The first 10 years of the Trojan Horse defence

Stephen Bowles and Julio Hernandez-Castro, University of Kent

Apprehended criminals throughout history have always attempted to put the blame on someone else, a strategy popularly known as a SODDI defence (Some Other Dude Did It). When this defence is used, the act of the crime (actus reus) and the guilty mind (mens rea) is blamed on another party. A Trojan Horse Defence (THD) is a type of modern SODDI defence, where the mens rea and actus reus are blamed on a piece of software, known as a trojan.

A trojan is a type of malicious software (globally known as malware) that is either packaged along with a useful piece of software or pretends to be a piece of useful software itself. Once a trojan is activated, which usually goes unnoticed by the user, it releases a payload such as a virus or a backdoor that may allow a remote user to gain access to the system. A popular means of infection is for online downloads to be packaged with a trojan, resulting in an infection when the user runs the downloaded application. The more complex trojans – those that capture typed keyboard characters or permit a remote user access to the system – can be used to create a credible legal defence. This is in contrast to simple trojans, which act as a nuisance by deleting files and changing user settings.

The THD is believed to have been first used in 2003, when it appeared in the Charles Schofield case. After a literature review of relevant publications over the past decade, this article discusses court cases where the THD has been used.

The prosecutor’s view

In 2004 Susan Brenner and her colleagues wrote a paper that primarily concentrates on the view of the prosecutor. As it was major news at the time, the paper focuses on the Caffrey case, regularly linking points to it. The paper reservedly states that the defence could be “empirically valid”. However, especially with recent advancements in malware, it should be stressed that it is entirely possible for a trojan to be the cause of a crime. The paper provides an introduction to four 2003 court cases and then continues on to a discussion of the legal and technical issues regarding the THD.

"The paper suggests that a thorough examination is only needed if malware is found – however it is not that simple. A thorough investigation is always required, partly because a quick analysis might miss things, but also because malware may still be responsible."

The forensics and technical side

A 2003 paper added a thorough discussion about digital forensic techniques and provides intricate details on how a forensic investigation should take place on a Windows- or Linux-based system. Although perhaps dated, many of the techniques and tools mentioned are still valid and used today. The paper primarily focuses on data recovery; however it also discusses viewing logs and tracking the activities of a hacker. The paper
makes a significant point that, in some cases (which we report on later), investigators failed to perform an important step — which is to make a copy of the hard drive and its contents and then proceed to carry out the investigation on that copy. This is important so the evidence is not altered and the investigation can be carried out by other parties.

“It is considerably easier today for a malicious party to modify a malicious program that, when hashed, matches a safe value or a hash of a lesser malware”

A reader should be aware of the current significance that recent developments bring to technology. For example, where the paper suggests using cryptographic checksums MD5 and SHA-1 to check a file system, we suggest that, in light of the multiple security vulnerabilities recently discovered, it would be better to use SHA-1 instead of MD5, or even SHA-2 or SHA-3. Although it has been recently stated that MD5 and SHA-1 are still valid for digital forensics, it is considerably easier today for a malicious party to modify a malicious program that, when hashed, matches a safe value or a hash of a lesser malware. Some authors have highlighted how a hex editor can be used to change the signature of malware, making it, in some cases, no longer recognisable by protection software.

The two-part 2005 paper (Haagman and Ghavalas) discusses techniques to develop trojans and scenarios of trojans compromising a system. The paper also discusses volatile and network evidence, areas barely touched by the previously mentioned ‘Issues in Computer Forensics’ paper. Evidence that can be collected from a system’s memory and evidence from the network, such as activity logs, could be invaluable in a case.

Finally, another recent 2012 paper (Sepec) provides the reader with an insight into the difficulties that forensic examiners face, and the steps they should take. The paper details trojans, their behaviour, how they work and the THD. Although it refers mostly to Slovenian law, the contents remain significant and relevant for other legal systems. In regards to THD cases, the paper mentions four key considerations and although they are not the only ones to take into account, they are nonetheless quite important. In brief they are:

1) Investigators should use several up-to-date investigative programs.
2) Digital evidence is not the only evidence to be considered in cybercrime offences.
3) Expert testimonies on the mental state of the accused is relevant.
4) Consider that the accused might have planted a trojan on their computer system as a defence tactic.

Statistics

A 2004 paper (Carney and Rogers) details an attempted stepwise discriminant analysis of four scenarios, all which resulted in five illicit images of children being created on the system. The goal was to see, “whether an investigator could determine if images were downloaded intentionally or without the owner’s knowledge based on characteristics located in the operating and file system”. The paper lists seven characteristics of an operating and file system, which it proposes an investigator should look at when performing an investigation. The results from this analysis proved interesting and showed that it might be possible to create an accurate statistical model, although much more work is required in this area before that becomes a reality.

Other authors have looked into plausibility metrics and posterior odds, where they discuss an Operational Complexity Model and an Enhanced Complexity Model (ECM). In a personal communication with one of the authors, Richard Overill provided the following summary: “On the basis of the ECM we are suggesting that the Trojan Horse Defence is a rather implausible defence, provided there is an up-to-date anti-virus/malware scanner installed. When compared with other accepted forensic techniques (fingerprints), the Enhanced Complexity Model shows a lower odds ratio against the THD, which implies that this kind of complexity based metric is not as clear-cut as a biometric like fingerprinting or DNA. Also, of course, the biometric identifies an individual whereas the ECM on its own does not; it simply says that it is 198 times more likely that an individual did it rather than a TH.”

“A 2010 paper suggests that an...”

However, we believe that this approach, while still having some merit, is no longer true, because malware is becoming increasingly equipped with tools to disable or hide from anti-virus software. Although more research is required here, we believe that anti-virus software provides very little protection, other than to avoid questions in court as to why there is no anti-virus software installed on the relevant computer.

Defending against a trojan

The usual defence against a trojan includes anti-malware, anti-virus and firewalls. However, some authors have recently suggested other measures to improve defences against a trojan.

A 2010 paper suggests that an ‘Education, Enforcement and Engineering’ or ‘Triple-E’ approach could be taken. For education, the paper suggests changing the perception of people towards hackers, teaching how to use a computer appropriately, and teaching the public safe habits when using a computer system and the Internet.
relating enforcement to procedures for fact finding and investigations, propose the MDFA (Multi-faceted Digital Forensics Analysis) strategy. MDFA is a range of procedures for collecting and linking characteristics from the four phases of a crime that the authors list as evidence, scene, victim and suspect. The M-N model and Ideal Log are discussed for the engineering part of Triple-E, where both focus on logs (such as audit logs) and the task of building a series of events by synchronising events that could be from different time zones, which can happen because the time on a client may not be the time on a server or the network devices in-between.

“It is important to note that, as mobile technology is becoming dominant, the next THD case could involve a smartphone”

Other papers suggest an intrusion detection system to work alongside other software, which will attempt to detect potentially malicious programs. Although these papers take different routes, the general principle is the same – having a system in place that monitors and tracks the execution of programs and the resources they attempt to use. By placing programs into different levels, where one level has higher trust or permissions than another, the system can alert a user if a program is attempting to access something it should not. The concepts discussed here are similar to User Access Control (UAC), a system in use by some of the more recent Microsoft Windows operating systems.

Mobile

When considering the THD, it is too easy to think only of the classical scenario – a home, possibly family-shared, desktop PC. However, it is important to note that, as mobile technology is becoming dominant, the next THD case could involve a smartphone. Although some papers have started to discuss the mobile malware topic, awareness needs to be raised regarding smartphone security.

Court cases

Table 1, along with the timeline shown in Figure 1, provides what we believe to be the currently most detailed and comprehensive list of court cases where the THD (by this or a similar name) has been used. The date within the table corresponds to the conviction or acquittal date. Where it was difficult to accurately gather the date, an educated guess is used, based on time of articles published, news stories or the appeals documents. This section will discuss some of the most notable cases listed in Table 1.

Aaron Caffrey

The first heavily publicised THD case was the 2003 Aaron Caffrey case. The Port of Houston, in the US, suffered a denial of service (DoS) attack, which was traced back to Caffrey’s machine. This case is interesting because no evidence of a trojan was found. Furthermore, Caffrey was a member of a hacking group and tools that could have carried out the attack were found on Caffrey’s system. During the case, Caffrey made false claims that were not disputed, such as an anti-virus cannot scan every file on a system, which is blatantly false. Even with a possible motive identified, Caffrey was acquitted on the argument that he was being framed by another hacker by means of a trojan.

Eugene Pitts

The 2003 Eugene Pitts case is different from any others, as the type of crime was very different. Pitts was accused of income tax evasion in the US and even though he had a history of troubles, where he was accused in previous years of under reporting income, he was acquitted under the claim that a virus was responsible for modifying his files. It was noted, however, that his customers’ tax returns, which were on the same system, were surprisingly not affected by the virus at all.

Michael Aaron O’Keefe

This case involves a defendant who apparently created two websites, modelquest and hctweens, to catch paedophiles. Michael Aaron O’Keefe was arrested and accused of advertising, receiving, and possessing indecent images of children. The websites that O’Keefe had apparently created to catch paedophiles were hosting child pornography images, something that O’Keefe said a virus must have done. Logs were also found that show O’Keefe was posing as a young girl, apparently in an attempt to lure
<table>
<thead>
<tr>
<th>Date/Place</th>
<th>Victim/Crime/Evidence</th>
<th>Defendant</th>
<th>Result</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003 April, UK</td>
<td>Possession of 14 indecent pictures of children</td>
<td>Karl Schofield</td>
<td>Acquitted: “unnamed trojan did it”</td>
<td>An analyst found and testified that a trojan was responsible.</td>
</tr>
<tr>
<td>2003 July, UK</td>
<td>Possession of 172 indecent pictures of children</td>
<td>Julian Green</td>
<td>Acquitted</td>
<td>Martin Gibbs, working at Vogon International, found 11 trojans on Green’s system.</td>
</tr>
<tr>
<td>2003 August, US</td>
<td>Evasion of tax/income tax evasion</td>
<td>Eugene Pitts</td>
<td>Acquitted</td>
<td>It was noticed that Pitts had previous a history of issues; also it was odd that the virus did not affect any customers’ tax returns.</td>
</tr>
<tr>
<td>2003 September, US</td>
<td>Possession of 2,000+ indecent images of children</td>
<td>Brian Bass</td>
<td>Convicted on five counts</td>
<td>Bass used history remover software, claimed to have a “morbid curiosity” with child porn, but blames a virus for downloading the images.</td>
</tr>
<tr>
<td>2003 October, UK</td>
<td>Denial of service attack on the Port of Houston, US</td>
<td>Aaron Caffrey</td>
<td>Acquitted: “trojan did it”</td>
<td>Caffrey was a known hacker, possessed hacking tools, and no evidence of a trojan was found.</td>
</tr>
<tr>
<td>2004 November, US</td>
<td>Buying, advertising and owning child pornography</td>
<td>Michael Aaron O’Keefe</td>
<td>Convicted on four counts</td>
<td>Claimed to have created websites to catch paedophiles; blamed the indecent imagery being hosted on a virus.</td>
</tr>
<tr>
<td>*2005 January, UK</td>
<td>Possession of 1,793 indecent images of children</td>
<td>Mark Craney</td>
<td>Convicted on 16 counts</td>
<td>Claimed a virus was to blame.</td>
</tr>
<tr>
<td>*2005 July, US</td>
<td>Possession of 320 indecent images and videos of children</td>
<td>Michael Shawn McCourt</td>
<td>Convicted on two counts</td>
<td>Claimed a hacker put the images on the system using peer-to-peer programs.</td>
</tr>
<tr>
<td>2006 September, UK</td>
<td>Possession of 3,000 pictures of child pornography</td>
<td>Julian Spencer</td>
<td>Guilty plea</td>
<td>When arrested, Spencer claimed a virus was to blame.</td>
</tr>
<tr>
<td>2006 November, US</td>
<td>Possession of nine indecent images of children</td>
<td>Matthew Bandy</td>
<td>Plea-bargain to avoid going to court</td>
<td>Over 200 viruses and malware were found on the system belonging to the 16 year-old.</td>
</tr>
<tr>
<td>2006 December, US</td>
<td>Possession of child pornography</td>
<td>Donald R Miller</td>
<td>Convicted</td>
<td>Unsuccessful claim that a virus downloaded the images, which got mixed in with other, legal pornography.</td>
</tr>
<tr>
<td>2007 July, US</td>
<td>Possessing and transporting indecent images of children</td>
<td>Gregory James Shiver</td>
<td>Convicted</td>
<td>During the case, a brief mention that ‘pop-ups’ or a virus had placed the images on the computer was made.</td>
</tr>
<tr>
<td>2007 July, UK</td>
<td>33,000+ indecent images and 1,000+ videos of children</td>
<td>Craig Geddes</td>
<td>Convicted on possession of child pornography</td>
<td>A detective in the case said she had never heard of a virus that could install child porn, even though defendants have been acquitted before on this reason.</td>
</tr>
<tr>
<td>*2007 August, South Africa</td>
<td>1,159 indecent images of children</td>
<td>Mark Rawlinson</td>
<td>Convicted on 330+ counts</td>
<td>Believed to be the first use of the THD in South Africa.</td>
</tr>
<tr>
<td>2008 February, US</td>
<td>Two counts of possession and one count of receiving child pornography</td>
<td>Ronald Vaughn</td>
<td>Convicted on all three counts</td>
<td>A post-trial attempt at the THD was made, but failed, as evidence was found on multiple systems with the only common user being Vaughn.</td>
</tr>
<tr>
<td>*2008 June, US</td>
<td>Government work laptop contained indecent pictures of children</td>
<td>Michael Fiola</td>
<td>Acquitted after investigation proved innocence</td>
<td>Forensic examiner Tami Loehrs found the system was infected months before it was given to Fiola, by the state. Malware visited 40 child porn websites a minute.</td>
</tr>
<tr>
<td>2008 November, US</td>
<td>Receiving and possession of child pornography</td>
<td>Nathaniel Solon</td>
<td>Convicted after initial guilty plea was withdrawn</td>
<td>Apparently an expert hired by Solon found evidence that could prove his innocence. There are claims that the judge behaved oddly, such as walking out without reason for a trivial matter.</td>
</tr>
</tbody>
</table>
paedophiles.23 Within this case, evidence of malware that allowed a remote user to gain access to the defendant’s system was found. As there was no evidence that O’Keefe ever made any effort to inform or help law enforcement of his actions or motives, the claim that he was trying to catch paedophiles was not believed.

Julie Amero

Julie Amero’s case is an example of the serious miscarriages of justice that have occurred in some THD cases. Here, the simplest forensic techniques were not used, the computer experts used had questionable knowledge and made false claims, and events were blown out of proportion by parents and prosecutors, as it is highly questionable if any damage was done to the pupils who saw the pornographic pop-ups.24,25

During the case, Amero’s system was not scanned for viruses or malware, and for what little investigation did happen, it did not follow best practices, as the investigator worked off the hard drive rather than making a copy.26 This case helps show why parties involved must be careful, as the stress of the case has been blamed for the miscarriage of Amero’s unborn child.27

Matthew Bandy

The 2006 Matthew Bandy case is another clear miscarriage of justice. Detectives in Bandy’s case did not seem aware of the steps that should be taken, which included requesting a digital forensic analysis, and as a result multiple mistakes were made.28 After being named a paedophile, a digital forensic investigation found that the anti-virus software was disabled, there was no firewall, and there were a myriad of infections and running malware, some of which had the capabilities to place pornography on the system.29 This case ended in a plea-bargain, but had it gone to a jury, the result would have been again unpredictable, as we have seen in other cases with the jury, even when the evidence is nearly identical.

Craig Geddes

Craig Geddes’ 2007 case is very similar to other THD cases where the defendant was acquitted, although in this particular case Geddes was convicted. Geddes asked, “Can a virus do that?”, the answer to which is yes. But during the case the opposite was stated.30 Detective Constable June McKay, of Strathclyde Police computer crime unit, said that they had never heard of a virus that could place child porn on someone’s system.31 This is a worrying claim by a computer crime detective, especially several years after similar cases. Other dubious claims include incorrect statements regarding changing IP address. During the case, no evidence was shown of any digital forensic investigation, making Geddes’ innocence conceivable.

Mark Rawlinson

Although information on Mark Rawlinson’s case is scarce, it is worth mentioning as it is believed to be the first use of the THD in South Africa.32 Cases outside of the UK and US are not as well documented, but the THD can happen and does get used in other countries, of which Rawlinson’s case is an example. Rawlinson’s THD claim failed and he was convicted of possessing over a thousand illicit images of children on his system.

Michael Fiola

Michael Fiola’s 2008 case stands out by its remarkable conclusion. Fiola was arrested and lost his job, as indecent images of children were found on his laptop. Unlike other cases, such as Amero’s, a digital forensic investigation was conducted by a qualified forensic examiner. The other major difference with other cases is that the laptop in question was a government issued computer, from the Department of Industrial Accidents in Massachusetts, US, where Fiola was working at the time.

The results from the forensic investigation found that the laptop was infected with malware months before it was given to Fiola. It was also found that the malware in question was visiting 40 child porn websites a minute. Although Fiola was acquitted with this evidence, the department never offered an apology or his job back.33,34,35

Nathaniel Solon

Nathaniel Solon’s 2008 case seems to be another miscarriage of justice example.36 Solon’s case initially followed a
path similar to other cases, which led to acquittals. The same digital forensic investigator from Amero’s case was involved in Solon’s, where some evidence of a virus that could have downloaded the pornography was found but no evidence of the material being viewed by Solon was found. The findings were seemingly ignored and as a result Solon was convicted only because of the material found on his system. A later appeal claimed that the behaviour of the judge swayed the jury, who disregarded the investigator’s findings and was unhappy with the fees.

Recent developments

With the huge increase in cyber-criminals using ransomware, the future looks grim. Ransomware is a type of malware, often the payload to a trojan, that is used by cyber-criminals to hold a user to ransom.

In 2013, Jay Riley’s system was infected with ransomware that pretended to be from the FBI and asked Riley to pay a fine. After Riley asked the police if there was any warrants out for him and volunteered for a search, he was arrested, because indecent images of children were found on his system. As well as possession, Riley was also charged on “1 count of indecent liberties with a minor”, as one of the images was of a 13-year-old girl from Minnesota, roughly 1,000 miles away from Riley in Woodbridge, Virginia, US. It is possible that Riley travelled that distance, but it is also likely that the malware planted those images.

In 2014, Marcel Datcu, a 36-year-old man from Romania, hanged himself and his four-year-old son in his home after a piece of ransomware infected his computer. A note left for his wife stated that a warning appeared on Datcu’s computer demanding a payment of 70,000 lei. This incident shows that not enough of the general public are aware of ransomware and the fake messages apparently coming from authorities. It also shows how serious the effect of ransomware can be on people’s lives.

Conclusions

This article has discussed the THD, the cases where it has been used, and related published material over the past decade, since its first use in 2003. By compiling an exhaustive list of cases, this work has shown that there are many occasions where serious miscarriages of justice have occurred. There have also been cases where clear and obvious mistakes have been made, either in the forensic investigation (or lack thereof) or from incorrect evidence given by incompetent experts.

As with many SODDI cases, the THD brings new challenges to the table, some that still need to be addressed by the forensic community. There should never be a worry that an innocent person has been put into jail or branded a criminal. Hopefully lessons would be learned from mistakes from the previous decade, so cases in the following decade will not see the same errors made.

Future work

Forensic techniques: It is possible that a criminal might successfully use the THD in their favour. To help combat this, forensic techniques need to improve and evolve. Not only this, but investigators need to know how they can investigate such a case. As a future work, a list of techniques and methods could be compiled, which can be followed by an investigator in a THD case. This work may help highlight additional steps, methods or techniques that could be used, that have not been considered before.

Court cases, match-and-learn, and data mining: Table 1 contains currently the most comprehensive list of THD cases; however, it is likely there are more that are currently not listed there. By researching cases, perhaps using data mining tools, it might be possible to not only add to that table but also create a set of match-and-learn profiles, which might predict the outcome of a THD case. As we have seen, this may not be easy, as often the defining and unpredictable factor in a THD case is the prosecution, jury or judge, rather than the evidence found.

About the authors

Stephen Bowles is a recent graduate of the University of Kent with a first class honours BSc. He is keen to progress in the computer security field, either in industry as a practitioner or in research to pursue a PhD.

Dr Julio Hernandez-Castro is a computer security lecturer at The School of Computing, University of Kent, UK. He was previously senior lecturer at Portsmouth and associate professor at Carlos III University in Madrid, Spain. His interests are cyber-security, cybercrime, steganography and steganalysis, malware and securing the Internet of Things.

References


About the authors

Stephen Bowles is a recent graduate of the University of Kent with a first class honours BSc. He is keen to progress in the computer security field, either in industry as a practitioner or in research to pursue a PhD.

Dr Julio Hernandez-Castro is a computer security lecturer at The School of Computing, University of Kent, UK. He was previously senior lecturer at Portsmouth and associate professor at Carlos III University in Madrid, Spain. His interests are cyber-security, cybercrime, steganography and steganalysis, malware and securing the Internet of Things.

References


15. R Di Pietro; F Lombardi; S Rossicone. ‘Modelling mobile resource security’.


22. Former Atlanta teacher sentenced to 17 years prison following a child pornography conviction’. 2005.


Resources

- ‘Computer virus causing some to be framed for child porn’. Naples News.
- FS Monterosso. ‘Protecting the children: challenges that result in, and consequences resulting from, inconsistent prosecution of child pornography cases in a technical world’. 2010.
The dangers in our trail of digital breadcrumbs

Tracey Stretton and Luke Aaron, Kroll Ontrack

What do Apple, eBay, Snapchat, Target and LinkedIn have in common? All of you with accounts at any of the aforementioned companies should know the answer to the question, as you will have likely been asked to change your password in the recent past, or at least read about it in the news. With varying degrees of exposure (in some cases quite literally) the users of these services have all suffered some form of data breach in recent history.

In a single attack on the US retailing giant Target, the credit and debit card details of an estimated 40 million people were stolen and made available on the black market to cyber-criminals able to exploit stolen data for financial gain. You cannot hear these stories without wondering ‘what information am I leaving online that people may access?’.

Need to communicate

The truth is we all leave a huge amount of personal data online. We worry about government surveillance and snooping yet ‘trust’ a corporate giant such as Facebook with personal information that is very sensitive – every private message you’ve written, every photo you’ve liked and every video you’ve watched. In fact, we go further – we tell them our relationship history, our thoughts, feelings, aspirations, worries and of course all of this information plus every disparaging message you may have written about your boss is logged, stored and made available to others for a price. The human need to communicate appears to trump concerns about privacy.

“We are giving away so much data on social media sites now that good old-fashioned data theft is perhaps no longer the major risk when it comes to data loss”

And as we can see from the above list, just because a company is big and conceivably has infinite resources available, certainly doesn’t mean your data is safe with it. Until last year Instagram was able to share your private information with Facebook and Facebook was able to sell information to advertisers without seeking your permission or offering compensation. That has apparently not differed following changes to Instagram’s privacy policy.

Of course, we haven’t even mentioned yet the risks attached to social media accounts that you no longer access (hand’s up – who doesn’t even know if they still have a MySpace or Bebo account and wouldn’t know how to login to it if they did?) And there are new chat tools, location-based dating apps and apps like Whisper that allow you to send messages anonymously, to contend with as well. According to news reports, Whisper was recently accused of tracking the approximate location of some of its users who opted out of geolocation services even though its privacy policy stated that access to location-based information was “purely voluntary”.

The bottom line is that we are giving away so much data on social media sites now that good old-fashioned data theft is perhaps no longer the major risk when it comes to data loss.