Is the Internet for Porn?
An Insight Into the Online Adult Industry

Gilbert Wondracek[1], Thorsten Holz[1], Christian Platzer[1],
Engin Kirda[2], and Christopher Kruegel[3]

[1]Secure Systems Lab, Technical University Vienna
[2]Institute Eurecom, Sophia Antipolis
[3]University of California, Santa Barbara

Abstract

The online adult industry is among the most profitable business branches on the Internet, and its web sites attract large amounts of visitors and traffic. Nevertheless, no study has yet characterized the industry's economical and security-related structure. As cyber-criminals are motivated by financial incentives, a deeper understanding and identification of the economic actors and interdependencies in the online adult business is important for analyzing security-related aspects of this industry.

In this paper, we provide a survey of the different economic roles that adult web sites assume, and highlight their economic and technical features. We provide insights into security flaws and potential points of interest for cyber-criminals. We achieve this by applying a combination of automatic and manual analysis techniques to investigate the economic structure of the online adult industry and its business cases. Furthermore, we also performed several experiments to gain a better understanding of the flow of visitors to these sites and the related cash flow, and report on the lessons learned while operating adult web sites on our own.

1 Introduction

"The Internet is for Porn" is the title of a satirical song that has been viewed several million times on YouTube. Its popularity indicates the common belief that consuming pornographic content via the Internet is part of the modern pop-culture. Compared to traditional media, the Internet provides fast, easy, and anonymous access to the desired content. That, in turn, results in a huge number of users accessing pornographic content. According to the Internet Pornography Statistics [14], 42,7% of all Internet users view pages with pornographic content. From the male portion of these users, 20% admittedly do it while at work.

With a total worth of more than 97 billion USD in 2006 [14], the Internet porn industry yields more revenue than the top technology companies *Microsoft, Google, Amazon, eBay, Yahoo!*, and *Apple* combined. Interestingly, however, to the best of our knowledge, no study has yet been published that analyzes the economical and technological structure of this industry from a security point of view. In this work, we aim at answering the following questions:

Which economic roles exist in the online adult industry? Our analysis shows that there is a broad array of economic roles that web sites in this industry can assume. Apart from the purpose of selling pornographic media over the Internet, there are much less obvious and visible business models in this industry, such as *traffic trading* web sites or *cliques* of business competitors who cooperate to increase their revenue. We identify, in this paper, the main economic roles of the adult industry and show the associated revenue models, organizational structures, technical features and interdependencies with other economic actors.

Is there a connection between the online adult industry and cyber-crime? According to web statistics, adult web sites regularly rank among the top 50 visited web sites worldwide [2]. Anonymous and free access to pornographic media appeals to a huge audience, and attracts large amounts of Internet traffic. In this paper, we show that this highly profitable business is an attractive target for cyber-criminals, who are mainly motivated by financial incentives [9, 13].

What specific threats target visitors of adult web sites? Common belief suggests that adult web sites tend to be more dangerous than other types of web sites, considering well-known web-security issues such as malware, or script based attacks. Our results verify this assumption, and in addition, we show that many adult web sites use aggressive marketing and advertisement methods that range from "shady" to outright malicious. They include techniques that clearly aim at misleading web site visitors and deceiving

business partners. We describe the techniques we identified, and their associated security risks.

Is there domain-specific malicious activity? To be able to assess the abuse potential of adult web sites, we describe how we created and operated two adult web sites. This enabled us to identify potential attack points, and participate in adult traffic trading. We conducted several experiments and performed a security analysis of data obtained from web site visitors, evaluating remote vulnerabilities of visitors and possible attack vectors. We also identified and experimentally verified scenarios involving fraud and mass infection that could be abused by adult site operators, showing that we could potentially exploit more than 20,000 visitors spending only about $160.

To summarize, we make the following contributions:

1. We provide a detailed overview of the individual actors and roles within the online adult industry. This enables us to better understand the mechanisms with which visitors are redirected between the individual parties and how money flows between them.

2. We examine the security aspects of more than 250,000 adult pages and study, among other aspects, the prevalence of drive-by download attacks. In addition, we present domain-specific security threats such as disguised traffic redirection techniques, and survey the hosting infrastructure of adult sites.

3. By operating two adult web sites, we obtain a deeper understanding of the related abuse potential. We participate in *adult traffic trading*, and provide a detailed discussion of this unique aspect of adult web sites, including insights into the economical implications, and possible attack vectors that a malicious site operator could leverage. Furthermore, we experimentally show that a malicious site operator could benefit from domain-specific business practices that facilitate click-fraud and mass exploitation.

Ethical and Legal Considerations

Studying the online adult industry and performing experiments in this area is an ethically sensitive area. Clearly, one question that arises is if it is ethically acceptable and justifiable to participate in adult traffic trading. Similar to the experiments conducted by Jakobsson et al. in [15, 16], we believe that realistic experiments are the only way to reliably estimate success rates of attacks in the real-world.

We also implemented several preventive measures to limit ethical objections during our study. First, in the traffic experiments we performed, we only collected user information that is readily available by the webserver we set up (such as for example the HTTP request headers) or information that can be queried from the browser via standard interfaces such as JavaScript or Flash. Second, we anonymized the information and only stored the data for the offline analysis we performed after collecting the information. Third, we did not withdraw any funds but forfeited our traffic trading accounts at the end of the experiments. Fourth, we made sure that during our crawling experiments the number of outgoing requests was so low that it could not influence the performance of any website we accessed.

We also consulted the legal department of our university (comparable to the IRB in the US), and we were informed that our experiments are approved.

2 Analysis Techniques

In this section, we describe the experimental setup that we used to perform the analysis that allowed us to gain insights into the online adult industry. As part of this study, we first manually examined about 700 pornographic web sites. This allowed us to infer a basic model of the industry's economic system. In the second step, we created a system that crawls adult web sites and extracts information from them to automatically gather additional data.

2.1 Manual Inspection

Given the minimal amount of (academic) information currently available for this very specific type of Internet content, we basically had to start from scratch by projecting ourselves into a "consumer" role. By using traditional search engines, we located 700 distinct web sites related to adult content. This initial sample set provided the first insights into the general structure of adult web pages. For example, we observed that many web sites contain parts that implement similar functionality, such as preview sections and sign-up forms. In addition, we also looked for specialized services and web sites that appeal to "producers" of pornographic web sites. We used information gained from industry-specific business portals [29] to identify business-to-business web sites, such as adult hosting providers and web payment systems.

We identified several web site "archetypes" that represent the most important business roles present in the online adult industry. The majority of web sites that we analyzed fits into exactly one of these roles. The economic relationships between these entities are shown in Figure 1. Whenever suitable, we named the roles according to the industry jargon. In the following section, we provide a detailed overview of each role. Based on these observations, we then created an automated crawling and analysis system to gain a broader insight into the common characteristics of adult web pages, operating on a large sample set of about 270,000 URLs (on more than 35,000 domains).

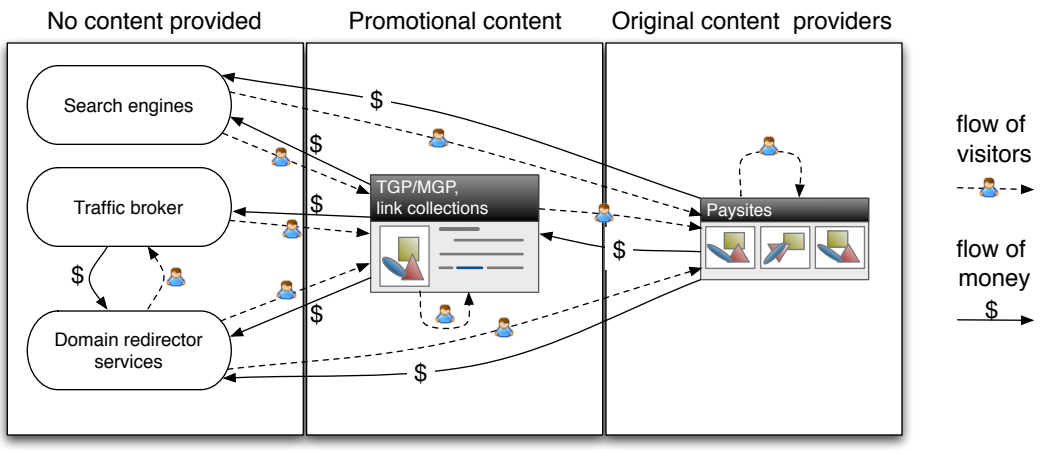


Figure 1: Observed traffic and money flows for different roles within the online adult industry.

2.2 Identified Site Categories

Based on our observations, we can classify the market participants in the following categories.

2.2.1 Paysites

This type of web sites constitutes the economic core of the online adult industry. These web sites typically act as "content providers", producing and distributing pornographic media such as images and videos via their web pages, charging money in return. Most common users would consider these sites to be representative for this genre.

2.2.2 Link Collections, TGP / MGP

Complementary to paysites, a large number of pornographic web sites promise free content. These sites often call themselves *link collections*, *thumbnail gallery posts* (TGPs) or *movie gallery posts* (MGPs), depending on the provided form of pornographic media. We use the term *free site* to denote these types of web sites.

Link collections typically consist of a series of hyperlinks (often adding textual descriptions of the underlying media) to other web sites. TGP and MGP sites are structurally similar, with the addition of displaying miniature preview (still) images next to each link. It is indicative for free sites that they do not produce their own content. Our evaluation shows that they receive media from other content providers, as their main economic role is marketing for paysites. A secondary role is *traffic trading*, as it will be explained in Section 2.2.6.

2.2.3 Search Engines

With the multitude of different providers, specialized search engines evolved to fit the need of every potential customer. Functionally similar to general purpose search engines such as Google, adult search engines [10] allow users to search for web sites that match certain criteria or keywords. Unlike traditional search engines, adult search engines claim to manually classify the web sites in their index, instead of relying on heuristics or machine learning techniques. However, this claim – suggesting that their results are more accurate than other search engines – is highly questionable, considering the fact that pornographic pages account for 12% of the total number of web pages on the Internet [14]. Search engines generally generate revenue by displaying advertisements and selling higher-ranked search result positions.

2.2.4 Domain Redirector Services

Interestingly, there are services that specialize in managing adult *domain portfolios*. They are similar to commercial *domain parking* services that display web pages with advertisements (which are often targeted towards the domain name) in lieu of "real" content [28].

Adult domain redirector services such as Domain Players Club [6] not only allow their clients to simply park their domains, but are rerouting any web traffic from their clients' domains to adult web sites. Adult sites that wish to receive traffic from the redirector service have to pay a fee for being registered as a possible redirection target. The exact destination of the redirections is typically based on the string edit distance between the domain name of the web site participating in the redirector service, and the domain name of the adult web sites which wish to receive traffic. For example, a user might browse to `www.freehex.com`, not knowing that this site participates in a redirector service. The user

will then be redirected to an adult web site with a domain name that has a low edit distance to this domain name. The destination adult web site initially has to pay a fee for being considered by the redirection service, while the domain owner is rewarded for any traffic that originates from his domains. Technically, these redirector services work by using a layer of HTTP redirections, giving no indication to the user that a redirection has occurred.

From a miscreant's point of view, these redirector services appear to be an ideal tool for *typo-squatting* [28]. Typo-squatting is the practice of registering domain names that are syntactically very close to the names of legitimate web sites. The idea behind typo-squatting is to parasitize web traffic from users that want to go to the legitimate site, but make a typographical error while entering the URL.

2.2.5 Keyword-Based Redirectors

Several businesses offer a service that aims at increasing the visibility and (traditional) search engine ranking of their clients (adult web sites). To this end, keyword based redirector services operate websites that have a large numbers of subdomains. The names of these subdomains consist of combinations of adult-related search engine keywords.

Similar to domain redirector services, these subdomains are configured to redirect visitors to "matching" web sites, e.g. the redirector's clients. Clearly, this technique is an attempt to exploit ranking algorithms to achieve higher search result positions, effectively subverting the search engine's business model of selling search result positions. Furthermore, it is an efficient way to prepare a web site for spam advertisement. Unsolicited bulk (spam) mails tend to yield a higher penetration rate when embedded links differ from mail to mail [23].

2.2.6 Traffic Brokers

This unique type of service provider allows its clients to directly *trade* adult web traffic for money, and vice versa (i.e., web traffic can be turned into real money with this kind of providers). Prospective clients who want to *buy* traffic can place orders (typically in multiples of 1,000 visitors) that will then be directed to a URL of their choice. Usually, the buyer can select the source of the web traffic according to several criteria, such as interest in certain niches of pornography or from specific countries. Available options also include traffic that originates from other adult sites, e-casinos, or from users who click on advertisements such as pop-up or pop-under windows, or even links in YouTube comments. Another option is traffic that is redirected from recently expired domains, which have been re-registered by the traffic broker.

On the other hand, clients who want to *sell* traffic can do so by redirecting their visitors to URLs that are specified by the traffic broker, receiving money in return. If the broker has no active orders from buyers for the type of traffic that is provided, the traffic is sent back to a link specified by the client. However, if the broker has an active order, the traffic is redirected to the site of the buyer's choice and the seller is credited a small amount of money. Figure 2 visualizes the flow of visitors and money for both scenarios.

Before a client can participate in traffic trading, brokers typically claim that they check the source or destination site of the traffic to prevent potential abuse. For example, many traffic brokers state that they do not tolerate hidden frames on target web sites. However, in our experiments with traffic brokers, we found this claim to be false: We successfully managed to buy large quantities of traffic for a web site that makes extensive use of hidden `iframes` and even performs vulnerability checks on its visitors (see Section 4 for more details).

2.3 Experimental Setup

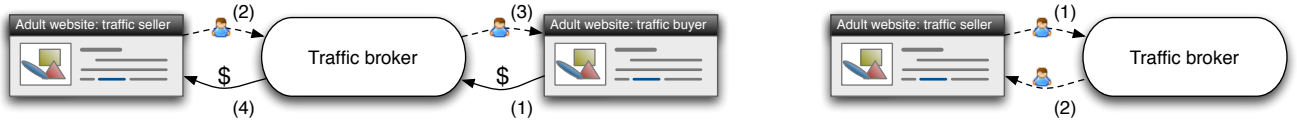
To acquire real-world data and to perform a large-scale validation of the initial results from our manual analysis, we created a web crawler system. Based on our observations, we added several domain-specific features. Our system consists of the following components.

2.3.1 Search Engine Mining

For our crawling system, it was necessary to acquire a set of adult web sites that were suitable as initial input. To mimic the way a consumer would look for adult web sites, we made use of search engines. We manually compiled a set of domain-specific search queries and automatically fed it as input to a set of 13 search engines. This included three general purpose search engines (Google, Yahoo, and Microsoft Live) and ten adult search engines. We then automatically extracted the URLs from the search results and stored them in a database. The result set consisted of 95,423 URLs from 11,782 unique domains. These URLs were the seed used in the crawling step.

2.3.2 Crawling Component

The core component of our system is a custom web crawler we implemented for this purpose. We configured it to follow links up to a depth of three for each domain. For performance reasons, we additionally limited the maximum amount of URLs for a single domain to 500. Starting from the previously-mentioned seed, we crawled a total of 269,566 URLs belonging to 35,083 web sites. For each crawled URL, we stored the web page source code, and the embedded hyperlinks. This formed the data set for our subsequent analysis. In addition to the crawling, we used the



(a) Traffic buyer is interested in receiving traffic and pays for it.

(b) No traffic buyer available, traffic broker returns visitor to a specific URL.

Figure 2: Schematic overview of traffic trading and the flow of visitors/money.

following heuristics to further classify the content, and detect a number of features.

Enter Page Detection. A characteristic feature of many adult web sites (unrelated to their economic role) are "doorway" web pages that require visitors to click on an `Enter` link to access the main web site. These *enter pages* often contain warnings, terms of use, or reminders of legal requirements (for example, a required minimum age for accessing adult material).

In order to automatically detect enter pages, we used a set of 16 manually compiled regular expressions to scan textual descriptions of links. Since some enter pages use buttons instead of text-only descriptions, we also checked the HTML alternative text for images. For example, if a link description matches $.*enter\ here.*$ or $.*over.*years.*$, we classify the page as an enter page.

Adult Site Classifier. Since we wish to avoid crawling non-adult web sites, and since not all outgoing links lead to adult web sites, we created a simple, light-weight keyword-based classifier to identify adult web sites. To this end, we first check for the appearance of 45 manually selected, domain-specific keywords in the web site's HTML meta description tags. In case no matches are found, we also extend our scan to the HTML body of the web page. If at least two matches are encountered, we consider the web site to contain pornographic content.

According to our experience, this *naïve* classification works surprisingly well, as porn sites usually promote their content openly. To evaluate the true positive (TP) and false positive (FP) rate of our classifier, we ran it on a hand-labeled subset of 102 web sites that we chose randomly our manual-analysis test set. It achieved rates of 81.5% TP and 18.5% FP. Moreover, a limitation of our current implementation is that it currently only works with English-language web sites. After excluding non-English web sites, the rate improved to 90.1% TP and 9.9% FP. We are aware that far more advanced classifiers for adult sites exist, for example systems that include image recognition techniques [11]. However, these classifiers are typically aimed towards filtering pornographic content and are not readily and freely available, and our current heuristic yields sufficiently accurate results for our purposes.

2.3.3 Client Honeypots

Malicious web sites are known to direct a multitude of different types of attacks against web surfers [21, 22, 27]. Examples include drive-by downloads, Flash-based browser attacks, or malformed PDF documents that exploit third-party software. To detect such attacks, we used two different client honeypots to check the web sites that we crawled in our study.

Capture-HPC. We used an adapted version of the Capture-HPC [25] *client honeypot*. The tool detects and records changes to the system's filesystem and registry by installing a special kernel driver. We set up Capture-HPC in virtual machines (VMs) with a fully patched Windows XP SP2, resembling a typical PC used for web browsing. We then instrumented the VMs to open the URLs from our crawling database using Internet Explorer 7 (including the popular Flash and Adobe PDF viewer plugins). This allowed us to detect malicious behavior triggered by (adult) web sites. In our experimental setup, we ran eight instances of the VMs in parallel, to achieve a higher throughput rate.

Wepawet. To complement the analysis performed by Capture-HPC, we used another client honeypot, namely Wepawet [18, 17], in parallel. The software features special capabilities for detecting and analyzing Flash-based exploits, and for handling obfuscated JavaScript, which is commonly used to hide malicious code. Wepawet also tries to match identified code signatures against a database of known malware profiles, returning human-readable malware names.

2.3.4 Economic Classification

To decide if paysites are more or less secure (i.e., trustworthy) than free sites, we created a heuristic for automatically classifying each web site depending on its economic role. Our classifier is limited to determining if a web site is either a paysite or a free site; otherwise, the web site's economic role remains undefined.

Paysite Indicators. We identity paysites based on manual observations and by using information we found on adult business-to-business web sites: we compiled a list of 96 adult payment processors, i.e., companies appointed by a

web site operator to handle credit card transactions on behalf of him. If a web site links to a payment service provided by one of these processors, we immediately mark it as a paysite. In case no payment processor is found, we look for additional features of paysites. To this end, we match the web site source code against a set of regular expressions to determine if it contains a "tour", "member section", or membership sign-up form. We assume these structural features to be indicative for paysites, as we did not find any counter-examples in our manual observations.

Free Site Indicator. To identify free web sites, we examine their hyperlink topology. For this classification, we only regard *outgoing links* as a reliable feature, as it is not feasible to recover (all) incoming links for a web site. We analyze the number of hyperlinks pointing to different domains for each web site, and additionally compare the Whois entries for both the source and destination domains. If a web site exceeds a threshold t of links to "foreign" domains (e.g., the Whois entries show different registrants), we label it as a free site. To evaluate this classifier and instantiate a value for t, we tested it on a hand-labeled set of 384 link collection web sites that we selected randomly from our database. Based on this experiment, we chose $t = 25$ for the evaluation.

3 Observations and Insights

During our crawling experiments, we observed several characteristics of adult sites. In this section, we provide an overview of the most interesting findings, and discuss how they are security-relevant.

3.1 Revenue Model

The ultimate goal for commercial web site operators is of course to earn a maximum amount of money, and the slogan "sex sells" is a clear testimony to this fact. In the following, we analyze the revenue model of the major categories identified in Section 2.2.

3.1.1 Paysites

We found the revenue model of paysites to be centered around selling *memberships* to customers. A membership grants the customer access to an otherwise restricted *member area* with username/password credentials. In the member area, an archive of pornographic media can be browsed or downloaded by the customer. Memberships typically have to be renewed periodically, causing recurring fees for the customer and, therefore, providing a steady cash-flow for the paysite. To appeal to customers and to create a stimulus for purchasing a membership, paysites rely heavily on a number of marketing and advertising techniques, like for example:

A "Tour" of the Web Site. Similar to traditional advertising methods (for example cinematic trailers for movies), preview media content is published for free on the paysites' web pages, eventually directing the user to membership sign-up forms.

Search Engines and Web Site Directories. Specialized promotion services, such as *adult search engines* and web site directories, allow users to submit hyperlinks to web sites. These links are then categorized (depending on the nature of the content), and made available on a web site where they can be searched and browsed. While these services are typically free of charge, higher ranked result positions can be purchased for a fee.

Affiliate Programs. The main purpose of an *affiliate program* is to attract more visitors to the paysite. The business rationale is that more visitors translates to more sales. To this end, paysites allow business partners to register as affiliates, thus giving them access to promotional media. This media is designated for marketing the paysite. It consists of hyperlinks pointing to the paysite and optionally includes a set of pornographic media files. In return for directing visitors to the paysite, affiliates are rewarded a fraction of the revenue that is generated by those customers that were referred by the affiliate.

By using affiliate programs, paysites are effectively shifting part of their marketing effort towards their affiliates. Additionally, those sites that distribute the media files (instead of just providing hyperlinks) can reduce their resource consumption (such as bandwidth costs) as an additional benefit. Many paysites even offer specialized services to their affiliates, for example, by providing preview images and textual descriptions of the content, or even creating administrative shell scripts. Also, Internet traffic statistics are made available to affiliates, so that they can optimize their marketing efforts.

3.1.2 Free Sites

Free sites typically participate in multiple affiliate programs. We found examples of sites participating in more than 100 different programs, generating revenue by directing visitors to paysites. To account for the origin of customer traffic, paysites usually identify their affiliates by unique tokens that are assigned on registration. These tokens are then used to associate traffic with affiliates, for example, by incorporating them as HTTP parameters in hyperlinks pointing from the affiliate site to the paysite. The same technique is used to identify links originating from spam mails, providing the site with the means to evaluate a spammers' advertising impact.

Often, affiliates can choose between two revenue system options:

- Pay-per-sign-up (PPS): The affiliate receives a one-time payment from the paysite for each paysite member that was referred by the free site.
- Recurring income: In contrast to PPS, the affiliate can choose to receive a fraction of each periodic fee as long as the membership lasts.

We found that the payment systems that are used to transfer money from paysites to affiliates offer a wide variety of options, including wire transaction, cheques, and virtual payment systems. In addition to affiliate programs, free sites display advertisements to increase their revenue.

3.2 Organizational Structure

Paysites We noticed that many paysites are organized in *paysite networks*. Such networks act as umbrella organizations, where each paysite contains hyperlinks to other members of its network. Additionally, networks often offer customers special membership "passes" that grant collective membership for multiple paysites.

Interestingly, however, upon inspection of the Whois [20] entries for member sites within several networks, we found the registration information to often match (e.g., the sites were belonging to the same owner). Apparently, the individual network members prefer to create the outward impression of representing different enterprises, when they are in fact part of the same organization. This indicates that a diversification among paysites, depending on the sexual specifics of the offered content, is advantageous for the owners. These specialized sites are called *niche sites* in the industry jargon.

Free Sites Similar to paysite networks, we found free sites to be also organized in networks. However, in contrast to paysites, free sites also frequently link to each other even if the site owners differ. This means that business competitors are collaborating. This appears counter-intuitive at first. However, one has to take into account that cross-linking between free sites is a search engine optimization method. Thus, the search engine ranking of all sites participating in a "clique" of free sites improves, as the sites are artificially increasing their "importance" by creating a large number of hyperlinks pointing towards them.

3.3 Economic Roles

From a consumer perspective, paysites and free sites are the most important types of adult web sites. To get an overview of the distribution of paysites and free sites with regard to the total population of adult web sites, we applied our classification heuristic to the 35,083 adult web sites (domains) in our data set.

Our classifier was able to determine the role of 87,7% of these web sites. For the remaining 12,3%, whose roles remained undefined, we found a high percentage of web sites that either served empty pages, returned HTTP error codes (for example, HTTP 403 "Forbidden"), or were parked domains. We assume that many of these sites are either still under construction or simply down for maintenance during our crawling experiment.

Our results indicate that 8.1% of the classified sites are paysites and 91.9% are free sites (link collections). This is consistent with the intuition that we gained from our initial, manual analysis, showing that most adult site operators make money by indirectly profiting from the content provided by paysites.

3.4 Security-Related Observations

For either economic role, we found a relatively large number of web sites that use questionable methods and techniques that can best be described as "shady." Unlike well-known web-based attacks and malicious activities (such as drive-by downloads [21, 27]), these practices directly aim at manipulating and misleading a visitor to perform actions that result in an economic profit for the web site operator. Overall, we found free sites to employ at least one of these techniques more often (34.2%) when compared to paysites (11.4%). In particular, we frequently found the techniques listed below on adult web sites.

3.4.1 JavaScript Catchers

These client-side scripts "hijack" the user's browser, preventing him from leaving the web site. To this end, usually JavaScript code is attached to either the `onunload` or `onbeforeunload` event handlers. Anytime the user tries to leave the web site (e.g., by entering a new address, using the browser's "Back" button, or closing the browser) a confirmation dialogue is displayed. The user is then asked to click on a button to leave the web site, while, at the same time, advertisements are displayed or popup windows are spawned. Apart from the obvious annoyance, this could easily be used in a *clickjacking* attack scenario [12]. We detected catcher scripts in 1.2% of the paysites and 3.9% of free sites.

3.4.2 Blind Links

This technique uses client-side scripting via JavaScript to obscure link destinations, effectively preventing the addresses from being displayed in the web browser's status bar. The most popular methods that we found in the wild either work by overwriting the `window.status` or `parent.location.href` variables. We scanned the

source code of the web sites for occurrences of these variable names, and found 10.9% of paysites, and 26.2% of free sites to use blind links.

While the destination addresses are still contained in the web page source code, we believe it is fair to assume that most users will be unable to extract them. This is problematic, as it not only leaves the user unaware of the link's destination (leading to different web sites), but could also potentially be used to mask malicious activities such as cross site scripting (XSS) or cross site request forgery (CSRF) attacks.

3.4.3 Redirector Scripts

Redirector scripts make use of server-side scripting (for example PHP scripts) to redirect users to different web sites. In contrast to blind links, the link targets are determined at the server at run-time, making it impossible for a client to know in advance where a link really points to.

Typically, these redirector scripts are presented in combination with pornographic media. For example, small preview images usually have links to full-size versions attached. Instead of this expected behavior, users are redirected with a probability p to different web sites (so called *skimming rate*). The rationale behind redirector scripts is that users will know from experience that by keeping on clicking on the preview image, the desired media will eventually be shown at some point. At the same time, they "generate" artificial outgoing traffic for the web site, even though the user originally never intended to leave the site.

In our crawler implementation, we use a simple, yet effective technique to detect redirector scripts. Whenever our system finds hyperlinks with a destination address that contains a server-side script (currently `*.php` and `*.cgi` scripts), it resolves the link 10 times. If there is more than one destination address, the script is regarded as a redirector script, and the set of targets is added to our crawling queue. We chose a value of 10, because in our initial tests, we observed this as an upper bound for the number of redirection targets. When tested on a sample of 100 redirector scripts, none of them exceeded this threshold.

We found examples of p ranging from 0 (no *random* redirection) to 1 (the promised content is never shown). Also, the number of possible target addresses n varied from 1 to 6 destinations. Interestingly, only 3.2% of paysites but 23.6% of free sites contained redirector scripts. This implies that free sites have an incentive for using this technique.

The most likely explanation of this phenomenon are traffic brokers (see Section 2.2.6). These services have specialized in (adult) *traffic trading* and allow visitor traffic to be sold, a unique feature available only in this type of online industry. This means that a miscreant could lure unsuspecting visitors who click on pornographic media to click on redirector links. The resulting traffic can then be sold to such a traffic trading service, which redirects it to targets of the buyer's choice. The web site operator earns money with every click, even if a single visitor clicks on one links many times – something not possible in traditional online advertisement.

3.4.4 Redirection Chains

If web sites which contain redirector scripts link to other sites with redirector scripts, we call this a *redirection chain*. This topology can be abused to further increase the revenue from artificial traffic generation.

We observed that JavaScript catchers are frequently used in conjunction with redirector chains, effectively "trapping" the user in a network of redirections. In our evaluation, we found 34.4% of those web sites that use redirector scripts to be part of redirector chains. Potentially, this could easily be abused for performing click-fraud or similar traffic-based cyber-crime because it enables the redirection operators to direct large amounts of "realistic" traffic to destinations of their choice. We study this phenomenon in more detail in Section 4.

3.5 Malware

To find more "traditional" web-based attacks, we applied our client honeypot analysis (see Section 2.3) to all 269,566 pages in our data set (which represents the adult web sites' main pages, subdomain pages, and enter page targets). Of these, 3.23% were found to trigger malicious behavior such as code execution, registry changes, or executable downloads. This percentage is significantly higher than what we expected based on related work [21], where slightly more than 0.6% of adult web sites were detected as malicious.

We used Anubis [5], a behavior-based malware analysis tool, to further analyze the malware samples that were collected by the honeypots. Also, Wepawet could successfully identify several families of exploit toolkits used by the malicious sites. This gave us human-readable malware names for the malware, showing that the most popular types of malware that we found are Spyware and Trojan downloaders (e.g., `rootkit.win32.tdss.gen` or `backdoor.win32.bandok`).

Whenever `iframes` were used as infection vectors, we extracted the hosting location of the injected code, finding the malicious code to be mostly (98.2%) not stored on the adult web sites themselves. We believe this is a clear indication that the web sites that distribute the malware were originally exploited themselves, and are not intentionally serving malware. This was also confirmed by results from Wepawet, which automatically attributed several exploits to the "LuckySploit" malware campaign [8].

4 Becoming an Adult Webmaster

The analysis methods and findings presented in the previous sections allow us to gain information from an external observer's point of view, enabling us to outline the online adult industry's business relationships and studying some security-related aspects. However, we are also interested in more technical, security relevant information that is only available to adult web site operators themselves, for example, data about the web site visitors or the mechanisms behind traffic trading. One of the goals of our research is to estimate the malicious potential of adult web sites, for example, as a mass exploitation vector. Therefore, we also need the *internal* point of view to understand this area of the Internet in detail.

Unfortunately, we are not aware of any available real-world data set that could be used for such an analysis. Therefore, we took over the role of an adult webmaster and created two adult web sites from scratch to conduct our experiments.

4.1 Preparation Steps

To be able to interact with the adult industry, we performed the following operations to mimic an adult web site. First, we created two relatively simple web sites. We designed both sites' layout to resemble existing, genuine adult web sites, allowing us to blend in with the adult web site landscape. We chose to mimic two popular types of free sites, one "thumbnail gallery" web site and one link collection web site. After registering domain names that are indicative for adult web sites, we put the sites online on a rented web hosting server.

Affiliate Programs. To receive promotional media, we then registered as an adult web site operator at eight adult affiliate programs. Surprisingly, the requirements for joining affiliate programs appear to be very low. In our case, only the web site URL, a contact name, and an email address had to be provided. There is no verification of neither the contact identity information nor is a proof of ownership required for the web site.

Immediately before signing up to an affiliate program, we created a snapshot of our web server access logs. As soon as an affiliate program accepted our application, we compared the current access logs to the snapshot. We found that six of the eight affiliate programs were accepting our application, even though no access to our web sites happened during the period between sign-up and acceptance. This means, that they were blindly accepting our application, performing no check of the web sites at all.

Traffic Brokers. Furthermore, we also registered our web sites at four traffic brokers that we chose due to their popularity among adult site operators, allowing us to participate in traffic trading. The registration procedure was almost identical to affiliate programs, and again, most brokers accepted our application without looking at the web sites. Only one broker checked our site and subsequently declined our application after detecting our analysis scripts (see next section).

Payment System. To be able to buy traffic, we had to send money to the traffic brokers. To this end, we used the "ePassporte" electronic payment system, that is popular among adult site operators, as it is widely accepted in the adult industry. We spent slightly more than $160 for our traffic trading experiments (including transaction fees).

4.2 Traffic Profiling

Our main goal in operating these web sites is to acquire as much security relevant information about web traffic coming to the sites as possible. To this end, we added several features to the web sites that allow us to collect additional information from each visitor. Since the collected data may contain detailed information about a unique visitor, and especially privacy related information, we implemented several precautions to protect the user's privacy (e.g., anonymization of the collected raw log data). This information is then used in subsequent *offline* analysis steps, for example to determine if a user is vulnerable to remote exploits like arbitrary code execution or drive-by downloads. Specifically, we collect the following information from each visitor:

Browser Profiling. First, we store general information for each visitor that is available through the web server log files, for example, the `User-Agent` string and the HTTP request headers that are sent by the user's browser.

Additionally, we added several JavaScript functions to the web site. These routines gather specific data about a visitor's web browser capabilities, for example, the supported data types or installed languages. We also collect information about any installed browser plugins, including their version numbers. This information is security relevant, as browser plugins are frequently vulnerable to remote exploits, and we can infer from this data if the visitor is potentially vulnerable to a drive-by download attack.

In particular, we are interested in the Flash browser-plugin [1], which is typically used to embed videos in web sites, as it is known for its bad security record [26]. Our intuition is that visitors to multimedia-rich adult web sites will most likely have Flash installed. Therefore, in addition to the plugin detection, we implemented a JavaScript-independent Flash detection mechanism that uses a small Flash script to check if the user has Flash installed. This allows us to detect vulnerable clients, even if they have JavaScript turned off (see Section 4.4). In addition to Flash, we also check for vulnerable versions of browser plugins

for the Adobe PDF document viewer and Microsoft Office as they are the most prevalent targets for malicious attackers [4].

Outgoing Links. To be able to verify statistics provided by affiliate program partners, we track all outgoing (i.e., leaving the web site) hyperlinks that a user has clicked. This is implemented by scripts that operate similar to redirector scripts often employed on adult web sites (see Section 3.4.3 for details).

4.3 Traffic Buying Experiments

After having prepared the web sites with our profiling tools, we placed orders for buying web site visitors at three different traffic brokers. We tested different brokers to study the differences in delivered traffic and to gain a better understanding of their intricacies. In total, we ordered almost 49,000 visitors at the three different traffic brokers during a period of seven weeks. We spent a total of $161.84 on these traffic orders (average $3.30 per thousand visitors). Surprisingly, each traffic broker redirected traffic to our site (almost) instantly after placing an order. This suggests that they have an automated traffic distribution system in place, capable of flexibly rerouting traffic to customers, and enough incoming traffic that they can handle orders in a timely manner. Checking our web server logs confirmed that we indeed received the correct amount of visitors (e.g., clients with unique IP addresses) at the correct rate for *all* orders.

In addition to the rate limit, we also chose the more expensive "high quality" option when buying traffic, which is regarded by traffic brokers as synonymous with traffic coming mostly from the US and Europe. To verify the geographical origin of traffic, we performed an IP to country lookup for the bought traffic. We found that 98.22% of the traffic really originates from the US and Europe, thus the origin is correct for the vast majority of visitors.

4.4 Profiling Results

After having received the ordered amount of traffic, we analyzed the output of the profiling steps outlined in Section 4.2. An overview of the results of this analysis is shown in Table 1. All brokers sent a similar type of visitors to our site and there are no major differences between the brokers. Therefore, we discuss the overall results in the following sections.

4.4.1 Browser Profiling

When a visitor accesses one of our web sites, we automatically start to collect information about him (e.g., all request headers and information about browser extensions). In certain cases, our system cannot obtain this profiling information for a web site visitor. The reasons can be manifold, for example a client can have JavaScript support disabled, it can be an "exotic" web browsers with reduced functionality, the visitor might stay for only a few seconds on our web site, or it might not be a human visitor but a bot. The most prevalent case were visitors that did not correctly execute our JavaScript-independent Flash detection: 18,794 (38.43%) of our overall visitors behaved in this way. In contrast, 30,106 (61.57%) visitors correctly performed the test, and of those 96.24% had Flash installed. Furthermore 10,214 visitors (about 20.89%) did not download any images, but just requested the HTML source code of the site. While we cannot coherently explain this behavior, we think that it is caused by bots (e.g., click-bots [7]), since the browser of a human visitor would start to download the complete content of the site.

For about 47% of all visitors we were able to build a *complete* browser profile, which includes all the information we are interested in. For the remaining visitors only certain types of information were collected (e.g., only HTTP headers and no other information since the visitor spent not enough time on our site). We opted to analyze only the cases in which we have collected the complete browser profile to be conservative in our analysis.

During our analysis we also detected some noteworthy anomalies that prohibit browser profiling. For example, about 0.53% of the visitors used browser versions typically found in mobile phones or video game consoles (such as Nintendo Wii, Playstation Portable, or Sony Playstation). These devices do not fully support JavaScript or have a limited set of features, preventing our profiling scripts from executing correctly. We also found that in about 0.14% of the cases our profiling did not work since the HTTP headers were purged, a fact that we could attribute to clients which have the Symantec Personal Firewall installed.

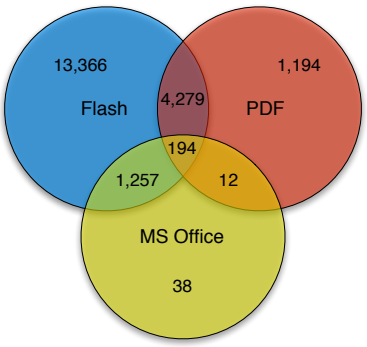
4.4.2 Vulnerability Assessment

We determine if a client is vulnerable to *known* exploits by matching the visitor's browser properties (e.g., version number of common plugins and add-ons) against a list of common vulnerabilities we compiled manually. We focussed on only the most prevalent browser plugins such as those related to Adobe Flash and PDF, and Microsoft Office. These three plugins had seven vulnerabilities in the recent past, and an attacker can buy toolkits that exploit these vulnerabilities to compromise a visitor [4]. Since realistically, additional exploits (even some that are not publicly known yet) exist in the wild, this provides us with a lower bound for the number of vulnerable systems among visitors to our web sites. Using this heuristic, we found that more

	Broker A	Broker B	Broker C	Total
Ordered Visitors	12,000	7,900	29,000	48,900
Performed Flash Detection	8,638 (71.98%)	5,010 (63.42%)	16,458 (56.75%)	30,106 (61.57%)
↦ Flash Found	8,401 (97.26%)	4,876 (97.33%)	15,697 (95.38%)	28,974 (96.24%)
Complete Browser Profiles	6,183 (51.53%)	3,682 (46.60%)	13,176 (45.43%)	23,041 (47.12%)
↦ Vulnerable	5,251 (84.93%)	3,242 (88.05%)	11,847 (89.91%)	20,340 (88.28%)
# Clicked Links	3,662	2,742	8,997	15,401

Table 1: Statistics about the visitors studied during our traffic buying experiments.

than 20,000 visitors had at least one vulnerable component installed and more than 5,700 visitors had multiple vulnerable components. Figure 3a shows a Venn diagram depicting the prevalence of different types of vulnerabilities.



(a) Distribution of the three vulnerability types we examined. Note that the display format is not proportional.

User-Agent	# Suspicious Clients	Type
FunWebProducts	260	Adware
SIMBAR	136	Adware
DesktopSmiley	93	Spyware
JuicyAccess	85	Nagware
Antivir XP 2008	52	Fake AV
Other	289	-
Total	915	-

(b) Overview of suspicious User-Agent strings that we observed frequently, indicating that these clients are presumably infected with some kind of malware, e.g., scareware or adware.

Figure 3: Results for vulnerability assessment of clients studied during traffic experiments.

A malicious site operator could take advantage of these vulnerabilities and compromise the visitor's browser with a drive-by download [21]. Besides the opportunity to build a botnet with only a small investment (e.g., we spent $160 and could potentially infect more than 20,000 machines), an operator could also earn money with the help of so called *Pay-Per-Install* (PPI) affiliate programs. In a PPI program, the "advertiser" pays the partner a commission for every install of a specific program by a user. The exact amount of this commission depends on the countries that the users come from. For example, we registered at one PPI program (note that we did *not* install any software to clients) and found the rate for 1,000 installs to computers located in the US and parts of Europe to be set to $130, while it would be as low as $3 for most Asian countries. This is consistent with information that we manually compiled from five other PPI program web sites. Related work that focusses on PPI (for example [24]) lists even higher prices per installation. Since we only bought US and European traffic in our experiments, we found a large fraction of traffic to fall into the highest selling PPI category (more than 95%). While an in-depth analysis of PPI programs is outside the scope of this work, these figures clearly show that it would be highly profitable for a malicious site operator to participate in PPI programs, and covertly trigger installs of unwanted software at vulnerable clients.

In addition to vulnerable browser versions and plugins, we also analyzed the `User-Agent` strings obtained from the visitor's browser. This enables us to detect certain cases of clients that are already compromised: While the `User-Agent` string can be arbitrarily set by a client, it is still a good indicator for clients that are infected with certain types of malware, which intentionally "mark" infected clients to avoid re-infection or change the behavior of web sites that act as an infection vector. We found 915 clients (1.87%) that contain known malware marker strings, such as for example the adware "Simbar", or scareware like "Fake Antivirus 2008". Figure 3b provides an overview of the most common suspicious `User-Agent` strings we observed in the visitor's browser.

4.5 Traffic Selling Experiments

Traffic brokers also allow their clients to *sell* web traffic, paying them for visitors that are redirected to the broker's web site; from there the visitors are forwarded to traffic buyers (see Section 2.2.6 for details). The commission a traffic seller receives mainly depends on the niche that is attributed to the traffic, and is influenced by the type of web site the seller operates. To explore the security aspects of traffic

selling, we included traffic selling links on our web sites and participated in this business.

4.5.1 Click Inflation Fraud Scenario

The first thing we noticed is the fact that traffic brokers do not require traffic selling web sites to include any content (for example a script) that is hosted by the traffic broker or by a third party. This stands in contrast to other types of web businesses that rely on partner web sites to publish information. For example, online advertisers such as Google typically require the inclusion of JavaScript code that is hosted by the advertiser himself on the publisher's web site. This code enables the content provider to acquire information about the publishing web site that can be used for abuse and fraud detection, for example by computing the click-through-ratio (CTR) or by checking the cookie information of a user that clicks on a link.

Since traffic brokers do not use this technique, they cannot implement these well-known techniques for fraud detection and are thus subject to specific abuses. However, we found that traffic brokers check the HTTP `Referrer` header of redirected traffic to see if it really originates from the seller's web site. If this is not the case, the traffic is either rejected (redirected back to the seller), or only a very low price is paid.

These observations led us to the assumption that the level of sophistication of anti-fraud techniques employed by traffic brokers is rather low. To verify this assumption, we devised a simple, yet effective fraud scenario to test the vulnerability of traffic brokers to click fraud. In this scenario an attacker (legitimately) buys traffic from at least one traffic broker, and then "resells" this traffic to n different traffic brokers in parallel by forwarding the incoming traffic. Figure 4 illustrates the concept of our attack, which is a variation of click inflation attacks [3].

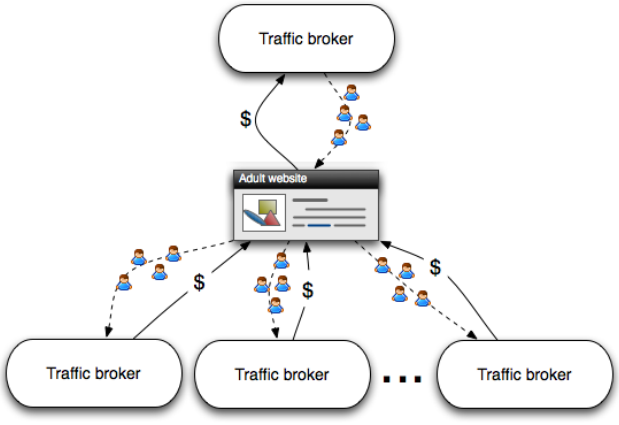


Figure 4: Overview of n-fold click inflation fraud scenario.

With the help of this *n-fold click inflation*, an attacker can earn money if the total earnings from selling the traffic n times exceeds the amount of money she needs to spend for buying the traffic. Furthermore, she could even earn more money by abusing each visitor she bought, for example by compromising vulnerable visitors with the help of a drive-by download.

From a technical point of view, we found that simple HTTP or JavaScript redirections to the traffic selling URLs would not suffice, as many popular web browsers such as Internet Explorer and Mozilla Firefox incorporate pop-up blocking features that prevent opening new browser windows without the user's interaction. However, during our traffic buying experiment, we noticed that a relatively high amount of visitors clicked on links on the web site: overall, we had more than 15,400 clicks based on just 48,900 visitors (see Table 1). From those users that clicked on at least one link, we received an average of 3.78 clicks. Based on this observation and the fact that pop-up blockers do not trigger if user interaction is involved in opening links, we were able to attach JavaScript code to the `onclick` event handler of hyperlinks. This allows us to perform n-fold traffic selling every time a user clicks on a hyperlink on the web site.

We performed an experiment that shows that this attack is effective against traffic brokers: We signed up as traffic seller at two different traffic brokers and bought visitors from a third broker. Each click on our web site was redirected to both brokers. We implemented only a 2-fold click inflation attack to test the setup in practice, but higher values for n can be implemented without problems. In total, we bought 17,000 visitors to our site, from which more than 1,800 visitors clicked at least one link and thus we could sell them to both brokers. In total, the visitors generated about 4,100 clicks. During our experiment, we successfully accumulated funds of slightly less than $10 (on average, we received $2.22 for 1,000 sold visitors). To prevent damage to the traffic broker, we did not withdraw any funds, but forfeited our traffic trading accounts.

Based on the insights gained from this experiment, we think that other, even more powerful types of click fraud (such as *clickjacking* [12]) would be equally easy to employ. The success of this experiment also suggests that traffic brokers do not share information among each other about traffic sold, and that no advanced fraud detection systems are in place.

5 Related Work

Little academic information is available about the online adult industry, yet it consists of thousands of web sites that generate a revenue of billions of dollars every year. Many publications that analyze general web security issues have

been published in recent years. For example, Wang et al. developed client honeypots to detect and capture web-based malware samples [27].

Existing work on web-based threats often targets specific types of malware. For example, Moshchuk et al. provide an analysis of web-based spyware [19]. Provos et al. focus on analyzing technical exploitation details [21]. Zhuge et al. studied malicious aspects of the Chinese Web [30]. The authors did mention the adult industry, however, the scope of their work is limited to drive-by downloads found on adult web sites and no other aspects were studied.

Several studies show parallels and draw connections between malicious Internet activity and the underground economy. For example, Provos et al. provide technical details on how cyber-criminals use web-based malware to their advantage [22]. The aspect of an underground economy that is fuelled by financially motivated cyber-criminals is highlighted by Franklin et al. [9]. In a recent paper, Holz et al. study the structure and profits of keyloggers [13].

To the best of our knowledge, this study is the first that combines an economic analysis of the online adult industry with a security analysis from a technical and a cyber-crime perspective.

6 Conclusion

In this paper, we presented novel insights into the online adult industry. We analyzed the economic structure of this industry, and found that apart from the expected "core business" of adult sites, more shady business models exist in parallel. Our evaluation shows that many adult web sites try to mislead and manipulate their visitors, with the intent of generating revenue. To this end, a wide range of questionable techniques are employed, and openly offered as business-to-business services. The tricks that these web sites employ range from simple obfuscation techniques such as relatively harmless *blind links*, over convenience services for typo-squatters, to sophisticated redirector chains that are used for traffic trading. Additionally, the used techniques have the potential to be exploited in more harmful ways, for example by facilitating CSRF attacks or click-fraud.

By becoming adult web site operators ourselves, we gained additional insights on unique security aspects in this domain. For example, we discovered that a malicious operator could infect more than 20,000 with a minimal investment of about $160. We conclude that many participants of this industry have business models that are based on very questionable practices that could very well be abused for malicious activities and conducting cyber-crime. In fact, we found evidence that this kind of abuse is already happening in the wild.

References

[1] Adobe Systems Incorporated. Adobe Flash Player. `http://www.adobe.com/de/products/flashplayer/`, 2009.

[2] Alexa. Top 500 Global Sites. `http://www.alexa.com/topsites`, 2009.

[3] V. Anupam, A. Mayer, K. Nissim, B. Pinkas, and M. K. Reiter. On the Security of Pay-per-click and Other Web Advertising Schemes. In *Proceedings of the Eighth Conference on World Wide Web (WWW)*, 1999.

[4] B. Stone-Gross and M. Cova and L. Cavallaro and B. Gilbert and M. Szydlowski and R. Kemmerer and C. Kruegel and G. Vigna. Your Botnet is My Botnet: Analysis of a Botnet Takeover. In *ACM Conference on Computer and Communications Security (CCS)*, 2009.

[5] U. Bayer, P. M. Comparetti, C. Hlauschek, C. Kruegel, and E. Kirda. Scalable, Behavior-Based Malware Clustering. In *Symposium on Network and Distributed System Security (NDSS)*, 2009.

[6] Beano Publishing. Domain Players Club. `http://www.domainplayersclub.com`, 2009.

[7] N. Daswani and M. Stoppelman. The Anatomy of Clickbot.A. In *First Workshop on Hot Topics in Understanding Botnets (HotBots)*, 2007.

[8] Finjan Inc. LuckySploit Toolkit Exposed. `http://www.finjan.com/MCRCblog.aspx?EntryId=2213`, 2009.

[9] J. Franklin, V. Paxson, S. Savage, and A. Perrig. An inquiry into the nature and causes of the wealth of internet miscreants. In *ACM Conference on Computer and Communications Security (CCS)*, 2007.

[10] Guywire, Inc. Booble. Adult Search Engine. `http://www.booble.com`, 2009.

[11] M. Hammami, Y. Chahir, and L. Chen. Webguard: A web filtering engine combining textual, structural, and visual content-based analysis. *IEEE Transactions on Knowledge and Data Engineering*, 18(2), 2006.

[12] R. Hansen and J. Grossman. Clickjacking. Technical report, SecTheory – `http://www.sectheory.com/clickjacking.htm`, 2008.

[13] T. Holz, M. Engelberth, and F. Freiling. Learning More About the Underground Economy: A Case-Study of Keyloggers and Dropzones. In *European Symposium on Research in Computer Security (ESORICS)*, 2009.

[14] Internet Filter. Internet Pornography Statistics. `http://internet-filter-review.toptenreviews.com/internet-pornography-statistics.html`, 2006.

[15] M. Jakobsson, P. Finn, and N. Johnson. Why and How to Perform Fraud Experiments. *Security & Privacy, IEEE*, 6(2):66–68, March-April 2008.

[16] M. Jakobsson and J. Ratkiewicz. Designing ethical phishing experiments: a study of (ROT13) rOnl query features. In *15th International Conference on World Wide Web (WWW)*, 2006.

[17] M. Cova and C. Kruegel and G. Vigna. Detection and Analysis of Drive-by Download Attacks and Malicious JavaScript

Code. In *19th International World Wide Web Conference (WWW2010)*, 2010. http://wepawet.iseclab.org.

[18] M. Cova and S. Ford. Wepawet: Detecting and Analyzing Web-Based Malware. http://wepawet.iseclab.org, 2009.

[19] A. Moshchuk, T. Bragin, S. D. Gribble, and H. M. Levy. A crawler-based study of spyware on the web. In *Symposium on Network and Distributed System Security (NDSS)*, 2006.

[20] Network Working Group. WHOIS Protocol Specification. http://tools.ietf.org/html/rfc3912, 2004.

[21] N. Provos, P. Mavrommatis, M. Abu Rajab, and F. Monrose. All Your iFRAMEs Point to Us. In *17th Usenix Security Symposium*, 2008.

[22] N. Provos, D. McNamee, P. Mavrommatis, K. Wang, and N. Modadugu. The Ghost In The Browser. In *First Workshop on Hot Topics in Understanding Botnets (HotBots)*, 2007.

[23] Spam Assassin. List of performed Tests. http://spamassassin.apache.org/tests_3_2_x.html, last accessed: 23.04.2009, 2009.

[24] Symantec Corporation. Misleading Applications. http://www.symantec.com/connect/blogs/misleading-applications-show-me-money-part-3, 2009.

[25] The Honeynet Project. Capture-HPC Client Honeypot. https://projects.honeynet.org/capture-hpc, 2009.

[26] Trusteer, Inc. Flash Security Hole Advisory. http://www.trusteer.com/files/Flash_Security_Hole_Advisory.pdf, 2009.

[27] Y.-M. Wang, D. Beck, X. Jiang, R. Roussev, C. Verbowski, S. Chen, and S. King. Automated Web Patrol with Strider HoneyMonkeys. In *Symposium on Network and Distributed System Security (NDSS)*, 2006.

[28] Y.-M. Wang, D. Beck, J. Wang, C. Verbowski, and B. Daniels. Strider Typo-Patrol: Discovery and Analysis of Systematic Typo-Squatting. In *2nd Conference on Steps to Reducing Unwanted Traffic on the Internet*, 2006.

[29] XBIZ. The Adult Industry Source for Business News and Information. http://www.xbiz.com, 2009.

[30] J. Zhuge, T. Holz, C. Song, J. Guo, X. Han, and W. Zou. Studying Malicious Websites and the Underground Economy on the Chinese Web . In *Proceedings of 2008 Workshop on the Economics of Information Security (WEIS'08)*, June 2008.