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

Gilbert Wondracek, Thorsten Holz, Christian Platzer
(Vienna University of Technology)
Engin Kirda (Institut Eurecom)
Christopher Kruegel (UC Santa Barbara)

# Motivation



- Online adult industry is big business – Sex sells!
  - Even low estimates are in the range of *billions of dollars*
  - 12% of all websites porn, 70% of men under 24 browse porn
  - 35% of all downloads, 8% of email are related to porn
  - 3,000$ / second spent on porn, conservatives spend most
- Few academic publications exist on subject
- Common sense and anecdotal evidence tell us that adult websites are more dangerous to web surfers
  - True? What means dangerous exactly?

# Where to Start?



- No reliable data from this industry available for research
  - Finding volunteers was no problem
- Manual analysis
  - We looked at 700 adult websites
    - Structural similarities between individual websites?
    - Allowed us to gain first insights and to infer a basic model on how the industry works
  - Most sites *look* different, but structurally, we found a relatively small number different types
- We identified two main types of adult websites (from a consumer's POV)

# Core Business / Website Types



- *Paysites*
  - Websites that offer lots of pornographic media
  - Often dedicated to (sexual) niche markets
  - Access to content is generally restricted to members (sign-up on website and pay a fee for access)
  - Often produce their own content, host affiliate programs, offer promotional content to business partners
- *Free sites*
  - Basically, link collections with pornographic content
  - Link to paysites or other free sites
  - Affiliate programs: Free sites get pornographic material from paysites, in return, they link to paysites → a commission is paid if users sign up

# B2B Website Archetypes

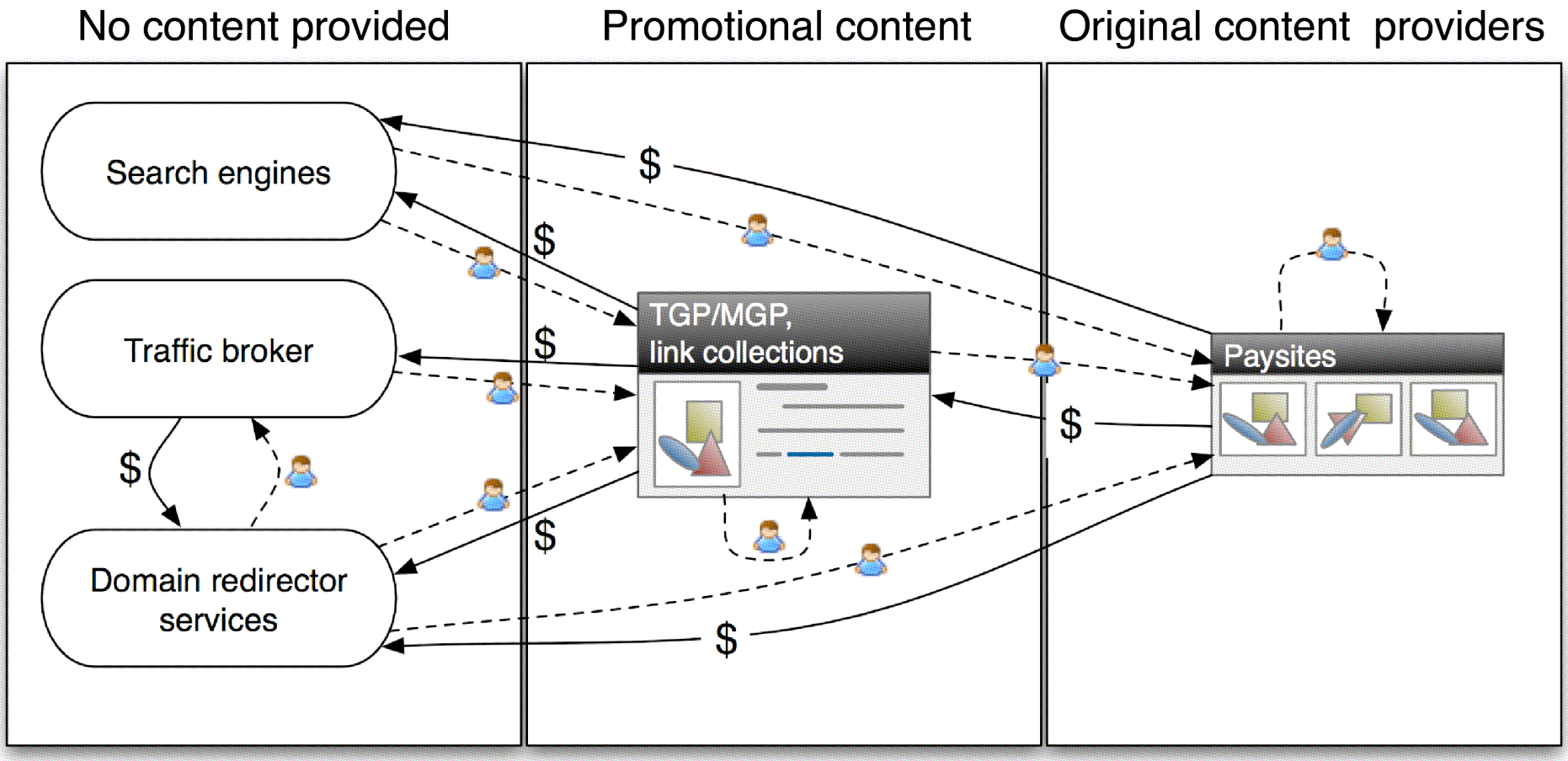


- *Traffic brokers*
  - Buy / Sell visitors, according to specific criteria
    - Country of origin, Type of sexual preference, “Is known to click often”, Likes videos / images, ....
  - Typically: 2-3$ for 1,000 visitors (buy)
  - Traffic comes from websites that redirect visitors
- See paper for more roles / details
  - Adult search engines
  - Redirection services / facilitation of domain and typo squatting

  ...

# Traffic and Money





# Getting More Data



- Manual analysis is nice, but we want more data
- Automate the data acquisition with a crawling system
- Use search engines to get crawling seeds
  - Idea is to mimic potential clients / victims
  - Three general purpose (Google, Yahoo, Bing), ten adult search engines
  - Used domain-specific search queries (*not* listed in paper)
  - Extracted the top results
- We crawled 269,000 URLs from 35,000 domains
  - Result: For each URL, we stored the source code and extracted the hyperlinks

# Crawler Data Evaluation



- We first performed a classification of the economic roles to label and compare paysites and free sites
  - Difference between free sites or paysites?
- Then, we checked the crawled websites for common scenarios that website operators use to trick users
  - Based on what we saw during manual analysis
- To detect malware, we checked each URLs with two client honeypots
  - Systems that detect unintended changes to file system, registry, ...

# Economic Classification



- Paysite detection
  - We compiled a list of (adult) payment processors
  - If a site links to a payment processor, we label it a paysite, if not, we look for member sections or sign-up forms
- Free site detection
  - Hyperlink topology reveals economic role
  - Many outgoing links to “foreign” domains? If yes, free site assumed (Whois entries used for ownership)
- Classification Results
  - 35,000 domains, 87.7% classified (12.3% undefined)
  - For these: 8.1% paysites, 91.9% free sites
  - Confirms assumptions from role description

# Shady Business Practices



- We frequently found these three methods of tricking users
- JavaScript catchers
  - Keep users from leaving the website with dialogs, pop-ups,...
  - On each click, send them somewhere else, load something in the background, display new ads → traffic multiplication scheme
- Blind links
  - Don't show link destinations in browser (simple Javascript)
  - Lure user to different websites (hide hyperlinks on images)
- Redirector scripts
  - Server-side scripts determine link destination, hidden to user
  - Used to send users to different websites

# Shady Business Practices



- We created detectors for each of these tricks
  - Analyzed website source code, Javascript, look for code fragments
  - Resolve suspicious links several times to find redirection targets

| | Free Sites | Paysites |
|---|---|---|
| JS Catcher | 3.9% | 1.2% |
| Blind Links | 26.2% | 10.9% |
| Redirectors | 23.6% | 3.2% |

  - Underlines the motto of (shady) free sites: bring traffic to paysites or traffic brokers at all costs

# Infectious Stuff



- Malware detection with client honeypots
  - About 269,000 URLs checked
  - **3.23%** trigger malicious behavior, trojans, bots, adware, ...
  - This is more than five times as much as expected
- Source code hints at compromised sites
  - Found traces of website mass exploits
  - >98% of `iframe` sources hosted somewhere else
  - Owners do not notice / care?

# Why? Some Ideas



- Traffic trading: Tricking users to go to certain websites means money for the source site
- “Skimming”: Instead of sending users to the affiliate program, send them somewhere else, for example:
  - Site F attracts users with promo material from site P, but redirects them to another site if it is more profitable (traffic broker)
  - SEO: Free sites (competitors) link to each other, boost search engine rankings → free sites profit, paysite loses
- Bad guy's perspective: Adult sites seem to be an ideal vector for distributing malware
  - Adult website visitors don't know where they will really end up when they click on links or images → They still click a lot
  - All that is needed is a porn site and visitors...

# The Insider Job



- How to get data that only site operators would have?
- We created two adult websites
  - Free sites: One link collection and one TGP type adult site
  - Problem: No content (pictures, videos)
- Signed up to eight affiliate programs
  - No real verification of site ownership etc. was necessary
  - 6 of 8 didn't even access our sites
  - Gave us access to lots of pornographic material
- Whole setup (excluding website programming) took only a few days
  - Ready for traffic trading – we signed up at 3 traffic brokers

# Traffic Buying Experiment

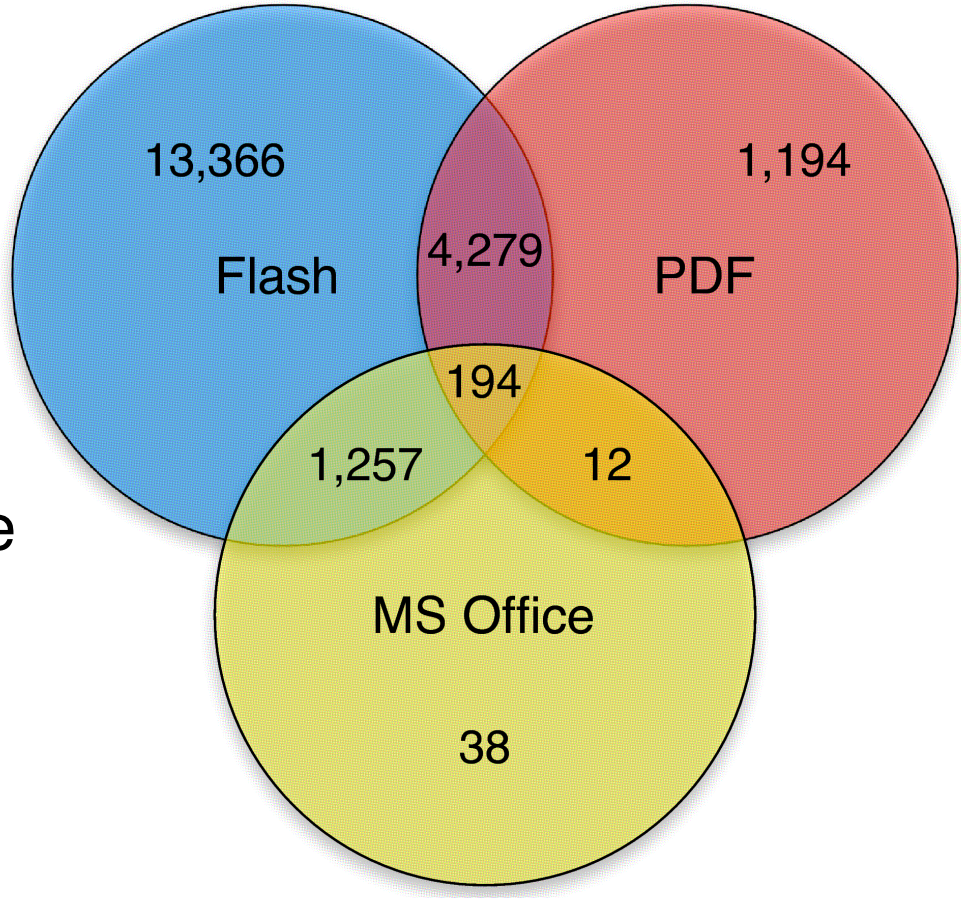


- We spent about **$160** to buy adult traffic from three different brokers for our websites
  - Resulted in about **49,000 visitors** to our websites
- Delivery was always instant, and orders were executed accurately
- To see the security threat to these users, we ran some Javascript code and Flash animations
- We used this data to identify vulnerable visitors
  - Collected data checked against a database of popular vulnerable software, browser plugins, etc.
  - Only worked as intended for 49% of these visitors (closed site immediately, JavaScript disabled, PS3, Wii, bots?)

# Vulnerability Results



- More than **20,000 visitors vulnerable** to at least one **known** exploit (LB)
- 160$ for a medium sized botnet?
- Could also be abused for PPI installs (install malware for a commission)
  - 130$ / 1,000 US installs
  - Fits with choosing visitor criteria when buying traffic...



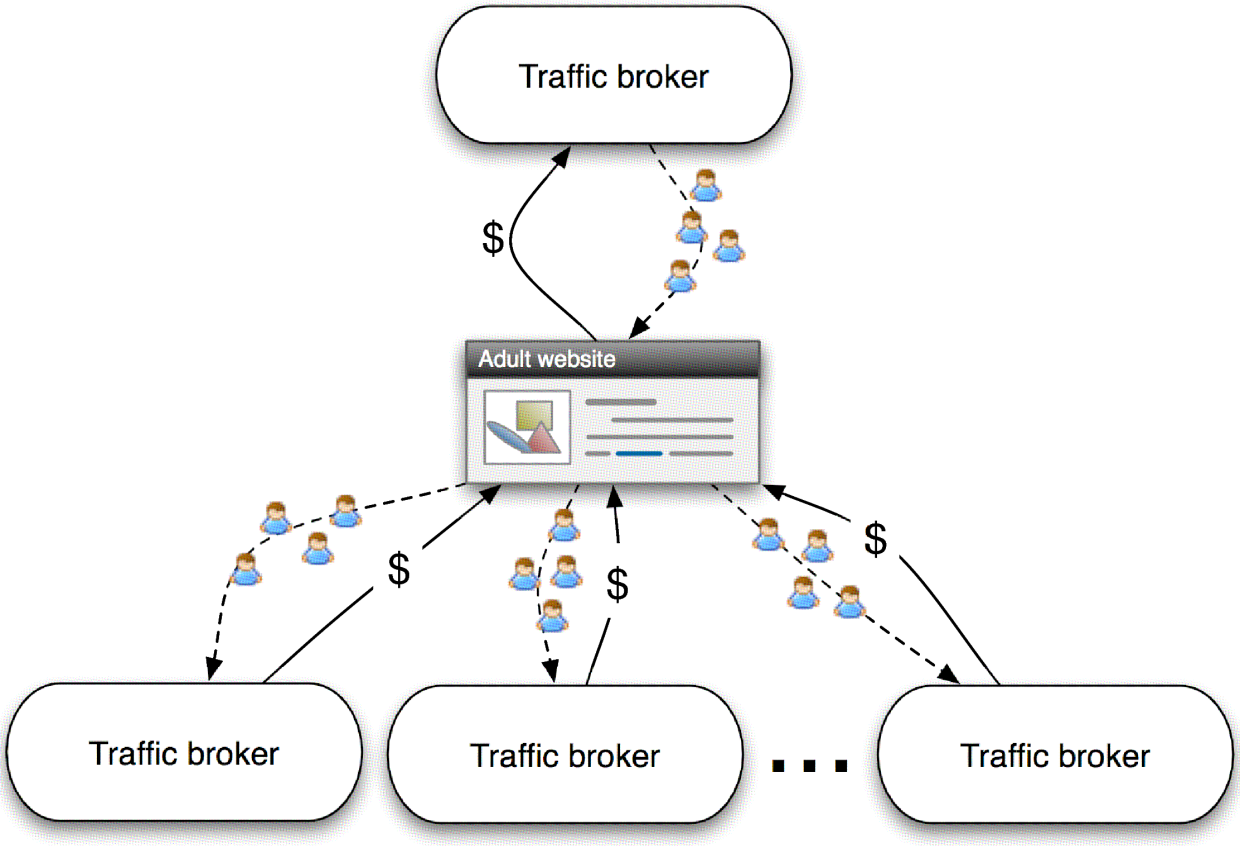
# Traffic Selling Experiment



- Traffic selling: Redirect users to traffic brokers, get money in return
- No questions asked
  - Embedding of scripts or other content was not necessary
  - Used by online advertising to detect fraud
- So, are there *any* security / anti-fraud measures in place?
  - We came up with a simple fraud scenario to check this

# Click Inflation Fraud Scenario





# Click Inflation Results



- We managed to use this technique to accumulate about 10$ in total earnings
  - We never withdrew any money and forfeited our accounts
  - Real criminals could install malware in addition to committing fraud against traffic brokers
  - Clickjacking scenarios also possible
  - We assume, that traffic brokers do not share data to prevent fraud
- Shows lack of technical sophistication
  - Security awareness?

# Summary



- We performed several experiments on adult websites, and created two websites to study traffic trading in this domain
- Basic outcome: No, not all porn sites are bad or dishonest - but many are
  - Still, more dangerous than other types of websites
- Shady services exist and are an accepted part of the industry
- Lack of security measures / checks & controls
  - Intention or just a big mess?
- Cyber-crime can use adult websites as cheap and effective vehicles for malware and fraud
- Traffic trading is scary

# Thank you!

# Ethical Considerations



- Anonymized collected data
- We did not withdraw any generated funds
- No content stored
- Automated access to websites was limited to prevent resource monopolization
- Contacted our university's law department (IRB equivalent)

# Crawling



- Custom crawler with domain-specific extensions
- Heuristics to improve crawling performance
  - Enter page detection (stay in adult domain)
  - Adult / Non-adult site classifier (limit scope of crawling)
- We crawled 269K URLs from 35K domains (sites)
  - For each URL, we stored the source code and extracted the hyperlinks
- These results are the input data used for our evaluation
  - What kind of threats can adult site visitors expect?

# Client Honeypots



- Virtual machines that browse the web and pretend to be a regular “user”
- Software detects changes to system
  - Files being written to disk, Registry changes, Unusual network traffic, ...
- Allows us to find websites that trigger malicious activity (e.g. drive-by downloads)
  - Match behavioral profile against malware database
- We used two different honeypots
  - Some malware detects honeypots and stays dormant
  - Flash and PDF vulnerabilities can be found too