

併買関係に基づく金銭的価値による商品ランキング

横田 祐介* 松井 藤五郎* 大和田 勇人*

東京理科大学理工学部経営工学科*

1 はじめに

現在、様々な商品から膨大な売上記録が蓄積されており、これらを分析する手法としてアソシエーションルールが注目を集めている。近年の研究で八木 [1] らは、アソシエーションルールと HITS やページランクなどのアルゴリズムを利用して各アイテムの重要度を分析し、併買されやすい商品のランキングを行っている。また、TOM BRJIS [2] らは、アソシエーションルールを用いた頻出のアイテムセットから最適な製品選択のためのモデルを構築している。このように商品の販売利益だけでなく、併買関係に注目して商品の価値を測る様々な研究が存在しているが、商品の併買関係のみで販売利益を考慮できていないため、実際に利益をもたらすのはどの商品なのかが不透明である。

そこで本論文では、併買関係を考慮に入れた金銭的価値に基づく商品のランキングを作る手法を提案する。これにより販売側にとってどの商品が有益であるかが視覚的にもわかりやすくなる。

2 提案手法

八木らを用いたページランク法とは、サーチエンジン Google のページ重要度の自動判定システムであり、「多くの良質なページからリンクされているページは、良質なページである」という再帰的な関係から、全てのページの重要度を判定している。これを商品に用いて、商品自体の販売利益が高いことかつ販売利益が高い多数の商品から併買されている商品は金銭的価値が高いといえる。併買関係のあるそれぞれの商品から分配される利益を併買利益とする。今回は併買利益の配分方法としてアソシエーションルールの信頼度を利用する。アソシエーションルール $X \Rightarrow Y$ の信頼度 $confidence(X \Rightarrow Y)$ はすべてのトランザクション (一人の顧客が一回の買い物で購

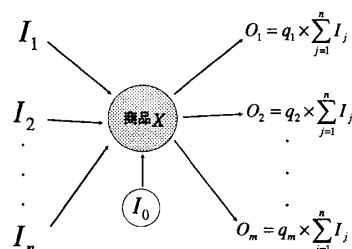


図 1: 商品 X と併買利益の関係

入したアイテムの集合) の中で X を含むトランザクションのうち、 X と Y を共に含むトランザクションの割合で計算される。計算手法としては、最初の販売利益からユーザが設定した併買寄与率 p の割合に応じて併買利益を分配する。残った販売利益を金銭的価値として固定し、次に自分が他の商品から併買利益をもらう。これを繰り返すことによって金銭的価値を評価する。他の商品に与える併買利益の合計と他の商品からもらう併買利益の合計は等しくなく、一部が金銭的価値として固定されていくこと、併買利益の配分は平等ではないこと、最初に用いる各商品の販売利益が異なることがページランクとの大きな違いである。この計算方法を以下に示す。

1. 各商品は固有の販売利益を持っている。
 2. ある商品 X に対して
 - X の販売利益を I_0 とする。
 - 他の商品から X へくる各々の併買利益を I_1, I_2, \dots, I_n とする。
 - X から他の商品に与える各々の併買利益を O_1, O_2, \dots, O_m とする。
 - X が他の商品に与える利益の割合値を併買寄与率 p とする。
 - X の金銭的価値を P とする。
- このとき、次が成り立つものとする。

$$P = (1-p) \sum_{i=0}^n I_i, \quad \sum_{j=1}^m O_j = p \sum_{i=0}^n I_i$$

Commodity ranking by pecuniary value based on closs buying
Yusuke YOKOTA*, Tohgoroh MATSUI*, and Hayato OHWADA*
Department of Industrial Administration, Faculty of Science and Technology,
Tokyo University of Science*

p の値は商品ごとに任意に決めることができる。今回は信頼度の値を基本として、販売個数の割合で正規化を行い、次のようにして決めた。

- 商品 X から利益が与えられる各々の商品を X_1, X_2, \dots, X_m とする。
- 各商品へ与える併買利益の割合を q_1, q_2, \dots, q_m とする。
- 商品の各々の売り上げ個数を w_1, w_2, \dots, w_m とする。
- アソシエーションルール $A \Rightarrow B$ の信頼度を $conf(A \Rightarrow B)$ とする。

このとき次が成り立つものとする。

$$p = \sum_{i=1}^m q_i, \quad O_i = q_i \sum_{j=0}^n I_j$$

$$q_i = \frac{w_i}{\sum_{j=1}^m w_j} \times conf(X_i \Rightarrow X)$$

商品 X と併買関係にある商品との利益の関係は図 1 のようになる。

p を決めることによって信頼度と販売個数に基づいて利益を分配できるようにしたが、変動する金額が小さいことが予想されるためユーザーがパラメータ α ($0 \leq \alpha < 1$) の値を与えることで配分する利益の値を全体的に引き上げられるようにした。各商品の p を p_1, p_2, \dots, p_k とし、 $\alpha = 0$ の時、 p_k の値は変わらず、 $\alpha = 1$ の時、 p の合計は 1 となるように設定している。

$$p'_k = \frac{(1 - \alpha) \times \sum_{i=1}^k p_i + \alpha}{\sum_{i=1}^k p_i} \times p_k$$

3 実験

実験は OS が Windows XP、ブラウザが Internet Explorer 7、データベースが Microsoft SQL Server Analysis Services のサンプルキューブである FoodMart、その他に eclipse 3.2、MySQL 5.1、TOMCAT 5.5 などを使い JAVA で実装した。実験は、FoodMart 中の sales1998dec 表のトランザクション 3781、商品 1559 個を用いて行った。すべての商品の併買関係を計算すると膨大な時間がかかるため、今回はアソシエーションルールの支持度を用い、0.005 を下回る商品は計算に加えないこととし 49 個の商品に絞って行った。実験結果の表示は順位、元の順位、商品名、利益(円)、元の利益(円)の 5 つを表示した。元の順位とは、それぞれの商品の販売利益の順位、元の利益はその総額である。順位とは、提案した手法で併買関係に基づく利益を各商品間で受け渡したあとの順位、利益はその総額である。 α を 0 とした時のランキングが表 1 で、 α を 0.5 とした時のランキングが表 2 である。灰色で表示されているところが販売利益の値が変化したところである。

表 1: α が 0 の場合のランキングの一部

順位	元の順位	商品名	利益(円)	元の利益(円)
1	1	ブッカー 低脂肪カッターチーズ	18111.22708	18378
2	2	ヒルトップ MG イブプロフェン	18095.83228	18095
3	3	エヴァーベター チーズボール	17270	17270
4	4	山本ドイツ産のワイン	16878.82024	16878
5	5	ベスト・ホイイス チョコレートナッツクッキー	16440	16440
6	6	オランダ村 レーズクッキー	16220	16220
7	7	ベスト・ホイイス 皮付チーズ	15528.94824	15780
8	8	赤い翼 プラスティックナイフ	15371.72056	15450
9	9	ぶたのしっぽ 冷凍コーン	15170	15170
10	10	オランダ村 クッキー	14910	14910
11	11	王様の冠 オレンジジャム	14430	14430
12	12	赤い翼 プラスティックナイフ	14400	14400
13	13	ヒルトップ ミントマウスウォッシュ	14390	14390
14	14	ハイマウンテン 新じゃが	14290	14290
15	15	メジャー トップ ワイン(シャブリ)	14150	14150

表 2: α が 0.5 の場合のランキングの一部

順位	元の順位	商品名	利益(円)	元の利益(円)
1	3	エヴァーベター チーズボール	17270	17270
2	2	ヒルトップ MG イブプロフェン	18095.83228	18095
3	5	ベスト・ホイイス チョコレートナッツクッキー	16440	16440
4	6	オランダ村 レーズクッキー	16220	16220
5	24	山本ドイツ産のワイン	15878.18024	16878
6	9	ぶたのしっぽ 冷凍コーン	15170	15170
7	10	オランダ村 クッキー	14910	14910
8	11	王様の冠 オレンジジャム	14430	14430
9	12	赤い翼 プラスティックナイフ	14400	14400
10	13	ヒルトップ ミントマウスウォッシュ	14390	14390
11	31	赤い翼 鋼製ボルトナット	13302.02228	13100
12	14	ハイマウンテン 新じゃが	14290	14290
13	15	メジャー トップ ワイン(シャブリ)	14150	14150
14	17	浜辺の恋人 ミンツ	13950	13950
15	18	ABC BEST りんごの砂糖漬け	13750	13750

4 考察

実験では併買関係に基づいた金銭的価値の計算を行った。 α が 0 のときは金銭的価値の大きな変動はなかった。 α が 0.5 のときは併買利益の変動によりランキングが変動した。例えば、ブッカー低脂肪カッターチーズは販売利益のみのランキングは 1 位だったが併買関係を考慮すると 73 位へ、ABC BEST なたね油は 26 位から 5 位へなど大きく変動している。ランキングが併買関係に基づいて変動したことから、提案手法は有効だと考えられる。また、 α を用いることでランキングがより併買関係に基づいて作成されたと考えられる。

今後の課題としてすべての商品を計算に含めて短時間で結果を出す手法や、 α を用いずに併買関係に基づいたランキングを作る手法の考案などが考えられる。

参考文献

- [1] 八木一光, 岡野慎吾, 森本康彦, 相関ルールの 2 部グラフを用いた重要アイテムの発掘, DEWS2006 6A-i7
- [2] TOM BRIJS, GILBERT SWINNEN, KOEN VAN-HOOF, GEERT WETS. Building an Association Rules Framework to Improve Product Assortment Decisions. Data Mining and Knowledge Discovery, 8, 7-23, 2004