Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 3, Issue. 7, July 2014, pg.942 – 947

RESEARCH ARTICLE

An Unwanted Messages Filtering System from OSN User Walls using Blacklist Mechanism

G. Lavanya[1], Sunil Kumar. V[2]

[1]M.Tech 2nd year, Department of CSE, PBR VITS, Kavali, Nellore, A.P, India

[2]Associate Professor, Department of CSE, PBR VITS, Kavali, Nellore, A.P, India

[1] lavanya.goli215@gmail.com; [2] sunil.vemula1981@gmail.com

Abstract- In This paper we proposes a content-based message filtering conceived and system enforcing machine learning as a key service for On-line Social Networks (OSNs). As we know, today everybody is using On-line Social Networks (OSNs) to communicate and share information. Then one important need in today On-line Social Networks (OSNs) is to give users the capability to control the messages posted on their own private space to avoid that unwanted content is displayed. OSNs provide little support to this requirement up to now. To provide this, we suggest a system allowing OSN users to have a direct control on the messages posted on their walls. This is accomplished through a flexible rule-based system, which allows users to customize the filtering criterion to be applied to their walls, and Machine Learning based soft classifier which automatically produces membership labels in support of content-based filtering. Finally OSN plays a vital role in day to day life. User can communicate with other user by sharing several types of contents like image, audio and video contents. Only the unwanted messages will be blocked not the user. To avoid this issue, BL (Black List) mechanism is proposed in this journal, which avoids undesired creators messages.

Keywords- Pattern matching, Information Filtering, On-line Social Networks, Text Classification, Policy-based Personalization and Blacklist

I. INTRODUCTION

At present the most popular interactive medium to communicate, share and disseminate a considerable amount of human life information are On-line Social Networks (OSNs) [1]. Daily and continuous communications imply the exchange of several types of content, including free text, audio, and image and video data. Then according to Facebook statistics average user creates 90 pieces of content each month, whereas more than 30 billion pieces of content are shared each month. Information filtering can therefore give users the ability to automatically control the messages written on their own walls, by filtering out unwanted messages. Truly, in the present day OSNs provide very little support to prevent unwanted messages on user walls. For example, Facebook lets users to state who is allowed to insert messages in their walls (i.e., friends, friends of

friends, or defined groups of friends and more….). However, content-based preferences are not supported. Then Wall messages are constituted by short text for which traditional classification methods have serious limitations since short texts do not provide sufficient word occurrences.

Hence the aim of the present work is to propose and experimentally evaluate an automated system, called Filtered Wall (FW), able to filter unwanted messages from OSN user walls. And the support for content based user preferences is the key idea of proposed system. This is possible thank to the use of a Machine Learning (ML) text categorization procedure [12] able to automatically assign with each message a set of categories based on its content. The aim of the present work is to propose and experimentally evaluate an automated system, called Filtered Wall (FW), able to filter out unwanted messages from social network user walls. The key idea of the proposed system is the support for content-based user preferences. This is possible thank to the use of a Machine Learning (ML) text categorization procedure [2] able to automatically assign with each message a set of categories based on its content. We believe that the proposed strategy is a key service for social networks in that in today social networks users have little control on the messages displayed on their walls. In contrast, by means of the proposed mechanism, a user can specify what contents should not be displayed on his/her wall, by specifying a set of filtering rules. Filtering rules are very flexible in terms of the filtering requirements they can support, in that they allow to specify filtering conditions based on user profiles, user relationships as well as the output of the ML categorization process. In addition, the system provides the support for user defined blacklist management, that is, list of users that are temporarily prevented to post messages on a user wall.

II. RELATED WORK

Filtering is based on explanations of individual or group information preferences that typically represent long-term interests. Users get only the data that is extracted. Information filtering systems are intended to categorize a stream of dynamically generated information and present it to the user that information that are likely to satisfy user requirements. The aim of the present work is to propose and experimentally evaluate an automated system, called Filtered Wall (FW), able to filter out unwanted messages from social network user walls. The key idea of the proposed system is the support for content-based user preferences. This is possible thank to the use of a Machine Learning (ML) text categorization procedure [2] able to automatically assign with each message a set of categories based on its content. We believe that the proposed strategy is a key service for social networks in that in today social networks users have little control on the messages displayed on their walls [3], [4]. For example, Face book 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.

For instance, it is not possible to prevent political or vulgar messages. In contrast, by means of the proposed mechanism, a user can specify what contents should not be displayed on his/her wall, by specifying a set of filtering rules. Filtering rules are very flexible in terms of the filtering requirements they can support, in that they allow to specify filtering conditions based on user profiles, user relationships as well as the output of the ML categorization process. In addition, the system provides the support for user defined blacklist management, that is, list of users that are temporarily prevented to post messages on a user wall. Exploited by the filtering mechanism and as well as by the language to express filtering rules [6]. In contrast, no one of the access control models previously cited exploit the content of the resources to enforce access control. We believe that this is a fundamental difference. Moreover, the notion of blacklists and their management are not considered by any of these access control models.

Machine learning text categorization technique is also used in proposed, to automatically assign the short text based on the content. J. Golbeck Offered an application, called Film Trust, to personalize access to the website. But, 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 filtering out unwanted information [5]. As far as privacy is concerned, current work is mainly focusing on privacy-preserving data mining skills, that is, protecting information related to the network, i.e., relationships/nodes, In micro blogging services such as Twitter, there may arrive a situation where the users may become overwhelmed by the raw data. One solution to this problem is the classification of short text messages [7]. The proposed approach effectively classifies the text to a

predefined set of generic classes such as News, Events, Opinions, Deals, and Private Messages. So, in this paper there was a focus on to classify news, opinions and other messages according to their categories.

III. SHORT TEXT CLASSIFIER

On datasets with large documents such as newswires corpora, established techniques used for text classification work well but suffer when the documents in the corpus are short. In this context, critical aspects are the definition of a set of characterizing and discriminate features allowing the representation of underlying concepts and the collection of a complete and consistent set of supervised examples [8].

From a ML point of view, we approach the task of short text categorization by defining a hierarchical two level strategy assuming that it is better to identify and eliminate "neutral" sentences, then classify "non-neutral" sentences. The first level task is considered as a hard classification where short texts are labeled with crisp Neutral and Non-Neutral labels. The second level soft classifier acts on the crisp set of non-neutral short texts and, for each of them, it "simply" produces estimated appropriateness or "gradual membership" for each of the conceived classes, without taking any "hard" decision on any of them. Such a list of grades is then used by the successive phases of the filtering process.

3.1 Text demonstration

The most appropriate feature set and feature representation for short text messages have not yet been sufficiently investigated. We consider three types of features, BoW, Document properties (Dp) and Contextual Features (CF). The first two types of features, already used in [9] [10], are endogenous. Text representation using endogenous knowledge has a good general applicability, though in operational settings it is appropriate to use also exogenous knowledge. We introduce contextual features (CF) modelling information that characterize the environment where the user is posting. These features play important role in deterministically understanding the semantics of the messages [12].According to Vector Space Model (VSM) for text representation, a text document d_j is represented as a vector of binary or real weights $d_j = w_{1j}, \ldots, w_{|T|j}$, where T indicates the set of terms that occur at least once in at least one document of the collection Tr, and $w_{kj} \in [0; 1]$ denotes how much term t_k contributes to the semantics of document d_j [11].

3.2 Machine Learning Classification

Short text categorization is a hierarchical two-level classification process. The first-level classifier does a binary hard classification that labels messages as Neutral and Non-Neutral. The first-level filtering task enables the subsequent second-level task in which a finer-grained classification is performed. The second-level classifier carries out a soft-partition of Non-neutral messages assigning a given message a gradual membership to each of the non-neutral classes. We select the RBFN model, among the variety of multi-class ML models well-suited for text classification for the experimented competitive behavior with respect to other state of the art classifiers. The first level classifier is then structured as a regular RBFN. In the second level of the classification stage we introduce a modification of the standard use of RBFN.

IV. FILTERED WALL CONCEPTUAL ARCHITECTURE

The main conceptual architecture of OSN services is a three-tier structure (Figure 1). The first layer is Social Network Manager (SNM), usually aims to provide the basic OSN functionalities like profile and relationship management; however the second layer provides the support for external Social Network Applications (SNAs). And then the supported SNAs may in turn need an additional layer for their desired Graphical User Interfaces (GUIs). By considering this reference architecture, the proposed system is placed in the second and third layers. Users interact with the system by means of a GUI to set up and manage their FRs/BLs. Furthermore, 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 main components of the proposed system are the Content-Based Messages Filtering (CBMF) and the Short Text Classifier (STC) modules. STC goals to categorize messages according to a set of categories.

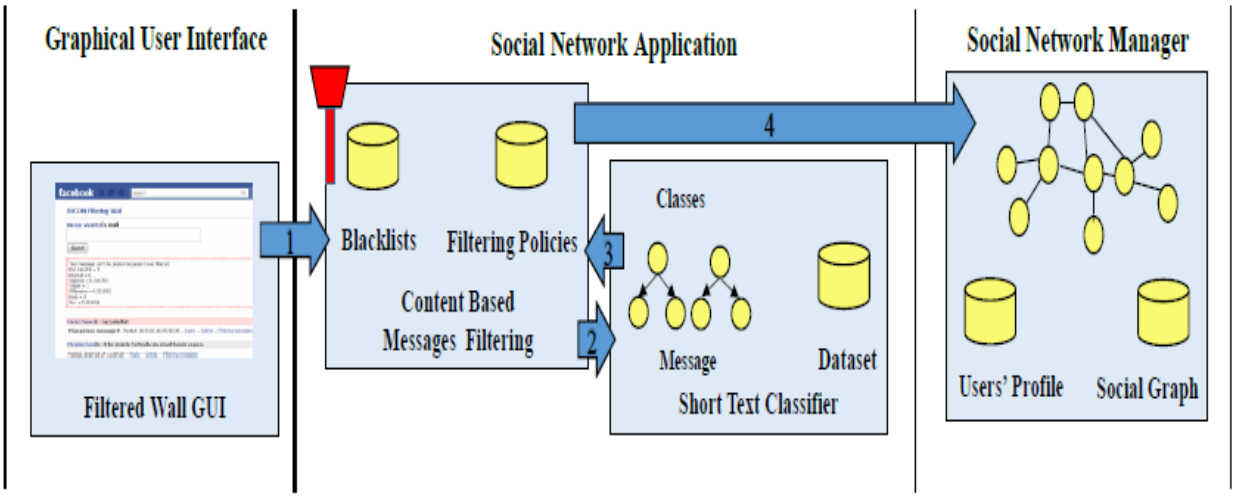

Fig. 1 Filtered Wall Conceptual Architecture

The first component exploits the message categorization provided by the STC module to enforce the FRs specified by the user. As shown in Figure 1, the path followed by a message, from its writing to the possible final publication can be given as follows:

1) The user attempts to post a message after entering the private wall of his/her contacts which is interrupted by FW.
2) A ML-based text classifier extracts metadata from the message content.
3) Metadata together with data extracted from the social graph and users' profiles provided by the classifier is used by FW, to enforce the filtering and BL rules.
4) The message will be published or filtered by FW Depending on the result of the prior step.

Blacklist:

The main implementation of our paper is to execute the Blacklist Mechanism, which will keep away messages from undesired creators. BL are handled undeviating by the system. This will capable to decide the users to be inserted in the blacklist. And then it also decides the user preservation in the Blacklist will get over. Set of rules are applied to improve the stiffness, such rules are called Blacklist rules. By applying the BL rule, owner can categorize which user should be blocked based on the relationship in OSN and the user's profile. And the user may have bad opinion about the users can be banned for an uncertain time period. In this we have information based on two bad attitudes of user, and for that two principles are stated. First one is within a given time period user will be inserted in BL for numerous times, he /she must be worthy for staying in BL for another sometime. This principle will be applied to user who inserted in BL at least once. Relative Frequency is used to find out the system, who messages continue to fail the FR. Two measures can be calculated globally and locally, which will consider only the message in local and in global it will consider all the OSN users walls.

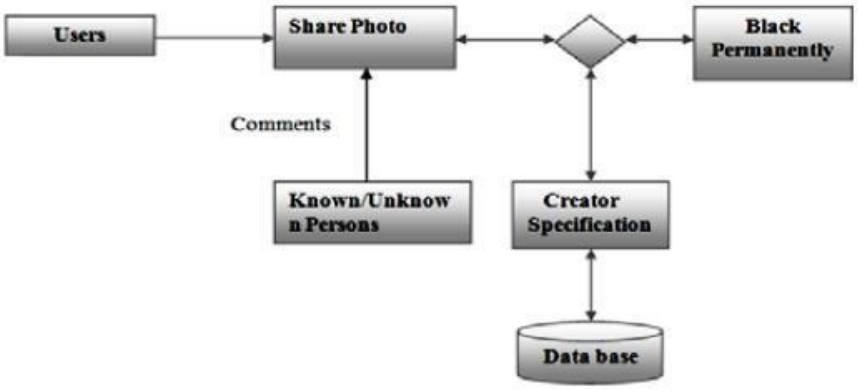

FIG 2: blacklist system

DEFINITION (BL rule) BL rule is a tuple (auth, CreaSpec, CreaB, t), where

- Auth is a user who states the rule.
- CreaSpec is a creator specification.
- CreaB have two components ,RF Blocked and min Banned

RFBlocked= (RF,mode,window) such that

RF=*bMessages/*tMessages

Where *tMessage is the total number of messages that OSN

User recognized using CreaSpec, whereas *bMessage is the number of message in *tMessage that have been blocked.

Window represents the time interval of message creation. minBanned= (min,mode,window)

min is the minimum number of times in the time interval enumerate in window that OSN user recognized using CreaSpec .mode indicates all OSN user.

- T signify the time period the user recognized by CreaSpec and CreaB which will be banned from auth wall.

V. CONCLUSION

In this paper, a system to filter unwanted message in OSN wall is offered. The first step of the paper is to classify the content using several rules. And then next step is to filter the undesired rules. The first concerns the extraction and / or selection of contextual features that have been shown to have a high discriminative power. The second task includes the learning phase. At the same time as the underlying domain is dynamically changing, the collection of pre-classified data may not be representative in the longer term. Finally Blacklist rule is implemented. So that owner of the user can insert the user who posts undesired messages. Better privacy is given to the OSN wall using our system. In future Work, we plan to implement the filtering rules with the aim of bypassing the filtering system, it can be used only for the purpose of overcome the filtering system.

REFERENCES

[1] 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.

[2] P. W. Foltz and S. T. Dumais, "Personalized information delivery: An analysis of information filtering methods," *Communications of the ACM*, vol. 35, no. 12, pp. 51–60, 1992.

[3] P. E. Baclace, "Competitive agents for information filtering," *Communications of the ACM*, vol. 35, no. 12, p. 50, 1992.

[4] Boykin, P.O., Roychowdhury, V.P.: Leveraging social networks to fight spam. IEEE Computer Magazine 38, 61– 67 (2005).

[5] J. Golbeck, "Combining provenance with trust in social networks for semantic web content filtering," *in Provenance and Annotation Data, ser. Lecture Notes in Computer Science*, L. Moreau and I. Foster, Eds. 2006

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

[7] B. Sriram, D. Fuhry, E. Demir, H. Ferhatosmanoglu, and M. Demirbas, "Short text classification in twitter to improve information filtering," 2010P. J. Denning, "Electronic junk," *Communications of the ACM*, vol. 25, no. 3, pp. 163–165, 1982.

[8] D. D. Lewis, Y. Yang, T. G. Rose, and F. Li, "Rcv1: A new benchmark collection for text categorization research," Journal of Machine Learning Research, 2004.

[9] M. Vanetti, E. Binaghi, B. Carminati, M. Carullo, and E. Ferrari, "Content-based filtering in on-line social networks," in Proceedings of ECML/PKDD Workshop on Privacy and Security issues in Data Mining and Machine Learning (PSDML 2010), 2010.

[10] S. Pollock, "A rule-based message filtering system," *ACM Transactions on Office Information Systems,* vol. 6, no. 3, pp. 232–254, 1988.
[11] Strater, K., Richter, H.: Examining privacy and disclosure in a social networking community. In: SOUPS '07: Proceedings of the 3rd symposium on Usable privacy and security, pp. 157– 158. ACM, New York, NY, USA (2007)
[12] F. Sebastiani, "Machine learning in automated text categorization," ACM Computing Surveys, vol. 34, no. 1, pp. 1– 47, 2002.

SHORT BIOGRAPHY

Ms. G. Lavanya received the **B.Tech** Degree in Computer Science and Engineering from Jawaharlal Nehru Technological University, Anantapur, in **2012.**He currently pursing **M.Tech (CSE) in Dept of Computer Science and Engineering** in PBR VITS Engg College, kavali, Nellore**,** under JNTUA University, Anantapur**.**

Vemula.V. Sunil Kumar has received his B.Tech in Electrical Communication Engineering and M.Tech degree in Computer science from JNTU, Hyderabad in 2002 and 2008 respectively. He is dedicated to teaching field from the last 11 years and he has 1 year industrial experience in BEL at Hyderabad. He has guided 12 P.G Students and 25 U.G students. His research areas included CN- MANETs, Neural Networks, and Image processing, embedded systems. At present he is working as Associate professor in PBR Visvodaya Institute of Technology & Science, Kavali, Andhra Pradesh, India.

