

# A Survey on an Intelligent System for Unwanted Message Filtering from OSN User Wall

Miss. Radhika S. Fulzele<sup>1</sup>, Prof. Vipul Dalal<sup>2</sup>

M. E. Student, Computer Engineering, VIT Mumbai, India<sup>1</sup>

Associate Professor, Department of Computer Engineering, VIT Mumbai, India<sup>2</sup>

**Abstract:** OSN often encounters the issue of users' control over their private messages. OSN does not completely support this requirement. This dissertation deals with the same issue, where a system is proposed, in which users directly can control the posted messages on their wall. For achieving this, a rule based system is designed which allows a user to manage their filtering criteria. Also the use of machine learning based soft classifier helps in automatically labeling the messages that supports content based filtering.

**Keywords:** On-line Social Networks, Machine learning, Information Filtering, Short Text Classification.

## I. INTRODUCTION

Now a days, online social networks (OSN) is mainly used as an interactive medium to communicate, share day to day information. OSN is used to share several types of content like text, image, audio and video data. OSN serves as a platform to build social relations with people. It also lets us find our old friends, colleagues who are active on social media website. It is a web based service that allows individuals to create a public profile and send a request to the one with whom sharing of information is required. Some of the social network sites are: face book, twitter, goggle+, etc.

OSN is generally used for web content mining strategies used to automatically discover useful and relevant information from a large amount of data. 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. Information filtering should enable the users to control the message written on their own walls by filtering out unwanted message. However, OSNs do not completely support this requirement. For example, Facebook allows user to state as to who is allowed to insert messages on user walls. No content

based preferences are supported hence it is not possible to filter political, vulgar or any type of unwanted messages. This service not only uses the earlier defined web mining techniques but it requires an adhoc classification strategy as wall messages consists of maximum use of short texts but the methods currently in use produce the poor results in evaluating such messages.

## II. LITERATURE SURVEY

To reach relevant products earlier recommender systems were used, that focused mainly on the user likes and dislikes whereas Raymond J. Mooney et al. [1] proposes content based method (LIBRA) where it uses the information of the item itself for access to relevant products. A content-based book recommending system is described that utilizes information extraction and a

machine-learning algorithm for text categorization. Initial experimental results describe that this approach can produce accurate recommendations.

M. Vanetti et al. [2] suggests that the user to have straight rule over their own wall to avoid the unwanted messages. This paper suggests that the user can directly control his/her private wall messages by using an automated system called filtering wall (FW). This system blocks unwanted messages but will not block the sender from writing unwanted messages, which is the only drawback of the system.

Victoria Bobicev et al. [3] used prediction by partial matching method (PPM). This method compresses the texts to capture text features and creates a language model adapted to a particular text. It shows that the method achieves a high accuracy of text classification.

Fabrizio Sebastiani [4] suggests an approach to the problem of text categorization based on machine learning technique which consists of a general inductive process that 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- improved effectiveness, prominent savings in terms of expert labor power, and straightforward portability to different domains.

Stephen Pollock [5] describes the ISCREEN prototype system for screening text messages. It is a rule-based system that was integrated with an existing information system.

It includes the following components:

- Friendly interface that allows unskilled users to specify instructions to the system in terms of rules;
- Conflict detection component that evaluates users' instructions for completeness and consistency;

- Screening component that intercepts messages as they are received, decides on actions based on users' rules, and takes the actions it has recommended.

It describes definitions of rules and message classifications, and explains conflicts that are found in user's rules. The system is currently limited in its ability to understand the content of messages. Thorsten Joachim's [6] explores the use of Support Vector machine (SVM) for learning text classification.

- Text Categorization

It has become one of the key techniques for handling and classifying text data. These techniques are used to categories news stories, to find interesting information on the web, and to guide a user's search through hypertext. Since building text

Classifiers is difficult and time-consuming, therefore it is use Support Vector Machines (SVMs) for text categorization.

- Support Vector Machine

It is a new learning method. SVMs are universal learners. In their basic form they learn linear threshold function. They can be used to learn polynomial classifiers, radial basic function (RBF) networks, and three-layer sigmoid neural nets. One remarkable property of SVMs is that their ability to learn can be independent of the dimensionality of the feature space. It measure the problem of hypotheses based on the margin with which they separate the data, not the number of features. SVMs acknowledge the particular properties of text:

- (1) High dimensional feature spaces,
- (2) Few irrelevant features, and
- (3) Sparse instance vectors.

The advantage of SVMs is their robustness and eliminated the need for feature selection, making the application of text categorization considerably easier. SVMs show good performance, avoiding catastrophic failure. It is used for text categorization.

Barbara Carminati et al. [7], describe filtering rule over their own private wall to avoid the unwanted messages. The users have a straight control over various messages posted on their own private space. So we are using the automated system called Filtered wall (FW), which has the capacity to filter unwanted messages by applying filtering rules. The system will blocks only the undesired messages send by the user. The drawback of this paper is user will not be blocked, means only the content posted by the user will block. Content based filtering is performed. Here the user can manually perform the blacklist process.

Jennifer Golbeck [8] introduces a social network is the popular movement on the network. Two level approaches are stated to integrated trust, gloss and origin. Here the algorithm is used for concluding trust relationship with

origin content and trust gloss in web social network. Firstly the algorithm computing personalized trust make acceptable using the origin of existing trust annotations in social networks. Then, we introduced two applications that join together the computed trust values with the origin of other annotations to personalize websites. In this paper it inferring the trust relationship within them. FilmTrust that combines the computed trust values with the origin of other gloss to personalize the website. Filmtrust application is introduced which uses trust to video ranking and ordering the review. We consider film trust give the good crop model.

### III. PROBLEM DEFINITION

Some people will use offensive and vulgar words in commenting on public posts. Providing this service is not only a matter of using previously defined web content mining techniques for a different application, rather, it requires to design ad-hoc classification strategies. Hence filtering rules are applied over unwanted messages that prevent vulgar messages. Generally a user is allowed to maintain a black list of users who keep posting unnecessarily. This is done manually. The present work proposes and experimentally evaluates 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 content. The major efforts in building a robust short text classifier are concentrated in the extraction and selection of a set of characterizing and discriminate features.

### IV. CONCLUSION

Some people might use offensive language while commenting on public posts. The system will filter unwanted messages by using filtering rules. The system handles black list manually but it is not automated. Different text categorization techniques enhanced with filtering are also studied.

### REFERENCES

- [1] R. J. Mooney and L. Roy, "Content-based book recommending using learning for text categorization," in *Libraries*. New York: ACM Press, 2000, pp.195–204..
- [2] 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.
- [3] 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.
- [4] F. Sebastiani, "Machine learning in automated text categorization," *ACM Computing Surveys*, vol.34, no.1, pp.1–47, 2002.
- [5] S. Pollock, "A rule-based message filtering system," *ACM Transactions on Office Information Systems*, vol. 6, no. 3, pp. 232–254, 1988
- [6] 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.
- [7] Marco Vanetti, Elisabetta Binaghi, Elena Ferrari, Barbara Carminati, and Moreno Carullo, "A System to Filter Unwanted Messages from OSN User Walls", 2013.
- [8] 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.
- [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] 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.
- [11] C. Apte, F. Damerau, S. M. Weiss, D. Sholom, and M. Weiss, "Automated learning of decision rules for text categorization," *Transactions on Information Systems*, vol. 12, no. 3, pp. 233–251, 1994.
- [12] S. Dumais, J. Platt, D. Heckerman, and M. Sahami, "Inductive Learning algorithms and representations for text categorization," in *Proceedings of Seventh International Conference on Information and Knowledge Management (CIKM98)*, 1998, pp. 148–155.
- [13] D. D. Lewis, "An evaluation of phrasal and clustered representations on a text categorization task," in *Proceedings of 15th ACM International Conference on Research and Development in Information Retrieval (SIGIR-92)*, N. J. Belkin, P. Ingwersen, and A. M. Pejtersen, Eds. ACM Press, New York, US, 1992, pp. 37–50.
- [14] Y. Zhang and J. Callan, "Maximum likelihood estimation for filtering thresholds," in *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2001, pp. 294–302.
- [15] E. D. Wiener, J. O. Pedersen, and A. S. Weigend, "A neural network approach to topic spotting," in *Proceedings of 4th Annual Symposium on Document Analysis and Information Retrieval (SDAIR-95)*, Las Vegas, US, 1995, pp. 317–332.
- [16] M. Carullo, E. Binaghi, and I. Gallo, "An online document clustering technique for short web contents," *Pattern Recognition Letters*, vol. 30, pp. 870–876, July 2009.
- [17] M. Carullo, E. Binaghi, I. Gallo, and N. Lamberti, "Clustering of Short commercial documents for the web," in *Proceedings of 19th International Conference on Pattern Recognition (ICPR 2008)*, 2008.
- [18] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. Cambridge, UK: Cambridge University Press, 2008.
- [19] 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