SNS のエンゲージメントに受信者のトライブが与える影響

早稲田大学　杉浦裕文，東京大学　川中孝章，早稲田大学　大野高裕

An Impact of Recipient's Tribe for Engagement of the SNS

Waseda Univ. Hirofumi SUGIURA, The University of Tokyo Takaaki KAWANAKA, Waseda Univ. Takahiro OHNO

1 はじめに

SNS マーケティングにおいて，アカウントの利用目的の多くは，商品や企業の認知向上や宣伝である[3]．そのため，「いいね」やリツイート(以下 RT)といったエンゲージメントをいかに獲得するかが重要である．その KPI としては，投稿自身が生み出す宣伝効果の測定が難しいため，エンゲージメントの獲得数そのものとする場合が多い．

また SNS からは顧客の投稿文自体も収集できるので，それを用いた顧客ターゲティングが行われている．そのためのセグメント作成のアプローチは大きく 2 つある．1 つ目はフォロー・フォロワーの関係から作成するもので，「コミュニティ」と呼ばれ，学術的にはこれを対象としたものが多い．2 つ目は，投稿のコンテンツから趣味嗜好の近い人を抽出する方法であり，これは「トライブ」と呼ばれ，主に実務的に用いられることが多い．近年，多様化した顧客を捉えるために，趣味嗜好のまとまりであるトライブに焦点を当てたマーケティングが行われつつあるが，トライブに関する学術的な検討を行った論文はほとんど見られず，モデリングや分析も経験則的に進められている．

そこで本研究では，日本で利用者数の最も多い Twitter を対象とし，受信者のトライブの違いによる，エンゲージメントへの影響をモデル化して学術的に明らかにする．この成果は，ターゲット顧客に最も効果的な投稿方法を推定することにつながり，今後のトライブマーケティングに活かすことができると考えられる．

2 従来の研究と本研究の特徴

トライブに関する先駆的な研究として，Patrick et al.[1]がある．この研究では，世界中の「痛み」「Pain」という単語が入っている投稿を取得し，それと共起する単語の繋がりからネットワークグラフを作成し，それを近い意味からなる単語のまとまりで分けて，トライブを作成している．結果として，トライブごとの特徴となるコンテンツが大きく異なることを明らかにした．

一方，エンゲージメント獲得に関する影響因子の研究では，津川[2]は先行研究で用いられている URL の有無や文字数などの投稿内容変数に加え，新しく「コミュニティの横断可否」という変数を用い，その結果，分析対象の変数全てが RT に正の影響を与えることを明らかにしている．

しかし所属するトライブの違いによって，エンゲージメントに対する説明変数である URL の有無やフォロワー数等の影響度合いが異なるという考えに至った論文はない．

3 本研究の方針

3．1 研究の流れ

本研究は大きく 3 つのステップから構成される．

まず分析対象である商材の単語を含む「投稿データの取得」である．各変数となるデータやトライブ作成のための投稿文を含むデータセットを一定期間分取得する．

次に，取得した投稿データから「トライブの作成」を行う．Patrick et al.[1]を基に，投稿データの頻出語が利用目的や特徴ごとにまとまりが出来るように作成する．

最後に，エンゲージメントの合算値を目的変数とし，フォロワー数等の投稿内容変数を説明変数とした「モデル式の構築」を行う．トライブを説明変数内の階層構造として組み入れることで，トライブごとの影響の差を推定する．

3．2 各ステップの詳細

3．2．1 投稿データの取得

2019 年 8 月 1 日～31 日の期間に Twitter API を用いて「バイク」「ハーレーダビッドソン」「カメラ」「一眼レフ」のキーワードを含む投稿を取得した．これらの商材を選択した理由は，高関与商材ほど，各トライブの属性が利用目的や特徴の違いによって明確になると考えたためである．取得したデータセットは，表 1 の通りである．

表 1．データ内容

ID	Contents	length	Favorite	Retweet	Time	media	Hashtag	URL
1.17E+18	洗車完了^	55	80	2	2019-08	1	0	0
1.16E+18	【#Rei】	128	19	5	2019-08	0	1	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

UserID	Contribu	Follow	Followers	Authorized
1.05E+18	masataka	553	639	FALSE
2.4E+09	umusicja	543	49087	TRUE
⋮	⋮	⋮	⋮	⋮

項目	0	1
media	画像なし	画像あり
Hashtag	なし	あり
URL	なし	あり

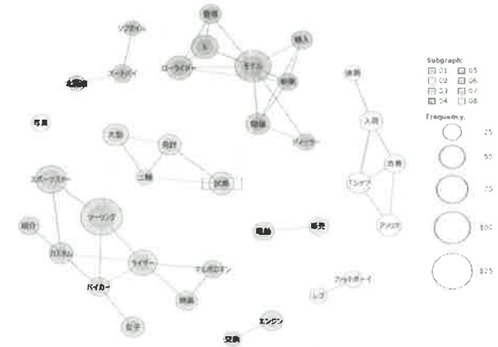

図 1．ハーレーダビッドソンの共起ネットワーク

3．2．2 トライブの作成

まず分析の際に単語の頻出度が過大評価されないためにデータ加工として投稿の除去を行った．方法としては単純な重複の除去，微妙な改変への対応として指定文字数抽出後の重複の除去，最後に目視による除去を行った．

トライブの作成には KH Coder を使用した．KH Coder は文章データを読み込み，データから語を抽出し，頻出語の確認や語の特徴の図示を行うことができるソフトである．本研究では利用対象となる名詞のみを分析の対象とし，キーワード自身や「笑」など意味を持たないが出現頻度が多いため効果のあるノイズは抽出されないように設定した．キーワード「ハーレーダビッドソン」のネットワークグラフは図 1 の通りであり，トライブは 8 つできた．

3．2．3 モデル式の構築

「いいね」と RT の合算値と投稿数を図示すると図 2 のようになっている．図 2 や津川[2]より，負の二項分布か 0 過剰ポアソン分布に従うと仮定する．またモデルの構造は，エンゲージメントと投稿内容変数の関係を表す回帰モデルと，回帰モデルにおける回帰係数パラメータを被説明変数とし，トライブを説明変数とした階層モデルを持つものである．想定モデル式は(1)の通りである．

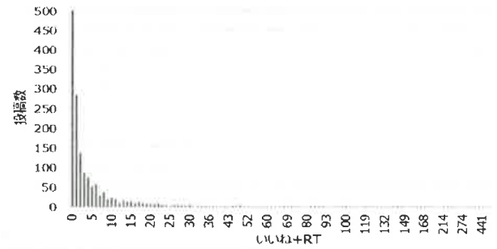

図 2．ハーレーダビッドソンのヒストグラム

$$y_i = \beta_0(1+\varepsilon_{0i}) + (\overline{\beta_1}+\varepsilon_{1i})x_1 + \cdots + (\overline{\beta_j}+\varepsilon_{ji})x_j \cdots (1)$$

y_i：目的変数　$\overline{\beta_j}$：説明変数 j における全体の効果
x_j：説明変数　ε_{ji}：トライブ i ごとの説明変数x_jへの効果
β_0：定数項　i=1~j(説明変数の数), j=1~n(トライブの数)

4 おわりに

現在，実務的にはトライブを用いたマーケティングが盛んに行われているものの，トライブの効果について学術的な検討はされていない．しかし SNS 広告による宣伝効果が明確に確認できない以上，学術的にトライブによるエンゲージメントの影響を推定することの価値は高いといえる．本研究では，トライブの違いによる投稿内容変数のエンゲージメントへの影響を推定するモデルを構築できた．

参考文献

[1] Patrick J Tighe, Ryan C Goldsmith, Michael G ravenstein, H Russell Bernard, Roger B Fillingi m「The Painful Tweet: Text, Sentiment, and Co mmunity Structure Analyses of Tweets Pertaini ng to Pain」, JOURNAL OF MEDICAL INTER NET RESEARCH, vol.17, iss.4, e84, pp.1-19 (2 015)

[2] 津川翔「ソーシャルネットワークのコミュニティ構造がソーシャルメディア上の投稿の拡散規模に与える影響の分析」, The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, pp1-4 (2018)

[3] 株式会社日経 BP「平成 27 年度経済産業省委託事業商取引適正化・製品安全に係る事業報告書」, http://www.meti.go.jp/policy/economy/consumer/consumer/pdf/sns_report.pdf, (2019 年 9 月 29 日閲覧)