

視聴者反応の時系列構造に基づく「刺さる」動画の定量化と早期予測

高橋 孝成(東京理科大学)^{†1}, 朝日 弓未(東京理科大学)²

キーワード: 動画プラットフォーム, コメント時系列, 期待値-残差モデル, ユーザエンゲージメント

1. はじめに

1.1. 背景と課題

現代のデジタル環境では、ユーザは膨大な選択肢の中でコンテンツを選択しており、単なる露出量だけではなく、記憶や感情に「刺さり」、自発的に語られ・保存されるコンテンツの重要性が高まっている。しかし、動画プラットフォームのヒット研究は、再生数や初動拡散といった量的指標に依存しており、「同じ露出条件の下で、なぜある動画は語られ続け、別の動画は忘れられるのか」という問いに十分に答えてこなかった。視聴規模の違いから独立して「刺さり」を測定する枠組みは未整備である。一方、消費者行動やマーケティングの研究では、動画コンテンツや動画に伴う能動的なユーザ行動（レビュー・ロコミ・評価・推奨）は、情報探索やブランド態度、購買意図といった意思決定プロセスに影響を与えることが報告されている。例えば、短尺動画コンテンツの有用性や楽しさが消費者の信頼や購買意図に正の影響を及ぼすことが示されているほか、動画広告と電子的ロコミ（eWOM）がブランド認知や購買意向に有意な影響を与えるという実証もある[1]。これらの知見は、視聴者の自発的な反応が単なる結果ではなく、消費行動やブランド評価の形成に関与する可能性を示しており、動画コンテンツの価値を再生数だけで測る限界を支持している

1.2. 提案

本研究は、こうした背景を踏まえ、ニコニコ動画における再生数と投稿経過時間から予測されるコメント数およびマイリスト数の期待値を回帰モデルにより推定し、その上振れ（期待度残差）を「エンゲージメント効率」として定義することで、「刺さる」動画を統計的に同定する枠組みを提案する。さらに、この指標を基準として、コメントの初動集中度、差し込み位置分布（vpos）、24 時間反応率といった時系列構造を分析し、視聴者に「刺さる」コンテンツに固有の反応パターンを明らかにする。本アプローチは、動画コンテンツそのものや視聴者のアクティブな反応が、単なる再生行動よりも深い消費者行動や評価形成に関わる

という先行研究[2]と整合しつつ、その構造を大規模ログから定量的に捉えることを目指すものである。

2. 方法

2.1. 期待値-残差法（Expectation-Residual）

動画のコメント数およびマイリスト数は、主として再生数と投稿からの経過時間によって決定される。そこで、各動画について

$$\log(1 + \text{comment}_i) = f(\log(1 + \text{watch}_i), \log(1 + \text{age}_i)) + \varepsilon_i,$$

$$\log(1 + \text{mylist}_i) = g(\log(1 + \text{watch}_i), \log(1 + \text{age}_i)) + \eta_i$$

を Huber 回帰により推定し ε_i, η_i を「期待からの逸脱（残差）」として解釈する。両残差を標準化し、

$$\text{score_resid}_i = z(\varepsilon_i) + z(\eta_i)$$

を定義することで、再生数や経過時間では説明できない「刺さり」を測定する指標とした。

分析は再生数上位 30% ($q_{\{70\}}$ 以上) の動画に限定し、**score_resid** の上位 5% を「刺さり動画 (True-top5%)」と定義した。対照群として、同一条件からランダムに抽出した動画群を用いた。

2.2. 視聴者行動の三軸（A, B, C）

刺さりを生む視聴者行動の構造を捉えるため、コメントログから以下の三軸を構成した。

A) （初動）

$$\text{share}_{24h} = \frac{\text{総コメント数投稿後}}{24\text{時間以内のコメント数}}$$

初動反応の集中度を表す。

B) （差し込み位置）

コメントの差し込み位置 **vpos** の Shannon エントロピーを計算し、値が小さいほど特定シーンへの集中が強いと解釈する。

C) （意味的一致性）

コメントを TF-IDF ベクトル化し、コサイン類似度の平均を算出することで、視聴者間で語られている内容の収束度を測定する。

刺さり群 (True-top5%) と対照群の分布差は、Mann-Whitney U 検定と Cliff's δ (順位ベースの効果量) により評価した。また、主要 5 ジャンル (anime, entertainment, game, music_sound, other) について、各ジャンル内で Top100 と

Random100 を抽出する準実験比較を行った。

2.3. 早期予測 (Early-stage Prediction)

上記とは独立に、別の動画ファイル群 (約 2 万本) を用いて、「投稿後 24 時間の視聴者行動だけで将来の刺さりをどれだけ予測できるか」を検証した。各動画について、

- share24h
- burst_30min_share (24 時間コメントのうち最初 30 分に集中した割合)
- vpos エントロピー

を計算し、それらの標準化和を

$$EarlyScore_i = z(\text{share24hi}) + z(\text{burst_30min_share}_i) + z(\text{vpos_entropy}_i)$$

として定義した。

この EarlyScore に基づき各動画を順位付けし、上位 5% を **Early-top5%** と定義する。これを、期待値-残差法により同じ母集団から定義された **True-top5%** と比較し、

- Precision@5%
- Recall@5%
- Jaccard 係数

を用いて一致度を評価した。さらに、True-top5% とそれ以外の動画の間で EarlyScore の分布差を Mann-Whitney U 検定および Cliff's δ により検証し、24 時間のコメント構造が長期的な刺さりをどの程度識別できるかを定量化した。

3. 結果

3.1. 「刺さり動画」に固有の視聴者行動

約 2 万本の動画メタデータとコメントログを用い、期待値-残差法により再生数および投稿経過時間を統制した上で「刺さり動画 (True-top5%)」を定義した。その結果、刺さり動画は単に再生数が多い動画とは異なり、視聴者のコメント行動に明確な構造的特徴を示した。三軸分析により、刺さり動画は

(A) 投稿後 24 時間のコメント集中度が高く、

(B) vpos の Shannon エントロピーが低く特定シーンに反応が集中し、

(C) コメント内容の意味的一貫性も高いことが確認された。これらの差はいずれも Mann-Whitney U 検定

(Wilcoxon 順位和検定) で有意であり、Cliff's δ も中程度以上の効果量を示した。すなわち、「刺さる」動画では視聴者反応が初動から時間的・意味的に強く組織化される。ジャンル別では、アニメ、エンターテインメント、ゲームでこの傾向が特に顕著であり、初期のコメント集中とシーン同期が長期的なエンゲージメント効率と強く結びついていた。一方、音楽ジャンルではこの差が相対的に弱く、後続の視聴や保存の蓄積によって評価が形成される「遅延型」の構造が示唆された。

3.2. 初動コメントによる刺さりの外部予測

3.3. 独立した約 2 万本の動画集合を用い、share24h,

burst_30min_share, vpos エントロピーから EarlyScore を構成し、上位 5% (Early-top5%) を抽出した。これを期待値-残差法による True-top5% と比較したところ、重複は 43 本となり、ランダム一致 (約 15 本) を大きく上回った (Precision@5% \approx 14.7%)。また、True-top5% と非 top5% の間で EarlyScore の分布は Mann-Whitney U 検定で強く分離され ($p = 2.7 \times 10^{-27}$)、Cliff's δ は約 0.37 であった。すなわち、投稿後 24 時間のコメント構造だけで、長期的な刺さりを有意かつ実質的に予測できる。ジャンル別でも EarlyScore はすべての主要ジャンルで True-top5% を有意に識別したが、その強さには差があった。アニメ、エンターテインメント、ゲームでは Precision@5% が約 15% と高く、初動反応がその後の刺さりを強く規定していた。一方、音楽では Precision@5% が約 7% にとどまり、初動よりも長期的な蓄積が重要であることが示された。

4. 考察

本研究は、「刺さる」動画を再生数や投稿時期を統制した上でのコメント・マイリストの上振れとして定義し、その視聴者行動構造を明らかにした。初動のコメント集中、シーン同期 (vpos)、および意味的一貫性が刺さりと結びつくことは、エンゲージメントが単なる量ではなく反応の組織化によって形成されることを示している。

さらに独立データによる検証から、投稿後 24 時間のコメント構造だけで将来のエンゲージメント効率を予測できることが確認された。

ジャンル別には、アニメ、エンターテインメント、ゲームでは初動の反応が刺さり強く規定する一方、音楽では反復視聴や時間をかけた蓄積によって評価が形成される傾向が見られ、「刺さり」の形成メカニズムがジャンルによって異なることが示唆された。

これらの知見は、広告やデジタルコンテンツ設計において、ジャンルごとのエンゲージメント特性に応じた戦略の必要性を示しており、今後はこうした差異を組み込んだ早期評価モデルがコンテンツ流通や広告最適化に貢献すると考えられる。

参考文献

- [1] Luo, C., Mohd Hasan, N. A., Ahmad, A. M. Z. b., & Lei, G. (2025). Influence of short video content on consumers' purchase intentions on social media platforms with trust as a mediator. Scientific Reports, 15, Article 16605. <https://www.nature.com/articles/s41598-025-94999-4>
- [2] Saleem, A., Mehmood, R., Taj, A., Khalid, M. U., Moiz, A., & Lakho, A. (2024). Impact of video content marketing on consumer engagement. Journal of Policy Research, 10(3), 83–95. <https://jprpk.com/index.php/jpr/article/view/627>