

A System to Filter Unwanted Messages from Online Social Network (OSN) User Walls

Ankush S. Nisal
DYPSOET, Pune

Yogesh Nimbalkar
DYPSOET, Pune

Niteen Pisal
DYPSOET, Pune

Ravikumar Pathare
DYPSOET, Pune

Satish Kumar Penchala
DYPSOET, Pune

ABSTRACT

Internet becomes more and more popular in the day to day activities of user's. In recent years use of online social networks (OSN) also increased rapidly. The users can communicate and share their views, ideas and content through online social networking (OSN). Several types of content like image, text, audio, video etc can be shared between the users. The main drawback of these Online Social Networking (OSN) services is the lack of privacy for the user's own private space. The users can not have ability to direct control to prevent the undesired messages posted on their own private walls. There are only the unwanted messages will be blocked not the user. To avoid this issue, BL (Black List) mechanism is used in this paper, which avoids undesired creators messages.

Keywords

Machine learning techniques, Online Social Networks (OSN) Short text classifier, Filtered Wall, Filtering rules and Black list(BL).

1. INTRODUCTION

Internet technology increases day by day life. Online social network is most interactive medium to communicate with others. In online social networking users share data or content in between, the type of contents are audio, video, image, text, etc. There are chances of posting unwanted data or content on particular public/private space in Online Social Networks (OSN), it is called public walls. In this paper we are mainly focus on the text based messages. OSN provide few amount of security in posting or commenting the unwanted messages. Ability of a user to automatically control the messages or text written on the user public wall, by filtering additional communication will be called as information filtering [1].

To naturally assign the short text based on the content or messages the Machine learning text categorization technique is also used [3]. This technique has done in many steps, by using short text classifier technique we represent the messages, data, text in machine learning based classification, radial basis function network. Filtering rules and black list management is Second step. Finally black list is added. To filter unwanted messages from user wall used automated system filter wall. Proposed system is supported to content based message filtering, but is not supported to existing system. Two level grouping is performed. Short messages are categorized as neutral and Non-neutral in first level called soft categorization. In second stage Neutral messages are classified called hard categorization. Blacklist rule is implemented in proposed system, it said as user who post the unwanted messages kept in black list for particular period of time, and spam tab also implemented. By using this rule more security is provided to Online Social Network.

Web content Mining is used to discover relevant and useful information from a very large amount of Data and Information. In OSN's, information filtering can be used for a many purpose. This is due to fact that in Online Social Networking (OSN) there is the chances of posting (or) commenting other posts on particular public space called user Walls. Information filtering is used to give the user ability to control the message written on their public walls by filtering out unwanted messages.

In this paper, Section 1 include Introduction, Section 2 Related works, Section 3 Existing and Proposed System, Section 4 Methodology, Section 5 Algorithm and Application, section 6 Conclusion.

2. RELATED WORK

In this section describe survey related to recent methods over the content-based filtering in Online Social Networking (OSN).

Macro Vanetti et al [1] in this paper provides the user to have a straight rule over their own public wall to avoid the unwanted messages. The aim of this paper is, users have a straight control over messages posted on their own private wall. So we are using the automated system called Filtered wall (FW), which have a capacity to filter unwanted messages. This system blocks only the unwanted messages posted by the user. Drawback of this paper is user will not block; only the content posted by the user are blocked. Short text classification and content based message filtering support by the system. To avoid this problem, Blacklist rule will be implemented as future enhancement.

L. Roy and R.J. Mooney [12] uses mutual filtering method, but in our proposed system content based filtering is used. It defines the content based book proposal system that develops information pulling out and machine learning technique for text categorization.

B. Carminative et al are define in this paper, the system can generally take decision about the message which is blocked, because of the acceptance depends up on statistical information.[7]

Bodicev and M. Sokolova [5] classification of text put in complex and specific terminology; need the application of learning process. Fractional Matching method is applied which expand the text for confining the text characteristic. Fractional equivalent develops a language model. The output of Fractional matching compression provides consistent care of text classification

J. Colbeck [6] OSN is the common concentration group in network. To make the faith many explanations are required. 2 level approaches are stated to combine gloss, trust and origin.

We state an algorithm for concluding trust relationship with origin information and gloss trust in web social network. Film trust application is introduced which uses trust to movie ranking and ordering the review. we can consider film trust give the good crop.

E. Ferrari [8] this paper proposes a system enforcing content-based message filtering for Online Social Networks (OSNs). The system allows OSNs users to have a direct control on the messages posted on user walls. This is achieved through a flexible rule-based system, that allows a user to customize the filtering criteria to be applied to their walls, and a Machine Learning based short text classifier automatically labeling messages in support of content-based filtering

3. EXISTING AND PROPOSED SYSTEM

3.1 Existing System

Today online social Network provides less protective support to prevent unwanted messages on user wall. For example face book allow users to post who is allowed to post messages in their wall, (like friends, friends of friends, or group of friends). However, no content based preferences are supported therefore it is not possible to prevent unwanted messages, such as vulgar once, and there is no matter who posts them.

3.1.1 Disadvantages of Existing System

1. Social Networks have restrictions on the users who post and comment on their user wall, they do not have restriction on what they post or comment. So, some people use the vulgar words and indecent in comment on the public users wall.
2. Providing this service is not only the matter of using previously defined web content mining techniques for different application, rather it requires to design ad-hoc grouping strategies.

3.2 Proposed System

Machine learning (ML) technique is used to automatically assign every single short text categories based on their contents as text categorization technique. The most effort in construction of robust Short Text Classifier (STC) concentrates on the extraction and selection of group of characterizing and tasteful features. Here, a database of the categorized words is built and it is used to check to if it has indecent words. If message contain any vulgar word, then it will be sent to Black list to filter out that words. The message without any indecent words that posted on user's public wall as the result of the content based filtering technique.

As far as the learning model has concerned, we confirm in the current paper the use of neural learning which is today recognized as one of the more efficient solutions in text classification. In particular, we base the overall short text classifying strategy on Radial Basis Function Networks (RBFN) for their proven capabilities in acting as soft classifiers, in managing unwanted data and intrinsically vague classes. We insert the neural model within a hierarchical 2 level classification strategy. In the first level, the RBFN categorizes short messages as Neutral and Non-neutral. In the 2nd stage, Non-neutral messages are classified producing

gradual estimates of appropriateness to each of the considered category.

Blacklist technique is used, where the user's blocked and avoided for instant to post user public wall. As per filtering rule included, additionally Blacklist (BL) rule is used.

Advantages of proposed work

1. A system mechanically filters unwanted messages using the blacklists on the base of both message content and the message creator interactions and characteristics.
2. Most important difference include , a different semantics for filtering rules to better suite the considered domain, to help the users Filtering Rules(FRs) specification, the extension of the set of features measured in the classifying process.

4. METHODOLOGY

The users have created as well as manage their own "groups" (like the new Face book groups' pages). Every group has a homepage that make available a place for subscribers to post and share (by posting messages, images, etc.) and a block that provides required data about the group. Users can also allow additional features in their public page like view friends list & add friends by using friend's requests as well as share their images with selected group's members. The status should be updated of their friends' requests here.

4.1.1 Filtering Process

In defining the language for Filter rules specification, we consider three main issues that, in our opinion, affect message filtering decision. 1st, in OSNs like in everyday life, the same message may have different meanings and relevance based on who writes it. As a consequence, Filter rules should allow users to state constraints on the message creators. Creators on which a Filter rules is applied can be selected on the basis of several different criteria; one of the most relevant is by imposing conditions on their profile attributes. In such a way it is, for instance, possible to define rules applying only to young creators or to those creators with a given religious/political view. Given the online social network scenario, identification of creators is by exploiting information on their social graph. This implies to state conditions on a type, depth and trust values of the relationship creators should be involved in order to apply them the specified rules.

The problem of setting thresholds to FRs is also addressed, by conceiving and implementing within FW, an Online Setup Assistant (OSA) procedure. For particular message, the user tells the system, the decision to accept or to reject the message. The collection & processing of user decisions on a sufficient set of messages distributed over all the classes allows computing customized thresholds for representation of the user attitude in accepting or rejecting certain contents. Such messages are selected on the basis of following process. (fig.1)A certain amount of non-neutral messages taken from a fraction of the dataset and not belonging to the test sets, are classified by the ML in order to have, for each message, the second level class membership values.

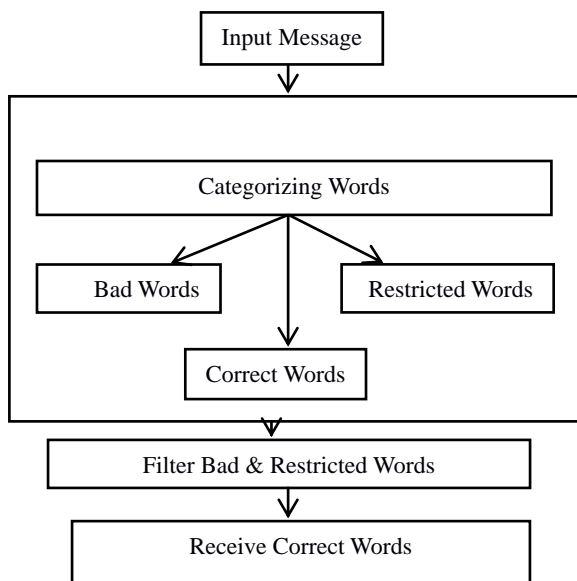


Fig.1. Filtering Process

4.1.2 Short Text Classifiers

The main task of the proposed system is the content-based message filtering (CBMF) and short text classifier. It additionally supports the classifying of message based up on the category set. Our study is aimed at designing and evaluating various representations of techniques in combination with a neural learning strategy to semantically categorize short text. From a ML point of view, we close with the task by defining the hierarchical two level strategies assuming that it is better to identify and elimination of “neutral” sentences, then classifying “Non Neutral” sentences by the class of interest instead of doing everything in 1 style. This choice is motivated by related work showing advantages in classifying text and short text using a hierarchical strategy. The first level task is conceived as hard classification in which short text are labeled with fresh Neutral and Non Neutral labels. The second level soft classifier acts on the fresh collection of non-neutral short text and, for each of them, it “simply” produces estimated appropriateness or “gradual membership” for each of the imagine classes, without taking any “hard” decisions on any of them. Subsequent phases use the list of grades of a filtering process.

4.1.3 Black List process

A further component of our system is a Blacklist mechanism to avoid messages from undesired creators, independent from their contents. Blacklists are directly managed by the system, which should be able to show who are the users to be inserted In the Blacklist and decide when user retention in the Blacklists is finished (fig.2). To enhance flexibility, such information is given to the system through a set of rules, hereafter called Blacklist rules. Such rules are not defined by the SNM, therefore they are not meant as general high level directives to be applied to the whole community. We decide to let the users themselves, i.e., the wall’s owners to specify Blacklist rules regulating who has to be banned from their walls and for how long. Therefore, a user might be banned from a wall, at the same time, being able to post in other walls.

Similar to Filter rules (fig.2); our BL rules make the wall owner able to identify users to be blocked according to their profiles as well as their relationships in the Online Social Network. Therefore, by means of a BL rule, wall owners are

for example able to banned from their walls users they do not directly know (i.e., with which they have only indirect relationships), or user that are friend of a given person as they may have a bad opinion of this person. This banning can be adopted for an undetermined time period or for a specific time window. Moreover, banning criteria may also take into account users’ behavior in the Online Social Network. More precisely, among possible information denoting users’ bad behavior we have focused on 2 main measures. The 1st is related to the principle that if within a given time interval a user has been inserted into a Blacklist for several times, say greater than a given threshold, he might deserve to stay in the Blacklist for another while, as his/her behavior is not improved. This principle works for those users that have been already inserted in the considered Blacklist at least one time. In contrast, to catch new bad behaviors

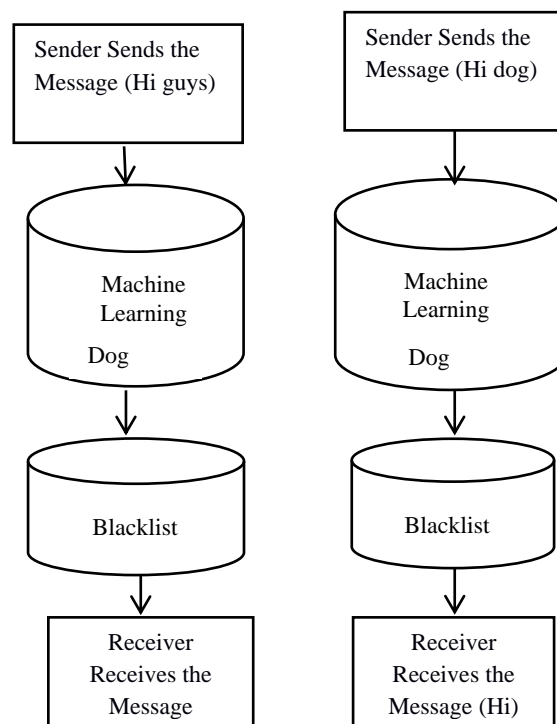


Fig.2. Blacklist Process

We can use the Relative Frequency (RF) that let the system be able to detect those users whose messages continue to fail the Filter rules. The 2 measures can be computed either locally, that is, by considering only the messages and/or the Blacklist of the user specifying the BL rule or globally, that is, by considering all online social network users walls and BLs

The admin manages all the user information including posting comments in the user status box. Each unwanted message has an alert from admin that provides a place for post & share for the respective user walls. And admin can see blocked message from the users & also that provides information about the user who used the blocked message. Admin can also enable some extended features in their owned page like user list, adding unwanted message, update unwanted messages, Blocked users list & lastly filter performance graph. And also in this module, we show the performance evaluation of the system in the graph.

1.1.4 Machine Learning Techniques

The Machine Learning is a system which can learn from the data and take decisions based on the learned data. For ex, a

Machine Learning System in the mail Inbox can be used to learn & distinguish the emails received in the inbox between spam or non-spam emails. In a similar way, the Machine Learning system traces the posted messages for the good and the unwanted words used in the wall by the public users. The above algorithm represents the concept of Machine Learning with the BLs. a user is showing his interest in posting and commenting in other person's wall regardless of their relationship. He can post any message.

First user shows his interest in posting or commenting in other person's wall regardless of their relationship. A user can post any message there without the filtering technique. But the Machine Learning system here learns the message which is yet to be posted & finds whether it contains any illegal words in it. If it can't find any illegal words, then the system allows the message to be posted on the wall. If it finds any illegal words in that message while learning it, then it will remove the illegal words from the message and then insert those words in the BL which stores the indecent words in it. Lastly, the system prints the message without the indecent words. This mechanism helps in preventing the users to get annoyed by the illegal words in a public wall of the Social Networking Sites. It doesn't prevent the unknown users from posting their messages rather; it helps to prevent the obscenity with the illegal words.

5. ALGORITHM AND APPLICATION

5.1 Application

1. This application is useful for common people who don't want to write any unwanted messages like vulgar, political, sexual messages on his/her own wall by any third person.
2. Mostly, this type of activities are happen with some famous personalities, So if this facility will provide with OSN sites then people can protect his wall from this type of malpractices.

5.2 Algorithm

- Step 1: Start
- Step 2: A User tries post the message in a wall.
- Step 3: Machine learning checks each word of the message.
- Step 4: If (Words == Good Words)
- Step 5: Message is posted on the wall.
- Step 6: Else if (Words == Bad Words)
- Step 7: Reject Bad Words using Blacklist
- Step 8: Stop

6. CONCLUSION

In this paper, a system in preventing the indecent messages from the Social Networking site walls has been presented. The

Usage of Machine Learning system has given higher results to the system to trace the messages and the users to distinguish between the good & bad messages and the authorized & unauthorized users in the Social Networking User Profiles. Thus the Machine Learning system Technique plays an important role in this paper in order to generate the blacklist of the bad words and the users which are unauthorized. The user has to update his privacy setting in his account in order to add this method in preventing the vulgarity in his public profile. In this context, a statistical analysis has been conducted to provide the usage of the good & bad words by the persons in the sites. Overall, the obscenity of the users has been prevented.

7. REFERENCES

- [1] Marco Vanetti, Elisabetta Binaghi, Elena Ferrari, Barbara Carminati, an Moreno Carullo, " A System to Filter Unwanted Messages from OSN User Walls",2013.
- [2] Hongyu Gao, Alok Choudhary, Yan Chen, Northwestern University Evanston, IL, USA.
- [3] F.Sebastiani, "Machine Learning Automated Text Categorization", ACM Computing surveys, vol.34, no.1, pp.1-47, 2002.
- [4] M.Carullo, E.Binaghi and I.Gallo, and N.Lamberti, "Clustering of Short commerical Documents for the web," Proc.19th Int'l conf. Pattern Recognition (ICPR '08), 2008.
- [5] V.Bobicev and M.Sokolova, "An Effective and Robust Method for Short Text Classification," Proc.23rd Nat'l Conf. Artificial Intelligence (AAAI), D.Fox and C.P.Gomes, eds., pp.1444-1445,2008.
- [6] J.Colbeck, "Combining Provenance with Trust in Social Networks for Semantic Web Content Filtering," Proc. Int'l conf. Provenance and Annotation of Data, L.Moreau and I.Foster, eds., pp.101-108, 2006.jejv
- [7] M.Vanetti, E.Binaghi, B.Carminati, M.Carullo, and E.Ferrari, "Content- Based Filtering in On-Line Social Networks", 2010.
- [8] M. Vanetti, E. Binaghi, B. Carminati, M. Carullo, and E. Ferrari, "Content-Based Filtering in On-Line Social Networks," Proc.ECML/PKDD Workshop Privacy and Security Issues in Data Mining and Machine Learning (PSDML '10), 2010.
- [9] R.J.Mooney and L.Roy, "Content-Based Book Recommending Using Learning for Text Categorization", 2000.
- [10] S.Zelikovitz and H.Hirsh, "Improving Short Text Classification Using Unlabeled Background Knowledge," Proc. 17th Int'l Conf. Machine Learning (ICML '00), P.Langley, ed.,pp.1183-1190, 2000