1

Text Mining and CyberCrime

April Kontostathis, Lynne Edwards, Amanda Leatherman

Ursinus College

This chapter describes the state of technology for studying Internet crimes against children, specifically sexual predation and cyberbullying. We begin by presenting a survey of relevant research articles that are related to the study of cybercrime. This survey includes a discussion of our work on the classification of chat logs that contain bullying or predatory behavior. Many commercial enterprises have developed parental control software to monitor these behaviors, and the latest version of some of these tools provides features that profess to protect children against predators and bullies. The chapter concludes with a discussion of these products and offers suggestions for continued research in this interesting and timely sub-field of text mining.

1.1 Introduction

According to the most recent 2008 online victimization research, approximately 1 in 7 youth (ages 10 to 17-years-old) experience a sexual approach or solicitation by means of the Internet [NCMEC (2008)]. In response to this growing concern, law enforcement collaborations and non-profit organizations have been formed to deal with sexual exploitation on the Internet. Most notable is the Internet Crimes Against Children (ICAC) task force [Internet Crimes Against Children (n.d.)]. The ICAC Task Force Program was created to help state and local law enforcement agencies enhance their investigative response to offenders who use the Internet, social networking websites, or other computer technology to sexually exploit children. The program is currently composed of 59 regional Task Force agencies and is funded by the United States Department of Justice, Office of Juvenile Justice and Delinquency Prevention.

The National Center for Missing and Exploited Children (NCMEC) has set up a CyberTipLine for reporting cases of child sexual exploitation including child pornography, online enticement of children for sex acts, molestation of children outside the family, sex tourism of children, child victims of prostitution, and unsolicited obscene material sent to a child. All calls to the tip line are referred to appropriate law enforcement agencies –

and the magnitude of the calls is staggering. From March, 1998, when the CyberTipLine began operations, until April 20th, 2009, there were 44,126 reports of "Online Enticement of Children for Sexual Acts," one of the reporting categories. There were 146 in the week of April 20th, 2009 alone [NCMEC (2008)].

The owners of Perverted-Justice.com (PJ) began a grass-roots effort to identify cyberpredators in 2002. PJ volunteers pose as youth in chat rooms and respond when approached by an adult seeking to begin a sexual relationship with a child. We are currently working with the data collected by PJ from these conversations in an effort to understand cyberpredator communications.

Cyberbullying, according to the National Crime Prevention Council, is using the Internet, cell phones, video game systems, or other technology to send or post text or images intended to hurt or embarrass another person – and is a growing threat among children. In 2004, half of U.S. youth surveyed stated that they or someone they knew had been victims or perpetrators of cyberbullying [National Crime Prevention Council (n.d.a)]. Being a victim of cyberbullying is a common and painful experience. Nearly 20 percent of teens had a cyberbully pretend to be someone else in order to trick them online, getting the victim to reveal personal information. Seventeen percent of teens were victimized by someone lying about them to others online. Thirteen percent of teens learned that a cyberbully was pretending to be them while communicating with someone else. Ten percent of teens were victimized by someone posting unflattering pictures of them online, without permission [National Crime Prevention Council (n.d.b)].

The anonymous nature of the Internet may contribute to the prevalence of cyberbullying. Kids respond to cyberbullying by avoiding communication technologies or messages all together. They rarely report the conduct to parents (for fear of losing phone/Internet privileges) or to school officials (for fear of getting into trouble for using cellphones or Internet in class [Agatston et al. (2007); Williams and Guerra (2007)].

As we analyzed cyberbullying and cyberpredator transcripts from a variety of sources, we were struck by the similar communicative tactics employed by both cyberbullies and cyberpredators – in particular, masking identity and deception. We were also struck by the similar responses of law enforcement and youth advocacy groups: reporting and preventing. Victims are physically and psychologically abused by predators and bullies who trap them in vicious communicative cycles using modern technologies; their only recourse is to report the act to authorities after it has occurred. By the time a report is made, unfortunately the aggressor has moved on to a new victim.

Cyberbullying and Internet predation frequently occur over an extended period of time and across several technological platforms (i.e., chat rooms, social networking sites, cell phones, etc.). Techniques that link multiple online identities would help law enforcement and national security agencies identify criminals, as well as the forums in which they participate. The threat to youth is of particular interest to researchers, law enforcement and youth advocates because of the potential for it to get worse as membership in online communities continues to grow [Backstrom et al. (2006); Kumar et al. (2004); Leskovec et al. (2008)] and as new social networking technologies emerge [Boyd and Ellison (2007)]. Much of modern communication takes place via online chat media in virtual communities populated by millions of anonymous members who use a variety of chat technologies to maintain virtual relationships based on daily (if not hourly) contact [Ellison et al. (2007); O'Murchu et al. (2004)]. MSN Messenger, for example, reports 27 million users and AOL Instant Messenger has the largest share of

the instant messaging market (52% as of 2006) [IM MarketShare (n.d.)]; however, Facebook, the latest social networking craze, reported over 90 million users worldwide [Nash (2008)]. These media, along with MySpace, WindowsLive, Google, and Yahoo all have online chat technologies that can be easily accessed by anyone who chooses to create a screen-name and log-on; no proof of age, identity, or intention required. A recent update to Facebook also allows users to post and receive Facebook messages via text messaging on their cell phones [FacebookMobile (n.d.)].

We describe the current state of research in the areas of cyberbullying and Internet predation in Section 1.2. In Section 1.3, we describe several commercial products which claim to provide chat and social networking site monitoring for home use. Finally in Section 1.4 we offer our conclusions and discuss opportunities for future research into this interesting and timely field.

1.2 Current research in Internet Predation and Cyberbullying

This section provides a summary of research into Internet predation and cyberbullying. We first review the technology that is available for capturing Internet Messager (IM) and Internet Relay Chat (IRC). Next we discuss the datasets that are currently available for research in the area. Finally we survey several research articles for both Internet predation and cyberbullying detection, as well as a summary of the literature as it relates to legal issues.

1.2.1 Capturing IM and IRC chat

Data collection is the first step in any research project in text mining. Data collection for the study of cybercrime needs to focus primarily on capturing data from chatrooms and social networking sites; however, there are both legal and technical issues that must be overcome. In this section we discuss the work by several research groups who have successfully captured online chat.

In [Dewes et al. (2003)], Dewes, et al. use a multi-layered approach for capturing web chat from various sources including IRC and Web-based (both HTTP and java) chat systems. They begin by casting a wide net, essentially capturing all network traffic that passes through a particular router. Several filters are then applied to separate the chat traffic from non-chat traffic. Early experiments show that 91.7% of the chat traffic can be identified (recall) and 93.7% of the traffic that is captured is indeed chat (precision).

Other research groups take a more direct approach. Gianvecchio, et al. signed into Yahoo chatrooms and logged all posts for a two week period in order to capture data for their bot detection study [Gianvecchio et al. (2008)]. Others set up host servers and monitor all activity directly at the server level [Cooke et al. (2005)]. Several low cost commercial products for capturing relevant network packets are also available [ICQ-Sniffer (2009)].

1.2.2 Current collections for use in analysis

There is very little reliable labeled data about predator communications; much of the work that has appeared in both computer science and communication studies forums is focused on anecdotal evidence and chat log transcripts from Perverted Justice (PJ) [Perverted-Justice.com (2008)]. Perverted-Justice.com began as a grass-roots effort to

identify cyberpredators. PJ volunteers pose as youth in chat rooms and respond when approached by an adult seeking to begin a sexual relationship with a minor. When these activities result in an arrest and conviction, the chat log transcripts are posted online. New chat logs continue to be added to the web site. There were 325 transcripts, representing arrests and convictions, on the site as of July 2009. Details about early research projects that use this data are described in Section 1.2.4.

The use of PJ transcripts for research into cyberpredation is controversial. The logs contain transcripts of conversations between a predator and a pseudo-victim, an adult posing as a young teenager. However, the predators who participated in these conversations were convicted based, at least in part, on the content of the chat logs, which provides a measure of credibility to the data. We have been in communication with several researchers who are working on related projects in Computer Science, Media and Communication Studies, Criminal Justice, and Sociology and have not been able to identify another source of data. We will continue to seek transcripts that contain conversations between predators and minors; however, it will be extremely difficult. Law enforcement agencies are rarely able to share chat log transcripts (when they have them), even for scholarly examination, because the logs are not stored in a central repository and only excerpts are used when cases go to trial [Personal Communication (2008)].

A second dataset was created by Dr. Susan Gauch, University of Arkansas, who collected chat logs during a chat room topic detection project [Bengel et al. (2004)]. Dr. Gauch's project included the development of a crawler that downloaded chat logs (ChatTrack). Unfortunately, the software is no longer available. This chat data, although somewhat dated, has been used in some of the preliminary studies involving an analysis of predator communications [Kontostathis et al. (2009)].

We have identified one additional publically available dataset which can be used for research on the communication styles of cybercriminals. In 2009, the Content Analysis for the Web 2.0 workshop (held in conjunction with WWW2009), proposed three independent shared tasks: text normalization, opinion and sentiment analysis, and misbehavior detection. The misbehavior detection task addressed the problems of detecting inappropriate activity in which some users in a virtual community are harassing or offensive to some other members of the community. A common training dataset was made available to all task participants. The provided dataset was intended as a representative sample of what can be found in Web 2.0. The data have been collected from five different public sites, including Twitter, MySpace, Slashdot, Ciao, and Kongregate. Interested parties should refer to the CAW 2.0 website for additional information [CAW2.0 (n.d.)]. This data is exclusively intended for research purposes. A research project which used this data to detect cyberbullying is discussed in Section 1.2.5.

1.2.3 Analysis of IM and IRC chat

Much of the social networking research in computer science has focused on chat room data [Jones et al. (2008); Muller et al. (2003)]. A lot of this work has centered on identifying discussion thread sub-groups within a chat forum [Acar et al. (2005); Camtepe et al. (2004)]; and some researchers focus on the technical difficulties encountered when trying to parse chat log data [Tuulos and Tirri (2004); Van Dyke et al. (2009)]. Surprisingly few researchers have attempted to deal with the creation of specific applications for analysis and management

of Internet predators or cyberbullies. The few that we have identified are described in the following sections.

1.2.4 Internet Predation Detection

We have identified articles that take two different approaches to detection of cyberpredator communications. The first uses a bag of words approach and a standard statistical classification technique. The second leverages research in Communications Theory to develop more sophisticated features for input to the classifier.

A Statistical Approach

Pendar used the Perverted Justice transcripts to separate predator communication from victim communication [Pendar (2007)]. In this study, the author downloaded the Perverted Justice transcripts and indexed them. After preprocessing to reduce some of the problems associated with Internet communication (i.e. handling netspeak), the author developed attributes for each chat log. The attributes consisted of word unigrams, bigrams and trigrams. Terms that appeared in only one log or in more than 95% of the logs were removed from the index. Afterwards approximately 10,000 unigrams, 43,000 bigrams, and 13,000 trigrams remained. The author describes using 701 log files [1]. Each log file was split into victim communication and predator communication, resulting in 1402 total input instances, each with 10,000-43,000 attributes, depending on the model being tested. Additional feature extraction and weighting completed the indexing process.

The data file was split into a 1122 instance training set and a 280 instance test set, stratified by class (i.e. the test set contained 140 predator instances and 140 victim instances). Classification was then attempted using both support vector machine (SVM) and distance-weighted k nearest neighbor (k-NN) classifiers. The F-measure reported by the author ranged from .415 - .943. The k-NN classifier was a better classifier for this task and trigrams were shown to be more effective than unigrams and bigrams. The maximum performance (f-measure=.943) was obtained when 30 nearest neighbors were used and 10000 trigrams were extracted and used as attributes.

An approach based on Communicative Theory

In contrast to the purely statistical methods employed by Pendar, Kontostathis, et al. used a rule-based approach in [Kontostathis et al. (2009)]. This project integrates communication and computer science theories and methodologies to develop tools to protect children from cyberpredators.

The theory of luring communication provides a model of the communication processes that child sexual predators use in the real world to entrap their victims [Olson et al. (2007)]. This model consists of 3 major stages:

1. gaining access to the victim,
2. entrapping the victim in a deceptive relationship,

[1] it appears as if the perverted-justice.com site has changed their method of presenting the chat data in recent years

3. initiating and maintaining a sexually abusive relationship.

During the gaining access phase, the predator maneuvers himself into professional and social positions where he can interact with the child in a seemingly natural way, while still maintaining a position of authority over the child. For example, gaining employment at an amusement park or volunteering with a community youth sports team. The next phase, entrapping the victim in a deceptive relationship, is a communicative cycle that consists of grooming, isolation and approach. Grooming involves subtle communication strategies that desensitize victims to sexual terminology and reframe sexual acts in child-like terms of play or practice. In this stage, offenders also isolate their victims from family and friend support networks before approaching the victim for the third phase: sexual contact and long-term abuse.

In previous work, we expanded and modified the luring theory to accommodate the difference between online luring and real world luring [Leatherman (2009)]. For example, the concept "gaining access" was revised to include the initial entrance into the online environment and initial greeting exchange by offenders and victims, which is different from meeting kids at the amusement park or through a youth sports league. Communicative desensitization was modified to include the use of slang, abbreviations, net speak, and emoticons in online conversations. The core concept underpinning entrapment is the ongoing deceptive trust that develops between victims and offenders. In online luring communications, this concept is defined as perpetrator and victim sharing personal information, information about activities, relationship details, and compliments.

Communications researchers define two primary goals for content analysis [Riffe et al. (1998)]:

1. To describe the communication
2. To draw inferences about its meaning

In order to perform a content analysis for Internet predation, we developed a codebook and dictionary to distinguish among the various constructs defined in the luring communication theoretical model. The coding process occurred in several stages. First, a dictionary of luring terms, words, icons, phrases, and netspeak for each of the three luring communication stages was developed. Second, a coding manual was created. This manual has explicit rules and instructions for assigning terms and phrases to their appropriate categories. Finally, software that mimics the manual coding process was developed (this software is refered to as ChatCoder below).

Twenty-five transcripts from the Perverted Justice website were carefully analyzed for the development of the dictionary. These 25 online conversations ranged from 349 to 1500 lines of text. The perpetrators span from 23 to 58 years of age, were all male, and were all convicted of sexual solicitation of minors over the Internet.

We captured key terms and phrases that were frequently used by online sexual predators, and identified their appropriate category labels within the luring model: deceptive trust development, grooming, isolation and approach [Leatherman (2009); Olson et al. (2007)]. The dictionary included terms and phrases common to net culture in general, and luring language in particular. Some examples appear in Table 1.1. The version of coding dictionary used in these experiments contained 475 unique phrases. A breakdown of the phrase count by category appears in Table 1.2.

Table 1.1 Sample excerpt from Codebook for Internet Predation

Phrase	**Coding Category**
are you safe to meet	Approach
i just want to meet	Approach
i just want to meet and mess around	Approach
how cum	Communicative Desensitization
if i don't cum right back	Communicative Desensitization
i want to cum down there	Communicative Desensitization
i just want to gobble you up	Communicative Desensitization
you are a really cute girl	Compliment
you are a sweet girl	Compliment
are you alone	Isolation
do you have many friends	Isolation
let's have fun together	Reframing
let's play a make believe game	Reframing
there is nothing wrong with doing that	Reframing

Table 1.2 Dictionary Summary - Phrase Count by Category

Category	**Phrase Count**
Activities	11
Approach	56
Communicative Desensitization	220
Compliment	35
Isolation	43
Personal Information	29
Reframing	57
Relationship	24

In order to provide a baseline for the usefulness of the code book for detecting online predation, we ran two small categorization experiments. In the first experiment, we coded 16 transcripts in two ways: first we coded the predator dialogue (so only phrases used by the predator were recorded), and then we coded for the victim. Thus, we had 32 instances, and each instance had a count of the phrases in each of the coding categories (eight attributes). Our class attribute was binary (predator or victim).

We used the J48 classifier within the Weka suite of data mining tools [Witten and Frank (2005)] to build a decision tree to predict whether the coded dialogue was predator or victim. The J48 classifier builds a C4.5 decision tree with reduced-error pruning [Quinlan (1993)]. This experiment is similar to [Pendar (2007)], but Pendar used a bag-of-words approach and an instance-based learner. The classifier correctly predicted the class 60 percent of the time, a slight improvement over the 50 percent baseline. This is remarkable when we consider the fact that we were coding individuals who were in conversation with each other, and therefore the terminology used was similar. Stratified three-fold cross validation, as implemented within Weka, was used to evaluate the results.

In a second experiment we built a C4.5 decision tree to distinguish between PJ and ChatTrack transcripts. The ChatTrack dataset is described in Section 1.2.2. We coded 15 PJ transcripts (both victim and predator dialogue) and 14 transcripts from the ChatTrack data set [Bengel et al. (2004)]. The classifier that was built was able to distinguish the PJ transcripts 93 percent of the time. We also used stratified 3-fold cross validation for evaluation in these experiments.

As we analyzed the PJ transcripts, we noticed recurring patterns within the dialogue used by the suspects and began to wonder if we could cluster different types of predators via their language pattern usage.

We chose the k-means [Hartigan and Wong (1979)] clustering algorithm because it is known to be both simple and effective. The k-means algorithm partitions a set of objects into k sub-classes. It attempts to find the centers of natural clusters in the data by assuming that the object attributes form a vector space, and minimizing the intra-cluster variance. Thus, k-means generally forms tight, circular clusters around a centroid, and the algorithm outputs this centroid. k-means is particularly applicable to numeric attributes, and all of our attributes are numeric.

In our experiments, we counted the number of phrases in each of the eight coding categories for the 288 transcripts that were available on the PJ website as of August 2008 (predator only), and created an 8-dimensional vector for each instance. Thus, we used the same attributes that were used in the categorization experiments, but we were able to use all of the PJ transcripts. The vectors were column normalized by dividing by the maximum value in each column (i.e., all *activities* values were divided by the maximum value for *activities*). These vectors were then input to the k-means algorithm, and a set of clusters was determined.

The user must provide a value of k to the k-means clustering tool, and we were unsure about the number of categories of suspects that we might find, so we tried various values for k. We found that $k = 4$ produced the best result (the minimum intra-cluster variance), suggesting the hypothesis that there are four different types of Internet predators. More work is needed to determine labels for these categories of suspects. The centroid for each cluster appears in Figure 1.1. This figure clearly shows that some suspects spend more time overall with the victim (lines that are higher on the graph) and also that suspects in different clusters used different strategies during their conversations (as determined by line shape).

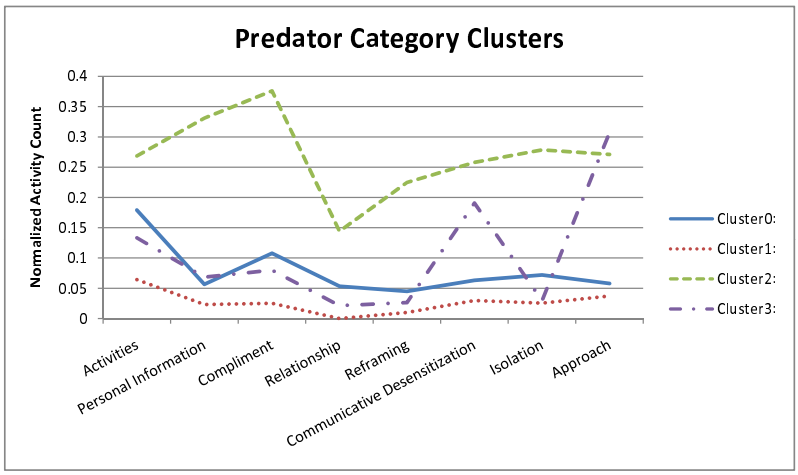

Figure 1.1 Initial clustering of predator type

For example, cluster 2 has a higher ratio of compliments vs. communicative desensitization as compared to cluster 3.

1.2.5 Cyberbullying detection

In 2006, the Conference on Human Factors in Computing Systems (CHI), ran a workshop on the misuse and abuse of interactive technologies, and in 2008 Rawn and Brodbeck showed that participants in first person shooter games had a high level of verbal aggression, although in general there was no correlation between gaming and aggression [Rawn and Brodbeck (2008)].

Most recently, in 2009 the Content Analysis for the Web 2.0 (CAW 2.0) workshop was formed and held in conjunction with WWW2009. As noted above, the CAW 2.0 organizers devised a shared task to deal with online harassment, and also developed a dataset to be used for research in this area. Only one submission was received for the misbehavior detection task. A brief summary of that paper follows.

Yin, et. al define harassment as communication in which a user intentionally annoys another user in a web community. In [Yin et al. (2009)] detection of harassment is presented as a classification problem with two classes: positive class for posts which contain harassment and negative class for posts which do not contain harassment.

The authors combine a variety of methods to develop the attributes for input to their classifier. They use standard term weighting techniques, such as TFIDF (term frequency - inverse document frequency) to extract index terms and give appropriate weight to each term. They also develop a rule-based system for capturing sentiment features. For example, a post that contains foul language and the word ‘you’ (which can appear in many forms in online communication) is likely to be an insult directed at someone, and therefore could be perceived as a bullying post. Finally, some web communities seem to engage in friendly banter or ‘trash talk’ that may appear to be bullying, but is instead just a communicative style.

The authors also were able to identify contextual features by comparing a post to a window of neighboring posts. Posts that are unusual or which generate a cluster of similar activity from other users are more likely to be harassing.

After extracting relevant features, the authors developed a SVM classifier for detecting bullying behavior in three of the six datasets provided by the CAW 2.0 conference organizers. They chose two different types of communities, Kongregate which captures IM conversations during game play, and Slashdot/MySpace which tend to be more asynchonous discussion-style forums where users write longer messages and discussion may continue over days or weeks. The authors manually labeled all three datasets. The level of harassment in general was very sparse. Overall only 42 of the 4802 posts in the Kongregate dataset represented bullying behavior. The ratio of bullying to nonbullying in Slashdot was similar (60 out of 4303 posts). MySpace was a little higher with 65 out of 1946 posts.

The authors employed an SVM to develop a model for classifying harassing posts. Their experiment results show that including the contextual and sentiment features improves the classification over the local weighting (TFIDF) baseline for all three datasets. The maximum recall was achieved with the chat style collection (recall .595 for Kongregate). Precision was best when the dataset contained more harassment (.417 for MySpace). Overall the F-measure ranged from .298 to .442, so there is much room for improvement. A random chance baseline would be less than 1%, however, so the experimental results show that detection of cyberbullying is possible.

1.2.6 Legal issues

Companies have long been aware of the potential for misuse of email for bullying and harassment. In [Sipior and Ward (1999)], the authors report on the increased litigation surrounding sexual harassment in the workplace, particularly harassment via email.

Internet predation and cyberbullying are relatively new crimes, and as such, the legal community is struggling to work with the technical community to protect victims while also protecting the civil rights of innocent users of Internet channels. Early attempts at collaboration between technicians and law enforcement, as described in a case study in [Axlerod and Jay (1999)], were initially frustrating. The collaborative work eventually paid off as the computer scientists learned what is (and is not) permitted under our legal system; and law enforcement officials learned to trust and use technical solutions to their best advantage.

In [Burmester et al. (2005)] the authors describe a combined hardware and software solution for providing law enforcement personnel with information in cases of cyberstalking. The article provides a profile of a technically advanced cyberstalker (who shares many traits with Internet predators and cyberbullies), as well as develops a solution that recognizes the very real constraints placed upon law enforcement officials, such as chain-of-custody issues, and providing proof of integrity of digital evidence.

1.3 Commercial software for monitoring chat

Many commercial products profess to provide parents with the tools to protect their children from Internet predators and cyberbullies. We provide a brief overview of several of popular products in this section.

Like most of the parental control products we identified, eBlaster™ records everything that occurs on a monitored computer and forwards the information to a designated recipient, but does not provide a mechanism for filtering or analyzing all the data it collects [eBlaster™ (2008)].

Net Nanny™ can also record everything, and offers multiple levels of protection for different users [Net Nanny™ (2008)]. The latest version of Net Nanny™ claims to send alerts to parents when it detects predatory or bullying interactions on a monitored computer. The alerts appear to be based on simple keyword matching [PC Mag (2008)].

IamBigBrother captures everything on the computer including chats, instant messages, email, and websites [IamBigBrother (n.d.)]. The program also records all Facebook and Myspace keystrokes, and captures all passwords typed. IamBigBrother can also take a picture of the screen when certain words are used. This feature allows parents to identify keywords that they are concerned about (personal information, foul language, sexual terms, etc.). Unfortunately, the program does not include pre-defined words; parents have to define problematic words themselves TopTenReviews (n.d.). The software also captures Internet activity from programs like America Online, MSN, and Outlook Express. The program can record incoming and outgoing Yahoo Mail, Hotmail, and Gmail. IamBigBrother can operate in stealth mode that cannot be detected by users. Users / children also cannot avoid IamBigBrother by clearing cache or history.

While IamBigBrother appears to focus primarily on keystroke capture and surveillance, Kidswatch Internet Security appears to focus more on blocking [TigerDirect (n.d.)]. The program allows parents to control their childrens' access to inappropriate web content and sends email notifications to parents when their children try to visit blocked or restricted sites. Parents can select content to be restricted from a list of over 60 categories. According to the Kidswatch web site: "Our dynamic content categorization technology attempts to categorize thousands, even millions, of websites based on content." Parents have the option to override restricted lists if they choose, and are encouraged to submit websites they think should be blocked to the software producer.

KidsWatch also supports chat protocols for Yahoo, MSN, ICQ, AIM, and Jabber. Parents receive email alerts when a "suspect phrase or word" is encountered in an online chat. The alert report can include the phrase or the entire conversation. The alerts are based on a customizable list of 1630 words and phrases. Although the surveillance and alert features are similar to the one featured in the NetNanny and IamBigBrother programs, Kidswatch takes this feature one step further by providing information about known sex offenders by providing the locations of sex offenders in the user's neighborhood.

Similar to other control programs, the Safe Eyes Parental Control program limits access to restricted sites that fall into 35 predetermined categories of website content [InternetSafety (n.d.)]. The program also prevents children from accidentally finding inappropriate sites. When restricted sites are accessed, parents are alerted by email, text message or phone call.

CyberPatrol provides filtering and monitoring features that can use the company's presets or can be customized by parents [CyberPatrol (n.d.)]. Several features that distinguish this program are the ability to customize settings for child, young teen, mature teen, or adult and the ability to block objectionable words and phrases commonly used by cyberbullies and predators. Parents receive weekly and daily reports on web pages visited and length of visits; however, there does not appear to be an alert feature.

Bsecure provides filtering — with "patent-pending technology and human review"

[Bsecure (n.d.)] that blocks offensive Websites from users' computers — and reporting options similar to other programs, but this program also offers an Application Control that allows parents to control music sharing, file sharing and instant messaging programs. The software appears to be similar to Cyberpatrol. Bsecure does not offer an alert feature.

The latest versions of Windows Vista and Apple's OS X 10.5 (Leopard) include integrated parental controls. Their features appear to be similar to most commercial monitoring and filtering products and neither operating system, unlike many commerical products, requires an annual subscription [Consumer Search (2008)]. Unfortunately neither product provides specific protection against predation or cyberbullying.

Finding information about AOL parental controls proved to be fairly difficult without an AOL userid and AOL installed. Like Windows Vista and OS X 10.5, AOL does not require installation of any additional software on the computer being monitored. There is no indication that AOL provides specific features for protection against Internet predators or cyberbullies.

McAfee and Norton are primary known as anti-virus and security software products. Both now offer parental control built in as well. As with the operating system products, the parental controls are designed to block specific web site and monitor online activity in general.

1.4 Conclusions and Future Directions

The Internet continues to grow and to reach younger audiences. Opportunities for connecting with classmates, friends, and people with shared interests abound. Email, online chat, and social networking sites allow us to interact with people in the same town and people on the other side of the world.

Unforunately, the opportunity for misuse comes with any new technology. There were sexual predators and bullies long before the advent of the Internet and chatrooms. Cyberbullying and Internet predation threaten minors, particular teens and tweens who do not have adequate supervision when they use the computer. As Internet connectivity moves to the cell phone, the portable gaming device and the multi-player gaming console, more avenues for contact and exploitation of youth become available.

Our literature review also shows that there are few scholars researching cyberpredation and cyberbullying. As more researchers enter this field, future research should attempt to be more proactive in addressing the role that newer technologies, particularly cellphones and peer-to-peer devices, play in new incarnations of cybercrime, like sexting. There is room for researchers in the fields of information retrieval and text mining to contribute solutions to these vexing problems. Classifiers that identify predatory behavior can be developed. New datasets can be collected, labeled, and distributed to other research groups. Collaborations with network engineers, psychologists, sociologists, law enforcement, and communications specialists can provide new insight into understanding, detecting, and stopping cybercrime.

Cybercrime continues to escalate and evolve as new technologies are introduced and as their popularity grows among young people. We have found only three research articles that use text mining techniques to classify cyberpredators and cyberbullies. This interesting and socially relevant subfield of text mining is begging for attention from the research community. The research to date provides a starting point for exploration - an exploration that moves away from solely focusing on the computer platform as the site of cybercrimes to studying the network level as bullying and predation move from text-only, and include audio and visual.

1.5 Acknowledgements

This work was supported in part by the Ursinus College Summer Fellows program. The authors thank Dr. Susan Gauch and her students for providing the ChatTrack data, and Dr. Nick Pendar for his helpful advice on acquiring the Perverted-justice.com transcripts. We also thank Fundación Barcelona Media (FBM) for compiling and distributing the CAW 2.0 shared task datasets. Our thanks extend to the many students and colleagues in both Mathematics and Computer Science and Media and Communication Studies departments at Ursinus College who have provided support and input to this project, as well as to the editors for their patience and feedback.

References

Acar E, Camtepe S, Krishnamoorthy M and Yener B 2005 Modeling and Multiway Analysis of Chatroom Tensors. *IEEE International Conference on Intelligence and Security Informatics*.

Agatston P, Kowalski R and Limber S 2007 Students perspectives on cyber bullying. *Journal of Adolescent Health*.

Axlerod H and Jay DR 1999 Crime and punishment in cyberspace: dealing with law enforcement and the courts. *SIGUCCS '99: Proceedings of the 27th annual ACM SIGUCCS conference on User services*, pp. 11–14.

Backstrom L, Huttenlocher D, Kleinberg J and Lan. X 2006 Group formation in large social networks: membership, growth, and evolution. *In Proceedings of the 12th ACM SIGKDD international Conference on Knowledge Discovery and Data Mining KDD '06*.

Bengel J, Gauch S, Mittur E and R.Vijayaraghavan. 2004 ChatTrack: Chat room topic detection using classification. *Second Symposium on Intelligence and Security Informatics*.

Boyd D and Ellison N 2007 Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication* **13**(1), 210–230.

Bsecure n.d. `http://www.bsafehome.com/Products/Family.aspx`.

Burmester M, Henry P and Kermes LS 2005 Tracking cyberstalkers: a cryptographic approach. *ACM SIGCAS Computers and Society* **35**(3), 2.

Camtepe S, Krishnamoorthy M and Yener B 2004 A tool for Internet chatroom surveillance. *Second Symposium on Intelligence and Security Informatics*.

CAW2.0 n.d. `http://caw2.barcelonamedia.org/`.

Consumer Search 2008 Parental control software review. `http://www.consumersearch.com/parental-control-software/review`.

Cooke E, Jahanian F and Mcpherson D 2005 The zombie roundup: Understanding, detecting, and disrupting botnets. *Workshop on Steps to Reducing Unwanted Traffic on the Internet (SRUTI)*, pp. 39–44.

CyberPatrol n.d. `http://www.cyberpatrol.com/family.asp`.

Dewes C, Wichmann A and Feldmann A 2003 An analysis of internet chat systems. *IMC '03: Proceedings of the 3rd ACM SIGCOMM conference on Internet measurement*, pp. 51–64.

eBlaster™ 2008. `http://www.eblaster.com/`.

Ellison N, Steinfield C and Lampe C 2007 The benefits of facebook "friends:" social capital and college students' use of online social network sites.. *Journal of Computer-Mediated Communication* **12**(4), 1143–1168.

FacebookMobile n.d. `http://www.facebook.com/mobile/`.

Gianvecchio S, Xie M, Wu Z and Wang H 2008 Measurement and classification of humans and bots in internet chat. *SS'08: Proceedings of the 17th conference on Security symposium*, pp. 155–169. USENIX Association, Berkeley, CA, USA.

Hartigan J and Wong MA 1979 A k-means clustering algorithm. *Applied Statistics* **28**(1), 100–108.

IamBigBrother n.d. `http://www.iambigbrother.com/`.

ICQ-Sniffer 2009. `icq-sniffer.qarchive.org/`.

IM MarketShare n.d. `http://www.bigblueball.com/forums/general-other-im-news/34413-im-market-share.html/`.

Internet Crimes Against Children n.d. `http://www.icactraining.org/`.

InternetSafety n.d. `http://www.internetsafety.com/safe-eyes-parental-control-software.php`.

Jones Q, Moldovan M, Raban D and Butler B 2008 Empirical evidence of information overload constraining chat channel community interactions. *Proceedings of the ACM 2008 Conference on Computer Supported Cooperative Work*.

Kontostathis A, Edwards L and Leatherman A 2009 Chatcoder: Toward the tracking and categorization of internet predators. *In Proc. Text Mining Workshop 2009 held in conjunction with the Ninth SIAM International Conference on Data Mining (SDM 2009)*.

Kumar R, Novak J, Raghavan P and Tomkins A 2004 Structure and evolution of blogspace. *Communications of the ACM* **47**(12), 35–39.

Leatherman A 2009 Luring language and virtual victims: Coding cyber-predators online communicative behavior. Technical report, Ursinus College, Collegeville, PA, USA.

Leskovec J, Lang KJ, Dasgupta A and Mahoney MW 2008 Statistical properties of community structure in large social and information networks *WWW '08: Proceeding of the 17th international conference on World Wide Web*, pp. 695–704.

Muller M, Raven M, Kogan S, Millen D and Carey K 2003 Introducing chat into business organizations: toward an instant messaging maturity model. *Proceedings of the 2003 international ACM SIGGROUP Conference on Supporting Group Work*.

Nash KS 2008 A peek inside Facebook. `http://www.pcworld.com/businesscenter/article/150489/a_peek_inside_facebook.html`.

National Crime Prevention Council n.d.a. `http://www.ncpc.org/topics/by-audience/teens/protect-yourself/cyberbullying`.

National Crime Prevention Council n.d.b. `http://www.ojp.gov/cds/internet_safety/NCPC/Stop%20Cyberbullying%20Before%20It%20Starts.pdf`.

NCMEC 2008 National center for missing and exploited children. `http://www.missingkids.com/en_US/documents/CyberTiplineFactSheet.pdf`.

Net Nanny™ 2008. `http://www.netnanny.com/`.

Olson L, Daggs J, Ellevold B and Rogers T 2007 Entrapping the innocent: Toward a theory of child sexual predators' luring communication. *Communication Theory* **17**(3), 231–251.

O'Murchu I, Breslin J and Decker S 2004 Online social and business networking communities. Technical report, Digital Enterprise Research Institute (DERI). `http://www.deri.ie/fileadmin/documents/DERI-TR-2004-08-11.pdf`.

PC Mag 2008 Netnanny 6.0. `http://www.pcmag.com/article2/0,2817,2335485,00.asp`.

Pendar N 2007 Toward spotting the pedophile: Telling victim from predator in text chats. *Proceedings of the First IEEE International Conference on Semantic Computing*, pp. 235–241.

Personal Communication 2008 Trooper Paul Iannace, Pennsylvania State Police, Cyber Crimes Division.

Perverted-Justice.com 2008 Perverted justice. `www.Perverted-justice.com`.

Quinlan R 1993 *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers.

Rawn RWA and Brodbeck DR 2008 Examining the relationship between game type, player disposition and aggression. *Future Play '08: Proceedings of the 2008 Conference on Future Play*, pp. 208–211.

Riffe D, Lacy S and Fico F 1998 *Analyzing Media Messages: Using Quantitative Content Analysis in Research*. Lawrence Erlbaum Associates.

Sipior JC and Ward BT 1999 The dark side of employee email. *Communications of the ACM* **42**(7), 88–95.

TigerDirect n.d. `http://www.tigerdirect.com/applications/SearchTools/item-details.asp?EdpNo=3728335&CatId=986`.

TopTenReviews n.d. `http://monitoring-software-review.toptenreviews.com/i-am-big-brother-review.html`.

Tuulos V and Tirri H 2004 Combining topic models and social networks for chat data mining. *Proceedings of the 2004 IEEE/WIC/ACM international Conference on Web intelligence*, pp. 235–241.

Van Dyke N, Lieberman H and Maes P 2009 Butterfly: a conversation-finding agent for Internet relay chat. *Proceedings of the 4th international Conference on intelligent User interfaces*.

Williams K and Guerra N 2007 Prevalence and predictors of Internet bullying. *Journal of Adolescent Health* **41**(6), S14–S21.

Witten I and Frank E 2005 *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann Publishers.

Yin D, Xue Z, Hong L, Davison BD, Kontostathis A and Edwards L 2009 Detection of harassment on Web 2.0. *Proceedings of the Content Analysis in the Web 2.0 (CAW2.0) Workshop at WWW2009*.

and Interaction. Article. Text Mining in Cybersecurity: Exploring Threats. and Opportunities. Maaike H. T. de Boer 1, * , Babette J. Bakker 2, Erik Boertjes 3, Mike Wilmer 1Â need to protect themselves against the overall harm of cybercrime, that is, the sum of the material. harms or costs, and the non-material harms of cybercrime. [2.]. As it is expected that the number of. Text Mining may be defined as the process of examining data to gather valuable information. Text mining, also known as text data mining involves algorithms of data mining, machine learning, statistics, and natural language processing, attempts to extract high quality, useful information from unstructured formats. The recent years have seen a tremendous increase in the adoption of text mining for business applications.Â The unidentified criminal soon becomes untraceable. Thanks to mining techniques, intelligence, and anti-crime applications are keeping cybercrimes at bay. Enterprise and law enforcement or intelligence agencies make use of text mining techniques to analyze the source and nature of data extraction. Customer Care Service According to Wikipedia, â€˜Text Mining is the discovery, by computer, of new previously unknown information, by automatically extracting information from different written resourcesâ€™. This mainly includes finding novel insights, trends or patterns from text-based data. Such novel insights can be highly essential in fields like business. The main sources of data for text mining is acquired from customer and technical support, emails and memos, advertising and marketing, human resources as well as other competitors. Index. Process of Text Mining. Relevance and Applications of Text Mining. Few Softwares Used in Text Mining. The Process of Text Mining. The process of text mining mainly involves five steps