



# INTERNATIONAL JOURNAL OF PURE AND APPLIED RESEARCH IN ENGINEERING AND TECHNOLOGY

A PATH FOR HORIZING YOUR INNOVATIVE WORK

## FILTERING OF UNWANTED MESSAGES FROM OSN USER WALL USING CONTENT-BASED FILTERING METHOD

SHISHIR THAKARE<sup>1</sup>, ARPIT PATOKAR<sup>1</sup>, PRANIT SHIRSATH<sup>1</sup>, CHETAN CHOUDHRY<sup>1</sup>, PROF. S.A. DHAWALE<sup>2</sup>

1. Student, Department of Information, Technology, S.G.B.A.U. Amravati.

2. IBSS COE, Amravati.

Accepted Date: 05/03/2015; Published Date: 01/05/2015

**Abstract:** One major problem in today's Online Social Networks (OSNs) is to give users skill to regulate the messages posted on their own personal space to avoid that unauthorized data is displayed. OSNs give small support to these needs. In this thesis, I propose a system permit OSN users to have a straight control on the messages posted on their walls. It is achieved through a flexible rule-based system, that permit users to customize the filtering criteria to be put to their walls, and a Machine Learning-based soft classifier automatically labelling messages in endure of content based filtering. First conduct a set of large-scale measurements with a collection of accounts observe the difference among human. Up to now OSNs provide little support to this requirement.

**Keywords:** Online Social Networks, Filtering Method

Corresponding Author: MR. SHISHIR THAKARE



PAPER-QR CODE

Access Online On:

[www.ijpret.com](http://www.ijpret.com)

How to Cite This Article:

Shishir Thakare, IJPRET, 2015; Volume 3 (9): 1462-1469

## INTRODUCTION

In the last years, On-line Social Networks (OSNs) have become a popular interactive medium to communicate, share and disseminate a considerable amount of human life information. Daily and continuous communication implies the exchange of several types of content, including free text, image, audio and video data. According to Facebook statistics 1 average user creates 90 pieces of content each month, whereas more than 30 billion pieces of content (web links, news stories, blog posts, notes, photo albums, etc.) are shared each month. A main part of social network content is constituted by short text, a notable example are the messages permanently written by OSN users on particular public/private areas, called in general walls. Facebook allows users to state who is allowed to insert messages in their walls. However, no content-based preferences are supported and therefore it is not possible to prevent undesired messages, such as violent or vulgar ones, no matter of the user who posts them.

The aim of the present paper is to propose and experimentally evaluate an automated system, called Filtered Wall (FW), able to filter out unwanted messages from social network user walls.

## LITRATURE AND SURVEY

In the OSN domain, interest in access control and privacy protection is quite recent. As far as privacy is concerned, current work is mainly focusing on privacy-preserving data mining techniques, that is, protecting information related to the network, i.e., relationships/ nodes, while performing social network analysis? Works more related to their proposals are those in the field of access control. In this field, many different access control models and related mechanisms have been proposed so far, which mainly differ on the expressivity of the access control policy language and on the way access control is enforced

### 1.1 Content Based filtering

Information filtering systems are designed to classify a stream of dynamically generated information dispatched asynchronously by an information producer and present to the user those information that are likely to satisfy his/her requirements. In content-based filtering each user is assumed to operate independently. As a result, a content-based filtering system selects information items based on the correlation between the content of the items and the user preferences as opposed to a collaborative filtering system that chooses items based on the correlation between people with similar preferences. Documents processed in content-based filtering are mostly textual in nature and this makes content-based filtering close to text classification. The activity of filtering can be modeled, in fact, as a case of single label, binary

classification, partitioning incoming documents into relevant and non relevant categories. More complex filtering systems include multi-label text categorization automatically labelling messages into partial thematic categories. Content-based filtering is mainly based on the use of the ML paradigm according to which a classifier is automatically induced by learning from a set of pre-classified examples.

## **RELATED WORK**

The aim of the present work is therefore to propose and experimentally evaluate an automated system, called Filtered Wall (FW), able to filter unwanted messages from OSN user walls. We exploit Machine Learning (ML) text categorization techniques [1] to automatically assign with each short text message a set of categories based on its content. The major efforts in building a robust short text classifier are concentrated in the extraction and selection of a set of characterizing and discriminate features. The solutions investigated in this work are an extension of those adopted in a previous work [11] from which we inherit the learning model and the elicitation procedure for generating pre-classified data. A remarkable variety of related work has recently appeared which differ for the adopted feature extraction methods, model learning, and collection of samples. The feature extraction procedure maps text into a compact representation of its content and is uniformly applied to training and generalization phases. Additional challenges given the short length of these messages other than the wide range of topics that can be discussed. Short text classification has received up to now few attentions in the scientific community. Most of these models express access control requirements in terms of relationships that the requestor should have with the resource owner. We use a similar idea to identify the users to which a filtering rule applies. We mainly deal with filtering of unwanted contents rather than with access control. As such, one of the key ingredients of our system is the availability of a description for the message contents to be exploited by the filtering mechanism as well as by the language to express filtering rules. In contrast, no one of the access control models previously cited exploit the content of the resources to enforce access control. Moreover, the notion of black- lists and their management are not considered by any of these access control models

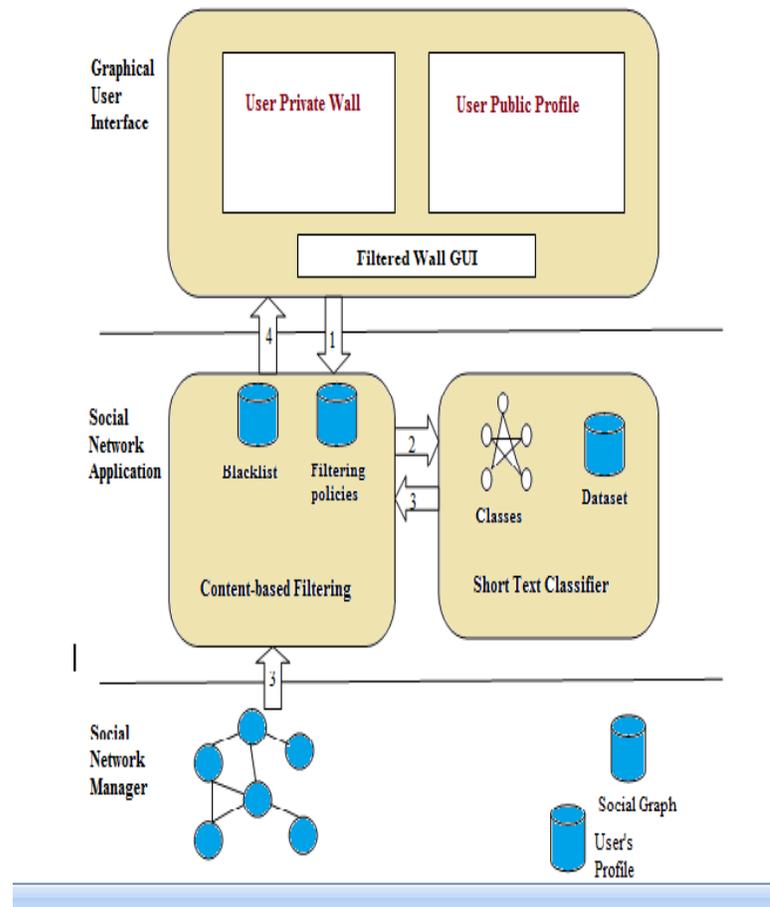


Figure 4.1: Filtered Wall architecture

### 1.2 Filtering rules

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

### 1.3 FRs threshold

We address the problem of setting thresholds to filter rules, by conceiving and implementing within FW, an Online Setup Assistant (OSA) procedure. OSA presents the user with a set of messages selected from the dataset. For each message, the user tells the system the decision to accept or reject the message. The collection and processing of user decisions on an adequate set of messages distributed over all the classes allows to compute customized thresholds representing the user attitude in accepting or rejecting certain contents. Such messages are selected according to the following process. A certain amount of non neutral messages taken from a fraction of the dataset and not belonging to the training/test sets, are classified by the ML in order to have, for each message, the second level class membership values.

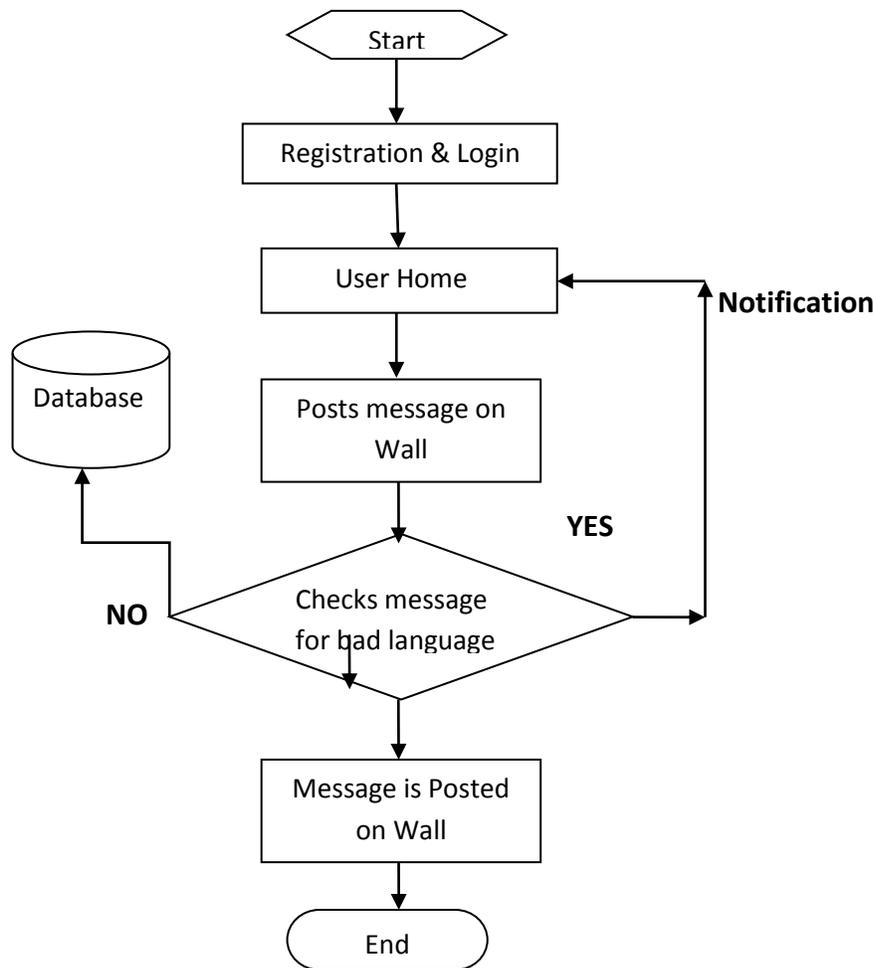


Figure: Flowchart

When any user tries to post a messages on other user wall it is intercepted by the filtered wall first.

Filtered wall check the messages for any restricted word and extracts the metadata and follows some procedure. If there is no restricted word found in that messages then that message is successfully sent to the destination user.

If there is restricted word occur in that message then the filtered wall apply the different filtering rules and at the same time the threshold value is increases in the database.

If the threshold value (TH) is equals to 4 and less than 8 then user gets notification that you crossed the first threshold value so be careful to send any type of message having bad word. If the threshold value is equals to 8 means. if user crossed the second threshold value after getting notification then he goes into the blacklist.

#### 1.4 Blacklists

A further component of our system is a BL mechanism to avoid messages from undesired creators, independent from their contents. BLs are directly managed by the system, which should be able to determine who are the users to be inserted in the BL and decide when users retention in the BL is finished. To enhance flexibility, such information are given to the system through a set of rules, hereafter called BL rules. Rather, we decide to let the users themselves, i.e., the wall's owners to specify BL rules regulating who has to be banned from their walls and for how long. Therefore, a user might be banned from a wall.

#### SOCIAL AWARENESS

Social Networking plays a very important role in social awareness in today's world. As now the number of users who uses social networking is increasing rapidly. So any type of bad comments which belongs to particular topic or any small topic can be spread on social networking in very bad manner. The proposed system provides solution for such problems in social networking. We proposed a new system called Filtered Wall which is able to filter out the unwanted messages from OSN User wall.

#### RESULT AND DISCUSSION

This section gives the performance of proposed system with the existing system and shows that the proposed system is secured than existing system. As the popularity of social networking sites continues to grow, so do the security risks associated with them increases.

According to the survey 80% of the organizations use social networking for their business. As the number of users increases for their product, the percentage of security risk will increase. The proposed system will provide a way to minimize the security risk for such scenarios.

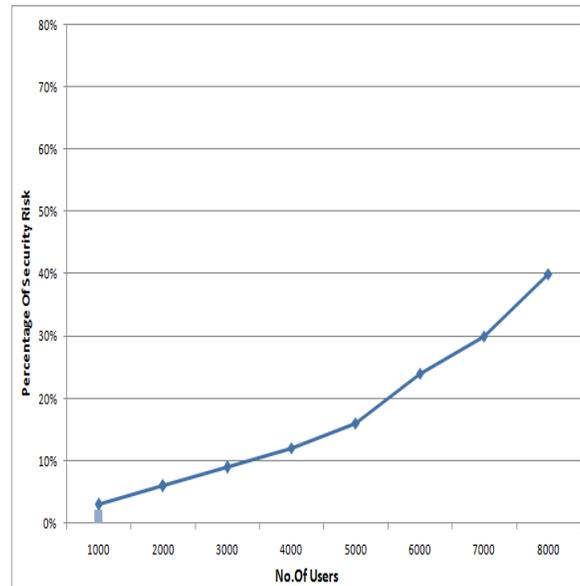


Figure: Security Risk

This system will have a way to block specified words from getting communicated over social networking. Above Figure shows the graph that represents that how the security risk decreases with the proposed system in which different filtering rules are there, that's why the security risk will be minimize as compare to the existing system.

## CONCLUSION

We proposed a new system called Filtered Wall which is able to filter out the unwanted messages from OSN User wall. We compared this system with the number online social networking sites and other existing system. The experimental results show that the proposed system outperforms many of the existing system.

Day-to-day life requires the Online Social Networking (OSN) sites. So OSN requires a way to handle any messages to which is having bad impact on day-to-day life. Filtered wall is the best solution to handle such situations. FW will provide a way to minimize risk of such messages on OSN. FW is the best way to block such messages and also the user to whom user don't want access his profile.

## REFERENCES

1. F. Sebastiani, "Machine learning in automated text categorization," *ACM Computing Surveys*, pp. 1–47, 2002, vol. 34, no. 1.
2. 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.
3. P. J. Hayes, P. M. Andersen, I. B. Nirenburg, and L. M. Schmandt, "Tcs: a shell for content-based text categorization," in *Proceedings of 6th IEEE Conference on Artificial Intelligence Applications (CAIA- 90)*. IEEE Computer Society Press, Los Alamitos, US, pp. 320–326, 1990.
4. G. Amati and F. Crestani, "Probabilistic learning for selective dissemination of information," *Information Processing and Management*, vol. 35, no. 5, pp. 633–654, 1999.
5. A. Adomavicius, G. and Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *IEEE Transaction on Knowledge and Data Engineering*, vol. 17, no. 6, pp. 734–749, 2005.
6. M. J. Pazzani and D. Billsus, "Learning and revising user profiles: The identification of interesting web sites," *Machine Learning*, vol. 27, no. 3, pp. 313–331, 1997.
7. 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, pp. 195–204, 2000.
8. M. J. Pazzani and D. Billsus, "Learning and revising user profiles: The identification of interesting web sites," *Machine Learning*, vol. 27, no. 3, pp. 313–331, 1997.