

ユーザの興味とコメント分布によるニコニコ動画の分析

Konishi, Atsuro / 小西, 敦郎

(出版者 / Publisher)

法政大学大学院情報科学研究科

(雑誌名 / Journal or Publication Title)

法政大学大学院紀要. 情報科学研究科編

(巻 / Volume)

16

(開始ページ / Start Page)

1

(終了ページ / End Page)

6

(発行年 / Year)

2021-03-24

(URL)

<https://doi.org/10.15002/00023867>

ユーザの興味とコメント分布によるニコニコ動画の分析

Analyzing Nico Nico Douga Videos by Using Users' Interests and the Distribution of Comments

小西 敦郎

Atsuro Konishi

法政大学情報科学研究科情報科学専攻

E-mail: atsuro.konishi.6b@stu.hosei.ac.jp

Abstract

Nico Nico Douga (NND) is one of the most famous social media platforms for sharing videos in Japan. There are 18 million videos posted to NND, for which users search by using keyword search and related video recommendation. However, it is difficult for the users to find interesting videos because videos generally are associated with only short texts and a few tags. In this paper, we present a method for analyzing videos in NND by using the distribution of time-synchronized comments. Our method regards a video as the set of the users' comments on it and enables the clustering of videos based on the users' shared interests. In our experiment, we applied the proposed method to videos posted to NND, and evaluated our method by quantitatively comparing it with existing text-based methods and by qualitatively performing the subjective evaluation of clustering results. In the result of the quantitative evaluation, the proposed method showed a higher score of normalized mutual information than the existing methods when category metadata were used as correct results. The experimental results of the qualitative evaluation showed that the proposed method was as good as or better than the existing text-based and image-based methods.

1. はじめに

近年、ユーザが動画を投稿するソーシャルメディア (YouTube, bilibili, ニコニコ動画等) の利用が盛んになっており、多数の動画が投稿されている。ニコニコ動画は日本で有名な動画共有プラットフォームの1つで、ニコニコ動画に投稿された動画はゲームやアニメ、音楽に関連するものを中心に約 1.8 千万件である。このような動画共有プラットフォームでの動画の探索にはキーワードによる検索と関連動画機能が用いられるが、これらでは不十分であると考えられる。キーワードによる検索では、動画のタイトルや説明文、タグにキーワードが含まれる動画を検索することができる。通常動画には少ない量の文字情報しか付与されず、またこれらは主に動画の投稿者が付与するものであるため、動画の内容に関係のない情報を加えることもできる。また、キーワードにより動画

の内容に着目して検索することは、動画の内容を明確に示す単語が存在し、それを動画投稿者とユーザ両方が認識している必要がある。関連動画機能では、選択された動画との関係性をもとに関連する動画を提示する。ニコニコ動画では1つの動画に対し29の動画が関連動画として提示される。関連動画機能のアルゴリズムは公開されていないが、同じカテゴリーの動画や文字情報が類似する動画が関連動画として提示される場合が多い。そのため、キーワード検索の結果と類似した動画や、関係のない動画が提示されることがある。

本論文では、ニコニコ動画のコメント機能の特徴に着目し、ユーザの共通する興味をもとにした分析手法を提案する。提案手法は動画のコメント分布をもとにユーザが興味を持った動画を推定しクラスタリングを行う。コメント分布はニコニコ動画の機能である Time-Synchronized Comment (TSC) を用いて取得する。TSC はリアルタイムで動画の内容に関連付ける独特なユーザインタラクションレビューであり、最大の特徴は vpos データを持つことである。vpos はユーザがコメントをした際に見ていた動画上の時間的位置を表す。図 1 のように、あるユーザが動画を視聴すると、すでに投稿されたコメントの内容が vpos に対応する動画上の同じ画面に表示される。



図 1 TSC を表示したニコニコ動画の動画再生画面

実験において、提案手法をニコニコ動画に投稿された動画に対して適用し、定量的評価と定性的評価を行った。定量的評価では、正解となるラベルデータとして既存のテキストベースや画像ベースのクラスタリング手法の結果を用い、正規化相互情報量 (NMI) を用いて評価を行

った。定性的評価では、被験者が1つの動画が2つのクラスターのどちらに所属するかを回答するタスクを繰り返す主観実験を行い、結果をもとに精度を評価した。定量的評価の結果より、提案手法はユーザベースの既存手法に対して高いNMIの値を示した。定性的評価より、提案手法は既存手法と同等かそれ以上の精度を示した。

2. 関連研究

2.1. 動画の分析と推薦

動画の分析や推薦を行う研究は数多くある。Zhouら[1]は複数のユーザが共有するコミュニティへの動画推薦を提案した。この手法ではクリックされた動画を文書として扱っている。動画は時間的に連続したキーフレーム上に構築されたキューボイドシグネチャの組と、それにコメントしているユーザの集合として記述された社会的つながりを表現した。彼らは動画間のシグネチャの社会的つながりを計算し、動画を推薦した。Daniel[2]はYouTubeのトランスクリプトドキュメントの集合からトピックを発見する手法を提案した。トランスクリプトドキュメントは動画の字幕を表すテキスト情報であり、トピックはLDAを用いて抽出された。Royら[3]は協調フィルタリングの問題はほとんど閲覧されていないアイテムを推薦することが難しい点にあると考えた。この問題を軽減するため、ユーザとアイテム間の感情のつながりをもとにモデル化した潜在因子を学習するvisual-CliMFを提案した。

2.2. TSC

TSCに関する研究は主に日本や中国の動画共有プラットフォームを対象に行われている。これはTSCを採用した有名な動画共有プラットフォームが日本のニコニコ動画や、中国のbilibiliやAcfunであることによる。Renら[4]はbilibiliに投稿された動画を、TSCを利用して分類している。彼らは多層ニューラルネットワークを組み合わせた深層ニューラルネットワークを用いている。この手法は各動画のTSCと各ユーザのTSCに着目し、動画の特徴とユーザの特徴、コメントの時間をもとに学習を行う。Tsukudaら[5]はTSCの感情を動的に検出し、サポートベクター回帰で学習するSmartVideoRankingを提案した。このシステムは正規化されたコメントから14の特徴を計算し、ランク付けを行う。Yangら[6]はレビューベースの推薦手法を提案した。彼らはTSCの、後にされるコメントが既にされているコメントに影響を受けるという文脈依存性(Herding Effect)の現象を活用して、画像とテキストを混合させたモデルを設計した。またコメントと動画上の画像を結び付けることで、TSCと画像を対応付けた。

3. TSCにおけるコメント分布

この章では、実験で用いたデータセット内(表1)のTSCにおけるコメントの分布がどのようなものか説明する。一般的にコメント分布は一様でなく、動画ごとに特徴がある。我々はコメント分布により動画の内容に反応しているユーザの認識ができると仮定する。図2は表1

に含まれる動画のコメント分布である。横軸は動画時間を10分割したものであり、縦軸は各動画での全体のコメント数に対する当該時間帯のコメント数の割合の平均である。図2より、コメントは動画の初めと終わりに集中しており、その他の時間では時間とともに減少していく。動画の初めと終わりにされているコメントの多くは挨拶のような、動画の内容に関係しないコメントである。この特徴は特にコメント数の少ない動画で顕著である。

表1 データセット

	Dataset5	Dataset50	Dataset100
投稿日時	2018/9/1~30		
コメントしたユーザの最低数	5	50	100
動画数	45274	9470	4628
コメント数	6134029	4761992	3477685
コメントの投稿日時	2018/8/23~2020/1/25	2018/9/1~2020/1/25	
コメントしたユーザ数	1945830	1467446	1198719

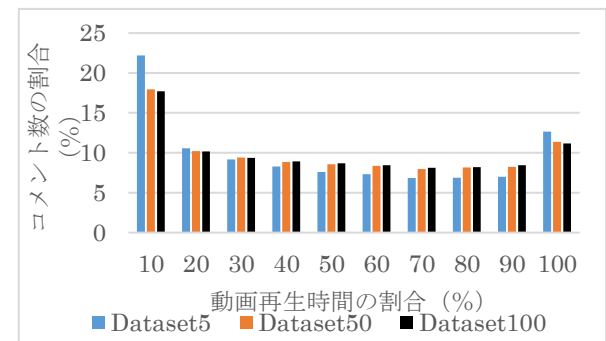


図2 各データセットにおけるすべての動画のTSCのコメント分布

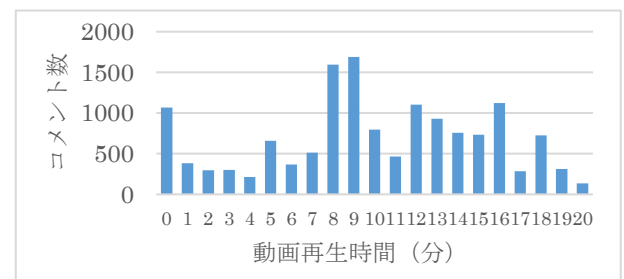


図3 特定の動画での特徴的なコメント分布

一方で個々の動画には、平均的なコメント分布とは異なる特徴的なコメント分布が存在する。図3の横軸は特定の動画について時間を1分区分けとしたものであり、縦軸はコメント数である。この動画は図2のように、多くのコメントが動画の初めと終わり集中しており、これらの中には挨拶に対応するコメントがそれぞれ36.7%と21.7%を占めている。しかし、動画の盛り上がり動画

の中盤にあるため、図 2 とは異なるコメント分布となっている。本研究では、このようなコメント分布の特徴の違いや共通点をもとに、動画の内容により強い興味を持つユーザを識別する。

4. 提案手法

本論文では、似たようなユーザがコメントした動画は共通する内容であるという仮定をもとに、コメント分布の特徴を用いた動画のクラスタリング手法を提案する。この手法は Uchida ら [7] の手法を拡張したものである。

本手法は以下の 3 つのステップからなる。

- 全ての動画の組に対して、2 つの動画両方に対してコメントとしたユーザの重複度を用いて動画間の類似度を計算する。
- 計算した類似度をもとに、重み付き無向グラフを構築する。
- Louvain 法を用いて重み付き無向グラフをクラスタリングする。

Louvain 法はクラスタ間のつながりの程度を表すモジュラリティを用いたクラスタリング手法で、分割したグラフのモジュラリティの最大化を目的とする。また、大規模なネットワークを高速で分割することができる。

ステップ 1 で、動画間の類似度を計算する必要がある。そのため、動画 i に対するユーザ k の値 $r_{i,k}$ を定義する。 B_k はデータセットにおいてユーザ k がしたすべてのコメントを表す多重集合である。

$$B_k = \bigcup_i \bigcup_n c_{i,k,n}$$

$c_{i,k,n}$ は動画 i に対するユーザ k の n 番目のコメントの v pos である。 $r_{i,k}$ は動画 i に対してユーザ k がしたすべてのコメントに重み関数 $w(c_{i,k,n}, t)$ を適用した値の和であり、 B_k により正規化される。

$$r_{i,k} = \frac{1}{|B_k|} \sum_n w(c_{i,k,n}, t)$$

$w(c_{i,k,n}, t)$ は $c_{i,k,n}$ の値とあらかじめ設定した時間幅 t により決定する重み関数である。 \mathbf{r}_i は動画 i に対応する行ベクトルであり、 U はデータセット内のユーザの総数である。

$$\mathbf{r}_i = (r_{i,0}, r_{i,1}, \dots, r_{i,U})$$

この行ベクトルを用いて動画 i と動画 j の類似度を、コサイン類似度をもとに計算する。

$$\text{sim}(\mathbf{r}_i, \mathbf{r}_j) = \frac{|\mathbf{r}_i \cdot \mathbf{r}_j|}{|\mathbf{r}_i| |\mathbf{r}_j|}$$

ここで、重み関数 $w(c_{i,k,n}, t)$ を以下の仮定のもとに定義する。動画におけるコメント分布は一様ではなく、動画の初めと終わりにコメントが集中する。それらの時間帯に投稿されたコメントの多くは挨拶など定型的なコメントである。そのようなコメントは動画の内容に言及するものでないため、その時間帯に投稿されたコメントは重要でない。また、動画の初めと終わり以外でコメント数が急激に増加している時間帯は動画の盛り上がりである

と判断する。その時間帯にコメントしたユーザは動画の内容に対して強い興味を持っており、そのコメントは重要であると考える。

これらの仮定より、重み関数 $w(c_{i,k,n}, t)$ を以下のように定義する。 $A_{c_{i,k,n}, t}$ は動画上の時間 $c_{i,k,n}$ を中心に幅 t の間に投稿されたコメントを表す多重集合である。

$$A_{c_{i,k,n}, t} = \{c_{i,k',n'} \mid c_{i,k,n} - \frac{t}{2} \leq c_{i,k',n'} \leq c_{i,k,n} + \frac{t}{2}\}$$

$w(c_{i,k,n}, t)$ の値は $c_{i,k,n}$ によって決定される。

$$w(c_{i,k,n}, t) = \begin{cases} 0.5 & \left(\text{if } c_{i,k,n} \leq \frac{3}{2}t \text{ or } v_i - \frac{3}{2}t \leq c_{i,k,n} \right) \\ \frac{2|A_{c_{i,k,n}, t}|}{|A_{c_{i,k,n}-t, t}| + |A_{c_{i,k,n}+t, t}|} & (\text{otherwise}) \end{cases}$$

v_i は動画 i の再生時間であり、 $w(c_{i,k,n}, t)$ は $c_{i,k,n}$ が動画の初めと終わりの $\frac{3}{2}t$ 以内であれば値を 0.5 とする。それ以外の時間であれば、動画上の時間 $c_{i,k,n}$ を中心に幅 t の間に投稿されたコメント数と、その時間帯の外側のそれぞれ幅 t の間に投稿されたコメント数の平均の比を値とする。本論文の実験では、時間幅 t の値として 10 と 20 を用いた。

5. 実験

5.1. データセット

本研究では、国立情報学研究所の IDR データセット提供サービスにより株式会社ドワンゴから提供を受けた「ニコニコ動画コメント等データ」を利用した [8]。また、ニコニコ動画 API を用いて追加のコメントデータを取得した。「ニコニコ動画コメント等データ」は動画とコメントのメタデータで構成されている。実験において、この動画のカテゴリデータ (37 種類) を正解のラベルとして用いている。また、ニコニコ動画 API を用いてコメントのユーザ ID や動画のサムネイルのデータを収集した。実験で対象とした動画の条件は表 1 であり、コメントしたユーザの最低数をもとに 3 つのデータセットを用いた。

5.2. 準備

提案手法の評価を行うため、NMI による定量的評価と主観実験による定性的評価を行った。定量的評価は提案手法の妥当性を評価するために行う。データセットには正解となるラベルデータが存在しないため、3 つの既存手法の結果 (LDA, GSDMM [9], IIC [10]) と動画のカテゴリメタデータを正解のラベルとして用いた。主観実験は提案手法が、ユーザの知識や背景をもとに動画分類することの利点を評価するために行う。以下は実験に用いた既存手法の説明である。

- Uchida らの手法は提案手法のもととなった手法である。本論文では、Uchida らの手法でのツイートにリツイートしたユーザが提案手法での動画にコメントしたユーザと対応している。本実験でも行ベクトル \mathbf{r}_i は $r_{i,k}$ により定義され、 $r_{i,k}$ はコメントしたか否かで 0 か 1 の値をとる。類似度はシンブソン係数により計算される。

- Latent Dirichlet Allocation (LDA)は1つの文書は複数のトピックからなると仮定した言語モデルであり、文書の集合に存在する隠れたトピックをもとに文書の分類を行う。LDAは確率分布を推定し、文書がどのトピックに所属するかを計算する。本実験では、動画のタイトルのテキストを文書とした方法(LDA-T)と動画のタイトルと動画説明文のテキストを文書とした方法(LDA-TD)の2種類を用い、各文書における単語の値はtf-idfを用いて計算した。
- A collapsed Gibbs Sampling algorithm for the Dirichlet Multinomial Mixture model (GSDMM)は短いテキストのクラスタリング手法であり、初めに分割するクラスタ数を大きく設定すると自動的に適切なクラスタ数に分割する。実験では、動画のタイトルを文書として扱った。
- Invariant Information Clustering (IIC)は相互情報量を最大化することで学習を行う、深層学習を用いた画像クラスタリング手法である。IICでは元の画像と、その画像に加工を加えた画像をペアの入力とし、出力の相互情報量をもとに学習する。そのため、ラベルのないデータに対して学習を行うことができる。本実験では、Dataset100の動画のサムネイルを入力としてIICを適用した。

LDAとGSDMMでは各文書に含まれる単語が必要となる。そのため、日本語形態素解析器であるMeCabを用いて文書の単語への分割を行った。LDAはあらかじめ分割するクラスタ数を決める必要がある。各データセットにおいてLDAは動画をおよそ100クラスタに分割し、それ以上の値を設定しても分割するクラスタ数が頭打ちとなった。そのため、LDAでの事前に定めたクラスタ数は100とした。LDAにおいて1つの動画が複数のクラスタへ同じ確率で所属する場合は、その動画を要素数1の新しいクラスタとみなした。

5.3. NMIによる定量的評価

提案手法の妥当性を評価するため、NMIによる評価を行った。NMIは2つのクラスタ間の相互情報量を0から1の値に正規化したものであり、1が完璧な関係を示す。NMIの値はクラスタの質を評価するために用いられ、クラスタ数の異なる2つの結果を評価することができる。2つのクラスタリング結果 λ_1 と λ_2 が与えられた場合、NMIの値は以下の式で表される。Nはデータセット内の要素数を、 n_X と n_Y はそれぞれ λ_1 のクラスタXと λ_2 のクラスタYの要素数を表す。 $n_{X,Y}$ は λ_1 のクラスタXと λ_2 のクラスタY両方に属する要素数である。

$$NMI(\lambda_1, \lambda_2) = \frac{\sum_{X \in \lambda_1} \sum_{Y \in \lambda_2} n_{X,Y} \log \frac{N n_{X,Y}}{n_X n_Y}}{\sqrt{\left(\sum_{X \in \lambda_1} n_X \log \frac{n_X}{N} \right) \left(\sum_{Y \in \lambda_2} n_Y \log \frac{n_Y}{N} \right)}}$$

表2、表3、表4はそれぞれDataset5、Dataset50、Dataset100のNMIを示している。実験では正しいラベルとする既存手法としてLDA、GSDMM、IICを選択し、それらのクラスタリング結果とカテゴリーメタデータを正解のラベルとして扱った。

表2 Dataset5のNMI

		Uchidaらの手法	提案手法 $t = 10$	提案手法 $t = 20$
	ラベル数	20779	21110	21120
LDA-T	100	0.520	0.554	0.555
LDA-TD	99	0.384	0.410	0.411
GSDMM	5089	0.788	0.831	0.832
Category	37	0.435	0.453	0.453

表3 Dataset50のNMI

		Uchidaらの手法	提案手法 $t = 10$	提案手法 $t = 20$
	ラベル数	2259	2268	2265
LDA-T	99	0.460	0.560	0.561
LDA-TD	62	0.344	0.411	0.414
GSDMM	1925	0.697	0.802	0.802
Category	37	0.388	0.437	0.438

表4 Dataset100のNMI

		Uchidaらの手法	提案手法 $t = 10$	提案手法 $t = 20$
	ラベル数	803	943	948
LDA-T	89	0.479	0.580	0.583
LDA-TD	55	0.356	0.432	0.432
GSDMM	414	0.659	0.773	0.775
IIC	191	0.322	0.422	0.427
Category	37	0.373	0.429	0.433

表2、表3、表4より、提案手法はUchidaらの手法に対してより高いNMIの値を示した。これにより、提案手法がUchidaらの手法よりも既存手法やカテゴリーメタデータをカバーする効果が高いことを示している。各結果において時間幅 t の値を変えても、提案手法で得られたNMIの値に大きな差は見られなかった。これはデータが疎であるためと考える。より多くのユーザがコメントした動画を含むデータセットを作成することで、より明確な違いが出てくると考えられる。

表2において、提案手法とUchidaらの手法では、既存の手法に比べてクラスタ数が非常に大きくなっている。これはコメント数に対してコメントしたユーザが多いためデータが疎となり、要素数の少ない動画が増えたことにあると考えられる。実際にDataset5では17053個の動画が要素数1のクラスタとなっている。これらの動画はデータセット内の他の動画と類似度が0であり、既存の手法と比較してクラスタ数に差が生じている。

5.4. 主観実験による定性的評価

Uchidaらの実験[7]をもとに、主観実験による定性的評価を行った。実験の概要は図4である。実験では、クラスタリング結果の適合率を以下のタスクにより評価した。はじめに、クラスタリング結果からクラスタXとクラスタYをランダムに選択する。その後、クラスタXより6つの動画 x_1 から x_6 を、クラスタYより5つの動画 y_1 から y_5 をそれぞれランダムに抽出する。 x_6 をターゲットの動画と

し、被験者は「 x_6 が x_1 から x_5 の動画群と、 y_1 から y_5 の動画群のどちらと同じクラスタに属するか」という質問に回答する。このタスクを繰り返し、正解した割合を再現率として扱う。

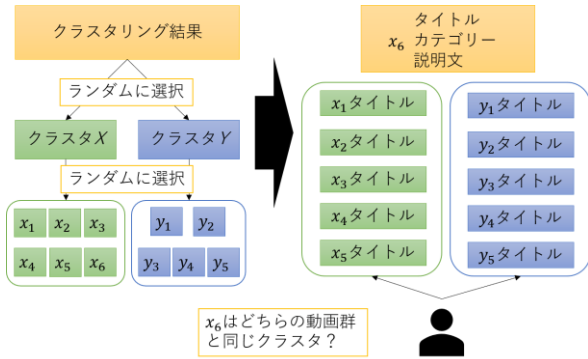


図4 主観実験の概要

この実験において、クラスタXとクラスタYはクラスタに属する動画の数とデータセットの動画数の比による確率でそれぞれ選択される。このタスクにおいて与えられる情報は動画のテキスト情報であり、これらの情報は被験者が素早く回答できるために付与される。被験者として、著者が Dataset5 と Dataset50 に関しては 6 つの手法 (LDA-T, LDA-TD, GSDMM, Uchida らの手法, 提案手法 $t = 10$, 提案手法 $t = 20$) のそれぞれに対して 500 タスクを行い、Dataset100 に関しては IIC を加えた 7 つの手法に対してそれぞれ 500 タスクを行った。

図 5 は各データセットの結果である。実験のタスクは動画のテキスト情報をもとに行われるため、テキストベースの手法の結果が良くなる傾向にある。その条件において、提案手法がテキストベースの手法と同程度かそれ以上の結果となったことは提案手法の良さを表している。提案手法はクラスタリングにテキスト情報を用いていないが、ユーザの共通する興味によってテキスト情報の似た動画を同じクラスタとして分割することができている。この実験では、提案手法における時間幅 t の値を変えたことによる結果の大きな違いは生じなかった。これも前述したデータが疎となることが関係すると考える。

5.5. 各クラスタの評価

提案手法のクラスタリング結果の詳細について説明する。表 5 は Dataset5 に対して提案手法 $t = 10$ を適用した結果の内、要素数が上位 9 クラスタの動画内容である。

表 5 Dataset5 での提案手法 $t = 10$ の上位 9 クラスタの動画内容 (手動で集計)

クラスタ No.	動画数	主な動画の内容
1	590	Fortnite (43.7%) Splatoon (6.1%)
2	283	新作ゲーム (50.1%) Fate/Grand Order (9.5%) 新作アニメ (7.4%)
3	202	Azurlane (26.3%) VOICELOID 実況プレイ (10.9%)
4	195	声優(79.5%)
5	150	車/バイクの車載動画(87.3%)
6	148	KPOP (58.9%) 歌ってみた (7.4%)
7	145	仮面ライダー (19.3%) Dead by Daylight (14.5%)
8	143	ガンダム (70.6%)
9	138	Splatoon (34.1%) KPOP (15.9%)

表 5 のクラスタ No. 2 は、動画の内容に関してユーザの共通する興味をもとにしなければ分割できない、提案手法に特有の例である。このクラスタでは、半分以上の動画が新作のゲームやアニメに関するものである。ゲームに関する動画はゲームのイベントで公開されたプロモーションビデオや制作会社が直接公開したものであり、その映像を視聴する人のリアクションが付随されているものもある。アニメに関する動画はすでに放送されているアニメのコマーシャルや、放送予定のアニメのビデオクリップなどである。また、同クラスタの Fate/Grand Order はゲームのタイトルであり、ゲームの新しいコンテンツについての動画である。これらの動画にはテキストや画像の類似はあまりないが、ユーザのゲームやアニメの新しいコンテンツに対する興味を反映し、1 つのクラスタとして分割することができている。

クラスタ No. 6 と No. 9 には韓国のポピュラーミュージック (KPOP) に関する動画が含まれるがクラスタの大部分は異なっている。クラスタ No. 6 は「歌ってみた」や「踊ってみた」の動画が含まれ、クラスタ No. 9 にはゲームに関する動画が含まれている。加えて、クラスタ No. 9 には KPOP アーティストが出演するバラエティー番組が含まれており、これらは直接音楽に関係しない動画であった。これらより、クラスタ No. 6 の動画に対してコメントしたユーザは音楽に対して強い興味があり、クラスタ

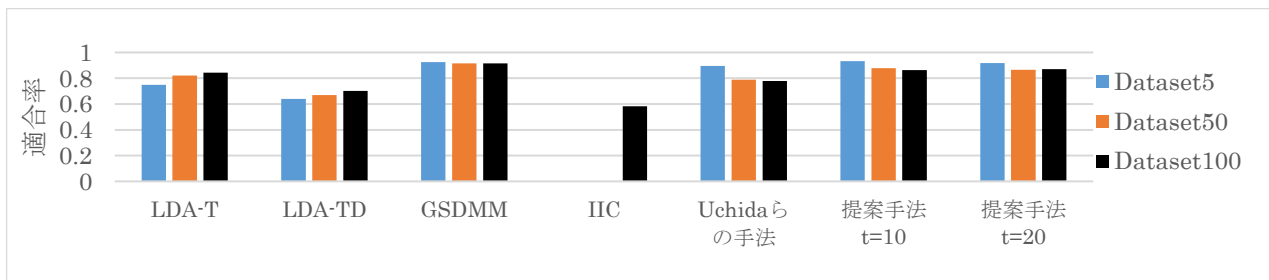


図5 主観実験の結果

No. 9 は動画制作者や動画の企画に対して興味を持ったユーザがコメントしたと考える。

6. 議論

5.4 節の主観実験では、著者のみが被験者として実験を行っている¹。主観実験は複数の被験者により行われることが好ましいが、被験者にはニコニコ動画に関する専門的な知識（アニメ、ゲーム、音楽など）が必要となるため難しい。例えば表 5 のクラスタ No. 1 であれば、被験者は「Fortnite」や「Splatoon」といった単語から、それらが主に若い年代に人気のある TPS というジャンルのゲームのタイトルだということを認識する必要がある。そのため、本論文では 1 人のみが実験を行った。

提案手法のようにユーザのレビューを用いる手法には、レビューの少ないアイテムを取り扱うことが難しいという問題がある。実験において、Dataset5 の 37.7%、Dataset50 の 19.5%、Dataset100 の 13.9% の動画がデータセット内の他のすべての動画と類似度が 0 となった。これは、データセット内のコメントしたユーザの多くが 1 度しかコメントしていないことによると考える。この問題を解決するために、提案手法に他の情報を組み合わせる必要がある。現在のデータセットでは、タイトルなどのテキスト情報や画像などを使うことができる。しかし、これらの情報を組み合わせることで結果が既存手法に大きく近づいてしまう可能性があるため、本研究では用いなかった。

他のクラスタリング手法を提案手法に用いることで、より正確な結果を得られる可能性がある。提案手法では、大きなネットワークを効率的に分割するために Louvain 法をクラスタリング手法として用いている。これは、一般的に動画にコメントしたユーザは数多く存在し、それに伴いネットワークが大きくなるためである。しかし近年、グラフのクラスタリングに対して多くのアプローチが提案されている。例えば、深層学習をもとにした DANMF [11] はオリジナルのネットワークと最終的なコミュニティとの階層的マッピングを学習することでネットワークのコミュニティを検出し、また効率よく学習することができる。行列分解をもとにした CNMMA [12] は、ネットワークのクラスタをノードの属性の共起から発見し、頂点数 n のネットワークを計算量 $O(n^2)$ で計算することができる。これらの手法 Louvain 法の代わりに用いることで、より精度のよい結果を得ることができる可能性がある。

7. おわりに

本論文では、TSC のコメント分布によるニコニコ動画に投稿された動画のクラスタリング手法を提案した。提案手法は動画をコメントしたユーザの集合として扱い、コメント分布の特徴によってユーザの重要度を計算した。定量的評価と定性的評価により、提案手法は既存のテキ

スト、画像ベースの手法や提案手法のもととなったユーザベースの手法に対して同等かそれ以上の結果を示した。今後の課題として、定性的評価に用いた主観実験の複数の被験者による実施や、コメントしたユーザ以外のデータを組み合わせることによるコメント数の少ない動画に生じる問題の解決、他のクラスタリング手法の適用による結果の評価が必要である。

文 献

- [1] X. Zhou, L. Chen, Y. Zhang, D. Qin, L. Cao, G. Huang and C. Wang, "Enhancing Online Video Recommendation Using Social User Interactions," *The VLDB Journal*, vol. 26, pp. 637-656, 2017.
- [2] C. Daniel, "Thematic Exploration of YouTube Data: A Methodology for Discovering Latent Topics," *Muma Business Review*, vol. 1, pp. 141-155, 2017.
- [3] S. Roy and S. C. Guntuku, "Latent Factor Representations for Cold-Start Video Recommendation," *Proc. 10th ACM Conference on Recommender Systems (RecSys)*, pp. 99-106, 2016.
- [4] H. Ren and D. Wang, "TRRS: Temporal Recurrent Recommender System based on Time-sync Comments," *Proc. 3rd International Conference on Machine Learning and Soft Computing (ICMLSC)*, pp. 123-127, 2019.
- [5] K. Tsukuda, H. Masahiro and M. Goto, "SmartVideoRanking: Video Search by Mining Emotions from Time-Synchronized Comments," *Proc. IEEE 16th International Conference on Data Mining Workshops (ICDMW)*, pp. 960-969, 2016.
- [6] W. Yang, W. Gao, X. Zhou, W. Jia, S. Zhang and Y. Luo, "Herd Effect Based Attention for Personalized Time-Sync Video Recommendation," *Proc. IEEE International Conference on Multimedia and Expo (ICME)*, pp. 454-459, 2019.
- [7] K. Uchida, F. Toriumi and T. Sakai, "Evaluation of Retweet Clustering Method: Classification Method Using Retweets on Twitter without Text Data," *Proc. IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, pp. 187-194, 2017.
- [8] 株式会社ドワンゴ, "ニコニコ動画コメント等データ," 国立情報学研究所 情報学研究データリポジトリ, 2018. [Online]. Available: <https://doi.org/10.32130/idr.3.1>.
- [9] J. Yin and J. Wang, "A dirichlet multinomial mixture model-based approach for short text clustering," *Proc. 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 233-242, 2014.
- [10] X. Ji, H. F. Joao and A. Vedaldi, "Invariant Information Clustering for Unsupervised Image Classification and Segmentation," *Proc. 2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 9864-9873, 2019.
- [11] F. Ye, C. Chen and Z. Zheng, "Deep Autoencoder-like Nonnegative Matrix Factorization for Community Detection," *27th ACM International Conference on Information and Knowledge Management (CIKM)*, pp. 1393-1402, 2018.
- [12] T. He, K. C. C. Chan and L. Yang, "Clustering in Networks with Multi-Modality Attributes," *Proc. IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, pp. 401-406, 2018.

¹ 小規模の実験（提案手法 $t = 10$, LDA-T, GSDMM のそれぞれに 100 タスク）を他の被験者 1 人が行い、著者と似た結果となった。