

アイテム利用時のユーザコンテキストを考慮した 情報推薦システムの提案

関 匠吾[†] 中島 伸介[†] 張 建偉[†]

[†] 京都産業大学コンピュータ理工学部 〒603-8555 京都府京都市北区上賀茂本山

E-mail: †{g0846697,nakajima,zjw}@cc.kyoto-su.ac.jp

あらまし インターネット上には膨大な量の情報が溢れており、この膨大な量の中から特定の情報を見つけるのは安易なことではない。このような背景から、ユーザに適切な情報の提供を可能にする情報推薦技術が注目されている。従来の情報推薦に関する研究で扱われるコンテキストの多くは、ユーザ毎に嗜好学習を行うため、初めて利用するユーザもしくは利用経験が浅いユーザに対してはコールドスタート問題が発生する。そこで我々は、ユーザがどのようなコンテキストの際にどのようなアイテムを好むのか、を学習するのではなく、アイテム側の特性として、どのようなコンテキストのユーザから支持を受けているのかということを示すデータを蓄積することで、新しいユーザに対してもその時のコンテキストに応じて適切なアイテムを推薦することが可能な、アイテム利用時のユーザコンテキストを考慮した情報推薦システムの提案を行う。

キーワード 情報推薦, ユーザコンテキスト

Information Recommendation System Considering Users' Contexts of Using Items

Shogo SEKI[†], Shinsuke NAKAJIMA[†], and Jianwei ZHANG[†]

[†] Faculty of Computer Science and Engineering, Kyoto Sangyo University Motoyama, Kamigamo, Kita-ku,
Kyoto-City, Kyoto, 603-8555 Japan

E-mail: †{g0846697,nakajima,zjw}@cc.kyoto-su.ac.jp

Abstract An enormous amount of information is overflowing in the Internet, and it is not easy to find specific information from it. Therefore information recommendation technique that provides proper information to users has attracted a lot of research interests. Because the preference is usually learnt for each user in the existing information recommendation technique, the cold-start problem occurs for the users who use the system for the first time or the users with shallow usage experience. The existing system learns what items the user selects in what context, while we consider users' contexts of using items and propose an information recommendation system accumulating the data that for each item users in what context select it. Our system can recommend proper items to a new user according to the context at the time.

Key words Information recommendation, Users' contexts

1. はじめに

インターネット上には膨大な量の情報が溢れており、この膨大な量の中から特定の情報を見つけるのは安易なことではない。このような背景から、ユーザに適切な情報の提供を可能にする情報推薦技術が注目されている。最近では、この情報推薦技術に対して、ユーザの状況(コンテキスト)を考慮することで、推薦精度の向上を目指した取り組みがなされている。従来の情報推薦に関する研究で扱われるコンテキストの多くは、時

間や場所など単純なコンテキストを扱うものであるが、奥らの研究 [1] では、24 次元のコンテキストを考慮したユーザの履歴データに SVM を用いて嗜好学習データを作成し、利用することで成果を挙げている。しかし、ユーザ毎に嗜好学習を行うため、初めて利用するユーザもしくは利用経験が浅いユーザに対してはコールドスタート問題が発生する。

そこで我々は、ユーザがどのようなコンテキストの際にどのようなアイテムを好むのか、を学習するのではなく、アイテム側の特性として、どのようなコンテキストのユーザから支持を

受けているのかということを示すデータを蓄積することで、新しいユーザに対してもその時のコンテキストに応じて適切なアイテムを推薦することが可能な、アイテム利用時のユーザコンテキストを考慮した情報推薦システムの提案を行う。推薦アイテムとして飲食店を対象とした場合で考えると、1人のユーザが利用した飲食店数に比べて、1軒の飲食店に来店するユーザ数ははるかに多いと推測できるため、蓄積データの収集という観点からも提案手法の利点は大きいと考えられる。提案手法では、蓄積データとして、アイテム利用時のユーザコンテキストと利用アイテムをデータベースに蓄積する。その後、蓄積されたデータを n 次元の空間座標にプロットし、 k 近傍法を用いることで推薦に適したアイテムを選定する。

2. 関連研究と関連技術

ユーザコンテキストとして位置情報を考慮した研究としては、Setten ら [2] は、コンテキストを考慮した携帯観光アプリケーション COMPASS というサービスを提案している。また、Asthana ら [3] は、店内限定であるがユーザの位置情報を考慮して商品を提供するサービスを提案している。さらに、奥ら [4] は、観光施設など地域限定性を考慮した情報推薦方式を提案している。しかし、これらはユーザのコンテキストとして位置情報のみを考慮しているものであるため推薦精度が高いとは言えない。

ユーザの嗜好学習に関する研究として、奥ら [5] は、ユーザコンテキストごとに変化する価値判断基準に基づいたランキング手法を提案している。このシステムでは、学習データからユーザコンテキストに依存する価値判断基準モデルを使用し、そのモデルに基づいてランキングを行っている。また、協調フィルタリングを利用したものとして、Amazon.com [6] が有名である。しかし、これらはユーザの嗜好を学習するものであるため新規ユーザに対してはコールドスタート問題が発生する。また、ユーザ毎に嗜好の学習を行うため効率も悪いと考えられる。

コンテキストを評価した技術として、食ベログ [7] がコンテキストの評価分布を掲載している。この技術は、コンテキスト毎に星の数を用いて採点を行っている。しかし、コンテキスト毎であるため「夜にデート」といったように複合的に見る事が出来ない。

このように、コンテキストを考慮した研究や嗜好学習により推薦精度が向上したものはあるが、新規ユーザには精度の高い推薦が行えず、またコンテキスト毎の評価も複合的に見る事ができない。したがって、我々が提案する手法は新規性が高いと言える。

3. 提案手法

本章では、3.1 節において、システムの動作の流れを説明する。3.2 節では、推薦に用いるユーザコンテキストについて述べる。3.3 節では、推薦システムに用いる手法と検討している手法を説明する。なお、本章での推薦対象アイテムを飲食店として説明する。

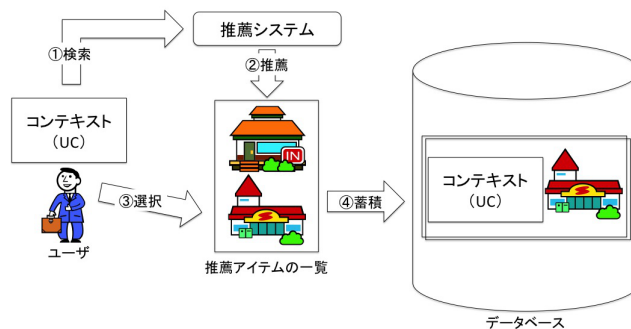


図1 データ蓄積の流れ

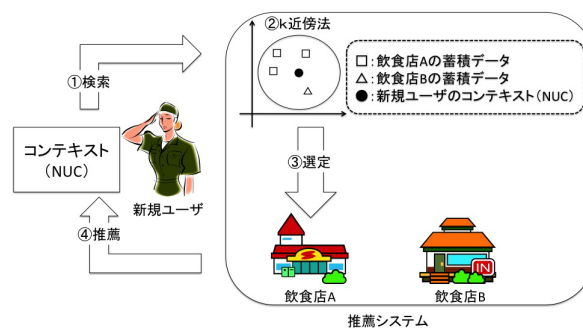


図2 アイテム推薦の流れ

3.1 システムの概要

3.1.1 データ蓄積

データの蓄積について説明するため、ユーザコンテキストによる検索からデータ蓄積の流れの例である図1を用いる。なお、ここで用いるユーザは新規ユーザでも構わないとし、コンテキストについては3.2節で述べるものを用いる。また、推薦システムについては3.1.2項で述べる。

- ① ユーザコンテキスト (UC) を取得してユーザが検索する。
- ② 推薦システムがユーザコンテキストを考慮して推薦アイテム一覧を提示する。
- ③ その推薦結果に対してユーザはアイテムを選択する。
- ④ ユーザが選んだアイテムとユーザコンテキストをデータベースへ蓄積する。

蓄積する例としては [アイテム：アイテム名，コンテキスト：UC] のようにアイテムにユーザコンテキストを持たせるイメージである。

3.1.2 アイテム推薦

アイテム推薦について説明するため、新規ユーザが利用した例である図2を用いる。

- ① ユーザコンテキスト (NUC) により新規ユーザが検索する。
- ② これまでデータベースに蓄積されてきたアイテムとユーザコンテキストを n 次元の空間座標で表す。また、その空間座標上に新規ユーザが検索に用いたコンテキスト (NUC) も表す。その後、 k 近傍法を用いて新規ユーザのコンテキスト (NUC) を中心に近傍を探す。

- ③ k 近傍法の k の個数を k=4 とした場合，飲食店 A のデータが多いことから飲食店 A を選定する．
- ④ 飲食店 A を新規ユーザへ推薦する．

また，新規ユーザの選んだアイテムとユーザコンテキスト (NUC) をデータベースへ蓄積する．

3.2 採用するパラメータ

本節では，ユーザのコンテキストとして採用するコンテキストについて述べる．

本手法では，ホットペッパー API の使用を検討しているためホットペッパーを参考に採用する．また，奥ら [1] の手法で用いていた特徴パラメータも参考にする．表 1 はアイテムの特徴で，表 2 はユーザコンテキストである．表の各項目を空間座標軸とする．しかし，経緯度については座標軸に用いないとする．経緯度の詳しい説明については 3.3.2 項で述べる．

表 1 アイテムの特徴

最寄り駅から	徒歩0分～30分
経度(東経)	122度56分01秒～153度59分11秒
緯度(北緯)	20度25分31秒～45度31分35秒
予算	1,000円～10,000円
ジャンル	居酒屋/ダイニングバー/創作料理/和食/洋食/イタリアン/フレンチ/中華/焼肉/韓国料理/アジア/各国料理/カラオケ/パーティ/バー/カクテル/ラーメン/お好み焼き/鉄板焼き/カフェ/スイーツ/その他
特徴	個室/座敷/貸切/駐車場/飲み放題/食べ放題/お子様連れ

表 2 ユーザコンテキスト

日時情報	月(2)	1月～12月
	曜日(2)	日曜日～土曜日
	時刻(2)	00:00～23:59
気象情報	天気(4)	晴れ/曇り/雨/雪
	休日(2)	休日/休日前
ユーザ情報	予算(1)	1,000円～10,000円
	年齢(1)	10歳～60歳
	経度(東経)(1)	122度56分01秒～153度59分11秒
	緯度(北緯)(1)	20度25分31秒～45度31分35秒
場面(6)		1人/家族/恋人/友人/先輩・後輩/上司・部下/ビジネス/観光/その他
人数(1)		1人～10人

3.3 推薦アイテムの選定とランキング

3.3.1 k 近傍法の利用

本手法では，ランキングにも応用でき，単純な学習アルゴリズムである k 近傍法の利用について述べる．まず，検索したユーザコンテキストを中心に空間座標上に円を作る．この円の大きさは指定した k の個数により異なる．ここでは，図 3 を例に用いるとする．なお，ここで用いる記号 ● は検索したユーザコンテキスト，▲×■ はこれまで蓄積したアイテム毎のユーザコンテキストとする．まず，ユーザが検索に用いたコンテキスト ● を中心として空間座標上に円と作る．この図の場合は k=5 とするので 5 個の蓄積したユーザコンテキストが含まれるまで円を大きくしていく．結果，5 個含んだ状態が図 3 となり，▲ が一番多いことがわかるので ▲ のアイテムを推薦する．この流れが k 近傍法である．

しかし，本手法ではただ k 近傍法により合致するものを推薦するのではなくランキングも行う．先程の図 3 を例に用いると，k=5 とした場合，5 個のデータの中で一番多いアイテム ▲ を推薦アイテムの一位として推薦する．その後，空間座標から推薦したアイテム ▲ を削除する (図 4) つづいて，また k=5 として 5 個のデータの中で一番多いアイテム × を二位として推薦する．そして，その推薦したアイテム × も空間座標から削除する．これを数回行うことで推薦アイテム一覧の提示を行う．

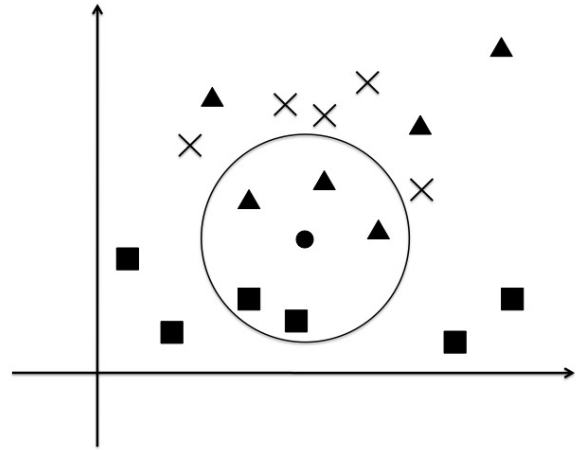


図 3 k=5 の場合の k 近傍法の例

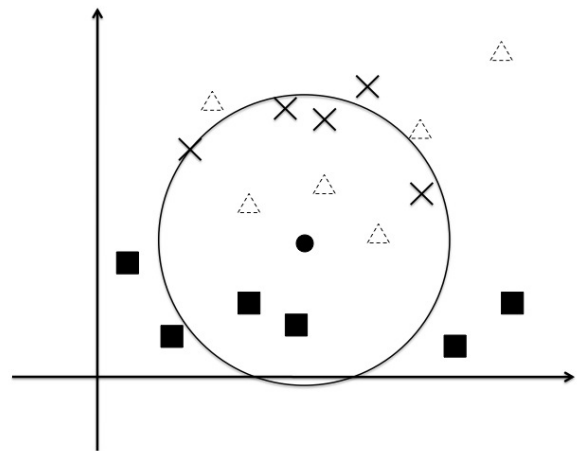


図 4 本手法での k 近傍法の利用法

3.3.2 事後フィルタリング

本手法でユーザコンテキストの経緯度を座標軸として用いない理由も含めて事後フィルタリングについて述べる．

経緯度を座標軸に用いない理由として，本手法では k 近傍法を利用するため実距離による範囲指定が不可能であるからだ．例えば，ある店 A に 100m 離れた場所からのみ人が行くとする．ここでの 100m は経緯度の計算によってわかったとする．この場合 100m の値を含むコンテキストが空間座標上へプロットされていく．その状態で，あるユーザが店 A から 10m 離れ

た場所から検索したとする．その時，k 近傍法を使用すると k 近傍法ではそのユーザに近いコンテキストが蓄積されている店を推薦するため店 A との実距離 10m の地点にいるにも関わらず 100m でのコンテキストの蓄積しか無いため推薦することが出来なくなる．そのため，我々は事後フィルタリングによって実距離を考慮することにした．

簡単な動作例として図 2 に事後フィルタリングを組み込んだ図 5 を用いる．

- ① ユーザがコンテキスト (UC) で検索を行う．
- ② k 近傍法を用いる．
- ③ 推薦アイテムを選定するが，まだユーザへ結果を提示しない．
- ④ 事後フィルタリングを行う．
- ⑤ その結果をユーザへ提示する．

この時，取得したユーザコンテキスト (UC) の経緯度はアイテムとの実距離を計算し，データベースへ蓄積する．

使用時の具体例として，ユーザコンテキストに天気雨がであることを取得して推薦アイテムを提示したとする．しかし，推薦アイテムとの実距離が遠い場合，雨の中を移動するのは困難であると考えられるので事後フィルタリングにより実距離を考慮する．また，実距離以外にもユーザの個別化についても考慮する必要があると考える．ユーザの個別化については 3.3.3 項にて述べる．

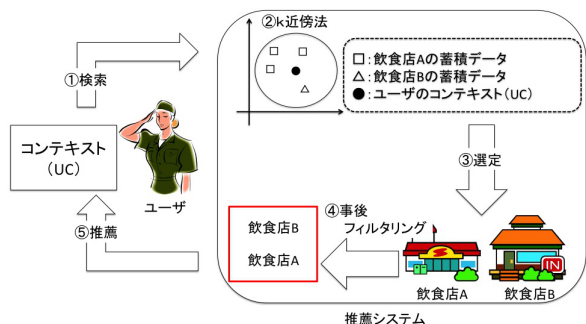


図 5 事後フィルタリングを行う場合の流れ

3.3.3 ランキングに対する個別化の検討

本システムをユーザが数回使うことによってユーザ毎にコンテキストのどの要素を重要視しているか学習する．その結果，ユーザが重みを置くコンテキストを判定し推薦の補助的なものとして考慮を検討する．

簡単な例としては，数回本システムでアイテムを検索したことがあるユーザ A が本システムを使用して検索したとする．3.3.2 項で述べた事後フィルタリングの流れにより，これまで蓄積されてきたユーザ A のコンテキストとその時選んだアイテムからどのコンテキストの要素を重要視しているかを判定する．その判定結果から予算を重要視していることがわかったとすると，今回の検索ではユーザコンテキストで取得した予算を重視して推薦結果一覧を提示する．

4. 評価実験

提案する手法の有効性を評価するために，簡易的な評価実験を行った．本実験では，アイテムを飲食店とし，ホットペッパー [8] から京都市四条河原町付近の 100 店舗を選んだ．また，使用したユーザコンテキストを 22 次元とし以下に示す．

月 (2) 時刻 (2) 天気 (4) 休日前か休日か平日 (3) 予算 (1) 年齢 (1) 場面 (8) 人数 (1)

学習フェーズの流れについて説明する．

- ① ユーザコンテキストをランダムに生成しユーザへ提示
- ② ユーザは提示されたコンテキストにおける最も気に入った飲食店を選択
- ③ コンテキストを数値に置き換えて選択された飲食店の ID を付与して DB へ蓄積

評価フェーズの流れについて説明する．

- ① 新たにランダムにコンテキストを生成
- ② システム内で①のコンテキストと蓄積データを用いて k 近傍法とランキングを行う
- ③ ユーザへ②の結果である上位 5 件と①のコンテキストを提示
- ④ ユーザは提示された飲食店毎に①のコンテキストを用いて 5 段階評価を行う

なお，この時にコンテキストを考慮しない場合の店舗ランキングを用意しておき，同様に④を行い比較対象とした．

表 3 に，本手法と比較対象の上位 5 件の平均スコア (5 段階) を提示する．コンテキストを考慮した提案手法の方が高いスコアを示した．しかし，スコアの数値が低かった理由として，評価実験を行った際のコンテキストに偏りがあったことや蓄積データが少なかったこと，飲食店の店舗数が少なかったことが挙げられる．今後は，さらに詳細な評価実験を行った結果を踏まえて，システム実装に向けた検討を行う．

表 3 ランキング上位の平均スコア

	コンテキスト無	コンテキスト有
上位 5 件の平均スコア	2.6	1

5. 実装に向けた検討

実装に向けた検討として，今回はアイテムを飲食店をして考える．

5.1 アイテムデータの収集と DB 蓄積

飲食店の特徴であるアイテムデータを集めるために，ホットペッパー [8] の API を使用する．集めたアイテムの特徴やユーザコンテキストを収集，蓄積する方法としては PHP と MySQL を用いる．集めるアイテムの特徴やユーザコンテキストに関しては 3.2 節で述べたものである．

5.2 ユーザコンテキストの取得方法

3.2節で述べたユーザコンテキストの取得方法として、経緯度が必要となるためモバイル端末でのGPSの使用を検討している。また取得した経緯度により、SimpleAPI [9] の『最寄り駅 Web サービス』を使用して最寄り駅を取得する。天気情報はlivedoorのWeather Hacks [10] が提供している『お天気 Web サービス』を使用する。その他、日時情報はPHPの関数を用いて取得する。また、場面に関してはGoogle [11] のサービスである『Google カレンダー』を用いることが出来るのではないかと考える。

5.3 インタフェースの設計

設計として、検索フォームページでは、コンテキストを自動取得とするので入力部分は予算などの限りなく少なく出来るようにする。推薦アイテム一覧ページでは、k近傍法や事後フィルタリングで一位になったアイテムから順に上から並べ、アイテムのサムネイルとアイテム名を表示する。アイテム個々のページでは、アイテムの特徴やアイテムの雰囲気であるサムネイルを表示。また、これ以外にもユーザに提示した方がいいものを検討する。

6. ま と め

本論文では、アイテム側の特性として、どのようなコンテキストのユーザから支持を受けているのかということを示すデータを蓄積することで、新しいユーザに対してもその時のコンテキストに応じて適切なアイテムを推薦することが可能な、アイテム利用時のユーザコンテキストを考慮した情報推薦システムを提案した。評価実験により提案手法の有効性を確認した。

今度の課題として、推薦システムの実装及び評価実験に基づく手法の改良を行う。

謝 辞

本研究の一部は、文部科学省科学研究費補助金若手研究(B)(課題番号:20700089)による。ここに記して謝意を表します。

文 献

- [1] 奥健太, 中島伸介, 宮崎純, 植村俊亮: 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, 情報処理学会, データベース, Vol.48, No.SIG 11(TOD 34), pp.162-176 (June 2007).
- [2] Mark van Setten, Stanislav Pokraev and Johan Koolwaaij: Context-Aware Recommendations in the Mobile Tourist Application COMPASS, Adaptive Hypermedia and Adaptive Web-Based Systems, Lecture Notes in Computer Science, Vol.LNCS 3137, pp.235-244 (2004).
- [3] Abhaya Asthana, Mark Cvavatts and Paul Krzyzanowski: An Indoor Wireless System for Personalized Shopping Assistance, Workshop on Mobile Computing Systems and Applications, IEEE Computer Society Press, pp.69-74 (1994).
- [4] 奥健太, 服部文夫: 地域限定性を考慮した情報推薦方式に関する基礎検討, WebDB Forum (2009).
- [5] 奥健太, 中島伸介, 宮崎純, 植村俊亮, 加藤博一: 情報推薦におけるユーザの価値判断基準モデルに基づくコンテキスト依存型ランキング方式, 情報処理学会論文誌, データベース, Vol.2, No.1(TOD 41), pp.57-80 (Mar 2009).

- [6] 『Amazon.com』. <http://www.amazon.com/> (2011-01-06 確認).
- [7] 『食べログ』. <http://tabelog.com/> (2011-01-06 確認).
- [8] 『ホットペッパー Web サービス』. <http://api.hotpepper.jp/> (2011-01-06 確認).
- [9] 『SimpleAPI』. <http://map.simpleapi.net/> (2011-01-06 確認).
- [10] 『Weather Hacks』. http://weather.livedoor.com/weather_hacks/ (2011-01-06 確認).
- [11] 『Google カレンダー』. <http://www.google.com/calendar/> (2011-01-06 確認).