International Journal on Applications in Information and Communication Engineering
Volume 1: Issue 8: August 2015, pp 23-25. www.aetsjournal.com ISSN (Online) : 2394-6237

Designing A System To Filter Out Unwanted Messages From Osn(Online Social Networks) User Walls Using Machine Learning Technique

P.Gowsalya , S.Brindha

Abstract— In recent years, Online Social Networks (OSNs) have become an important part of daily life. Users build explicit networks to represent their social relationships. Users can upload and share information related to their personal lives. The potential privacy risks of such behavior are often ignored. And the fundamental issue in today On-line Social Networks is to give users the ability to control the messages posted on their own private space to avoid that unwanted content is displayed. Today OSNs provide very little support to prevent unwanted messages on user walls. For that purpose, we proposed a new system allowing OSN users to have a direct control on the messages posted on their walls. This is achieved through a flexible rule-based system, that allows users to customize the filtering criteria to be applied to their walls, and a Machine Learning (ML) based soft classifier automatically labeling messages in support of content-based filtering. The system exploits a ML soft classifier to enforce customizable content-dependent Filtering Rules. And the flexibility of the system in terms of filtering options is enhanced through the management of Blacklists. The proposed system gives security to the On-line Social Networks.

Keywords — Online Social Networks, Machine Learning, Filtering Rules, Content-based filtering, Filtering system.

I. INTRODUCTION

Information and communication technology plays a significant role in today's networked society. It has affected the online interaction between users, who are aware of security applications and their implications on personal privacy. There is a need to develop more security mechanisms for different communication technologies, particularly online social networks. OSNs provide very little support to prevent unwanted messages on user walls. With the lack of classification or filtering tools, the user receives all messages posted by the users he follows. In most cases, the user receives a noisy stream of updates. In this paper, an information Filtering system is introduced. The systemFocuses on one kind of feeds: Lists which are a manually selected group of users on OSN. List feeds tend to be focused on specific topics, however it is still noisy due to irrelevant messages. Therefore, we propose an online filtering system, which extracts the such topics in a list, filtering out irrelevant messages[1].

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. In the proposed system 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. The aim of the present 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 [2] to automatically assign with each short text message a set of categories based on its content. The major efforts in building a robust short text classifier are concentrated in the extraction and selection of a set of characterizing and discriminant features.

P.Gowsalya , Student,CSE Department, K.L.N College of Engg, Madurai.(Email: devikowsalya10@gmail.com)
S.Brindha, Professor, CSE Department, K.L.N College of Engg, Madurai.(Email: brindha1963@gmail.com)

II. LITERATURE SURVEY

A distinction is made between two types of text filtering systems: content-based and social filtering systems. In content-based systems, filtering is done by exploiting the information extracted from the text of documents. In social filtering systems, documents are filtered based on annotations made by prior readers of the documents. With respect to this framework, our system is closer to content-based filtering systems, however we utilize other sources of information next to the text of documents. With respect to this framework, our system is closer to content-based filtering systems, however we utilize other sources of information next to the text of documents. We use social features of the users to identify the ones who are more likely to post relevant content, however it is different from the social filtering systems where other users' feedbacks are used. We believe that this is a key OSN service that has notbeen provided so far. Indeed, OSNs provide very little support to prevent undesired 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. Providing this service is not only a matter of using previously defined web content mining techniques for a different application, rather it requires to design ad-hoc classification strategies. This is because wall messages are constituted by short text for which traditional classification methods have serious limitations since short texts do not provide sufficient word occurrences.

The main contribution of this is the design of a system providing customizable content-based message filtering for OSNs, based on ML techniques. Our work has relationships both with the state of the art in content-based filtering, as well as with the field of policy-based personalization for OSNs

and, more in general, web contents. Therefore, in what follows, we survey the literature in both these fields.

A. Content-Based Filtering

Information filtering systems are designed to classify a stream of dynamically generated information dispatched asynchronously by an information producer and present to the user those information that are likely to satisfy his/her requirements [3]. In content-based filtering each user is assumed to operate independently. As a result, 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 [4]. While electronic mail was the original domain of early work on information filtering, subsequent papers have addressed diversified domains including newswire articles, Internet"news" articles, and broader network resources [5], [6].

Documents processed in content-based filtering are mostly textual in nature and this makes content-based filtering close to text classification. The activity of filtering can be modeled, in fact, as a case of single label, binary classification, partitioning incoming documents into relevant and non relevant categories [7]. More complex filtering systems include multi-label text categorization automatically labeling messages into partial thematic categories.

In [4] a detailed comparison analysis has been conducted confirming superiority of Boosting-based classifiers [10], Neural Networks [11] and Support Vector Machines [12] over other popular methods, such as Rocchio and Naive Bayesian. However, it is worth to note that most of the work related to text filtering by ML has been applied for long-form text and the assessed performance of the text classification methods strictly depends on the nature of textual documents.

B. Policy-Based Personalization Of OSN Contents

There have been some proposals exploiting classification mechanisms for personalizing access in OSNs. For instance, in [8] a classification method has been proposed to categorize short text messages in order to avoid overwhelming users of micro blogging services by raw data. The user can then view only certain types of tweets based on his/her interests. In contrast, Golbeck and Kuter [9] propose an application, called FilmTrust, that exploits OSN trust relationships and provenance information to personalize access to the website. However, such systems do not provide a filtering policy layer by which the user can exploit the result of the classification process to decide how and to which extent filtering out unwanted information. In contrast, our filtering policy language allows the setting of FRs according to a variety of criteria, that do not consider only the results of the classification process but also the relationships of the wall owner with other OSN users as well as information on the user profile. Moreover, our system is complemented by a flexible mechanism for BL management that provides a further opportunity of customization to the filtering procedure.

The approach adopted by MyWOT is quite different. In particular, it supports filtering criteria which are far less flexible than the ones of Filtered Wall. Content filtering can be considered as an extension of access control, since it can be used both to protect objects from unauthorized subjects, and subjects from inappropriate objects. In the field of OSNs, the majority of access control models proposed so far enforce topology-based access control, according to which access control requirements are expressed in terms of relationships that the requester should have with the resource owner. We use a similar idea to identify the users to which a FR applies. However, our filtering policy language extends the languages proposed for access control policy specification in OSNs to cope with the extended requirements of the filtering domain. Indeed, since we are dealing with filtering of unwanted contents rather than with access control, one of the key ingredients of our system is the availability of a description for the message contents to be exploited by the filtering mechanism. In contrast, no one of the access control models previously cited exploit the content of the resources to enforce access control. Moreover, the notion of BLs and their management are not considered by any of the above-mentioned access control models. Finally, our policy language has some relationships with the policy frameworks that have been so far proposed to support the specification and enforcement of policies expressed in terms of constraints on the machine understandable resource descriptions provided by Semantic web languages. Examples of such frameworks are KAoS and REI, focusing mainly on access control, Protune [13], which provides support also to trust negotiation and privacy policies, and WIQA [14], which gives end users the ability of using filtering policies in order to denote given "quality" requirements that web resources must satisfy to be displayed to the users. However, although such frameworks are very powerful and general enough to be customizedand/or extended for different application scenarios they have not been specifically conceived to address information filtering in OSNs and therefore to consider the user social graph in the policy specification process.

III. BACKGROUND

All current OSNs adopt the client-server architecture. The OSN service provider acts as the controlling entity. It stores and manages all the content in the system. On the other hand, the content is generated by users spontaneously from the client side. The OSN service provider offers a rich set of well-defined interfaces through which the users can interact with others. Currently two popular ways of interaction exist. Facebook is representative of OSNs that adopt the interaction between a pair of sender and recipient as their primary way of interaction, although they also support other ways. Twitter is representative of OSNs that adopt broadcasting as their primary way of interaction. In both the service provider mediates all the interactions. The generated messages are first

stored at the service provider's side, and will be relayed when the corresponding recipient signs in. Unfortunately, all the content in OSNs is generated by users and is not necessarily legitimate. The posted messages could be spam. So it is necessary to restrict that unwanted message.

IV. GOAL

Our goal is to design an online message filtering system that is deployed at the OSN service provider side. Once deployed, it inspects every message before rendering the message to the intended recipients and makes immediate decision on whether or not the message under inspection should be dropped.

V. WORKING MODULES:

A. Filtering rules

The system provides a powerful rule layer exploiting aflexible language to specify Filtering Rules (FRs), by whichusers can state what contents should not be displayed ontheir walls.

B. Online setup assistant for FRs thresholds:

OSA presents the user with a set of messages selected from the dataset. For each message, the user tells the system the decision to accept or reject the message.

C. Blacklists

A further component of our system is a BL mechanism to avoid messages from undesired creators, independent from their contents.

VI.CONCLUSION

In this paper, we describe our work to provide unwanted message filtering for social networks. we have presented a system to filter undesired messages from OSN walls. The system exploits a ML soft classifier to enforce customizable content-dependent FRs. Moreover, the flexibility of the system in terms of filtering options is enhanced through the management of BLs. we would like to remark that the system proposed in this paper represents just the core set of functionalities needed to provide a sophisticated tool for OSN message filtering. Additionally, we studied strategies and techniques limiting the inferences that a user can do on the enforced filtering rules with the aim of bypassing the filtering system, such as for instance randomly notifying a message that should instead be blocked, or detecting modifications to profile attributes that have been made for the only purpose of defeating the filtering system.

REFERENCES

[1] Measuring semantic similarity between words using web search engines. In WWW '07: Proceedings of the 16th international conference on World Wide Web, pages 757{766, New York, NY, USA, 2007. ACM.

[2] F. Sebastiani, "Machine learning in automated text categorization," ACM Computing Surveys, vol.34, no. 1, pp. 1–47, 2002.

[3] N. J. Belkin and W. B. Croft, "Information filtering and information retrieval: Two sides of the same coin? " Communications of the ACM, vol. 35, no. 12, pp. 29–38, 1992.

[4] P. J. Denning, "Electronic junk," Communications of theACM, vol. 25, no. 3, pp. 163–165, 1982.

[5] P. S. Jacobs and L. F. Rau, "Scisor: Extracting information from online news," Communications of the ACM, vol. 33, no. 11, pp. 88–97, 1990.

[6] S. Pollock, "A rule-based message filtering system," ACM Transactions on Office Information Systems, vol. 6, no. 3, pp. 232–254,1988.

[7] P. J. Hayes, P. M. Andersen, I. B. Nirenburg, and L. M. Schmandt, "Tcs: a shell for content-based text categorization," in Proceedings of 6th IEEE Conference on Artificial Intelligence Applications (CAIA- 90). IEEE Computer Society Press, Los Alamitos, US, 1990, pp. 320–326.

[8] B. Sriram, D. Fuhry, E. Demir, H. Ferhatosmanoglu, and M. Demirbas, "Short text classification in twitter to improve information filtering," in Proceeding of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2010, 2010, pp. 841–842.

[9] J. Golbeck, "Combining provenance with trust in social networks for semantic web content filtering," in Provenance and Annotation of Data, ser. Lecture Notes in Computer Science, L. Moreau and I. Foster, Eds. Springer Berlin / Heidelberg, 2006, vol. 4145, pp. 101–108.

[10] R. E. Schapire and Y. Singer, "Boostexter: a boosting-based system for text categorization," Machine Learning, vol. 39, no. 2/3, pp. 135–168, 2000.

[11] H. Schiutze, D. A. Hull, and J. O. Pedersen, "A comparison of classifiers and document representations for the routing problem," in Proceedings of the 18th Annual ACM/SIGIR Conference on Resea. Springer Verlag, 1995, pp. 229–237.

[12] 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.

[13] P. Bonatti and D. Olmedilla, "Driving and monitoring provisional trust negotiation with metapolicies," in 6th IEEE International Workshop on Policies for Distributed Systems and Networks (POLICY 2005). IEEE Computer Society, 2005, pp. 14–23

[14] C. Bizer and R. Cyganiak,"Quality-driven information filtering using the wiqa policy framework," Web Semantics: Science, Services and Agents on the World Wide Web,vol.7, pp.1 jan 2009.