

Review Article

Filteration of Unwanted Text Messages from Online Social Networks: A Review

Garima Singh¹, Prof. Amit Yerpude², Prof. Toran Verma³¹⁻³Rungta College of Engineering and Technology, Dept. of Computer Science and Engineering, Bhilai, Chhattisgarh, India

*Corresponding author

Dr. Garima Singh

Email: garima93singh@gmail.com

Abstract: Now-a-days, Online Social Networking sites are important part of human society. In other way a person maintains a social networking site account which is then used for the following purposes like building relations, for business, exchange of data etc. The use of OSN's is therefore tremendously increasing day by day and its demanding for concentration over protection of private information of user's, their likes and dislikes etc. There are different filtering criteria's provided by various OSN's but the role of user to prevent his/her privacy is very less with the help of those criteria's. This paper represents a model that helps to allow the user to take over the control of their profile's and walls. The key parameters of the proposal illustrate are that the message posted on user wall's will be filtered for unwanted content in any form and type and will be posted with the users consent only. The feasibility of this model is presented considering OSN's scenarios.

Keywords: Filtering System, Online Social Networks, Information Filtering, Clustering.

INTRODUCTION

The internet is filled with almost every types and kind of information. To achieve or say to use and gain knowledge from this data there are equal numbers of brains i.e. users exist on internet. The same users can share, disseminate as well as communicate multiple variety of information among them. The information is in different types say texts, audio, video, images etc. The medium for this type of data exchange can be in the form of mails, messages and social networks on the internet.

Online Social Networks: OSN's is an internet service that helps the user to increase social networks and relations with other users for sharing of backgrounds, interests, discover real life connections and participate in multiple activities. OSN services basically are web based group centered services in which users can share almost every type of information.

According to the statistics there are 1.28 billion users of Face book and on an average more than 30 billion pieces of content are shared each month [1]. Similar is the case with Twitter comprising of 255 million users and sending of 500 million tweets every day [8]. The same trend is followed by other OSN's like Instagram, Google plus, Linked In, etc. this type of huge and with

dynamic characteristics the data that concludes for a strategy with the help of web content mining to establish and extract useful information from the data.

Therefore in a way OSN's has transformed the way of interaction and exchange of information among users over the internet. The aim at which OSN's lack its concern is filtering the Information. There is less support for prevention of unwanted messages on users profile and walls. A small support system is arranged by some of the OSN's.

For an illustrate consider Face book but it also allows filtering up to a certain level i.e. friends, this may happen sometimes that friends of friends or some groups may posts message on user's profile which is also termed as a wall. But the variety of types and contents of message is not explained and is not taken into consideration by support system. User who is permitted to post a message can post any type of content in the message regardless of the category associated with the message. The message content may be vulgar, political or personal and thus the data contained by the message no longer left private on public platform.

Filtered Wall Architecture

The architecture of OSN services is a three-tier structure. It mainly consists of three layers

- Social Network Manager (SNM)
- Social Network Application (SNA)
- Graphical User Interface (GUI)

Social Network Manager:

It persists the basic OSN related functionalities i.e., the user profiles and relationship management and also provide the information to the social network application layer.

Social Network Application:

To provide external social network application, the layer apply filtering rules (FR) and blacklists (BL). SNA layer composed of short text classifier and Content Base Message Filtering (CBMF).

It is very important layer because the classifier distinguish each message according to its content and CBMF filters the message according to filtering criteria and blacklist provided by the user. CBMF is used to select data item based on the correlation between the content of item and the user preferences.

Graphical User Interface:

It provides the interaction between the users and the system. Which user gives her input and is able to see published wall messages on social media. Moreover GUI also provides user the facility to apply filtering rules for his/her wall messages and helps to give list of BL user who are temporally prevented to publish messages on user's wall. The GUI also include of Filtered Wall (FW) where the user is able to see her desirable messages.

- After entering one of the user to the private wall, the user tries to post a message which has been captured by Filtered wall.
- A ML-based text classifier obtain data from the message content.
- Filtered wall uses data given by the classifier, along with data obtained from the user's profiles, to implement the filtering rules and blacklists techniques.
- Considering the result of the previous step, the result message will be filtered.

LITERATURE SURVEY

Priyanka Salunkhe [1], focused on today various social networking sites are available which make people remain in constant touch with each other. Sharing any type of data has become easy. There are great advantages of such social networking sites excepting a few minor drawbacks like poor security which create huge problems to people when they were active on such sites. As we have seen Face book allows users to post comment on another users wall even when they were

unknown to each other. But if that comment is a vulgar one then it may cause serious problem to user reputation. To avoid such a problem Information filtering is used to filter the content of the message. So we have analyzed various Information filtering methods like content based filtering, policy-based filtering, and collaborative filtering in this paper. Content-based filtering method is best filtering method than any other methods, because it has filtered out bad or non-neural words from the input message and allows posting only pleasant comment to be posted on a user's wall. This will help us to avoid unwanted messages from ever spoiling reputation which carries the utmost importance in the world of socialization.

R. Yasotha [2], focused on a new approach to automatically categorize text documents. Instead of the common practice of string matching approach used for automatic text document categorization this paper proposes a Latent Dirichlet allocation (LDA) based approach. Here the clusters identified by LDA are labeled based on the underlying natural clusters on the domain in concern. Proposed model was able to categorize unseen documents with an accuracy of 66.66%. The lack of accuracy is due to the limitation of classification on one level. The proposed approach can be extended to further lower levels and would be able to achieve greater accuracy.

Sunil Yadav [3], the goal of this paper was to provide an initial base line to build a filtered discussion forum on the most critical issues related to the online social networks. The development of the sample learning websites and blacklisted words table made the filtering easier. As blogs comment and feedback comments are concerned, our proposed work has worked successfully.

Swapnali V. Jadhav [4], have discussed the literature survey of the filtering system. We are developing a system to filter undesired messages from OSN walls. The wall that restricts the unwanted message called as the Filtered Wall (FW). In this report we discussed the idea about the system. Additionally, we studied strategies and techniques limiting the inferences that a user can do on the enforced filtering rules with the aim of bypassing the filtering system, such as for instance randomly notifying a message that should instead be blocked.

Marco Vanetti [5], have presented a system to filter undesired messages from OSN walls. The system exploits a ML soft classifier to enforce customizable content-dependent FR's. Moreover, the flexibility of the system in terms of filtering options is enhanced through the management of BLs. This work is the first step of a wider project. The early encouraging results we have obtained on the classification procedure prompt us to continue with other work that will aim to improve the

quality of classification. In particular, future plans contemplate a deeper investigation on two interdependent tasks. The first concerns the extraction and/ or selection of contextual features that have been shown to have a high discriminative power. The second task involves the learning phase. Since the underlying domain is dynamically changing, the collection of pre classified data may not be representative in the longer term. The present batch learning strategy, based on the preliminary collection of the entire set of labeled data from experts, allowed an accurate experimental evaluation but needs to be evolved to include new operational requirements. In future work, we plan to address this problem by investigating the use of online learning paradigms able to include label feedbacks from users. Additionally, we plan to enhance our system with a more sophisticated approach to decide when a user should be inserted into a BL.

L. Roy and R.J. Mooney [6] utilizes shared filtering technique, however in our proposed framework content based filtering is utilized. It characterizes the content based book proposal framework that creates data pulling out and machine learning procedure for text classification.

B. Carminative *et al.* [7] are characterizing in this paper, the framework can by and large take decision about the message which is blocked, in light of the acceptance depends up on statistical information.

Bodicev and M. Sokolova [8] classification of content put in complex and particular phrasing; require the utilization of learning procedure. Fractional Matching technique is connected which extend the content for confining the content trademark. Fractional equivalent builds up a language demonstrate. The yield of Fractional coordinating compression gives predictable care of content classification

J. Colbeck [9] OSN is the basic concentration groups in network. To make the confidence numerous clarifications are required. Two level methodologies are expressed to join gloss, trust and starting point. We express an algorithm for finishing up trust association with inception data and gloss confide in web social network. Film trust application is acquainted which utilizes trust with motion picture positioning and requesting the review. We can consider film trust give the good product.

E. Ferrari [10] this paper proposes a framework implementing content based message filtering for Online Social Networks (OSNs). The framework enables OSNs clients to have an immediate control on the messages posted on user walls. This is accomplished through an adaptable manage based framework, that enables a user's to modify the filtering

criteria to be connected to their walls, and a Machine Learning based short content classifier naturally marking messages in support of content based filtering.

Gediminas Adomavicius [11] gives a diagram of the field of recommender frameworks and depicts the present era of suggestion techniques that are generally characterized into the accompanying four fundamental classifications: content-based, collective, Policy based personalization and hybrid recommendation approaches. This paper likewise depicts different impediments of current suggestion strategies and examines conceivable expansions that can enhance proposal capacities and make recommender frameworks material to a significantly more extensive scope of utilizations. In this paper, they evaluated different impediments of the present suggestion techniques and talked about conceivable augmentations that can give better recommendation capacities. This extension incorporate among others, the enhanced displaying of clients and items, incorporation of the relevant data into the suggestion procedure, bolster for multi-criteria evaluations, and arrangement of a more adaptable and less intrusive recommendation process.

Bharath Sriram *et al.* [12] states micro blogging sites, for example, Twitter, the clients may progress toward becoming overpowered by the raw information. One answer for this issue is the classification of short instant messages. As short messages don't give adequate word events, conventional classification strategies, for example, —Bag-Of-Words have restrictions. To address this issue, they propose to utilize a little arrangement of space particular elements removed from the creator's profile and content. The proposed approach successfully characterizes the content to a predefined set of non-specific classes, for example, News, Events, Opinions, Deals, and Private Messages. They have proposed a way to deal with group tweets into general however imperative classes by utilizing the creator data and elements inside the tweets. With such a framework, clients can subscribe to or see just certain sorts of tweets in view of their advantage.

Michael Beye *et al.* [13] talked about, as of late, Online Social Networks (OSNs) have turned into an imperative piece of day by day life for some. Clients construct unequivocal systems to speak to their social relationship, either existing or new. Clients additionally frequently transfer and offer a plenty of data identified with their own lives. The potential security dangers of such conduct are regularly belittled or overlooked. For instance, clients frequently unveil individual data to a bigger gathering of people than expected. Clients may even post data about others without their assent. An absence of experience and mindfulness in clients, and additionally appropriate apparatuses and outline of the OSNs, sustain the circumstance. This paper means to

give knowledge into such security issues and takes a gander at OSNs, their related protection chances, and existing research into solutions.

Josie Maria [14] examined Effective Web content filtering is a need in instructive and work environment conditions, however current methodologies are a long way from great. They talk about a model for content based shrewd Web content filtering, in which shallow linguistic examination plays a key part. With a specific end goal to demon strate how this model can be acknowledged, they have built up a lexical Named Entity Recognition framework, and utilized it to enhance the viability of measurable Automated Text Categorization strategies. They have played out a few analyses that affirm this reality, and energize the incorporation of other shallow linguistic preparing procedures in keen Web content filtering. They examined that shallow linguistic analysis as a rule, and Named Entity Recognition specifically, can be utilized to enhance the viability of content classification in the system of intelligent Web content filtering.

Michael Chau *et al.* [15] proposes a machine learning-based approach that consolidates Web content investigation and Web structure examination. They represent each Web page by an arrangement of substance based and link-based highlights, which can be utilized as the contribution for different machine learning algorithms. The proposed approach was actualized utilizing both a feed forward/back propagation neural network and a support vector machine. Two investigations were planned and led to analyze the proposed Web-include approach with two existing Web page filtering techniques — a keyword based approach and a lexicon based approach. The test comes about demonstrated that the proposed approach as a rule performed superior to anything the benchmark approaches, particularly when the quantity of preparing reports was little. The proposed methodologies can be connected in point particular internet searcher advancement and other Web applications, for example, Web content administration.

R.J Mooney *et al.* [16] depict a substance based book prescribing framework that uses data extraction and a machine-learning algorithms for content classification. Initial test comes about exhibit that this approach can deliver precise suggestions. These trials depend on appraisals from arbitrary samplings of things and they examine issues with past tests that utilize skewed samples of user selected cases to evaluate execution.

Fabrizio *et al.* [17] demonstrated that the computerized arrangement (or classification) of writings into predefined classifications has seen a blasting enthusiasm for the most recent 10 years, because of the expanded accessibility of reports in

advanced shape and the resulting need to arrange them. In the exploration group the prevailing way to deal with this issue is in light of machine learning systems: a general inductive procedure consequently constructs a classifier by learning, from an arrangement of pre-grouped reports, the attributes of the classifications. The upsides of this approach over the information building approach (comprising in the manual meaning of a classifier by space specialists) are a decent viability, extensive funds as far as master work control, and clear compactness to various areas. This review talks about the primary ways to deal with content arrangement that fall inside the machine learning worldview. They will talk about in detail issues relating to three distinct issues, in particular, report portrayal, classifier development, and classifier evaluation.

S. Pollok *et al.* [18] portrays the ISCREEN model framework for screening instant messages. ISCREEN incorporates an abnormal state interface for clients to characterize rules, a segment that screens instant messages, and a contention identification segment that inspects rules for inconsistencies. A clarification part utilizes content era to answer client inquiries about past or potential framework activities in view of Grice's conversational maxims.

F. Sebastian *et al.* [19] portrays the mechanized arrangement (or classification) of texts into predefined classifications has seen a blooming interest for the most recent ten years, because of the expanded accessibility of reports in computerized frame and the following need to sort out them. In the research community, the dominant way to deal with this issue depends on machine learning strategies: a general inductive process consequently manufactures a classifier by learning, from an arrangement of pre-ordered records, the qualities of the classifications. The upsides of this approach over the learning designing methodology (comprising in the manual meaning of a classifier by space specialists) are a decent adequacy, extensive reserve funds as far as master work control, and clear convenience to various areas. This overview talks about the primary ways to deal with content arrangement that fall inside the machine learning worldview. We will talk about in detail issues relating to three distinct issues, to be specific archive portrayal, classifier development, and classifier evaluation.

Adomavicius and Tuzhilin [20] examined an advanced recommendation frameworks idea, additionally talked about strategies for current proposal era are arranged into three fundamental classes cooperation, and hybrid recommendation proposal approaches in light of the contents. Yet, these methodologies don't consider the way toward comprehension of clients and items.

CONCLUSION

This survey paper presents various approaches for designing of a system to filter not required messages from OSN walls using different classifiers. Additionally the flexibility of a system can be increased through filtering rules as well as blacklist management. In this context, since the underlying domain is dynamic, the collection of pre-classified data which are provided for training purpose which is may not valid for longer time. Future scope lies in the use of any of the classifiers or hybridization of these classifiers can get more accurate results.

REFERENCES

1. Salunkhe P, Bharme S, Padiya P. Filtering unwanted messages from OSN walls. In Innovation and Challenges in Cyber Security (ICICCS-INBUSH), 2016 International Conference on 2016 Feb 3 (pp. 261-264). IEEE.
2. Yasotha R, Charles EY. Automated text document categorization. In Intelligent Computing and Information Systems (ICICIS), 2015 IEEE Seventh International Conference on 2015 Dec 12 (pp. 522-528). IEEE.
3. Yadav S, Das S, Rudrapal D. A system to filter unsolicited texts from social learning networks. In Computing, Communications and Networking Technologies (ICCCNT), 2013 Fourth International Conference on 2013 Jul 4 (pp. 1-5). IEEE.
4. Vanetti M, Binaghi E, Ferrari E, Carminati B, Carullo M. A system to filter unwanted messages from OSN user walls. IEEE Transactions on Knowledge and data Engineering. 2013 Feb; 25(2):285-97.
5. Vanetti M, Binaghi E, Ferrari E, Carminati B, Carullo M. A system to filter unwanted messages from OSN user walls. IEEE Transactions on Knowledge and data Engineering. 2013 Feb; 25(2):285-97.
6. Mooney RJ, Roy L. Content-based book recommending using learning for text categorization. In Proceedings of the fifth ACM conference on Digital libraries 2000 Jun 1 (pp. 195-204). ACM.
7. Vanetti M, Binaghi E, Carminati B, Carullo M, Ferrari E. Content-based filtering in on-line social networks. In International Workshop on Privacy and Security Issues in Data Mining and Machine Learning 2010 Sep 24 (pp. 127-140). Springer, Berlin, Heidelberg.
8. Bobicev V, Sokolova M. An Effective and Robust Method for Short Text Classification. In AAAI 2008 Jul 13 (pp. 1444-1445).
9. Golbeck J. Combining provenance with trust in social networks for semantic web content filtering. IPAW. 2006 May 3; 2006:101-8.
10. Vanetti M, Binaghi E, Carminati B, Carullo M, Ferrari E. Content-based filtering in on-line social networks. In International Workshop on Privacy and Security Issues in Data Mining and Machine Learning 2010 Sep 24 (pp. 127-140). Springer, Berlin, Heidelberg.
11. Adomavicius G, Tuzhilin A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE transactions on knowledge and data engineering. 2005 Jun; 17(6):734-49.
12. Sriram B, Fuhry D, Demir E, Ferhatosmanoglu H, Demirbas M. Short text classification in twitter to improve information filtering. In Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval 2010 Jul 19 (pp. 841-842). ACM.
13. Beye M, Jeckmans A, Erkin Z, Hartel P, Lagendijk R, Tang Q. Literature overview-privacy in online social networks. Centre for Telematics and Information Technology, University of Twente; 2010.
14. Hidalgo JM, García FC, Sanz EP. Named entity recognition for web content filtering. In International Conference on Application of Natural Language to Information Systems 2005 Jun 15 (pp. 286-297). Springer, Berlin, Heidelberg.
15. Chau M, Chen H. A machine learning approach to web page filtering using content and structure analysis. Decision Support Systems. 2008 Jan 31; 44(2):482-94.
16. Mooney RJ, Roy L. Content-based book recommending using learning for text categorization. In Proceedings of the fifth ACM conference on Digital libraries 2000 Jun 1 (pp. 195-204). ACM.
17. Sebastiani F. Machine learning in automated text categorization. ACM computing surveys (CSUR). 2002 Mar 1; 34(1):1-47.
18. Pollock S. A rule-based message filtering system. ACM Transactions on Information Systems (TOIS). 1988 Jul 1; 6(3):232-54.
19. Sebastiani F. Machine learning in automated text categorization. ACM computing surveys (CSUR). 2002 Mar 1; 34(1):1-47.
20. Adomavicius G, Tuzhilin A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE transactions on knowledge and data engineering. 2005 Jun; 17(6):734-49.