An Effective Approach for Cyber intimidate Detection

G.Archana Assistant Professor in Department of Information Technology in Teegala Krisha Reddy Engineering college.
P.Sanjana UG Scholar in Department of Information Technology in Teegala Krisha Reddy Engineering college.Telangana
K.Sai Teja UG Scholar in Department of Information Technology in Teegala Krisha Reddy Engineering college.Telangana
D.Shirisha UG Scholar in Department of Information Technology in Teegala Krisha Reddy Engineering college.

Abstract: **:** The rapid growth of social networking is supplementing the progression of cyber intimidate activities. Most of the individuals involved in these activities belong to the younger generations, especially teenagers, who in the worst scenario are at more risk of suicidal attempts. We propose an effective approach to detect cyber intimidate messages from social media through a weighting scheme of feature selection. We present a model to extract the cyber intimidate network, which is used to identify the most active cyber intimidate predators and victims through ranking users.

Keywords- *Social Networks; Cyber*intimidate*; Text-Mining.*

I. INTRODUCTION

With the proliferation of the Internet, cyber security is becoming an important concern. While Web 2.0 provides easy,interactive, anytime and anywhere access to the online communities, it also provides an avenue for cybercrimes like cyberintimidate . A number of life threatening cyberintimidate experiences among young people have been reported internationally,thus drawing attention to its negative impact. In the USA, the problem of cyberintimidate has become increasingly evident and it has officially been identified as a social threat . There is an urgent need to study cyberintimidate in terms of its detection, prevention and mitigation.Traditional intimidate is any activity by a person or a group aimed at a target group or individual involving repeated emotional, physical or verbal abuse. Intimidate as a form of social turmoil has occurred in various forms over the years with the WWW and communication technologies being used to support deliberate, repeated and hostile behaviour by an individual or group, in order to harm others . Cyberintimidate is defined as an aggressive, intentional act carried out by a group or individual, using electronic forms of contact, repeatedly and over time, against a victim who cannot easily defend him or herself .Recent research has shown that most teenagers experience cyberintimidate during their online activities including mobile phone usage , and also while involved in online gaming or social networking sites. As highlighted by the National Crime Prevention Council, approximately 50% of the youth in America are victimised by cyberintimidate . The implications of cyberintimidate become serious (suicidal attempts) when the victims fail to cope with emotional strain from abusive, threatening, humiliating and aggressive messages . The impact of cyberintimidate is exasperated by the fact that children are reluctant to share their predicament with adults (parents/teachers), driven by the fear of losing their mobile phone and/or Internet access privileges . The challenges in fighting cyberintimidate include: detecting online intimidate when it occurs; reporting it to law enforcement agencies, Internet service providers and others (for the purpose of prevention, education and awareness); and identifying predators and their victims.

II. OBJECTIVE

We proposed a cyberintimidate network, which is a weighted directed graph model. This graph model can be used to
critically analyse and answer user queries regarding predators and victims. Based on the weighted arcs between two users, the model iteratively computes the predator and victim scores for each user, and accurately identifies the most active predator and its target. From Table IV, we observed that some of the users identified as predators are also identified as victims, with different ranks. This shows the involvement of a user in intimidate activities as a predator and a victim. There could be several reasons for this. For example, suppose a user is involved in a discussion on a topic and that discussion may lead to an aggressive discussion, where users in a discussion thread started using aggressive language. Another reason could be that a receiver of the intimidate message replied through a intimidate message. The strategy of finding most active predators and victims can be adopted to classify users in various categories of victimization based on the predator and victim ranking of a user, for example, severe, moderate and normal intimidate cases. The severe category could be the case when a user ranked high as a victim is not ranked (or ranked lower than threshold) as a predator. Thus it can be argued that the victim is unable to defend himself. Accordingly, victims identified at the Rank II may not be considered as victims because they are also the top ranked predators, which shows that these victims were able to defend themselves, hence cannot be considered to be victims. Therefore this case can be discarded for further investigation. Moreover, human interference can be employed; for example, consultation with social scientists to examine cases where users appear at a severe level.

Dataset
For this work, we considered the datasets described below for the experiment on cyberintimidate detection, which are available from the workshop on Content Analysis for the Web 2.0 and we obtained the manually-labelled data from

as a ground truth dataset. The dataset contains data collected from three different social networks: Kongregate, Slashdot and MySpace. Kongregate is an online gaming site, which provides data in the chat-log style. Being gaming, the site players are likely to use aggressive words during their conversation. In Slashdot, a discussion-based site, users broadcast their message. MySpace is a popular social networking website. Datasets were provided in the form of XML files, where each file represented a discussion thread containing multiple posts. We extracted and indexed each post as one document. Each message is considered as one document and indexed through the inverted file index; thus assigning an appropriate weight to each term.

III. Proposal

Being a cyberintimidate victim entails; being subjected to personal feelings. It is when a cyberintimidate target is unable to defend oneself. Therefore, in identifying cyberintimidate predators and victims we determine the most active predators and the most attacked users' 'victims' through the sent and received intimidate messages, and the density of the badness of the message. A predators' and victims' identification graph is developed for a given scenario. Only the posts identified as intimidate were considered each user. In the experiments, each user is indexed and a userID is generated, which represents a node.Thus the username is represented by a user ID. The user information was extracted to analyse predators' and victims' data in the matrix form as depicted in Table I. The rows indicate message senders and the columns outline receivers of the post. The matrix values are the summation of intimidate messages posted and received. To examine the data content from the forum-based website, we considered every user involved in a topic discussion as both a sender and a receiver of the post. However, we assumed that the individuals will not be posting messages to themselves. Therefore we excluded the self-loop and hence assigned the post value as zero. However, in future work, similarity measures between two posts will be considered to find the reply (or a receiver) of a particular post. The chatlog dataset consists of direct conversations between two users, so for every paper, we have compared the identified top ranked predators and victims against expert judgement. However, in both cases, density of the post was not considered. In this paper, we identified the most active predators and victims, and rank are grouped together. We also noted that predators flagged at Rank I are also identified as a victim at Rank II. Similarly.

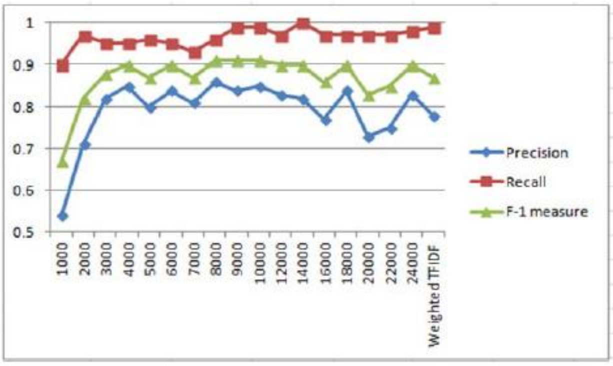

Figure: measuring the data set of users.

This work proposes a text classification model, which is helpful in identifiying suspecious: harmful and intimidate -like posts from the online conversations. Therefore, it is significant to focus on recall , because it is important to reduce mislabeling intimidate -like posts as normal posts i.e. reducing false negative. Thus, compares false positive and false negative based on weighted features because of its better performace, on individual and combine datasets. Though dataset is imbalanced, reasonable performance was obtained on individual datasets. False negative cases are very low. It indicates that system is robust in identifying cyberintimidate posts. However, number of false positive cases are still high, which is because of overfitting. Although oversampling of positive posts was adopted, high number of false positive result indicates more sophisticated learning methods need to be devised, which are able to deal with a few positive trainings. This is because in the real world problem, it is almost impossible to get a sufficient number of positive samples for training. Techniques like oversampling are subject to offline training.

IV. Conclusions

In this paper we propose an approach for cyberintimidate detection and the identification of the most active predators and victims. To improve the classification performance we employ a weighted TFIDF function, in which intimidate -like features are scaled by a factor of two. The overall results using weighted TFIDF outperformed other methods. This captures our idea to scale-up inductive words within the harmful posts. However, intimidate -like feature sets are limited to a static set of keywords. Therefore, dynamic strategies are required to be implemented to find emerging harmful and abusive words from the streaming text. To improve classifier's training in the absence of a sufficient number of positive examples, oversampling of positive posts is used. Also, throughout our experiments, we note that comparatively better performance was observed for false negative compared to false positive cases in individual and combined datasets. This is because of the fewer positive cases available for classifier's training. Therefore advance methods, which are capable of dealing with a few training sets in automatic cyber intimidate detection, and to reduce false positive and false negative cases need to be developed, In addition, we proposed a cyber intimidate graph model to rank the most active users (predators or victims) in a network. The proposed graph model can be used to answer various queries regarding the intimidate activity of a user. It can also be used to detect the level of cyber intimidate victimization for decision making in further investigations. Our future research in cyber intimidate detection will continue to reduce false cases and train classifiers with fewer positive examples. We also plan to continue the in-depth analysis of cyberintimidate victimization and its emerging patterns in stream text, to help the detection and mitigation of the cyberintimidate

REFERENCES

[1] Cyberintimidate . Available:http://en.wikipedia.org/wiki/Cyberintimidate

[2] B. Belsey. (6th July 2011). cyberintimidate .org. Available: http://www.cyberintimidate .org/

[3] P. K. Smith, J. Mahdavi, M. Carvalho, S. Fisher, S. Russell, and N. Tippett, "Cyberintimidate : Its nature and impact in secondary school pupils," Journal of Child Psychology & Psychiatry, vol. 49, pp. 376-385, 2008.

[4] M. A. Campbell, "Cyber intimidate : An old problem in a new guise?," Australian Journal of Guidance and Counselling, vol. 15, pp. 68-76,2005.

[5] NCPC.org. Cyberintimidate . Available: http://www.ncpc.org/cyberintimidate

[6] NCPC.org. Stop Cyberintimidate . Available: http://www.stopcyberintimidate .org/what_is_cyberintimidate _exactly.

[7]NCH. (2005). Putting U in the picture - Mobile intimidate survey 2005. Available: http://www.filemaker.co.uk/educationcentre/downloads/articles/Mobile_intimidate _report.pdf

[8] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet Allocation," Journal of Machine Learning Research, vol. 3, pp. 993-1022,2003.

[9] D. Butler, S. Kift, and M. Campbell, "Cyber Intimidate In Schools and the Law: Is There an Effective Means of Addressing the Power Imbalance?," eLaw Journal: Murdoch University Electronic Journal of Law, vol. 16, 2009.

[10] CAW2. (April 2009, 10 November 2010). CAW 2.0 training datasets, in Fundacion Barcelona Media (FBM). Available: http://caw2.barcelonamedia.org/

[11] D. Yin, B. D. Davison, Z. Xue, L. Hong, A. Kontostathis, and L. Edwards, "Detection of Harassment on Web 2.0," In Proceedings of The Content Analysis In The Web 2.0 (CAW2.0) Workshop at WWW2009, 2009.

[12] M. Dadvar, F. d. Jong, R. Ordelman, and D. Trieschnigg, "Improved cyberintimidate detection using gender information," In Proceedings of the Twelfth Dutch-Belgian Information Retrieval Workshop (DIR 2012), pp. 23-25, February 2012.

[13] K. Reynolds, A. Kontostathis, and L. Edwards, "Using Machine Learning to Detect Cyberintimidate ," In Proceedings of the 2011 10th International Conference on Machine Learning and Applications Workshops (ICMLA 2011), vol. 2, pp. 241-244, December 2011.

[14] K. Dinakar, R. Reichart, and H. Lieberman, "Modeling the Detection of Textual Cyberintimidate ," International Conference on Weblog and Social Media - Social Mobile Web Workshop, Barcelona, Spain 2011, 2011.

[15] A. Kontostathis, L. Edwards, and A. Leatherman, "ChatCoder: Toward the Tracking and Categorization of Internet Predators," In Proceedings of Text Mining Workshop 2009 held in conjunction with the Ninth SIAM International Conference on Data Mining (SDM 2009) 2009.

[16] I. Mcghee, J. Bayzick, A. Kontostathis, L. Edwards, A. Mcbride, and E. Jakubowski, "Learning to Identify Internet Sexual Predation," International Journal on Electronic Commerce 2011, vol. 15, pp. 103-122, 2011.

[17] Bsecure. Available: http://www.safesearchkids.com/BSecure.html

[18] Cyber Patrol. Available: http://www.cyberpatrol.com/cpparentalcontrols.asp

[19] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," Journal of Statistical Mechanics: Theory and Experiments, vol. P10008, pp. 1-12, 2008.