Volume 5, Issue 2, February 2015 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

# An Advanced Approach for Avoiding Unwanted Messages on OSN

**Prof.Devendra Shamkuwar, Ravindra Tambe, Trupti Amburle, Rohit Vishwakarma, Rahul Sarode**

IT Department, SPCOE, Pune University, Pune,

Maharashtra, India

***Abstract— In today's world everyone familiar with social network and most of daily routine communication has carried out through such online social network. This is achieved through a flexible rule-based system, that allows users to customize the filtering criteria to be applied to their walls, but the observation show that as use of OSN get enhanced, an unwanted messages and contents came in front of us and that affect on entire private wall and system. Since such contents are harmful to wall and system through this paper, we are trying to avoid such unwanted messages and contents from unknown entities in a network.***

***Keywords—Social Network; Text Identification; Filtering text based contets; Image Processing; classification;***

## I. INTRODUCTION

Today On-line Social Networks (OSNs) are one of the most popular interactive medium to communicate, disseminate and Share a considerable amount of human life information. Now in daily routine communications imply the exchange of several types like as Content, including (free text, image, and audio and video data). 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. In that OSNs include person- specific information creates both interesting opportunities and challenges. A clean observation shows that OSNs provide very little support to prevent unwanted messages on user walls. For example, data available on social network is useful for marketing products to the right customers. However, preferences content based are not supported. To address some of these limitations, propose an extensible, fine-grained OSN access control model based on semantic web technologies. Information filtering can be used for a different, more sensitive, purpose. Information Filtering can therefore be used to automatically control the messages written on their own walls, by filtering out unwanted messages. This is due to the fact that in OSNs there is the possibility of posting or commenting other posts on particular public/private areas, called in general walls, called Filtered Wall (FW), able to filter unwanted messages from OSN user walls. The Filtered wall (FW) has defined to filter an unwanted contents and messages from own private wall that can used through any OSN. Text categorized technique which is called as Machine Learning (ML), 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.

## II. LITERATURE SURVEY

Here in literature survey discussion of recent methods over the Content-based Filtering in On-line Social Networks. 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 is likely to satisfy user requirements.

Carminati, B., Ferrari[1] all have proposed an extensible fine-grained OSN(Online Social Network) access control model based on word(semantic) web tools. In extra ,They proposed authorization filtering policy and Administration which is model using SWRL & OWL .Within Bellcore, nearly 150 new TMs are published each month, it's an very few are related to any single person's interests. The previous related abstracts feedback using provided a very simple and efficient way of pree-demonstrating (confidence) people's interests. Further this type of work could be attended in the area which is determining a minimum set of access policies which can be used in evaluating access requests in a further attempt to improve the efficiency of these requests.

Churcharoenkrung N., Kim, Y.S., Kang, B.H.[2] all has addressed diversified domains including newswire articles, Internet "news" articles, and broader network resources. That means them focuses on maintainable information filtering system. Here working goes on just prioritizing information by using rating values. The simple and efficient solution to this problem is to block the Web sites by URL, including IP address.

Fang, L., LeFevre, K.[4] has defined Privacy is an important emerging problem in online social networks. While these sites are growing rapidly in popularity, existing policy configuration tools are difficult for average users to understand and use. Template presented for the design of a privacy wizard, which removes much of the burden from individual users. At a high level, the wizard solicits a limited amount of input from the user. Using this input, and other information already visible to the user, the wizard infers a privacy-preference model describing the user's personal privacy preferences. This model, then, is used to automatically configure the user's detailed privacy settings, that offers an automated anti-spam tool that, exploiting the properties of social networks, can recognize unsolicited commercial e-mail, spam and messages related with people the user knows. However, it is important to note that the strategy just stated

 

does not exploit ML content-based techniques. J. Golbeck Offered an application, called FilmTrust, 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.

As work is mainly focusing on privacy-preserving data mining skills, that is, protecting information related to the network, i.e., relationships/nodes, while performing social network analysis. Fong, P.W.L., Anwar,[5] They have formalized the distinct access control paradigm behind the Facebook privacy preservation mechanism into an access control model, which delineates the design space of protection mechanisms under this paradigm of access control. They have also demonstrated how the model can be instantiated to express access control policies that possess rich and natural social significance. In microblogging 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.

As above agenda effectively classifies the text to a predefined set of generic classes such as News, Events, Opinions, Deals, and Private Messages. So, there was a focus on to classify news, opinions and other messages according to their categories.

## III. FILTERED WALL CONCEPTUAL ARCHITECTURE

Basically to define filtering and processing of efficient information speedly a three tier structure has utilized for effective handling of information which may in different format and semantics.As conceptual architecture of OSN services can be define with the help of three-tier structure which is shown in 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.

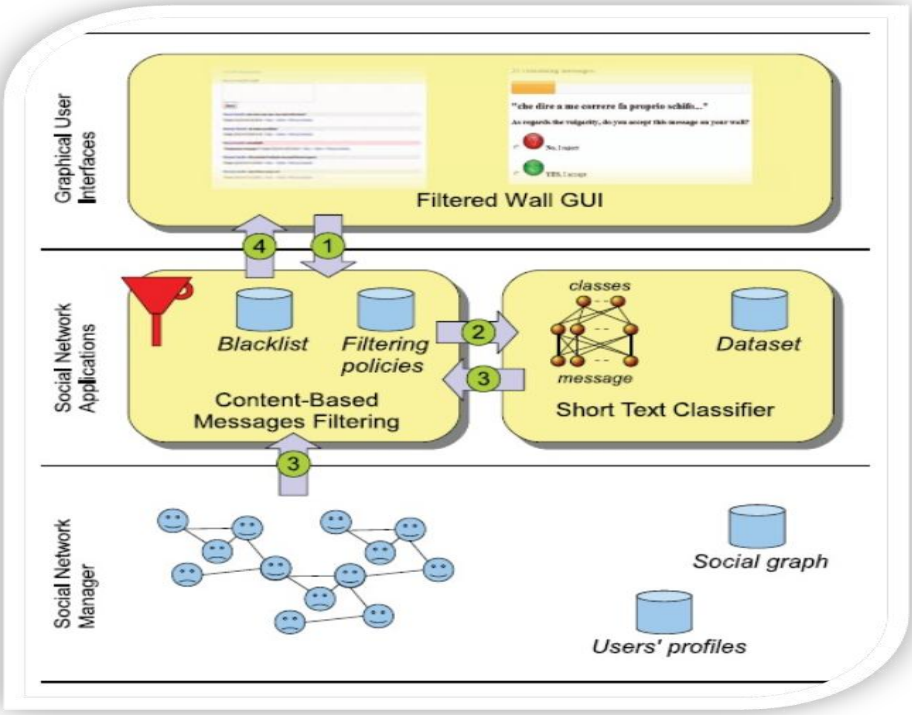


Figure 1: Filtered Wall Conceptual Architecture

The first component exploits the message categorization provided by the STC module to enforce the FRs specified by the user. As shown in Figure1, the path followed by a message, from its writing to the possible final publication can be given as follows:

- The user attempts to post a message after entering the private wall of his/her contacts which is interrupted by FW.
- A ML-based text classifier extracts metadata from the message content.
- 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.
- The message will be published or filtered by FW Depending on the result of the previous step.

 

1) **Social Network manager**- This is a first tier of conceptual architecture as it's having ability to provide basic fulfillment of online social network entities. As social network increasingly goes, on now days since everyone can interacting with other with the help of his/her private wall which insist for effective communication on social site. A statistical graph structure shows the relationship of entities of different objects in a respective network. Social graph also represent the public and private connections of specific object with others.
2) **Social Network Applications**-This tier consist of Black List (BL), Filtering Policy (FP), and Short Text classifier. A Black list contains a queue of words that have to avoid from private wall. Filtering policies are applied as consideration of black list since from black list the selection of prevented words has done. Semantically defined rules used in co-ordination with enlisted words in black list and such contents may change according to the criteria of private wall defined by entity. In Filtering policies specifications affect on a message filtering decision. Filtering policies should allow users to state constraints on message creator. The social network scenario also can be identified by exploiting information on their graph.
3) **Graphical User Interface**- As all knows graphical user interface is an effective platform to interact with any system. It reduces complexity since system becomes user friendly. By considering the view of easy implementation of such complex system possible when efficient graphical terms are used. Here Filtered wall has utilized to filter contents by using GUI terms that shows effective system.

## IV. SHORT TEXT CLASSIFIER

On the Short Text classifier techniques used for the purpose of text classification work on dataset with big document like as newswires corpora. The aimed of our study is designing various representation techniques with a neural learning strategy to categorize short text. 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.

From a ML point of view, 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.

### *A. Text Representation*

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).In Text Representation there are only first two types of features, already used in the number of fifth paper, 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) modeling information that characterize the environment where the user is posting. These features play important role in deterministically understanding the semantics of the messages.

According to Vector Space Model (VSM) for text representation, a text document *dj* is represented as a vector of binary or real weights $d_j = w_{1j}, \ldots\ldots\ldots\ldots\ldots, w_{|T|j}$,where *T* indicates the set of terms that occur at least once in at least one document of the collection *Tr*, and *wkj* € [0; 1] denotes how much term *tk* contributes to the semantics of document *dj*. In the BoW representation, terms are identified with words. For non-binary weighting, the weight *wkj* of term *tk* in document *dj* is computed according to the standard term frequency - inverse document frequency (tf-idf) weighting function, defined as (1) Where # (*tk*, *dj*) indicates the number of times *tk* occurs in *dj*, and *#Tr* (*tk*) indicates the document frequency of term *tk*, i.e., the number of documents in *Tr* in which *tk* occurs. Dp features are heuristically calculated; their definition stems from intuitive considerations, domain specific criteria and in some cases required trial and error procedures.

$$tf - idf(t_k, t_j) = \#(t_k, d_j).\log\ |T_r|/T_r(T_k)$$

a) Correct words: it represents the amount of terms *tk* € *T* ∩ *K*, where *tk* is a term of the considered document *dj* and *K* is a set of known words for the domain language. This value is normalized by #(*t* | *T* | *k*=1 k, *dj*).

b) Bad words: they are determined similarly to the correct words feature, where the set *K* is a collection of "dirty words" for the domain language.

c) Capital words: it represents the amount of words mostly written with capital letters, calculated as the percentage of words within the message, having more than half of the characters in capital case. For example, the value of this feature for the document "To be OR Not to BE" is 0:5 since the words "OR" "Not" and "BE" are considered as capitalized ("To" is not uppercase since the number of capital characters should be strictly greater than the characters count).

d) Punctuations characters: it is computed as the percentage of the punctuation characters over the total number of characters in the message. For example, the value of the feature for the document "Hello!!! How are you doing?" is 5/24.

e) Exclamation marks: it is computed as the percentage of exclamation marks over the total number of punctuation characters in the message. Referring to the aforesaid document, the value is 3/5.

f) Question marks: it is computed as the percentage of question marks over the total number of punctuations characters in the message. Referring to the aforesaid document, the value is 1/5.

 

### *B. Machine Learning-Based Classification*

Short text categorization is a hierarchical two-level classification process from our address. 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. In a classification method has been proposed to categorize short text messages in order to avoid overwhelming users of micro blogging services by raw data. The system described in focuses on Twitter and associates a set of categories with each tweet describing its content. The user can then view only certain types of tweets based on his/her interests. In contrast, an application, called Film Trust, that exploits OSN trust relationships and provenance information to personalize access to the website. However, 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. In contrast, our filtering policy language allows the setting of FRs according to a variety of criteria that do not consider only the results of the classification process but also the relationships of the wall owner with other OSN users as well as information on the user profile. Moreover, our system is complemented by a flexible mechanism for BL management that provides a further opportunity of customization to the filtering procedure.

## V. MANAGEMENT OF FILTERING RULES AND BLACKLIST

In this section, we introduce the rules adopted for filtering unwanted messages. We model a social network as a directed graph, where each node represents a network user and edges represent relationships between two different users. Each edge is labeled by the type of the established relationship (e.g., friend of, colleague of, parent of) and, possibly, the corresponding trust level, which represents how much a given users considers trustworthy with respect to that specific kind of relationship the user with whom he/she is establishing the relationship. We assume that trust levels are rational numbers in the range [0; 1]. There exists a direct relationship of a given type *RT* and trust value *X* between two users, if there is an edge connecting them having the labels *RT* and *X*. Moreover, two users are in an indirect relationship of a given type *RT* if there is a path of more than one edge connecting them.

### *A. Filtering Rules*

Consideration of three main issues involve in defining the language for FRs specification. First is related to the fact that, the same message may have different meanings and relevance based on who writes it, therefore the identification of sender entity is an essential task has carried out then actual message can read. The content of message can show whether this have to display on the wall or not. Message 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. The attributes of private wall or profile are major factors for identification of what kind of user this is? 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 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.

## VII. CONCLUSION

In this paper, we have presented a system to filter unwanted messages from OSN walls, and also we have represent filter Vulgar Images and words .The system exploits a ML (Machine Learning)soft classifier to enforce customizable content-dependent FRs(Filtering Rules). Furthermore, the Efficiency of the system in terms option of filtering is enhanced through the management of BLs(Black List). 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. As the underlying domain is dynamically changing, the collection of pre-classified data may not be representative in the longer term. The present batch learning strategy, based on the preliminary collection of the entire set of labeled data from experts, permitted an accurate experimental evaluation but needs to be developed to include new operational requirements. We plan to address this problem by investigating the use of on-line learning paradigms able to include label feedbacks from users and also we plan image processing in future work. The proposed system may suffer of problems similar to those encountered in the specification of OSN privacy settings. We plan to investigate the development of a GUI and a set of related tools to make easier BL and FR specification, and images as usability is a key requirement for such kind of applications.

## REFERENCES

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

[2] Carminati, B., Ferrari, E., Perego, "*Enforcing access control in web-based social networks*,"ACM Trans. Information System Security, pp. 1–38, 2009.

[3] Churcharoenkrung N., Kim, Y.S., Kang, B.H., "*Dynamic web content filtering based on user's knowledge*," International Conference on Information Technology, Coding and Computing 1, pp. 184– 188 2005.

 

[4] Fang, L., LeFevre, K., "*Privacy wizards for social networking sites*;" In: WWW '10: Proceedings of the 19th international conference on World Wide Web, pp. 351–360, ACM, New York, NY, USA, 2010.

[5] Fong, P.W.L., Anwar, M.M., Zhao, Z., "*A privacy preservation model for facebook-style social network systems*;" In: Proceedings of 14th European Symposium on Research in Computer Security (ESORICS), pp. 303–320, 2009.

[6] Ali B., Villegas W., Maheswaran M., "*A trust based approach for protecting user data in social networks*;" In: Proceedings of the 2007 conference of the center for advanced studies, on Collaborative research, pp. 288–293. ACM, New York, NY, USA, 2007.

[7] Amati, G., Crestani, F.: Probabilistic learning for selective dissemination of information. Information Processing and Management 35(5), 633–654 (1999)

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

[9] 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.

[10] Tootoonchian, A., Gollu, K.K., Saroiu, S., Ganjali, Y., Wolman, A. "*Lockr: social access control for web 2.0*;" In WOSP '08: Proceedings of the first workshop on Online social networks, pp. 43–48. ACM, New York, NY, USA, 2008.

[11] 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, pp.285-297, February 2013.

[12] J . Moody and C. Darken, "*Fast Learning in Networks of Locally- Tuned Processing Units*," Neural Computation, vol. 1, no. 2, pp. 281-294, 1989.

[13] M.J.D. Powell, "*Radial Basis Functions for Multivariable Interpolation: A Review*," Algorithms for Approximation, pp. 143-167, Clarendon Press, 1987.

[14] E.J. Hartman, J.D. Keeler, and J.M. Kowalski, "*Layered Neural Networks with Gaussian Hidden Units as Universal Approximations*," Neural Computation, vol. 2, pp. 210-215, 1990.

[15] J. Park and I.W. Sandberg, "*Approximation and Radial-Basis- Function Networks*," Neural Computation, vol. 5, pp. 305-316, 1993.

[16] A.K. Jain, R.P.W. Duin, and J. Mao, "*Statistical Pattern Recognition: A Review*," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 1, pp. 4-37, Jan. 2000.

 