



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 12, December 2014

A Survey on Filtering Unwanted Messages from Online Social Network Users Wall Using Text Classification

Akshay Bagal , Shriniwas Gadage

P.G.Student, Dept. of Computer Engineering, G.H.R.C.E.M Pune, Savitribai Phule Pune University, India.

Professor, Dept. of Computer Engineering, G.H.R.C.E.M Pune, Savitribai Phule Pune University, India.

ABSTRACT: -Today one of the best communication between people is using the Online social network to which they can share information. Therefore the users who are using Online Social Networks(OSN) requires control over the unwanted messages that are posted on thier walls and to avoid the unwanted content which is displayed on private space of user. OSN provide us little support to the users requirement. To provide this, we propose a system which allows OSN users to have a direct control on the messages posted on users walls. This is achieved through a flexible rule based system in which users to customize the filtering criteria to be applied to their walls, and a Machine Learning (ML) based soft classification and short text classification which automatically produce membership labels in support of content-based filtering of a unwanted messages.

KEYWORDS: Online Social Networks (OSNs); Short text classification; Content based filtering; Filtering Rules

I. INTRODUCTION

Today the most popular interactive medium to communicate, share and disseminate information of a human life are On *line* Social Networks(OSNs). On Daily and continuous communications imply the exchange of several types of content, which includes free texts, images, audio and video data. According to Facebook statistics 90 pieces of content each month is created by a user, where as more than 30 billion pieces of content are shared by each month. Information filtering introduced technique can give users the ability to automatically control the messages written or posted on their own wall, by filtering out unwanted messages. Where, today OSNs provide very little support to prevent unwanted messages on user walls. For example, Facebook allows user who can insert messages on their walls (i.e., friends, friends of friends, or defined group of friends). However, Content based filtering preferences are not supported. A wall message contains the short text for which traditional classification methods have serious limitations since short texts do not provide sufficient word occurrences.

So the aim of the present work is to propose and experimentally evaluate an automated system, called as Filtered Wall (FW) which is able to filter unwanted messages from OSNs users wall. The content based user preferences is the key idea of proposed system [12]. We can use the machine learning (ML) text categorization techniques [4] able to automatically assign with each message a set of categories based on its users content. And by introducing filtering rules (FRs), user can state what contents should not be displayed on thier walls which consists different filtering criteria. Text classification is to be done by extraction and selection of a set of characterizing and discriminant features which is solution for classification of short text. Section 2 reviews related work, whereas Section 3 introduce the proposed system and filtered wall conceptual architecture system[1]. Section 4 describes the Short text classification methods used to categorize text contents, whereas Section 5 explains management of FRs and BLs. Section 6 concludes the paper.

II. RELATED WORK

Marco Vanetti, Moreno Carullo, Elisabetta Binaghi, Barbara Carminati, and elena Ferrari [1] provides the user a system customizable content-based message filtering over their own wall to avoid the unwanted messages. Aim of this



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 12, December 2014

paper is, user have a direct control over messages posted on their own wall and privacy preserving content. Therefore, automated system called Filtered wall (FW) is introduced, which have a capacity to filter unwanted message and blocks the message posted by users .L.Roy and R.J.Mooney[12] uses Collaborative filtering method which is the system that chooses items based on the correlation between people with similar preferences, but in the proposed system Content based filtering method is used. Which has content based recommending system that develop information extraction and machine learning algorithm for text classifications. B.Carminati, M.Vanetti, E.ferrari, M.Carullo, and E.Binaghi[7] Performance of classification includes different semantics for filtering rules is considered as the main aim. This system can usually take decision about the messages which has to block, due to the tolerance depends on preclassified data set. F.Sebastiani[4] The main approach used here is ML text categorization technique. Which automatically assigns with each short text message from a set of categories based on its content. H.Schutze, D.A.Hull, and J.O. The feature selection and indexing uses number of approaches in filtering and clasification[8]. Comparision analysis is to be done on approaches where better performance will be taken. M.Chau and H.Chen[2] Relevant data are very complicated to find on web content. Web page is represented with content based and link based feature in proposed system. Neural network approach is used for proposed system to avoid useless data. So Proposed approach can be applied for web content management. A.Adomavicius and G.Tuzhilin[3] Recommender system's uses three approach content-based recommendation, collaborative and hybrid recommendation. Using this appraoches we can enlarge recommendation system using contextual features. B.Sriram, D.Fuhry, E.Demir, H.Ferhatosmanoglu, and M.Demirbas[6] in online services like twitter, users may grown to be problematic development of a reliable data. Solution of this crisis is short text messages classification.To solve this problem , we suggest a small set of categories domain specific features from each tweet describes its content. This approach successfully classifies the text into certain types of interest. V.Bobicev and M.Sokolova[5] provides robust method for short text classification by using a statistical model,named as Prediction Partial Matching.However, the study is oriented to text containing complex and specific terminology. Partial Matching(PM) compression provides consistent precision of text classification. J.Golbeck and kuter[9] propose a application on social network called as Film Trust that exploits the particular OSN relationship and provenance.Film trust application is introduced where each user trust to movie reviews and ratings of a film. For subscribing rating the criteria is introduced i.e trustworthiness, privacy, vendor reliability, safety and preferences of users.so on the basis of the specified rating gives flexible trust output to end users. M.Carullo, E.Binaghi, and I.Gallo[10] proposes clustering of document is useful in many fields. Their are two categories of clustering general purpose and text oriented contextual features, these both will be used for clustering of data. The result will indicate the power of proposed system. C.D. Manning, P. Raghavan, and H. Schutze[11]to which information retrieval is to be done which hasmodels for text representation vector space model(VSM). However text should be represente by binary or real wieghts on the Document properties(Dp) which characterizes the environment where messages are posted.

III. PROPOSED WORK

The aim of this paper and related work defines to propose and experimentally evaluate an automated system, called Filtered Wall (FW),which is able to filter unwanted messages from OSN user walls. We exploit Machine Learning (ML) text categorization techniques[4] to automatically assign with each short text message classification on a set of categories based on its content. The major efforts is to build a robust Short text classifier (STC) [5], which concentrate in the particular extraction and selection of a set which characterize and discriminate features. In this current paper considering as a learning model we use the neural learning which is recognized as one of the most efficient solution in text classification. In short text classification we use strategy Radial Basis Function Networks(RBFN) for which it proves the capability in acting as a soft classifiers to managing noisy data and intrinsically vague classes. Using neural model in classification strategy the RBFN categorizes short messages as neutral and non neutral messages. Besides classification facilitates, the system provide a powerful rule layer exploiting a flexible language to specify Filtering Rules (FRs), by which users can customize what contents should not be displayed on their wall. FRs supports a variety of different filtering criteria that can be combined and customized according to the user needs where the undesired messages are filtered. More precisely Filtering Rules exploit user profiles, user activities and user relationships as well as the output of the ML categorization process to state the filtering criteria to be enforced. Additionally the system also introduces support for user defined Black lists, where the list of users that are prevented to post any kind of messages on a wall temporarily.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 12, December 2014

A. Proposed System Filtered Wall Architecture:

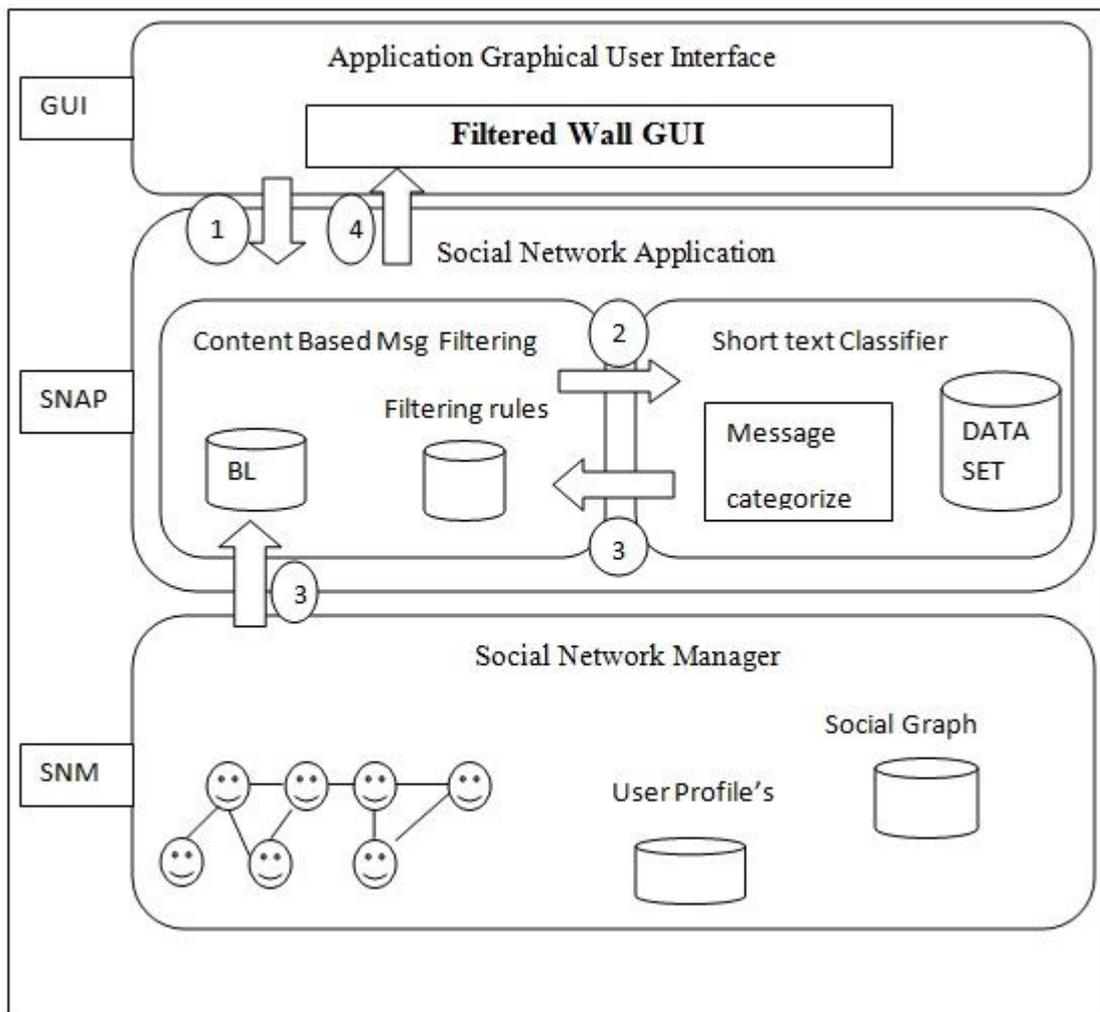


Fig. 1. Filtered Wall Conceptual Architecture

The Filtered wall architecture in support of OSN is a three tier structure where, first layer is called Social Network Manager(SNM),which provides the basics of OSN functionalities(i.e profile and Relationship management). Second Layer provides the support for external Social Network Applications(SNAs) and the supported SNAs may in turn require an additional layer for their needed Graphical user Interfaces(GUIs). Whereas the proposed system is placed in the second and third layer. GUI provides users with a Filtered Wall(FW) where the authorized messages are published according to FRs/BLs, Particularly user interacts through GUI to manage FRs/BLs. The main components of our proposed system are the Content-Based Messages Filtering (CBMF) and the Short Text Classifier (STC) modules. STC classifies messages according to a set of categories. First component exploits the message categorization which is provided by Short text classifier module to enforce the FRs specified by user and BLs also used to enhance the filtering process. From fig.1 the path followed by messages, from its writing to the possible final publication can optimized as follows:

- Whenever user enters the one of the friends or his contacts private wall and wants to post message, Filtered wall intercepts the message.
- Extraction of metadata from the content of message is to be done by using ML based text classifier.



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 12, December 2014

- c. Filtering policies and BL rules are applied on the Metadata which is provided by text classifier and the extracted data available from the user's profiles and social graphs.
- d. FW uses these results from the above step and takes the decision whether to be published or not

IV. SHORT TEXT CLASSIFIER

Proposed study introduces short text classifier where it characterizes classification on the small data sets and short texts. Our goal is designing and representing various discriminant features with a neural learning strategy which categorizes short texts. A hierarchical two level strategy is introduced for better to identify and eliminate "neutral" sentences and classifies "non-neutral" sentences by the class of particular interest, where the short texts are labeled as neutral or non-neutral and further non-neutral texts are classified and filtering process are applied on the texts.

A. Text Representation:

On the basis of extraction of features for a given document Representation of text is an important task where the performance affecting the classification strategy is measured. The survey suggest three types of features considerations for text representation. They are Bag of words (BoW), Document Properties(Dp) and Contextual Features(CF). The BoW, Dp types of features already uses entirely derived from information contained within the text of message which is endogenous whereas the contextual features are exogenous. A contextual features (CF) modelling information introduces the characterizing the environment where the user is posting. Vector space model (VSM) is used for analyze the experimental evaluation of features for the text representation by which the text document is represented as vector of binary or real weights.

B. Description of the Proposed Algorithm:

As short text classification categorizes as hierarchical two level classification process, Where first level classifier performs a binary hard categorization that labels messages as neutral and neutral. First level filtering task facilitates the subsequent second level task in which a finer grained classification is performed. The second level performs the soft partition of non neutral messages. Machine Learning(ML) chooses RBFN model for text classification which have single hidden layer of processing units with local, restricted activation domain. RBFNs advantages are that classification function is non-linear to the model may produce confidence values and it may be robust to outliers.

V. MANAGEMENT OF FILTERING RULES AND BLACKLIST MANAGEMENT

A. Filtering rules:

In this section, we introduce the rules used for filtering unwanted messages which defines the language for FRs specification. We consider three main issues that are affected a message filtering in our opinion. First issue is, user uses OSN in everyday life, the message which posts may have different meanings and relevance based on who writes it. As a consequences, FRs should give the permissions to state constraints on message creators. Where creators on which a FR applied can be selected on different criteria, which is one of the most relevant on the conditions of user profiles attributes. In this way we can apply rules only to young creators or to creators with given religious or political views. From the social network scenario creators may also be identified by exploiting information on social graph. This all are implied to state conditions on the type, depth and trust values of relationships of creators should be involved in order to apply them in specified rules. Where the notion of creator specification can be defined as; A filtering rule FRs is depend on these factors, author, creatorSpec, contentSpec, action, where, the filtering evaluation is to be done.

B. Blacklists:

A further component of our proposed system is a BL mechanism which states that to avoid undesired messages from undesired creators, autonomous from their substances. BLs is straightly supervised by the system, which should be able to check who are the users to be introduced in the BL and makes the decision about users that at what time user is retend from BL. To improve flexibility of system, such information is provided to the system during a set of rules which are set is called BL rules. Such rules are not defined by the SNMP; thus, they are not meant as commonly high-level directives to be practical to the entire society. Rather, we choose to allow the users themselves, i.e., the wall's owners to indicate BL rules that regulates who should be banned from their walls and for how long time he should



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 12, December 2014

remain in BL, and banned to post any kind of message at same time on the other walls. A Blacklist Rule (BL) depends upon these factors, author, creatorSpec, creatorBehavior, where the BL list is to be maintained.

VI. CONCLUSION AND FUTURE WORK

In this paper, A system to filter unwanted messages in OSNs wall is presented, which provides the customizable content based filtering to the user. The first step of the project is to classify the content using several rules applied to available data. Next step is to filtering the undesired messages using rules. Finally, Blacklist rule is introduced so that owner of the user wall can insert the friends who post unwanted messages. Providing this proposed system better privacy is given to the OSN user wall.

REFERENCES

1. Marco Vanetti, Elena Ferrari, Moreno Carullo, Barbara Carminati, Elisabetta Binaghi, and Barbara Carminati, "A System to Filter Unwanted Messages from OSNs User Walls", 2013.
2. M.Chau and H.Chen, "A Machine Learning Approach to Web Page Filtering Using Content and Structural Analysis", *Decision Support Systems(dss)*, vol.44, no.2, pp.482-494, 2008.
3. A.Adomavicius and G.Tuzhilin, "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions", *IEEE Trans. Knowledge and Data Engg.*, vol.17, no.6, pp.734-749, June 2005.
4. F.Sebastiani, "Machine Learning Automated Text Categorization", *ACM Computing survey*, vol.34, no.1, pp.1-47, 2002.
5. V.Bobicev and M.Sokolova, "An Effective and Robust Method for Short Text Classification", *Proc.23rd National Conf. of Artificial Intelligence (AAAI)*, C.P. Gomes and D. Fox, eds., pp.1444-1445, 2008.
6. B.Sriram, D.Fuhry, H.ferhatatosmanoglu, M.Demirbas, and E.Demir, "Short Text Classification in Twitter to Improve Information Filtering", *Proc.33rd Int'l ACM Conf. Research and Development in Information Retrieval(sIGIR '10)*, pp.841-842, 2010.
7. M.Vanetti, M.Carullo, E.Binaghi, B.Carminati, and E.Ferrari, "Content Based Filtering in On-Line Social-Networks", 2010.
8. H.Schutze, J.O.Pedersen, and D.A.Hull, "A Comparison of Classifiers and Document Representations for the Routing Problem", 1995.
9. J.Colbeck, "Combining Provenance with Trust in Social Networks for Semantic Web Content-Based Filtering", *Proc. Int'l conf. Provenance and Annotation of Data*, I. Fosters and L. Morea, eds., pp.101-108, 2006.
10. M.Carullo, E.Binaghi, and I.Gallo, "An Online Document Clustering Technique for Short Text Web Contents", *Pattern Recognition Letter*, vol.30, pp.870-876, July 2009.
11. C.D. Manning, H. Schütze and P. Raghavan, "Introduction to Information Retrieval(IR)", Cambridge Uni. Press, 2008.
12. R.J.Mooney and L.Roy, "Content-Based Book Recommending Using Learning for Text Categorizations", 2000.
13. S.Zelikovitz and H.Hirsh, "Improving Short Text Classification Using Unlabeled Background Knowledge", *Proc. 17th(ICML '00) Int'l Conf. Machine Learning*, P. Langley, ed., pp.1183-1190, 2000.
14. J.Nin, E.Ferrari, B.Carminati, and V.Torra, "Computing Reputation for Collaborative of Private Networks", *Proc.33rd Ann. IEEE Int'l computer software and application Conf.*, Vol.1, pp. 246-253, 2009.