An Automated System to Filter Unwanted Message from OSN User wall

S.B.Madankar[1], Anupama Mishra[2], Apurva Bhandari[3], Reshma Chaudhari[4], Aishwarya Kulkarni[5]

Assistant Professor, Information Technology, PVPIT, Pune, India [1]

Student, Information Technology, PVPIT, Pune, India [2,3,4,5]

Abstract: One fundamental issue in today On-line Social Networks (OSNs) is to give users the ability to control the messages posted on their own private space to avoid that unwanted content is displayed. Up to now OSNs provide little support to this requirement. To overcome this problem, we propose a system allowing OSN users to have a direct control on the messages posted on their walls. This is achieved through a flexible rule-based system, that allows users to customize the filtering criteria to be matter-of-fact to their walls, and a Machine Learning based soft classifier automatically labelling messages in content-based filtering.

Keywords: Facebook, Filtered walls, Machine Learning, Filtering Rules, Text Categorization.

I. INTRODUCTION

Today's modern life is completely based on Internet. Now a day's people cannot imagine life without Internet. From last few years people share their views, ideas, information with each other using social networking sites. Such interchanges might include diverse sorts of substance such as text, image, audio and video data. According to Facebook statistics 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[1]. Information Filtering has been widely used and employed for the textual documents and web contents. However, the goal of this proposal is mainly to provide categorization techniques to give the security to user walls from useless and meaningless data [2]. This is especially for that in OSNs, the users can comment the post in public/private areas of another user walls [6].

These comments can be useless or meaningless or unwanted messages. So, here information Filtering plays a vital role to protect the user walls in OSNs from undesired messages and give the authority to user to automatically control the undesired data on their walls [3]. A System which will give ability to users to control the messages posted on their own private space to avoid unwanted messages displayed. Customizable Filtering Rules are used to filter the unwanted messages from OSNs users' wall as well as Machine Learning approach, Short Text Classification and Black list techniques are applied on Users Wall [6]. The huge and dynamic character of these data creates the premise for the employment of web content mining strategies aimed to automatically discover useful information dormant within the data [2]. This application is useful for common people who don't want to write any unwanted messages like vulgar, political, sexual messages on his / her wall by any third person [3]. OSNs provide very little support to prevent unwanted messages on user walls. For example, Facebook allows users to state who is allowed to insert messages in their

walls (i.e., friends, friends of friends, or defined groups of friends) [1]. However, no content-based preferences are supported and therefore it is not possible to prevent undesired messages, such as political or vulgar ones, no matter of the user who posts them [2]. However, no content-based preferences are supported and therefore it is not possible to prevent undesired messages, such as political or vulgar ones, no matter of the user who posts them. Providing this service is not only a matter of using previously defined web content mining techniques for a different application, rather it requires to design ad hoc classification strategies. This is because wall messages are constituted by short text for which traditional classification methods have serious limitations since short texts do not provide sufficient word occurrences [1] [2].

II. LITERATURE SURVEY

In paper [1], Information filtering is the process of providing appropriate information to the people who need it. It significantly searches for what actually concerns the textual document, specifically web contents, and offers a user with classification mechanism to avoid the unnecessary information. This information filtering process is used in the online social network for insightful objective. To facilitate the content based filtering, this article introduces the filtered wall architecture. It will filter the incoming post based on the content. The main goal of this system is to provide customizable content based message filtering for online social networks, based on machine learning techniques. Information Filtering Systems are designed to categorize the information which are generated dynamically and offer the information to the user fulfil their requirement. In the content Based Filtering system, each user is assumed to operate separately. So the filtering system selects the information based on the correlation between the content of the items and user preferences. To support the content based filtering in online social network, Filtered wall architecture is

introduced. In this architecture, text mining techniques are employed to categorize the incoming messages. Traditional text classification methods have major inadequacy in classifying the short text message. An automated system called filtered wall is designed in this paper to filter unwanted messages from user walls.

In this paper [2], Author have presented a system to filter undesired messages from OSN walls. The system exploits a ML soft classifier to enforce customizable content-dependent FR's. Moreover, the edibility of the system in terms of filtering options is enhanced through the management of BLs. The development of a GUI and a set of related tools to make easier BL and FR specification is also a direction we plan to investigate, since usability is a key requirement for such kind of applications. In particular, we aim at investigating a tool able to automatically recommend trust values for those contacts user does not personally known. We do believe that such a tool should suggest trust value based on users actions, behaviors and reputation in OSN, which might imply to enhance OSN with audit mechanisms. However, the design of these audit-based tools is complicated by several issues, like the implications an audit system might have on users privacy and/or the limitations on what it is possible to audit in current OSNs. A preliminary work in this direction has been done in the context of trust values used for OSN access control purposes. However, we would like to remark that the system proposed in this paper represents just the core set of functionalities needed to provide a sophisticated tool for OSN message filtering. Even if we have complemented our system with an online assistant to set FR thresholds, the development of a complete system easily usable by average OSN users is a wide topic which is out of the scope of the current paper. As such, the developed Facebook application is to be meant as a proof-of-concepts of the system core functionalities, rather than a fully developed system. Moreover, we are aware that a usable GUI could not be enough, representing only the first step. Indeed, the proposed system may suffer of problems similar to those encountered in the specification of OSN privacy settings. In this context, many empirical studies have shown that average OSN users have difficulties in understanding also the simple privacy settings provided by today OSNs. To overcome this problem, a promising trend is to exploit data mining techniques to infer the best privacy preferences to suggest to OSN users, on the basis of the available social network data.

In this paper [3], A system to prevent the indecent messages from the Social Networking site walls has been presented. The Usage of Machine Learning has given higher results to the system to trace the messages and the users to distinguish between the good and bad messages and the authorized and unauthorized users in the Social Networking User Profiles automatically. Thus the Machine Learning Technique plays a vital role in this paper in order to generate the blacklist of the bad words and the unauthorized users. The user has to update his privacy setting in his account in order to add this method

to prevent the obscenity in his public profile. In this context, a statistical analysis has been conducted to provide the usage of the good and bad words by the persons in the sites. Overall, the obscenity of the users has been prevented. The Machine Learning is a system which can learn from the data and take decisions based on the learned data. the Machine Learning here traces the posted messages for the good and the illegal words used in the wall by the public users. FRs should allow users to state constraints on message creators. The 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. A further component of our system is a Blacklist (BL) mechanism to avoid messages from undesired creators, independent from their contents. BL is 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.

A system automatically filters unwanted messages using the blacklists on the basis of both message content and the message creator relationships and characteristics. Major difference include , a different semantics for filtering rules to better fit the considered domain, to help the users Filtering Rules(FRs) specification, the extension of the set of features considered in the classification process.

In this paper [4], A system to filter unwanted message in OSN wall is presented. The first step of the project is to classify the content using several rules. Next step is to filter the undesired rules. Finally Blacklist rule is implemented. So that owner of the user can insert the user who posts undesired messages. Better privacy is given to the OSN wall using our system. In future Work, we plan to implement the filtering rules with the aim of bypassing the filtering system, it can be used only for the purpose of overcome the filtering system. In this paper, Blacklist mechanism is used, where the user's list will be avoided for the moment to post on user wall. In this paper, all classification and filtering rules will be included, additionally BL rule is used. Based on the user wall and relationship, the owner of the wall can block the user. This prohibition can be approved for an uncertain period of time.

Aim of the short text classifier is to recognize and eradicate the neutral sentences and categorize the non neutral sentences in step by step, not in single step. This classifier will be used in hierarchical strategy.

Representing the text of a document is critical, which will affect the classification performance. Many features are there for representation of text, but we judge three types of features. BOW, Document properties (DP) and contextual features. Filtering rules will be applied, when a user profile does not hold value for attributes submitted by a FR. BL rule, owner can identify which user should be blocked based on the relationship in OSN and the user's profile. The user may have bad opinion about the users can be banned for an uncertain time period.

III. MOTIVATION AND PROBLEM STATEMENT

Indeed, today OSNs provide very little support to prevent unwanted messages on user walls. For example, Facebook allows users to state who is allowed to insert messages in their walls [2] [3]. However, no content-based preferences are supported and therefore it is not possible to prevent undesired messages, such as political or vulgar ones, no matter of the user who posts them [1]. However, no content-based preferences are supported and therefore it is not possible to prevent undesired messages, such as political or vulgar ones, no matter of the user who posts them [6]. Providing this service is not only a matter of using previously defined web content mining techniques for a different application, rather it requires to design ad hoc classification strategies. This is because wall messages are constituted by short text for which traditional classification methods have serious limitations since short texts do not provide sufficient word occurrences. It inspects every message before rendering the message to the intended recipients and makes immediate decision on whether or not the message under inspection should be dropped [2] [3]. The application of content-based filtering on messages posted on OSN user walls poses further challenges given the short length of those messages apart from the wide range of topics that may be mentioned. Short text classification has received up to currently little attention within the scientific community. Recent work selects difficulties in shaping robust options, basically result of the very fact that the description of the short text crisp, with several wrong spellings, non-standards terms, and noise [8]. Our work is additionally galvanized by the various access management's models and connected policy languages and social control mechanisms that are projected to date for OSNs since filtering shares many similarities with access management. 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 model [7].

IV. PROPOSED SYSTEM

An automated system called filtering wall that is able to filter unwanted messages from OSN user walls. We exploit machine learning text categorization techniques to automatically assign with each short text message a set of categories based on its content. Our contribution is that we are going to implement real time system using facebook app. The project is to develop a system that is going to block the unwanted messages from OSN user's wall. Now we are implementing the software which is going to work for filtering messages/comments in the form of a text, so in future we can extend our project scope to filter images, audio, video format or filtering. Paragraphs must be justified, i.e. both left-justified and right-justified.

A. Objectives
The objective of our system investigates the utility of linguistic features for detecting the sentiment of the posts done on person's time line. We will use Information

filtering techniques to remove unwanted contents by using customizable content based filtering rules, Machine learning approach
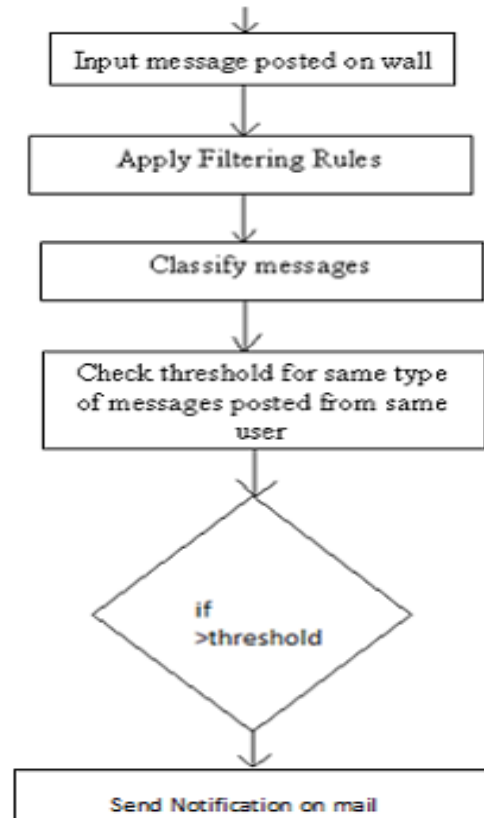
V. DATA FLOW DIAGRAM

Fig 1.Data Flow Diagram

In DFD, the figure shows that the users messages is posted on facebook by using facebook 4jAPI.Here,the users details and the messages get stored into database. Then the OSN message filtering system will fetch the data from the database and perform analysis and pre-processing on data and will find sense of post and according to that the message will be posted on hidden [3].
All these levels or activities are explained briefly in the system architecture.

VI. SYSTEM ARCHITECTURE

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 to automatically assign with each short text message a set of categories based on its content [1] [2] [3]. Fig.2 shows the system architecture also known as general block diagram.

The major efforts in building a robust short text classifier (STC) are concentrated in the extraction and selection of a set of characterizing and discriminates features [2]. The solutions investigated in this paper are an extension of those adopted in a previous work by us from whom we

inherit the learning model and the elicitation procedure for generating reclassified data. The original set of features, derived from endogenous properties of short texts, is enlarged here including exogenous knowledge related to the context from which the messages originate [7]. As far as the learning model is concerned, we confirm in the current paper the use of neural learning which is today recognized as one of the most efficient solutions in text classification.
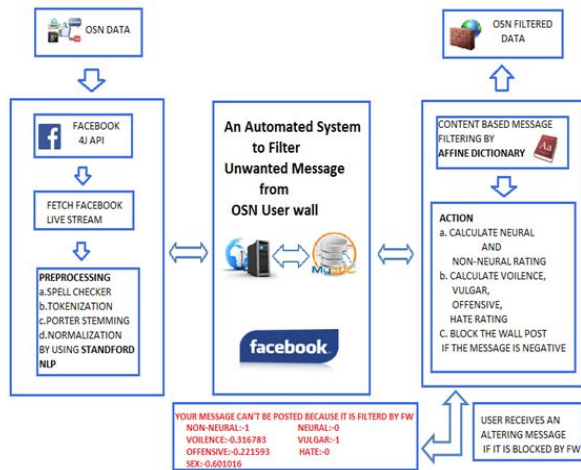

Fig 2 System Architecture

In particular, we base the overall short text classification strategy on Radial Basis Function Networks (RBFN) for their proven capabilities in acting as soft classifiers, in managing noisy data and intrinsically vague classes [6]. Moreover, the speed in performing the learning phase creates the premise for an adequate use in OSN domains, as well as facilitates the experimental evaluation tasks. We insert the neural model within a hierarchical two level classification strategy [3]. In the first level, the RBFN categorizes short messages as Neutral and Non-neutral; in the second stage, Non-neutral messages are classified producing gradual estimates of appropriateness to each of the considered category. Besides classification facilities, the system provides a powerful rule layer exploiting a flexible language to specify Filtering Rules (FRs), by which users can state what contents, should not be displayed on their walls [2] [3] [6]. FRs can support a variety of different filtering criteria that can be combined and customized according to the user needs. More precisely, FRs. exploit user profiles, user relationships as well as the output of the ML categorization process to state the filtering criteria to be enforced [6]. In addition, the system provides the support for user-defined Blacklists (BLs), that is, lists of users that are temporarily prevented to post any kind of messages on a user wall [7].

OSN have five important components:

- **Content-Based Messages Filtering (CBMF):** For content-Based Massages filtering, we first filter out duplicate tweets and facebook comments, non-English tweets and non English facebook comments, and tweets that do not contain hash tags. From the remaining set (about 4 million), we investigate the distribution of hash tags and identify what we hope will be sets of frequent hash tags that are indicative of positive, negative and neutral messages. These hash tags are used to select the tweets that will be used for development and training [1].

- **Short Text Classifier:** Designing and evaluating various representation techniques in combination with a neural learning strategy to semantically categorize short texts.

- **Integrate the System with facebook:** The system will integrate with facebook and able to read the real time posts from users wall.

- **Access Token Generation:** As soon as the user logs in to facebook, the access token will be generated for that particular user.

- **Post Reading from User Wall and Analysis :**
a. With the help of that access token, the system will be able to read all the posts from users timeline.
b. Preprocessing and NLP

- **Pre Processing :**
Tokenization: First of all we did the tokenization by which sentences are split into the words.
Normalization: After that we used Stanford NLP to remove stop words from all the words.
Part-of-speech (POS) tagging: Detects if the word token is noun, verb, and adjective [2].

- **NLP and Feature Extraction :**
1. Apply Stanford NLP to separate part of speech from the sentence.
2. Porter Stemmer Algorithm will be applied for getting root of the word for adjectives.
3. After getting root of the word, we will compare weight / sense of each word with the affine dictionary.
4. Finding negative annotations in the sentence and reverse the weight.
5. Calculate overall weight using emoticons approach.
6. Sum up both to draw final conclusion.
7. Finally, positive, negative or neutral count for that particular post will be calculated.

- **Action on Post:** After the analysis, the action on the posts will be taken accordingly, whether to publish the post or not on the users wall. If found negative sense, the system won't allow user make the posts his/her friends wall. In case of real time fetched posts, the system will either delete or hide the posts depending on the user's choice.

- **Maintenance:** As said above, the access token will get expired after two months, the user will just needs to log in with facebook once in two months

- **Design and Implementation Constraints :**
1. FB Login: User should login with his facebook account through the system for getting the access token needed by the system.
2. Access Token Renewal: As per FACEBOOKs constraints, User needs to login with the facebook once in two months two renew it.

VII. ALGORITHM

Algorithm: Porter Stemming Algorithm

Here an algorithm for suffix stripping

Input: Plurals words and -ed or -ing suffixes.
Output: Words suffix stripping
Begin
Step 1: Gets rid of plurals and -ed or -ing suffixes.
Step 2: Turns terminal y to i when there is another vowel in the stem.
Step 3: Maps double suffixes to single ones:
    -ization, -ational, etc.
Step 4: Deals with suffixes, -full, -ness etc.
Step 5: Takes o_ -ant, -ence, etc.
Step 6: Removes a final -e.
End

Removing suffixes by automatic means is an operation which is especially useful in the field of information retrieval. In a typical IR environment, one has a collection of documents, each described by the words in the document title and possibly by words in the document abstract. Ignoring the issue of precisely where the words originate, we can say that a document is represented by a vector of words, or terms. Terms with a common stem will usually have similar meanings, for example:

CONNECT
CONNECTED
CONNECTING
CONNECTION
CONNECTIONS

Frequently, the performance of an IR system will be improved if term groups such as this are conflated into a single term. This may be done by removal of the various suffixes -ED,
-ING, -ION, IONS to leave the single term CONNECT. In addition, the suffix stripping process will reduce the total number of terms in the IR system, and hence reduce the size and complexity of the data in the system, which is always advantageous.

VIII. CONCLUSION

The In this paper, we have presented a system to filter out unwanted messages from OSN user walls. The system exploits a Machine Learning soft classifier to enforce customizable content-depended filtering rules. The flexibility of the system in terms of filtering options is enhanced trough the management of BLs
This is the first step of a wider project. The early encouraging results we have obtained on the classification procedure prompt us to continue with other work that will aim to improve the quality of classification. Additionally, we plan to enhance our filtering rule system, with a more sophisticated approach to manage those messages caught just for the tolerance and to decide when a user should be inserted into a BL. In this paper, we proposed a system with the flexible rules to filter the unwanted messages posted on user wall. After crossing threshold value the notification message is send to that user. This allows users to customize the refining criteria to be applied to their walls, and a machine learning-based classifier automatically classifies the messages and labelling messages in support of content-based filtering.

REFERENCES

[1] N. Thilagavathi, R. Taarika "Content Based Filtering in Online Social Network using Inference Algorithm", 2014 International Conference on Circuit, Power and Computing Technologies [ICCPCT].

[2] K.Babu , P.Charles "A System to Filter Unwanted Words Using Blacklists in Social Networks", (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (2) , 2014.

[3] Marco Vanetti, Elisabetta Binaghi, Elena Ferrari, Barbara Carminati, and Moreno "A System to Filter Unwanted Messages from OSN User Walls" IEEE Transaction on knowledge and data engineering, vol. 25, no. 2, February 2013.

[4] Facebook development platform. http://developers.facebook.com

[5] Ms. Shruti C. Belsare, Prof. R.R. Keole, "OSN user filtered walls for unwanted messages using content mining" march-2014.

[6] Dipali D. Vidhate, Ajay P. Thakare,"To Avoid Unwanted Messages from OSN User Wall: Content Based Filtering Approach", IJCSMC, vol.3, Issue.4, April 2014.

[7] M. Vanetti, E. Binaghi, B. Carminati, M. Carullo and E. Ferrari "Content-based Filtering in On-line Social Networks" 2010.

[8] Dactchayani Sukumar1, Maniraj .S.P "Context Aware Message Filtering in OSN user Walls" IJIRCCE, Vol. 2, Issue 3, March 2014.

[9] Rashmi R. Atkare, P.D.Soni "Survey of Filtering System For OSN (Online Social Networks)", IJCTA, Vol 4 (6), Dec 2013.

[10] Aarti Supe, R.N. Phursule "A System to Clarify Unwanted Messages from OSN User Surface", International Journal of Scientific Engineering and Research (IJSER), Vol. 3, Issue 5, May 2015.

[11] Bhushan Jagtap, Rahul Rathod, Vitthal Anpat, Utsav "A System to Filter Unwanted Messages from OSN User Walls", International Journal of Advance Foundation and Research in Computer (IJAFRC), vol. 2, Special Issue, January 2015.

[12] Vikrant Sanghavi, Amol Nanaware, Divya Nadar, Chitra Bhole "A System to Filter Unwanted Messages from OSN User Wall", International Journal of Research in Advent Technology, Vol. 1, Issue 5, December 2013.

[13] Mayuri Uttarwar, Yogesh Bhute "A Review on Customizable Content-Based Message Filtering from OSN User Wall", International Journal of Computer Science and Mobile Computing (IJCSMC), Vol. 2, Issue 10, October 2013.

[14] Aishwarya Kashyup, Himanshu Kothari, Ravina Kothari, Geetika Narang "Filtering Mechanism for Unwanted Messages on OSN User Wall Using CBMF(Content Based Message Filtering)", International Journal of Advanced Research in Computer and Communication Engineering(IJARCCE), Vol. 4, Issue 2,February 2015.

[15] Amruta Kachole, S.D.Jondhale "Unwanted Messages Filtering System from OSN User's Wall Using Customizable Filtering Rules and Black list Techniques", International Journal of Emerging Technology and Advanced Engineering(IJETAE), Vol. 4, Issue 2, February 2014.

[16] Sruthi. T, Greeshma T.R. "Filtered Wall-An Online Social Network Filter", International Journal of Science, Engineering and Technology Research (IJSETR), Vol. 3, Issue 4, April 2014.

[17] K. Arulmurugan, P. Ranjitkumar "An Analysis of Unwanted Messages Filtering Methods from OSN User Walls", International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE), Vol. 2, Issue 1, March 2014.

[18] J.Anishya Rose, A.Pravin "Machine Learning Text Categorization in OSN to Filter Unwanted Messages", International Journal of Computer Science and Information Technologies (IJCSIT), Vol 5(1), 2014.