ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

A System to Filter Unwanted Messages from OSN User Walls

B.Janaiah
Department of Computer Science and Engineering
MVSR Engineering College,
Hyderabad, Telangana - 501510, India.

Abstract:
The very basic requirement in present days Online Social Networks (OSNs) or Mobile Social Networks (MSNs) is to enable users with the facility to filter and moderate the messages posted on their own private space(like Profile pages, walls etc) to evade that unnecessary content is displayed. As of now Online Social Networks (OSNs) do not provide any kind options in order to fill this user requirement. To address this issue, in this paper, we propose a system allowing Online Social Networks (OSNs) users to have a direct control on the messages posted on their profiles and walls. This is accomplished through a flexible rule-based system, that permits registered users to modify the filtering criteria to be applied to their walls, and Machine Learning based soft classifier automatically labeling messages in support of content-based filtering.

Keywords: ***Online Social Networks, Information Filtering, Short Text Classification, Policy-based Personalization, User Profile Page.***

Introduction:

A social networking service is a platform to build social networks or social relations among people who share interests, activities, backgrounds or real-life connections. A social network service consists of a representation of each user (often a profile), his or her social links, and a variety of additional services. Social networks [1-5] are web-based services that allow individuals to create a public profile, to create a list of users with whom to share connections, and view and cross the connections within the system. Most social network services are web-based and provide means for users to interact over the Internet, such as e-mail and instant messaging. Social network sites are varied and they incorporate new information and communication tools such as mobile connectivity, photo/video/sharing and blogging.

More and more, the line between mobile and web is being blurred as mobile apps use existing social networks to create native communities and promote discovery, and web-based social networks take advantage of mobile features and accessibility. As mobile web evolved from proprietary mobile technologies and networks, to full mobile access to the Internet, the distinction changed to the following types:
1) Web based social networks being extended for mobile access through mobile browsers and Smartphone apps, and

2) Native mobile social networks with dedicated focus on mobile use like mobile communication, location-based services, and augmented reality, requiring mobile devices and technology.

However, mobile and web-based social networking systems often work symbiotically to spread content, increase accessibility and connect users from wherever they are.

Users of these online networking sites form a social network, which provides a powerful means of organizing and finding useful information. This communication involves exchange of several types of content including text, image, audio and video data. Therefore in Online Social Networks (OSN), there is a chance of posting unwanted content on particular public/private areas, called in general walls [3].

Information filtering has been greatly explored for what concerns textual documents and, more recently, web content. It can be used to give users the ability to automatically control the messages written on their own walls, by filtering out unwanted messages. In this paper, our main aim is to survey the classification technique and to study the design of system to filter the undesired messages from OSN user wall [3], [9-13].

Existing System

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, 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 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. This is because wall messages are constituted by short text for which traditional classification Methods have serious limitations since short texts do not provide sufficient word occurrences [7].

Disadvantages of Existing System:

- 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.
- This is because wall messages are constituted by short text for which traditional classification methods have serious limitations since short texts do not provide sufficient word occurrences.

Proposed System:

The aim of the present work is therefore to propose and experimentally evaluate an automated system, called Filtered Wall (FW), able to filter unwanted messages from OSN user walls. We exploit Machine Learning (ML) text categorization techniques to automatically assign with each short text message a set of categories based on its content [8].

The major efforts in building a robust short text classifier (STC) are concentrated in the extraction and selection of a set of characterizing and discriminant features. The solutions investigated in this paper are an extension of those adopted in a previous work by us from which we inheritthe 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.

In particular, we base the overall short text classification strategy on Radial Basis Function Networks (RBFN) for their proven capabilities in acting as soft classifiers, in managing noisy data and intrinsically vague classes.

Moreover, the speed in performing the learning phase creates the premise for an adequate use in OSN domains, as well as facilitates the experimental evaluation tasks. We insert the neural model within a hierarchical two level classification strategy.

In the first level, the RBFN categorizes [10] short messages as Neutral and Nonneutral; in the second stage, Nonneutral messages are classified producing gradual estimates of appropriateness to each of the considered category. Besides classification facilities, the system provides a powerful rule layer exploiting a flexible language to specify Filtering Rules (FRs), by which users can state what contents, should not be displayed on their walls. FRs can support a variety of different filtering criteria that can be combined and customized according to the user needs.

More precisely, FRs exploit user profiles, user relationships as well as the output of the ML categorization process to state the filtering criteria to be enforced. In addition, the system provides the support for user-defined Blacklists (BLs), that is, lists of users that are temporarily prevented to post any kind of messages on a user wall.

Advantages of Proposed System:

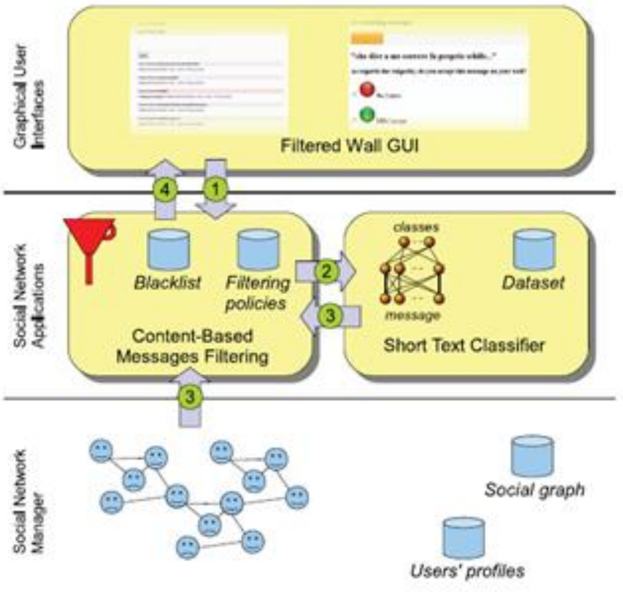
- A system to automatically filter unwanted messages from OSN user walls on the basis of both message content and the message creator relationships and characteristics.
- The current paper substantially extends for what concerns both the rule layer and the classification module.
- Major differences include, a different semantics for filtering rules to better fit the considered domain, 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.

Implementation

Implementation is the stage of the project when the theoretical design is turned out into a working system. Thus it can be considered to be the most critical stage in achieving a successful new system and in giving the user, confidence that the new system will work and be effective [11].

The implementation stage involves careful planning, investigation of the existing system and it’s constraints on implementation, designing of methods to achieve changeover and evaluation of changeover methods.

System Architecture:



Modules:

1. Filtering rules

In defining the language for FRs specification, we consider three main issues that, in our opinion, should affect a message filtering decision. First of all, in OSNs like in everyday life, the same message may have different meanings and relevance based on who writes it. As a consequence, FRs should allow users to state constraints on message creators. Creators on which a FR applies can be selected on the basis of several different criteria; one of the most relevant is by imposing conditions on their profile’s attributes. In such a way it is, for instance, possible to define rules applying only to young creators or to creators with a given religious/political view.

Given the social network scenario, creators may also be identified by exploiting information on their social graph. This implies to state conditions on type, depth and trust values of the relationship(s) creators [6] should be involved in order to apply them the specified rules. All these options are formalized by the notion of creator specification, defined as follows.

2. Online setup assistant for FRs thresholds:

As mentioned in the previous section, we address the problem of setting thresholds to filter rules, by conceiving and implementing within FW, an Online Setup Assistant (OSA) procedure. OSA presents the user with a set of messages selected from the dataset discussed in Section VI-A. For each message, the user tells the system the decision to accept or reject the message. The collection and processing of user decisions on an adequate set of messages distributed over all the classes allows computing customized thresholds representing the user attitude in accepting or rejecting certain contents. Such messages are selected according to the following process. A certain amount of non neutral messages taken from a fraction of the dataset and not belonging to the training/test sets, are classified by the ML in order to have, for each message, the second level class membership values.

3. Blacklists:

A further component of our system is a BL mechanism to avoid messages from undesired creators, independent from their contents. BLs [12] are directly managed by the system, which should be able to determine who are the users to be inserted in the BL and decide when users retention in the BL is finished. To enhance flexibility, such information are given to the system through a set of rules, hereafter called BL rules. Such rules are not defined by the SNM, therefore they are not meant as general high level directives to be applied to the whole community. Rather, we decide to let the users themselves, i.e., the wall's owners to specify BL rules regulating who has to be banned from their walls and for how long. Therefore, a user might be banned from a wall, by, at the same time, being able to post in other walls.

Similar to FRs, our BL rules make the wall owner able to identify users to be blocked according to their profiles as well as their relationships in the OSN. Therefore, by means of a BL rule, wall owners are for example able to ban from their walls users they do not directly know (i.e., with which they have only indirect relationships), or users that are friend of a given person as they may have a bad opinion of this person. This banning can be adopted for an undetermined time period or for a specific time window. Moreover, banning criteria may also take into account users' behavior in the OSN. More precisely, among possible information denoting users' bad behavior we have focused on two main measures. The first is related to the principle that if within a given time interval a user has been inserted into a BL for several times, say greater than a given threshold, he/she might deserve to stay in the BL for another while, as his/her behavior is not improved. This principle works for those users that have been already inserted in the considered BL at least one time. In contrast, to catch new bad behaviors, we use the Relative Frequency (RF) that let the system be able to detect those users whose messages continue to fail the FRs. The two measures can be computed either locally, that is, by considering only the messages and/or the BL of the user specifying the BL rule or globally, that is, by considering all OSN users walls and/or BLs.

Conclusion:

In this paper, we have presented a system to filter undesired messages from OSN walls. The system develops a ML soft classifier to implement customizable content-dependent FRs. In particular, we aim at investigating a tool able to automatically recommend trust values for those contacts user does not individually identified. We do consider that such a tool should propose expectation assessment based on users procedures, performances, and reputation in OSN, which might involve enhancing OSN with assessment methods. Though, the propose of these assessment-based tools is difficult by several concerns, like the suggestions an assessment system might have on users' confidentiality and/or the restrictions on what it is possible to audit in present OSNs. An introduction work in this direction has been prepared in the context of expectation values used for OSN access control purposes. However, we would like to remark that the system proposed in this paper represents just the core set of functionalities needed to provide a sophisticated tool for OSN message filtering.

Still if we have balanced our system with an online associate to set FR thresholds, the improvement of a absolute system effortlessly exploitable by average OSN users is a wide topic which is out of the scope of the present paper.

References:

[1] Marco Vanetti, Elisabetta Binaghi, 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] 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] 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.

[4] 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.

[5] K. Nirmala, S. Satheesh kumar, "A Survey on Text Categorization in Online Social Networks," in Proceedings ofInternational Journal of Emerging Technology and Advanced Engineering,Volume 3, Issue 9, September 2013.

[6] A. K. Jain, R. P.W. Duin, and J. Mao, "Statistical pattern recognition:A review," IEEE Transactions on Pattern Analysis and MachineIntelligence, vol. 22, pp. 4–37, 2000.

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

[8] 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.

[9]P.W. Foltz and S.T. Dumais, "Personalized Information Delivery: An Analysis of Information Filtering Methods," Comm. ACM,vol. 35, no. 12, pp. 51-60, 1992.

[10] S. Zelikovitz and H. Hirsh, "Improving short text classification using unlabeled background knowledge," in Proceedings of 17th International Conference on Machine Learning (ICML-00), P. Langley, Ed.Stanford, US: Morgan Kaufmann Publishers, San Francisco, US,2000

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

[12] J. Moody and C. Darken, "Fast learning in networks of locally-tuned processing units," Neural Computation, vol. 1, p. 281, 1989.

[13] W. B. Frakes and R. A. Baeza-Yates, Eds., Information Retrieval:Data Structures & Algorithms. Prentice-Hall, 1992.