Web-Based Child Pornography: Quantification and Qualification of Demand

Chad M.S. Steel, Virginia Polytechnic Institute and State University, USA

ABSTRACT

While the supply of child pornography through the World Wide Web has been frequently speculated upon, the demand has not adequately been explored. Quantification and qualification of the demand provides forensic examiners a behavioral basis for determining the sophistication of an individual seeking child pornography. Additionally, the research assists an examiner in searching for and presenting the evidence of child pornography browsing. The overall search engine demand for child pornography is bounded as being between .19 and .49%, depending on the inclusion of ambiguous phrases, with the top search for child pornography being “lolita bbs”. Unlike peer-to-peer networks, however, the top child pornography related query ranks only as the 198th most popular query overall. The queries on search engines appear to be decreasing as well, and the techniques employed are becoming less reliant direct links to content.

Keywords: Child Pornography, Forensics, Search Engines

INTRODUCTION

One of the most frequent cases encountered by digital forensic examiners is the possession of child pornography. A major challenge for the examiner is characterizing the activities of the suspect – understanding how the suspect seeks out child pornography helps to meet that challenge. Because the web activities of a suspect are one of the most common features analyzed in an examination, understanding the demand for web-based child pornography is important.

A second reason for the study of web-based demand for child pornography is to provide an academic basis for statements related to the quantity of the activity. Those proposing particular legislation targeted at combating child pornography may tend to exaggerate the activities, whereas defense counsels may downplay the relationships between child pornography and particular search terms. By performing similarity analysis, we can provide accurate information to both parties without prejudice.

Major Contributions

The primary contribution of this article is the overall quantification of the demand for child pornography. By analyzing both Google and Dogpile data, the demand was found to be between .19 and .49% of all searches. This cannot be extrapolated to a particular percentage of users, but does show that one out of every 200 to 500 searches is related to child pornography.
Secondarily, the top search terms used by child pornographers were identified and quantified. The terms showed a distinct dichotomy between sophisticated searchers who are fluent in the language and nature of the subculture and unsophisticated users. This dichotomy can be exploited by forensic examiners to focus their searches based on user classification.

In terms of trends, the overall demand for child pornography using web-based searches appears to be declining and not increasing. While this does not mean that the actual demand for all child pornography is on the decline, it does show a potential shift toward other methods of acquisition.

Other areas that are frequently linked to child pornography were also explored—specifically the linkage between the terms associated with naturism/nudism and child pornography and that between bbs-associated searches and child pornography. Many naturism/nudism-related searches were found to focus on visual depictions, but not necessarily of children. On the other hand, the now antiquated term “bbs” was found to be strongly associated with child pornography.

Prior Art

Originally, child pornography on the Internet was traded largely over newsgroups through Usenet. Users were able to upload files semi-anonymously and from around the globe, and others could similarly download the content semi-anonymously. Initial studies by Rimm (Rimm, 1995) showed vast amounts of Usenet posts were dedicated to the trafficking of child pornography, though later analysis revealed significant flaws in his approach (Sigel & Sauer, 1999) (Hoffman & Novak, 1999). A follow-on study by Mehta was comprehensive and showed a much more accurate picture of Usenet-based trafficking in child pornography—specifically citing a figure of 15% of traffic being child-pornography related (Mehta, 2001). While Usenet is still active and growing—recent traffic rates are up to 4.5TB a day (Giganews)—it has been eclipsed by peer-to-peer and web-based traffic (Sandvine, 2008).

Early studies of peer-to-peer traffic were performed by the US Government Accountability Office (GAO) consisted of 42-44% child pornography (GAO, File-Sharing Programs: Peer-to-Peer Networks Provide Ready Access to Child Pornography, 2003) (GAO, File-Sharing Programs: Child Pornography is Readily Accessible over Peer-to-Peer Networks, 2003), though there are methodological issues with the study.

In 2006, Hughes et al conducted a study of the traffic being shared on the Gnutella network, a popular peer-to-peer network using the Gnutella protocol, focused on illicit content. The Hughes study used humans to classify search content, and used a small subset of queries (Hughes, Walkerdine, Coulson, & Gibson, 2006). A follow-on study by Hughes et al identified 1% of the queries on the Gnutella network as having child pornography. This study used an automated analysis on the same dataset and found a prevalence of child pornography amongst the queries of approximately 1% (Hughes, 2008). Another larger study of peer-to-peer networks confirmed the prevalence statistics of the Hughes study and found the single most prevalent query on the Gnutella network was child-pornography related (Steel, 2008).

No comprehensive studies of child-pornography demand on the web currently exist, but studies of the demand for pornography overall have been published by Beitzel et al. Beitzel et al analyzed a very large query log from a major search provider and categorized the content of the queries. The initial study found that approximately 10% of the queries were related to pornography (Beitzel, Jensen, Chowdhury, Grossman, & Frider, 2004). A follow-on study using automated analysis concluded that 7.9% of queries in 2007 were seeking pornography (Beitzel, Jensen, Lewis, Chowdhury, & Frieder, 2007). As child pornography is a subset of all pornography, it stands to reason that the prevalence of child pornography would be substantially less than the 7.9% figure.
Related Content

Two Methods for Active Detection and Prevention of Sophisticated ARP-Poisoning Man-in-the-Middle Attacks on Switched Ethernet LANs
[www.igi-global.com/article/two-methods-active-detection-prevention/58408?camid=4v1a](www.igi-global.com/article/two-methods-active-detection-prevention/58408?camid=4v1a)

Spam Image Clustering for Identifying Common Sources of Unsolicited Emails
[www.igi-global.com/article/spam-image-clustering-identifying-common/3906?camid=4v1a](www.igi-global.com/article/spam-image-clustering-identifying-common/3906?camid=4v1a)

A Forensic-as-a-Service Delivery Platform for Law Enforcement Agencies
[www.igi-global.com/chapter/forensic-service-delivery-platform-law/73968?camid=4v1a](www.igi-global.com/chapter/forensic-service-delivery-platform-law/73968?camid=4v1a)
Data Mining and Privacy
www.igi-global.com/chapter/data-mining-privacy/60946?camid=4v1a