

A Classifier System for Unwanted Message Filtering From O.S.N User Space

Sneha R. Harsule¹, Mininath K. Nighot²

¹Savitribai Phule Pune University, Pune, Maharashtra, India

²Professor, Savitribai Phule Pune, Pune, Maharashtra, India

Abstract: As the use of On-line Social Networks (OSNs) rapidly increases, today there is need to provide users with the facility of controlling the messages posted on their private wall to avoid some unwanted content to be displayed. Up till now such requirement is shortly fulfilled by OSNs. But for user convenience, in this paper, a system which we propose is basically to provide users a classification mechanism to avoid useless data. Information filtering technique can also be used for some different and more sensitive, purpose as well. So, mainly our propose system allowing OSN users to have a direct control on the messages posted on their walls. This is possible to achieve through an experimental evaluation an automated system known as Filtered Wall (FW) which has capability to filter unwanted messages from OSN user walls on the basis of both message content and the message creator relationships and characteristics of it. We also insert the neural model in the hierarchical form of two level classification strategies. Besides these classification facilities, our system also provides a powerful rule layer enhancing a flexible language to specify Filtering Rules (FRs), through which users can mention, what contents they do not want to be displayed on their walls and We exploit automatically assign Machine Learning (ML) text categorization techniques for each short text message a set of categories based on its content. This strategy is based on Radial Basis Function Networks (RBFN), managing noisy data and intrinsically vague classes. More- over, the flexibility of the system in terms of filtering options is enhanced through the management of Black lists (BLs).

Keywords: On-line Social Networks, Information Filtering , Filtered Wall, Machine Learning text categorization , Radial Basis Function Networks, Black Lists

1. Introduction

The most popular interactive medium to share the information On-line Social Networks (OSNs). For last few years, Social networks became as hottest online trend. It is now easy to communicate and share information among people situated around the globe. Social networks formed by users of these online networking sites, which provides a powerful means of organizing and finding useful information. Several types of content including text, image, audio and video data are possibly exchanged. Today there is need to improve access control over the contents displayed on the web in order to preserve privacy. Today's smart devices are increasingly becoming an essential part of human life, with the help of these devices today's users wants to keep in touch with others and it is not so difficult right now.

Dream of "Information sharing at your fingertips anywhere anytime ". OSN users still lack in functionality of access control to avoid unwanted messages posted on their wall. Our classifier system is the best solution to this. Therefore one important need in today OSNs is to give users the ability to have control over the messages posted on their own private / public areas that known as general walls, to avoid that unwanted content is displayed. OSNs provide little support to this requirement up to now. Information filtering has been greatly explored for the concept called concerns textual documents and, more recently, web content.

So we propose a system allowing OSN users to have a direct control on the messages posted on their walls. This is achieved by using flexible rule-based system. This system allows users to customize the filtering criteria which are then

to be applied to their walls, and we also using a Machine Learning based soft classifier which automatically produces membership labels in support of content-based filtering. At last we are introducing a Black List mechanism which allows user to state who can posts the messages on their walls. In this paper, our main goal is to study the various classification techniques and also the design of system to filter the undesired messages from OSN user wall.

2. Literature Review

By Studying different research papers, Journals related to this topic provided lot of useful information. This information consist of filtering techniques like Web page filtering, current issues related to such systems, possibility to proceed further with different solutions to the existing problems. Hence it helps a lot to enhance future Social networking. Here are some Papers that we have referred for making literature survey.

In [1] a survey for not only different access control models but also for main contribution correspond to each model is made and studied the basic aspects of access control and classical ACMs. Paper also highlighted privacy problems in OSNs and mentioned solutions to these problems. Finally, suggested solutions to the problems in current ACMs in order to facilitate fine-grained access control and better privacy preservation in OSNs.

Paper [2] have exploit an access control model and related enforcement mechanism for WBSNs, which takes a rule-based approach for describing its policies on the resources owned by network participants, and where authorized users

are denoted in terms of the type, depth, and trust level of relationships. Instead of using traditional ACMs, it preferred to use a semi-decentralized architecture.

Important problem in OSN is privacy. A template to design of a privacy wizard is provided in [3] paper. Input for wizard is from the user and infers a privacy-preference model describing the user's personal privacy preferences and other information already visible to the user. This model, then, is used to automatically configure the user's detailed privacy settings.

Paper [4] mapped Facebook privacy preservation mechanism into an access control model, which delineates the design space of protection mechanisms under this paradigm of access control. Paper worked on the fact that how the model can be instantiated to express access control policies that possess rich and natural social significance.

In [5] paper showed the similarity between Information filtering and Information retrieval. Paper studied tested methods for purpose of prediction that which Technical Memos match people's technical interests. Feedback was taken from previous related abstracts which help in an efficient and simple way of demonstrating people's interests.

In [6] paper addressed different domains including newswire articles, Internet "news" articles, and other network resources. And also prioritizing information on the basis of rating values. The Amalthea, an ecosystem of multiple agents and its application to personalized information filtering and discovery were described. Their approach, through its market-based dynamics, implies that agents that are of service to users or other agents will run more often, reproduce, and survive while in-competent agents will be removed.

This [7] paper offered an automated anti-spam tool that, improving the properties of social networks, can recognize unsolicited commercial e-mail, spam and messages related with people the user knows. Also Offered an application, called Film Trust. But, such systems do not provide a filtering policy layer to decide how and to which extent filtering out unwanted information.

Paper [8] mainly work is mainly focusing on privacy-preserving data mining skills, that is, protecting information related to the network, means relationships / nodes, while performing social network analysis.

3. Proposed Work

Designing and implementing a filtering System to provide users with classification Mechanism to avoid they are overpowered by useless data.

For our convenience, in [13] paper, a system which propose is basically to provide users in the OSNs with the classification mechanism to avoid useless data. Information filtering technique can also be used for some different and more sensitive, purpose as well. This is possible to achieve

through an experimental evaluation an automated system known as Filtered Wall (FW) which has capability to filter unwanted messages from OSN user walls on the basis of both message content and the message creator relationships and characteristics of it. We also insert the neural model in the hierarchical form of two level classification strategies.

Besides these classification facilities, system also provides a powerful rule layer enhancing a flexible language to specify Filtering Rules (FRs), through which users can mention, what contents they do not want to be displayed on their walls and We exploit automatically assign Machine Learning (ML) text categorization techniques for each short text message a set of categories based on its content. This strategy is based on Radial Basis Function Networks (RBFN), managing noisy data and intrinsically vague classes. Moreover, the flexibility of the system in terms of filtering options is enhanced through the management of Black lists (BLs).

We are also planning to enhance other classification techniques to achieve our goal. Our aim is to build automated customizable filtering system that will enhance the utilization of Online Social Network (O.S.N). It takes into account user profile and compares it with referred characteristics or properties. Recommender systems have become popular in recent years. It includes user interest and recommends an item. Our proposed systems works in one of following ways:

- Content based filtering
- Collaborative filtering
- Policy-Based Personalization

User can state what contents should be blocked or displayed on filtered wall by the phenomenon called Filtering rules. These are specified on the basis of user profile as well as user social relationship. Users whose messages are prevented independent from their contents are the Black Lists users.

Our future approach might be like trying to use some another classification technique that will make the enhancement in the Online Social Networking, especially related to sensitive data on the user's wall. We investigate the use of N-grams for data classification purposes. Our method will be based on using the N-grams frequency to classify documents in order to detect and prevent leakage of sensitive data. This will be definitely going to beneficial for our proposed system.

4. Research Methodology

- General Algorithm

Step 1 : Start
Step 2 : A User tries to enter and post the message in a wall.
Step 3 : Apply Machine learning technique.
Step 4 : Machine Learning Checks each word of the user's message.
Step 5 : If (Word = Good Word) then User is able to posted message on the wall.
Step 6 : Else if (Word = Bad Word)
Step 7 : Bad Words are rejected using Blacklist and filtered message posted on the wall.

Step8 : Stop.

• **Method**

The path followed by a message, from its writing to the possible final publication can be summarized as follows:

- 1) User tries to post a message after entering into private wall, which is intercepted by FW.
- 2) A ML-based text classifier then applied automatically on that message and it extracts metadata from the content of the message.
- 3) FW uses metadata provided by the classifier, used to enforce the and BL rules.
- 4) Based on the outcome of our classifier system, the message will be either published or filtered by FW Different strategies and algorithm we are used in our project at different stages of implementation such as for text representation we used vector space model (V.S.M). For short text classifier we studied different access control methods. We used neural machine learning technique and prototype application called DICOMFW and many more.

5. Conclusion

We have successfully proposed our customizable wall. In this paper, we offered A system that filters unwanted message in OSN user's space. Primarily we classified the content using several rules nothing but content based classification. In the next step we performed policy based. The first step related with the extraction and / or selection of contextual features that have been shown to have a high discriminative power. And second concerns with the learning phase.

As the underlying domain is dynamically changing, so we simultaneously deal with the collection of pre-classified data that may not be representative in the longer term. At the end Blacklist rule is implemented. Blacklist rule provides the ability that owner of the user can insert the user who posts undesired messages. Efficient privacy is given to the OSN user's wall using our proposed system. In future Work, for the purpose of overcome problems with the system, we plan to implement the rules with the aim of bypassing the system.

References

- [1] "Access Control Models For Online Social Networks " : Rula Sayaf and Dave Clarke, IBBT- Distrinet Dept. of Computer Science, Katholieke Universiteit Leuven, Belgium, [2011].
- [2] "Enforcing Access Control in Web-based Social Networks" : Barbara Carminati, Elena Ferrari, And Andrea Perego, DICOM, Università degli Studi dell'Insubria, Varese, Italy, ACM Journal Name, Vol. V, No. N, Month 20YY, [2008].
- [3] "Privacy Wizards for Social Networking Sites " : Lujun Fang and Kristen LeFevre, Electrical Engineering Computer Science, University of Michigan, Electrical Engineering Computer Science, University of Michigan 2260 Hayward Ave. Ann Arbor, MI 48109 USA, ljfang@umich.edu , klefevre@umich.edu , April 26-30 , Raleigh , NC , US, [2010].
- [4] "A Privacy Preservation Model for Facebook-Style Social Network Systems" : Philip W. L. Fong and Mohd Anwar, Department of Computer Science University of Calgary, Calgary, Alberta, Canada, (pwl.fong, manwar)@ucalgary.ca; Zhen Zhao, Department of Computer Science University of Regina Regina, Saskatchewan, Canada, zhao112z@uregina.ca, [2009].
- [5] "Information and information retrieval: Two sides of the same coin?" : Nicholas J. Belkin and W. Bruce Croft, Communications of the ACM, v35 n12 p29(10), Dec [1992].
- [6] "Evolving a Multi-agent Information Filtering Solution in Amalthea" : Alexandros Moukas and Giorgos Zacharia, Autonomous Agents Group MIT Media Laboratory 20 Ames street, Cambridge MA 02139-4307, (mouk, lysi)@media.mit.edu, [1996].
- [7] "Leveraging Social Networks to Fight Spam" : P. Oscar Boykin, University of Florida; Vwani P. Roychowdhury, University of California, Los Angeles, 0018-9162-05-20.00 IEEE , April [2005].
- [8] "Combining Provenance with Trust in Social Networks for Semantic Web Content Filtering" : Jennifer Golbeck , University of Maryland, College Park, College Park MD 20742, USA, golbeck@cs.umd.edu, [2006].
- [9] "An Effective and Robust Method for Short Text Classification" : Victoria Bobicev, Technical University of Moldova Studentilor, 7, Chisinau, Moldova, vika@rol.md; Marina Sokolova, CHEO Research Institute 401 Smyth Road, Ottawa, Ontario, Canada, msokolova@ehealthinformation.ca , Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence [2008].
- [10] "Towards Online Spam Filtering in Social Networks" : Hongyu Gao , Northwestern University Evanston , IL , USA, hygao@u.northwestern.edu ; Yan Chen, Northwestern University Evanston, IL , USA , ychen@northwestern.edu ; Kathy Lee , Northwestern University , Evanston, IL , USA , kml649@eecs.northwestern.edu ; Diana Palsetia , Northwestern University Evanston , IL , USA palsetia@u.northwestern.edu ; Alok Choudhary , Northwestern University Evanston , IL , USA choudhar@eecs.northwestern.edu , [2011].
- [11] "SafeChat: A Tool To Shield Children's Communication From Explicit Messages" : Gunter Fahrnberger, Deveeshree Nayakt, Venkata Swamy Martha+ and Srini Ramaswamy, University of Hagen, North Rhine-Westphalia, Germany, guenter.fahrnberger@fernuni-hagen.de; University of Memphis, Tennessee, USA, nayak.deveeshree.2013@ieee.org; @WalmartLabs, Mountain View, California, USA, martha@walmartlabs.com; Industrial Software Systems, ABB Corporate Research Center, Bangalore, Karnataka, India, srini@ieee.org, 978-1-4799-5350-9-14-31.00 IEEE, [2014].
- [12] "Online Social Networks: Threats and Solutions" : Michael Fire, Member, IEEE, Roy Goldschmidt, and Yuval Elovici, Member, IEEE, IEEE Communication Surveys Tutorials, Vol. 16, No. 4, Fourth Quarter [2014].
- [13] "A System to Filter Unwanted Messages from OSN User Walls" : Marco Vanetti, Elisabetta Binaghi, Elena

Ferrari, Barbara Carminati, Moreno Carullo, Department of Computer Science and Communication, University of Insubria , 21100 Varese, Italy, IEEE Transactions On Knowledge And Data Engineering Vol:25,[2013].

Author Profile

Sneha R. Harsule, Student of K.J.College of Engineering & Research Pune,

Prof. Mininath K. Nighot, Faculty of K.J.College of Engineering & Research Pune.