

Análisis automático de conversaciones para determinar el comportamiento de pederastas

Beatriz Beltrán, Darnes Vilarino, David Pinto

Benemérita Universidad Autónoma de Puebla,
Facultad de Ciencias de la Computación, México

{bbeltran, darnes, dpinto}@cs.buap.mx

Resumen. Debido a que, en la actualidad, el acceso a medios masivos se encuentra al alcance de cualquier persona, en especial de los niños y sobre todo al no existir restricciones para su acceso, personas sin escrúpulos, tales como los pederastas y pedófilos andan en búsqueda de nuevas víctimas en estos sitios. Es por eso que encontrar formas de identificar acciones que llevan a cabo los pederastas, tales como, propuestas indebidas toma relevancia. En el presente trabajo, se realiza una investigación acerca de la forma en que los pedófilos convencen a sus posibles víctimas, tal análisis se realiza a partir de las intervenciones que tienen los pederastas en salas de conversaciones.

Palabras clave: Pederastas, víctimas, reconocimiento de patrones.

Automatic Analysis of Conversations to Determine Pederast Behaviors

Abstract. Nowadays, access to mass media is available to everyone, especially to children, there exist an absence of restrictions for accessing those media, therefore, unscrupulous persons such as pedophiles and pederasts take advantage of this issue for looking new victims in these sites. This is why finding ways to identify actions carried out by pederasts, such as, inappropriate proposals becomes to be relevant. In this paper we present an investigation about the manner in which pederasts convince their potential victims; such an analysis is performed by using conversations established by pederasts in chat rooms.

Keywords. Pederasts, victims, pattern recognition.

1. Introducción

Se han desarrollado diversas investigaciones para la búsqueda de depredadores sexuales en la red. Con el desarrollo de internet cada día es más fácil que un individuo

trate de interactuar con otros en búsqueda de algún favor sexual, sin que la víctima lo detecte de manera directa.

La mayoría de los trabajos que abordan esta temática toman como punto de referencia la investigación presentada en Pendar [1], donde se realiza un estudio piloto sobre el uso de técnicas de clasificación automática de textos para identificar depredadores sexuales en línea. Se trabaja con un corpus obtenido del sitio Perverted Justice¹, el cual recopila conversaciones de pedófilos con personas que se hacen pasar por niños o adolescentes en la región de Estados Unidos. Este estudio se presenta como una investigación inicial para la detección del *grooming attack*, definido en Harms [2] como “proceso de comunicación por el cual un autor aplica estrategias de búsqueda de afinidad, mientras que simultáneamente adquiere información sobre sus víctimas con el fin de desarrollar las relaciones que resulten en cumplimiento de su necesidad, por ejemplo, acoso sexual físico”.

En los experimentos realizados, se identifican los textos de los depredadores sexuales de las víctimas (clasificación en dos clases), por lo tanto, se separan los diálogos por autor, posteriormente se extraen las palabras cerradas, además de hacer una corrección a las conversaciones, como quitar las letras repetidas de las palabras, remover signos de puntuación, emoticones, entre otros y se obtienen unigramas, bigramas y trigramas de palabras. Los experimentos fueron realizados utilizando SVM y k-NN. Se hicieron varias pruebas con diferentes características (de 5,000 a 10,000), obteniendo los mejores resultados con un conjunto de 10,000 características con k-NN (k=30), con un *F-measure* de 94%.

En los experimentos realizados en Michalopoulos & Mavridis [3] también se intenta reconocer el *grooming attack* utilizando técnicas de clasificación de documentos para la creación de patrones. Se comparan siete diferentes algoritmos para un corpus que contiene tres clases:

1. Ganarse a la víctima.
2. Cultivo de una relación amorosa.
3. Petición de favor sexual o abuso.

Estas clases describen el nivel de relación existente entre el depredador y la víctima, y permiten hacer una clasificación de los tipos de depredadores. Se utilizó Naïve Bayes, el algoritmo “Esperanza-Maximización” (EM), k-NN (con k=24), EM-SIMPLE, TF-IDF, máxima entropía y SVM, obteniendo una mejor precisión con Naïve Bayes (0.96), además de un tiempo menor de clasificación.

En Miah et al. [4] se identifica cuando ocurre explotación infantil en una conversación. Realizan una comparación entre el uso de características basadas en términos con las extraídas utilizando la herramienta *Linguistic Inquiry and WordCount* (LIWC) [5], el cual es un sistema para calcular el grado en que las personas usan ciertas categorías de palabras para analizar aspectos como decepción, honestidad, entre otros. LIWC acepta un texto de entrada y produce variables de salida, las cuales se agrupan en 4 categorías:

¹ <http://www.perverted-justice.com/>

1. **Procesos lingüísticos:** Palabras, pronombres, verbos, entre otros.
2. **Procesos psicológicos:** Sociales, afectivos, cognitivos, de percepción, biológicos, entre otros.
3. **Preocupaciones personales:** Trabajo, casa, dinero, religión, etc.
4. **Categorías del habla:** Asentimientos, influencias, rellenos.

Las conversaciones fueron obtenidas de los sitios perverted.justice.com, chatdump.com y www.fugly.com, y fueron etiquetadas de acuerdo a las siguientes categorías:

- EI: Explotación infantil (200 chats).
- FS: Fantasías sexuales entre adultos (85 chats).
- CG: Chat general sin contenido sexual (107 chats).

Se construyeron 4 distintos conjuntos de datos con las diferentes categorías (EIFS, EI-CG, FS-CG y EI-FS-CG). Estos se clasificaron con Naïve Bayes, árboles de decisión y clasificación vía regresión. Se reporta el número de conversaciones clasificadas correctamente en cada caso, obteniendo los mejores resultados con el uso de LIWC con Naïve Bayes, particularmente en la detección de la categoría “EI”, con alrededor de 185 conversaciones detectadas en los 4 conjuntos contruidos.

La presente investigación intenta encontrar patrones de escritura en las conversaciones entre los depredadores y sus víctimas, partiendo de la hipótesis de que cuando la longitud de una de las intervenciones (ya sea depredador o víctima) es más larga que la otra, entonces, asumimos que se está aportando más información y, por tanto, esa fase podría ser de importancia en la detección de acoso sexual, por ejemplo.

El trabajo está dividido como sigue. En la Sección 2, se describe la metodología que se llevó a cabo para el análisis de la hipótesis planteada. En la sección 3 se analizan los resultados, partiendo de la metodología sobre un análisis de las conversaciones. Finalmente, en la sección 4 se presentan las conclusiones del presente trabajo.

2. Metodología

Se trabajó con un corpus obtenido del sitio Perverted Justice, de donde se extrajeron conversaciones de pederastas. Perverted Justice tiene como objetivo el prevenir a las familias del acoso sexual de pederastas, debido a que en redes sociales y salas de chat es muy simple para estos individuos conectarse con mujeres y niños; estas personas por lo regular se hacen pasar por menores para iniciar una plática, hasta lograr su objetivo. El sitio mantiene información actualizada de las personas que fueron capturadas.

En este trabajo se propone detectar patrones de comportamiento en las conversaciones que mantienen los pederastas con una víctima. El procedimiento propuesto se describe en la Fig.1; se inicia con la descarga de conversaciones desde el sitio web Perverted Justice, de donde se obtiene un archivo por cada pederasta. Posteriormente se procede a obtener el grado de interés dentro de la plática, obteniendo una gráfica para cada conversación de cada uno de los pederastas, para finalmente realizar un análisis basado en grado de interés en la conversación.

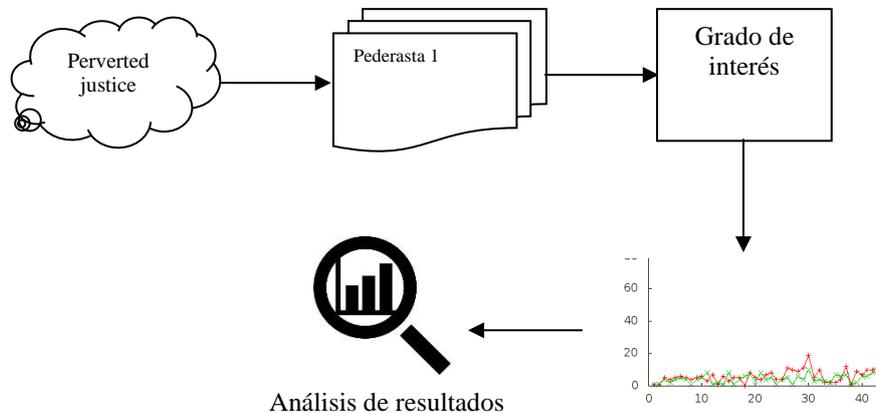


Fig. 1. Esquema del Sistema en general.

En total se descargaron 20 conversaciones correspondientes a 20 diferentes depredadores sexuales. Las características de las conversaciones se presentan en la Tabla 1.

Tabla 1. Características de las 20 conversaciones.

	Total de Intervenciones	Longitud Promedio	Tamaño Vocabulario	Cantidad de Palabras
Depredador 1	360	9.744	529	3836
Victima 1	360	4.369	299	1657
Depredador 2	101	1.762	26	155
Victima 2	101	5.425	153	540
Depredador 3	279	9.358	405	2665
Victima 3	279	6.602	333	1901
Depredador 4	321	15.336	619	4957
Victima 4	321	7.579	394	2450
Depredador 5	142	13	209	1856
Victima 5	142	7.352	214	1044
Depredador 6	773	10.126	897	8749
Victima 6	773	6.598	653	5388
Depredador 7	770	5.906	711	4598
Victima 7	770	5.210	671	4063
Depredador 8	4921	7.011	3679	35164
Victima 8	4921	5.036	2662	24876
Depredador 9	68	6.235	63	568
Victima 9	68	7.897	78	559
Depredador 10	139	6.510	207	1525
Victima 10	139	7.676	218	1130
Depredador 11	367	8.678	545	3191
Victima 11	367	5.896	412	2169

Depredador 12	2258	10.680	2829	28216
Victima 12	2258	4.604	1348	22852
Depredador 13	70	7.471	117	526
Victima 13	70	5.728	110	401
Depredador 14	2146	7.461	2292	16210
Victima 14	2146	5.486	1602	12066
Depredador 15	1296	7.158	1112	8126
Victima 15	1296	6.968	1103	8517
Depredador 16	191	8.005	321	2023
Victima 16	191	4.340	216	997
Depredador 17	394	6.398	410	2741
Victima 17	394	5.111	373	2096
Depredador 18	735	9.997	877	7722
Victima 18	735	5.382	570	4045
Depredador 19	1362	7.253	1264	9880
Victima 19	1362	7.718	1405	10550
Depredador 20	157	5.095	224	1009
Victima 20	157	5.859	247	977

3. Análisis de resultados

Para la búsqueda automática de los patrones y poder detectar el momento en que el acosador inicia el ataque se realizó lo siguiente:

1. Cuantificar el tamaño de cada intervención.
2. Graficar el tamaño de las intervenciones.
3. Analizar las gráficas obtenidas.

A continuación, se realiza un análisis de 2 conversaciones a partir de las gráficas obtenidas, considerando que se detectan los momentos donde se incrementa la interacción entre el acosador y la víctima, en aras de comprobar la hipótesis planteada.

En la Fig. 2, se puede observar que en el período de intervenciones de la 75 a la 100, la víctima tiene más intervenciones que el pederasta, y dentro de la conversación se tiene que durante ese tiempo es cuando la víctima ofrece información personal, como por ejemplo, que actividades lleva a cabo de manera cotidiana, si su mamá se encuentra trabajando, a qué hora puede salir de su casa, etc.

Así mismo, en el intervalo de 125 al 250, es cuando el pederasta tiene una mayor intervención, durante este tiempo es en el cual el pederasta inicia el convencimiento para llegar a tener un contacto físico, realizando comentarios de índole sexual.

Parte de la conversación analizada se muestra en la Fig. 3, cabe hacer mención que las conversaciones son realizadas en el idioma inglés.

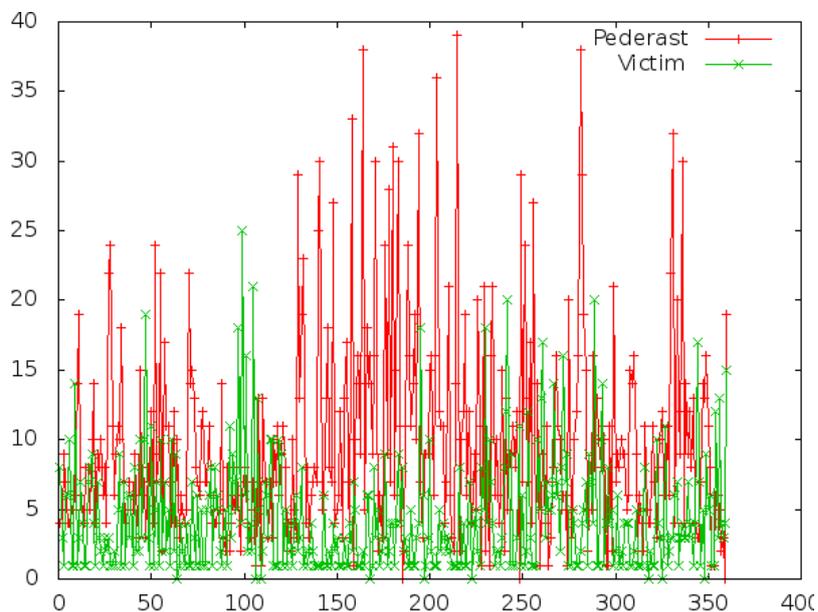


Fig. 2. Conversación del pederasta con el sobrenombre=Arthinice.

189 arthinice (6:52:47 PM): know how it has that silky slit in the middle?
189 arthinice (6:53:31 PM): the part between the pussy lips? (A description. Great)
189 sadlilgrll (6:53:31 PM): yeah
190 arthinice (6:53:48 PM): i'd want to stroke my tongue up and down that part (I bet you would)
190 sadlilgrll (6:54:00 PM): oh
190 sadlilgrll (6:54:03 PM): does that feel good?
191 arthinice (6:54:05 PM): especially if it is really wet there
191 arthinice (6:54:16 PM): feel good to me? or to you?
191 sadlilgrll (6:54:29 PM): either?
192 arthinice (6:54:37 PM): i like the feel and the taste
192 arthinice (6:54:47 PM): but it is you that would be feeling good
192 arthinice (6:54:54 PM): that tingly feeling
192 sadlilgrll (6:55:04 PM): cool
193 arthinice (6:55:04 PM): but a lot lot lot stronger
193 arthinice (6:55:47 PM): think you'd like that?
193 sadlilgrll (6:56:28 PM): i think so

Fig. 3. Conversación analizada.

En otra conversación analizada, se puede ver un comportamiento muy similar, en las primeras intervenciones de la conversación, el pederasta se concreta a preguntar cosas acerca de la vida cotidiana de la víctima, y es a partir de la intervención 100, cuando inicia el ataque con comentarios de posesión y en donde en base a una pregunta de la

víctima, acerca de un sitio, que se hace llamar “*Legion of master*”, inicia una conversación para controlar a su víctima a tal grado que lo llama “*master*”.

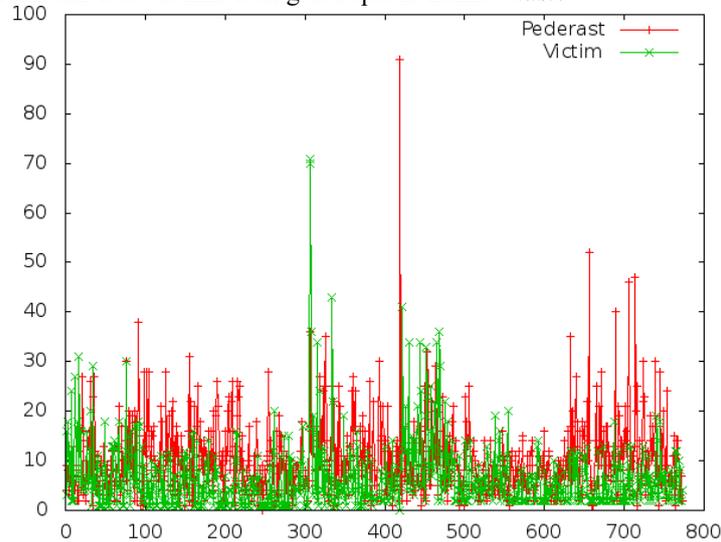


Fig. 4. Conversación del pederasta con el sobrenombre=*Fleet_captain_jaime_wolfe*.

En el punto 306, la víctima tiene que disculparse por lo que el pederasta llama una ofensa y por lo cual la víctima se ve en la necesidad de pedir perdón además de explicar situaciones familiares.

```

617 fleet_captain_jaime_wolfe (6:02:35 PM): Hello girl... How does my girl feel? (He's very
possessive.)
617 sadlilgrrl (6:02:39 PM): much better master
618 fleet_captain_jaime_wolfe (6:02:51 PM): Good... And, is she looking forward to to-
night?
618 sadlilgrrl (6:02:59 PM): very much so master
619 fleet_captain_jaime_wolfe (6:03:09 PM): Good... Does she need what will happen?
619 sadlilgrrl (6:03:15 PM): yes master
620 fleet_captain_jaime_wolfe (6:03:28 PM): Good, & do you have anything special you
want me to do to you?
620 sadlilgrrl (6:03:42 PM): whatever you want master
620 sadlilgrrl (6:03:47 PM): im really excited master
621 fleet_captain_jaime_wolfe (6:03:51 PM): Good... So you are up for anything?
621 sadlilgrrl (6:03:56 PM): yes master
622 fleet_captain_jaime_wolfe (6:04:00 PM): How excited? Are you nice & wet for me?
622 sadlilgrrl (6:04:06 PM): yes master
623 fleet_captain_jaime_wolfe (6:04:26 PM): Good... And, are you looking forward to feel-
ing your nipples clamped?
    
```

Fig. 5. Conversación analizada.

Al final de la conversación, el control ya llega a tal grado que la víctima solo contesta con *yes/no* *master*. En todo momento la conversación gira alrededor del tema sexual y sobre el control que él tiene sobre ella. Además de que demuestra complacencia por la actitud que la víctima toma, al comportarse en todo momento de manera sumisa y con actitud de querer complacerlo en todo.

4. Conclusiones

El problema del acoso a menores en redes sociales ha aumentado cada año, por lo cual el poder identificar patrones que se encuentren asociados a los pederastas, ayudaría a los diferentes elementos de seguridad y a la sociedad en sí misma.

Este trabajo contribuye con el análisis automático de comportamientos para encontrar dichos patrones, y es así como se pudo comprobar que en los momentos en que el acosador requiere información personal de su víctima, se incrementan las intervenciones de la víctima, incluso, cuando ya tiene dominada a la víctima y el pederasta no le gusta lo expresado por su víctima y entonces, esta se tiene que disculpar, se ve un comportamiento semejante de mayor participación de la víctima.

Sin embargo, este comportamiento se invierte completamente cuando el pederasta está realizando la labor de consentimiento, convencimiento e incluso cuando logra algo, como el obtener fotografías, sometimiento o hasta llegar a una cita.

Referencias

1. Pendar, N.: Toward spotting the pedophile telling victim from predator in text chats. In: Proceedings of the International Conference on Semantic Computing, ICSC '07, IEEE Computer Society, 235–241 (2007)
2. Harms, C. M. Grooming: An operational definition and coding scheme. 8(1), 1–6 (2007).
3. Michalopoulos, D. & Mavridis, I.: Utilizing document classification for grooming attack recognition. In: Proceedings of the 2011 IEEE Symposium on Computers and Communications, ISCC '11, IEEE Computer Society, 864–869 (2011)
4. Miah, M. W. R., Yearwood, J., Kulkarni, S.: Detection of child exploiting chats from a mixed chat dataset as a text classification task. In: Proceedings of the Australasian Language Technology Association Workshop 2011, 157–165 (2011)
5. Pennebaker, J. W., Chung, C. K., Ireland, M., Gonzales, A., Booth, R. J.: The Development and Psychometric Properties of LIWC2007. 3–14 (2007)