Predicting Cyber Bullying On Social Media In The Big Data Era Using Machine Learning Algorithms Review Of Literature And Open Challenges

N.SRINIVASA RAO [1] **SIDDAMSETTI SANGEETHA**[2].

[1] Assistant Professor, **DEPT OF MCA, SKBR PG COLLEGE , AMALAPURAM, Andhra Pradesh**

Email:- naagaasrinu@gmail.com

[2]**PG Student of MCA, SKBR PG COLLEGE , AMALAPURAM, Andhra Pradesh**

Email:- sangeethasiddamsetti24@gmail.com.

ABSTRACT

Prior to the innovation of information communication technologies (ICT), social interactions evolved within small cultural boundaries such as geo spatial locations. The technologies have considerably transcended the temporal and spatial limitations of traditional communications. These social technologies have information, online human networks, and rich human behavior-related data. However, the misuse of social technologies such as social media (SM) platforms, has introduced a new form of aggression and violence that occurs exclusively online. A new means of demonstrating aggressive behavior in SM websites are highlighted in this paper. The motivations for the construction of prediction models to fight aggressive behavior in SM are also outlined. We comprehensively review cyberbullying prediction models and identify the main issues related to the construction of cyberbullying prediction models in SM. This paper provides insights on the overall process for cyberbullying detection and most importantly overviews the methodology. Though data collection and feature engineering process has been elaborated, yet most of the emphasis is on feature selection algorithms and then using various machine learning algorithms for prediction of cyberbullying behaviors. Finally, the issues and challenges have been highlighted as well.

I.INTRODUCTION

.1 Overview of the project

Machine or deep learning algorithms help researchers understand big data . Abundant information on humans and their societies can be obtained in this big data era, but this acquisition was previously impossible . One of the main sources of human-related data is social media (SM). By applying machine learning algorithms to SM data, we can exploit historical data to predict the future of a wide range of applications. Machine learning algorithms provide an opportunity to effectively predict and detect negative forms of human behavior, such as cyber bullying. Big data analysis can uncover hidden knowledge through deep learning from raw data . Big data analytics has improved several applications, and forecasting the future has even become possible through the combination of big data and machine learning algorithms. An insightful analysis of data on human behavior and interaction to detect and restrain aggressive behavior involves multifaceted angles and aspects and the merging of theorems and techniques from multidisciplinary and interdisciplinary. The accessibility of large-scale data produces new research questions, novel computational

methods, interdisciplinary approaches, and outstanding opportunities to discover several vital inquiries quantitatively. However, using traditional methods (statistical methods) in this context is challenging in terms of scale and accuracy. These methods are commonly based on organized data on human behavior and small-scale human networks (traditional social networks). Applying these methods to large online social networks (OSNs) in terms of scale and extent causes several issues. On the one hand, the explosive growth of OSNs enhances and disseminates aggressive forms of behavior by providing platforms and networks to commit and propagate such behavior. On the other hand, OSNs offer important data for exploring human behavior and interaction at a large scale, and these data can be used by researchers to develop effective methods of detecting and restraining misbehavior and/or aggressive behavior. OSNs provide criminals with tools to perform aggressive actions and networks to commit misconduct.

Therefore, methods that address both aspects (content and network) should be optimized to detect and restrain aggressive behavior in complex systems.

RISE OF AGGRESSIVE BEHAVIOR ON SM

Prior to the innovation of communication technologies, social interaction evolved within small cultural boundaries, such as locations and families [5]. The recent development of communication technologies exceptionally transcends the temporal and spatial limitations of traditional communication. In the last few years, online communication has shifted toward user-driven technologies, such as SM web- sites, blogs, online virtual communities, and online sharing platforms. New forms of aggression and violence emerge exclusively online [6]. The dramatic increase in negative human behavior on SM, with high increments in aggressive behavior, presents a new challenge [6], [7]. The advent of Web 2.0 technologies, including SM websites that are often accessed through mobile devices, has completely transformed functionality on the side of users [8]. SM characteristics, such as accessibility, flexibility, being free, and having well-connected social networks, provide users with liberty and flexibility to post and write on their platforms. Therefore, users can easily demonstrate aggressive behavior [9], [10]. SM websites have become dynamic social communication websites for millions of users worldwide. Data in the form of ideas, opinions, preferences, views, and discussions are spread among users rapidly through online social communication. The online interactions of SM users generate a huge volume of data that can be utilized to study human behavioral patterns [11]. SM websites also provide an exceptional opportunity to analyze patterns of social interactions among populations at a scale that is much larger than before. Aside from renovating the means through which people are influenced, SM websites provide a place for a severe form of misbehavior among users. Online complex networks, such as SM websites, changed substantially in the last decade, and this change was stimulated by the popularity of online communication through SM websites. Online communication has become an entertainment tool, rather than serving only to communicate and interact with known and unknown users. Although SM websites provide many benefits to users, cyber criminals can use these websites to commit differ- ent types of misbehavior and/or aggressive behavior. The common forms of misbehavior and/or aggressive behav- ior on OSN sites include cyberbullying [3], phishing [12], spam distribution [13], malware spreading [14], and cyberbullying [15]. Users utilize SM websites to demonstrate different types of aggressive behavior. The main involvement of

SM web- sites in aggressive behavior can be summarized in two points [9], [15]. 1)[I.] OSN communication is a revolutionary trend that exploits Web 2.0. Web 2.0 has new features that allow users to create profiles and pages, which, in turn, make users active. Unlike Web 1.0 that limits users to being passive readers of content only, Web 2.0 has expanded capabilities that allow users to be active as they post and write their thoughts. SM websites have four particular features, namely, collaboration, participation, empowerment, and timeliness [16]. These characteristics enable criminals to use SM websites as a platform to commit aggressive behavior without confronting victims [9], [15]. Examples of aggressive behavior are committing cyberbullying [17]–[19] and financial fraud [20], using malicious applications [21], and implementing social engineering and phishing [12]. 2)[II.] SM websites are structures that enable information exchange and dissemination. They allow users to effortlessly share information, such as messages, links, photos, and videos [22]. However, because SM websites connect billions of users, they have become delivery mechanisms for different forms of aggressive behavior at an extraordinary scale. SM websites help cybercriminals reach many users [23].

MOTIVATIONS FOR PREDICTING AGGRESSIVE BEHAVIOR ON SM WEBSITES

Many studies have been conducted on the contribution of machine learning algorithms to OSN content analysis in the last few years. Machine learning research has become crucial in numerous areas and successfully produced many models, tools, and algorithms for handling large amounts of data to solve real-world problems [24], [25]. Machine learning algorithms have been used extensively to analyze SM web- site content for spam [26]–[28], phishing [29], and cyber- bullying prediction [19], [30]. Aggressive behavior includes spam propagation [13], [31]–[34], phishing [12], malware spread [14], and cyberbullying [15]. Textual cyberbullying has become the dominant aggressive behavior in SM websites because these websites give users full freedom to post on their platforms [17], [35]–[39]. SM websites contain large amounts of text and/or non-text content and other information related to aggressive behavior. In this work, a content analysis of SM websites is performed to predict aggressive behavior. Such an analysis is limited to textual OSN content for predicting cyberbullying behavior. Given that cyberbullying can be easily committed, it is considered a dangerous and fast-spreading aggressive behavior. Bullies only require willingness and a laptop or cell phone with Internet connection to perform misbehavior without confronting victims [40]. The popularity and proliferation of SM websites have increased online bullying activities. Cyberbullying in SM websites is rampant due to the structural characteristics of SM websites. Cyberbullying in traditional platforms, such as emails or phone text messages, is per- formed on a limited number of people. SM websites allow users to create profiles for establishing friendships and communicating with other users regardless of geographic location, thus expanding cyberbullying beyond physical location. Anonymous users may also exist on SM websites, and this has been confirmed to be a primary cause for increased aggressive user behavior [41]. Developing an effective prediction model for predicting cyberbullying is therefore of practical significance. With all these considerations, this work performs a content-based analysis for predicting textual cyberbullying on SM websites. The motivation of this review is explained in the following section. ideation [45], [46]. Other studies [45], [46] reported an association between cyberbullying victimization and suicidal ideation risk. Consequently, developing a cyberbullying

prediction model that detects aggressive behavior that is related to the security of human beings is more important than developing a prediction model for aggressive behavior related to the security of machines. Cyberbullying can be committed anywhere and anytime. Escaping from cyberbullying is difficult because cyberbullying can reach victims anywhere and anytime. It can be committed by posting comments and statuses for a large potential audience. The victims cannot stop the spread of such activities [47]. Although SM websites have become an integral part of users' lives, a study found that SM websites are the most common platforms for cyberbullying victimization [48]. A well-known characteristic of SM websites, such as Twitter, is that they allow users to publicly express and spread their posts to a large audience while remaining anonymous [9]. The effects of public cyberbullying are worse than those of private ones, and anonymous scenarios of cyberbullying are worse than nonanonymous cases [49], [50]. Consequently, the severity of cyberbullying has increased on SM websites, which support public and anonymous scenarios of cyberbullying. These characteristics make SM websites, such as Twitter, a dangerous platform for committing cyberbullying [43]. Recent research has indicated that most experts favor the automatic monitoring of cyberbullying [51]. A study that examined 14 groups of adolescents confirmed the urgent need for automatic monitoring and prediction models for cyberbullying [52] because traditional strategies for coping with cyberbullying in the era of big data and networks do not work well. Moreover, analyzing large amounts of complex data requires machine learning-based automatic monitoring

CYBERBULLYING ON SM WEBSITES

Most researchers define cyberbullying as using electronic communication technologies to bully people [53]. Cyberbullying may exist in different types or forms, such as writing aggressive posts, harassing or bullying a victim, making hateful posts, or insulting the victim [54], [55]. Given that cyberbullying can be easily committed, it is considered a dan- gerous and fast-spreading aggressive behavior. Bullies only require willingness and a laptop or cell phone connected to the Internet to perform misbehavior without confronting the victims [40]. The popularity and proliferation of SM websites have increased online bullying activities. Cyberbullying on SM websites is performed on a large number of users due to the structural characteristics of SM websites [48]. Cyberbullying in traditional platforms, such as emails or phone text messages, is committed on a limited number of people. SM websites allow users to create profiles for establishing friendships and interacting with other online users regardless of geographic location, thus expanding cyberbullying beyond physical location. Moreover, anonymous users may exist on SM websites, and this has been confirmed to be a primary cause of increased aggressive user behavior [41]. The nature of SM websites allows cyberbullying to occur secretly, spread rapidly, and continue easily [54]. Consequently, developing an effective prediction model for predicting cyberbullying is of practical significance. SM websites contain large amounts of text and/or non-text content and information related to aggressive behavior.

METHODOLOGY

This section presents the methodology used in this work for a literature search. Two phases were employed to retrieve published papers on cyberbullying prediction models. The first phase included searching for reputable academic databases and search engines. The search engines and academic databases used for the retrieval of relevant papers were as follows: Scopus, Clarivate Analytics' Web of Science, DBLP Computer Science Bibliography, ACM

Digital Library, Science Direct, SpringerLink, and IEEE Xplore. The major keywords used for the literature search were coined in relation to social media as follows: cyberbullying, aggressive behavior, big data, and cyberbullying models. The second phase involved searching for literature through Qatar University's digital library. The articles retrieved from the search were scrutinized to ensure that the articles met the inclusion criteria. According to the inclusion criteria, for an article to be selected for the survey, it must report an empirical udy describing the prediction of cyberbullying on SM sites. Otherwise, the article would be excluded in the selection. Many articles were rejected based on titles. The abstract and conclusion sections were examined to ensure that articles satisfied the screening criteria, and those that did not satisfy the criteria were excluded from the survey.

II.EXISTING SYSTEM

• State-of-the-art research has developed features to improve the performance of cyber bullying prediction.

• For example, a lexical syntactic feature has been proposed to deal with the prediction of offensive language; this method is better than traditional learning-based approaches in terms of precision. Dadvar et al. examined gender information from profile information and developed a gender-based approach for cyber bullying prediction by using datasets from My space as a basis. The gender feature was selected to improve the discrimination capability of a classifier. Age and gender were included as features in other studies, but these features are limited to the information provided by users in their online profiles.

• several studies focused on cyber bullying prediction based on profane words as a feature. Similarly, a lexicon of profane words was constructed to indicate bullying, and these words were used as features for input to machine learning algorithms. Using profane words as features demonstrates a significant improvement in model performance.

• For example, the number of ``bad" words and the density of ``bad" words were proposed as features for input to machine learning in a previous work.The study concluded that the percentage of ``bad" words in a text is indicative of cyber bullying. Another research expanded a list of pre-defined profane words and allocated different weights to create bullying features. These features were concatenated with bag-ofwords and latent semantic features and used as a feature input for a machine learning

Disadvantages

• The System is not much affective due to Semi supervised machine learning techniques. • The system doesn't have sentiment classification for cyber bullying

III.PROPOSED SYSTEM

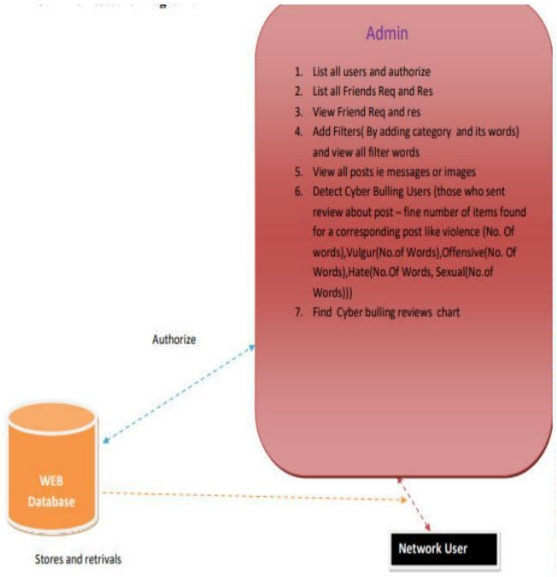
• The proposed system is constructing cyber bullying prediction models is to use a text classification approach that involves the construction of machine learning classifiers from labeled text instances. • Another means is to use a lexicon-based model that involves computing orientation for a document from the semantic orientation of words or phrases in the document. • Generally, the lexicon in lexicon-based models can be constructed manually or automatically by using seed words to expand the list of words. However, cyberbullying prediction using the lexicon-based approach is rare in literature. • The primary reason is that the texts on SM websites are written in an unstructured manner, thus making it difficult for the lexicon-based approach to detect cyberbullying based only on lexicons. • However, lexicons are used to extract

features, which are often utilized as inputs to machine learning algorithms. • For example, lexicon based approaches, such as using a profane-based dictionary to detect the number of profane words in a post, are adopted as profane features to machine learning models. • The key to effective cyber bullying prediction is to have a set of features that are extracted and engineered.

Advantages

• The system is more effective due to LOGISTIC REGRESSION CLASSIFICATION and UNSUPERVISED MACHINE LEARNING. • An effective cyber bullying prediction models is to use a text classification approach that involves the construction of machine learning classifiers from labeled text instance and also is to use a lexicon-based model that involves computing orientation for a document from the semantic orientation of words or phrases in the document.

IV.SYSTEM SRCHITECTURE



V. MODULES

• Admin

In this module, the Admin has to login by using valid user name and password. After login successful he can perform some operations such as view and authorize users, view all friends request and responses, Add and View Filters, View all posts, Detect Cyber Bullying Users, Find Cyber Bullying Reviews Chart. Viewing and Authorizing Users In this module, the admin views all users details and authorize them for login permission. User Details such as User Name, Address, Email Id, Mobile Number. Viewing all Friends Request and Response In this module, the admin can see all the friends' requests and response history. Details such as Requested User Name and Image, and Requested to User Name and Image, status and date. Add and View Filters In this module, the admin can add filters (like Violence, Vulgar, Offensive, Hate, and Sexual) as Categories with the words those related to corresponding filters. View all posts In this module, the admin can see all the posts added by the users with post details like post name, description and post image. Detect Cyber Bullying Users In this module, the admin can see all the Cyber Bullying Users (The users who had posted a comment on posts using cyber bullying words which are all listed by the admin to detect and filter). In this, the results shown as, Number of items found for a corresponding post like Violence (no. of words belongs to Violence Filter used in comments by the users), Vulgar (no. of words belongs to Vulgar Filter used in comments by the users), Offensive (no. of words belongs to Offensive Filter used in comments by the users), Hate (no. of words belongs to Hate Filter used in comments by the users), Sexual (no. of words belongs to Sexual Filter used in comments by the users). Find Cyber Bullying Reviews Chart In this module, the admin can see all the posts with number of cyber bullying comments posted by users for particular post.

User

In this module, there are n numbers of users are present. User should register before performing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user can perform some operations like viewing their profile details, searching for friends and sending friend requests, Posting Your Messages as Posts by giving details, View and Comment on Friend Posts, viewing all friends posts and comment, view all your cyber bullying comments on your friend posts. Viewing Profile Details, Search and Request Friends In this module, the user can see their own profile details, such as their address, email, mobile number, profile Image. The user can search for friends and can send friend requests or can accept friend requests. Add Posts In this, the user can add their own posts by giving post details such as, post title, description, uses, and image of post. View and Comment on Your Friends Post In this, the user can see his entire friend's post details (post title, description, uses, creator and image of post) and can comment on posts. View all Friends Posts and Comment (Cyber bullying Related) In this, the user can see his all friend's post details (post title, description, uses, creator and image of post) and can comment on posts. Don't Post If the comment consists of Cyber bullying words and Shows the reason why comment is not posted by indicating Detected Cyber Bullying Words like Numbers of Cyber Bullying words Related to Filter Violence found in comment, Numbers of Cyber Bullying words Related to Filter Vulgar found in comment, Numbers of Cyber Bullying words Related to Offensive found in comment, Numbers of Cyber Bullying words Related to Hate found in comment, Numbers of Cyber Bullying words Related to Sexual found in comment, View all Your Cyber bullying comments on your friend posts The user can see all his posted cyber bullying comments on their friend created posts.

VI. IMPLEMENTATION

6.1 Machine learning algorithms: The machine learning field focuses on the development and application of computer algorithms that improve with experience [75], [76]. The objective of machine learning is to identify and define the patterns and correlations between data. The importance of analyzing big data lies in discovering hidden knowledge through deep learning from raw data [1]. Machine learning can be described as the adoption of computational models to improve machine performance by predicting and describing meaningful patterns in training data and the acquisition of knowledge from experience [77]. When this concept is applied to OSN content, the potential of machine learning lies in exploiting historical data to detect, predict, and understand large amounts of OSN data. For example, in supervised machine learning for classification application, classification is learned with the help of suitable examples from a training dataset. In the testing stage, new data are fed into the model, and instances are classified to a specified class learned during the training stage. Then, classification performance is evaluated. This section reviews the most common processes in the construction of cyberbullying prediction models for SM web-sites based on machine learning. The review covers data collection, feature engineering, feature selection, and machine learning algorithms.

A.DATA COLLECTION

Data are important components of all machine learning-based prediction models. However, data (even "Big Data") are useless on their own until knowledge or implications are extracted from them. Data extracted from SM websites are

used to select training and testing datasets. Supervised prediction models aim to provide computer techniques to enhance prediction performance in defined tasks on the basis of observed instances (labeled data) [78]. Machine learning models for a certain task primarily aim to generalize; a successful model should not be limited to examples in a training dataset only [79] but must include unlabeled real data. Data quantity is inconsequential; what is crucial is whether or not the extracted data represent activities on SM websites well [80]–[82]. The main data collection strategies in previous cyberbullying prediction studies on SM websites can be categorized into data extracted from SM websites by using either keywords, that is, words, phrases, or hash- tags (e.g., [19], [43], [83]–[85]), or by using user profiles (e.g., [38], [62], [70], [86]). The issues in these data collection strategies and their effects on the performance of machine learning algorithms are highlighted in the Data Collection section (related issues).

B.FEATURE ENGINEERING

Feature is a measurable property of a task that is being observed [87]. The main purpose of engineering featurevectors is to provide machine learning algorithms with a set of learning vectors through which these algorithms learn how to discriminate between different types of classes [76]. Feature engineering is a key factor behind the success and failure of most machine learning models [79]. The success and failure of prediction may be based on several elements. The most significant element is the features used to train the model [78]. Most of the effort in constructing cyberbullying prediction models using learning algorithms is devoted to this task [61], [62], [72]. In this context, the design of the input space (i.e., features and their combinations that are provided as an input to the classifier) is vital. Proposing a set of discriminative features, which are used as inputs to the machine learning classifier, is the main step toward constructing an effective classifier in many applications [76]. Feature sets can be created based on human-engineered observations, which rely on how features correlate with the occurrences of classes [76]. For example, recent cyberbullying studies [88]–[94] established the correlation between different variables, such as age, gender, and user personality, and cyberbullying occurrence. These observations can be engineered into a practical form (feature) to allow the classifier to discriminate between cyberbullying and non-cyberbullying and can thus be used to develop effective cyberbullying prediction models. Proposing features is an important step toward improving the discrimination power of prediction models [76], [79]. Similarly, proposing a set of significant features of cyberbullying engagement on SM web- sites is important in developing effective prediction models based on machine learning algorithms [68], [95]. State-of-the-art research has developed features to improve the performance of cyber bullying prediction. For example, a lexical syntactic feature has been proposed to deal with the prediction of offensive language; this method is better than traditional learning-based approaches in terms of precision [18]. Dadvar et al. examined gender information from profile information and developed a gender-based approach for cyberbullying prediction by using datasets from My space as a basis. The gender feature was selected to improve the discrimination capability of a classifier. Age and gender were included as features in other studies [17], [61], but these features are limited to the information provided by users in their online profiles. Several studies focused on cyberbullying prediction based on profane words as a feature [35], [68], [70], [95], [96]. Similarly, a lexicon of profane words was constructed to indicate bullying, and these words were used as features for input to machine learning algorithms [97], [98]. Using profane words as features

demonstrates a significant improvement in model performance. For example, the number of ‘‘bad’’ words and the density of ‘‘bad’’ words were proposed as features for input to machine learning in a previous work [70]. The study concluded that the percentage of ‘‘bad’’ words in a text is indicative of cyberbullying. Another research [85] expanded a list of pre-defined profane words and allocated different weights to create bullying features. These features were concatenated with bag-of-words and latent semantic features and used as a feature input for a machine learning algorithm. Reference [19] proposed features, such as pronouns and skip grams, as additional features to traditional models, such as bag of words (n-gram n 1). The authors claimed that adding these features improved the overall classification accuracy. Another study [62] analyzed textual cyberbullying associated with comments on images in Instagram and developed a set of features from text comprising traditional bag-of-words features, comment counts for an image, and post counts within less than one hour of posting the image. Features mined from user and media information, including the number of followers and likes, and shared media and features from image content, such as image types, were added.

VII.CONCLUSION

This study reviewed existing literature to detect aggressive behavior on SM websites by using machine learning approaches. We specifically reviewed four aspects of detecting cyberbullying messages by using machine learning approaches, namely, data collection, feature engineering, construction of cyberbullying detection model, and evaluation of constructed cyberbullying detection models. Several types of discriminative features that were used to detect cyberbullying in online social networking sites were also summarized. In addition, the most effective supervised machine learning classifiers for classifying cyberbullying messages in online social networking sites were identified. One of the main contributions of current paper is the definition of evaluation metrics to successfully identify the significant parameter so the various machine learning algorithms can be evaluated against each other. Most importantly we summarized and identified the important factors for detecting cyberbullying through machine learning techniques specially supervised learning. For this purpose, we have used accuracy, precision recall and fmeasure which gives us the area under the curve function for modeling the behaviors in cyberbullying. Finally, the main issues and open research challenges were described and discussed.

REFERENCES:

[1] V. Subrahmanian and S. Kumar, ``Predicting human behavior: The nextfrontiers," Science, vol. 355, no. 6324, p. 489, 2017.

[2] H. Lauw, J. C. Shafer, R. Agrawal, and A. Ntoulas, ``Homophily in thedigital world: A LiveJournal case study," IEEE Internet Comput., vol. 14,no. 2, pp. 15_23, Mar./Apr. 2010.

[3] M. A. Al-Garadi, K. D. Varathan, and S. D. Ravana, ``Cybercrime detection in online communications: The experimental case of cyberbullyingdetection in the Twitter network," Comput. Hum. Behav., vol. 63,pp. 433_443, Oct. 2016.

[4] L. Phillips, C. Dowling, K. Shaffer, N. Hodas, and S. Volkova, ``Using social media to predict the future: A systematic literature review," 2017, arXiv:1706.06134. [Online]. Available: https://arxiv.org/abs/1706.06134

[5] H. Quan, J. Wu, and Y. Shi, ``Online social networks & social network services: A technical survey," in Pervasive Communication Handbook. Boca Raton, FL, USA: CRC Press, 2011, p. 4.

[6] J. K. Peterson and J. Densley, ``Is social media a gang? Toward a selection, facilitation, or enhancement explanation of cyber violence," Aggression Violent Behav., 2016.

[7] BBC. (2012). Huge Rise in Social Media. [Online]. Available: http://www.bbc.com/news/uk20851797

[8] P. A.Watters and N. Phair, ``Detecting illicit drugs on social media using automated social media intelligence analysis (ASMIA)," in Cyberspace Safety and Security. Berlin, Germany: Springer, 2012, pp. 66_76.

[9] M. Fire, R. Goldschmidt, and Y. Elovici, ``Online social networks: Threats and solutions," IEEE Commun. Surveys Tuts., vol. 16, no. 4, pp. 2019_2036, 4th Quart., 2014.

[10] N. M. Shekokar and K. B. Kansara, ``Security against sybil attack in social network," in Proc. Int. Conf. Inf. Commun. Embedded Syst. (ICICES), 2016, pp. 1_5.

[11] J. Ratkiewicz, M. Conover, M. Meiss, B. Gonçalves, A. Flammini, and F. Menczer, ``Detecting and tracking political abuse in social media," in Proc. 5th Int. AAAI Conf. Weblogs Social Media, 2011, pp. 297_304.

[12] A. Aggarwal, A. Rajadesingan, and P. Kumaraguru, ``PhishAri: Automatic realtime phishing detection on Twitter," in Proc. eCrime Res. Summit (eCrime), Oct. 2012, pp. 1_12. article/view/2793/2431