

Classification of Messages and Content based Filtering from OSN User Walls

S. M. Talekar
Research Student
Department of CSE
Government College of Engineering
Aurangabad

M. B. Nagori
Assistant Professor
Department of CSE
Government College of Engineering
Aurangabad

ABSTRACT

Now a day Online Social Networks (OSNs) have become more popular among the internet users. But it has one major problem i.e. to prevent unwanted messages posted on user's personal space. OSNs provide very little support to this requirement. So we have proposed a system which makes the classification of messages and content based filtering from OSN user walls.

This system permits OSN users to have a direct control on the messages posted on their walls by using a flexible rule-based system, that permit users to customize the filtering criteria to be put to their walls and a Machine Learning-based soft classifier that automatically classify and labels the messages in endure of content based filtering.

Keywords

Content based filtering, bag of words, Black list, Machine learning, Filter Wall.

1. INTRODUCTION

Online Social Networks (OSNs) are today one of the most popular interactive medium to communicate, share, and spread a considerable amount of information. Daily and continuous communication leads to exchange of several types of information such as text, image, audio and video. Exchange of information includes free text, image, and audio and video data. The messages posted by OSN users on particular public/private areas, called in general walls are the main part of social network contents. Therefore there is a chance of posting unwanted content on particular public/private areas.

In this paper we have proposed an automated system, called Filtered Wall (FW), able to filter out unwanted messages from social network user walls. This system is the support in content based user preferences. For text classification Machine Learning (ML) text categorization procedure [1] is used. ML is used automatically with each message to set the categories based on its content.

In today's social network, users have little support to control on the messages displayed on their walls. For example, In Facebook little security is provided i.e. who is allowed to insert messages in their walls such as friends, friends of friends, or defined groups of friends. It does not support for content-based preferences. For example, it is not possible to prevent political or vulgar messages. Using this proposed system user can specify a user can specify what contents should not be displayed on his/her with the help of filtering rules. These rules are very flexible in terms of the filtering requirements. Filtering conditions will be based on user profiles and relationship between them.

This paper is organized as follows: in Sect. 2 we describe literature survey related to this paper, Sect. 3 explains the architecture of the proposed system. Sect. 4 describes details of short text classifier. Section 5 describes content based message filtering Sect. 6 concludes the overall proposed system.

2. RELATED WORK

The aim of this paper is to filter the unwanted messages from OSN and to make the content based classification. There are number of system which does the information filtering. Recently Recommender systems [2] have used to predict the preference of user. Recommender system contain two part i.e.

1. Collaborative filtering
2. Content based filtering

2.1 Collaborative filtering

In Collaborative filtering system, preferences, actions and predictions of similarities between likes and dislikes will be considered. According to that the item rating [11] will be determined. In filtering information there is collaboration of multiple agents hence it requires large dataset. This technique faces problems like Cold Start, Sparsity, First Rater and Popularity Bias.

2.2 Content based filtering

Content based technique is based on User's interest .It selects the item based on the correlation between the content of the items and the user preferences. This filtering is based on ML paradigm.

ML techniques [3,9, 10] and other strategies [8, 7] have been used to exploit the Content-based filtering technique However the problem occurred is that content-based filtering on the contents of social network users has achieved few attention in the scientific community. For example Boykin and Roychowdhury [4] have proposed an automated system, exploiting the properties of social networks, able to find out the unsolicited commercial e-mail, spam and messages. But it was not that much useful in Content-based filtering. Content-based filtering is useful in text classification as it is textual in nature. In text classification text classifiers will be created for the classification of new dataset.

In existing technique no content-based preference is supported. Due to this the undesired messages, such as political or vulgar ones cannot be prevented. It follows the following algorithm:

1. In find similar method the Rocchio's formulation is used in which the weight assigned to a term is a combination of its weight in an original query. It

judges the relevant and irrelevant documents from that query.

2. In second step recursive greedy splitting will be used to create the decision_trees. The splits will be chosen using the Bayesian posterior probability of model structure. In this class probability will be given at each node rather than a binary decision.
3. In next step Naive Bayes theorem is used. The training data will be used to calculate the probability of each category which will give the document feature values of a new instance.
4. In the Bayes Nets a limited form of dependence between feature variables is allowed. Due to these very restrictive assumptions of the Naïve Bayes classifier will be considered. 2-dependence Bayesian classifier is used to find out the probability of each feature x_i which is influenced by the appearance or non-appearance of at most two other features.
5. Now days Support Vector Machines (SVM) have become most popular in the learning community. SVM is a hyper plane that distinguishes between set of positive examples from a set of negative examples with maximum margin.

3. PROPOSED SYSTEM

The goal of this paper is to prevent the unwanted messages posted on OSNs user walls using Content based criteria. Thus we have proposed an automated system, called Filtered Wall (FW) able to filter the unwanted messages. Machine Learning (ML) text categorization techniques [5] will be applied automatically to each short text message to find out a set of categories based on its content. In general the social network model acts like directed graph, in which each node indicates a network user and edges indicate relationships between two different users and particular each edge will be labeled by the *type* of the established relationship (e.g., friend of, colleague of, parent of) . Hence it shows a direct relationship between two users. Users can have an indirect relationship if there is a path of more than one edge connecting them of a particular type.

The architecture of an automated system, called Filtered Wall (FW) in support of OSN services is a three-tier structure. Following are the three layers of OSN Filtered Wall.

1. Social Network Manager (SNM)
2. Social Network Application (SNA)
3. Graphical User Interface (GUI)

The first layer is Social Network Manager (SNM) which provides the basic OSN functionalities such as profile and relationship management and it keeps the record of user profile. Social Network Manager provides the data to the second layer for applying filtering rules (FR) and blacklist. Second layer is the Social Network Application (SNA).SNA provides support for external social network applications. It consist of Black List, Content Based Message Filtering (CBMF) [12], Image comparison and Short Text Classifier (STC) modules [13]. The Short Text Classifier classifies the messages based on its contents CBMF filter messages

according to filtering criteria. The third layer is Graphical User Interface (GUI). In this layer user provide the input and user is able to see the posted messages. Graphical User Interface provides the facility for OSN users to apply filtering rules and black list to avoid unwanted message displayed on the wall.

According to this Filtered wall Conceptual Architecture, the proposed system will be placed in second and third layers. User can interact with the system through the GUI in order to apply Filter rule or Black List.

Following flow shows how the message will be published and filtered out in case of unwanted messages.

1. A user will enter to the private wall of his/her contact list. Then he/she will try to post some messages, it may be text or image on his/her wall on which filter rule will be applied.
2. Then we will check whether user belongs to Black List or not. If that user is in BL, then we will block that posted message without considering the contents of the message else go to next step.
3. In this step message (text or image) will be passed to STC or Image comparison module based on its format. The STC extract content of the message using content based filtering. The Image comparator check the image match with the images that already in database or not.
4. FW will use metadata provided by the STC and data extracted from the social graph in order to apply the filtering criteria
5. This the last step in which FW will publish or filter the message will be depend on previous step.

4. SHORT TEXT CLASSIFIER

There are two main phases in STC i.e. Text representation and Machine Learning-based classification

4.1 Text representation

It is a critical task because of it affects classification process. Many features are there but we consider the two types of features i.e. Bag of Words (BOW) and Document properties (DP). Correct words, Bad words, Capital words, Punctuations characters, Exclamation mark and Question marks will be evaluated.

4.2 Machine Learning based classification

In this section we use the machine learning based text classification method for short text classification. The machine learning based text classification method calculates the Counting number of words, Stop Word Removal Process, Removal of Special Character, Removal of Repeated Words, and Multilabel classification.

5. CONTENT BASED MESSAGE FILTERING

Content based message filtering (CBMF) consists of Filtering Rules and Black List.

5.1 Filtering Rules

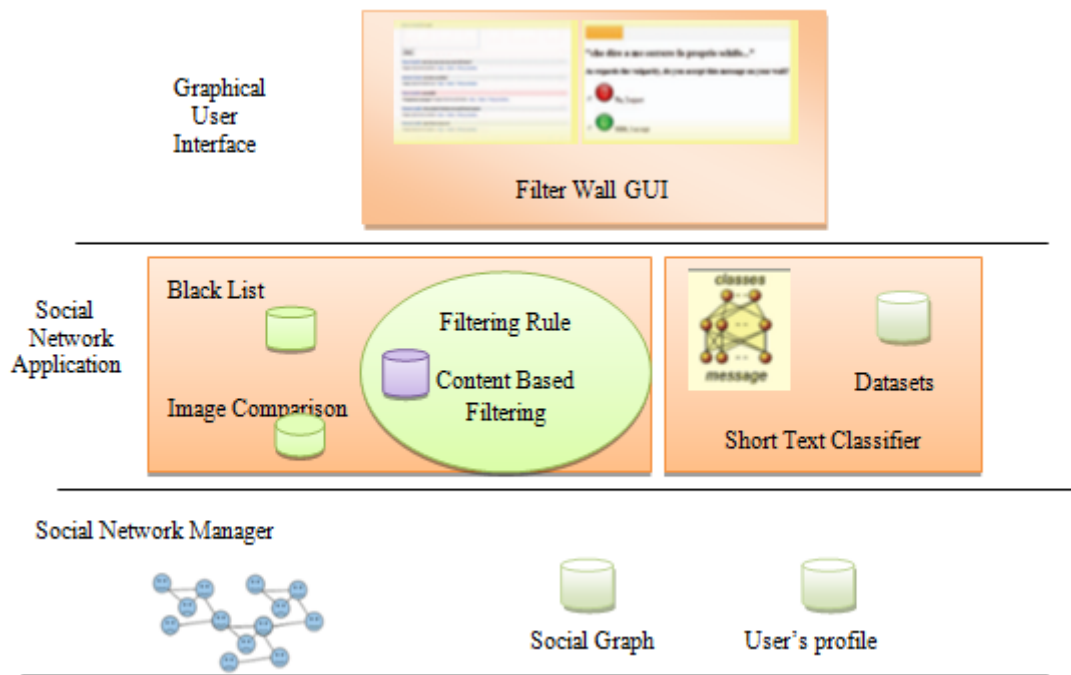


Fig. 1. Filter Wall Architecture

Filtering rules (FRs) are the set of rules, states what content should be displayed by the user. In FR we consider some issues which affect a message filtering decision. While doing communication same message can have different meaning which is depend upon who writes it.

In FRs there should constraints on message creators. Using different several criteria creator should be selected. One of them is using conditions on user's profile attributes.

5.2 Black List

In this Black List (BL) mechanism will be used to find out the undesired user. Following strategy is used to create BL:

1. Which user should be add in BL
2. All messages of that user will be blocked.

6. CONCLUSION

In this paper, we have presented a proposed system which provides the direct control on the messages posted on their private walls using a flexible rule based system. This system provides flexibility to the Online Social Network through filtering rules, blacklist management and an image comparison technique.

Using this system user can prevent their walls through the application of filtering criteria and able to classify the messages using a Machine Learning (ML) based soft classifier. This proposed system provides security to the Online Social Networks by preventing display of unwanted text and image contents. Using this proposed system user can specify a user can specify what contents should not be displayed on his/her with the help of filtering rules.

In future we will enhance an idea of content based image filtering in large scale based on similarity between images

Used as in Online Social Networks.

7. REFERENCES

- [1] Sebastiani, F.: Machine learning in automated text Categorization. *ACM Computing Surveys* 34 (1), 1–47 (2002)
- [2] Marco Vanetti, Elisabetta Binaghi, Elena Ferrari, Barbara Carminati, a Moreno Carullo, "A System to Filter Unwanted Messages from OSN User Walls", 2013.
- [3] Amati, G., Crestani, F.: Probabilistic learning for selective dissemination of information. *Information Processing and Management* 35(5), 633–654 (1999)
- [4] Boykin, P.O., Roychowdhury, and V.P.: Leveraging social networks to fight spam. *IEEE Computer Magazine* 38, 61–67 (2005).
- [5] F. Sebastiani, "Machine Learning in Automated Text Categorization," *ACM Computing Surveys*, vol. 34, no. 1, pp. 1-47, 2002.
- [6] B.Sriram, D.Fuhry, E.Demir, H.ferhatatosmanoglu, and M.Demirbas, "Short Text Classification in Twitter to Improve Information Filtering,"
- [7] Churcharoenkrung, N., Kim, Y.S., Kang, and B.H.: Dynamic web content filtering based onuser's knowledge. *International Conference on Information Technology: Coding and Computing* 1, 184–188 (2005)
- [8] Hanani, U., Shapira, B., Shoval, and P.: Information filtering: Overview of issues, research and systems. *User Modeling and User-Adapted Interaction* 11, 203–259 (2001)
- [9] Kim, Y.H., Hahn, S.Y., Zhang, B.T.: Text filtering by boosting naive bayes classifiers. In: *SIGIR '00: Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval*. pp. 168–175. ACM, New York, NY, USA (2000)

- [10] P´erez-Alc´azar, J.d.J., Calder´on-Benavides, M.L., Gonz´alez-Caro, and C.N.: Towards information filtering system in the web integrating collaborative and content based techniques. In: LA-WEB '03: Proceedings of the First Conference on Latin American Web Congress. p.222. IEEE Computer Society, Washington, DC, USA (2003)
- [11] 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
- [12] 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.
- [13] B.Sriram, D.Fuhry, E.Demir, H.ferhatatosmanoglu, and M.Demirbas, "Short Text Classification in Twitter to Improve Information