CONTENT BASED MESSAGE FILTERING IN ONLINE SOCIAL NETWORKS USING ONTOLOGY

A.R. Josena
Assistant Professor, Easwari Engineering College, Chennai
arjosena@gmail.com

Abstract

The fundamental issue in today's Online Social Networks (OSNs) is to give users the ability to control the messages posted on their own private space to avoid that unwanted content is displayed. Up to now, OSNs provide little support to this requirement. To fill the gap, this paper propose a system allowing OSN users to have a direct control on the messages posted on their walls. Traditional content-based filtering methods usually utilize text extraction and classification techniques for building user profiles as well as for representations of contents, i.e. item profiles. These methods have some disadvantages e.g. mismatch between user profile terms and item profile terms, leading to low performance. The disadvantages are overcome by incorporating an ontology which enables the users to filter the content based on a set of pre-defined categories. The aim of this paper is to create an automated system, called Filtered Wall (FW), able to filter unwanted messages from OSN user walls using ontology.

Index Terms: Content Based Filtering, Online Social Network, Ontology, Text Classification.

I. INTRODUCTION

In content-based filtering, the representations of the content of items (e.g. documents, News), i.e. the items' profiles, are compared with the representation of the users, i.e. the users' profiles. In order to enable matching and measuring the similarity between the profiles, it is assumed that a user's profile and an item's profile share a common method of representation (e.g., by keywords). The output of the matching process can be expressed as a ranking score, indicating the similarity between the user's profile and each item.

A user profile can be generated in various ways, including explicit definition by the user, or implicit analysis of the user's behavior (e.g. by logging and analyzing what the user read). An item's profile too can be generated in various ways, e.g. explicitly, by asking the originator (author) to specify proper index terms, or automatically, using a text classification algorithm which extracts terms representing the item's content in the best way. At any rate, no matter which method is used for creating either type of profile, content-based filtering has drawbacks due to well-known problems of term ambiguity. For example, different terms may be used to represent the same content or the same user (synonymy); or, the same term may be used to represent different contents or different users (homonymy).

A possible way to overcome such problems of ambiguity might be through the use of ontology, i.e., a controlled vocabulary of terms or concepts, and semantic relationships among them. Ontology can bridge the gap between the terms in the users' profile and the terms representing the items. Ontology can be organized in various ways. For example, a taxonomy is a hierarchical structure with is-a relationships; in a thesaurus there are a few more types of relationships, e.g. BT/NT (broader-term; narrower terms) and general relatedness. Note that a thesaurus is a graph, not

a hierarchy, because a term may have many NTs and more than one BT. In the Social Networks domain, which is exemplified in this study, there is an ontology specifically generated for classification of messages.

Assuming that there exists ontology of a specific domain, which is used for representing users (user profiles), and contents of items (item profiles), the research question we deal with is how exactly to match and measure the similarity between a user's profile and an items' profile. Obviously, if a user's post includes exactly the same concept (terms) as an item's profile, there is some similarity between them; but the two profiles may include different concepts and still be similar to a certain degree – depending on if and how "close" the concepts are in the two profiles with respect to the common ontology. This research is conducted within the framework of a social networking website, which is aimed to provide a filtered wall able to filter out the unwanted messages from the user's wall. In this domain, instant filtering of messages is important.

The content-based filtering algorithm can perform the necessary matching with the users' profiles and determine the degree of relevancy of each item to the potential users.

The remaining of this paper is structured as follows: The next section provides a background on content-based filtering and on ontological modeling, and reviews related research on conceptual and ontological modeling employed in content-based filtering. Section 3, the main section of the paper, presents the proposed method for the ontology- content-based filtering, along with an example. Section 4 describes the evaluations conducted with the proposed method, and the last section summarizes and proposes further research and extensions to the proposed method.

II. RELATED WORK

In content-based systems, filtering is done by exploiting the information extracted from the text of documents. It has been investigated by exploiting ML techniques [2, 9, 10] as well as other strategies [8, 5]. However, the problem of applying content-based filtering on the contents exchanged by users of social networks has received up to now few attentions in the scientific community. A best example which focuses on this work is by Boykin and Roy chowdhury [3].The paper proposes an automated anti-spam tool that, exploiting the properties of social networks, can recognize unwanted commercial e-mail, spam and messages associated with people the user knows. However, it is important to note that the strategy just mentioned does not exploit content-based techniques.

The advantages of using Ontology based Content filtering strategies over other engineering approaches are a very good effectiveness, flexibility to changes in the domain and portability in differ applications. However difficulties arise in finding an appropriate set of features by which to represent short, grammatically ill formed sentences and in providing a consistent training set of manually classified text. Focusing on the OSN domain, interest in access control and privacy protection is quite recent. As far as privacy is concerned, current work is mainly focusing on privacy-preserving data mining techniques, that is, protecting information related to the network, i.e., relationships/nodes, while performing social network analysis? Work more related to our proposals is those in the field of access control. In this field, many different access control models and related mechanisms have been proposed so far (e.g., [4, 1, 6]), which mainly differ on the expressivity of the access control policy language and on the way access control is implemented. Most of these models express access control requirements in terms of relationships that the requestor should have with the resource owner. We use a similar idea to identify the users to which a filtering rule applies. However, the main goal of this paper is completely different, since this paper mainly focus with filtering of unwanted contents rather than with access control. The disadvantages with the other system are overcome by incorporating an ontology which enables the users to filter the content based on a set of pre-defined categories.

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 classification has received up to now little attention in the scientific community. Recent work highlights difficulties in defining robust features, essentially due to the fact that the

description of the short text is concise, with many misspellings, nonstandard terms, and noise. Zelikovitz and Hirsh[9] 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 documents.

This solution is inapplicable in the proposed domain in which short messages are not summary or part of longer semantically related documents. A different approach is proposed by Bobicev and Sokolova[5] that circumvent the problem of error-prone feature construction by adopting a statistical learning method that can perform reasonably well without feature engineering. However, this method, named Prediction by Partial Mapping, produces 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

The aim of this paper is to create an automated system, called Filtered Wall (FW), able to filter unwanted messages from OSN user walls using ontology. The disadvantages with the other system are overcome by incorporating an ontology which enables the users to filter the content based on a set of pre-defined categories. The aim of this paper is to create an automated system, called Filtered Wall (FW), able to filter unwanted messages from OSN user walls using ontology.

III. 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 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. 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. 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. More complex filtering systems include multi label text categorization automatically labeling messages into partial thematic categories.

Content-based filtering is mainly based on the use of the ML paradigm according to which a classifier is automatically induced by learning from a set of pre-classified examples. A remarkable variety of related work has recently appeared which differ for the adopted feature extraction methods, model learning, and collection of samples. The feature extraction procedure maps text into a compact representation of its content and is uniformly applied to training and generalization phases. Several experiments prove that Bag-of-Words (BoW) approaches yield good performance and prevail in general over more sophisticated text representation that may have superior semantics but lower statistical quality. As far as the learning model is concerned, there are a number of major approaches in content-based filtering and text classification in general showing mutual advantages and disadvantages in function of application dependent issues.

. In this paper, blacklist content based filtering algorithm is proposed to filter unwanted user messages from the OSN user walls.

IV. FILTERED WALL ARCHITECTURE

The architecture in support of OSN services is a three-tier structure (Fig. 1). The first layer, called Applications Graphical User Interface, commonly aims to provide the basic OSN functionalities (i.e., profile and relationship management), whereas the second layer provides the support for external Social Network Applications (SNAs). The supported SNAs may in turn require an additional layer for their needed Graphical User Interfaces (GUIs). The third layer is the Social Network Manager (SNA) which contains the Filtered Wall GUI.

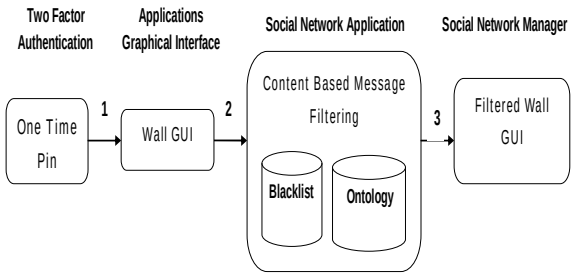

Fig.1 Filtered Wall Architecture

4.1 **Two-factor Authentication**

Social networks are unequivocally the most used application for communication and information sharing in the 21st century. As growth of this technology increases, there is a need to implement a more secure authentication mechanism to protect users as well as the platform providers from various social engineering attacks. This end-user security lapse often paves way for phishing and malware attacks and undermines the overall integrity of the system. In this study, we review the rise of social networks and the underlying concepts of two factor authentication. Furthermore, we propose a novel, feasible, cost effective and secure technique of applying an email based password tokenization as a second factor authentication in social networking sites.

Two factor authentication is an extra layer of authentication added to the conventional single factor authentication to an account login, which requires users to have additional information before access to a system is granted. The traditional method of authentication requires them user to enter only a username and password before being granted access to a closed software or application, whereas two factor authentication requires the user to have additional information known only to the user before access to the system is granted. Knowledge based i.e. a piece of information the user knows such as a Personal Identification Number (PIN) or password.

4.2 Blacklist

Blacklist rule is used to avoid unwanted messages created by the users. This is the mechanism which is managed by the system, such that the system may able to decide what type of messages can be present in the system. That is the wall owner must decide who are the users enter into his/her private wall and the kind of words that they can post in our user walls.

This algorithm is able to filter out the messages based on its content. The content is a short text where in the filtering is done based on the blacklist word count. A set of words which are considered as blacklist words, are stored. In the short text each word is checked for, if it appears in the blacklist and the count is incremented. If the word count exceeds the minimum threshold then the text is blocked, else it passes on for further filtering.

4.3 Ontological Modeling

Ontology is a specification of a conceptualization. It can be described by defining a set of representational concepts. These definitions are used to associate the names of entities in the universe (e.g., classes, relations, functions or other objects) with human-readable text, describing what the names mean, and formal axioms that constrain the interpretation and focus the well-formed use of these concepts. When constructing ontology, not only concepts and relationships are defined, but also the context in which the concept (relationship) applies. Therefore, ontology defines a set of representational terms which are called concepts, and the inter relationships among the concepts.

Linguistic ontology's (e.g., Word Net) and thesauri express various relationships between concepts (e.g. Synonyms, antonyms, is-a, contains-a), and have a hierarchical structure based on the relations between concepts. But they do not explicitly and formally describe what a concept means. Word Net, for example, is an electronic lexical database that contains nouns, verbs, adjectives and adverbs which are organized into synonym sets (synsets), each representing one underlying lexical concept. It is offering two distinct services: a vocabulary which describes the various word senses, and an ontology which describes the semantic relationships among senses.

This ontological modeling will help to increase the performance in the filtering process.

V. EXPERIMENTAL RESULTS

A prototype social network application emulates a personal wall where the user can apply a simple combination of the proposed FRs. Throughout the development of the prototype; this paper focused attention only on the FRs.

However, the implemented functionality is critical, since it permits the Ontology components to interact. Since this application is

conceived as a wall and not as a group, the contextual information (from which CF are extracted) linked to the name of the group are not directly accessible. Contextual information that is currently used in the prototype is relative to the group name where the user that writes the message is most active.

It is important to stress that this type of contextual information is related to the environment preferred by the user who wants to post the message; thus, the experience that you can try using the prototype is consistent with what described and evaluated..

To summarize, the application permits to

1. view the list of users' FWs;
2. view messages and post a new one on a FW;
3. filter using ontology

When a user tries to post a message on a wall, he/she receives an alerting message (see Fig. 2) if it is blocked by FW.

Fig 2. Prototype

VI. CONCLUSION

In this paper, the proposed system can able to create a filtered wall overcoming the drawbacks of content based filtering. Traditional content-based filtering methods usually utilize text extraction and classification techniques for building user profiles as well as for representations of contents, i.e. item profiles. These methods have some disadvantages e.g. mismatch between user profile terms and item profile terms, leading to low performance. The disadvantages are overcome by incorporating a common ontology which enables representing both the users and the items profiles with concepts taken from the same vocabulary. The method can be enhanced in various aspects. One possible enhancement is to assign more importance to concepts co-occurring in items read in the past by the user.

REFERENCES

[1] Marco Vanetti, Elisabetta Binaghi, Elena Ferrari, Barbara Carminati, and Moreno Carullo , A System to Filter Unwanted Messages from OSN User Walls"- IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 25, NO. 2, FEBRUARY 2013.

[2] 2. "Spam Filtering in Twitter using Sender-Receiver Relationship"- Jonghyk song, Sangho Lee and John Kim, Proceedings of the 14th international conference on Recent Advances in Intrusion Detection, 2011

[3] Boykin, P.O., Roychowdhury, V.P.: Leveraging social networks to fight spam. IEEE Computer Magazine 38, 61–67 (2005)

[4] "An Otology – Content- based filtering method" Peretz Shoval,Veronica Maidel,Bracha Shapira,International Journal "Information Theories Applications" Vol.15/ 2008

[5] Bobicev and Sokolova "Recognition of Sentiment Sequences in Online Discussions"

[6] Towards Online Spam Filtering in Social Networks Hongyu Gao, Yan Chen, Kathy Lee, Diana Palsetia, Alok Choudhary, NDSS Symposium, 2012

[7] Carminati, B., Ferrari, E.: Access control and privacy in web-based social networks. International Journal of Web Information Systems 4, 395–415 (2008)

[8] Carminati, B., Ferrari, E., Perego, A.: Enforcing access control in web-based social networks.ACM Trans. Inf. Syst. Secur. 13(1), 1–38 (2009)

[9] Sarah Zelikovitz, Haym Hirsh "Integrating Background Knowledge Into Text Classification"

[10] 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, 184–188 (2005)

[11]Fang, L., LeFevre, K.: Privacy wizards for social networking sites. In: WWW '10: Proceedings of the 19th international conference onWorld wide web. pp. 351–360. ACM, New York, NY, USA (2010)

[12]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)