに対するユーザのフィードバックに基づいて新たな学習データを追加する。

5.2 コンテキスト依存型協調フィルタリング (C-CF)

コンテキスト依存型協調フィルタリング (以降、C-

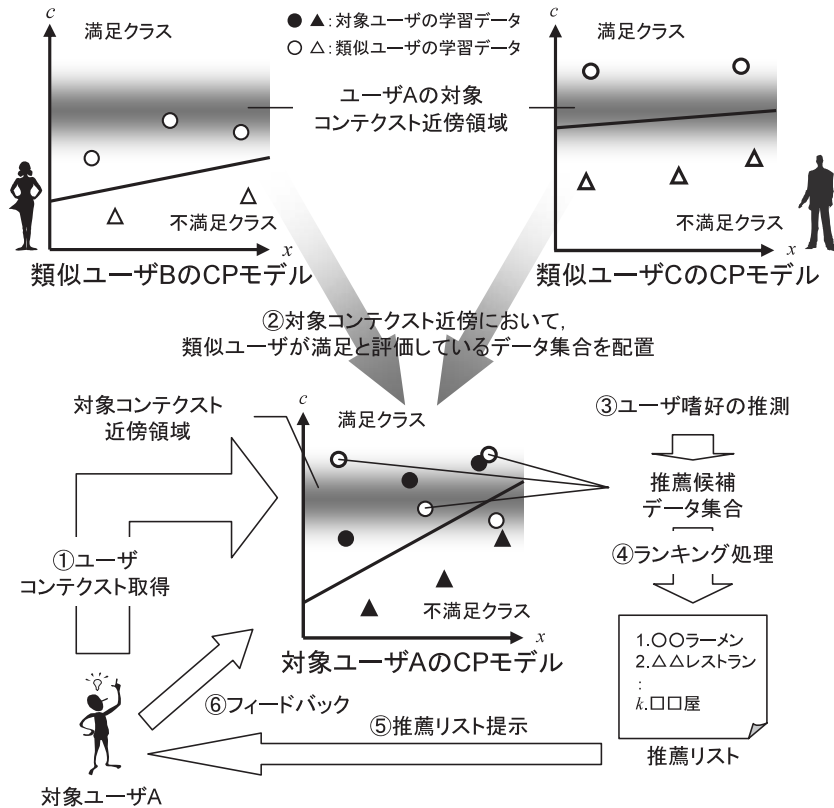


図 4 コンテキスト依存型協調フィルタリング (C-CF) によるシステムの動作手順
 Fig. 4 Procedure of system based on Context-Aware Collaborative Filtering (C-CF).

CF と称する) では、ユーザ同士の CP モデルの類似度に基づいて、類似ユーザを探し出し、対象ユーザのコンテキスト近傍において、その類似ユーザが満足と評価しているコンテンツを対象ユーザに推薦しようというものである。

以下、システムの動作手順を示す (図 4 参照)。ここでは、まず 4.2 節で説明したユーザ類似度の算出手順に従って、あらかじめ対象エリアに存在する類似ユーザを探索しておき、ユーザ類似度が最も高い m 人のユーザの CP モデルを参照することにする。図 4 では、対象ユーザをユーザ A、そのユーザ A の類似ユーザとして、ユーザ B およびユーザ C が発見されているとする。

- (1) 対象ユーザのコンテキストを自動的に、もしくはユーザによる明示的な入力により取得する。
- (2) m 人の類似ユーザの CP モデルに基づいて、対象ユーザのおかれているコンテキストの近傍において、満足と評価されているコンテンツデータ集合を対象ユーザの CP モデル上に配置する。
- (3) 対象ユーザの CP モデルに基づいて、配置された各コンテンツデータに対するユーザの嗜好

(“満足”・“不満足”) を推測する。

- (4) ユーザの現在おかれているコンテキストにおいて満足されると推測されたデータ集合を推薦候補のコンテンツデータ集合とし、このデータ集合に対してランク付けを行う。
 - (5) そのランキングに従って、上位 k 件のコンテンツデータを推薦リストとして、ユーザに提示する。
 - (6) 推薦データに対するユーザのフィードバックに基づいて新たな学習データを追加する。
6. ユーザ嗜好モデリング手法として SVM を適用することの妥当性の検証

情報推薦システムは、ユーザが把握しきれていないコンテンツを適切に推薦することに価値がある。したがって、情報推薦手法に対して適用するモデリング手法は、既知 (学習済み) データに対して正しく識別できる能力も当然必要であるが、むしろ未知 (未学習) データに対して高い汎化性能を示すことが望ましい。本研究では、一般に汎化性能に優れているといわれている理由から、ユーザ嗜好をモデリングするためのク

ラス分類手法として SVM³⁾ を用いている。

SVM のほかには、主なクラス分類手法として、Neural Network (NN)¹⁵⁾、 k -Nearest Neighbor (k NN)¹⁶⁾、Decision Tree (DecTree)^{17),18)}、Bayesian Filtering (BF)¹⁹⁾ があげられる。本実験では、汎化性能の観点から、これら 5 つの手法を比較対象とした定量評価実験を行った。本章では、この定量評価実験の結果を示し、それに対する考察を行う。

6.1 実験方法

6.1.1 データセット

まず飲食店情報推薦を対象とした実データセットを用いて、各クラス分類手法の汎化性能について比較評価を行う。また、データセットの特性の違いに対する汎用性を分析するために、次元数およびデータ数が異なる人工データセットに対しても比較評価を行う。

我々はすでに提案手法の有効性を示すために、飲食店情報推薦を対象として、被験者を用いた検証実験を行っている^{1),2)}。この検証実験において用いた実データセットを本実験で用いる。実データセットは、飲食店データの特徴パラメータ 17 次元とコンテキストの特徴パラメータ 14 次元の合計 31 次元から構成される。データ数は 100 個であり、1 つ 1 つのデータに対して被験者による評価値 (“満足”・“不満足”) が与えられている。

人工データセットとして、表 1 に示す 5 つのデータセットを用いる。各データセットは、LIBSVM のサイト^{20),21)} で提供されている 2 クラス用のデータセットのうち、それぞれ表 1 に示すようにデータ数および次元数が異なるように選択した。

6.1.2 評価指標

性能評価のための指標として識別性能を用いる。まず、学習データ集合に基づいて各クラス分類手法による識別モデルを構築する。つづいて、この識別モデルによりテストデータ集合に対して識別判定を行う。このとき、そのテストデータ集合に対してどの程度正しく識別できたかを示す指標が識別性能であり、式 (2) で表される。

$$\text{識別性能} = \frac{\text{正しく識別されたデータ数}}{\text{総テストデータ数}} \times 100(\%) \quad (2)$$

本実験では、未知データに対する識別性能すなわち汎化性能について検証する。これは、データセットの一部のデータを学習データとし、残りのデータをテストデータとしたときの識別性能である。具体的には、5-fold cross-validation により、総データ数の 4/5 を学習データ、残りの 1/5 をテストデータとし、この 5 通りの組合せに対する平均をとっている。これは、

表 1 クラス分類手法の性能分析に用いる人工データセット

Table 1 Artificial data sets used for classifier's performance analysis.

データセット名	データ数	次元数
fourclass	862	2
heart	270	13
australian	690	14
german	1,000	24
sonar	208	60

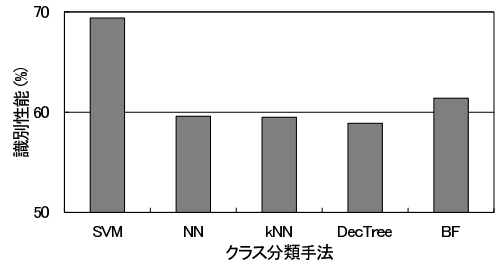


図 5 実データセットを対象としたときの未知データに対する識別性能の比較

Fig. 5 Model accuracy comparison for untrained data using real data sets.

学習データおよびテストデータの選び方により生じるデータ分布の偏りをなくすためである。

なお、各クラス分類手法による識別性能の取得には、データマイニングツール WEKA²²⁾ を用いた。

6.2 実験結果と考察

図 5 に実データセットを対象としたときの未知データに対する識別性能を示す。

この結果から、NN、 k NN、DecTree、BF においては、未知データに対する識別性能が 60% 程度であったのに対し、SVM は 70% 近い性能を示した。本実験ではまた、5 つの人工データセットいずれに対しても、SVM が比較的高い識別性能を示すことを確認した。

実データセットを対象としたときには、SVM は他の各手法に対して比較的高い汎化性能を示した。したがって、飲食店を対象コンテンツとした情報推薦システムにおいては、SVM を情報推薦手法に対して適用することは有用性があるといえる。

今回は、実データセットとして飲食店情報推薦を対象としたものだけを扱ったが、データ数および次元数が異なるさまざまな人工データセットに対しても、比較的高い汎化性能を示した SVM は、データセットの特性にとらわれず汎用性があるといえる。したがって、推薦対象コンテンツを音楽や映画などに変えた場合でも、汎用的に有用性を示すことが期待できる。

7. 被験者満足度によるシステムの評価

本章では、5章で説明したプロトタイプシステムの評価を行う。具体的には、飲食店を推薦対象コンテンツとしたとき、提案方式による推薦に対する被験者の満足度を用いてシステムの有用性を評価する。

7.1 実験条件

7.1.1 データセット

本実験では、飲食店情報を対象としたデータセットを用いた。ここでは、一般向けグルメサイトである『Yahoo!グルメ²³⁾』を参考に、表2に示すエリアを対象に選択した。『Yahoo!グルメ』に登録されているデータ件数のうち、本実験で用いることができる有効なデータだけを用いる。その有効データ数は表2に示すとおりである。さらに、『Yahoo!グルメ』を参考に、表3に示す28次元の飲食店特徴パラメータを設定した。

また、コンテキスト特徴パラメータについても、同様に表4に示す24次元の特徴パラメータを設定した。これらは、ユーザの飲食店情報の選択に対する意思決定に影響を及ぼすという観点から、著者が主観的に決めたものである。なお、これら各特徴パラメータの選定においては、特徴ベクトル次元の直交性確保の観点から、できる限り独立関係になるよう心がけた。

7.1.2 被験者と評価指標

前項で示したコンテキスト特徴パラメータの組合せにより表現されたユーザコンテキストにおいて、飲食店データに対する評価(“満足”・“不満足”)を被験者に行ってもらい、被験者は、研究室の学生および教官で、22歳から35歳の9人(被験者 a, b, ..., i)である。いずれの被験者も、表2に示したエリアには何度か訪れたことがあり、これらのエリアの雰囲気などについてはある程度把握している。このように被験者が把握しているエリアを実験対象として選択したのは、研究室内で仮想的に実験を行うため、本実験を行うにあたって、できるだけ被験者にユーザコンテキストを想像しやすくするためである。

なお、本実験では、被験者満足度を用いてシステムの評価を行う。満足度とは、被験者が情報推薦に対してどの程度満足したかを表す指標であり、式(3)で表される。

$$\text{満足度} = \frac{\text{被験者が満足した回数}}{\text{推薦回数}} \times 100(\%) \quad (3)$$

7.2 ランダム提示およびコンテキスト非依存の推薦方式との比較

本節では、C-IF および C-CF による情報推薦の有

表2 飲食店データセット

Table 2 Restaurant data sets.

エリア	登録件数	有効データ数
I 心斎橋, 堀江	2,069	650
II 天王寺, 阿倍野	416	121
III 梅田, 中津	1,508	445

表3 飲食店特徴パラメータ(28次元)

Table 3 Restaurant feature parameters (28 dimensions).

パラメータ項目	パラメータ値
ジャンル (8)	[和食] / [洋食] / [中華] / [居酒屋] / [アジア] / [ラーメン] / [カフェ, スイーツ] / [その他]
予算(昼夜)(2)	0~10,000円
店の特徴 (11)	[個室] / [駐車場] / [禁煙席] / [カラオケ] / [深夜営業] / [食べ放題] / [ランチ] / [カード可] / [ライブ] / [テイクアウト] / [貸し切り可]
雰囲気 (7)	[デート向き] / [接待向き] / [宴会・パーティー向き] / [合コン向き] / [ファミリー向き] / [1人でも可] / [少人数向き]

表4 コンテキスト特徴パラメータ(24次元)

Table 4 Context feature parameters (24 dimensions).

パラメータ項目	パラメータ値
時間情報	月 (2) 1~12月 曜日 (2) 日~土 時刻 (2) 00:00~23:59
気象情報	天気 (4) [晴れ] / [曇り] / [雨] / [雪] 気温 (1) 0~40°C
ユーザ情報	休日 (2) [休日] / [休日前] 予算 (1) 0~10000円 余裕時間 (1) 0~4時間
同伴者情報	グループ (7) [1人] / [家族] / [恋人] / [友人・先輩・後輩] / [上司・部下] / [ビジネス] / [該当なし] 人数(男女)(2) 0~10人

効性を、ランダム提示およびコンテキストに依存しない推薦方式との比較を行うことにより評価する。

7.2.1 実験方法

本実験で用いるデータセットの対象エリアは、表2に示すうち、「心斎橋, 堀江」エリアとした。

まず、あらかじめ下記のとおり各被験者のユーザ嗜好モデルを構築しておく。

- (1) 被験者には、各自「心斎橋, 堀江」エリアにおいて、現実起こりうるコンテキストを想像してもらい、あらかじめ10パターンのコンテキストを設定してもらう。
- (2) 被験者には、1つのコンテキストあたり、デー

タセットの中からランダムに選択された 10 件の飲食店データに対して，“満足”・“不満足”の評価を与えてもらう。

- (3) この評価に基づいて、被験者 1 人あたり合計 100 個 (=10 コンテキスト ×10 飲食店データ) の学習データを作成する。
- (4) 作成された学習データに基づいてコンテキスト依存/非依存型のユーザ嗜好モデルを構築する (4.1 節参照)。
- (5) コンテキスト依存型ユーザ嗜好モデルに基づいて各被験者間のユーザ類似度を算出する (4.2 節参照)。ここで、各被験者に対してユーザ類似度が最も高い 3 人の被験者を C-CF における類似ユーザとして扱う。

つづいて、以下の各方式で被験者に対して飲食店を推薦する。

- ランダム提示：被験者の嗜好およびコンテキストに関係なく、ランダムに飲食店を提示する。
- コンテキスト非依存型推薦：被験者のコンテキストは考慮しないが、ある程度の嗜好に基づいて飲食店を推薦する。
- C-IF による推薦：コンテキストごとに変化する被験者の嗜好を考慮して飲食店を推薦する。
- C-CF による推薦：コンテキストも考慮した類似ユーザの嗜好に基づいて飲食店を推薦する。

被験者は、各方式により推薦された飲食店に対して，“満足”・“不満足”の評価を与える。この評価に基づいて、被験者満足度を算出する。

7.2.2 実験結果と考察

図 6 に各方式に対する被験者満足度を示す。これは、9 人の被験者の満足度の平均を示している。C-CF に対する満足度が最も高く、次いで、C-IF に対する満足度が高かった。

ランダム提示は、そもそも提示する飲食店に根拠がないため、最も満足度が低かったということは自明の結果である。コンテキスト非依存型推薦は、コンテキストごとに変化するユーザの嗜好を正しくモデル化できていないことから、ときにコンテキストにふさわしくない飲食店が推薦されたことなどが満足度が低くなった原因であると考えられる。これに対して、C-IF は、コンテキストごとに変化するユーザの嗜好を正しくモデル化しているため、コンテキストに合った飲食店が適切に推薦されたため、比較的高い満足度が得られたといえる。さらに、C-CF は、コンテキストに適合し、かつ 1 度他のユーザが満足と評価した飲食店を推薦しているため、推薦情報に対する信頼度が高く、

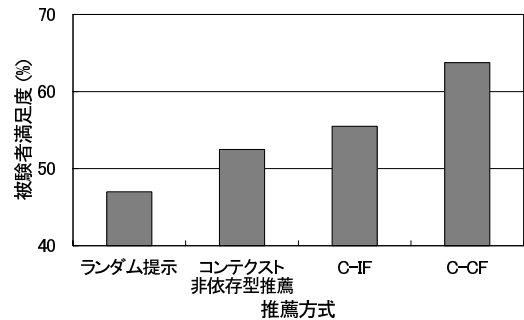


図 6 各推薦方式に対する被験者満足度の比較

Fig. 6 Users' satisfaction rates comparison for each recommendation method.

結果、最も高い満足度が得られたといえる。

以上のように、C-IF および C-CF の有効性を示した。なお、本実験では結果的に C-CF に対する満足度が最も高くなったが、これは、推薦対象エリアが固定であり、また各被験者がそのエリアにおいてあらかじめ学習データを作成しているという条件下での結果である。そもそも周りに類似ユーザが存在しない場合など、システムの利用場面の違いにより、C-IF および C-CF の有効性も変わってくる可能性もある。次節では、これをふまえて、推薦対象エリア別の C-IF および C-CF の特長の違いを検証する。

7.3 C-IF および C-CF の推薦対象エリア別の有効性の比較

本節では、C-IF および C-CF による有効性を推薦対象エリア別に比較することによって、各々のシステム利用場面に応じた特長の違いを分析する。

7.3.1 実験方法

実験方法は以下のとおりである。

- (1) 9 人の被験者を 3 人ずつの 3 つの被験者グループ ($A\{a, b, c\}$, $B\{d, e, f\}$, $C\{g, h, i\}$) に分ける。
- (2) 各被験者グループはそれぞれ対応するエリア (グループ A → エリア I, グループ B → エリア II, グループ C → エリア III) で学習データを作成する。
- (3) 3 つのエリアにおいて、C-IF および C-CF により各被験者に対して飲食店を推薦する。
- (4) 被験者は、推薦された飲食店に対する評価 (“満足”・“不満足”) を与える。

被験者 1 人あたりの学習データ数は 100 個 (=10 コンテキスト ×10 飲食店データ) とする。推薦回数は、各被験者に対して、1 エリアあたり C-IF および C-CF 各 25 回 (=5 コンテキスト ×5 飲食店データ) ずつの 50 回、3 エリアで合計 150 回とする。ここで、

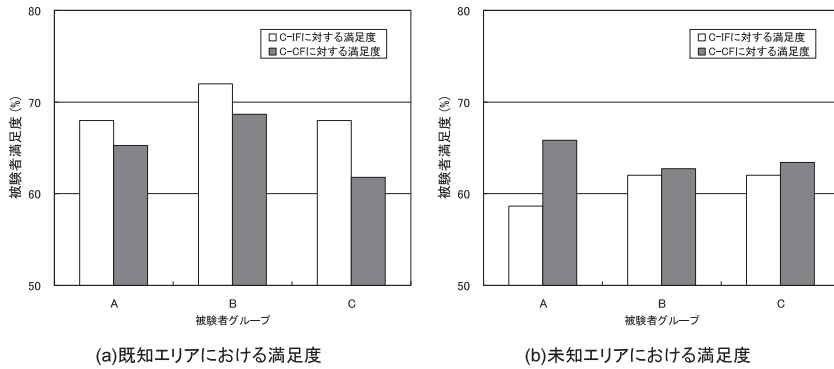


図 7 既知/未知エリアにおける C-IF および C-CF に対する満足度の比較
Fig. 7 Users' satisfaction rates comparison for C-IF and C-CF in trained/untrained area.

C-CF において用いる類似ユーザは、推薦対象エリアを学習エリアとしたグループに属する被験者すべてとする。たとえば、被験者 a がエリア II を推薦対象エリアとしていた場合、そのエリアにおいて学習データを作成した被験者 d, e, f が、被験者 a の類似ユーザとなる。

なお、本実験では、被験者が学習データを作成したエリアを、その被験者にとって「既知エリア」とし、既知エリア以外のエリアを「未知エリア」とよぶ。たとえば、被験者グループ A の場合、エリア I が既知エリアであり、エリア II および III が未知エリアである。

7.3.2 実験結果と考察

図 7 に、対象エリア別の被験者満足度を示す。(a) は既知エリアを推薦対象としたときの、(b) は未知エリアを推薦対象としたときの被験者満足度をそれぞれ示している。本実験より以下の結果が得られた。

- 既知エリアを推薦対象とした場合には、C-IF に対する満足度の方が高い。
- 未知エリアを推薦対象とした場合には、C-CF に対する満足度の方が高い。

この結果は、C-IF は対象ユーザ自身の嗜好が学習されたモデルに基づく推薦方式であるため、その学習データが十分に蓄積されたエリアを推薦対象としたときには、C-IF による情報推薦の方が有効であるということを示唆している。一方で、推薦対象エリアにおける学習データが乏しいときには、代わりにそのエリアにおいて十分な学習データをもつ他の類似ユーザのモデルに基づいた推薦方式、つまり C-CF による推薦方式の方が有効であるということも示唆している。

C-IF は対象ユーザの嗜好に忠実に従った、正確性を重視した推薦方式であるといえ、対して、C-CF は類似ユーザの嗜好や経験に基づくため、意外性を重視した推薦方式であるといえる。このような C-IF の正

確性、C-CF の意外性といった各々の特長点をふまえて、推薦対象エリアなどの利用場面に応じて、推薦方式に対する優先度を変えることにより、正確性と意外性を考慮した情報推薦が可能であるといえる。

7.4 推薦対象エリア特性に基づく C-IF の有効性の分析

さらに、C-IF による情報推薦に焦点を当てて、前節で得られた被験者満足度を被験者グループ別および推薦対象エリア別の傾向を分析する。

図 8 に C-IF に対する推薦対象エリア別の被験者満足度を示す。被験者グループ別にグラフを分けて示しており、(a)、(b)、(c) がそれぞれ被験者グループ A、B、C の被験者満足度に対応する。

この結果から、前節の結果で示したとおり、いずれの被験者グループも学習対象エリアにおいて高い満足度を示していることが分かる。被験者グループ A とグループ C について、より深くみると、被験者グループ A はエリア III を推薦対象としたときにも高い満足度が得られ、逆に被験者グループ C はエリア I を推薦対象としたときにも高い満足度が得られたことが分かる。エリア I 「心斎橋、堀江」およびエリア III 「梅田、中津」は、ともに大阪都心部の歓楽街であり、店舗分布や利用者の活動内容などの地域特性が類似している²⁴⁾。この結果は、推薦対象エリアと学習エリアとの地域特性が類似している場合には、C-IF による情報推薦もある程度有効に働くということを示唆している。

8. 特徴パラメータの最適化に関する検討

すでに説明したように、本実験では特徴パラメータとして、表 3、表 4 に示すものを用いた。ただし、これら特徴パラメータは、著者が主観的に選択したものであり、必ずしも最適なものであるという保証はない。

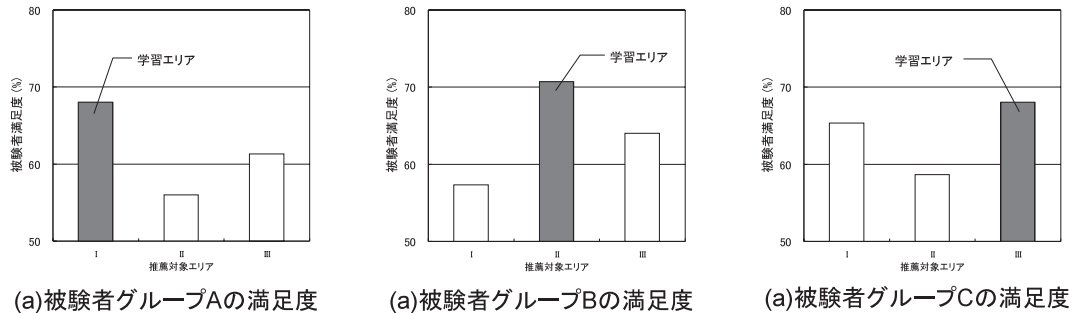


図 8 各被験者グループの推薦対象エリア別の C-IF に対する満足度の比較

Fig. 8 Users' satisfaction rates comparison for C-IF in each recommendation area.

特徴パラメータの中には、ユーザの情報選択の判断に対する意思決定に強く影響を与えるものもあれば、逆にモデル構築に際してノイズとなってしまうものもあると考えられる。こういったノイズとなるような特徴パラメータは、システムの推薦精度を低下させてしまうため、システム設計の際にはこれらの特徴パラメータを最適化する必要がある。本実験では、各特徴パラメータがモデルの識別性能に寄与している度合いを分析することによって、特徴パラメータの最適化に関する検討を行う。

8.1 実験方法

表 3, 表 4 に示した特徴パラメータを対象に分析を行う。分析対象のモデルは、前章の実験において構築した被験者 9 人それぞれの CP モデルとする。

以下、実験手順を示す。

- (1) あらかじめ前章の実験で構築した CP モデルの識別性能を得ておき、これを標準時の識別性能とする。
- (2) 表 3 に示す飲食店特徴パラメータを「ジャンル」、「予算」、「店の特徴」、「雰囲気」の 4 組に、表 4 に示すコンテキスト特徴パラメータを「月」、「曜日」、「時刻」、「気象情報」、「ユーザ情報」、「同伴者情報」の 6 組に、それぞれ分ける。
- (3) 学習データを構成する特徴パラメータから、上にあげた特徴パラメータを 1 組ずつ排除する。
- (4) 1 組の特徴パラメータが排除された学習データ集合から各被験者の CP モデルを新たに構築し、そのモデルの識別性能を得る。
- (5) 特徴パラメータ排除後の識別性能と標準時の識別性能とを比較する。
- (6) これをすべての組について行う。

以上の実験の結果、排除されることにより識別性能を標準時よりも低下させた特徴パラメータは、もともとモデルの識別性能の向上に寄与していた特徴パラ

メータであるといえる。一方で、排除されることにより識別性能を標準時よりも向上させた特徴パラメータは、モデル構築に際してノイズとなっていた特徴パラメータであるといえる。このような考えに基づいて特定された、ノイズとなった特徴パラメータを排除していくという方法によって、特徴パラメータの最適化を行う。

8.2 実験結果

図 9 に、飲食店特徴パラメータとコンテキスト特徴パラメータを 1 組ずつ排除したときのモデルの識別性能を示す。ここでの識別性能は、未知データに対する識別性能すなわち汎化性能であり、9 人の被験者の CP モデルの識別性能の平均を示している。横軸は排除した特徴パラメータの組を表す。図中の破線は特徴パラメータを排除しないとき、つまり標準時の識別性能を示す。

図 9 から、特徴パラメータを排除することによって、識別性能を標準時よりも低下させたものとして、低下率が大きいものから順に、

- (1) 「ジャンル」
- (2) 「同伴者情報」
- (3) 「時刻」、「ユーザ情報」
- (4) 「雰囲気」
- (5) 「予算」

の 6 組が特定された。これらは、モデルの識別性能の向上に寄与していた特徴パラメータであるといえる。ここで、この低下率が大きいものほど、識別性能の向上に大きく寄与していたものであり、特に「ジャンル」パラメータは、本実験においては、その向上に大きく寄与していたといえる。

一方、特徴パラメータを排除することによって、識別性能を標準時よりも向上させたものとして、向上率が大きいものから順に、

- (1) 「店の特徴」

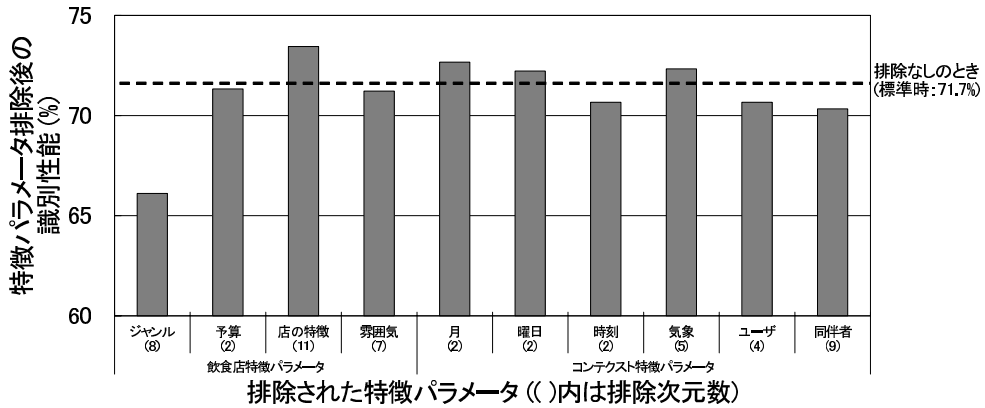


図 9 特徴パラメータがモデルの識別性能に与える影響の分析
Fig. 9 Analysis of influence of feature parameters on model accuracy.

表 5 特徴パラメータを累積的に排除したときの識別性能の変化
Table 5 Change of model accuracy when removing feature parameters cumulatively.

特徴パラメータの組 (×: 排除パラメータ)					識別性能 (%)
店の特徴	月	気象情報	曜日	予算	
-	-	-	-	-	71.7
×	-	-	-	-	73.4
×	×	-	-	-	73.8
×	×	×	-	-	74.1
×	×	×	×	-	74.7
×	×	×	×	×	73.6

- (2) 「月」
- (3) 「気象情報」
- (4) 「曜日」

の 4 組が特定された。これらは、モデル構築に際してノイズとなっていた特徴パラメータであるといえる。

さらに、この 4 組のノイズとなった特徴パラメータと「予算」パラメータとを累積的に排除していったときの識別性能の変化を表 5 に示す。この結果から、各被験者の CP モデルから、「店の特徴」、「月」、「気象情報」、「曜日」の 4 組を排除したときに、最大の識別性能 (74.7%) が得られた。

8.3 考察

以上の実験結果に基づいて、特徴パラメータ最適化に関する考察を行う。

本実験の被験者に対して、「情報選択の判断において、参考にした項目の順位」に関してヒアリング調査を行った。この調査の結果、被験者が意識的に参考にした項目の順位は、飲食店情報の項目に関して、

- (1) 「ジャンル」
- (2) 「予算」
- (3) 「雰囲気」

コンテキストの項目に関して、

- (1) 「同伴者情報」
- (2) 「時刻」
- (3) 「ユーザ情報」

となり、この結果と本実験結果から得られた順位とがほぼ一致することを確認した。

この結果から、特徴パラメータがモデルの識別性能に与える影響を分析することによって、ユーザの情報選択の判断基準となる特徴パラメータを抽出することが可能であるといえる。本来、人の意思決定過程をモデル化するためには、行動心理など複雑な事象に基づいた分析が求められると考えられるが、本実験結果は、このような複雑な分析を、今回のような簡単な方法により行うことができることを示唆している。また、表 5 に示すように、ノイズとなった特徴パラメータを排除していくことによって、識別性能を高めていくことが可能であるという結果が得られたことから、本実験で行った分析方法は、特徴パラメータ最適化手法として、適用しうるものであるといえる。

ただし、本実験で特定した最適パラメータは、あくまでも今回の実験環境および実験条件によるものであるため、この最適パラメータをそのまま実用システムに適用できるとは限らない。システムを実用化するためには、たとえば、システム運用前に一定のモニタ期間を設置するなど、より現実に近い環境において、特徴パラメータを最適化する必要がある。

また、扱う特徴パラメータに関しては、各特徴ベクトル次元の直交性が保証されていることが望ましい。本研究で採用した特徴パラメータに関しても、できる限り独立関係になるよう選定したが、主観的な要素も含んでいるため、完全な直交性を保証することは困難である。したがって、今後も、できる限り各パラメータの直交性を確保することを目指すことで、推薦精度

の低下を防ぐことが必要であると考えている。

9. おわりに

本論文では、コンテキストに応じて変化するユーザの嗜好をモデル化する手法、コンテキスト依存型ユーザ嗜好モデリング手法 (CP モデリング) について説明し、これを適用したコンテキスト依存型情報推薦システムのプロトタイプを説明した。このプロトタイプは以下の 2 つの情報推薦方式から構成される。

- (a) コンテキスト依存型情報フィルタリング (C-IF)
- (b) コンテキスト依存型協調フィルタリング (C-CF)

本論文ではまた、このプロトタイプを飲食店情報推薦に対して適用した検証実験を行った。この検証実験より得られた成果と知見を以下に示す。

- 他のクラス分類手法との比較評価の結果、SVM は未学習データに対する汎化性能に優れていることが実証でき、情報推薦システムに対して適用することの妥当性を裏付けた。
- 2 つの情報推薦方式、C-IF および C-CF に対して以下の特長を明らかにした。
 - 既知エリアを推薦対象とした場合には、C-IF による情報推薦が有効に働く。
 - 未知エリアを推薦対象とした場合には、C-CF による情報推薦が有効に働く。
- 推薦対象コンテンツおよびユーザコンテキストの特徴パラメータの最適化に関する検討を行い、各特徴パラメータが識別性能に与える影響を分析することによって、特徴パラメータの最適化が可能であるということを示した。

今後は、ユーザの意思決定に影響を及ぼすコンテキストとして、周辺環境や時系列的要因も考慮に入れた情報推薦システムについて検討する。また、飲食店だけでなく音楽情報や映画情報など、さまざまなコンテンツを推薦対象とした情報推薦システムの開発を行う。

謝辞 本研究の一部は、科学研究費補助金基盤研究 (A)(2)(課題番号: 15200010)、若手研究 (B)(課題番号: 17700132) の支援による。ここに記して謝意を表す。

参考文献

- 1) 奥 健太, 中島伸介, 宮崎 純, 植村俊亮: Context-Aware SVM に基づく状況依存型情報推薦方式の提案, 日本データベース学会論文誌 (DBSJ Letters), Vol.5, No.1, pp.5-8 (2006).
- 2) Oku, K., Nakajima, S., Miyazaki, J. and Uemura, S.: Context-Aware SVM for Context-Dependent Information Recommendation, *Proc. International Workshop on Future Mobile and Ubiquitous Information Technologies (FMUIT 2006)*, pp.119-122 (2006).
- 3) Cortes, C. and Vapnik, V.: Support-Vector Networks, *Machine Learning*, Vol.20, No.3, pp.273-297 (1995).
- 4) 土方嘉徳: 情報推薦・情報フィルタリングのためのユーザプロファイリング技術, 人工知能学会誌, Vol.19, No.3, pp.365-372 (2004).
- 5) 岩濱数宏, 土方嘉徳, 西田正吾: 決定木を用いた音楽情報フィルタリングシステムとその有効性の検証, 電子情報通信学会論文誌 D, Vol.J88-D1, No.3, pp.642-656 (2005).
- 6) 杉山一成, 波多野賢治, 吉川正俊, 植村俊亮: ユーザからの負担なく構築したプロフィールに基づく適応的 Web 情報検索, 電子情報通信学会論文誌 D, Vol.J87-D1, No.11, pp.975-990 (2004).
- 7) Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P. and Riedl, J.: GroupLens: An Open Architecture for Collaborative Filtering of Netnews, *Proc. ACM 1994 Conference on Computer Supported Cooperative Work (CSCW'94)*, pp.175-186 (1994).
- 8) Konstan, J.A., Miller, B.N., Maltz, D., Herlocker, J.L., Gordon, L.R. and Riedl, J.: GroupLens: Applying collaborative filtering to Usenet news, *Comm. ACM*, Vol.40, No.3, pp.77-87 (1997).
- 9) MovieLens. <http://movielens.umn.edu/login> (2007-4-17 確認)
- 10) Amazon.com. <http://www.amazon.com/> (2007-4-17 確認)
- 11) Chen, G. and Kotz, D.: A Survey of Context-Aware Mobile Computing Research, Paper TR2000-381, Department of Computer Science, Dartmouth College (2000).
- 12) van Setten, M., Pokraev, S. and Koolwaaij, J.: Context-Aware Recommendations in the Mobile Tourist Application COMPASS, *Adaptive Hypermedia and Adaptive Web-Based Systems*, Lecture Notes in Computer Science, Vol.LNCS3137, Springer Berlin/Heidelberg, pp.235-244 (2004).
- 13) Asthana, A., Cravatts, M. and Krzyzanowski, P.: An Indoor Wireless System for Personalized Shopping Assistance, *Proc. IEEE Workshop on Mobile Computing Systems and Applications*, Santa Cruz, California, pp.69-74, IEEE Computer Society Press (1994).
- 14) Abowd, G.D., Atkeson, C.G., Hong, J., Long, S., Kooper, R. and Pinkerton, M.: Cyberguide: A Mobile Context-Aware Tour Guide, *Wireless*

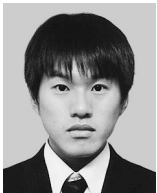
Networks, Vol.3, No.5, pp.421-433 (1997).

- 15) Haykin, S.: *NEURAL NETWORKS: A comprehensive foundation*, Prentice-Hall Engineering/Science/Mathematics, international edition (1998).
- 16) Duda, R.O., Hart, P.E. and Stork, D.G.: *Pattern Classification and Scene Analysis*, 2nd ed., John Wiley & Sons, New York (1993).
- 17) Quinlan, J.R.: Induction of decision trees, *Machine Learning*, Vol.1, No.1, pp.81-106 (1986).
- 18) Quinlan, J.: *C4.5: Programs for Machine Learning*, Morgan Kaufmann (1993).
- 19) Hand, D.J., Mannila, H. and Smyth, P.: *Principles of Data Mining*, The MIT Press (2001).
- 20) Chang, C.-C. and Lin, C.-J.: *LIBSVM: A Library for Support Vector Machines* (2006).
- 21) LIBSVM. <http://www.csie.ntu.edu.tw/~cjlin/libsvm/> (2007-4-17 確認)
- 22) WEKA. <http://www.cs.waikato.ac.nz/ml/weka/> (2007-4-17 確認)
- 23) Yahoo! グルメ. <http://gourmet.yahoo.co.jp/gourmet/> (2007-4-17 確認)
- 24) 李 知雄, 西井和夫, 近藤勝直, 佐々木邦明: 都市観光回遊行動パターンの基礎分析: 大阪キタ・ミナミ地区の比較, 第 34 回土木計画学研究発表会講演集 (CD-ROM) (2006).

(平成 18 年 12 月 20 日受付)

(平成 19 年 2 月 9 日採録)

(担当編集委員 江口 浩二)



奥 健太 (学生会員)

1981 年生. 2004 年大阪市立大学工学部土木工学科卒業. 2006 年奈良先端科学技術大学院大学情報科学研究科博士前期課程修了. 同年同大学院同研究科博士後期課程入学, 現

在に至る. 日本データベース学会, 電子情報通信学会, ACM 各学生会員.



中島 伸介 (正会員)

奈良先端科学技術大学院大学情報科学研究科助教. 博士 (情報学). 1997 年神戸大学大学院自然科学研究科博士前期課程修了. 2004 年京都大学大学院情報科学研究科博士後期課程修了. 主に Web マイニングおよびマルチメディアコンテンツ検索の研究に従事. 日本データベース学会, IEEE CS 各会員.



宮崎 純 (正会員)

奈良先端科学技術大学院大学情報科学研究科准教授. 1992 年東京工業大学工学部情報工学科卒業. 1997 年北陸先端科学技術大学院大学情報科学研究科博士後期課程修了. 博士 (情報科学). 同大学助手を経て, 2003 年より現職. 2003 ~ 2007 年科学技術振興機構さきがけ研究員 (兼務). 2000 ~ 2001 年テキサス大学アーリントン校客員研究員. 高性能・高機能データ工学システムの研究に従事. 電子情報通信学会, 日本データベース学会, IEEE CS, ACM SIGMOD 各会員.



植村 俊亮 (フェロー)

奈良産業大学情報学部情報学科教授. 1964 年京都大学大学院工学研究科修士課程修了. 同年電気試験所 (現産業技術総合研究所). マサチューセッツ工科大学電子システム研究所客員研究員, 東京農工大学教授, 奈良先端科学技術大学院大学教授を経て, 2007 年から現職. データ工学, データベースシステムの研究に従事. 工学博士. IEEE Fellow, 電子情報通信学会フェロー.