International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

BULLYNET: UNMASKING CYBERBULLIES ON SOCIAL NETWORKS

Manan Narmada Bai (MCA Scholar), B V Raju College, Vishnupur, Bhimavaram, West Godavari District, Andhra Pradesh, India, 534202.
G. Ramesh Kumar, B V Raju College, Vishnupur, Bhimavaram, West Godavari District, Andhra Pradesh, India, 534202.

Abstract
One of the most harmful consequences of social media is the rise of cyberbullying, which tends to be more sinister than traditional bullying given that online records typically live on the internet for quite a long time and are hard to control. In this paper, we present a three-phase algorithm, called BullyNet, for detecting cyberbullies on Twitter social network. We exploit bullying tendencies by proposing a robust method for constructing a cyberbullying signed network. We analyze tweets to determine their relation to cyberbullying, while considering the context in which the tweets exist in order to optimize their bullying score. We also propose a centrality measure to detect cyberbullies from a cyberbullying signed network, and we show that it outperforms other existing measures. We experiment on a dataset of 5.6 million tweets and our results shows that the proposed approach can detect cyberbullies with high accuracy, while being scalable with respect to the number of tweets.

1. INTRODUCTION

The Internet has created never before seen opportunities for human interaction and socialization. In the past decade, social media, in particular, has had a popularity explosion. From MySpace to Face book, Twitter, Flickr, and Instagram, people are connecting and interacting in a way that was previously impossible. The widespread usage of social media across people from all ages created a vast amount of data for several research topics, including recommender systems [1], link predictions [2], visualization, and analysis of social networks [3].

While the growth of social media has created an excellent platform for communications and information sharing, it has also created a new platform for malicious activities such as spamming [4], trolling [5], and cyber bullying [6]. According to the Cyber bullying Research Center (CRC) [7], cyber bullying occurs when someone uses the technology to send messages to harass, mistreat or threaten a person or a group. Unlike traditional bullying where aggression is a short and temporary face to- face occurrence, cyber bullying contains hurtful messages which are present online for a long time. These messages can be accessed worldwide, and are often irrevocable. Laws about cyber bullying and how it is handled differ from one place to another. For example, in the United States, the majority of the states incorporate cyberbullying into their bullying laws, and cyber bullying is considered a criminal offense in most of them [8]. Popular social media platforms such as Face book and Twitter are very vulnerable to cyber bullying due to the popularity of these social media sites and the anonymity that the internet offers to the perpetrators. Although strict laws exist to punish cyber bullying, there are very

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
IJARST
ISSN: 2457-0362

less tools available to effectively combat cyber bullying.

2. EXISTING SYSTEM

The first method of determining bullying messages was done using a combination of text-based analytics and a mix of text and user features. Zhao et al. [18] proposed a text based Embeddings-Enhanced Bag-of-Words (EBoW) model that utilizes a concatenation of bullying features, bag-of-words, and latent semantic features to obtain a final representation, which is then passed through a classifier to identify cyberbullies.

Xu et al. [21] used textual information to identify emotions in bullying traces, as opposed to determining whether or not a message was bullying. Singh et al. [19] proposed a probabilistic socio-textual information fusion for cyberbullying detection. This fusion uses social network features derived from a 1.5 ego network and textual features, such as density of bad words and part-of-speech-tags. Hosseinmardi et al. [20] used images and text to detect cyberbullying incidents. The text and image features were gathered from media sessions containing images and the corresponding comments, which was then fed into various classifiers. Chen [25] proposed an novel method in identifying cyberbullies within a multi-modal context. To understand cyberbullying Kao et al. [26] proposed a framework by studying social role detection. By using words and comments, temporal characteristics, and social information of a session as well as peer influence Cheng et al. [27], [28] proposed frameworks for detecting cyberbullies.

The second method was aimed at identifying the person behind the cyberbullying incidents. Squicciarini et al. [22] used MySpace data to create a graph, which integrated user, textual, and network features. This graph was used to detect cyberbullies and predict the spreading of bullying behavior through node classification. Gal´an-Garc´ıa et al. [23] used supervised machine learning to detect the real users behind troll profiles on Twitter, and demonstrated the technique in a

real case of cyberbullying. In a recent paper on aggression and bullying in Twitter, Chatzakou et al. [24] found cyberbullies and aggressors using user, text, and network-based features.

Disadvantages

The system is less effective due to lack of Constructing bullying signed network.

The system doesn't effective due to lack of training large scale datasets.

3. PROPOSED SYSTEM

In the proposed system, the system studies the problem of cyberbullying in social media in an attempt to answer the following research question: Can tweet contexts (conversations) help improve the detection of cyberbullying in Twitter?. Our intuition is that each tweet should be evaluated not only based on its contents, but also based on the context in which it exists. The system calls such a context a conversation, which is a set of tweets between two or more people exchanging information about a certain subject. Thus, our solution consists of three parts. First, for each conversation, a conversation graph is generated based on the sentiment and bullying words in the tweets. Second, we compute the bullying score for each pair of users in a conversation graph, and then combine all graphs to create an SSN

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
IJARST
ISSN: 2457-0362

called bullying signed network (B). The inclusion of negative links can bring out information that would otherwise be missed with only positive links [16]. Finally, we propose a centrality measure called attitude & merit (A&M) to detect bullying users from the signed network B.

Our main contributions are organized as follows:

1) Collected, preprocessed and labelled the Twitter dataset.

2) Proposed a novel efficient algorithm for detecting cyberbullies on Twitter.

a) Built conversation.

b) Constructed Bullying Signed Network.

c) Proposed Attitude and Merit Centrality.

3) Experimented on 5.6 million tweets collected over 6 months. The results show that our approach can detect cyberbullies with high accuracy, while being scalable with

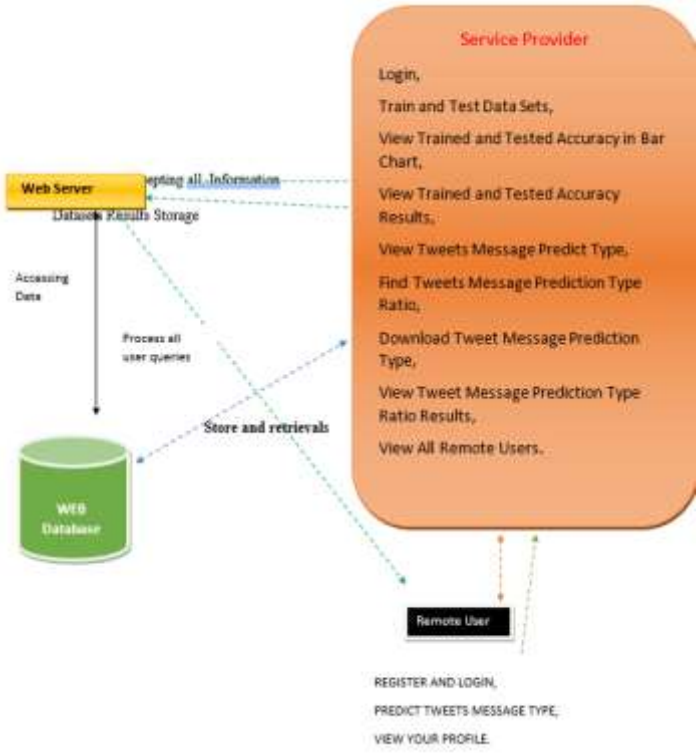
respect to the number of tweets.

Advantages

The system is more effective due to presence of Conversation Graph Generation Algorithm, Bullying Signed Network Generation Algorithm, and Bully Finding Algorithm.

The system is more effective due to the techniques to analyze large number of datasets.

4. ARCHITECTURE DIAGRAM



5. INPUT DESIGN

Input Design plays a vital role in the life cycle of software development, it requires very careful attention of developers. The input design is to feed data to the application as accurate as possible. So inputs are supposed to be designed effectively so that the errors occurring while feeding are minimized. According to Software Engineering Concepts, the input forms or screens are designed to provide to have a validation control over the input limit, range and other related validations.

6. CONCLUSION

Although the digital revolution and the rise of social media enabled great advances in communication platforms and social interactions, a wider proliferation of harmful behavior known as bullying has also emerged. This paper presents a novel

framework of Bully Net to identify bully users from the Twitter social network. We performed extensive research on mining signed networks for better understanding of the relationships between users in social media, to build a signed network (SN) based on bullying tendencies. We observed that by constructing conversations based on the context as well as content, we could effectively identify the emotions and the behavior behind bullying. In our experimental study, the evaluation of our proposed centrality measures to detect bullies from signed network, we achieved around 80% accuracy with 81% precision in identifying bullies for various cases.

There are still several open questions deserving further investigation. First, our approach focuses on extracting emotions and behavior from texts and emojis in tweets. However, it would be interesting to investigate images and videos, given that many users use them to bully others. Second, it does not distinguish between bully and aggressive users. Devising new algorithms or techniques to distinguish bullies from aggressors would prove critical in better identification of cyber bullies. Another topic of interest would be to study the relationship between conversation graph dynamics and geographic location and how these dynamics are affected by the geographic dispersion of the users? Are the proximity increase the bullying behaviour?

7. REFERENCES

1] J. Tang, C. Aggarwal, and H. Liu, "Recommendations in signed social networks," in Proceedings of the International Conference on WWW, 2016, pp. 31–40.

[2] D. Liben-Nowell and J. Kleinberg, "The link-prediction problem for social networks," Proceedings of the ASIS&T, vol. 58, no. 7, pp. 1019– 1031, 2007.

[3] U. Brandes and D. Wagner, "Analysis and visualization of social networks," in Graph drawing software, 2004, pp. 321–340.

[4] X. Hu, J. Tang, H. Gao, and H. Liu, "Social spammer detection with sentiment information," In Proceedings of IEEE ICDM, pp. 180—189, 2014.

[5] E. E. Buckels, P. D. Trapnell, and D. L. Paulhus, Trolls just want to have fun, 2014, pp. 67:97–102.

[6] S. Kumar, F. Spezzano, and V. Subrahmanian, "Accurately detecting trolls in slashdot zoo via decluttering," in Proceedings of IEEE/ACM ASONAM, 2014, pp. 188–195.

[7] J. W. Patchin and S. Hinduja, "2016 cyberbullying data," 2017.

[8] C. R. Center, "https://cyberbullying.org/bullying-laws."

[9] D. Cartwright and F. Harary, "Structural balance: a generalization of heider's theory." Psychological review, vol. 63, no. 5, p. 277, 1956.

[10] J. Leskovec, D. Huttenlocher, and J. Kleinberg, "Signed networks in social media," in Proceedings of the SIGCHI CHI, 2010, pp. 1361–1370.