Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 4, Issue. 5, May 2015, pg.940 – 944

RESEARCH ARTICLE

Protecting Social Network Users from Spam Messages Using Machine Learning Algorithm

Nagasree K[1], Asst.Prof. Anita Patil[2]

[1]M.Tech 4th semester, Department of CSE, BITM, Bellary, Karnataka, India
nagasreek70@gmail.com
[2]Assistant Professor, Department of CSE, BITM, Bellary, Karnataka, India
Anitha.bijapur@gmail.com

ABSTRACT: Online Social Networks (OSNs) have become an important part of daily life. Users build networks to represent their social relationships. Users can upload and exchange the information related to their personal lives. Online Social Networks is to give users the ability to control the messages posted on their own private space to avoid unwanted content is displayed. The project proposes a new system allowing OSN users to have a direct control on the messages posted on their walls. This can be done by a flexible rule-based system that allows users to provide the filtering criteria to be applied to their user walls, and a classifier automatically labeling messages in support of content-based filtering. The system exploits a classifier to enforce customizable content-dependent Filtering Rules.

INTRODUCTION

Online Social Networks (OSNs) are today one of the most popular interactive medium to share, communicate, and distribute an important amount of human living information. On a daily basis and continuous messages involve the swap of several types of content, including free content, image, audio, and video information. A main part of social network content is constituted by short text, a notable example are the messages permanently written by OSN users on particular public/private areas, called in general walls.

In OSNs, information filtering can also be used for a different, more sensitive, purpose. 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. 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. Today OSNs provide very little 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). However, no content-based preferences are supported and therefore it is not possible to prevent undesired messages, such as political or vulgar ones, no matter of the user who posts them. The aim of the present work is to propose and experimentally evaluate an automated System, called Filtered Wall (FW), able to filter out unwanted messages from social network user walls. It exploits Machine Learning(ML) text categorization techniques to automatically assign with each short text message a set of categories based on its content.

The most important efforts in building a robust short text classifier (STC) are concentrated in the extraction and selection of a set of characterizing and discriminant aspects. The original set of aspects, derived from endogenous assets of short texts, is inflamed here including exogenous information associated to the context from which the messages begin.

In particular, it presents the overall short text classification strategy on Radial basis Function Networks (RBFN) for their proven capabilities in acting as soft classifiers, in administration noisy information and essentially unclear classes. Furthermore, the speed in achieving the learning stage creates the premise for an adequate use in OSN fields.

It inserts the neural model within a hierarchical two level classification strategy. In the first level, the RBFN categorizes short messages as Neutral and Non-neutral; in the second stage, Non-neutral messages are classified producing gradual estimates of appropriateness to each of the considered category.

The system provides a powerful rule layer exploiting a flexible language to specify Filtering Rules (FRs), by which users can state what contents should not be displayed on their walls. FRs can support a variety of different filtering criteria that can be combined and customized according to the user needs. More precisely, FRs exploit user profiles, user relationships as well as the output of the ML categorization process to state the filtering criteria to be enforced.

RELATED WORKS

Carminati [1] proposed an extensible fine-grained online social network access control model based on semantic web tools. In addition, they propose authorization, administration and filtering policies that are modeled using OWL and SWRL. The architecture of a framework in support of this model has also been presented. Further, they have implemented a version of this framework and presented experimental results for the length of time access control can be evaluated using this scheme. Further work could be conducted in the area of determining a minimal set of access policies that could be used in evaluating access requests in a further attempt to increase the efficiency of these requests.

Carminati [2] have proposed an access control model and related enforcement mechanism for WBSNs, which adopts a rule-based approach for specifying access control 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. Differently from traditional access control systems, our mechanism makes use of a semi-decentralized architecture, where the information concerning users' relationships is encoded into certificates, stored by a certificate server, whereas access control enforcement is carried out client-side.

Churcharoenkrung et al [3] focuses on the development of a maintainable information filtering system. The simple and efficient solution to this problem is to block the Web sites by URL, including IP address. However, it is not efficient for unknown Web sites and it is difficult to obtain complete block list. Content based filtering is suggested to overcome this problem as an additional strategy of URL filtering. The manual rule based method is widely applied in current content filtering systems, but they overlook the knowledge acquisition bottleneck problems. To solve this problem, we employed the multiple classification ripple-down rules (MCRDR) knowledge acquisition method, which allows the domain expert to maintain the knowledge base without the help of knowledge engineers. Throughout this study, they prove the MCRDR based information filtering system can easily prevent unknown Web information from being delivered and easily maintain the knowledge base for the filtering system.

Fang et al [4] explains 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. This paper presented a template 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.

Fong et al [5] 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.

PROPOSED SYSTEM

This project proposes an automated system called filtered wall where the unwanted messages will be prevented. The core components of the proposed system are the Content-Based Messages Filtering (CBMF) and the Short Text Classifier modules. Content based message filtering exploits the message categorization provided by the STC module to enforce the FRs specified by the user. Blacklists can also be used to enhance the filtering process.

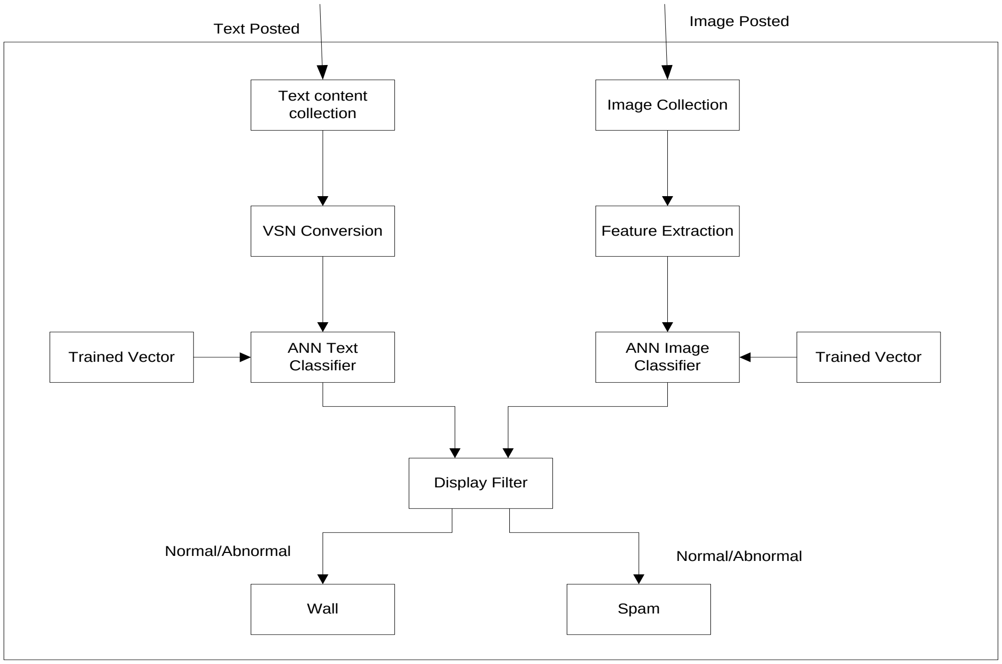

Fig 1: System Architecture

CONCLUSION

This project presents an automated system to filter unwanted messages from user walls. It provides ML soft classifier to enforce content based filtering rules. The flexibility of the system depends on the management of blacklists.

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.