A System to Filter Unwanted Messages from OSN User Walls Using Inference Attacks on Social Networks

S. R. Pandit, Pooja Bhoyar, Kiran Sarode, Supriya Kadam, Sayali More
Department of Computer Engineering,
AVCOE Sangamner,Savitribai Phule Pune University,
Maharashtra.

ABSTRACT

Now a day's social Networks are more popular. Users are using multiple applications for social media. Users post their comments on their private space to avoid that unwanted content is displayed. To overcome this problem, we propose a system allowing OSN users to have a direct control on the messages posted on their walls. We explore how to prevent personal information using learning algorithm. This paper describes how to launch inference attacks using released social networking applications data to predict private information. Then we have 3 different techniques which can be used in such situations

Keywords: *Online Social Network, Information Filtering, Short Text Classification, Policy-based Personalization*

Introduction

Online Social Networks (OSN) is the most popular way to access information about any one. Communication can be easily done through OSN. OSN is also used today widely for sharing information, images, links, documents, video, audio, text messages, etc. According to Facebook statistics1 average user creates 90 pieces of content each month, whereas more than 30 billion pieces of content (web links, news stories, blog posts, notes, photo albums, etc.) are shared each month. They are instrumental to provide an active support in complex and sophisticated tasks involved in OSN management, such as for instance access control or information filtering. Classification is done to avoid overwhelmed data for user is the main concept of this paper. In OSN information filtering is also done. 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. Because these sites gather extensive personal information, social network application providers have a rare opportunity: direct use of this information could be useful to advertisers for direct marketing. Privacy policy of user can be classified into two categories: privacy after data release, private information leakage. As data get leakage the personal information of any user get release to general public. Perhaps the most illustrative example of this type of privacy breach (and the repercussions thereof) is the AOL search data scandal.

1. Related Work

Computer scientists have recently undermined our faith in the privacy-protecting power of anonymization, the name for techniques for protecting the privacy of individuals in large databases by deleting information like names and social security numbers. These scientists have demonstrated they can often "reidentify" or "deanonymize" individuals hidden in anonymized data with astonishing ease. By understanding this research, we will realize we have made a mistake, labored beneath a fundamental misunderstanding, which has assured us much less privacy than we have assumed. This mistake pervades nearly every information privacy law, regulation, and debate, yet regulators and legal scholars have paid it scant attention. We must respond to the surprising failure of anonymization, and this Article provides the tools to do so.

This paper presents an overview of the field of recommender systems and describes the current generation of recommendation methods that are usually classified into the following three main categories: content-based, collaborative, and hybrid recommendation approaches. This paper also describes various limitations of current recommendation methods and discusses possible extensions that can improve recommendation capabilities and make recommender systems applicable to an even broader range of applications. These extensions include,

among others, an improvement of understanding of users and items, incorporation of the contextual information into the recommendation process, support for multi criteria ratings, and a provision of more flexible and less intrusive types of recommendations.

As the Web continues to grow, it has become increasingly difficult to search for relevant information using traditional search engines. Topic-specific search engines provide an alternative way to support efficient information retrieval on the Web by providing more precise and customized searching in various domains. However, developers of topic-specific search engines need to address two issues: how to locate relevant documents (URLs) on the Web and how to filter out irrelevant documents from a set of documents collected from the Web. This paper reports our research in addressing the second issue. We propose a machine-learning-based approach that combines Web content analysis and Web structure analysis. We represent each Web page by a set of content-based and link-based features, which can be used as the input for various machine learning algorithms. The proposed approach was implemented using both a feed forward/back propagation neural network and a support vector machine. Two experiments were designed and conducted to compare the proposed Web-feature approach with two existing Web page filtering methods a keyword-based approach and a lexicon-based approach. The experimental results showed that the proposed approach in general performed better than the benchmark approaches, especially when the number of training documents was small. The proposed approaches can be applied in topic-specific search engine development and other Web applications such as Web content management.

The automated categorization (or classification) of texts into predefined categories has witnessed a booming interest in the last 10 years, due to the increased availability of documents in digital form and the ensuing need to organize them. In the research community the dominant approach to this problem is based on machine learning techniques: a general inductive process automatically builds a classifier by learning from a set of pre classified documents, the characteristics of the categories. The advantages of this approach over the knowledge engineering approach (consisting in the manual definition of a classifier by domain experts) are a very good effectiveness, considerable savings in terms of expert labor power, and straightforward portability to different domains. This survey discusses the main approaches to text categorization that fall within the machine learning paradigm. We will discuss in detail issues pertaining to three different problems, namely, document representation, classifier construction, and classifier evaluation.

This paper proposes a system enforcing content-based message filtering conceived as a key service for On-line Social Networks (OSNs). The system allows 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 a user to customize the filtering criteria to be applied to their walls, and a Machine Learning based soft classifier automatically producing membership labels in support of content-based filtering.

2. Analysis

Indeed, 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.

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. In this system user is consider as a unique. 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. 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.

B. Policy-based personalization of OSN contents

To personalizing access in OSN, there have been some proposals exploiting classification mechanisms. To categorized short text messages in order to avoid overwhelming users of micro blogging; a classification method has been Proposed. The system described focuses on Twitter2 and associates a set of categories with each tweet describing its content. The user view information on their own interest. In opposite to this, Golbeck and Kuter proposed an application in which user can exploits OSN trust relationship and provenance information to personalize access to 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.

C. Detail Removal

Our method is successful in removing the details. This is the most highly connected with a class to remove details of links. The accuracy of links classifier is also decreased as we remove details. The details of two nodes are compared to find a similarity. As we remove details from the network, the set of "similar" nodes to any given node will also change. This can account for the decrease in accuracy of the links classifier.

D. Link Removal

We have a generally more stable downward trend, with only a few exceptions in the "political affiliation" experiments.

E. Combined Removal

While each measure provides a decrease in classification accuracy, we also test what happens in our data set if we remove both details and links. To do this, we conduct further ex
periments where we test classification accuracy after removing 0 details and 0 links (the baseline accuracy), 0 details and 10 links, 10 details and 0 links, and 10 details and 10 links. We choose these numbers because after removing 12 links, we found that we were beginning to create a number of isolated groups of few nodes or single, disconnected nodes.

3. Proposed System

There are three phases in this proposed system:

1) Social Network Manager:

Social Graph: The social graph work on the OSN functionalities. It also maintain the relationships.

User Profile: User profile is actually the account of user on social site. It contain the details of user.

2) Social Network Application:

Content Based Message Filtering:

There are two functions in content based message filtering:

Blacklist:

Blacklist rule make the wall owner able to be blocked according to their profile as well as their relationship in OSN. In addition, the systemprovides 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.

Filtering Policies:

Filtering policies define the rules which are used for avoid/filter the text. Using filtering rules systemcan publish authorized messages on OSN wall.

$$tf - idf(tk,dj) = \#(tk,dj) \cdot \log (|Tr| / \#Tr(tk))$$

where, T = Set of term

#(tk,dj)=no. of times tk occur in dj

#Tr(tk)=document frequency of term tk

Short Text Classifier:

Short text classifier includes classes, massages, dataset. Dataset contains the database which is given to the systemand with the help of database the systemperform the classification.

3) Graphical User Interface:

It provide filter wall to the user. On the filter wall the massages which are authorize are published.

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 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 (STC) are concentrated in the extraction and selection of a set of characterizing and discriminate features. The solutions investigated in this paper are an extension of those adopted in a previous work by us from which we inherit the learning model and the elicitation procedure for generating pre classified data. The original set of features, derived from endogenous properties of short texts, is enlarged here including exogenous knowledge related to the context from which the messages originate. As far as the learning model is concerned, we confirm in the current paper the use of neural learning which is today recognized as one of the most efficient solutions in text classification. 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 managing noisy data and intrinsically vague classes. Moreover, the speed in performing the learning phase

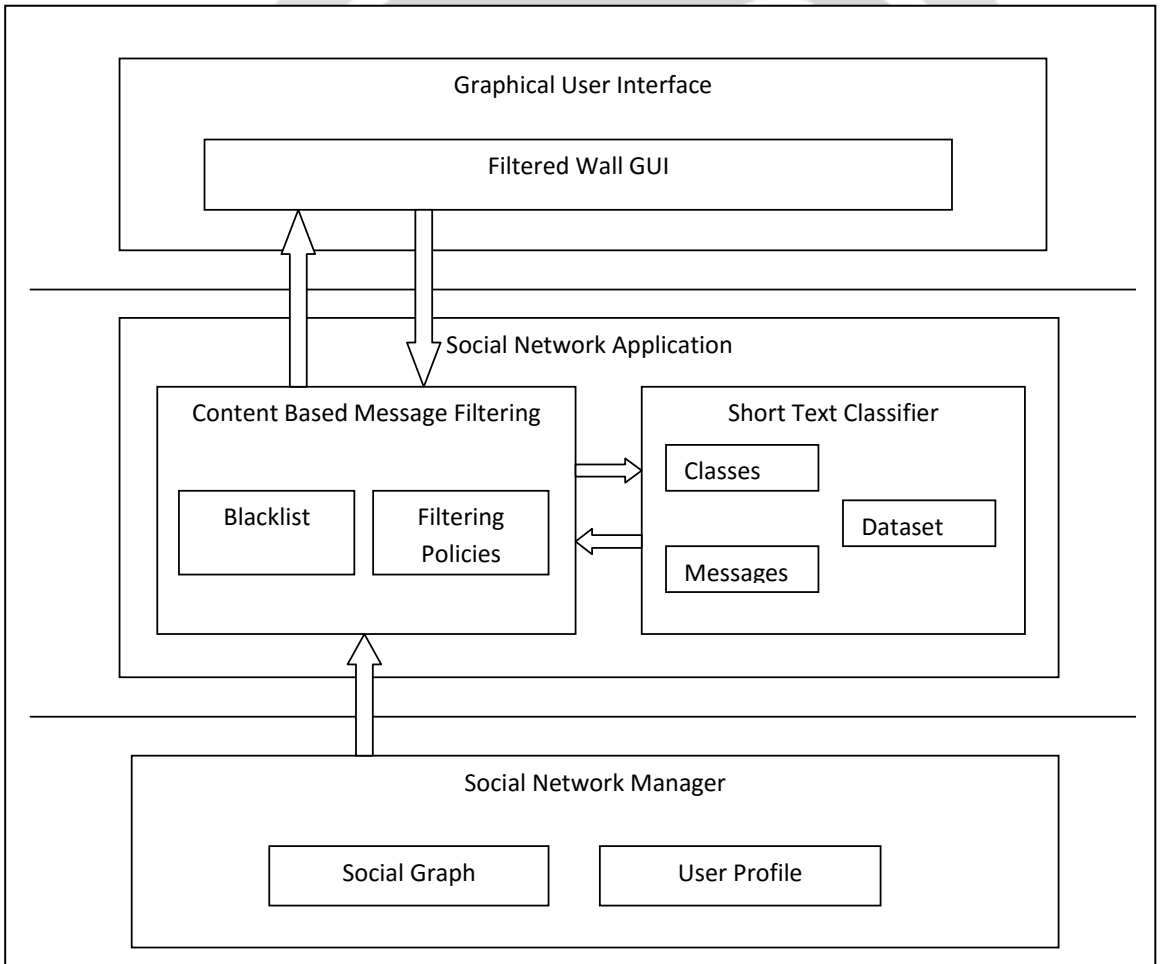

Fig: Filtering Wall Architecture And Message Flow

creates the premise for an adequate use in OSN domains, as well as facilitates the experimental evaluation tasks. We insert 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. Besides classification facilities, 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.

4] Conclusion

In this paper, we have presented a system to filter unwanted messages from OSN walls. In this system uses the ML soft classifier to remove the unwanted messages. BL is used to enhanced the flexibility of system for filtering. We will be design the system which will more sophisticated approach to decide when a user should be inserted into the BL. As future work, we intend to exploit similar techniques to infer BL rules and FRs.

5] References

[1] M. Vanetti, E. Binaghi, B. Carminati, M. Carullo, and E. Ferrari, "Content-based filtering in on-line social networks," in Proceedings of ECML/PKDD Workshop on Privacy and Security issues in Data Mining and Machine Learning (PSDML 2010), 2010.

[2] K.M. Heussner, "'Gaydar' n Facebook: Can Your Friends Reveal Sexual Orientation?" ABC News, http://abcnews.go.
com/Technology/gaydar-facebook-friends/story?id=8633224#. UZ939UqheOs, Sept. 2009.

[3] 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. [4] T. Zeller, "AOL Executive Quits After Posting of Search Data," The New York Times, no. 22, http://www.nytimes.com/2006/08/22/technology/22iht- l.2558731.html?pagewanted=all&_r=0, Aug.
2006.

[5] A. Adomavicius, G.and Tuzhilin, "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.

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

[7] R. J. Mooney and L. Roy, "Content-based book recommending using learning for text categorization," in Proceedings of the Fifth ACM Conference on Digital Libraries. New York: ACM Press, 2000, pp. 195–204.

[8] 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.

[9] P. W. Foltz and S. T. Dumais, "Personalized information delivery: An analysis of information filtering methods," Communications of the ACM, vol. 35, no. 12, pp. 51–60, 1992.

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

[11] Facebook Beacon, 2007.