Detecting malicious users in the social networks using machine learning approach

Article *in* International Journal of Social Computing and Cyber-Physical Systems · January 2021

DOI: 10.1504/IJSCCPS.2021.10041246

U. Tanuja
Vidya Vardhaka College Of Engineering
6 PUBLICATIONS 17 CITATIONS
SEE PROFILE

Gururaj H L
Manipal Institute of Technology
101 PUBLICATIONS 290 CITATIONS
SEE PROFILE

Detecting malicious users in the social networks using machine learning approach

H.L. Gururaj*, U. Tanuja and V. Janhavi

Department of Computer Science and Engineering,
Vidyavardhaka College of Engineering,
Mysuru – 570002, India
Email: gururaj1711@vvce.ac.in
Email: tanuumashankar715@gmail.com
Email: janhavi.v@vvce.ac.in
*Corresponding author

B. Ramesh

Department of Computer Science and Engineering,
Malnad College of Engineering,
Hassan – 573202, India
Email: sanchara@gmail.com

Abstract: Social networking plays a very important role in today's life. It helps to share ideas, information, multimedia messages and also provides the means of communication between the users. The popular social medias such as Facebook, Twitter, Instagram, etc., where the billions of data are being created in huge volume. Every user has their right to use any social media and a large number of users allowed malicious users by providing private or sensitive information, which results in security threats. In this research, they are proposing an natural language processing (NLP) technique to find suspicious users based on the daily conversations between the users. They demonstrated the behaviour of each user through their anomaly activities. Another machine learning technique called support vector machine (SVM) classifiers to detect the toxic comments in the comments blog. In this paper, the preliminary work concentrates on detecting the malicious user through the anomaly activities, behaviour profiles, messages and comment section.

Keywords: social networks; malicious users; Naïve bayes; NLP; natural language processing; comments; social media; SVM; support vector machine.

Reference to this paper should be made as follows: Gururaj, H.L., Tanuja, U., Janhavi, V. and Ramesh, B. (2021) 'Detecting malicious users in the social networks using machine learning approach', *Int. J. Social Computing and Cyber-Physical Systems*, Vol. 2, No. 3, pp.229–243.

Biographical notes: H.L. Gururaj received his Bachelor's, Master's and Doctral degrees in Computer Science and Engineering from the Visveswaraiah technological university, Karnataka in 2010, 2013 and 2020, respectively. He is the author of more than 40 national/international journals and 26 national/international conferences. His research interests include computer networks, network security; cloud computing, IOT and machine learning. He is the member of IEEE and The Indian society for technical education. He awarded as Young Scientist international conference on internet of things,

data and cloud computing. Best project guide award 2015, 2016, and 2017 for the project entitled "An optimal TCP congestion control method for high speed data transfer in mobile Ad-hoc networks", "Intensifying the battery life and reliability of mobile devices using RMECR protocol reducing the power consumption", "Born baby synthesis". Best paper awards in various national and international conferences.

U. Tanuja received her BE in Computer Science and Engineering from the Government Engineering College, Hassan, in 2017 and she is pursuing her MTech degree in Vidyavardhaka College of engineering, Mysuru. Her research interests include computer networks and machine learning.

V. Janhavi received her BE in Computer Science and Engineering from the Visveswaraiah Technological University, Karnataka in 2006 and MS degree in Computer networks from DeVry University in 2010. She is the author of two national/international journals and one national/international conferences. Her research interests include computer networks, network security, information security and wireless networks. She is the member of ISTE.

B. Ramesh completed his BE in Computer Science and Engineering from Mysore University, Karnataka, India in 1991 and MTech degree in Computer Science from DAVV, Indore, Madhya Pradesh, India in 1995 and PhD degree from Anna University in 2009. Currently, he is working as Professor and the Head in the Department of Computer Science and Engineering at Malnad College of Engineering, Hassan, India. His current research interests lie in the areas of congestion control QoS-aware routing algorithms in ad hoc networks and multimedia networks.

1 Introduction

Social media has become one of the largest epidemics in the World. Some of the social networks like Facebook, Instagram, and Twitter etc… have become more popular means for communication. Social media is expanding & growing like something trending in the world (Conti et al., 2018). It allows the users to create their own profiles, communication between the user and sharing their information's, status, photos and videos. Social media application have become daily activities to the user and also helps in the business world, advertisement, journalism and these are all indirectly depending upon the user's opinions (Perez et al., 2017). This popularity of social media became black market that spoils the trust in between users .These social media applications become illegal services for the malicious user or the fake users through the daily conversations, comments, lives tagging and sharing URL's as depicted in Figure 1.

Figure 1 describes that the number of users depending upon the social media application. Now a days, one of the most usable social media applications, in which people rely on Facebook (Rahman et al., 2016). As per corporate social and environmental responsibility (CSER) survey, there are millions of fake accounts being created in social media applications. Some of the fake information is being spreading through un proper channels or links. We need to avoid all these unwanted accounts and harmful information from social media to save high volume space (Lescovec and Krevel, 2014). To avoid these fake accounts, we are implementing natural language processing (NLP) and text classification in between the texts. Another machine learning algorithm is used, i.e., Naïve bayes, which helps to identify the fake accounts based on user's friends list. Support vector machine (SVM) algorithm is implemented in between the comments section. As per a survey there are almost 800 million fake accounts in Facebook. It is very difficult to identify the malicious user through the status and posts. In this paper, we have some of the architecture and concepts which helps to identify the malicious users in social networks.

Figure 2 describes the proposed methodology for detecting malicious users in social networks using login authentication, comment section, chat between users using NLP and SVM classifiers.

Figure 1 Number of users depending on social media applications (see online version for colours)

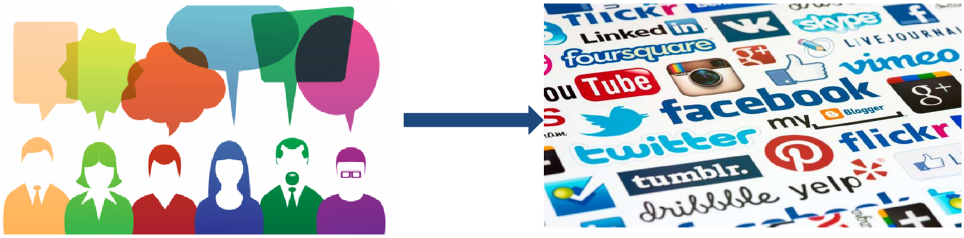
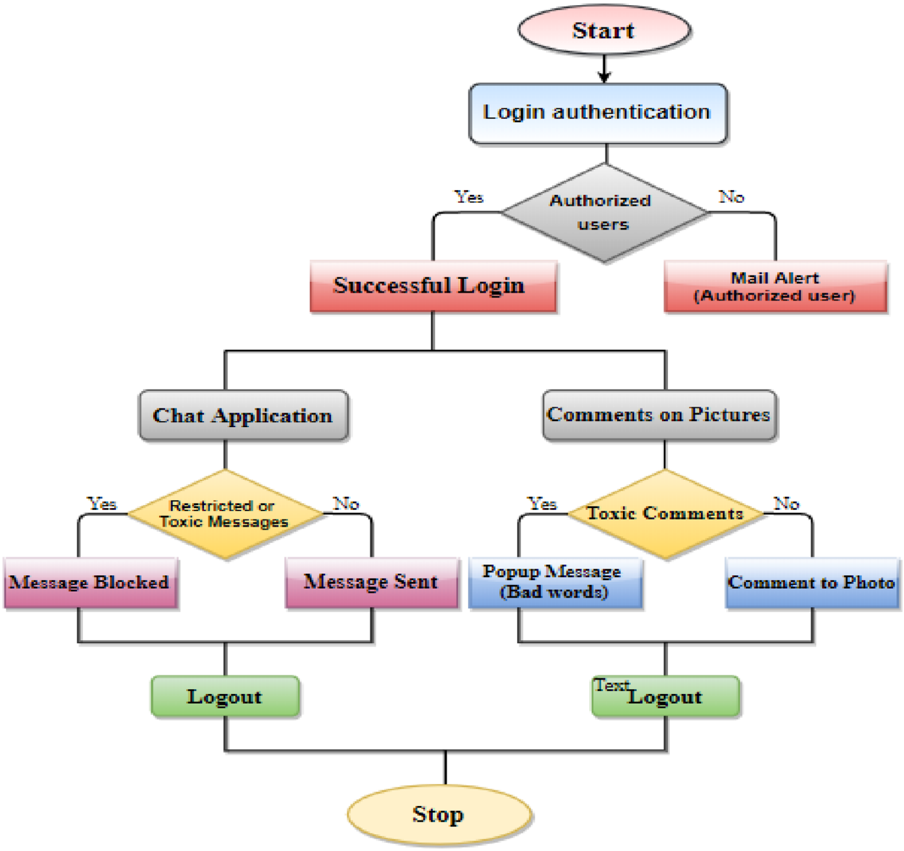


Figure 2 Processing of detecting malicious users in social network (see online version for colours)



The rest of the paper is systematised as follows: Section 2 describes the related work of social networks. Identifying the malicious users in social network based on login authentication, message characteristics, comments section and friends list is discussed in Section 3 and Section 4 provides the results of the proposed methodology. Finally the last section presents conclusion and future direction.

2 Related work

- *Facebook: detecting fake profiles in on-line social networks*: In this paper the preliminary work is to concentrate on detecting fake profiles in online social networks like Facebook. The proposed methodology in this paper is to detect fake profiles based on the following parameter are, Evolution over time of the number of OSN friends. Real social interaction. Evolution over time of the structure of OSN graph.
- *A dynamic approach to detecting suspicious profiles on social platforms*: In this paper they are proposed a method for detecting spam users in twitter by using an algorithm for scoring profiles dynamically in social networks.
- *Detecting malicious Facebook applications*: In this paper they are detecting malicious applications of Facebook; they are using two methods FRAppE Lite and FRAppE.
- *Classification of malicious and legitimate nodes for analysing the users' behaviour in heterogeneous online social networks*: In this paper, detecting of malicious users based on the behavioural profiles. They are taking datasets from the weka tool.
- *An Enhanced system to identify mischievous social malwares on Facebook applications*: This paper concentrate on identifying mischievous behaviour on social networks based on the datasets used in the Facebook application using classification algorithm and FRAppE.

3 Identifying malicious users in social network

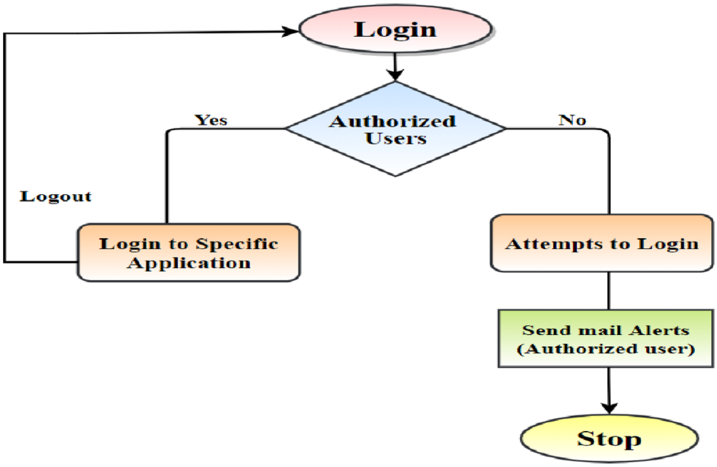
Social networks are an important part of today's internet and used by billions of people in the world wide. One of the most popular applications Facebook, can have access to the public information, posting the news feeds, access the friend's lists and access the data at any time. Identifying the malicious users in social network based on their anomaly activities in the application.

3.1 Login authentication and registration

In Social network registration and login process for each application is necessary for the security purpose. There are two factors for login authentication i.e., User name and Password. The username and password must be strong and the password must contain some special characters, letters and numbers (Dev et al., 2016). If your password is complex, then the malicious user/fake user cannot get any of the information regarding to the login authentication. It is a security feature given to the user, which helps to protect their profile. Registration process is required for using the social network applications, because the valid email id will be given for the verification process. If any malicious/fake user tried to attempt the user id, the original user will get the alerts to their mail.

Figure 3 represents the login authentication of the authorised users. If any of the malicious users tries to login, then the system will sends the alert message to the authorised users.

Figure 3 Processing of login authentication (see online version for colours)


Login
Authorized Users
Yes
No
Logout
Login to Specific Application
Attempts to Login
Send mail Alerts (Authorized user)
Stop

While registering to any of the social media application, need to give the authorised email id. It helps for securing the user's profile and it can also expose someone's hack activity. When malicious user/fake user trying to hack your social media profile, an alert message will sent to your email that someone is trying to hack your account (Egele et al., 2017; Vatrapu et al., 2016). In this paper, we are trying to alert the users, when the malicious user tries to hack your account. Number of attempts they try to hack your account; proportionally the number of messages will appear as a mail alerts.

Let 'x' be the number of attempts to login, if the first attempt to login then

$$p(x=n)=(1-p)^{n-1}p$$

The average number of attempts to login, then the expectation is,

$$\begin{aligned} x=E(x)&=\sum_{1}^{\infty}n(1-p)^{n-1}p \\ &=p\sum_{1}^{\infty}n(1-p)^{n-1}=p(1/p^2) \qquad \text{(i)} \\ &=1/p \end{aligned}$$

The above equation (i) shows the number of attempts to login. *P* is the probability that the number of times malicious user attempting to login the application.

3.2 Message characteristics

In Specific social networks, the malicious user can be detected based on the communication between the users. There are different categories included in messages.

(a) Message topic

Many of the social networks like Facebook, Instagram… etc., provides a chat application/text messages to communicate with each other. Users may send many messages that contain several information (Ramkumar et al., 2016). But we can also expect that the number of users frequently talk about some of their topics such as TV shows, Movies, Sports (Pasini et al., 2014; Sang and Xu, 2015). When the user suddenly make changes in their daily conversation or some of the messages they communicate with each other may not related to their daily conversation, this message can be rated as anomalous/fake messages.

Based on their daily conversation, the message which is unrelated to the topic may not be considered as anomalous, without any context is difficult. Some of the social networking application allows the users to post messages explicitly based on the topics which they are talked about (Bhise and Shishupal, 2018; Farasat et al., 2016). For example, there is a tagging mechanism using with hash character. If a user would use #Cricket to associate his or her tweet with cricket, it is also become more popular in Facebook. These techniques can be used to identify malicious users.

(b) Multimedia messages

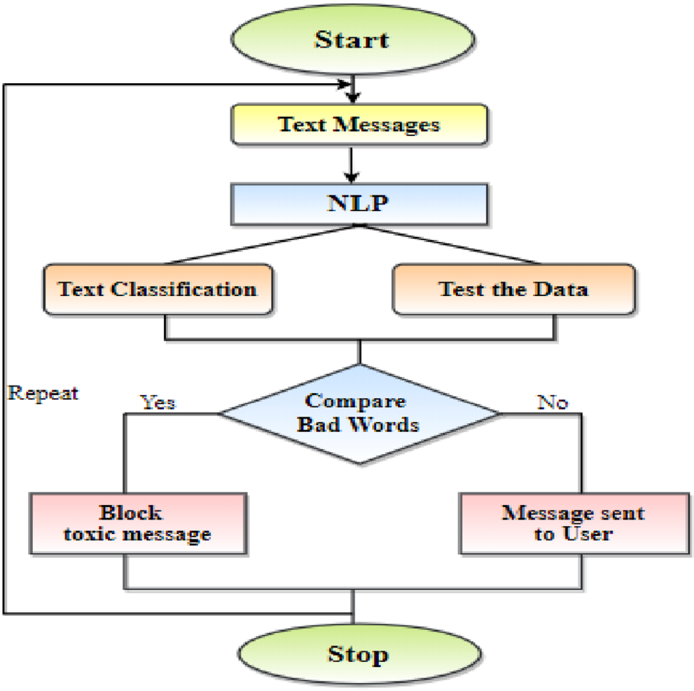
There are different categories included in the multimedia messages such as text messages, images, audio or video. Any of the user can interact with other user, these technique provided by the social network platforms. Based on the different types of bad images can considered as a malicious user. If the audio sent by the other user which may blackmail or other types of audio can be considered as a fake user (Gupta and Kaushal, 2016). Based on the videos or spam videos & spam links which is used by the user can be considered as malicious/fake user.

(c) Links in messages

In social networks there are some of the application, which includes links that are intended to harm, mislead the user. When a user click on to the URL pr links that triggered to rob the personal information of the users. These links helps to determine whether the messages are malicious or not (Swarna Sudha et al., 2018). These links can be posted as a message or can be used as comments, links which are posted as a message in the social network (Stieglitza et al., 2017), if any of the user clicks on to that link the malware will automatically downloaded & the personal information of the user will be hacked by the malicious user.

Figure 4 shows that how the restricted words will find out in the conversation between the users, If any of the words that matches with the trained datasets then that message will be blocked to both the users and considered as a malicious user.

Figure 4 Processing of text messages (bad words) using NLP method (see online version for colours)



3.2.1 Methodology

Social media applications become a primary tool for instant messaging. It plays a very important role in sharing personal information, emotional soothe and sensitive information (Egele et al., 2016). In this paper, we are proposing a method called natural language processing (NLP) for identifying the malicious users based on user's conversations. Natural language processing allows the machines to analyse text in real world application like text mining, automatic text summarisation, machine translation etc. In NLP we are focusing on word tokenisation and prediction, it implicates in breaking a sentences into individual words. Then the next word in which the user is typed can be predicted by using N-gram model. N-gram model finds the probability score of each word in a sentence. In addition we can also observe the whole word sequence in a sentence by using chain rule probability,

$$\begin{aligned} P\left(Y_1 \ldots\ldots Y_n\right) &= P\left(Y_1\right) P\left(Y_2 \mid Y_1\right) P\left(Y_3 \mid Y_1^2\right) \ldots\ldots P\left(Y_n \mid Y_1^{n-1}\right) \\ &= \prod_{k=1}^{n} P(Y_k \mid Y_1^{k-1}) \end{aligned} \tag{ii}$$

Now, applying the chain rule to words in equation (ii), we get

$$\begin{aligned} &P\left(W_1 \ldots\ldots W_n\right) = P\left(W_1\right) P\left(W_2 \mid W_1\right) P\left(W_3 \mid W_1^2\right) \ldots\ldots\ldots P\left(W_n \mid W_1^{n-1}\right) \\ &P\left(W_1^n\right) = \prod_{k=1}^{n} P(W_k \mid W_1^{k-1}) \end{aligned} \tag{iii}$$

where W_1^n is the present word and W_k is the *K*th instance of a word. The equation (iii) shows that, we can find out the probability of sequence of words in sentence. In this rule sequence of words (bad words) compared with the trained datasets, if it is matched then the message will be blocked. In this paper, we are using maximum likelihood estimation (MLE) to find out the probability of individual words. The probability of a particular word in N-gram computed by equation:

$$P(W_n \mid W_{n-1}) = \frac{C(W_{n-1}W_n)}{\sum_w C(W_{n-1})} \tag{iv}$$

Generalisation of MLE n-gram can be written as,

$$P(W_n \mid W_{n-N+1}^{n-1}) = \frac{C\left(W_{n-N+1}^{n-1} W_n\right)}{C\left(W_{n-N+1}^{n-1}\right)} \tag{v}$$

The above equation (v) finds out the n-gram probability by dividing the sequence of words (bad words) in a sentence. In MLE, the data resulting parameter of words likelihood of trained datasets *T* in a given model *M*, i.e., *P*(*T*/*M*).

3.3 Comment section

In social network the comment section also plays an important role, because each and every user is curious about the comments for their posts. Many of the news websites, blogs, social medias also provides the comments section to the users to comments on their posts or published content (Xin et al., 2018). In social media applications there are different types of comments blogs are provided such as videos, photos, tagged photos, news websites, status, profile pictures etc, each & every users of social media has a rights to comment. There are two types of comments included, Gated comments sections will post some of the information in their website, then the user will post a comment regarding the information (Tanuja et al., 2019). Non-Gated comments section also provides information and it allows the user/people to post the comments and discussed about the information according to their view points.

In social media toxic comments are also includes, that can harm the original users. Toxic Comments are bad words, some of the bad emoji's and GIF's based on these comments we can identify malicious user or fake user.

Figure 5 describes the processing of the comments in the comments blogs, finding the malicious users based on the toxic comments. If any toxic comments appears in the comments blog first it classify and compared with trained datasets using SVM classifiers, if any toxic comments appears pop message will be displayed as bad words.

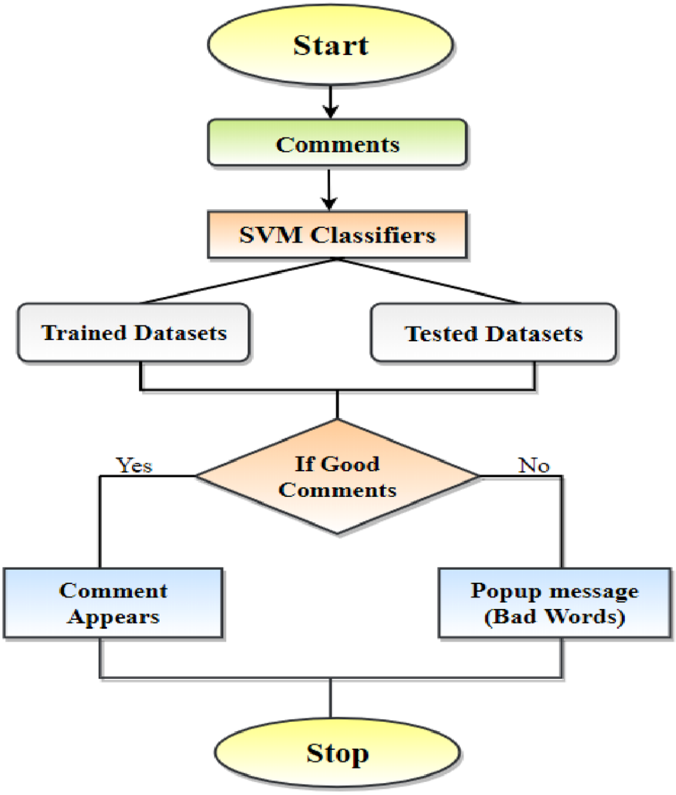
3.3.1 Methodology

In this paper, we are emphasis on SVM for identifying restricted comments in social media application. SVM is a supervised classification algorithm which is used for categorising the documents or data (Torkyl et al., 2018). The main goal of SVM algorithm is to generate the decision boundary that can separate input and output by a hyperplane. To classify the document, we are focusing on linear SVM. It is used for segregating the data into two classes by using straight line. Here the word frequencies and the vectors are classified in multidimensional space. The equation of SVM:

$$f(W) = \frac{\|W\|^2}{2} + C\left(\sum_{i=1}^{n} \zeta i\right) \quad \text{(vi)}$$

Where in equation (vi) W represents the weight of the comments and *C* represents the penalty for misidentifying the data. The above equation (vi) of SVM helps to identify the number of restricted words in a comment section which are harmful. The restricted words are collected as a dataset and divide it into test and training dataset. If any harmful words appear in comment section then SVM algorithm compares it with trained datasets and blocks that comment. In this process, SVM also calculate the number of restricted words appears in individual account and warning them to provide authentication proofs.

Figure 5 Describes the processing of the comments using SVM classifiers (see online version for colours)



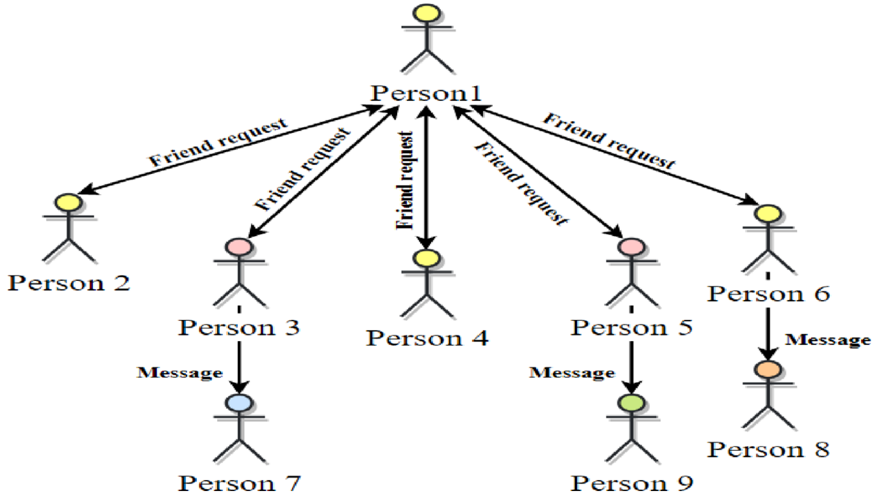
3.4 *Friends list*

In social network, detecting of malicious based on their friends lists. The users should be more careful while accepting request from the unknown users. Identifying malicious users through their anomaly activities such as searching for different celebrities profiles, girl's profile (Xiao et al., 2016; Hassan et al., 2017. Sending requests only for the girls.

Make a group of friends, posting the toxic comments to the user's profiles. For making friends through online is for sometimes is good, but don't try to make a friend who is unknown (Ivaschenko et al., 2018; Amato et al., 2016). Do not make friends for the sake of gaining more followers, like and comments. Malicious users may harm you or blackmailed you through your personal information. Malicious users can also create fake profiles through the information in social websites. These are the anomaly activities of malicious users.

Figure 6 represents the network of single user who can have more number of friends and can communicate with each other.

Figure 6 Network of single user in social network (see online version for colours)



3.4.1 *Methodology*

In this work Naïve bayes classifier is used to identify the malicious users based on friends list. In Naïve bayes classifier, there are multiple features or behaviour of friends say $B1, B2, \ldots, Bi, \ldots, Bn$ for $1 \le i \le n$ and friends list say $F1, F2, \ldots, Fj, \ldots, Fk$ for $1 \le j \le k$, where Fj is the class of friends and Bi is the behaviour that belong to all friends list classes. Since Naïve bayes classifiers calculate the conditional probability by classifying the person's behaviour belonging to a particular friend list class Fj is given by the equation:

$$P(Fj \mid B1, B2, \ldots, Bn)\,\alpha\left(\prod_{j=1}^{n} P(bi \mid Fj)\right)\left(P(Fi)\right) \text{ for } 1 \le j \le k \qquad \text{(vii)}$$

Naïve bayes classifier helps to determine the probabilistic model as shown in equation (vii). In this work, we are taking only two classes: good and bad person. Then the decision rule is:

$$P(Fj \mid B1, B2, B3, \ldots, Bn) \ \ge\ P\left(\overline{Fj}\,\middle|\,\overline{B1}\,\middle|\,\overline{B2} \mid \overline{B3}, \ldots, \overline{Bn}\right)$$

We assume the cut off probability will be 80% to determine the person either good or bad. It helps to measure the probability of the friend's list classes.

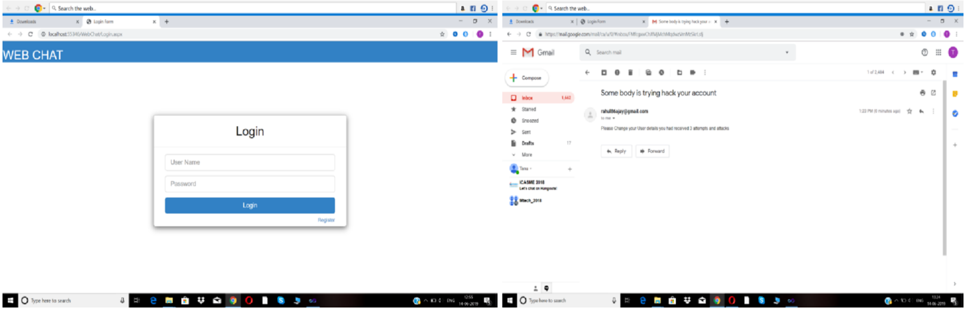
4 Results

The below snapshots describes the comment section of the web chat application, all the users have a rights to comment to the picture blogs, video blogs, news updates (Vigneshwari and Aramudhan, 2015). Based on these comments we can find the malicious users in the social media. If a person or user comment is bad, then the popup message will be displayed to them.

All the snaps describe the proposed methodology of this paper. We took four parameter to identify the malicious users in social networks. It is one of the web chat application like Facebook. Login authentication, registration form, comments section, uploading pictures, finding friends, friend requests, chat with others are the parameters used in this web chat application (Stein et al., 2018; Baltazar et al., 2009). Detecting malicious users based on the daily conversation of the users if the bad words appear it will block that message. In the comment section, if the toxic comments appear popup message will be displayed to the user.

Figure 7 represents the login page to the web application. If the authorised user login they continued to use web services otherwise considered as a malicious users. When an unauthorised users attempt to login to others the profile, mail alerts to authorised users.

Figure 7 Login page and mail alerts of the chat application (see online version for colours)



Some body is trying hack your account

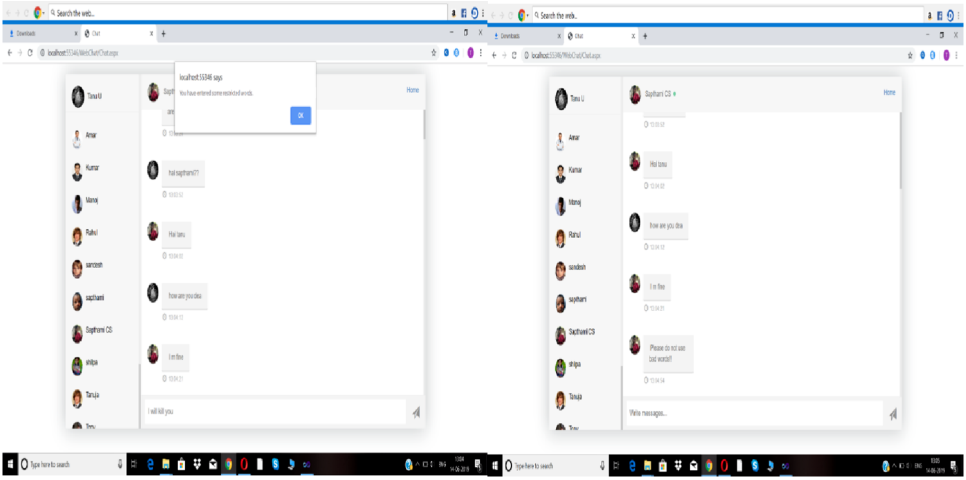
rahul86ajay@gmail.com

to me

1:23 PM (0 minutes ago)

Please Change your User details you had received 3 attempts and attacks

Figure 8 Shows the chating between the users (see online version for colours)



localhost:55346 says

You have entered some restricted words.

hai sapthami??

13:03:52

Hai tanu

13:04:02

how are you dea

13:04:12

I m fine

13:04:21

I will kill you

Sapthami CS

Please do not use bad words!!

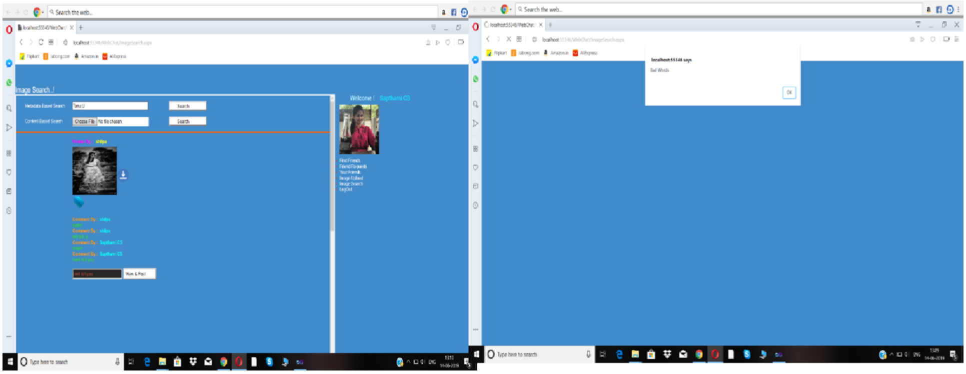
13:04:54

Write messages...

Figure 8 shows that communication of users in this webchat application like Facebook, Twitter, Instagram, etc. Detecting malicious users based on the restricted words used in their daily conversations. If the bad words used by the users a popup message will appears and the message doesnot appears to both of the users but the message “donot use bad words!!!” will displayed to the users.

Figure 9 represents the comment section of the web chat application. SVM classifier is used to block the toxic comments in this application. SVM classifiers compared the toxic comments with trained datasets, if it is matched then the comment will be blocked and popup message will be displayed.

Figure 9 Shows that the comment section of the webchat application (see online version for colours)



Posted By : shilpa

Comment By : shilpa
super
Comment By : shilpa
will kill u
Comment By : Sapthami CS
super
Comment By : Sapthami CS
I will kill you

localhost:55348 says
Bad Words

Figure 10 shows the number of active users in the social network. The graph represents that the total population, internet users, active social users and active mobile social users.

Figure 10 Comparison of active social media users (see online version for colours)

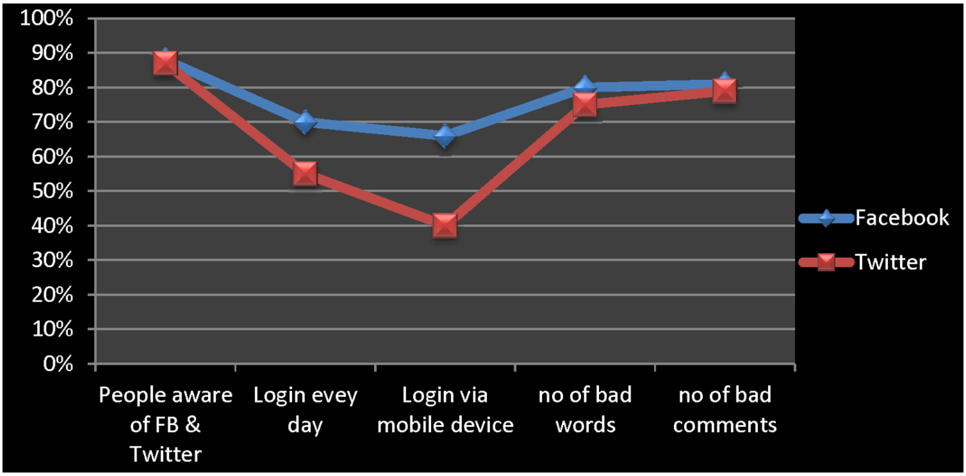


Figure 11 represents the number of malicious users in social networks. The graph describes the number of increasing malicious users in social networks.

Figure 11 Number of malicious users in social networks (see online version for colours)

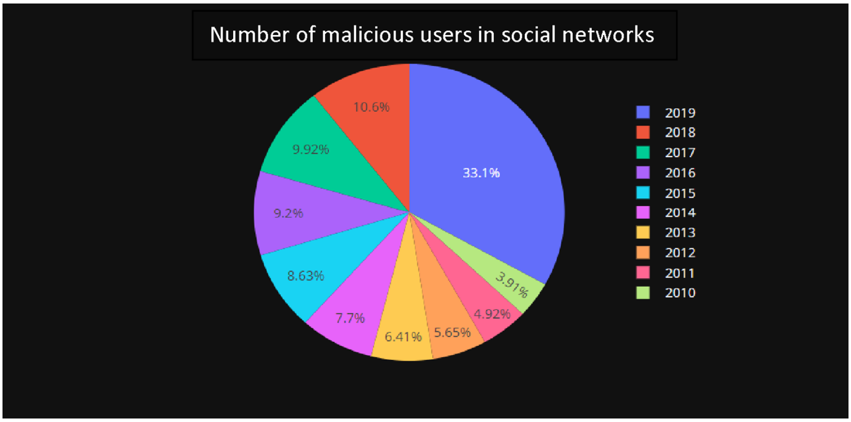


Figure 12 represents the number of active members in diffferent social medias like facebook, whatsapp, instagram, twitter, snapchat, telegram, etc.

Figure 12 Number of active members in different social medias (see online version for colours)

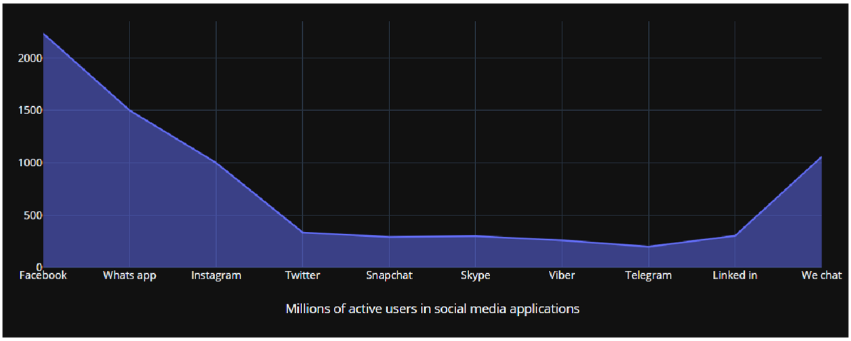
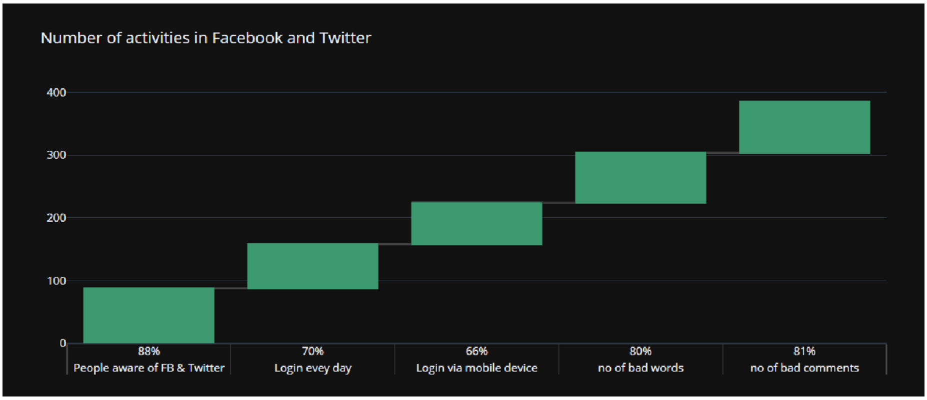


Figure 13 represents the daily activities in social networks. The above graph depicts the activities (login, no of bad word and comments) in Facebook and twitter.

Figure 13 Number of activities in Facebook and twitter (see online version for colours)



5 Conclusion and future direction

Social networks become an important communication media in today's society. The proposed methodology in our paper indicates that there is a mechanism to identify the malicious users from login authentication, i.e., number of attempts to login equal to number of mail alerts to the authenticated user. Using NLP technique to determine malicious users based on the bad words they use in their daily conversation. The SVM algorithm is used to classify the toxic comments in the comment section. SVM splits the data into training and testing data, and then the present data is compared with training and testing data for prediction. If the predictive results matched then it blocks that data. The Naïve bayes classifier is used to detect and accurately predict the malicious users through their friends list. Our simulation results demonstrate the detection of malicious users in social media applications. As a future work, this research can be extended to improve other social networks also. Our proposed methodology is extended for the future privacy in the social network.

References

Amato, F., Moscato, V., Picariello, A. and Sperl, G. (2016) *Multimedia Social Network Modeling: A Proposal*, IEEE, Laguna Hills, CA, USA.

Baltazar, J., Costoya, J. and Flores, R. (2009) 'KOOBFACE: the largest Web 2.0 Botnet explained', *Trend Micro Research*, Vol. 5, No. 9, pages 10.

Bhise, K. and Shishupal, R.S. (2018) *A Method For Recognize Malignant Facebook Application*, IEEE, Greater Noida, India.

Conti, M., Poovendran, R. and Secchiero, M. (2018) *FakeBook: Detecting Fake Profiles in On-line Social Networks*, IEEE, Istanbul, Turkey.

Dev, P., Singh, K. and Dhawan, S. (2016) 'Classification of malicious and legitimate nodes for analysing the users' behaviour in heterogeneous online social networks', *International Conference on Futuristic Trend in Computational Analysis and Knowledge Management*, Greater Noida, India, pp.359–363.

Egele, M., Stringhini, G., Kruegel, C. and Vigna, G. (2016) 'COMPA: detecting compromised accounts on social networks', presented at the *Network and Distributed System Security Symp.*, February, San Diego, CA, USA, pp.1–17.

Egele, M., Stringhini, G., Kruegel, C. and Vigna, G. (2017) 'Towards detecting compromised accounts on social networks', *IEEE Transactions on Dependable and Secure Computing*, Vol. 14, No. 4, July–August, pp.447–460, doi: 10.1109/TDSC.2015.2479616.

Farasat, A., Gross, G., Nagi, R. and Nikolaev, A.G. (2016) 'Social network analysis with data fusion', *IEEE Transactions on Computational Social Systems*, Vol. 3, No. 2, June, pp.88–99, doi: 10.1109/TCSS.2016.2613563.

Gupta, A. and Kaushal, R. (2016) *Towards Detecting Fake User Accounts in Facebook*, IEEE.

Hassan, A.U., Hussain, J., Hussain, M., Sadiq, M. and Lee, S. (2017) 'Sentiment analysis of social networking sites (SNS) data using machine learning approach for the measurement of depression', *2017 International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju, Korea (South), pp.138–140, doi: 10.1109/ICTC.2017.8190959.

Ivaschenko, A., Khorina, A., Isayko, V., Krupin, D., Bolotsky, V. and Sitnikov, P. (2018) *Modeling of User Behavior for Social Media Analysis*, IEEE, Moscow, Russia.

Lescovec, J. and Krevel, A.J. (2014) *SNAP Datasets: Stanford Large Network Dataset Collection*, [Online], Available: http://snap.stanford.edu/data (Consult June 2015).

Pasini, C., Tagliasacchi, M., Fraternali, P. and di MilanoMilan, P. (2014) *histoGraph – A Visualization Tool for Collaborative Analysis of Networks from Historical Social Multimedia Collections*, IEEE, Paris, France.

Perez, C., Lemercier, M. and Birregah, B. (2017) *A Dynamic approach to Detecting Suspicious Profiles on Social Platforms*, IEEE, Budapest, Hungary.

Rahman, S., Huang, T-K., Madhyastha, H.V. and Faloutsos, M. (2016) *Detecting Malicious Facebook Applications*, IEEE, Madurai, India.

Ramkumar, G., Vigneshwari, S. and Roodyn, S. (2016) 'An enhanced system to identify mischievous social malwares on Facebook applications', *International Conference on Circuit, Power and Computing Technologies*, Nagercoil, India pp.1–5.

Sang, J. and Xu, C. (2015) 'On analyzing the 'variety' of big social multimedia', *IEEE International Conference on Multimedia Big Data*, Beijing, China, pp.5–8.

Stein, T., Chen, E. and Mangla, K. (2018) 'Facebook immune system', *Proceedings of the 4th Workshop on Social Network Systems*, ACM, Europe, pp.1–8.

Stieglitza, S., Mirbabaiea, M., Rossa, B. and Neubergerb, C. (2017) 'Social media analytics – challenges in topic discovery, data collection, and data preparation', *International Journal of Information Management*, Elsevier, Vol. 39, April, pp.156–168.

Swarna Sudha, M., Arun Priya, K., Kanaka Lakshmi, A., Kruthika, A., Lakshmi Priya, D. and Valarmathi, K. (2018) *Data Mining Approach for Anomaly Detection in Social Network Analysis*, IEEE, Coimbatore, India.

Tanuja, U., Gururaj, H.L. and Janhavi, V. (2019) 'A machine learning algorithm for classification, analyzation and prediction of multimedia messages in social networks', *Proceedings of First International Conference on Computing, Communications, and Cyber-Security (IC4S 2019)*, Springer, Singapore, pp.485–499.

Torkyl, M., Meligy, A. and Ibrahim, H. (2018) *Recognizing Fake Identities in Online Social Networks Based on a Finite Automaton Approach*, IEEE, Cairo, Egypt, pp.1–7.

Vatrapu, R., Mukkamala, R.R., Hussain, A. and Flesch, B. (2016) 'Social set analysis: a set theoretical approach to big data analytics', *IEEE Access*, Vol. 4, pp.2542–2571, doi: 10.1109/ACCESS.2016.2559584.

Vigneshwari, S. and Aramudhan, M. (2015) 'Personalized cross ontological framework for secured document retrieval in the cloud', *National Academy Science Letters-India*, Vol. 38, No. 5, pp.421–424.

Xiao, C., Freeman, D.M. and Hwa, T. (2016) *Detecting Clusters of Fake Accounts in Online Social Networks*, IEEE, Denver, Colorado, USA.

Xin, Y., Zhao, C., Zhu, H. and Gao, M. (2018) *A Survey of Malicious Accounts Detection in Large-Scale Online Social Networks*, IEEE, Omaha, NE, USA.