

E-016

ニコニコ動画におけるコメント分布の相関に着目した不正動画検出 Cheating Detection by Using Correlation of Comment Distribution in Nicovideo

清水 隼人†
Hayato Shimizu

松本 和幸†
Kazuyuki Matsumoto

吉田 稔†
Minoru Yoshida

北 研二†
Kenji Kita

1. はじめに

近年の情報処理技術およびネットワーク技術の発達に伴いスマートフォンやタブレット端末が普及し手軽に動画を視聴することが可能となった。動画共有サイトはプライベートの公開から企業の宣伝活動や最新のニュースなど個人・法人を問わず様々な用途で活用されてきている。しかし、この仕組みを悪用し不正に再生数を稼ぐための動画を作成するユーザーが増えていることも事実である。ニコニコ動画では、再生数(視聴したユーザ数)やマイリスト数(お気に入り登録したユーザの数)、コメント数(ユーザがコメントした数)により動画の評価を行なってランク付けされているが[1],これらの数はアカウントを多用することにより一時的に増やすことが可能である。ある程度の評価を得てランキング上位に浮上すれば、動画の注目度が上がり、どんな内容の動画であっても支持が強くなるバンドワゴン効果の問題が生じてしまう。そのためそういった数値を用いるのではなく新たな動画評価方法が必要となる。

本研究では、前述の問題により不正に評価が伸びている動画をコメントの時系列解析を用いて検出する手法を提案する。具体的には人気動画と不正動画のコメントの投稿分布の違いに着目した不正動画検出を行う。

2. コメント投稿日時とコメント投稿再生位置の関係

動画に対するコメント投稿日時を縦軸、コメント再生位置を横軸としてコメントの分布を表示させたグラフを図1に示す。

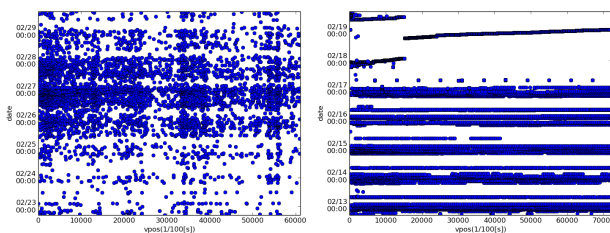


図1: 人気動画(左)と不正動画(右)

図1では人気動画のコメント分布を左図に、不正動画のコメント分布を右図に示す。2つのグラフを見比べればわかるように人気動画ではコメントが幅広く散らばっており空白が少ないのに対して不正動画ではコメントが横に固まっており空白が多いことがわかる。これらのコメント投稿日時とコメント再生位置の特性を考える。2.1, 2.2で述べる特徴を素に人気動画に含まれる不正動画を

検出するため、人気動画と不正動画の2つのデータを収集し分類する方法を考える。

2.1 コメント投稿日時による違い

コメント投稿日時による判別では図を見ればわかるように人気動画では1日ごとにコメントの量が徐々に増加し減少していくのに対して、不正動画ではコメント投稿数がある時間帯に多く集中し、1日ごとに多い時と少ない時がはっきり分かれている。

2.2 コメント投稿再生位置による違い

コメント投稿再生位置の関係では人気動画では再生位置によってコメントの多いところ少ないところが定まっているのに対して、不正動画ではコメント再生位置とコメントの量に関連性を見出せない。

3. 提案手法

3.1 時系列データによる判別

本研究では1日おきにどのようなコメントの変動があるのかを調べる。これには、再生時間を一定の区間に区切り、その区間のコメント数を1日おきに集計し1日毎の相関係数を計算することで不正動画を判別することを考える。

3.2 相関係数を用いた不正動画の検出方法

不正動画検出に必要な7日分の相関係数 $r_{x,x+1}$ ($x = 1, 2, 3, \dots, 6$) の求め方は以下の手順で行う。

1. x 日の再生位置によるコメント数を得るために図2のように再生時間を N 等分しそれぞれ四角の領域 (x, n) ($n = 1, 2, 3, \dots, N$) のコメント数 $C_{x,n}$ を集計する(本研究では $N=10$ で相関係数を求める)。
2. 1日ごとの情報を得るために $C_{x,n}$ の平均値を $\overline{C_x}$ とし、式(1)から $C_{x,n}$ と $C_{x+1,n}$ からなる相関係数 $r_{x,x+1}$ を求める。
3. 動画投稿後7日分の計6つの相関係数 $r_{x,x+1}$ を計算し、人気動画と不正動画の相関係数の違いを見つける。

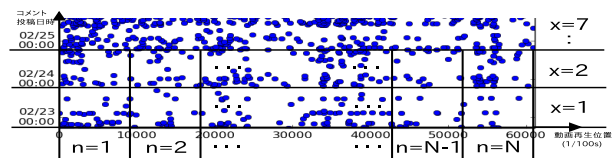


図2: 領域の分割方法

$$r_{x,x+1} = \frac{\frac{1}{N} \sum_{n=1}^N (C_{x,n} - \overline{C_x})(C_{x+1,n} - \overline{C_{x+1}})}{\sqrt{\frac{1}{N} \sum_{n=1}^N (C_{x,n} - \overline{C_x})^2} \sqrt{\frac{1}{N} \sum_{n=1}^N (C_{x+1,n} - \overline{C_{x+1}})^2}} \quad (1)$$

†徳島大学, Tokushima University

4. コメント内容に基づく手法

ニコニコ動画コメント内容から単語の重要性に基づく不正動画検出方法を比較手法として用いる。動画コメントの単語の重要性を扱っている研究として江端ら [2] の tf-idf 法を用いた分析がある。本研究もそれに従いニコニコ動画におけるコメントデータから tf-idf 法を用いて調査データから一つの動画のコメント文の集合を特徴ベクトルとして表現し、不正動画の特徴語とする。その特徴ベクトルからサポートベクターマシン (SVM) を用いて判別を行う。数式に tf-idf の計算式を示す。\$d\$ を一つの動画すべてのコメント文、\$t\$ を文章 \$d\$ 内に含まれる単語 \$n_{t,d}\$ をある単語 \$t\$ の文章 \$d\$ 内での出現回数、\$N\$ を総動画数、\$df(t)\$ をある単語 \$t\$ が出現する文章の数とすると以下の式より tf-idf の値 \$w(t,d)\$ と動画の特徴ベクトル \$\vec{m}_d\$ の計算を行い、これを SVM により学習させた二値分類器を用いて判別を行う。

$$tf(t,d) = \frac{n_{t,d}}{\sum_{s \in d} n_{s,d}} \quad (2)$$

$$idf(t) = \log\left(\frac{N}{df(t)}\right) + 1 \quad (3)$$

$$tfidf(t,d) = tf(t,d) * idf(t) \quad (4)$$

$$\vec{m}_d = (w(t_1,d), w(t_2,d), \dots, w(t_N,d)) \quad (5)$$

5. 実験データの抽出

表 1: 実験に使用したデータ

	訓練データ	テストデータ
人気動画	100	65
不正動画	100	65

今回実験に使用したデータとその抽出方法を説明する。使用した動画数を表 1 に示す。人気動画の抽出ではニコニコ動画公式より動画ランキング上位 50 の動画から抽出した。不正動画の抽出では以下の方法を用いた。

- 再生数 < コメント数
- コメント数 ≥ 10000
- マイリスト数 ≤ 1000

この条件より大まかに動画を抽出し、さらに不正動画の条件として動画に映像も音も含まれない、同じシーンしか映らない、または同じ音声の繰り返しであるといった条件に合致するか否かを目視で確認し抽出する。

6. 実験と考察

訓練データの人気動画不正動画から相関係数を計算し、それぞれの全体平均、最大値平均、最小値平均を表 2 に示す。最大値平均、最小値平均とは動画一つ一つの相関係数の最大値と最小値を抽出し、それらを平均した値である。この表を見てわかるとおり人気動画と不正動画の最小値が大きく違うことがわかる。この最小値を特徴として用いてテストデータの判別を行った結果が表 3 である。

結果から相関係数の最小値の閾値が 0.6 から 0.7 の間では不正動画は 10% 未満で検出されず人気動画では 90% 以上で検出されているのがわかる。またコメント内容に基

づく判別手法と比べて相関係数を用いたほうが不正動画検出率が高くなることを確認できた。

提案手法の実験結果において、相関係数の最小値の閾値が 0.6 の時に不正動画が 4.2% 検出されていなかった。この理由として、相関係数が全体的には高くないが低い値が続いている場合、例えばすべて 0.60 より大きく 0.65 より小さい場合は不正動画の疑いがあるが、提案手法では検出されない。

また、人気動画が誤検出される理由として、ニコニコ動画ランキングから人気動画を抽出したがランキングにランクインした一部の動画には不正な動画も含まれていることがあげられる。

表 2: 相関係数の平均

動画の種類	全体平均	最大値平均	最小値平均
人気動画	0.979	0.993	0.848
不正動画	0.778	0.972	0.496

表 3: 不正動画の検出率と誤検出の割合

	閾値	不正動画の検出率	誤検出された人気動画の割合
相関係数に基づく判別	0.5 以下	84.8%	0%
	0.6 "	93.8%	0%
	0.7 "	98.5%	1.5%
	0.8 "	100%	10.7%
コメント内容に基づく判別		85.6%	0%

7. おわりに

本研究では、動画のコメント情報を用いて不正動画を検出する方法を提案した。評価実験の結果、提案手法により、コメント数の大小に関係なくコメントデータの時系列解析から動画の特徴を読み取り不正動画を検出することが可能なことが分かった。

今後の課題として動画の解析を行う際に式 (1) で用いる \$N\$ の値を 10 として定めたが、相関係数の値は \$N\$ の値に大きく影響されるため明確な基準により求める必要がある。また実験では動画の判別基準としてコメントの時系列解析を行ったが、その他の素性を用いることで、判別精度を向上させることができるかを検討したい。

謝辞

本研究の一部は、科学研究費補助金 (基盤研究 (C)15K00425, 若手研究 (B)15K16077) の補助を受けて行った。

参考文献

- [1] ニコニコ大百科「ランキング」(2015年7月20日参照) <http://dic.nicovideo.jp/a/ランキング>
- [2] 江端 祐介, 川村 秀憲, ユーザコメントの tf-idf 法による分析を用いたインタラクティブな関連動画の提示, 電子情報通信学会, Vol.109, No.439, pp.7-10 (2010-02)
- [3] Nii 国立情報学研究所 ニコニコ動画データセット <http://www.nii.ac.jp/cscenter/idr/nico/nico.html>