A Survey on OSN Message Filtering

Kalpesh Gandhi
D.Y. Patil Institute of Engineering
and Technology,
Ambi, Pune

Rahul Panditrao
D.Y. Patil Institute of Engineering
and Technology,
Ambi, Pune

Vibha B. Lahane
Asst. Professor
D.Y. Patil Institute of Engineering
and Technology, Ambi, Pune

ABSTRACT

In the specified documents data mining technique has been deprecated for filtering the OSN wall with unwanted messages or any type of vulgar messages. OSN is Online Social Network which has become an important part of the people life these days. People communicate over it with friends, relatives over a OSN wall. Thus to provide a feel of security to users personal stuff it is important to filter the OSN wall for any unwanted message .But the questions Arises, how to filter the OSN wall with a facility provided of blocking unwanted messages on the user's private wall. This can be gained through the flexible rule-based system which implements filtering criteria that can be customized by the user and a Machine Learning-based soft classifier which automatically labels messages in the support of content-based filtering .This paper consist of a literature survey paper of the existing system with proposed system as a technique to filter similar meaning words using Ontology along with the basic functionality to filter the OSN wall for unwanted message. In this paper a technique to build a social network with filtered message is elaborated.

General Terms

On-line Social Network, short text classifier.

Keywords

Online-Social Networks, Content-based filtering, Machine Learning, Filtering Rules, Data Mining, Text mining.

1. INTRODUCTION

From past few years, Online Social Networks (OSNs) has become an important Asset of daily life. In Online social networks particular user can upload and share information related to their personal lives with the family, relatives and friends. As the personal information is shared online it is important to handle the risk of privacy being tampered .The most important issues in using social networks are security of the personal information of the user and the user's secretive space where anyone can post the message. Thus it is important to provide user with the mechanism which can filter the unwanted content from his secretive space. Present OSN gives a very limited support for such type of filtration .Thus for that purpose new system is proposed that can filter the users wall automatically and give user a direct control on the message that are being posted on his/her wall. OSN produces a huge data daily and classifying a useful data is called as Information Filtering e.g. [1-3] .This proposed however just gives a support to user to classify the useful data from the raw data.

A broader perspective of information filtering in OSN can be different, more specific as in the OSN any one is allowed to post, comment and share on the public/private area called general Wall. Thus in the proposed system, information filtering can provide the system that gives the user an ability to filter the unwanted messages even if the word used in the message is not under vulgar category but has a same meaning as of vulgar meaning from his/her wall. The aim is to propose and experimentally evaluate an automated system, called Filtered Wall (FW), which is able to filter unwanted messages from OSN user walls. The exploitation of Machine Learning (ML) text categorization techniques [4] is done to automatically assign each short text message to a set of categories based on its content.

2. LITERATURE SURVEY

There are two types in which text filtering process can be carried out: content –based and social filtering based. Content-based systems the process of filtering is carried out by exploiting the information which is extracted from the text of documents while in social filtering systems the annotations made by the prior reader's documents are filtered .OSN system can be customized as content-based filtering systems. It is been noticed that the present OSN does not have any content based policy provided so far. OSNs provide a mere support to prevent any type of undesired message on user's wall. For example, in Facebook users are only featured to restrict their wall post by stating who is allowed to post messages (i.e friends or their friends, or mutual friends). There are no content-based preferences thus it is hard to stop undesired messages, such as political or vulgar ones. Thus providing these services is not that using old classification mining techniques as they can only filter text but wall messages consist of small text for old classification technique stands with limitation .Thus the ML techniques stand as a main contributor in designing the OSN message filtering system. OSN filtering work has been thus residing in both the content based as well as the policy based personalization and more over in web contents too. Therefore the literature in both these fields has been followed.

2.1 Content-based Filtering

W. B. Croft and N. J. Belkin in 2002 acknowledged that Information filtering systems which was designed to classify a stream of dynamically generated information dispatched asynchronously by an information producer and present it to the user those information that are likely to satisfy his/her requirements [5]. In such content-based filtering system each user is supposed to operate independently. The result of this is a content-based filtering system selects information items on the basis of correlation between the content of the items and the preferences of the users as contrast to a collaborative filtering system that selects items based on the correlation between people with similar preferences [6,7].Electronic mails were the earlier domain for the work on information filtering ,subsequent papers documented diversified domains including the newswire articles, Internet "news "articles and broader network resources[8-10].Mostly the Documents processed in content-based filtering are textual in nature and thus makes content-based filtering near to text classification. Filtering process can be modeled as a binary classification, single label, partitioning incoming documents into relevant

and non-relevant categories [11]. In enhancement some complex filtering systems includes multi-label text categorization in which automatically labeling of messages into partial thematic categories is done.

Detailed comparison analysis is specified in [4] which confirmed superiority of Boosting-based classifiers [12], Neural Networks [13] and Support Vector Machines [14] over other popular methods, such as Rocchio [15] and Naive Bayesian[16]. However, it is a profit to note that maximum of the work in 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. The Short length messages that are posted on the user's wall have been a challenge for the application of content based techniques. In the Scientific community short text classification has been taking attention these days. From the recent work it is clearly seen that there are difficulties in detecting the short text messages which defines a robust features as there description is concise or they are not spelled properly or they are not up to the standard terms and noise. To improve the classification of short text strings Zelikovitz and Hirsh [17] attempted to develop a semi supervised learning strategy which was based on a combination of labeled training data plus a secondary corpus of unlabeled but related longer documents. This solution was inapplicable in OSN domain in which short messages are not summary or part of longer semantically related documents. Thus in OSN classes are considered to match the key features for defining a policy-based personalization strategy.

2.2 Policy-based Personalization

Some proposals have exploited classification mechanisms for personalizing access in OSNs. To categorize short text messages, in [18] a classification method has been proposed in order to avoid overwhelming users of micro blogging services by raw data which focused on the tweets and associated the categories relating to content of the tweets. Therefore providing user to view certain types of tweets based on his/her interests. In contradictory, Golbeck and Kuter [19] proposed an application FilmTrust that exploits OSN trust relationships and provenance information to personalize access. Hence such filtering policy have some issues in providing the user adequate result analysis for classification process to decide how and to which extent filtering of unwanted messages is possible. In contrast [20] proposed filtering policy specifies that the settings of FRs according to the variety of criteria needed are provided so that it considers both the results of the classification process and also the relationships of the wall owner with other OSN users as well as information on the user profile. Moreover in [20], as the complementary the BL management which is the flexible mechanism that provides a further opportunity of customization.

The MyWOT a only social networking site which provided filtration has an quite different approach .This filtration as compared to filtered wall stands low. An extension to the access control can be considered as a content filtering, since protection is given from both objects from unauthorized subjects, and subjects from inappropriate objects. Till now in the field of OSN the techniques used for access control is that the requestor should have some relation with the user. Proposed system in [20] uses a similar idea to identify the users to which a FR applied. But there filtering policy language in access control has been extended to cope with the extended requirements of the filtering domain. The system checks for the content rather then the access control .Thus

extending one step ahead the system proposed in [20] maintains a Blacklist BL which are not implemented in any of the other specified modules till date. Talking about the Policy language specified in this paper has some relationship with the other policy framework which supports the specification and enforcement of constraints for the machine resources provided by rthe semantic web. Some example of such frameworks are KAoS and REI, which focuses mainly on access control, Protune [21], provides a support and to trust negotiation and privacy policies, and WIQA [22], which enables end users with the ability of using filtering policies which gains a quality demanded by the user from web resources. Eventually all the above frameworks are very powerful still the proposed system in [20] uses its own policy for filtering the OSN .The proposed system by us will work on same policy with an extended feature of filtering the word related to it using ontology.

3. PROPOSED METHODOLOGY

The aim of the present work is to experimentally evaluate an automated system which filters the unwanted messages from the user's wall of his/her social network account. Thus present work has implemented ML text classification techniques which automatically assign's each text message a set of categories based on its content. The main purpose behind the present system was to implement a text classifier which would concentrate in extracting and selecting from a set of categories. The proposed methodology introduces a system which has an add-on feature where the text having a meaning of any such unwanted message will also be filtered using Ontology based algorithm. The main goal of the proposed system is to filter all the text that has any type of similar meaning that an unwanted message has. So to achieve this type of system the proposed system combines Ontology based algorithm and the Ml based content filtering algorithms .The System will also extend the Present work component Black list by adding a feature to it of permanent banning of such blacklisted user if he/she tries to breach warnings.
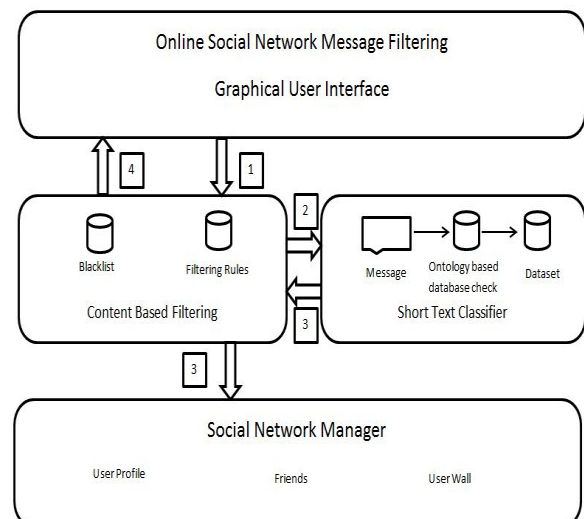
3.1 Proposed Architecture

Fig.1. Proposed Conceptual Algorithm for filtering the user's wall in OSN

Figure 1 illustrates the Conceptual architecture of the technique to filter any type of unwanted message on user OSN wall. As from the figure it is understood that the Module are

included to form a conceptual algorithm. Text Classification, Flltering rules, Blacklisting, Ontology based Data sets are some models that are combined. As these Modules stand the best option at a larger prospect when combined it will give the satisfactory results.

3.2 Working Modules

3.2.1 Filtering rules

The powerful rule based layer provides a system where user can specify a filtering rule (FR) where the user can select what content should be displayed on his/her wall.

3.2.2 Ontology based Data base Check

The system includes a checking criteria where the data base is checked for the words which resemble the same meaning as that of the vulgar meaning, or any undesired meaning using the ontology based classification algorithm.

3.2.3 Blacklists

The Blacklist (BL) are the extra add–on feature given to the user for blocking the person who is trying to again and again post any unwanted /vulgar message on users wall. The system automatically blocks such user if the person many times tries to post undesirable message on user wall.

4. CONCLUSION

The paper contains the working system to filter the unwanted message from online social network .The Proposed system is capable of filtering the undesired message or any type of text that has a similar meaning of unwanted message from the users wall. The exploitation system for ML soft classifier to enforce customizable content-dependent FRs is developed. In a greater term, the flexibility of the system is enhancement of filtering options is done through the management of BLs. The newly analyzed encouraging results thus obtained by proposed system will improve the quality of classification in the future. Proposed system is based on the short text classification strategy which will classify the unwanted messages with respective meanings also classified. Thus proposed system checks for the unwanted messages and filters it.

This work is the first step of a upcoming wider project as the work that are used in present osn can merely filter any such type of message .The proposed system can be used to update the present osn so that all the messages can be filtered providing the secured environment for the user.

5. REFERENCES

[1] 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.

[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," in Proceedings of the Fifth ACM Conference on Digital Libraries. New York: ACM Press, 2000, pp.195–204

[4] F. Sebastiani, "Machine learning in automated text categorization,"ACM Computing Surveys, vol. 34, no. 1, pp. 1–47, 2002.P. J. Denning, "Electronic junk,"

Communications of the ACM, vol. 25, no.3, pp. 163–165, 1982.

[5] 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.

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

[7] 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.

[8] 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.

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

[10] P. E. Baclace, "Competitive agents for information filtering," Communications of the ACM, vol. 35, no. 12, p. 50, 1992.

[11] P. J. Hayes, P. M. Andersen, I. B. Nirenburg, and L. M. Schmandt,"Tcs: a shell for content-based text categorization," in Proceedings of6th IEEE Conference on Artificial Intelligence Applications (CAIA-90). IEEE Computer Society Press, Los Alamitos, US, 1990, pp.320–326.

[12] 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.

[13] H. Sch¨utze, 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.

[14] 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.

[15] "A probabilistic analysis of the rocchio algorithm with tfidf for text categorization," in Proceedings of International Conference on Machine Learning, 1997, pp. 143–151.

[16] S. E. Robertson and K. S. Jones, "Relevance weighting of search terms," Journal of the American Society for Information Science, vol. 27, no. 3, pp. 129–146, 1976.

[17] 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.

[18] 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.

[19] 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.

[20] Marco Vanetti, Elisabetta Binaghi, Elena Ferrari, Barbara Carminati, an Moreno Carullo, "A System to Filter Unwanted Messages from OSN User Walls*",*2013.

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

[22] 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. 1–10, January 2009.