

Want list を用いた購買支援と情報推薦

楊斐[†] 清光英成[†] 大月一弘[†] 森下淳也[†]

ユーザの購買希望を列挙した Want list を利用し、個別のユーザに適切な個別化された情報推薦を議論する。従来の情報推薦は、主にコンテンツフィルタリングと協調フィルタリングを利用して、より多くのユーザに適切な情報を提供することを目的とする。これに加えて個別のユーザが必要とする情報の推薦も求められるようになってきた。本研究は特に、インターネットを通じての商品購入に注目し、価格や納期、店舗の評価などと共にユーザの置かれている状況などに適応し、ユーザの購買欲を効率よく満足する適応型 Wantlist の枠組みを提案する。

Supporting and Recommendation for Purchasing by Using Want Lists

Fei Yang[†], Hidenari Kiyomitsu[†], Kazuhiro Ohtsuki[†], and
Junya Morishita[†]

A want list of an individual user expresses his desire for purchasing. Conventional recommender systems have been aimed to guess what their users want to purchase by using techniques from content filtering and collaborative filtering. These systems enable users to explore goods what they want implicitly. We wish, in addition, users could express their wants to the recommender systems explicitly. Our recommender system uses a want list of a user to guess some stores which he wants to buy. In this paper, we discuss about information recommendation with personalization and adaptation by using want lists. Also, we propose a framework of want list which satisfies a user's interest.

1. はじめに

Want list はユーザが欲する個別の商品やサービスを want として表現し、それらを列挙したものである。Want list にはユーザが自ら購入する能動的な want を列挙した Want list と、他人に贈与されたい want を列挙した受動的 Want list がある。近年、ネット上のショッピングモールにも欲しい物リスト、Wish list などの名称で Want list を良く見かけるようになった。Amazon の欲しい物リスト²⁾は前者、BABIESRUS の wishlist³⁾は後者である。本論文では、主に前者を扱うことにする。

Want list は個別の want や、want をある側面からまとめた want group から構成される。want group は複数の want group で構成される場合や、want と want group から構成されても良い。

一般的な購買行動ではユーザはまず、商品・サービスを同定する。次に、どこで買う（いつ買う）を決定し購入する。提案する Want list はどこで買うべきかというユーザの意思（行動）決定を支援することが目的であるため Want を以下の複数の属性で表現する。

価格、店舗評価、納期、在庫状況、ポイント、機能、
キャンペーン、新作、値下げ幅、ブランド名

本研究は、ユーザが欲する商品やサービスなどを Want list として明示的に表明することで、ユーザのすべての want の効率的な獲得を支援することを目的とする。ユーザが want を入力すると、want の属性の優先順から購入すべき店舗の候補を推薦する。ここで属性の優先順とは、価格や納期などの属性の内どの属性をより重視するかという順序を意味する。詳細は3節で述べる。本研究は want を購入すべき店舗の候補を推薦する情報推薦、個別のユーザに適するコンテンツとその見せ方を提供するパーソナライゼーション、ユーザの時間的・空間的状况に適したコンテンツとその見せ方を提供する環境適応化の分野にまたがる。従来の情報推薦はユーザが何を欲するかを推測して提供するものが主であったが、本研究は、ユーザが何を欲するかは明示的に示してもらい、どこで買いたいかを推測して提供する点が異なる。

以降、本論文は2節で関連研究について述べ、3節で Want list の表現方法と推薦度の計算方法を提案する。4節は提案方式を用いた実験とその考察であり、5節でまとめと今後の展望を示す。

2. 関連研究

電子モールなどでユーザの嗜好を分析し、ユーザ毎に興味のありそうな商品やサービスをおすすめする情報推薦サービス（レコメンド）が行われている。レコメンドはユーザにとって潜在的に好む商品などに受動的にアクセスできる利点ならびに電子モールや店舗にとって、顧客の購買率や将来的な累積利益額(lifetime value)を向上させるなどの利点がある。

情報推薦は、コンテンツフィルタリングや協調フィルタリングなどの技術が利用されている。コンテンツフィルタリングは各ユーザのプロファイルや購入履歴などからユーザの嗜好情報を蓄積して利用する。ユーザの嗜好に類似した商品やサービスが推薦される。例えば、購買履歴や閲覧履歴にある商品と類似する商品をレコメンドするサービスがこれにあたる。履歴に存在しない種類の商品の推薦が困難であり、ユーザが未知のコンテンツをレコメンドできない。嗜好情報の蓄積と管理は適合フィードバックを用いて評価、修正される。協調フィルタリングは、プロファイルや履歴が類似したユーザの集約情報を用いてユーザが属する集合の嗜好を推測する手法であり既存の情報推薦システムの多くが何らかの形で利用している。コンテンツフィルタリングに対して、ユーザが未知のコンテンツをレコメンドできる特徴を持つ。協調フィルタリングを利用した情報推薦の欠点としてコールドスタート問題1)を指摘することが少なくないが、Amazon では「いまもっともクリックされている商品」2)、楽天市場では「お問合せの多い商品」4)を表示するなどの集計値に基づいた推薦を行っており、ほぼすべての電子モールで何らかの情報推薦が使われている。

個別ユーザのプロファイルや集約情報から特定のユーザに推薦するコンテンツを決定する Web 上の例として、

- 閲覧しているコンテンツに類似したコンテンツ
- ユーザと似ているとされるユーザ集合の閲覧履歴から頻度の高いコンテンツ

を選ぶなどがある。類似度の計算はベクトル空間モデルなど、多くの方法が利用されている。情報推薦に関連する研究では、コンテンツフィルタリングにおける商品やサービスなどのアイテムの類似度、ユーザプロファイルや履歴の類似度を計算する方法ならびにその評価・検証が主な課題である。つまり、ユーザが直接的に嗜好情報を入力することなく情報推薦を享受できる高文脈(high context)なレコメンドサービスを提供しようとしている。レコメンドがユーザの推薦されたいコンテンツを表示できたかどうかを別途評価・検証する必要がある。このような高文脈のレコメンドはユーザが情報を入力する労力を要求しないことが利点とされるがその反面、ユーザはプロファイルを更新することによってどのようにレコメンドが変化するかについても関心があ

る。本研究は従来のレコメンドサービスに加えてユーザがレコメンドに何らかの関与をすることで、より効果的な情報推薦の実現を目的とする。そこで、電子モールのインターネット上の購買行動に注目し、ユーザが効率よく購買欲を満足できるような支援システムについて研究を始めた。ユーザが Want list を作成することで明示的に want を表明し、ユーザに適した購入先を推測してレコメンドすることを試みた。提案する情報推薦サービスのゴールは Want list 内のすべての want を購入させることである。また、提案方式の評価尺度はユーザが Want list を記述してからすべての want を購入するまでの時間、支払総額、重視するパラメータの順序と購入店舗とのマッチングが考えられる。本論文では主にユーザが重視するパラメータの順序と店舗が提供する商品属性とのマッチングについて実験と考察を行なっている。これを評価する尺度については、ユーザの満足度を調査するあるいは購入した店舗のリピート利用率などを候補とするが本論文では今後の展望として議論している。

3. Want list と推薦度

買い物メモのような Want list は、ユーザの思いどおりに記入できる自由度が高いものである。実用上 Want list に構造はあってもなくてもいいわけであるが、want がある程度具体的に表現できる場合は構造化表現が可能であるため半構造的な Want list の記法を 3.1 節で設計した。また、ユーザが作成した Want list を用いて商品を買うべき店舗を推測する方法を 3.2 節で提案する。ここでは、商品やサービスを提供する店舗の評価方法として店舗の各パラメータから推薦度を計算してユーザが重視するパラメータの順序を用いた推薦度の合成（総合推薦度）を求める方法について述べる。

3.1 Want list の記法

各 want は Want list 内で他の want と区別するための属性 want name、品番や型番など商品やサービスを同定するための属性 object、ユーザが重視するパラメータの順序を表現するための属性 order などからなる。Want list の記述法(図 1)の一部を示す。<want list> は 0 個以上の <want> からなる。各 <want> が一つの want を表現する。<want> は want の強度を順位表現した <priority>、商品・サービスなどを同定するための <objects>、want を獲得するために重視するパラメータの順序 <order> からなる。<order> の値は価格 p、在庫 t、評価 e を重視する順に左から並べて記述する。

一眼レフデジタルカメラの Nikon D7000 を価格>店舗の評価>在庫(pet)というパラメータの順序で want として記述した例を示す。

(want name=D7, object=Nikon D7000, order=pet)

```

    -----
    < want list > ::= (< want > ";") *
    < want > ::= want name=" <w name>","<priority>","<order>","<object>"," ...
    < w name > ::= string
    < name > ::= string
    < priority > ::= int
    < objects name > ::= string
    < order > ::= pet|ept|tpe|pte|etp|tep
    < object > ::= "object=" <object name>";" < num > ";" (< object cond > ";")? < name > | < id > ...
    < num > ::= int
    < object cond > ::= "o-cond=" (< brand > ";")? ...
    < brand > ::= "brand = "string
    < name > ::= string
    < id > ::= < ISBN > | < JAN > | < brand > < Model Number > | ...
    < ISBN > ::= int
    < JAN > ::= int
    < Model Number > ::= "model = "string
    -----
    
```

図 1: Want list の記法

3.2 推薦度の計算

ユーザの want と店舗が扱う商品やサービスとのマッチングを評価するために、価格や在庫などの商品の属性毎の推薦度を計算する方法と、属性ごとの推薦度を利用してユーザが重視するパラメータの順序を考慮した総合推薦度の計算方法について述べる。

3.2.1 属性推薦度の計算

ある want で指定された商品やサービスを扱う店舗の属性毎の推薦度を計算する。ここで、ある want で指定された商品やサービスを扱うすべての店舗から属性値を収集し、

- S_{ij} : 店舗 S_i の商品の属性 j の値
- B_j : 属性 j の最良値
- M_j : 属性 j の最頻値

とする。このとき、店舗 S_i の属性 j における推薦度 R_{sij} を

$$R_{S_{ij}} = \begin{cases} \left(\frac{M_j - S_{ij}}{M_j - B_j} \right)^2 & ((M_j - S_{ij}) \cdot (M_j - B_j) > 0) \\ -\left(\frac{M_j - S_{ij}}{M_j - B_j} \right)^2 & ((M_j - S_{ij}) \cdot (M_j - B_j) < 0) \\ 0 & M_j = B_j \end{cases}$$

とする。各店舗の商品の各属性の最頻値と最良値を利用することで、ある商品の店舗毎の各属性の推薦度を計算する。上式は店舗間の属性の値の差を増幅するために二乗をした。推薦度の最大値は 1 である。総合推薦度が正の値をとるのは、属性値が最頻値と最良値の間に存在するときである。最頻値と最良値が同じ場合は 0 である。

3.2.2 総合推薦度

前節で示した各属性の推薦度を利用してユーザの want と店舗が扱う商品やサービスとのマッチングを評価する総合推薦度を以下のように定義した。

- R_{S_i} : 店舗 S_i の商品における総合推薦度
- S_{ik} : 店舗 S_i の商品の k 属性の値
- w_k : 店舗 S_i の商品の k 属性に対応する重み

$$R_{S_i} = \sum_{k=0}^{n-1} S_{ik} \cdot w_k$$

(ただし、属性に付与された重み $w = \{w_0, \dots, w_{n-1}\}$ は $w_0 + \dots + w_{n-1} = 1$)

我々は、ユーザの want と店舗が扱う商品やサービスの各属性の合計を 1 つの総合推薦度と捉えるため、店舗の商品属性の重みの合計を 1 とする。総合推薦度の値の範囲は 1 以下とする。次節で重み付けに AHP(階層分析法)5)6)を利用して実験を行った。

4. 実験

4.1 データセット

価格比較サイトからある商品の価格、在庫、評価を収集して提案方式を適用した。使用したデータセットを表 1 に示す。価格は送料込みと送料別が混在していたため、送料込みの価格に統一した。在庫の表記は△が問合せ（在庫の有無は不明）、○は在庫あり、数値は待ち日数をそれぞれ表している。便宜上、在庫ありは 1、問合せは 0、待ち日数はその逆数をそれぞれ属性値とした。評価は過去 6 ヶ月の店舗の評価（注文のし易さ、連絡・対応、納期、発送・梱包、リピート希望）の顧客満足度を 0~1 の尺度で表したものである。店舗数は 22 店舗である。

表 1: 実験に使用したデータセット (kakaku.com, May 9, 2011)

店名	価格	在庫	評価	店名	価格	在庫	評価
イトレンド	230600	△	0.9	ラディカルベース	247280	△	0.96
digicame-online	230600	△	0.96	EC-JOY	248815	△	0.8
富士カメラ	230600	○	0.96	中京カメラ	253000	△	0
BONZ	231230	○	1	TSUKUMO	253491	△	0.84
カデンオンライン	230800	○	0.96	カメラのキタムラ	255000	△	0.64
カメヨシ	237440	3	1	カメラ会館	255940	○	0.9
Caravan YU	237400	10	0.86	ビックカメラ.com	271400	△	0.62
Joshin web	240800	10	0.64	ソフマップ.com	271400	△	0.56
三星カメラ	246100	○	0.9	エディオン e.com	283500	△	0.8
カメラドーム	246100	○	0.94	アライカメラ	283980	△	0.96
Amazon.co.jp	246703	○	0.66	ムラウチ	285552	5	0.64

4.2 属性推薦度

実験データセットの各属性について推薦度を計算した結果を表 2 に示す。価格属性の最悪値は 285552、最良値は 230600 であった。これを 5 階級 (230000~242000, 242001~254000, 254001~266000, 266001~278000, 278001~290000) 分けた、それぞれの度数は 8, 7, 2, 2, 3 であった。最頻階級が 230000~242000 であるため 242000 を最頻値とした。他の属性の最頻値についても同様に求めた。

図 2 に各店舗の価格推薦度をグラフ表現した。属性値が最頻値と等しい場合推薦度は 0 となり、最頻値よりも悪い場合は推薦度が加速的に悪化する。実験データセットの価格属性値は最良値を含む階級が最頻階級であり、度数分布が散乱していないため、推薦度の最低が -1 を下回っていない。図 3, 4 に在庫推薦度と評価推薦度のグラフをそれぞれ示した。属性推薦度を説明するために便宜上、推薦度に基づいて店舗を整理している。評価の最良値は 1、最悪値は 0 を除いて 0.56、最頻値は 0.912、最頻階級の度数は 13 であった。最良値が最頻階級に属し、階級幅が 0.088 と狭隘かつ最頻階級の度数が多いために評価の属性値が 0.9 のとき -0.0186, 0.8 のとき -1.61983, 0.64 のとき -9.55372 と差異増幅の効果が大きい。

表 2: 推薦度

店名	価格	在庫	評価	店名	価格	在庫	評価
イトレンド	1.00000	-0.06250	-0.01860	ラディカルベース	0.18719	-0.06250	0.29752
digicame-online	1.00000	-0.06250	0.29752	EC-JOY	0.14474	-0.06250	-1.61983
富士カメラ	1.00000	1.00000	0.29752	中京カメラ	0.05669	-0.06250	-107.40496
BONZ	0.95760	1.00000	1.00000	TSUKUMO	0.04902	-0.06250	-0.66942
カデンオンライン	0.98644	1.00000	0.29752	カメラのキタムラ	0.02892	-0.06250	-9.55372
カメヨシ	0.58882	0.02778	1.00000	カメラ会館	0.01907	1.00000	-0.01860
Caravan YU	0.59091	-0.01563	-0.34917	ビックカメラ.com	-0.15035	-0.06250	-11.01033
Joshin web	0.42649	-0.01563	-9.55372	ソフマップ.com	-0.15035	-0.06250	-16.00000
三星カメラ	0.22353	1.00000	-0.01860	エディオン e.com	-0.63891	-0.06250	-1.61983
カメラドーム	0.22353	1.00000	0.10124	アライカメラ	-0.66528	-0.06250	0.29752
Amazon.co.jp	0.20456	1.00000	-8.20041	ムラウチ	-0.75536	0.00000	-9.55372

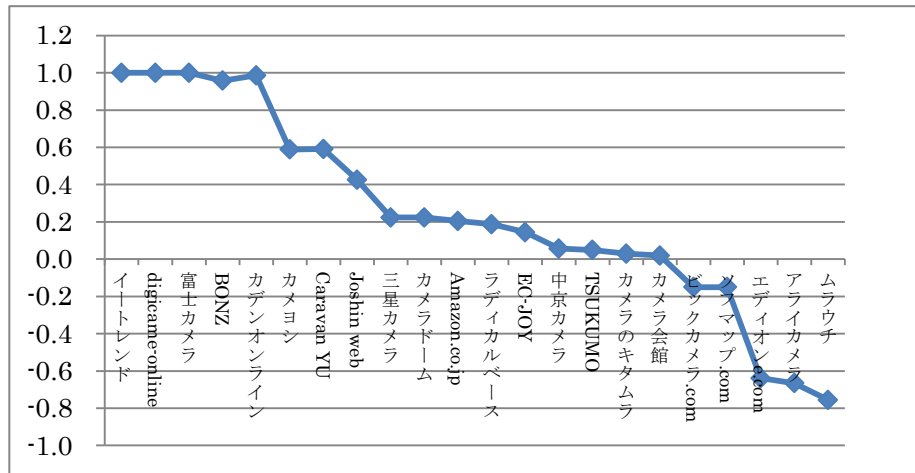


図 2: 価格推薦度

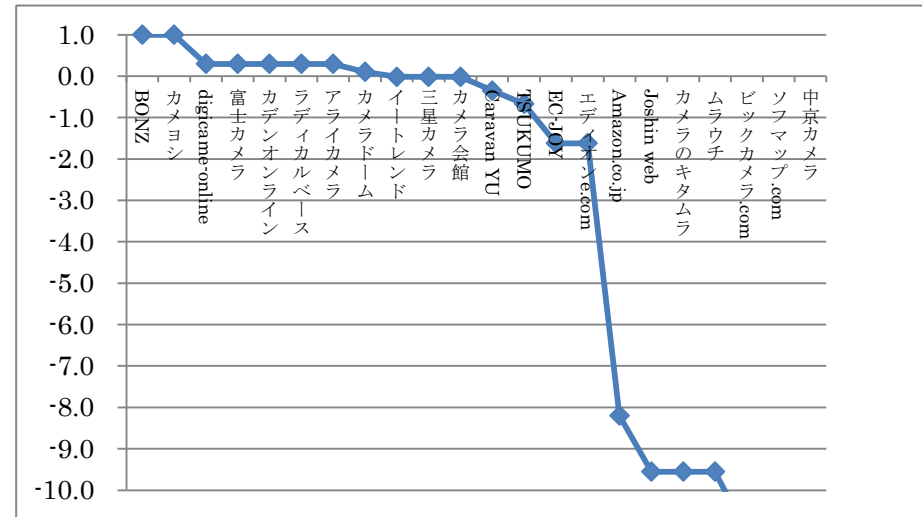


図 4: 評価推薦度

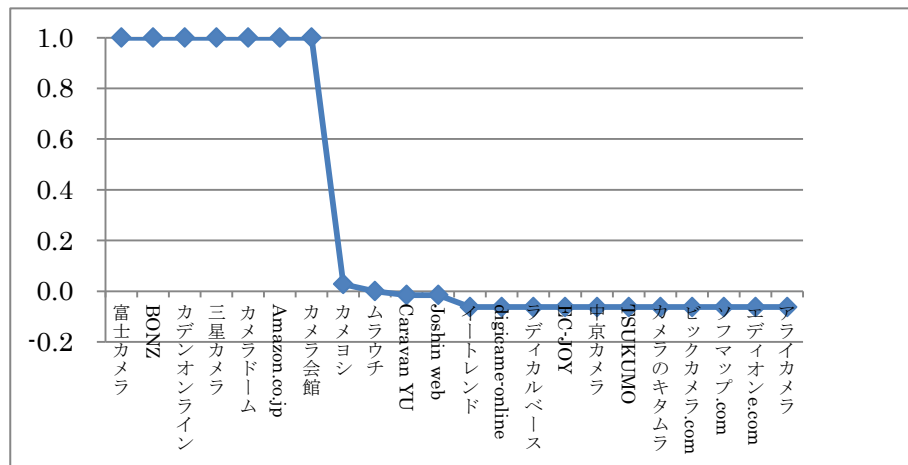


図 3: 在庫推薦度

4.3 総合推薦度

表2の各属性推薦度と want のパラメータの順序から実験データセットの各店舗について総合推薦度を求めた (表 3)。実験データセットの属性数が 3 であるからパラメータの順序は 6 種類である。

価格(p)>納期(t)>評価(e)のパラメータ順序を表す図 3 では, 上位 3 店舗の推薦度が他店舗と比較して良好であるため Recommend できる。在庫>評価>価格のグラフと評価>価格>在庫のグラフでは, パラメータの順序により推薦度の順位が変化していることがわかる。在庫>価格>評価のグラフでは, 他のパラメータの順序では上位にならない店舗が上位に現れている。ユーザが want に重視するパラメータの順序を記述することで, Recommend するコンテンツの選択や Recommend する順番を変化させることが期待できる。

表 3: 総合推薦度

店名	pte	tep	ept	pet	tpe	etp
イトレンド	0.62394	0.02855	0.26642	0.63311	0.24144	0.04436
digicame-online	0.64733	0.11801	0.46968	0.72257	0.26483	0.24762
富士カメラ	0.94802	0.80120	0.54831	0.80120	0.94802	0.54831
BONZ	0.93090	0.99205	0.96959	0.93090	0.96959	0.99205
カデンオンライン	0.92565	0.79862	0.53846	0.77883	0.93817	0.54573
カメヨシ	0.18474	0.31270	0.69034	0.38794	0.13714	0.66270
Caravan YU	0.07443	-0.09681	-0.17960	0.00472	0.01019	-0.21689
Joshin web	-0.70427	-2.71293	-6.14106	-2.69773	-0.71389	-6.14664
三星カメラ	0.19845	0.62817	0.02544	-0.01443	0.60502	0.26147
カメラドーム	0.20732	0.66208	0.10249	0.01948	0.61389	0.33853
Amazon.co.jp	-0.43326	-1.69031	-5.24703	-2.35615	-0.01200	-5.00246
ラディカルベース	-0.13360	0.02814	0.12597	-0.05836	-0.07888	0.15774
EC-JOY	-0.36735	-0.52505	-1.14732	-0.69283	-0.26119	-1.08569
中京カメラ	-8.56432	-30.50469	-69.32950	-30.99890	-8.25164	-69.14797
TSUKUMO	-0.72053	-0.30482	-0.72260	-0.84738	-0.37726	-0.52331
カメラのキタムラ	-1.56082	-2.84012	-6.51568	-3.54448	-1.11518	-6.25696
カメラ会館	-0.67983	0.52709	-0.36111	-0.89271	0.21847	0.16039
ビックカメラ.com	-5.10903	-3.64828	-8.96649	-7.39712	-2.73718	-7.58950
ソフマップ.com	-5.47826	-5.06036	-12.17485	-8.80920	-3.10641	-10.79786
エディオン e.com	-8.65869	-1.47926	-4.79654	-8.98417	-3.91041	-2.03990
アライカメラ	-8.71506	-0.95946	-3.65094	-8.63982	-3.85579	-0.82986
ムラウチ	-10.09161	-3.78374	-10.27345	-12.08834	-4.83738	-7.22308

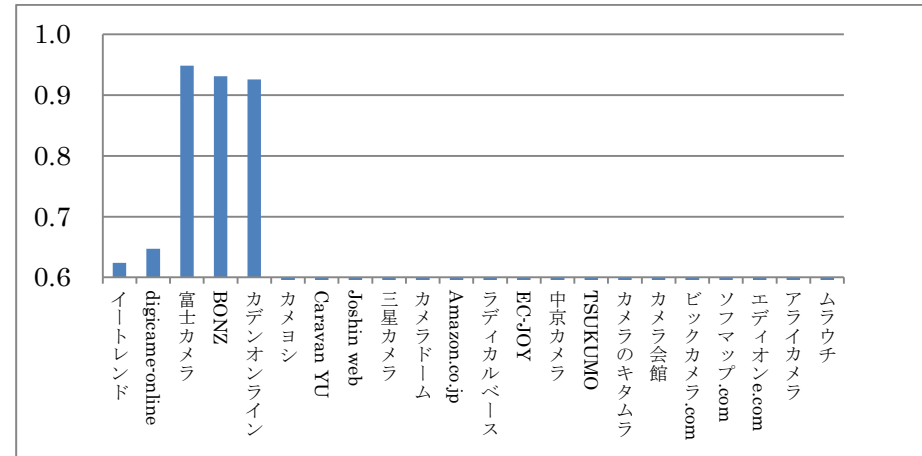


図 5: 総合推薦度 (価格>在庫>評価)

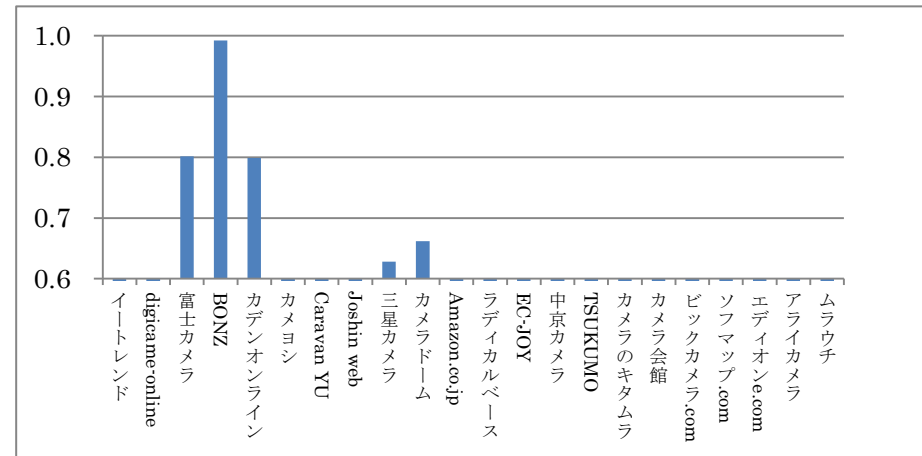


図 6: 総合推薦度 (在庫>評価>価格)

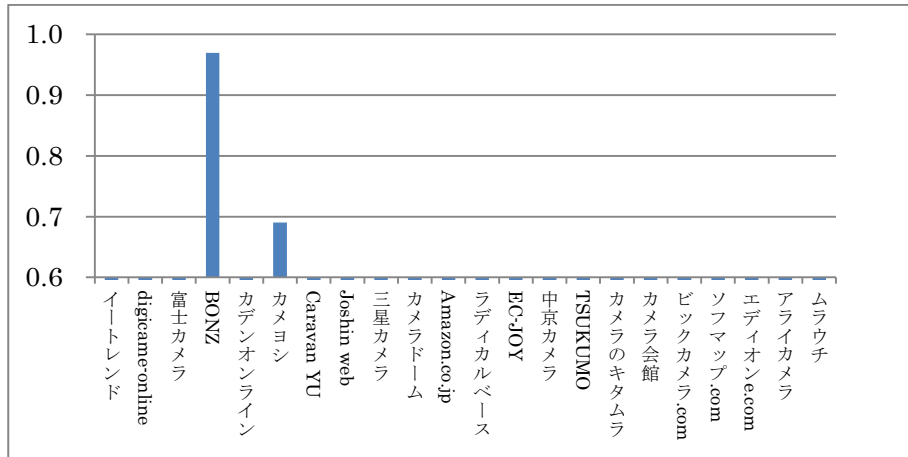


図 7: 総合推薦度 (在庫>価格>評価)

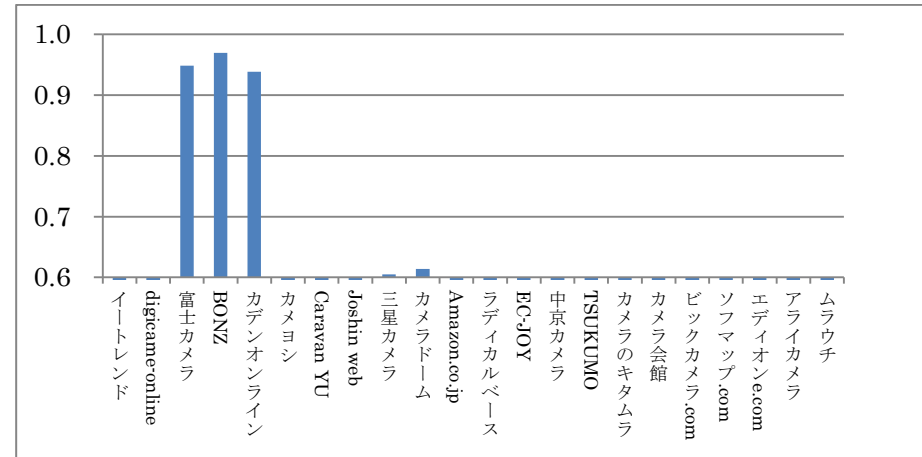


図 9: 総合推薦度 (在庫>価格>評価)

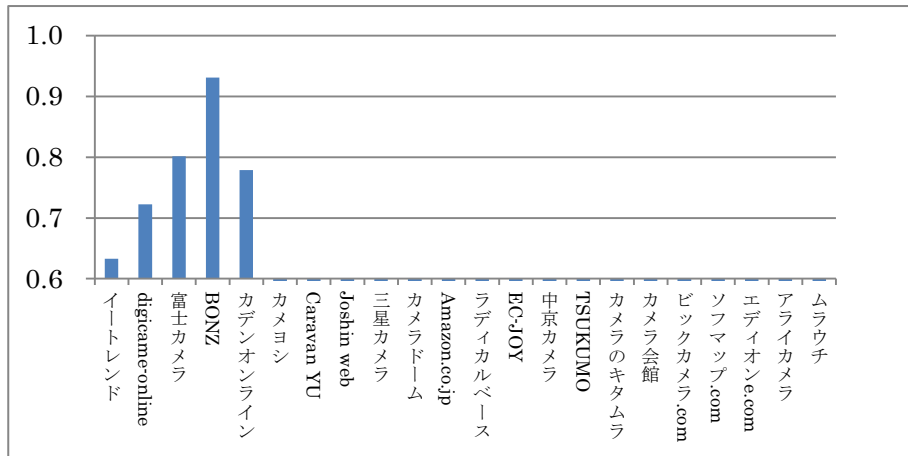


図 8: 総合推薦度 (価格>評価>在庫)

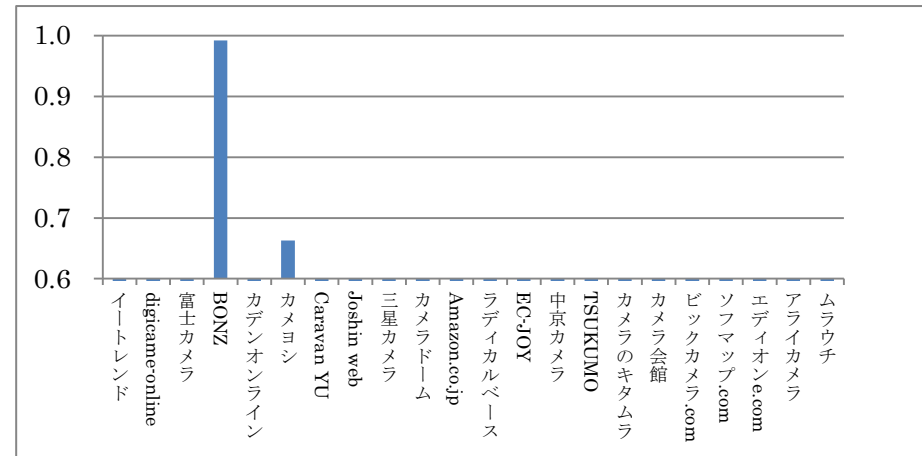


図 10: 総合推薦度 (評価>在庫>価格)

5. まとめ

従来の情報推薦サービスはユーザが直接嗜好情報を入力することなく個別のユーザが潜在的に好むと推測できるコンテンツを提供する高文脈のサービスを目指してきた。本研究はこれに加えてユーザが何らかの関与をすることで、より効果的な情報推薦サービスが出来るのではないかというアイデアから出発した。具体的なアプリケーションとして電子モールのインターネット上の購買行動に注目し、ユーザが効率よく購買欲を満足できるような支援システムについて研究を始めた。

本論文ではユーザが Want list を直接的に作成することで明示的に want を表明し、ユーザの要望を最も満たす購入先を推測してレコメンドすることを試みた。まず、原始的な want を容易に記述できるように Want list の簡単な syntax を定義した。さらに、want に対応する商品やサービスを提供する店舗から価格や納期などの属性値を収集し、どの店舗をユーザに推薦すべきかを判別するための方法を考察した。具体的には、各属性の最頻値を利用して、ある店舗の属性値が他の店舗の属性値とどのような優劣の関係にあるのかを、直感的には、少し大きく差異の表現ができるように属性推薦度の計算式を工夫した。最頻値を普通と捉えて、属性値が普通ならば推薦度が 0 付近、最も良いならば 1、普通より劣っているならばユーザの目に触れないように推薦度を加速的に不の方向に増幅させることを意図した。次に、計算した各属性推薦度とユーザが want に指定した重視するパラメータの順序を用いて、ユーザの店舗に対する要求と店舗が提供するサービスとのマッチングを図った。実験結果として、異なるパラメータの順序を与えることで推薦できる店舗が変化する事が確認できた。

実験データセットでは店舗数が 22 であったが、商品やサービスによっては扱うべき店舗数が 100 以上になる場合も少なくない。ユーザに提供するコンテンツの数としては優良な 3~10 店舗程度が良いのではないかと考えられる。また、複数の want 間の優先度、つまり、ユーザが欲しい順番を記述することも想定しているが、購入する順序と必ずしも一致しない。ユーザが重視するパラメータの順序から総合推薦度を計算する我々の方式は何らかの拡張を経てユーザがお買い得な商品を提供する店舗を推薦するあるいは、購入する順序を決定する支援が可能になると期待できる。

本論文で提案した属性推薦度の計算方法は実験データセットではある程度の有用性を確認できたが、極端な属性値を持つ商品やサービス集合ではそのような値を除外するなどの前処理が必要な場合があった。属性値を階級に分割した時の階級の度数が 2 つ以上のピークを持つような場合にも対応できるような方法も想定しておく必要がある。総合推薦度の計算では、推薦度が広く分布する属性と狭い範囲に分布する属性の組み合わせで導出される推薦度の妥当性の検証も必要と考えている。

本研究の今後の展望として、上述の検討課題に加えて、関連する複数の want をグループ化する、グループ化された want の同時購入や順次購入におけるグループの分割や

購入順序の推薦があげられる。購入順序は want 間の関連性に基づいて A は B を所有していても利用できかつ、B は A を所有していなければ利用できないならば A から購入を促すというような応用が考えられる。

参考文献

- 1) 土方嘉徳, 神嶋敏弘, 市川裕介, 河合由起子, 村上知子, 小野智弘, 本村陽一, 麻生英樹, 乾孝司, 奥村学, 金山博, ” 利用者の好みをとらえ活かす-嗜好抽出技術の最前線-,” 情報処理, Vol. 48, No. 9, pp. 955-107, 2007.
- 2) Amazon.co.jp, <http://www.amazon.co.jp/>
- 3) ベビーザラス, <http://www.toysrus.co.jp/b/>
- 4) 楽天市場, <http://www.rakuten.co.jp/>
- 5) 刀根 薫: 「ゲーム感覚意志決定法」, 日科技連, 1986.
- 6) 刀根 薫, 真鍋 龍太郎: 「AHP 事例集」, 日科技連, 1990.