

# 動画共有サイトにおける協創性を用いたタグ推薦手法の提案

池田拓生† 倉田陽平†

首都大学東京 大学院都市環境科学研究科 観光科学域†

## 1. はじめに

近年 Web 上でのマスコラボレーションや協調的創造活動と呼ばれるような協創が注目を集めている。これら協創の盛り上がる条件として、D.Tapscott ら[1]は「参加者で共有できる明確な目標の設定」を挙げている。先駆的に発展した Wikipedia[2]のような百科事典やオープンソースプロジェクトについては、それらの目的は明確であり、この条件に一致している。また、日本の動画共有サイト「ニコニコ動画」[3]における「初音ミク」を始めとした娯楽コンテンツの大規模な協創についても、ニコニコ動画の持つ特徴的なアノテーション機能であるコメントやタグ機能によって擬似的に目的共有が可能になっていることが大規模な協創の発生要因として濱野[4]は取り上げている。またニコニコ動画については、タグやコメントを、動画コンテンツの新たな発見を実現する検索ツール開発のためのメタデータとして活用する試みも見られる。村上ら[5]は、タグ同士の共起性を用いて、タグの階層化を行い、類似したタグを発見できるシステムを提案している他、中村ら[6]は動画に付与されているコメントを印象別に分類し、どういった印象のコメントが多い動画であるかのデータベースを用意した上で、感情的な条件を加味した検索を可能としている。しかしながら、濱野が指摘するタグの協創における擬似的な目的共有の媒体としての効果について注目し、そのタグの付与された動画の本数と、その投稿者の数といった、タグの協創の規模と質を用いた検索手法については開発されていない。また、タグの協創の規模と質を用いた検索手法を提供した場合、利用者が推薦されるタグから、協創されているジャンルや文脈を発見することが可能となり、協創のための「参加者で共有できる明確な目標の設定」の構築を支援することができると考えられる。そこで本研究では、ニコニコ動画から抽出した動画に付与されているタグデータから、タグの協創の規模と質を指標化し、この指標によってタグを発見する検索システムを開発・提案する。

Suggestion of the tag recommendation technique using co-creation disposition in the video sharing site.

†Takumu Ikeda, Yohei Kurata, Dept. of Tourism Science, Tokyo Metropolitan University

## 2. 取得データセット

ニコニコ動画の検索機能を利用し、2014年11月5日に、「車載動画」「旅行」「聖地巡礼」「Vocaloid」をタグとして持つ動画を抽出し、抽出した動画のメタデータを NicoNicoAPI で取得した。このメタデータには、動画に付与されたタグ、動画 ID、投稿者 ID が含まれる。これらデータのうち、タグに注目し、各タグが付与された動画の本数とそのタグを付与された動画を投稿した投稿者の数を数えた。この結果、「車載動画」には 80,253 本の動画に対し 11,215 人の投稿者と 56,400 種類のタグ、「旅行」には 34,174 本の動画に対し 6,608 人の投稿者と 34,828 種類のタグ、「聖地巡礼」には 3,125 本の動画に対し 1,327 人の投稿者と 6,357 種類のタグ、「Vocaloid」には 336,455 本の動画に対し 50,460 人の投稿者と 236,844 種類のタグがそれぞれ存在した。また、2つの異なるタグが同時に付与されている場合(以下タグセット)についても、そのタグセットが付与されている動画の本数とその動画を投稿した投稿者の数をそれぞれ数えた。

## 3. 協創の規模と質を評価する指標

今回提案するタグの検索手法は、タグを協創の対象となっているジャンルとみなし、協創の規模と質によってタグを推薦するものである。そこで、協創の規模と質を評価する基準を定義する必要がある。今回タグ(セット)検索システムに利用する協創の規模と質を評価する指標を 3つ定義した。

### Index1. 協創規模値

タグ(セット) $u$  が与えられている動画群の総本数  $D_u$  を協創規模値とする。この指標によって、検索するタグ(セット)の規模を指定することが可能となる。

### Index2. 協創ジニ係数

タグ(セット) $u$  が与えられている動画群について、それらの投稿者を動画投稿数の少ない順に並べる。そのとき、順番毎にその投稿者までの累積投稿本数をもとめ、X 軸に累積投稿者数比、Y 軸に累積投稿本数比をとると、図 1 のようなグラフを描くことができる。

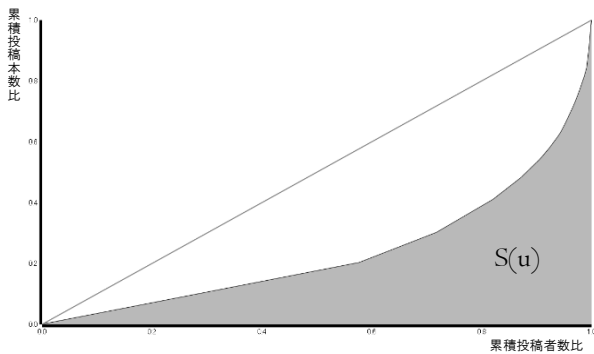


図1. 投稿者毎の投稿本数のローレンツ曲線例

このグラフ中の直線の下側の面積  $S(u)$  を用いて「協創ジニ係数  $Z(u)$ 」を次のように定義する。

$$Z(u) = 1 - 2S(u)$$

この協創ジニ係数  $Z(u)$  は、0~1の間を取り、0に近いほどタグ(セット)  $u$  における投稿者ごとの投稿数が平等であり、1に近いほど投稿者ごとの投稿数に格差がある状態を示す。この係数により、検索するタグ(セット)の示すジャンルの投稿者間の格差の範囲を指定することが可能となる。

### Index3. 協創指数

タグ(セット)  $u$  が与えられている動画群の動画本数  $D(u)$  と、その投稿者数  $U(u)$  を用い、以下のように「協創指数  $C(u)$ 」を定義した。

$$C(u) = \frac{U(u)}{D(u)}$$

この協創指数  $C(u)$  は、0~1の間を取り、0に近いほど少数の投稿者が同じタグ(セット)を持つ動画を大量に投稿しており、1に近いほど、タグ(セット)が付与された動画が多数の投稿者の投稿した動画によって成り立っている状態を示す。この指数によって、検索するタグの示すジャンルの動画数に対する投稿者数の比率の範囲を指定することが可能となる。

これら3つの指標を「車載動画」「旅行」「聖地巡礼」「Vocaloid」で出現したタグ(セット)についてそれぞれ求め、各タグ(セット)とこれら指標を提案システムのためのメタデータとして整理した。なお、出現回数が10回以下のタグ(セット)は、①規模が小さすぎる、②処理の軽量化の2つの理由から、今回システムの検索対象のタグ(セット)からは対象外とした。

### 4. 提案タグ(セット)検索システム

上記に定義した、協創規模値、協創ジニ係数、協創

指数の3つの指標の範囲を利用者が選択することでタグ(セット)の検索を可能とするWebシステムを構築した。システムの概要を図2に示す。

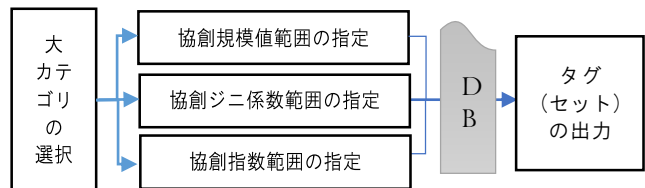


図2. 推薦システムの概要

利用者は、まず画面中でタグ検索をするカテゴリ(「車載動画」「旅行」「聖地巡礼」「Vocaloid」のいずれか)を選択すると検索するデータセットが決まる。次に、協創状況を3つの指標の範囲で指定する。この範囲によってDBから該当するタグ(セット)が判定され、画面上に表示される。最後に表示されたタグ(セット)のいずれかを選択すると、画面下部にこのタグを持つ動画が表示され、リンクから試聴することが可能となる。

### 5. おわりに

本研究では、ニコニコ動画のメタデータの1つであるタグデータを利用し、カテゴリ内のジャンルの協創状況から動画検索を可能とするタグ(セット)検索システムを提案した。このタグ検索システムでは、協創状況からタグ(セット)を発見することが可能となり、利用者が共有するジャンルやその文脈を視覚化することに貢献できる。今後はこのタグ(セット)検索システムの評価実験を行っていく予定である。また、各投稿者の投稿数だけでなく、再生数やマイリスト数、コメント数といった個別の動画の評価数値についても、各指標の計算式に取り入れていくことで、よりタグ(セット)の示すジャンルの協創状況の実態に則したタグ(セット)の推薦が可能となると考えられる。

### 参考文献

- [1] Don Tapscott and Anthony D. Williams (著)・井口耕二(訳)、ウィキノミクス マスコラボレーションによる開発・生産の世紀へ、クニメディア株式会社(2007)
- [2] Wikipedia: <http://ja.wikipedia.org>
- [3] ニコニコ動画: <http://www.nicovideo.jp/>
- [4] 濱野智史、アーキテクチャの生態系—情報はいかに設計されてきたか、NTT出版(2008)
- [5] 村上直至・伊東栄典、動画共有サイトで付与された動画タグの階層化、情報処理学会研究報告、バイオ情報学、Vol.23, No.17, pp.1-6(2010)
- [6] 中村聡史・田中克己、ソーシャルアノテーションに基づく動画検索手法、電子情報通信学会・日本データベース学会・情報処理学会第1回データ工学と情報マネジメントに関するフォーラム(DEIM2009), D6-1, (2009)