

深夜アニメ番組に関するつぶやきに着目した Twitter 分析 Analysis of Tweets of Midnight Anime Programs through Twitter

相川 実穂¹ 仙海 かおり² 安田 智亜紀³ 酒井啓江⁵ 小川 祐樹⁴ 梅原 英一⁵
Miho Aikawa Kaori Senkai Chiaki Yasuda Hiroe Sakai Yuki Ogawa Eiichi Umehara

1. はじめに

近年、フェイスブックやツイッター(<http://twitter.com/>)などのミニブログが注目を集めている。中でもツイッターは、1回の投稿につき140文字の制限があり、これが手軽なコミュニケーションを促進している。中桐(2014)[10]は、9月19日に株式会社ビデオリサーチがテレビ番組の反応を測る新たな指標サービス「ツイッターTV エコー」を開始すると報道した。これは投稿ユーザー数、インプレッション数とインプレッションユーザー数からなる指標である。ツイッターは利用者だけでなく様々な分野において関心が高まっており、無視できない存在であることは明らかである。

さらに最近では、ツイートの内容に着目し、テレビ番組の視聴率の指数にツイッターを取り入れる動きまでもみられている。株式会社ビデオリサーチ(2014)[1]は、ツイッターをテレビ番組に利用している割合は26.1%であり、また、テレビ番組放送中にインターネット上に書き込む場合の98.6%がツイッターであるという結果を得ている。このことから、視聴率への注目度が増している動きがうかがえる。また、ジャンルごとの特性も分かっており、1分あたりのツイート件数平均で最も多いジャンルはアニメである。そこで本研究ではアニメ番組に関するツイートについて分析を行う。ツイート数、ツイートから得られる社会ネットワーク、ツイート内容の3つと視聴率の関係性を明らかにすることが、本論文の目的である。

本論文の構成は2節で先行研究をレビューする。3節で研究の目的や問題意識・仮説を提示する。4節では本論文で扱ったツイートと視聴率のデータについて述べる。5節では分析手順を示す。6節で分析結果を示す。7節で考察を述べ、8節で結論と今後の課題について述べる。

2. 先行研究

視聴率とSNSの関係では、齋藤ら(2014)[2]は、大きな話題となった朝ドラ『あまちゃん』が、視聴者にどう見られたかを知るため視聴率分析、世論調査、インターネット・アンケート調査、ソーシャル・リスニング調査の4つの調査を実施した。しかし、『あまちゃん』では形容詞について大きな流れや変化はみられなかった。今後の展望として、彼らは個々の発言を精査し、視聴者全体のトレンドの移り変わりをつかむことが出来れば、SNSでの発言が新しい指標となり得る可能性を指摘した。そこで本論文では、ツイッターに投稿されたツイートを形態素解析して単語を抽出し、発言内容と視聴率の関係を検証することとする。

ツイッターの社会ネットワーク分析では、濱岡(2012)[3]は、2011年10月上旬に公開された15本の映画のツイッ

ターにおけるコミュニケーションの社会ネットワークを分析した。彼らはツイッターにおけるリツイート(eクチコミ(eWOM))と捉えて、被リツイート回数を規定する要因を分析した。これらの要因はメッセージの内容、メッセージ送信者の特性などである。メッセージの形態のうち、「http」「@」の符号は負で有意であり、これらを含むものはRTされにくくなる。逆に「RT(拡散願い)」を含むものは実際にRTされる傾向があることがわかった。「値引き情報」はRT回数への影響がないが、「キャンペーン、プレゼント情報」「試写会、トークイベント開催」「公開日、公開時間」「出演者情報」「監督、ストーリー、主題歌情報」は正の相関があった。内容では「(観る前の)ネガティブ評価」のみが正で有意となった。これは、負のクチコミの方が広まりやすいという先行研究と一致する結果である。また、映画公式アカウントや出演タレントよりも、映画関係者などからの情報の方がRTされる傾向にあった。映画では公式サイトばかりでなく、監督などのプラスアルファの情報が重要である可能性を指摘した。そこで本論文でも社会ネットワーク分析を行い、アニメ番組の公式アカウント以外に番組関係者等その他のアカウントと視聴率との関係を検証する。

齊藤ら(2013)[4]は、形態素解析で障害となる表記揺れの修正方法について提案した。彼らは、インターネット上の投稿文から抽出した表記揺れの文章に正規表記を付け加えたデータを用いて、もっとも正規確率が高い経路を求め、表記の崩れパターンを自動抽出した。実験により従来法に比べ約30%の表記揺れの解析を改善できた。本論文では、形態素解析後に人力で単語の確認を行い、表記揺れによる不正確さを補うこととする。

小林ら(2010)[5]は、マイクロブログサービスTwitterにおいて実際にスポーツを観戦したりしながらリアルタイムに書き込む“実況書き込み”が増加している現象に目をつけ、これを検出する手法を提案した。イベント開始直後の書き込みは、「なう」や「はじまた」などイベント内容に依存しないものが多い。そこでイベント内容に依存しない表現を抽出し、開始直後の単語の出現頻度を学習することで識別器構築した後「観覧」「不観覧」のラベル付けを行った。風間(2013)[6]は、SNSの機能と人の繋がりが情報検索行動に対して持つ役割と意義に関して調査した。Webからの情報収集は、能動的と受動的の2つである。情報収集に用いるという観点で、良いSNSの条件は3つある。1つはサービス規模が大きいこと。2つ目はサービスの方向性が明確であることであること。3つ目はうまく機能が制限されていることである。この例としてTwitterを挙げた。Twitterはリアルタイム性の確保されており、情報拡散のためのRT機能や、コメントするためのリプライ機能が提供されている。本論文でもTwitterのメッセージを分析することで深夜アニメの視聴率との関係を見ることとする。

石原ら(2013)[7]は、東日本大震災時にTwitterにおいて情報発信の起点となった拡散アカウントを、次数中心性

1 日本電気	NEC
2 マイソフト	MySoft co.ltd.
3 MS&AD システムズ	MS&AD Systems
4 立命館大学	Ritsumeikan University
5 東京都市大学	Tokyo City University

と媒介中心性によって表現した。取得した総ツイートは3億3235万8199件である。分析の結果、震災後に中心性が高まったアカウントは、震災に関するアカウントや信頼できるアカウント(東大放射能チーム、首相官邸災害、東日本大震災の情報センターなど)であることがわかった。本論文においても次数中心性と媒介中心性を求めることにより、情報拡散の起点となるアカウントや情報を仲介するアカウントを特定し、これらの動きが視聴率の予測に生かせるか否かを検証する。

中桐(2014)[3]は、株式会社ビデオリサーチがテレビ番組の反応を測る新たな有料指標サービス「ツイッターTV エコー」を開始すると報道した。これはテレビ番組に関するつぶやき(ツイート)がいかにか拡散し、多くの人に見られているかを調べるサービスである。主な指標は「投稿ユーザー数」、「インプレッション数」、「インプレッションユーザー数」である。投稿ユーザー数は番組名や出演者の入ったフレーズを投稿した人数である。「インプレッション数」はユーザーが投稿したコメントが見られた延べ人数、「インプレッションユーザー数」は投稿を実際に見た人数である。これを放送時間中、前後一定の時間内で計測している。ビデオリサーチによると投稿数よりもインプレッションユーザー数が適していると述べている。本論文では、簡単に取得可能なツイッターのデータを用いて、視聴率の予測が可能かを調査する。そのために、ツイッター分析の先行研究で利用されている投稿数、社会ネットワーク分析、形態素解析による意味解析を用いて、視聴率との関係を調査する。

3. 仮説

インターネット上に溢れる評価情報を利用することが出来れば、より良い番組作りやマーケティング方法を提案できると考えられる。また、株式会社ビデオリサーチ[1]はツイッターの反応を番組評価に活かす取り組みが紹介している。しかし次週の視聴率を予測する指標としては明確には述べられてない。そこで、本研究ではアニメ番組に関するツイート数、ツイートから得られる社会ネットワーク、ツイート内容の3つと視聴率の関係性を明らかにする。

本稿の仮説は以下の6点である。

(仮説1) ツイート数の量とアニメ番組の視聴率は相関関係がある。

視聴という行動は関心があるからこそ起こる。そのため視聴した感想を吐き出すツールとしてツイッターの利用が盛んになり、視聴率が高いほどツイート数が多く、視聴率が低いほどツイート数が少ないと考えた。

(仮説2) 視聴率の高いテレビ番組は、番組公式アカウントへの興味が高い。

最新の情報を得られる番組公式アカウントは、視聴者からの興味を集めやすいため、番組関連アカウントの中でリツイートされる回数が最も多いと考えた。

(仮説3) 視聴率の高いテレビ番組は、番組関係者などの公式でないアカウント(非公式アカウント)への興味も高い。

番組自体への興味が高ければ、公式アカウント・非公式アカウントの両方に反応しリツイートする人も多いと思われる。よって、視聴率の高いテレビ番組ほど、非公式アカウントへのリツイートも多いと予測した。

(仮説4) 前週の実況ツイート内にプラス表現が多い番組ほど、次週の視聴率が高くなる。

前週にプラス表現のツイートが多ければ、番組の評価が上昇するはずである。好印象を受けるテレビ番組は継続して視聴されやすいため、視聴率が高くなると考えた。

(仮説5) 前週の実況ツイート内にマイナス表現が多いほど、次週の視聴率が高くなる。

濱岡(2012)のTwitterの分析結果によると、(観る前の)ネガティブ評価が重要であることが示されている。そこで本論文もプラス表現だけではなく、マイナス表現にも注目することとする。

4. 分析データ

4.1 取得したデータ

ツイッターメッセージは深夜アニメ番組に関するものに限定した。深夜アニメのクールの切り替わりである2014年7月3日から、10話が終わる2014年9月14日までとした。収集対象テレビ番組は11タイトルあり、それぞれのハッシュタグを含むツイートを対象とした。収集したデータの詳細を表1に示す。また、公式推奨ハッシュタグ以外にもツイッター上で独自に作られたハッシュタグも存在したが、情報の正確性を高めるため公式推奨ハッシュタグを含まないツイートは取得しないものとした。

結果、取得した総ツイート数は1,598,593件に上った。その内、リツイートとリプライは1,108,615件、ツイートは489,978件であった。ツイッターメッセージはツイート、リツイート、リプライの3種類に分類される。

本研究での投稿文の分類は、本文のみの投稿を「ツイート」、文頭にRTが付き@ユーザIDが含まれる投稿文を「リツイート」、文頭が@ユーザIDのみでそれ以降は本文になっているものを「リプライ」とした。また、URLを含むツイートは悪質なアダルト画像を載せた業者や自動発言システムによる投稿など迷惑ユーザが多かったため、番組公式アカウントのツイート以外は研究対象から除外した。

本研究では番組に関連の高いツイートを集めるため、番組公式が推奨しているハッシュタグを含んだツイートのみを収集した。収集したハッシュタグは表1の通りである。

表1 番組名とハッシュタグ一覧

アニメ番組のタイトル	公式ハッシュタグ	ツイート数
DRAMAtical Murder	#アニメドラマダ	7,414件
PSYCHO-PASS 新編集版	#pp_anime	38,457件
RAIL WARS!	#rw_anime	90,677件
Re: ハマトラ	#ハマトラ	13,201件
ばらかもん	#brkmn_anime	71,634件
ハナヤマタ	#ハナヤマタ	53,307件
ペルソナ4ザ・ゴールデン	#P4G_A	77,303件
黒執事 Book of Circus	#kuroshitsuji	33,991件
残響のテロル	#zantero	57,673件
戦国BASARA Judge End	#BASARA_JE	25,792件
普通の子供が【ろこどる】をやってみた。	#ろこどる	20,529件

4.2 視聴率データ

株式会社ビデオリサーチの関東地区で調査された表1に示した11タイトル各10話分の平均世帯視聴率のデータを用いた。世帯視聴率とは、テレビ保有世帯のうち番組を視

聴していた世帯の割合であり、一般的に使われている視聴率のことである。また株式会社ビデオリサーチ (2014) [8]によると、関東地区の場合、視聴率1%は約 18 万 1 千世帯である。

5. 分析手順

5.1 ツイート数

ツイート数とアニメ番組の視聴率を用いて相関分析を行った。今回使用したツイートデータは表 2 に示す。

表 2 視聴率との相関分析に用いたデータ

収集期間	2014 年 7 月 3 日～2014 年 9 月 14 日
データ概要	番組公式ツイッターが推奨するハッシュタグを含んだツイートのうち、各タイトルの各話が放送している間に投稿されたツイート (実況ツイート) のみを分析対象とした。また、「RT」や「@」の文字列を含んだ、リツイートやリプライと思われるツイートは排除した。
ツイート数	489,978 件

ツイートの収集は TwitterAPI を使用した。番組公式ツイッターが推奨するハッシュタグを検索キーワードとし、全番組の 1 話から 10 話のツイートを収集した。なお各タイトルが放送中に投稿されたツイート数の合計を表 1 に示す。各話ごとのツイート数と視聴率の相関は 3 つに分けて行った。

- (1) 各話のツイート数と各話の視聴率
- (2) 各話のツイート数と前週の視聴率
- (3) 各話のツイート数と次週の視聴率

5.2 社会ネットワーク分析

リツイート・リプライを用いて社会ネットワーク分析を行った。使用したツイートデータを表 3 に示す。

表 3 社会ネットワーク分析に用いたツイート

収集期間	2014 年 7 月 3 日～2014 年 9 月 14 日
データ概要	番組公式ツイッターが推奨するハッシュタグを含んだツイートのうち、上記期間中すべてのリツイートとリプライを分析対象とした。「RT」や「@」を含まない本文のみのツイートや、URL は削除した。
リツイート・リプライ数	1,108,615 件

表 4 各番組のクラスター係数

アニメ番組名のタイトル	次数中心性	媒介中心性
DRAMAtical Murder	0.00965	0.00000538
PSYCHO-PASS	0.01817	0.00000111
RAIL WARS!	0.03222	0.00000117
Re:ハマトラ	0.00286	0.00000120
ばらかもん	0.02991	0.00001686
ハナヤマタ	0.00415	0.00014932
ペルソナ4 ザ・ゴールデン	0.02647	0.00000291
黒執事	0.02925	0.00000786
残響のテロル	0.03599	0.00000423
戦国 BASARA	0.02616	0.00000315
普通的女子校生が【ろこどる】やってみた。	0.00128	0.00000480

本研究では次数中心性と媒介中心性という 2 つの指標を用いた。次数中心性とは、辺が集まっているノードほど中心性が高いという尺度である。媒介中心性とは、他のノードとノード点を結ぶ線が多いほど中心性が高いという尺度である。本研究ではツイートの関係は常に双方向であるという考えから、無方向性のネットワークとした。最後に次数中心性、媒介中心性のクラスター係数を求めた。これを表 4 に示す。クラスター係数とは、点をつないでいる辺の密度の平均値である。なお、計算には R を使用した。

5.3 形態素解析

表 5 評価情報の重みづけの例

プラス感情(+1)	中立(0)	マイナス感情(-1)
かわいい	先生	修羅場
イケメン	ピカチュウ	変
おもしろい	今期	酷い
愛	チョコレート	死ぬ

形態素解析に使用したデータは表 2 と同じである。形態素解析ツールは Igo¹を使用した。形態素解析で抽出した単語から、TF/IDF 法で単語の重要度を (1) 式で計算する。

$$w(t, d) = \ln(t_{f,t,d} + 1) \cdot \ln(N/df_d) \quad (1)$$

N: すべての投稿数

T_{f,t,d}: 投稿 t に単語 d が出現する頻度

df_d: 単語 d が出現する投稿の数

次に TF/IDF で計算した重要度が大きい単語から順に上位 50 件抜粋し、プラス感情とマイナス感情を判別して数値化した。数値化した例を表 5 に示す。

数値化は目視で行った。「かわいい」や「おもしろい」などのプラス感情の単語には 1、「死ぬ」、「嫉妬」などのマイナス感情の単語は -1、キャラクター名などのプラス感情とマイナス感情どちらにも分類されないものは中立とし 0 を割り付けた。最後にプラス感情とマイナス感情を算出した結果 (1 話分) と、それぞれの次週の視聴率との相関を見た。ここで、次週の視聴率を利用した理由は、前週のツイートの感情値によって、次週の視聴率に影響を及ぼすかどうか検証するためである。この作業を各 11 タイトル 10 話分行った。

6. 視聴率の分析結果

6.1 ツイート数の時系列推移

投稿されたツイート数の時系列推移、そしてリツイートとリプライ数の時系列推移を図 1 に示す。なおこの図は、ツイート数は「ハナヤマタ」、リツイート・リプライ数は「ばらかもん」のデータを用いて作成した。10 話まで同等の変化が見られたため 5 話までを記載した。また、他のタイトルについても同様にグラフを作成したが、全て同等のグラフとなったため割愛する。この 2 つの図から、ツイート数はアニメ番組放送中にのみ急激に伸びる傾向があり、一方リツイートとリプライ数は放送時間に関係なく増減する傾向にあることが分かった。このため以降の分析では、ツイート数を用いる場合はアニメ番組放送中に投稿されたデータ (以降、実況ツイートと呼ぶ) のみを使用する。リツイートとリプライ数を用いる場合はアニメ番組放送期間中に投稿されたすべてのデータを使用するものとする。

¹ <http://igo.sourceforge.jp/>

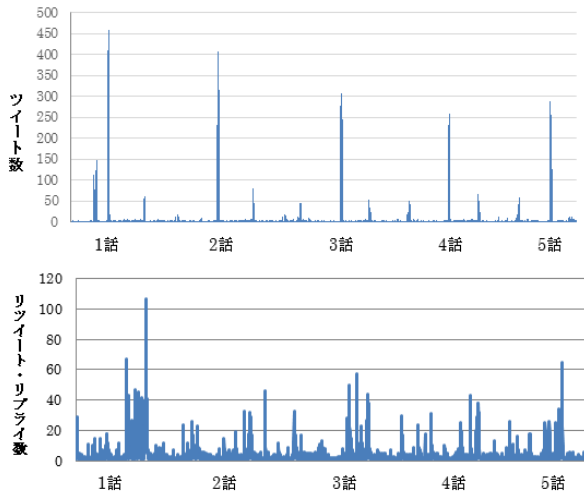


図1 ツイートとリツイート・リプライの時系列推移

6.2 ツイート数と視聴率

表1の11タイトルを分析対象とし、視聴率との相関分析を行った。(1)各話のツイート数と各話の視聴率(同時指標)、(2)各話のツイート数と前週の視聴率(遅行指標)、(3)各話のツイート数と次週の視聴率(先行指標)の3つの結果を表6に示す。また、それぞれの散布図を図2、図3および図4に示す。

この結果、仮説1(ツイート数の量とアニメ番組の視聴率は相関関係がある)は成立する。つまり、視聴者によるツイッターへの投稿が多い深夜アニメは、視聴率も高いと言える。さらに各話のツイート数は、翌週の視聴率とも正の相関がみられた。これにより、アニメ番組放送中に投稿されたツイート数は、翌週の番組視聴率の予測指数となる可能性が得られた。

表6 ツイート数と視聴率の相関分析の結果

	相関係数	p値	結果
(1) 同時指標	.4226**	.0000	1%有意
(2) 遅行指標	.3825**	.0001	1%有意
(3) 先行指標	.4396**	.0000	1%有意

** : 1%有意

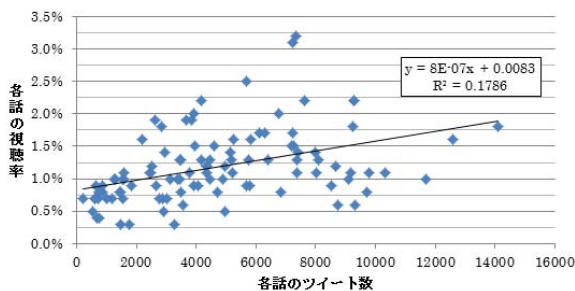


図2 (1) ツイートと視聴率(同時指標)

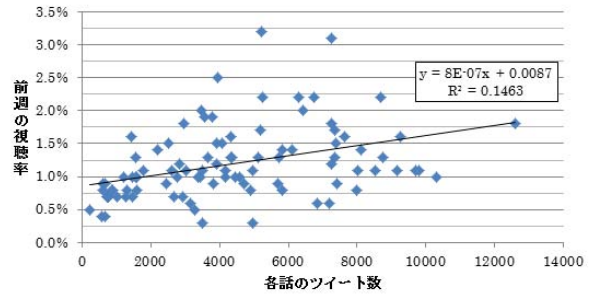


図3 (2) ツイートと前週の視聴率(遅行指標)

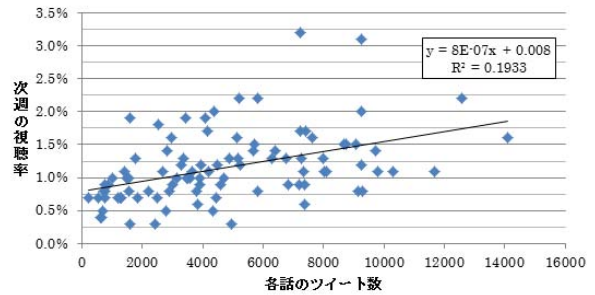


図4 (3) ツイート数と次週の視聴率(先行指標)

6.3 社会ネットワーク分析

視聴率と社会ネットワークの関係性を明らかにするため、R[9]を用いて各番組の平均視聴率と各番組の次数中心性・媒介中心性のクラスター係数で相関をとった。

6.3.1 次数中心性

表7 次数中心性のクラスター係数と平均視聴率

相関係数	p値	結果
.6485**	.0028	1%有意

** : 1%有意

次数中心性のクラスター係数と平均視聴率で相関分析した結果を、表7と図5に示す。分析した結果、正の相関が見られた。この結果、仮説2(視聴率の高い番組は、番組公式アカウントへの興味が高い)は成立する。つまり、リツイートやリプライに関する次数中心性のクラスター係数の値が大きいアニメ番組ほど、平均視聴率が高いという関係が成立する可能性がある。

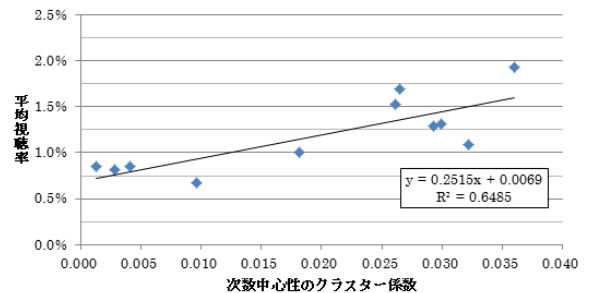


図5 次数中心性のクラスター係数と平均視聴率

参考までに cytoscape²を用いてネットワーク図を作成した。この結果、2種類の特徴が現れた。ここで中心となるアカウントが少ない一極集中型を「打ち上げ花火型」と名付ける。一方、情報を媒介するアカウントが多く存在し、全体的に分散して見えるものを「線香花火型」と名付けた。

² <http://www.cytoscape.org/>

次数中心性のクラスター係数が最も高い番組（残響のテロル）と最も低い番組（普通的女子校生が【ろこどる】やってみた。）でネットワーク図を6と図7に示す。

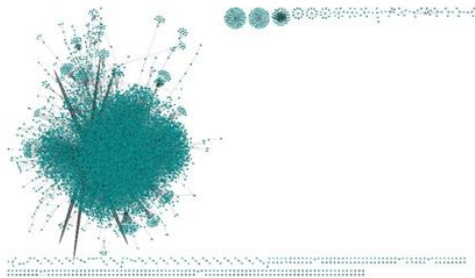


図6 次数中心性が高い打ち上げ花火型（残響のテロル）

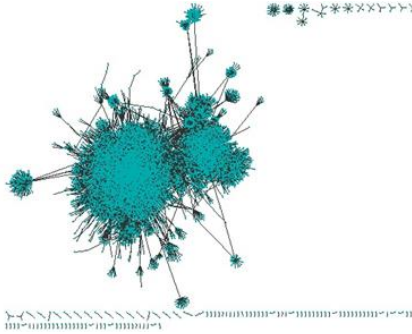


図7 線香花火型

（普通的女子校生が【ろこどる】やってみた。）

打ち上げ花火型は、ネットワークの中心にいる番組公式アカウント (@zankyono_terror) から放射線状に線が広がっている。ひとつのアカウントを個人それぞれがリツイートしているため、このような形になる。一方線香花火型は中心となるアカウントが複数あり、番組公式アカウント (@locodol_anime)、情報を拡散する bot (@Rokodoru_bot) や個人アカウントなど、それぞれが同程度リツイートされていた。

そこで、次は非公式アカウントと視聴率の関係を確認するために、媒介中心性と視聴率の関係を検証する。

6.3.2 媒介中心性

媒介中心性のクラスター係数と平均視聴率で相関分析した結果を、表8と図8に示す。また、他の値から大きく外れた番組（ハナヤマタ）は外れ値として扱うこととした。

表8 媒介中心性のと平均視聴率の相関分析

相関係数	p 値	結果
.0132	.4335	相関なし

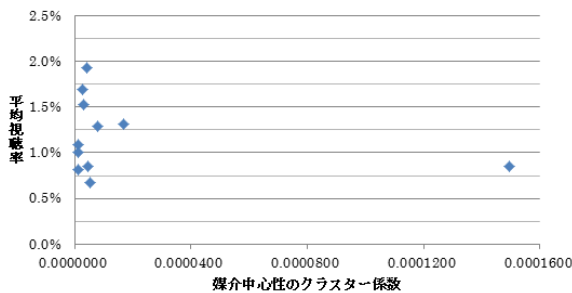


図8 媒介中心性と平均視聴率

表8および図8より、媒介中心性のクラスター係数と平均視聴率に統計的に有意な関係は確認できず、仮説3（視聴率の高いテレビ番組は、番組関係者などの公式でないア

カウント（非公式アカウント）への興味も高い）は棄却された。つまり、深夜アニメにおいては、非公式アカウントへのリツイート数と視聴率には関係がない。

また、媒介中心性のクラスター係数が最も高い番組（ばらかもん）と最も低い番組（PSYCHO-PASS 新編集版）でネットワーク図を作成する。こちらも同様に打ち上げ花火型と線香花火型に分かれた。その図を図9と図10に示す。打ち上げ花火型は、ネットワークの中心にいる番組公式アカウント (@brkmm_anime) から放射線状に線が広がった。ひとつのアカウントを個人それぞれがリツイートしているため、このような形になる。一方線香花火型は、中心となるアカウントが複数あり、番組公式アカウント (@psychopass_tv)、公式ショップ (@noitamina_shop)、他番組の公式アカウント (@hamatora_PR)、個人のアカウント等、それぞれが同程度リツイートされていた。

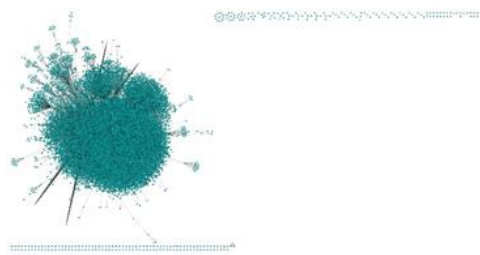


図9 媒介中心性の高い打ち上げ花火型（ばらかもん）

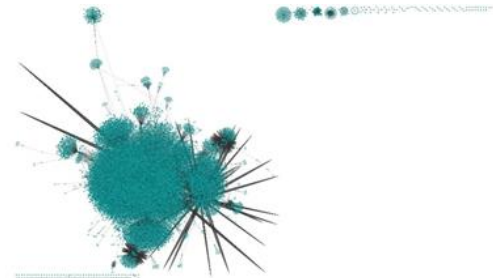


図10 媒介中心性の低い線香花火型（PSYCHO-PASS 新編集版）

また、外れ値としたハナヤマタで図を作成したところ、図11のようなネットワーク図が確認された。この番組は媒介中心性が高く次数中心性が極端に低いため、外れ値となった。しかし外れ値であっても、ネットワーク図の形状にはさほど差異は見られなかった。中心となるアカウントは番組公式アカウント (@hanayamata_PR) 以外にも、オープニング主題歌の作曲・編曲者 (@MONACA_tanaka)、画像 bot (@hanayamata_)、出演声優3名 (@iRis_w_yuki、@ayanon_y、@takaichan18) など複数みられた。しかしながら、この番組の平均視聴率はそれほど高くなかった。

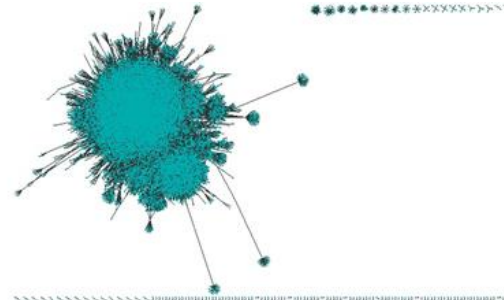


図11 媒介中心性の外れ値（ハナヤマタ）

6.3.3 社会ネットワーク分析の結論

社会ネットワーク分析から、視聴率と関係性があるのは次数中心性であった。一方、媒介中心性と視聴率には相関が見られなかったため、視聴率と媒介中心性は関係性が薄いことが分かった。ネットワーク図に関しては、次数中心性・媒介中心性ともにクラスター係数が低ければ線香花火型になり、高ければ打ち上げ花火型になるという特徴を見つけれられた。

平均視聴率との関係では、仮説2（視聴率の高いテレビ番組は、番組公式アカウントへの興味が深い）は成立したが、仮説3（視聴率の高いテレビ番組は、番組関係者などの公式でないアカウント（非公式アカウント）への興味も高い）は棄却された。つまり平均視聴率は、番組公式アカウントへの注目度から得られる次数中心性と強い関係があることが明らかとなった。

6.4 形態素解析と視聴率

表9 プラス感情の単語個数と視聴率の相関分析の結果

相関係数	p値	結果
.0075	.9429	相関なし

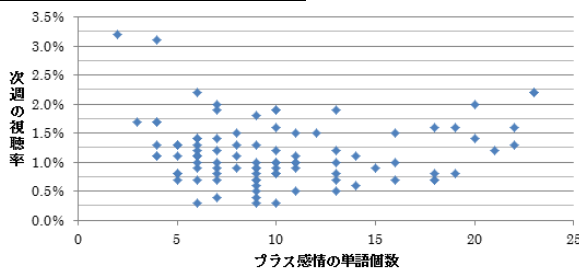


図12 プラス感情と視聴率の相関

実況ツイートにおけるTF/IDF法による上位50単語を抽出し、プラス感情とマイナス感情の単語の重みづけを行った。なお、プラス感情の単語/マイナス感情の単語の分類は、著者らの判断で行った。その結果、視聴率が高い番組も低い番組も出現単語に大きな流れや差は見られなかった。特徴としては、全ての放送期間を通じて「かわいい」「面白い」「良い」「楽しい」といったプラス感情の単語が上位に並ぶことが多い点である。また、マイナス感情の単語が上位に並ぶ時はその回の内容に沿ったものが多いという結果になった。

プラス感情であると評価した単語の個数と視聴率で相関分析をした結果を表9、図12に示す。

この結果、プラス感情と評価した単語の個数と視聴率は相関無しということが分かった。故に、仮説4（前週の実況ツイート内にプラス表現が多い番組ほど、次週の視聴率が高くなる）は棄却された。「かわいい」「良い」といったプラス感情の単語は、ほぼ全ての番組で共通して常に形態素解析の上位に食い込んでいたため、次週の視聴率の変動に影響しづらかったようだ。

次に、マイナス感情と評価した単語の個数と視聴率で相関分析をした結果を表10、図13に示す。

表10 マイナス感情の単語個数と視聴率

相関係数	p値	結果
.2150*	.0364	5%有意

*:5%有意

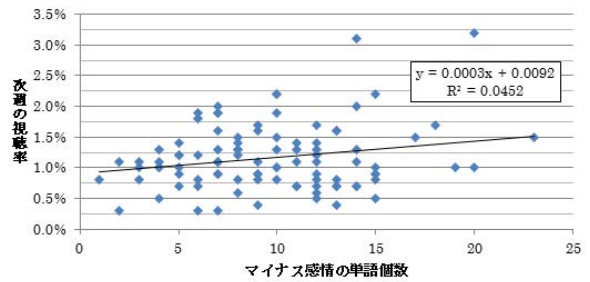


図13 マイナス感情と視聴率

分析の結果、5%水準で有意となり、相関ありという結果が得られた。故に、仮説5（前週の実況ツイート内にマイナス表現が多い番組ほど、次週の視聴率が高くなる）は成立した。

番組にもよるが「死ぬ」「酷い」といったマイナス感情の単語は、形態素解析の上位に入ることが少なかった。そのため、常に上位にあるプラス感情の単語をマイナス感情の単語が追い抜くような回では、マイナス感情で煽ることによって一種のスリルを味わわせ、次回への期待を高めていると考えられる。それにより、必然的に次週の視聴率への影響が大きくなったようだ。

7. 考察

今回の分析で検証した仮説の結果のまとめを、表11に示す。

表11 仮説検証の結果

仮説	仮説内容	成/棄
1	ツイート数の量とアニメ番組の視聴率は相関関係がある。	1%有意
2	視聴率の高いテレビ番組は、番組公式アカウントへの興味が深い。	1%有意
3	視聴率の高いテレビ番組は、番組関係者などの公式でないアカウント（非公式アカウント）への興味も高い。	棄却
4	前週の実況ツイート内にプラス表現が多い番組ほど、次週の視聴率が高くなる。	棄却
5	前週の実況ツイート内にマイナス表現が多いほど、次週の視聴率が高くなる。	5%有意

7.1 ツイート数

ツイート数の時系列推移については、常に一定のツイートが投稿されていたが、アニメ番組放送中になると倍以上の量に跳ね上がることが分かった。このことから、アニメ番組のハッシュタグを用いてツイートする時は、ほとんどがアニメ番組放送中の時間帯であると言える。次に、リツイートとリプライ数の時系列推移について述べる。ツイート数とは異なり、アニメ番組の放送時間帯でなくとも増減があることが分かった。これは、アニメ番組公式アカウントによる新しい情報やbotなど、時間の規定がないツイートが拡散されることで起こった現象だと考える。

第2に、ツイート数と視聴率の関係について述べる。相関係数の数値を見ると、(2)各話のツイート数と前週の視聴率は弱い相関だということが分かる。つまり、前週放送時の視聴率と、次に放送された際のツイート数との関係性は弱いことが読み取れた。一方、(1)各話のツイート数と各話の視聴率と、(3)各話のツイート数と次週の視聴率は相関ありという結果だった。これは、中桐(2014)[10]によるとアニメ番組はコアな視聴者が多いため、ツイ

ート数が増加する傾向にあることが要因だと考えられる。視聴中、なんらかの刺激が多いとツイート数が増加し、視聴率も高まり、次回の放送への期待が膨らむ。逆に刺激が少なければツイート数は減少し、視聴率は低下、次回の放送への意欲が生まれにくい。このことから、仮説 1 で述べた、ツイート数の量とアニメ番組の視聴率は相関関係があることは説明できると考えられる。つまりアニメ番組放送中に投稿されたツイート数は視聴率の予測に利用できる可能性があることが分かった。

7.2 社会ネットワーク分析

濱岡 (2012) [3]によれば、映画公式アカウントや出演者のリツイート数よりもプラスアルファの情報が得られるであろう bot や監督・演出家 (制作側) のリツイート数が増えるのではないかと述べられていたが、深夜アニメは若干異なる傾向にある。

まずは次数中心性について述べる。ネットワーク図から、次数中心性の中央にいるのは番組公式アカウントであることが分かった。次数中心性のクラスター係数が大きくなるほどはっきりとした打ち上げ花火型になることから、番組公式アカウントが宣伝に力を入れるほど視聴率は上がるといえる。一方その他のアカウントがどれだけ宣伝したとしても視聴率の上昇に貢献しているという結果は得られなかった。自分が面白いと感じたものを広めたい、自分も番組の宣伝に関わっているという一体感を得たい、フォロワーに知らせたいという点から、番組公式アカウントの情報を拡散する人が多いと考えられる。次数中心性のクラスター係数により、視聴率の高低の予測ができる可能性があると考えられる。

一方、媒介中心性のクラスター係数と視聴率には相関が見られなかったため、媒介中心性のクラスター係数を求めると視聴率の予測はできないといえる。このことから、視聴率の高いテレビ番組は、番組関係者などの公式でないアカウント (非公式アカウント) への興味も高いという仮説 3 は立証できなかった。

最後にネットワーク図について、媒介中心性の外れ値が現れた理由を述べていく。これには以下 4 点が挙げられると考えられる。

(1) 番組公式アカウント (@hanayamata_PR) 以外にも、オープニング主題歌の作曲・編曲者 (@MONACA_tanaka)、画像 bot (@hanayamata_)、メインヒロイン以外の出演声優 3 名 (@iRis_w_yuki、@ayanon_y、@takaichan18) など、番組制作側のアカウントが多いため、リツイート対象が他アニメと比べて複数あった。

(2) 本作の舞台は鎌倉・江の島と明確なため、「聖地巡礼」と呼ばれる舞台巡りに関する個人のツイートが多く、それがリツイートされた。

(3) アニメ番組放送開始の一週間ほど前に同番組のゲーム化が発表され、ゲームショップの公式アカウントの販売促進ツイートが多くリツイートされた。

(4) 株式会社ドワンゴが設立した動画共有サービス「ニコニコ動画」のコンテンツである「ニコニコ生放送」で本作出演声優 5 名による番組が放送された際に投稿された実況ツイートが多い。

社会ネットワーク分析についてまとめると、ネットワーク図と相関分析の結果から、番組公式アカウントの影響力が強くリツイートが増えると視聴率も上がり、反対に影響力が弱く視聴者のリツイート元がばらけるとリツイートが

視聴率に影響しにくいという傾向を得た。

7.3 形態素解析

分析の結果、視聴率の高い番組も低い番組も、形態素単位では大きな流れや変化は見られないことがわかった。理由としては、投稿されるツイート内容は番組内容に依存し、視聴者数には左右されなかったためであると推察される。しかしながら、仮説 4 (前週の実況ツイート内にプラス表現が多い番組ほど、次週の視聴率が高くなる) 及び仮説 5 (前週の実況ツイート内にマイナス表現が多い番組ほど、次週の視聴率が高くなる) を検証すると、別の側面が分かった。そこでプラス感情とマイナス感情の単語に注目してみると、それぞれ異なる特徴が見られた。つまり、仮説 4 は棄却されたが、仮説 5 は成立した。

プラス感情の「かわいい」「良い」「かっこいい」などの単語は、視聴率の高低に関係なく、放送期間を通じてほぼすべての番組で常に上位に並んでいた。しかし分析の結果、相関が見られなかったため、プラス感情の単語は視聴率にさほど影響を与えないということがわかった。

逆に、「死ぬ」や「つらい」などのマイナス感情の単語が上位に並ぶと、次週の視聴率が上がる傾向が見られた。そこでアニメ番組を詳しく調べてみると、マイナス感情のツイートが増えるような回は、長期間明かされていなかった謎が解明されたり、登場人物が絶体絶命に陥っていたりするなど、緊張感のあるシーンや危機的なシーンが次週に引き継がれる内容が多いことがわかった。つまり、このようなネガティブな内容で視聴者の危機感を煽ることで次週の内容に関心が集まり、その結果、次週の視聴率が上がったと考えられる。実際、複数の話に渡ってポジティブな内容を放送している番組よりも、時折ネガティブな内容を流して視聴者に一種のスリルを味あわせ、次週の内容を気にかけてさせるような番組の方が視聴率も高く、ツイートの伸びも高い傾向があった。これらのことから、視聴者の危機感を煽る番組ほど、視聴率が高くなると考えられる。次回視聴率の予測指標として、マイナス表現の単語の出現頻度は重要であることが分かった。

7.4 視聴率予測指標

本論文の分析結果より、以下の 3 つ指標が視聴率の予測に有効である可能性があることが明らかとなった。

- ① アニメ番組放送中に投稿されたツイート数
- ② アニメ番組放送期間中のリツイートとリプライに関する次数中心性
- ③ アニメ番組放送中に投稿されたツイートにおけるマイナス感情を持つ単語の重要度

この中で実況ツイートに着目して、①投稿されたツイート数、③マイナス感情を持つ単語の出現数の 2 変数を独立変数、次回視聴率を従属変数とし、ステップワイズ重回帰分析を行った。その結果を表 12 に示す。2 変数が抽出され、決定係数は.227 となった。故に、この実況ツイートに関する 2 つの変数は視聴率の予測に利用できる可能性があると考えられる。

表 12 重回帰分析の結果

	相関係数	決定係数	F 値
①ツイート数	0.213	.045	20.3799
③マイナス単語の個数	0.424	.180	4.3982
①、③	0.476	.227	13.4847

8. 結論と今後の課題

本研究ではアニメ番組に関するツイートについて分析を行い、ツイート数、ツイートから得られる社会ネットワーク、ツイート内容の3つと視聴率の関係性を明らかにすることを目的とした。ツイート数においては、実況ツイート数が増加すると、視聴率も高まり、次回の放送への期待が膨らむということが分かった。社会ネットワーク分析においては、番組公式アカウントの注目度が高くリツイートが増えれば視聴率が上がるということが分かった。形態素解析と視聴率の関係においては、視聴者の危機感を煽るなどマイナス感情のツイートが多い番組ほど、視聴率が高くなることが分かった。以上のことから、深夜アニメ番組に関するツイートと視聴率の関係性は以下3点の知見を得た。

(1) 前話の実況ツイート数が増えるほど、次回の視聴率が伸びる傾向がある

(2) 回数中心性のクラスター係数が大きいほど、平均視聴率が高い傾向がある

(3) 前話の実況ツイートにマイナス表現が多いほど、次回の視聴率が伸びる傾向にある

以上の3つの指標は、視聴率の予測指標として利用できる可能性のあることが分かった。

しかし、今回の研究ではいくつかの課題が見つかった。

(1) 形態素解析の際、品詞ごとに分類してその移り変わりなどを調査すれば、視聴者がどのタイミングで何に感情移入していたかわかる可能性がある。

(2) 本研究は夏季アニメだけを対象としたため、他の時期に放送されているアニメやアニメ以外の番組にも同じ調査を行い、今回の結果が共通か否か探る必要がある。

(3) 中田ら(2012)[11]によると、1ツイートあたりの品詞の出現頻度の平均はリツイート及びお気に入り登録における各項目の値が大きくなったことがわかっている。つまり、実況ツイートのお気に入りの数をカウントすることで、視聴率の指標を増やすことができる可能性がある。

(4) 共起分析を行いさらに詳しく単語を調べることで、より、キャラクター名や、「可愛い」「かわいい」「カワイイ」などの同じ意味を持つ単語の表記揺れを修正することができる可能性がある。

(5) 社会ネットワーク分析において、本論文では1話から10話の全放送を通じた回数中心性のクラスター係数と平均視聴率を検討し、回数中心性のクラスター係数と平均視聴率には強い相関があることを明らかにした。しかし本研究では、前話までのクラスター係数を用いて次回視聴率を予測できるか否かについて確認できなかった。このため、各話ごとのクラスター係数と各回の視聴率との相関を求め、回数中心性は次回視聴率が予測できる指数であるかの可能性を検討する必要がある。

(6) ツイート収集の際、公式番組の推奨ハッシュタグのついたメッセージを収集対象にしたために、打ち上げ花火型が多く出た可能性がある。今後ツイートを収集する場合は、公式推奨のもの以外のハッシュタグも併せて収集し、この可能性について検証したい。

(7) 今回の分析では、前話の実況ツイート数、回数中心性のクラスター係数、前話の実況ツイートの出現上位単語のマイナス表現単語数の3パラメータが視聴率に影響を与えていることが分かった。この3パラメータを使った視聴率推計モデルを構築することは今後の課題である。

(8) 本論文の分析では、11タイトルを同一の方法で分析した。しかし、11タイトルの中には青春、スポーツ、ス

リルなど様々なジャンルが含まれている。ジャンルごとに詳細な分析を行うことも今後の課題である。

謝辞

本研究はJSPS 科研費 25380481 の助成を受けたものです。

参考文献

- [1] 株式会社ビデオリサーチ, “テレビ×ビデオリサーチ 実例! Twitter からみるテレビ番組評価”, URL (<http://www.video.co.jp/casestudies/research/tv/2013/04.htm>), 2014.
- [2] 齋藤建作・二瓶互・関口聡・三矢恵子, “朝ドラ『あまちゃん』はどう見られたか〜4つの調査を通して探る視聴のひろがり視聴熱”, 放送研究と調査 MARCH 2014, pp.12-41, 2014.
- [3] 濱田豊, “Twitter におけるコミュニケーションの社会ネットワーク分析”, 情報処理学会第74回全国大会要録集, pp.4-521-522, 2012.
- [4] 齊藤いつみ, 貞光九月, 浅野久子, 松尾義博, “正規崩れ表記のアライメントに基づく表記崩れパターンの抽出と形態素解析への導入”, 情報処理学会研究報告自然言語処理 (NL) 2013-NL-214(5), pp.1-9, 2013.
- [5] 小林尊志, 野田雅文, 出口大輔, 高橋友和, 井手一郎, 村瀬洋, “Twitter における実況書き込み検出手法の検討”, 映像情報メディア学会技術報告 34(25), pp.129-130, 2010.
- [6] 風間一洋, “ソーシャルネットワークによる Web からの情報収集”, 情報の科学と技術 63 巻 1 号, pp.28-33, 2013.
- [7] 石原裕規, 諏訪博彦, 鳥海不二夫, 太田敏澄, “震災前後における中心性に基づく Twitter ネットワーク分析”, 第19回社会情報システム学シンポジウム(2013.1.23), pp.77-82, 2013.
- [8] 株式会社ビデオリサーチ, “TV RATING GUIDE BOOK”, pp.22, 2014.
- [9] 鈴木努, “R で学ぶデータサイエンス 8 ネットワーク分析”, 共立出版, pp.1, 2-3, 2009.
- [10] 中桐基喜, “ビデオリサーチの新サービス ツイッターTV エコーとは?”, 日経エンタテインメント! 2014 年 12 月号, pp.69, 2014.
- [11] 中田侑輝, 上岡英史, “形態素解析を用いた Twitter における影響力のあるツイート抽出”, 2012 年電子情報通信学会通信ソサイエティ大会, pp.505, 2012.