



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 1, January 2014

Spam Filtering in Online Social Networks Using Machine Learning Technique

T.Suganya¹, T.Hemalatha²

ME (final year), Department of Computer Science and Engineering, Karpagam University, Coimbatore, India¹

Assistant Professor, Department of Computer Science , Mar Gregorious College, Chennai, India ²

ABSTRACT: Internet has become a popular medium today. The usage of internet has increased widely. Internet reduces the manual work and also more time consuming .These is the reason why people today are connected to internet. In recent years Online Social Networks (OSNs) also developed well and plays an equal role. People use Online Social Networks (OSNs) to share their view of life. Online Social Networks (OSNs) enables them to connect with their friends and families. Online Social Networks (OSNs) has become an important part of many people life today. Therefore Online Social Networks (OSNs) should be highly secure to protect the individual's privacy. Though (OSNs) provides the security measures but it was limited. To filter the spam, in this paper, we propose an enhanced filtering measure by using a machine learning technique based on a content filtering.

Keywords: Online social networks, Machine Learning, short text classifier, Filtered Wall.

I. INTRODUCTION

Web mining is the use of data mining techniques to extract knowledge from Web . Web mining is an application of data mining techniques to discover knowledge from Web content, structure, and usage, is the collection of technologies to fulfill this potential.

Today internet has become a part of people's daily life. People use social network to share pictures, music, video etc., social network allows user to connect to various other pages on the web, including some use full sites like education, marketing, online shopping, business, e-commerce. Social networks such as Facebook , LinkedIn, MySpace, Twitter are more popular recently. LinkedIn allows user to connect to Professional group of peoples and provide more related connection to the user to endorse the use full people in education field. Facebook connects user to friends , friends of friends , public figure and a group of community. Twitter allows a user to follow their friends and post the comments. Twitter is very popular among celebrities and public figure. Online social network is a key to keep in touch with friends and families.

Spam is the major problem of today's internet. The spam is usually transformed by the spammers. Spammers spread useless information in social networks .spammers spread the spam from one user to another which leads to decrease the respect of individual in online community. Therefore spam should me identified and detected to protect the users' priority. The detection of spam is carried out by applying Machine learning technique to provide the secure online social networks to the user.

Section 2 describes the related work that includes the study of the features and content classification. Section 3 deals with the existing system and section 4 describe the proposed system. Section 5 focuses on result and Section 6 presents the conclusions.

II. RELATED WORK

Online social network is daily habit of more number of users today. The use of content-based filtering on messages appear on user walls have some challenges given the short text messages than the long text can be discussed. Short



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 1, January 2014

text classification obtained only limited focus by the scientific community today. Our work is focused on the difficulties in analyzing the robust features including the short text and misspelling of the data and even also the noise of the data.

The goal of the work is to evaluate the system called Filtered Wall (FW), which is used to filter the Spam message from the user profile. Also we use Machine Learning (ML) text categorization application to allocate the short messages based on the content. The main use of utilizing this technique is to provide a detail solution to the approach [1].

This is the first work to automatically filter spam messages from OSN user walls based on the content of the message and relationship of the message posted by the person and its characteristics [5].

The system is used to extract the message based on the structure of the online social network and also the type of the message posted on the wall of the OSN. The structure is analyzed based on the short text and it predicts the message as spam message or normal message. The clear solution of detecting the spam is provided based on the content based features [1].

III. EXISTING SYSTEM

Online Social Networks (OSNs) are very popular now days to make a instant chat, share and communicate the part of their life. Daily usage of Online Social Networks (OSNs) leads to share different types of data including text, video, music, images and multimedia data. Online social network will have the control on who can able to post the message on the walls of the user. Therefore the control is given only to the people who post message on the wall .They can either be friends , friends of friends ,group of friends. Also online social network allows the user only to post the message on the wall this type of control is named as (only me) option. But the online social network does not provide control over the content of the message (like vulgar, violence, sex, offensive etc..) therefore the chance of preventing the unwanted message is limited by the online social network. This is due to the short word occurrence.

A. Information filtering and information retrieval

Information filtering system is designed to analysis the data sent to the user . information filtering deals with data which are sent from remote control and other resources. Information filtering not only deals with text data but also process the multimedia data like image, music video etc.,

Filtering is based on the type of control set by the users on their wall in online social networking. The control can be either single or group of friends as desired by the user. Filtering mainly deals with filtering the data which is transmitted by the source rather than concentrating on the type of the data. User can only able to visualize the filtered data on their wall disregarding of the content type of the data. More information regarding the filtering work is presented in [1], [6].

B. Maximum Likelihood Estimation for Filtering Thresholds

Information filtering use the statistical retrieval models for calculating the number score which is used for matching the document with the user profile. Document with good profile scores are delivered by using the dissemination threshold . dissemination threshold improves the utility by using the number score detected which is used to analysis the use full and unwanted data. Use full data is analyzed by the parameter distribution.

For increase the function of dissemination thresholds , In this paper, a new approach is proposed for enhancing the function of dissemination thresholds by using a algorithm which is based on the Maximum Likelihood concept. A separate parameter is maintained for use full and unwanted data[8].

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 1, January 2014

C. Policy-Based Personalization Of Osn Contents

Many proposal are given for exploiting classification mechanism in online social network recently. For instance, in [9], classification method is used to categorize the short message obtained from online social network from the raw data. The system in [9] concentrate on Twitter and each of the tweet content in it. The priority of view of tweets by user is based on the privacy setting of the twitter user. Golbeck and Kuter [10] gives an application namely Film trust that detect the true online social network relationship. But filter of the raw data is limited so the result is focused on the relevant data from the OSN [1].

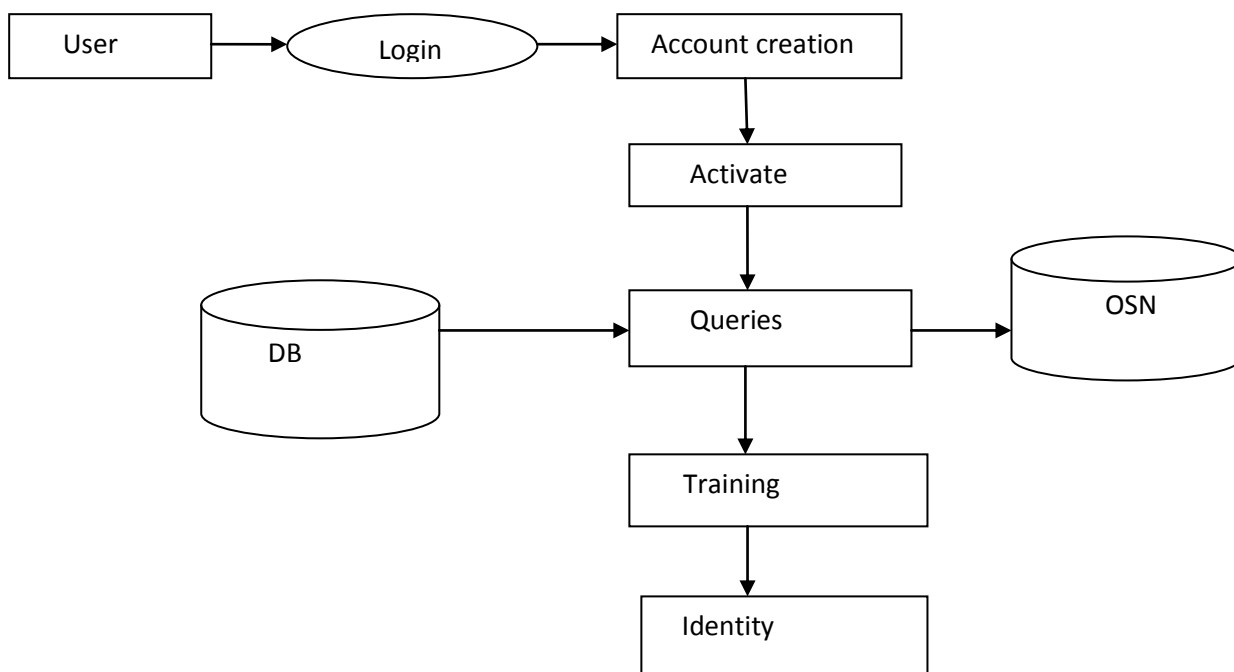


Fig. 1 system block diagram

The existing system works provides minimal security to the users so it is recommended to propose a Machine learning technique to filter the spam in the online social network.

IV. PROPOSED SYSTEM

Effective spam detection in OSN, is achieved by the use Machine learning technique. Content based features can be mined from the messages. The machine learning categorization can be used to classify the messages based on its contents. People involving in online social network are interested in posting their views and ideas mostly in the form of text. And the users in OSN environment are communicating through short messages. Content based features are extracted from the short messages posted from the user walls.



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 1, January 2014

A. Short Text Classifier

TABLE I
Long Words Substitution

Short words	Long words
Abt	About
Lemme	Let me
Tmrw	Tomorrow
Yup	Yes

The main task of the proposed system are the Content-Based Messages Filtering (CBMF) and Short Text Classifier .It additionally supports the classification of message based on the category set. Also it focus on the following:

- Filtered wall (FW) is used to intercept the message posted on the private walls of the user.
- From the content of the message ,meta data are extracted based on the Machine learning(ML).
- The extracted meta data are used by Filtered wall (FW) based on the classification and users' profile.
- Based on the result obtained , the messages are filtered by Filtered wall (FW).

TABLE III
Spam Words Samples

Hot	Cash	Offer
trip	Free	Win
Coin	Prize	Call

B. Machine Learning-Based Classification

We address short text categorization as a hierarchical two level classification process. The first-level classifier performs a binary hard categorization that labels messages as Neutral and Non-neutral. The first-level filtering task facilitates the subsequent second-level task in which a finer-grained classification is performed.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 1, January 2014

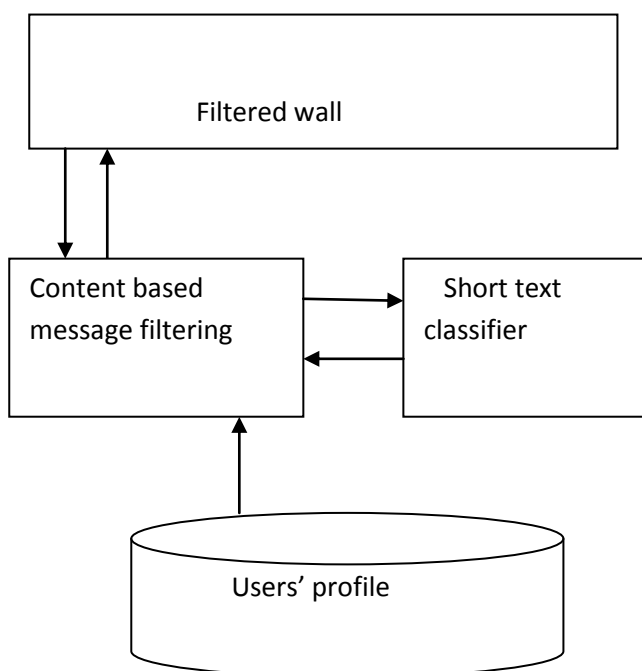


Fig. 2 Block diagram of proposed system

The second-level classifier performs a soft-partition of Non-neutral messages assigning a given message a gradual membership to each of the non-neutral classes. Among the variety of multiclass ML models well suited for text classification, we choose the RBFN model [11] for the experimented competitive behavior with respect to other state-of-the-art classifiers. RBFNs have a single hidden layer of processing units with local, restricted activation domain: a Gaussian function is commonly used, but any other locally tunable function can be used. They were introduced as a neural network evolution of exact interpolation [12], and are demonstrated to have the universal approximation property [13], [14].

As outlined in [2], RBFN main advantages are that classification function is nonlinear, the model may produce confidence values and it may be robust to outliers; drawbacks are the potential sensitivity to input parameters, and potential overtraining sensitivity. The first-level classifier is then structured as a regular RBFN. In the second level of the classification stage, we introduce a modification of the standard use of RBFN. Its regular use in classification includes a hard decision on the output values: according to the winner-take-all rule, a given input pattern is assigned with the class corresponding to the winner output neuron which has the highest value. In our approach, we consider all values of the output neurons as a result of the classification task and we interpret them as gradual estimation of multi membership to classes.

V.RESULT

The classification of the messages as spam can be evaluated using extracted features by employing metrics such as precision, recall and F measures.



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 1, January 2014

TABLE IIIII
Result of Proposed Model (Level 1)

Metrics	Neutral	Non-neutral
Precision	98%	77%
Recall	89%	55%
F-Measure	93%	52%

TABLE IVV
Result of Proposed Model (Level 2)

Metrics	Violence	Vulgar	Sex	Offensive
Precision	82%	87%	68%	58%
Recall	88%	95%	71%	76%
F-Measure	93%	52%	81%	88%

The Result of the proposed system is categorized into two levels. The first level analysis the neutral and non-neutral content with the help of precision, recall, F-measure. The second level of result classifies the content into vulgar, sex, violence, offensive and etc., from the results it can be observed that the performance of the current system which includes content based features is better.

V.CONCLUSION

Online Social Networks (OSNs) is a medium to develop a strong social relationship between the people to share and communicate a part of life. The factors which affect the Online Social Networks (OSNs) is spam .Efficient Spam classification techniques are needed for effective usage of OSNs. In this work a spam classification method based on machine learning and content features has been implemented. From the results obtained, it is observed that the performance measures of spam classification such as precision, recall and F-measure have increased because of the combination of machine learning and content based features. The current work can be extended by using additional content features based on word combinations for better classification of messages.

REFERENCES

- [1] Sujapriya, S , G. Immanual Gnana Durai , Dr. C.Kumar Charlie Paul “Filtering Unwanted Messages from Online Social Networks (OSN) using Rule Based Technique”, *IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661, p- ISSN: 2278-8727*Volume 16, Issue 1, Ver. 1 (Jan. 2014), PP 66-70 www.iosrjournals.org www.iosrjournals.org .



ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 1, January 2014

- [2] A.K. Jain, R.P.W. Duin, and J. Mao, "Statistical Pattern Recognition: A Review," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 1, pp. 4-37, Jan. 2000.
- [3] C. Cleverdon, "Optimizing Convenient Online Access to Bibliographic Databases," Information Services and Use, vol. 4, no. 1, pp. 37-47, 1984.
- [4] F. Sebastiani, "Machine Learning in Automated Text Categorization," ACM Computing Surveys, vol. 34, no. 1, pp. 1-47, 2002.
- [5] M. Vanetti, E. Binaghi, B. Carminati, M. Carullo, and E. Ferrari, "Content-Based Filtering in On-Line Social Networks," Proc. ECML/PKDD Workshop Privacy and Security Issues in Data Mining and Machine Learning (PSDML '10), 2010.
- [6] N.J. Belkin and W.B. Croft, "Information Filtering and Information Retrieval: Two Sides of the Same Coin?" Comm. ACM, vol. 35, no. 12, pp. 29-38, 1992.
- [7] N.J. Belkin and W.B. Croft, "Information Filtering and Information Retrieval: Two Sides of the Same Coin?" Comm. ACM, vol. 35, no. 12, pp. 29-38, 1992.
- [8] Y. Zhang and J. Callan, "Maximum Likelihood Estimation for Filtering Thresholds," Proc. 24th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval, pp. 294-302, 2001.
- [9] B. Sriram, D. Fuhry, E. Demir, H. Ferhatosmanoglu, and M. Demirbas, "Short Text Classification in Twitter to Improve Information Filtering," Proc. 33rd Int'l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR '10), pp. 841-842, 2010.
- [10] J. Golbeck, "Combining Provenance with Trust in Social Networks for Semantic Web Content Filtering," Proc. Int'l Conf. Provenance and Annotation of Data, L. Moreau and I. Foster, eds., pp. 101-108, 2006.
- [11] J. Moody and C. Darken, "Fast Learning in Networks of Locally-Tuned Processing Units," Neural Computation, vol. 1, no. 2, pp. 281-294, 1989.
- [12] M.J.D. Powell, "Radial Basis Functions for Multivariable Interpolation: A Review," Algorithms for Approximation, pp. 143-167, Clarendon Press, 1987.
- [13] E.J. Hartman, J.D. Keeler, and J.M. Kowalski, "Layered Neural Networks with Gaussian Hidden Units as Universal Approximations," Neural Computation, vol. 2, pp. 210-215, 1990.
- [14] J. Park and I.W. Sandberg, "Approximation and Radial-Basis-Function Networks," Neural Computation, vol. 5, pp. 305-316, 1993.

BIOGRAPHY



Myself T. Suganya Thangavel doing my final year M.E (CSE) in Karpagam University, coimbatore. I have completed my B.E(CSE) in Anna University, Chennai in 2010. I have worked as a Senior Lecturer for 2years. Am a holder of yuva kala bharathi award for my academic skill. I have presented more than 2 national and international conference. My area of interest include data mining and mobile computing.