

口コミ指数による事例類型化に基づく複数メディアの ヒット前の露出を先行指標とした情報拡散過程の分析

大知 正直 (東京大学大学院 工学系研究科) 長濱 憲 (東京大学大学院学際情報学府, 株式会社 電通パブリック
リレーションズ) 榎 剛史 (東京大学大学院 工学系研究科, 株式会社ホット
リンク) 森 純一郎 (東京大学大学院 工学系研究科) 坂田 一郎 (東京大学大学院 工学系研究科)

1. はじめに

近年、ソーシャルメディア、マスメディアが人々に及ぼす影響について盛んに分析されている。その背景の一部として、広報や広告の実務において、商品やサービスを多くの人々に認知・利用してもらうための効果的なコミュニケーション手段が求められている状況がある。以前は、新聞、雑誌、テレビなどのマスメディアを通じてしか、多くの人々の認知を促し関心を喚起することができなかった。しかし近年はソーシャルメディアの発達にともなって、マスメディアに留まらず様々なメディアを通して人々への認知が広がるようになってきた。また、インターネットの普及によって、商品やサービスが幅広い消費者の間で話題になり、国境を越えて情報が共有されることも増えつつある。これによって人々は単に消費するという意味での消費者ではなく、消費した商品・サービスを他人に対し発信する役割を以前より果たしていると言えるだろう。本稿では、消費だけでなく発信を行う人々という意味で以降、対象とする人々を生活者と呼ぶこととする。

最近では、生活者がソーシャルメディアを通して発信した情報が爆発的に拡散し、消費者の商品・サービス利用に影響を及ぼすことが認識されてきた。そして、その過程について、ダイナミックな分析手法が求められるようになってきている。しかしこれまで、このような爆発的な情報拡散（以降、ヒット現象と呼ぶ）の発生過程について経験則で語られることはあっても、データに基づいた検証がなされることは少なかった。一方、情報工学の分野では、ソーシャルメディア上での情報拡散を動的なモデルを用いて分析を行う試みがなされている。本稿では、情報拡散モデルと、複数のメディア上で露出した実際のヒット現象のデータを利用することで、これまでより早い段階でその後の情報拡散過程を説明することを目的とする。また、「口コミ指数」を提案することで、定量的にヒット現象を分類し、どのような情報発信手段を用いれば効果的な情報拡散を可能にするかを明らかにする。

本研究では、商品やサービスに対する生活者の関心が増大していく過程に、マスメディアやソーシャルメディア上での商品・サービス情報の露出が、どのような影響を及ぼしているのかを明らかにする。そのために、商品・サービスに対する生活者の関心をあらわす指標として、Google の検索数を用いる。近年の研究成果では、例えば選挙時の有権者の投票行動は、事前の検索結果から一定程度予測可能なことが示されている¹⁾。そのため、検索数をもって生活者の関心のみならず、実際の商品・サービス利用状況の指標とすることは妥当であると考えられる。

また、商品・サービスに対する生活者の関心が増大していく過程を分析するに際して、感染症の流行過程をモデル化した手法を用いてソーシャルメディアを利用するユーザ数の推移を予測する研究 (John and Joshua 2014) の考え方を援用する。この研究は、ある一時点までの検索数の時系列データを元に、その後の検索数の変化を予測するものであり、モデル内部の既知の変数を事後の予測に用いているという特徴がある。そのため、精度の高いモデルを構築するためにはヒット現象が起きた後での予測を行う必要がある。それに対して、本研究では、ヒット現象が起きる前段階で、その予兆がマスメディア、ソーシャルメディア上で表れていると考えられることで、より早期の段階でのヒット現象の説明を行うことを目指した。ただし、ヒット商品は、リリースさ

れた商品のほんの一部であるので、ヒット商品と同様の予兆を示していたが、実際にはヒットしなかった商品が多数存在することが考えられる。ヒットの予兆を示した商品がその後どの程度ヒットするかを予測するには、ヒットの予兆を示したがヒットしなかった商品群を含めた上での分析が必要であろう。しかし、ヒットの予兆というものが明らかになっていない段階で、ヒットしなかった商品群を特定することは難しい。本研究は、このような理由から、ヒットの程度を予測することに主眼を置くのではなく、ヒットした商品群に共通する、ヒットの予兆とは何かを明らかにすることを目的とする。さらに本研究では商品・サービスの情報拡散について、「口コミ指数」を提案し利用することで、個別の流行現象をマスメディア上の露出によって流行したタイプと、生活者の口コミによって流行したタイプに分類した。これにより、商品やサービスの持つ性質に応じた商品PRの方策を検討するための指針とすることができるだろう。

本論文は以下のように構成される。まず2章で関連研究について述べ、本研究の位置づけや、有用性や新規性について議論を行う。3章で提案手法の背景と根拠について議論し、提案手法について説明を行う。その後、4章で使用するデータと実験について説明する。5章で結果を示す。結果について6章で考察し、本研究で明らかになったことを述べる。最後に本研究についてまとめ、貢献と今後の研究方針について述べる。

2. 関連研究

ここでは、関連する研究分野として、インターネットによる情報拡散への影響の研究、情報拡散に関する研究について述べ、検索行動と実世界の関連性について議論したあと、本研究の位置づけについて明確にする。

2.1. インターネットによる情報拡散への影響

1990年代からのインターネットの普及によって、生活者は自身の持つ情報を全世界に発信できるようになった。遠藤はネットメディアの出現によって誕生した社会を「間メディア社会」と述べている(遠藤 2010)。「間メディア社会」では、ネットメディアの誕生によって従来のマスメディアが消滅するのではなく、両メディアが並立することで、情報拡散の回路が拡張され、多層化していく(遠藤 2014)。

さらに、インターネットによる情報拡散への影響として特筆すべき点は、不特定のユーザによってインターネット上に投稿される口コミの発生である。口コミ(Word of Mouth、以下 WoM)とは、会話を通じた人から人への情報の伝達を指す。特にインターネットを通じた口コミは電子的口コミ情報(electronic Word of Mouth、以下 eWoM)と呼ばれ(Kietzmann and Canhoto 2013)、近年のインターネット利用機会の増大とともに重視されるようになってきた。

電子的口コミ情報に関する研究は、大別すると2つの内容に集約される。1つはネット上の口コミの内容に関する研究であり、もう1つはネット上に口コミを投稿した人物同士のネットワークに関する研究である。

まず、ネット上の口コミの内容に関する研究としては、投稿された口コミの内容を評判に関する情報として捉え、例えば、肯定的であるか否定的であるかというようなセンチメントを検出する研究が挙げられる。Hatzivassiloglou らや Turney らによって行われたセンチメントを検出する研究は、初期に予め設定した“excellent”や“poor”といった種となる単語を元に、何らかの指標に類似した語、共起する語彙を検出し、文章全体が肯定的か否定的かを判定する(Hatzivassiloglou and McKeown 1997, Turney 2002)。また、商品情報のどの内容に興味をもたれているかを検出する研究分野もあり、Ghose らは、Amazon の本のレビューに対し“helpful”であると認めたユーザの数を利用し、レビューをランキングすることで、読み手(購入を検討するユーザ)にとって有用なレビューを検出する試みを行っている(Ghose and Ipeirotis 2007)。さらに、実際に投稿された内容を細かく分析することで、分析対象とする商品のどの部分に関心が持たれているかを検出する研究もある。こうした分析は、意見要約に関する研究の1分野とみなすことができ、多くの研究が蓄積されている(Liu 2012)。

口コミのネットワークに注目した研究は、Lazarsfeld や Katz らによって提唱された2ステップフローコ

コミュニケーションモデルに基づいている (Lazarsfeld et al. 1948, Katz and Paul 1955)。このモデルは、メディアがオピニオンリーダーを通じて人々に影響を与えるという仮説である。後述するネットワーク研究の進展に伴い、この考え方を発展させたモデルに基づいてネットワーク上で影響力を持つ人物を検出しようという研究が行われ (Bakshy, Hofman, Mason and Watts 2011)、ウェブ サービスとしても公開されるに至った²⁾。口コミのネットワークに関する分析結果を利用したウェブ 上のサービスは、マーケティング担当者のソーシャルリスニングのツールとして重要性を増しており、近年注目されている³⁾。

しかし、これまで述べてきた従来の研究は、マスメディアや口コミ、インターネット上の口コミの影響についてある一時点や一定期間の分析は可能でも、情報が拡散する動的な過程をモデル化し分析するには限界があった。

2.2. 情報拡散過程に関する研究

本研究では、商品・サービスに関する情報がどのように拡散し、ヒット現象を生み出しているのかについて分析を行う。こうした研究は情報工学の研究対象の一つである、情報拡散に関する研究の1つと捉えることができる。ここでは、特にネットワークに関する分野を中心に、情報工学分野を中心として、情報拡散に関するこれまでの研究と本研究の関連性について述べる。

情報拡散モデルはさまざまな研究で提案されており、消費者行動論の文脈で、商品の購買行動が増える過程をモデルとして定式化した研究も多い (Robertson 1967, Frank 1969)。特にバスマodelとして知られる購買予測モデルでは、耐久消費財などの新商品の購入意欲を、他人にまどわされない購入意欲 (Innovation 効果) と他の購買者数からの影響による効果 (Imitation 効果) の和で表現できると提案している (Frank 1969)。しかし、この形式のモデルは人々の持つネットワークについて十分に考慮していない。

近年、情報拡散過程のモデル化に関連して、世界的にネットワーク分野の研究の進展が見られている。もともと同分野はグラフ理論の研究を基盤とし、近年はコンピュータを利用し実データに応用することで、インターネットのみならず生態系や電力供給網など、様々な分野で進展をみせている。特に、Watts らによる「スモール・ワールド・ネットワーク」の研究 (Watts 1999) と、(Albert and Barabási 2002) らによる「スケール・フリー・ネットワーク」の研究によって、同分野の研究が一気に進展することとなった。例えば、ツイッターという SNS の持つメディア性、リツイートネットワークについて議論を行った報告がされている (Kwak, Lee, Park and Moon 2010)。関連して、ツイッターに注目し、ツイッターを利用しているユーザをさまざまに分類することで属性を明らかにしようとする研究も存在している (Alan, Sune, Yong-Yeol, Jukka-Pekka and Rosenquist 2011, Java, Song, Finin and Tseng 2007)。

人々の持つネットワークに関するこのような研究の蓄積によって、情報拡散過程のモデル化も進展することになった。SIR モデル (Kermack and McKendrick 1927) は、感染症の流行過程をモデル化したもので、情報拡散過程のモデル化にも応用されている。例えば、John らの研究 (John and Joshua 2014) では、SNS の利用者数の変化を、『Google トレンド⁴⁾』を元に取得した人々の検索数の履歴から、SIR モデルを拡張したモデルを利用して予測を行っている。また、Takayasu らは、東日本大震災後にネット上で拡散した情報について観察し、報告を行っている (Takayasu, Sato, Sano, Yamada, Miura and Takayasu 2015)。このように SIR モデルは免疫を獲得することが前提で二度と同じ感染症に罹患しないことが前提となっており、これらの分析も一度だけ発生する現象に対して適用されている。

感染モデルに関する調査文献 (Britton 2010) によると、SIR モデルを始めとした感染症モデルにはさまざまなものがあり、免疫を獲得しない病気に適用可能な SIS モデルや潜伏期間を考慮する SEIR モデルなどが代表的なものとして紹介されている。しかし、本研究では個々の流行現象を 3.2 節において観察した結果、全く同一の内容での流行は起こっていないことがわかった。そのため、同一の流行 (一種の病気とする) には二度とかからない仮定を置くこととし、SIR モデルを適用することとした。

2.3. 実際の現象と検索行動の相関

本研究では、商品に関する生活者の持つ関心の変化のデータとして「Googleトレンド」の人気度の時系列の推移を利用している。ここでは、こうしたインターネット上の人々の検索行動と実世界上の人々の行動に相関があることを示す。

代表的な研究として、Ginsbergらが報告している（Ginsberg, Mohebbi, Patel, Brammer, Smolinski and Brilliant 2009）インフルエンザの流行過程に関する研究（“Google Flu Trends”、以下、GFT）があげられる。この研究では、実際のインフルエンザの流行と検索数の変化は一致していることが示されている。ただし、この報告では、病院を訪問したインフルエンザの患者数の推移と一致させるために自動的に抽出された複数の検索クエリの推移を多変量解析の手法を用いて最適化させている。これによって、人々の検索行動が実際のインフルエンザの発病と相関していることが認められる。一方で、この手法の精度にはLazerらによって疑問が提示されている（Lazer, Kennedy, King and Vespignani 2014）。これによると、前の年に作成されたモデルでは、その翌年は高い精度の予測ができなかったことが示されている。これはモデルフィッティングの段階で高い精度を達成するために過剰適合が起こっていることや、メディアでの過剰な露出（ニュースでの報道）と実際の発症数との乖離が起こっていることが原因である。このように人々の検索行動と実際の行動の相関で汎用的なモデルを構築することは未だ困難を伴うが、少なくともある時点での相関を説明することはかなり高い精度で達成できることが示されている。

今回、我々は商品名を表す単一の検索語のみの「Googleトレンド」の人気度の推移を人々の実際の人気を表すものとして仮定している。GFTでは複数の検索語の組み合わせでフィッティングさせることで高い精度を達成しているため、単一の検索語では精度の低下が懸念される。しかし、Mohebiらによって報告された、インフルエンザ以外にも含めたさまざまな現象と実際の検索語群の推移との多変量解析の結果を示す「Google Correlate⁵⁾」サービスによる分析結果によると、インフルエンザの場合、“influenza type a”、“symptoms of flu”等のかかなり直接的な検索クエリが0.9の相関係数を示していることが報告されている（Mohebbi, Vanderkam, Kodysh, Schonberger, Choi and Kumar 2011）。よって、商品名を表す単一の検索語の「Googleトレンド」の結果は、実際の人々の持つ関心と高い相関を持っていると考えられる。

本研究では、感染モデルと情報拡散に関する研究で得られたモデル構築手法を元に、国内のヒット現象に関する複数のメディア上に露出しているデータを用いることで、ヒット前の段階の様々なメディアへの露出が、人々の関心にどの程度影響を与えるか、または、ヒット現象がどのように推移するか説明できることを示す。

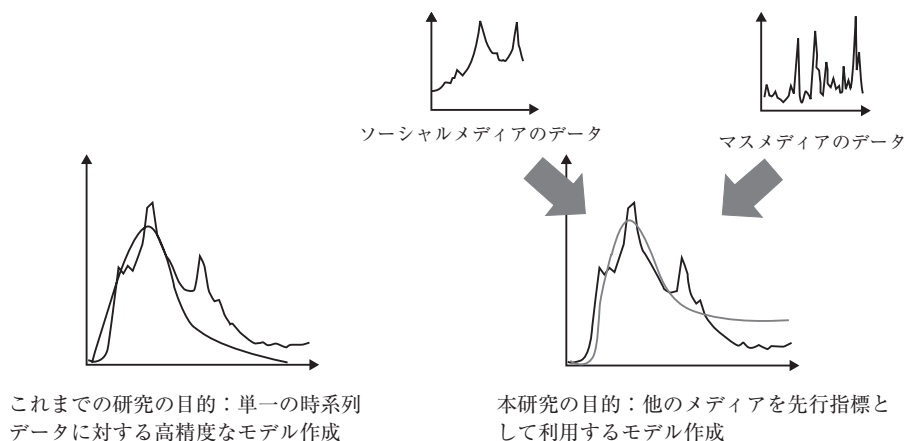
3. 提案手法

本研究では、『日経トレンディ』が発表した“2014年ヒット商品ベスト30⁶⁾”の上位10商品を取り上げ、ヒット前のメディアの露出の中から、商品のヒット現象の過程に早い段階で影響を及ぼす指標を明らかにする。以下では、そうした分析を可能とする原理について説明を行い、指標を提案する。

3.1. 手法の動機

商品・サービスに関する情報の拡散過程に関するモデルとして、バスモデル（Frank 1969）やSIRモデル（Ginsberg et al. 2009）が挙げられる。これらのモデルは、商品・サービスの購入には、ユーザ間でのコミュニケーションが影響を及ぼしているという仮定を伴っている。石井ら（石井・吉田 2005）はユーザ間のコミュニケーションの効果を2つに分けて、バスモデルを拡張している。1つはユーザ同士の直接の口コミ等の影響である「直接コミュニケーション」、もう1つは、ユーザ同士が様々なメディアから受ける影響である「間接コミュニケーション」である。間接コミュニケーションは、対象への評判やうわさを表し、ソーシャルメディア上の口コミの効果を包含している。彼らはこの「間接コミュニケーション」の重要性を主張しており、この効果を表す項を追加することによって、売上の推移データを高い精度でモデル化できることを示した。

図1 本研究の目的と過去の研究との違い



しかし、我々は予測したいデータのみを用いて必ずしもモデリングを行う必要は無い。例えばある映画の入場者数を前もって予測したいと考えた場合、ツイッターやブログでの言及、TVでの露出量は先行指標として有用であろう。特にソーシャルメディア上では公開前にファンの人々のコミュニティで盛り上がっているだろうことは想像に難くない。そこで、本研究では、こうしたマスメディアやソーシャルメディアでの露出量を用いて、予測したい対象への情報拡散現象の発生をモデル化することを目指す。本研究の目的と過去の研究の目的との違いを図1に示す。我々は図に示すようにソーシャルメディアやマスメディアでの露出を先行指標として利用することで、対象とするデータの推移の説明を可能にするモデルの作成を目指す。

説明したい対象は購買数や売上高の推移であるが、実際のデータを自由に取得することは難しい。そこで、Kotlerらの提唱した5段階の購買プロセスのモデル (Kotler and Turner 1979) やAIDMA、AISASなどの購買意思決定モデルに基づいて、購買の一段階前のデータを利用する。AISASモデルは、近年のインターネット環境の普及を前提としたモデルであるので、今回はこのモデル内の「Action (購買)」の前段階である、「Search (検索)」の数を実際のユーザーの関心を表す指標と捉え、この推移を説明する指標を提案することを目指す。つまり、特定の商品・サービスについて検索を行った人の数を、当該商品に一定の関心を持った人の数 (と相関関係の強い変数) と想定して議論を進める。

本研究では、人々への関心の拡散を感染モデルとして捉える。そして、ヒット現象はたくさんの人々が同時に興味・関心を持っている (いわば、ある種の病気に感染している) 状態と考える。感染モデルでは、たくさんの人々が同時に興味・関心を持っているヒット現象が起こっている状態は、興味を持ちやすく、飽きのこない商品を多くの人々が認知していることで発生していると解釈することができるだろう。つまり、「興味を持つ可能性のある人の総数」、「興味・関心の持ちやすさ」、「飽きやすさ」という3つのパラメータによって決定される。そして、本研究では、これら3つのパラメータを決定づける要因が、他のメディア上での情報の露出に表れていると仮説を設定し、早い段階でのメディアの露出から提案する指標とこれらのパラメータとの相関の分析を行う。これによって、より早い段階でのヒット商品・サービスの情報拡散過程を説明できるモデルの構築を図る。

なお、感染現象を説明するモデルには、SISモデルという同じ病に繰り返し罹患するモデルもある。しかし、本研究ではSIRモデルを適用した。これは、1つの内容に関するヒット現象は一度しか起こらないという仮定にもとづいているからである。この理由を以下に説明する。

3.2. SIRモデルを適用する理由

例として「妖怪ウォッチ」のGoogleトレンドにおける人気度の変化を図2に示す。図2に示したとおり、「妖怪ウォッチ」は複数的人气度のピークを持っており、長い期間に、繰り返し流行していることがわかる。さて、それぞれ的人气度のピークにおいて、人々がどういう興味を持っているのだろうか。簡易的に確認する

図2 “妖怪ウォッチ”のGoogleトレンド上の人気度の時系列の変化を表したグラフ

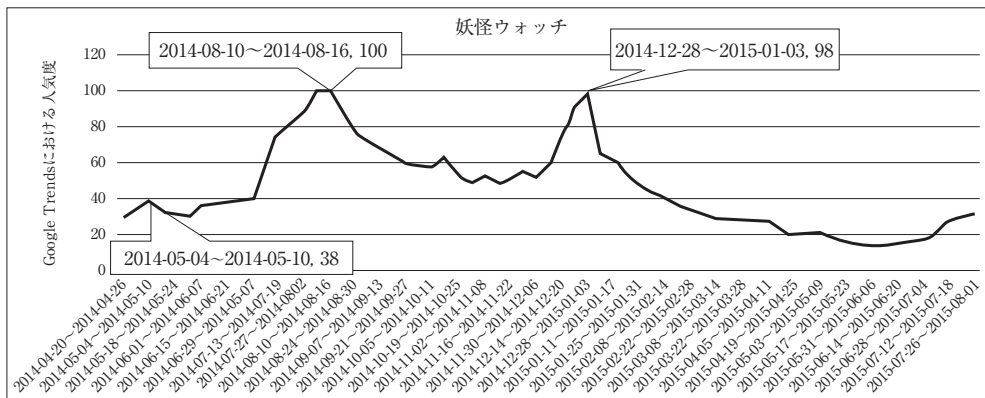


表1 Googleトレンド上で急上昇ワードとして上がっているクエリ

順位	2014年05月の上昇クエリ	上昇率	順位	2014年08月の上昇クエリ	上昇率	順位	2014年12月の上昇クエリ	上昇率
1	ゲラゲラボーのうた	+160%	1	ガブニャン	over +5000%	1	ジバコマ	over +5000%
2	ゲラゲラボー	+150%	2	マクドナルド妖怪ウォッチ	+300%	2	ダークニャン	+600%
3	妖怪ウォッチ時計	+80%	3	マック	+250%	3	妖怪ウォッチ真打	+300%
4	キャラ弁	+70%	4	妖怪ウォッチラムネ	+250%	4	真打	+300%
5	妖怪ウォッチキャラクター	+70%	5	マック妖怪ウォッチ	+190%	5	真打妖怪ウォッチ	+300%

ために、流行時点で人々がどのような検索語を実際に入力したかの調査を、Googleトレンドを利用して行った。Googleトレンドでは検索語と共起する語を確認することができる。ここでは、そのときどきの人々の関心を調査するために、ピークの日を含む月に検索語の「妖怪ウォッチ」とともにそれ以前とくらべて検索されている語を確認した。その結果を表1に示す。表1を見ると、2014年5月では、「ゲラゲラボーのうた」（妖怪ウォッチの主題歌のこと）や「キャラ弁」（コンテンツに登場するキャラクターを模した弁当のこと）について興味を持たれていることがわかる。次に2014年8月では、「ガブニャン」（ゲーム「妖怪ウォッチ2 元祖 本家（2014年7月10日発売）」に登場するキャラクターのこと、希少性が高い）、「マクドナルド」（当時、マクドナルドの商品が妖怪ウォッチ関連のグッズとタイアップするキャンペーンを行っていた）に対して興味を持たれていることがわかる。2014年12月では、「ジバコマ」、「ダークニャン」（ともに2014年12月13日発売のゲーム「妖怪ウォッチ2 真打」に登場するキャラクターのこと、ゲーム内で主人公と友人関係を形成する方法が特殊）に対して興味を持たれていることがわかる。このように繰り返し流行している「妖怪ウォッチ」において、そのときどきにおいて人々が関心を持っている内容はすべて異なっている。

この調査を今回対象とする商品群に行った結果、すべての商品の持つそれぞれのピークに対して、すべて異なる内容で流行が起こっていることを確認した。この調査結果を元に、本研究では1つの商品に対して繰り返し起こる流行（人気度のピーク）を1つの現象として捉えるのではなく、すべて異なる現象として捉えることとした。つまり、1つの流行は1つの人気度のピークを持つもので、厳密な意味（商品のある側面に関心を持たれている流行現象）で同一の流行現象は起きないと考え、それぞれの流行に対して、SIRモデルを適用することとした。

3.3. SIRモデル

本研究では、SIRモデルによって人々の関心の変化の近似を行う。ここでは、SIRモデルの説明とパラメータの推定方法について説明を行う。

SIRモデルは、Kermackら（Kermack and McKendrick 1927）によって1927年に発表された一階微分方程式によって記述されるモデルで、感染症の流行過程を近似させることを目的としたものである。SIRモデルは以下の3つの一階微分方程式によって表現される。

$$\frac{d}{dt}S(t) = -\beta S(t)I(t) \quad (1)$$

$$\frac{d}{dt}I(t) = \beta S(t)I(t) - \gamma R(t) \quad (2)$$

$$\frac{d}{dt}R(t) = \gamma R(t) \quad (3)$$

SIR モデルでは、ある時点 t での一定数の感受性保持者 $S(t)$ が感染者 $I(t)$ から一定の割合で感染し、感染者 $I(t)$ から一定の割合で回復し、免疫保持者 $R(t)$ となる仮定をモデル化したものである。本研究では、3.2 節の分析の結果、人々は同一商品・サービスについて一度感染した（関心を持ち検索した）後に、同じ話題で再び感染しない（関心を持ち検索しない）ことが観察されたため、繰り返し感染することの無い SIR モデルを採用している。ここで、各式中の β を感染率、 γ を回復率、 $S_0 = S(0)$ を初期感受性保持者とする。本研究では、人々がある商品・サービスの情報に関心を持ち検索した状態にある種の病気に感染している状態とみなす。つまり、ある時点 t で関心を持ち検索を行う人々の数を感染者数 $I(t)$ とする。同様に類推すると、感染率 β は値が大きいほど急激に人々の関心を得ることになるので、情報の拡散力を示す。また、回復率 γ は値が大きいほど急激に人々の関心が無くなることになるので、情報拡散の減衰力を表す。そして、初期感受性保持者数はその商品に関心を持ち検索する可能性のある人々の総数を表す。この SIR モデルを利用して、実際の観測データから最小二乗法を利用してパラメータの推定を行う。

本研究では、「Google トレンド」上でのそれぞれの商品名をクエリとして、ある期間の検索量の推移を取得し、その値を $I(t)$ の観測データとする。このデータを元に、パラメータとして、初期感受性保持者数、感染率 β 、回復率 γ の推定を行う。

3.4. マスメディア上の露出による情報拡散と生活者の口コミによる情報拡散

ここでは、情報拡散の過程について、マスメディア上の露出による情報拡散と生活者の口コミによる情報拡散の2種類のタイプに分けられることを、具体的な事例分析をまじえて説明する。

例として、図 3.4 に「アナと雪の女王」、「クロワッサンドーナツ」の『Google トレンド』上の人気度の変化と SIR モデルのパラメータ最適化を行った結果を示す。「アナと雪の女王」は、2014 年 4 月 6 日～2014 年 7 月 19 日の期間で、「クロワッサンドーナツ」は、2014 年 10 月 5 日～2015 年 1 月 17 日の期間の検索量の推移を抽出したものである。「アナと雪の女王」に関しては、この期間が開始したのは映画の全国公開から 4 週目にあたり、4 月 26 日からは 3D 日本語吹替版が約 138 劇場で、映画館の観客と一緒に挿入歌を歌える限定公開イベント「みんなで歌おう♪」歌詞付き」版が約 85 劇場で上映開始されたことが報道されている。一方、「クロワッサンドーナツ」に関しては、この期間に 2014 年度のヒット商品ランキングが発表されている。SIR

図 3 “アナと雪の女王” の人気度と SIR モデルのパラメータを最適化したグラフ

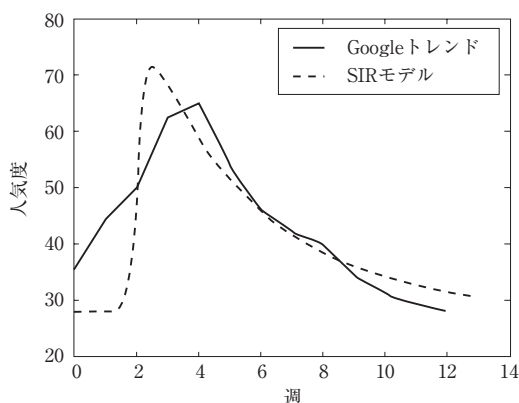
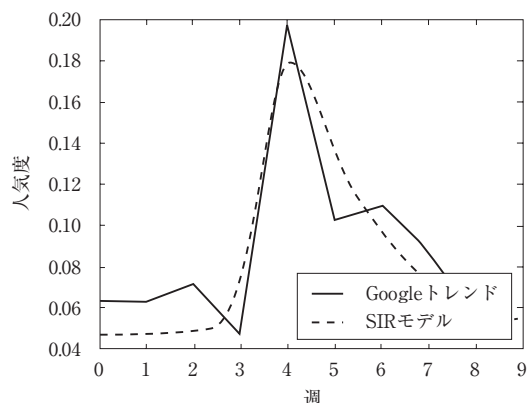


図 4 “クロワッサンドーナツ” の人気度と SIR モデルのパラメータを最適化したグラフ



モデルによって推定されたパラメータは、「アナと雪の女王」が、 $S_0=52.1$, $\beta=7.05$, $\gamma=0.270$ で、「クロワッサンドーナツ」は、 $S_0=0.231$, $\beta=4.56$, $\gamma=0.670$ であった。

観測された人気度の変遷をグラフから読み取ると、「アナと雪の女王」は「クロワッサンドーナツ」と比較すると、人気度の総和が高く、感染のカーブが急で減衰のカーブが緩やかであることが読み取れる。この違いは推定したパラメータに表れており、「アナと雪の女王」は「クロワッサンドーナツ」と比較して、人気度の総数を表す S_0 、拡散の速度を表す β が高く、人気の減衰速度を表す γ が低くなっている。

実際、「アナと雪の女王」のヒットのピーク前のブログの平均露出数は、18,407.0であったのに対し、「クロワッサンドーナツ」のヒットのピーク前のブログの平均露出数は、268.8であった。このことから、ブログへのたくさんの露出は、 S_0 に正の影響を与えることが考えられるだろう。また、それぞれのブログでその期間に見られるようになった語を調査してみると表2, 3のようになった。これによると、「アナと雪の女王」では主題歌に好感を示す単語が見られ、イベントが成功していることが観察されるのに対し、「クロワッサンドーナツ」ではミスタードーナツでクロワッサン生地を使用している他の商品が多く観察される。この違いから「クロワッサンドーナツ」が、マスメディアや広告の露出が元で話題になっている可能性が考えられるだろう。ここで、マスメディア上の露出数とソーシャルメディア上での露出数を元に「口コミ指数 (I_{WoM})」を以下の式のように提案する。

$$I_{WoM} = \frac{N_S \times N_M}{\sum_i^{N_S} \sum_j^{N_M} \frac{1}{I_{S_i, M_j}}}$$

$$I_{S_i, M_j} = \frac{C_{S_i} + 1}{(C_{M_j} + 1) \times R_{S_i, M_j}}$$

表2 blog上で「アナと雪の女王」に関連する上昇ワード

順位	クエリ	上昇率
1	アナ	+1.3%
2	雪の女王	+1.3%
3	主題歌	+1.1%
4	海外	+1.0%
5	アナ雪	+0.8%
6	好き	+0.8%
7	話題	+0.7%
8	子ども	+0.5%
9	大ヒット	+0.5%
10	ネット	+0.5%

表3 blog上で「クロワッサンドーナツ」に関連する上昇ワード

順位	クエリ	上昇率
1	ミスド	+2.4%
2	カップケーキ	+1.9%
3	ドーナツ	+0.7%
4	マロン	+0.6%
5	アップル	+0.5%
6	楽しい	+0.5%
7	ミルク	+0.5%
8	うまい	+0.5%
9	美味しい	+0.4%
10	おやつ	+0.4%

この式において、 i は N_S 種類あるソーシャルメディアのうち、 i 番目のソーシャルメディア S_i を表し、 j は N_M 種類あるマスメディアのうち、 j 番目のマスメディア M_j を表している。そして、 I_{S_i, M_j} はソーシャルメディア S_i とマスメディア M_j 間の露出量 C_{S_i} , C_{M_j} の比率を表している。ただし、式中の R_{S_i, M_j} はソーシャルメディア S_i とマスメディア M_j 間の差を調整する係数で、手動で設定を行う。そして、最終的な口コミ指数 I_{WoM} は、複数のソーシャルメディア、マスメディア間の露出量の比率の調和平均を取る。

$$I_{blog, MetaTV} = \frac{C_{blog} + 1}{(C_{MetaTV} + 1) \times R_{blog, MetaTV}} \tag{4}$$

$$I_{Twitter, MetaTV} = \frac{C_{Twitter} + 1}{(C_{MetaTV} + 1) \times R_{Twitter, MetaTV}} \tag{5}$$

$$I_{WoM} = \frac{2}{\frac{1}{I_{blog, MetaTV}} + \frac{1}{I_{Twitter, MetaTV}}} \quad (6)$$

今回、ソーシャルメディアとして Blog、Twitter、マスメディアとして TV を採用した。よって、 $I_{Si, Mj}$ 、 I_{WoM} は式 4～6 のようになる。式 4, 5 はそれぞれ Blog と Twitter の露出量と TV での露出量との比である。露出量の尺度は様々なものが考えられ、適切な尺度については議論の余地があるが、今回は単純に露出数とする。つまり、Blog の場合は該当の記事数、Twitter の場合は該当のツイート数、TV の場合は該当の番組数とした。 $R_{blog, MetaTV}$ 、 $R_{Twitter, MetaTV}$ は、TV での露出数がソーシャルメディア上での露出数と比較すると極端に少ないため、調整するために $R_{blog, MetaTV} = 750$ 、 $R_{Twitter, MetaTV} = 950$ を代入している。また最終的に算出する I_{WoM} は、 I_{blog} と $I_{Twitter}$ の調和平均を取った値となっている。これによって極端に Twitter と Blog で異なった値となっても小さい値の方に近い値となる。この口コミ指数が小さい値の場合、その流行に関して、Blog でも Twitter でも TV での露出量と比較して、あまり話題になっていないことを示す。逆に大きい値の場合、その流行は、Blog でも Twitter でも TV での露出量と比較して、大きな話題になっていることを示す。

3.5. 商品情報の拡散過程分析

前述のように、ヒット商品・サービスの中で、同一の情報発信内容についてとりあげられた期間のデータのみを抽出し、『Google トレンド』のデータを人々の関心の広がりを示すデータとして用いて、SIR モデルのパラメータの推定を行う。次に、検索数が最も多くなったピークの時点（これを流行が頂点に達した状態とする）より以前の一定期間について、各メディアでのその商品・サービスの情報の露出数のデータを取り出し、指標化する。ここでは、以下の指標を定義する。

$$\bar{C} = \sum_{t=1}^N \frac{C(t)}{N} \quad (7)$$

式 7 は人々の関心が最も高くなった時点までの前 N 時点でのあるメディアでの露出量 $C(t)$ の平均を示したものである。

最後に提案した指標と、商品情報の拡散過程を表すパラメータとの相関を示すことで、類型化した商品・サービスに関する情報の拡散過程を説明する指標を明らかにする。相関係数 r にはピアソンの積率相関係数を採用し、以下のように算出する。

$$r(x, y) = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2} \sqrt{\sum (y_i - \bar{y})^2}} \quad (8)$$

この式内で、 \mathbf{x}, \mathbf{y} はそれぞれ指標や推定したパラメータを表し、 x_i, y_i は抽出した期間の流行現象での指標、推定したパラメータの値を示すものとする。 \bar{x}, \bar{y} はそれぞれの指標、パラメータの期間全体での平均を表すものとする。

具体的には、5.1 節に示す手法で同一の話題となっている期間を同定し、その期間内でパラメータ (S_0, β, γ) の推定を行い、それぞれ \mathbf{x} の各要素 x_i とする。次に (7) 式に示したあるメディア (ブログ、ツイッター、新聞、TV) でのヒット前露出量に関する、期間 i における指標の値を y_i とする。そしてパラメータ、メディアごとの組で、全体の相関係数 r を (8) 式によって算出する。

4. 使用するデータ、実験設定

本研究の分析で使用するデータ、実験設定について説明する。

ヒット商品・サービスについては、『日経トレンド』2014 年ヒット商品ベスト 30 の上位 10 商品を選定した。日経トレンドは日経 BP 社が発刊する月刊誌で、毎年 12 月号にその年にトレンドになったものを選

出する“ベスト30”を掲載している。2014年度のトップ10の商品・サービスとその説明を表4に示す。

本研究では、さまざまなサービスからデータを取得しているが、すべてのデータを、2010年07月04日から2015年08月01日の期間で取得し、1週間ごとに集計したものを使用する。そして、集計したデータを時系列データとして取り扱っている。また、データ取得には各サービスにおいて商品名を直接のクエリとして入力し、出力されたデータを取得している。

本研究では、商品・サービスへの人々の関心をあらわす指標として、『Googleトレンド』のデータを使用した。ただし、05位の“Ban汗ブロックロールオン”は十分な検索量がなかったため『Googleトレンド』でデータを取得できなかった。このため、分析対象から外すこととした。また、個別の商品についての検索量は、『Googleトレンド』を用いた単独商品・サービスの分析結果では0～100の範囲の人気度として示されてしまうため、2つの商品の結果を同時に取得することで、商品間での相対的な検索量を比較することを可能とした。こうした比較を行い導いた、それぞれの商品・サービスの検索量の相対的な比較値を表4に示した。本研究で使用する検索量データは、それぞれの商品・サービスの単独での『Googleトレンド』上での人気度の結果に対し、この表に示した比較値を乗じた値を使用する。

表4 日経トレンドィ2014年ヒット商品ベスト30の上位10商品。
(ただし、表内のWwoHPは、ウィザーディング・ワールド・オブ・ハリー・ポッターを表す。)

順位	比較値	商品	説明
01位	6.50×10^{-1}	アナと雪の女王	ウォルト・ディズニー・アニメーション・スタジオ制作の映画。
02位	1.000	妖怪ウォッチ	レベルファイブが開発したゲームおよび様々なメディアで展開される一連の作品。
03位	7.80×10^{-2}	WwoHP	テーマパーク施設のユニバーサルスタジオジャパンにあるアトラクション。
04位	6.24×10^{-3}	ジェルボール線材	ゼリー状の球形の新型洗剤。
05位	—	Ban汗ブロックロールオン	直塗りタイプの制汗剤。ボール付き容器が特徴。
06位	2.43×10^{-3}	伊右衛門 特茶	“特定保健用食品”の許可証票の表示許可を受けた飲用茶。
07位	4.49×10^{-1}	TSUM TSUM	LINE株式会社が公開しているスマートフォン向けアプリケーション。
08位	7.90×10^{-3}	クロワッサンドーナツ	ニューヨーク発祥のクロワッサン生地のだーナツ。
09位	1.88×10^{-2}	格安スマホ	MVNO(仮想移動体通信事業者)が提供する通信サービスを利用するスマートフォン。
10位	5.53×10^{-2}	あべのハルカス	大阪市阿倍野区に立地する超高層ビル。

さらに、人々の検索量(関心)の変化を説明できる可能性のある変数として、マスメディア、ソーシャルメディアにおける露出数のデータを利用する。ソーシャルメディアとして、ブログ、ツイッターを選定し、これらの過去のデータを利用する⁷⁾。ブログは、個人または団体が比較的長い文章を書くもので、本研究ではAmebaブログ⁸⁾やlivedoorブログ⁹⁾のデータを利用している。ツイッターは、個人または団体が140字以内の短い文章を書くもので、本研究では10%サンプリングしたデータを利用している。ソーシャルメディアは、他にもさまざまなものが考えられる。例えば、Facebook, YouTube, LINEがあげられる。今回これらのデータは用いなかった。その理由は、これらのソーシャルメディアにおける投稿情報は取得や解析が困難であること、生活者の関心を表す先行指標としてソーシャルメディアが有用であるかを明らかにするには今回取得したデータでも代用できること、の2点があげられる。もちろん、多くのソーシャルメディアのデータを利用することによってより精度の高い分析、議論ができるため、この点については今後の課題としておきたい。ブログ、ツイッターのデータは株式会社ホットリンクが運営している“口コミ@係長”というサービスから提供されたデータを用いている。マスメディアとしては、新聞、テレビを選定し、これらの過去のデータを利用した。

例えば、日経新聞の露出データを日経テレコン 21¹⁰⁾ から取得した。なお、日経新聞を選択したのは、経済メディアが商品・サービスのヒット現象に及ぼす影響を分析するためである。テレビの露出データは株式会社エム・データが提供しているテレビ情報検索システム MetaTV¹¹⁾ から取得したデータを使用した。対象は在京のNHK 及び民放のすべての番組である。ただし、MetaTV のデータには番組での露出情報は含まれるがCMの放映や視聴率は含まれていない。そのため、本稿でのマスメディアの露出に関する分析ではCMや視聴者数に関するものは行っていない。

5. 結果

ここでは、提案手法を取得したデータに対して用いて行った実験結果について議論する。

5.1. 同一の話題となっている期間の抽出

まず、収集したデータを用いて、同一の話題で流行となっている期間の抽出を行った。これは Google トレンドの推移でピークとなっている週を抽出し、その週の4週間前から10週間後までの、合計15週間を1つの流行の期間とした。そして、それぞれの話題について、各メディアの露出状況を調査し、3.4節で提案した「口コミ指数 (I_{WoM})」を算出し、主にマスメディア上の露出によって流行しているものと生活者の口コミによって流行しているものに分類した。ここでは、 $I_{WoM} \geq 1.5$ なるものを生活者の口コミによって流行したもの、 $I_{WoM} \leq 0.5$ なるものをマスメディア上の露出によって流行したものと分類した。

表5, 6に結果の一部を示す。まず、すべての商品・サービスにおいて、生活者による検索数は時間と共に変化していることが確認された。抽出した期間におけるメディアそれぞれに対して、SIR モデルの最適化によるパラメータ推定を行い、後の分析を行った。次に、生活者の口コミによって流行したと「口コミ指数」によ

表5 生活者の口コミによって流行になったと思われる商品と話題の組 (全14組) 一部

順位	口コミ指数	人気度	期間	商品	主な話題
1	8.32	393.33	2015/03/22~2015/07/04	TSUM TSUM	めざましテレビで紹介される。
2	8.21	99.57	2014/02/09~2014/05/24	TSUM TSUM	攻略サイトでの高得点取得方法の紹介。
3	7.60	14.10	2013/05/12~2013/08/24	あべのハルカス	6/16に近鉄百貨店開業と報道。
4	6.47	365.30	2014/06/15~2014/09/27	アナと雪の女王	DVD 発売開始。
5	4.19	635.05	2014/04/06~2014/07/19	アナと雪の女王	レリゴー現象 (主題歌が話題になる)。
6	3.76	203.17	2014/05/11~2014/08/23	TSUM TSUM	テレビ番組で紹介。
7	3.31	354.32	2014/11/30~2015/03/14	TSUM TSUM	年末年始でアプリダウンロードが伸びたため。
8	3.31	85.15	2014/11/30~2015/03/14	アナと雪の女王	年度末ヒット商品ランキング1位の効果。
9	2.76	2.88	2012/07/29~2012/11/10	あべのハルカス	地上高300m に到達の発表。
10	2.26	553.15	2014/02/16~2014/05/31	アナと雪の女王	3/14公開。
平均	4.13	271.85			

表6 マスメディア上での露出によって流行になったと思われる商品と話題の組 (全18組) の一部

順位	口コミ指数	人気度	期間	商品	主な話題
1	0.035	1.86	2013/10/13~2014/01/25	クロワッサンドーナツ	テレビ番組で流行している商品として紹介される。
2	0.070	1.45	2014/10/05~2014/01/17	クロワッサンドーナツ	ヒット商品ランキング発表。
3	0.080	4.34	2014/04/13~2014/07/26	ジェルボール洗剤	誤飲による報道。
4	0.102	3.52	2014/06/22~2014/10/04	ジェルボール洗剤	ヒット商品ランキング発表。
5	0.117	0.747	2014/08/31~2014/12/13	伊右衛門 特茶	Twitter, FB 1周年キャンペーン。
6	0.138	15.59	2014/09/28~2015/01/10	格安スマホ	楽天が参入。
7	0.154	0.876	2014/05/04~2014/08/16	伊右衛門 特茶	Twitter フォローキャンペーンの発表。
8	0.160	40.99	2014/02/02~2014/07/17	あべのハルカス	3/7に全面開業と発表。
9	0.184	34.63	2014/03/30~2014/07/12	あべのハルカス	GW のレジャースポットとして話題になる。
10	0.234	0.284	2013/05/26~2013/09/07	クロワッサンドーナツ	パン屋で販売されたものを食べた感想。
平均	0.236	11.49			

て分類された商品・サービスと話題の内容を確認する。実際のブログの記事を読むと、「TSUM TSUM」は、ゲームで高得点を取る方法がブログ上で書きこまれており、そうした内容を求めて人々が検索を行なった結果と考えられる。また、表2でも示した通り、「アナと雪の女王」は主題歌が評判になり、様々な著名人、一般の個人が歌う動画が投稿され大きな流行となっていた。このように生活者の口コミによって流行した商品・サービスの話題には、生活者が体験した感想や共感できる内容が含まれていることがわかる。一方で、マスメディア上の露出によって流行した商品は、表3で一部見られたようにマスメディアで接触した話題に対する期待感や、ニュースを拡散しようとするブログ記事が多かった。

5.2. SIR モデルを利用した実験結果

SIR モデルを利用した全体の分析結果を表7に示す。同一の話題の期間について、SIR モデル近似によるパラメータ推定と各指標の相関分析を行った結果、『商品に関心を持ち検索する可能性のある人々の総数 (S_0)』はブログ・テレビの露出量に正に相関していた。一方で、情報の『拡散力 (β)』、『減衰力 (γ)』はブログ、新聞、テレビの露出量に弱く正に相関する結果となった。この結果は、例えばテレビの場合、露出が増えるほど「テレビでの露出に接することで興味を持つ可能性を持つ人の総数」を増やし、実際に興味を持ちやすくなりつつも、飽きやすくなる、と解釈することができる。

生活者の口コミによって拡散した情報に関するデータの分析結果を表8に示す。期間中の人々の『人気度の総量』はブログ、テレビの露出量に正に相関する結果となった。マスメディア上での露出によって拡散した情報に関するデータの分析結果を表9に示す。期間中の人々の『人気度の総量』はブログ、新聞、テレビの露出量に正に相関する結果となった。情報の『拡散力』、『減衰力』はブログ、テレビの露出量に正に相関する結果

表7 推定したパラメータと指標との相関係数（表中の“*”は無相関検定の検定結果で、それぞれ*： $p < 0.05$ ，**： $p < 0.01$ ，***： $p < 0.001$ で有意であることを示す。）

パラメータ	指標			
	C_{blog}	$C_{Twitter}$	C_{nikkei}	C_{MetaTV}
S_0	0.955***	0.390**	0.340*	0.761***
β	0.378**	0.0950	0.401**	0.349*
γ	0.379**	0.0940	0.400**	0.349*

表8 生活者の口コミによって流行したデータでの推定したパラメータと指標との相関係数（表中の“*”は無相関検定の検定結果で、それぞれ*： $p < 0.05$ ，**： $p < 0.01$ ，***： $p < 0.001$ で有意であることを示す。）

パラメータ	指標			
	C_{blog}	$C_{Twitter}$	C_{nikkei}	C_{MetaTV}
S_0	0.919***	0.109	0.386	0.809***
β	0.522	0.018	0.351	0.303
γ	0.522	0.019	0.351	0.305

表9 マスメディア上での露出によって流行したデータでの推定したパラメータと指標との相関係数（表中の“*”は無相関検定の検定結果で、それぞれ*： $p < 0.05$ ，**： $p < 0.01$ ，***： $p < 0.001$ で有意であることを示す。）

パラメータ	指標			
	C_{blog}	$C_{Twitter}$	C_{nikkei}	C_{MetaTV}
S_0	0.635**	0.424	0.664**	0.697**
β	0.502*	0.231	0.455	0.479*
γ	0.502*	0.230	0.453	0.479*

となった。

6. 考察

本章では考察を述べる。

6.1. 商品・サービスの類型化について

ヒット現象は口コミ指数を用いることで、生活者の口コミによって起きるものとマスメディア上の露出によって起きるものの2つに類型化することができる。そして、この2つの類型について、情報拡散の過程に差異を見出すことができる。例えば、「伊右衛門 特茶」に関するヒット現象は、ソーシャルメディアを利用したキャンペーンによって起きているが、メディア上の露出に対し、実際のソーシャルメディア上の露出はあまり多くないことが口コミ指数によって観察される。このタイプのヒット現象は、表4に示した通り人気度の平均が小さい。マスコミ上への露出のみで大きなヒット現象を作り出すのは難しいようだ。なお、4章で言及したとおり、今回のマスメディアの露出数はTV番組への露出数で、CMの放映数、視聴者数を考慮していない。また、一般的にメディア露出やソーシャルメディア上の露出の後押しをするために、費用をかけて広告出稿を行う方法も考えられる。本研究ではこうした手法について分析対象とはしていないが、この点についても踏み込んだ分析を行うことで、また異なる知見を得られる可能性もある。

一方で、「アナと雪の女王」のように生活者の口コミによって大ヒットする商品が認められた。このタイプのヒット現象は、表5に示した通り人気度の平均が大きく大ヒット現象となっているものが多い。しかし、「アナと雪の女王」もメディア上の露出数は非常に大きく、生活者の口コミのみで流行したわけではない。口コミによる評判の良さとメディアの露出の相乗効果によって大きなヒット現象になっていることについては注意しておきたい。また、メディア間での密接な相互作用については、本論文では取り扱っておらず、今後の課題としたい。

また、「TSUM TSUM」に関しては、どの流行現象についても口コミ指数が高いものとなっている。これは、スマートフォン向けのゲームアプリケーションという性質上、スマートフォン上のアプリケーションプラットフォームで、人気アプリとして多くのユーザーに提示されているという原因も考えられる。ユーザーが接触するメディアの情報としてスマートフォンの画面の重要性は非常に高まっており、今後の分析が必要と考えられる。

そして、ソーシャルメディアによる「口コミ効果」の重要性は、すでに多くの関係者に認識されている。本稿でも、ソーシャルメディアを利用したキャンペーンなどの事例が見られたが、世間では関係者がいかにも生活者としてブログ投稿を行う例も話題になっている。本研究では、ブログの個々の記事について詳細な分類は行っておらず、こうした作為的な口コミ効果を含んでいる可能性があることにも言及しておく。

6.2. 情報発信手法に関する考察

表7に示した結果から、ヒット現象の生活者による『人気度の総量』には、ブログやテレビでの露出量が関係していることがわかった。また、ブログ、新聞、テレビでの露出は、多くの人々に対してより急速に情報を拡散させる上で有効である可能性が示された。

流行した話題の内容ごとのタイプ分けの結果を比較すると、生活者の口コミによって流行したタイプについては、ブログ、テレビでの露出数が、『人気度の総量』と密接に相関していることが明らかになった。これは、ソーシャルメディア上で話題となる商品は、魅力的であるので、他の生活者の関心をひきやすく、マスメディアで露出させることによって、より多くの人々へ情報拡散を促進すると考えることができる。つまり、もともと流行するだけのポテンシャルがあり、すでに特定のコミュニティ内に口コミで流行しているものをマスメディアで多くの人々に認知させることで、さらに情報拡散を促すことができることを示している。また、「アナと雪の女王」の事例で見られるように、ソーシャルメディア上で話題の動画をマスメディアに取り上げても

らうことでも、同様の現象が発生する可能性があると考えられる。

一方で、マスメディア上での露出によって流行したタイプではブログとTVの露出量が『拡散力』と一定の相関があることが認められた。ブログやTVでの露出は、生活者に一定の関心を持たせるような影響を与える。

ここまでの考察をまとめると、メディアがヒットに与える影響としては、どのメディアでも露出量が多ければ多いほど、人気の合計（検索数の総量）は上昇する。また、ツイッター以外のメディアでの露出は、『拡散力』、『減衰力』に正の相関がある。つまり、これらのメディアでの露出は急速な拡散が見込めると同時に早期の減衰も引き起こす可能性が高まることを示唆している。また、生活者の口コミによって情報が拡散した商品・サービスは、テレビや新聞で露出させることで、大きな『人気の総量』を得ている。一方、マスメディアの露出によって情報が拡散した商品・サービスは、ツイッター以外のメディアによって、大きな『人気度の総量』を得ることができる。

また、本研究の成果の制限・限界について言及する。今回は、ある商品・サービスについて、同一内容の情報が話題となった期間のみを抽出して分析を行った。しかし、実際のヒット商品・サービスは長期間にわたってヒットが継続しているが、その期間の中で発信する内容を変化させることで、継続して生活者の関心が増大させている。本稿での分析では、次々と発信内容を変えることが人気度の変化にどのような影響を与えるかは分析していない。また、同一の話題で情報発信が行われている期間の同定については、各メディアへの情報の露出状況を確認することでやっている。同一の話題の時系列の分布を取得する手法として Dynamic Topic Models (Blei and Lafferty 2006) といった手法が提案されている。こうした手法を利用し正確に商品の情報拡散の話題を捉えることで、より精緻な分析やモデル構築を行うことができるだろう。

6.3. 本論文の貢献

ここまでの議論から、情報拡散（ヒット現象）の発生にはブログによる情報発信と、テレビを通じた情報発信が比較的大きな影響を及ぼしている可能性が示された。また、ヒット商品・サービスに関する事例を「口コミ指数 (I_{WOM})」を用いて類型化したところ、マスメディアによって流行した商品は全体的に人気度の総量が小さく、キャンペーンや単発のイベント、ニュースに関連するものが多くみられた。一方で、生活者の口コミによってヒット現象となった商品は、人気度の総量が大きく、大きなニュースや生活者が参加しやすい内容のものが多くみられた。また、それぞれのタイプにおいてメディアが及ぼす影響力も異なることから、目指すべきヒットのタイプを明確にし、情報発信手法を設計することが必要なことが明らかになった。

本研究の成果を要約すると、以下の貢献が挙げられる。

- 人々による商品・サービスへの検索数（関心度）の推移が、ソーシャルメディアやマスメディアの露出量と相関していることを示した。
- 商品・サービスに関する情報が拡散し生活者の関心が増減する過程には、「口コミ指数 (I_{WOM})」を用いることで、マスメディア上の露出によるものと生活者の口コミによるものに分類でき、それぞれ異なる情報拡散過程を示すことがわかった。
- マスメディア上の露出による情報拡散は人気度が小さく、大量の人気度の獲得には生活者の口コミが重要であることがわかった。

今後、本研究の成果を利用することで、ある商品・サービスの広報活動の成果を最大化するために、どのようなメディアを活用すべきか指針を得ることが可能になる。一方、今回はある商品・サービスの中でも同一の内容の情報が拡散している期間を同定してデータを抽出して分析を行った。今回の分析をもとに今後は、長期間にわたってヒット現象が続く商品・サービスについても説明可能なモデルを構築し、生活者への影響を拡大する情報発信手法の実証を目指していきたい。

7. おわりに

本研究ではヒット商品・サービスに関する情報の拡散過程を分析し、マスメディアやソーシャルメディアでの情報の露出によって、人々による該当商品・サービスへの検索数が変化する過程を明らかにした。また、商品・サービスに関する情報拡散（ヒット現象）に、本論文で提案した「口コミ指数」を用いることで、マスメディア上の露出による流行を示しているか、生活者による口コミによる流行を示しているか判断することができる。そして、それぞれのタイプにメディアが与える影響が異なることがわかった。本研究の成果は、インターネット上の情報拡散の重要性が増している現在、商品・サービスの情報を伝えるために広報実務家がどのような情報発信を行うべきか、検討するための一つの指標を提示している。

なお、一つの商品・サービスに関する長期間のヒット現象については、実際には複数の話題による情報発信の結果が組み合わさり、生じている。今回は、複数の話題の相乗作用によるヒット現象の分析は行っていない。また、単一のメディアの影響のみならず、複数のメディアの影響を同時に折込んだ動的な情報拡散モデルの構築も必要となるだろう。今回の研究を嚆矢として、引き続き商品・サービスの情報拡散過程に関する研究に取り組み、詳細な情報拡散過程を明らかにすることで、広報の実務にも役立つことが出来れば喜びである。また、本研究への支援を頂いた株式会社ホットリンク、株式会社電通パブリックリレーションズの関係者にこの場を借りて御礼申し上げたい。

脚注

- 1) 2015年1月9日ヤフー株式会社「第47回衆院選の議席数予測を振り返る」
本レポートでは、選挙前の検索履歴のデータを用い、過去の選挙結果における検索履歴との獲得議席数の相関を利用する手法で92%の精度で政党ごとの議席数を予測できることを示している。また、これはアンケート結果をもとに行う議席数の予測と同程度の的中率であると述べられている。
<http://docs.yahoo.co.jp/info/bigdata/election/2014/03/>
- 2) Klout
ソーシャルメディア上での影響力を測定するウェブサービス。ユーザの情報発信がネットワーク上でどれくらいの影響力を持つかを0～100の値で指標化する。
<https://klout.com/>
- 3) 口コミ@係長
ソーシャル・ビッグデータ分析ツール。Twitter, Facebookを始め、ウェブ上の掲示板、ブログサイトから収集したデータをユーザが自由に把握・分析するためのマーケティングツール。
<https://www.hottolink.co.jp/service/kakaricho>
- 4) Googleトレンド
ある単語がGoogleでどれだけ検索されているかというトレンドをグラフで見ることができるツール。検索数の絶対量は提示されず、複数の単語間の検索量の相対値を0～100の値で指標化して表示する。
<https://www.google.co.jp/trends/>
- 5) Google Correlate
ユーザが時系列のデータをアップロードすると検索クエリ群の時系列データとの多変量解析を行った結果を提示するウェブサービス。それぞれの検索語と入力データとの相関係数を確認できる。
<https://www.google.com/trends/correlate>
- 6) 2014年ヒット商品ランキング
2013年10月から2014年9月までに発表された商品やサービスを日経トレンド誌編集部が「売れ行き」「新規性」「影響力」の3項目で総合的に判定。それぞれのヒットの度合いを評価し、1位から30位までのランキングに集計したもの
<http://trendy.nikkeibp.co.jp/article/pickup/20141030/1061085/?rt=nocnt>
- 7) 口コミ情報の一つとしてブログを用いているが、一般にウェブサイトの運営者は、検索誘導でのトラフィック獲得のため、検索数の大きいワードの記事を書くことが多いことにも注意を払う必要がある。
- 8) Amebaブログ
サイバーエージェントが提供するレンタルブログサービス。会員数は2012年時点で約2000万人。ブログのアクティブユーザーは2009年11月時点で約115万人/月と報告されている。
<http://ameblo.jp/>
- 9) livedoorブログ
韓国のIT企業ネイバーの子会社、LINEが運営するブログサービス。国内有数のブログサービスで、アクティブユーザーは2011年3月時点で約16万人/月と報告されている。

<http://blog.livedoor.com/>

10) 日経テレコン 21

日本経済新聞社が提供する日経新聞データベースサービス。過去の新聞記事を検索可能なデータベース機能を提供する。

<https://t21.nikkei.co.jp/g3/CMN0F11.do>

11) MetaTV

エム・データが提供する TV 番組や CM の放送内容をテキスト化、データベース化を行い、検索可能にするサービス。

<http://mdata.tv/metatv/>

参考文献

- Alan, M., L. Sune, A. Yong-Yeol, O. Jukka-Pekka, J. N. Rosenquist *Understanding the Demographics of Twitter Users*, pp. 554-557. AAAI Press 2011.
- Albert, R., A. Barabasi, Statistical mechanics of complex networks, *Reviews of Modern Physics*, 74, pp. 47-97 2002.
- Bakshy, E., J. M. Hofman, W. A. Mason, D. J. Watts, Everyone's an Influencer: Quantifying Influence on Twitter, *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*, WSDM '11, pp. 65-74 2011.
- Blei, D. M., J. D. Lafferty, Dynamic Topic Models, *Proceedings of the 23rd International Conference on Machine Learning*, ICML '06, pp. 113-120, New York, NY, USA: ACM 2006.
- Britton, T. Stochastic epidemic models: a survey, *Mathematical biosciences*, Vol. 225, 1, pp. 24-35 2010.
- Frank, M. B. A New Product Growth for Model Consumer Durables, *Management Science*, Vol. 15, 5, pp. 215-227 1969.
- Ghose, A., P. G. Ipeirotis, Designing Novel Review Ranking Systems: Predicting the Usefulness and Impact of Reviews, *Proceedings of the Ninth International Conference on Electronic Commerce*, ICEC '07, pp. 303-310 2007.
- Ginsberg, J., M. H. Mohebbi, R. S. Patel, L. Brammer, M. S. Smolinski, L. Brilliant, Detecting influenza epidemics using search engine query data, *Nature*, Vol. 457, 7232, pp. 1012-1014 2009.
- Hatzivassiloglou, V., K. R. McKeown, Predicting the Semantic Orientation of Adjectives, *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics*, ACL '98, pp. 174-181 1997.
- Higie, R. A., L. F. Feick, L. L. Price, Types and amount of word-of-mouth communications about retailers., *Journal of Retailing*, 63(3), pp. 260-278 1987.
- Java, A., X. Song, T. Finin, B. Tseng, Why we twitter: understanding microblogging usage and communities, *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis*, pp. 56-65, ACM 2007.
- John, C., A. S. Joshua, Epidemiological modeling of online social network dynamics, *CoRR*, Vol. abs/1401.4208 2014.
- Kermack, W. O., A. G. McKendrick, A contribution to the mathematical theory of epidemics, *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, Vol. 115, pp. 700-721 1927.
- Kietzmann, J., A. Canhoto, Bittersweet! Understanding and Managing Electronic Word of Mouth, *Journal of Public Affairs*, Vol. 13, 2, pp. 146-159 2013.
- Kotler, P., R. E. Turner *Marketing management: analysis, planning, and control*, Upper Saddle River, NJ, USA: Prentice-Hall 1979.
- Kwak, H., C. Lee, H. Park, S. Moon, What is Twitter, a Social Network or a News Media? *Proceedings of the 19th International Conference on World Wide Web*, WWW '10, pp. 591-600, New York, NY, USA 2010.
- Lazer, D., R. Kennedy, G. King, A. Vespignani, Google Flu Trends Still Appears Sick: An Evaluation of the 2013-2014 Flu Season, Available at SSRN 2408560 2014.
- Liu, B. Sentiment analysis and opinion mining, *Synthesis Lectures on Human Language Technologies*, Vol. 5, 1, pp. 1-167 2012.
- Robertson, T. S. Consumer Innovators: The Key to New Product Success, *California Management Review*, Vol. 10, 2, pp. 23-30 1967.
- Rogers, E. M. *Diffusion of Innovations*, New York, NY, USA: The Free Press 1962.
- Takayasu, M., K. Sato, Y. Sano, K. Yamada, W. Miura, H. Takayasu, Rumor Diffusion and Convergence during the 3.11 Earthquake: A Twitter Case Study, *PLoS ONE*, Vol. 10, 4, p. e0121443, 04 2015.
- Turney, P. D. Thumbs Up or Thumbs Down?: Semantic Orientation Applied to Unsupervised Classification of Reviews, *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, ACL '02, pp. 417-424 2002.
- Watts, D. J. *Small worlds: the dynamics of networks between order and randomness*, Princeton, NJ, USA: Princeton University Press 1999.
- 遠藤薫 「「ネット世論」という曖昧: 〈世論〉, 〈小公共圏〉, 〈間メディア性〉 (〈特集〉世論と世論調査)」, 『マス・コミュニケーション研究』 77 号, 105-126 頁 2010.
- 遠藤薫 「「序章 なぜいまジャーナリズムを考えるか」, 『間メディア社会の〈ジャーナリズム〉 ソーシャルメディアは公共性を変えるか』, 1-17 頁 2014.
- 石井晃・吉田就彦 「ヒット現象の数理モデル」, 『鳥取大学工学部研究報告』, 36 巻 39 号, 71-80 頁 2005.
- 濱岡豊 「口コミの発生と影響のメカニズム」, 『消費者行動研究』 2 巻 1 号, 29-73 頁 1994.

口コミ指数による事例類型化に基づく複数メディアの ヒット前の露出を先行指標とした情報拡散過程の分析

大知 正直
(東京大学大学院
工学系研究科)

長濱 憲
(東京大学大学院学際情報学府,
株式会社 電通パブリック
リレーションズ)

榊 剛史
(東京大学大学院
工学系研究科,
株式会社ホット
リンク)

森 純一郎
(東京大学大学院
工学系研究科)

坂田 一郎
(東京大学大学院
工学系研究科)

インターネットを利用したマーケティングの重要性が提起されて以来すでに久しい。しかし、実際に商品・サービスの情報が拡散していく過程と影響について、十分な解明がなされてきたとは言いがたい。特に、マスメディアとソーシャルメディア間での情報拡散の分析の重要性については、遠藤らが主張している「間メディア性」という概念として知られているものの、定量的なデータに基づく実証研究はまだ十分に行われていない。本論文では、マスメディアとソーシャルメディアでのヒット商品情報から算出した指標が、実際の情報の拡散過程を説明することに有効であることを示す。また、ヒット現象を提案する「口コミ指数」を用いて類型化を行い、情報拡散過程に及ぼすメディアによる影響の大きさ、性質を明らかにした。本研究は学術的意義のみならず、商品・サービスの広報戦略の立案に資する実務的な意義も有している。

キーワード：間メディア性、情報流通構造、情報拡散、SIR モデル、商品パブリシティ

The Dynamic Analysis of the Process of Information Diffusion Using Media Exposure Indication

Masanao OCHI
(The University
of Tokyo,
Graduate School of
Engineering)

Ken NAGAHAMA
(The University of Tokyo,
Graduate School of Interdisciplinary
Information Studies, and Dentsu
Public Relations Inc.)

Takeshi SAKAKI
(The University of
Tokyo, Graduate
School of Engineering,
and Hottolink Corp.)

Junichiro MORI
(The University
of Tokyo,
Graduate School of
Engineering)

Ichiro SAKATA
(The University
of Tokyo,
Graduate School of
Engineering)

We already know the importance of internet marketing. However, we do not know how much people are influenced by the Internet and other media. In particular, regarding the influence between mass media and social media, only few papers used quantitative data in their research. In this paper, we analyze the process of information diffusion using data of coverage on popular products from mass media or social media. We show how the presentation from media is important to people. We also believe that the findings are useful for the person who wants to make a publicity plan.

Keywords: Intermediality, Information Interinfluence, Information Diffusion, SIR model, Product publicity