

# Twitterにおけるテレビ番組の実況ツイートの可視化

菅野 遥菜<sup>1</sup> 竹島 由里子<sup>1</sup> 加納 徹<sup>1</sup>

東京工科大学

## 1. はじめに

近年、ソーシャルメディアの普及により、ソーシャル・ネットワークング・サービス (SNS) の利用者が増加している。中でも、Twitter の日本国内でのツイート数は月に 10 億にのぼり、さまざまな人が発信した身の周りの出来事や情報をリアルタイムに知ることができる。また、最近ではテレビ番組を見ながらハッシュタグを付加した実況ツイートをして、同じテレビ番組を見ている視聴者同士で感想を共有することが、新たな視聴の楽しみの一つとなっている。さらに、視聴率を上げるため、そして盛り上げるために、テレビ番組と SNS を連携させている企画も多々ある [1]。しかし、Twitter 上で検索して表示される一覧から、視聴者の反応を把握することは難しい。

この問題を解決するため、本研究では実況ツイートをした視聴者の反応を、感情値の変化とともに可視化するシステムを提案する。まず、「ハッシュタグ+番組名」が付加されたツイート取得し、番組前、番組中、番組後の時間帯別に感情分析を行う。さらに、取得したツイートの中で最も使用された単語や印象語に対し、視聴者がどのような感情を持っていたのかを同時に可視化する。これによって、Twitter のツイート一覧や、視聴率の数字からは得られない感情の変化や視聴者の反応を把握することが容易になる。

## 2. 処理の流れ

### 2.1 ツイート取得

ツイート取得には、Twitter 社が提供しているツイートの参照や検索などを行うアプリケーション開発のための Twitter API と、Java 用のライブラリである Twitter4J を用いる。特定のハッシュタグが付加されたテレビ番組の実況ツイートは、番組前、番組中、番組後の各時間帯で取得する。

### 2.2 形態素解析

取得したツイートを、言語で意味を持つ最小単位に分割するために、日本語形態素解析エンジンである MeCab と、

MeCab を R から使用する RMeCab のパッケージを利用する。形態素解析の処理を行い、取得した実況ツイート全体から、頻出度の高い、番組のイメージを掴むキーワードになる単語を取得する。本研究では、可視化するキーワードを 16 個とした。その際、番組全体、前、中、後の時間帯別に番組のキーワードを含む実況ツイートのテキストファイルを、それぞれのキーワードごとに 16 個作成する。

### 2.3 感情値算出

熊本ら [2] が開発した感情値算出ツールを利用し、番組のキーワードを含む実況ツイートの感情値を算出する。感情値は 1～7 の実数値で文章ごとに算出され、感情値が 1 に近い場合は「楽しい」「うれしい」「のどか」のポジティブな感情を感じることを表し、7 に近い場合は「悲しい」「怒り」「緊迫」のネガティブな感情を感じることを表す。また、感情値が 4 のときは、どちらも感じないニュートラルの感情であることを表す。

### 2.4 可視化

ツイート取得、形態素解析、感情値算出の処理を行ったデータを使用し、番組全体、番組前、番組中、番組後の時間帯別に、キーワードの頻出度と感情値がどのように変化しているかを可視化する。感情値の変化は、キーワードのフォントの色を変化させることで表現する。色相が赤に近いほどポジティブであることを、青に近いほどネガティブであることを示す。また、頻出度の値の変化に応じてキーワードのフォントサイズを変化させる。キーワードは、頻出度が高いものを中心に、頻出度が低くなるにつれて時計回りに配置する。

## 3. 結果

### 3.1 「逃げるは恥だが役に立つ」の可視化結果

本研究では、地上波映画の番組、テレビアニメ、連続テレビドラマの 3 種類の番組を可視化した。2016 年 9 月 7 日放送「超高速！参勤交代」、2016 年 9 月 9 日放送「クレヨンしんちゃん」、2016 年 11 月 15 日放送「逃げるは恥だが役に立つ 第 6 話」の 3 番組である。図 1 に、「逃げるは恥だが役に立つ 第 6 話」の番組全体の可視化結果を示す。こ

\*Visualization of Twitter comments for TV shows

<sup>1</sup> Tokyo University of Technology

の可視化結果から、ドラマタイトルや、出演者名などの単語が頻出度の高いキーワードとして抽出されていることがわかる。また、全体的にどのキーワードも、ポジティブからネガティブにかけてほぼ全ての感情値で出現している。番組中では、「旅行」「セーラー服」「アイマスク」という各シーンに関する単語が多く、番組後では「死人」「最後」「やばい」といった単語が多く見られた。



図 1 「逃げるは恥だが役に立つ 第 6 話」番組全体の可視化結果

### 3.2 ユーザ実験

#### 3.2.1 概要

本システムの可視化結果から、視聴者の評価を把握することが出来るかについてのユーザ実験を実施した。

- ・対象：大学生 8 人
- ・実施内容：ユーザに「超高速！参勤交代」, 「クレヨンしんちゃん (冷やし中華が食べたいゾ、手作りブローチをあげるゾ)」, 「逃げるは恥だが役に立つ 第 6 話」の実況ツイートの可視化結果を見てもらい、使用後に各番組の可視化結果についてアンケートを行う。

- ・質問内容：(質問 1) 番組を視聴したか (質問 2) 可視化結果を見てどのような番組だと思ったか (質問 3) 可視化結果から視聴者がどのような評価をしているか、把握することが出来るか

#### 3.2.2 アンケート結果

可視化結果から番組の評価を把握することが出来るか問う質問の結果を、表 1 に示す。回答は 4 段階評価とし、以下のように定義する。

- ・評価 1：出来ると思う
- ・評価 2：どちらかと言えば出来ると思う
- ・評価 3：どちらかと言えば出来ないと思う
- ・評価 4：出来ないと思う

番組を視聴したと回答したユーザは、いずれの場合も「出来ると思う」と回答した。しかし、「超高速！参勤交代」と「クレヨンしんちゃん」では、「出来ると思う」という回答が多いものの、「出来ないと思う」といった回答も少数得られた。また、質問 3 では「恋ダンス」を連想させるキーワードや、共演者の名前があることから「流行の番組」, 「ガッ

キー (新垣結衣) と星野源が出ている番組」といった回答が得られた。「超高速！参勤交代」では「歴史を題材にした映画」「江戸時代に基づいた話」といった回答が得られ、「クレヨンしんちゃん」では、「ひろし」に関しての回答が多く見られた。

表 1 質問 2 の集計結果

	評価 1	評価 2	評価 3	評価 4
超高速！参勤交代	0 人	6 人	2 人	0 人
クレヨンしんちゃん	2 人	5 人	1 人	0 人
逃げるは恥だが役に立つ	4 人	4 人	0 人	0 人

## 4. 考察

可視化を行った 3 番組全てにおいて、視聴していないと回答したユーザが多かった。しかし、番組の評価を把握することが出来るかに関しては、「どちらかと言えば出来ると思う」「出来ると思う」といった前向きな結果が多数得られた。これは、視聴していないユーザでも実況ツイートをした視聴者の感情を表す、「楽しみ」「泣き」「面白い」といった言葉が見られたためであると考えられる。反対に、「超高速！参勤交代」では、「参勤交代」「超高速」「徳川」といった言葉が多く、「クレヨンしんちゃん」では、「ひろし」「クレヨンしんちゃん」「声」といった言葉が多く見られ、視聴者の感情を表す言葉が少なかった。そのため、視聴者の反応を把握することが難しくなったと考えられる。以上のことから、キーワード中に含まれる名詞や形容詞などの品詞の比率によって、評価の把握の容易さが左右されることが考えられる。

## 5. おわりに

本稿では、Twitter から特定のテレビ番組の実況ツイートを取得し、頻出単語とその感情に応じてどのような変化をしているか可視化した。さらに、その可視化結果から、視聴者の反応や評価の把握が容易になるかを、ユーザ実験により確認した。番組の評価の把握に関しては、「出来ると思う」という意見が大多数を占めた結果となった。しかし、各番組で可視化したキーワードの品詞の違いによって評価の把握の容易さが左右されてしまうという課題が残った。今後は、抽出するキーワードの品詞に注目し、可視化システムの改善を行う。

### 参考文献

[1] テレビ×ビデオリサーチ 実例！ Twitter からみるテレビ番組評価：ビデオリサーチ, 入手先 (<http://www.videor.co.jp/casestudies/research/tv/2013/04.htm>) (参照 2016-12).

[2] 熊本 忠彦, 河合 由起子, 田中 克己: 新聞記事を対象とするテキスト印象マイニング手法の設計と評価, 電子情報通信学会論文誌 (2011).