# A PROBABILISTIC APPROACH TO METADATA FILTER IN SOCIAL NETWORK

S.Lavanya, Student,
Department of Computer Science & Engineering
IFET College of Engineering,
Villupuram, India
lavanya2193@gmail.com

P.Kanimozhi, Asst. Prof.
Department of Computer Science & Engineering
IFET College of Engineering,
Villupuram, India
ifetcse198@gmail.com,

***ABSTRACT-*** **The proposed application model for meta data keyword filter is a technique to monitor the user activities in social networks such as facebook and forum. The application has a background watcher which has the set of keywords added by the admin. The admin can add the list of crude or unkind words. When the user post a message the background watcher monitors the post and checks whether any crude or unkind word is in the message. If any appropriate content is deducted the message is banned by the background wall filter. The proposed approach is very accurate and efficient improving upon existing methods in terms of accuracy and efficiency in different settings. The application raises a warning message to the user who forward the unkind words to some other. If the user continues such behavior the respective user is blocked permanently.**

***Keywords— Online social networks, filter blacklisted words, warning information and background watcher***

## I. INTRODUCTION

The best entertainment for the younger generation now is given in the form of Social Networking sites. The Online Social Networks (OSN) mainly helps an individual to share new experiences with their family, friends and society. Nowadays, the OSNs are facing the crisis of the people posting the indecent messages on any individual's wall which annoys other people on seeing those indecent words. In order to sort and ban those awful messages a system called Background Watcher is introduced. The aim of the present work is therefore to propose and experimentally evaluate an automated system, called Background Watcher (BW), able to filter unwanted messages from OSN user walls. The major efforts in building a robust Background Watcher (BW) is to automatically banned crude words by blacklisting process.

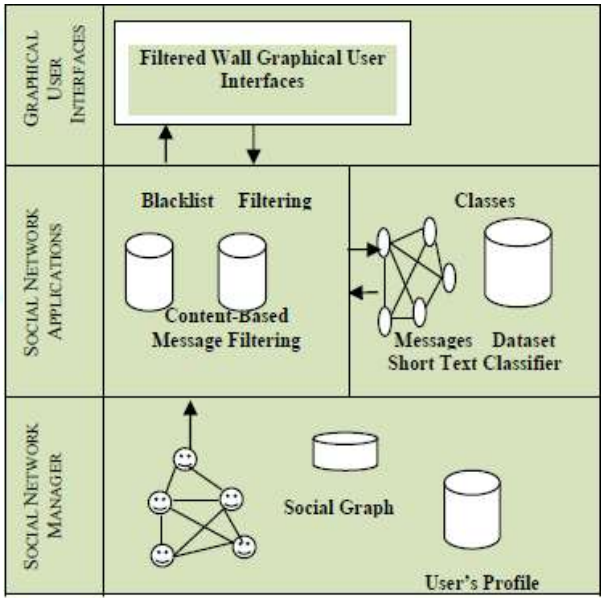


Fig 1: OSN general architecture

OSN as a possibility for posting message on others wall. Background watcher as a ability to control the message written on their wall by blacklisting the unwanted message. For example face book allows the user to post any kind of message in friend's wall. To avoid posting crude words the background watcher tracks the user message and blocks the message containing blacklisted words

With the flooding of crude information on the

Internet, the most important research areas in network information security is how to keep people away from that disgusting information. To filter the hurtful text the keyword matching is used in content based. According to the results, a new crude text filtering model based on reconfirming is put forward in metadata based.

## II. RELATED WORK

F.Sebastiani[1] describes the automatically categorizing text into predefined categories. This is due to the booming interest in last ten years, the documents in digital form are increasingly available and as to be organized. This survey discusses that categorizing text fall within machine learning paradigm. In research community the machine learning as a problem in the process of automatically building a classifier from a set of predefined document, further it as a problem in document representation, classifier construction, and classifier evaluation .the advantage of this approach are very good effective in knowledge engineering approach by saving labor power and straightforward process in different domain

M.Vanetti, E.Binaghi, B. Carminati, M. Carullo, and E. Ferrari [2] this paper proposes a system that describes the content-based message filtering in on-line social networks . In this system the users have a direct control on the messages posted on their friends walls. This is attained by flexible rule-based system, that allows a user to filter the customized words in message and a machine is automatically in support of content-based filtering for labeling the messages.

Fabrizio silvestri[3] describes the mining query logs for turning search usage data into knowledge . The primary focus of this survey is on introducing query mining and to extract useful information from the infinite source of information. The web search engine stores log information about the user since the start of the operation. This shows how search application may benefit the analysis by analyzing the application of query log mining and influence on user experience.

K.babu[4] describes the categorizing of words by machine learning and to check the message containing any crude words. The messages are classified by short text classifier. The short text classifier (STC) concentrate on the set characterizing and discriminating features by extracting and selecting the text. If the message contains any crude words the machine learning sent it to blacklist to filter the crude words from the message and finally post the message in user wall without any crude words , it is based on the results of content-based-technique.

## III. EXISTING SYSTEM

In the existing system the problem is sophisticated, implementation cost too high and limited input. Today, technology is moving towards the future environment. There is no predefined method or control to watch or to track the user messages while chatting. The current technique involves designing a new Script and code the new script which will consume more time.

### *A. Drawbacks*

1. The social network today as a restriction on the user that who can post and comment on any user's wall, but do not have any restriction on the message posting. So some people may send crude to the other user wall, which cause annoy to the user reading on those crude words.
2. Providing this service the user can only share their message with their friends, family and society. But it does not block or filter any crude words in the message.

## IV. PROPOSED SYSTEM

The proposed system is designed to eliminate the drawbacks of the existing system. The application uses the naive bayesian method, when the user chat with their friends and make a post in their wall. The background watcher automatically tracks the user post, if the message contains any crude words, then the message is blocked automatically and warning message is send to the sender. If the user continues to post the crude message for more than three times, then the user account is blocked permanently. The primary goal of the new system is to reduce the time and cost cutting.

### *A. Blacklisting Process*

A user shows interest in posting or commenting in other person's wall regardless of their relationship. But the machine learning learns the messages that are to be posted and finds whether it contains any crude words. If it can't find any crude words, then the system allows the message to be posted on the wall. If it finds any crude words in that message, then it will blacklist the message and sends warning message to sender. If the warning message received by sender exceeds more than three times then it automatically blocks the user account.

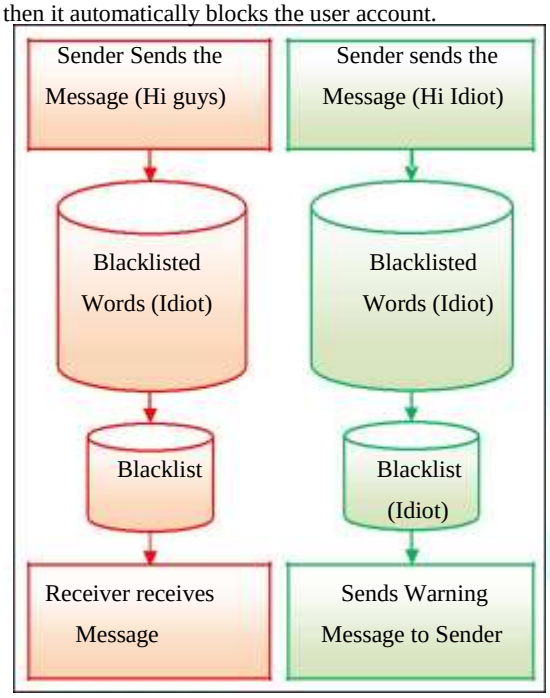


**Fig 3:Blacklisting process**

The admin can manage and update the crude or adult contents words and the chat message of the user can also be viewed by the admin at any time. The blocked users and their message are listed separately so that the admin can view and update the new crude words in the blacklist.

### *B. Advantages of proposed work*

1.Using the Background watcher the message content is compared and checked with the blacklisted words to avoid posting crude words.

2. It increases more number of users to use social networks if the posting of crude words are banned.

## VI. RESULTS:

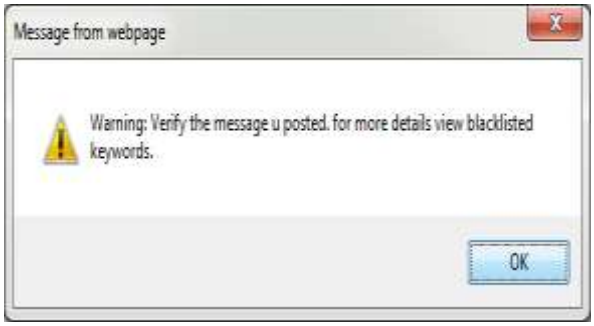


**Fig 5: Warning message**


**Fig 6:Warning message in user wall**

Fig 5 shows The background watcher tracks the message automatically and compare with the blacklisted words stored in database .if the user sends a message containing crude words to others wall then the user receives a warning message .Fig 6 shows that the user can also view the number of warning message received.

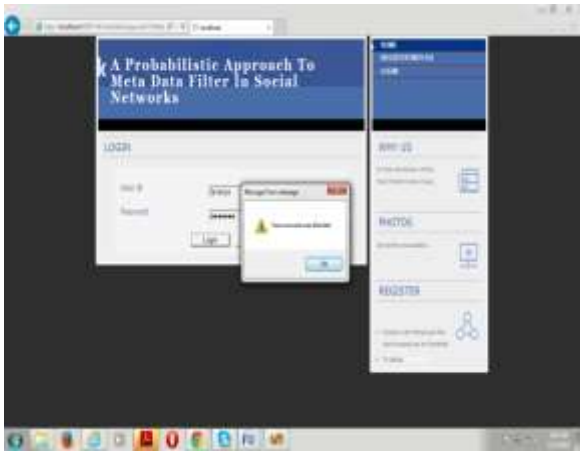


**Fig 7: Account Blocked**

Fig 7 shows the background watcher checks the user post. If the user continues to send crude words to other user more than three than the user account is blocked permanently.

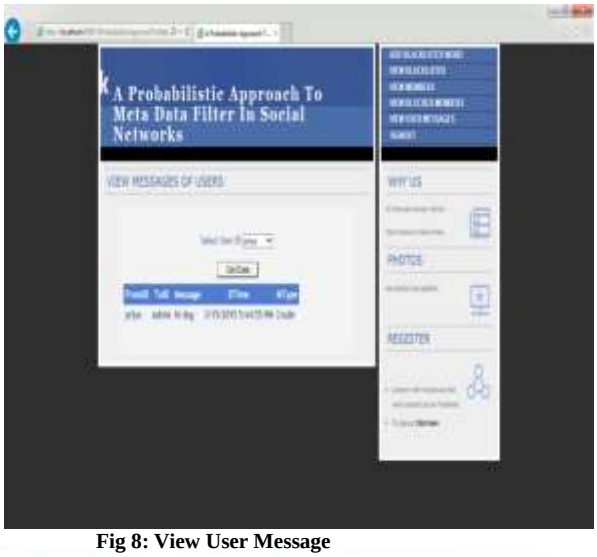


**Fig 8: View User Message**

Fig 8 shows that the admin can see the user chats to update the crude words.The admin is the only authorized person who can add or delete the blacklisted words.The blocked members and messages can also be seen by the admin to manage the user post by updating the crude words in the database.


**Fig 9:Add new Blacklisted Words**

Fig 9 shows that the admin can add the new blacklisted words to the database at any time. If the blacklisted words are saved wrongly it can also be deleted. The admin can update crude words by viewing the crude message of blocked members. Since the user may send more crude content words in a message, the admin can update the crude words from the blocked message.

## VII. CONCLUSION

A system to avoid posting indecent messages from the social networking site walls has been presented. The usage of background watcher is to trace the user message and distinguish between good and bad words. The background watcher plays a vital role in this paper to generate the blacklist of bad words and send warning message to the sender sending crude content message. Overall , the obscenity of the user has been prevented and usage of social network sites by the user will be increased.

## REFERNCES

[1] F. Sebastiani, “Machine Learning in Automated TextCategorization,” ACM Computing Surveys, vol. 34, no. 1, pp. 1-47, 2002.

[2] 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.

[3] Fabrizio Silvestri **“**Mining Query Logs: Turning Search Usage Data Into Knowledge” ISTI - CNR, via G. Moruzzi, 156124 Pisa, Italy

[4] 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, 1748-1753

[5] P.J. Denning, “Electronic Junk,” Comm. ACM, vol. 25, no. 3, pp. 163-165, 1982.

[6] P.W. Foltz and S.T. Dumais, “Personalized Information Delivery: An Analysis of Information Filtering Methods,” Comm. ACM, vol. 35, no. 12, pp. 51-60, 1992.

[7] P.S. Jacobs and L.F. Rau, “Scisor: Extracting Information from On-Line News,” Comm. ACM, vol. 33, no. 11, pp. 88-97, 1990.