

# スレッド生成による映像・音声放送番組におけるツイートの補完手法の提案

A Study of Thread Generation Methods for Analyzing Program Hashtag Tweets in Audio Broadcasting

森田 陸斗<sup>1</sup> 栗 達<sup>1</sup> 河合 由起子<sup>1</sup>  
Morita Rikuto Da Li Yukiko Kawai

京都産業大学情報理工学部 〒 603-8555 京都府京都市北区上賀茂本山<sup>1</sup>

## 1 はじめに

近年、Twitter など SNS の普及に伴い、テレビやラジオなどの放送番組を視聴しながらリアルタイムに感想等を投稿するソーシャルビューイング (SV) という形態が一般化してきている。放送番組の内容とツイート内容は時系列に沿ってある程度の同期性があるが、既に過ぎ去った話題にも関わらずツイートされるケースは多く存在するため時系列順のツイートを見ても番組の流れの概要を掴むことは難しい。また、放送内容とツイート内容の話題の時系列の差が大きくなるにつれ、ツイートの内容が過去のどのトピックかを理解することは困難である。そこで本研究では、番組を視聴しているユーザーのツイートを取得し、トピックに対するスレッドをリアルタイムに生成する手法を提案する。また、番組内容と同期するツイートにおいて欠損しがちな情報に対して、番組と非同期でも理解可能なツイートの補完手法を提案する。

これまで番組に対するツイートを収集し、解析する研究は広く取り組まれおり、平岡ら [1] は、番組ハッシュタグを含むツイートを収集し、それらと共に共起する単語を抽出し、抽出した単語を含むツイートも追加する手法を提案している。それらハッシュタグを含めたツイートを番組放送中に投稿したユーザーを視聴ユーザーとして、視聴率を可視化している。しかしツイート数に基づく視聴率の提供にとどまっており、盛り上がりは明らかとなるが、ツイートの内容から番組のトピックを理解することは困難である。切原ら [2] は、収集したツイートを BERT を用いてベクトル化し、EmbedRank アルゴリズムによって番組内容を示すキーワードを抽出し、番組のシーン分類を行う手法を提案している。ツイートからシーンに対するキーワード抽出は有用であるが、シーンに出現する内容とツイート内容との関連性は分析されておらず、シーンと非同期のツイート分析の有用性は明らかではない。本研究は、番組のトピックと同期するツイートだけでなく、非同期のツイートとのトピックの関連性を明らかにし、リアルタイムにスレッドを生成しツイートを補完する点が特異であり、提案手法によりツイートと番組の双方の興味喚起の実現を目指す。

## 2 放送番組に関するツイートのスレッド生成・情報補完手法

図 1 から従来のツイートの時間軸による提供手法 (左) とスレッド生成と補完による提案手法 (右) を示す。提案手法では、まず、放送番組で提供されているハッシュタグを取得する。ハッシュタグの指定がなければ「#番組



図 1 従来のアプリケーションと提案手法の比較  
組名」とする。次に、以下を再帰処理する。

指定したハッシュタグを含むツイートを Twitter API より重複を除き取得する。一定量ツイートを取得後、単語の TfIdf の出現頻度と感情値を算出し、k-means 等よりクラスタリングし、スレッドを 1 つ生成する。スレッド生成後、新たに取得したツイートがスレッドに分類されるか否か判定する。スレッドに分類される場合は、ツイートの文字数が閾値以下かつ主語・被修飾語 (名詞) を伴わない述語・修飾語 (動詞・サ変名詞・形容詞) に対して、主語を補完する。補完語は、スレッドごとに TfIdf より抽出した上位を特徴語とし、共起頻度から判定する。スレッドに分類されない場合は、一定量のツイート取得後に再度クラスタリングを行い、新たにスレッドを 1 つ生成する。複数に分類される場合は複数で提示する。

## 3 おわりに

本研究では、番組に同期・非同期するツイートの出現頻度と感情値に基づきクラスタリングし、スレッドの生成および主語を補完する手法を提案した。今後システムを構築し、同期・非同期ツイートのリアルタイムなスレッド生成ならびに補完手法を検証する。

## 謝辞

本研究の一部は、科研費基盤 (C)(19K12240) および京産大 HMD センターの研究活動による。ここに記して謝意を表す。

## 参考文献

- [1] 平岡照久, 中村和裕, 荒木伸也, 阿部倫之. ツイートを利用したテレビ視聴傾向の評価方法. 第 77 回全国大会講演論文集 3M-07, 2015.
- [2] 切原大海, 松本和幸, 吉田稔, 北研二. テレビ番組視聴者のツイートからのキーワード抽出および分類. 第 34 回全国大会 4Rin1-47, 2020.