Comparison of Rule-based to Human Analysis of Chat Logs

April Kontostathis, Lynne Edwards, Jen Bayzick, India McGhee,
Amanda Leatherman, and Kristina Moore

Ursinus College
Collegeville, PA USA
{akontostathis,ledwards}@ursinus.edu
http://www.ursinus.edu

Abstract. Key words: cyber-crime, internet predation, chat log analysis

1 Introduction

The National Center for Missing and Exploited Children reported in a 2008 survey that approximately 1 in 7 youth (ages 10-17) are approached or solicited for sexual purposes through the Internet. The NCMEC established a CyberTipLine for reporting cases of child sexual exploitation, and the magnitude of calls received is staggering, with 40,353 reports of "Online Enticement of Children for Sexual Acts" from March 1998 to October 2008. In this single category, 122 were reported in the week of October 27th, 2008 alone [7].

We are continuing a study that attempts to evaluate and categorize the strategies used by online sexual predators in their attempts to develop relationships with children using the Internet. Working with a program called Chat-Coder, we redesigned the method employed by the program to decide which lines in the chat log contain luring language. Communications researchers define two primary goals for content analysis [12]: to describe communication, and to draw inferences about its meaning. In this paper we describe a rule-based approach for identifying predatory communication within a chat log and compare our algorithmic effort to the results produced by trained communications analysts. The fifteen chat logs we have been working with come from the Perverted Justice (PJ) website [11] and are discussions between a sexual predator, who has been convicted of soliciting a minor over the Internet for sexual activity, and a trained PJ volunteer posing as an adolescent.

Like most of the commercial products we have seen, ChatCoder 1 was based on a simple keyword matching technique. This approach has many shortcomings. We have found that the rule-based technique used in ChatCoder 2 provides an overall improvement in intercoder reliability by a maximum of 13.21% and an average of 5.81%. Intercoder reliability is a well-known metric used by researchers in communication studies for determining the consistency of coding between two coders or coding systems.

A number of social networking sites involve chats or forum discussions, and as these sites draw younger users, they also draw cyber criminals. Our analysis of chat data can provide interesting insights and may inform the development of new tools for site monitoring.

2   Related Work

This project integrates both communication theories and computer science algorithms to create a program that can detect the occurence of predation in an online social setting. Though there has been much work exploring social media, this aspect of online social interaction remains largely untapped. Previous work by Kontostathis, et al. includes developing an optimized model of predation and a phrase-based program, ChatCoder 1, to analyze chat logs [5]. This surface-level analysis provided an excellent base from which to work with ChatCoder 2.

Though there have been a number of research projects involving the parsing of chat logs, there have been few that do so with predation in mind. To our knowledge, there are three such projects. Pendar has had some success analyzing chat transcripts to differentiate between the victim and the predator [10]. The study by Hughes, et al. focuses on the distribution of child pornography through peer-to-peer networks [4]. Recently, Adams and Martel conducted research to create a program that can detect and extract the topic of discussion [1]. There are a few computer programs that attempt to police chat conversations, but they are generally lacking in true analysis capabilities as none of them are based on communication theory[8][2].

3   Modeling Predation

Olson, et al. established a luring communication theoretical (LCT) model that defines five phases of predation: gaining access, deceptive trust development, grooming, isolation, and approach [9]. This model was expanded for online predation, and operationalized by Leatherman, et al. [6]

Gaining access, as defined by Olson, et al. represents the first step in the luring process wherein the predator must be "motivated and able to gain access to potential victims and their families." Primarily, then, gaining access involves exchanging personal traits of both the predator and the victim, as well as the strategic placement of the predator. In online predation, a predator gains access to a minor through mediums such as instant messaging forums, chat rooms, and social networking sites such as MySpace and Facebook. For instance, a predator would place himself in a chat room frequented by minors.

Olson, et al. define deceptive trust development as "a perpetrator's ability to cultivate relationships with potential victims and possibly their families that are intended to benefit the perpetrator's own sexual interest." In the arena of online predation, this is divided into four sub-categories: personal information, relationship details, activities, and compliments. The exchange of *personal information* involves details about the victim and predator's actual locations, ages,

names, computer locations (i.e. in the bedroom, basement, etc), birthdays, cell or home phone numbers, and pictures of themselves. *Relationship* information includes discussion of feelings and attitudes toward maintaining, building, and dismantling their relationships with each other, friends, significant others, and family members. *Activities*, a broad category, is defined primarily as preferred social behaviors shared by both the predator and victim, including but not limited to music, movies, books, sports, hobbies, and favorite foods. *Compliments* involve the predator or victim offering praise for one another's appearance, activities, and personal information with the intention of making the victim view the predator in a positive, appreciative light.

Deceptive trust occurs and is affirmed throughout the entire communication between the predator and the victim. Indeed, incurring the trust of the victim is essential to the success of the later stages of the entrapment cycle: isolation and approach. When the victim trusts the predator, the offender begins to groom the minor to accept offers of sexual contact. Grooming is "the subtle communication strategies that sexual abusers use to prepare their potential victims to accept the sexual conduct [9]." Thus communication that functions as grooming does not directly lead to sexual contact, but instead desensitizes the victim to sexual remarks or foul language. Successful grooming leaves the victim unaware that any process is underway. There are two sub-categories of grooming - communicative desensitization and reframing.

Communicative desensitization refers to the offender purposefully and frequently using vulgar sexual language in an attempt to desensitize the victim to its use. Additionally, the perpetrator will often attempt to encourage the minor's interest in sexual subjects with the eventual goal of perpetrating future abuse. In terms of online predation, this can be achieved by sending pornographic images and using sexual slang terms or netspeak in lieu of every day words (i.e. "cum" instead of "come").

Reframing occurs when sex offenders endeavor to make the victim comfortable with experiencing sexual advances over the internet. From Olson, et al., reframing is "contact or sex play between victim and adults that may be communicated in ways that would make it beneficial to the victim later in life." To this end, sexual conversation is presented in a positive light and is often referred to as a learning experience, a game to be played, or an important skill to learn in order to participate in loving relationships in the future.

Beyond grooming, physically and emotionally *isolating* the victim is essential to the sexual predator, be it online or in the real world. Physical isolation is defined as arranging to spend time alone with the victim, and mental isolation is increasing emotional dependency upon the predator for things like friendship and guidance. While complete physical isolation cannot occur over the Internet, the predator achieves isolation by making sure the victim chats without supervision. Predation is most successful with minors who are isolated from support networks, be it by low paternal or maternal relationships or by having very few friends. This information is gleaned through online communications by asking questions about the minor's social life, by providing sympathy and support in reaction to

their situation, and by questioning the strict rules of the parent. The predator seeks to isolate the victim and then to fill the social gaps in the victim's life as a tool to facilitate abuse and gain control of the victim.

When the predator has established the victim's trust, commenced grooming, and isolated the minor from support networks, the predator attempts to *approach* the victim by suggesting that they meet for sexual purposes. In the LCT, Olson, et al., define approach as "the initial physical contact or verbal lead-ins that occur prior to the actual sexual act." In the online model of luring communication, approach is the final step when the predator requests to meet the victim offline with the intent of beginning a sexual relationship.

4   Data Collection and Processing

When beginning this project, chat transcripts in text files were downloaded from the Perverted Justie website [PJ]. The files contain entire conversations between adults posing as young teens and convicted sexual offenders. At present, ChatCoder works with one "post" at a time, meaning one line of chat posted by one user, including the username, timestamp, and any other information the instant message program would include about the post. We processed each post by separating the username (to determine whether the user is a predator or victim) and the body of the the post (what the user actually typed and sent). The next two sections will discuss how ChatCoder determines danger in the conversation using the body of each post.

5   ChatCoder 1

The original ChatCoder program (ChatCoder 1) was developed to operationalize the LCT model. ChatCoder 1 utilized a dictionary containing 454 unique words and phrases. The phrase list was developed by a communication studies student working on the predation project. Each phrase was assigned to one of the nine luring subcategories. Some samples appear in Table 1.

When ChatCoder 1 found one of these phrases in a line, the line was coded with the appropriate luring subcategory assigned. Many posts which should have been coded were missed by ChatCoder 1. The phrase-matching system required that a phrase appear precisely the way it appeared in the dictionary. Extra words or typos would result in a false negative (a phrase was not coded when it should have been).

ChatCoder 1 often over-coded as well, because some of the dictionary entries were short, single words. For example, *age* was a dictionary entry assigned to the *personal information* category, but *age* also appears in many other words (ex. *image, average*) and the appearance of *age* as part of a word would cause a line to be coded as personal information, a false positive. Checking for spacing around a word was an insufficient response to this problem because running words together is one of the most frequent typographic errors in chat communications.

Table 1. Sample Excerpt from ChatCoder 1 Dictionary

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

6   ChatCoder 2

The goal for ChatCoder is to mimic a trained communications analyst. Thus we want the software to identify not only dangerous words or phrases, but also to identify a luring category for each post (if there is one). To compare the efficacy of our software, we have carefully analyzed the coding done by human coders who used a codebook that was developed during the initial operationalization of the luring theory. In this section we describe the rules given to the human coder, along with the rules employed by ChatCoder 2. Eighteen tokens were created to represent items in each LCT category - for example, approach nouns (*hotel*), action verbs (*think*), and complimentary adjectives (*pretty*). The dictionary was redeveloped by computer science students and the new dictionary contains 305 words, each assigned to a token type.

The computer science students closely re-analyzed the same fifteen chat transcripts from the Perverted-Justice.com website [11], and they identified common linguistic patterns for lines that fit into one of the nine subcategories in the LCT model. Rules were written when patterns present in multiple chat logs were identified. For example, the rule-based approach will flag all permutations of *you* and *pretty* anywhere in the line, such as *you pretty, you're pretty,* or *you are pretty.* ChatCoder 2 will flag all of the variations of the phrase, including contractions such as *you're pretty* and netspeak like *u r pretty.*

Each of the nine luring subcategories is also assigned a nominal value. The nominal value helps the system mimic the human coder when a post contains multiple luring categories. For example, if a line contains both *communicative desensitization* and *personal information* language, the line will be coded as *communicative desensitization* because it has a higher nominal value. Higher values are assigned to subcategories further along in the luring model, indicating

that the predator has moved to a more dangerous phase in his communication with the victim.

In the following subsections we describe the instructions provided to the human coders and also our algorithmic implementation for each subcategory in ChatCoder 2. The categories are described in order by their nominal values (i.e. *Gaining Access* has nominal value '1', and *Approach* has nominal value '9').

6.1   Gaining Access

Gaining access is defined as greetings or stated presence in online social networking sites. Human coders were told to code the line as *gaining access* for any formal or informal greeting such as *hi, hello,* or *what is up.* They were also told to code *gaining access* for any mention of a social networking site where the perpetrator and victim met / talked. A greeting term was coded only the first time it appeared, not each time the predator leaves and returns to the conversation. ChatCoder 2 coded *gaining access* for the first predator post in the transcript.

6.2   Personal Information

Personal information exchange is defined as victim and predator revealing, exchanging, requesting, or sending real information about themselves. Human coders were instructed to code the line as *personal information* when non-screen names, actual ages, hometown and computer locations (e.g. bedroom, office, etc.), cell or home phone numbers, email addresses, profiles and personal pictures are exchanged or discussed. They also coded requests for information about general, non-specific likes and dislikes: *what do you like* or *what do you like to do* when there is no clear context for the question or when the context is clearly non-sexual. They were told to only code personal data that could be reasonably used to physically locate and identify the victim. Making arrangements to meet was explicitly excluded from this category (it was coded as *approach*) as was exchanging information about general town or state locations. After one post is coded as *approach*, all subsequent *personal information* exchanges were also coded as *approach* by the analysts.

To identify personal information exchange, ChatCoder 2 looks first for an information noun (*pic, birthday*) combined with an action verb (*look, think*) or a question word (*when, ever*), which would include lines such as *look at my pic* or *when is your birthday.*

6.3   Relationship

Relationship exchange is defined as predators and victims talking about, sharing, or soliciting feelings and attitudes (positive, negative, or neutral) regarding romantic relationships, familial relationships or friendships. Analysts code the post as *relationship* if it includes language about building, maintaining, dismantling or ambivalence regarding their relationship with each other, with family,

friends, and romantic partners. Relationship was also coded if a post included references to the specific experiences and practices that were a part of their past or current associations with others.

To code a line as relationship information, ChatCoder 2 looked for a relationship noun (*boyfriend, divorce*) and an approach verb (*meet, call*) or action verb (*think, like*). Thus lines like *how did you meet your boyfriend* or *did you like your boyfriend* were coded.

6.4   Activities

Activities is defined as sharing information about activities and hobbies. Analysts coded the line as *activities* if there was discussion of favorite or shared social behaviors, like preferred musicians, movies and hobbies. They also coded requests for information about general, non-specific activity likes and dislikes: *what do you like to do for fun* or *where do you like to go to hang out* when there is a reasonably clear non-sexual context for the question.

ChatCoder 2 could only account for the most broad discussions by flagging all lines containing an activity noun (*book, movie, favorite*) which would catch lines such as *what is your favorite color* and *did you like that book.*

6.5   Compliments

Compliments is defined as offering praise. Analysts coded the line as a compliment if it contained language offering praise about appearance, activities, or personality.

When looking for a compliment, ChatCoder 2 first determines if there is a complimentary adjective (*pretty, mature, sexy*) in the line. If the complimentary adjective is present, the program then looks for a second person pronoun (*you, your*) or an information noun (*girl, pic*) and the line is flagged as a compliment. Thus phrases like *you are pretty* and *nice pic* are caught and the post is coded into the compliment category of the luring theory.

6.6   Communicative Desensitization

Communicative desensitization is defined as the use of vulgar sexual terms or the discussion of sexual acts or the willingness to experiment sexually. Analysts coded the line as communicative desensitization if it contained any vulgar language, any discussion of sexual acts, or any demonstration of sexual acts. This code also includes innuendos or vague references to sexual arousal, sex as *doing things* or *doing anything,* and the exchange of general photos of naked others, but not predator or victim. Words like *dammit, shit,* and *hell* were excluded.

At present, ChatCoder 2 successfully finds communicative desensitization in a number of situations. If it finds a communicative desensitization noun (*body, lover, sex*) and an action verb (*think, have*), then lines like *you have a hot body* and *think you'd like sex* are flagged as communicative desensitization. Communicative desensitization is also flagged when a communicative desensitization

noun and a communicative desensitization verb (*want, kiss*) or a linking verb (*am, be*) are found together, meaning lines like *kiss your body* and *you will be my lover* are flagged. A communicative desensitization verb with a second person pronoun or question word will also be flagged to include phrases like *do you shave* and *ever kissed*. If there is a communicative desensitization adjective (*sexy, hard, naked*) and a first or second person pronoun or action verb, ChatCoder 2 will flag the line, finding *get naked* and *you are sexy.*

6.7   Reframing

Reframing is defined as redefining sexual behaviors in non-sexual terms. Analysts were instructed to code the line as *reframing* if the language attempts to describe or connect sexual acts to *messing around, playing, learning, helping, assisting,* or *practicing.*

Reframing is coded by ChatCoder 2 when a reframing verb (*play, learn, teach*) is found with a first or second person pronoun. Thus *you want to play with it* and *i could teach you how* are flagged in the reframing category.

6.8   Isolation

Isolation is defined as separating the victim from his or her family or friends. A post was coded as isolation if it includes a question about the physical location of the victim's friends or family or the victim's location. Also includes discussion about lying to, or concealing from, parents and friends.

To flag isolation, ChatCoder 2 had to find a family noun (*mom, dad*) and an approach verb or question words, such as *did your dad leave* or *what does your dad do.* Isolation was also flagged when an isolation adjective (*alone, lonely*) was found with a second person pronoun, like *you must be lonely.*

6.9   Approach

Approach is defined as an attempt to meet the victim in person. Analysts coded the line as approach if it contained language asking the victim to speak on the telephone or to meet in person. Approach also includes specific requests for a victim's location, address or appearance after an approach has been coded. It includes statements or requests for meeting time and references to *coming over,* as well as discussions of bringing items to the victim.

In ChatCoder 2 an *approach* line had to contain an approach verb (*meet, see, come*), no information nouns, and a first or second person pronoun or approach noun (*hotel, highway, car*). This included phrases such as *I want to meet you* and *could we go to a hotel.*

7   Evaluation Methodology and Results

Intercoder reliability compares the number of shared coding decisions between two coders doing content analysis. The online predation project has data from

four different coders: two student hand coders, ChatCoder 1 and ChatCoder 2. In this paper we compute intercoder reliability using Holsti's method [3]. The formula is shown in Equation 1; N is the number of matches, n_1 is the number of lines coded by coder 1, and n_2 is the number of lines coded by coder 2. In this section coder 1 and coder 2 can be either human or computer coders.

$$Reliability = 2N/(n_1 + n_2), \tag{1}$$

Intercoder coder reliability for two trained undergraduate student coders ranged from 61% to 94%. The students improved their intercoder reliability over time, after results from previous attempts were compared and the manual coding process was discussed (with a faculty mentor). Human-to-human intercoder reliability was initially in the 60-70% range, but eventually improved to 94%.

After human-to-human coding reliability stabilized, we calculated intercoder reliability between ChatCoder and the human coders (with the data we already had available from the humans). Computer-to-human intercoder reliability ranged from 30% to 58% (see Table 2).

ChatCoder 1 vs. human intercoder reliability ranged from 24.29% to 56.56%. ChatCoder 2 ranges from 31.94% to 58.74% for the same transcripts. Overall, ChatCoder's ability to correctly identify predation when compared to a human coder improved by an average of 5.81% with a maximum improvement of 13.21%. Intercoder reliability increased in 16 of the 17 test cases.

Table 2. Intercoder Reliability Statistics

Predator User Name	ChatCoder 2 vs. Human 1	ChatCoder 1 vs. Human 1	ChatCoder 2 vs. Human 2	ChatCoder 1 vs. Human 2
AbraxisReborn	31.94	24.29	34.59	24.81
Bendix632	–	–	48.16	43.05
DaddyWants2PlayNokc	–	–	40.80	38.81
Dick_HungWell64	33.33	32.35	–	–
Fuddster88	–	–	33.86	24.53
funjxn1	58.74	56.56	–	–
Good_Boy_B88	45.38	34.90	–	–
Hardblackdick	–	–	16.67	15.38
jc_bah	32.35	31.50	–	–
Johnchess2000	35.48	30.71	–	–
Michaelmeyer2448	–	–	41.54	40.16
ndguy58746	46.13	47.67	–	–
Wesleybrannen82	50.00	40.23	55.49	44.03
x_cums_in_the_night_x	–	–	35.11	32.61
your_lil_nene	–	–	39.70	26.49

Intercoder reliability can also be measured at the category level. For personal information, we improved by an average of 13.35%. Because of the large range of

questions and answers that personal information contains, the phrase matching strategy that ChatCoder 1 employed was largely unsuccessful because it could not possibly match all permutations of the ways in which a predator could ask about or offer personal information. With our more flexible strategy, we were able to match the human coders more successfully.

Posts in the relationship category were very difficult for ChatCoder 1 to detect because there are numerous way to ask or talk about a relationship. It also is difficult to separate discussion of a relationship that should be categorized as communicative desensitization from a discussion involving a relationship that is purely relationship information. For example, the line *I am divorced and have no girlfriend* does not match any phrase in the ChatCoder 1 dictionary, but it is clearly relationship information. ChatCoder 2 codes the aforementioned line as relationship information, based on the fact that the line contains the relationship nouns *divorced* and *girlfriend*, and also contains a first person pronoun *I*, and an action verb, *have*. We improved relationship coding by an average of 21.08%. There were some difficulties in coding for relationships, as in six out of seventeen cases, no relationship lines were flagged by either the old or new ChatCoder, a statistic that results largely from disagreements between the human coders about what constitutes discussion of a relationship. For instance, the line *may I ask how far you've gone* is coded by one hand coder as relationship discussion and by a second as communicative desensitization.

As a category, activities represents an extraordinarily broad range of topics and inquiries that are difficult to narrow down to a few rules. Phrase matching produced practically no results in the activities category, but by utilizing a few common types of activities, such as movies and music, we could code for the more prevalent activity-related discussions. However, when the specificity of the activities-related discussion became too high, such as *have you ever gone kayaking,* we were unable to account for it. Thus our average improvement in coding for activities was 14.24%.

Compliments improved by an average of 25.29%, with a maximum improvement of 42.06%. Compliments were easy to improve by adding any complimentary adjectives found to the appropriate category. This resulted in giving ChatCoder 2 the ability to find a number of different phrases that ultimately have the same meaning, such as *you cute, you are so cute,* and *I think you are cute.*

Communicative desensitization was a difficult category to code because so much depends on context, which is difficult to teach a computer. By implementing rules based on parts of speech, our average intercoder reliability decreased by 2.23%. We believe that this is largely based on the fact that words such as *cum* are coded as communicative desensitization by the hand coders regardless of context. In addition, since categories are coded on a hierarchical basis, if the phrase contains an approach verb, such as *come,* then it is coded incorrectly as approach instead of communicative desensitization. To remedy this problem, the next version of ChatCoder will include more specific hierarchical rules to avoid this miscoding.

Our working definition of the reframing category is open to interpretation, primarily due to a lot of inconsistency in hand coding. It was difficult to identify specific phrases to be flagged as reframing. ChatCoder 1 consequently coded reframing very poorly, only receiving positive intercoder reliabilities in three of our seventeen chat transcripts. We were able to catch most occurrences of reframing, with an average increase of 23.18%, peaking at a maximum improvement of 85.71%.

Another difficult category to pinpoint was isolation. Our average intercoder reliability decreased by 4.13%. Future improvements would include coding for an isolation adjective and an information noun to catch phrases such as *home alone*. Previous attempts to catch these exceptions have vastly over-coded, so more careful analysis of instances of isolation must be undertaken.

Finally, the approach category had an average improvement of 1.05%, improving in seven transcripts and declining in ten. A case by case analysis of the intercoder reliability for approach showed that although our rules caught many instances of approach, it vastly over-codes because any phrase like *I go to the pool* is marked as approach, since it contains a first person pronoun and an approach verb.

8   Conclusions

We made significant progress by improving upon the methodology of ChatCoder 1. We increased the average overall intercoder reliability by 5.81% with a maximum improvement of 13.21%. Furthermore, we improved significantly in the categories of personal information, relationship information, activities, and compliments, peaking with values solidly in the range of reliability that a human coder achieves.

In the categories in which we did not see improvement across the board - communicative desensitization, reframing, isolation, and approach - we learned a great deal about the integrity of the categories themselves. At this point, we can confidently say that much of the information that ChatCoder 2 misses is due to a lack of clarification within the categories. Refinement of the luring model or updates to the coding instructions given to the human coders may be warranted.

There are a number of improvements that we believe would further increase the efficiency of ChatCoder. Typos and netspeak continue to be major issues and statistical methods may be needed to account for these discrepancies. We are currently working to integrate a translate function that will utilize an online dictionary of slang words. Future generations of ChatCoder can also be expanded to consider a window of posts rather than just single lines in order to better catch insinuations.

Overall, the intercoder reliability of ChatCoder has markedly increased, and we are well on our way to automating the content analysis process for this difficult and important application.

References

1. P. H. Adams and C. H. Martell. Topic detection and extraction in chat. In *Proceedings of the 2008 IEEE International Conference on Semantic Computing*, pages 281–588, 2008.
2. eBlaster™, December 2008. `http://www.eblaster.com/`.
3. O. R. Holsti. *Content Analysis for the Social Sciences and Humanities*. Addison-Wesley, Reading, MA, 1969.
4. D. Hughes, P. Rayson, J. Walkerdine, K. Lee, P. Greenwood, A. Rashid, C. May-Chahal, and M. Brennan. Supporting law enforcement in digital communities through natural language analysis. In *Proceedings of the Second International Workshop on Computational Forensics (IWCF'08)*, 2008.
5. A. Kontostathis, L. Edwards, and A. Leatherman. ChatCoder: Toward the tracking and categorization of Internet predators. In *Proceedings of the Text Mining Workshop*, 2009.
6. A. Leatherman. Luring language and virtual victims: Coding cyber-predators online communicative behavior. Technical report, Ursinus College, Collegeville, PA, USA, 2009.
7. NCMEC. National center for missing and exploited children, October 2008. `http://www.missingkids.com/en_US/documents/CyberTiplineFactSheet.pdf`.
8. Net Nanny™, Dec 2008. `http://www.netnanny.com/`.
9. L. Olson, J. Daggs, B. Ellevold, and T. Rogers. Entrapping the innocent: Toward a theory of child sexual predators' luring communication. *Communication Theory*, 17(3):231–251, 2007.
10. N. Pendar. Toward spotting the pedophile: Telling victim from predator in text chats. In *Proceedings of the First IEEE International Conference on Semantic Computing*, pages 235–241, 2007.
11. Perverted-Justice.com. Perverted justice, August 2008. `www.perverted-justice.com`.
12. D. Riffe, S. Lacy, and F. Fico. *Analyzing Media Messages: Using Quantitative Content Analysis in Research*. Lawrence Erlbaum Associates, 1998.