

メディアの報道に対する ソーシャルメディアユーザの感情の可視化手法

西上 貴雅[†] 風間 一洋[†] 吉田 光男^{††} 土方 嘉徳^{†††}

[†] 和歌山大学システム工学部 〒640-8441 和歌山県和歌山市栄谷 930

^{††} 筑波大学ビジネスサイエンス系 〒112-0012 東京都文京区大塚 3-29-1

^{†††} 関西学院大学商学部 〒662-8501 兵庫県西宮市上ヶ原一番町 1-155

E-mail: [†]takamasa.nishigami@gmail.com, ^{††}kazama@ingrid.org, ^{†††}mitsuo@gssm.otsuka.tsukuba.ac.jp,
^{††††}contact@soc-research.org

あらまし メディアが報じるニュースのトピックによっては、ソーシャルメディア上でユーザの意見の分極化が発生するが、その現象を直感的に捉えることは難しい。本論文では、各種メディアの報道に対するソーシャルメディア上のユーザの感情の傾向を、色の分布として捉えるマクロな視点と、個々のユーザの割合として観察するミクロな視点の両方から可視化・分析する手法を提案する。まず、特定のトピックを表すクエリで検索したニュース記事とそれを引用・言及したポスト（旧：ツイート）の共引用関係を利用し、注目しているニュース記事が類似するユーザが近くに集まるように配置する。そして、Louvain 法で分割したコミュニティごとにノードの外側の色を変え、難波らの T5 感情検出器で抽出した感情を統計処理して、ノードの内側を感情ごとに異なる色の円グラフとして表示する。これにより、個々のユーザに焦点を当てると感情の割合が分かり、一方でコミュニティや全体を俯瞰する際には感情の色が視覚混合され単一の色としてユーザの分布が把握できる。さらに、実際のデータに適用して複数のトピックを分析することで、提案手法の有効性を示す。その結果、ユーザが言及したニュース記事が似ているコミュニティ間でユーザの感情に偏りがあることを明らかにした。また、提案した可視化手法は、コミュニティ全体の感情を直感的に捉えるだけでなく、個々のユーザに注目することでユーザが抱く感情を理解でき、グラフ構造の可視化方法において感情の分布を視覚的に理解する手段として有用であることが確認できた。

キーワード ソーシャルメディア、マイクロブログ、ニュース・マスコミ、情報可視化、ユーザ理解

1 はじめに

1.1 研究の背景

メディアのニュース報道は、そのトピックの内容によっては X（旧：Twitter）などのソーシャルメディア上で広く拡散されて、多くのユーザに大きな影響を与えることが知られている [1]。その影響力の強さゆえに、ユーザの意見を操作するために事実とは異なるフェイクニュースや釣り記事が生み出されたり、政府の報道機関への介入や報道機関の利害関係者への忖度により都合が悪い情報が隠蔽される事例が見られるようになった。さらに、ユーザの個人情報や学習するアルゴリズムによって異なる意見や価値観の情報が隠されるフィルターバブル [2] や、自分と類似したユーザとばかり繋がることで情報の多様性が失われてしまうエコーチェンバー現象 [3] により、ユーザの意見の分極化を招くようになった。

そこで、中立的な視点から多様性のある情報を理解するためには、異なる視点からの報道の全体構造を俯瞰して、ソーシャルメディア空間における意見の分布や偏りを把握することが必要になる。

1.2 研究の目的

本論文では、あるトピックに関する各種メディアのニュース報道に対してソーシャルメディア上のユーザが持つ感情の分布や偏りを、色の分布として捉えるマクロな視点と、個々のユーザの感情の割合として観察するミクロな視点の両方から可視化することで、直感的に理解できるようにする手法を提案する。

まず、特定のトピックを表すクエリで検索したニュース記事とそれを引用・言及したポスト（旧：ツイート）の共引用関係を利用し、注目しているメディアやニュース記事が類似するユーザが近くに集まるように配置する。そして、Louvain 法で分割したコミュニティごとにノードの外側の色を変え、難波らの T5 感情検出器で抽出した感情を統計処理して、ノードの内側を感情ごとに異なる色の円グラフとして表示する。これにより、個々のユーザに焦点を当てると感情の割合が分かり、一方でコミュニティや全体を俯瞰する際には感情の色が視覚混合 [4] され単一の色としてユーザの分布が把握できるようになる。さらに、実際のデータに適用して複数のトピックを分析することで、提案手法の有効性を示す。

2 関連研究

ソーシャルメディア分析では、ユーザの感情を知るために、

テキストから怒りや喜びといった感情 (emotion) を分析する感情分析器や検出された感情がポジティブかネガティブかニュートラルかという極性 (sentiment) を分析する極性分析器が用いられることが多い。これらは、辞書ベースの手法と機械学習ベースの手法に大きく分類できる。以下では特に日本語に限定して説明する。

辞書ベースの感情分析器として、ML-Ask (eMotive eLement and Expression Analysis system) がよく用いられる [5] [6]。これは中村の 10 種類の感情分類 [7] に基づいた感情語辞書を用いて、マッチした感情表現の数から感情成分を推定するシステムであり、CVS (Contextual Valence Shifters) [8] に基づいて否定表現も考慮できる。辞書ベースの極性分析には、日本語評価極性辞書がよく用いられる [9] [10]。これは名詞に評価極性情報を手付けした辞書である。辞書ベースの手法は簡単に利用できるが、検出結果が辞書に大きく左右されるだけでなく、感情語が含まれない場合や複雑な状況の正確な判定が難しいなどの問題が存在する。

機械学習ベースの感情・極性分析器としては、福田らによる BERT を用いた感情分析器が存在する [11]。これは事前学習済み日本語 BERT モデルを Kajiwara らの日本語感情分析データセット WRIME [12]¹ でファインチューニングした感情分析器である。WRIME は、SNS 上の投稿に対して、Plutchik の基本 8 感情 [13] (喜び, 悲しみ, 期待, 驚き, 怒り, 恐れ, 嫌悪, 信頼) を 4 段階 (0, 1, 2, 3), 極性を 5 段階 (-2, -1, 0, 1, 2) で人手でラベル付けしたデータセットである。また、難波らはその後継として T5 を用いた感情分析器を提案した [14]。これは約 100GB の日本語コーパスで事前学習した T5 モデルを、さらに WRIME でファインチューニングした感情分析器であり、特に感情別に T5 モデルをファインチューニングした場合に、BERT モデルや全感情を一度にファインチューニングした T5 モデルよりも高い分類性能を示した。

実際に、感情分析器をソーシャルメディア分析に用いた事例として、鳥海らはソーシャルメディア上の新型コロナウイルスに関する投稿を収集し、ML-Ask を用いて各感情成分の時系列変化と日毎の感情成分の相関を分析した [15]。笹原らは、Twitter のコロナ禍における転売現象に関するポストに対して、感情辞書 ML-Ask に加えて、日本語版の心理学的カテゴリ辞書 J-LIWC と道徳基盤辞書 J-MFD を用いて、消費者心理・行動を感情成分と動因カテゴリ成分、モラル成分の時系列変化を分析した [16]。福田らは BERT を用いた感情分析器を用いて、Twitter Search API で「ワクチン」をクエリとして収集したポストデータセットに対して、各感情ラベルが付与されたポスト割合の時系列変化を分析した [11]。難波らは、感情分析器を用いて「恐怖」または「不安」が 0 から 3 の 4 段階のうち 3 と判定された不安ポストをさらに 10 カテゴリーに分類し、各カテゴリーのポスト数を時系列で可視化することで、技術の社会的受容性を分析した [14]。すなわち、今までは感情分析器の結果は、統計的または時系列的な分析に用いられることが多かった。

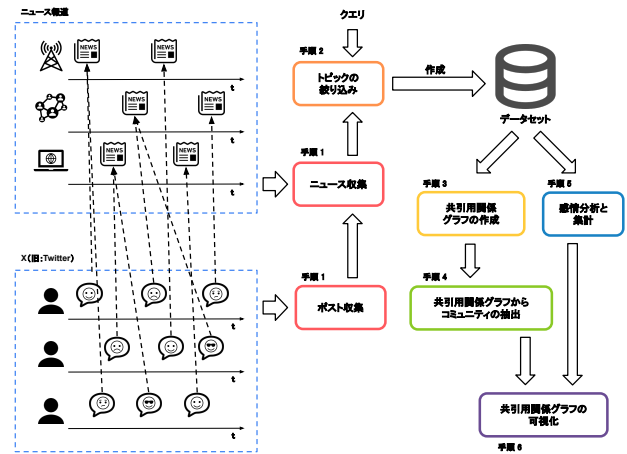


図1 分析手法の概要

3 提案手法

3.1 ソーシャルメディア空間の感情の分析

今まで感情分析器がソーシャルメディアの情報空間的な分析にほとんど用いられてこなかった理由は、大きく分けて2つある。まず、今まで用いられていた ML-Ask のような辞書ベースの感情分析器は、辞書に登録されている単語(感情語)しか判別できず、単独の単語ではなく文章として表される感情の検出もできないために、感情検出率が低い。さらに、各感情の検出は、辞書への感情語の登録状況に大きく左右される。例えば「給料が振り込まれた!」という文章は明らかに「喜び」の感情を含むが、「喜び」を表現する単語が文章中に存在しないため ML-Ask による感情分析結果は「無感情」となる。これらの問題から、全体的な傾向の把握程度にしか利用できず、情報空間の全体的な可視化は難しかった。

しかし、BERT や T5 のような Transformer ベースの言語モデルと感情ラベルが付与されたデータセットの登場により、現在は感情検出率が高い感情分析器が構築できるようになった。先ほどの例において、難波らの手法を用いると「喜び」が 3, 「極性」が 2 と検出される。

本論文では、あるトピックに関する各種メディアのニュース報道に対してソーシャルメディア上のユーザが持つ感情の情報空間における分布を可視化することで、全般的な感情の傾向や偏りを分析する手法を提案する。提案手法の手順を以下に、概要を図3.1に示す。図中の手順番号は以下の手順番号と一致する。

- 手順1: ニュース記事とポストの抽出
- 手順2: トピックの絞り込み
- 手順3: ユーザの共引用関係グラフの構築
- 手順4: 共引用関係グラフからのコミュニティ抽出
- 手順5: ポストの感情分析と集計
- 手順6: 共引用関係グラフの可視化

3.2 ニュースとポストの抽出

ある期間に投稿されたニュース記事とポスト(リポストを含

¹ : <https://github.com/ids-cv/wrime>

む)を収集し、Elasticsearch に登録する。

3.3 トピックの絞り込み

藤兼らの論争度 [17] を用いて、論争度上位 20 件のニュース記事を求める。その後、トピッククエリを手手で選定する。藤兼らの論争度は、論争を招きやすいニュース記事の発見に使用される指標で、あるニュースが言及された期間の逆順位と言及したユーザ数の逆順位の調和平均で求めることができる。論争度は 0 から 1 の値を取り、値が大きければ大きいほど論争を招きやすいことを表す。

続いて、選定したトピッククエリを用いて Elasticsearch を全文検索して関連するニュース記事集合を抽出し、次にそれらのニュース記事の URL を用いてニュース記事を引用したポスト集合を抽出する。

3.4 ユーザの共引用関係グラフの構築とコミュニティ抽出

ユーザの共引用関係の強さはユーザに引用された記事集合の類似度として扱えると仮定して、Jaccard 係数を用いて定量化し、閾値を越える場合にエッジを張ることでユーザの共引用関係グラフを構築する。ただし、1 回しか引用していないユーザやグラフの最大連結成分に含まれないユーザは除外する。この理由は、前者は興味本位で偶然ニュース記事を引用しただけである可能性が高く、後者は、一般的なユーザとは興味・関心のベクトルが異なるために孤立しているからである。

3.5 共引用関係グラフからのコミュニティ抽出

共引用関係グラフにおける類似するユーザ集合を把握するために、Blondel らが提案した Louvain 法 [18] を用いて、共引用関係グラフを複数のコミュニティに分割する。Louvain 法は Modularity に基づくグラフ分割手法の一つであり、局所最適化と同一コミュニティに属するノードの集約により、大規模なグラフ構造を高速に処理でき、高い Modularity が得られる。

3.6 ポストの感情分析と集計

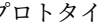
共引用関係グラフのノードであるユーザの引用ポストを感情分析する。ただし、引用ポストにはニュースカテゴリ、タイトル、メディア名、メディアのユーザ名、ハッシュタグなどのニュース記事やメディアに関する情報が同時に埋め込まれていることが多く、ユーザの発言だけを分析したい場合にはノイズとなる。そこで、中北らの手法 [19] を用いて、ユーザ自身が書いたコメント部分だけを抽出した。

さらに、ニュース記事に難波らの T5 を用いた感情分析器を適用し、Plutchik の基本 8 感情を求め、さらにそれを各ユーザごとに集計することで、ユーザがそのトピック全体に抱いた感情の割合を求める。

3.7 共引用関係グラフの可視化

得られた共引用関係グラフを可視化する。この可視化システムは Unity²を用いて実装した。Unity を採用した理由は、大量のオブジェクトのレンダリング処理能力が優れているから、イ

表 1 感情とカラーコード

感情	カラーコード	サンプル
怒り	#d94638	
期待	#eb8b2f	
喜び	#eec937	
信頼	#72af4c	
恐れ	#1a824e	
驚き	#049fd9	
悲しみ	#375fac	
嫌悪	#7e4699	
無感情	#7f7f7f	

ンタラクティブなシステムを簡単にプロトタイピング出来るから、20 種類以上のプラットフォームでプログラムを動作させることができるからである。

各ユーザはノードとして表す。ノードの位置は、Gephi³のレイアウトアルゴリズム ForceAtlas2 [20] を用いて決定する。これにより引用するニュース記事が類似するユーザ同士は近くに、類似しないユーザ同士は遠くなるように配置される。ノードの外枠は、共引用関係グラフにおける類似ユーザ集合を容易に把握できるように、Louvain 法で抽出されたコミュニティごとに異なる色で着色する。ノードの内側は、そのユーザが抱く Plutchik の基本 8 感情の割合を表 1 に示す色を用いて円グラフとして可視化する。これは Plutchik が論文で用いた配色と同じである。円グラフとして可視化することで、個々のユーザに焦点を当てると感情の割合が分かり、一方でコミュニティや全体を俯瞰する際には感情の色が視覚混合され、単一の色としてユーザの分布が把握できる。視覚混合とは、隣接する 2 色以上の色彩が、遠くから見ると混じり合ってひとつの色に見える光学現象であり、クロード・モネやジョルジュ・スーラをはじめとする印象主義の画家たちによって絵画に応用された。

実際に提案手法の可視化結果として、グラフ全体を俯瞰できるように縮小表示した例を図 2(a) に、また個々のユーザに注目できるように拡大表示した例を図 2(b) と図 2(c) に示す。図 2(a) の右下のノード群は、喜びの黄色と期待の橙色の割合が多く、明るい暖色系のノードとして見る事ができる。また、左上のノード群は、嫌悪の紫色と驚きの水色、悲しみの紺色の割合が多く、寒色系のノードとして見る事ができる。ここで、各ノード群に注目して拡大表示をすると、ユーザの感情の割合を個々に理解することができる。

4 分 析

4.1 データセット

ニュースに言及したポストデータセットとして、Twitter API で取得した 2022 年 1 月 1 日から 2022 年 12 月 31 日までの 1,489,667 件のニュース記事と、それらのニュース記事の URL が含まれる 3,466,472 件のポストとリポスト（旧：リツイート）を収集して分析に使用した。また、以降の章で感情分

2 : <https://unity.com/ja>

3 : <https://gephi.org/>

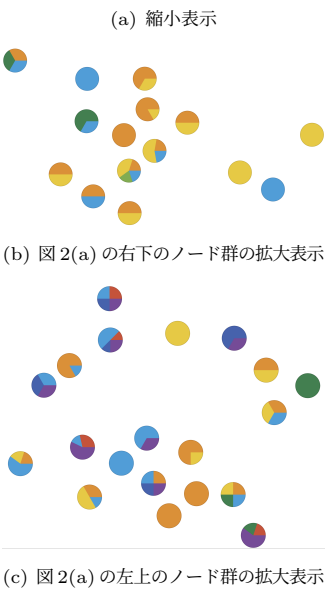
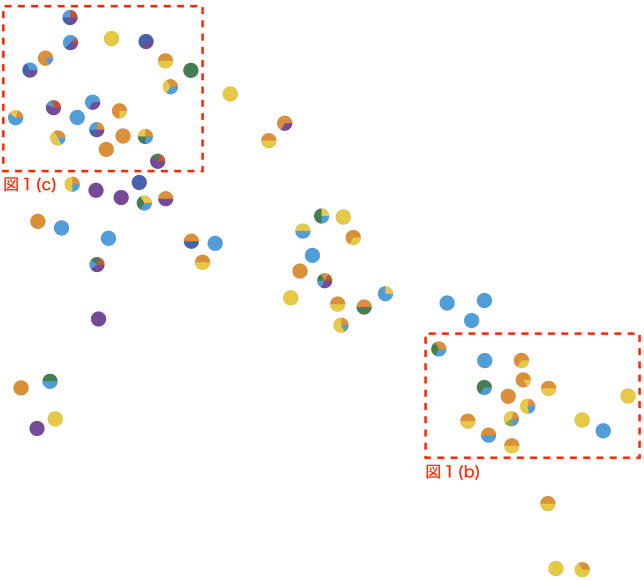


図 2 拡大・縮小表示の例

表 2 感情分析に用いたポスト数		
トピック番号	トピックを表すクエリ	ポスト数
1	太地町	232 件
2	コロナ予備費 使途	196 件
3	岸田首相 留学生	596 件
合計		1,024 件

析を行ったニューストピックのポスト数を表2に示す。

4.2 トピックの選定

論争度上位 20 件のニュース記事を表3に示す。これらの記事の中から 3 件の記事を選定し、そのトピックを表す次のクエリを人手で作成した。

- トピック 1: 太地町
- トピック 2: コロナ予備費 使途
- トピック 3: 岸田首相 留学生

トピック 1 は表3の 5 位、トピック 2 は 1 位、トピック 3 は 12 位のニュース記事に該当する。

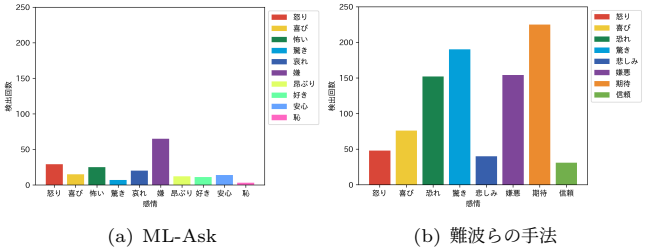


図 3 感情別の検出総数

4.3 ML-Ask と難波らの手法の比較

本論文で提案する可視化手法は、ポストからの感情検出率が高いことを前提にしている。難波らの手法の性能は論文 [14] で示されているので、本論文では辞書ベースの ML-Ask 感情検出機と機械学習ベースの難波らの手法で感情検出率がどの程度異なるかを評価した。

4.3.1 感情検出率の分析

全体的な感情検出率について分析した。ML-Ask と難波らの手法の感情検出率のトピック別と全体の結果を表 4 に示す。「感情あり」はいずれかの感情が検出されたポスト数を表すことに注意されたい。この結果を見ると、どの場合も ML-Ask よりも難波らの手法のほうが感情検出率が明らかに高いことがわかる。

4.3.2 感情別の検出総数の分析

今回選定した 3 つのトピックの全ポストから ML-Ask と難波らの手法で検出された感情の総数を比較した。感情別の検出総数を 図3 に示す。ただし、それぞれ異なる感情分類を用いており、一致する感情は怒り、喜び、恐れ（怖い）、驚き、悲しみ（哀れ）、嫌悪（嫌）であり、ML-Ask では昂り、好き、安心、恥、難波らの手法では期待と信頼は一致しないので、一致する感情を左側に配置していることに注意すること。また中北らの手法 [19] を適用後にテキストが残らなかったポストは対象外とした。

まず検出総数に関しては、難波らの手法の方がどの感情においても明らかに多かった。次に感情ごとのばらつきに関しては、ML-Ask は嫌だけが突出しているのに対して、難波らの手法は嫌悪以外に恐れ、驚き、期待も同程度に検出されており、特に嫌悪よりも、ポジティブな感情である期待や ML-Ask ではほとんど検出されていない驚きの方が多く検出されていることがわかる。これらの結果から、難波らの手法は単に感情検出率が高いだけではなく、より多くの感情をあまり偏らずに検出できていることがわかる。

4.3.3 ポストの感情割合の分析

ポストの感情割合を分析した。個々のポストの検出された感情の割合を、積み立て棒グラフを用いて、図4、図5、図6 に示す。横軸はポストを時系列に従って並べた時の投稿番号を表す。X はリアルタイム性が高いので、投稿順に並べることで、同一ニュースを引用しているポスト同士は近くに配置されると共に、時系列に伴う変化が判別しやすくなると期待できる。縦軸はポストの感情割合である。比較しやすいように、一致する

表 3 2022 年の論争度上位 15 件のニュースタイトル

	タイトル	論争度
1	コロナ予備費 12 兆円、使途 9 割迫えず 透明性課題	0.5000
2	日本会議系に統一教会系…安倍新内閣はまるで“カルト内閣”	0.1250
3	宮城県が水道運営権を 10 億円で売却 全国初	0.0454
4	国内死亡数が急増、1〜3 月 3.8 万人増 コロナ感染死の 4 倍	0.0425
5	「水族館の人気者」イルカが迎える恐ろしい結末 太地町のイルカ猟を毎日記録してわかったこと ―…	0.0416
6	高額医療費負担、財務省「廃止を」 省庁の無駄、予算執行調査	0.0400
7	政府、マイナンバー「全口座ひも付け」義務化検討 来年の法改正目指す	0.0392
8	旧統一教会側と自民党、改憲案が「一致」 緊急事態条項、家族条項…濃厚な関係が影響？	0.0392
9	ワクチン 2 回の陽性率、半数世代で未接種上回る 厚労省再集計で判明	0.0377
10	安倍元首相銃撃の山上容疑者の背後に 2 つの”反アベ団体”か 捜査当局が重大関心	0.0338
11	水道橋博士を訴訟恫喝し囂聲、松井市長が引き起こしたトンデモ事態！ 保健所逼迫で感染者入力を外…	0.0176
12	岸田首相「留学生 30 万人」見直し さらに増やす計画策定を指示	0.0175
13	SNS 上ひぼう中傷対策強化 侮辱罪に懲役刑導入 刑法改正案決定	0.0149
14	30 代半ば〜50 代半ばの世帯所得 20 年前と比べ 100 万円超減少	0.0136
15	田崎史郎、橋下徹、東国原英夫、古市憲寿、三浦瑠麗…安倍元首相と統一教会の“関係”をなかったこ…	0.0136

表 4 感情検出率

トピック番号	感情検出器	感情あり	感情検出率
1	ML-Ask	28 件	12.06 %
	難波らの手法	146 件	62.93 %
2	ML-Ask	32 件	16.32 %
	難波らの手法	101 件	51.53 %
3	ML-Ask	94 件	15.77 %
	難波らの手法	330 件	55.36 %
全体	ML-Ask	154 件	15.04 %
	難波らの手法	577 件	56.35 %

感情は下から同じ順で積み上げ、その上に一致しない感情を配置した。

この結果から、感情が検出されるポストの数は難波らの手法の方が明らかに多い。また検出された場合の感情の対応を見ると、ML-Ask と同じ感情が難波らの手法でも検出されているが、難波らの手法の方が検出される感情の種類が多い。これらが感情検出率が高くなる理由だと推測できる。

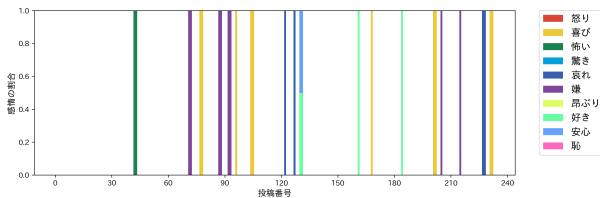
さらに時系列的な特徴に着目すると、難波らの手法は連続するポストで感情の種類や割合が類似する傾向が高いことから、棒グラフが太く表示される傾向がある。同一ニュースを引用しているポストが連続しているという仮定が正しいならば、異なるテキストのポストであっても感情が安定して検出されていると推測できる。また、ML-Ask では長い感情検出不能期間が存在するが、難波らの手法では短く、時系列的にも網羅的に検出できることがわかる。

4.4 可視化結果の分析

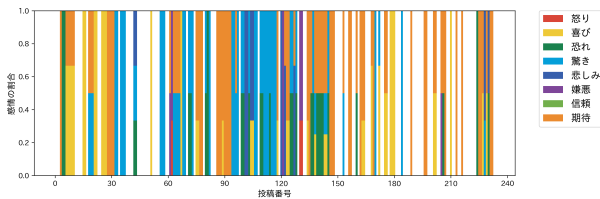
提案手法で3つのトピックをグラフ構造、コミュニティ分布、コミュニティ別の感情割合、感情分布の4種類の観点から可視化して、その結果を分析した。

4.4.1 トピック 1: 太地町

トピック 1 のグラフ構造を図 7(a) に示す。抽出されたノード数は 20、エッジ数は 105 であった。

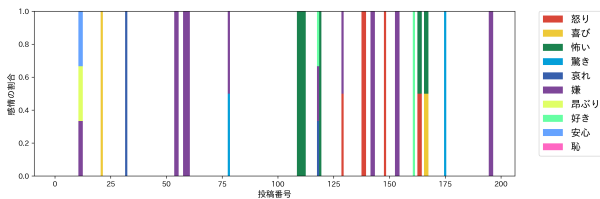


(a) ML-Ask

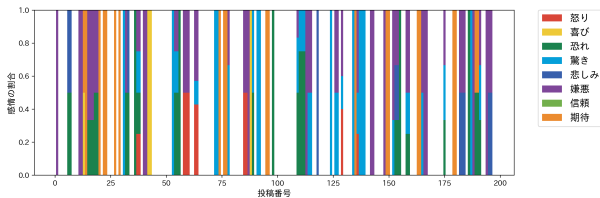


(b) 難波らの手法

図 4 トピック 1 の感情検出結果



(a) ML-Ask



(b) 難波らの手法

図 5 トピック 2 の感情検出結果

次にコミュニティ分布を図 7(b) に示す。なおコミュニティの違いが明確になるようにエッジは描画しない。抽出されたコミュニティ数は 2 であり、明確に分離していた。

次にコミュニティ別の感情の割合を図 7(c) に示す。横軸はコ

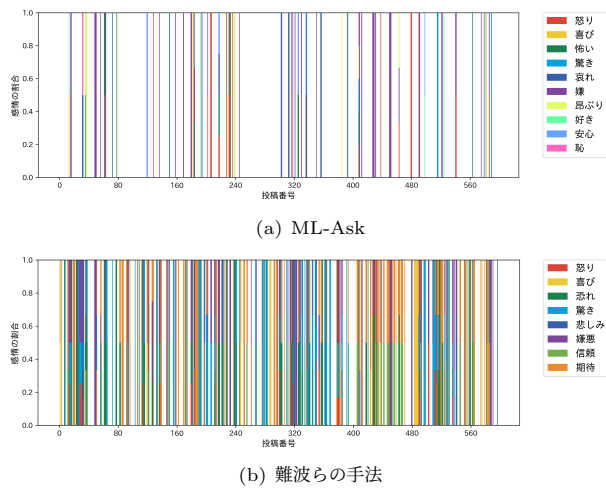


図 6 トピック 3 の感情検出結果

コミュニティ番号を、縦軸はクラスタ内の感情の割合を表す。なお無感情は、テキストは存在しても感情分析の結果が無感情であったことを意味する。このトピックは、主に捕鯨文化やその周辺問題についてのポストであった。コミュニティ0では、喜び・希望の感情が相対的に多く検出された。捕鯨の豊漁を願うポストや太地町立くじらの博物館で行われた鯨との触れ合いを通じた教育報告のポストがあり、捕鯨文化に関して肯定的な意見が多くみられた。コミュニティ1では、驚き・悲しみ・嫌悪の感情が相対的に多く検出された。クジラやイルカの生態販売を問題視するポストがあり、捕鯨文化への否定的な意見が多くみられた。また、コミュニティ1で検出された喜び・期待の感情は、捕鯨問題に言及するニュース記事が発表されたことに対する喜びや、状態が改善されることを期待するポストだった。

最後に感情分布を図7(d)に示す。左下のコミュニティ0は期待・喜びの感情を持っているために暖色系に、右上のコミュニティ1の右下は驚きの感情を持っているために寒色系に感じられ、コミュニティごと、またはコミュニティ内で類似した感情を持つノードが近くに配置されていることがわかる。

4.4.2 トピック 2: コロナ予備費 使途

トピック 2 のグラフ構造を図8(a)に示す。抽出されたノード数は37、エッジ数は290であった。

次にコミュニティ分布を図8(b)に示す。抽出されたコミュニティ数は3であり、2つの大きなコミュニティと2個のノードからなる小さなコミュニティであった。

次にコミュニティ別の感情の割合を図8(c)に示す。このトピックは、主に予備費の使途が不透明なことに対して疑問を投げかけるポストであった。コミュニティ0では、「怒り・嫌悪」の感情が相対的に多く検出され、予備費の使途を管理できなかった与党（自民党）への怒りや、飽きれに近い嫌悪を含むポストがあった。コミュニティ2では、「恐れ」の感情が相対的に多く検出され、予備費の使途を追跡できない、管理できない政府の対応への不透明感の恐れを含むポストがあった。

最後に感情分布を図8(d)に示す。コミュニティ0の右側に怒りの感情を、コミュニティ1の左側には驚きの感情を、コミュニティ2とその近辺には恐れ

の感情を持つノードが集中的に分布しており、グラフの部分領域ごとに色の傾向が異なるように感じられることがわかる。

4.4.3 トピック 3: 岸田首相 留学生

トピック 3 のグラフ構造を図9(a)に示す。抽出されたノード数は94、エッジ数は1,155であった。

次にコミュニティ分布を図9(b)に示す。抽出されたコミュニティ数は4であり、3つの大きなコミュニティと1つの小さなコミュニティであった。

次にコミュニティ別の感情の割合を図9(c)に示す。このトピックは、岸田首相が留学生への支援を増やすことを検討したことに対するポストであった。コミュニティ1と2は、怒りの感情が検出されると共に、嫌悪の感情の割合が他よりも多かった。これは、自国民の学生を優先するべきといった、自民党の対応についての嫌悪感であった。また、コミュニティ0はコロナ水際対策の緩和によるニュースに多く言及されており、期待・恐れが相対的に多く検出された。期待の感情は、コロナ水際対策の緩和による留学生の増加への期待であった。恐れ

の感情は、コロナ水際対策の緩和による海外からのコロナ感染者の流入への恐れであった。最後に感情分布を図9(d)に示す。右側のコミュニティ0では、恐れ

5 おわりに

本論文では、ニュース記事の共引用関係に基づいて注目するメディアやニュース記事が類似するユーザを近くに配置し、難波らの T5 感情検出器で抽出した感情の割合を円グラフ表示することで、あるトピックに関する各種メディアのニュース報道に対してソーシャルメディア上のユーザが持つ感情の分布や偏りを可視化する手法を提案した。

まず難波らの T5 感情検出器を辞書ベースの ML-Ask と比較し、提案手法で前提とする高い感情検出率が得られることを示した。さらに実際に3つのトピックのニュース記事群に対するポストデータを用いて、Louvain 法で抽出されたコミュニティ間の違いやコミュニティ内の局所的な偏り、さらには複数のコミュニティにまたがる類似を判別できることを示した。

今回は感情の分布や偏りを直感的に把握できるように可視化しただけだが、今後の課題として感情の偏りの定量化指標と、感情的な類似性を考慮した新しいコミュニティ抽出法を検討する予定である。

謝 辞

本研究は JSPS 科研費 21H03557 の助成を受けた。

文 献

- [1] Roja Bandari, Sitaram Asur, and Bernardo Huberman. The

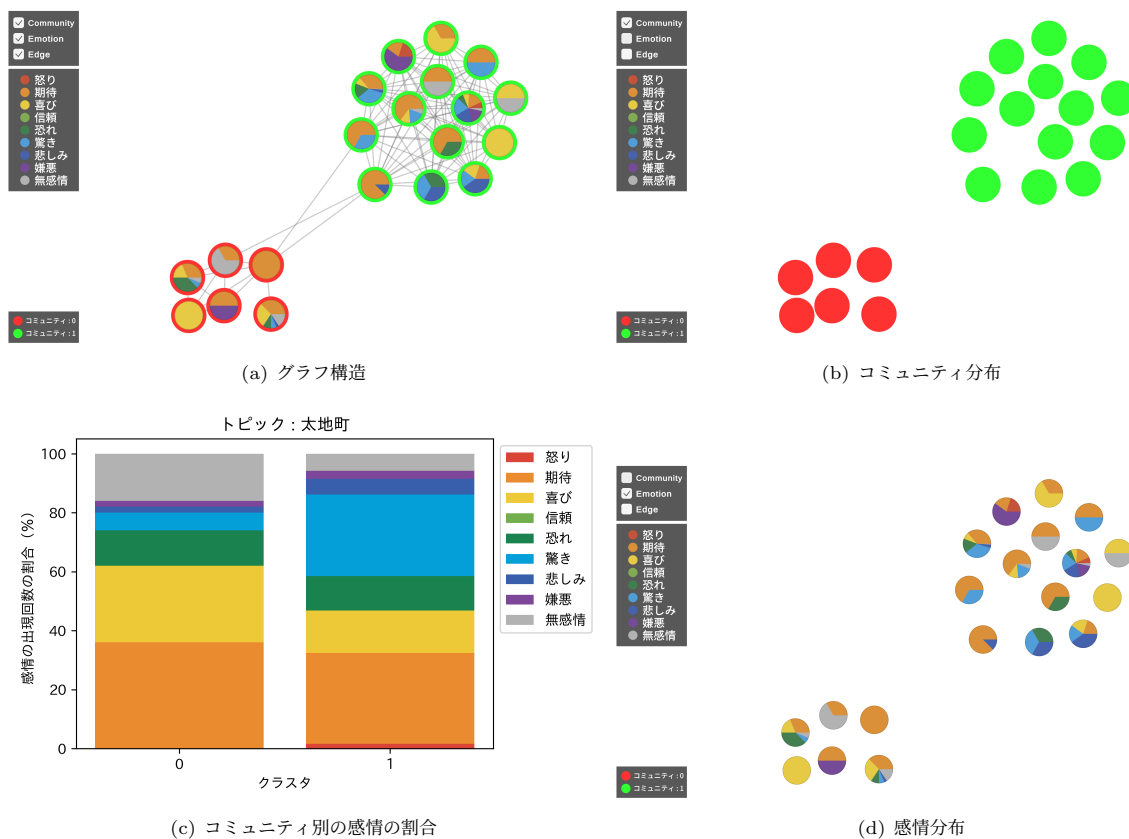


図7 トピック1の可視化結果

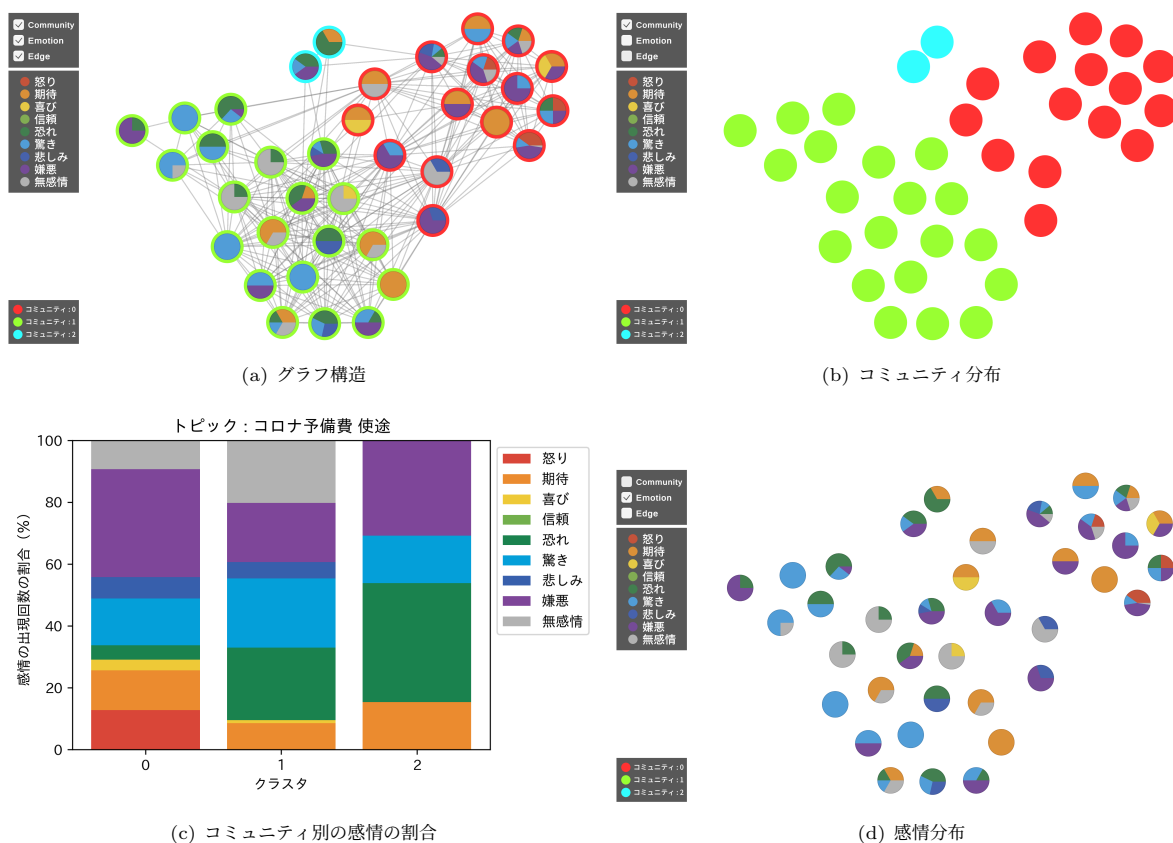


図8 トピック2の可視化結果

pulse of news in social media: Forecasting popularity. In *Proceedings of the International AAAI Conference on Web*

and Social Media, Vol. 6, pp. 26–33, 2012.

[2] Eli Pariser. *The Filter Bubble: How the New Personal-*

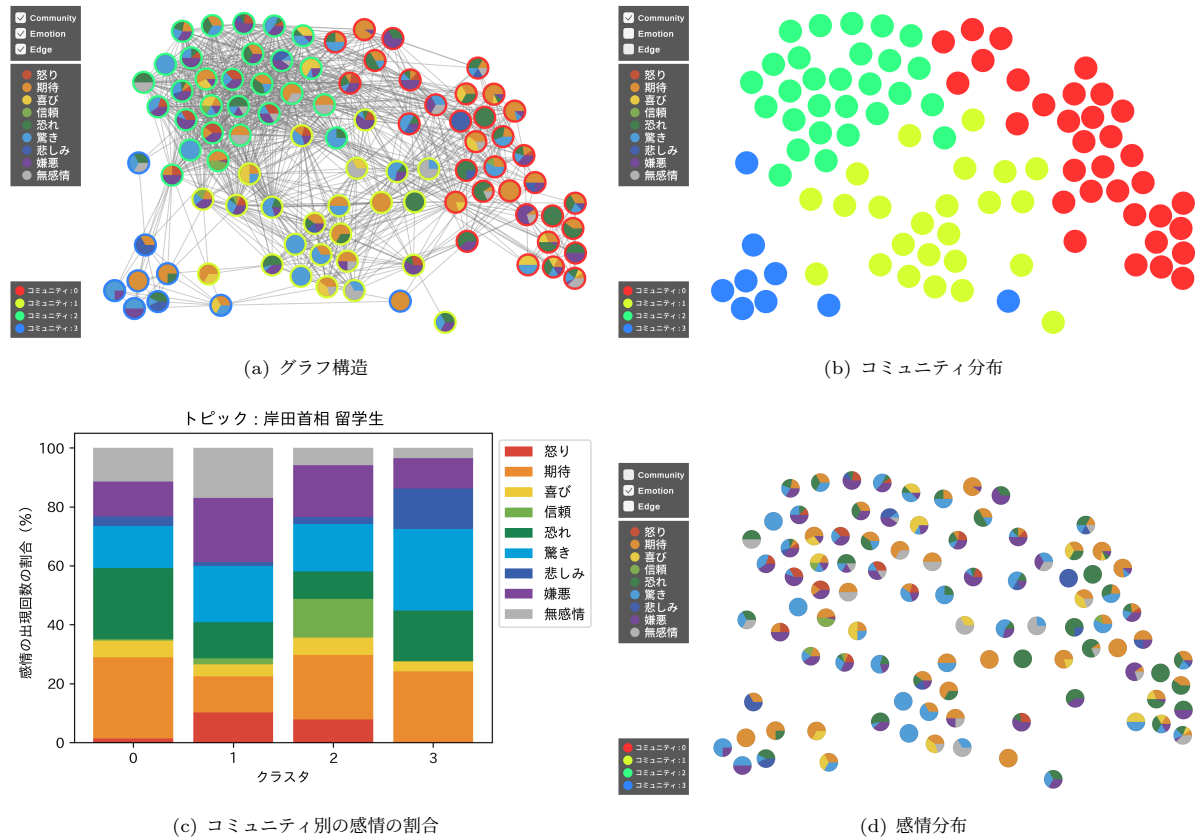


図9 トピック3の可視化結果

- ized Web Is Changing What We Read and How We Think. Penguin Books, 2011.
- [3] 笹原和俊. フェイクニュースを科学する: 拡散するデマ、陰謀論、プロパガンダのしくみ. 株式会社化学同人, 2019.
- [4] ハーヨ・デヒティング. ジョルジュ・スーラ: 1859-1891: 点に要約された絵画. タツシユンジャパン, 2000.
- [5] Michal Ptaszynski, Pawel Dybala, Wenhan Shi, Rafal Rzepka, and Kenji Araki. A System for Affect Analysis of Utterances in Japanese Supported with Web Mining. 知能と情報, Vol. 21, No. 2, pp. 194-213, 2009.
- [6] Michal Ptaszynski, Pawel Dybala, Rafal Rzepka, Kenji Araki, and Fumito Masui. ML-Ask: Open source affect analysis software for textual input in Japanese. *Journal of Open Research Software*, Vol. 5, No. 1, p. 16, 2017.
- [7] 中村明. 感情表現辞典. 東京堂出版, 1993.
- [8] Livia Polanyi and Annie Zaenen. Contextual valence shifters. In *Computing attitude and affect in text: theory and applications*, pp. 1-10. Springer, 2006.
- [9] 小林のぞみ, 乾健太郎, 松本裕治, 立石健二, 福島俊一. 意見抽出のための評価表現の収集. 自然言語処理, Vol. 12, No. 3, pp. 203-222, 2005.
- [10] 東山昌彦, 乾健太郎, 松本裕治. 述語の選択選好性に着目した名詞評価極性の獲得. 言語処理学会第14回年次大会論文集, pp. 584-587, 2008.
- [11] 福田悟志, 難波英嗣, 庄司裕子. コロナ禍におけるワクチンに対する人々の感情変化とその要因の分析. 知能と情報, Vol. 34, No. 3, pp. 592-600, 2022.
- [12] Tomoyuki Kajiwar, Chenhui Chu, Noriko Takemura, Yuta Nakashima, and Hajime Nagahara. WRIME: A new dataset for emotional intensity estimation with subjective and objective annotations. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL 2021)*, pp. 2095-2104, 2021.
- [13] Robert Plutchik. The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American Scientist*, Vol. 89, No. 4, pp. 344-350, 2001.
- [14] Hidetsugu Nanba, Katsuya Yamamoto, Satoshi Fukuda, Hiroko Shoji, Masayoshi Tanishita, Yasushi Kyutoku, and Mitsuru Yamashina. Modeling the social acceptability of technologies using Twitter data. In *Proceedings of the 2023 IEEE Conference on Systems*, 2023.
- [15] 島海不二夫, 榎剛史, 吉田光男. ソーシャルメディアを用いた新型コロナウイルス禍における感情変化の分析. 人工知能学会論文誌, Vol. 35, No. 4, pp. F-K45.1-7, 2020.
- [16] 笹原和俊, 奥田慎平, 五十嵐祐. テキストマイニングによるコロナ禍の消費者心理・行動の定量化. 第35回人工知能学会全国大会 (JSAI2021), 1D3-OS-3b-04, 2021.
- [17] 藤兼由生, 風間一洋, 吉田光男, 土方嘉徳. ツイートからの論争を招きやすいニュース記事の発見手法. 第35回人工知能学会全国大会 (JSAI2021), 1D3-OS-3b-02, 2021.
- [18] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, Vol. 2008, No. 10, p. P10008, October 2008.
- [19] 中北雄大, 風間一洋, 吉田光男, 土方嘉徳. 感情とトピックに注目したメディアの報道姿勢の分析. 第14回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2022), B24-2, 2022.
- [20] Mathieu Jacomy, Tommaso Venturini, Sebastien Heymann, and Mathieu Bastian. ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PLoS One* 9(6): e98679, 2014.