International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 07, July 2015
Available at http://internationaljournalofresearch.org

A Frame Work for Filtering Unwanted Messages from OSN User Walls

S Satya Avinash Kumar [1] & N Shirisha [2]

[1] Department of Computer Science & Engineering, MLR institute of Technology, Hyderabad, Telangana, India

[2] Assistant Professor ,Department of Computer Science & Engineering, MLR institute of Technology, Hyderabad, Telangana, India

ABSTRACT

The user's can communicate and share their views and content through online social networking services (OSN). The sharing between the users should be several types of content like image, audio, video etc.. In recent years online social networks (OSN) also increased rapidly. Online Social Networks (OSNs) are one of the most wanted and popular social networks where users can share, exchange and communicate with the large amount of information to each other. One basic issue in these OSNs is to give users a protected private space from the undesirable and unwanted messages displayed. So the main goal of the present work is to propose and experimentally evaluate Filtered Wall (FW) which is an automated system can filter noisy messages from OSN user space. This is possible through a flexible rule-based system that allows users to customize the filtering options to be used to their surfaces and a Machine Learning (ML) depended soft classifier involuntarily labeling messages in support of content-based filtering.

Keywords: Online social network (OSN); Short text classification; Message Filtering

INTRODUCTION

Today the most interactive medium to communicate with others is online social network. In online social networking information or content will be shared between the users, the type of contents are audio, video, image. Therefore in Online Social Networks (OSN), there is chance of posting unwanted content on particular public/private area, called in general walls. In this paper we are mainly focus on the text based messages [1]. OSN provide very less amount of security in posting the unwanted messages. Ability of a user to automatically control the messages written on the user wall, by filtering additional communication will be termed as information filtering.

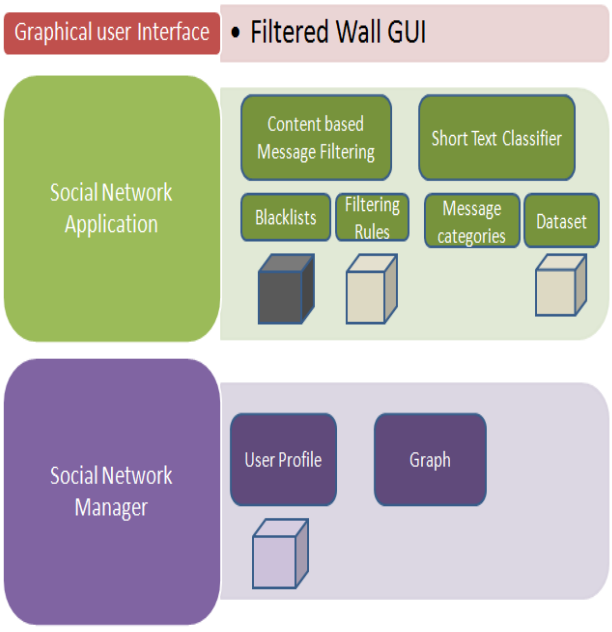
Social networking accounts for 1 of every 6 minutes spent online and as MySpace declines, LinkedIn, Twitter and Tumbler have grown at impressive rates. Social media are becoming increasingly important to recruiters and jobseekers alike. In Online Social Networks (OSNs), there is the possibility of posting or commenting unwanted messages on particular public/private areas, called in general "walls" [2]. Unnecessary posts could spam a user's wall, thus disabling the user from viewing relevant messages. Information filtering can therefore be used to give users the ability to automatically control the messages written on their own walls, by filtering out unwanted messages.

Information Filtering has been widely used and employed for the textual documents and web contents. However, the goal of this proposal is mainly to provide categorization techniques to give the security to user walls from useless and meaningless data [4]. This is especially for that in OSNs, the users can comment the post in

International Journal of Research
International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 07, July 2015
Available at http://internationaljournalofresearch.org

public/private areas of another user walls. These comments can be useless or meaningless or unwanted messages. So, here information filtering plays a vital role to protect the user walls in OSNs from undesired messages and give the authority to user to automatically control the undesired data on their walls.

OSNs today do not provide much support to prevent unwanted messages on user 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). Existing Filters as browser extensions and add-ons are: Spoiler Shield app to filter feeds from both Facebook and twitter, for iPhone and Open Tweet Filter which is a filter for twitter on Chrome [3]. These filters only take keywords and filter out messages that contain the specific words. Applications that are trying to solve this issue through machine learning techniques are either in the beta phase or do not perform efficiently due to poor learning curves used to analyze the messages.



Most of the work related to text filtering by Machine Learning has been applied for long-form text. Wall messages are constituted by short text for which traditional classification methods have serious limitations since short texts do not provide sufficient word occurrences. Thus, a suitable text representation method is proposed in this paper along with a neural-network based classification algorithm to classify each message as neutral or non-neutral, based on its content. Besides classification facilities, Filtering Rules (FRs) can support a variety of different filtering criteria that can be combined and customized according to the user needs [6][4]. More precisely, FRs exploit user profiles, user relationships as well as the output of the classification process to state the filtering criteria to be enforced. In addition, the system provides the support for user-defined Black Lists (BLs), that is, lists of users that are temporarily prevented to post any kind of messages on a user wall.

RELATED WORK

In our daily life continuous communications imply the exchange of several types of contentment, including free text, picture, audio, and video data. According to Facebook statistics average user creates 90 pieces of contentment each month, whereas more than 30 billion pieces of contentment (web links, news stories, notes, picture albums, etc.) are overlapped every month.. The immense and impulsive character of these data creates the premise for the employment of web content mining strategies aimed to automatically discover useful information dormant within the content. They are subservient to render an active support [8].

In complex and sophisticated tasks involved in OSN management, such as for example recover control or data separation. Data filtering has been widely diagnosed for what concerns textual documents in recent, web contentment. More ever, the objective of the priority of these commitments is mainly to provide users a classification mechanism to avoid they are overwhelmed by unnecessary data. In OSNs, data separation can also be used for a unique, more erogenous, cause [9]. This is because of the fact that in OSNs there is the possibility of posting or

International Journal of Research
International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 07, July 2015
Available at http://internationaljournalofresearch.org

commenting other posts on particular open areas, called in open walls. Data separation can therefore be used to give users the ability to automatically control the messages written on their own walls, by separating out unnecessary data. We trust that this is a key note OSN function that has not been rendered so far. In our recent days OSNs render very few support service to prevent unwanted messages on user walls [12].

The main part the system is an information filter which discards the unwanted information. There are number of applications use the concept of information filtering. In recent years Recommender systems have become popular which is a type of information filtering system that predicts the preference of user [13]. It give importance to user interest and recommends an item. Recommender system can works in two ways

- **Collaborative filtering**
- **Content based filtering**

A.**Collaborative filtering:** Collaborative filtering system mainly based on user's preferences, actions and predicts what users will like based on his similarities to other users. User likes and dislikes determine the item rating [2]. A collaborative filtering involves collaboration of multiple agents while filtering information and it requires large dataset. Cold Start, Sparsity, First Rater and Popularity Bias are problems related to collaborative filtering [10].

B. **Content based filtering:** Content based filtering focus on user interest and select items based on it. It suggests the best matched item based on previously chosen item. In content based system each user act as independently. In content based filtering there is no Cold Start, Sparsity and First rater problem. There are some disadvantages also present in content based filtering such as, it requires contents that can be encoded as meaningful features and user taste must be represented as learnable function of the content features [10].

Challenges in the existing Systems:

- However, no content-based preferences are supported and therefore it is not possible to prevent unwanted information, such as policy-making or unrefined ones, doesn't 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 various methods, rather it needs to make ad hoc classification strategies.

- This is due to wall information 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 main motto of the 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 contentment. The chief efforts in developing a strapping short text classifier (STC) are concentrated in the extraction and selection of a set of characterizing and unique features. The solutions inquired in this paper are an extension of those adopted in a previous work by us from which we inherit the learning model and the elicitation procedure for generating pre classified data. FRs can support a variety of different filtering criteria that can be combined and customized according to the user necessity.

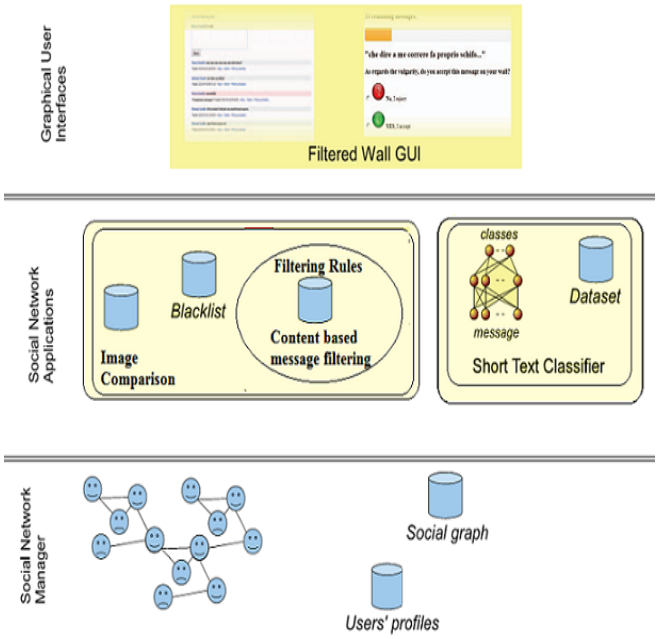
More accurately, FRs utilizes user incites, user associations as well as the creation of the ML categorization process to state the filtering criteria to be implemented. In excess, the system renders the support for user-defined Blacklists (BLs), that is, lists of users that are temporarily prevented to post any kind of messages on a user wall.

International Journal of Research
International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 07, July 2015
Available at http://internationaljournalofresearch.org

In general, the architecture in support of OSN services is a three-tier structure shown in figure. These three layers are

- Social Network Manager (SNM)
- Social Network Application (SNA)
- Graphical User Interface (GUI)

Social Network Manager (SNM) is the first layer, which are commonly aims to provide the basic OSN functionalities such as profile and relationship management. And it maintains the data related to user profile and provides the data to the second layer for applying filtering rules (FR) and blacklist. The second layer is Social Network Application (SNA) which provides support for external social network applications. SNA composed of Black List, Content Based Message Filtering (CBMF), Image comparison and Short Text Classifier (STC) modules.



The STC classify the messages based on its contents CBMF filter messages According to filtering criteria. Graphical User Interface (GUI) is the third layer through which user provide the input and is able to see published messages. In addition to this GUI provide the facility to apply filtering rules and black list for the user in order to avoid unwanted message display. According to this architecture the system is placed in second and third layer.

The short classification module composed of two main phases: Text representation and Machine Learning-based classification

A. Text representation

Text representation is a critical task because of it affects classification process. Many features are there for used in representation of text, but here we judge three types of features .we consider the two types of features, Bag of Words (BOW) and Document properties (DP), that are used in experimental evaluation to determine the combination that is most appropriate for short message classification are considered to be endogenous. Here we introduce Contextual Features (CFS) modeling information that are exogenous in nature and also characterizes the environment where the user is posting.

1. Correct words: It states the amount of terms. Correct words will be calculated.
2. Bad words: comparison to the correct words will be evaluated. Collection of dirty words will be determined.
3. Capital words: It will say about the amount of words written in message. Percentage of words in capital case will be calculated.
4. Punctuations characters: Percentage of punctuation character over the total number of character will be calculated.
5. Exclamation mark: Percentage of exclamation marks over the total number of punctuation characters will be calculated.
6. Question marks: Percentage of question marks over the total number of punctuation character will be evaluated.

The definitions which are used for CFS also will be used for BOW so that CFS and BOW are almost similar.

International Journal of Research
International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 07, July 2015
Available at http://internationaljournalofresearch.org

B. Machine Learning Based classification

In this section we can use any one the machine learning based text classification method which is best for short text classification. Here we suggest a Multi label classifier based on Bayesian network [14].

1. Counting number of words

The word counter algorithm implemented should find out the number of words (short texts) in the message.

2. Stop Word Removal Process

After find out the number of short text we remove the stop words associated with message. In this step we reduce the content size but improve the quality of classification process when all the stop words present in the document are removed.

3. Removal of Special Character

After Stop word removal we go to the process removing Special characters such as —.‖,‖?‖etc. Which again reduce the size of message that is it reduces number of short texts.This improves the quality of STC.

4. Removal of Repeated Words

After the removal of Stop words, Special characters we perform the removal of repeated or duplicate words this also increase the efficiency of Short Text Classifier. In this we keep frequency of occurrence of removed text for future use in computation probability of occurrence.

CONCLUSION

In this paper, we propose a new system called Filtered Wall that allows OSNs users to have a direct control on the messages posted on their private walls by means of a flexible rule based system. In addition to this a OSNs user can enhance the flexibility of the system through filtering rules, blacklist management and an image comparison technique. The system allows the users to customize their walls through the application of filtering criteria and a Machine Learning (ML) based soft classifier automatically labeling messages in support of content-based filtering. Here flexibility of the system can be enhanced through filtering rules and blacklist management. Our proposed system gives security to the Online Social Networks by means of preventing display of unwanted text and image contents. As a future work we are proposing an idea of content based image filtering in large scale based on similarity between images used as in Online Social Networks.

REFERENCES

[1] Marco Vanetti, Elisabetta Binaghi, Elena Ferrari, Barbara Carminati, a Moreno Carullo, "A System to Filter Unwanted Messages from OSN User Walls", 2013.

[2] 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 Eng., vol.17, no.6, pp.734-749, June 2005

[3] T. Joachims, ―Text categorization with support vector machines: Learning with many relevant features,‖ in Proceedings of the European Conference on Machine Learning. Springer, 1998, pp. 137–142.

[4] F.Sebastiani, "Machine Learning Automated Text Categorization", ACM Computing surveys, vol.34, no.1, pp.1-47, 2002

[5] M.Vanetti, E.Binaghi, B.Carminati, M.Carullo, and E.Ferrari, "Content-Based Filtering in On-Line Social Networks", 2010

[6] B.Sriram, D.Fuhry, E.Demir, H.ferhatatosmanoglu, and M.Demirbas, "Short Text Classification in Twitter to Improve Information Filtering," Proc.33rd Int'l ACM SIGIT

Conf. Research and Development in Information Retrieval(sIGIR '10), pp.841-842,2010.

[7] V.Bobicev and M.Sokolova, "An Effective and Robust Method for Short Text Classification," Proc.23rd Nat'l Conf. Artificial Intelligence (AAAI), D.Fox and C.P.Gomes, eds., pp.1444-1445, 2008.

[8] J. A. Golbeck, ―Computing and applying trust in web-based social networks,‖ Ph.D. dissertation, PhD thesis, Graduate School of the University of Maryland, College Park, 2005.

[9] P.S. Jacobs and L.F. Rau, "Scisor: Extracting Information from On- Line News," Comm. ACM, vol. 33, no. 11, pp. 88-97, 1990.

[10] S. Pollock, "A Rule-Based Message Filtering System," ACM Trans. Office Information Systems, vol. 6, no. 3, pp. 232-254, 1988.

[11] P.E. Baclace, "Competitive Agents for Information Filtering," Comm. ACM, vol. 35, no. 12, p. 50, 1992.

[12] P.J. Hayes, P.M. Andersen, I.B. Nirenburg, and L.M. Schmandt, "Tcs: A Shell for Content-Based Text Categorization," Proc. Sixth IEEE Conf. Artificial Intelligence Applications (CAIA '90), pp. 320- 326, 1990.

[13] G. Amati and F. Crestani, "Probabilistic Learning for Selective Dissemination of Information," Information Processing and Management, vol. 35, no. 5, pp. 633-654, 1999.

[14] M.J. Pazzani and D. Billsus, "Learning and Revising User Profiles: The Identification of Interesting Web Sites," Machine Learning, vol. 27, no. 3, pp. 313-331, 1997.