A System to Filter Unwanted Messages on Social Networking Site

Nitin Pondhe[1], Prof. H.B.Jadhav[*],

[1]*Student of Master of Engineering (Computer Engineering) VACOE, Ahmednagar, Maharashtra, India*
[1]*pondhenitin007@gmail.com*

[*]*Asst. Prof. (Computer Engineering) VACOE, Ahmednagar, Maharashtra, India*
[*]*hem3577@gmail.com*

Abstract:

Today we mostly using on-line Social Networks (OSNs) for sending the messages to one another but there is no any limitations of any type of message flow. In this project we will give the users the ability to control the message posted on their own private space to avoid that unwanted contend is displayed. This will be achieved through a flexible rule-based system, that allows user to customize the filtering criteria which will be applied on their wall and machine based soft classifier automatically labeling messages in support of contend based filtering. If such kind of posting of unwanted messages on user wall is done many times then system will be put automatically that user in to blacklist. This is achieved through a flexible rule-based system, that allows 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. In content-based filtering each user is assumed to operate independently. As a result, a content-based filtering system selects information items based on the correlation between the content of the items and the user preferences as opposed to a collaborative filtering system. Content-based filtering is mainly based on the use of the ML paradigm according to which a classifier is automatically induced by learning from a set of pre-classified examples. 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. The strategy underlying this module is described in Section IV. 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.

Keyword: *On-Line Social Network, Content-Based Filtering, Short Text Classifier, Machine Learning.*

I.   Introduction:

In the last years, On-line Social Networks (OSNs) have become a popular interactive medium to communicate, share and disseminate a considerable amount of human life information. Daily and continuous communication implies the exchange of several types of content, including free text, image, audio and video data. 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 within the data and then provide an active support in complex and sophisticated tasks involved in social networking analysis and management. A main part of social network content is constituted by short text, a notable example are the messages permanently written by OSN users on particular public/private areas, called in general walls.

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 because using of a Machine Learning (ML) text categorization procedure [4] 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. 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. 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 [5]. In addition, the system provides the support for user defined blacklist management as well as manual blocking, that is, list of users that are temporarily prevented to post messages on a user wall.

II.   FILTERED WALL ARCHITECTURE:
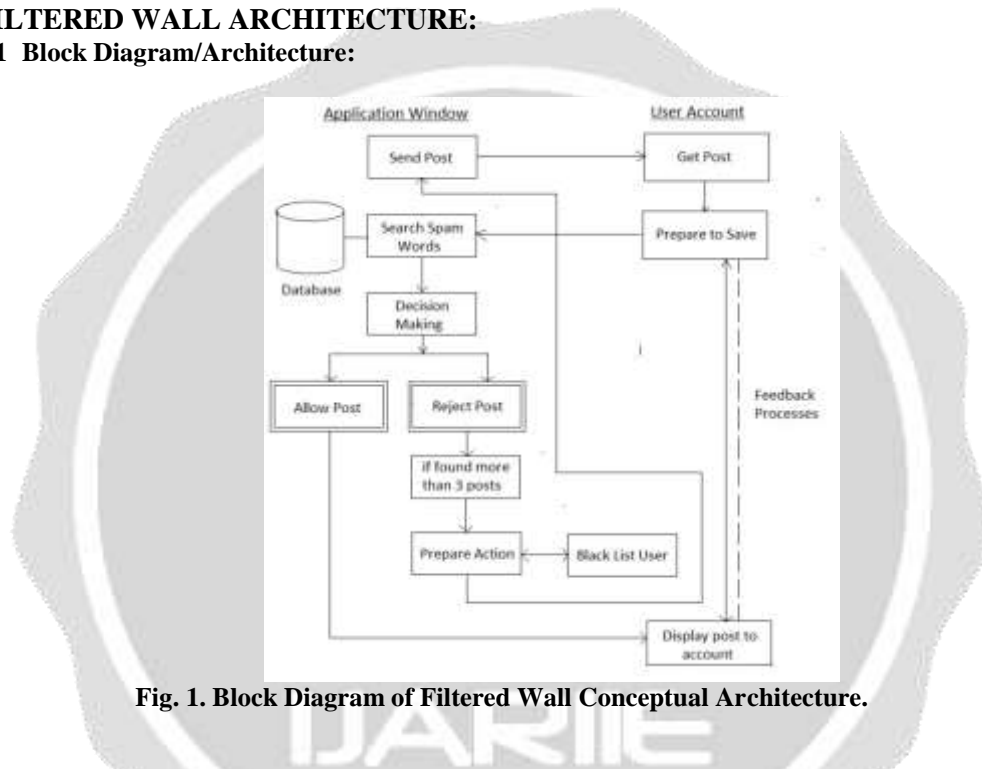### 2.1 Block Diagram/Architecture:

Fig. 1. Block Diagram of Filtered Wall Conceptual Architecture.

In this block diagram the send post method window will allow to send the messages to the user on the user account then the system will filter the spam words using machine learning technique and short text classifier which is present in the database. Then the decision will be takes place that will allows or reject post on user wall also if the same user was trying to post unwanted messages of same users wall number of times then the system will automatic puts user in the black list (i.e. temporarily unavailable). The system also provides manual blocking option to the user.

The conceptual architecture of OSN services is a three-tier structure (Figure 1). The first layer is Social Network Manager (SNM), commonly aims to provide the basic OSN functionalities (i.e., profile and relationship management), however the second layer provides the support for external Social Network Applications (SNAs). 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 classify messages according to a set of categories.

2.2 Filtering rule:
A filtering rule FR is a tuple (author, creatorSpec, contentSpec, action), where:

- author is the user who specifies the rule.
- creatorSpec is a creator specification, specified according to setting rules.
- ContentSpec is a user message which he want to post on another user wall.
- Action denotes the action to be performed by the system on the messages matching contentSpec and created by users identified by creatorSpec which will block or notify the message.

III.   Content-based filtering:

In content-based filtering each user is assumed to operate independently. As a result, a content-based filtering system selects information items based on the correlation between the content of the items and the user preferences as opposed to a collaborative filtering system that chooses items based on the correlation between people with similar preferences [7], [8]. Content-based filtering is mainly based on the use of the ML paradigm according to which a classifier is automatically induced by learning from a set of pre-classified examples. In this section, we introduce the rules adopted for filtering unwanted messages. Another relevant issue to be taken into account in defining a language for filtering rules specification is the support for content-based rules. This means filtering rules identifying messages according to constraints on their contents. In order to specify and enforce these constraints The last component of a filtering rule is the action that the system has to perform on the messages that satisfy the rule. The possible actions we are considering are "block", "publish" and "notify", with the obvious semantics of blocking/publishing the message, or notify the user about the messag11e so to wait him/her decision.

IV.   Short Text Classifier:

Established techniques used for text classification work well on datasets with large documents such as newswires corpora 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.

Our study is aimed at designing and evaluating various representation techniques in combination with a neural learning strategy to semantically categorize short texts. From a ML point of view, we approach the task by defining a hierarchical two level strategy assuming that it is better to identify and eliminate "neutral" sentences, then classify "non neutral" sentences by the class of interest instead of doing everything in one step. This choice is motivated by related work showing advantages in classifying text and/or short texts using a hierarchical strategy. The first level task is conceived as a hard classification in which 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 subsequent phases of the filtering process.

V.   Conclusion:

In this paper, we have presented a system to filter undesired messages from OSN walls. The system exploits a ML soft classifier to enforce customizable content-dependent FRs. Moreover, the flexibility of the system in terms of filtering options is enhanced through the management of Black List. The development of a GUI and a set of related tools to make easier BL and FR specification is also a direction.

In this paper, we have presented a system to filter out undesired messages from OSN walls. The system exploits a ML soft classifier to enforce customizable content-depended filtering rules. Moreover, the flexibility of the system in terms of filtering options is enhanced trough the management of BLs. In this paper we proposed a system with a new framework to solve the user's problem. This new framework can be represented by a text classifier is able to classify a receive post messages on their walls. In this a system to filter out undesired messages from OSN walls. The system exploits a ML soft classifier to enforce customizable content-depended filtering rules. Moreover, the flexibility of the system in terms of filtering options is enhanced trough the management of BLs.

Therefore 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. 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 able to automatically assign with each message a set of categories based on its content.

VI.   Acknowledgement:

VII.   References:

[1] Macro Vanetti, Elisabetta Binaghi, Elena Ferrari, Barbara Carminati, Moreno Carullo,   "A System to filter Unwanted Messages From OSN User Walls" Department of Computer Science and Communication IEEE transaction on knowledge and data engineering vol.25, 2013 .

[2] A. Adomavicius, G.and Tuzhilin, "Toward the next generation of recommender   systems:A survey of the state-of-the-art and possible extensions," IEEE Transaction on Knowledge and Data Engineering,vol. 17, no. 6, pp. 734–749, 2005.

[3] F. Sebastini, "Machnie learning in automated text categorization", ACM Computing Surveys, vol.34, no. 1 pp-47, 2002.

[4] 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.

[5] 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.

[6] S. Pollock, "A rule-based message filtering system," ACM Transactions on Office Information Systems, vol. 6, no. 3, pp. 232–254,1988.

[7] 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.

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