

商品の販売実績に与える 消費者生成型メディア上における コミュニケーションの影響に関する研究

[継続研究]

常勤研究者の部



代表研究者 鶴 見 裕 之

横浜国立大学大学院
国際社会科学研究院 経営学専攻
准教授



共同研究者 増 田 純 也

株式会社 電通マーケティングインサイト



共同研究者 中 山 厚 穂

首都大学東京大学院
社会科学研究科 経営学専攻
准教授

1章 研究の背景と目的

本研究の目的は消費者生成型メディア（CGM : Consumer Generated Media）におけるコミュニケーションの内容とその規模が、一般消費財の販売実績に及ぼす影響をマクロ的な視点から分析し把握することにある。

今日、人口に対する普及率、ネットユーザーに対する普及率

の双方から見て、SNS の 1 つに過ぎない twitter 単体が既に大きな影響力を有するメディアとなりつつあると言えよう。またこのような SNS が既に複数存在する。この現状を鑑みたとき、マーケターはマーケティングを取り巻く重大な環境変化の 1 である CGM 利用者数の拡大に対応することが必要となっていると言えよう。しかし、CGM を通じて情報を発信する、情報を収集する、いずれの立場に立ったとしても、予め確かめるべき事がある。それはマーケティングの最終的な目的のひとつである商品の売り上げと SNS 上のやり取りに関連があるのか？という点である。情報を発信する立場に立つとき、SNS に情報に拡散したとして、その情報拡散の成果と売上が結びついていないとしたら、その様な活動は企業活動にとって無駄な投資と言わざるを得ない。また情報を収集する際にも、新商品に関するマーケティング・アクションの開発のためにソーシャル・リスニングを通じて消費者を理解する、という目的であれば SNS 上のやり取りと売上の間にある関連について議論する必要はないだろう。しかし、既存商品に対する評価やその評価を用いて今後のマーケティング計画に関する修正を行うといった目的のためにソーシャル・リスニングを行うのであれば SNS 上のやり取りと売上の間にある関連の有無について知っておく必要があるだろう。

SNS はマーケターにとって直接的にはコントロールすることが不可能なメディアである。しかしながら、もし SNS におけるやり取りが、売上に影響しているとしたら、コントロールすることは出来なくとも、無視することは出来ない。また、SNS におけるやり取りが、売上とは無関係であることが確認されれば、SNS におけるコミュニケーションを無視したマーケティングを展開すべきであると言えよう。このように、量的・質的な変化をとげたインターネットとどのようにマーケターが向き合うかを考える上で、SNS と売上の関係を知ることは重要な研究テーマになり得ると思われる。

なお、本要旨では紙幅の都合、6 章の内容を中心に記載する。

2章 研究計画

本研究計画は大きく 7 つの項目によって構成される。その内容は①NPI データ（市場データ）の収集、②ブログ・Twitter データ収集プログラム改良（および収集）、③広告出稿量データの収集、④モデル開発・実証分析、⑤オムニバス調査、⑥調査結果分析、⑦学会等での成果報告・論文作成、となっている。

なお、実施スケジュールは①②③および⑤⑥を 2011 年度中、④を 2012 年度

前期、⑦を 2012 年度後期にそれぞれ実施した。

3章 CGM上のテキスト・データの収集

本研究では Twitter, ブログ上のテキスト・データ収集を、独自のプログラム (API) を Ruby 言語をベースに開発し、実施した。

Twitter 社は検索 API における出力を過去 7 日間に限っている。そのため、定期的にツイートを取得する環境が必要である。Twitter 上のテキスト・データ収集にあたり、本研究ではキーワードが含まれるツイートを 5 分ごと検索・収集するプログラムを開発することで対応した。

またブログ上のテキスト・データ収集は、Yahoo！ブログ検索を利用して、Ruby 上でブログ・データを取得するプログラムを作成した。また、記事に関してはアフィリエイトやバナー広告といった、記事とは本来関係ないテキストも含まれてしまう。そのような文字列を除外して記事のみを抽出する extractcontent.rb を利用し、不要なテキストを排除するなどの配慮を行った。

4章 分析対象カテゴリーとブランドの決定

本研究では分析に用いる商品カテゴリーを決定するために、オムニバス調査を実施した。その結果を分析し、各カテゴリーの購買に際して、Twitter 情報、ブログ情報がどの程度購買行動に影響を与えていたかを考察するため、Twitter、ブログの情報参照度を比較し、対象商品カテゴリーを決定した。

5章 出現単語の週次頻度変化に注目したTwitter上のトピック分類

5 章では「サッポロ金のオフ」「キリン午後の紅茶パンジェンシー」「サントリー伊右衛門グリーンエスプレッソ」のそれぞれにおいて、ユーザーの書き込み(文書) × 単語のデータ行列を作成し非負行列因子分解 NMF (Lee & Seung, 2000)により分析した。その結果、各製品についての週次での製品時系列を考慮した Twitter 上での書き込みのトピックが抽出された。そのトピックの内容は製品属性(アルコール飲料か茶系飲料か)により異なるものと、飲料として共通のトピックとが存在していることが分かった。また、抽出されたトピックの傾向は製品の時系列にそって変化していくことが示された。時系列にそった変化の中でも GRP の変動と広告に関するトピックの変動が連動していることが確認された。この点に関して、次章では GRP と Twitter の書き込み件数の連動につ

いてより詳しく検証する。

6章 商品の販売実績とCGM上のコミュニケーションの関連性分析

本研究では飲料や酒類等の一般消費財の Twitter 上の書き込みデータを利用し, Web コミュニケーション・データと商品の売上の関係性について分析した.

データは大きく, POS データ, GRP データ, Twitter データ, ブログ・データから構成される. 被説明変数である売上変数は POS データから, 説明変数である価格変数も POS データから, テレビ広告出稿量データは GRP データから, Twitter の書き込み件数は Twitter データから, ブログの書き込み件数はブログ・データからそれぞれ生成される. またいすれも週次のデータである.

POS データには NPI (全国 POS データ・インデックス) を用いた. NPI は財団法人流通経済研究所が提供する総合スーパー・マーケット, スーパー・マーケットを中心とした全国 360 店舗の POS データである. この NPI の関東・中部・関西のアイテム別週次販売実績データを用いた.

販売実績の変数としては点数 PI (来店客 1000 人あたりの販売点数) を用いた. GRP の調査対象地域と合致する京浜圏・中京圏・阪神圏のアイテム別週次 POS データから販売点数を抽出し, 分析対象店舗のレジ通過客数を用いて PI を算出した. 期間は 2011 年 9 月 12 日-2012 年 2 月 5 日である. なお対象商品は上述の様に既存のコミュニケーション活動の累積的な効果が与える影響を排除するために, 2011 年秋に市場投入され新商品を対象とした. 対象ブランドは「サッポロ金のオフ」「キリン午後の紅茶パンジェンシー」「サントリー伊右衛門グリーンエスプレッソ」「サントリーのんある気分」である. なお 2011 年 10 月前後の新商品販売直後の CM と, 翌年の新 CM の投入では, その内容や効果が異なる可能性があり, その事が分析結果に影響を与える危険性を考慮し, 2 月中までのデータを用いた.

価格の変数は同じく, POS データから算出した. 京浜圏・中京圏・阪神圏毎の当該商品の総販売金額を総販売点数で割ることで当該週における価格を算出した. なお分析に際してはブランド間で比較が出来る様に, 当該の商品の期間最大販売価格を 1 とする価格掛け率を算出し, 用いた.

Twitter データとして 3 章で述べた方法により収集した Twitter 上の書き込みデータを利用した. 対象商品名が記載された書き込みを Twitter 上から定期的に検索し, 期間中の対象商品に関する書き込みデータを収集するプログラム

を開発し、2011年9月から収集を開始した。ブログ・データも同様である。

広告出稿量データには株式会社ビデオリサーチ社が調査した関東・関西・中部の15秒換算の世帯GRP (Gross Rating Point: 延べ視聴率) をテレビ広告の出稿量データとして利用した。

本稿では、全部で4つの実証分析を実施した。実証分析Iでは、濱岡・里村(2009)の研究をベースに、ブログではなく新たにTwitterのデータを活用した分析結果について報告する。また実証分析ではTwitter書き込み数、GRP、売上実績の3点のバランスが良いブランドであった「金のオフ」を分析対象とするデータ分析をパイロットケースとして実施する。その結果を踏まえてモデルを調整し、実証分析I'ではモデルの内容を一般化するために他の複数ブランドのデータにモデルを適用する。

実証分析IIでは、実証分析Iの内容を踏まえ、パス解析を用いた分析を実施する。同様に実証分析II'では、実証分析IIのモデルの内容を一般化するために複数のブランドにパス解析を実施する。

なお紙幅の都合、ここでは実証分析I、実証分析IIの内容を掲載する。

実証分析Iの内容は下記の通りである。

はじめに一連の既存研究、並びに濱岡・里村(2009)における研究の課題に基づき、本研究におけるモデルを検討した。

本モデルにおける第1のポイントは、一般消費財を対象にした分析モデルである、という点である。既存研究における問題点は大半の研究が映画の興業成績を分析対象とした点にある。しかし、今後はより多くの商品分野における分析事例が蓄積されるべきである。そこで本研究では、映画だけではなく一般消費財を分析対象とする分析モデルの開発、および実証分析を実施した。

第2のポイントは、分析対象のCGMを既存研究が対象としたブログではなく、Twitter等のミニ・ブログタイプのSNSを対象にする点である。理由には1) ブログよりもTwitterなどのSNSがCGMの主体となりつつある点。また高関与商品を対象としたブログに対して、低関与商品を対象としたブログの数は少数だが、一方Twitterの場合、低関与商品を対象としたツイート数が一定数確保しやすい点。2) 本報告書の3章で述べたように、ブログからアフィリエイト広告などの文字情報を取り除き、本文のみを取り出すなどの加工が必要になる、また、字数に幅があり、写真などの非言語的な要素を用いた記載が多くテキスト・データとして取り扱いが比較的難しい点がある。一方、Twitterの場合、140

字という制限があり、本文以外の情報はほとんど無い。また、字数の幅は小さく、写真などを用いた場合もほぼ1枚程度であり、非言語的な要素も少なく、分析者として扱いやすいというメリットがある。

第3のポイントは、感情情報の考慮である。濱岡・里村（2009）における研究課題は前週のブログの書き込み件数に与える影響が有意になりながらも、その符号がマイナスを示すなど、書き込み件数と売上に関する正の関係を明らかにするには至っていない点にあると言えるだろう。この原因については、いくつかの理由が考えられるが、本研究ではこのような負の係数となった原因として、ポジティブな記述を行ったブログと、ネガティブな記述を行ったブログを識別せずに分析した結果であるという仮説を立てた。そこで本モデルにおいてはTwitterの書き込み内容に基づき、ポジティブな記述を行った書き込みと、ネガティブな記述を行った書き込みを識別したモデルを構築する。

以上の3点を加味して、次の式のようにモデルを構築した。

$$\begin{aligned}\ln \hat{y}_i = & b_0 + b_1 \ln x_1 + b_2 \ln x_2 + b_3 \ln x_3 \\ & + b_4 \ln x_4 + b_5 x_5 + b_6 x_6\end{aligned}$$

ここで

y_i ：当該週のブランド i の販売実績（点数 PI）

x_1 ：当該週の価格（価格掛け率）

x_2 ：当該週のTVCM出稿量（GRP）

x_3 ：当該週のポジティブなTwitter書き込み数

x_4 ：当該週のネガティブなTwitter書き込み数

x_5 ：阪神圏エリアダミー

x_6 ：中京圏エリアダミー

b_0-b_6 ：パラメータ

である。

濱岡・里村（2009）との違いは x_3 と x_4 が共にブログではなく Twitter の書き込み件数に変わっている点、またそれぞれがポジティブ、ネガティブに分かれている点である。

以上の分析モデルをデータに適用する。なお、テキスト・データ収集を行つ

たいくつかのブランドうち、最初に「金のオフ」の分析結果を分析する。複数のブランドの中で売上の面で成功しており、またツイートの中身に関してキャンペーンなどの企業側のアクションによる影響が少なく、消費者の生の声に近いツイートが多く存在する点により当該ブランドを最初の分析対象として選択した。本ブランドを用いた推定結果を踏まえてモデルを検討し、その後、各種ブランドのデータに適用する。

なお、テキストの感情の判定には SPSS Text Analysis for Surveys を利用し「良い」「悪い」「中立」の3つの感情に分け、「良い」「悪い」のテキスト・データの件数をカウントした。ただしいくつかの誤判定は目視にて調整を行った。

また、ブログではなく Twitter データの利用、感情情報の考慮の効果を検証するため、比較モデルとして、

- ・「価格・GRP」を用いたモデル 1
- ・「価格・GRP・ブログ書き込み件数」を用いたモデル 2,
- ・「価格・GRP・ツイート数」を用いたモデル 3,
- ・「価格・GRP・ツイート数（良い）」を用いたモデル 4,
- ・「価格・GRP・ツイート数（良い）・ツイート数（悪い）」を用いた本モデルを構築した。その上で、各モデルをデータに適用し、最小自乗法によってパラメータを推定した。その結果から、調整済 R2 乗値を比較し、変数投入の影響を評価することで、モデルを選択する。調整済 R2 乗値を用いる理由は、各モデル間で変数の数が異なるため、モデルにおける変数の数に対する適合度を比較する必要があるためである。

最小二乗法を用いたモデル毎の調整済 R2 乗値は下記の通りである。

- ・モデル 1（価格・GRP）・・・0.594
- ・モデル 2（価格・GRP・ブログ）・・・0.589
- ・モデル 3（価格・GRP・ツイート数）・・・0.639
- ・モデル 4（価格・GRP・ツイート数（良い））・・・0.640
- ・本モデル（価格・GRP・ツイート数（良い）・ツイート数（悪い））
・・・0.635

この結果について、比較検討する。最も調整済 R2 乗値が悪いのはモデル 2 であった。ブログの件数のデータを投入したが、投入する以前のモデル 1 の方が良い指標となっている。ブログ変数の投入に見合う適合度の向上が見られない、という結果になっている。また、その次に調整済 R2 乗値の指標は CGM の変

数を投入しなかったモデル 1, 本モデルが続いている。想定した結果と異なり, Twitter の「悪い」書き込み件数に関する変数を含まないモデル 3, モデル 4 の方が変数の数に対して適合度が高いという結果になっている。最終的にはモデル 3, モデル 4 がほぼ同程度に適合度が高いという結果になっている。若干ではあるがモデル 4 が高くなっているため本研究ではモデル 4 を採択する。

モデル 4 におけるパラメータの推定結果は表 1 の通りである。

表 1 実証分析 I パラメータ推定結果

	非標準化係数	標準化係数	t 値	有意確率
定数項	38.893	-	2.956	0.005
x_1 価格	-6.197	-0.304	-3.055	0.003
x_2 GRP	-0.014	-0.094	-0.819	0.417
x_3 ツイート（良い）	0.192	0.324	2.841	0.006
x_5 阪神圏	-0.871	-0.672	-7.450	0.000
x_6 中京圏	-0.111	-0.085	-0.792	0.432

この結果に関して、5%水準で各変数の評価を行う。

x_1 価格は 5%水準で有意となっており、係数の符号は負を示している。通常価格が下がれば、売上が上がる所以、符号は想定通りの結果となっている。また x_5 , x_6 の地域ダミー変数を除くマーケティング変数の中では標準化係数の値が最も大きくなっている、売上に対して大きな影響を与えていていることが分かる。

x_2 GRP は 5%水準で非有意となっている。一般に多くのマーケティング・サイエンス研究の論文においても当期の広告効果変数は非有意となる場合があるが、本研究も同様の結果となっている。

x_3 ツイート（良い）書き込み変数は 5%水準で有意となっている。符号は正を示しており、濱岡・里村（2009）モデルでは実証に至らなかった Web コミュニケーションと売上実績における正の影響を捕捉した。標準化係数の値は、価格と大きく変わらない数値を示しており、価格に次いで売上との関連の強い変数であることが分かる。

x_5 , x_6 の地域ダミー変数は阪神圏が有意、かつ負の符号を示しており、京浜

圏を基準にすると「金のオフ」の商品ならびにコミュニケーションを含むマーケティング展開があまり支持されていない傾向にあることが分かる。中京圏については京浜圏との有意な差が無い、という結果になっている。

以上の様に、パイロットケースとしての「金のオフ」の分析結果では、一般消費財を対象とした分析を実施し、Twitter の書き込み変数を用いて濱岡・里村（2009）モデルでは実証に至らなかった Web コミュニケーションと売上実績における正の影響を捕捉した。しかし、感情情報を用いた結果は、感情情報を用いなかった調整済 R2 乗値に大きな違いはないという結果に至った。統計上はモデルにおける変数の数に対する適合度を比較して、若干、調整済 R2 乗値の良いモデル 4 を採択したが、実務的には Twitter の感情判定に目視を含む多くの作業工数がかかる事から調整済 R2 乗値がほとんど変わらないモデル 3 が最も優れたモデルであると言えよう。そこで「金のオフ」によるパイロットケースの分析結果を他のブランドに展開する実証分析ではモデル 3 をデータに適用した。実証分析 I' では I とほぼ同様の分析結果が得られた。

ここで、実証分析 I と実証分析 I' から抽出された課題について検討したい。

まず第 1 の課題は Twitter 書込み変数の意味合いについてである。Twitter の書き込み件数は売上に有意の影響を与えている。しかし、Twitter の原文には購入したことをつぶやく、もしくは消費したことをつぶやくツイートが散見される。つまり、多数の購入が発生することで、つぶやきが生まれるという、モデルが想定する結果とは逆の書き込みが多く見られている。このことからつぶやくから売上が伸びるのか、売上が伸びるからつぶやくのか、どちらの構造にあるかを特定する必要が出てくる。

また第 2 の課題は GRP の影響に関する抽出である。GRP の係数は 4 ブランドのうち「のんある気分」を除く、パイロットケースの「金のオフ」、及び「パンジエンシー」「グリーンエスプレッソ」においていずれも非有意という推定結果になった。しかしながら、Twitter の原文にはテレビ広告を見た事に関するつぶやきが散見される。5 章での Twitter 上のトピック分類においても CM に関連したトピックが抽出されている。このことからも当期の GRP の売上に対する有意な影響を検出できなかつたが、何らかの影響を及ぼしていると想定される。よって、売上と有意な関連を有する Twitter 上での書き込み促進要因としてのマス・コミュニケーション効果について確認する必要がある。

実証分析 II の内容は下記の通りである。

以上の 2 点の課題を解決するために、パス解析によるモデルを検討した。
パス解析では実証分析 I の回帰モデルをベースとしつつ、パス解析により 1) 双方向モデル、2) 間接効果モデルを加味したモデルを構築する。

まず 1) 双方向モデルを構築することで、課題である Twitter の書き込みは原因と結果のいずれであるかを特定する。そのためにパス解析において「点数 PI→Twitter 書き込み件数」「Twitter 書き込み件数→点数 PI」の双方を仮定したモデルを構築し検証する。また 2) 間接効果モデルを構築することで売上と有意な関連を有する Twitter 上での書き込み促進要因としてのマス・コミュニケーション効果を検出する。そのためにパス解析において「GRP→Twitter 書き込み件数」「Twitter 書き込み件数→点数 PI」の間接効果を仮定したモデルを構築する。

実証分析 I と同様にパイロットケースとして「金のオフ」のデータを分析し、その結果を踏まえ他のブランドのデータにモデルを適用する。

パス解析は実証分析 I で用いた変数をそのまま利用する。「金のオフ」のデータを対象にした回帰モデルにおいて採択されたモデル 4 において利用した「価格」「当該週の TVCM 出稿量」「当該週のポジティブな Twitter 書き込み数」「阪神圏エリアダミー」「中京圏エリアダミー」を用いる（従って「当該週のネガティブな Twitter 書き込み数」は除く）。なお、上記の 2 つの仮説以外に回帰分析と異なる点として、パス解析では、外生変数間の相関関係と内生変数に誤差項を含めるのが一般的であるため、そのようにモデルを構築している。

以上のモデルをデータに適用した結果が図 1 である。

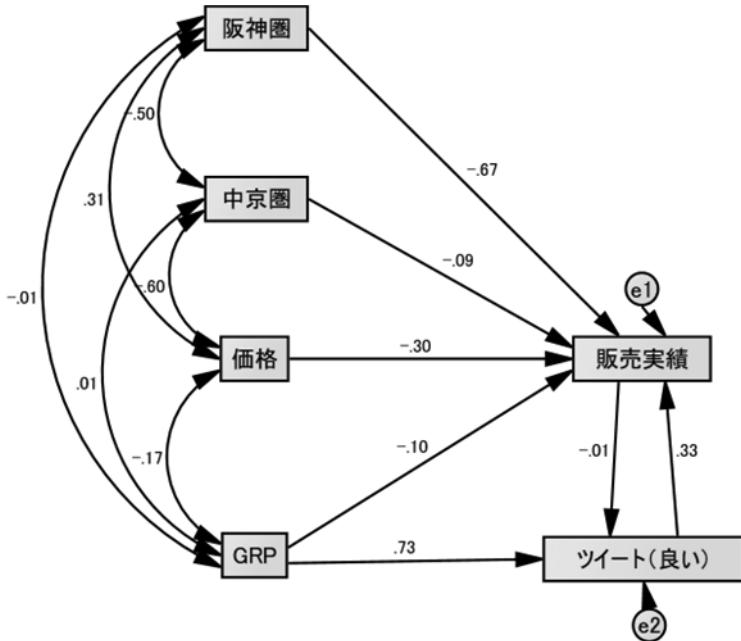


図1 パス解析：標準化係数（金のオフ）

なお GFI : 1.000, AGFI : 1.000, RESEA : 0.000 となっており、高い適合度を示している。

この結果に関して、GRP、ツイート書き込み件数と販売実績の関係を中心に整理したい。

まず 1) 双方向モデルの係数から Twitter の書き込みは原因と結果のいずれであるかを特定する。「販売実績からツイート（良い）」へのパス係数は5%水準で非有意であった。逆に「ツイート（良い）から販売実績」へのパス係数は5%水準で有意であり、0.33と正の符号を示している。このことから、販売実績が伸びるからツイート数が増えるのではなく、ツイート数が増えることで販売実績が伸びる、と言う因果の方向が確認された。

また、2) 間接効果モデルを構築することで売上と有意な関連を有するTwitter 上の書き込み促進要因としてのマス・コミュニケーション効果を検出する。「GRP から販売実績」のパス係数は回帰モデルの結果と同様に5%水準で非有意であった。しかし「GRP からツイート」のパス係数は5%水準で有意であり、

係数は 0.73 となっている。この事から GRP はツイートの発生に有意に影響を与えていていると言えよう。また、上述の様に「ツイート（良い）から販売実績」へのパス係数は 5% 水準で有意であることから、Twitter を経由した TV 広告が売上に与える間接効果が確認された。

以上の実証分析の結果を受けて、マーケターが Twitter 上のコミュニケーションをどのように捉えるべきかを考察したい。本研究で分析対象とした一般消費財の各種ブランドに関するツイート数は多くても 1 ヶ月あたり数千ツイートである。今日では日本語だけでも 1 ヶ月あたり 10 億超のツイートがあり、その様な膨大なツイートの中で、商品に関する数千ツイートが販売実績を動かすほどの情報伝達機能を有するとは想定しづらい。しかし、分析結果はツイート数が販売実績と正の有意な関係を有していることを示している。この結果から、本研究では Twitter における各種コミュニケーションは当該商品や広告が持つ「話題性の代理指標」とである、という仮説を導出した。Twitter 上に現れる各種商品・広告に関するつぶやきは、リアル、ネットを問わず、消費者間で交わされるクチコミや評価が膨大な商品に対する話題性が表面化した、いわば「話題性の氷山の一角」であるという仮説が成り立つと考えられる。商品に対する話題性が豊富であれば、販売実績が増える可能性は大きくなり、乏しければ販売実績が増える可能性は小さくなる。この関係は至極当然の関係であるが、その話題性に関して従来は捕捉することが困難であった。しかし、今日 Twitter の登場によってその全てではないが、一部を把握することが可能になった。今後のマーケターは、Twitter を単なるコミュニケーション・チャネルとしてだけではなく、商品の販売実績と連動した消費者の生の声を収集する事が出来るマーケティング調査チャネルとしても位置づけるべきであると言えよう。

7 章 結論と課題

もし仮に、ツイートと商品の販売実績に何ら関係が無ければ、マーケティング関係者は CGM 全般に関して、特別な注意を払わずとも良いだろう。しかし、本研究の成果が示す様に、ツイート数と販売実績には有意な正の関係がある。また CGM 上の消費者のやり取りの中では一定規模の広告に関するトピックが含まれていることが確認され、更にパス解析によって TV 広告出稿量がツイートを通じて、有意に売上に影響を与えていることが確認された。つまり、ツイートは商品の販売実績と関連があり、またそのツイートも既存のマーケティング活

動と密接に関連している。以上の結果から、次の様に結論付けたい。これからマーケターは売上と有意に連動する CGM 上の消費者のやり取りに注意を払い、その声を聞くべきである。CGM 上のやり取りに対して、積極的に関与する必要がある。

今後、マーケターは CGM を、1) 消費者の動向を知る、もしくは広告や商品の話題性を測る調査チャネルとして、2) 販売実績と関連するコミュニケーション・チャネルとして見なし、各種マーケティング計画を計画すべきである。

<引用文献>

- 濱岡豊, 里村卓也, (2009), 「消費者間の相互作用についての基礎研究—クチコミ, e クチコミを中心に」慶應義塾大学出版会.
- Lee, D.D. and Seung, H.S. (2000). Algorithms for Non-Negative Matrix Factorization. In K. T. Leen, T. G. Dietterich and V. Tresp (Eds.), *Advances in Neural Information Processing Systems* (vol. 13, 556–562). MIT Press.

謝辞

本研究は公益財団法人 吉田秀雄記念事業財団の第 45 次研究助成を受けて実施したものである。関係者の皆様に厚く御礼申し上げます。

また、テレビ廣告出稿データに関しては株式会社ビデオリサーチからご提供頂いた。ここに記して厚く御礼申し上げます。