

# Web 上の製品評価サイトのクチコミデータの特徴解析と ブランドマネジメントへの応用

石塚 隆男

亜細亜大学経営学部

## 1. はじめに

今日、企業に限らず多くの意思決定主体にとってブランドの構築とともにブランド力を高めることが最重要課題となっている。しかしながら、多くの場合、“ブランド力”は明確な定義がなされぬまま感覚的、雰囲気的に用いられているのが実状である。従来、ブランド力の測定は、ブランド知覚に関するアンケート調査による測定が中心である（たとえば、恩蔵(1995)）。本来、製品のブランド力は時間とともに変動するものであり、リアルタイムとまではいかななくてもルーチ的に随時測定可能であることが望ましい。POS データは商品のシェア情報の把握には役立つが、単純なシェア＝ブランド力ではないことは明らかであり、シェアの質を考慮したブランド力の構成が必要である。

そこで、本研究では Web 上の製品評価サイトに書き込まれたクチコミデータを用い、ブランド力の定量化を行うことを目的とする。本研究では、ブランド力はある製品カテゴリーにおけるブランド間で相対的に定量化され、比較可能であるとの立場をとる。こうした観点からわれわれが用いた方法は、データ包絡分析 (DEA) の発想に近いことを示す。今回、シャンプー市場を対象にクチコミデータからブランド力の定量化を試みた結果、いくつかの知見が得られたので報告する。

## 2. ブランド力の構成

アーカー(1997)は、ブランド力なる表現は用いず、ブランド・エクイティとして概念構成し、その構成要素として、「ブランド認知」、「ブランド・ロイヤルティ」、「知覚品質」、「ブランド連想」、その他の資産、の5つを挙げている(恩蔵(前掲))。

アーカー(前掲)は、各構成要素について測定尺度を具体的に列挙しているが、ブランド力について理論的な構成を明らかにし実証まで行った研究は見当たらない。

本研究では、アーカーの主張に基づき、個別製品のブランド力を、「ブランド認知」、「ブランド・ロイヤルティ」、「知覚品質」、「ブランド・イメージ」の4つの要素により構成されるとした。

しかし、クチコミコメントデータからこれらの要因を直接的に測定することはできないので、各要因について以下の代替特性を用いた。

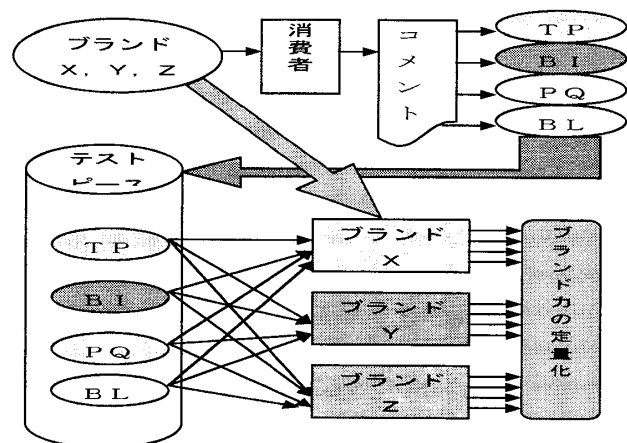
- ①ブランド認知：クチコミのコメント数とタッチ・ポイントの種類<TP>
- ②ブランド・イメージ：ブランドのイメージに関する記述内容<BI>
- ③知覚品質：製品の品質並びにその評価に関する記述内容(+0-の記述がある)<PQ>
- ④ブランド・ロイヤルティ：リピート購買に関する記述内容(+の記述がある)<BL>

多くのコメント文は、TP、BI、PQ、BLに関する文から構成されているが、これらの要素の一部が欠落したコメント文も少なからず存在する。そこで、ある製品ブランドのコメント文集合の中でこれらの各要素を含むコメント文の割合が高いほど対応する要素特性のブランド力が高いと考えられる。

## 3. DEAの発想に基づくブランド力モデル

ブランド力が、前章で述べたTP、BI、PQ、BLの4つの要素により構成されるとすれば、任意の製品ブランドがこれらの4要素の“試薬”あるいは“テストピース”にどう反応するかを調べることによりその製品のブランド力を定量化することができる。消費者は、刺激(S)となる製品ブランドにたいしてさまざまな反応(R)を示すが、その反応を記録したデータが製品評価サイトのクチコミデータである。図1は、ブランド力の構成要素と各ブランドとの関係を示している。

図1. ブランド力とブランドとの関係モデル



データ包絡分析 (DEA) は、多入力多出力の意思決定主体の効率を相対的に比較し、改善するための手法である (Charnes et al (1978))。しかし、入出力変数は計量または計数データとして入手できる必要があり、本研究が対象とする自然言語データに適用するためには各入力変数の数量化を行わなければならない。

図1に示すように、各ブランドは多入力多出力の意思決定主体 (DMU) と見なすことができる。各ブランドに最も有利なテストピースを用いれば、そのブランドのブランド力が最大になることが期待されるが、同じテストピースを他のブランドに適用したとき、より高いブランド力特性を示すことがありうる。したがって、テストピースの種類 (バリエーション) が十分であれば、各ブランドに共通のテストピースを適用することができる。そこで、4要素のテストピースの作成方法について説明する。

#### 4. 研究方法

##### 対象データの抽出

@COSME の製品ランキングページにアクセスし、シャンプー・コンディショナー製品のランキングの中から、Asience (A)、Lux (L)、TSUBAKI (T) の3ブランド含む 10 ブランドに関する 2006 年 7 月までの期間最長 2 年間のコメント文をダウンロードし、インデックスに相当するコメントプロフィールデータとコメント文リストに分割し、書き込みの時系列順に保存した。

各コメント文には、1～7までの星の数により評価ポイントが付与されているが、「評価なし」のコメント文も若干あり、整合性を保つため、これらのコメントは対象外とした。

##### テストピースとしてのフレーズリストの作成

コメント文は、省略形や短縮形を多用した日本語で表現されており、リポートする、しないに関しても多様な表現があり機械的に判断するのは難しい。そこで、テストデータとしてある製品に関する大量のコメント文をダウンロードし、その中から目視によりブランド力を構成する4つの要素に該当するフレーズ (句) を抽出する。見つかったフレーズは、XML 形式のタグでくくり、重複抽出や抽出漏れがないようにコメントテキスト全体の中で同じ記述のフレーズを自動的に置換するようにした。2004 年 8 月～2006 年 7 月までの 2 年間の Lux のコメント文 1307 件から、以下の件数のフレーズを抽出した。

<TP>: 54 種類 <BI>: 49 種類  
<PQ>: 285 種類 <BL>: 445 種類

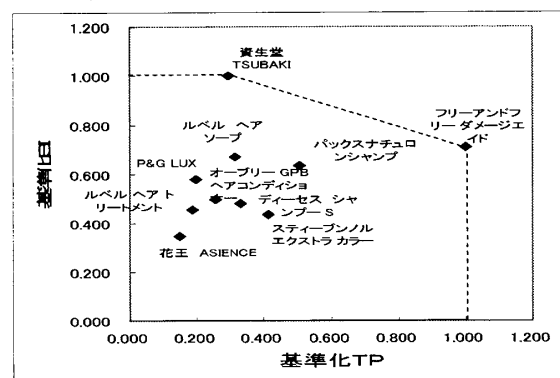
こうして抽出された4つの区分のフレーズリストを他の製品ブランドのコメントデータに適用し、その出現頻度の割合の比較を行った。

#### 5. 結果

シャンプー市場における 10 ブランドに Lux のコメント文から抽出した要素別フレーズリストをテストピースとして適用し、コメント数に対する TP、BI、PQ、BL を含むコメントの比率をブランド力特性として定量化

した。しかし、比率はコメント数に影響されないため、ブランド認知力が反映されているとはいえない。そこで、比率  $p$  をその標準偏差  $\{p(1-p)/n\}^{0.5}$  (ただし、 $n$  はコメント数) で割り、 $\{pn/(1-p)\}^{0.5}$  を特性値とした。これにより、比率  $p$  が高く、コメント数  $n$  が大きいほど、特性値が大きくなり、ブランド認知度であるコメント数が反映された値となる。各特性について、最大値で割ることにより基準化し、相対的な効率として解釈することが可能である。図2に、TP と BI についてそれぞれの特性値を [0, 1] に基準化した 10 ブランドの散布図を示す。図2より、TSUBAKI とフリーアンドフリーがフロンティアを構成していることがわかる。

図2. TP×BIのブランド力散布図



次に、シャンプーの各ブランドのクチコミデータが当該ブランドに特有な表現をどの程度用いているか、並びにクチコミ内容が時間の経過に伴い、どのように変化しているかを調べるため、クチコミ中の単語×ブランドの頻度マトリクス並びに単語×月別の頻度マトリクスの集計を行った。それにより、消費者は企業が発信するブランド・アイデンティティを受け止めているとともに、特にマイナスの評価を表す単語の時系列変化が明らかとなった。

#### 6. 考察並びに今後の課題

本研究では、製品評価サイトのコメントデータからブランド力の定量化を試み、自然言語データにもDEAの考え方が応用できることを示した。また、単語×ブランドマトリクスの解析がブランドマネジメントにも有効であることが確認された。今後の課題として、ブランド体系のオントロジーの調査やブランド間のポートフォリオにクチコミデータの活用が考えられる。

##### 参考文献

- Charnes, A. W. et al (1978) "Measuring the Efficiency of Decision Making Units," *European Journal of Operations Research*, Vol. 2, No. 6, pp. 429-444  
 アーカー, デービッド A (陶山計介他訳) (1997) 『ブランド優位の戦略』ダイヤモンド社  
 恩蔵直人 (1995) 『競争優位のブランド戦略』日本経済新聞社