2019 Third IEEE International Conference on Robotic Computing (IRC)

Exploring Automatic Identification of Fantasy-Driven and Contact-Driven Sexual Solicitors

Tatiana Ringenberg, Kanishka Misra, Kathryn Seigfried-Spellar, Julia Taylor Rayz
Computer Information Technology
Purdue University
West Lafayette, United States
{tringenb, kmisra, kspellar, jtaylor1}@purdue.edu

***Abstract*—In an increasingly computer-mediated world, it is easier to start conversations with people online, making it convenient to build trust with people based on their online persona. This encourages people with malevolent intentions to take advantage of trustful youngsters by means of various internet-based messaging platforms and subsequently inflicting harm to people below the legal age if a physical meeting is arranged. While previous research has prioritized the automatic identification of online solicitors by means of analyzing online chat conversations, little work has been done in terms of studying the conversational features of chats involving fantasy-driven and contact-driven online solicitors. This study focuses on 271 conversations extracted from Perverted-Justice, an online repository of chats between a predator and a volunteer pretending to be underage and presents results from training a model to triage between fantasy and contact-driven online solicitors.**

***Index Terms*—Online Predator Identification, Machine Learning, Natural Language Processing**

I. Introduction

In 2013, Sweetie, a virtual agent who mimicked the appearance of a 10 year old girl, was used by researchers to communicate with individuals interested in webcam child-sex [1]. While Sweetie was only live for 10 weeks, over 1,000 predators provided enough information to Sweetie to be personally identified, and the information was passed on to Interpol. Though the Sweetie sting was considered a success, the agent still required a human to guide the conversation with the solicitor behind the scene. An autonomous multi-agent system that incorporates behavioral modeling, biometrics, and image analysis as inputs can greatly aid the movement towards a safer Internet experience. Such a tool will also help law enforcement who struggle to keep up with the large case loads of sexual solicitation [2].

The authors are currently developing a tool which will be used to model solicitor behavior online, with the goals of (1) understanding behavioral characteristics necessary for triage of physical meetings and (2) mimicking behavior of minors online to proactively identify solicitors in social media [3]. This paper presents a study encompassing the textual inputs from internet conversations involving online solicitors and examines the various behavioral and lexical features for the purpose of modelling solicitor behavior. Information gleaned from this analysis will be incorporated as one component in the multi-agent system described above. The presented research lies at the intersection of Humans and Computational Agents.

In 2017 alone, the National Center for Missing and Exploited Children (NCMEC) reports 10.2 million instances of suspected child exploitation [4]. Furthermore, researchers have noted this area suffers from under-reporting [5], [6]. Less is known about the incidence of online sexual solicitation specifically, however NCMEC labeled the goals of 3,592 of the offenders reported to the organization in 2017 and found 32% of offenders wanted to meet a child for sexual purposes and eight percent wanted to engage in cybersex with a child [4].

Online sexual solicitation is defined as the pursuit of an under-aged individual, by an adult, for sexual gratification in either the digital or physical domain [7]. Researchers further divide online sexual solicitors into fantasy-driven and contact-driven categories based on the solicitor's primary goal during an interaction with a child [2], [8], [9]. Contact-driven sexual solicitors seek to meet a minor in person for sexual purposes while fantasy-driven sexual solicitors seek to engage in cybersex with a minor [2], [8].

There are several works that exist in the domain of identifying sexual solicitors and the grooming stages prevalent in online conversations using behavioral and lexical features. These features are either manually [10], [11] or automatically derived either using Linguistic Inquiry Word Count (LIWC) [2], [11]–[14] or word/character n-grams [15]–[18] . While these works address the identification of predation, or grooming, little work exists on presenting automatic triage into fantasy and contact goals of the offenders. It stands to reason, then, the methods previously used to identify differences between non-offenders and offenders might also be useful in differentiating the offenders by goal (fantasy or contact) as well.

This study seeks to address this gap in research by examining behavioral and linguistic features used in previous studies on identify online sexual solicitation for the automatic identification of contact-driven and fantasy-driven sexual solicitors. Specifically, the authors are using the behavioral features

978-1-5386-9245-5/19/$31.00 ©2019 IEEE
DOI 10.1109/IRC.2019.00110

IEEE computer society

identified in the reference [15], as well as exploring all 93 categories (for previously reported categories see [2], [11]–[14]).

In order to identify salient features and salient feature categories (behavioral, LIWC, and n-grams) the authors use the forward selection technique [19]. The following LIWC categories will be present within the final model: pronouns, sexual words, body words, positive emotion words, and negative emotion words. Adding word unigrams, word bigrams, character bigrams, and character trigrams will improve identification, as in previous work on solicitation [16], [18], [20]. Finally, the authors hypothesize the ratio of lines within the chat will contribute to the model, as researchers have found that contact offenders are more likely to engage in self-disclosures which involves the mutual exchange of personal information between both solicitor and victim [2].

The remainder of the paper is divided into the background, methodology, data analysis, and conclusion. In the background section, the authors describe the feature categories and algorithms used in previous studies on automatically identifying sexual solicitors in general. The authors then relate these studies to the ways in which researchers believe fantasy and contact-driven offenders may differ. In the methodology section, corpus creation, re-sampling data, and feature selection are discussed. In the data analysis section, the results are presented and discussed. In the conclusion, closing remarks are made and limitations are summarized.

II. Background

To the best of our knowledge, little to no work has pursued the problem of the automatic triage or risk assessment of contact-driven sexual solicitors versus fantasy-driven sexual solicitors. The following review of background literature revolves around methods and features previously used in studies to identify conversations with sexual solicitors and characteristics which have been found to differ between contact-driven solicitors and fantasy-driven solicitors.

A. Identification of Online Sexual Solicitors and Grooming

Research on automatic identification of sexual solicitors has focused on labeling conversations which contain an offender and on differentiating between offenders, decoys, and children [3], [20]–[22].

Current methods of detecting online sexual solicitors, or the grooming stages in which they engage, revolve around the use of machine learning (generally logistic regression [23], support vector machines (SVM) [20], [24], [25], or neural networks [22], [26]) in combination with lexical and chat-based features [11], [15], [20], [22]. The reference [13] used the Language Inquiry Word Count (LIWC) tool to investigate the different stages of the grooming process throughout a conversation between decoys and sexual solicitors. The reference [12] used the Perverted Justice (PJ) chat corpus along with the LIWC tool to differentiate offenders and decoys on the use of sexual words, word usage, and clout. In another study, the authors used the PAN 2012 dataset [27] along with convolutional neural networks (CNN) in order to categorize chats with sexual solicitors from non-sexual solicitor chat conversations [22]. The paper used the F1-score as the feature selection criteria and Matthews Correlation Coefficient (MCC) [28] as a measure of success. Additionally, authors have found models trained outside of the domain were not able to represent the vocabulary in the chats [22]. Another paper [21] was able to differentiate chats containing solicitors from chats without solicitors using nave bayes networks along with features related to stylometry and linguistic features related to emotions, and family. The authors were able to achieve 94% accuracy with linguistic and stylometric features [21].

Each of these studies provides a unique view of the solicitor versus non-solicitor problem. However, none of these studies have approached the problem of the automatic triage of the various groups within the solicitor hierarchy.

B. Differences Between Contact-Driven and Fantasy-Driven Solicitors

Contact-driven sexual solicitors seek to converse with a child online with the goal of meeting the child in person for sexual purposes [2], [8]. Fantasy-driven sexual solicitors seek to engage in cybersex with a minor. While a fantasy-driven offender may discuss a physical meeting with a child, the fantasy-driven offender will not show up for a meeting [2], [8].

Previous works have suggested contact-driven and fantasy-driven offenders differ in several meaningful ways including demographics, deviant sexual interests, and chat progression [29], [30]. Fantasy offenders tend to be younger than contact-driven offenders. Fantasy offenders also appear to have more sexually deviant interests than contact-driven offenders [30]. Furthermore, fantasy-driven offenders declare their intent early on in a conversation with a victim [5]. Solicitors differ in their use of language in chats as well. The reference [2] reports contact offenders are more likely than fantasy offenders to use self-disclosure techniques related to emotion words and first-person pronouns. The authors also report self-disclosure messages were matched by self-disclosure messages of the adolescent [2]. Another study finds sexual solicitation offenders differ greatly in usage of sexual words [12]. The reference [31] finds the two groups differed significantly on discussions of adult sexual relationships. When the reference [32] used a different but related typology that grouped solicitors into cybersex-only, schedulers, cybersex/schedulers, and buyers, the authors found associations between offender type and cancellation of meetings and offender type and likelihood of mention child-specific and incest themes.

Finally, fantasy-driven solicitors have fewer criminogenic behaviors than other forms of sexual offenders, include contact-driven solicitors [33]. Additionally, contact-driven offenders are more likely than fantasy-driven offenders to re-offend [34]. As such, contact-driven offenders appear to pose a larger threat to children.

While researchers have identified several factors which appear to distinguish the two groups of solicitors, most of

TABLE I
FEATURE AND CATEGORY SELECTION

Feature	Category	Description
Message Ratio	Behavioral	Ratio of message count of solicitor:decoy
Word Ratio	Behavioral	Ratio of words used solicitor:decoy
Max Streaks	Behavioral	Maximum number of repeated messages by a user in the conversation
Mean Streaks	Behavioral	Average number of repeated messages by a user in the conversation
Question Count	Behavioral	Number of messages that are questions
LIWC buckets	Lexical	Percentage of each words that matches either of the LIWC categories
Word Unigrams	Lexical	tf-idf of words in the conversation (top 1000)
Word Bigrams	Lexical	tf-idf of bigrams in the conversation (top 1000)
Character Bigrams	Lexical	tf-idf of character bigrams in the conversation (top 100)
Character Trigrams	Lexical	tf-idf of character trigrams in the conversation (top 100)

the studies are small. Additionally, the studies which exist on fantasy-driven and contact-driven solicitors do not focus on automatic prediction of whether or not a solicitor will show up to meet a child.

III. METHODOLOGY

A. Corpus Selection

The corpus chosen for this analysis was the Perverted Justice (PJ) corpus of conversations between online sexual solicitors and adult volunteer decoys [35].

The PJ corpus comes from the non-profit organization Perverted Justice Foundation Inc. Volunteers at PJ are trained to imitate children and act as decoys in chat conversations with individuals soliciting minors for sex [35]. Often, the volunteers are overseen by members of Law Enforcement and, once the suspect is prosecuted or the case is abandoned, the transcripts of all chat conversations occurring with an offender are posted on the PJ website. The conversations start in regional chat rooms and transition into private means of communication (text messages, video chats, email, etc).

At the time of this writing, the PJ corpus contained 623 chat conversations in total, occurring as early as 2004. The PAN 2012 corpus uses a subset of the PJ corpus along with non-offender chats. Law enforcement corpora vary in size but are generally smaller than the PJ corpus and are more difficult to acquire due to the sensitive nature of the domain.

Both PAN [27] and PJ [12], [36] were used within previous literature. However, chats were pulled directly from PJ instead of PAN because the current study focused on identifying differences between offenders - conversations outside of offender conversations were superfluous and outside of the scope of this work.

Due to inconsistent formatting on the PJ website, only 379 chats were extracted along with the summaries and annotations created by the decoy for the purposes of this study. The authors then manually labeled the chats into fantasy-driven, contact-driven, or unknown based on certain criteria as described. Fantasy-driven individuals were categorized by the motivation of engaging in cybersex and not meeting in-person [8]. As such, the authors labeled a chat as fantasy-driven if the chat met one of the following conditions:

- The decoy stated the offender did not show up in the decoys summary.
- The decoy stated the chat was terminated early (either by the decoy or the solicitor).
- There was no evidence in the chat showing that the offender was leaving to meet the decoy.

The authors labeled a chat as contact-driven if the chat met one of the following conditions:

- The decoy stated the offender did show up in the summary.
- There was evidence in the chat showing the offender was leaving to meet the decoy.

Of the 379 chats the authors examined, 109 chats were labeled as unknown and were not used in the study. 59 chats were labeled as fantasy-driven and 212 chats were labeled as contact-driven, resulting in a corpus of 271 offender chats with a 21:79 split between classes. While small, the dataset was consistent with corpora within the sexual solicitation domain [13], [32], [33]. Additionally, the imbalanced nature of the data was consistent with the reference [36] who found the majority of their Perverted Justice corpus contained contact-driven chats.

B. Feature Set Extraction

The authors divided the extracted features into ”Lexical” and ”Behavioral” as done in studies involving the PAN 2012 corpus [15], [16], [18], [27]. Table I shows the summarized list and descriptions of the various behavioral and lexical features chosen in this study.

The authors used the behavioral features prevalent in studies regarding the detection of sexual solicitation [15]. While lexical features describe what the author of a message writes, behavioral features focus on the manner in which the authors use his/her words [15]. We examined all behavioral features we found in the literature and removed all time-based features due to the lack, or inconsistency, of time stamps within the corpus. Ultimately, the five behavioral features shown in Table I remained.

To create lexical features, the authors used the Linguistic Inquiry Word Count (LIWC) tool. The LIWC tool contained a total of 93 word/stem buckets which are organized by category. Categories included emotions, linguistic categories

TABLE II
CROSS-VALIDATED RESULTS IN APPLYING SVM WITH VARIOUS COSTS AND FEATURES

Features	SVM cost	f1 score	MCC
LIWC_social + LIWC_male	1	0.836	NA
LIWC_relativ + mean_streaks	10	0.800	-0.086
LIWC_verb	100	0.800	-0.086
LIWC_I + LIWC_hear + LIWC_anx	1000	0.727	0.046
LIWC_social + LIWC_male + word unigrams	1	0.791	NA
LIWC_social + LIWC_male + word bigrams	1	0.791	NA
LIWC_social + LIWC_male + character bigrams	1	0.791	NA
LIWC_social + LIWC_male + character trigrams	1	0.791	NA
LIWC_relativ + mean_streaks + word unigrams	10	0.776	-0.063
LIWC_relativ + mean_streaks + word bigrams	10	0.761	-0.090
LIWC_relativ + mean_streaks + character bigrams	10	0.716	-0.056
LIWC_relativ + mean_streaks + character trigrams	10	0.746	0.064
LIWC_verb + word unigrams	100	0.731	-0.130
LIWC_verb + word bigrams	100	0.776	0.066
LIWC_verb + character bigrams	100	0.716	-0.056
LIWC_verb + character trigrams	100	0.642	-0.144
LIWC_I + LIWC_hear + LIWC_anx + word unigrams	1000	0.746	-0.111
LIWC_I + LIWC_hear + LIWC_anx + word bigrams	1000	0.776	-0.063
LIWC_I + LIWC_hear + LIWC_anx + character bigrams	1000	0.716	-0.056
LIWC_I + LIWC_hear + LIWC_anx + character trigrams	1000	0.612	-0.240

(e.g., pronouns), and psychologically meaningful categories (e.g., drives). Category names were unique, however the words in each category were not mutually exclusive. Some words existed in multiple categories because they were homonyms. Some other words existed in multiple categories because the categories formed hierarchies one category could subsume part or all of another category.

The LIWC tool was chosen because of its support within the literature for the detection of solicitors versus non-solicitors [2], [11]–[14]. All 93 of the features within LIWC were examined due to the wide array of features which authors had studied in relation to sexual solicitation in general and to the grooming process. For instance, the reference [13] mapped 26 categories from LIWC to the various stages of the grooming process, while the reference [11] used 17. The study [14] performed a similar analysis in which the authors analyzed grooming stages. However, the reference [14] used 62 dimensions of LIWC to explore the grooming stages.

Another set of lexical features used were the word unigrams and bigrams [16], [18], as well as character n-grams [20] which had previously been successful in detecting sexual solicitation.

Before creating the features, the corpus had to undergo pre-processing. Annotations which had been provided by decoys in the chats were removed. URLs and emails were also removed from the corpus, as their existence would likely not affect LIWC results but would affect the creation of n-grams.

C. Feature Selection

The feature selection methodology chosen for this research was based on the reference [19]. The authors chose to examine SVM over neural networks because of the small amount of data in the corpus. While the reference [22] was successful in using a CNN for the detection of solicitors, the analysis was done on the PAN 2012 dataset which has a large quantity of negative examples of chats involving two or more non-predatory conversations. In this study, the authors were only interested in classifying the types of solicitors instead of differentiating between predatory and non-predatory conversations.

In the reference [19], the authors trained the SVM classifier on each of the features individually, using error-rate as a measure of performance. The top three features were used together as the starting point for forward selection, and each remaining candidate feature was subsequently examined. In our study, we examined each of the behavioral and lexical features individually, using the F1-measure. The feature with the highest F1 score was selected. Once a feature was selected, models were incrementally created by the addition of each of the remaining features individually to the first feature. The best model of all of the candidate models was selected and the process repeated with the remaining features. This incremental feature selection continued until adding additional features no longer improved the model's F1 score.

While [19] used 10-fold cross validation, the authors used five-fold cross validation due to the small size of the corpus. Additionally, the authors chose the F1-measure over error-rate for this analysis because a balance between precision and recall was preferable in real-world application in this domain as in [22]. Balance of prediction between classes was important because the authors ultimately wished to use this to help inform Law Enforcement - it would not be helpful if the classifier was able to predict one class with high accuracy and the other with low.

IV. RESULTS AND EVALUATION

The authors sought to determine whether or not the individual behavioral or Language Inquiry Word Count (LIWC) features improved the classification of the solicitors, as in the reference [19]. Forward selection was utilized, with a support

TABLE III
BEST MODELS BASED ON MCC CRITERION

Model	MCC	F1
LIWC_I + LIWC_hear + LIWC_anx	0.046	0.727
LIWC_relativ + mean_streaks + ch_trigrams	0.064	0.746
LIWC_verb + word_bigrams	0.066	0.776

vector machine (SVM), on the behavioral and LIWC features using various cost values (1, 10, 100, 1000). Multiple cost values were assessed due to the imbalanced nature of the data, however the cost values were chosen opportunistically. N-grams were then added to the final models for each cost value. The chosen metric for assessing the candidate models was the F1 score. For evaluation of how the models performed on the evaluation set, the authors also report each model's Matthews Correlation Coefficient (MCC) [28]. The authors chose MCC as a reporting metric to consider the proportion of each class of the confusion matrix created upon application of the model on the evaluation set. The F1 measure and MCC are computed as follows:

$$F1 = \frac{2TP}{2TP + FN + FP} \tag{1}$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{2}$$

Where TP is the number of True Positive, TN is the number of True Negatives, FP is the number of False Positives, and FN is the number of False Negatives.

Table II shows the results of the SVM model using behavioral, LIWC, and n-gram features along with varying costs. The common features here are the features that gave the best F1 scores in the forward selection experiment and these have been augmented by adding character and word level n-grams. It can be observed that the F1 scores decrease by adding word and character n-grams, except in the case of the model with *LIWC_I, LIWC_hear, LIWC_anx* features, where it increased in terms of the F1 scores. Augmenting the model with unigrams also results in better MCC values.

Only three out of the 20 models show positive MCC value, while the rest show either negative or undefined indicating the poor performance of this model on imbalanced data, especially in its ability to detect true negatives (fantasy-driven). The three models with a positive MCC scores are included in Table III.

While the third model appears to be the best, the MCC value implies the result is little better than random.

Additionally, from Table II, there is evidence to suggest behavioral features may not be effective predictors of whether or not an individual is contact-driven or fantasy-driven. Though five behavioral features were examined, mean_streaks was the only behavioral feature to show up in the top models. It is possible time-based features would have contributed positively, however we did not have the data to test this. Future work should focus on identifying the effectiveness of such features.

V. CONCLUSION

This work explores the automatic triage of contact-driven and fantasy-driven sexual solicitors in online chats with decoys. This research has implications for law enforcement, offender treatment, offender typology, and future research directions. A model that can successfully utilize conversational features can be used as a textual component towards a multi-agent system that triages between fantasy and contact solicitors.

The ability to triage fantasy-driven and contact-driven sexual solicitation cases could result in earlier apprehension of individuals who intend to sexually abuse children in person. Additionally, the triage process may assist in the identification of individuals who are likely to re-offend.

The authors found LIWC features, word and character n-grams, and behavioral features were not enough to build a predictive model for differentiating between fantasy and contact-driven solicitors. This is an important finding in the sexual solicitation domain because it suggests the variability between types of offenders is too nuanced for the features typically used in solicitation research.

One limitation of this research which may have contributed to the lack of predictive power is the size of the corpus. Corpora in the sexual solicitation domain are small and even more difficult to work with once divided into yet smaller class subsets. Additionally, like most corpora in the area, the Perverted Justice corpora appears to be imbalanced with a higher proportion of contact-driven offenders. This is unsurprising considering in 2017 NCMEC found only eight percent of sexual solicitors wished to engage in cybersex with youths while 32% desired physical interaction [1].

Though the LIWC features were sparse due to a lack of data and a large dimensional space, the authors believe lexical features may not be sufficient to address salient linguistic components of the text. By only considering LIWC word categories, the models lose all of the relationships between words and concepts in the text. Future work should focus on features which take into account not only word distribution but semantic context. However, the authors acknowledge the difficulty of this task given the lack of publicly available data within the domain.

The labeling of sexual solicitors may also have been a limitation. The authors chose an approach in which fantasy and contact solicitation chats are distinct classes. However, based on the poor performance of the models presented in the paper and the low MCC values, the authors believe the fantasy versus contact distinction is likely fuzzy. In future work, the application of a fuzzy classifier may contribute to better performance as well as better examination of features and their contribution to the membership of a conversation within each class.

REFERENCES

[1] "'sweetie 2.0: Stop webcam childsex'," Jul 2018. [Online]. Available: https://www.terredeshommes.nl/en/programmes/sweetie-20-stop-webcam-child-sex

[2] M. M. Chiu, K. C. Seigfried-Spellar, and T. R. Ringenberg, "Exploring detection of contact vs. fantasy online sexual offenders in chats with minors: Statistical discourse analysis of self-disclosure and emotion words," *Child Abuse & Neglect*, vol. 81, pp. 128–138, Jul. 2018.

[3] K. Seigfried-Spellar, T. Ringenberg, M. Chiu, and M. Rogers, "Distinguishing contact child sex offenders vs. non-contact solicitors: Toward a digital forensics tool for automatic analysis of their chats with minors," in *Presented at the LEIU/IALEIA annual training conference in Bloomington, MN*, 2017.

[4] "Online Enticement of Children: An In-Depth Analysis of CyberTipLine Reports." National Center for Missing and Exploited Children, 2017.

[5] J. A. Kloess, A. R. Beech, and L. Harkins, "Online Child Sexual Exploitation: Prevalence, Process, and Offender Characteristics," *Trauma, Violence, & Abuse*, vol. 15, no. 2, pp. 126–139, Apr. 2014.

[6] H. C. Whittle, C. Hamilton-Giachritsis, and A. R. Beech, "Victims voices: The impact of online grooming and sexual abuse," *Universal Journal of Psychology*, p. 13, 2013, read w/ comments.

[7] S. Webster, J. Davidson, A. Bifulco, P. Gottschalk, V. Caretti, T. Pham, J. Grove-Hills, C. Turley, C. Tompkins, S. Ciulla *et al.*, "European online grooming project final report," *European Commission Safer Internet Plus Programme, Tech. Rep*, 2012.

[8] K. J. Mitchell, L. M. Jones, D. Finkelhor, and J. Wolak, "Trends in Unwanted Online Experiences and Sexting : Final Report," p. 59.

[9] H. L. Merdian, N. Moghaddam, D. P. Boer, N. Wilson, J. Thakker, C. Curtis, and D. Dawson, "Fantasy-Driven Versus Contact-Driven Users of Child Sexual Exploitation Material: Offender Classification and Implications for Their Risk Assessment," *Sexual Abuse*, vol. 30, no. 3, pp. 230–253, Apr. 2018.

[10] I. McGhee, J. Bayzick, A. Kontostathis, L. Edwards, A. McBride, and E. Jakubowski, "Learning to identify internet sexual predation," *International Journal of Electronic Commerce*, vol. 15, no. 3, p. 103122, 2011, read with comments.

[11] A. Gupta, P. Kumaraguru, and A. Sureka, "Characterizing pedophile conversations on the internet using online grooming," *arXiv preprint arXiv:1208.4324*, 2012.

[12] M. Drouin, R. L. Boyd, J. T. Hancock, and A. James, "Linguistic analysis of chat transcripts from child predator undercover sex stings," *The Journal of Forensic Psychiatry & Psychology*, vol. 28, no. 4, pp. 437–457, Jul. 2017.

[13] P. J. Black, M. Wollis, M. Woodworth, and J. T. Hancock, "A linguistic analysis of grooming strategies of online child sex offenders: Implications for our understanding of predatory sexual behavior in an increasingly computer-mediated world," *Child Abuse & Neglect*, vol. 44, pp. 140–149, Jun. 2015.

[14] A. Cano Basave, M. Fernandez, and H. Alani, *Detecting Child Grooming Behaviour Patterns on Social Media*, Nov 2014, vol. 8851.

[15] C. Morris and G. Hirst, "Identifying sexual predators by svm classification with lexical and behavioral features." in *CLEF (Online Working Notes/Labs/Workshop)*, vol. 12, 2012, p. 29.

[16] G. Eriksson and J. Karlgren, "Features for modelling characteristics of conversations: Notebook for pan at clef 2012," in *CLEF 2012 Evaluation Labs and Workshop, Rome, Italy, 17-20 September 2012*, 2012.

[17] M. Popescu and C. Grozea, "Kernel methods and string kernels for authorship analysis notebook for pan at clef," in *Conference and Labs of the Evaluation Forum, Rome-Italy*, 2012.

[18] 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.

[19] L. M. Stuart, S. Tazhibayeva, A. R. Wagoner, and J. M. Taylor, "On Identifying Authors with Style," in *2013 IEEE International Conference on Systems, Man, and Cybernetics*, Oct. 2013, pp. 3048–3053.

[20] D. Bogdanova, P. Rosso, and T. Solorio, "Exploring high-level features for detecting cyberpedophilia," *Computer Speech & Language*, vol. 28, no. 1, pp. 108–120, Jan. 2014.

[21] ——, "On the impact of sentiment and emotion based features in detecting online sexual predators," in *Proceedings of the 3rd Workshop in Computational Approaches to Subjectivity and Sentiment Analysis*. Association for Computational Linguistics, 2012, pp. 110–118.

[22] M. Ebrahimi, C. Y. Suen, and O. Ormandjieva, "Detecting predatory conversations in social media by deep Convolutional Neural Networks," *Digital Investigation*, vol. 18, pp. 33–49, Sep. 2016.

[23] H. Pranoto, F. E. Gunawan, and B. Soewito, "Logistic models for classifying online grooming conversation," *Procedia Computer Science*, vol. 59, p. 357365, Jan 2015.

[24] F. E. Gunawan, L. Ashianti, and N. Sekishita, "A simple classifier for detecting online child grooming conversation," *TELKOMNIKA; Yogyakarta*, vol. 16, no. 3, p. 12391248, Jun 2018.

[25] S. J. Pandey, I. Klapaftis, and S. Manandhar, "Detecting predatory behaviour from online textual chats," in *International Conference on Multimedia Communications, Services and Security*. Springer, 2012, pp. 270–281.

[26] D. Liu, C. Y. Suen, and O. Ormandjieva, "A novel way of identifying cyber predators," *arXiv preprint arXiv:1712.03903*, 2017.

[27] G. Inches and F. Crestani, "Overview of the international sexual predator identification competition at pan-2012." in *CLEF (Online working notes/labs/workshop)*, vol. 30, 2012.

[28] B. W. Matthews, "Comparison of the predicted and observed secondary structure of t4 phage lysozyme," *Biochimica et Biophysica Acta (BBA)-Protein Structure*, vol. 405, no. 2, pp. 442–451, 1975.

[29] D. Shannon, "Online Sexual Grooming in SwedenOnline and Offline Sex Offences against Children as Described in Swedish Police Data," *Journal of Scandinavian Studies in Criminology & Crime Prevention*, vol. 9, no. 2, pp. 160–180, Sep. 2008.

[30] K. M. Babchishin, R. K. Hanson, and C. A. Hermann, "The Characteristics of Online Sex Offenders: A Meta-Analysis," *Sexual Abuse: A Journal of Research and Treatment*, vol. 23, no. 1, pp. 92–123, Mar. 2011.

[31] M. A. McManus, L. Almond, B. Cubbon, L. Boulton, and I. Mears, "Exploring the Online Communicative Themes of Child Sex Offenders," *Journal of Investigative Psychology and Offender Profiling*, vol. 13, no. 2, pp. 166–179, Jun. 2016.

[32] D. DeHart, G. Dwyer, M. C. Seto, R. Moran, E. Letourneau, and D. Schwarz-Watts, "Internet sexual solicitation of children: a proposed typology of offenders based on their chats, e-mails, and social network posts," *Journal of Sexual Aggression*, vol. 23, no. 1, pp. 77–89, Jan. 2017.

[33] P. Briggs, W. T. Simon, and S. Simonsen, "An Exploratory Study of Internet-Initiated Sexual Offenses and the Chat Room Sex Offender: Has the Internet Enabled a New Typology of Sex Offender?" *Sexual Abuse*, vol. 23, no. 1, pp. 72–91, Mar. 2011.

[34] J. A. McCarthy, "Internet sexual activity: A comparison between contact and non-contact child pornography offenders," *Journal of Sexual Aggression*, vol. 16, no. 2, pp. 181–195, Jul. 2010.

[35] "Perverted-Justice.com - The largest and best anti-predator organization online." [Online]. Available: http://perverted-justice.com/

[36] G. M. Winters, L. E. Kaylor, and E. L. Jeglic, "Sexual offenders contacting children online: an examination of transcripts of sexual grooming," *Journal of Sexual Aggression*, vol. 23, no. 1, pp. 62–76, Jan. 2017.