Measuring eWhoring

Sergio Pastrana\textsuperscript{1}, Alice Hutchings\textsuperscript{2},
Daniel Thomas\textsuperscript{3}, Juan Tapiador\textsuperscript{1}

\textsuperscript{1}Universidad Carlos III de Madrid
\textsuperscript{2}University of Cambridge
\textsuperscript{3}University of Strathclyde

19th Internet Measurement Conference (IMC)
October 23, 2019
Outline

1. Introduction
2. Analysis and provenance of images
3. Analysis of financial gains
4. Analysis of actors
5. Conclusions
Outline

1. Introduction
2. Analysis and provenance of images
3. Analysis of financial gains
4. Analysis of actors
5. Conclusions
eWhoring is an online fraud where scammers imitate partners in (cyber)sexual encounters.

Steps:

1. Learn techniques
2. Obtain images
3. Create an alias, prepare a backstory, open accounts
4. Source traffic
5. Negotiate
6. Receive payment
7. Send images
8. Block or continue to milk customers
9. Exchange funds
In this work we measure three main aspects of eWhoring:

**Provenance and type of Images**
- What are the main hosting services used for sharing images?
- Do packs contain illegal images?
- Where are the images taken from?

**Monetizing and financial profits**
- How much money are fraudsters making?
- What are the preferred payment platforms/currencies?

**Characterization and social relations of actors**
- What are the different roles of actors?
- Who are the key actors? How they have evolved?
In this work we measure three main aspects of eWhoring

- **Provenance and type of Images**
  - What are the main hosting services used for sharing images?
  - Do packs contain illegal images?
  - Where are the images taken from?

- **Monetizing and financial profits**
  - How much money are fraudsters making?
  - What are the preferred payment platforms/currencies?

- **Characterization and social relations of actors**
  - What are the different roles of actors?
  - Who are the key actors? How they have evolved?

We rely on information gathered from underground forums
Ethics are in forefront
This measurement used data gathered from various underground forums.

We use the CrimeBB dataset from the Cambridge Cybercrime Centre.

<table>
<thead>
<tr>
<th>Forum</th>
<th>#Threads</th>
<th>#Posts</th>
<th>First post</th>
<th>#TOPs</th>
<th>#Actors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hackforums</td>
<td>42,292</td>
<td>596,827</td>
<td>11/08</td>
<td>4,027</td>
<td>64,035</td>
</tr>
<tr>
<td>OGUUsers</td>
<td>1,744</td>
<td>23,974</td>
<td>04/17</td>
<td>76</td>
<td>5,586</td>
</tr>
<tr>
<td>BHW</td>
<td>258</td>
<td>2,694</td>
<td>04/08</td>
<td>0</td>
<td>1,420</td>
</tr>
<tr>
<td>V3rmillion</td>
<td>95</td>
<td>1,348</td>
<td>02/16</td>
<td>6</td>
<td>697</td>
</tr>
<tr>
<td>MPGH</td>
<td>62</td>
<td>922</td>
<td>07/12</td>
<td>12</td>
<td>341</td>
</tr>
<tr>
<td>RaidForums</td>
<td>48</td>
<td>405</td>
<td>03/15</td>
<td>10</td>
<td>318</td>
</tr>
<tr>
<td>Others (4):</td>
<td>21</td>
<td>614</td>
<td>05/15</td>
<td>6</td>
<td>586</td>
</tr>
<tr>
<td>TOTAL:</td>
<td>44,520</td>
<td>626,784</td>
<td>11/08-03/19</td>
<td>4,137</td>
<td>72,982</td>
</tr>
</tbody>
</table>

Table 1: Number of eWhoring related conversations per forum. (TOPs = Threads Offering Packs.)
Outline

1. Introduction
2. Analysis and provenance of images
3. Analysis of financial gains
4. Analysis of actors
5. Conclusions
The methodology is designed to minimize the exposure to indecent images.
The methodology is designed to minimize the exposure to indecent images
The methodology is designed to minimize the exposure to indecent images
The methodology is designed to minimize the exposure to indecent images.
The methodology is designed to minimize the exposure to indecent images.
Images are typically shared using third party services for cloud storage and image sharing

- Identified 4,137 TOPs
  - Pack URLs from 774
- 1,255 packs
  - 111,288 images
  - 53,948 unique
- 5,788 preview images

<table>
<thead>
<tr>
<th>Site</th>
<th>#Links</th>
</tr>
</thead>
<tbody>
<tr>
<td>imgur</td>
<td>3,297</td>
</tr>
<tr>
<td>gyazo</td>
<td>1,006</td>
</tr>
<tr>
<td>imageshack</td>
<td>679</td>
</tr>
<tr>
<td>prnt</td>
<td>383</td>
</tr>
<tr>
<td>photobucket</td>
<td>311</td>
</tr>
<tr>
<td>imagetwist</td>
<td>105</td>
</tr>
<tr>
<td>imagezilla</td>
<td>97</td>
</tr>
<tr>
<td>minus</td>
<td>51</td>
</tr>
<tr>
<td>postimage</td>
<td>47</td>
</tr>
<tr>
<td>imagebam</td>
<td>44</td>
</tr>
<tr>
<td>Others</td>
<td>700</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>7,314</strong></td>
</tr>
</tbody>
</table>

Table 3: Number of links per image sharing site.

<table>
<thead>
<tr>
<th>Site</th>
<th>#Links</th>
</tr>
</thead>
<tbody>
<tr>
<td>MediaFire</td>
<td>892</td>
</tr>
<tr>
<td>mega</td>
<td>284</td>
</tr>
<tr>
<td>Dropbox</td>
<td>130</td>
</tr>
<tr>
<td>oron</td>
<td>95</td>
</tr>
<tr>
<td>depositfiles</td>
<td>46</td>
</tr>
<tr>
<td>filefactory</td>
<td>37</td>
</tr>
<tr>
<td>drive.google</td>
<td>31</td>
</tr>
<tr>
<td>ge.tt</td>
<td>28</td>
</tr>
<tr>
<td>zippyshare</td>
<td>25</td>
</tr>
<tr>
<td>filedropper</td>
<td>24</td>
</tr>
<tr>
<td>Others</td>
<td>94</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1,719</strong></td>
</tr>
</tbody>
</table>

Table 4: Number of links per cloud storage service.
Most images depict nudity or explicit sexual content. Packs included child abuse material

- 36 images matched the PhotoDNA service, and thus are *known* illegal images
  - Immediately reported to the Internet Watch Foundation (UK Inhope) and removed
  - 61 URLs were actioned
  - These images have been downloaded from at least 476 actors
  - Forum users discuss and are aware of the age of models
- From the 5,788 preview images, 3,496 (60%) are classified as Non-Safe-For-View (NSFV).
Most images depict nudity or explicit sexual content. Packs included child abuse material.

- 36 images matched the PhotoDNA service, and thus are known illegal images.
  - Immediately reported to the Internet Watch Foundation (UK Inhope) and removed.
  - 61 URLs were actioned.
  - These images have been downloaded from at least 476 actors.
  - Forum users discuss and are aware of the age of models.
- From the 5,788 preview images, 3,496 (60%) are classified as Non-Safe-For-View (NSFV).
  - Others include chats, screenshots or removed files.
Most images were taken from adult sites, but also from others like social networks, news or blogs.
Outline

1. Introduction
2. Analysis and provenance of images
3. Analysis of financial gains
4. Analysis of actors
5. Conclusions
Financial earnings are estimated from *Proof-of-Earning* images and the *Currency Exchange* board.

- Some actors share their alleged earnings
  - Advertise product or service, brag about their skills, etc.
- Images are posted to prove the claimed earnings
- We have manually analysed 2,067 images (1,868 proofs)
- Annotated info include:
  - Payment platform
  - Currency
  - Total amount
  - Number of transactions (payments received)
  - Language
Financial earnings are estimated from *Proof-of-Earning* images and the *Currency Exchange* board.
eWhoring is a low level fraud. A few portion actors have larger earnings.

- Average income per transaction is $42
- 661 actors reported earnings
  - Total of US$511k
  - 22 (3.3%) report more than $5k
- The more ‘proofs’ are posted, the higher earnings can be estimated
  - Though actors might be forging images
Amazon Gift Cards (934) and PayPal (795) are the main payments platforms to receive payments.
Measuring eWhoring
Analysis of actors

Outline

1. Introduction
2. Analysis and provenance of images
3. Analysis of financial gains
4. Analysis of actors
5. Conclusions
The analysis of actors is based on forum activity and social interactions

- The analysis focuses on hackforums.net
- SNA to measure popularity and influence of actors
  - Directed graph, nodes are actors and edges are public interactions
  - Eigenvector for influence
  - H-index, i-10, i-50 and i-100 indexes for popularity
- Forum activity for the analysis of interests and evolution
  - Activity before, during and after interacting with the eWhoring community
  - Interests are based on the category of the boards where users post
eWhoring is a large community

<table>
<thead>
<tr>
<th>#Posts</th>
<th>#Actors</th>
<th>Avg. posts</th>
<th>%ewhor.</th>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥ 1</td>
<td>72,982</td>
<td>8.8</td>
<td>23.3</td>
<td>165.3</td>
<td>474.2</td>
</tr>
<tr>
<td>≥ 10</td>
<td>13,014</td>
<td>37.6</td>
<td>22.8</td>
<td>142.7</td>
<td>449.7</td>
</tr>
<tr>
<td>≥ 50</td>
<td>2,146</td>
<td>126.9</td>
<td>26.0</td>
<td>133.8</td>
<td>293.8</td>
</tr>
<tr>
<td>≥ 100</td>
<td>815</td>
<td>222.4</td>
<td>29.1</td>
<td>132.8</td>
<td>210.1</td>
</tr>
<tr>
<td>≥ 200</td>
<td>263</td>
<td>402.3</td>
<td>34.9</td>
<td>153.6</td>
<td>165.7</td>
</tr>
<tr>
<td>≥ 500</td>
<td>46</td>
<td>930.8</td>
<td>40.6</td>
<td>157.4</td>
<td>157.8</td>
</tr>
<tr>
<td>≥ 1,000</td>
<td>13</td>
<td>1,566.8</td>
<td>37.5</td>
<td>412.6</td>
<td>137.3</td>
</tr>
</tbody>
</table>

Table 8: Number of actors, mean posts made, percentage made in eWhoring and mean days posting before and after eWhoring grouped by the number of posts made in eWhoring.
Previous interests for some key actors was on hacking/gaming

- Selection of 195 actors. Key actor if included in either one of these groups:
  - Offering TOP
  - Reporting earnings
  - Popular actors (H-Index)
  - Requesting currency exchange
  - Influencing (eigenvector)
- Some actors from various groups
Previous interests for some key actors was on hacking/gaming
Outline

1. Introduction
2. Analysis and provenance of images
3. Analysis of financial gains
4. Analysis of actors
5. Conclusions
So, why concerning about eWhoring?

- eWhoring is a popular activity in the underground world
  - Attractive for users looking for easy money-making methods
  - Until now it has been unexplored

- Is this illegal?
  - Fraud and scam
  - Copyright violation
  - Redistribution of images of another person
  - Distribution and possession of illegal images
So, why concerning about eWhoring?

- eWhoring is a popular activity in the underground world
  - Attractive for users looking for easy money-making methods
  - Until now it has been unexplored

- Is this illegal?
  - Fraud and scam
  - Copyright violation
  - Redistribution of images of another person
  - Distribution and possession of illegal images

- Definitely, it is *unhealthy* and *unethical*
  - What are our kids doing at their room?
  - Visiting porn sites and gathering material to create packs?
  - Going to chat rooms and scamming people by mimicking prostitutes?
  - Learning mechanisms to bypass PayPal anti-fraud detection?
So, why concerning about eWhoring?

- eWhoring is a popular activity in the underground world
  - Attractive for users looking for easy money-making methods
  - Until now it has been unexplored
- Is this illegal?
  - Fraud and scam
  - Copyright violation
  - Redistribution of images of another person
  - Distribution and possession of illegal images
- Definitely, it is *unhealthy* and *unethical*
  - What are our kids doing at their room?
  - Visiting porn sites and gathering material to create packs?
  - Going to chat rooms and scamming people by mimicking prostitutes?
  - Learning mechanisms to bypass PayPal anti-fraud detection?

spastran@inf.uc3m.es
@serpastrana

THANK YOU!

+ TineEye, IWF, Cambridge Cybercrime Centre
Measuring eWhoring

Sergio Pastrana\textsuperscript{1}, Alice Hutchings\textsuperscript{2},
Daniel Thomas\textsuperscript{3}, Juan Tapiador\textsuperscript{1}

\textsuperscript{1}Universidad Carlos III de Madrid
\textsuperscript{2}University of Cambridge
\textsuperscript{3}University of Strathclyde

19th Internet Measurement Conference (IMC)
October 23, 2019
Pornhub Finally Removes Girls Do Porn

After months of knowingly working with a porn production company that was accused of lying to women, Pornhub finally removes its official page from its platform.

By Samantha Cole and Emanuel Malberg
Oct 14 2019, 11:12pm  Share  Tweet

Feds hit GirlsDoPorn owners with criminal sex trafficking charges

"I was in a state of panic when I first found out," one alleged victim testified.

TIMOTHY B. LEE - 10/11/2019, 3:51 AM
BACKUP SLIDE: Tools and services

- Extraction of Thread Offering Packs (TOP)
  - Machine Learning and Heuristics for classification
  - NLP and thread statistics as features

- URL extraction and downloader
  - Regular expressions to get the URLs
  - Custom crawler for the different sites

- Filter out illegal images
  - PhotoDNA
  - Strict policy if image matches: report and remove

- Image classification
  - Yahoo Non-Safe-For-View (NSFW) classifier
  - Tesseract OCR

- Reverse search and classification
  - TinEye reverse image search
  - Three domain classifiers: McAfee, VirusTotal and OpenDNS
For analysis of financial activity, we also processed threads posted for Currency Exchange:

Table 7: Number of threads offering and asking for currencies by forum actors with more than 50 posts in eWhoring. ‘?’ means unclassified.
Table 5: Number of matches for the reverse image search. Seen Before means that the image was URL was online before the image was posted in the forum. Max and ratio refer to the maximum and average number of matches per image.