Quantitative Plausibility of the Trojan Horse Defence against Possession of Child Pornography

Richard E Overill & Jantje A M Silomon

Department of Informatics,
King's College London

K P Chow & Y W Law

Department of Computer Science,
University of Hong Kong

Synopsis

- Trojan Horse Defence
- Possession of Child Pornography
- Digital Forensic Sub-hypotheses
- Evidential Traces Recovered
- Enhanced Complexity Model
- Trojan Horse Model
- Complexities & Posterior Odds
- Conclusions & Further Work

Trojan Horse Defence

- First reported use in the UK October 2003
  (Aaron Caffrey, 19, Port of Houston hack)
- It concedes that the offence was committed, but contends that it was not by the defendant
  (Some Other Dude Did It - SODDI)
- In the absence of other evidence (*e.g.* DNA, fingerprint) tying defendant to crime scene, it requires the prosecution to prove a negative – that there was no Trojan Horse in operation at the material time

Possession of Child Pornography

- Trojan Horse Defence is highly successful globally in countering prosecutions of various e-crimes, including possession of child pornography (CP)
- (HK) law enforcement generally requires at least **five** items of digital CP before bringing charges

Digital Forensic Sub-hypotheses (Prosecution)

- Downloading of CP has been performed
  - three alternative possibilities: browser, email, peer-to-peer (P2P)
  - this study models browser download
- Copying of CP has been performed
  - two alternative possibilities: USB and CD/DVD
  - this study models USB device
- Viewing of CP has been performed

Evidential Traces Recovered (I)

- CP (image/video) on computer
- Internet history / cache from downloading
- Credit card payment to CP website
- Metadata on computer matched CP website
- USB device was plugged into the computer
- CP on computer matched that on USB device

Evidential Traces Recovered (II)

- Modified timestamp predates created timestamp of CP
- Image / video viewing tools on computer
- CP displayed by image / video viewing tools
- Access timestamp postdates created timestamp of CP

Enhanced Complexity Model (Hypotheses & Model)

Hypotheses:

- the more complex a process is, the less likely it is to happen without user awareness
- effort to implement and integrate TH software components must be taken into account

Model process complexity using:

- computational complexity (CC)
- GOMS Keyboard Level Model (KLM)
- Halstead's Effort (E) metric

Halstead's Effort (E) metric

- $\boldsymbol{n_1}$ – number of distinct operators
- $\boldsymbol{n_2}$ – number of distinct operands
- $\boldsymbol{N_1}$ – total number of operators
- $\boldsymbol{N_2}$ – total number of operands

- Program vocabulary $\boldsymbol{n} = n_1 + n_2$
- Program length $\boldsymbol{N} = N_1 + N_2$
- Program volume $\boldsymbol{V} = N \times log_2 n$
- Programming difficulty $\boldsymbol{D} = (n_1 \times N_2)/(2 \times n_2)$
- Programming effort $\boldsymbol{E} = D \times V$

Enhanced Complexity Model (Processes)

- For process *i*:

$$p_i \propto [\, CC_i + KLM(CC)_i + E_i + KLM(E)_i \,]^{-1}$$

- For two mutually exclusive processes *i* and *j*, the 'posterior odds' of process *i* over process *j*:

$$O(i:j) = \Pr(H_i \mid E) \,/\, \Pr(H_j \mid E) = p_i / p_j$$

Trojan Horse Model

- Simplest possible system that produces all of the requisite evidential traces and no others: an ***electronic, random framing attack***
- Lower bound on complexity implies upper bound on plausibility of Trojan Horse defence
- Consists of:
  - Dropper
  - Installer / Uninstaller
  - Payload (inc. keylogger, string search algorithm)

Complexities & Posterior Odds

	OCM		ECM	
	Non-Trojan	Trojan	Non-Trojan	Trojan
CC	11,569,216	19,232,355	11,569,216	19,232,355
KLM(CC)	1,730	—	1,730	—
E	—	—	—	13,850,047
KLM(E)	—	—	—	1,381,959
Total	**11,570,946**	**19,232,355**	**11,570,946**	**34,464,361**

	OCM	ECM
Unprotected computer	1.367	2.979
98% protected computer	117.4	197.9

Conclusions

- Potential significance for both prosecution and defence sides when assessing their own worst case scenario and their opponents' best case scenario
- For an **unprotected** computer, posterior odds do not favour a successful criminal prosecution
- For a **protected** computer, posterior odds strongly favour a successful criminal prosecution
- Off-the-shelf Trojan models (OCM) are not much harder to prosecute than bespoke ones (ECM)

Further Work

- **DOCM** – ‘de-parameterised’ OCM, independent of file size N
- **BOCM** – ‘buffered’ OCM, with file buffer of length N_B, filled/flushed with N_F operations, so a block-copy requires $N_F \lceil N/N_B \rceil$ operations
- Modelling degree of motivation, capability/skill level, opportunity/lack of deterrence
- Expand model beyond current computer platform (PC with Win-XP & IE browser)

Acknowledgement

- Testwell for the grant of an evaluation licence for their *CMT++ Complexity Measures Tool for C/C++* to calculate the Halstead E metric
- US ONR MINERVA programme “Strategy and the Network Society” research grant
- UK EPSRC Overseas Travel Grant

Thank you!

Questions? Comments?

Richard E Overill
{richard.overill@kcl.ac.uk}