

Prediction and Clustering using content based mobile (OSN) Online Social networks

¹Chandralekha . M, PG Scholar, Computer Science Engineering, Bharathidasan Engineering College, Natrampalli, Vellore, Tamil Nadu.

²R. Dhandapani, Asst Professor, Computer Science Engineering, , Bharathidasan Engineering College Natrampalli, Vellore, Tamil Nadu.

³A.Sudhakar, HOD, Computer Science Engineering, , Bharathidasan Engineering College ,Natrampalli, Vellore, Tamil Nadu.

Abstract:

Internet has a big influence in the life of the people in positive way. The usage of internet has increased enormously. In current years Online Social Networks also developed and plays an equivalent role. Online social networks are ideal for exchanging ideas, views and public opinion. Hence Online Social Networks should be highly secure and should protect the individual's privacy. The Online Social Network provides the security measures but they were limited. During Socialization the user can access the profile of other members involved in social sites and even share data such as text, images etc. One crucial issue in user wall is to give users the ability to control the messages posted on their own confidential space in order to avoid unwanted content to be displayed. In this paper, we review filtering methods by using machine learning technique to filter the unwanted messages on the user walls.

Keywords: Internet, filtering, online social network, security, socialization.

1. INTRODUCTION

Social networking sites are a part of daily life and they have brought inventive changes in communication between people. These sites provide diverse resources such as instant messages and email at single place. Miscellaneous type of data can be shared such as images, music, video etc., through the social networking sites. Social network allows user to connect to a variety of pages on the network that includes some useful sites like business, online shopping, marketing, e-commerce, education. Availability of these resources makes the communication easier and faster. Social networking sites help in developing connection with people, friends and relatives. Those People having similar professions can make groups like students, writers, lawyers, social workers, doctors etc. An online social networking site is a platform to create social networks between people who share interests, activities or connections. An online social networking site consists of a profile of each user, links of the user, and a variation of additional services. Online Social networks are web-based services which allow individuals to create a public profile, to create a list of users with whom they can share and view the connections in the system. An Information filtering is a composition that removes repetitive unwanted data from immense collection of information using automated and semi automated methods before the presentation of a human user. In order to execute this, the user's profile is compared to some associating characteristics. Information filtering is used here to manage the massive data from the online social networks. Nowadays online social network provide

restricted support to avoid unwanted messages on user walls. For example, Facebook permits users to manage which user is allowed to insert messages on their walls on the basis of relationship based filtering (i.e., friends, friends of friends, or defined groups of friends). However, content-based techniques are not used and therefore it is not possible to avoid undesired messages, such as religious ones, without taking into consideration about the user who posts them. Content based filtering is preferable for the short texts that occur in messages. In this paper, our main aim is to analyze the classification technique and to design the system to filter the undesirable messages from OSN user wall.

2. RELATED WORK

The aim of the present work is to suggest and experimentally estimate an automated system which is called as Filtered Wall (FW) that should be able to filter unwanted messages from OSN user walls. Machine Learning (ML) text categorization techniques are evolved to automatically assign with each short text message based on its content by using a set of categories. In addition, the system will use a flexible language to demonstrate the filtering rules (FRs), with the help of it the users can decide what contents should be displayed on their walls. The FRs can be personalized according to the users need. Along with it there are user defined blacklists (BLs) which will temporary intercept users to post any type of message on user walls. Content based filtering which is also known as cognitive filtering, recommends items for a user based on the representation of previously evaluated items and information available from the content. The content of each item is represented as a set of descriptors or terms, typically the words that occur in a document. 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 [1]. Selection of item is based on user interest. In the content based filtering, the systems is capable of learning from user's actions related to a particular content and use them for other content types which is the main advantage. Filtering concept is enforced to the Online Social Network user wall using rule based text categorization technique. The latest experiments emphasize complexities, efficiently as short text is brief, with a variety of misspellings, nonstandard conditions, and noise. Zelikovitz et.al tried to improve the classification of short text strings by developing a semi-supervised learning policy based on a combination of labelled training data and a secondary amount of unlabeled but related longer essays.

3. PROPOSED SYSTEM

The aim is to propose and experimentally assess an automated system, Filtered Wall (FW) which enable to filter unwanted messages from OSN user walls. We make full use of Machine Learning (ML) text categorization which automatically assigns with each short text message a set of categories based on its content. Using a hierarchical two stage classification strategy, we insert the neural model. In the first stage, the RBFN categorizes short messages into Neutral and Non-neutral. In the second stage, Non-neutral messages are classified to produce gradual evaluations of appropriateness to each of the reviewed category. Besides classification facilities, the system provides a dominant rule layer exploiting a flexible and workable language to specify filtering rules (FRS), using these users can state what contents, will not be exhibit on their walls. FRS can support different filtering standards that can be combined and modified according to the user needs. FRS exploits user profiles, the output of the ML categorization procedure to

state the filtering standards to be enforced as well as user relationships. In addition, the system calls for black lists (BLS) defined by users, lists of users that are temporarily blocked to post any kind of messages on a user wall and provide the spam. The main function of the proposed system is the content-based message filtering (CBMF) and short text classifier. In addition it supports the classification of message based up on the category set. On datasets with large documents like newswires corpora, already established techniques used for text classification work well but fails when the documents in the corpus are short. In this context, critical aspects are the definition of a set of characterizing and different features allowing the delineation of underlying concepts and the collection of a complete and consistent set of governed examples. From a ML point of view, we approach the task of short text categorization by defining a hierarchical two stage strategy assuming that it is better to identify and eliminate neutral sentences, then classify non neutral sentences. The first stage task is considered as a hard classification where short texts are labeled with crisp Neutral and Non-Neutral labels. The second stage 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.

4. ANALYSIS

In defining the language for FRs specification, we examine three main issues that, should affect a message filtering decision. First of all, the same message may have non-identical meanings and closely connected based on who writes it. As a result, FRs should allow users to impose constraints on message creators. FR applied creators can be selected on the basis of different standards; one of the most promising is by forcing conditions on their profile’s attributes.

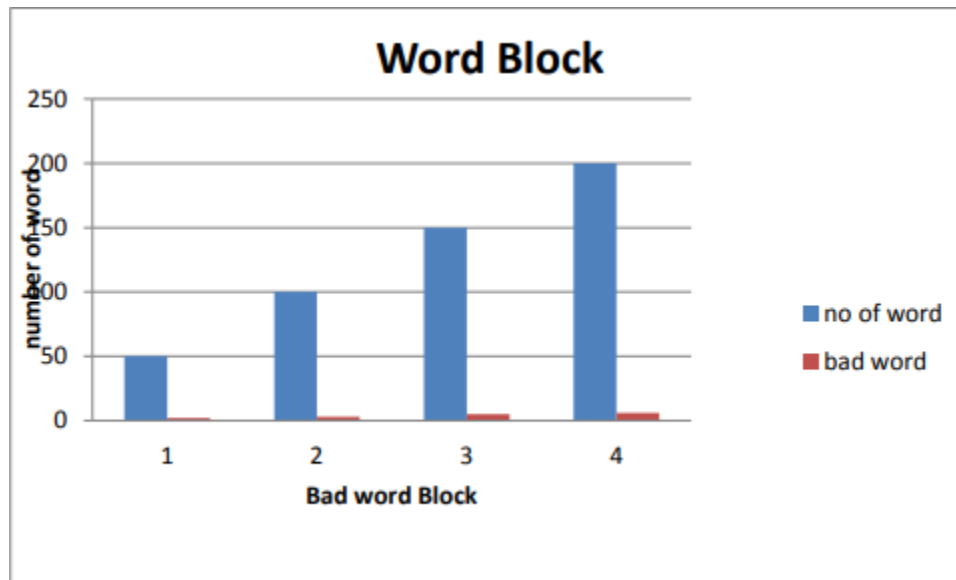


Fig.1. Analysis

For example, possible to specify rules applying only to creators with a given religious or political view or young creators. Given the social network scenario, creators may also be identified by misusing information on their social graph. This indicate to state conditions on depth, type and trust values of the relationships creators should be involved in order to apply them the defined rules. All these issues are formalized by the notion of creator specification. As mentioned in the previous section, we deal with the problem of setting thresholds to filter rules, by formulating and implementing within FW, an Online Setup Assistant (OSA) procedure. OSA confers the user with a set of messages which has to select from the dataset. For each message, the user tells the system his decision to reject or accept 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 defined process. A certain amount of non neutral messages taken from a portion of the dataset, not belonging to the training or test sets, are arranged by the ML in order to have the second stage class membership values for each message. The user interface login form is designed for new user to register by filling the details in the form. The user who is already registered can login by entering the details such as username and password. This form is connected to the database so that the data entered while registering is stored in database. When the registered user enters the correct login details which gets matched with the database, it will display a message showing login successful. Whenever the user tries to post unwanted message on the wall of another user, the word or the message will be filtered and it will display a message. The message will show the unwanted text which the user does not want it to be posted on his wall.

CONCLUSION

In this paper, we have proposed a system to filter undesired messages from OSN walls. The system exploits a ML soft text classifier to impose customizable content-dependent FRS. Besides, the flexibility of the system in terms of filtering criteria is enhanced through the management of BLs. This work is the further step of a wider project. The early motivating results we have obtained on the classification procedure instigate us to continue with other work that will aim to enhance the quality of classification. In particular, future plans contemplate a extensive investigation on two interdependent tasks. The current batch learning strategy, based on the preparatory collection of the entire set of labeled data from experts, allowed an accurate experimental evaluation but needs to be developed to include new operational requirements The development of a GUI and a set of related tools make easier BL and FR specification is also we plan to investigate, since usability is a key requirement for such kind of applications. The future work may include Image Filtering Techniques. In our system we can only filter the text messages. So Image filtering can be tried in future system. Furthermore, the flexibility of the system in terms of filtering options is enhanced through the management of BLs. We can address this problem by investigating the use of on-line learning paradigms able to include label feedbacks from users in future work. We can use the same concept in other social forums. It never block all social medias instead we can apply this on selected space.

REFERENCES

- [1] Marco Vanetti et. Al "A System to Filter Unwanted Messages from OSN User Walls" University of Insubria, Italy IEEE Transactions On Knowledge And Data Engineering Vol:25 Year 2013

- [2] Adomavicius et. Al, "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. Sebastiani, "Machine learning in automated text categorization," ACM Computing Surveys, vol. 34, no. 1, pp. 1– 47, 2002.
- [4] N. J. Belkin et. Al, "Information filtering and information retrieval: Two sides of the same coin?" Communications of the ACM, vol. 35, no. 12, pp. 29–38, 1992
- [5] A. D. Swami et. Al, "A Text Based Filtering System for OSN User Walls" International Journal of Advanced Research in Computer Science and Software Engineering Volume 4, Issue 2, February 2014
- [6] Amruta Kachole et.Al, "Unwanted Message Filtering System from OSNs User's Wall Using Customizable Filtering Rules and Black list Techniques" IJETAE Volume 4, Issue 2, February 2014.
- [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.