



Unwanted Message Filtration From Online Social Network (OSN)

Swati Tidke, SaritaGangbhoj, AnkitaBadwaik

Department of Computer Engineering, RTMN University, M.I.E.T. Shahapur, Bhandara, Maharashtra, India

Department of Computer Engineering, RTMN University, M.I.E.T. Shahapur, Bhandara, Maharashtra, India

Department of Computer Engineering, RTMN University, M.I.E.T. Shahapur, Bhandara, Maharashtra, India

ABSTRACT: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 connect with their friends, family and the society online in order to gather and share new experiences with others. Now- a-days, the OSNs are facing the problem of the people posting the indecent messages on any individual's wall which annoys other people on seeing them. In order to filter those unbearable messages a system called Machine Learning is introduced. 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.

KEYWORDS:Online social networks, information filtering, short text classification, and policy-based personalization, flexible rule-based system.

I. INTRODUCTION

Online Social Networks (OSNs) is mainly used as an interactive medium to communicate, share a considerable amount of human related information. OSN is commonly used to share many types of information, such as text, image, audio, and video data. Online Social Network is a platform to build social networks or social relations among people, for example, share interest, picture, text and real time connections. A social network service consists of each user having his own profile, his social links, and variety of additional services. It Some of the social networks which are mainly used to connect with friends are: Face book, Google+, YouTube, Twitter widely used worldwide.

Today OSN's provide little support to prevent unwanted messages on user profile. For example face book permit users to state who is allowed to insert messages in their walls i.e. friends, friends of friends, defined group of friends. Filtered wall is used to filter unwanted messages from OSN user walls. This wall is a public writing space so others can view what has been written on wall. Therefore in online social network there is possibility of posting bad or undesirable messages on wall which is visible to others also. To detect this problem, users wall messages should be classified and the unwanted messages should be filtered out on the wall owner. The short text classification is a difficult task as it does not have sufficient word occurrences. Additionally the attacker who want to write offensive message on his/her friend's wall will not write the message using direct presence of bad word but will use variety of patterns using special characters so that the message is not easily traceable to the system. Therefore we use a short text classifier to classify short text words.

II. RELATED WORK

By F. Sebastiani [2] describes a method " Machine Learning in Automated Text Categorization". A method describes the automated categorization (or classification) of texts into predefined categories has witnessed a booming interest in the last ten years, due to the increased availability of documents in digital form and the ensuing need to organize them. In the research community the dominant approach to this problem is based on machine learning techniques: a general inductive process automatically builds a classifier by learning, from a set of pre-classified documents, the characteristics of the categories.



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

By Robert E Schapire and Yoram Singer [3] describes “A Boosting-based System for Text Categorization”. A method that can be used to adopt a different approach in which the method use two extensions of Boost that were specifically intended for multiclass, multi-label data. In the first extension, the goal of the learning algorithm is to predict all and only all of the correct labels. Thus, the learned classifier is evaluated in terms of its ability to predict a good approximation of the set of labels associated with a given document. In the second extension, the goal is to design a classifier that ranks the labels so that the correct labels will receive the highest ranks. we introduce the use of a machine-learning technique called boosting to the problem of text categorization. The main idea of boosting is to combine many simple and moderately inaccurate categorization rules into a single, highly accurate categorization rule. The simple rules are trained sequentially; conceptually, each rule is trained on the examples which were most difficult to classify by the preceding rules.

By H. Schutze, D.A. Hull, J.O. Pedersen [4] proposed a technique “A Comparison of Classifiers and Document Representations for the Routing Problem”. A method that compare two approaches to document routing, relevance feedback via query expansion and statistical classification with error minimization. A system show that advanced classification algorithms perform 10-15% better than relevance feedback on the Tipster document collection. Since learning algorithms based on error minimization and numerical optimization are computationally intensive and prone to over fitting in a high dimensional feature space, it is necessary to apply some method of dimensionality reduction. The compare learning techniques based on statistical classification to traditional methods of relevance feedback for the document routing problem.

By Raymond J. Mooney, Loriene Roy [5] describes a method “Content-Based Book Recommending Using Learning for Text Categorization”. Recommender systems improve access to relevant products and information by making personalized suggestions based on previous examples of a user's likes and dislikes. Most existing recommender systems use social filtering methods that base recommendations on other user's preferences. By contrast, content-based methods use information about an item itself to make suggestions. We describe a content-based book recommending system that utilizes information extraction and a machine-learning algorithm for text categorization.

III. PROPOSED SYSTEM

In our proposed system there are three methods message filtration by admin, message filtration by user and short text classifier. In message filtration by admin method the messages filtered by admin, Admin sets the word category. In message filtration by user messages are filtered by user, user sets the word category. In short text classifier short text word are set by admin in the data base.

[3.1] Message Filtration by Admin

In this method all the filtering analysis is done by the admin .whenever any user sends or chats with the other, the information is stored in the database. The admin accesses this database and apply filtering techniques on each message content. In this system first the admin is login and set the word category, which type of messages he doesn't want to posted on the user wall such as vulgar, abusive, violence, hate and offensive, then new users sign up and registered and login. Whenever any user chats with the other and use the word which is restricted by admin then that word can't be displayed and give the message, 'your message can't be posted because it was filtered' which are shown in the figure 3.1.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

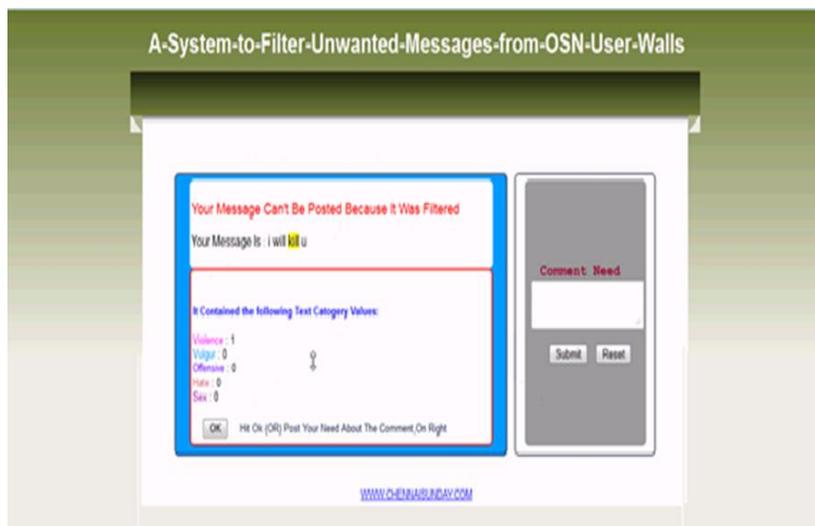


Figure 3.1.: Message filtered by admin

[3.2] Short Text Classifier

Short text classifier categorizes messages according to a set of categories. In this method short text words are set by admin in the data base. When any user sends any short text word which is set by admin the full form of that corresponding word is displayed or shown in the receiver wall. The machine learning mechanism is used to classify the short text.

[3.2.1] Machine Learning Mechanism

Machine learning (ML) is used as text categorization techniques to automatically assign each short text message within a set of categories based on its content. In machine learning approach, the problem of classification is an activity of supervised learning because the learning step is supervised by the knowledge of the categories. Figure 3.2 shows how admin adds the short text word 'gd' and its full form 'good day' in the database. Whenever any user sends that short text word 'gd', the full form of that word 'good day' is displayed on the receiver's wall, which is shown in Figure 3.3.



Figure 3.2: Add short-text word

In this figure 3.2 we set the word 'gd' to mean 'good day' to add to the database set by the admin and then logout.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

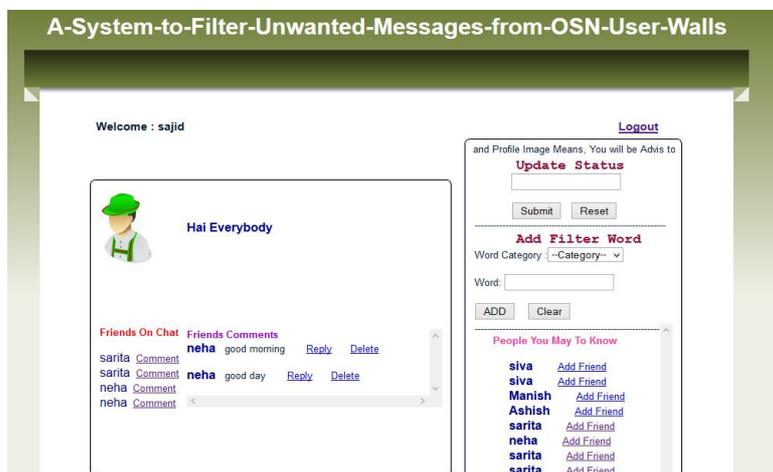


Figure 3.3: Show the full form of short text word

In this figure 3.3 we show the full form of the word is gd that means good day.

[3.3] Message Filtration by User

We propose a system allowing OSN users to have a direct control on the messages posted on their walls. Message filtration by user technique is done through a flexible rule-based system that allows users to customize the filtering criteria to be applied to their walls, in support of content-based filtering. In this method messages are filtered by user. User sets the word categories. First the user is login and add the filter word, which type of word he doesn't want to posted on their own wall such as violence, offensive or any type of messages. Whenever any user chats with the other and use that word which is set by user then that word can't be displayed and give the message to sender, 'your message can't be posted because it was filtered' which are shown in the figure 3.4.

[3.3.1] Flexible Rule Based System

Flexible rule based system allow information originator, administrators, and requesters to control and influence the flow of and access to information. The originator generate message then optionally specify rules indicating the type recipient they would like to reach. Recipients define rules that specify what types of messages and from what types of originators they want to receive.

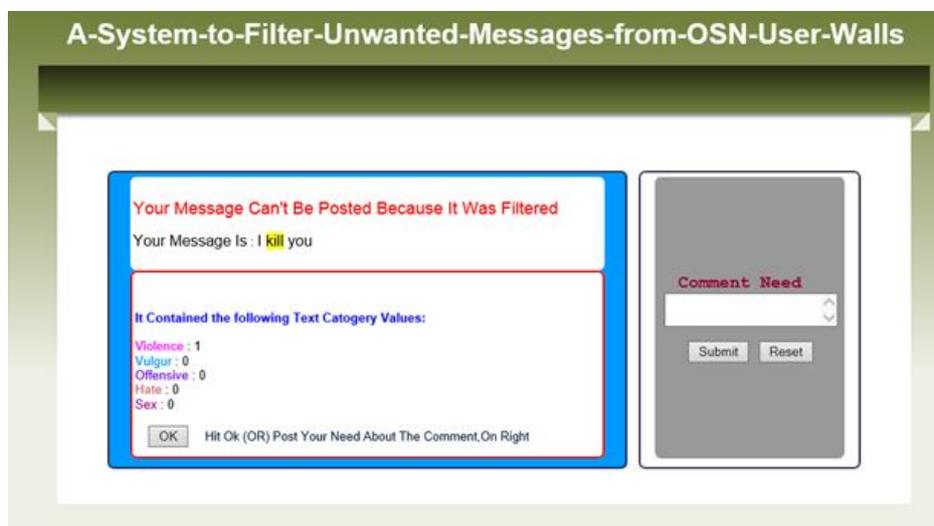


Figure 3.4: Message filtered by user



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

In this figure 3.4 we show the message of your message can not be posted because it was filtered.

IV. SIMULATION RESULTS

Here we have proposed three methods i.e. message filtration by admin, short text classifier and message filtration by user. By studying these three methods we get information that according to 'usability', the message filtration by admin method is user friendly with all related task, the message filtration by user gets high usability with user understandable GUI and short text classifier method is understandable language for user. But according to security purpose the message filtration by admin is highly secure, message filtration by user is moderate and short text classifier is highly secure.

The following formulae are used for calculating the performance of all three methods i.e. message filtration by admin, short text classifier and message filtration by user.

$$1) \text{ Precision} = \frac{\text{Number of documents correctly classified as positive by the system}}{\text{Number of all documents classified as positive by the system}}$$

$$2) \text{ Recall} = \frac{\text{Number of documents correctly classified as positive by the system}}{\text{Number of positive documents in the testing set}}$$

$$3) \text{ Accuracy} = \frac{\text{Number of documents currently classified by the system}}{\text{Number of all documents in the testing set}}$$

In message filtration by admin method, we have provided 30 input samples, for testing we have taken 20 input. In message filtration by user method, we have provided 28 input samples, for testing we have taken 23 input. In short text classifier method, we have provided 35 samples, for testing we have taken 27. In performance analysis we have calculated six parameters that is Accuracy, Precision, Recall, Usability, Implementation and Security. The outputs by using of all these three methods are shown in the table below:

Parameter	Admin login (Message Filtration for System Users)	user login (Message Filtration for User Side)	short text classifier (Classify and convert short text to full text)
Accuracy	95%	87%	96%
Precision	85%	78%	89%
Recall	90%	70%	93%
Usability	User friendly with all related tasks	High usability with user understandable GUI	Technique to use with user understandable language
Implementation	Complicated Backend Side	Complicated	Easy
Security	Highly Secure	Moderate	Highly Secure

V. CONCLUSION AND FUTURE WORK

In this system to filter undesired message from OSN walls. The use of Machine learning has provided higher results to the system to trace the messages and the users to distinguish between good and bad messages in the social networking user profiles automatically. Thus two methods message filtration by admin and message filtration by user



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

are used for filtration of unwanted messages. In filtration by admin method restrictions are applicable to all users and in message filtration by user method restriction are applicable to that particular user which apply restriction on their own wall. short text classifier is used to classify the short text words.

Future scope of this system is that Image Filtering Techniques. In our system we can only filter the text messages. So Image filtering will be tried in our future system. I plan to study techniques limiting the inferences that a user can do on the enforced filtering rules with the aim of bypassing the filtering system, such as for instance randomly notifying a message that should instead be blocked, or detecting modifications to profile attributes that have been made for the only purpose of defeating the filtering system.

REFERENCES

- [1] Sujapriya. S, G. ImmanuelGnanaDurai, Dr. C.Kumar Charlie Paul in "Filtering Unwanted Messages from Online Social Networks (OSN) using Rule Based Technique". IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661, p- ISSN: 2278-8727Volume 16, Issue 1, Ver I (Jan. 2014), PP 66-70
- [2] F. Sebastiani, "Machine Learning in Automated Text Categorization," ACM Computing surveya, vol.34.no.1, pp.1- 47, 2002.
- [3] Robert E Schapire and Yoram Singer in "A Boosting-based System for Text Categorization". Robert E.Schapire AT&T Labs, ShannonLaboratory,180 Park Avenue, Room A279, Florham Park, NJ 07932-0971 Yoram Singer AT&T Labs, ShannonLaboratory,180 Park Avenue, Room A277, Florham Park, NJ 07932-0971 Machine Learning, 39(2/3):135-168, 2000.
- [4] H. Schutze, D.A. Hull, J.O. Pedersen in "A Comparison of Classifiers and Document Representations for the Routing Problem" y Rank Xerox Research Center 3333 Coyote Hill Road6 Chemin de Maupertuis Palo Alto, CA 94304, USA 38240 Meylan, Francefschuetze.pederseng.
- [5] R. J. Mooney and L .Roy, "content-Based Book Recommending Using Learning for Text categorization" proc .Fifth ACM conf. Digital Libraries, pp.195-20categorization," ACM computing survey, vol.34, no.1, pp.1-47, 2002.
- [6] A. Adomavicius and G. Tuzhilim, "Toward the Next Generation of Recommender system: 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.
- [7] M. Chau and H. chen, "A Machine Learning Approach to web page Filtering Using content and structure Analysis," Decision support system, vol.44, no.2, pp.482-494, 2008.
- [8] M. Vanetti, E. Binagi, B. Carminati , M. Carullo and E. Forarri, "Content-Based Filtering in On- Line Social Networks, " pro. ECML/PKDD Workshop privacy and Security Issues in Data Mining and Machine Learning (PSDML'10), 2010.
- [9] N.J. Belkin and W.B. Croft, "Information Filtering and Information Retrieval: Two sides of the same coin?" comm. ACM, vol. 35, no .12. pp.29-38, 1992.
- [10] 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.
- [11] P.S. Jacobs and L .F. Rau, "Scion Extracting Information from On-Line News," comm. ACM, vol.33, no.11, pp.88-97, 1990.
- [12] S. Pollock, "A Rule-Based Message Filtering System," ACM Trans .Office Information Systems, vol.6, no.3, pp.232-254, 1988.

BIOGRAPHY

Swati Tidke Department of Computer Engineering, College of M.I.E.T Shahapur, Bhandara, RTMN University.

Sarita Gangbhoj Department of Computer Engineering, College of M.I.E.T Shahapur, Bhandara, RTMN University.

Ankita Badwaik Department of Computer Engineering, College of M.I.E.T Shahapur, Bhandara, RTMN University.