

Wantlist を用いた商品検索のための演算法考察 - 店舗集合候補提示のための演算を例として -

清光 英成¹

概要: ユーザが購入したい商品やサービスなどを明示的に Want list として表明することにより効率的な商品購入支援を目的とする。現状の商品検索は、単一の商品に対してある側面（価格や評価など）から整列することができる。しかしながら、複数種類の商品を同じ店舗で購入したい、あるいは指定した店舗数で購入したいといった要求を満足するような整列や検索結果の表示方法は普及していない。そこで、本稿はこのような要求に応えるための演算法を考察する。

キーワード: Want list, 情報推薦, 関係代数, SQL

A Calculation Method Consideration for Product Search - Example Operations for Presenting the Candidates of Store Sets by Using Want List -

HIDENARI KIYOMITSU¹

Abstract: We have discussed about a calculation method for product search by using want lists. A want list of an individual user expresses his desire for purchasing. Users' explicit desire for purchasing should be used widely. Conventional product search engines can return lists of individual pair of item and store. Also, they provide some useful presentations about sort by price, store rating and so on. However, it has not become popular that system respond a demand that a user wants to buy more than one items at the same store, or wants to buy items from his specified numbers of stores. In this paper, we report a consideration about a calculation method to adapt the demands above.

Keywords: Want list, Information Recommendation, Relational Algebra, SQL

1. はじめに

WWW 上のショッピングモールには、ユーザが購入したい商品を能動的に見つける仕組みが充実している。例えば、カテゴリやジャンルといったディレクトリ型の構造をリンクをたどることで分類上の絞り込みができる。また、全文検索技術の応用によりキーワードや商品名の一部、俗称などで商品を特定することも容易になっている。これに加えて、ユーザが潜在的に購入したいであろう商品と受動的に出会う仕組みも普及しつつある。ユーザが強い購入意

思を持たないウィンドウショッピングでは、表示する商品や陳列方法の選択により訴求や告知といった店舗側の意図を反映する手法が試されている。ユーザのプロファイルや購入履歴、閲覧履歴などを積極的に利用した情報推薦手法による商品提示も珍しくなくなってきた。

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が Want list である。Want list にはユーザが能動的に 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。ユーザが購入を検討する商品を選択してブックマークが設定されると商品と店舗の組となる。この組を保存しておく仕組みが多くのショッピングモールで「お気に入り」、

¹ 神戸大学大学院国際文化学研究所
Graduated School of Intercultural Studies, Kobe University

「欲しい物リスト」, 「Wish list」などの名称で提供されている。Amazon の欲しい物リストは前者, BABIESRUS の wishlist は後者である。本研究では, 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購入すべき商品・サービスを同定する。次に, 購入店舗と購入時期を決定し購入する。本研究で提案する 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を目的としている。本研究は,

- (1) Wantlist の構造と記法
 - 商品列挙とグループ化
 - 優先順位
- (2) 店舗推薦
 - 店舗評価方法
 - 候補の提示方法
- (3) 広告表示支援
 - Wantlist 内の商品との関連
 - 協調的手法の応用

を当面の課題とし, 本稿では(2)の第2項目を中心に考察する。第2節で商品検索と情報推薦分野の関連研究を紹介し, 本研究の位置づけを明らかにする。第3節で店舗検索における演算を関係代数の演算子と対応付け, 第4節で店舗検索機能を実現するための演算法について整理並びに考察する。第5節はまとめである。

2. 関連研究

2.1 商品検索

ユーザが所望の商品を見つけるために, 多くのショッピングモールやサイトが商品をジャンルやカテゴリで分類して階層化したディレクトリと商品名や通称, 関連語をもとに全文検索できるサーチという二つの方法を提供している。ディレクトリとサーチはそれぞれ単体でも組み合わせでも利用することができる。ディレクトリはユーザが購入する商品に対して詳細な特定ができていない場合に効率的である。ユーザは購入目的を満たす商品を見つけるために商品情報の収集や条件の列挙・整理をディレクトリを移動しながら行うことができる。そして適切なディレクトリの位置から商品購入のための明示的な意図を表現するために検索キーワードを入力して能動的に商品と店舗の組の集合を得る。

全文検索はサーチを実現する一技術であり広く普及している。全文検索は言語を利用した検索手法であるためユーザの要求を単語あるいはフレーズで表現できる場合は最も有効な手法と考えられる。しかしながら, ファッションなどのカテゴリでは感性表現による検索も必要とされており, 興味深い研究が行なわれている。商品の色に基づい

て並べ替えたり, 感性語を利用してユーザとインタラク션을しながら購入すべき商品候補集合を得るものである。Shopping@nifty では商品を色で整列することができた[1]。村上らはインタラクションデザインを商品検索に導入し, 検索中に柔軟に変化するユーザの感性尺度に追随しながら画像から商品を選択する手法を提案している[2]。インタラクションデザインは認知心理学, 社会学, 経済学などの知見をユーザとの対話方法に援用したものである。

商品検索で重要な要素の一つに商品と店舗との組のリスト表示がある。ショッピングモールのディレクトリやサーチは, 単なる商品検索のツールではなく, 価格や店舗評価(レビューのスコアなど), 商品登録時期(新着・定番)などに基づきユーザの購買行動を促すべく整列されている。整列の方法は数多あるが, ユーザが購入すべき商品を特定したならば, 次に決定すべきは購入店舗と購入時期である。本研究は, ユーザが購入すべき商品を決定し want list に追加した以後の商品購入支援について議論する。

2.2 情報推薦

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。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[3]。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 [4]．

3. 店舗検索

ユーザが購入すべき商品を特定し want list を作成すると，want list 内の商品を購入する店舗を決定するフェーズへ移行する．want list 内の商品を購入する店舗候補は，ユーザの要求を次元毎に数値化することにより評価されて提示される．ここでいう次元とは，価格，商品レビュー数，商品レビューの平均スコア，店舗レビュー数，店舗レビューの平均スコアなどユーザの要求とを比較するための軸である．店舗の評価値は購入店舗候補の整列に利用される．整列方法は本研究の課題の一つであるが，本稿では扱わない．

商品購入支援のための演算法についての本質的な議論をするために便宜上，want list 内の商品 ID とショッピングモールの商品 ID は同一のドメイン (domain) の要素とする．ユーザの want list のリレーション表現 R は商品 ID (item_ID) を属性として含んでいる．ショッピングモールの商品情報を記録したリレーション S に商品 ID (item_id) と店舗 ID (store_id) の二つの属性が含まれている．このとき， R の射影 $R[item_ID]$ と S の射影 $S[item_ID, store_ID]$ との自然結合

$$S[item_ID, store_ID] \bowtie R[item_ID] \quad (1)$$

が，可能な want list 内の商品と購入する店舗との組を表現するリレーションである．

ユーザの want list 内には複数の購入予定商品が登録されている．購入店舗決定に関するパターンを整理すると

- (i) m 個の商品を単一の店舗から購入
- (ii) m 個の商品を n 店舗から購入
- (iii) m 個の商品をできるだけ多くの店舗から購入

が考えられる．ここで m は，ユーザの want list 中の商品の個数ではなく同時あるいは比較的近い時期に購入する商品の個数である．(i) m 個の商品を単一の店舗から購入する動機は，送料や支払い時の手数料を節約することなどが考えられる．特定の購入金額以上は送料の減額あるいは免除，支払いに価格手数料の優遇という特典が用意されている店舗は少なくない．また，店舗側も客単価向上や注文あたりのコスト削減といったメリットがあり，普及したビジネスモデルでもある．直観的には R の射影 $R[item_ID]$ と S の射影 $S[item_ID, store_ID]$ との商が購入店舗候補を表現すると考えられるが，詳細は次節で考察することにする．

(ii) m 個の商品を n 店舗から購入する動機は，商品購入の総支出を最小化することなどが考えられる．各店舗での購入にかかる費用は商品単価と購入数量との積を商品ごとに和をとった商品の合計金額，取扱手数料や送料，支払いにかかわる手数料などとの合計である．各店舗での購入金額の和が商品購入の総支出となる．「何をどこで買えば最も安く済むか」という本質的な要求であるが，今日のネットショッピングでは購入店舗数に応じて還元される特典が変化するサービスも存在している．つまり，従来の総支出の最小化に加えて (支払金額-還元額) の最小化という要求に答える必要がある．

(iii) m 個の商品をできるだけ多くの店舗から購入する動機は，それによる何らかのメリットがあることである．これは， m 個の商品を m 店舗から購入するという (ii) の特化問題であるため，(ii) の解決を先に行い，(iii) 固有の問題が明らかになり次第着手することになっている．

4. 演算法考察

want list 内の同時に購入する m 個の商品を単一の店舗から購入するために， m 個の商品を同時に扱っている店舗を検索する．第 3 節と同様に，ユーザの want list のリレーション表現 R は商品 ID (item_ID) を属性として含んでおり，ショッピングモールの商品情報を記録したリレーション S に商品 ID (item_id) と店舗 ID (store_id) の二つの属性が含まれている． R から同時に購入する m 個の商品を表現するタプルのみを抽出したリレーションを R' とする．このとき， S の射影 $S[item_ID, store_ID]$ と R' の射影 $R'[item_ID]$ との商

$$(S[item_ID, store_ID]) \div (R'[item_ID]) \quad (2)$$

が，購入店舗候補集合を表現するリレーションである．この SQL 表現は以下ようになる．一行目の DISTINCT は，ある店舗が同一の商品を異なる条件で販売している場合に store_ID が重複して表示されないために必要である．

商の SQL 例

```
SELECT DISTINCT SX.store_ID
FROM S AS SX
WHERE NOT EXISTS
  (SELECT R'.item_ID
   FROM R'
   WHERE NOT EXISTS
     (SELECT SY.store_ID
      FROM S AS SY
      WHERE SY.store_ID = SX.store_ID
      AND SY.item_ID = R'.item_ID))
```

m 個の商品を n 店舗から購入するための演算を考える． $m \geq n$ である．商品をどの店舗で購入するかを表現

するために属性に購入する商品と対応する商品 ID を ($item_1, \dots, item_m$) 並べ、各属性のドメインを店舗 ID とするリレーション T をリレーション R とリレーション S から導出する。 $item_i (1 \leq i \leq m)$ を扱う店舗を表現するリレーションは、制限と射影

$$(S[item_id = 'item_i'])[store_ID] \quad (3)$$

によって導出される。これは、リレーション T の射影 $T[item_i]$ である。リレーション T は $T[item_1], \dots, T[item_m]$ の直積

$$T[item_1] \times \dots \times T[item_m] \quad (4)$$

で求められる。

ここで、 m も n も任意の整数である。つまり、何個の商品を何店舗で購入するのかという特定がなされていないので最も制約のない状態である。実際の応用を考慮すれば、何店舗で買うかは指定しないが店舗と商品の組み合わせの組を合計金額で昇順に整列する例が考えられる。

— 合計金額で整列する SQL 例 —

```
SELECT S1.store as Item_1, S2.store as Item_2,
       S3.store as Item_3, S4.store as Item_4,
       S5.store as Item_5,
       S0.price+S1.price+S2.price+S3.price
       +S4.price as total
FROM S S1, S2, S3, S4, S5
WHERE S1.item = 'item_1'
      AND S2.item = 'item_2'
      AND S3.item = 'item_3'
      AND S4.item = 'item_4'
      AND S5.item = 'item_5'
ORDER BY total
```

— 上記 SQL に追加する制限 (重複店舗なし) —

```
AND S1.store_ID <> S2.store_ID
AND S1.store_ID <> S3.store_ID
AND S1.store_ID <> S4.store_ID
AND S1.store_ID <> S5.store_ID
AND S2.store_ID <> S3.store_ID
AND S2.store_ID <> S4.store_ID
AND S2.store_ID <> S5.store_ID
AND S3.store_ID <> S4.store_ID
AND S3.store_ID <> S5.store_ID
AND S4.store_ID <> S5.store_ID
```

購入店舗が重複しない、つまり $m=n$ の場合はタプル内で同じ店舗 ID が使われないように制限

$$T[item_i <> item_j](i \neq j) \quad (5)$$

によって条件を満たすタプルのみを抽出する。SQL では、WHERE 句で制限することになる。

m 個の商品を店舗数を指定せずに購入する応用を実現する代数表現も SQL 表現も記述することができた。また、 m 個の商品を m 店舗で購入する応用を実現する代数表現も SQL 表現も同様に示した。

それでは、 m 個の商品をユーザが指定した店舗数から購入する要求に応える代数表現や平易な SQL はどのように記述するのであろうか。リレーション T の各タプル中の店舗 ID の重複数を m から減算することで利用する方法やタプルを集合表現して集合の要素数を店舗数とすることが考えられる。リレーション T の次数はのべ店舗数である。タプル中の実店舗数を表現する演算子を the_number としてリレーション T を拡張

$$EXTEND TADD the_number AS NoS \quad (6)$$

してリレーション T' を得る。これにより、リレーション T' の属性 NoS (the Number of Stores) に対する制約

$$T'[Nos \theta p] \quad (7)$$

で m 個の商品をユーザが指定した店舗数から購入する要求を満たすリレーションが得られる。 θ は一般的な比較演算子 ($<$, \leq , $=$, $>$, \geq , $<>$) である。

5. まとめ

ユーザが購入したい商品を want list に登録し、ユーザ所望の条件で購入支援を行うための演算法を考察した。本稿では特に、 m 個の商品を n 店舗で購入するという条件を満足するための演算方法を関係代数と SQL を用いて示し、必要となる機能を明らかにした。今回提案のみに終わってしまった演算子 the_number の形式的な定義と SQL 表現は喫緊の課題として取り組む。

参考文献

- [1] 椎谷秀一, 遠藤進, 上原裕介, 増本大器, 長田茂美, 画像の色特徴を用いた商品検索サービス, 情報処理学会研究報告グループウェアとネットワークサービス, Vol. 2001, No. 6(2001-GN-042), pp. 37-42, 2001.
- [2] 村上裕一, 中村真吾, 橋本周司, インタラクショナルデザインを取り入れた感性商品検索システム, 第 74 回全国大会講演論文集, Vol. 2012, No. 1, pp. 641-643, 2012.
- [3] 土方嘉徳, 神鳥敏弘, 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.
- [4] 楊斐, 清光 英成, 大月 一弘, 森下 淳也, Want list を用いた購買支援と情報推薦, 情報処理学会研究報告データベースシステム (DBS), Vol. 2011-DBS-153, No. 15. pp. 1-8, 2011.