

# **A Review on Message Filtering from Online Social Network**

**Alpa A. Kothari**  
Computer Department  
MET's BKC Institute of Engineering, Nashik  
Savitribai Phule Pune University

**M.U. Kharat**  
Computer Department  
MET's BKC Institute of Engineering, Nashik  
Savitribai Phule Pune University

## **ABSTRACT**

In OSNs (On-line Social Networks), the ability to avoid the display of unwanted content of the posted messages on their own private space; this is one of the most fundamental issue to give users. At the present online social networks provide dissatisfactory support to this requirement. Author built a system which allows online social network users to have a direct control to post messages on their walls. This is possible by using flexible rule-based system which permits user to modify the filtering criteria to be applied to user walls, and a based on soft classifier of Machine Learning, mechanically category messages in favor of content-based filtering.

## **Keywords**

Policy-based personalization, Short Text Classification, On-line Social Networks, Information Filtering.

## **1. INTRODUCTION**

In existence, On-line Social Networks have become a admired interactive medium to communicate, share and broadcast information about human life. Daily and constant communications imply the swap of several types of content, with image, audio, free text, and video data. The vast and lively character of these data creates the basis for the employment of web content mining program aimed to automatically find substantive information inactive within the data and then give an active support in complex and critical tasks included in social networking analysis and management. A noticeable example are the messages permanently written by online social network users on public/private areas, called in general walls was part of social network content and constitute by short text [1]. The current work is supposed to propose and also experimentally evaluate an computerized system which is called as Filtered Wall (FW) which is able to filter unnecessary messages from social network user walls. The basic idea of this system is the hold for content- based user preference. Online Social Networks are nowadays one of the most admired interactive medium to communicate, share and broadcast information about human life. Daily and constant communications imply the swap of several types of content, with free text, image and also audio and video data. Face-book statistics says that one average user creates ninety pieces of content each month, while more than thirty billion pieces of content (blog posts, notes, photo albums, web, links, news stories, etc.) are shared each month. The vast and lively character of these data gives the foundation for the employment of web content mining strategy aimed to routinely detect useful information inactive within the data. However the intend of these proposals is mainly to give users a classification mechanism to keep away from inundated by useless data. In Online social network, information filtering can also be used for more sensitive reason. This is due to the fact that in online social network there is the chance of

messaging or commenting other posts on particular public/private areas, called in general walls.

Information filtering can be used to give users the ability to routinely control the posts written on user general wall, by filtering out unnecessary messages. It believes that, from so far the basic online social network service has not been provided. In fact, nowadays online social networks provide very slight support to filter or remove or avoid unnecessary messages on their own walls. For example, Face-book provide facility to users that, they decide who is allowed to post messages on their walls (i.e., friends of friends, friends, or defined groups of friends) but it does not provide facility to prevent undesired messages such as political issue related or vulgar, or any other unnecessary message, as no content-based preferences are supported without considering who is going to posts them. This service is not only for using earlier defined web content mining procedure for a dissimilar application, but it requires to designs ad-hoc categorization strategies. This is due to that short text organizes wall messages by using traditional classification methods which have serious restrictions since short texts do not provide enough word occurrences. The aim of current work is to give a system which is able to filter unnecessary messages from online social network user which is called as Filtered wall (FW). System uses Machine Learning (ML) text categorization procedure [2] able to automatically assign to each short text message as a set of categories based on its content. To the best on our data this is the first proposal of a system which automatically filter unnecessary messages from online social network user walls by using message content along with the message creator relationships and characteristics.

## **2. LITERATURE SURVEY**

Nicholas J. Belkin and W. Bruce Croft has been discussed relationship between information filtering and information retrieval and they come to the conclude that both are two sides of the same coin [3][10]. The previous suggested systems use social filtering methods that base on other users' preferences recommendations.

By contrast R. J. Mooney and L. Roy define a content-based system that utilizes extracted information and a machine-learning algorithm for Short text classification. This way they improve access to relevant products and information [8][10]. The assignment of natural language texts categorization is an main component in several information organization and management tasks.

J. Platt, S. Dumais, D. Heckerman, and M. Sahami compare the effectiveness of 5 different automatic learning algorithmic rule for text classification in terms of learning speed, real-time classification speed, and classification accuracy and they conclude that Linear Support Vector Machines (SVMs) are most accurate classifier, fastest to train, and quick to evaluate.

They used SVMs for categorizing email messages and Web pages.

They also hope to extend their work by including the additional structural information about documents, as well as knowledge-based features for classification accuracy and automatically classify items into hierarchical category structures [4][10]. The widely use text representation techniques for text retrieval are phrase indexing and clustering. D. D. Lewis studied the properties of phrasal and clustered indexing, to isolation from query interpretation issues. He worked on same number of features for each category and there was no automated feature selection.[5][10].

R. E. Schapire and Y. Singer, describe in detail an application, called BoosTexter, which consists of boosting algorithms for text classification tasks. And also they compare its performance with a number of other text-classification algorithms on a Flavour of tasks.[6][10].

Neural network allows us to model the higher order communication between document terms and to simultaneously guess various topics using shared hidden futures. A. S. Weigend, J. O. Pedersen, and, E. D. Wiener presents an application of nonlinear neural network to topic spotting [7]. Whereas topic spotting is the problem of identifying which of the set of predefined topics are existing in a natural network document [7].

Sarah Zelikovitz, Haym Hirsh describe a method for improving the classification of short text series using a combination of labeled and unlabeled but related longer documents [9][10].

### 3. METHODOLOGY

The aim of the current work is to evaluate an automated system, called Filtering Wall (FW), capable to filter unnecessary messages from online social network user walls. It make use of Machine Learning (ML) text classification techniques to assign with each short text message a set of categories based on its content [2]. The construction of robust short text classifier is on the basis of withdrawal and choice of a set of characterizing and also on discriminating features.

Earlier work are also refer to find the solution in this paper and also elicitation procedure and learning model are also inherited in this work for giving pre-classified data.

Properties of short texts gives the original set of features, and are enlarged here which include knowledge related to the context from which the messages initiated. Learning model is concerned with authentication of the use of neural learning from so far but nowadays it recognized as one of the best solutions in text classification [2]. The short text classification strategy based on Radial Basis Function Networks (RBFN) and capable to act as soft classifiers, manage noisy data and intrinsically vague classes.

#### 3.1 System Design

The design of online social network services has a three-tier structure (Figure 1).

1. The 1<sup>st</sup> layer, called Social Network Manager (SNM), the main aims is to provide the basic online social network functionalities (i.e., profile and relationship management).
2. The 2<sup>nd</sup> layer, Social Network Applications (SNAs) which support external social network. The supported SNAs may be called additional layer which is connected Graphical User Interfaces (GUIs).

3. The 3<sup>rd</sup> layer is Graphical user interface which interact with System.

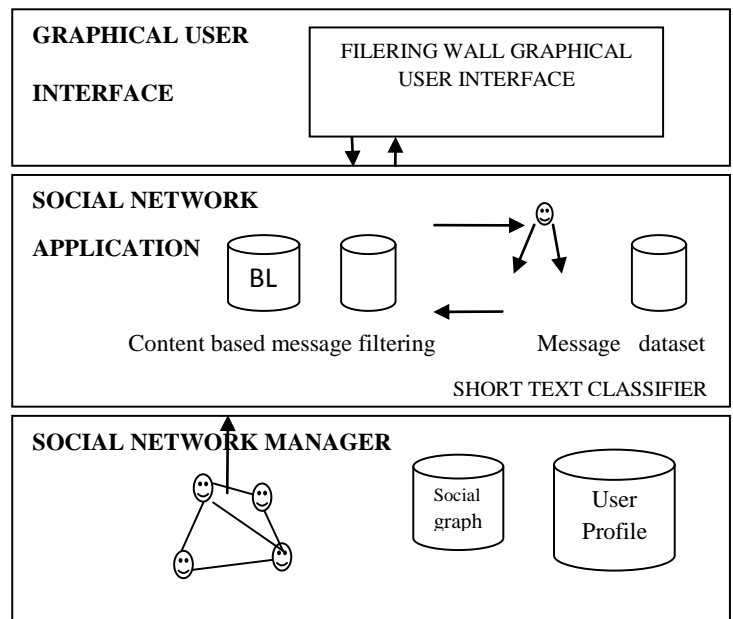


Figure 1: Filtered Wall Design

Proposed system is placed in the second and third layer according to the reference architecture. By using GUI users interact with the system to set up and manage their filtering rules and black lists. Moreover, the GUI provides users with a filtering wall which publish messages that are certified according to their filtering rules and black lists. The core components of the system are the Content-Based Messages Filtering (CBMF) and the Short Text Classifier (STC) modules [1]. The first component refers the STC module which uses message categorization to enforce the FRs specified by the user. Black lists can also be added to enhance the filtering process. Fig1 shows the system execution path, and can be summarized as follows:

- Step 1- A user post the message on general wall.
- Step 2- Machine learning module checks each word of the message.
- Step 3- If word is good message is posted on wall else it will be rejecting using Blacklist.
- Step 4- then filtered message is posted on general walls.

#### 3.2 Filtering Method

It considers three main issues in defining the language for Filtering rules specification. The one is related to the fact that, the different meaning with the same message and decided on relevance based on who writes it. FR applied message creators can be selected on the basis of various different criteria out of which first one of the most applicable is by considering their profile's attributes. Users state constraints on message creators by using filtering rule. Creators may also uses exploiting information on their social graph for identification. For applying the specified rule it considers depth and trust values of the relationship specified.

The filtering process is as follows:

1. Input message.
2. Machine Learning Module.

Categorizing Words (bad, correct, restricted)

3. Filter bad and restricted words.

Receive correct words.

### **3.3 Online Setup Assistant for FRs Threshold**

By conceiving and implementing within filtering wall, an Online Setup Assistant (OSA) procedure, we find the solution of setting thresholds to filter rules. Every time for every message, the user permitted the system to take decision to accept or reject the message. After that user decision are collected and processed on an adequate set of messages which are distributed over all the classes then after that it allows computing customized thresholds which represent the user decision that certain content should going to be accepted or rejected. Then message selection can be according to the following process. The non-neutral messages are taken from a part of the dataset and not from to the training/test sets, are classified with the help of machine learning for getting each message.

### **3.4 Blacklists Method**

Blacklist mechanism is another component of this system to avoid messages from unwanted creators, which is not depend on the message content. Blacklists are managed by the system, and should be able to determine the users to be in the Blacklist and decide user's retention in the Blacklist is finished. Then let the wall's owners to specify Blacklist rules controlling who has to be restricted or block listed from their walls and for how much time. Therefore, a user might be restricted from a wall but at the same time user are free to post in other walls. The Following process can be followed my Blacklist method:

1. Sender sends the message
2. Message is checked in ML dataset.
3. And then in Blacklist dataset.
4. Receiver receives the message.

This rules are not describe by the Social network Manager, so they are not consider to be apply on whole community as general high level directives. Rather, Author decides that user have been able to decide by self that who has to be banned from user general walls and for how much time by specifying Blacklist rules and regulation, so that user might be banned from a one wall but at the same time user able to post in other walls. And owner can also be able to identify users to be blocked according to their relationships in the online social network as well as their profiles by using Blacklist rule. This banning time period can be undetermined or for a specific time window also. Banning criteria may also consider into user's account behavior in the online social network. There is various possible information denoting users' bad behavior, among those author has to decide to consider two main measures. Out of two main measures the first consideration is that within a particular time period user entered into Blacklist for several time that is suppose greater than given threshold the user might responsible to stay in the Blacklist for another as soon as behavior is not improved. If the users already inserted in the considered Blacklist at least one time then and then this principle works on this user. And the second measure in contrast to Blacklist author uses Relative Frequency (RF) to catch new bad behaviors, which help the system to detect those users whose messages continuously blocked in the Filtering Methods. These two measures are

considered locally or globally. Locally refers only the messages and/or the blacklist of the user specifying the Blacklist rule and globally refers all online social network users' walls and/or Blacklists.

## **4. CONCLUSION**

In the paper, a system to filter undesired message in online social networks wall is offered. The first step is to categorize the content using number of rules. Next step is to filter the undesired rules. Last but not the least Blacklist rule is implemented. So that user can filter undesired messages. In future Work, the filtering rules can be implemented with the aim of bypassing the filtering system, it can be used for the overcoming the purpose of the filtering system.

## **5. ACKNOWLEDGEMENTS**

The Author wish to thank MET's BKC Institute of Engineering Nashik, HOD of computer Department and Parents for supporting and motivating for this work. And without their blessing this won't be possible softly.

## **6. REFERENCES**

- [1] Marco Vanetti, Elisabetta Binaghi, Elena Ferrari, Barbara Carminati, and Moreno Carullo "A System to Filter unwanted Messages from OSN User Walls"- IEEE Transactions on Knowledge and data engineering, vol. 25, no. 2, Feb 2013.
- [2] F. Sebastiani, "Machine learning in automated text categorization," ACM Computing Surveys, vol. 34, no. 1, pp. 1–47, 2002.
- [3] N. J. Belkin and W. B. Croft, "Information filtering and information retrieval: Two sides of the same coin?" Communications of the ACM, vol. 35, no. 12, pp. 29–38, 1992.
- [4] S. Dumais, J. Platt, D. Heckerman, and M. Sahami, "Inductive learning algorithms and representations for text categorization," in Proceedings of Seventh International Conference on Information and Knowledge Management (CIKM98), 1998, pp. 148–155.
- [5] D. D. Lewis, "An evaluation of phrasal and clustered representations on a text categorization task," in Proceedings of 15th ACM International Conference on Research and Development in Information Retrieval (SIGIR-92), N. J. Belkin, P. Ingwersen, and A. M. Pejtersen, Eds. ACM Press, New York, US, 1992, pp. 37–50.
- [6] R. E. Schapire and Y. Singer, "Boostexter: a boosting-based system for text categorization," Machine Learning, vol. 39, no. 2/3, pp. 135–168, 2000.
- [7] E. D. Wiener, J. O. Pedersen, and A. S. Weigend, "A neural network approach to topic spotting," in Proceedings of 4th Annual Symposium on Document Analysis and Information Retrieval (SDAIR-95), Las Vegas, US, 1995, pp. 317–332.
- [8] R. J. Mooney and L. Roy, "Content-based book recommending using learning for text categorization," in Proceedings of the Fifth ACM Conference on Digital Libraries. New York: ACM Press, 2000, pp.195–204.
- [9] S. Zelikovitz and H. Hirsh, "Improving short text classification using unlabeled background knowledge," in Proceedings of 17th International Conference on Machine Learning (ICML-00), P. Langley, Ed. Stanford, US: Morgan Kaufmann Publishers, San Francisco, US, 2000, pp. 1183–1190.
- [10] Vidya Alone and .R.B.Talmale, "Message Filtering Techniques for On-Line Social Networks: A Survey" ISSN 2319 – 4847, Volume 3, Issue 3, March 2014.