

顧客の成長に着目したネットワーク分析による重要商品の抽出に関する一考察

伊藤 寛彬[†] 雲居 玄道[§] 山下 遥[‡] 後藤 正幸[¶]
 早稲田大学[†] 理工学術院総合研究所[§] 早稲田大学[‡] 早稲田大学[¶]

1. 研究背景と目的

近年の小売業においては、ICカードを用いた会員システムやポイントシステムを利用し、顧客がいつ、どこで、何を買ったのかについての詳細なデータを蓄積することが可能となった。このようなデータを活用したマーケティング分析も一般的になりつつあり、優良顧客をどのように獲得するのか、どのような施策を講じれば、離反顧客を抑止できるのかといった様々な観点からの解析が数多く行われている [1],[2]。

本研究では共同研究先である小売チェーンの1年間の購買履歴データを対象事例として扱う。この小売チェーンにおける特徴の1つとして会員ステージ制度がある。この制度では、年間の累積購買金額ごとに5段階(0~4)の会員ステージに分割され、それぞれのステージに到達した会員に特典が設けられている。また、毎年2月末日にステージがリセットされ、毎年3月に全ての顧客が初期ステージ(ステージ0)から再スタートするという特徴がある。2月末の最終的な到達ステージは、優良顧客の度合いと捉えることもでき、どのような商品を購入している顧客の会員ステージが向上しているのかについて分析することは重要である。

そこで、本研究では最高ランクのステージ4において重要度の高い商品を特定し、その商品を最高ランク以外の会員に購入してもらうために重要な商品をステージ毎に特定する方法を検討する。こうした商品をその他のランクの会員に購入してもらうことにより、最高ランクの会員ステージへの成長が期待できる。

具体的には、購買における商品間の共起の度合い(類似度)を用いて、商品をいくつかのクラスタへと分割する。さらに重要度分析によって、最高ランクの会員に人気のある商品との関連性が高い他商品の重要度を算出する。この分析により、各ステージの会員に購入を促すことによって最高ランクの会員に成長させることが可能となるような商品を特定する。

同様に、短期的な顧客成長のための施策として、各ステージの1段階上のステージの会員に人気がある商品を特定することで、ステージ0から1、ステージ1から2、ステージ2から3への段階的な顧客成長を促進するために重要度の高い商品をそれぞれ抽出することを目的とする。

2. 準備

2.1 分析事例

事前分析により、2月の最終時点でステージ0からステージ4であった顧客の購入商品を集計した結果、各ステージ間での売上個数や購買金額が多い商品に差異があることが明らかになった。したがって、各ステー

ジによって商品推薦などにおいて会員ステージを最高ランクまで引き上げるために重要度が高い商品も各ステージで異なると考えられる。そこで最終的にステージ0~3に留まった顧客に着目し、各ステージで重要度が高い商品の中で、ステージ4の重要度の高い商品と関連度が高いものを抽出することができれば、顧客の成長を促す商品の特定が可能となり、マーケティング施策の立案に対する助けとなるものと思われる。

2.2 分析対象データ

本研究では、株式会社良品計画から提供頂いた購買履歴データを分析対象とする。購買履歴データの期間は2014年3月1日から2015年2月28日の1年間である。商品の購買は各店舗またはインターネット上のECサイトのどちらか、あるいはその両方で可能である。本研究では、全顧客のうち、最頻利用店舗が東京都の店舗である360,923人に限定し、購買商品についての分析をおこなう。ここで、ECサイトを含めた実店舗の総数は447店舗(うち東京都にある店舗の総数は83店舗)、アイテムは、当該店舗で販売されている商品の中分類669カテゴリを用いて解析する。

3. 分析手法

3.1 概要

本研究ではNTTソフトウェアイノベーションセンタが開発したグラフマイニングアプリケーションであるGraponを使用し、クラスタ分析 [3] と重要度分析 [3] を行った。

3.2 クラスタ分析

クラスタ分析は、与えられたグラフ構造から、ノードをつながりの密なグループへと分割する手法である。具体的には、グラフのリンク構造から、部分グラフ内のノード間のエッジ密度が、相対的に部分グラフ外へのエッジの密度よりも高くなるような部分グラフを切り出す。尺度としてモジュラリティ値 Q を用いる。このモジュラリティ値が高いほど、よいクラスタ分割ができていると解釈できる。 C を抽出したコミュニティ集合、 e_{ij} をコミュニティ i からコミュニティ j へ接続される重み付エッジ数、 m をグラフ構造全体に含まれる重み付エッジ数とするとモジュラリティ Q は式(1)で表現される。

$$Q = \sum_{i \in C} \left(e_{ii}/2m - \left(\sum_{j \in C} e_{ij}/2m \right)^2 \right) \quad (1)$$

A Study on Extraction of Important Items
 Focused on Customer Growth
 Based on Network Analysis

[†] Hiroaki Ito · Waseda University

[§] Gendo Kumoi · Waseda Research Institute for
 Science and Engineering

[‡] Haruka Yamashita · Waseda University

[¶] Masayuki Goto · Waseda University

3.3 重要度分析

3.3.1 ステージ毎の重要度分析

重要度分析とは、リンクのつながり方の重要度に応じてノードをランク化することである。尺度として page rank p を用いる。この page rank p をノード毎に逐次更新し、得られた p^* をノード (N 個) における重要度とする方法である。 O_i をノードのエッジ数、 A を正規化された隣接行列 ($A_{ij} = 1/O_i$)、 q を各ノードの重みの値とすると page rank p は式 (2) で表現される。

$$p = (1 - c)Ap + cq \quad (2)$$

この時、重み q をすべて 1 と設定したもとの各ステージごとに重要度分析を行う。

3.3.2 ステージ間の重要度分析

重要度分析においての各商品の重み q を 3.3.1 節の重要度分析の結果より、現ステージと目標とするステージを比較して、目標とするステージの上位 20 位に含まれ、重要度の順位が高い商品のみを 1、それ以外を 0 と設定したもとの各ステージごとに重要度分析を行う。

4. 分析結果

4.1 クラスタ分析

クラスタ分析を各ステージごとにおこなった結果を表 1 に示す。クラスタ数はモジュラリティ値が最小となるときの数である。クラスタ 2 に含まれる商品はステージ 0 (非優良顧客) では食品のみであるのに対し、ステージ 4 (優良顧客) になると食品・衣類・生活雑貨と 1 つのクラスタに違う種類のアイテムが混在している。

表 1. 各ステージのクラスタ分析結果

クラス	Stage0	Stage1	Stage2
1	衣類	雑貨, 衣類 A	衣服 A
2	食品	食品	衣服 B
3	子供用品	子供用品, 衣類 B	食品, 子供用品
4	雑貨		雑貨

	Stage3	Stage4
	衣類 A 食品, 衣類 B, 雑貨 子供用品 雑貨	衣類 A 食品, 衣類 B, 雑貨 子供用品 雑貨

4.2 重要度分析

4.2.1 ステージ毎の重要度分析

以下にステージ 0~4 の会員について、重要度の高い商品の抽出の結果を表 2 に示す。

表 2. 各ステージの重要度が高い商品

stage0		stage1		stage2	
商品名	重要度	商品名	重要度	商品名	重要度
食品 A	0.01004	食品 A	0.00816	食品 A	0.00702
食品 B	0.00900	雑貨 A	0.00736	雑貨 B	0.00653
雑貨 A	0.00880	食品 B	0.00731	雑貨 A	0.00635
食品 C	0.00837	雑貨 B	0.00718	食品 B	0.00630
雑貨 B	0.00795	食品 C	0.00682	食品 C	0.00592

stage3		stage4	
商品名	重要度	商品名	重要度
食品 A	0.00606	食品 A	0.00474
雑貨 B	0.00578	雑貨 B	0.00471
食品 B	0.00552	雑貨 J	0.00459
雑貨 A	0.00544	雑貨 M	0.00459
食品 C	0.00514	雑貨 K	0.00459

表 2 より、ステージが上がると主に生活雑貨の重要度の順位が上昇する傾向にあることがわかる。ステージ 4 の会員は重要度の偏りが少なく、幅広く商品を購入する傾向にあるといえる。

4.2.2 ステージ間の重要度分析

ステージ 0 とステージ 4 の重要度分析より、ステージ 0 から 4 で重要度の順位が上昇している生活雑貨 B,D,E,M,K,J の 6 商品を重要商品として設定し、ステージ 0 に対してステージ 4 とのステージ間の重要度分析を行った。結果として生活雑貨 I,R,O,U が新たにステージ 0 における重要度ランキングの上位 20 位に出現した。すなわち、長期的な視点でステージを向上させるためには生活雑貨 I,R,O,U も重要な商品と考えられる。同様にステージ 0 と 1 間において分析をおこなった結果、生活雑貨 A,B,D,E,J を重要商品として、生活雑貨 F,G,L,I,M,R も新たにステージ 0 における重要度ランキングの上位に出現し、短期的にステージ 0 の会員を成長させるためにはこれらの商品も重要と考えられる。

4.3 考察

今回分析した小売チェーンでは、優良顧客であるほど幅広く商品を購入する傾向にある。また、重要度分析より、ステージ 0 からこの小売チェーンで買い物をするきっかけになるのは「食品」であり、会員ステージが向上するにつれて身のまわりの「生活雑貨」の重要度が高くなる傾向にあることが明らかとなった。さらに、短期と長期の両視点から、生活雑貨の中でも特に生活雑貨 O や U が長期的な優良顧客に導く可能性の高い重要商品であることがわかった。

5. 今後の課題

今回会員ステージ別に行ったクラスタ分析と重要度分析を、ネット/実店舗別に層別したデータに対しても行ない、重要度の高い商品を特定した。また、性別・年代別で層別したデータに対しても同様の分析を行っており、その結果を発表する予定である。今後の課題として、現段階では商品の中分類レベルでの分析だが、最終的には商品の小分類レベルでの分析をおこなうことが挙げられる。

謝辞

本研究に際し、ご協力を頂いている株式会社良品計画に深く感謝いたします。

参考文献

- [1] 伊藤孝太郎, 澤邊 剛, 保坂佳佑, 松下亮祐, 雪島正敏: “顧客のセグメンテーションと商品のスコアリングによる購買予測”, オペレーションズ・リサーチ: 経営の科学, Vol.60, No.2, pp.75-80, 2015
- [2] 高野祐一ほか: “ファジィクラスタワイズ回帰を用いた共同購入型クーポンサイトの閲覧傾向分析”, オペレーションズ・リサーチ: 経営の科学, Vol.59, No.2, pp.81-87, 2014
- [3] 飯田恭弘ほか: “大規模グラフ構造データからのコミュニティ抽出と重要度計算”, 人工知能, Vol.29, No.5, pp.472-479, 2014