Overview of the International Sexual Predator Identification Competition at PAN-2012

Giacomo Inches and Fabio Crestani
Faculty of Informatics
University of Lugano (USI)
Switzerland
Motivations

• Instant Messaging software: facilitate **new connections** between persons or **reinforce existing** ones but...

• allow for **misbehaviours** or **cybercriminal** acts!

• aligned with my Ph.D. work ;)

Thursday, September 20, 12
Motivations

Facebook analyzes relationships and chats to flag up sexual predators

Social networks scan for sexual predators, with uneven results
International Sexual Predator Identification Competition

• Exploratory (sub)task in PAN at CLEF 2012

• Challenging (reusable) collection

• Challenging goals

• Unique place for researchers to discuss this task
“sexual predator” ?!

• “a person or group that ruthlessly exploits others”, New Oxford American Dictionary

• “is used pejoratively to describe a person seen as obtaining or trying to obtain sexual contact with another person in a metaphorically ‘predatory’ manner”, Wikipedia.
“sexual predator” ?!
“sexual predator” ?!
“sexual predator” ?!
Previous works

• **Kontostathis** et. al [9, 12]
  - PJ website

• **Pendar** [16]
  - I) Predator/Other
    - (Ia) Predator/Victim (victim is underage);
    - (Ib) Predator/Pseudo-Victim (volunteer posing as child);
    - (Ic) Predator/Pseudo-Victim (law enforcement officer posing as child)
  - II) Adult/Adult (consensual relationship)

• **Other domains**
Previous works

- **Kontostathis et. al** [9, 12]
  - PJ website

- **Pendar** [16]
  - I) Predator/Other
    - (Ia) Predator/Victim (victim is underage)
    - (Ib) Predator/Pseudo-Victim (volunteer posing as child)
    - (Ic) Predator/Pseudo-Victim (law enforcement officer posing as child)
  - II) Adult/Adult (consensual relationship)

- **Other domains**
Collection Composition

Looking for applicants to our training program
We have undertaken a massive effort to rework how we bring on new volunteers to learn how to do the chat-logs that you see on the website. From learning how to collect evidence in a professional and credible manner to learning how to chat with these individuals to secure an arrest, our online e-learning program will help out.

Anyone is eligible to apply given that they're over 21 years of age with a clean criminal background. Full details can be found at the following link: [Application Information]

Wednesday, March 28th, 2012
Chatlog Conviction #549
- [Chatlog](#) - Ogden, Utah
Felipe Koldalschi received 57 months for his chat.
[Read Report] Links to Chat Log:
[With Commentary] [No Commentary]

Wednesday, October 19th, 2011
Chatlog Conviction #548
- [Chatlog](#) - Wales, Michigan
Joseph Faronino isn't really 18, but he is definitely a predator.
[Read Report] Links to Chat Log:
[With Commentary] [No Commentary]

Monday, August 22nd, 2011
Chatlog Conviction #547
- [Chatlog](#) - Ypsilanti, Michigan
Joshua Blake Hoe will no longer be running summer camp.
[Read Report] Links to Chat Log:
[With Commentary] [No Commentary]

New Canadian resource profiles pedophiles
Over the years, Canadians have been asking how to help out, or for links to resources regarding pedophiles and other sexual offenders against children. Recently a pretty good resource was launched in Canada as detailed in this Ottawa Sun article: Convicted Pedophile Outing Website Goes Live.

The website, StopPedophiles.ca, is set up with resources for Canadian parents and links to identities of convicted pedophiles set up by province, and each listing has articles to media stories regarding the individuals abuse case.

Canadians should check out the site and support greater efforts to drag the culprits of such abuse out into the sunlight.

Conviction Counter
Number of predators convicted due to Perverted-Justice.com since June 2004:

550
Chatlog Convictions: 549
Research Convictions: 1

Real Stories Project
For the next six months, we're documenting actual cases of internet abductions and molestations. One case a day for at least six months. Beware, these stories can be disturbing.

02/28/10: Man, 27, rapes girl, 12
A 27 year old Delaware man was arrested for raping a 12 year old girl he met online.

05/25/10: Man, 22, rapes girl, 13
A 22 year old North Carolina registered sex offender was arrested for raping a 13 year old girl he met through Facebook.

06/24/10: Man, 26, assaults teen girls
A 26 year old Texas man was arrested for sexually assaulting a 13 year old girl. During the investigation police learned that the man had sexually assaulted at least five other girls. He met his victims on Facebook and MySpace.
Collection Composition

<table>
<thead>
<tr>
<th>True</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRUE</td>
<td>Perverted Justice</td>
<td>Regular conversations</td>
</tr>
<tr>
<td>FALSE</td>
<td>Sexual Related Conversations</td>
<td></td>
</tr>
</tbody>
</table>
Collection Composition

<table>
<thead>
<tr>
<th>TRUE</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perverted Justice</td>
<td>✔️</td>
<td>Regular conversations</td>
</tr>
<tr>
<td>Sexual Related Conversations</td>
<td>❓</td>
<td>❓</td>
</tr>
</tbody>
</table>
Collection Composition

Krjin (http://krijnhoetmer.nl/irc-logs)
#whatwg #html-wg #html5 #css #fx #webapps
#microformats #accessibility #developers

Irclog (www. irclog.org)
#aix #apache #azureus #blender #c #c++
cisco #csharp #css #debian #fedora #flood
#freebsd #gentoo #gentoo-dev #gtk #hardware
#html #iptables #irix #java #javascript #linux-bg
#macosx #mysq #netbsd #openbsd
#opensolaris #oracle #php #python #qt
#reactos #samba #solaris #suse #tomcat
#ubuntu #vim #windows #wireless

Omegle (http://omegle.inportb.com)
“abusive language and general silliness online” and sometimes users “engage in cybersex”

Conversation:
messages exchange separated by a break of 25+ minutes and with a duration \( \leq 150 \) exchanges

Training
~ 67k conv.
~ 97k users

Testing
~ 155k conv.
~ 218k users.
# Collection Statistics

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>#conversations</td>
<td>11350</td>
<td>50510</td>
<td>28501</td>
<td>267261</td>
</tr>
<tr>
<td>#conv. length ≤ 150</td>
<td>9076</td>
<td>48569</td>
<td>21896</td>
<td>265747</td>
</tr>
<tr>
<td>(% all)</td>
<td>(80%)</td>
<td>(96%)</td>
<td>(77%)</td>
<td>(99%)</td>
</tr>
</tbody>
</table>
## Collection Statistics

<table>
<thead>
<tr>
<th></th>
<th>PJ perverted-justice.com</th>
<th>krjin krijnhoetmer.nl/irc-logs</th>
<th>irclog irclog.org</th>
<th>omegle omegle.inportb.com</th>
</tr>
</thead>
<tbody>
<tr>
<td>#conversations</td>
<td>11350</td>
<td>50510</td>
<td>28501</td>
<td>267261</td>
</tr>
<tr>
<td>#conv. length ≤ 150</td>
<td>9076</td>
<td>48569</td>
<td>21896</td>
<td>265747</td>
</tr>
<tr>
<td>(% all)</td>
<td>(80%)</td>
<td>(96%)</td>
<td>(77%)</td>
<td>(99%)</td>
</tr>
</tbody>
</table>

### Training set

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>#conv. length ≤ 150</td>
<td>2723</td>
<td>14571</td>
<td>6569</td>
<td>43064</td>
</tr>
<tr>
<td>” and exactly 2 user</td>
<td>984</td>
<td>2420</td>
<td>1146</td>
<td>41067</td>
</tr>
<tr>
<td>(% training)</td>
<td>(36%)</td>
<td>(17%)</td>
<td>(17%)</td>
<td>(95%)</td>
</tr>
<tr>
<td>unique (perverted)</td>
<td>291 (142)</td>
<td>2660</td>
<td>10613</td>
<td>84131</td>
</tr>
</tbody>
</table>
## Collection Statistics

<table>
<thead>
<tr>
<th></th>
<th>PJ perverted-justice.com</th>
<th>krjin krijnhoetmer.nl/irc-logs</th>
<th>irclog irclog.org</th>
<th>omegle omegle.inportb.com</th>
</tr>
</thead>
<tbody>
<tr>
<td>#conversations</td>
<td>11350</td>
<td>50510</td>
<td>28501</td>
<td>267261</td>
</tr>
<tr>
<td>#conv. length ≤150</td>
<td>9076</td>
<td>48569</td>
<td>21896</td>
<td>265747</td>
</tr>
<tr>
<td>(% all)</td>
<td>(80%)</td>
<td>(96%)</td>
<td>(77%)</td>
<td>(99%)</td>
</tr>
</tbody>
</table>

### Training set

<table>
<thead>
<tr>
<th></th>
<th>PJ perverted-justice.com</th>
<th>krjin krijnhoetmer.nl/irc-logs</th>
<th>irclog irclog.org</th>
<th>omegle omegle.inportb.com</th>
</tr>
</thead>
<tbody>
<tr>
<td>#conv. length ≤150</td>
<td>2723</td>
<td>14571</td>
<td>6569</td>
<td>43064</td>
</tr>
<tr>
<td>” and exactly 2 user</td>
<td>984</td>
<td>2420</td>
<td>1146</td>
<td>41067</td>
</tr>
<tr>
<td>(% training)</td>
<td>(36%)</td>
<td>(17%)</td>
<td>(17%)</td>
<td>(95%)</td>
</tr>
<tr>
<td>unique (perverted) users</td>
<td>291 (142)</td>
<td>2660</td>
<td>10613</td>
<td>84131</td>
</tr>
</tbody>
</table>

### Testing set

<table>
<thead>
<tr>
<th></th>
<th>PJ perverted-justice.com</th>
<th>krjin krijnhoetmer.nl/irc-logs</th>
<th>irclog irclog.org</th>
<th>omegle omegle.inportb.com</th>
</tr>
</thead>
<tbody>
<tr>
<td>#conv. length ≤150</td>
<td>5321</td>
<td>33998</td>
<td>15327</td>
<td>100482</td>
</tr>
<tr>
<td>” and exactly 2 user</td>
<td>1887</td>
<td>5648</td>
<td>2673</td>
<td>95648</td>
</tr>
<tr>
<td>(% testing)</td>
<td>(35%)</td>
<td>(17%)</td>
<td>(17%)</td>
<td>(95%)</td>
</tr>
<tr>
<td>unique (perverted) users</td>
<td>440 (254)</td>
<td>4358</td>
<td>17788</td>
<td>196130</td>
</tr>
</tbody>
</table>
Task goal

• given a collection containing chat logs involving two (or more) persons:

  • **identify** the *predators* among all users in the different conversations - problem 1

  • **identify** the *part* (the lines) of the conversations which are the *most distinctive* of the predator behavior - problem 2
Participants

16 submissions
12 notebooks
Problem 1

“identify the predators among all users in the different conversations”

"Were you serious about... the hard-core ----ing?"

icetruckkiller103

"We'll see what happens."
Results

![Bar chart showing precision for different runs](chart.png)
Results
Results

- villatorotello-run-2012-06-15-2157g
- snider12-run-2012-06-16-0032
- parapar12-run-2012-06-15-0959j
- morris12-run-2012-06-16-0752-main
- eriksson12-run-2012-06-15-1949
- peersman12-run-2012-06-15-1559
- grozea12-run-2012-06-14-1706b
- sitarz12-run-2012-0615-1515
- vartapetiance12-run-2012-06-15-1411
- kontostathis-run-2012-06-16-0317e
- kang12-run-2012-06-15-0904b
- kern12-run-2012-06-18-1827b
- bogdanova12-run-2012-06-14-1117
- prasath12-run-2012-06-15-2122
- vilarino12-run-2012-06-14-2121b
- gomezhidalgo12-2012-06-15-1900

صلاح
Approaches

• pre-filtering
  • presence of 1 participants only
  • less than 6 interventions per user
  • 3 long sequences of unrecognised characters

• features
  • lexical
  • behavioral
  • character 5-gram

• classifiers
  • Neural Network classifier [19]
  • Maximum-Entropy [4,8]
  • SVMs [13,14,15,19]
  • decision trees[10]
  • k-NN [7,17]
  • random forest [17]
  • Naïve Bayes [6,1]

• line based / author based
Approaches

- **pre-filtering**
  - presence of 1 participants only
  - less then 6 interventions per user
  - 3 long sequences of unrecognised characters

- **features**
  - lexical
  - behavioral
  - character 5-gram

- **classifiers**
  - Neural Network classifier [19]
  - Maximum-Entropy [4,8]
  - SVMs [13,14,15,19]
  - decision trees[10]
  - k-NN [7,17]
  - random forest [17]
  - Naïve Bayes [6,1]
  - line based / author based
Problem 2

“identify the **part** (the lines) of the conversations which are the **most distinctive** of the predator behaviour”
Results

- grozea12-run-2012-06-14-1706b
- kontostathis-run-2012-06-16-0317e
- peersman12-run-2012-06-15-1559
- sitarz12-run-2012-0615-1515
- morris12-run-2012-06-16-0752-main
- kern12-run-2012-06-18-1827b
- eriksson12-run-2012-06-15-1949
- prasath12-run-2012-06-15-2122
- parapar12-run-2012-06-15-0959j
- vartapetiance12-run-2012-06-15-1411
- vilarino12-run-2012-06-14-2121b
- bogdanova12-run-2012-06-14-1117
- villatorotello-run-2012-06-15-2157g
- gomezhidalgo12-2012-06-15-1900

Precision
Results

- grozea12-run-2012-06-14-1706b
- kontostathis-run-2012-06-16-0317e
- peersman12-run-2012-06-15-1559
- sitarz12-run-2012-0615-1515
- morris12-run-2012-06-16-0752-main
- kern12-run-2012-06-18-1827b
- eriksson12-run-2012-06-15-1949
- prasath12-run-2012-06-15-2122
- parapar12-run-2012-06-15-0959j
- vartapetiance12-run-2012-06-15-1411
- vilarino12-run-2012-06-14-2121b
- bogdanova12-run-2012-06-14-1117
- villatorotello-run-2012-06-15-2157g
- gomezhidalgo12-2012-06-15-1900
Results

F beta 1

0.2
0.4
0.6
0.8
1

grosea12-run-2012-06-14-1706b
kontostathis-run-2012-06-16-0317e
peersman12-run-2012-06-15-1559
sitarz12-run-2012-0615-1515
morris12-run-2012-06-16-0752-main
kern12-run-2012-06-18-1827b
eriksson12-run-2012-06-15-1949
prasath12-run-2012-06-15-2122
parapar12-run-2012-06-15-0959j
vartapetiance12-run-2012-06-15-1411
vilarino12-run-2012-06-14-2121b
bogdanova12-run-2012-06-14-1117
villatorotello-run-2012-06-15-2157g
gomezhidalgo12-2012-06-15-1900
Results

groza12-run-2012-06-14-1706b
kontostathis-run-2012-06-16-0317e
peersman12-run-2012-06-15-1559
sitarz12-run-2012-0615-1515
morris12-run-2012-06-16-0752-main
kern12-run-2012-06-18-1827b
eriksson12-run-2012-06-15-1949
prasath12-run-2012-06-15-2122
parapar12-run-2012-06-15-0959j
vartapetiance12-run-2012-06-15-1411
vilarino12-run-2012-06-14-2121b
bogdanova12-run-2012-06-14-1117
villatorotello-run-2012-06-15-2157g
gomezhidalgo12-2012-06-15-1900
Approaches

• All the **conversations lines** of all the **identified predators** from problem 1 [17] (characters 5-gram + SVM)

• Filtering
  
  • dictionary of “perverted” terms
  
  • terms with particular score (e.g. TF-IDF weighting) [15,13,14,4]
  
  • LM [19]

• **Lines** already labelled as predatory for problem 1
# Discussions

<table>
<thead>
<tr>
<th>Issues</th>
<th>Possible Improvements</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Performance Evaluation</strong></td>
<td></td>
</tr>
<tr>
<td>• unranked list of arbitrary number of suspects (1)</td>
<td>• ordered list of k suspects - prioritization (1)</td>
</tr>
<tr>
<td>• 1 expert only (2)</td>
<td>• more experts (2) - any volunteer?</td>
</tr>
<tr>
<td><strong>Collection</strong></td>
<td><strong>Task</strong></td>
</tr>
<tr>
<td>• too easy? - precision 98% (1)</td>
<td>• include <strong>true negative</strong> documents on topic with true positive</td>
</tr>
<tr>
<td></td>
<td>• hidden collection, access with API</td>
</tr>
<tr>
<td><strong>....</strong></td>
<td><strong>....</strong></td>
</tr>
<tr>
<td></td>
<td>• include both predator and not predator lines</td>
</tr>
<tr>
<td></td>
<td>• ...</td>
</tr>
</tbody>
</table>
Conclusions

Sexual Predator Identification

- Interesting & realistic topic
- Automatic methods applicable
- Room for improvements
I'm looking for new data!

Uncle Sam, PAN@CLEF 2013
Giacomo Inches - PhD Student

giacomo.inches@usi.ch
www.giacomo.inches.ch

Faculty of Informatics
University of Lugano (USI)
Via G. Buffi 13
CH - 6904 Lugano
Switzerland
Thanks for participating!

Giacomo Inches - PhD Student

giacomo.inches@usi.ch
www.giacomo.inches.ch
<table>
<thead>
<tr>
<th>Participant run</th>
<th>RETR, REL</th>
<th>P</th>
<th>R</th>
<th>$F_{\beta=1}$</th>
<th>$F_{\beta=0.5}$</th>
<th>Official run rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>villatoro12-run-2012-06-15-2157g</td>
<td>204</td>
<td>200</td>
<td>0.9804</td>
<td>0.8754</td>
<td>0.9346</td>
<td>1</td>
</tr>
<tr>
<td>snider12-run-2012-06-16-0032</td>
<td>186</td>
<td>183</td>
<td>0.9839</td>
<td>0.7205</td>
<td>0.8318</td>
<td>2</td>
</tr>
<tr>
<td>villatoro12-run-2012-06-15-2157c</td>
<td>211</td>
<td>200</td>
<td>0.9479</td>
<td>0.7854</td>
<td>0.8610</td>
<td>3</td>
</tr>
<tr>
<td>parpar12-run-2012-06-15-0959j</td>
<td>181</td>
<td>170</td>
<td>0.9392</td>
<td>0.6693</td>
<td>0.7816</td>
<td>4</td>
</tr>
<tr>
<td>morris12-run-2012-06-16-0752-main</td>
<td>159</td>
<td>154</td>
<td>0.9686</td>
<td>0.6063</td>
<td>0.7458</td>
<td>5</td>
</tr>
<tr>
<td>eriksson12-run-2012-06-15-1949</td>
<td>265</td>
<td>227</td>
<td>0.8566</td>
<td>0.8937</td>
<td>0.8748</td>
<td>6</td>
</tr>
<tr>
<td>parpar12-run-2012-06-15-0959g</td>
<td>171</td>
<td>162</td>
<td>0.9474</td>
<td>0.6378</td>
<td>0.7624</td>
<td>7</td>
</tr>
<tr>
<td>morris12-run-2012-06-17-0126</td>
<td>152</td>
<td>147</td>
<td>0.9671</td>
<td>0.5787</td>
<td>0.7241</td>
<td>8</td>
</tr>
<tr>
<td>parpar12-run-2012-06-15-0959l</td>
<td>173</td>
<td>161</td>
<td>0.9306</td>
<td>0.6339</td>
<td>0.7541</td>
<td>9</td>
</tr>
<tr>
<td>parpar12-run-2012-06-15-0959k</td>
<td>182</td>
<td>164</td>
<td>0.9011</td>
<td>0.6457</td>
<td>0.7523</td>
<td>10</td>
</tr>
<tr>
<td>peersman12-run-2012-06-15-1559</td>
<td>170</td>
<td>152</td>
<td>0.8941</td>
<td>0.5984</td>
<td>0.7170</td>
<td>11</td>
</tr>
<tr>
<td>parpar12-run-2012-06-15-0959d</td>
<td>175</td>
<td>151</td>
<td>0.8629</td>
<td>0.5395</td>
<td>0.7040</td>
<td>12</td>
</tr>
<tr>
<td>parpar12-run-2012-06-15-0959c</td>
<td>169</td>
<td>145</td>
<td>0.8580</td>
<td>0.5709</td>
<td>0.6586</td>
<td>13</td>
</tr>
<tr>
<td>villatoro12-run-2012-06-15-2157a</td>
<td>108</td>
<td>103</td>
<td>0.9537</td>
<td>0.4055</td>
<td>0.5691</td>
<td>14</td>
</tr>
<tr>
<td>parpar12-run-2012-06-15-0959b</td>
<td>205</td>
<td>160</td>
<td>0.7805</td>
<td>0.6299</td>
<td>0.6972</td>
<td>15</td>
</tr>
<tr>
<td>grozea12-run-2012-06-14-1706c</td>
<td>215</td>
<td>163</td>
<td>0.7381</td>
<td>0.6417</td>
<td>0.6951</td>
<td>16</td>
</tr>
<tr>
<td>parpar12-run-2012-06-15-0959f</td>
<td>202</td>
<td>154</td>
<td>0.7624</td>
<td>0.6063</td>
<td>0.6734</td>
<td>17</td>
</tr>
<tr>
<td>siar12-run-2012-06-15-1515b</td>
<td>218</td>
<td>159</td>
<td>0.7294</td>
<td>0.6260</td>
<td>0.6737</td>
<td>18</td>
</tr>
<tr>
<td>parpar12-run-2012-06-15-0959h</td>
<td>223</td>
<td>161</td>
<td>0.7220</td>
<td>0.6339</td>
<td>0.6751</td>
<td>19</td>
</tr>
<tr>
<td>parpar12-run-2012-06-15-0959a</td>
<td>200</td>
<td>128</td>
<td>0.6400</td>
<td>0.5039</td>
<td>0.5639</td>
<td>20</td>
</tr>
<tr>
<td>vartapetiance12-run-2012-06-15-1411</td>
<td>160</td>
<td>99</td>
<td>0.6188</td>
<td>0.3898</td>
<td>0.4783</td>
<td>21</td>
</tr>
<tr>
<td>villatoro12-run-2012-06-15-2157j</td>
<td>269</td>
<td>143</td>
<td>0.5316</td>
<td>0.3680</td>
<td>0.5468</td>
<td>22</td>
</tr>
<tr>
<td>grozea12-run-2012-06-14-1706a</td>
<td>322</td>
<td>142</td>
<td>0.4410</td>
<td>0.2591</td>
<td>0.4931</td>
<td>23</td>
</tr>
<tr>
<td>kontostathis-run-2012-06-16-0317q</td>
<td>475</td>
<td>170</td>
<td>0.3879</td>
<td>0.6693</td>
<td>0.4664</td>
<td>24</td>
</tr>
<tr>
<td>kortostathis-run-2012-06-16-0317d</td>
<td>688</td>
<td>172</td>
<td>0.2500</td>
<td>0.6772</td>
<td>0.3652</td>
<td>25</td>
</tr>
<tr>
<td>kang12-run-2012-06-15-0904b</td>
<td>930</td>
<td>203</td>
<td>0.2183</td>
<td>0.7992</td>
<td>0.3429</td>
<td>26</td>
</tr>
<tr>
<td>kou12-run-2012-06-15-0904a</td>
<td>1049</td>
<td>202</td>
<td>0.1926</td>
<td>0.7953</td>
<td>0.3101</td>
<td>27</td>
</tr>
<tr>
<td>kern12-run-2012-06-18-1827d</td>
<td>1172</td>
<td>177</td>
<td>0.1510</td>
<td>0.6969</td>
<td>0.2482</td>
<td>28</td>
</tr>
<tr>
<td>villatoro12-run-2012-06-15-2157d</td>
<td>240</td>
<td>36</td>
<td>0.1500</td>
<td>0.1417</td>
<td>0.1457</td>
<td>29</td>
</tr>
<tr>
<td>kontostathis-run-2012-06-16-0317c</td>
<td>3969</td>
<td>206</td>
<td>0.0557</td>
<td>0.8110</td>
<td>0.1043</td>
<td>30</td>
</tr>
<tr>
<td>villatoro12-run-2012-06-15-2157b</td>
<td>204</td>
<td>12</td>
<td>0.0588</td>
<td>0.0472</td>
<td>0.0524</td>
<td>31</td>
</tr>
<tr>
<td>kontostathis-run-2012-06-16-0317a</td>
<td>5225</td>
<td>266</td>
<td>0.0394</td>
<td>0.8110</td>
<td>0.0752</td>
<td>32</td>
</tr>
<tr>
<td>vilarino12-run-2012-06-14-2121b</td>
<td>5625</td>
<td>222</td>
<td>0.0393</td>
<td>0.8071</td>
<td>0.0752</td>
<td>33</td>
</tr>
<tr>
<td>vilarino12-run-2012-06-14-2121a</td>
<td>9071</td>
<td>236</td>
<td>0.0260</td>
<td>0.9291</td>
<td>0.0506</td>
<td>34</td>
</tr>
<tr>
<td>bogdanova12-run-2012-06-14-1117</td>
<td>2109</td>
<td>55</td>
<td>0.0261</td>
<td>0.2165</td>
<td>0.0466</td>
<td>35</td>
</tr>
<tr>
<td>prasath12-run-2012-06-15-2122</td>
<td>10289</td>
<td>207</td>
<td>0.0201</td>
<td>0.8150</td>
<td>0.0393</td>
<td>36</td>
</tr>
<tr>
<td>vilarino12-run-2012-06-14-2121b</td>
<td>5225</td>
<td>98</td>
<td>0.0188</td>
<td>0.8358</td>
<td>0.0358</td>
<td>37</td>
</tr>
<tr>
<td>villatoro12-run-2012-06-15-2157e</td>
<td>305</td>
<td>6</td>
<td>0.0197</td>
<td>0.2236</td>
<td>0.0215</td>
<td>38</td>
</tr>
<tr>
<td>gonzegatal12-run-2012-06-15-1900</td>
<td>150</td>
<td>1</td>
<td>0.0067</td>
<td>0.0039</td>
<td>0.0050</td>
<td>39</td>
</tr>
<tr>
<td>gonzegatal12-run-2012-06-15-1900</td>
<td>150</td>
<td>1</td>
<td>0.0067</td>
<td>0.0039</td>
<td>0.0050</td>
<td>40</td>
</tr>
</tbody>
</table>
Social networks scan for sexual predators, with uneven results

1,186 people recommend this. Be the first of your friends.
Facebook's software likewise depends on relationship analysis and archives of real chats that preceded sex assaults, Chief Security Joe Sullivan told Reuters in the company's most expansive comments on the subject to date.

Like most of its peers, Facebook generally avoids discussing its safety practices to discourage scare stories, because it doesn't catch many wrongdoers, and to sidestep privacy concerns. Users could be unnerved about the extent to which their conversations are reviewed, at least by computer programs.

CATCHING ONE IN 10?

In part because of its massive size, Facebook relies more than some rivals on such technology.

"We've never wanted to set up an environment where we have employees looking at private communications, so it's really important that we use technology that has a very low false-positive rate" he said. In addition, Facebook doesn't probe deeply into what it thinks are pre-existing relationships.

A low rate of false positives, though, also means that many dangerous communications go undetected.

[...]