

携帯電話通信販売における商品レコメンドシステム

Recommendation System of Merchandise for Cellular Phone Direct Marketing

関口 雄太郎 宮田 秀明 田中 謙司 藤田 健 田村 泰一
Yutaro Sekiguchi, Hideaki Miyata, Kenji Tanaka, Ken Fujita, Tamura Taichi

東京大学工学系研究科
School of Engineering, The University of Tokyo

携帯電話を利用した通信販売において商品提案が有効である。なぜならば、ユーザーごとの購買データを収集して分析することで、提案する商品をユーザー一人一人に対して選択でき、また、通販の個人ページやメールを通して商品提案をユーザーに直接提示できるからだ。一方で、携帯端末の計算能力や記憶容量は PC に比べて劣るため、現行の PC 通販の商品提案と比較してシンプルなアルゴリズムが求められる。我々は、購買データを利用してユーザーのグルーピングと購買傾向の分析を行うことにより、計算量が少なく精度の高いアルゴリズムを提案する。また、アルゴリズムを実際の購買データに適用して有効性を検証する。

1. 序言

通販はユーザーごとの購買データを取得できるため、個々のユーザーの購買傾向を把握することができる。また、個人ページやメールを通して個々のユーザーにアクセスすることができる。これらの理由から、通販においては個々のユーザーに対応した商品提案が有効である。本稿では商品提案をレコメンドと呼ぶことにする。

レコメンドシステムを携帯電話に搭載する場合、携帯電話は演算能力や記憶容量に限界があるため、シンプルなアルゴリズムが求められる。ユーザーの負担を減らすレコメンドシステムの研究[中村 2007]は進んでいるが、携帯端末向けの、ハードの負担を減らすレコメンドシステムの研究は行われていない。

本稿では携帯電話における使用にも耐え得る、シンプルかつ効果の高いレコメンドアルゴリズムを提案し、実購買データに適用して効果を検証する。なお、分析に用いる購買データの関係上、BtoB 携帯電話通販のレコメンドシステムを扱う。

顧客のグルーピング方法は頻繁に購入するキー商品によるグルーピング・登録業種によるグルーピング・発注するチャンネルによるグルーピングなどが考えられる。レコメンドする商品のタイプは、定番の売れ筋商品・最近売上が伸びている商品・話題や季節を反映させたおすすめ商品などが考えられる。アルゴリズムによる自動化の程度は、完全に自動化するか、中程度の商品分類レベルでシステムにより絞り込みを行い、具体的な商品は人が選ぶ方法がある。

(2) 本稿で扱うレコメンドシステム

本稿では業種により顧客のグルーピングを行う。業種情報は顧客企業が通販サイトに登録する際に得られる情報であり、簡便に入手できるためである。[田村 2008]によるキー商品グルーピングは煩雑な計算が必要のため、ハード性能が低い携帯通販には適さない。

レコメンドする商品のタイプは、定番の売れ筋商品・最近売上が伸びている商品・話題や季節を反映させたおすすめ商品すべてを扱う。これは、レコメンドには見る者を楽しませるという狙いもあるためである。多彩な商品をレコメンドすることで、新鮮さが失われないようにする。

自動化の程度は、完全自動化する商品タイプと具体的な商品は人が選ぶ商品タイプを設ける。定番商品は過去のデータから分析できるが、話題性のある商品は先取りすることが重要であり、購買データに表れる前にラインナップに加える必要があるためだ。

なお、以後の分析は、大手 BtoB 通販企業 X 社から提供された購買データを用いて行う。分析に用いるのは関東圏における、X 社の 2006 年度の半年間の購買データである。また、購買データは売上ではなく購買数を用いる。これは、売上は商品単価の影響を大きく受けてしまうため、購買数の方が商品の売れ行きを判断するのに適しているからだ。

2. レコメンドシステム

2.1 レコメンドシステムの概要

(1) レコメンドシステムの種類

図 1 にレコメンドシステムのパターンを示す。顧客のセグメント分けの基準・レコメンドする商品のタイプ・どの程度までアルゴリズムにより自動化するかにより区別することができる。

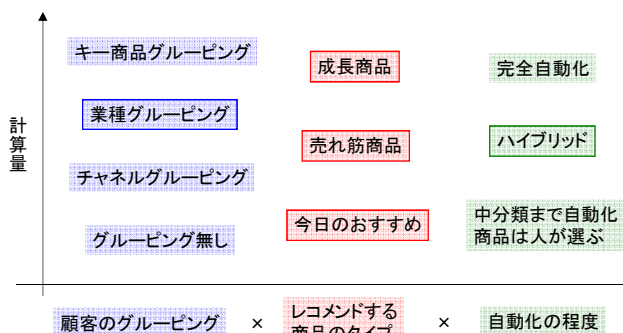


図 1. レコメンドシステムのパターン

2.2 顧客ユーザーのグルーピング

(1) グルーピングの概要

顧客ユーザーのグルーピング方法は、グルーピング重要度の高い業種を抽出し、抽出した業種を、購買傾向を確認しながら手作業でグループ分けするものとする。また、ここで扱う業種区分は、NTT の業種 64 分類である。

連絡先: 関口雄太郎, 東京大学大学院工学系研究科
sekiguchi@triton.naoe.t.u-tokyo.ac.jp

(2) グルーピング重要度

全業種の平均的な購買傾向からの乖離度が大きく、かつ購買量が多い業種を、グルーピングの重要性が高い業種とする。これは、購買傾向が平均から離れている業種ほど個別のレコメンドを行う必用があり、また、購買量が多い業種ほど通販企業から見た優先度が高いからである。

業種ごとに購買傾向に関して、図2に、商品を中分類単位で捉えた際の、業種ごとの売上にも占める割合を示す。赤い円で囲んだ部分は購買傾向が特に顕著に表れた部分である。業種ごとの、平均的な購買傾向からの乖離度を式1により定義する。

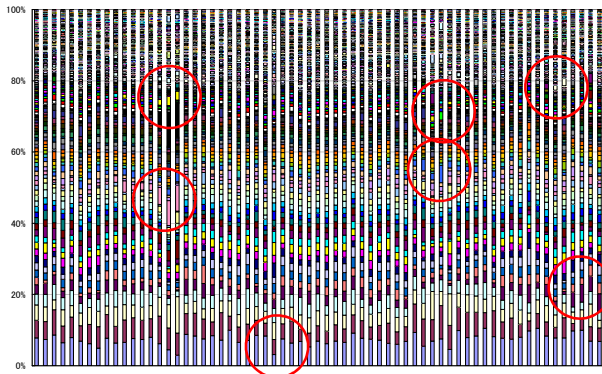


図2. 業種ごとの購買傾向

$$\text{乖離度} = \sum_{\text{全中分類}} (x_i - \bar{x})^2$$

式1. 乖離度

また、業種が X 社の受注回数に占める割合を業種の購買量とする。式2より、業種の乖離度と規模を掛け合わせ、業種ごとにグルーピング重要度を求める。

$$\text{グルーピング重要度} = \text{乖離度} \times \text{購買量}$$

式2. グルーピング重要度

図3に、全64業種のグルーピング重要度を求め、グルーピング重要度が大きい順に並べた。上位15業種のグルーピング重要度が大きく、16位以下は徐々に減少してゆく。よって、上位15業種を購買傾向に応じてグルーピングを行い、16位以下の顧客業種はまとめて一つのグループとする。

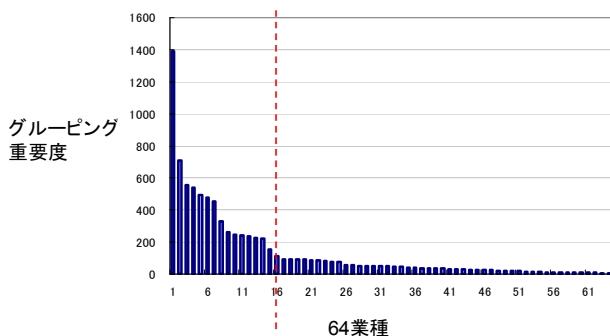


図3. 全64業種のグルーピング重要度

(3) グルーピング結果

表1に最終的なグルーピング結果を示す。上位15業種に関しては、業種ごとに、式で求めた乖離度が大きい商品と比較し、類似した商品が大きな乖離度を示す業種を同一グループとして括った。その結果、上位15業種を、医療、介護、オフィス、情報、教育、飲食、工事、接客の8つのグループにまとめた。16位以下のグループをその他とし、計9つのグループとなる。

表1. グルーピング重要度の上位30業種

順位	業種	乖離度	規模	重要度
1	医療機関	40706114	341675	1391
2	医療及び医療機械器具	48304014	147105	711
3	専門サービス(コンサルタント等)	39318235	141070	555
4	金融・保険・証券	51167542	105957	542
5	総合工事及び測量・調査・設計	21975975	225107	495
6	スナック・バー・酒場・喫茶店	212579445	22570	480
7	洋風・中華飲食店	148647340	30530	454
8	施設・機関	27901654	117928	329
9	クリーニング・理容・浴場	59743444	44338	265
10	教育	18058344	137776	249
11	不動産業	21258126	113455	241
12	情報・調査・広告	16762240	140069	235
13	組合・団体	27283292	83533	228
14	設備工事	25142552	87960	221
15	和風飲食店	49252627	31749	156
16	人材紹介・代行サービス	14903353	75176	112
17	官公庁	27605967	34081	94
18	趣味娯楽及びその関連産業	11846188	79195	94
19	音響及び通信・コンピュータ機器	10120513	91682	93
20	印刷・出版・書籍	11056250	82711	91
21	写真・デザイン・装飾	23175434	37722	87
22	運輸・倉庫	9343123	90895	85
23	一般機械器具	8643317	94360	82
24	輸送用機械器具	7867339	96492	76
25	衣服・呉服・小物	11158055	67767	76
26	放送・通信・報道	14769430	38155	56
27	鉄鋼	9491219	57322	54
28	化粧品・装飾品・民芸品	10188918	51671	53
29	離別工事	8275541	63586	53
30	建設資材	8557743	61347	52

医療
介護
オフィス
情報
教育
飲食
工事
接客

各グループの、乖離度が大きい商品を以下に記す。

- ・医療グループ:メディカル用品・医療事務用品
- ・介護グループ:メディカル用品・入浴用品・ペーパータオル・おむつ
- ・オフィスグループ:コピー用紙・クリアフォルダ・フラットファイル・スタンプ
- ・情報グループ:CD-R・クリップ・ふせん・ボトル飲料・お茶・ミネラルウォーター
- ・教育グループ:コピー用紙・マーカー、飲食グループは飲食用消耗品・キッチン用品・伝票
- ・工事グループ:フラットファイル・パイプ式ファイル・ラベルライター・梱包資材・テープ類
- ・接客グループ:コンディメント・紅茶・紙コップ

これらの商品は各グループの業務内容から無理なく連想されるものであり、本グルーピング手法は妥当と言える。

2.3 レコメンドする商品のタイプとアルゴリズム

(1) レコメンドする商品のタイプ

本研究では以下の6種類のレコメンドを考案した。日替わりでレコメンドの種類を切り替え、各レコメンドにつき5個の商品を選択できるようにする。選択する商品数が5個なのは、携帯電話の画面が小さく、5個の商品を表示するのが限界だからである。

- ・ユーザーがよく買う商品の類似品
 - A. ユーザーの所属グループ内の売れ筋商品
ユーザーが頻繁に買う商品と類似しており、しかもユーザーの所属グループ内で売れている商品。

- ▶ B. ワンランク上の商品
 - ユーザーが頻繁に買う商品と類似した高級商品。エコ商品や来客用など。
- ユーザーが買ったことがないタイプの商品
 - ▶ C. 全体で見た売れ筋商品
 - 全体的に売れているが、ユーザーが買ったことがないような商品。
 - ▶ D. ユーザーの所属グループ内の売れ筋商品
 - ユーザーの所属グループ内で売れているが、ユーザーが買ったことがないような商品。
- 最近売れている商品
 - ▶ E. 最近になって出荷数が伸びている商品
- マーケティング担当者のおすすめ
 - ▶ F. 話題性のある商品や、季節の商品
 - データ分析からは得られない、人間の感性を活かしたレコメンドを行う。

(2) レコメンドシステムのフロー

レコメンドシステムのフローチャートを図4に示す。本システムは購買データの管理とレコメンド機能から成る。

購買データに関しては、あるユーザーの購買行動が起こるたびに、購買データを本社のデータベースとユーザーの携帯端末の双方にストックする。ユーザー自身の購買データは携帯端末の中で常に最新の状態に保たれ、他ユーザーの購買データに関しては半年に一度、本社データベースから各ユーザーの携帯端末に送信され、端末内のデータが更新される。

レコメンド機能は携帯端末の中でアプリケーションとして完結する。これは本社にレコメンドシステム用のサーバーを構築するよりも携帯端末内にアプリケーションを導入する方が簡易であることと、X社が今後予定するビジネスモデルに起因する。

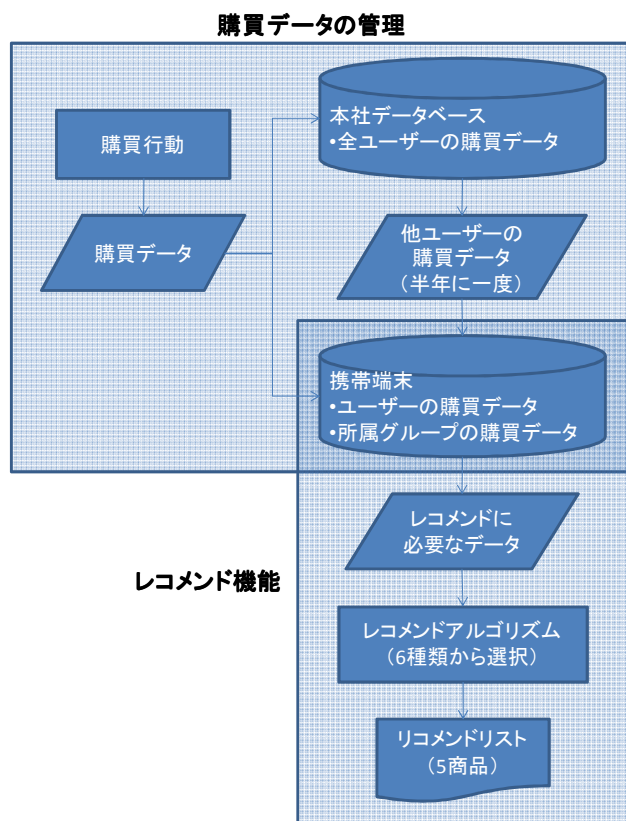


図4. レコメンドシステムのフローチャート

図5に、本社データベースと携帯端末内にストックされるデータの詳細を示す。本社データベースにストックされるのは、ユーザーCD・購買商品と数・購買日時などの生の購買データであるのに対し、携帯端末内にストックされるのは、ユーザー・ユーザーの所属グループ・全ユーザーの商品中分類と商品CDの購買数のみ加工されたデータである。レコメンドに用いない余計なデータを省くことで記憶容量を節約し、また、レコメンドに用いるデータ形式に加工して保存することで計算負荷を低減する。

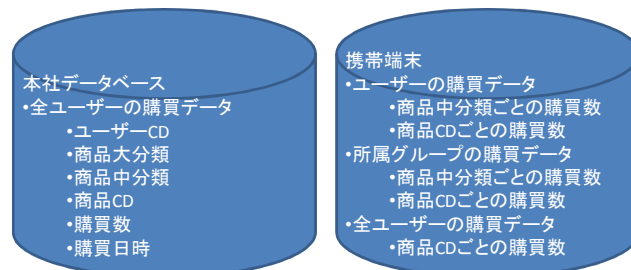


図5. 本社データベースと携帯端末内データの詳細

(3) レコメンド A のアルゴリズム

表2はあるユーザーの購買データを商品中分類単位で集計したリストのうち、上位5中分類を示したものである。レコメンドAにおいては、ユーザーの購買数上位5中分類と同じ商品中分類に関して、各々、ユーザーが所属するグループ内で購買数が最も多い商品をレコメンドする。ただし、その商品をユーザーが過去に購買したことがある場合は、所属グループ内で2番目に購買数が多い商品をレコメンドする。

表2. ユーザーが頻繁に買う上位5商品中分類

順位	中分類名称	カウント
1	洗剤	34
2	メディカル(衛生用品)	18
3	トイレトペーパー	18
4	コピー用紙	17
5	マーカー	15

(4) レコメンド B のアルゴリズム

商品中分類ごとに、マーケティング担当者が売り出したい高級商品をあらかじめ選択しておく。ユーザーの購買数上位5中分類に対応する高級商品をレコメンドする。

(5) レコメンド C のアルゴリズム

表3に、同ユーザーに対してレコメンドCを適用した様子を示す。まず、全ユーザーが購買した商品を中分類単位で集計したリストから、ユーザー自身が購買したことのある商品を除き、出荷数が多い順に5個の商品中分類を選択する。次に、5個の各中分類につき、中分類内で最も出荷数が多い商品を選択し、レコメンドする。

表 3. 全ユーザーの中分類購買データとレコメンド C

ユーザーが買った中分類 レコメンドCの中分類		
順位	中分類名称	カウント
1	コピー用紙	412242
2	インクジェットカートリッジ	264500
3	洗剤	232706
4	マーカー	164097
5	フラットファイル	135582
6	ラベルライターテープ	118598
7	ボールペン	114396
8	ティッシュペーパー	114246
9	ふせん	111803
10	印章・スタンプ用品	105928
11	クリップ	102335
12	コンディメント	100958
13	作業向粘着テープ	95396
14	クリアホルダー	93059
15	トイレトペーパー	90569

(6) レコメンド D のアルゴリズム

同ユーザーに対してレコメンド D を適用した様子を示す。表 4 はユーザーが所属するグループ内で購買された商品の中分類単位で集計したものである。まず、ユーザー自身が購買したことがある商品中分類とレコメンド C で選択した商品中分類を除き、出荷数が多い 5 個の商品中分類を選択する。次に、レコメンド C と同様に、5 個の各中分類につき、中分類内で最も出荷数が多い商品を選択し、レコメンドする。

表 4. 所属グループの中分類購買データとレコメンド D

ユーザーが買った中分類 レコメンドCの中分類 レコメンドDの中分類		
順位	中分類名称	カウント
1	コピー用紙	33297
2	インクジェットカートリッジ	32807
3	洗剤	31256
4	マーカー	12938
5	袋・包装紙類	12671
6	メディカル(衛生用品)	10953
7	医療事務用品	10737
8	印章・スタンプ用品	10579
9	トイレトペーパー	9545
10	ティッシュペーパー	9013
11	ゴミ袋	8330
12	ボトル飲料	8094
13	ラベルライターテープ	8051
14	のり	7804
15	ふせん	7617
16	キッチン用品	6604
17	日本茶/中国茶	6586
18	結束用品	6517
19	接着テープ	6308
20	清掃用品	6090

(7) レコメンド E のアルゴリズム

全ユーザーの購買データを用いて、商品ごとに、直近 3 か月間とその前の 3 か月間の購買数から購買数増加率を求め、増加率の大きい 5 商品レコメンドする。

(8) レコメンド F のアルゴリズム

マーケティング担当者が、おすすめ商品を 5 個選択する。他ユーザーの購買データが携帯端末内のデータに反映されるのは半年ごとなので、実際の売れ筋商品とレコメンド結果にタイムラグが生じる恐れがある。マーケティング担当者はその点を考慮し、その瞬間の流行や話題、季節を読んでレコメンドを行う。

3. 実購買データによる検証

3.1 検証方法

本章ではレコメンドシステムを X 社の顧客の購買データに適用し、レコメンドの効果を検証する。

代表として、最もグルーピング重要度が大きい医療グループのレコメンド結果を検証する。ユーザー数が最も多い、半年間の注文回数が 150 回前後のユーザー群の中からランダムにユーザー Y を選択する。

表 5 に、ユーザー Y の購入回数が多い商品中分類 5 個と、具体的な商品を示す。

表 5. ユーザー Y が頻繁に購買する商品

商品名称	中分類名称
ハイターE業務用	洗剤
ディステック ラテックスグローブ L/D-1125D	メディカル(衛生用品)
オリジナルイレットロールシングル業務用(8個)	トイレトペーパー
マルチペーパースーパーエコノミーA3	コピー用紙
ホワイトボードマーカー タフ 補充インク 黒	マーカー

3.2 検証結果

(1) レコメンド結果

ユーザー Y に対しレコメンドを行った結果を表 6 に示す。前述の通り、各レコメンドにつき 5 個の商品をレコメンドする。レコメンド B とレコメンド F のメーカーによる商品選択は我々が行った。

表 6. ユーザー Y に対するレコメンド結果

レコメンド	商品名称	中分類名称
A	オリジナルイレクリーナー3個パック	洗剤
	ティスポーザブルマスク ホワイト ES3-00175W	メディカル(衛生用品)
	オリジナルイレットロール リサイクル100 ダブル12R	トイレトペーパー
	マルチペーパースーパーエコノミーA4(5000枚)	コピー用紙
	マッキー極細 1パック(10本入) 黒	マーカー
B	善玉パイオ洗剤“浄”3個セット	洗剤
	3層式ティスポーザブルマスク	メディカル(衛生用品)
	トイレトペーパー・ホットケーキの香り	トイレトペーパー
	上質コピー用紙	コピー用紙
	蛍光OAマーカー<プリフィクス>(シングル)	マーカー
C	キヤノン インクC BCI-3eBK ブラック	インクジェットカートリッジ
	テブラPROテープカートリッジ 白ラベル 黒文字12mm	ラベルライターテープ
	オリジナル再生紙ふせん 75×25mmカラー4色	ふせん
	ネーム9専用補充インキ	印章・スタンプ用品
	PKAS ゼムクリップ 大(10袋)	クリップ
D	レセプト用紙A4(2500枚入)	医療事務用品
	アスクル半透明ゴミ袋45リットル	ゴミ袋
	サッポロ 恵比寿茶房 玉露入りお茶 2.0L 1箱(6本)	ボトル飲料
	伊藤園 おーいお茶緑茶ティーバッグ(100バッグ入)	日本茶/中国茶
	ラバーバンド#16 500g	結束用品
E	エナージェル0.5mmニードル黒インク BLN75-A	ボールペン
	エナージェル0.5mmニードル赤インク BLN75-B	ボールペン
	アースジェット2本パック	雑貨
	セレクション・ザ・グロム春夏	お菓子
	うがい用紙コップ(どうぶつ5種)3000個入	メディカル(衛生用品)
F	耐震用固定ベルト	防災用品
	森永製菓 ミルクココア 370g	飲料
	エアークイックプレッシャー	オフィス備品
	60cmジャンプ傘 エンボス白	日用品
	大特価 プレンディーベアパック	飲料

(2) レコメンド A に関して

ユーザーが頻繁に買う商品に類似した商品をレコメンドしていることが確認できる。BtoB 通販で扱う商品数は一企業あたり最大で 7 万点にも及ぶので、類似した他商品の存在に気付いていないユーザーの存在が想定され、そのようなユーザーに代替的な選択肢を提供することができる。

一方、コピー用紙に関してはサイズ違いの商品をレコメンドしている。必要ならば購入しているはずなので、的外れなレコメンドと言える。サイズ違いや容量違いの商品は除外するなどの工夫が考えられる。

(3) レコメンド B に関して

ユーザーが頻繁に購入する商品の高付加価値な代替品がレコメンドされている。香りのするトイレトペーパーなど、個人医院であれば来客用に揃えると印象が良くなる。また、環境に優しいエコ商品をレコメンドしても良い。

(4) レコメンド C に関して

3 位から 5 位にふせん、スタンプ、クリップなど、どの業種でも共通して使う商品が選択されていることがわかる。外れの無い無難なレコメンドと言える。

一方、1 位と 2 位は補充品が選択されている。対応する商品を使っていないユーザーには無意味なレコメンドとなってしまうので、補充品は除外するか、補充品が選択される場合は本商品をレコメンドする改善方法が考えられる。

(5) レコメンド D に関して

1 位に医療事務用品がレコメンドされている。これは、このユーザーが医療グループであるにも関わらず X 社で医療事務用品を購入していないことを意味する。レコメンドがきっかけとなりユーザー Y が医療事務用品を購入する効果が期待される。

(6) レコメンド E に関して

4 位の菓子は季節を反映した流行商品であり、5 位のうがい用紙コップはデザイン入りのおしゃれな商品である。最近になって注目され始めた商品をうまく選択している。

一方、1 位から 3 位には、商品の性格上、販売量の変化が小さいと考えられる商品が選択されている。これは X 社が最近になって 3 つの商品を扱い始めたからだと推測される。レコメンドを通して新商品の宣伝効果が期待される。

(7) レコメンド F に関して

耐震用品、冬に温まるココア、マイナスイオン商品、雨が降っている日であれば傘、大特価品を選択した。これらは、世の中の変化に応じて選択する商品であり、分析対象のデータに反映される頃には流行遅れとなる。ただし、ここに示したレコメンドはあくまで一例であり、実際には商品ラインナップのマーケティング担当者が商品を選択する。

4. 考察

6 種類のレコメンドに関して、シンプルなアルゴリズムでありながら、外れないレコメンドが可能であることが確認できた。さらに、各レコメンドで全く異なる商品をレコメンドしているため、レコメンドのタイプを定期的に切り替えて新鮮さを保つという狙いも達成された。

一方で、補充品やサイズ違いの商品のレコメンドに関しては前述した改善策を実行する必要がある。

5. 結言

- ハード性能が PC に劣る携帯電話でも利用可能な、シンプルかつ精度の高いアルゴリズムを提案した。
- 実際の購買データにレコメンドシステムを適用し、レコメンドの精度を検証した。また、改善方法を提案した。

参考文献

[中村 2007] 中村美穂: 商品購入時期に対するユーザ志向を考慮したレコメンド方式の提案, 第 21 回人工知能学会全国大会, 2007.

<http://www.ai-gakkai.or.jp/jsai/conf/2007/data/pdf/100046.pdf>

[田村 2008] 田村泰一: 通信販売ビジネスのコンシェルジュシステムの構築, 東京大学大学院技術経営戦略学専攻修士論文, 2008.

