Contextual Feature Based One-Class Classifier Approach for Detecting Video Response Spam on YouTube

Vidushi Chaudhary
MTech (IS) 2011-2013 (vidushi1116@iiitd.ac.in)

Ashish Sureka
MTech Thesis Adviser (ashish@iiitd.ac.in)

PST 2013 Presentation : 11/07/2013
Presentation Outline

• Research Motivation & Aim
• Related work & Research Contributions
• Methodology
  – Empirical Analysis
  – Proposed Solution Approach
• Performance Evaluation & Results
• Limitations
• Conclusion and Future Work
YouTube

- Over 800 million unique users visit YouTube each month
- Over 4 billion hours of video are watched each month on YouTube
- 72 hours of video are uploaded to YouTube every minute
- 100 million people take a social action on YouTube (likes, shares, comments, etc) every week
- Millions of videos are favorited every day

Statistics

[http://www.youtube.com/t/press_statistics]
Video Response and Related Video

Nursery Rhymes - Baa Baa Black Sheep

Video Responses

Related Videos
General Definition:
• Spam is an irrelevant or unsolicited messages sent over the internet, typically to a large number of users, for the purpose of advertising, phishing, spreading malware etc.

Ref: [http://oxforddictionaries.com/definition/english/spam](http://oxforddictionaries.com/definition/english/spam)

Disadvantages:
• Spam is annoying
• Damage company’s reputation
• Bandwidth wastage
• Time wastage

In our problem context
• Spam is the irrelevant, unwanted, unsolicited video posted as a video response to a YouTube video.

Types of Spam
- Spam Video
- Pornographic Video
- Promotional Video
- Botnet Video
Examples

**Cartoon gaming video for children**

Happy Wheels - STAR WARS
by TobyGames • 587,076 views

**Porn video response**

**VIDEO RESPONSES**

**happy wheels 2**
by sargent nick • 29 views
Happy.

**How to Grow on YouTube | The Best...**
by CoffeeCupGamer • 5 views
This is just a simple video telling you, the...

**Super Mario, opening with lyrics - A...**
by sonajo • 1,122 views
No aito da montanha tem um bosque Onde tem...

**happy wheels Santa does a flip!**
by sargent nick • 24 views
happy wheels

**Left 4 Dead 2 montage**
by duffosaur games • 4 views
A montage of Left 4 dead

**chris smoove inspired me to make c...**
by SKxALABAMA • 33 views
chris smoove thats a noob move

**Death of the Dream 2 - IT HUNGERS...**
by Ilovebanannas • 35 views
Thanks for watching and I hope you enjoyed! Be...

**FREE PORN VIDEO (18+!) I WATCH N...**
by XxDimkon96Xx • 43,425 views
Hello guys! This is XxDimkon96XX! Today I am...

**Happy Wheels Let's get throwin**
by VideoJaybird Fish • 34 views
alot of throwng levels if you don't like those...

**Im A Robot!?- Cave Story Part 6 - Th...**
by TheJuicyGameCorner • 42 views
Oh yea the story is started to get good. This...
Examples

**Educational dance video**

Arabic Belly Dance Basic Moves Part 01 of 04 by Zizzy Raciq - 8,256,162 views
Learn Arabic Style Belly dance in the easy steps.

**Pornographic video**

SEX ON THE BEACH - sexo en la playa by ingridott - 12,181 views
sex on the beach is a caribbean drink sexo en la...

VIDEO RESPONSES

- **belly dance improvisation - Ho Lan** by Ho Lan - 3,431 views
  hello! Now an improve belly dancing with song...

- Shahedah Belly Dancing to Alihan Sa...
  by shahedah Ab - 277 views
  https://www.facebook.com/ShahedahDesignDanfr...

- **My New Belly Dancing Solo to an Ara...**
  by TheSammichLove - 1,109 views
  New outfit, New Dance..If you leave some...

- **Shahedah Bellydancing to----Alihan ...**
  by shahedah Ab - 684 views
  Shahedah Bellydancing to----Alihan Samedov...
Examples

Most viewed music video

Commercial, Botnet Video

VIDEO RESPONSES

Party Rock Anthem - LMFAO ft. Lauren Bennett, GoonRock
by LMFAOVEVO - 527,567,552 views
Buy now http://glnk.it/6I. Music video by LMFAO performing Party Rock Anthem featuring Lauren Bennett and GoonRock. (c) 2011 Interscope VEVOCertified on July 1, 2011. http://www.vevo.com/certified...

$5 For FREE!
by Joanna T - 635 views
Everyday I'm shuffling... With a harp? P.S....

$5 For FREE!
by Jesus Christos - 4 views
http://goo.gl/5x8HK CHECK IT OUT...

$5 For FREE!
by Jesus Christos - No views
http://goo.gl/5x8HK CHECK IT OUT...

$5 For FREE!
by Jesus Christos - No views
http://goo.gl/5x8HK CHECK IT OUT...

Local Finding Millions!
by Jesus Christos - No views
http://goo.gl/5x8HK CHECK IT OUT...

RUN!!!!!!1
by Dillon Griswold - 73 views
Told You

$5 For FREE!
by Jesus Christos - No views
http://goo.gl/5x8HK CHECK IT OUT...

Local Finding Millions!
by Jesus Christos - No views
http://goo.gl/5x8HK CHECK IT OUT...
## Research Aim

### Broad Objective

- Increase understanding of video response spam on YouTube.
- Investigate effective solution to combat the video response spam problem.

### Specific Objective

- Characterization study on real world dataset.
- Examine one class classifier for recognizing spam video responses.

### Advantages:

- Less bandwidth wastage
- No time wastage
- Better User experience
- No compromise on system reputation
**Related work**

- **3 Line of Research:**
  1. Video Response Interactions and Video Response Spam.
     - Fabricio et. al – Characterization of video based interaction from video response feature
     - Fabricio et. al – Finding video spammers by analyzing their social behavior.
     - Fabricio et. al – Finding Video spammers and content promoters on YouTube
  2. Social media Spam detection.
     - Sureka – Detect comment spammers on YouTube forum by mining comment feed.
     - Benjamin et al. – Study of automatic detection of spammers in a social system.
     - Christian et. al – Detect porn videos by analyzing image features and motion information.
  3. Classification of YouTube videos based on contextual features.
     - Yuan et. Al – Propose contextual based analysis to automatically detect Forum spamming [6].
Research Contributions

✓ First step in the direction of applying a one-class classifier based approach using contextual features to detect video response spam on YouTube.

✓ Presented the intuition behind each discriminatory feature and an empirical analysis demonstrating its influence or impact on the classification task.
Research Framework

1. Positive Class Training Dataset
   - PV
   - BV
   - CV
   - Visual Inspection
   - Empirical Analysis

2. Characterization and Training
   - PV
   - BV
   - CV
   - Discriminatory Features
     - Linguistic, Contextual Popularity
     - Temporal, Social Networking

3. One Class Classifier
   - New Video
   - Distance or Resemblance
   - Class Description
   - Verification or Recognition

4. Performance Evaluation
   - Target Class
   - Unknown
   - Classifier Accuracy
   - Feature Influence
Research Framework

1. Positive Class Training Dataset
   - PV
   - BV
   - CV
   - Visual Inspection
   - Empirical Analysis

2. Characterization and Training
   - PV
   - BV
   - CV
   - Discriminatory Features
     - Linguistic, Contextual Popularity
     - Temporal, Social Networking

3. One Class Classifier
   - New Video
   - Class Description
   - Distance or Resemblance

4. Performance Evaluation
   - Target Class
   - Unknown
   - Classifier Accuracy
   - Feature Influence

- Methodology
Positive Class Training Dataset

**Input:** A list $L$ of most popular and discussed videos

**Result:** Experimental dataset of video response spam initialization;

```plaintext
for each video $V$ in $L$ do
    if $V$ is a responded video then
        fetch all video response of $V$;
        forall the video response $V_r$ do
            Add $V_r$ in the experimental dataset; 
            manually assign label to each video;
        end
    end
    if $V$ is a responsive video then
        Add $V$ in the experimental dataset;
        Manually assign label to the video;
    end
end
```

<table>
<thead>
<tr>
<th></th>
<th>Training Dataset</th>
<th>Testing Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>PVRD</td>
<td>250</td>
<td>10018</td>
</tr>
<tr>
<td>CVRD</td>
<td>250</td>
<td>9256</td>
</tr>
<tr>
<td>BVRD</td>
<td>61</td>
<td>3389</td>
</tr>
</tbody>
</table>
Research Framework

Positive Class Training Dataset
- PV
- BV
- CV

1. Visual Inspection
   Empirical Analysis

Characterization and Training

2. Discriminatory Features
   Linguistic, Contextual Popularity
   Temporal, Social Networking

One Class Classifier
- New Video
- Class Description

3. Distance or Resemblance

Verification or Recognition

Performance Evaluation

4. Classifier Accuracy
   Feature Influence

- Target Class
- Unknown
Research Framework

Positive Class Training Dataset

- PV
- BV
- CV

Visual Inspection
Empirical Analysis

Characterization and Training

- PV
- BV
- CV

Discriminatory Features
Linguistic, Contextual Popularity
Temporal, Social Networking

One Class Classifier

New Video

Class Description

Distance or Resemblance

Verification or Recognition

Performance Evaluation

- Target Class
- Unknown

Classifier Accuracy
Feature Influence

Precision & Recall

16
• In one class classification problem, either negative class is not present or it is not properly sampled.
Research Framework

1. Positive Class Training Dataset
   - PV
   - BV
   - CV
   - Visual Inspection
   - Empirical Analysis

2. Characterization and Training
   - PV
   - BV
   - CV
   - Discriminatory Features
     - Linguistic, Contextual Popularity
     - Temporal, Social Networking

3. Verification or Recognition
   - New Video
   - Class Description
   - Distance or Resemblance

4. Performance Evaluation
   - Classifier Accuracy
   - Feature Influence
   - Target Class
   - Unknown
## Characterization: PVRD

### Pornographic Video Response Detection

<table>
<thead>
<tr>
<th>Abbr.</th>
<th>Feature Title</th>
<th>Feature Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPTT</td>
<td>Percentage of pornographic terms in title</td>
<td>Linguistic, Numerical</td>
</tr>
<tr>
<td>PPTD</td>
<td>Percentage of pornographic terms in description</td>
<td>Linguistic, Numerical</td>
</tr>
<tr>
<td>CatV</td>
<td>YouTube category of the video</td>
<td>YouTube Basic, Categorical</td>
</tr>
<tr>
<td>WTRL</td>
<td>Web of Trust rating of the links</td>
<td>Trust, Categorical</td>
</tr>
<tr>
<td>RSBV</td>
<td>Ratio of subscriber by view</td>
<td>Temporal + Popularity, Numerical</td>
</tr>
<tr>
<td>RLBV</td>
<td>Ratio of likes by view</td>
<td>Temporal + Popularity, Numerical</td>
</tr>
<tr>
<td>DYTV</td>
<td>Duration of YouTube video</td>
<td>Time based, Numerical</td>
</tr>
<tr>
<td>PVARF</td>
<td>Percentage of videos with age-restricted flag</td>
<td>Trust, Categorical</td>
</tr>
</tbody>
</table>
## Characterization: PVRD

### Pornographic Video Response Detection

<table>
<thead>
<tr>
<th>Abbr.</th>
<th>Feature Title</th>
<th>Feature Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPTT</td>
<td>Percentage of pornographic terms in title</td>
<td>Linguistic, Numerical</td>
</tr>
<tr>
<td>PPTD</td>
<td>Percentage of pornographic terms in description</td>
<td>Linguistic, Numerical</td>
</tr>
<tr>
<td>CatV</td>
<td>YouTube category of the video</td>
<td>YouTube Basic, Categorical</td>
</tr>
<tr>
<td>DYTV</td>
<td>Duration of YouTube video</td>
<td>Time based, Numerical</td>
</tr>
<tr>
<td>RSBV</td>
<td>Ratio of subscriber by view</td>
<td>Temporal + Popularity, Numerical</td>
</tr>
<tr>
<td>RLBV</td>
<td>Ratio of likes by view</td>
<td>Temporal + Popularity, Numerical</td>
</tr>
<tr>
<td>PVARF</td>
<td>Percentage of videos with age-restricted flag</td>
<td>Trust, Categorical</td>
</tr>
<tr>
<td>WTRL</td>
<td>Web of Trust rating of the links</td>
<td>Trust, Categorical</td>
</tr>
</tbody>
</table>
Linguistic Features

- Empirical Analysis

<table>
<thead>
<tr>
<th>Percentage of Pornographic Terms in Title</th>
<th>Percentage of Pornographic Terms in Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>32.8% videos</td>
<td>No pornographic terms present</td>
</tr>
<tr>
<td>55% videos</td>
<td>Greater than 50% porn terms</td>
</tr>
<tr>
<td>6.4% videos</td>
<td>More than 80% porn terms</td>
</tr>
<tr>
<td>53.6% videos</td>
<td>No pornographic terms present</td>
</tr>
<tr>
<td>22% videos</td>
<td>Greater than 20% porn terms</td>
</tr>
<tr>
<td>0.4% videos</td>
<td>More than 80% porn terms</td>
</tr>
</tbody>
</table>
## Characterization: PVRD

### Pornographic Video Response Detection

<table>
<thead>
<tr>
<th>Abbr.</th>
<th>Feature Title</th>
<th>Feature Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPTT</td>
<td>Percentage of pornographic terms in title</td>
<td>Linguistic, Numerical</td>
</tr>
<tr>
<td>PPTD</td>
<td>Percentage of pornographic terms in description</td>
<td>Linguistic, Numerical</td>
</tr>
<tr>
<td>CatV</td>
<td>YouTube category of the video</td>
<td>YouTube Basic, Categorical</td>
</tr>
<tr>
<td>DYTV</td>
<td>Duration of YouTube video</td>
<td>Time based, Numerical</td>
</tr>
<tr>
<td>RSBV</td>
<td>Ratio of subscriber by view</td>
<td>Temporal + Popularity, Numerical</td>
</tr>
<tr>
<td>RLBV</td>
<td>Ratio of likes by view</td>
<td>Temporal + Popularity, Numerical</td>
</tr>
<tr>
<td>PVARF</td>
<td>Percentage of videos with age-restricted flag</td>
<td>Trust, Categorical</td>
</tr>
<tr>
<td>WTRL</td>
<td>Web of Trust rating of the links</td>
<td>Trust, Categorical</td>
</tr>
</tbody>
</table>
YouTube basic features

- **Entertainment**: 48% videos
- **People and Blogs**: 28% videos
- **Travel & Events, Auto**: 0% videos

**Empirical Analysis**

- **Less than 50 seconds duration**: 31% videos
- **Less than 100 seconds duration**: 48% videos
# Characterization: PVRD

## Pornographic Video Response Detection

<table>
<thead>
<tr>
<th>Abbr.</th>
<th>Feature Title</th>
<th>Feature Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPTT</td>
<td>Percentage of pornographic terms in title</td>
<td>Linguistic, Numerical</td>
</tr>
<tr>
<td>PPTD</td>
<td>Percentage of pornographic terms in description</td>
<td>Linguistic, Numerical</td>
</tr>
<tr>
<td>CatV</td>
<td>YouTube category of the video</td>
<td>YouTube Basic, Categorical</td>
</tr>
<tr>
<td>DYTV</td>
<td>Duration of YouTube video</td>
<td>Time based, Numerical</td>
</tr>
<tr>
<td>RSBV</td>
<td>Ratio of subscriber by view</td>
<td>Temporal + Popularity, Numerical</td>
</tr>
<tr>
<td>RLBV</td>
<td>Ratio of likes by view</td>
<td>Temporal + Popularity, Numerical</td>
</tr>
<tr>
<td>PVARF</td>
<td>Percentage of videos with age-restricted flag</td>
<td>Trust, Categorical</td>
</tr>
<tr>
<td>WTRL</td>
<td>Web of Trust rating of the links</td>
<td>Trust, Categorical</td>
</tr>
</tbody>
</table>
Temporal + Popularity based

- Empirical Analysis

<table>
<thead>
<tr>
<th>Videos</th>
<th>Subscriber by view ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>92.4%</td>
<td>less than 0.01</td>
</tr>
<tr>
<td>77%</td>
<td>less than 0.001</td>
</tr>
<tr>
<td>3%</td>
<td>No subscriber</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Videos</th>
<th>Likes by view ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>78%</td>
<td>less than 0.001</td>
</tr>
<tr>
<td>1.6%</td>
<td>Likes is 0.001</td>
</tr>
</tbody>
</table>
## Characterization: PVRD

<table>
<thead>
<tr>
<th>Abbr.</th>
<th>Feature Title</th>
<th>Feature Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPTT</td>
<td>Percentage of pornographic terms in title</td>
<td>Linguistic</td>
</tr>
<tr>
<td>PPTD</td>
<td>Percentage of pornographic terms in description</td>
<td>Linguistic</td>
</tr>
<tr>
<td>CatV</td>
<td>YouTube category of the video</td>
<td>YouTube Basic</td>
</tr>
<tr>
<td>DYTV</td>
<td>Duration of YouTube video</td>
<td>Time based</td>
</tr>
<tr>
<td>RSBV</td>
<td>Ratio of subscriber by view</td>
<td>Temporal + Popularity</td>
</tr>
<tr>
<td>RLBV</td>
<td>Ratio of likes by view</td>
<td>Temporal + Popularity</td>
</tr>
<tr>
<td>PVARF</td>
<td>Percentage of videos with age-restricted flag</td>
<td>Trust</td>
</tr>
<tr>
<td>WTRL</td>
<td>Web of Trust rating of the links</td>
<td>Trust</td>
</tr>
</tbody>
</table>
Trust features

- Empirical Analysis

**91% videos** | Marked as age-restricted video
---|---
**9% videos** | Not age-restricted

**94.8% videos** | No dirty link present in description
---|---
**5.2% videos** | No dirty link present
BotNet Video Response Detection

Constant duration of the uploaded videos

Don’t buy Blackberry before you watch this video
by medoselhou - 2 years ago - 320 views
Torch today! Don’t buy Blackberry before you watch this video...

how to hack walmart gift cards-Free $1000 WalMart Gift Card
by medoselhou - 2 years ago - 1,734 views
Card how to hack walmart gift cards-Free $1000 WalMart Gift Card...

walmart clown commercial-Free $1000 WalMart Gift Card
by medoselhou - 2 years ago - 129 views
Gift Card walmart clown commercial-Free $1000 WalMart Gift Card...

review blackberry torch - Free BlackBerry Torch
by medoselhou - 2 years ago - 89 views
Torch today! review blackberry torch - Free BlackBerry Torch...

Format of writing description is same for all uploaded videos
## Characterization: Botnet

<table>
<thead>
<tr>
<th>Botnet Video Response Detection</th>
<th>TDUV: Time difference between uploaded videos</th>
<th>Time based</th>
<th>Numerical</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDUV: Constant duration of uploaded videos</td>
<td>Time based</td>
<td>Numerical</td>
<td></td>
</tr>
<tr>
<td>NSUB: Number of subscribers of the user</td>
<td>Temporal + Popularity</td>
<td>Numerical</td>
<td></td>
</tr>
<tr>
<td>NCYV: Number of comments of the YouTube video</td>
<td>Temporal + Popularity</td>
<td>Numerical</td>
<td></td>
</tr>
</tbody>
</table>
## Characterization: Botnet

<table>
<thead>
<tr>
<th>Botnet Video Response Detection</th>
<th>Time difference between uploaded videos</th>
<th>Time based</th>
<th>Numerical</th>
</tr>
</thead>
<tbody>
<tr>
<td>TDUV</td>
<td>Time difference between uploaded videos</td>
<td>Time based</td>
<td>Numerical</td>
</tr>
<tr>
<td>CDUV</td>
<td>Constant duration of uploaded videos</td>
<td>Time based</td>
<td>Numerical</td>
</tr>
<tr>
<td>NSUB</td>
<td>Number of subscribers of the user</td>
<td>Temporal + Popularity</td>
<td>Numerical</td>
</tr>
<tr>
<td>NCYV</td>
<td>Number of comments of the YouTube video</td>
<td>Temporal + Popularity</td>
<td>Numerical</td>
</tr>
</tbody>
</table>
Time Based Features: TDUV

Empirical Analysis
Time Based Features: TDUV

<table>
<thead>
<tr>
<th>25 videos</th>
<th>Time interval 10 seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 videos</td>
<td>0 second time interval</td>
</tr>
</tbody>
</table>
Time Based Features: CDUV

Empirical Analysis

![Graph showing duration vs. number of videos](image)
Time Based Features: CDUV

42% videos 30 seconds in duration
68% videos Range between 20-40 seconds
# Characterization: Botnet

## Botnet Video Response Detection

<table>
<thead>
<tr>
<th>Metric</th>
<th>Description</th>
<th>Time Basis</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>TDUV</td>
<td>Time difference between uploaded videos</td>
<td>Time based</td>
<td>Numerical</td>
</tr>
<tr>
<td>CDUV</td>
<td>Constant duration of uploaded videos</td>
<td>Time based</td>
<td>Numerical</td>
</tr>
<tr>
<td>NSUB</td>
<td>Number of subscribers of the user</td>
<td>Temporal + Popularity</td>
<td>Numerical</td>
</tr>
<tr>
<td>NCYV</td>
<td>Number of comments of the YouTube video</td>
<td>Temporal + Popularity</td>
<td>Numerical</td>
</tr>
</tbody>
</table>
Time Based Features

- 24% videos: No comment present
- 47% videos: Less than 5 comments
- 4.6% videos: No subscriber
- 77% videos: RSBV value less than 0.002
# Characterization: CVRD

## Commercial Video Response Detection

<table>
<thead>
<tr>
<th>Abbr.</th>
<th>Feature Title</th>
<th>Feature Type</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PCTT</td>
<td>Percentage of commercial terms in title</td>
<td>Linguistic</td>
<td>Numerical</td>
</tr>
<tr>
<td>PCTD</td>
<td>Percentage of commercial terms in description</td>
<td>Linguistic</td>
<td>Numerical</td>
</tr>
<tr>
<td>NWLD</td>
<td>Number of web links present in description</td>
<td>Trust</td>
<td>Numerical</td>
</tr>
<tr>
<td>DYTV</td>
<td>Duration of YouTube video</td>
<td>Time based</td>
<td>Numerical</td>
</tr>
<tr>
<td>RSBV</td>
<td>Ratio of subscriber by view</td>
<td>Temporal + Popularity</td>
<td>Numerical</td>
</tr>
<tr>
<td>NCYV</td>
<td>Number of comments of the YouTube video</td>
<td>Temporal + Popularity</td>
<td>Numerical</td>
</tr>
</tbody>
</table>
Linguistic Features: PCTT

<table>
<thead>
<tr>
<th>Percentage of Commercial Terms in Title</th>
<th>Percentage of Videos</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-20%</td>
<td>65% videos</td>
</tr>
<tr>
<td>20-40%</td>
<td>No commercial term present in title</td>
</tr>
<tr>
<td>40-60%</td>
<td>10% videos</td>
</tr>
<tr>
<td>60-80%</td>
<td>More than 40% commercial terms present</td>
</tr>
<tr>
<td>80-100%</td>
<td></td>
</tr>
</tbody>
</table>
Linguistic Features: PCTD

- 16% videos: No commercial term present
- 72% videos: Some commercial terms present in description
- 12% videos: More than 80% commercial term present
## Characterization: CVRD

<table>
<thead>
<tr>
<th>Abbr.</th>
<th>Feature Title</th>
<th>Feature Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCTT</td>
<td>Percentage of commercial terms in title</td>
<td>Linguistic</td>
</tr>
<tr>
<td>PCTD</td>
<td>Percentage of commercial terms in description</td>
<td>Linguistic</td>
</tr>
<tr>
<td>NWLD</td>
<td>Number of web links present in description</td>
<td>Trust</td>
</tr>
<tr>
<td>DYTV</td>
<td>Duration of YouTube video</td>
<td>Time based</td>
</tr>
<tr>
<td>RSBV</td>
<td>Ratio of subscriber by view</td>
<td>Temporal + Popularity</td>
</tr>
<tr>
<td>NCYV</td>
<td>Number of comments of the YouTube video</td>
<td>Temporal + Popularity</td>
</tr>
</tbody>
</table>
Trust & Time Based Features: NWLD

- Empirical Analysis

<table>
<thead>
<tr>
<th>Percentage of Videos</th>
<th>Number of Web Links Present in Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>88% videos</td>
<td>0-5</td>
</tr>
<tr>
<td>27% videos</td>
<td>More than 5 web links</td>
</tr>
</tbody>
</table>

88% videos have web links present in the description. 27% of videos have more than 5 web links.
1. Around 23% videos are of less than 50 seconds in duration.
# Characterization: CVRD

## Commercial Video Response Detection

<table>
<thead>
<tr>
<th>Abbr.</th>
<th>Feature Title</th>
<th>Feature Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCTT</td>
<td>Percentage of commercial terms in title</td>
<td>Linguistic, Numerical</td>
</tr>
<tr>
<td>PCTD</td>
<td>Percentage of commercial terms in description</td>
<td>Linguistic, Numerical</td>
</tr>
<tr>
<td>NWLD</td>
<td>Number of web links present in description</td>
<td>Trust, Numerical</td>
</tr>
<tr>
<td>DYTV</td>
<td>Duration of YouTube video</td>
<td>Time based, Numerical</td>
</tr>
<tr>
<td>NCYV</td>
<td>Number of comments of the YouTube video</td>
<td>Temporal + Popularity, Numerical</td>
</tr>
<tr>
<td>RSBV</td>
<td>Ratio of subscriber by view</td>
<td>Temporal + Popularity, Numerical</td>
</tr>
</tbody>
</table>
Temporal + Popularity based features

1. 24% commercial videos does not contain any comment.
2. 56% videos contain less than 5 comment.

1. Around 4% videos does not have any subscriber.
2. 77% videos have RSBV value less than 0.002.
Proposed Solution Approach

- Video response spam detection problem is divided into 3 Sub-Problems:
  - Pornographic Video Detection.
  - Botnet Video Detection.
  - Promotional Video Detection.
Classification Approach

Where $t$ is the threshold used to separate a spam video from a legitimate video.

$$C_{value} = \sum_{i=0}^{n} W_i \times S_i , \ 0 \leq C_{value} \leq 1$$
Weight Computation

Assign equal weights to each feature s.t. \( \sum w_i = 1, 0 < i < n \)

Run the classifier and check the accuracy of the system.

Remove one feature \( (i) \) and adjust weights of other s.t. \( \sum w_j = 1, j \neq i \)

Run the classifier and check the accuracy of the system.

If change in accuracy < threshold

- NO
  - Removed feature is not important feature
  - \( \text{Featureweight} = \text{featureweight} + \delta \)

  - YES
  - Removed feature is an important feature
  - \( \text{Featureweight} = \text{featureweight} - \delta \)
Effect of Individual feature on the accuracy of the system

- Proposed Solution Approach

<table>
<thead>
<tr>
<th>Feature</th>
<th>PPTT</th>
<th>PPTD</th>
<th>CatV</th>
<th>WTRL</th>
<th>RSBV</th>
<th>RLBV</th>
<th>DYTV</th>
<th>SSAR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.16</td>
<td>0.2</td>
<td>0.12</td>
<td>0.2</td>
<td>0.09</td>
<td>0.08</td>
<td>0.1</td>
<td>0.05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Feature</th>
<th>TDUV</th>
<th>NSUB</th>
<th>NCVY</th>
<th>CDUV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.05</td>
<td>0.4</td>
<td>0.35</td>
<td>0.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Feature</th>
<th>PCTT</th>
<th>PCTD</th>
<th>NWLD</th>
<th>RSBV</th>
<th>DYTV</th>
<th>NCVY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.15</td>
<td>0.05</td>
<td>0.075</td>
<td>0.325</td>
<td>0.15</td>
<td>0.25</td>
</tr>
</tbody>
</table>
Score Computation

• Score of the feature is a unique value which represents that feature in comparison to the training dataset.

• Let $S_i = \text{Score of feature } i \text{ s.t. } 0 \leq s_i \leq 1$

• For numerical features, average difference of feature value with training dataset is calculated.

• Lower the difference, higher the chance that the new object is similar to the existing dataset.

$$\text{Score}_{nf} = \sum_{i=0}^{j} \left( \frac{\text{new}_{value} - TS[i]}{j} \right) , j = \text{size of training dataset}$$

$$\text{Score}_{nf} = 1-(x/100) \quad , \quad x = \text{percentage of respective terms in title or description.}$$
Score Computation

• For categorical features:

Let \( y = \) percentage of videos fall under a particular category

\[
\text{Score}_{cf} = 1 - \left( \frac{y}{100} \right)
\]

\[
C_{\text{value}} = \sum_{i=0}^{n} Wi \times Si , \quad 0 \leq C_{\text{value}} \leq 1
\]

Resemblance of the feature with the target class.
# Accuracy Results

<table>
<thead>
<tr>
<th></th>
<th>PVRD</th>
<th>BVRD</th>
<th>CVRD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Predicted</td>
<td>Predicted</td>
<td>Predicted</td>
</tr>
<tr>
<td>True</td>
<td>True</td>
<td>True</td>
<td>True</td>
</tr>
<tr>
<td>Unkn</td>
<td>84.1%</td>
<td>88.5%</td>
<td>86.3%</td>
</tr>
<tr>
<td>own</td>
<td>15.9%</td>
<td>11.4%</td>
<td>13.6%</td>
</tr>
<tr>
<td>AC</td>
<td>True</td>
<td>True</td>
<td>True</td>
</tr>
<tr>
<td>True</td>
<td>4.06%</td>
<td>1.402%</td>
<td>7.1%</td>
</tr>
<tr>
<td>Unknown</td>
<td>95.9%</td>
<td>98.59%</td>
<td>92.8%</td>
</tr>
</tbody>
</table>
Conclusion

- Analysis of discriminatory features.

- Present an approach based on one-class classifier.

- Evaluation results (more than 80% accuracy) show the feasibility of the proposed solution approach.
Future Directions

- Analysis of more contextual features
- Related video response detection
QUESTIONS
Play, Learn and Grow... Together!
References


References


• YouTube statistics. www.youtube.com/t/press_statistics