IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 1, Issue 1, March, 2013
ISSN: 2320 - 8791
www.ijreat.org

Controlling and Blocking of Spammers and Content Promoters Using User Profile Attributes in Online Video Sharing Platforms

Vijayalakshmi.K[1], Amala Ezhilarasi.M[2] and Gayathri.S[3]

[1,2&3]Anand institute of higher technology, Chennai, Tamil Nadu, India

Abstract

In the online video sharing systems, the provision to videos as responses are often used to assist the user in seeking all the connected content. But this feature is frequently misused by posting irrelevant content by the class of users known as spammers. These video responses are further popularized by increasing the views and likes of the particular video and by sharing it in the other social networks. However the spammers can be detected by conducting a series of steps. This paper concentrates on detecting the spammers and content promoters from the legitimate users using attributes based on the user's profile, the user's social behavior in the system, and the videos posted by the user as well as the target videos and investigate the feasibility of applying a supervised learning method to identify polluters. Then it controls or neutralizes the polluted content by deleting the uploaded irrelevant videos and blocking the content promoters.

Keywords—*Promoter, social media, social networks, spammer, video promotion, video response, video spam.*

1. Introduction

Web services are open standard (XML, SOAP, HTTP etc.) based Web applications that interact with other web applications for the purpose of exchanging data. Web Services can convert existing applications into Web-applications. Few definitions are given here and all the definitions are correct. 1) A web service is any piece of software that makes itself available over the internet and uses a standardized XML messaging system. XML is used to encode all communications to a web service. For example, a client invokes a web service by sending an XML message, then waits for a corresponding XML response. Because all communication is in XML, web services are not tied to any one operating system or programming language--Java can .talk with Perl; Windows applications can talk with Unix applications. 2) Web Services are self-contained, modular, distributed, dynamic applications that can be described, published, located, or invoked over the network to create products, processes, and supply chains. These applications can be local, distributed, or Web-based. Web services are built on top of open standards such as TCP/IP, HTTP, Java, HTML, and XML.3) Web services are XML-based information exchange systems that use the Internet for direct application-to-application interaction. These systems can include programs, objects, messages, or documents.4) A web service is a collection of open protocols and standards used for exchanging data between applications or systems. Software applications written in various programming languages and running on various platforms can use web services to exchange data over computer networks like the Internet in a manner similar to inter-process communication on a single computer. This interoperability (e.g., between Java and Python, or Windows and Linux applications) is due to the use of open standards.

1.1 Video spams and detection mechanism

Spam is the use of electronic messaging systems to send unsolicited bulk messages, especially advertising, indiscriminately. While the most widely recognized form of spam is e-mail spam, the term is applied to similar abuses in other media. Face book and Twitter are not immune to messages containing spam links. Most insidiously, spammers hack into accounts and send false links under the guise of a user's trusted contacts such as friends and family.[6] As for Twitter, spammers gain credibility by following verified accounts; when that account owner follows the spammer back, it legitimizes the spammer and allows him or her to proliferate. In this context we focus on the spam caused due to advertisements in the online video sharing websites.

Video sharing sites, such as YouTube, are now being frequently targeted by spammers. The most common technique involves people (or spambots) posting links to sites, on the comments section of random videos or people's profiles. Another frequently used technique is using bots to post messages on random users' profiles to a spam account's channel page, along with enticing text and images. These pages may include their own or other users' videos, again often suggestive. The main purpose of these accounts is to draw people to their link in the home page section of their profile. YouTube has blocked the posting of such links. In addition, YouTube has implemented a CAPTCHA system that makes rapid posting of repeated comments much more difficult than before, because of abuse in the past by mass-spammers who would flood people's profiles with thousands of repetitive comments.

The systems usually offer three basic mechanisms for video retrieval: 1) a search system; 2) ranked lists of top videos; and 3) social links connecting users and/or videos. Although appealing as mechanisms to facilitate content location and enrich online interaction, these mechanisms open opportunities for users to introduce polluted content into the system. As an example, video search systems can be fooled by malicious attacks in which users post their videos with several popular tags . Opportunistic behavior on the other two mechanisms for video retrieval can be exemplified by observing a YouTube feature that allows users to post a video as a response to a video topic. Some users, which we call *spammers*, post unrelated videos as responses to popular video topics aiming at increasing the likelihood of the *responses* being viewed by a larger number of users. Other users, to whom we refer as *promoters*, may try to gain visibility toward a specific video by posting a large number of (potentially unrelated) responses to boost the rank of the *video topic* among the most responded videos, making it appear in the top lists maintained by YouTube. Promoters and spammers are driven by several goals, such as spread advertisements to generate sales, disseminate pornography, or simply compromise system reputation. Polluted content may compromise user patience and satisfaction with the system since users cannot easily identify the pollution before watching at least a segment of it, which also consumes system resources, particularly bandwidth. Additionally, promoters may further negatively impact system mechanisms related to content distribution, since promoted videos that quickly reach high rankings are strong candidates to be kept in caches or in content distribution networks. First, crawling of a large user data set from YouTube site, containing more than 260 thousand users take place. Second step is sampling of user data set to create a labeled test collection of users, which were “manually” classified as legitimate, spammers, and promoters. Sampling was performed to capture different profiles of users in each category. Third, the analysis of a variety of video, individual and social attributes that reflect the behavior of our sampled users, aiming at drawing some insights into their relative discriminatory power in distinguishing legitimate users, promoters, and spammers is done. Fourth, using the same set of attributes, which are based on the user’s profile, the user’s social behavior in the system, and the videos posted by the user as well as her target (responded) videos, the investigation of the feasibility of applying supervised learning methods for identifying the two envisioned types of polluters is carried out. Finally a state-of-the-art supervised classification algorithm namely, support vector machine (SVM) is considered.

2. Modules

2.1 User Login Process

2.1.1 Registration

User should register in our application in order accessing our social networking Video Site.

2.1.2 Login

User may login through their registered credentials.

2.2. User test collection

We say a You Tube video is a *responded video* or a *video topic* if it has at least one video response. Similarly, we say a You Tube user is a *responsive user* if she has posted at least one video response, whereas a *responded user* is someone who posted at least one responded video.

2.2.1 Crawling

Our strategy consists of collecting a sample of users who participate in interactions through video responses, i.e., who post or receive video responses. The sampling starts from a set of 88 seeds, consisting of the owners of the top-100 most responded videos of all time, provided by You Tube. The crawler follows links of responded videos and video responses, gathering information on a number of different attributes of their contributors (users), including attributes of all responded videos and video responses posted by her.

2.2.2 Building test collection

The main goal of creating a user test collection is to study the patterns and characteristics of each class of users. Our user test collection aims at supporting research on detecting spammers and promoters. Since the user classification labeling process relies on human judgment, which implies in watching a significantly high amount of videos, the number of users in our test collection is somewhat limited.

2.3 User reviews

2.3.1 Post a Review:

User can post their reviews about the videos. We say a You Tube video is a *responded video* or a *video topic* if it has at least one video response. Similarly, we say a You Tube user is a *responsive user* if she has posted at least one video response, whereas a *responded user* is someone who posted at least one responded video.

2.3.2 Collecting user’s information

We consider three separate groups of videos owned by the user. The first group contains aggregate information of *all videos* uploaded by the user, being useful to capture how others see the (video) contributions of this user. The second group considers only *video responses*, which may be pollution. The last group considers only the *responded videos* to which this user posted video responses (referred to as *target* videos). In order to obtain a representative sample of the You Tube video response user graph, we build a crawler that implements Algorithm 1. The sampling starts from a set of 88 seeds, consisting of the owners of the top-100 most responded videos of all time, provided by You Tube.

2.3.3 Analyzing user behavior attributes

Our next step is to analyze a large set of attributes that reflect user behavior in the system aiming at investigating their relative discriminatory power to distinguish one user class from the others. We considered three attribute sets, namely, video attributes, user attributes, and social network (SN) attributes.

2.4. Detecting Spammers and promoters

Once we have understood the main tradeoffs and challenges in classifying users into spammers, promoters and legitimate, we now turn to investigate whether competitive effectiveness can be reached with fewer attributes. We report results for the flat classification strategy, considering two scenarios. In this approach, each user is represented by a vector of values, one for each attribute. It is worth noting that some of the most expensive attributes such as User Rank and between ness, which require processing the entire video response user graph, are among these attributes.

2.4.1 Calculating User behavior

The most discriminative user and social network attribute are the average time between video uploads and the User Rank, respectively. In spite of appearing in lower positions in the ranking, particularly for the User Rank attribute, these two attributes have potential to be able to separate user classes apart.

2.4.2 Detecting and blocking of spammers and promoters

Once we have understood the main tradeoffs and challenges in classifying users into spammers, promoters and legitimate, we now turn to investigate whether competitive effectiveness can be reached with fewer attributes. We report results for the flat classification strategy, considering two scenarios. In this approach, each user is represented by a vector of values, one for each attribute. It is worth noting that some of the most expensive attributes such as User Rank and between ness, which require processing the entire video response user graph, are among these attributes.

2.4.3 Classification using SVM

We use a Support Vector Machine (SVM) classifier, which is a state-of-the-art method in classification and obtained the best results among a set of classifiers tested. The goal of a SVM is to find the hyper plane that optimally separates with a maximum margin the training data into two portions of an N-dimensional space. The Training Data which compares the testing data. As a result, we are getting a complete filtering of video responses.

3. System architecture

The architectural diagram of the system performing classification among legitimate users, spammers and content promoters is shown (Figure 1). The system collects n number of users in defined sets from social networking sites that are stored in the administrator’s database system. The first step carried out in the process is sampling of the defined set of users participating in the interaction. This technique is known as *crawling*. The sampling starts from a set of 88 seeds, consisting of the owners of the top-100 most responded videos of all time, provided by You Tube. The crawler follows links of responded videos and video responses, gathering information on a number of different attributes of their contributors (users), including attributes of all responded videos and video responses posted by him. This result collected from the social networking site is in turn filtered into two classifications. SVM classification to filter the video responses by mapping input vectors into an *N*-dimensional space and checking in which side of the defined hyper plane the point lies. Unless otherwise noted, the classification experiments discussed in this section are performed using a fivefold cross validation. In each test, the original sample is partitioned into five subsamples, out of which four are used as training data, and the remaining one is used for testing the classifier. The process is then repeated

Strategies (e.g., one against all [29]). When the (training) data is not completely linearly separable, one may parameterize SVM to assign a cost to possible misclassifications. By tuning this cost, one may exploit the tradeoff between allowing training errors and forcing rigid margins, or, in other words, allowing *soft margins*. There is also the possibility of defining more complex boundaries for separations using *kernel functions* (e.g., polynomial or radial basis functions—RBF) which map the data points into a different space in which the data become more separable. The choices of the kernel and cost value, two parameters of the classifier that maximize classification effectiveness are data dependent. Unless otherwise noted, the classification experiments discussed in this section are performed using a fivefold cross validation.

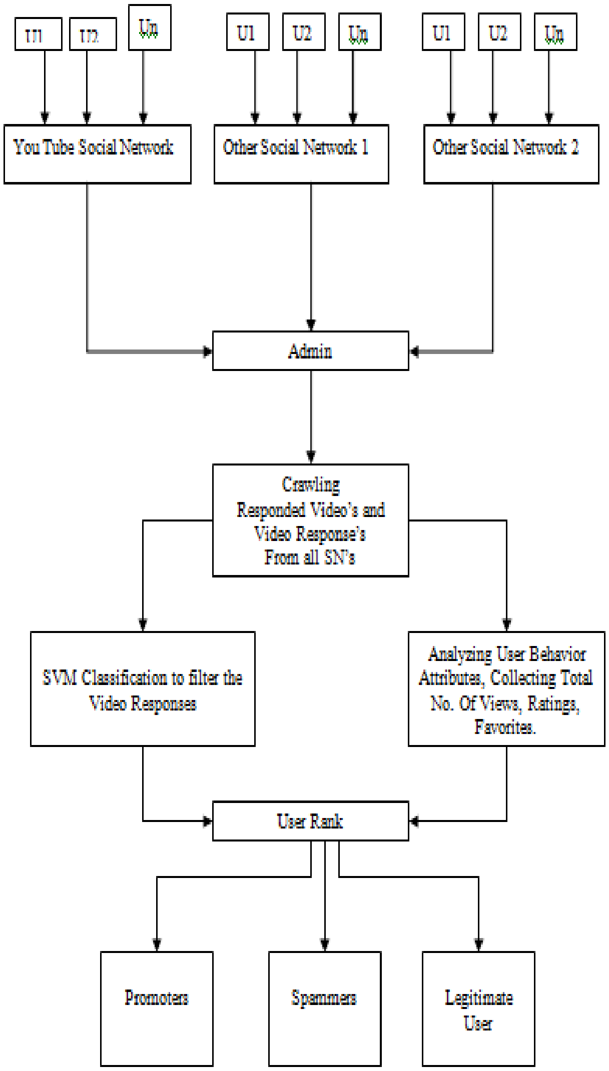
U1
U2
Un
U1
U2
Un
U1
U2
Un
You Tube Social Network
Other Social Network 1
Other Social Network 2
Admin
Crawling Responded Video's and Video Response's From all SN's
SVM Classification to filter the Video Responses
Analyzing User Behavior Attributes, Collecting Total No. Of Views, Ratings, Favorites.
User Rank
Promoters
Spammers
Legitimate User

Figure 1. Proposed architecture for classification of users based on ranking using behavioral patterns

In each test, the original sample is partitioned into five subsamples, out of which four are used as training data, and the remaining one is used for testing the classifier. The process is then repeated five times, with each of the five subsamples used exactly once as test data, thus producing five results. The entire fivefold cross validation is repeated five times with different seeds used to shuffle the original data set, thus producing 25 different results for each test. The results reported are averages of the 25 runs. We also report error intervals with 95% of confidence level [30]. We used a nonlinear SVM with RBF kernel. The implementation of SVM used in our experiments is provided with lib- SVM [17], an open source SVM package that allows searching for the best classifier parameters, namely type of kernel and cost, using the *training data*, a mandatory step in the classifier setup. In particular, we used the *easy* tool from lib SVM, which provides a series of optimizations, including normalization of all numerical attributes. In case of LAC, best parameters were also obtained using cross validation in the training set, being the maximum size of the rules set to five (i.e., at most four attribute values in the antecedent of the rule) and the minimum confidence set to 0.01.

Algorithm 1 Video Response Crawler (run by slave nodes)

Input: A list L of users received from master node

```
for each user $U$ in $L$do
    Collect $U$'s information and video list;
    for each video $V$ in $U$'s video list do
        Collect information of $V$ ;
        if$V$ is a responded video then
            Collect information of $V$ 's video responses;
            Insert the responsive users in list of new users $NL$;
        end if
        if$V$ is a video response then
            Insert the responded user in list of new users $NL$;
        end if
    end for
end for
Return $NL$ to the master node;
```

In the next two sections, we discuss the results obtained with the two classifiers using all 60 attributes, since, as discussed in Section IV, even attributes with low ranks according to the employed feature selection methods (e.g., User Rank) may have some discriminatory power. Moreover, both classifiers are known for dealing well with high dimensional spaces. For instance, SVM is able to properly choose the weights for each attribute, giving low weights to attributes that are not helpful for classification. The crawler ran for one week (01/11–18, 2008), gathering a total of **264 460** users, **381 616** responded videos, and **701 950** video responses. The crawler followed links of responded videos and video responses, gathering information on various attributes of their contributors (users), including attributes of all responded videos and video responses posted by them. Particularly, for each video that was crawled, we collected a number of pieces of information, including video identifier, video owner (i.e., contributor) identifier, title, category, description, tags, upload time, video duration, number of ratings, average rating, number of views, number of users who set the video as favorite, number of comments received, and number of video responses received. We also collected statistics about the author of the video responses of each video and the sequence order in which the video responses were posted.

5. Conclusion

Promoters and spammers can pollute video retrieval features of online video SNs, compromising not only user satisfaction with the system, but also the usage of system resources and the effectiveness of content delivery mechanisms such as caching and content delivery networks. We here proposed an effective solution that can help system administrators to detect spammers and promoters in online video SNs. Relying on a sample of pre classified users and on a set of user behavior attributes, our supervised classification approaches are able to correctly detect the vast majority of the promoters and many spammers, misclassifying only a very small number of legitimate users. Thus, our proposed approach poses a promising alternative to simply considering all users as legitimate or to randomly selecting users for manual inspection. Moreover, given that the cost of the labeling process may be too high for practical purposes, we also propose an active learning approach, which was able to produce results very close to the completely supervised solutions, but with a greatly reduced amount of labeled data.

6. References

[1] comscore: Americans viewed 12 billion videos online in may 2008. http://www.comscore.com/press/release.asp?press=2324.

[2] The New York Times: Search ads come to YouTube. http://bits.blogs.nytimes.com/2008/10/13/search-ads-come-to-youtube.

[3] YouTube fact sheet. http://www.youtube.com/t/fact_sheet.

[4] Y. Ahn, S. Han, H. Kwak, S. Moon, and H. Jeong. Analysis of topological characteristics of huge online social networking services. In *Int'l World Wide Web Conference (WWW)*, 2007.

[5] F. Benevenuto, F. Duarte, T. Rodrigues, V. Almeida, J. Almeida, and K. Ross. Understanding video interactions in YouTube. In *ACM Multimedia (MM)*, 2008.

[6] F. Benevenuto, T. Rodrigues, V. Almeida, J. Almeida, C. Zhang, and K. Ross. Identifying video spammers in online social networks. In *Int'l Workshop on Adversarial Information Retrieval on the Web (AIRWeb)*, 2008.

[7] F. Douglis. On social networking and communication paradigms. *IEEE Internet Computing*, 12, 2008.

[8] R. Fan, P. Chen, and C. Lin. Working set selection using the second order information for training svm. *Journal of Machine Learning Research (JMLR)*,6, 2005.

[9] D. Fetterly, M. Manasse, and M. Najork. Spam, damn spam, and statistics: Using statistical analysis to locate spam web pages. In *Int'l Workshop on the Web and Databases (WebDB)*, 2004.

[10] P. Gill, M. Arlitt, Z. Li, and A. Mahanti. YouTube traffic characterization: A view from the edge. In *Internet Measurement Conference (IMC)*, 2007.