

An Approach for Filtering of Unwanted Messages from Social Networks wall using Blacklist Technique

Minal S. Mahure, Prof. P. L. Ramteke

CSIT, Shree Hanuman Vyayam Prasarak Mandal College of Engineering & Technology, Amravati, Maharashtra, India

ABSTRACT

In recent years, Online Social Networks (OSNs) have become an important part of daily life. Users build explicit networks to represent their social relationships. Users can upload and share information related to their personal lives. The potential privacy risks of such behavior are often ignored. In addition, the fundamental issue in today Online Social Networks is to give users the ability to control the messages posted on their own private space to avoid that unwanted content is displayed. Today OSNs provide very little support to prevent unwanted messages on user walls. For that purpose, we proposed a new system allowing OSN users to have a direct control on the messages posted on their walls. We design the system where message are filter on the basses of its contents by using contents base filtering and blacklist technique. So that unwanted messages will not be post on user wall, also that unwanted message will be block.

Keywords : Filter Wall, Safe Mode, Content Base Filtering, Filtering of Images

I. INTRODUCTION

Information and communication technology plays a significant role in today's networked society. It has affected the online interaction between users, who are aware of security applications and their implications on personal privacy. There is a need to develop more security mechanisms for different communication technologies, particularly social networks. social networks provide very little support to prevent unwanted messages on user walls. With the lack of classification or filtering tools, the user receives all messages posted by the users he follows. In most cases, the user receives a noisy stream of updates. We were designing an information Filtering system. Information filtering has been greatly explored for what concerns textual documents and, more recently, web content (e.g., [2], [3], [4]).

In social network, information filtering can also be used for a different, more sensitive, purpose. This is due to the fact that in social networking there is the possibility of posting or commenting other posts on particular public/private areas, called in general walls. Information filtering can therefore be used to give users the ability to

automatically control the messages written on their own walls, by filtering out unwanted messages.[1]

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 social network user walls. We exploit Machine Learning (ML) text categorization techniques [5] to automatically assign with each short text message a set of categories based on its content. 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 [6]. 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 [7], [8].

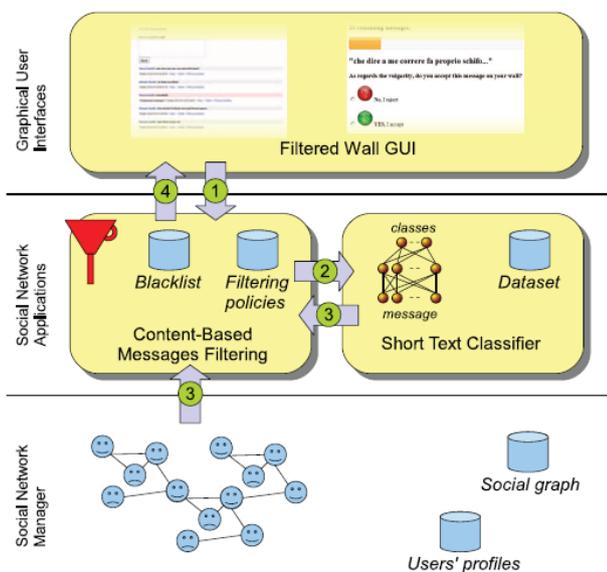


Figure 1. Filtered wall conceptual architecture and the flow messages follow, from writing to publication [1]

II. METHODS AND MATERIAL

Literature Review

Marco Vanetti, Elisabetta Binaghi, Elena Ferrari, Barbara Carminati, and Moreno Carullo propose technique “A System to Filter Unwanted Messages from OSN User Walls” in which 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 applied to their walls, and a Machine Learning-based soft classifier automatically labeling messages in support of content-based filtering. 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 [4] to automatically assign with each short text message a set of categories based on its content[1].

K.Babu , P.Charles Department of Computer Science, MRK Institute of Technology Kattumannarkoil propose technique “A System to Filter Unwanted Words Using Blacklists In Social Networks” in which We used Machine Learning text categorization technique to automatically categorize each short text messages based on its content. We base the overall short 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. We use the neural model RBFN categorizes as Neural and Non-neural FR filtering rules by which it can state what contents should not be displayed on their walls. In addition, the system provides the user defined Blacklists that is mainly used to prevent to post any kind of message on a user wall[9] In paper [10], 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.

Proposed work:

We are going to propose a system in which only register user can make communication. If anyone posts an unwanted message then the system automatically detect it and restrict the user from post such messages. Now new implementation is system will give limited chance to the user if he or she crosses the restriction policies then system will deactivate the post and on one able to comment on that post. This deactivated post will transfer to the account of admin. Then admin will decide where the post is unwanted or its good. Admin will delete the post or reactive the post.

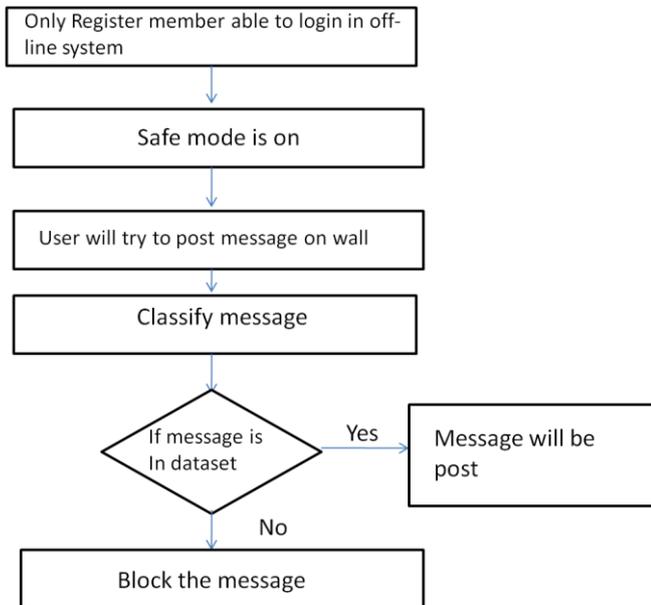


Figure 2. Flow of proposed system

General Flow

- Registration of Users.
- Login Authentications.
- Posting Messages and Contents.
- Checking of vulgarness and obstacles within conversions.
- Noticing to user in case of vulgar messages.
- Blocking Mechanisms.
- Admin will check Post which is an image post.
- Admin will delete or reactive post.

Techniques:

Content Based Filters

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. The application of content-based filtering on messages posted on OSN user walls poses additional challenges given the short length of these messages other than the wide range of topics.

Text Classifier

Text classification has received up to now little attention in the scientific community. Established techniques

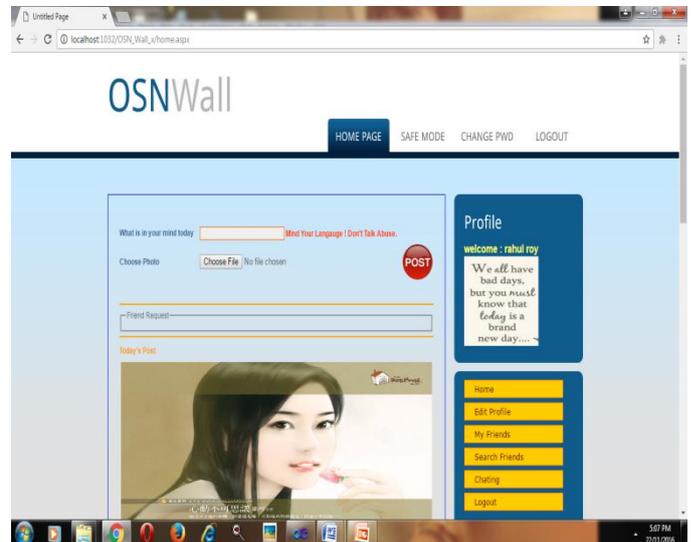
used for text classification work well on data sets with large posts. Recent work highlights difficulties in defining robust features, essentially due to the fact that the description of the short text is concise, with many misspellings, nonstandard terms, and noise.

Blacklist Rule

Here we r defining some rules for filtering of private wall area of user in which user can access filtering wall faction. Its depends on user where to on or off the functioning of filtering wall, here we give the option of safe mode to the when the safe mode is on no one can able to write any unwanted messages on the wall. If user off the safe mode anyone can post unwanted messages. Also we have the list of various shorts text which is consider as unwanted or abusing store in dataset. Where every time any user try to post message on wall it will check the blacklist rule store in dataset.

III. RESULTS AND DISCUSSION

If anyone try to post unwanted message system will block that message and there will be message appear as shown in fig below.



When any user try post unwanted comment on image it will be block the comment but at the same time if the blacklist rule limit will excide the post will deactivate and will go to the admin. Admin will reactive or delete the post base on the content of the image.



IV. CONCLUSION

We describe our work to provide unwanted message filtering for social networks. We have presented a system to filter undesired messages from OSN walls. We would like to remark that the system proposed represents just the core set of functionalities needed to provide a sophisticated tool for social network message filtering. Additionally, we studied strategies and techniques limiting the inferences that a user can do on the enforced filtering rules with the aim of bypassing the filtering system. We were applying this filtering rule on English language. As well as we are going to filter unwanted picture from user wall as wall. In future by using filtering rule can be used for regional language like Hindi Marathi Etc.

V. 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, FEBRUARY 2013

[2] A. Adomavicius and G. Tuzhilin, "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions," IEEE Trans. Knowledge and Data Eng., vol. 17, no. 6, pp. 734-749, June 2005.

[3] 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.

[4] R.J. Mooney and L. Roy, "Content-Based Book Recommending Using Learning for Text

Categorization," Proc. Fifth ACM Conf. Digital Libraries, pp. 195-204, 2000

[5] F. Sebastiani, "Machine Learning in Automated Text Categorization," ACM Computing Surveys, vol. 34, no. 1, pp. 1-47, 2002

[6] 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.

[7] B. Sriram, D. Fuhry, E. Demir, H. Ferhatosmanoglu, and M. Demirbas, "Short Text Classification in Twitter to Improve Information Filtering," Proc. 33rd Int'l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR '10), pp. 841-842,2010.

[8] K. Strater and H. Richter, "Examining Privacy and Disclosure in a Social Networking Community," Proc. Third Symp. Usable Privacy and Security (SOUPS '07), pp. 157-158, 2007.

[9] K.Babu , P.Charles Department of Computer Science, MRK Institute of Technology Kattumannarkoil "A System to Filter Unwanted Words Using Blacklists In Social Networks" International Journal of Computer Science and Information Technologies,1748-1753,2014.

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