IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p- ISSN: 2278-8727Volume 16, Issue 1, Ver. I (Jan. 2014), PP 66-70
www.iosrjournals.org

Filtering Unwanted Messages from Online Social Networks (OSN) using Rule Based Technique

Sujapriya. S [1], G. Immanual Gnana Durai [2]., Dr. C.Kumar Charlie Paul [3]

Abstract: *Online Social Networks (OSNs) are today one of the most popular interactive medium to share, communicate, and distribute a significant amount of human life information. In OSNs, information filtering can also be used for a different, more responsive, function. This is owing to the fact that in OSNs there is the possibility of posting or commenting other posts on particular public/private regions, 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. OSNs provide very little support to prevent unwanted messages on user walls. For instance, Facebook permits users to state who is allowed to insert messages in their walls (i.e., friends, defined groups of friends or friends of friends). Though, no content-based partialities are preserved and therefore it is not possible to prevent undesired communications, for instance political or offensive ones, no matter of the user who posts them. To propose and experimentally evaluate an automated system, called Filtered Wall (FW), able to filter unwanted messages from OSN user walls.*
Index terms-*Information filtering, online social networks, Short text classification, policy-based personalization*

I. 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. Along with Facebook information1 average user creates 90 pieces of substance every month, while more than 30 billion quantity of substance (web links, news stories, notes, blog posts, photo albums, etc.) are distributed every month. The vast and dynamic character of this information produces the premise for the employment of web content mining strategies aimed to automatically discover useful information dormant contained by the information. They are instrumental to give a dynamic support in complex and sophisticated tasks involved in OSN administration, for example such as access power or information filtering. Information filtering has been significantly searched for what concerns textual documents and, more recently, web content. However, the aim of the majority of these proposals is mainly to provide users a classification mechanism to avoid they are overwhelmed by unsuccessful information. In OSNs, information filtering can also be exploited for a dissimilar, more responsive, purpose. This is due to the fact that in OSNs there is the possibility of posting or commenting other posts on exacting public/private regions, called in common walls. Information filtering can therefore be used to provide users the capability to automatically control the messages written on their individual walls, by filtering out surplus communication. We believe that this is a key OSN service that has not been offered so far. Certainly, in the present day OSNs provide very tiny maintain to prevent unwanted messages on user walls. For instance, Facebook permits users to status who is allowed to insert messages in their walls (i.e., friends, defined groups of friends or friends of friends). Though, no content-based preferences are maintained and therefore it is not possible to prevent undesired messages, for instance political or offensive ones, no matter of the user who posts them. Providing this service is not only a topic of using previously defined web content mining methods for a different purposes, rather it entails to propose adhoc categorization strategies. This is because wall messages are represented by tiny text for which traditional classification methods have serious limitations since short texts do not provide sufficient word occurrences.

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 [4] to automatically assign with each short text message a set of categories based on its substance. The most important efforts in building a robust small text classifier (STC) are concentrated in the extraction and selection of a set of characterizing and discriminant aspects. The resolutions examined in this paper are an extension of those adopted in a previous work by us [5] from which we inherit the learning model and the elicitation procedure for generating preclassified information.

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. As far as the learning model is concerned, we authenticate in the present paper and utilize of neural learning which is today recognized as one of the most efficient solutions in text classification [4]. In particular, we base 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, as well as makes possible the experimental estimation tasks.

Besides categorization capabilities, the system offers a powerful rule layer utilizing a flexible language to specify Filtering Rules (FRs), by which users can state what substances, should not be showed on their walls. FRs can maintain a variety of different filtering criteria that can be combined and customized according to the user requirements. In particular, FRs utilizes user profiles, user relations as well as the production of the ML categorization process to state the filtering criteria to be forced. Additionally, the system gives the support for user-defined BlackLists (BLs), that is, lists of users that are temporarily prevented to post any kind of messages on a user wall. Main dissimilar includes a different semantics for filtering rules to best fit the measured domain, an OSA to aid users in FR specification, the extension of the set features considered in the classification process, a more deep performance evaluation plan and an update of the prototype implementation to reflect the changes made to the classification methods.

RELATED WORK

The main contribution of this paper is the design of a system providing customizable content-based message filtering for OSNs, based on ML methods. Since we have pointed out in the beginning, to the top of our facts, we are the first proposing such kind of purpose for OSNs. Though, our effort has relationships equally with the state of the ability in content-based filtering, as fit as with the field of procedure-based personalization for OSNs along with, more in common, web substances. All the techniques and procedures have been referred from some survey papers in both these fields.

FILTERING BASED CONTENTS 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. Focusing on the OSN domain, interest in access control and privacy protection is relatively recent. As future as confidentiality is disturbed, current work is essentially focusing on privacy-preserving data mining methods, that is, protecting data associated to the network, i.e., relations/nodes, while performing social network study. Effort more associated to our schemes is those in the field of access control. In this field, various dissimilar access control models and associated mechanisms have been proposed so far which essentially differ on the expressivity of the access control policy language and on the way access control is enforced (e.g., centralized vs. decentralized). The majority of these models convey access control requirements in terms of relationships that the requestor should have with the resource holder. We use a related idea to classify the users to which a filtering rule applies. Though, the general purpose of our suggestion is absolutely different, while we effectively agreement with filtering of unwanted substances rather than with access control. For itself, one of the key elements of our scheme is the availability of an explanation for the message contents to be exploited by the filtering mechanism as well as by the language to express filtering rules. In distinguish no one of the access control models previously cited exploit the content of the resources to enforce access control. We consider that this is an essential difference. Furthermore, the concept of blacklists and their administration are not believed by any of these access control models. The application of content-based filtering on messages posted on OSN user walls poses additional challenges given the short length of these messages other than the wide range of topics that can be discussed. Short text categorization has acknowledged up to now few attentions in the scientific community.

Using rule base engine components, filtering concept is applied to the Online Social Network user wall. Latest effort highlights complexities in significant robust aspects, effectively due to the fact that the explanation of the short text is brief, with various misspellings, nonstandard conditions, and noise. Zelikovitz and Hirsh attempt to improve the classification of short text strings developing a semi-supervised learning strategy based on a combination of labeled training data plus a secondary corpus of unlabeled but related longer essays.

This resolution is inappropriate in our field in which short messages are not summary or part of longer semantically associated documents. A different approach is planned by Bobicev and Sokolova that circumvent the problem of error-prone feature construction by adopting a statistical learning method that can perform reasonably well without aspect production.

Though, this technique, named Prediction by Partial Mapping, generates a language model that is used in probabilistic text classifiers which are hard classifiers in nature and do not easily integrate soft, multi-membership paradigms. In our development, we think gradual membership to programs a key feature for defining flexible policy-based personalization strategies.

OSN Contents for Policy-Based Personalization

Recently, there have been some proposals exploiting classification mechanisms for personalizing access in OSNs. For instance, in [7], 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 system described, focuses on Twitter2 and associates a set of categories with each tweet relating its substance. The user can then examination only certain categories of tweets based on his/her interests. In contrast, Golbeck and Kuter suggest a purpose, called FilmTrust that develops OSN trust relationships and provenance information to personalize access to the website. Though, such systems do not offer a filtering procedure 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 distinguish; our filtering policy language permits 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. Furthermore, our system is matched by a flexible mechanism for BL administration that provides a further opportunity of customization to the filtering procedure.

Our work is also inspired by the many access control models and related policy languages and enforcement mechanisms that have been proposed so far for OSNs, since filtering shares several similarities with access control. It can researches all the individual profiles so it is based on the profiling concepts. Really, content filtering can be considered as and expansion of access control, because it can be used equally to protect objects from not permitted subjects, and subjects from improper objects. In the field of OSNs, the greater part of access control models planned so far enforce topology-based access control, along with which access control conditions are expressed in terms of relationships that the requester should have with the resource holder. We utilize a similar thought to categorize the users to which a FR applies. Though, our filtering policy language enlarges the languages planned for access control policy requirement in OSNs to cope with the extended requirements of the filtering field. Certainly, as we are dealing with filtering of unwanted substances rather than with access control, one of the key elements of our system is the accessibility of a description for the message contents to be exploited by the filtering method. In compare, no one of the access control models before cited develop the content of the resources to enforce access control. Additionally, the concept of BLs and their administration are not considered by any of the above-mentioned access control models.

To finish, our policy language has some associations with the policy structures 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.

ARCHITECTURE OF FILTERED WALL

In general, the architecture in support of OSN services is a three-tier configuration. The initial layer generally aims to offer the essential OSN functionalities (i.e., profile and relationship administration). In addition, some OSNs offer an extra layer allowing the support of external Social Network Applications (SNA)1. Finally, the supported SNA may require an additional layer for their needed graphical user interfaces (GUIs). According to this orientation layered structural plan, the planned system has to be positioned in the second and third layers (Figure 1), as it can be considered as a SNA. Particularly, users cooperate with the system by means of a GUI setting up their filtering laws, along with which messages have to be filtered out. In addition, the GUI offers users with a FW that is a wall where only messages that are authorized according to their filtering rules are published.

The core components of the proposed system are the Content-Based Messages Filtering (CBMF) and the Short Text Classifier elements. The latter element aims to categorize messages according to a set of categories. In compare, the first element exploits the message categorization offered by the STC module to implement the FRs specified by the user.

As graphically illustrated in Fig. 1, the path pursued by a message, it can be summarized as follows:

1. After entering the private wall of one of his/her associates, the user attempts to post a message, which is captured by FW.

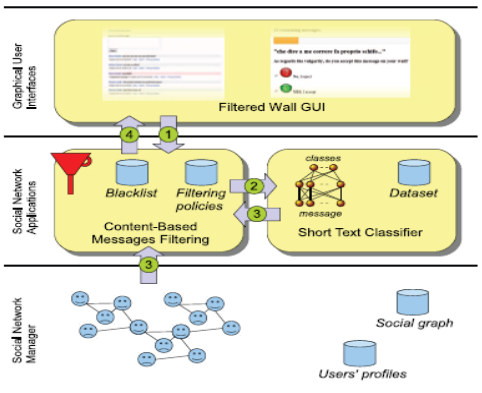

Fig.1. Architecture of Filtered wall

2. A ML-based text classifier extracts metadata from the content of the message.

3. FW uses metadata provided by the classifier, mutually with data extorted from the social graph and users' profiles, to implement the filtering and BL rules.

4. Depending on the result of the previous step, the message will be available or filtered by FW.

SHORT TEXT CLASSIFIER

Established techniques used for text classifications work well on datasets with large documents such as newswires corpora [16] but suffer when the documents in the quantity are tiny. In this perspective critical features are the description of a set of characterizing and discriminant features allowing the representation of underlying concepts and the collection of a complete and consistent set of supervised examples. Our study is aimed at designing and evaluating various representation techniques in combination with a neural learning strategy to semantically categorize short texts.

BLACKLIST AND MANAGEMENT FILTERING RULES

In this section, we introduce the rules adopted for filtering unwanted messages. In essential the language for filtering laws requirement, we consider three main concerns that, in our estimation, should influence the filtering assessment.

Filtering Rules

A filtering rule FR is a tuple (author, creatorSpec, contentSpec, action), where,

- author is the user who identifies the rule;
- creatorSpec is a creator specification,
- contentSpec is a Boolean expression defined on content constraints of the form (C, ml), where C is a class of the first or second level and ml is the minimum membership level threshold required for class C to make the constraint satisfied;
- action Є {block, notify} denotes the action to be performed by the system on the messages matching contentSpec and created by users identified by creatorSpec.

In that container, the system is not able to estimate whether the user profile matches the FR. Because how to agreement with such messages depend on the considered circumstances and on the wall owner approaches, we request the wall owner to choose whether to block or notify messages originating from a user whose profile does not match against the wall owner FRs because of missing attributes.

Blacklists

A further component of our system is a BL mechanism to avoid messages from undesired creators, autonomous from their substances. BLs is straightly supervised by the system, which should be able to establish who are the users to be introduced in the BL and decide when users retention in the BL is completed. To improve flexibility, such information is providing to the system during a set of rules, after this called BL rules. Such rules are not defined by the SNMP; thus, they are not meant as common high-level directives to be practical to the entire society. Rather, we choose to permit the users themselves, i.e., the wall's owners to indicate BL rules regulating who has to be banned from their walls and for how lengthy. Consequently, a user might be eliminated from a wall, by, at the same time, being capable to post in other walls.

A BL rule is a tuple (author, creatorSpec, creatorBehavior, T), where

- author is the OSN user who identifies the rule, i.e., the wall owner;
- creatorSpec is a creator requirement,
- CreatorBehavior consists of two components RFBlocked and minBanned. RFBlocked = (RF, mode, window) is defined such that

- RF ¼ #bMessages/#tMessages , where #tMessages is the total number of messages that each OSN user identified by creatorSpec has tried to publish in the author wall (mode ¼ myWall) or in all the OSN walls (mode ¼ SN); whereas #bMessages is the number of messages among those in #tMessages that have been blocked; window is the time period of making of those messages that have to be considered for RF computation; minBanned = (min, mode, window), where min is the minimum number of times in the time interval specified in window that OSN users identified by creatorSpec have to be inserted into the BL due to BL rules specified by author wall (mode = myWall) or all OSN users (mode = SN) in order to satisfy the constraint.

- T denotes the time phase the users recognized by creatorSpec and creatorBehavior have to be banned from author wall.

II. Conclusion

In this paper, we have presented a system to filter undesired messages from OSN walls. The system develops a ML soft classifier to implement customizable content-dependent FRs. In particular, we aim at

investigating a tool able to automatically recommend trust values for those contacts user does not individually identified. We do consider that such a tool should propose expectation assessment based on users procedures, performances, and reputation in OSN, which might involve enhancing OSN with assessment methods. Though, the propose of these assessment-based tools is difficult by several concerns, like the suggestions an assessment system might have on users' confidentiality and/or the restrictions on what it is possible to audit in present OSNs. An introduction work in this direction has been prepared in the context of expectation values used for OSN access control purposes. However, 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. Still if we have balanced our system with an online associate to set FR thresholds, the improvement of a absolute system effortlessly exploitable by average OSN users is a wide topic which is out of the scope of the present paper.

References

[1] 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.

[2] 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.

[3] R.J. Mooney and L. Roy, "Content-Based Book Recommending Using Learning for Text Categorization," Proc. Fifth ACM Conf. Digital Libraries, pp. 195-204, 2000. [4] F. Sebastiani, "Machine Learning in Automated Text Categorization," ACM Computing Surveys, vol. 34, no. 1, pp. 1-47, 2002.

[5] M. Vanetti, E. Binaghi, B. Carminati, M. Carullo, and E. Ferrari, "Content-Based Filtering in On-Line Social Networks," Proc. ECML/PKDD Workshop Privacy and Security Issues in Data Mining and Machine Learning (PSDML '10), 2010.

[6] N.J. Belkin and W.B. Croft, "Information Filtering and Information Retrieval: Two Sides of the Same Coin?" Comm. ACM, vol. 35, no. 12, pp. 29-38, 1992.

[7] P.J. Denning, "Electronic Junk," Comm. ACM, vol. 25, no. 3, pp. 163-165, 1982.

[8] P.W. Foltz and S.T. Dumais, "Personalized Information Delivery: An Analysis of Information Filtering Methods," Comm. ACM, vol. 35, no. 12, pp. 51-60, 1992.

[9] P.S. Jacobs and L.F. Rau, "Scisor: Extracting Information from On- Line News," Comm. ACM, vol. 33, no. 11, pp. 88-97, 1990.

[10] S. Pollock, "A Rule-Based Message Filtering System," ACM Trans. Office Information Systems, vol. 6, no. 3, pp. 232-254, 1988.

[11] P.E. Baclace, "Competitive Agents for Information Filtering," Comm. ACM, vol. 35, no. 12, p. 50, 1992.

[12] P.J. Hayes, P.M. Andersen, I.B. Nirenburg, and L.M. Schmandt, "Tcs: A Shell for Content-Based Text Categorization," Proc. Sixth IEEE Conf. Artificial Intelligence Applications (CAIA '90), pp. 320-326, 1990.

[13] G. Amati and F. Crestani, "Probabilistic Learning for Selective Dissemination of Information," Information Processing and Management, vol. 35, no. 5, pp. 633-654, 1999.

[14] M.J. Pazzani and D. Billsus, "Learning and Revising User Profiles: The Identification of Interesting Web Sites," Machine Learning, vol. 27, no. 3, pp. 313-331, 1997.