

A Review on Filter Undesired Text from Social Networks

Ujwala S.Tambe
University of Pune
Department of Computer Engg.
GES's R. H. Sapat College of Engg

Archana S. Vaidya
University of Pune
Department of Computer Engg.
GES's R. H. Sapat College of Engg

ABSTRACT

Online social Network (OSN) is a social networking service which is a platform to build relations among people who share comfort, actions, backgrounds or real-life connections. Hundreds of thousands of people are using these social networking services for personal use, marketing, Entertainment, Business purpose. User security is the main issue in present days from person to person interaction. Message filtering is main task of proposed system. An online social network provides the little support to the user to avoid unwanted messages displayed on their own private space. In this paper, we propose system which gives ability to user to control the unwanted messages posted on their wall. To filter undesired messages propose three tier architecture containing message classifier based on content and using machine learning techniques. User is able to customize the filtering rule as per his/her preferences.i.e grant access to allow user to insert messages on his/her wall.

Keywords

Online Social Networks, Message Filtering, Machine Learning Techniques.

1. INTRODUCTION

Social networking sites are today's one of the most popular interactive medium to communicate, split and disseminate a considerable amount of human life information. Daily communications is done by the exchange of several types of content, including textual data, image, audio and video data. [1]As per Facebook statistics, average user creates 90 pieces of content each month, whereas more than 30 billion pieces of content are shared each month.

In online social network (OSN) like facebook, twitter, etc., there is possibility of posting any kind of data on user wall. Such data may contain undesired post for example: political statement, vulgar data, personal teasing statement etc. User can view such data and comment on such post. Such post may affect user social image. So security of such personal wall is an important issue. Up to certain point some existing systems like 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). But no content-based preferences are supported and therefore it is not possible to prevent posting of such undesired messages. It may overwhelm useless data present in post of user wall.

The objective of present work is to protect user wall by filtering undesired messages and to protect user social image in networking site .Introduce three tier architecture containing message classifiers based on content and using machine learning technique. User must be able to customize the filtering rule as per his/her preferences and also set the filter on different user i.e. grant privileges to allow user to insert messages in his/her wall. Therefore we develop a system to provide Filtered Wall (FW) mechanism that will be able to filter unwanted messages from wall present in online social network. We used machine learning text categorization

techniques [2] to automatically assign with each short text message a set of categories based on its content.

2. LITERATURE SURVEY

The main contribution of this paper is the design of a system providing customizable content-based message filtering for OSNs, based on ML techniques. As we have pointed out in the introduction, to the best of our knowledge, we are the first proposing such kind of application for OSNs. However, 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 [1].

2.1 Content-based Filtering

Content Based filtering system approve a document by matching the document profile with the user profile [3], using traditional information retrieval techniques such Term Frequency and Inverse Document frequency (TF-IDF).[3] User characteristics are gathered over time and profiled automatically based upon a user's prior feedback and choices. The system needs item to item correlation in recommending the document to the user. The system initialize with the process of gathering the content details about the item, like treatments, indications etc. for disease related item .Then, the system test the user to rate the items. In final step,, system matches unrated item with the user profile item and assign score to the unrated item and user is presented with items ranked according to the scores assigned.

2.2 Collaborative Filtering

Collaborative filtering systems filters information based on the interests of the user (past history), and the ratings of other users with similar interests. It is widely used in many filtering systems or recommender systems, especially in ecommerce applications. One of the examples of such system are Amazon.com and e-Bay, where a user's past shopping history is used to make recommendations for new products[3].

2.3 Policy-based Personalization of OSN Contents

A Recently, there have been some proposals operator classification mechanisms to customize access in OSN. For example, in [4] a classification method was proposed to classify short text messages to avoid overwhelming users of micro blogging services in the raw data. The system described in [4] and Twitter2 focuses on a set of categories associated with each describing updates its contents. The user can then see that certain types of tweets on the basis of his / her interests. However, Kuter and Golbeck [5] propose an application called Film Trust, which operates OSN relationships of trust and provenance information to customize access to the site. However, these systems do not provide a layer of filtering policy by which the user can exploit the results of the classification process and decide how to filter unwanted information. Another social networking service, we are aware provide filtering capabilities for its

users is MyWOT. This service that allows its subscribers the ability to: 1) the rate of resources in relation to four criteria: reliability, vendor reliability, privacy and child safety; 2) set preferences that determine whether the browser should block access to a particular resource, or should send a warning message based on the specified value.

Table 1. Previous Related Work

Related Work	Objective
P. J. Hayes, P. M. Andersen, I. B. Nirenburg, and L. M. Schmandt [6].	In this paper new idea is discussed regarding text categorization. It explains that a text-categorization application developed with TCS consists of the TCS run-time system and a rule base. The rule base defines what categories the application can assign to texts and contains rules that make the categorization decisions for particular texts.
N. J. Belkin and W. B. Croft[7].	In this paper technique to filter information is applied to unstructured or semi-structured data, as opposed to database applications, which use very structured data. such as email. where format for sending email is semi structured but body of email may content unstructured data.
P. W. Foltz and S. T. Dumais[8]	In this paper technique to analyze the structured data is discussed. This technique analyzes the user's interest by analyzing the user feedback and keywords provided by the user. Keyword matching and Latent Semantic Indexing technique are used in it. In this user with similar preference are considered.
Zelikovitz and Hirsh [9]	In this paper, Author 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.
Bobicev and Sokolova [10]	This paper focus on 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..

J. Golbeck and kuter[5]	This paper proposes 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.
-------------------------	---

3. MACHINE LEARNING TECHNIQUES

Machine learning approach is used for classify the set of training data and automatically create the classifiers for the training data. Text classification is the task of automatically assign the texts into the predefined categories. In these text classification accomplished on the basis of endogenous collection of data. Text categorization mostly depends on the information retrieval technique such as indexing, inductive construction of classifiers and evaluation technique. In this machine learning, classifier learns how to classify the categories of documents based on the features extracted from the set of training data.

Following methods are applied to the text classification and predicting the accuracy of the classifiers.

3.1 Bayesian Network Model

Bayesian network principles can be used in the collaborative filtering system. [3]The rating data can be used to learn a Bayesian network, [11] where each item is represented by a node, and directed arrows between items signify user's interest on items influence interest of other items. The Bayesian network is used to create probabilistic decision trees for each item, where leaf nodes are likelihoods of the target user's interest on the target item, and intermediary decisions are based on the target user's view on the parent items of the target item from the network. Bayesian networks is useful when the user preferences changes slowly with respect to the time needed to build the model, but are not suitable for environments in which user preference models needs to be updated frequently[3].

3.2 Artificial Neural Network Models

Artificial neural networks models in the information filtering systems [12],[13], are used to automatically create the terms in user profiles by the training of neural network models through examples. The ability of the ANN to model non-linear relationships can be applied to the matching of the documents to the user profile. Neural networks can be used to represents the user's preferences where words from documents are represented as nodes and strength of association between words in the same document [3]. Supervised learning and unsupervised learning are types of neural networks.

3.3 Support Vector Machines

The support vector machine classifiers analyze data and recognize pattern in it [14]. They are based on supervised learning model and are able to perform nonlinear classification in addition to linear classification. The support

vector machine classifier is suitable for large amount of unlabeled data and small amount of labeled data. The high dimensional input space, irrelevant features, sparse document vectors and linearly separable text classification makes support vector machine classifier suitable for text categorization.

4. PROPOSED SYSTEM

In this section, we introduce three tier architecture for filtered unwanted messages post on user wall.

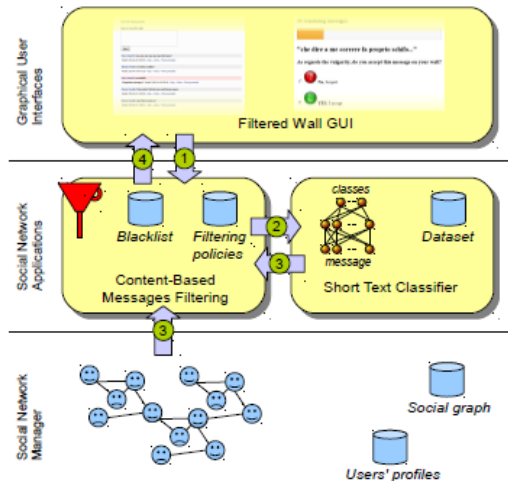


Fig1.Filtered Wall Architecture

The architecture Of OSN services is a three-tier structure (Fig1). Social Network Manager (SNM) (First Layer) is aims to provide the basic OSN functionalities. Social Network Applications (SNAs) (Second Layer) provides the support for external activities. It may require an additional layer for their needed Graphical User Interfaces (GUIs).Social Network Applications layer and Graphical User Interfaces layer plays important role for proposed system. In this system, users interact with the help of GUI to set up and manage their FRs (Filtering Rules)/BLs (Blacklists). The GUI provides users with a filtered wall, where only authorized messages are published according to their FRs.The core components of the proposed system are the Content-Based Messages Filtering (CBMF) and the Short Text Classifier (STC) modules. The later component aims to classify messages according to a set of categories [1].

Following are some basic step from writing to publications of message on user’s wall

- [A] The user when trying to post a message on the wall details, then it is first introduced into the wall filtered.
- [B] A machine learning text-based classifier to retrieve the metadata of the message content .
- [C] After this process filtered wall returns the metadata provided by the classifier. With metadata and user profile Black List rules and filtering criteria is applied.
- [D] According to this forum post the output will be filtered or filtered published by wall.

5. FILTERING RULES AND BLACKLIST MANAGEMENT

In this section, we introduce the rule layer adopted for filtering unwanted messages. We start by describing FRs, and then we illustrate the use of BLs.

5.1 Filtering Rules

In filtering rule, proposed system build model a social network as a directed graph, where each node represent a user and edges denote relationships between two different users. In graph, edge is labelled by the type of the relationship established (e.g., friend of, colleague of, parent of) and, the corresponding trust level, which represents how much a given user trustworthy with respect to that specific kind of relationship the user with whom he/she is establishing the relationship.

5.2 Blacklists

Blacklists users are those users whose messages are banned from their contents. BL rules allow the wall owner to decide users to be blocked on the basis of their profiles and relationship with wall owner. This prevention can be done for a specified period or forever according wall owner’s desire. BL is dependent on author, creator specification and creator behaviour.

6. CONCLUSION

In this article, we explain techniques to filter undesired messages in online social network. Additionally, we studied different machine learning techniques for classification purpose. In this article, proposed system represents the core set of functionalities needed to provide a proper tool for OSN message filtering. This system approach when the user decides to be inserted into a blacklist. The system developed GUI and a set of tools that make BLS easier and simpler specifications. Moreover, the flexibility of the system in terms of filtering options is enhanced by the direction of BLS. In future work, we can filter unwanted images from user wall.

7. REFERENCES

- [1] Vanetti, Elisabetta Binaghi, Elena Ferrari, Barbara Carminati, Moreno Carullo Department of Computer and Communication, University of Insubria "a system to filter out unwanted messages walls OSN user" IEEE Transactions on Knowledge And Engineering Flight Data: 25 Year 2013.
- [2] F. Sebastiani, "Machine learning in automated text categorization," ACM Computing Surveys, vol. 34, no. 1, pp. 1–47, 2002. J. Leskovec, D. P. Huttenlocher, and J. M. Kleinberg, "Predicting positive and negative links in online social networks," in Proc. 19th Int. Conf. World Wide Web, 2010, pp. 641–650.
- [3] Vinaitheerthan Renganathan1*, Ajit N Babu2 and Suptendra Nath Sarbadhikari3 " A Tutorial on Information Filtering Concepts and Methods for Bio-medical Searching".
- [4] 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.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] 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.
- [9] S. Zelikovitz and H. Hirsh, “Improving short text classification using unlabeled background knowledge,” in Proceedings of 17th International Conference on Machine Learning (ICML-00), P. Langley, Ed. Stanford, US: Morgan Kaufmann Publishers, San Francisco, US, 2000, pp. 1183–1190.
- [10] V. Bobicev and M. Sokolova, “An effective and robust method for short text classification,” in AAAI, D. Fox and C. P. Gomes, Eds. AAAI Press, 2008, pp. 1444–1445.
- [11] Pennock DM, Horvitz E, Lawrence S, Giles CL (2000) Collaborative filtering by personality diagnosis: A hybrid memory-and model-based approach. Proceedings of the Sixteenth conference on Uncertainty in artificial intelligence (UAI-2000), Morgan Kaufmann Publishers Inc 473-480.
- [12] Boger Z, Kuflik T, Shapira B, Shoval P (2000) Information filtering and automatic keyword identification by artificial neural networks. Proceedings of the 8th European Conference on Information Systems.
- [13] Jennings A, Higuchi H (1993) A user model neural network for a personal news service. *User Modeling and User-Adapted Interaction* 3: 1-25.
- [14] S. Venkata Lakshmi, K. Hema “Filtering Information for Short Text Using OSN” *International Journal of Advanced Research in Computer Science & Technology (IJARCST 2014)*317 Vol. 2, Issue 2, Ver. 2