International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 02 | February - 2023 Impact Factor: 7.185 ISSN: 2582-3930

A Review of Cyberbullying Detection

Vinayak Kalwar, Dr. Pushpinder Singh Patheja*
MTech Student Vellore Institute of Technology, Bhopal, India
Division Head Cyber Security and Digital Forensics, Vellore Institute of Technology, Bhopal, India

***Abstract*—** With the expansion of Web 2.0, online communication and social networks are emerging. This alternation helps users to share their information and collaborate easily. additionally, these internet services help establish new connections between persons or reinforce existing ones. However, they'll also result in misbehaviors or cybercriminal acts, for instance, cyberbullying. At the identical time, it can make children and adolescents use the technologies to harm another person. because of the negative effect of cyberbullying, some techniques and methods are proposed to beat this problem. This paper illustrates a survey covering some methods and challenges in cyberbullying. Next, we provide suggestions for continued research in this area.

Keywords- cyberbullying; cybercrime; cyber predator; text mining.

I. INTRODUCTION

Nowadays, people around the world use different online forums, blogs, social networking sites, and forums as a foundation for their networking, sharing, and transfer of data. Although online communities and social networks became more common, some users use these communities in illegal and unethical ways, which results in teens and youth people bullied over the net. National Crime Prevention Council delimits cyberbullying because of the following: 'when the web, cell phones or other devices are accustomed send or post text or images intended to harm or embarrass another person [22]. Dana Boyd as a man of science has expressed four phrases online. These phrases change the very dynamics of bullying. They also magnificent bullying to new levels: searchability, replicability, persistence, and invisible audiences [23].

Recent studies reported almost 43% of adolescents within the u. s. alone been bullied at some point in time [20]. Like traditional kinds of bullying, cyberbullying has an intensely negative effect on children and young adults. in line with the American Academy of kid and Adolescent Psychiatry, victims of cyberbullying usually suffer from emotional and psychological experiences [21].

Social science has been studied extensively for understanding various attributes and therefore the prevalence of the cyberbullying problem. Prevention measures include human interference, deleting offensive terms, blacklisting or scoring the authors' cyber performance, and academic awareness. However, because of the dearth of existing datasets, few studies are focused on online cyberbullying detection. The most course of action in fighting cyberbullying is the detection and provision of subsequent preventive measures. The issues in preventing cyberbullying are finding cyberbullying when it happens; revealing it to police agencies, internet service suppliers et al (for the item of avoidance, education, and awareness); and distinguishing between predator and victims [24].

In this paper, we are visiting review some kinds of literature on cyberbullying to grasp when it occurs and to spot between predators and victims.

This paper is structured into the following: section 2 describes cybercrime and illustrates two aspects of it i.e. cyberbullying and cyber predator detection. Section 3 presents the information source used for cyberbullying and cyber predators. Section 4 describes some applications and tools that are employed in this area. The last section concludes our review and discusses some future directions for research.

II. CYBERCRIME ANALYSIS

Cybercrime is said as any criminal activity which makes employing a computer the primary means of commission. This definition was expended by the U.S. Department of Justice, for any criminal activity to use a computer as a storage of evidence.

Based on [25], cybercrime might be categorized in two ways: content-based and technology-based crimes. the previous is managed by any specific political movement associated with the article of threatening, national security, porn, harassment, etc., and also the latter involves hacking, injecting malicious code, incidents of espionage, etc. Those that are involved in both types should have some technical knowledge. Cybercriminals tend to be residing in various varieties of the world and revel in getting the privilege of varied citizens.

In this paper, we consider content-based crime which incorporates cyber predators and cyberbullying detection.

A. Cyberbullying Detection

Cyberbullying was defined by Patchin and Hinduja as "willful and repeated harm inflicted through the medium of text [3]." The overview of the adolescent psychology literature shows nine various varieties of cyberbullying which might be recognized in [2], [3], [4]. These categories are flooding, masquerade, flaming, trolling, harassment, cyberstalking, denigration, outing, and exclusion. the kinds of bullying are defined as follows :

Flooding involves the bully frequently sending an identical comment, nonsense comments, or pressing the enter to prevent the victim from contributing to the conversation [2].
Masquerade involves the bully pretending to be someone they're not. this might make it appear to bully a victim directly [4].
Flaming or bashing may be a reasonable online right. The bully sends or posts electronic mail which is enticingly insulting and vulgar to 1 or several persons either privately or publicly to a web group [4].
Trolling, also called baiting, includes purposely publishing comments that all other comments. The poster intended to incite an argument or arouse emotions, however, the comments don't seem to be necessarily personal, vulgar, or emotional [1].
Harassment is the reason that the bully frequently sends insulting and rude messages to the victim [4].
Cyberstalking and threats occur the poster sends intimidating or offensive messages [4].
Denigration also called “dissing” happens when an electronic bully sends or publishes gossip or untrue statement to a few victims to wreck the victim’s friendship or reputation [4].
Outing occurs when a private sends or publishes private or embarrassing information in a very very public chat room or forum. this type of cyberbullying is analogous to denigration. However, in the outing, the link between bullies and victim are close [4].
Exclusion involves intentionally excluding someone from a web group. this sort of cyberbullying happened among youth and teenage generally [3].
For a variety of issues associated with cyberbullying recognition, research has been done to support the text mining paradigm like online sexual predator recognition [7] and spam detection [8]. Nevertheless, little study has been done on technical solutions, which is why there are insufficient proper training datasets. Moreover, privacy issues and ambiguities may be the explanations for describing cyberbullying.
References [5] proposed a supervised learning approach for determining harassing posts in chat rooms and discussion forums. Three features, namely content, sentiment, and contextual, were used to train a support vector machine classifier. They also used N-grams, TFIDF weighting, and foul word frequency because of the baseline. Although their results indicate enhancements over the baselines, the temporal or user information hasn’t been utilized. Moreover, they employed only supervised methods. Nevertheless, unsupervised methods can also encourage be valuable. In another study with an identical dataset, the authors tried to spot clusters containing cyberbullying by employing a rule-based algorithm [9].
For detecting cyberbullying in YouTube comments, researchers [6] described a technique. In their method, they used a range of binary and multiclass classifiers on a manually labeled dataset. Also, they applied good judgment knowledge for detecting cyberbullying. Using logic can help provide information about people’s goals and emotions and objects’ properties and relations which will help disambiguate and contextualize language. They also used two varieties of features: 1) general features that contain a term frequency-inverse document frequency (TF-IDF) weighted uni-grams, the Ortony lexicon of words denoting negative connotation, an inventory of profane words and often occurring part-of-speech (POS) bigram tags and 2) label specific features. Their study indicated that binary classifiers can outperform the popularity of textual cyberbullying as compared to multiclass classifiers. Their results illustrate using such features under consideration going to be more useful and may cause better modeling of the matter. the restrictions of their study are that they didn't consider the pragmatics of dialogue and conversation and therefore the social networking graph.

Reference [14] improved the work that was done by the Massachusetts Institute of Technology’s approach [6]. They proposed a machine learning approach for detecting cyberbullying from Formspr ing. me. They applied the amount of "bad" words (NUM) and also the density of "bad" words (NORM) features that were devised by assigning a severity level to the bad words list (nosewaring.com). They employed replication of positive examples up to 10 times and accuracy on the range of classifiers was reported. Their results illustrated that the C4.5 decision tree and an instance-based learner could recognize actuality positives with 78.5% accuracy.
Recently some work are done on recognizing users by interaction theine. While providing profile information for social networks, and browsing the web, users leave an amount of traces. This distributed user data may be utilized as some way to get information for systems that provide personalized services for their users or have to find more information about their users [12]. Connecting data from different sources have been used for various purposes, like standardization of APIs (e.g. OpenSocial1) and personalization [10].

Previous works in cyberbullying detection have mostly focused on the conversations' content though they failed to attend to the characteristics of the actors involved in cyberbullying. Social studies demonstrate that males and feminine bully one another in numerous ways. as example, while women with aggressive communication styles, like excluding someone from a gaggle of conspiracy against them, men tend to use more words and phrases that threatened outrage [11]. Reference [26] illustrates that more pronouns like “I”, “you”, “she”, etc. are utilized by females and more noun specifiers like,“, “, and “that” are employed by males. These findings have motivated a recent study [13] of the effect of linguistic features supported gender within the diagnosis of cyberbullying on social networks.

Reference [24] in 2013, proposed a good approach to detect cyberbullying from social media. and that they also presented a graph model to extract cyberbullying networks. This has led to identifying the foremost active predators and victims through a ranking algorithm. Their proposed graph model can be wont to recognize the amount of cyberbullying victimization for higher cognitive processes in further studies. they might improve the classification performance by applying a weighted TF-IDF function, during which bullying-like features are scaled by an element of two.
As we've mentioned before, there are four major tasks in cyberbullying detection: detecting online bullying; reporting it to enforcement agencies, Internet service providers et al. (for the aim of prevention, education, and awareness); and identifying predators and their victims. In part A, we review some techniques in cyberbullying detection. Next session we describe some papers on the detection of cyber predators and victim detection.

B. Cyber predator

"A cyber predator could use the web to search for victims to form the foremost of in any way, including sexually, emotionally, psychologically or financially. Cyber predators understand how to govern kids, creating trust and friendship where none should exist" [27].

National Center for Missing and Exploited Children (NCMEC), explained about 1 in 7 youth (ages 10- to 17 years- old) experience a sexual approach or appeal through the web [18]. Online sexual predator studies [16], [17] relate to the speculation of communication and text-mining techniques to tell apart between predator and victim conversations, as applied to one-to-one communication.

Recognizing the predator problem is split into two sub-problems, namely identifying predators and recognizing predators' lines for identifying predators. supported [39] we will divide the summary of existing approaches in identifying predators into three steps namely: pre-filtering approach, feature extraction approach, and classification approach. one of the foremost effective methods for pre-filtering of all conversations is finished by [31]. They displayed some specific patterns, for example, the existence of 1 participant only, those with but 6 interventions per user, or those including 3 long sequences of unrecognized characters. Although other researchers [38] proposed similar tasks, they applied a rule-based approach to different features for various methods.

For the second task i.e. feature extraction, the features are categorized into two principal groups: "lexical" features and "behavioral" features. the previous are those which might be derived from the raw text of the conversation, as an example unigram or bigram [31,37,35,38] features, emoticons counting, and therefore the weighting applied TF-IDF or the cosine similarity. Recognizing the name of the participants within the conversation (self, other, group) is another example [35].

The latter features are within a conversation [43,36]: the number of questions asked, intention (grooming, hooking, ...), and its capture of the "action" of the users. The creation of a single set of features for every author is one of the important approaches. This approach can describe and develop his predator potential. Some researchers used the Language Model (LM) for 2 participants within the chat [35]. Some approaches used LM at the line level or the conversation level. They applied this strategy, to sum up, the score of all the lines or conversations to seek out a singular set of features of every author [41,42,43,40,45]. For classifying predator and non-predator many alternative approaches were proposed, for instance, decision trees [42], random forest [46] furthermore and Naïve Bayes [43,41], and Maximum-Entropy [35,45]. compared among existing classifiers, Support Vector Machines (SVM) were used plenty [37,38,40,31]. Some researchers have shown the opposite approach performed better than SVM, for example, after they applied a Neural Network classifier [31].

For the second issue, identifying predator's lines, the proposed solution was associated with all the relevant conversation lines of all predators.

These predators are obtained from the primary problem [46]. one of the foremost used techniques was a filtering of all the conversations of predators via a thesaurus of "perverted" phrases or with a particular score (e.g. TF-IDF weighting) [40,37,38,35].

DATA SOURCE

In this section, we present a summary of the gathering and labeling of the info which is employed in previous approaches.

A) Dataset Origin

In predator communications there's little reliable labeled data; where plenty of the add both communication studies and computing is devoted to anecdotal evidence and chat logs transcripts from Perverted Justice (PJ)[47].

Using the PJ transcript for cyber predator detection is contentious. The PJ contains transcripts of the conversation. These conversations have involved a predator and pseudo-victim, an adult posing as a youth.

Dr. Susan Gauch, University of Arkansas made the second data set to acknowledge a predator [48]. She developed a brand new software called ChatTrack for crawling and downloading chat logs. Although ChatTrack isn't accessible now, the chart data are still employed in a number of the first research. The researchers have included analyses of predator communication [15].

In the Content Analysis for Web 2.0 workshop (CAW 2.0-2009), a shared task for misbehavior detection was proposed. Misbehavior detection is to acknowledge improper activity in the virtual community when some users harass or offend other members. CAW 2.0 provides datasets for online harassment. The dataset contains five various public sites; Ciao, MySpace, Twitter, and Kongregate [19]. These data may be categorized into two forms of communities; chat-style and discussion-style communities. Among the mentioned datasets, Kongregate is one of the samples of chat-style communities. during this style of the dataset, posts are contained in short messages. the most characteristic of those messages is having some words with many misspellings. as compared with chat-style communities, discussion-style communities (MySpace and Slashdot) have rather longer posts. These posts are still shorter than full websites. The terms in these posts also are more formal.

In chat-style communities like Kongregate, posts are usually short online messages containing only some words with many misspellings. In discussion-style communities, like Slashdot and MySpace, posts are relatively longer (but still shorter than full web pages) and also the usage of the terms in these posts is more formal, as compared with other chat-style communities.

A) Labeling the information

For generating a coded dataset, few numbers of devoted content coders were employed by previous researchers. Amazon Mechanical Turk (MTurk) has been utilized by recent researchers because of its crowdsourcing services. MTurk is an internet market that helps researchers to decompose asking post jobs into an oversized amount of small tasks.

Tukers (MTurk workers) have offered a quick description of accessible tasks and determine which task to accomplish. Usually used time for accomplishing each task is between 5 and 20 seconds. Also, workers paid about 5 cents for every single task. Besides its function as a source of labor, MTurk may be a good place to conduct user studies in human-computer interaction (Kittur, Chi, & Suh, 2008) and large-scale economic experiment (Mason & Watts, 2009).

In comparison to the traditional model, MTurk has several significant advantages. Firstly, Turkers are an on-demand proletariat. Using the MTurk system we could also rapidly engage an oversized number of programmers, without the much overhead related to hiring dedicated employees.

Secondly, MTurk workers are derived from various perspectives and supply several different experiences with online interaction and comments. a very important factor of MTurk is quality. Some studies
[49] illustrated utilizing MTurk for analyzing content was faster and cheaper than utilizing devoted raters. Other researchers [50] also demonstrated that using various non-expert workers could make prime-quality results.
Lastly, [51] proposed that it's possible to get prime quality, efficient coding through the employment of several non-expert coders while individual coders don't always agree (i.e. the info is "noisy").

APPLICATION AND TOOLS

With the expansion of cyberbullying among children and teenagers, the foremost important question that may be expected from a teen is to discern the degree between right and wrong. So responsible parents must protect their children from internet predators. In this regards some available commercial and networks are eBlaster TM, Net Nanny TM, and, ImBigBrother [32,33,34].

Packet sniffing is the most prevalent alternative to Safe Chat. Packet sniffers scan all the outgoing and ingoing traffic during a network and then apply a filter to only see the useful blocks of knowledge. Although many packet sniffers and tools are easy to use, parents may have problems with these tools. the foremost important reason is that tools that are now available to detect predation are supported an easy keyword matching and have not been studied. This brings the accuracy of those tools into question [15].

To beat the boundaries of other tools, SafeChat was proposed. This software is additionally better than currently available tools. the primary version of SafeChat as stand-alone software. SafeChat 1.0 used the WinpCap library. WinpCap could be a library that helps programmers to own high-level control over the retrieval and transmission of packets within the Windows environment. This library is employed by many widely used commercial products like Wireshark [29]. SafeChat 1.0 was designed to figure with AIM Instant Messaging because AIM has the most important market share among IM tools. AIM uses a protocol called Open System for Communication in Realtime (OSCAR). Despite the name, OSCAR isn't an Open Source System. In 2008, documentation on the OSCAR protocol was released [28]. Like AIM, many other chat clients didn't have proper documentation. SafeChat, to be successful, should be compatible with many protocols. SafeChat 2.0 is the recreate of SafeChat. it's a 3rd party plugin for Pidgin, an open-source instant messaging system.

It uses detection algorithms to classify chat participants as potential. Pidgin is one of the foremost popular open-source instant messaging systems. It works on any Windows or Unix-based environment and supports multiple protocols including AIM, MSN, ICQ, IRC, and Yahoo. Unsupported protocols, like Facebook Chat, will be employed in Pidgin with the employment of third-party plugins [30]. There are multiple reasons for selecting the Pidgin platform. the first reason is that we would like SafeChat to be available to help as many families as possible. Therefore, SafeChat has to support as many IM protocols as possible. Second, SafeChat can benefit from the event efforts of the Pidgin community. When new protocols are made or existing protocols are changed, the Pidgin community will update Pidgin. This allowed us to target the predation algorithms for SafeChat rather than infrastructure issues.

III. CONCLUSION

With the ascent of the net, more and more people interact with people within the same town or on the opposite side of the globe. However, the possibility of misuse comes with any new technology. Unfortunately, these techniques result in misbehavior or cybercriminal acts like cyberbullying. Our literature review illustrates that there is research on cyber predators and cyberbullying detection. In the future, addressing the role of newer technologies, especially peer-to-peer devices, and cell phones should be considered for further research.
Also, collaborations with text mining and data retrieval research groups help us to seek out an honest solution to detect this annoying problem. As we mentioned before, there's no labeled dataset, so future researchers can work on collecting new labeled datasets for the long-run study. Working with psychologists, sociologists, communications and enforcement experts can improve awareness of understanding, recognizing, and preventing cybercrime. And developing a decent classifier to acknowledge predator's behavior is additionally needed.

REFERENCES

[1] Glossary of cyberbullying terms. (2008, January). Retrieved from http://www.adl.org/education/curriculum_connections/cyberbullying/glossary.pdf
[2] D. Maher, "Cyberbullying: an ethnographic case study **of 1** Australian upper **elementary school** class". Youth Studies Australia, 27(4), 5057,2008.
[3] J. Patchin, & S. Hinduja, "Bullies move beyond the schoolyard; a preliminary **observe** cyberbullying." Youth violence and juvenile justice. 4:2 (2006). 148-169.
[4] N.E. Willard, "Cyberbullying and Cyberthreats: Responding to the Challenge of Online Social Aggression, Threats, and Distress." Champaign, IL: Research, 2007.
[5] D. Yin, Z. Xue, L. Hong, B.D. Davison, A. Kontostathis, L. Edwards, " Detection of harassment on Web 2.0." In: Proceedings of CAW2.0, 2009.Madrid, April 20-24.
[6] K. Dinakar, R. Reichart, H. Lieberman, "Modelling the Detection of Textual Cyberbullying." In: ICWSM 2011, Barcelona, Spain, July 17-21 2011.
[7] A. Kontostathis, "ChatCoder: Toward the tracking and categorization of internet predators." In: Proceedings of SDM 2009, Sparks, NV, May 2, 2009.
[8] P.N. Tan, F. Chen, A. Jain, "Information assurance: Detection of web spam attacks in social media." Proceedings of Army Science Conference, Orland, Florida. 2010.
[9] J.F. Chisholm, "Cyberspace violence against girls and adolescent females." Annals of the Academy of Sciences 1087, 2006. pp. 74–89.
[10] F. Nola, F. Cena, "User identification for cross-system personalization. Information Sciences" 179, 2009. pp. 16–32.
[11] J.F. Chisholm, "Cyberspace violence against girls and adolescent females." Annals of the **NY** Academy of Sciences 1087,2006. pp. 74-89.
[12] F. Abel, N. Henze, E. Herder, D. Krause, "Linkage, aggregation, alignment and enrichment of public user profiles with Mypes." In: Proceedings of I-SEMANTICS, Graz, Austria. 2010. pp. 1–8
[13] M. Dadvar, F. d. Jong, R. Ordelman, and D. Trieschnigg, "Improved cyberbullying detection using gender information," In Proceedings of the Twelfth Dutch-Belgian Information Retrieval Workshop (DIR 2012), February 2012. pp. 23-25,
[14] K. Reynolds, A. Kontostathis, and L. Edwards, "Using Machine Learning to Detect Cyberbullying," In Proceedings of the 2011 10th International Conference on Machine Learning and Applications Workshops (ICMLA 2011), vol. 2, December 2011. pp. 241-244,
[15] A. Kontostathis, L. Edwards, A. Leatherman, "Text mining and cybercrime." Text Mining: Applications and Theory. John Wiley & Sons, Ltd, Chichester, UK.2009.
[16] A. Kontostathis, L. Edwards, and A. Leatherman, "ChatCoder: Toward the Tracking and Categorization of Internet Predators," InProceedings of Text Mining Workshop 2009 held in conjunction with the Ninth SIAM International Conference on **data processing** (SDM 2009) 2009.
[17] I. Mcghee, J. Bayzick, A. Kontostathis, L. Edwards, A. Mcbride, and E. Jakubowski, "Learning **to spot** Internet Sexual Predation," International Journal on Electronic Commerce 2011, vol. 15, pp. 103-122, 2011.
[18] "NCMEC. National center for missing and exploited children", [online] October 2008, http://www.missingkids.com/en_US/documents/CyberTiplineFactSheet. pdf.
[19] Fundaci´on Barcelona Media (FBM). Caw 2.0 training datasets. [online] 2009, http://caw2.barcelonamedia.org.
[20] M. Ybarra, "Trends in technology-based sexual and non-sexual aggression over time and linkages to non-technology aggression." Presentation at the National Summit on Interpersonal Violence and Abuse Across the Lifespan: Forging a Shared Agenda. Houston, TX.2010.
[21] "Facts for families, the American Academy **of kid** Adolescent Psychiatry", [online], http://www.aacap.org/galleries/ FactsFor- Families/80_bullying.pdf.
[22] "Cyberbullying, The National Crime Prevention", [online] http://www.ncpc.org/cyberbullying
[23] D. Boyd, "Why youth (heart) social network sites: The role of networked publics in teenage social life." 2009.
[24] V. Nahar, X. Li, C. Pang, "An Effective Approach for Cyberbullying Detection". Journal of Communications in **IP** and Management Engineering .2013.
[25] M. Thangiah, S. Basri, S. Sulaiman, "A framework to detect cybercrime **within the** virtual environment". Computer & **informatics** (ICCIS), 2012 International Conference on, 2012.
[26] S. Argamon, M. Koppel, J. Fine, A. R. Shimoni, "Gender, genre, and **literary genre** in formal written texts". Text-The Hague Then Amsterdam Then Berlin-, 23(3), 321-346,2003.
[27] [online], http://cybersafety.wikispaces.com
[28] [online], http:/dev.aol.com/aim
[29][online],http://web.archive.org/web/20080308233204/http://dev.aol.com/aim/oscar
[30] [online], http://pidgin.im
[31] E. Villatoro-Tello, A. Juárez-González, H. J. Escalante, M. Montes-y-Gómez, and L. V. Pineda, "A Two-step Approach for Effective Detection of Misbehaving Users in Chats" . In CLEF (Online Working Notes/Labs/Workshop) 2012.
[32] eBlasterTM2008. [online], http://www.eblaster.com/
[33] Net NannyTM2008. [online], http://www.netnanny.com/
[34] IamBigBrother n.d. [online], http://www.iambigbrother.com/
[35] G. Eriksson, and J. Karlgren, " Features for modeling characteristics of conversations: Notebook for PAN at CLEF 2012." In CLEF 2012 Evaluation Labs and Workshop Online Working Notes. 2012.
[36] L. Gillam, and A. Vartapetiance, "Quite Simple Approaches for Authorship Attribution, Intrinsic Plagiarism Detection and Sexual Predator Identification".notebook for pan at clef 2012.
[37] C. Morris, and G. Hirst, "Identifying sexual predators by SVM classification with lexical and behavioral features" - notebook for pan at clef 2012.
[38] J. Parapar, D.E. Losada, A. Barreiro, "A learning-based approach for the identification of sexual predators in chat logs" - notebook for pan at clef 2012.
[39] G. Inches, and F. Crestani, "Overview of the international sexual predator identification competition at PAN-2012." In CLEF 2012 Evaluation Labs and Workshop — Working Notes Papers. Rome, Italy, 2012.
[40] Peersman, C., Vaassen, F., Asch, V.V., Daelemans, W.: Conversation level constraints on pedophile detection in chat rooms - notebook for pan at clef 2012.
[41] D.V. Ayala, E. Castillo, D. Pinto, I. Olmos, and S. León, "Information retrieval and classification based approaches for the sexual predator identification" - notebook for pan at clef 2012
[42] A. Kontostathis, W. West, A. Garron, K. Reynolds, and L. Edwards, "Identify predators using chat code 2.0" - notebook for pan at clef 2012.
[43] J.M.G. Hidalgo, and A.A.C. Díaz, "Combining predation heuristics and chat-like features in sexual predator identification" - notebook for pan at clef 2012.
[44] I.S. Kang, C.K. Kim, S.J. Kang, and S.H. Na, "Ir-based k-nearest neighbor approach for identifying abnormal chat users" - notebook for pan at clef 2012.
[45] R. Kern, S. Klampfl, and M. Zechner, "Vote/veto classification, ensemble clustering and sequence classification for author identification" - notebook for pan at clef 2012.
[46] M. Popescu, and C. Grozea, "Kernel methods and string kernels for authorship analysis "-notebook for pan at clef 2012.
[47] [online], http://www. Perverted-Justice.com .2008.

[48] J. Bengel, S. Gauch, E. Mittur, and R. Vijayaraghavan," ChatTrack: Chat room topic detection using classification". In Intelligence and Security Informatics, pp. 266-277. Springer Berlin Heidelberg, 2004.

[49] J. R. Tetreault, E. Filatova, and M. Chodorow, "Rethinking grammatical error annotation and evaluation with the Amazon Mechanical Turk."In Proceedings of the NAACL HLT 2010 Fifth Workshop on Innovative Use of NLP for Building Educational Applications, pp. 45-48. Association for Computational Linguistics, 2010.

[50] C. Callison-Burch, "Fast, cheap, and creative: evaluating translation quality using Amazon's Mechanical Turk". In Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 1-Volume 1. pp. 286-295. Association for Computational Linguistics,2009.

[51] V.S. Sheng, F. Provost, and P.G. Ipeirotis, "Get another label? improving data quality and data mining using multiple, noisy labelers." In Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 614-622. ACM, 2008.