International Journal of Ethics in Engineering & Management Education
Website: www.ijeee.in (ISSN: 2348-4748, Volume 2, Issue 12, December 2015)

Filtering Unwanted Messages from Online Social Networks User Walls

V. Ramesh
Assistant Professor , Dept of CSE
Sri Indu Institute of Engineering and Technology
Ibrahimpatan, TS, India

G. Raju
Assistant Professor , Dept of CSE
Sri Indu Institute of Engineering and Technology
Ibrahimpatan, TS, India

***Abstract* - Users have ability to keep in touch with his/her friends by exchanging different types of information or messages like text, audio and video data. Today's OSNs (Online Social Network System) do not provide much support to the users to avoid unwanted messages displayed on their own private space called in general wall. So, in this paper we present OSNs system which gives 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.**

***Index Terms* - On-line Social Network, Information filtering, short text classification.**

I. INTRODUCTION

Today's modern life is totally based on Internet. Now a days people cannot imagine life without Internet. Also, OSNs are just a part of modern life. From last few years people share their views, ideas, information with each other using social networking sites. Such communications may involve different types of contents like text, image, audio and video data. But, in today's OSN , there is a very high chance of posting unwanted content on particular public/private areas, called in general walls. So, to control this type of activity and prevent the unwanted messages which are written on user's wall we can implement filtering rules (FR) in our system. Also, Black List (BL) will maintain in this system .We present this system as www.winow.in on the internet. It can be used to give users the ability to automatically control the messages written on their own walls, by filtering out unwanted messages. 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. OSNs provide 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). 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.

II. RELATED WORK

In www.winow.in information filtering techniques are used to remove unwanted contents by using customizable content based filtering rules, Machine learning approach; according to user's interest and recommends an item. Recommender systems works in following ways

- Content based filtering
- Collaborative filtering
- Policy based filtering

A. ***Content-based filtering:*** In content based filtering to check the user's interest and previous activity as well as item uses by users best match is found [10]. For example OSNs such as Facebook, Orkut used content based filtering policy. In that by checking users profile attributes like education, work area, hobbies etc. suggested friend request may send. The main purpose of content based filtering, the system is able to learn from user's actions related to a particular content source and use them for other content types.

B. ***Collaborative filtering:*** In collaborative filtering information will be selected on the basis of user's preferences, actions, predicts, likes, and dislikes. Match all this information with other users to find out similar items. Large dataset is required for collaborative filtering system. According to user's likes and dislikes items are rated.

C. ***Policy-based filtering:*** In policy based filtering system users filtering ability is represented to filter wall messages according to filtering criteria of the user. Twitter is the best example for policy based filtering.[1] In that communication policy can be defines between two communicating parties.

We believe that this is a key OSN service that has not been provided so far. Indeed, today OSNs provide very little support to prevent unwanted messages on user walls. For example, Face book allows users to state who is allowed to insert messages in their walls (i.e., friends, friends of friends, or defined groups of friends). However, no content-based preferences are supported and therefore it is not possible to prevent undesired messages, such as political or vulgar ones,

International Journal of Ethics in Engineering & Management Education
Website: www.ijeee.in (ISSN: 2348-4748, Volume 2, Issue 12, December 2015)

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.

III. ARCHITECTURE OF FILTERED WALL

In general, the architecture in support of OSN services is a three-tier configuration. The initial layer generally aims to offer the essential OSN functionalities (i.e., profile and relationship administration). In addition, some OSNs offer an extra layer allowing the support of external Social Network Applications (SNA)1. Finally, the supported SNA may require an additional layer for their needed graphical user interfaces (GUIs). According to this orientation layered structural plan, the planned system has to be positioned in the second and third layers (Figure 1), as it can be considered as a SNA. Particularly, users cooperate with the system by means of a GUI setting up their filtering laws, along with which messages have to be filtered out. In addition, the GUI offers users with a FW that is a wall where only messages that are authorized according to their filtering rules are published. The core components of the proposed system are the Content-Based Messages Filtering (CBMF) and the Short Text Classifier elements. The latter element aims to categorize messages according to a set of categories. In compare, the first element exploits the message categorization offered by the STC module to implement the FRs specified by the user. As graphically illustrated in Fig. 1.

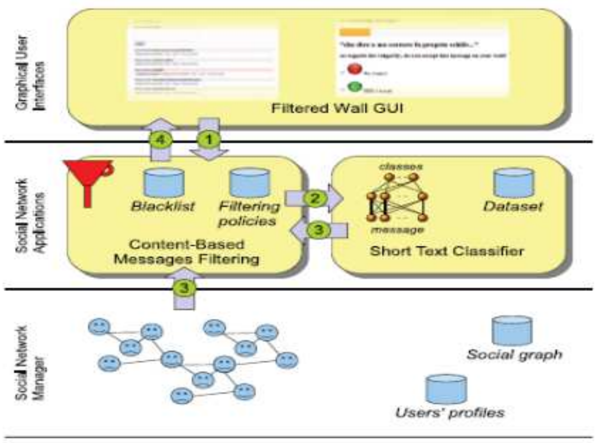


Fig.1. Architecture of Filtered wall

The path pursued by a message, it can be summarized as follows:

1. After entering the private wall of one of his/her associates, the user attempts to post a message, which is captured by FW.
2. A ML-based text classifier extracts metadata from the content of the message.
3. FW uses metadata provided by the classifier, mutually with data extorted from the social graph and users' profiles, to implement the filtering and BL rules. 4. Depending on the result of the previous step, the message will be available or filtered by FW.

IV. MODULES DESCRIPTION

Modules:

1. Filtering rules
2. Online setup assistant for FRs thresholds
3. Blacklist

1. *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. All these options are formalized by the notion of creator specification, defined as follows.

2. *Online setup assistant for FRs thresholds:* As mentioned in the previous section, 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 discussed in Section VI-A. 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.

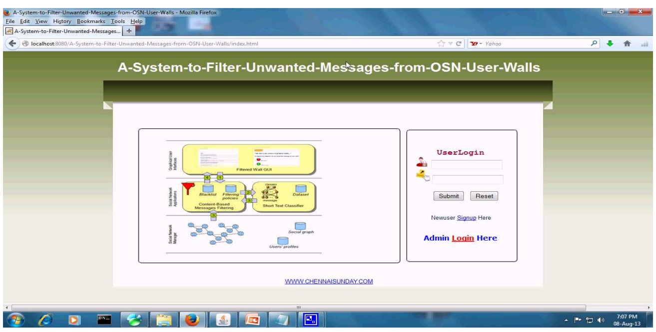
3. *Blacklist:* A further component of our system is a BL mechanism to avoid messages from undesired creators, independent from their contents. BLs are directly managed by the system, which should be able to determine who are the users to be inserted in the BL and decide when users retention in the BL is finished. To enhance flexibility, such information are given to the system through a set of rules, hereafter called BL rules. Such rules are not defined by the SNM, therefore

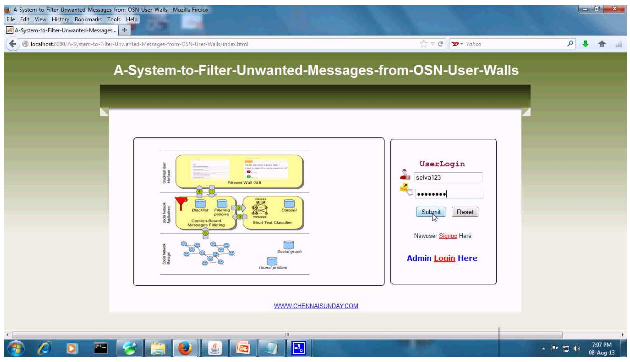
International Journal of Ethics in Engineering & Management Education
Website: www.ijeee.in (ISSN: 2348-4748, Volume 2, Issue 12, December 2015)

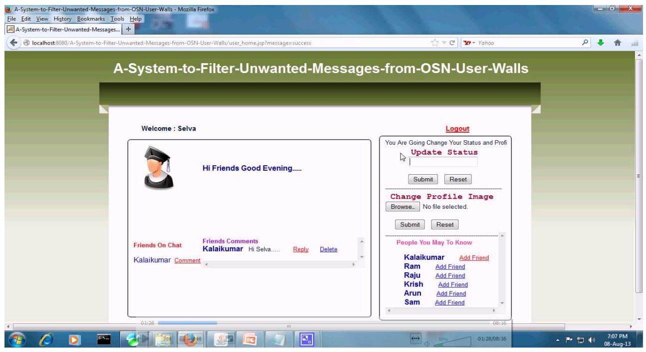
they are not meant as general high level directives to be applied to the whole community. Rather, we decide to let the users themselves, i.e., the wall's owners to specify BL rules regulating who has to be banned from their walls and for how long. Therefore, a user might be banned from a wall, by, at the same time, being able to post in other walls. Similar to FRs, 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 OSN. Therefore, by means of a BL rule, wall owners are for example able to ban from their walls users they do not directly know (i.e., with which they have only indirect relationships), or users 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 OSN. More precisely, among possible information denoting users' bad behavior we have focused on two main measures. The first is related to the principle that if within a given time interval a user has been inserted into a BL for several times, say greater than a given threshold, he/she might deserve to stay in the BL for another while, as his/her behavior is not improved. This principle works for those users that have been already inserted in the considered BL at least one time. In contrast, to catch new bad behaviors, we use the Relative Frequency (RF) that let the system be able to detect those users whose messages continue to fail the FRs. The two measures can be computed either locally, that is, by considering only the messages and/or the BL of the user specifying the BL rule or globally, that is, by considering all OSN users walls and/or BLs.

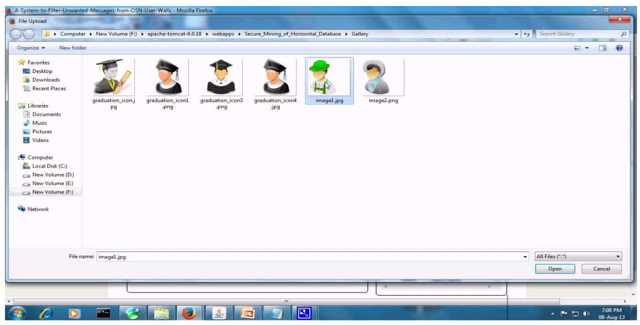
V. EXPERIMENTAL RESULTS

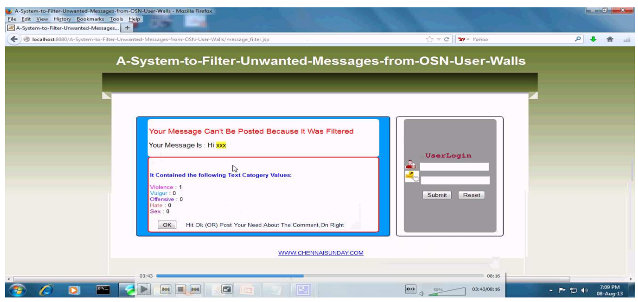
- We present www.winnow.in OSN site with basic functionalities of OSNs.
- In this system, using Filtering Rules we can make Filter Wall for preventing unwanted messages. Initially, we focus on Violance, Vulgar, Sextual Offensive, Hate type of messages and filter these messages.
- Also, maintain Black list for the user who will send the prevented type of messages more than three times then that user will automatically put into Black List.
- Administer can see monthly and yearly reports as well as Graphs like which category of messages are filtered (in percentage), who is message creator .



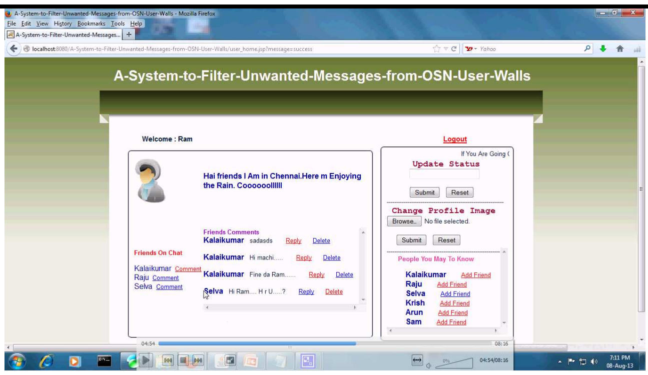


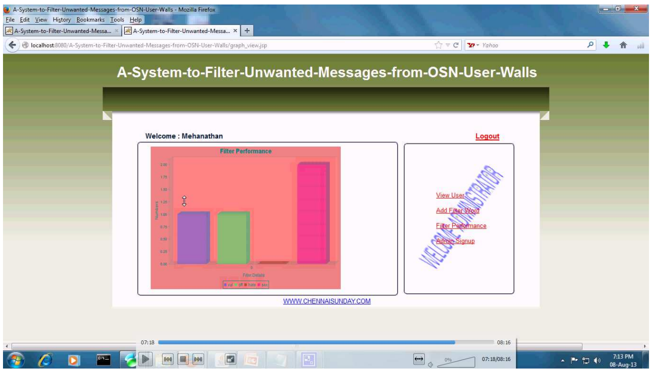






International Journal of Ethics in Engineering & Management Education
Website: www.ijeee.in (ISSN: 2348-4748, Volume 2, Issue 12, December 2015)





VI. CONCLUSION

Existing system is used to filter undesired messages from OSNs wall using customizable filtering rules (FR) enhancing through Black lists (BLs). In present system (www.winow.in), we are more focus on an investigation of two interdependent tasks in depth. This system approach decides when user should be inserted into a black list. The system developed GUI and a set of tools which make BLs and FRs specifications more simple and easy. Investigation tools may be able to automatically recommend trust value of the user. The primary work of this system is to find out trust values used for OSN access control. In this system we will provide only core set of functionalities which are available in current OSNs like Facebook, Orkut, Twitter,etc. In existing OSNs have some difficulties in understanding to the average users regarding privacy settings. But this problem will be overcome in present OSNs system.

REFERENCES

[1]. 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.
[2]. M. Chau and H. Chen, "A machine learning approach to web page filtering using content and structure analysis," Decision Support Systems, vol. 44, no. 2, pp. 482–494, 2008.
[3]. 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.
[4]. F. Sebastiani, "Machine learning in automated text categorization,"ACM Computing Surveys, vol. 34, no. 1, pp. 1–47, 2002.
[5]. M. Vanetti, E. Binaghi, B. Carminati, M. Carullo, and E. Ferrari, "Content-based filtering in on-line social networks," in Proceedings of ECML/PKDD Workshop on Privacy and Security issues in Data Mining and Machine Learning (PSDML 2010), 2010.
[6]. N. J. Belkin and W. B. Croft, "Information filtering and informationretrieval: Two sides of the same coin?" Communications of the ACM, vol. 35, no. 12, pp. 29–38, 1992.
[7]. P. J. Denning, "Electronic junk," Communications of the ACM, vol. 25, no. 3, pp. 163–165, 1982.
[8]. P. W. Foltz and S. T. Dumais, "Personalized information delivery: An analysis of information filtering methods," Communications of the ACM, vol. 35, no. 12, pp. 51–60, 1992.
[9]. P. S. Jacobs and L. F. Rau, "Scisor: Extracting information from online news," Communications of the ACM, vol. 33, no. 11, pp. 88–97,1990.
[10]. S. Pollock, "A rule-based message filtering system," ACM Transactions on Office Information Systems, vol. 6, no. 3, pp. 232–254,
1988.

About the authors:

V. Ramesh, Currently working as an Assistant professor in Dept of CSE in Sri Indu College of Engineering & Technology, Ibrahimpatan, Hyderabad, TS, India.

G. Raju, Currently working as an Assistant professor in Dept of CSE in Sri Indu College of Engineering & Technology, Ibrahimpatan, Hyderabad, TS, India.