International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2017 IJSRCSEIT | Volume 2 | Issue 2 | ISSN : 2456-3307

Defending Mechanism for Cyber Bullying

Naveen Kandlapalli, Shobha Shinde, Priyanka Shriramoji, Pooja Uke, Prof. Supriya Chaudhary

Department of Information Technology, PVPPCOE, Mumbai, Maharastra, India

ABSTRACT

The popularity and wide growth of the social networking sites over the communication world has resulted in its tremendous use. By using social networking sites, people are connected to each other in the world, usually they express their feelings, opinions and emotions. Popularity of the social networking sites cause major rise in offensive behavior, giving birth to one of the most critical problem called Cyberbullying and Online Grooming. The victims of Cyberbullying, broadly being the youngsters, undergo a deep scar which has led to suicidal attempts in many cases. Online harassment has become a common problem where the youngsters are highly targeted. Cyberbullying and Online Grooming are one of the attacks on social networking websites. Cyberstalking, child pornography, online sexual predators, sexual solicitation of children and compromising text and images with violent or sexual content are happening in online grooming attacks. Agenda for a defending mechanism for Cyberbullying is to detect and identify the above mentioned threats and safeguard the users of the social networking websites. Threat indications are determined by image analysis, social media analysis and text mining techniques in order to raise alertness about enduring attacks and to grant backing for further actions.

Keywords : Social Networking, Online Grooming, Cyberbullying, Facebook, Text Analysis, Image Analysis

I. INTRODUCTION

Web applications 2.0, especially social networking websites like Facebook and Twitter are gaining widespread attention due to tremendous increase and use of technology. A study reveals that Facebook has majority of active users and has reportedly croossed the mark of 1 billion active users since 2013 . An important aspect which should be considered in the context of social networking websites and its usage are the negative experiences which are faced by a lot of people. Social networks exposes young people to new threats and perils such as Cyberbullying and Online Grooming. According to the 'National Crime Prevention Council (NCPC)', the threats of Cyberbullying attack a large section of the younger generation. Cyberbullying is an attack based on deliberately insulting, threatening, embarrassing or harassing people on the internet or via mobile phones. Even though it may not take place in person, the emotional and psychological effects of cyber-bullying are just as destructive as physical and verbal bullying. Information such as private images, sexually abusing messages and photoshoped images are shared on online platforms and therefore it deeply affects the victims. The victims of Cyberbullying undergo deep mental trauma and develop a sense of isolation from the society. Once the information is shared online, the data spreads immediately within a few minutes and a large section of people access the same. Once published, private information is spread and it is more or less impossible to delete it, because of the pervasiveness of digital media. Even if it is possible to remove information it still remains in the minds of readers. Bullied kids are likely to experience anxiety, depression and unhappiness. Many a times victims shy away from sharing the trauma of Cyberbullying due to fear of embarassment. More often than not, victims respond passively to bullying. They tend to act anxious and appear less confident. Techniques used by perpetrators vary. Besides slanderous text messages, posts or comments, video and photo functions are also used in social networks for cyberbullying attacks. Therefore, compromising pictures and videos with violent or sexual content, such as Happy Slapping or Sexting, can be shared. Happy Slapping videos show

disputes and scuffles between teenagers which have been filmed and distributed over the internet. Sexting refers to young people, taking photos of themselves or others (voluntarily or under coercion) and publishing these pictures for a large audience.

For the protection of children against such hazards, research efforts for eradication of Cyberbullying has to be implemented.. Therefore, the major contribution of this paper is the introduction of a research agenda paving the way for detecting and preventing grooming and bullying activities. The aim is to establish the foundation for the development of an automatic approach which can be ideally integrated into a social networking website based on text mining, image analysis and social media analytics. By implementing this analytical techniques, one can feel safe and not be bullied in the world of online social networking.

II. EXISTING SYSTEM

The existing systems in the world of online social media include websites such as Facebook and Twitter which are used on a large scale. They provide features and functions such as posting comments and sharing pictures on their time line. Personal chats between two users are also provided which enables people to share private information among each other. However, with the tremendous evolution of social media and its role in everyday life, it has also resulted in serious life threatening crimes especially among the youth. Some important points of Existing System are mentioned below:

- Report/Spam user post.
- Messages can be sent and received across the world.
- Saving and E-Mailing Chat with the date and time.
- Manual action in case of unethical and unacceptable behavior.

The existing system provides a positive aspect of sharing images and posting comments. However, it has been observed that majority of the users are using it for their sexual needs and as a tool for online harassment. The drawbacks of the existing system are as follows:

- Posting offensive and slang language on their time line.
- Sharing of nude pictures of unknown people.
- Serious crimes and online trafficking.
- Impersonation and false identity.
- Blocking of an abusive and offensive image after it has been posted.
- Reporting it to the authentic center after the image which has been posted.

III. PROPOSED SYSTEM

A combination of techniques like image analysis and text analysis will be used in order to curb the Cyberbullying attacks. Initially, a user will be provided an account registration web page where account registration will be carried out. It is necessary to register to use the features and functions of the online social networking website. Once this is completed, the user will be provided features such as searching a friend in the search tab, sending friend request to a particular registered user, adding images on the time line, posting comments and messages in chats. The user will also be provided with different categories of interests such as entertainment, news, politics, sports, cultural, environmental, science and technology etc. The architecture of the Proposed System is shown in Fig. 1

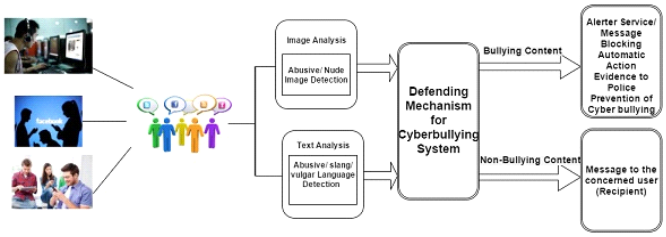

Figure 1. Proposed System Architecture

The proposed system will highly concentrate on the behavior of the user. It will keep a check on the content of the images and text which has been posted by a user. The system will also keep a track of the user actions which has been done by the user. If the system encounters an action where a user posts an abusive image/ pornographic image/ offensive image, the user will be intimated for such an action. This will hence block the posted image and a copy of the action will be saved in the database. Similarly, the system will be developed for the detection of abusive/ slang/ vulgar language. Subsequently, such text input will also be blocked. The State Transition Diagram of the Proposed System is as shown in Fig. 2 and the Activity Diagram of the Proposed System is as hown in Fig. 3.

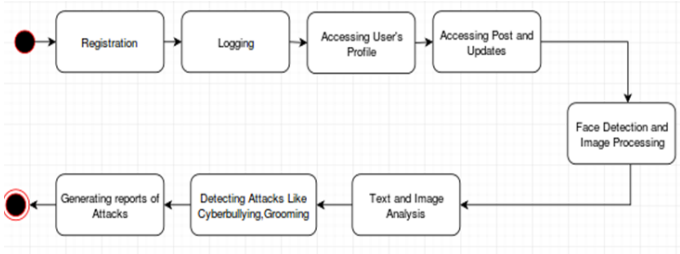

Figure 2. State Transition Diagram of the Proposed System

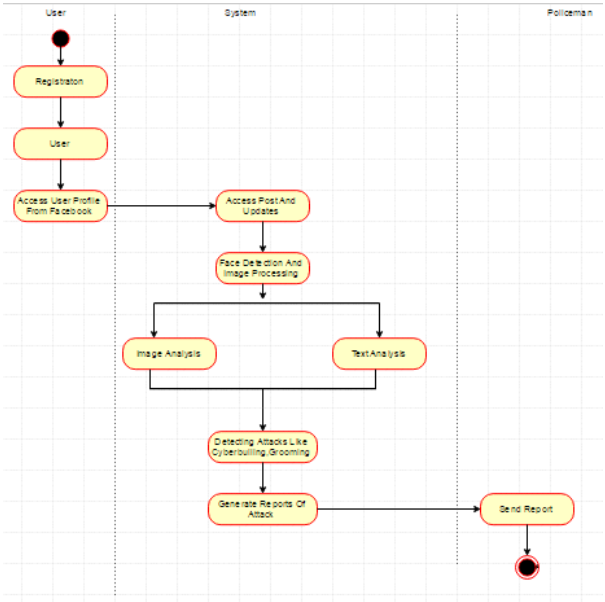

Figure 3. Activity Diagram of the Proposed System

IV. IMPLEMENTATION AND DESIGN

In order to implement the above discussed Proposed System, a combination of techniques for image and text analysis has been carried out for abusive image detection and abusive text detection respectively.

Abusive Image Analysis:

Image analysis is essential to be carried out to stop online Cyberbullying. Most of the online social networking websites allow sharing of pictures and hence abusive or nude pictures are posted many times by the offenders. Therefore, techniques are used to detect the type of the image - abusive or not? Skin color detection technique is used here to classify the image. The image is initially fetched in terms of the pixel value. And total number of pixels are found. The value of every pixel is found out using the cordinates (x,y) and also the RGB value. The skin tone values are also stored in the RGB. Skin color delta value is set to 50 and Abusivedelta value is set to 5. The mathematics associated is as follows:

The delta value of current pixel:

r1, g1, b1 = Skin Tone value

r2, g2, b2 = Pixel value

If delta value of the current pixel is less than or equal to the Skin color delta value, then increase the Skin Tone vale by using the following formula:

sqrt of $\{(r1-r2)^2, (g1-g2)^2, (b1-b2)^2\}$

Lastly, the threshold value is calculated:

Threshold Value = (Skin Tone value / Total Pixel value X 100)

The threshold value is compared with the AbusiveDelta value and if there are more number of skin colored in the image, then the image is blocked.

Detection of abusive images is also carried out by using the concept of BoVW model and classifier. Initially, the important points from the image are detected using the Local Binary Pattern (LBP). The centre pixel is labelled against the threshold of the neighbouring pixels. Every pixel in the cell is compared with the neighbouring pixel. When the center pixel's value is greater than the neighbor's value, the value is 1 and otherwise it is 0. Thus an 8 digit binary number is obtained. Calculation of Histogram and after its normalizatio, the concatenation is carried out for every cell. This gives the feature vector for the input image using local binary pattern (LBP).

Visual vocabulary - (BoVW) model:

BoVW (Bag of visual word) model represents the histogram with independent features and is also applied in image classification. Since each image is represented by local patches, numerical vectors are generated accordingly to represent the patches. After the processing is carried out, the image is a collection of vectors where all vectors have same dimension and order of vectors is not important. Ultimately, the BoVW functions by mapping the vector represented patches to that of visual words which belongs to a particular visual vocabulary. A numeric digit is associated with the visual words that represents the visual vocabulary size. Visual vocabulary has plenty of visual words where a particular visual word acts as a representative of sevral image patches which have similar resemblance. can be considered as a representative of several similar image patches. Thus, each patch in an image is mapped to a particular

visual word and the K-means clustering technique is used for mapping. The image can be represented in terms of the histogram of the related visual words of a image. Finally, a classifier classifies the image and tells whether it is abusive or not?

Abusive Text Analysis:

The techniques for abusive text detection categorize the text input by the user into abusive category and non-abusive category. The initial stage here is of pre-processing. The text input is processed so as to make it into a uniform format which can be understood by the learning algorithms. Spell corrrection algorithms are used to remove meaningless symbols, correction of words etc. Once this is completed, feature extraction is the next important task. The text is represented in terms of a feature vector because it has to be used by the classifier. The technique of BoW (Bag of Words) is used here to create a bag of words which are offensive and abusive. This Bag can be fed with all the abusive words. In this approach, every text message is represented as a feature vector comprising of binary attributes for every word which is occuring in the message. Let {w1,…,wm} represent a set of m feature vector (vocabulary of words) which appears in a message. Let ni(d) represent the number of times wi occurs in a message d. Thus every message is represented by the message vector d = (n1(d), n2(d),…, nm(d)). If the word input matches with that of the vocabulary bag, then it's attribute is set to 1, else it is set to 0. Finally, Bernoulli Naive Bayesian model is used to classify where the extracted features are represented as independent Boolean.

The results of both text analysis and image analysis is passed on to the Boolean system. The Boolean system opearates only in 0 and 1 and decides the bullying contents. Boolean system signals indicate true for bullying contents accordingly.

V. CONCLUSION AND FUTURE SCOPE

The Proposed Sytem has been achieved by using the above discussed techniques of image and text analysis. These techniques will categorize the text and image input of the user as abusive or non-abusive. Thus, by blocking the bullying content, the Cyberbullying attacks can be curbed to a large extent. This will also safeguard the users of the online social networking webistes from life threatening situations and abusive offenders. Other than just text and image analysis, sound and video analysis can be carried out in future. It can also be deployed on apps and can be trained to detect abusive videos in future.

VI.REFERENCES

[1] Margaret Anne Carter,Third party observers witnessing cyber bullying on social media sites, Procedia - Social and Behavioral Sciences 84 (2013) 1296 – 1309 www.s ciencedirect.com.

[2] Paridhi Singhal and Ashish Bansal, Improved Textual Cyberbullying Detection Using Data Mining, International Journal of Information and Computation Technology. ISSN 0974-2239 Volume 3, Number 6 (2013), pp. 569-576,© International Research Publications House, http://www. irphouse.com /ijict.htm 2013.

[3] Charles E.Notar *, Sharon Padgett, Jessica Roden, Cyberbullying: A Review of the Literature http://www.hrpub.org, Universal Journal of Educational Research 1(1): 1-9, 2013 DOI: 10.13189/ujer.2013.010101

[4] Dinakar, K., Jones, B., Havasi, C., Lieberman, H., and Picard, R., Common Sense Reasoning for Detection, Prevention, and Mitigation of Cyberbullying,ACM Trans. Interact. Intell. Syst. 2, 3, Article 18(September 2012), 30 pages. DOI = 10.1145/2362394.2362400, http://doi.acm.org/10.1145/2362394. 2362400, 2012.

[5] NaliniPriya. G and Asswini. M.,―A Dynamic Cognitive System For Automatic Detection And Prevention Of Cyber-Bullying Attacks, ARPN Journal of Engineering and Applied Sciences ©2006-2015 Asian Research Publishing Network (ARPN). VOL. 10, NO. 10, JUNE 2015

[6] Marco Vanetti, ElisabettaBinaghi, Elena Ferrari, Barbara Carminati, Moreno Carullo: A system to filter unwanted messages from OSN user walls ― Department of Computer Science and Communication ,University of Insubria 21100 Varese, Italy, 2013