情報処理学会論文誌 Vol.54 No.1 214–222 (Jan. 2013)

ニコニコ動画のログデータに基づくソーシャルノベルティのある動画の発見手法の提案

平澤 真大[1] 小川 祐樹[2] 諏訪 博彦[3,a)] 太田 敏澄[3]

受付日 2012年3月31日, 採録日 2012年10月10日

概要：ニコニコ動画等の動画共有サイトでは，多くの人が興味を持つ価値ある動画が日々蓄積されていく動画の中に埋もれていると考えられる．このような埋もれている動画の発見は，ユーザに新たな話題を提供し，ユーザ間のコミュニケーションを促進させると考える．そこで，先行研究におけるノベルティ（知らなかったが，推薦されて興味を持つ）の定義に基づき，本研究では，「社会（集団）的に未知であるが，推薦されて興味を持つ」動画をソーシャルノベルティ（Social-Novelty）のある動画と定義し，発見することを目的とする．我々は，分析により「もっと評価されるべき」タグがソーシャルノベルティのある動画の判定に活用できることを確認している．また，先行研究の整理により，動画を特徴付ける特性として，コメント特性，動画内容特性，視聴行動特性の 3 特性が存在していることを確認している．「もっと評価されるべき」タグが付けられた動画を集団にとってノベルティの高い動画であると考え，この 3 特性を特徴量として，ソーシャルノベルティのある動画を Support Vector Regression（SVR）を用いて発見する手法を新たに提案する．評価実験を行った結果，多くの人が未知かつ興味を持つ動画を発見できること，高い興味を持つ動画を発見できることを確認した．

キーワード：ソーシャルノベルティ，ニコニコ動画，動画推薦，SVR

Finding and Evaluating Social-Novelty Videos in NicoNico Douga

Masahiro Hirasawa[1] Yuki Ogawa[2] Hirohiko Suwa[3,a)] Toshizumi Ohta[3]

Received: March 31, 2012, Accepted: October 10, 2012

Abstract: In NicoNico Douga, many video contents that many people are very intersting is buried in increasing video contents. Discovery of such videos recommends users with new topic, and communication between users is more promoted. Then, based on the definition of the novelty (Not known, but interested) in precedence research, we define Social-Novelty as “In social not known, but intersted” and try to recommend users it. By analysis, we checked that the tag which “would be higher rated” can utilize for the judgment of Social-Novelty videos. Moreover, as the video characteristic, it checked that the 3 characteristics of the Comment-Characteristic, the Content-Characteristic, and Audience’s behavior-Characteristic exist by arrangement of precedence research. We proposed that “would be higher rated” tag and the 3 characteristics in Support Vector Regression (SVR) can discover Social-Novelty videos. To evaluate our proposed algorithm, the results showed the videos were not known but were interesting by many people and higher levels of interseted videos.

Keywords: Social-Novelty, Nico Nico Douga, video recommendation, SVR

1 株式会社日立製作所
Hitachi Ltd., Chiyoda, Tokyo 100–8280, Japan
2 産業技術総合研究所
National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki 305–8568, Japan
3 電気通信大学
The University of Electro-Communications, Chofu, Tokyo 182–8585, Japan

1. はじめに

ニコニコ動画や Youtube 等の動画共有サイトでは，投稿されてから長い時間が経過した後にランキングにランクインする動画が存在する．このような動画はブログや掲示板

a) h-suwa@is.uec.ac.jp

© 2013 Information Processing Society of Japan

で紹介されることで，ランキングにランクインしたと考えられる．日々蓄積されていく動画の中には，このように紹介されれば多くの人が興味を持つ動画が多数埋もれていると考えられる．本研究では，このような埋もれている動画を発見することを目的とする．

近年，インターネットの普及によって大容量の動画を誰でも簡単に視聴・投稿できるようになった．また，ニコニコ動画や Youtube 等の動画共有サイトを国民の 7 割以上が利用しているという調査結果 [1] も発表されており，ニコニコ動画等の動画共有サイトは一般の人々に身近な存在となってきている．しかし，このような動画共有サイトでは，大量の情報を誰もが簡単に取得できるという状況が生じたために，かえって欲しい情報が発見できない，あるいは欲しい情報があるにもかかわらず，それを利用できない状態が生じている．このような状況を情報過多（information overload）の状況ととらえ，推薦システムが考案されている [2]．ニコニコ動画においては投稿動画数が 600 万件を超え，毎日 5,000〜10,000 件の動画が投稿されている現状がある．そのため，興味を持つ動画が日々蓄積されていく動画の中に埋もれてしまうという問題が生じていると考える．

ニコニコ動画では，再生数（視聴したユーザの数）やマイリスト数（お気に入りに登録したユーザの数）によって動画の重み付けを行っており，この重み付けを利用したランキング機能によって，動画が推薦されている．しかし，ランキングに掲載されるためには，多数のユーザに視聴される必要があるため，少数のユーザに視聴されている動画はランキングでは提示されない．動画を視聴するユーザ数と動画の面白さは必ずしも比例関係にあるわけではなく，少数のユーザに視聴されている動画であっても，多数のユーザが興味を持つ面白い動画は多く存在していると考えられる．また，ランキングの構造上の問題点として，元々多くのユーザに視聴されていた動画がさらに注目を浴び，掲載された動画に対する支持がいっそう強固なものになってしまうという問題点をかかえている．そのため，1 度人気になった動画は繰り返しランキングに掲載されやすいと考えられる．同じような動画が繰り返しランキングに掲載されてしまうと，ユーザは飽きを感じ，システムの持続性を考慮すると，ランキングには負の面があると考えられる．ニコニコ動画でも，有名な動画投稿者のシリーズものの動画や，ランキングに頻繁に掲載される人気のカテゴリの動画や，今巷で流行っている動画が繰り返しランキングで提示される傾向があり，新規性や意外性が失われていると考える．

我々は，多くの人は知らないが，視聴した際に多くの人が興味を持つ動画の発見手法を新たに提案することを目的とする．このような埋もれてしまっている動画の発見手法を提案することで，ランキングに飽きを感じているユーザや，動画を投稿したにもかかわらず再生数が伸びない動画投稿者の満足度の向上につながると考える．そこで，ノベルティ [3], [4]（知らなかったが，推薦されて興味を持つ）の定義に基づき，「社会（集団）的に未知であるが，推薦されて興味を持つ」動画をソーシャルノベルティ（Social-Novelty）のある動画であると定義し，このような動画を発見する．ソーシャルノベルティのある動画は，個々のユーザにとってノベルティのある動画の積集合であると考えられる（**図 1**）．ソーシャルノベルティのある動画は，個々のユーザに則した推薦ではなく，社会に適用できる動画である．そのため，そのような動画の発見をすることで，ユーザ全体の満足度を向上させ，ひいてはユーザ同士のコミュニケーションの活性化やシステムの活性化につながると考えている．

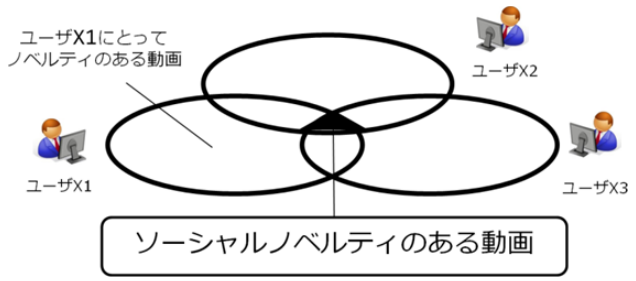

図 1　ソーシャルノベルティのある動画
Fig. 1　Social-Novelty videos.

本研究の構成を以下に示す．2 章では，従来の推薦システムに関する研究について述べる．3 章では，ソーシャルノベルティのある動画を発見するための方法を提案する．4 章では，評価実験について述べる．5 章で実験結果を述べ，6 章で結果に対する考察を行う．7 章で結論と今後の課題を述べる．

2.　関連研究

本章では提案手法と推薦システムに関する研究を比較することで，提案手法の有用性と新規性を示す．

推薦システムのとらえ方の 1 つとして，推薦システムを個人に合わせた個人化の推薦システムと，すべての人に同じアイテムを推薦する非個人化の推薦システムに分けることができる [5]．

よく知られる個人化の推薦システムの 1 つとして，コンテンツベースフィルタリングによる推薦がある [5], [6], [7], [8]．コンテンツベースフィルタリングではアイテムの属性情報から抽出した特徴情報と，ユーザの嗜好を表現した特徴情報を比較し，それらが類似するアイテムを推薦対象アイテムとしてユーザに推薦する．この推薦手法では主に正確さを評価指標として用いているが，正確さのみを評価指標とするだけでは不十分であることが指摘されている [3], [5]．コンテンツベースフィルタリングによる推薦では，アイテムの特徴のみに基づいて推薦対象を決めるため，推薦アイ

© 2013 Information Processing Society of Japan

テムが似かよったものになりやすいという問題点がある．個人に合わせた推薦が行われても，ユーザにとってありきたりなアイテムや，すでに知っているアイテムが推薦されてしまっては，それは有用な推薦とはいえない．そこで近年では，正確さ以外の評価指標を用いた推薦システムとして，ノベルティやセレンディピティ [5] を高める推薦手法が提案されている．

セレンディピティは，ノベルティに思いがけなさ，予見のできなさ，または意外さの要素が加わった概念として示されているが，ユーザの感情的な面に影響される部分が大きいため，定量的に示すことが難しい評価指標とされている [9]．清水らは，ユーザの「知らないアイテム」を推薦することを推薦の発見性ととらえ，ユーザのアイテムへの既知・不既知の情報を用いて発見性を向上させる推薦手法を提案した．この結果，既存の推薦システムよりも，ユーザの未知かつ好みのアイテムを多く推薦できることをオフラインの Novelty 精度 [3] の評価実験により示した [10]．小川らは，ユーザのアイテムの評価履歴からアイテムの評価関係をネットワーク化し，このネットワークをクラスタリングすることにより嗜好傾向を反映させたカテゴリを作成している．このカテゴリによって嗜好傾向の異なるアイテムを複数推薦することで推薦リストの嗜好傾向を多様化し，ノベルティを向上させる手法を提案している [4]．秋山らは，ユーザの嗜好に関するモデルを構築し，そのモデルに基づいてセレンディピティのある推薦を実現している [11]．嗜好の有無を確認する大規模なアンケートをユーザに実施することにより，モデルの構築を行っている．

本研究では，ユーザ全体にノベルティのあるアイテムを推薦することで，ユーザ全体が高く満足する動画の発見を目指す．このような動画をユーザに推薦することで，ユーザ全体の満足度向上によるコミュニケーションの活性化を図り，ひいてはシステムの活性化に期待できると考える．既存の研究では，個々のユーザにノベルティのあるアイテムを推薦することで個々のユーザの満足度の向上を狙っているため，本研究とは異なる．また，ユーザの事前情報を得ることなく，新規性・意外性のある動画を推薦できる点は本研究の強みでもあり，既存の研究と異なる点である．

一方，非個人化の推薦システムの代表的なものとして，ランキングによる推薦がある．しかしランキングでも同じようにありきたりなアイテムが推薦されがちであるという問題がある．ニコニコ動画では，今流行っている動画，多くのユーザが知っているカテゴリの動画，人気のある投稿者が投稿した動画が掲載されることが多く，ユーザがすでに知っている動画である可能性が高い．ユーザが興味を持つ動画であっても，すでに知っている動画が推薦されては，それは有用な推薦とはいえない．また，ランキングに掲載される動画の上限数の関係で，たまたまランキングに入ることができず，動画が埋もれてしまうという問題も考えられる．

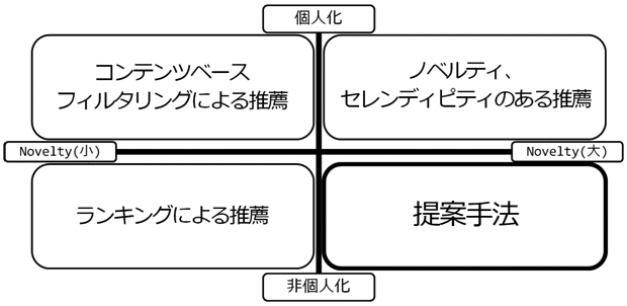

図 2 提案手法と既存の推薦システムの位置づけ
Fig. 2 The proposed recommender system and the existing recommender system.

ソーシャルノベルティのあるアイテムは，アイテムを介して作られるコミュニティ（ニコニコ動画では，動画の視聴者が 1 つのコミュニティを作っている）の活性化を促進させる効果があると考える．なぜなら，多くの人が興味を持つアイテムが推薦されるため，コミュニケーションが活発に行われると考えるからである．個人化の推薦システムでは，あくまで個人の嗜好とアイテムの適合度を測るため，コミュニティの活性化を促進させる効果は薄い．提案手法ではこれから盛り上がっていくアイテムを推薦するため，すでに盛り上がっているアイテムを推薦するランキングに比べ，よりコミュニティの活性化が期待できると考える．提案手法は，関連研究では言及されていないコミュニティの活性化の促進が望めるという点で，有用性がある．また，我々は新たにソーシャルノベルティという概念を定義した．したがって，このようなアイテムを推薦する手法には新規性がある．

以上で述べた提案手法と既存の推薦システムの位置づけを個人化の軸とノベルティの軸によってまとめると，**図 2** のように分類できると考える．

3. 提案手法

ニコニコ動画には感想等を自由に書き込むことができるコメント機能と，動画の特徴を表す言葉を自由に登録・編集することができるタグ機能がある．また，動画が投稿されてから現在までの累計再生数，コメント数，マイリスト数も記録されている．これらの要素に注目して動画を特徴付けている先行研究がある [12], [13], [14], [15], [16]．本研究でも，これらの要素を組み合わせて用いることで，動画を特徴付けられると考える．具体的には以下の 3 つの特性を機械学習手法の 1 つである Support Vector Regression (SVR)[*1]の特徴量として用いる（**図 3**）．

1. コメント特性：コメントに注目
2. 動画内容特性：タグに注目

[*1] SVR は，Support Vector Machine（SVM）を回帰問題に拡張した手法であり，SVM と同様に過学習やデータスパースネスに強いといわれている手法である [17]．

© 2013 Information Processing Society of Japan

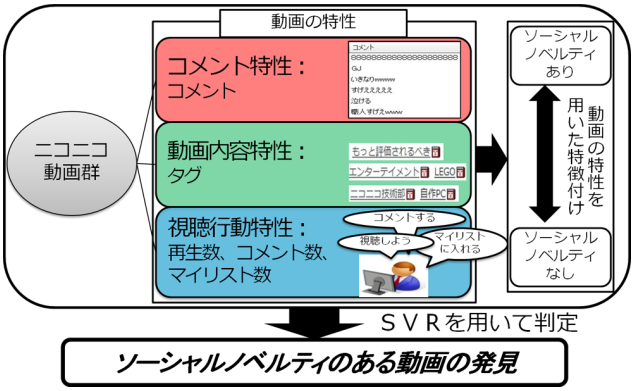

図 3　提案手法の全体図
Fig. 3　The outline of proposed recommender system.

3.　視聴行動特性：再生数，コメント数，マイリスト数に注目

これらの 3 つの特性を組み合わせて用いることで，ソーシャルノベルティのある動画を発見する．SVR による判定は以下の式で表され，訓練事例を $\{x_i, y_i\}, i = 1, \ldots, l, y_i \in \{-1, 1\} x_i \in R^d$ と表し，目的変数であるソーシャルノベルティの程度 L_p を算出する．ここで，x_i は動画の特徴であり，y_i はその動画がソーシャルノベルティのある動画かどうかを表す．

$$L_p \equiv \sum_i \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j x_i^T x_j$$

本章では，ソーシャルノベルティのある動画の判定に必要な y_i については 3.1 節で説明を行い，y_i の説明変数である x_i については 3.2 節で説明を行う．なお，L_p が高い動画であるほど，ソーシャルノベルティの高い動画であると判断し，順位付けを行う．

3.1　ソーシャルノベルティのある動画の判定における正解データ

SVR でソーシャルノベルティの程度を判定するための正解データとして，「もっと評価されるべき」タグに注目する．なぜなら，「もっと評価されるべき」タグは「動画が良質で面白いにもかかわらず再生数等の評価が低い，あるいは存在自体認知されていないと感じられた場合に，視聴者側から自然発生的に貼り付けられるタグ」[18] であり，ソーシャルノベルティの高い動画に付与されると考えられるからである．

我々は「もっと評価されるべき」タグを，ソーシャルノベルティを判定する際の正解データとして用いることが妥当か確認するために，「もっと評価されるべき」タグについて分析を行い，分析結果をまとめた（**表 1**，**表 2**，**図 4**）．なお，分析は 2009 年 9 月 1 日から 2009 年 9 月 5 日に投稿された動画 16,770 件を対象としている．表 1 における「もっと評価されるべき」タグ以外のタグは，ニコニコ動画公式のカテゴリ分けに用いられているタグである．その

表 1　タグの平均マイリスト数の上位 10
Table 1　The average number of mylist top 10 in tag.

タグ	マイリスト数の平均値
演奏してみた	530.0
作業用 BGM	249.9
初音ミク	213.6
Vocaloid	212.2
東方	161.0
もっと評価されるべき	153.8
歌ってみた	148.5
アニメ	140.4
音楽	139.7
スネ夫が自慢話をするときに流れている曲	134.3

表 2　タグの平均マイリスト数と動画数（表の値は平均マイリスト数，括弧の値は動画数）
Table 2　The average number of mylist and the number of videos in tag.

タグ	ランキングに入っていない動画 75(15,636)		ランキングに入っている動画 12,436.7(34)	
	当該タグが付いていない	当該タグが付いている	当該タグが付いていない	当該タグが付いている
もっと評価されるべき	72.6(15,182)	153.8(454)	12,436.7(34)	該当動画なし
音楽	42.8(14,291)	130.7(1,345)	12539.0(31)	11,380.3(3)
ゲーム	121.5(8,202)	23.7(7,434)	19,038.1(20)	2,942.0(14)

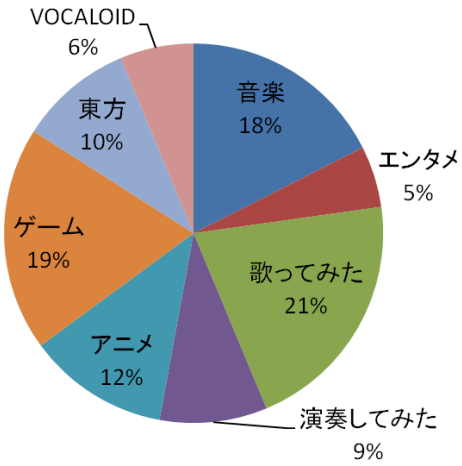

図 4　「もっと評価されるべき」タグの共起関係
Fig. 4　Cooccurrence relation in “would be higher rated” tag.

ため，どのタグもメジャーなタグであるために，マイリスト数が高くなっていると考えられる．表 2 ではニコニコ動画でメジャーなタグである「音楽」「ゲーム」タグを例に，「もっと評価されるべき」タグの有無とランキングの有無によってマイリスト数の比較を行った．図 4 では「もっと評価されるべき」タグの共起関係を分析した．以上の結果をまとめると，「もっと評価されるべき」タグは i〜iii の性質を持っていることを確認している．

i.　マイリスト数が多い．つまり，「もっと評価されるべき」タグが付いた動画に興味・関心を持つユーザが多いと考えられる（表 1）．

ii.　ランキングに掲載されない．すなわち，再生数が低い．タグが付与されない理由は，ランキングに掲載される

© 2013 Information Processing Society of Japan

とユーザが「もっと評価されるべき」タグを削除する自浄作用が働いているからと考えられる（表 2）.

iii. 特定のカテゴリに依存していない．そのため，社会的に知られていない動画をカテゴリにしばられず発見したいという本研究の目的と合致する（図 4）.

以上の分析結果をふまえまとめると，「もっと評価されるべき」タグが付与された動画は，興味を持つユーザの割合が多い一方で，再生数は低いため他の多くのユーザは知らないといえる．また，どのカテゴリにも使われているため，汎用性がある.

我々はこれらの分析結果をふまえて，「もっと評価されるべき」タグをソーシャルノベルティの動画の分類に活用できると判断した．なお，「もっと評価されるべき」タグが付与されていない多くの動画の中にも，ソーシャルノベルティのある動画が多数埋もれていると考えられるため，SVR を用いて「もっと評価されるべき」タグが付与されていない動画の中からも，ソーシャルノベルティのある動画を発見することを目指している.

3.2 ソーシャルノベルティのある動画の判定における説明変数

SVR でソーシャルノベルティの程度を判定するための説明変数として，「コメント特性」「動画内容特性」「視聴行動特性」に注目する．本節では，これらの動画の特性について詳細な説明を行う.

3.2.1 コメント特性

コメント機能はニコニコ動画の特徴的な機能であり，ユーザの生の声を表している．したがって，ソーシャルノベルティのある動画にはそれを表す特徴的な表現があると考える．よってその特徴的な表現を抽出し，動画を特徴付ける特性として用いることで，ソーシャルノベルティのある動画を発見できると考える.

ランダムに抽出した動画 10,000 件のコメント 7,038,296 件を，Mecab を用いて形態素解析し，tf-idf 値の上位 1,000 語をソーシャルノベルティのある動画の分類に活用する．形態素解析を行う際の辞書には，NAIST Japanese Dictionary，表記揺れ判別用辞書「たんし」，wikipedia のタイトルリスト（2011/7/25 時点），ニコニコ大百科 IME 辞書（2011/8/4 時点，ニコニコ動画でしか使われていない独特の言い回しに対応）を用いて，特徴的な単語を発見できるよう工夫した.

3.2.2 動画内容特性

タグ機能もニコニコ動画の特徴的な機能であり，ユーザは誰でも好きなタグを作成し，自由に追加・削除を行える．したがって，タグは，動画の内容を端的に表していると考えられる．そのため，ソーシャルノベルティのある動画には，それと関連するタグが付与されていると考える．よって，タグの頻度と共起パターンを抽出し，動画を特徴付ける特性として用いることで，ソーシャルノベルティのある動画を発見できると考える.

ランダムに抽出した動画 10,000 件のタグの中で，5 回以上使われているタグを抽出した．それらのタグの中で，頻度の上位 1,000 をソーシャルノベルティのある動画の分類に活用する．上位 1,000 のタグは，5 回以上使われているタグの 91%を占めている．タグについては頻度を分類に活用した.

3.2.3 視聴行動特性

再生数，コメント数，マイリスト数は，ユーザが視聴行動をすることで増減する．TV 番組における視聴率のように，ソーシャルノベルティのある動画には，それを表す特徴的な視聴行動特性があると考える．よって，そのような特徴を抽出し，動画を特徴付ける特性として用いることで，ソーシャルノベルティのある動画を発見できると考える.

ソーシャルノベルティのある動画は，再生数の割にコメント数が多い，あるいは再生数の割にマイリスト数が多いと考えられるため，正規化を行う必要がある．そこで，コメント率とマイリスト率を以下の式で規定し，ソーシャルノベルティのある動画の分類に活用した.

$$\text{コメント率} = \text{コメント数}/\text{再生数} \tag{1}$$

$$\text{マイリスト率} = \text{マイリスト数}/\text{再生数} \tag{2}$$

4. 評価実験

「もっと評価されるべき」タグをソーシャルノベルティを判定する際の正解データとし，コメント特性・動画内容特性・視聴行動特性を SVR の特微量とすることで，ソーシャルノベルティのある動画の発見を行った．発見した動画を用いて，評価実験を行う.

4.1 実験目的

提案手法によって発見された動画がソーシャルノベルティのある動画であることを，実際にニコニコ動画に投稿されている動画を用いた被験者実験によって明らかにする.

4.2 実験に用いる動画

実験に用いる動画は以下の A〜C の動画である.

A：提案手法を用いて抽出された動画

B：A と再生数が同程度で提案手法では抽出されなかった動画

C：ニコニコ動画のランキングに掲載されている動画

A と B の動画は，2010 年（1/1〜12/31）に投稿された動画 1,789,979 件のうち，ランダムに抽出した 1 万件の動画から 10 動画ずつ抽出している．A の動画は，SVR[*2] による predict 値が高く，著作権違反とならない上位 10 動画

[*2] SVR の結果, Mean absolute error：0.258, Root mean squared error：0.386 となっている.

© 2013 Information Processing Society of Japan

を抽出している．このAの動画と再生数が最も近く，提案手法では発見されずかつ著作権違反とならない動画をBとして抽出している．なお，この1万件の動画のうち，「もっと評価されるべき」タグが付与された動画は929動画だった．Cの動画は，2010年のニコニコ動画年間総合ランキングの上位10動画を抽出している．A〜Cの動画の再生数の最小値，最大値，平均値を**表 3**に示した．AとBの動画は同程度の再生数を持つ動画であり，Cのランキングから抽出した動画の再生数はAとBに比べ大きな値を持っていることが分かる．

本実験の狙いは，「もっと評価されるべきタグ」が付いていない再生数の低い動画の中から，ソーシャルノベルティのある動画を発見することである．そのため，「もっと評価されるべきタグ」が付いていない動画から提案手法を用いて抽出された動画と，提案手法で抽出された動画と再生数が同程度であり提案手法で抽出されなかった動画を比較し，「もっと評価されるべきタグ」が付いていない動画からソーシャルノベルティのある動画が発見できることを確認する．また，提案手法とランキングを比較することで，ランキングのように知られていないが，見たら興味を持つ動画が発見できることを確認する．

ニコニコ動画では，再生数，コメント数，マイリスト数ごとにランキングが存在するが，これらの数を総合的に判断してランキング表示する総合ランキングがニコニコ動画では最もポピュラーなランキングであると考える．また，2010年に投稿された動画からAの動画を抽出したため，比較対象に用いるランキングは2010年の総合ランキングとした．なお，抽出した動画の再生時間は，約1〜15分程度であり，A，B，Cで極端な再生時間の偏りがないことを確認している．

4.3　実験方法

実験の被験者として，21名に協力を依頼した．被験者の属性は，20代〜30代であり，性別は男性19名，女性2名である．ニコニコ動画を使用したことがない被験者が2名いたが，他の動画サイトは利用していると回答を得ている．

被験者には，A〜Cの各10動画，計30動画をそれぞれ最初の1分程度視聴してもらい[*3]，認知と興味の程度について回答を求めた．認知は，どの程度その動画を知っていたかをについて，「視聴したことがある」「視聴したことはないが，知っていた」「視聴したことはなく，知らなかった」の3段階で評価している．興味は，視聴してどの程度興味が湧いたかを「まったく興味がない」「あまり興味がない」「やや興味がある」「とても興味がある」の4段階で評価している．動画の提示順序については，A，B，Cの動画がランダムに出現するようにし，被験者からはどれが提案手法の動画か分からないように工夫している．

表 3　実験に用いる動画の概要
Table 3　The outline of videos in experiment.

	再生数		
	最小値	最大値	平均値
A：提案手法の動画	1,020	8,365	3,462
B：Aと再生数が同程度の動画	997	8,370	3,460
C：ランキングの動画	2,296,120	5,642,952	3,803,528

*3　明らかに最初の1分が動画内容と無関係の場合は，著者の判断により途中から再生している．

5.　実験結果

視聴したことはなく，知らなかったと回答し，かつやや興味がある・とても興味があると回答された動画数の合計を**図 5**に示す．また，興味の値の平均値を**図 6**に示す．図5においてはライアン法によって，図6においては一元配置分散分析によって，有意差検定を行っている．図5においては提案手法によって発見された動画と他の動画間で1%の有意差を確認している．また，図6においては提案手法によって発見された動画とランキングから抽出した動画において，再生数が同程度の動画間でそれぞれ1%の有意差を確認している．

図5から，提案手法が未知かつ興味の値が高い動画を多く発見できていることが確認できる．図6から，提案手法が，ランキングと同程度に高く興味を持たれる動画を発見できていることが確認できる．

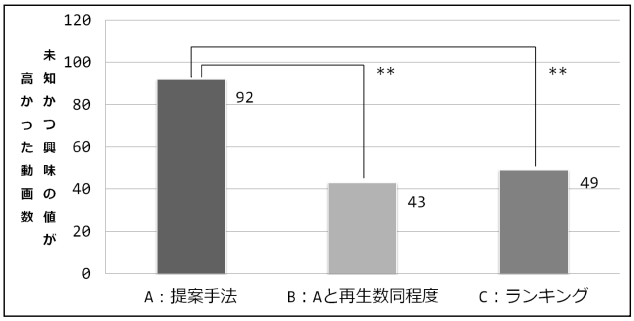

図 5　未知かつ興味の値が高かった動画数（n=21, **p < 0.01，ライアン法により検定）
Fig. 5　The number of unknown and no interested videos.

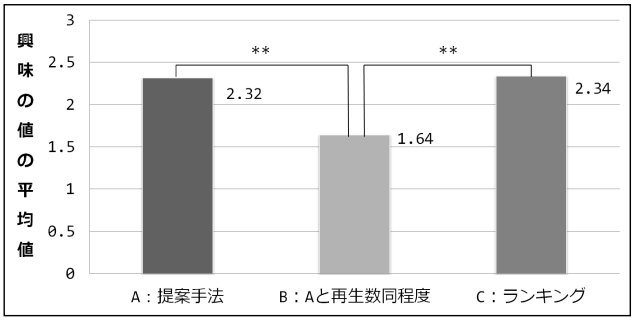

図 6　興味の値の平均値（n=21, **p < 0.01，一元配置分散分析により検定）
Fig. 6　The average of interested value.

© 2013 Information Processing Society of Japan

6. 考察

本章では，提案手法の有効性およびソーシャルノベルティのある動画を発見することの効果，本研究の限界について考察する

6.1 提案手法の有効性

図 5 の結果から，提案手法が未知かつ，興味の高い動画を多く発見できていることが分かる．図 6 の結果から，提案手法によって発見された動画は，ニコニコ動画ランキングで上位の動画と同程度に高く興味を持たれる動画であることが分かる．このことから，提案手法がソーシャルノベルティのある動画を多く発見する手法として有効であると考える．

3.1 節で示したとおり，「もっと評価されるべき」タグが付いた動画に興味・関心を持つユーザは多いと考えられ，「もっと評価されるべき」タグのみでも，ソーシャルノベルティのある動画の発見に寄与すると考える．「もっと評価されるべき」タグがなく predict 値の高い上位 30 動画[*4]を抽出し，それと視聴回数が同程度の「もっと評価されるべき」タグが付いている 30 動画と，「もっと評価されるべき」タグが付いていない 30 動画のマイリスト数を比較した．その結果，predict 値の高い上位 30 動画は平均 58.9，もっと評価されるべき」タグが付いている動画は平均 56.4 とほぼ同等であったのに対し，「もっと評価されるべき」タグが付いていない動画は平均 37.2 と低い値となった．このことから，提案手法は，「もっと評価されるべき」タグと同様にソーシャルノベルティのある動画の発見に寄与すると考える．提案手法の優位点は，「もっと評価されるべき」タグが付いていないにもかかわらず，ソーシャルノベルティのある動画の発見できることである．これにより，自動的にソーシャルノベルティのある動画を発見することができる．このことは，ランキング等に入らず埋もれてしまう動画の発見につながると考える．

6.2 ソーシャルノベルティのある動画の効果

本節では，ソーシャルノベルティのある動画を提案手法によって推薦する利点を，動画の投稿者と視聴者の立場からそれぞれ述べる．

動画の投稿者の立場では，自身の投稿した動画がマイナなカテゴリであった場合や，たまたまランキングに掲載されることなかった場合に，提案手法によって推薦されることで人気が出て再生数やマイリスト数が伸びる，あるいは多くのコメントが付与される場合が考えられる．投稿者にとって，自身の投稿した動画の再生数等が伸びることは，投稿者の満足度の向上につながると考えられるため，投稿者にとって提案手法は貢献できると考える．視聴者側の立場では，これから盛り上がっていく動画を他のユーザに先駆けて視聴することで，“他のユーザにおすすめ動画を紹介できるユーザ” となる．インターネットの世界でのコミュニケーションに限らず，オフラインの世界でのコミュニケーションを行うきっかけにもなり，コミュニケーションをとることが苦手だがコミュニケーションをとりたいと考えるユーザにとって有意義であると考える．また，まだ盛り上がっていない動画がソーシャルノベルティのある動画として推薦されるため，ユーザはこれから一緒に盛り上がっていくことができる．ニコニコ動画を「動画を介したコミュニケーションを行うサイト」ととらえると，一緒に盛り上がっていくことによってコミュニケーションが生じることは，コミュニティの活性化を促し，ひいてはシステムの活性化にもつながってくるという利点があると考えられる．実際にニコニコ動画で推薦システムとして運用することを考えると，ソーシャルノベルティのある動画推薦とランキングによる動画推薦とノベルティのある動画推薦は，それぞれ推薦する動画の性質が異なる．そのため，ユーザビリティの観点から，これらの推薦システムは並列して用いることが望ましいと考える．また，ニコニコ大百科（仮）の動画紹介等，外部サイトへの掲載も考えられる．ニコニコ大百科（仮）等で紹介される動画は，ニコニコ動画のランキングを変化させる動画として機能することがあり，コミュニティの活性化を促すと考える

ソーシャルノベルティのある動画は，ランキングに飽きてきたニコニコ動画ユーザに対して，特に有意義な推薦システムであると考える．なぜなら，ランキングではありきたりな動画が掲載されがちであるためである．特に，利用頻度の高いユーザにとって，この傾向は強いと考えられる．

6.3 本研究の限界

研究の制約として，本研究では「もっと評価されるべき」タグを正解データとしてソーシャルノベルティのある動画を発見したが，他の正解データを組み合わせて用いることを検討していく必要があると考える．「もっと評価されるべき」タグはカテゴリに依存することなく用いられているタグであったが，ほかにもこのようなタグの存在を確認している．たとえば，「もっと再生されるべき」タグや「本当にもっと評価されるべき」タグ等が該当する．また，特定のカテゴリに依存していながら，ソーシャルノベルティのある動画を発見するために有用と思われるタグの存在も確認している．たとえば，「もっと浄化されるべき」タグは，テレビゲーム「エルシャダイ」シリーズに特化しているタグである．タグだけでなく，動画のタイトルや説明文に含まれるテキストデータも組み合わせて用いることで精度の

*4 被験者実験では，著作権侵害の恐れがある動画を除外したため，実質的には 29 位（predict 値 1.32）の動画を視聴動画として用いている．当該動画は，ソーシャルノベルティのある動画として選択されている．そこで，30 位（predict 値 1.31）の動画を加え比較対象としている．

© 2013 Information Processing Society of Japan

向上を狙いたい．

2つ目の限界として，対象がニコニコ動画に限定されていることがあげられる．他の書籍検索サイトや音楽サイトにおいても，訓練事例を設定できれば，本提案と同様の手法でソーシャルノベルティのあるアイテムを発見できると考える．通常のサイトでは，「もっと評価されるべき」タグのようなフラグが存在しないため，どのようにフラグを立てるかが課題となると考える．

3つ目の限界として，被験者の年代のほとんどが20代であることがあげられる．ドワンゴによれば，ニコニコ動画の利用者は10代19.6%，20代44.1%，30代21.6%であり，被験者の年代を考慮して実験を行う必要がある [18]．最も利用者の多い20代を対象に実験を行っているとはいえ，他の世代においてもソーシャルノベルティのある動画が推薦可能であるか検証する必要があると考える．また，被験者の男女比に偏りがあることも問題である．ニコニコ動画の利用者の3分の1は女性であり，被験者の性別を考慮して実験を行う必要がある．今回の実験に用いた動画のジャンルは，エンタメ，歌ってみた，動物，音楽，VOCALOID等であり，性別による大きな影響はないと考えるが，性差による影響は検証できていない．今後，異なる推薦対象・多様な年代・性別の被験者に対して本研究の結果の比較を行うことで，推薦するアイテムや年代・性別による違いを分析していきたい．

4つ目の限界として，本手法の有効範囲の確認があげられる．被験者実験を用いたため，視聴できる動画数に制限があった．そのため，predict値がどのくらいのものまで，実用上有効なのかを確認することはできていない．また，6.1節で行った提案手法で抽出した動画と「もっと評価されるべき」タグが付いている動画の比較についても，マイリスト数の比較にとどまらず，利用者の直接的な評価が望ましい．被験者の負担にならない検証方法の検討が今後の課題である．

7. 結論

本研究では，「社会的に未知であるが，推薦されて興味を持つ動画」をソーシャルノベルティと定義し，そのような動画を発見するための手法を提案した．さらに，提案手法の有用性を明らかにするために被験者実験を行った．実験結果から，提案手法を用いることで，ソーシャルノベルティのある動画の推薦が行えることを示した．

今後は，よりソーシャルノベルティの高い動画を発見するために，コメントやタグの特徴後の抽出に工夫をしていきたい．具体的には共起関係の解析や主成分分析を行うことで，精度の向上を目指したいと考える．

参考文献

[1] Recording Industry Association of Japan：「動画サイトの利用実態調査検討委員会」報告書，入手先 ⟨http://www.riaj.or.jp/release/2011/pdf/20110808report.pdf⟩.

[2] 神嶌敏弘：推薦システム—情報過多時代をのりきる，情報の科学と技術，Vol.56, No,10, pp.452–457 (2006).

[3] Herlocker, J., Konstan, J. and Riedl, J.: Evaluating collaborative filtering recommender systems, *ACM Trans. Information Systems*, Vol.22, No.1, pp.5–23 (2004).

[4] 小川祐樹，諏訪博彦，山本仁志，岡田 勇，太田敏澄：動的なトピック分類に基づくNoveltyを考慮した推薦アルゴリズムの提案，情報処理学会論文誌，Vol.50, No.6, pp.1636–1648 (2009).

[5] 神嶌敏弘：推薦システムのアルゴリズム（1），人工知能学会論文誌，Vol.22, No.6, pp.826–837 (2007).

[6] 土方嘉徳：嗜好抽出・情報推薦の基礎理論1，嗜好抽出と情報推薦技術，*IPSJ Magazine*, Vol.48, No.9, pp.957–965 (2007).

[7] Pazzani, M., Muramatsu, J. and Billsus, D.: Syskill and webert: Identifying interesting web sites, *Proc. 13th National Conf. Artficial Intelligence*, pp.54–61 (1996).

[8] Mooney, R.J. and Roy, L.: Content-Based Book Recommending Using Learning for Text Categorization, *Proc. ACM SIGIR'99 Workshop Recommender Systems: Algorithms and Evaluation* (1999).

[9] McNee, S.M., Riedl, J. and Konstan, J.: Being accurate is not Enough: How Accuracy Metrics have hurt Recommender Systems, *A Work-In-Progress Paper in the Extended Abstracts of the 2006 ACM Conf. Human Factor in Computing Systems* (*ACM SIGCHI*), pp.1097–1101 (2006).

[10] 清水拓也，土方嘉徳，西田正吾： 発見性を考慮した協調フィルタリングアルゴリズム，電子情報通信学会論文誌，Vol.93-D, No.3, pp.538–550 (2008).

[11] 秋山高行，小原清弘，谷崎正明： ユーザの選択履歴に依存しない指標を利用したserendipityのある推薦方式の提案と評価，情報科学技術フォーラム講演論文集，Vol.9, No.4, pp.157–164 (2010).

[12] 青木秀憲，宮下芳明：ニコニコ動画における映像要約とサビ検出の試み，情報処理学会研究報告，ヒューマンコンピュータインタラクション研究会報告，Vol.50, pp.37–42 (2008).

[13] 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な協調的創造活動の創発のネットワーク分析，人工知能学会論文誌，Vol.25, No.1, pp.157–167 (2010).

[14] 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，人工知能学会研究会資料，知識ベースシステム研究会，Vol.80, pp.13–18 (2008).

[15] 村上直至，伊東栄典：動画投稿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).

[16] 森 友則，朝香卓也，高橋達郎： 動画共有サイトにおけるコンテンツ人気度推移の分析，電子情報通信学会技術研究報告，ネットワークシステム，Vol.108, No.286, pp.87–92 (2008).

[17] Smola, A.J. and Schölkopf, B.: A tutorial on support vector regression, NeuroCOLT2 Technical Report NC2-TR-1998-030 (1998).

[18] ニコニコ大百科「もっと評価されるべき」，入手先 ⟨http://dic.nicovideo.jp/a/%E3%83%8B%E3%82%B3%E3%83%8B%E3%82%B3%E5%A4%A7%E7%99%BE%E7%A7%91⟩（参照 2012-03-01）.

© 2013 Information Processing Society of Japan

[19]　株式会社ドワンゴ：平成 23 年 9 月期決算説明資料，ドワンゴ（オンライン），入手先
⟨http://info.dwango.co.jp/pdf/ir/news/2010/101111_pt.pdf⟩（参照 2012-07-29）.

平澤 真大

2010 年電気通信大学情報通信工学科卒業．2012 年同大学院情報システム学研究科博士前期課程修了．現在，（株）日立製作所勤務．

小川 祐樹　（正会員）

2006 年大分大学工学部卒業．2008 年電気通信大学大学院情報システム学研究科博士前期課程修了．2011 年同大学院博士後期課程修了．博士（工学）．現在，産業技術総合研究所特別研究員．推薦システム，社会シミュレーションの研究に従事．

諏訪 博彦　（正会員）

1998 年群馬大学社会情報学部卒業．2006 年電気通信大学大学院情報システム学研究科博士後期課程修了．博士（学術）．現在，電気通信大学大学院情報システム学研究科社会知能情報学専攻社会情報システム学講座助教．ソーシャルメディアに関する研究に従事．

太田 敏澄　（正会員）

1970 年東京工業大学経営工学科卒業，1972 年同大学大学院理工学研究科修士課程修了．1977 年工学博士．現在，電気通信大学大学院情報システム学研究科教授．社会情報システム学，組織知能工学，ソーシャルメディアに関する研究に従事．社会情報学会（SSI），日本ソフトウェア科学会，経営情報学会，日本 OR 学会，IEEE 等．

© 2013 Information Processing Society of Japan