

Content based Message Filtering from Online Social Walls

Snehal D'mello

Department of Computer Engineering
St. Francis Institute of Technology
Mumbai, India

Prof. A. K. Sen

St. Francis Institute of Technology
Mumbai, India

Asst. Prof. Dakshata Panchal

Department of Computer Engineering
St. Francis Institute of Technology
Mumbai, India

Abstract - Online Social Networks (OSNs) are very common in today's world. OSNs typically have an area where users post messages or comment on posts written on public / private areas commonly known as walls. Users have the ability to view the public posts made by other users. However OSN users have no direct control on the content of messages posted on their walls, apart from blocking another user entirely from writing on his wall. Users may not be interested in viewing all the messages that are posted on their walls, and might not wish to block another user entirely from writing any message. In this work, a flexible rule-based system is proposed which gives the users the ability to control the messages posted on their walls through customizable filtering rules applied to user walls. Also recommendation of filtering rules is given to users with similar interests. Messages can be classified into different classes based on their content which is achieved through Machine Learning Techniques based on soft classification of OSN messages.

Keywords - Online social networks, information filtering, text classification, RBPNN.

I. INTRODUCTION

Online social network is an online site where people establish social relationship with each other. This relation can also be a result of offline relations. People who have similar interest, similar work profile and background build a social relationship on OSNs. OSN provides the feature to create a personal profile where you can upload a photo, some personal information such as age, gender, religion, location, qualification, hobbies, likes, dislikes, favourite books, TV shows, movies etc. Today Online Social Networks (OSNs) like Facebook, Twitter and LinkedIn etc. are very popularly used. A large amount of electronic data gets generated and shared on these OSNs. Users can share photos, videos, post messages privately or publicly on user walls or comment on posts. The users can decide the privacy of each photo, video, message or post in his/her profile. In order to achieve this, the options available with the user are to share Publicly, Friends Only, Friend of Friend, Me only etc. Each OSN user sets a connection with other OSN users, and the contents of his or her profile the user wants to share with others to manage the privacy. Once a user joins an OSN he/she can search for friends or other people with similar interest and establish an online relationship. Different tags are used for different

relationships in OSNs like "Friends", "Contacts", "Fans", "Followers" etc. The user can search his/her friends list and can view their profiles [1]. Most of the OSN allow users to send messages to the profiles of the friend list in two ways,

- Private message that only the recipient can see.
- Public message that appear in the recipient's profile wall and all his or her friends/contacts can read.

Also messages can be sent privately to users not in the friends list.

OSN users usually have hundreds of online social relationships. Most of these relations are from offline relations for example friends, relatives, colleagues etc. Users usually tend to increase their social network. OSN users can organize their friends or contacts into groups. Messages or other multimedia content like photos, videos can be shared with entire group. Users can share their content with their friends/contacts or with specific group of friends/contact. Users posts messages and/or updates on OSN walls of users. OSNs like Facebook allow users to block certain users completely. However, if someone tries to post unwanted messages like political or vulgar ones then in such cases OSNs do not allow filtering of such messages without blocking the user entirely.

An enormous and dynamic data gets generated by these OSNs which lead to the employment of web mining strategies that help in automatic extraction of useful information from the data. Web mining helps in OSN management tasks like access control and information filtering [2]. Information filtering is removal of unwanted information from a stream of data. It is of two types: Content based filtering and Collaborative filtering. Content based filtering system selects information based on the correlation between the content of the information and the user preferences whereas a collaborative filtering system selects data based on the correlation between people with similar preferences. Content based filtering and collaborative filtering can be used to block unwanted messages from OSN walls. The collaborative filtering is a technique used in recommender systems that generates recommendations based on the preferences given by other users of the system [3]. The collaborative filtering technique assumes that if a person X has the same opinion or

preference on an issue as person Y, then person X is likely to have the same opinion as person y on another issue.

The aim of this work is to develop a system which provides OSN users the ability to directly control messages posted on their walls. Also the users with similar interest will be provided with automatic recommendation of filtering rules.

II. RELATED WORK

The importance of OSNs and the need for information filtering in OSN have been discussed in [1]-[3]. Information filtering systems classify the stream of data generated into appropriate categories and present only that data to the user that he/she is interested in. In content-based filtering, each user is assumed to operate irrespective of other. Content-based filtering is mainly based on the use of the Machine Learning (ML) paradigm according to which a classifier is automatically induced by learning from a set of pre-classified examples. Content-based filtering is used for recommender systems. Content-based recommendation systems try to recommend items similar to those a given user has liked in the past.

Text classification classifies text into a set of categories. The categories provided by the text classifier are used for content-based message filtering. The different text classification techniques are Naive Bayes, SVM (Support Vector Machines), K-Nearest Neighbors (KNN), Neural Network, Boosting based classifiers and Rocchio. In [4], a detailed comparison analysis based on the effectiveness measure of precision and recall has been conducted confirming superiority of Boosting-based classifiers [5], Neural Networks [6], [7], and Support Vector Machines [8] over other popular methods, such as Rocchio [9] and Naive Bayes Bayesian [10]. However, it is worth to note that most of the work related to text filtering by ML has been applied for long-form text and the assessed performance of the text classification methods strictly depends on the nature of text documents. In [11], it is proved that, the RBPNN is better than the RBFNN, in the following several aspects: the contribution of the hidden center vectors to the outputs of the neural networks, the training and testing speed and the pattern classification capability.

The system proposed in [12] exploits classification techniques for personalizing access in OSNs. This system focuses on Twitter. The tweets are classified into different categories based on its content in order to avoid overwhelming users of micro-blogging services by raw data the user is then able to view only those tweets in which he/she is interested in. In an application called FilmTrust, by Golbeck and Kuter OSN trust relationships and provenance information to personalize access to the website is used [13]. This system uses TidalTrust algorithm for Inferring Trust.

However, such systems do not provide a filtering policy layer by which the user can exploit the result of the classification process to decide how and to which extent he/she can filter messages. Also these systems do not provide any recommendations to apply filtering rules to other users of the system with similar preferences.

III. SYSTEM ARCHITECTURE

This system consist of two important modules first ,the classification module and second the suggestion module. This system provides OSN users the ability to directly control messages posted on their walls. Customizable filtering rules (FRs) developed help for this purpose. FRs can support a variety of different filtering criteria that can be combined and customized according to the user needs. The criteria's for filtering are based on Age and Gender. The text classification technique used is RBPNN to classify messages posted on user walls. The set of classes considered for classification of text in are Normal, Sexual, Political, Vulgar and Racists. The system will also recommend filtering rules to other users with similar preferences and interests. Also a list of users is maintained called as Blacklist. These blacklisted users are not allowed to post on the users wall.

The figure below illustrates the basic high level architecture of the system. Each of the sub-system is explained further.

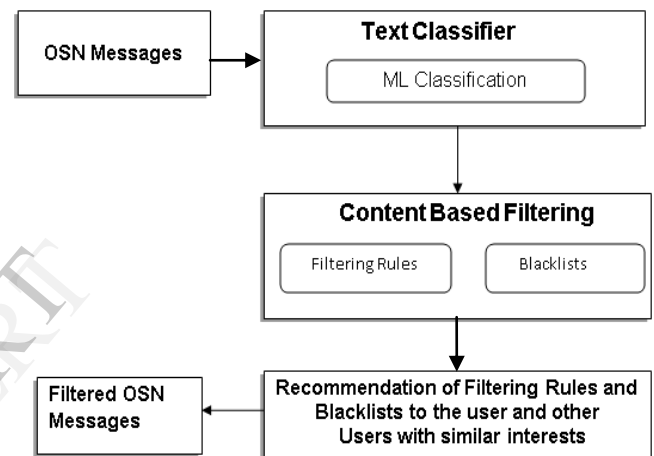


Figure 3.1.Proposed System

The implementation of the proposed system contains the following core modules:

- **Text Classifier Module:** This module is responsible for ‘understanding’ the crux of the message. It will parse through the message and will identify a set of keywords that will be a part of the metadata for that message. It will classify the data using Machine Learning text classification techniques like neural network. A ML-based text classifier (RBPNN) extracts metadata from the content of the message. It takes care of regular update of classification data that is used to classify messages.
- **Content Based Message Filtering Module:** The metadata provided by the classifier is used to enforce the filtering and Blacklist rules. The filtering takes place by comparison of the metadata with a dump of classification data. After comparison, an index for

each classification category is created. A higher index indicates the message being closer to that classification category. The result is then published based on the highest index for the classification. Depending on the results of the index, messages are published or filtered out.

- **Recommendation Module:** The users with similar preferences and interests are given recommendations to apply filtering rules. The similarity between users is calculated using the various demographic properties of users such as location, gender, religion etc. Also the user will be recommended other filtering rules based on previous rules applied by that user.

IV. RESULTS

The system classifies the messages into different categories and accordingly takes appropriate actions.

Following are the results of text classification for given data set using RBFN and RBPNN:

Comment	Expected Output	RBFN		RBPNN	
Support rahul gandhi n pay 150 per litre of petrol in 2015.	POLITICAL	24	POLITICAL	20	POLITICAL 18
are all khsatrayas maharashatrians?	RACIST	16	VULGAR	11	RACIST 13
the brahmins and the khsatrayas have never been there together	RACIST	24	RACIST	17	RACIST 16
I'm your slave for the night. Tell me what you want.	SEXUAL	23	SEXUAL	14	SEXUAL 16
brahmin and shatrayas both exist in the city	RACIST	18	RACIST	15	RACIST 18
Congrats our leader Narendra Modi for his Excellent victory.Hope the face of India will change within few years of time in all areas.He is the leader for those who want change and modernity.Jai Hind Narendra Modi Ji	POLITICAL	60	POLITICAL	53	POLITICAL 75
secular means not violence what modi had done on gujrat .	POLITICAL	19	POLITICAL	19	POLITICAL 18

Get on your hands and knees, sweetheart... and wait like a good girl.	SEXUAL	20	NORMAL	19	SEXUAL	23
Modi is a very Cheap type of politician always making fun of others he dont ve even a least decency in him .He only trying to fool ppl with his fakeism and media propaganda just like Hitler done to germans ppls.	POLITICAL	57	RACIST	59	POLITICAL	65
Faster! Deeper! Harder!	SEXUAL	11	SEXUAL	10	SEXUAL	13
Lets ignore kejrival and make the world forget him.. lets not give him importance .. afterall he is a barking cockroach	POLITICAL	30	RACIST	30	RACIST	27
I love having your body on top of mine in bed. It feels incredible.	SEXUAL	21	SEXUAL	21	SEXUAL	18
look how ready I am. Don't you want to put your dick in there?	SEXUAL	29	SEXUAL	23	SEXUAL	20
I want to fuck you everytime I see your nice tits	SEXUAL	29	SEXUAL	14	SEXUAL	16
hi rahul please try to learn basics of politics not from digvijay from pranab mukharjee you are a good hoice for pm best of luck	POLITICAL	60	VULGAR	27	VULGAR	37
kshatrayas will be there	RACIST	39	POLITICAL	11	RACIST	12
brahmins and kshatrayas have never stayed together	RACIST	18	RACIST	18	RACIST	20
I want you so bad	SEXUAL	9	SEXUAL	10	SEXUAL	8
modi is a killer of innocent people and congress is killing the country some new one should come	POLITICAL	24	POLITICAL	27	POLITICAL	25
Kiss me there... Lick every inch of me.	SEXUAL	13	SEXUAL	16	SEXUAL	17
rahul gandhi is spineless	POLITICAL	11	POLITICAL	11	POLITICAL	14

I'm going to fuck you till you can't walk! Ready?	SEXUAL	15	SEXUAL	16	SEXUAL	17
Come over here and ride me hard!	SEXUAL	12	RACIST	14	SEXUAL	18
God bless Dr. Manmohan Singh He is a great leader and an even greater statesman and human being. I am in awe of his humility, his dignity and his ability to be kind to even the most undignified personal attacks of the opponents.	POLITICAL	54	VULGAR	53	POLITICAL	58
Spray your juice all over my tits.	SEXUAL	11	SEXUAL	12	SEXUAL	18
Give me that come, honey. I want it in my mouth. Come on, give it to me.	SEXUAL	25	SEXUAL	24	SEXUAL	26
	TOTAL	672	18 / 26	564	24 / 26	626
	PERCENTAGE	69.230774%		92.30769%		

The above results show that text classification using RBPNN gives better results compared to classification using RBFN.

V. CONCLUSION

Content based message filtering from OSN walls is a useful service that will be provided to OSNs. With this service the users of OSNs will get the ability to control the messages posted on their walls and thereby avoid the nuisance created by unwanted messages posted on user walls.

This system can be modified for use in future for numerous other applications. For example, OSNs can have different walls for different contents based on this approach. A user can have a wall for political messages and another wall for religious messages.

REFERENCE

- [1] S. M. Maria, "Collaborative Filtering in Social Networks. Similarity analysis and feedback techniques", *Report*, May 2010.
- [2] Marco Vanetti, Elisabetta Binaghi, Elena Ferrari, Barbara Carminati, and Moreno Carullo, "A System to Filter Unwanted Messages from OSN User Walls," *IEEE Trans. Knowledge And Data Engineering*, Vol. 25, No. 2, pp.285-297, 2013.
- [3] 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.
- [4] F. Sebastiani, "Machine Learning in Automated Text Categorization," *ACM Computing Surveys*, vol. 34, no.1, pp. 1-47, 2002.
- [5] R.E. Schapire and Y. Singer, "Boostexter: A Boosting-Based System for Text Categorization," *Machine Learning*, vol. 39, nos. 2/3, pp. 135-168, 2000.
- [6] H. Schutze, D.A. Hull, and J.O. Pedersen, "A Comparison of Classifiers and Document Representations for the Routing Problem," *Proc. 18th Ann. ACM/SIGIR Conf. Research and Development in Information Retrieval*, pp. 229-237, 1995.
- [7] E.D. Wiener, J.O. Pedersen, and A.S. Weigend, "A Neural Network Approach to Topic Spotting," *Proc. Fourth Ann. Symp. Document Analysis and Information Retrieval (SDAIR '95)*, pp. 317-332, 1995.
- [8] T. Joachims, "Text Categorization with Support Vector Machines: Learning with Many Relevant Features," *Proc. European Conf. Machine Learning*, pp. 137-142, 1998.
- [9] T. Joachims, "A Probabilistic Analysis of the Rocchio Algorithm with TFIDF for Text Categorization," *Proc. Int'l Conf. Machine Learning*, pp. 143-151, 1997.
- [10] 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.
- [11] W. B. Zhao, D. S. Huang and L. Guo, "Comparative Study Radial basis probabilistic neural network and radial basis function neural network", *Intelligent Data Engineering and Automated Learning, Springer*, Volume 2690, pp 389-396, 2003.
- [12] 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.
- [13] J. Golbeck, "Combining Provenance with Trust in Social Networks for Semantic Web Content Filtering," *Proc. Int'l Conf. Provenance and Annotation of Data*, L. Moreau and I. Foster, eds., pp. 101-108, 2006.