A WEB-BASED APPLICATION FOR FILTERING OSN USER WALL

Pallavi Dani, Madhuri Andhale, Aarti Darade,Surbhi Gupta

Madhuri Andhale, Student,Dept. of Information Technology ,MET BKC IOE,Maharashtra,India

Pallavi Dani, Student,Dept. of Information Technology ,MET BKC IOE,Maharashtra,India

---***---

Abstract - Social networking sites are very popular. It has become a routine for individuals to keep checking messages on their online social wall. Social Networking sites are best entertainment for youngsters. Online social network helps you to connect with your family,friends, society to share your views and ideas on any topic. As social networking sites are open for all, anyone can post message on their own wall or others wall. Sometime people posts inappropriate messages on other's wall which annoys people viewing them. Considering this problem our paper work is for filtering OSN wall messages before they reach to the users wall and determine trust related to people on OSN.As OSN messages are short for that we used Short text classification method with trust calculation.

Key Words:

OSN : Online Social Networks

API : Application Programming Interface

OSA : Online Setup Assistant

CBMF : Content-Based Messages Filtering

STC   : Short text categorization

SNM : Social Network Manager

SNA:Social Network Applications

1 .INTRODUCTION

A social networking service or website is a platform to interact among the peoples, for example, The people can shares their ideas , messages , posts , interests , activities in worldwide network. A social network service includes the representation of each user and his/her profiles, his/her active participation on social platforms , Chats , Gossip , Stories , Advertisements , Audios ,Videos and a different types of web services. Many sites allow users to upload images or pictures, and can also add multimedia content to modify the veiw of the profile. e.g., Facebook, allow users to improve their profile by providing facility to insert , update , delete or add their personal information. Many social networking sites allows the users to post their status or blogs, posts or messages, search for other peoples with similar interests ,Preferences stated while creating public profile and

share contacts. For e.g. Sites like Google+ ,Facebook ,Twitter allows the user to choose the circle that is allowed to post messages on their walls(i.e. Friends, Friends of Friends, and any predefined group),but no content based filtering or user preferences are given into support to the existing systems, which doesn't prevent undesired and spam messages like political statements, advertisements, Vulgar messages. So it has needed to avoid the displaying of unwanted words on user's wall .This is achieved by using Rule Based Filtering System.

2 .RELATED WORK

Ms. Shruti C. Belsare, Prof. R.R. Keole, OSN user filtered walls for unwanted messages using content mining[13], proposed a system that will automatically filter unwanted messages from OSN user walls on the basis of both message content and message creator relationships,their preferences and characteristics. The limitation of this paper is that the users will have no privilege to access or modify the FR specification.

Dipali D. Vidhate, Ajay P. Thakare,to avoid unwanted messages from osn user wall:content based filtering approach[14], concerns about both the rule layer and the classification module. The limitation of this paper is the analysis of related work has highlighted the lack of a publicly available benchmark for comparing different approaches to content based classification of user walls short texts.

R. J. Mooney and L. Roy, Content-based book recommending using learning for text categorization, [6], uses Collaborative filtering method, but in the proposed system content based recommendation is used. It explains a content based book recommending system that develops information extraction and machine learning algorithm for text categorization.

M. Vanetti, E. Binaghi, B. Carminati, M. Carullo, and E. Ferrari,Content-based filtering in on-line social networks,[7], Quality of classification is considered as the main aim. This system can usually take decision about the messages which is blocked, due to the tolerance depends on statistical data.

F. Sebastiani,Machine learning in automated text categorization, [8], Efficiency is good, labour power will be saved is the advantage of this paper. The main approach used here is text categorization. Comparison will be performed between human expert and labour power expert.

H. Schutze, D. A. Hull, and J. O. Pedersen,A comparison of classifiers and document representations for the routing problem,[12], Latent semantics indexing and feature selection used as an approach comparison of this approaches will be done. Better performance will be taken.

R. E. Schapire and Y. Singer,Boostexter: a boostingbased system for text categorization, [11],Consist of two extensions, specially planned for Multi-class, multi labelled data. In first extension, learned classifier is evaluated to predict a good approximation of sets.

Adomavicius , G. and Tuzhilin , Toward the next generation of recommender systems: A survey of the state of-the-art and possible extensions,[9] ,Recommender system's Overview is explained. Main three approach used in present generation of recommendation system is hybrid, content based and collaborative recommendation.

Sriram, D. Fuhry, E. Demir, H. Ferhatosmanoglu , and M. Demirbas ,Short text classification in twitter to improve information filtering,[1], In online services like twitter users may grown to be plagued by the rare data.

V. Bobicev and M. Sokolova ,An effective and robust method for short text classification,[10], Classification of text encloses complex and specific terminology; need the application of learning method. Partial Matching method is applied which condense the text for confining the text feature.

J. Golbeck ,Combining provenance with trust in social networks for semantic web content filtering,[2],Social network is the common interest grouping web. To make the trust many explanations are required. Two level approaches are stated to combine annotation, trust and provenance. We state an algorithm for concluding trust relationship with provenance information and trust annotation in web social network.

3.PROBLEM DEFINITION

The major issue in social network is user does not have a control over their walls because it does not support content based preferences.Therefore it is not possible to prevent undesired messages such as political or vulgar ones which is posted on the private space of the users. Likewise, huge volumes of data are extracted and posted to the social sites, so it becomes a sophisticated task to social network management.

4.PROPOSED SYSTEM

We have implemented the system called ConnectifyMe, which contains all the features of Online Social Networking such as Register, Login, Upload Image, Upload post, Friend request etc. In this system we are providing the user with user defined filtering patterns which are given by the user itself to the system and according the user wall is filtered, the sender sending such message are blacklisted automatically by the system and if it exceeds more the three times by the same user he/she is blocked.
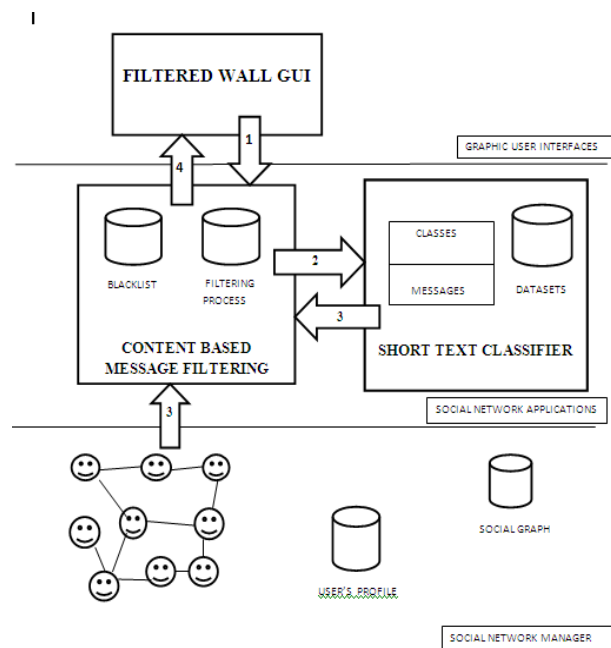
Fig:Filtered Wall Architecture

4.1 Stemming and Lemmatizing

The basic function of both the methods– stemming and lemmatizing is similar. Both of them reduce a word

variant to its 'stem' in stemming and 'lemma' in lemmatizing. There is a very subtle difference between both the concepts. In stemming the 'stem' is obtaining after applying a set of rules but without bothering about the part of speech (POS) or the context of the word occurrence. In contrast, lemmatizing deals with obtaining the 'lemma' of a word which involves reducing the word forms to its root form after understanding the POS and the context of the word in the given sentence.

In stemming, conversion of morphological forms of a word to its stem is done assuming each one is semantically related. The stem need not be an existing word in the dictionary but all its variants should map to this form after the stemming has been completed. There are two points to be considered while using a stemmer:

·Morphological forms of a word are assumed to have the same base meaning and hence should be mapped to the same stem.

·Words that do not have the same meaning should be kept separate.

For example, the word inflations like gone, goes, going will map to the stem 'go'. The word 'went' will not map to the same stem. However a lemmatizer will map even the word 'went' to the lemma 'go'.

Lemmatizing:

introduction, introducing, introduces –introduce gone, going, goes, went –go

4.2 Stopwords Algorithm

Sometimes, some extremely common words which would appear to be of little value in helping select documents matching a user need are excluded from the vocabulary entirely. These words are called *stop words* .

An example of a stop list is as follows:

a	an	and	are	as	at	be	by	for	from
has	he	in	is	it	its	of	on	that	the
to	was	were	will	with					

5. CONCLUSIONS

We have proposed a Filter wall to filter unwanted messages from Online Social Network private walls. We have used blacklist management and filtering rule to provide rules to the system. Blacklist rule is used to avoid unwanted messages created by the unauthorized users. So, we can avoid unauthorized person and we can keep our private wall secure.

REFERENCES

[1] B. Sriram, D. Fuhry, E. Demir, H. Ferhatosmanoglu, and M. Demir-bas, Short text classification in twitter to improve information filtering.

[2] J. Golbeck, Combining provenance with trust in social networks for semantic web content filtering.

[3] P. Bonatti and D. Olmedilla, Driving and monitoring provisional trust negotiation with metapolicies,

[5] C. Bizer and R. Cyganiak, Quality-driven information filtering using the wiqua policy framework" Semantics: Science, Services and Agents on the World Wide Web.

[6] R. J. Mooney and L. Roy,Content-based book recommending using learning for text categorization.

[7] M. Vanetti, E. Binaghi, B. Carminati, M. Carullo, and E. Ferrari, Content-based filtering in on-line social networks.

[8] F. Sebastiani,Machine learning in automated text categorization

[9] A. Adomavicius, G.and Tuzhilin,Toward the

next generation of recommender systems.

[10] V. Bobicev and M. Sokolova,An effective and robust method for on, n AAAI, D. Fox and C. P. Gomes, Eds. AAAI Press, 2008 .

[11] R. E. Schapire and Y. Singer, Boostexter: a boosting-based system text categorization, Machine Learning.

[12] H. Schutze, D. A. Hull, and J. O. Pedersen, A comparison of classifiers and document representations for the routing problem.

[13]Ms. Shruti C. Belsare, Prof. R.R. Keole, OSN user filtered walls for unwanted messages using content mining.

[14]Dipali D. Vidhate, Ajay P. Thakare,to avoid unwanted messages from osn user wall:content based filtering approach.