The Use of Random Sampling in Investigations Involving Child Abuse Material

Michael Wilkinson – Champlain College
Brian Jones – NSW Police Force
Syd Pleno – Australian Federal Police
Overview

• Putting things in context
  – NSW Police Force
  – Relevant NSW & Australian Laws

• Sampling & Statistics

• Implementation

• Benefits
About NSW

- Pop: 6.5 million
- Area: 309,432mi² (area of Texas: 268,600mi²)
- Police Force: ~20,000 (~17,000 sworn)
State Electronic Evidence Branch (SEEB)

• Created in 2003
• Funded in response to terrorism threat presented by 9/11 & Bali bombings
• Provide a response to “serious major crime”

• NSW Police also have:
  – Computer Crime Unit
  – Internet Child Protection Unit
SEEB Capabilities

- Computers (duh)
- Mobile phones
- GPS
- CCTV
- Embedded devices (ECM etc)
- Hardware level recovery
- Covert activities
Child Pornography 19%
Homicide 17%
Sex Offence - Assault 12%
Drug Detection 10%

Other Offences Against Person 2%
Miscellaneous 2%
Fire - Breach 2%
Assault 3%
Telecommunications 5%
Fraud 6%
Robbery 8%

Judicial 2%
Located Person 2%
Firearm Offences 1%
Other 7%

Child Pornography – investigation issues

- Data volumes
- High plea rate (~80%)
- ‘Criminality’ concepts – material quantity and severity
- Investigation risks,
  - OH&S
  - child at risk
  - prosecution time frames.
Australian Offences relating to CAM

• Federal
  – Telecommunications
  – Travel
  – Access overseas

• NSW
  – Possession
  – Disseminate
  – Produce
Definitions - Federal

(a) material that **depicts** a person, or a representation of a person, who:

(i) is, or appears to be, under **18** years of age;

(b) material that **describes** a person who:

(i) is, or is implied to be, under **18** years of age;
Definitions – NSW State

"child" means a person who is under the age of 16 years.

"child abuse material" means material that depicts or describes, in a way that reasonable persons would regard as being, in all the circumstances, offensive:

(a) a person who is, appears to be or is implied to be, a child as a victim of torture, cruelty or physical abuse, or

(b) (b) a person who is, appears to be or is implied to be, a child engaged in or apparently engaged in a sexual pose or sexual activity (whether or not in the presence of other persons),
Guilty

1 Picture
or
1 Movie
or
1 Story
Sentence

Determined by:
  • Severity
  • Quantity
<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No Sexual Activity</td>
</tr>
<tr>
<td></td>
<td>Depictions of Children with No Sexual Activity - Nudity, surreptitious</td>
</tr>
<tr>
<td></td>
<td>images showing underwear nakedness, sexually suggestive posing,</td>
</tr>
<tr>
<td></td>
<td>explicit emphasis on genital areas, solo urination</td>
</tr>
<tr>
<td>2</td>
<td>Child Non-Penetrative</td>
</tr>
<tr>
<td></td>
<td>Non-Penetrative Sexual Activity Between Children or Solo</td>
</tr>
<tr>
<td></td>
<td>Masturbation By A Child</td>
</tr>
<tr>
<td>3</td>
<td>Adult Non-Penetrate</td>
</tr>
<tr>
<td></td>
<td>Non-Penetrative Sexual Activity between Child(ren) and Adult(s).</td>
</tr>
<tr>
<td></td>
<td>Mutual masturbation and other non-penetrative sexual activity</td>
</tr>
<tr>
<td>4</td>
<td>Child/Adult Penetrate</td>
</tr>
<tr>
<td></td>
<td>Penetrative Sexual Activity Between Child(ren) or between Child(ren)</td>
</tr>
<tr>
<td></td>
<td>and Adult(s) - Including, but not limited to, intercourse, cunnilingus</td>
</tr>
<tr>
<td></td>
<td>and fellatio.</td>
</tr>
<tr>
<td>5</td>
<td>Sadism/Bestiality/ Child Abuse</td>
</tr>
<tr>
<td></td>
<td>Sadism, Bestiality or Humiliation (urination, defecation, vomit,</td>
</tr>
<tr>
<td></td>
<td>bondage etc.) or Child Abuse as per CCA 199</td>
</tr>
<tr>
<td>6</td>
<td>Animated or Virtual</td>
</tr>
<tr>
<td></td>
<td>Anime, cartoons, comics and drawings depicting children engaged in</td>
</tr>
<tr>
<td></td>
<td>sexual poses or activity.</td>
</tr>
</tbody>
</table>
Quantity

- 500,000+ porn collections common
- Movies take time to watch
- Each offending file must be categorised!
- Health risk to SEEB staff
- SEEB staff != reasonable people
  – Due to over exposure
Solution = Take Sample
Challenges

- Must be representative
- Must be reliable
- Must be able to convince NSW Police officers
- Must be able to convince the court...
Unbiased Sample

• Assign random numbers to each picture & movie
• Sort by random number
• Select $x$ files where $x$ gives 99% confidence that sample is representative
What about error?

- Influenced by % of CAM files found
- Tested 3 formulas
  - Adjusted Wald
  - Jefferys
  - Wilson
Testing Objective

- Does sampling work?
- Which error rate is most accurate?
Testing

• Picture populations between 5,488 and 520,610.

• The recommended sample or a selection of 300 (whichever comes first)

• Actual picture of interest counts between 3 and 47,000.

• Actual percentage of pictures of interest compared to population between 0.0057% and 50%
Test results

• It worked!
• Adjusted Wald gave best results
Implementation

• View files in random order
• As files selected display confidence & error rates
• When Confidence level & error rate acceptable generate report
• Implemented with enscript
Discovery Process

Exhibit “booted” into a controlled environment to facilitate forensic interrogation by room supervisor.

All valid allocated images, movies and documents are:
- Identified through signature analysis
- Counted,
- Randomised

Registry files exported automatically and a system/registry report created for SEEB.

A statistically representative sample is isolated for examination.

Sample is then examined by case officer for prohibited material against the CETS scale.

Report and statement generated and printed at conclusion of examination.

The report given to the investigator contains:
- Basic job, device, OS and user account information.
- Selected files with file information.
- Statistical analysis for each file type.

NO IMAGING!
Discovery room

© Michael Wilkinson
Please enter the details of the exhibit below

Client Exhibit Reference (e.g. D12345) / Item Number

Client Exhibit Location (e.g. Surry Hills Police Station)

SEEB Exhibit Reference (OPTIONAL)

☐ Exhibit contains other items.

Description of Exhibit

Exhibit model number

Exhibit serial number
Electronic Evidence Discovery Report
SEEB Task Number E01/0001

<table>
<thead>
<tr>
<th>Investigator Name</th>
<th>Detective Senior Constable John Smith</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accused Name</td>
<td>John DOE</td>
</tr>
<tr>
<td>Accused Date of Birth</td>
<td>12 March 1980</td>
</tr>
</tbody>
</table>

WARNING: This report contains offensive material. Disclosure is subject to the Sensitive Evidence Provisions under Chapter 8, Part 2A of the Criminal Procedure Act 1986
2 Exhibit Item X1234567890

2.1 Description

<table>
<thead>
<tr>
<th>Exhibit Reference</th>
<th>X1234567890</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exhibit Location</td>
<td>Surrey Hills</td>
</tr>
<tr>
<td>SEEB EXHIBIT Reference</td>
<td>A silver HP EliteBook laptop computer</td>
</tr>
<tr>
<td>Date/Time of examination</td>
<td>03/07/11 10:13:45</td>
</tr>
</tbody>
</table>

2.2 Device Summary

2.2.1 RS005

<table>
<thead>
<tr>
<th>Name</th>
<th>RS005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td></td>
</tr>
<tr>
<td>Serial Number</td>
<td>664025561665</td>
</tr>
<tr>
<td>Drive Type</td>
<td>DVD/CD-ROM</td>
</tr>
<tr>
<td>Total Size</td>
<td>19.1GB</td>
</tr>
</tbody>
</table>

2.2.2.1 C

| Volume Label | C |
| Volume Name | |
| File System | NTFS |
| Total Capacity | 2063490080, 19.1GB |
| Unused Space | 14GB |

2.3 Operating System No 1

<table>
<thead>
<tr>
<th>Operating System Name</th>
<th>Microsoft Windows XP Service Pack 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Path</td>
<td>RS006C</td>
</tr>
<tr>
<td>TimeOut Information</td>
<td>(GMT +10:00) Canberra, Melbourne, Sydney</td>
</tr>
<tr>
<td>Registered Owner</td>
<td>John Smith</td>
</tr>
<tr>
<td>Registered Organisation</td>
<td></td>
</tr>
<tr>
<td>Installed Date</td>
<td>09/04/09 11:19:38</td>
</tr>
</tbody>
</table>

2.3.1 User Account 1

<table>
<thead>
<tr>
<th>Account Name</th>
<th>Administrator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Username</td>
<td></td>
</tr>
<tr>
<td>Last Logon</td>
<td>21/05/09 09:34:00</td>
</tr>
<tr>
<td>Number of Logons</td>
<td>8</td>
</tr>
<tr>
<td>Password present</td>
<td>Yes</td>
</tr>
<tr>
<td>Screen saver Password</td>
<td>disabled</td>
</tr>
</tbody>
</table>

2.3.2 User Account 2

<table>
<thead>
<tr>
<th>Account Name</th>
<th>Guest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Username</td>
<td></td>
</tr>
<tr>
<td>Last Logon</td>
<td></td>
</tr>
<tr>
<td>Number of Logons</td>
<td>0</td>
</tr>
<tr>
<td>Password present</td>
<td>No</td>
</tr>
<tr>
<td>Screen saver Password</td>
<td>disabled</td>
</tr>
</tbody>
</table>
2.4.2 CETS Level 1: Depictions of children with no sexual activity

File Index: 2
File Name: Bliss.bmp
File Path: C:\Program Files\Microsoft Office\Office14\PART\PU\800\COR\F0\D\03936\W
File Created: 08/04/09 11:24 AM
File Modified: 08/04/09 11:24 AM
MD5 Hash: 210E260B7F25AF3453505350E500051C46B49
CETS Scale: 1

File Index: 3
File Name: F000336_WM
File Path: C:\Program Files\Microsoft Office\Office14\PART\PU\800\COR\F0\D\03936\W
File Created: 09/03/08 16:57 PM
File Modified: 09/03/08 16:57 PM
MD5 Hash: BBF4FC8018EF95239B58135D8561BC04
CETS Scale: 1

File Index: 4
File Name: J0009167.JPG
File Path: C:\Program Files\Microsoft Office\Office14\PART\PU\800\COR\F0\D\03936\W
File Created: 06/02/08 02:00 AM
File Modified: 06/02/08 02:00 AM
MD5 Hash: EC7C2AF979E21F93C5DE59EFD44E2F04
CETS Scale: 1

File Index: 5
File Name: J0228073_WM
File Path: C:\Program Files\Microsoft Office\Office14\PART\PU\800\COR\F0\D\03936\W
File Created: 09/07/02 17:38 PM
File Modified: 09/07/02 17:38 PM
MD5 Hash: 0C2334C7725FD1F89A039CABD856241A
CETS Scale: 1
## 2.5.2 CETS Level 1 - Depictions of children with no sexual activity

<table>
<thead>
<tr>
<th>File Index</th>
<th>14</th>
<th>Description:</th>
</tr>
</thead>
<tbody>
<tr>
<td>File Name</td>
<td>Sample.mov</td>
<td>A description of a video viewed by the case officer</td>
</tr>
<tr>
<td>File Path</td>
<td>X1234567890\Program Files\QuickTime\Sample.mov</td>
<td></td>
</tr>
<tr>
<td>Created</td>
<td>05/01/09 17:18:48</td>
<td></td>
</tr>
<tr>
<td>Modified</td>
<td>05/01/09 17:18:48</td>
<td></td>
</tr>
<tr>
<td>MD5 Hash</td>
<td>0E242AB2D25193887BFFDF12813E6E67</td>
<td></td>
</tr>
<tr>
<td>CETS Scale</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
3 Statistical Analysis

The files of interest identified as child pornography in this report are the result of a randomised sampling of a filtered set of files found on each exhibit item.

The statistical breakdown of each exhibit item compares the number of files of interest for each CETS level with the total number of files examined for each file type.

3.1 Exhibit Item X1234567890

3.1.1 Picture files of interest

The percentage values in the above chart have been rounded to the nearest whole number. Refer to the following table for an accurate quantity value.

<table>
<thead>
<tr>
<th>Description</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid Picture files on this exhibit</td>
<td>30789</td>
</tr>
<tr>
<td>Picture files in sample examined</td>
<td>7548</td>
</tr>
<tr>
<td>CETS Level 1 - Depictions of children with no sexual activity</td>
<td>85</td>
</tr>
<tr>
<td>CETS Level 2 - Non-penetrative sexual activity between children or solo masturbation by a child</td>
<td>39</td>
</tr>
<tr>
<td>CETS Level 3 - Non-penetrative sexual activity involving children and adult(s)</td>
<td>14</td>
</tr>
<tr>
<td>CETS Level 4 - Penetrative sexual activity involving children, or both children and adult(s)</td>
<td>31</td>
</tr>
<tr>
<td>CETS Level 5 - Sadism or bestiality</td>
<td>58</td>
</tr>
<tr>
<td>Picture files of interest selected from sample</td>
<td>227</td>
</tr>
<tr>
<td>Percentage of Picture files of interest in sample</td>
<td>3.01%</td>
</tr>
<tr>
<td>Estimated number of valid Picture files of interest on this exhibit</td>
<td>730 to 1081</td>
</tr>
</tbody>
</table>
Limitations

• Only suitable for estimating quantities

• Additional victims and offenders may not be identified
Results

• Backlog down from 3 months to 24 hours
• Reduced exposure of SEEB staff to CAM
• Favourable feedback from court
• Criminal Procedures Act amended to accept sampling in CAM cases
Questions?

Contact: Mike Wilkinson

wilkinson@champlain.edu
@mikewilko

www.writeblocked.org