

統計モデルに基づく楽曲動画の評価推移分析

An Evaluation Transition Analysis of Music Videos Based on a Statistical Model

山岸 祐己*¹ 齊藤 和巳*¹
Yuki YAMAGISHI Kazumi SAITO

静岡県立大学大学院経営情報イノベーション研究科
Graduate School of Management and Information of Innovation, University of Shizuoka

The evaluation of a video in Nico Nico Douga is strongly dependent on the ability to attract users which contributors and contents have. Therefore, it is difficult to use the videos which have high ability to attract users, and the other videos as the same data. In contrast, we assume that a user's action which registers a video into his or her favorites is a Bernoulli trial, and convert a registration rate of each video to an evaluation value in consideration of reliability by view counts of the video. Furthermore, we try clustering the videos using transitions of the evaluation value.

1. はじめに

近年、国内の大規模動画共有サービスである「ニコニコ動画*¹」において、音声合成ソフトウェア「VOCALOID*²」を用いた楽曲動画の投稿が非常に盛んである。しかし、ニコニコ動画内で VOCALOID 楽曲動画が日に日に増加する反面、注目される投稿者は限定されていく傾向が見られている。人気投稿者が投稿している楽曲動画は、キャラクター、ストーリー、歌詞といった演出面で工夫をこらせたものが多いので、楽曲動画には動画としての質や話題性も求められていることが伺える。逆に言えば、それらが欠如した楽曲は、十分な評価を得られていない可能性が高いということである。

上記のような状況が形成され、加速し続けている最大の原因として、動画が再生数等の単純な推移を基にランキングされていることが挙げられる。Salganik らの大規模な実験 [1] においても、再生数のような社会的情報に基づく楽曲の降順一列表示は、個々の意思決定に多大な影響を与え、市場の不平等性を大いに増加させるとして明確に示されている。さらに、この実験結果は、全く同じ楽曲群であっても、社会的情報信号を増幅させるか否かで、最終的な楽曲に対する評価は大きく変動するという結論付けている。つまり、圧倒的な影響力を持っている社会的情報は、個人の既存の嗜好すらも捻じ曲げてしまうということである。

そこで我々は、動画の再生数に依存しない新たな動画の評価手法を提案する。本論文では、ユーザが動画を再生してマイリストに登録する行動をベルヌーイ試行であると仮定し、動画がマイリストに登録された割合を、再生数による信頼性を考慮した評価値に変換することを試みる。さらに、その評価値の推移の類似度を計算することによって、動画の集客能力に依存しない評価パターンのクラスタリングを試みる。

本論文の構成は以下となる。まず、楽曲動画データセットの説明を述べる。次に、ニコニコ動画のランキングの詳細を示すと共に、提案法について説明する。最後に、従来法と提案法の比較と、本研究のまとめについて述べる。

連絡先: 山岸 祐己, 静岡県立大学大学院経営情報イノベーション研究科, 静岡県静岡市駿河区谷田 52-1, 054-264-5436, yamagissy@gmail.com

*1 <http://www.nicovideo.jp/>

*2 <http://www.vocaloid.com/>

2. 楽曲動画データセット

2.1 データセットの概要

今回使用するデータセットは、ニコニコ動画から取得した VOCALOID オリジナル曲*³の楽曲動画データである。本論文で使用するデータは、2013年4月3日から2014年1月3日まで24時間毎に取得したものであり、最終日付で106343曲の動画情報を有する。最終日付の基本統計量を表1に示す。なお、ここで Length は動画の再生時間(秒)を意味する。

表1: データセットの基本統計量

	Views	Comments	Mylists	Length
Max	10785852	11720493	292950	6132
Minimum	14	0	0	0
Average	12947.07	714.66	527.00	235.86
Median	869	22	24	235
Mode	220	4	3	240
S.D.	112285.78	50413.40	4019.01	118.51

2.2 データ項目の説明

ここでは、データセットへの理解を深めるべく、各データ項目の説明について述べる。

「再生数」とは、その動画が再生された回数を示すものである。再生数は、あるユーザが動画ページを読み込んだ時点で加算され、十分な時間*⁴が経過すれば、同一ユーザによって再び加算が可能となる。なお、プレイヤー機能によるリピート再生では再生数の加算は行われぬ。因みに、cookie を削除しつつ再生を行う等の不正アクセス行為によって再生数を意図的に増加させることを一般に「工作」と呼ぶ。

「コメント」とは、動画に書き込むことができる文字列のことであり、このコメント機能がニコニコ動画の最大の特徴と言われている。各動画における「コメント数」とは、それらコメントが動画に付けられた回数を示すものである。連投制限*⁵さえ守れば、同一ユーザによる加算がいくらかでも可能であり、

*3 VOCALOID を用いて演奏された投稿者オリジナルの楽曲のこと。

*4 公表はされていないが、一般に60分とされている。

*5 公表はされておらず、20秒で10コメントが目安とされている。

投稿者や運営が意図的にコメントを削除しない限り、値が減少することは無い。

「マイリスト」とは、ニコニコ動画内にあるお気に入りの動画をブックマークとして保存することができる機能のことである。各動画における「マイリスト数」とは、その動画をマイリストに登録しているユーザ数を意味している。マイリスト数は累積ではなく、即時的な値であり、同一ユーザによる重複も許されていないため、減少することも大いにあり得る。再生数同様、不正アクセス行為や、意図的な一斉登録によって「工作」が行われることが多々ある。

3. 従来の動画ランキング

3.1 ニコニコ動画における総合ポイント

ニコニコ動画は、2014年1月時点で約1000万という膨大な動画を有しており、日々様々なランキングがサイトの至るところで展開されている。ニコニコ動画において「ランキング」といった場合、「ニコニコ動画の運営が公式に公開している公式ランキング」か、「ニコニコ動画の運営が公式に発表するデータ等を元にユーザが独自に作成したランキング動画」のいずれかを指すことが多い。どちらのランキングも、基本的には「集計基準」「対象とする動画カテゴリ」「集計期間」の組み合わせで成り立っている。

ニコニコ動画が公式として扱っている集計基準は、再生数、コメント数、マイリスト数、ニコニ広告*6 宣伝ポイントを総合評価した「総合ポイント」である。「総合ポイント」の計算方法は明らかにされていないが、一般には『総合ポイント = 再生数 + (コメント数 × 補正值) + マイリスト数 × 15 + ニコニ広告宣伝ポイント × 0.3』『補正值 = (再生数 + マイリスト数) ÷ (再生数 + コメント数 + マイリスト数)』で算出できるとされている。現在のニコニコ動画のデフォルトランキングは、この総合ポイントによるランキングであるため、ユーザにより作成される「ランキング動画」でも、同じような計算方法が採用されることが多い。

公式ランキングでは、全てのカテゴリを扱う「カテゴリ合算ランキング」の他に、約30種類の動画カテゴリごとの「カテゴリランキング」が存在する。集計期間は、毎時、24時間、週間、月間、合計（投稿されてから現在まで）があり、デフォルトでは24時間ランキングが表示される。

3.2 総合ポイントの問題点

前節で示したように、総合ポイントの計算方法は、主に再生数、コメント数、マイリスト数の三項目の足し算であり、「補正值」等は気休めにすぎない。従って、集計期間において評価されるのは、「どれだけのユーザが動画に群がったか」という単純な社会的情報であるため、楽曲動画の質やユーザの既存の嗜好よりも、話題性や集客能力が先行して公式ランキングの結果に影響を与えていると考えることができる。

一方、ユーザが独自で公式ランキングとの差別化を図ったランキングを作成することは珍しくなく、その手のランキング動画やランキングサイトも数多く存在している。ユーザが作成するランキングの総合ポイント計算式では、再生数やコメント数よりもマイリスト数やマイリスト率*7 を重視していることが多く見受けられる。これは、「マイリストに登録する」という

ユーザのアクションが、ニコニコ動画において、唯一ユーザの嗜好を明確に示すものであるということに起因していると思われる。しかし、結局のところユーザが作成したランキングは、公式ランキングの結果に似通っているもの、又は、評価の信頼性に疑問を感じるものが殆どである。

ここで、図1に取得期間中の再生数の推移とマイリスト数の推移のプロットを、図2に最終日付のデータにおけるマイリスト率と再生数のプロットを示す。図1から分かるように、マイリスト数の増加には、相応の再生数の増加も必要となる。故に、マイリスト数の推移を重視したところで、公式ランキングとの明確な差別化にはならないことが示唆される。加えて、図2の下部から見て取れるように、再生数が少ないうちは、偶然マイリスト率が高くなってしまいがちである。よって、マイリスト率に重きを置き過ぎると、評価の信頼性が薄れてしまうことが危惧される。以上のことから、マイリスト率

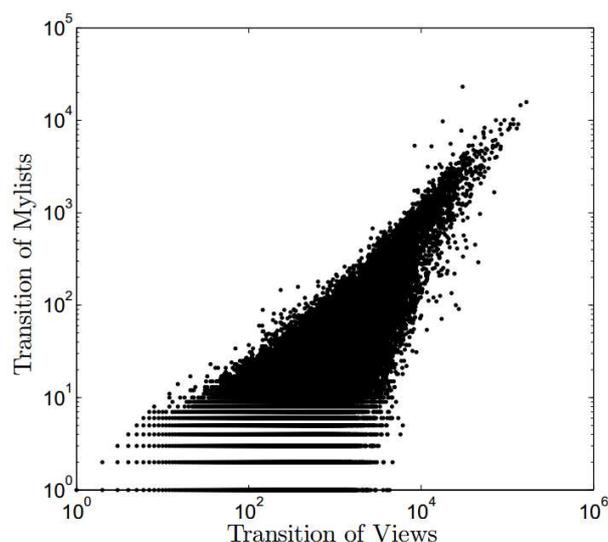


図1: 再生数の推移とマイリスト数の推移のプロット

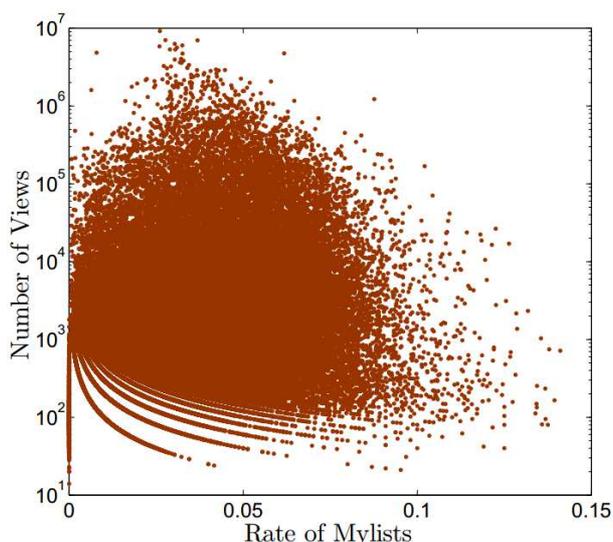


図2: マイリスト率と再生数のプロット

*6 ニコニコポイント（有料コンテンツのための仮想通貨）を使って、ユーザが動画または生放送を宣伝することができるニコニコ動画のサービス。

*7 動画がマイリストされた割合。「マイリスト数/再生数」で求められる。

を動画の評価指標として認めつつも、動画の再生数からその信頼性を割り出す必要があることが分かる。さらに、再生数とマイリスト率の両方を考慮した評価値は、動画の話題性や集客能力に左右されないということも期待できる。

4. 提案法

楽曲動画 $n \in \{1 \dots N\}$ の再生数を v_n 、マイリスト率を μ_n とすると、楽曲動画データ集合 \mathcal{D} は以下のように書ける。

$$\mathcal{D} = \{(v_1, \mu_1), \dots, (v_N, \mu_N)\}. \quad (1)$$

ここで「あるユーザがある動画 n を再生した際、マイリストに登録するか否か」というベルヌーイ試行を考える。試行結果 $j \in \mathcal{J} = \{0, 1\}$ の $j = 0$ を「再生のみ」、 $j = 1$ を「再生とマイリスト登録」とし、マイリスト率の平均 $\mu = 1/N \sum_{n=1}^N \mu_n$ を試行の成功確率、即ちマイリスト登録確率とすると、各結果の確率 $p(j)$ は $p(0) = (1 - \mu)$ 、 $p(1) = \mu$ となる。それぞれの試行が $p(j)$ に従って独立に行われたと仮定すると、 Q 回の試行結果 $S = \{j_1 \dots j_Q\}$ により得られたマイリスト率の期待される誤差の標準偏差（二乗平均平方根誤差）は以下となる。

$$\sqrt{\sum_{j_1 \in \mathcal{J}} \dots \sum_{j_Q \in \mathcal{J}} \left(\mu - \frac{1}{Q} \sum_{q=1}^Q j_q \right)^2 \prod_{q=1}^Q p(j_q)} = \frac{\sigma}{\sqrt{Q}}. \quad (2)$$

このとき、楽曲動画 n のマイリスト率 μ_n の Z-score は、試行回数を v_n 、 $\sigma = \sqrt{\mu(1 - \mu)}$ とおいて、以下のように考えることができる。

$$z_n = \frac{\mu_n - \mu}{\sqrt{\frac{\mu(1 - \mu)}{v_n}}}. \quad (3)$$

提案法は、この z_n を楽曲動画 n の即時的な評価値として扱うものである。

5. 従来評価値との比較

図 3 に提案法による評価値の分布を示す。横軸が評価値 z_n 、縦軸が再生数 v_n 、色がマイリスト率 μ_n である。図より、再生数が増加するにつれて、マイリスト率に対する評価が大きく反映されていることが見て取れる。

5.1 再生数別出現頻度比較

従来のような総合ポイントを想定した評価値（再生数 + マイリスト数 $\times 15$ ）と、提案法による評価値の比較を行う。図 4 は、24 時間毎に評価値の推移順で降順ソートしたときの、上位 100 位での再生数別出現頻度である。図より、従来ランキングではランクインし難い低再生数の動画が、提案ランキングの方で頻繁に出現しているため、提案法は従来法との明確な差別化ができていけると言える。なお、提案法のランキングには、恒常的に再生数とマイリスト数が増え続ける高再生数の動画もランクインしていることに注意されたい。

5.2 K-medoids クラスタリング結果比較

K-medoids 法によって楽曲動画集合 $\mathcal{N} = \{1, \dots, N\}$ を K 個のクラスタに分割することを試みる。まず、任意の動画ペア $\alpha, \beta \in \mathcal{N}$ 間の類似度を $\rho(\alpha, \beta)$ とする。そして、動画集合の中から他の動画との類似度の和が高い動画を選定し、類似度の高い動画ペアは同じクラスタに、類似度の低い動画ペアは異なるクラスタに属するように分割する。一般的に、centroid より medoid の方が頑健であることが知られている。K-medoids の

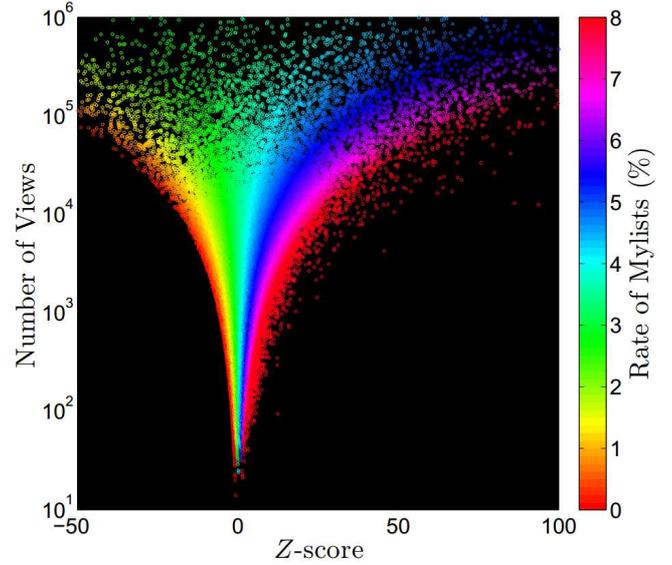


図 3: 提案法による評価値の分布

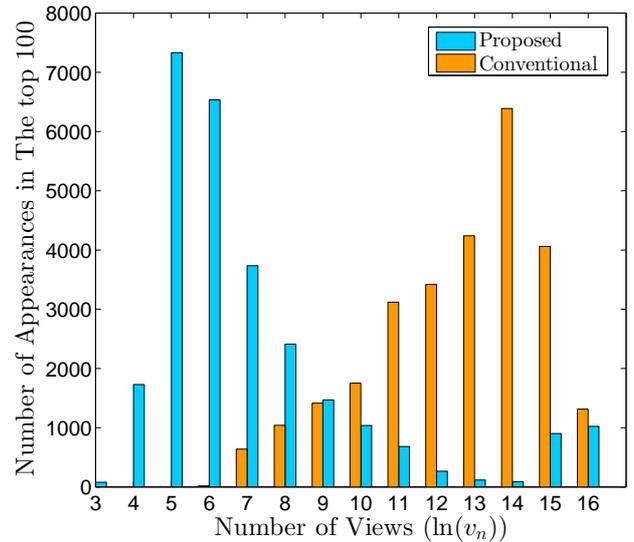


図 4: 24 時間ランキング上位 100 位での再生数別出現頻度

解法には反復法や貪欲法があるが、本研究では解の一意性が保証される貪欲法を採用する。貪欲法は、目的関数のサブモジュラ性により、厳密解ではないものの、ある程度妥当な精度で最悪ケースの解品質が理論的に保証されている [2]。貪欲法とは、既に選定した代表動画を固定し、評価関数値を最大にする動画を求め、目的関数が増加するならば代表動画集合に追加する方法である。代表動画が追加される度に、各動画は最も類似度の高い代表動画と同じクラスタに割り当てられる。既に選定した代表動画集合を \mathcal{G} とし、新たに追加を試みる動画を w とするとき、以下の目的関数を考える。

$$f(\mathcal{G} \cup \{w\}) = \sum_{n \in \mathcal{N}} \max\{\lambda(n; \mathcal{G}), \rho(n, w)\}. \quad (4)$$

ここで、 $\lambda(n; \mathcal{G})$ は既に選定された代表動画との類似度の最大値を表し、 $\lambda(n; \mathcal{G}) = \max_{g \in \mathcal{G}} \{\rho(n, g)\}$ で定義される。以下に貪欲法による K -medoids 法のアルゴリズムを説明する。

1. $k \leftarrow 1, \mathcal{G}_0 \leftarrow \emptyset$, 各動画 $n \in \mathcal{N}$ に対し、 $\lambda(n; \emptyset) \leftarrow 0$ と初期化する；
2. $\hat{g}_k = \arg \max_{w \in \mathcal{N} \setminus \mathcal{G}_{k-1}} \{f(\mathcal{G}_{k-1} \cup \{w\})\}$ を求め、 $\mathcal{G}_k \leftarrow \mathcal{G}_{k-1} \cup \{\hat{g}_k\}$ とする；
3. $k = K$ ならば $\hat{\mathcal{G}}_K = \{\hat{g}_1, \dots, \hat{g}_K\}$ を出力し終了する；
4. 各動画 $n \in \mathcal{N}$ に対し、 $\lambda(n; \mathcal{G}_k)$ を求め、 $k \leftarrow k+1$ としステップ 2. へ戻る。

最後に、各動画 n は、 $\lambda(n; \hat{\mathcal{G}}_K)$ を取り得る代表動画 $\hat{g}_k \in \hat{\mathcal{G}}_K$ のクラスター C_k に割り当てられるものとする。

今回、類似度 $\rho(\alpha, \beta)$ は動画ペアの評価値推移のコサイン類似度を定義する。観測期間 $[1, T]$ における、動画 n の観測日 $\tau \in \{1, \dots, T\}$ の評価値を $E(n, \tau)$ とし、この期間の n の評価値推移を T 次元ベクトルで

$$\mathbf{x}_n = \{E(n, 1), E(n, 2), \dots, E(n, T)\}^T, \quad (5)$$

と書き表せば、任意の動画ペア $\alpha, \beta \in \mathcal{N}$ 間のコサイン類似度は

$$\rho(\alpha, \beta) = \frac{\mathbf{x}_\alpha^T \mathbf{x}_\beta}{\|\mathbf{x}_\alpha\| \cdot \|\mathbf{x}_\beta\|}, \quad (6)$$

のように求めることができる。

図 5, 6 に従来評価値と提案評価値の推移クラスタリング結果 ($K = 10$) をそれぞれ示す。緑線は各クラスタリングの代表ベクトル $\hat{\mathcal{G}}_{10}$ であり、観測期間は各動画投稿日の 2 日後から 90 日間である。両図より、従来評価値よりも提案評価値の方が多種多様な変化をすることが見て取れるため、提案評価値は動画の動的な分析に幅を持たせることができると言える。

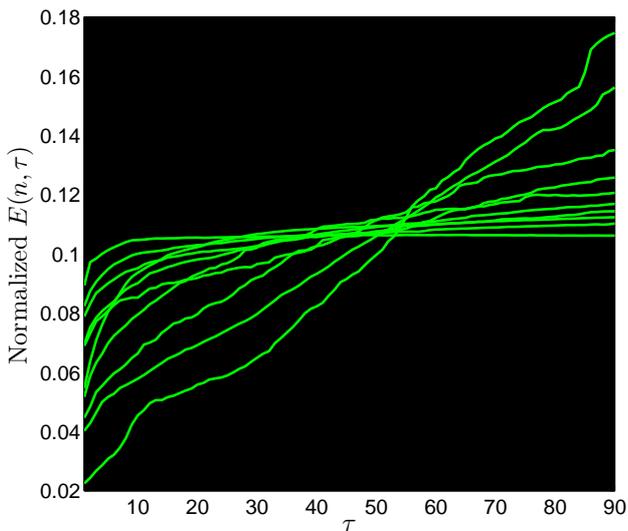


図 5: 従来評価値の推移クラスタリング ($K = 10$)

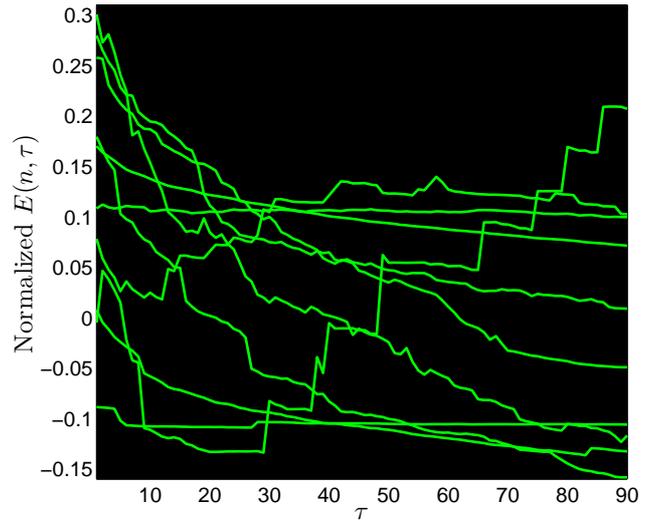


図 6: 提案評価値の推移クラスタリング ($K = 10$)

6. おわりに

再生数を信頼性として反映させたマイリスト率の Z -score を用いて、楽曲動画の話題性や集客能力の影響を受けにくいランキングを生成することを試みた。提案評価値によるランキングは、上位での再生数別出現頻度という点で、従来ランキングと明確な差別化ができていたことを示した。また、 K -medoids クラスタリング結果から、提案評価値は従来評価値よりも多種多様な推移をすることが分かった。今後は「同一ユーザによる再生数の加算」や「マイリスト解除」も考慮した厳密なモデルを追求していきたい。

謝辞

本研究は、科学研究費補助基金基盤研究 (C)(No.25330635) の支援を受けて行ったものである。

参考文献

- [1] M.J. Salganik, P.S. Dodds, and D.J. Watts, “Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market,” *Science*, vol.311, pp.854–856, 2006.
- [2] G.L. Nemhauser, L.A. Wolsey, and M.L. Fisher, “An analysis of approximations for maximizing submodular set functions - I,” *Mathematical Programming*, vol.14, pp.265–294, 1978.