iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 4 **Issue:** IV **Month of publication:** April 2016

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Filter unwanted messages from walls and blocking non-legitimate users in OSN

Miss. Gaikwad Rupali B[1], Miss. Kharat Mohini K[2], Mr.Tamhane Ajinkya V[3], Mr. Harde Pradeep S.[4]
Department of Computer Department, SCSCOE, Rahuri Factory

Abstract: ***The best entertainment for younger generation is given in the form of social Networking sites. The online social network (OSN) helps on individual to connect with their friends, family and the society to collect and share information with others. Now days, the OSN are facing the problem of people posting the indecent messages on any individuals wall which annoys other people on seeing them. The OSN provide little support to prevent unwanted message. So the propose system allowing OSN users to have a direct control on the message posted on their wall. This is achieved through filter wall (FW) able to filter unwanted message from OSN user walls. We propose a system using machine learning (ML) for filtering unwanted message. In our project Blacklists are used for more flexibility by which filtering choices are increased. Proposed system provides security to the online social Networks.***
Keywords: Online social Networks, Machine Learning, Information Filtering, Content Based Filtering.

I. INTRODUCTION

Today On-line Social Networks (OSNs) are the most popular interactive medium to communicate, share and disseminate a considerable amount of human life information. In OSN's are Daily and continuous communications imply the exchange of several types of content, including free text, image, and audio and video data. Today internet is the basic need of individuals. Now a days people cannot imagine life without Internet. Information and communication technology plays vital role in today's online networked society. Everyone is deeply connected with online social networks. It has affected the online interaction between different users. Online Social networks (OSNs) provide platform to all for sharing human life information. Machine learning text categorization technique is used in proposed system, It automatically assigns the short text based on the content. Techniques include some steps, first step is short text classifier, it includes text representation, radial basis function network, machine learning based classification. And the Second step is of filtering rules and management of blacklists. Filtering rules includes creator specification and filtering rule .and after that finally, Blacklist is included.

Therefore the primary aim of the current work is to propose and experimentally design an automated system, called as Filtered Wall (FW), which is able to filter unwanted messages from OSN user walls. The support for content based user preferences is the main idea of our proposed system. In proposed system Blacklist rule is implemented. An automated system called as filtered wall is designed. Filter wall is used to filter unwanted messages from OSN user wall, the content may be text images etc. content based message filtering is not supported to existing system.

II. LITERATURE SURVEY

A. Text classification

In our paper we proposed system, to filter unwanted messages text and images by using text classification method. In this two main approaches used Content based classification and policy personalization.

1) *Content Based Classification:* In content based filtering system content are matched with user profile using information retrieval techniques. Term Frequency and Inverse Document frequency(TF-IDF) docu ment filtered in content based filtering same as text classification. In this documents are classified into relevant and non-relevant categories[1]. In this paper various methods are used like Boosting based classifiers, Neural Networks, Support Vector Machine and Naïve Bayesian[2].
2) *Policy Based Personalization:* In some classification method provide short text message classification. In Twitter, some category provides which categorise each tweet. Also in FilmTrust application that provide trust relationship. But no application can provide filtering criteria. In contrast, our system provide Filtering Rule's(FR's)[5],which not only decide result of classification but also person which use OSN or owner of wall with other person. In policy based system, some FR's are decided and by which classification done. For more simplicity use BlackList.

©IJRASET: All Rights are Reserved

B. Image Classification

In image classification method porn images are filtered, and vulgar image are also filtered. For purpose of filtering pornography Detection Algorithm and Skin Detection algorithm used[4].

III. PROPOSED SYSTEM

The Architectural diagram for proposed System is as follows:

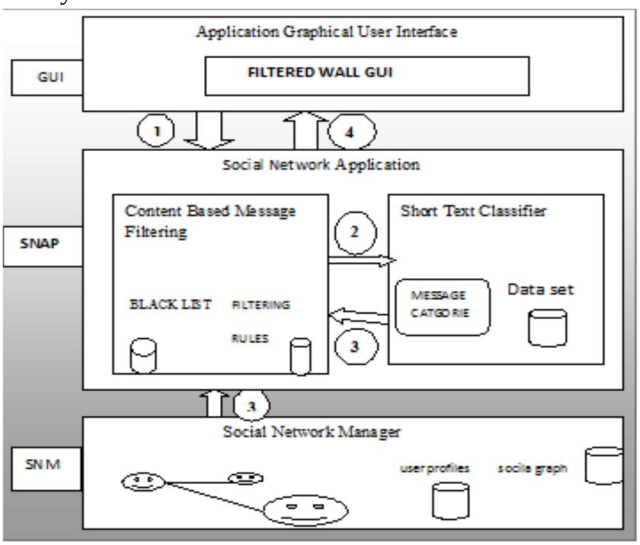


Fig: Filter Wall Architecture

The main aim of the system is to provide a facility to filter the messages which are to be posted by other users on his wall. This post can be in text form or in an image form. The system should be able to analyse the text or image content and allow desired content on the wall. In this system, when

Particular post is ready to be published on his wall, all customised settings are considered and wall post is filtered accordingly on that basis. While filtering the text messages, the thing which is checked first is that whether it is from an authentic user or not. If the user is authentic, then its content is analysed and properly categorizes using text classification techniques. After that system checks whether user preference is matching with derived post category. If it is matched, and the content of the post is neutral then that particular post is published as it is on hold till user permits it. In this paper, Blacklisting mechanism is used, where the user's list will be blocked for the moment to post on user wall. This paper is the extension of previous papers, we have included all the filtering and classification rules, additionally BL rule is used. Based on the user wall, if a user makes five attempts of posting vulgar posts or it may be depend on the users preference, on particular users wall then he may get blocked and he is not able to post anything on that particular users wall until that particular user unblocks him.

A. Message Filtering

We have to extract textual data from user's wall For message filtering purpose. Whenever someone upload textual data or textual content on general wall, then with the help of short text classification phase, system can classify that data into different categories with the help of dataset which is already provided to it.

B. Image Filtering

As part of contribution, we are also going for implementing the vulgar image filtering. To filter the vulgar images we are using two algorithms, which are Skin detection algorithm and pornography detection algorithm.

IV. ALGORITHM

A. Naïve Bayes

Naïve Bayes is a simple technique for constructing classifiers: model that assign class labels to problem instances represented as a vector of a feature values, where as a class label drawn from finite set.All Naïve Bayes classifier assume that the value of particular feature is independent of the value of any other feature, given the class variable.

©IJRASET: All Rights are Reserved

B. Probabilistic model

Naïve Bayes is a conditional probability model: given problem instance to be classified, represented by a vector **X** =(x1,….,x_n) representing some n feature (independent variables),it assigns to this instance probabilities.

$p(C_k|x_{1,}.....x_n)$

for each of K possible outcomes or classes

The problem with the above formulation is that if the number of feature n is large or if a feature can take on a large number of values, then basing such a model on probability tables is infeasible. We therefore reformulate the model to make it more tractable. Using Bayes' theorem ,

The conditional probability can be decomposed as

$p(C_k|x)= p(C_k)\ p(x|C_k)/p(x)$

In practice, there is interest only in the numerator of that fraction , because the denominator does not depend on C and the values of the features F_i are given , so that the denominator is effectively constant. The numerator is equivalent to the joint probability model.

$p(\ C_k,\ x_{1,}.....,x_n)$

Which can be rewritten as follows, using the chain rule for repeated applications of the definition off conditional probability:

$p(C_k,x_{1,}.....,x_n)= p(x_{1,}.....x_n,C_k)$

$= p(x_{1|}\ x_{2,}.....x_n,\ C_k\ )p(x_{2,}.....,x_n,C_k)$

$= p(x_1|x_{2,}.....,x_n,C_k)p(x_2|x_{3,}.....,x_n,C_k)p(x_{3,}.....,x_n,C_k)$

$= \ldots\ldots\ldots$

$= p(x_1|x_{2,}.....,x_n,C_k)p(x_2|x_{3,}.....,x_n,C_k)$

$\ldots\ldots.p(x_{n-1}|x_n,C_k)p(x_n|C_k)p(C_k)$

V. FILTERING PROCESS

To define the language for FR specification ,many problems are considered. First issue may arise when the message with different meaning and significance based on who writes it. Filtering rules will be applied ,when a user profile does not hold value for attributes submitted by a FR. This type of situation will be handled by with asking the user to choose whether to block or notify the messages initiating from the profile which does not match with the wall owners FRs, due to missing of attributes. Makers on which a FR applies can be chosen on the premise of a few criteria; a standout amongst the most significant is by forcing conditions on their profile's attribute. For example, conceivable to characterize rules applying just to youthful makers or to inventors with a given religious/political perspective

VI. BLACKLISTING PROCESS

To improve the effectiveness of the system such information are given to the system through a set of rules called as BL rules. A further segment of our system is a Blacklist (BL) system. It is used to avoid messages to be posted from undesired makers, free from their substance. The user should define his own rules and depending upon that it is illustrated by the system and if that user crosses that limit then he may be blocked or blacklisted. BL is is clearly administered by the by the system, the system has the ability to figure out, who are the users to be added in the BL and choose when user's maintenance in the BL is done. To improve adaptability, such type of data are given to the system through an arrangement of instructions, for this purpose the rules are used are called as BL rules. Maybe, we choose to let the clients themselves, i.e., the wall proprietors to indicate BL principles directing who must be banned from their wall and for to what extent, means that particular user decides which user should be banned from posting on his wall. Consequently, when a user gets blocked then he is no more able to post messages on that particular users wall until that articular user unblocks him.

VII. EXPERIMENTAL RESULT

The main aim of our project is to protect the user private wall's from the unwanted text or vulgar messages and also protect from the unwanted images to be posted on their own walls by filtering out the content of the messages. If any user can post normal or regular text on another user's wall, firstly this system checks this text is vulgar or normal. If this text is normal then system appose that user to post message and this is shown in following figure.

©IJRASET: All Rights are Reserved

Volume 4 Issue IV, April 2016
ISSN: 2321-9653

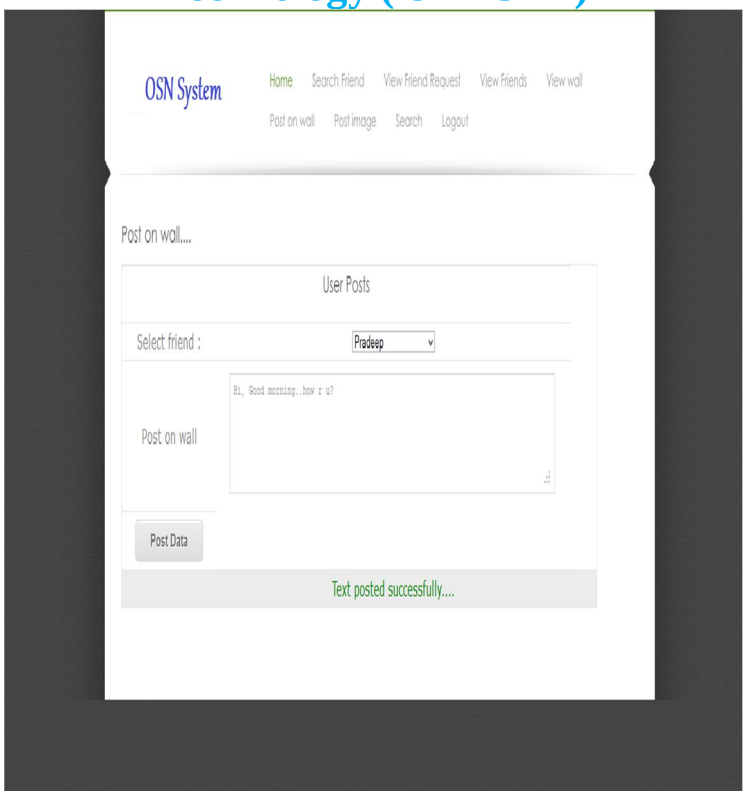


Fig: Normal text is successfully posted.



Fig: System oppose user to post vulgar messages.

This system is also blocking users who try to attempt five times to post unwanted text or image on another wall's. This is shown in following figure.

©IJRASET: All Rights are Reserved

Volume 4 Issue IV, April 2016
ISSN: 2321-9653



Fig: User is blacklisted.



Fig: Normal image successfully uploaded.

©IJRASET: All Rights are Reserved

Volume 4 Issue IV, April 2016
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

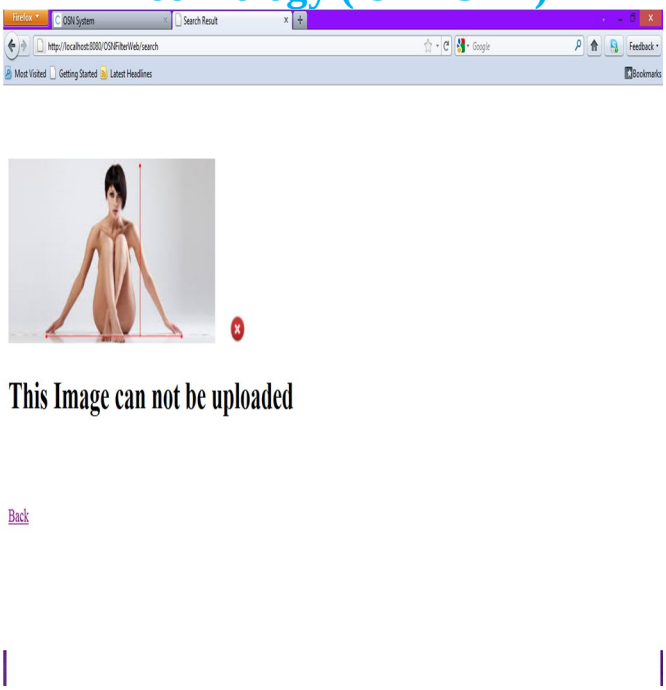


Fig: Vulgar images cannot be posted.

VIII. RESULT ANALYSIS

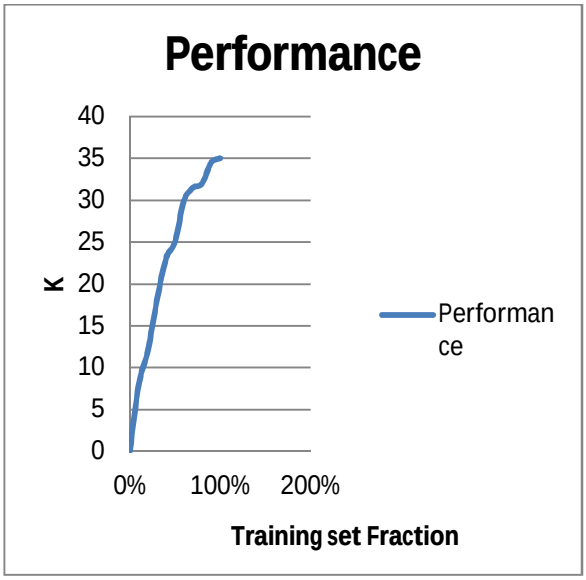


Fig: K value obtained training the model with different fractions of the original training set.

The results are obtained by performing 50 different distributions of messages between training set and test set. Improvement in the classification can be increased with increase in data set.

IX. CONCLUSION

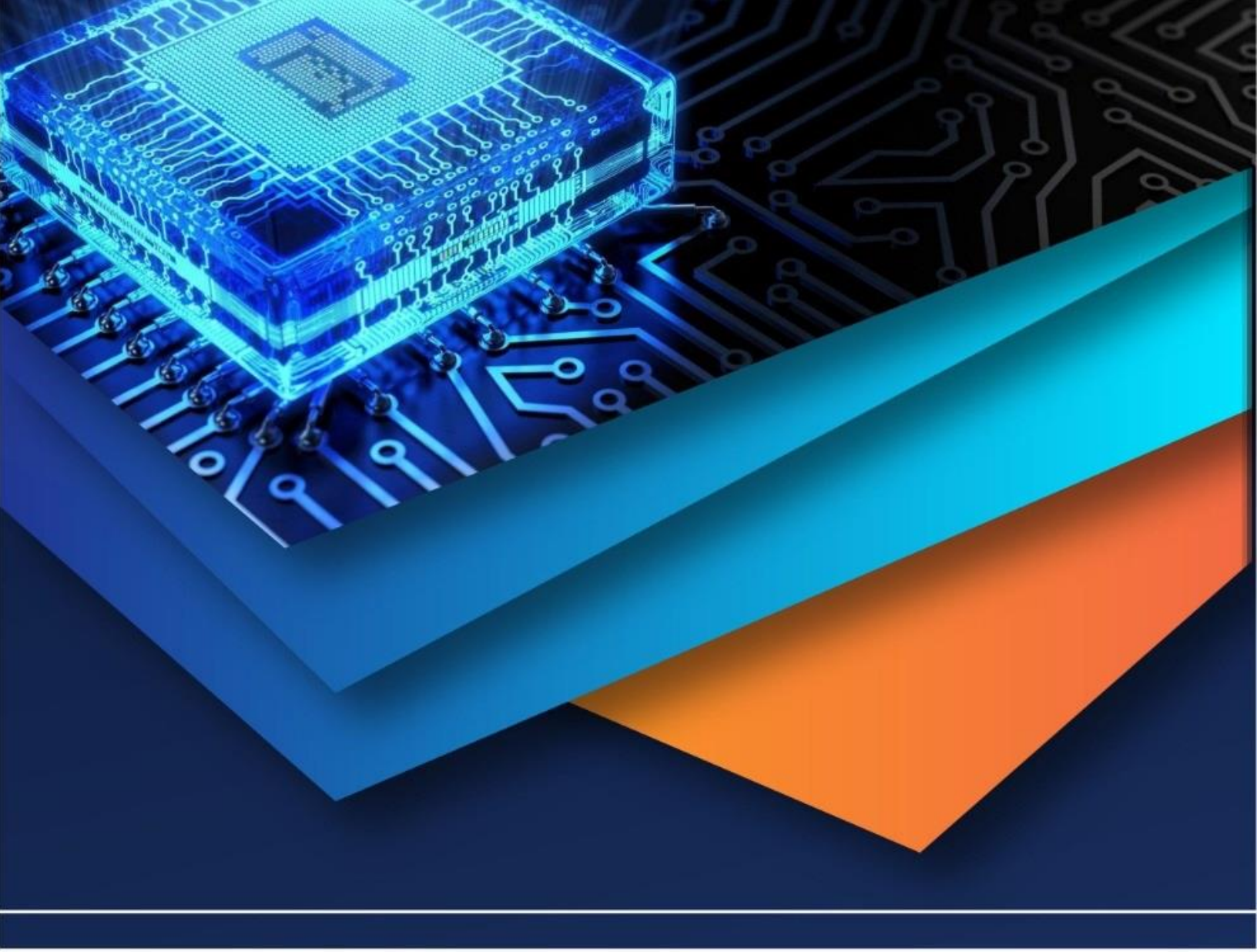
In this paper, a system to filter unwanted message in OSN wall is presented. The first step of the this project is to classify the content which is to be posted using several rules. Next step is to filter the undesired rules. And after that we have implemented the Blacklist rule. So that owner of the user can insert the user who post undesired messages. Better privacy can be given to the OSN user walls by using our system. In our system, filtered wall is a system to filter undesired messages from OSN walls. This system approach decides when user should be inserted into a blacklist. Filtered wall has a wide variety of applications in OSN wall . In future, more work is needed on further improving the performance measures.

©IJRASET: All Rights are Reserved

REFERENCES

[1] P.J Hayes,P.M. Andersen, I.B.Nirenburg, and L.M. Schmandt "Tcs: A Shell for Content-Based Text Categorization," Proc. Sixth IEEE Conf. Artificial Intelligence Application (CAIA '90),pp. 320-326, 1990.
[2] S.E. Robertson and K.S. Jones , "Relevance Weighting of search Terms," J. Am. Soc. For Information Science, vol. 27, no.3, pp. 129-146, 1976.
[3] Filtering Unwanted Contents from User's Wall in Online Social Networks, Miss.Ujwala S.Tambe, prof. Archana S. Vaidya.
[4] 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 Machin Learning (PSDML'10),2010.
[5] S.Pollock,"A Rule-Based MessageFiltering Methods " ACM Trans. Office Information Systems, vol. 6, no 3,pp. 88-97, 1990.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)