Message Filtering System

Sandip Bankar[#1], D.K.Chitre[#2]

Department Of Computer Engineering
Terna Engineering College,Nerul,Navi Mumbai
[1]sandipbankar1@gmail.com
[3]dkchitre@rediffmail.com

***Abstract*— One fundamental issue in today On-line Social Networks (OSNs) is to give users the ability to control the messages posted on their own private space to avoid that unwanted content is displayed. Up to now OSNs provide little support to this requirement. To fill the gap, in this paper, we are going to propose a system allowing OSN users to have a direct control on the messages posted on their walls. This is achieved through a flexible rule-based system, that will allow users to customize the filtering criteria to be applied to their walls, and a Machine Learning based soft classifier automatically labeling messages in support of content-based filtering.**

***Keywords*— On-line Social Networks, Information Filtering, Short Text Classification, Policy-based, Personalization Introduction**

A.INTRODUCTION

Online Social Networks (OSNs) are today one of the most popular interactive medium to communicate, share, and disseminate a considerable amount of human life information. Daily and continuous communications imply the exchange of several types of content, including free text, image, audio, and video data. According to Facebook statistics1 average user creates 90 pieces of content each month, whereas more than 30 billion pieces of content (web links, news stories, blogposts, notes, photo albums, etc.) are shared each month. The huge and dynamic character of these data creates the premise for the employment of web content mining strategies aimed to automatically discover useful information dormant within the data.

Information filtering can therefore be used to give users the ability to automatically control the messages written on their own walls, by filtering out unwanted messages. We believe that this is a key OSN service that has not been provided so far. Indeed, today OSNs provide very little support to prevent unwanted messages on user walls. For example, Facebook allows users to state who is allowed to insert messages in their walls (i.e., friends, friends of friends, or defined groups of friends). However, no content-based preferences are supported and therefore it is not possible to prevent undesired messages, such as political or vulgar ones, no matter of the user who posts them. Providing this service is not only a matter of using previously defined web content mining techniques for a different application, rather it requires to design ad hoc classification strategies

B. REVIEW OF THE LITERATURE

In paper [1] authors presents an overview of the field of recommender systems and describes the current generation of recommendation methods. for recommendation system. Recommender systems help in addressing the information overload problem by retrieving the information desired by the user based on his or similar users' preferences and interests. His entire document should be in Times New Roman or Times font. Type 3 fonts must not be used. Other font types may be used if needed for special purposes.

Methods that are usually classified into the following three main categories:

1. content-based Recommender System: It performs item recommendations by predicting the utility of items for a particular user based on how "similar", the items are to those that he/she liked in the past. E.g. In a movie recommendation application, a movie may be represented by such features as specific actors, director, genre, subject mattered.

2. Collaborative Recommender Systems: Collaborative filtering methods are based on collecting and analyzing a large amount of information on users' behaviors, activities or preferences and predicting what users will like based on their similarity to other users. A key advantage of the collaborative filtering approach is that it does not rely on machine analyzable content and therefore it is capable of accurately recommending complex items such as movies without requiring an "understanding" of the item itself.
3 Hybrid Recommender Systems: Combining collaborative filtering and content-based filtering could be more effective in some cases. Hybrid approaches can be implemented in several ways: by making content-based and collaborative-based predictions separately and then combining them.

In the paper [2], the Authors have provided a review of the field from the perspectives of Machine Learning and Information Retrieval and how they have been applied in Web mining

Since 1940's, many knowledge-based systems have been built. Most systems acquire knowledge manually from human experts, which is very time-consuming and labour-intensive. To address this problem, Machine Learning algorithms have

© 2015 IJAIR. All Rights Reserved

been developed to acquire knowledge automatically from examples or source data

The[3] paper introduces a technique for incorporating the vast amount of human knowledge accumulated in Wikipedia into text representation and classification. The simplest approach to represent the text semantics is to treat the text as an unordered bag of words where the words themselves (possibly stemmed) become features of the textual object The bag of words (BOW) method is very effective in easy to medium difficulty categorization tasks where the category of a document can be identified by several easily distinguishable keywords

The authors proposed an intuitive approach to determine the class labels and the set of features with a focus on user intentions on Twitter Such as daily chatter, conversations, sharing information/URLs, and reporting news They classified incoming tweets into categories such as News (N), Events (E), Opinions (O), Deals (D), and Private Messages (PM) based on the author information and features within the tweets .Categorization of tweets into the selected classes requires the knowledge of the source of information. They selected the authorship information as primary feature. If user defines an event as "something that happens at a given place and time", the presence of participant, place, and time information could determine the existence of an event in the text. Hence, they extracted the date/time information and time-event phrases which are collected from a set of tweets based on general observation of users and set the presence of them as a feature. Participant information is also captured via the presence of the '@' character [4].

C.A PROPOSED SYSTEM

1. Architecture:

The aim of this work is to propose and experimentally evaluate a system, called Filtered Wall (FW), which able to filter unwanted messages from user walls. We making use of Machine Learning (ML) text categorization techniques to automatically assign with each short text message a set of categories based on its content. The major efforts in building a robust short text classifier are concentrated in the extraction and selection of a set of characterizing and discriminate features. We are investigating solutions which are an extension of those adopted in a reviewed work from whom we are using the learning model and the elicitation procedure for generating preclassified data. The original set of features, derived from endogenous properties of short texts, is enlarged here including exogenous knowledge related to the context from which the messages originate. As far as the learning model is concerned, we confirm in the current paper the use of neural learning which is today recognized as one of the most efficient solutions in text classification.

The first layer, called Social Network Manager (SNM), commonly aims to provide the basic OSN functionalities (i.e., profile and relationship management), whereas the second layer provides the support for external Social Network Applications (SNAs). The supported SNAs may in turn require an additional layer for their needed Graphical User Interfaces (GUIs). According to this reference architecture, the proposed system is placed in the second and third layers. In particular, users interact with the system by means of a GUI to set up and manage their FRs/BLs. Moreover, the GUI provides users with a FW, that is, a wall where only messages that are authorized according to their FRs/BLs are published The core components of the proposed system are the Content-Based Messages Filtering (CBMF) and the Short Text Classifier (STC) modules. The latter component aims to classify messages according to a set of categories. In contrast, the first component exploits the message categorization provided by the STC module to enforce the FRs specified by the user. BLs can also be used to enhance the filtering process. Notification manager (NM) it does processing of notifications according to the posting of unwanted messages. Refer following diagram 1.

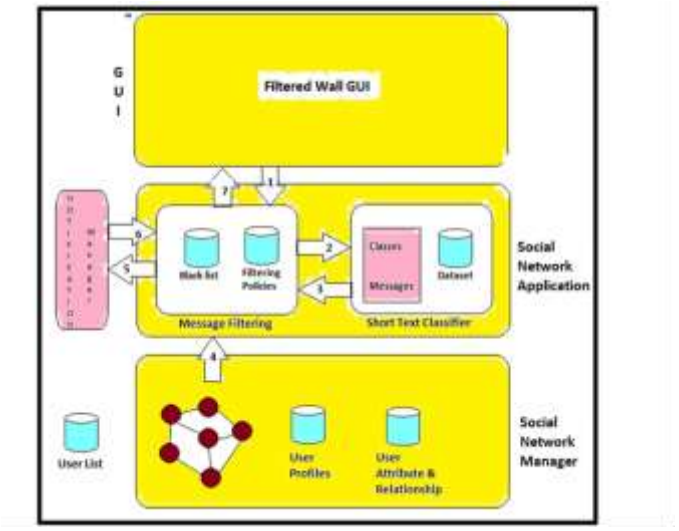

Figure 1: Architecture of proposed system.

Algorithm

Step 1: Start
Step 2: A User tries post the message in a wall.
Step 3: Machine learning checks each word of theMessage.
Step 4: If (Words = = Good Words)
Step 5: Message is posted on the wall.
Step 6: Else if(Words = = Bad Words)
Step 7: Reject Bad Words using Blacklist and notify user
Step 8: Update user levels and data base.
Step 9 stop.

Algorithm of the process of rejecting message

D.ADVANTAGES

A system to automatically filter unwanted messages from OSN user walls on the basis of both message content and the message creator relationships and characteristics.

This propoed system largely extends for what concerns both the rule layer and the classification module.

© 2015 IJAIR. All Rights Reserved

An online setup assistant (OSA) to help users in FR specification, the extension of the set of features considered in the classification process, a more deep performance evaluation study and an update of the prototype implementation to reflect the changes made to the classification techniques. Notification manager initially notifies user about posting of unwanted messages instead of blocking.

PERFORMANCE EVALUATION.

On-line Social Networks (OSNs) are today one of the most popular interactive medium to communicate, share and disseminate a considerable amount of human life information. Daily and continuous communications imply the exchange of several types of content, including free text, image, and audio and video data. Whereas more than 30 billion pieces of content (web links, news stories, blog posts, notes, photo albums, etc.) are shared each month.

Problem Raised For example, Facebook allows users to state who is allowed to insert messages in their walls (i.e., friends, friends of friends, or defined groups of friends). However, no content-based preferences are supported and therefore it is not possible to prevent undesired messages, such as political or vulgar ones, no matter of the user who posts them.

All the basic facility available on social networking site like creating new account, sending friend request, accepting friend request, searching for new friends, block existing friend, unblock friends, block lists, post messages on users wall etc. these facility are provided for our social networking site along with this, System has the main purpose of filtering the posts to check whether it contains any spam words or not.

This system is having 5 methods to analyze the post.

- .Violence :

This method is used to filter out violence type of words present in the post. It calculates the percentage of this type of words in the post.

- Hate :

This method is used to filter out Hate type of words present in the post. It calculates the percentage of this type of words in the post.

- Offensive:

This method is used to filter out Offensive type of words present in the post. It calculates the percentage of this type of words in the post.

- Vulgar :

This method is used to filter out vulgar type of words present in the post. It calculates the percentage of this type of words in the post.

- Sexual :

This method is used to filter out Sexual type of words present in the post. It calculates the percentage of this type of words in the post.

After getting all percentages it will get the sum of percentages of these type of spam words and it compare this value with the threshold.Threshold for the filtering is 0% to 0.5%.

It filters the message by comparing each and every word in the post with words stored in the database tables. There are 5 tables for 5 types of spam words. Every table contains vast amount of possible words of respective type. System is tested with all type of spam words of above types and works fine. It calculates the percentage accurately. If sum of percentage is more than 0.5%, then system will not allow you to send that post. If sum of percentage is more than 0% and less than 0.5% then system provide facility to receiver to accept or decline the post. If sum of percentage is 0% then system send the post to the receiver directly without any permission.

If any user sent post with spam words and receiver decline his post 3 times then system blocks the sender. System also provides facility to manually block or unblock any friend.

E.CONCLUSION

The fundamental issue in today On-line Social Networks is to give users the ability to control the messages posted on their own private space to avoid that unwanted content is displayed

The Review of The Literature gives study of the available methods to filter messages advantages as well as disadvantages and motivates towards proposing another new approach which will minimize the available disadvantages and maximize or retains advantages.

The system to be proposed, to filter unwanted message in OSN wall. System creates content based machine learning classifiers to enforce filtered rules. Also filtering will enhanced through use of blacklist management. Initially it will classify the content using several rules. Next is to filter the unwanted messages. Then Blacklist will be implemented. This blacklist updation will take place according to user wall post history. Owner of the user can insert the user who posts undesired messages. Better privacy is given to the OSN wall using our system.

User can specify what kind of messages he wants to receive by using policy based personalization. Finally proposed system is going to notify user, after posting unwanted message according to user attributes and history.

F.REFERENCE

[1]Marco Vanetti, ElisabettaBinaghi, Elena Ferrari, Barbara Carminati, and Moreno Carullo,"A System to Filter Unwanted Messages from OSN User Walls " IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 25, NO. 2, FEBRUARY 2013

[2] GediminasAdomavicius , Alexander Tuzhilin, "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and PossibleExtensions ", IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 17, NO. 6,

[3]M.Chau and H.Chen ,"A Machine Learning Approach to Web Page Filtering Using Content and Structure Analysis", Decision Support Systems, vol.44, no.2, pp.482-494, 2008.

© 2015 IJAIR. All Rights Reserved

[4]Hiroki Yamakawa, Jing Peng, and Anna Feldman ,"Semantic Enrichment of Text Representation with Wikipedia for Text Classification ", 978-1-4244-6588-0/10/$25.00 ©201 0

[5] Sujapriya. S ,G. Immanual Gnana Durai , Dr. C.Kumar Charlie Paul ,"Filtering Unwanted Messages from Online Social Networks (OSN) using Rule Based Technique ",, IOSR p- ISSN: 2278-8727Volume 16, Issue 1, Ver. I (Jan. 2014), PP 66-70

[6]M. Vanetti, E. Binaghi, B. Carminati, M. Carullo, and E. Ferrari,"Content-based Filtering in On-line Social Networks ".

[7] Carminati, B., Ferrari, E."Access control and privacy in web-based social networks",International Journal of Web Information Systems 4, 395–415 (2008)

[8] B.Sriram, D. Fuhry, E. Demir, H. Ferhatosmanoglu, and M. Demirbas, "Short text classification in twitter to improve information filtering," in Proceeding of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2010, 2010, pp. 841–842.

[9]F. Sebastiani, "Machine learning in automated text categorization," ACM Computing Surveys, vol.34, no. 1, pp. 1–47, 2002.

[10]N. J. Belkin and W. B. Croft, "Information filtering and information retrieval: Two sides of the same coin? " Communications of the ACM, vol. 35, no. 12, pp. 29–38, 1992.

[11]F. Sebastiani, "Machine Learning in Automated Text Categorization," ACM computing Surveys, vol. 34, no. 1, pp. 1- 47, 2002.

[12]R.J. Mooney and L. Roy, "Content-Based Book Recommending Using Learning for Text Categorization," Proc. Fifth ACM Conf. Digital Libraries, pp. 195-204, 2000

© 2015 IJAIR. All Rights Reserved