OSN FILTERING WALL USING HUMAN SENSE

[1] Rasal Rohit Sunil, [2] Pawar Kunal Dattatray,[3] Chavan Aniket Subhash, [4] Dalvi Sandip Bhausaheb

[1,2,3,4]Student, Computer Engineering, SGOI COE, Maharashtra, India

ABSTRACT

Abstract- With the large amount of growth in the data and use of social networking sites users share free text, image, audio and video data daily. Today's OSNs not provide much support to the users to avoid unwanted messages displayed on their wall. In system, if unwanted words are found, system directly blocks the message. But the problem is important articles which include unwanted message but having positive or good topics, gets blocked. So, to overcome the above problem, we proposed a system in which data will be divided into three categories: Normal, Low and High risk. Then filtering rules (FRs) are applied to block corresponding data using human intervention.

Keyword: - *Machine Learning Techniques (MLT), Online Social Network, Filtering Rules, VSM.*

---

1. INTRODUCTION

Now day's modern life is totally depending on Internet. The people cannot visualize life without Internet. Also, social networking is also part of modern life. Now a day's people share their ideas, views, information with each other uses social sites. Such communications media may include different types of contents like text, image, audio and video data [1]. Whereas large amount of data such as web links, news stories, blog posts, notes, photo albums, etc. are shared daily. In today's social network, there is a very large chance of posting unwanted content on particular public/private areas, called in walls. Information filtering (IF) can be used to give users the ability to automatically control the content of post written on their social walls, by filtering unwanted messages. Today OSNs offer very small to avoid unwanted messages or post on user walls. For example, In Facebook users are only permitted to state who is allowed to put messages in their walls i.e. direct to friends and indirect or other friends. Content-based preferences are not supported. The OSN wall messages are representing by short text for which traditional classification methods have some limitations. Therefore the goal of the present work is to propose and demonstration evaluates an automated system, called Filtered Wall, able to filter messages from OSN. It assigns the message automatically to each short text message, a set of categories based on its content we utilize Machine Learning (ML) text categorization methods [6].

In proposed system we classify the data in following types: Normal Risk (NR), Low Risk (LR) and High Risk (HR). In NR it consists of normal words which can be post on GUI directly. LR in which data is first processed as a black list from where it is identified. The data is of LR or HR, if data is of LR we select randomly three users from friend list among which if the given data is accepted by any of two users which has been taken into consideration randomly then the given data is post otherwise it would be blocked. In HR if the given data is accepted all of the three users then it would be allow to published or it would be blocked. The system provides a powerful rule layer a flexible language to specify FRs, by which users can state which contents shouldn't be displayed on their walls. FRs can support a variety of different filtering criteria that can be combined. The rules of filtering exploit user profiles, user relationships as well as the output of the Machine Learning categorization [6] process to state the criteria of filtering to be performed. In addition, the system will support for user-defined Blacklists (BL), the lists of users that are temporarily prevented to post any kind of messages on a user wall.

To the best of our information this is the first proposal of a system to automatically filter bad messages from OSN user walls on the basis of both messages content and the messages creator relationships and characteristics.

2. LITERATURE SERVEY

N.J.Belkin et al. [2] designed information filtering systems to classify a stream for generating information is dynamically too dispatched asynchronously by an information producer to satisfy according to user considerations. Nicholas J. Belkin has been talk about the relationship between information filtering and retrieval and they come to the conclude that both are two sides of the same coin the previous recommended systems use social filtering methods.

P.J.Hayes et al. [3] modeled filtering activity to perform binary classification to partitioning incoming documents into relevant and non-relevant categories. But that was difficult to include multi-label text categorization automatically labeling messages into partial thematic categories. By this system describe that a content-based book recommended system that utilizes information extraction method and a ML algorithm for text categorization. This way they improve access to relevant information and products.

M. J. Pazzani et al. [4] adopted feature extraction methods and collection of samples. This procedure maps text into a compact representation of generalization this phases. In this system real-time classification accuracy, and classification speed and they conclude that Linear Support Vector Machines are most unique classifier, fastest to train, and quick to evaluate.

D.D. Lewis et al. [5] had performed some experiments which proved that Bag of Words (BoW) approaches had good performance and prevail in general over more revealing text presentation that have superior semantics but lower statistical quality.

F. Sebastiani et al. [6] had given comparative analysis of Boosting-based classifiers, Neural Networks and Support Vector Machines over other popular methods, such as Rocchio and Na ̈ive Bayesian. The application of content-based filtering on messages posted on OSN user walls poses additional task or situation that given the short length of these messages other than the wide range of topics that can be discussed. A different approach is proposed by Bobicev and Sokolova that circumvent the problem of error-prone feature construction by adopting a machine learning method that can perform reasonably well without feature engineering.

B. Shriram et al. [7] proposed a classification method to categorize short text messages in order to avoid overwhelming users of micro blogging services by raw data. They focus on Twitter 2 and associate a set of categories with each argument describing its content. The user can then view only certain types of tweets based on his/her interests.

S. E. Robertson and K. S. Jones,[8]proposed “Relevance weighting of search terms,” Journal of the American Society for Information Science, In this paper they examine the statistical techniques for exploiting relevance information to weight search term. These methods are presented as a neural extension of weighting method using information about the distribution of index term in documents in general .A series of relevance weighting functions is derived and is justified by theoretically.

S. Zelikovitz and H. Hirsh,[9]proposed “Improving short text classification using unlabeled background knowledge,” We describe a method for improving the classic- fiction of short text strings using a combination of labeled training data plus a secondary corpus of unlabeled but related longer documents. We show that such unlabeled background knowledge can greatly decrease error rates, particularly if the number of examples or the size of the strings in the training set is small. This is particularly useful when labeling text is a labor-intensive job and when there is a large amount of information available about a particular problem on the World Wide Web. Our approach views the task as one of information integration using WHIRL, a tool that combines database functionalities with techniques from the information-retrieval literature.

3. RELATED WORK

In this paper the various information filtering techniques are used to remove unwanted contents by using content base filtering rules, Machine learning approach according to user interest an item. In this system we classify the data in following types: Normal Risk, Low Risk, and High Risk. The classification of data will down with the help of various technique, Such as Vector space model. This will use for text representation. Various filtering rules will use and blacklist will created.

3.1 Filtering Rules and Blacklist Techniques:

Filtering Rules:-

In defining the language for rules specification, we consider some issues that, in our opinion, should affect a message filtering decision. Firstly, in OSNs everyday life, the same post or message may have different meanings and based on who writes it. As effect, FRs should allow users to state constraints on post creators. Creators on which a rules applies can be selected on the basis of several different criteria; one of the most relevant is by massive conditions on their profile's attributes. In such a way it is, for instance, possible to define rules applying to young creators.

Given the social network scenario, creators may also be identified by providing the information on their social graph. This implies to state conditions on type, trust and depth values of the relationship creators should be involved in order to apply them the specified rules. This make the system able to automatically perform tasks, the BL mechanism has to be instructed with some rules, hereafter BL rules.

Blacklists:-

In this component of the system is a mechanism to prevent spam BL creators, regardless of their content. BL'S are directly managed by the OSN system, which should be able to determine who the words are inserted in the BL To improve flexibility, this information is in the system by a set of rules, the rules on BL. These rules are not defined by the user, so they are not intended as guidelines for high level will be applied to the entire community. As per our condition the black list will be decided. Means as per the risk of message or post. For deicide the there are some rules as per that rules risk will decided an create a BL. Therefore, using a rule of BL, will able to prevent the walls from unwanted message they do not know directly (i.e. with whom they have indirect relations).

3.2 Vector Space Model:

The vector space model procedure can be divided in to three stages. The first stage is the document indexing where content bearing terms are extracted from the document text. The second stage is the weighting of the indexed terms to enhance retrieval of document relevant to the user. The last stage ranks the document with respect to the query according to a similarity measure. The definition of *term* depends on the application. Typically terms are single words, Keywords, or longer phrases. If words are chosen to be the terms, the dimensionality of the vector is the number of words in the vocabulary

3.3 Text Representation:

The extraction of an appropriate set of features which represents the text of a given document is a crucial task to perform in overall classification strategy. Different sets of features vortex categorization have been proposed in the literature. We consider two types of features, Bag of Words (BoW) and Document properties (Dp), that are used in the experimental evaluation. It is the combination that is most appropriate for short message classification. We introduce Content Filter modelling information that characterizes the user is posting. These features play a key role to understanding the semantics of the messages. The underlying model for text representation is the Vector Space model (VSM) to which a text document djis represented as a vector of binary or real weights dj= W1j , . . . , W|T|j, where T is the set of terms (sometimes also called features) that occur at least once in at least one document of the collection of document Tr, and wkj∈[0; 1] represents how much term tk contributes to the semantics of document dj. In the BoW representation, terms are identified with words.

4. FILTERED WALL ARCHITRCTURE

This is three layers architecture such as follows, Social Network Manager, Social Network Applications, and Graphical User Interfaces.

A. Social Network Manager (SNM):- As shown in the fig-1 first layer is a Social Network Manager (SNM), which provides the OSN functionalities i.e., profile and relationship management. All data of user profile is maintains in this layer an provide to next layer.

B. Social Network Applications (SNA):- The second layer provides the support for external Social Network Applications (SNAs). In this layer the message categorization is down according to filters rules it is very important part in this layer. In this layer there is three types of classification Normal, Low and High Risk. The supported SNAs need an additional layer for their desired Graphical User Interfaces.

C. Graphical User Interfaces (GUI):- Third layer provides GUI to the user who wants to post his message as an input.

By considering this architecture our system is placed in the second and third layers. To set up and manage FRs/BLs users interact with the system through GUI. The filtered wall is shown to the user through the GUI, where only authorized messages are post.

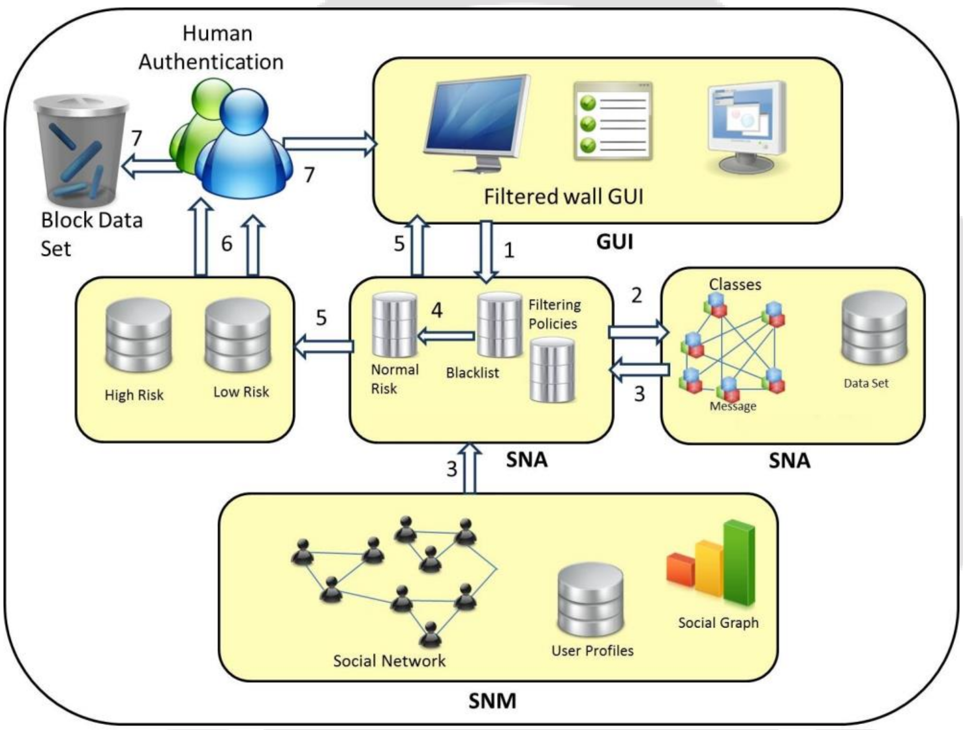

Fig-1: Filtered Wall Conceptual Architecture

As shown in Figure 1, the path followed by a message, from its writing to the possible final publication can be given as follows:

- The user post a message after entering the private Wall of his/her contacts which is interrupted by wall.
- From the message content a ML based text classifier extracts Meta data.
- A meta-data of user profile will be provided to system for the filtering various post or message.
- The category of message will be decided. Normal Risk, Low Risk and High Risk.
- Suppose message should be Low risk or High Risk then Human authentication will be check for posting the message.
- The message will be published or filtered by FW depending on the result of the previous step.

5. EXPERIMENTAL RESULTS

- We present OSN site with basic functionalities of OSNs.
- In this system, using Filtering Rules we can make Filter Wall for preventing unwanted messages.
- Initially, we focus on Violence, Vulgar, Sexual, Offensive, Hate type of messages and filter these messages.
- Also, maintain Black list for the word that will send the prevented type of messages more times than that word will automatically put into Black List.
- Administer can see monthly and yearly reports as well as Graphs like which category of messages are filtered, who is message creator.

6. CONCLUSIONS

Presented system in this paper filtering unwanted messages from OSN user walls. This system describes a ML soft classifier to enforce customizable content-dependent FRs. The flexibility of this system in terms of filtering options is enhanced through the management of BLs. Here the batch learning strategy which is based on the preliminary collection of the whole set of labeled data articles from authors allowed an accurate experimental evaluation but needs to be evolved to include new operational requirements. Our strategies and techniques limiting the inferences that a user can do on the enforced filtering rules with the goal of bypassing the filtering system, such as for instance randomly notifying a message that should be blocked, or detecting modifications to profile attributes that have been made for the only purpose of defeating the filtering system.

7. ACKNOWLEDGEMENT

We offer special thanks to the Prof. R. R. Rathod who have guided towards development of our system paper. Also thanks all who helped in the development of system and giving their valuable suggestions. So that we are able to improve our system.

8. REFERENCES

[1] Marco Vanetti, Elisabetta Binaghi, Elena Ferrari, Barbara Carminati, Moreno Carullo Department of Computer Science and Communication University of Insubria 21100 Varese, Italy “A System to Filter Unwanted Messages from OSN User Walls” IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING VOL:25 YEAR 2013

[2] 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.

[3]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.

[4] M. J. Pazzani and D. Billsus, “Learning and revising user profiles:The identification of interesting web sites,” Machine Learning,vol. 27, no. 3, pp. 313–331, 1997.

[5] 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.

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

[7] 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.

[8] 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.

[9] S. Zelikovitz and H. Hirsh, “Improving short text classification using unlabeled background knowledge,” in Proceedings of 17th Interna- tional 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]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.
[12] C. Apte, F. Damerau, S. M. Weiss, D. Sholom, and M. Weiss, "Automated learning of decision rules for text categorization," Trans- actions on Information Systems, vol. 12, no. 3, pp. 233–251, 1994.
[13] 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.
[14] 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.
[15] 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.
[16] 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.
[17] P. J. Denning, "Electronic junk," Communications of the ACM, vol. 25, no. 3, pp. 163–165, 1982.
[18] 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.
[19] P. S. Jacobs and L. F. Rau, "Scisor: Extracting information from on- line news," Communications of the ACM, vol. 33, no. 11, pp. 88–97, 1990.
[20] S. Pollock, "A rule-based message filtering system," ACM Transactions on Office Information Systems, vol. 6, no. 3, pp. 232–254,1988.
[21] P. E. Baclace, "Competitive agents for information filtering," Communications of the ACM, vol. 35, no. 12, p. 50, 1992.
[22] G. Amati and F. Crestani, "Probabilistic learning for selective dissemination of information," Information Processing and Management, vol. 35, no. 5, pp. 633–654, 1999.
[23] 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.
[24] 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.