

動画共有サイトにおける動画タグの有効度の検証

中村 翔太[†] 山口 実靖[†]

[†]工学院大学工学部情報通信工学科

1. はじめに

Web 上での動画共有サービスが普及しており、日々、ユーザによって動画がアップロードされている。動画共有サービスに投稿された動画には、その動画の特徴を表す「タグ」がユーザによって付与される。同一のタグがついた動画群はグループ化されるため、タグは動画の属性を表すだけでなく、動画の検索にも有効な情報となっている。

本稿では各タグの有効度の算出を最終的な目的とし、そのための基礎調査として動画共有サイトのタグ機能が効果的に機能しているかの検証方法に関する考察を行う。

2. 動画タグ

動画共有サイトでは、投稿された動画にその動画の特徴を表すタグを付加することできる。多くの場合、タグ名はユーザが自由に指定することができるが、タグ編集権限に関しては動画投稿者のみがタグ編集を行える例と、一般閲覧者もタグ編集が可能な例が存在している。

また、同一タグをもつ動画群はグループ化されて管理されるため、タグ情報をもとに関連する動画、類似の動画を発見することが可能となり、動画検索の目的においては重要な意味を意味を持っている。

動画共有サイトにおけるタグの影響、関連性を考察した既存の研究として伊藤らのタグ共起ネットワークの特徴抽出の研究[1]がある。当該研究では共起情報を元にタグのネットワークを作成し、動画の関連性に関する特徴の抽出を行っている。

また、タグ情報を分析に使用した既存の研究として、濱崎らの大規模協調的創造活動の分析[2]がある。当該研究では、異なるタイプの動画作成者がお互いに影響を与えながら大規模に協調的創造活動を行って様子の分析を行っており、タグ情報を創作カテゴリの推定に使用している。

A Study of Effectiveness of Movie Tags in Video Sharing Websites

Shota NAKAMURA[†], Sneyasu YAMAGUTI[†]

[†]Department of Information and Communications Engineering, Kogakuin University

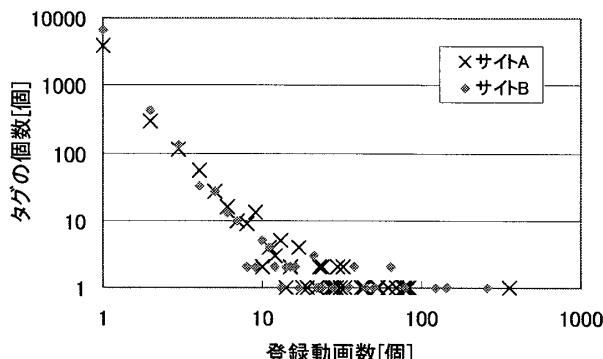


図 1. 登録動画数とタグの個数の関係

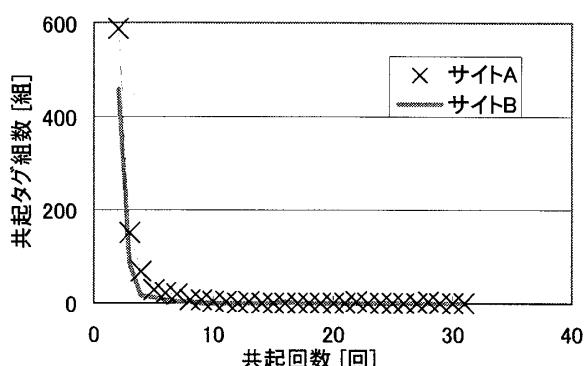


図 2. タグの共起回数と共起タグ組数 (1)

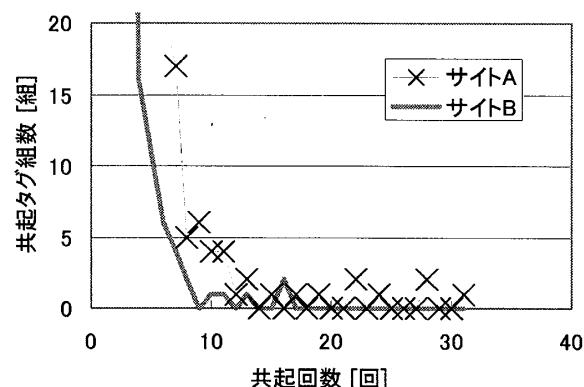


図 3. タグの共起回数と共起タグ組数の関係 (2)

3. タグの使用状況

動画共有サイトより「タグ編集を閲覧者が行える動画共有サイト A」と「タグ編集を動画投稿者のみが行える動画共有サイト B」の 2 つを選び、無作為に選んだ 1000 件の動画より、動画 ID とタグを取得した。

以下における“タグが登録されている動画数”とは本 1000 件の動画内での登録数であり、動画共有サイト全体での登録数ではない。この様な評価方法を用いることにより、動画共有サイトの規模、保有動画数に依存せずにタグの有効性の考察ができると考えられる。

3.1. タグ数

まず、“タグが登録されている動画数”と“タグ数”的関係を調査した。その結果を図 1 に示す。同図における x 座標が a で y 座標が b のプロットは“登録件数 a のタグ(a 件の動画に登録されているタグ)が、b 個ある”ことを意味している。

動画共有サイト Aにおいては、1000 件の動画に付与されているタグの種類は 4505 種であった。登録されている動画が 1 件しかないタグは 3922 個で全体のタグ数の 87.1% であった。また、各タグが登録されている動画数の平均は 1.61 個であった。

これに対して、動画共有サイト B では付与されているタグの種類が 7109 種であり、登録されている動画が 1 件しかないタグが 6420 個で全体の 90.3% であった。また、各タグが登録されている動画数の平均は 1.33 個であった。

以上より、同一動画数で比較した場合サイト B の方がサイト A より多くのタグが登場していることが分かる。

3.2. タグの重複率

次に、動画に付与されているタグの中で、同一タグが重複して使用されている(すなわち、同一タグが複数の動画で使用されている)確率を“重複率”として、この重複率を調べた。

動画共有サイト Aにおいて 1000 件の動画に対する“のべタグ登録回数”は 7253 回であり、重複を除いたタグの種類は 4505 種で、タグの重複率は 37.9% であった。一方、動画共有サイト B では、のべタグ登録回数が 9462 回、タグの種類が 7109 種であり、タグの重複率は 24.9% であった。

以上より、動画共有サイト A の方がより少ない種類のタグを多数回使用する傾向が強く、動画共有サイト B の方が多くの種類のタグを積極的に使用する傾向が強いことが分かった。

3.3. タグの共起数

両サイトの動画 1000 件における、タグ組の共起回数とタグ組数の関係を調査した。結果を図 2, 3 に示す。図 1 同様に x 座標 a, y 座標 b のプロットは、“共起回数 a のタグ組数が b 組ある”ことを意味している。ただし、共起回数が 1 のものはグラフから除いている。

動画共有サイトサイト A での総共起タグ組数は 904 組であり(共起回数 1 を除く)、動画共有サイト B での総共起タグ組数は 585 組あった。動画共有サイト B よりも動画共有サイト A の方がタグ数が少ないとわかるが、共起タグ組数はサイト A の方が多いことが分かった。

以上より、同一動画数で比較した場合、動画共有サイト A ではタグ名が分散せず比較的少ないタグにて動画が整列されており、動画共有サイト B では多様な表現のタグを積極的につけていた傾向があると考えることができる。これは、動画共有サイト A では閲覧者がタグを編集できるため、各動画に付与されているタグが他の動画と同一のものに編集されやすいことが原因であると考えられる。

4. おわりに

本稿では、動画タグの有効度の算出を目的として、2 つの動画共有サイトにおけるタグの登録動画数とタグの個数の関係、タグの重複度、タグ共起数を調査した。その結果、タグ編集を閲覧者が行える動画共有サイト A は、タグ編集を動画投稿者のみが行える動画共有サイト B よりもタグ数が少なく、再利用される回数が多い傾向が見られた。

今後は、タグによる動画検索効率をもとにしたタグ有効度の算出に関する考察、付与されているタグの妥当性の算出に関する考察を行う予定である。

参考文献

- [1] 伊藤聖修 鈴木育男 山本雅人 古川正志，“ニコニコ動画におけるタグ共起ネットワークの特徴抽出”，第 80 回知識ベース研究会 SIG-KBS-A703, pp. 13–18
- [2] 濱崎雅弘 武田英明 西村拓一，“動画共有サイトにおける大規模協調的な創造活動の分析”，情報処理学会研究報告, Vol. 2009-MUS-80 No. 11