A Two-step Approach for Effective Detection of Misbehaving Users in Chats*

Notebook for PAN at CLEF 2012

Esaú Villatoro-Tello[2], Antonio Juárez-González[1], Hugo Jair Escalante[1], Manuel Montes-y-Gómez[1], and Luis Villaseñor-Pineda[1]

[1] Laboratorio de Tecnologías del Lenguaje, Coordinación de Ciencias Computacionales
Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), Mexico.
{antjug, hugojair, mmontesg, villasen}@ccc.inaoep.mx
[2] Information Technologies Department, Universidad Autónoma Metropolitana (UAM),
Unidad Cuajimapla, Mexico City, Mexico.
evillatoro@correo.cua.uam.mx

Abstract This paper describes the system jointly developed by the Language Technologies Lab from INAOE and the Language and Reasoning Group from UAM for the Sexual Predators Identification task at the PAN 2012. The presented system focuses on the problem of identifying sexual predators in a set of suspicious chatting. It is mainly based on the following hypotheses: *(i)* terms used in the process of child exploitation are categorically and psychologically different than terms used in general chatting; and *(ii)* predators usually apply the same course of conduct pattern when they are approaching a child. Based on these hypotheses, our participation at the PAN 2012 aimed to demonstrate that it is possible to train a classifier to learn those particular terms that turn a chat conversation into a case of *online child exploitation*; and, that it is also possible to learn the behavioral patters of predators during a chat conversation allowing us to accurately distinguish victims from predators.

1 Introduction

It is well known that the World Wide Web (WWW) has vastly penetrated into social living and allows connecting people from different geographic regions through new forms of communication. Examples of such communication forms are instant messaging services, chat rooms, social networks (*e.g.*, Facebook, Twitter, etc.) and blogs. These services have become very popular tools for personal as well as for group communication, as they are cheap, easy to use, virtual and private in nature [6].

Such online services allow users to hide their personal information behind the monitor; which, on the one hand, makes this type of communication a source of fun, but on the other hand, it also represents a threat. The privacy and virtual nature of these services augment the possibilities of some heinous acts which one may not commit in the real world. Examples of such acts are the online paedophiles who "groom" children,

* This work was done under partial support of CONACYT (project grants 134186 and 106013). We also thank SNI-Mexico, INAOE and UAM for their assistance.

that is, who meet underage victims online, engage in sexually explicit text or video chat with them, and eventually convince the children to meet them in person. According to the National Center for Missing & Exploited Children [9] and the Office of Juvenile Justice and Delinquency Prevention[3], one out of every seven children receives an unwanted sexual solicitation online.

Traditionally, a term that is used to describe such malicious actions with a potential aim of sexual exploitation or emotional connection with a child is referred as "Child Grooming" or "Grooming Attack" [3], which has been defined by [1] as: a communication process by which a perpetrator applies affinity seeking strategies, while simultaneously acquiring information about and sexually desensitizing targeted victims in order to develop relationships that result in need fulfilment (*e.g.* "physical sexual molestation.").

Nowadays, the usual way to catch these sexual predators is for trained law enforcement officers or volunteers to pose as children in online chat rooms, thus predators fall into the trap and are identified. However, online sexual predators always outnumber the law enforcement officers and volunteers. An organization that employs this methodology to catch sexual predators is the *Perverted Justice* group, located in the United States. This organization has been able to convict more than 500 predators since 2004[4]. Nevertheless, there is a great need for software applications that can flag suspicious online chats automatically, either as a tool to aid law enforcement officials or as parental control features offered by chat service providers.

In this paper we propose a novel methodology that faces the problem of sexual predators identification as a text classification task by means of a supervised approach. The identification process is divided in two main stages: the *Suspicious Conversations Identification* (SCI) stage and the *Victim From Predator disclosure* (VFP) stage. Performed experiments showed that it is possible to train a classifier to learn those particular terms that turn a chat conversation into a case of *online child exploitation*; and, that it is also possible to learn the behavioral patters of predators during a chat conversation allowing us to accurately distinguishing victims from predators.

The rest of the paper is organized as follows. Section 2 presents recent work on the task of sexual predators identification. Section 3 describes the proposed methodology for identifying both suspicious chat conversations and the sexual predator within a chat conversation. Section 4 presents the experimental settings as well as the results achieved in the context of the PAN 2012 competition. Finally, Section 6 depicts our conclusions and formulates directions for future work.

2 Related Work

Traditionally, the problem of identifying grooming attacks has been tackled through text classification strategies. In [5] the problem is reduced to the task of distinguishing predators from victims within chat conversations. Such conversations are known to be cases of grooming attacks, particularly they used a set of 701 conversations obtained from the perverted-justice web page. In order to solve the problem, Pendar et al.

[3] *http://www.ojjdp.gov/*

[4] *http://www.perverted-justice.com/*

separated those interventions that belong to the set of victims from those that belong to the set of predators, *i.e.*, a two-class problem. Next, authors remove the stopwords, and then computed word unigrams, bigrams and trigrams. Subsequently, they processed them with the classification algorithms Support Vector Machine (SVM) and k-nearest neighbors (k-NN). Authors performed several experiments varying the number of features from 5000 to 10000, and concluded that the k-NN algorithm with k equal to 30 and employing a feature vector of 10000 elements provides the most effective classification resulting in a value of f-measure that equals to 0.943.

Similarly, in [3], Michalopoulos et al. proposed a decision making method to be used for recognizing potential grooming threats by extracting information from captured dynamic dialogues. Their proposed method makes use of the following three classification classes: *i) Gaining Access*: indicate predators intention to gain access to the victim; *ii) Deceptive Relationship*: indicate the deceptive relationship that the predator tries to establish with the minor, and are preliminary to a sexual exploitation attack; and *iii) Sexual Affair:* clearly indicate predator's intention for a sexual affair with the victim. The classification process computes the probability that a captured dialogue belongs to each one of the above classes. At the end, their system decides if there is a threat in the conversation by means of a linear combination of the computed probabilities. For their experiments, authors employed a set of 219 chat conversations (73 for each class), they removed all the stopwords and applied a spelling correction strategy. Michalopoulos et al. concluded that Naïve Bayes is the most appropriate technique, not only for reaching the highest average classification score of 96%, but also for being fastest than all the other algorithms that were evaluated.

Finally, the work proposed by RahmanMiah et al. faces the problem of grooming attack identification from a more general point of view [6]. Contrary to the work proposed in [5,3], the authors of [6] define a method for identifying when a conversation is a case of child exploitation instead of detecting which user is the predator directly. The system proposed by RahmanMiah et al. defines three classes of conversations: *i)* Child Exploitation: cases of grooming attacks; *ii)* Sexual Fantasies: conversations between adults with a high degree of sexual content; and *iii)* general chatting: general conversations with no sexual content. The proposed system applies traditional text categorization techniques in combination with psychometric and categorical information provided by LIWC (Linguistic Inquiry and Word Count [4]). For their experiments, authors employed a set of 392 conversations, they do not apply any pre-processing to the texts neither a spelling correction process. Authors conclude that psychometric and categorical information can be used by classifiers as a feature set to predict the suspected child exploitation in chats. These psychometric features significantly improve the performance of Naïve Bayes classifiers to predict child-exploitation type chats.

Our proposal differs from previous works in that it attacks both problems at once, *i.e.*, we are able to identify when a chat conversation is a case of child exploitation and subsequently we are able to tell which user is the sexual predator. Thus our proposal can be used for both purposes (detecting suspect conversations and identifying predators); besides, we show that the two-step approach outperforms a single-stage method.

3 Proposed Method

The proposed method for detection of misbehaving users in chats is based on two main hypotheses: *(i)* terms used in the process of child exploitation are categorically and psychologically different than terms used in general chatting; and *(ii)* predators usually apply the same course of conduct pattern when they are approaching a child. Accordingly, we propose a new methodology for solving the problem of sexual predators identification, which is divided in two main stages: the *Suspicious Conversations Identification* (SCI) stage and the *Victim From Predator disclosure* (VFP) stage. Figure 1 shows the general architecture of the proposed system.

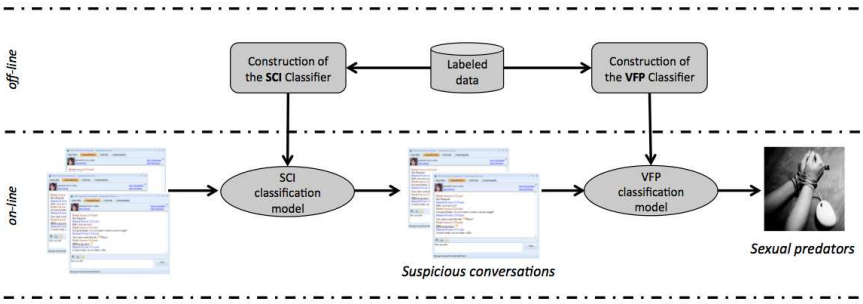

Figure 1. General overview of the proposed sexual predators identification system.

Notice that the goal of the first stage is to act as a filter, *i.e.*, it helps distinguishing between general chatting and possible cases of *online child exploitation*; in this way, the set of conversations to be analyzed by the VFP module will be reduced. Hence we can focus only on conversations that potentially include sexual predators for a more fine grained analysis. Consequently, the goal of the second stage is to identify (to point at) the potential predator from a possible case of child exploitation. The associated classification problem is less complex than trying to discriminate between predators and typical users directly, see Section 4.4.

3.1 Pre-processing stage

Our proposed system does not include any module for preprocessing the texts, *i.e.*, we did not remove any punctuation mark, stopwords and, neither apply a stemming process. The main reason for not applying any preprocessing was because the text in chat conversations had unique characteristics that distinguish them from any other type of text [2,6,7], for example, chat conversations do not follow any grammar rules (*i.e.*, are grammatically informal and unstructured), plus the frequent orthographical errors and the common use of abbreviations and emoticons.

We believe that using emoticons and intentional misspelled words may contain valuable contextual information in a chat text. For example, in the grooming phase the perpetrator may amend the relationship by an emphasized “soryyyyyyyyy” when the child

felt threatening by any obtrusive language. Another example may be the emoticon for "hug (>:d<)" and "kiss (:-*)" for a soft introduction of sexual stage. Preserving such information makes traditional language processing tools, such as stemmers and POS taggers, unsuitable for processing the chat texts [2,6].

3.2 Filtering stage

Although we did not apply any pre-processing stage, it is important to mention that we did apply a pre-filtering stage to all the conversations that were given for the PAN 2012 sexual predator competition. The goals of the pre-filtering stage were: *i)* to help us focusing only in the most important cases and, *ii)* to reduce the computational cost for automatically processing all the information.

This pre-filtering stage consisted in removing all the conversations that accomplish at least one of the following conditions: *1*) Conversations that had only one participant, *2)* Conversations that had less than 6 interventions per-user and *3)* Conversations that had long sequences of unrecognised characters (apparently images). Table 1 shows information about the training data before and after applying this pre-filtering stage.

Number of...	***Original data***	***Filtered data***
Chat conversations	66,928	6,588
Users	97,690	11,038
Sexual Predators	148	136

Table 1. Number of chat conversation, users and sexual predators that remain after applying the pre-filtering stage on the training data.

As it can be seen, by means of the pre-filtering stage we are able to reach a substantial reduction ratio (90% approximately) of conversations/users. It is important to notice that by doing this pre-filtering, we also removed a few sexual predators. Thus, even if our proposed system works perfectly we will not be able to identify the 100% of the sexual predators. Nevertheless, we think the information from interventions of removed predators was not enough to effectively recognize them as predators anyways.

3.3 Suspicious Conversations Identification

As we have mentioned before, our system faces the problem of sexual predators identification as a text classification task (TC). Accordingly, for training the SCI classifier (Figure 1) we employed traditional TC techniques to construct a model that distinguish between general chatting and cases of child exploitation.

In order to properly train our SCI classifier, we labeled as *suspicious* conversations all the chat conversations were at least one predator appears, resulting in 5790 non-suspicious conversations and 798 suspicious ones. During the experimentation phase, the SCI classifier represents the conversations by means of the bag of words representation (BOW) employing either a boolean or a TF-IDF weighting scheme. Since we

did not apply any preprocessing stage to the chat texts, we obtained features vectors of 117015 elements.

3.4 Victim From Predator disclosure

Similarly to the SCI classifier, our VFP stage was designed using traditional TC techniques. The goal of the VFP classifier was to recognize sexual predators in suspicious chat conversations, as detected by the SCI method. For training the VFP classifier we divided all the text conversations, where a predator was involved, into *interventions*. This means that if a text chat involved two different users, we had two sets of interventions. Therefore, we used as examples of victims the interventions of the users that had a conversation with a predator, resulting in 194 examples of victims; and as examples of predators we used the interventions of the 136 users already labeled as predators. For the experiments performed, the VFP classifier employs a BOW representation using either a boolean or a TF-IDF weighting scheme. For the VFP classifier we obtained features vectors of 16709 elements.

4 Experimental Setup

4.1 Data set

For our experiments we used the data set provided in the context of the PAN 2012 Lab: Sexual Identification competition. As we mentioned in Section 3.2 we were given for training a total of 66928 different chat conversations, where 97690 different users are involved and only 148 are tagged as sexual predators (See Table 1). Additionally, a test data set for evaluation was provided. Such corpus contained 155129 chat texts, where 218702 different users are involved and only 250 are tagged as sexual predators. For a more detailed explanation on how the *training* and *test* corpora were constructed please visit *http://pan.webis.de/*.

4.2 Classification methods

Two classifiers from the CLOP toolbox [8] are used in the text classification experiments; these are Neural Networks (NN) and Support Vector Machines (SVM) classifiers. The NN classifier was set as a two layer neural network with a single hidden layer of 10 units. For the SVM we tried linear and polynomial kernels.

During the development phase we adopted two-fold cross validation to estimate the performance of our methods using training data only. This validation technique was used for all of our experiments. For the final evaluation of our system we used the test data provided by organizers of PAN 2012, see Section 4.1. An analysis of the results using both training and test data sets is given in the following section. The evaluation of training-set results was carried out mainly by means of the classification accuracy, which indicates the overall percentage of text chats correctly classified. Additionally, due to the class imbalance, we also report results in terms of F1 measure. Regarding the final evaluation of the system on test data, we used the measures proposed by organizers, namely: F-0.5 measure, precision and recall.

4.3 Baseline definition

As baseline we used the traditional paradigm for solving the problem of sexual predators identification. Figure 2 provides a general view of the baseline definition.

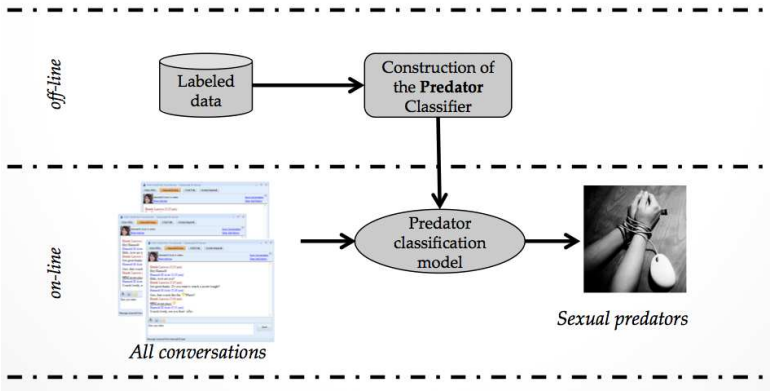

Figure 2. General overview of the baseline system for identifying Sexual Predators.

As can be noticed, the problem of identifying sexual predators is performed in one single step. For the baseline experiment we employed a BOW representation using either a boolean or a TF-IDF weighting scheme. By following the same procedure established for the SCI and the VFP classifiers, under this configuration we obtained features vectors of 117015 elements.

4.4 Experimental results

In this section we report experimental results obtained by the components of our two-step approach, as well as the results obtained with the baseline method, using training data. Next, in Section 4.5, we report the performance of the system in the test data set.

Baseline results. Table 2 shows results obtained by the baseline configuration. In order to evaluate if a dimensionality reduction strategy could be helpful for the classifier, we performed several experiments varying the size of the features vectors that were used to train the classifier. For this purpose we employed the well known information gain (IG) method to rank and preserve the most distinctive features.

Recall that the size of the features vector is 117015 elements, hence using only 10% of the features means that the classifiers employed only 11,702 features to represent the 11,038 chat conversations. The baseline configuration faces the problem of learning a model that helps to classify between normal users and sexual predators from a highly unbalanced corpora, *i.e.*, 10,902 normal users and only 136 sexual predators.

Table 2 indicates that using a binary weighting scheme allows a better performance for the classifier. It is also possible to observe that reducing the dimensionality of the

algorithm	Weighting	Num. of features	Accuracy	F-measure
SVM	*binary*	100%	**0.9935**	**0.6869**
SVM	*binary*	70%	0.9935	0.6869
SVM	*binary*	40%	0.9922	0.6587
SVM	*binary*	10%	0.9716	0.2737
SVM	*tf-idf*	100%	**0.9926**	**0.6611**
SVM	*tf-idf*	70%	0.9926	0.6611
SVM	*tf-idf*	40%	0.9907	0.5253
SVM	*tf-idf*	10%	0.9846	0.1747
NN	*binary*	100%	0.9936	**0.6697**
NN	*binary*	70%	0.9928	0.6153
NN	*binary*	40%	**0.9939**	0.5955
NN	*binary*	10%	0.9933	0.6696

Table 2. Results obtained with the baseline configuration. Several experiment applying a dimensionality reduction strategy were performed.

feature's vector does not allow a significant improvement. On the contrary, it decreases the ability of the classifier to accurately distinguish between normal users and sexual predators. Because of these results we decided to not use any dimensionality reduction strategy for both SCI and VFP classifiers.

SCI results. Table 3 shows results obtained from the SCI classifier. As mentioned in previous sections, the aim of this classifier is to distinguish between general chatting and possible cases of online child exploitation. It is worth mentioning that the training data employed by the SCI classifier represented an unbalanced corpus, since there are 5790 conversation labeled as general chatting and only 798 text chats labeled as cases of online child exploitation, however, obtained results showed that it is possible to accurately distinguish suspicious conversations.

Algorithm	Weighting	Accuracy	F-measure
SVM	*binary*	0.9848	0.9361
SVM	*tf-idf*	**0.9883**	**0.9516**
NN	*binary*	**0.9874**	**0.9464**
NN	*tf-idf*	0.9825	0.9254

Table 3. Results obtained by the SCI module.

Experimental results showed that both classifiers (*i.e.*, SVM and NN) are suitable for solving the problem of classifying suspicious conversations. Contrary to the results obtained with the baseline configuration, the SCI classifier using SVM as classification method obtained better results when chat conversations are represented by means of a BOW considering a tf-idf weighting scheme.

We concluded from these experiments that, terms used in the process of child exploitation are categorically and psychologically different than terms used in general chatting, which allows to train a classifier to learn those particular terms and accurately detect cases of *online child exploitation*.

VFP results. Table 4 shows results obtained from the VFP classifier. As we mentioned in Section 3.4, the aim of this particular module is, once a conversation has been tagged as a suspicious, to point at the sexual predator, *i.e.*, to tell which user is the victim and which one is the predator.

Algorithm	Weighting	Accuracy	F-measure
SVM	*binary*	0.9148	0.9138
SVM	*tf-idf*	**0.9259**	**0.9305**
NN	*binary*	**0.9407**	**0.9424**
NN	*tf-idf*	0.9296	0.9337

Table 4. Results obtained for the VFP module.

Obtained results showed that the proposed methods are adequate for solving the problem of classifying victims and predators. Similarly to the results obtained in the SCI classifier, using SVM as classification method obtained better results when a tf-idf weighting scheme was employed. Nevertheless, for the case of the VFP classifier, the best results were obtained when using NN algorithm and a binary weighing scheme.

From the experiments performed in this section we were able to show evidence that suggest predators apply the same course of conduct pattern when they are approaching a child, and that our proposed method it is able learn these behavioral patters of predators during a chat conversation allowing us to accurately distinguish victims from predators.

4.5 PAN 2012 competition results

Previous sections described obtained results employing the training data set, and all experiments were performed in a controlled scenario. However, the main goal of the PAN 2012 Lab was to evaluate in real scenarios the proposed method for sexual predators identification. This section reports official results obtained by our system on test data from the PAN 2012 competition.

In order to apply our proposed methodology, the first step consisted on applying the filtering stage (Section 3.2) to test data. Table 5 shows some statistics of the test data before and after applying the filtering stage. From this table we can observe similar reduction ratios as in training data. A total of 28 predators were removed by our filtering approach. We manually analyzed the interventions of removed predators and found that most of them contained a few characters only; we consider that with such little information it is not possible to effectively identify to removed predators.

After filtering the test data, we were able to apply our proposed method (Figure 1). The next step was to represent all the remaining conversations (*i.e.*, 15,330) into the

Number of...	Original data	Filtered data
Chat conversations	155,129	15,330
Users	218,702	25,120
Sexual Predators	254	222

Table 5. Number of chat conversation, users and sexual predators that remain after applying the pre-filtering stage on the test data.

generated model for the SCI classifier. Following, the chat conversations that the SCI classified as *suspicious* were divided into interventions and represented accordingly to the model generated for the VFP classifier.

Our team submitted three different runs: *i)* Baseline: it corresponds to the configuration showed in Figure 2 employing as classification method a Neural Network and using a binary weighting scheme with no dimensionality reduction; *ii)* SCI(NN-B)& VFP(NN-TF-IDF): it means that the SCI module was configured for using a NN and a binary weighting scheme, whereas the VFP module used a NN with a tf-idf weighting scheme; and *iii)*SCI(NN-B)& VFP(NN-B): it means that both the SCI and the VFP modules were configured for using a NN and a binary weighting scheme.

Official results of submitted runs are showed in Table 6. The leading evaluation measure was F-0.5 measure, which emphasizes the importance of the precision of the system, which is particularly important for sexual predator detection as mentioned by the organizers of the PAN 2012 competition.

Run	Recall	Precision	F-measure	F-measure ($\beta = 0.5$)
Baseline	0.4055	0.9537	0.5691	0.7507
SCI(NN-B)& VFP(NN-TF-IDF)	0.7874	0.9479	0.8602	0.9107
SCI(NN-B)& VFP(NN-B)	**0.7874**	**0.9804**	**0.8734**	**0.9346**

Table 6. Official evaluation results from the submitted runs.

As it is possible to observe, the best configuration was the third one, *i.e.*, using in both modules a binary weighting scheme. Thus showing that the hypotheses of our work were right. Indeed the configuration $SCI(NN-B)\&VFP(NN-B)$ obtained the highest performance among the 16 teams that participated in the sexual predator identification track of PAN 2012. The closest entry (*snider12-run-2012-06-16-0032*) achieved an F-0.5 measure of 0.9168, whereas the average was of 0.5105. The results obtained by our group are promising and motivate us to pursuing several future work directions.

5 Identifying predators' bad behavior

An additional task to that of sexual predator detection, proposed by organizers of PAN 2012, was to search for those lines (interventions) that reveal predator's bad behav-

ior. Traditionally, such lines are manually identified and used as evidence to convict paedophiles. PAN 2012 participants were encouraged to propose new ideas to automatically solve this issue.

We approached the line-detection task with a language-models based approach. Our main idea was based in the following statement; it is well known that every predator follows three main stages when approaching a child [3]: *i)* gain access to the victim, *ii)* involve the victim in a deceptive relationship and, *iii)* launch and prolong a sexually abuse relationship. Based on these facts, we believe that if we can generate language models from each one of the stages mentioned above, we will be able to find those lines that represent a bad behavior. We were particularly interested in the 2nd and the 3rd stages, since from our point of view these should be the most critical sections within a child exploitation chat conversation.

Our proposed solution works as follows: first we automatically divide all the conversations where a predator appears in three sections, such division is made without considering any type of contextual frontiers, *i.e.*, we did not identify where a child approaching stage begins or ends. Next we generated the language model (*lm*) of the 2nd and the 3rd parts[5]. Finally, for a user that is tagged as predator, we compute the perplexity against the *lm* of each one of its interventions, and we delivered as the most distinctive lines of bad behaving those with the minor perplexity value.

For the competition, we proceed as follows: from the set of users labeled as sexual predators by our system (SCI(NN-B)& VFP(NN-B)), we select the 50 most distinctive lines and delivered as examples of predators bad behaving. Table 7 shows examples of the lines that were delivered. Official evaluation results indicated that by following this procedure we were able to identify just 1 revealing line. We believe that our proposed idea could be very effective, although more work is necessary in order to improve its performance. Although this result seems absolutely negative we have to mention that the winner participant of this sub-task submitted $63,290$ lines (*grozea12-run-2012-06-14-1706b*).

From the 2nd parts	**From the 3rd parts**
<text>what do you want me to be?</text>	*<text>what do u want me to say</text>*
<text>what do u want me to say</text>	*<text>what do u want me to wear</text>*
<text>what do u want me to wear</text>	*<text>what do you want me to be?</text>*
<text>do u want to talk to me too</text>	*<text>do u want me to do it</text>*

Table 7. Examples of lines with the minor perplexity values.

6 Conclusions

We have proposed a new methodology for detecting sexual predators in text chats. Our proposal differs from traditional approaches in that it divides the problem in two stages:

[5] We used the SML toolkit http://svr-www.eng.cam.ac.uk/ prc14/toolkit.html to this end.

the *Suspicious Conversations Identification* (SCI) stage and the *Victim From Predator disclosure* (VFP) stage. The goal of the first stage is to work as a filter, *i.e.*, it helps distinguishing between general chatting and possible cases of *online child exploitation*; in this way, the set of conversations to be analyzed will be reduced, hence we can focus only on conversations that potentially include sexual predators for a more fine grained analysis. Consequently, the goal of the second stage is to identify (to point at) the potential predator from a possible case of child exploitation.

Performed experiments showed that it is possible to train a classifier to learn those particular terms that turn a chat conversation into a case of *online child exploitation*; and, that it is also possible to learn the behavioral patters of predators during a chat conversation allowing us to accurately distinguish victims from predators. Our participation in the PAN 2012 forum showed that the proposed methodology is able to produce very good results in a realistic scenario, obtaining an F-measure ($\beta = 0.5$) of 0.8936, which was the best ranked result among all of the participants.

As future work we plan to include some linguistic features, such as proposed by [6]. We believe that the inclusion of such type of features can be helpful for increasing the recall levels of our proposed system. In addition, we also believe that in the process of identifying the interventions that depict the predator's bad behavior this type of information could be very helpful.

References

1. Harms C. Grooming: An operational definition and coding scheme. In *Sex Offender Law Report*, Vol. 8, Num. 1. pp. 1-6, 2007.
2. Kucukyilmaz T., Cambazoglu B. B., Aykanat C. and Can F. Chat mining: predicting user and message attributes in computer-mediated communication. In *Information Processing and Management* Vol. 44(4), pp. 1448-1466. 2008.
3. D. Michalopoulos and I. Mavridis. Utilizing document classification for grooming attack recognition. In *IEEE Symposium on Computers and Communications (ISCC 2011)*, pp. 864-869, 2011.
4. O'Connell R. A Typology of Child Cyber- sexploitation and Online Grooming Practices. In *Cyberspace Research Unit, University of Central Lancashire*. 2003. Retrieved from http://image.guardian.co.uk/sys-files/Society/documents/2003/07/17/Groomingreport.pdf (accessed August 2012).
5. Pendar N. Toward Spotting the Pedophile Telling victim from predator in text chats. In *IEEE International Conference on Semantic Computing.* Irvine California USA, pp. 235-241, 2007.
6. RahmanMiah M. W., Yearwood J., and Kulkarni S. Detection of child exploiting chats from a mixed chat dataset as text classification task. In *Proceedings of the Australian Language Technology Association Workshop*, pp. 157-165, 2011.
7. Rosa K. D., and Ellen J. Text classification methodologies applied to micro-text in military chat. In *Proceedings of the eight IEEE International Conference on Machine Learning and Applications (ICMLA '09)*, pp. 710-714, 2009.
8. A. Saffari and I Guyon. Quick Start Guide for CLOP. Technical report, Graz-UT and CLOP-INET, May, 2006.
9. Wolak J., Mitchell K., and Finkelhor D. Online victimization of youth: Five years later. In *National Center for Missing & Exploited Children Builletin 07-06-025*, National Center for Missing & Exploited Children, Alexandia, VA, 2006.