



# What is an untrustworthy supply chain costing the US digital advertising industry?

IAB US benchmarking study

November 2015



**EY**

Building a better  
working world

## Table of contents

Study background 1

Cost impact summary 2

Key findings 3

Infringed content 4

Malvertising+ 16

Invalid traffic 26

# Study background

In 2015, the Interactive Advertising Bureau (IAB) commissioned EY's Media & Entertainment Advisory practice to perform a comprehensive study that estimated the cost impact of an untrustworthy digital advertising supply chain in the US. MediaLink, a strategic consulting firm, assisted the IAB in organizing and administering the study.

A supply chain is a complex economic system of people, processes and resources from different companies involved in moving a product from the start of the system through the delivery to the consumer. For the purposes of this study, we considered the digital advertising supply chain (i.e., moving an advertising creative through the internet until it reaches a consumer's browser) and the digital media supply chain (moving content through the internet until it reaches a consumer's browser).

The IAB wanted to better understand the impact of deliberate activities designed to exploit the current state of the supply chain for illicit gain. It also wanted to know more about the repercussions of unintentional activities by businesses that have put digital advertising as a legitimate business in jeopardy.

EY conducted part one of the study between March and September 2015. This included areas that have a high degree of illegal activity – infringed content, malvertising and invalid traffic. We will undertake part two of the study in early 2016, when we will focus on media transparency, reputational impact and brand safety.

Based on the results of phase one of the survey, EY has:

1. Identified areas of corruption in the digital advertising supply chain
2. Estimated the commercial cost impact to the ecosystem

The research methodology for the study included:

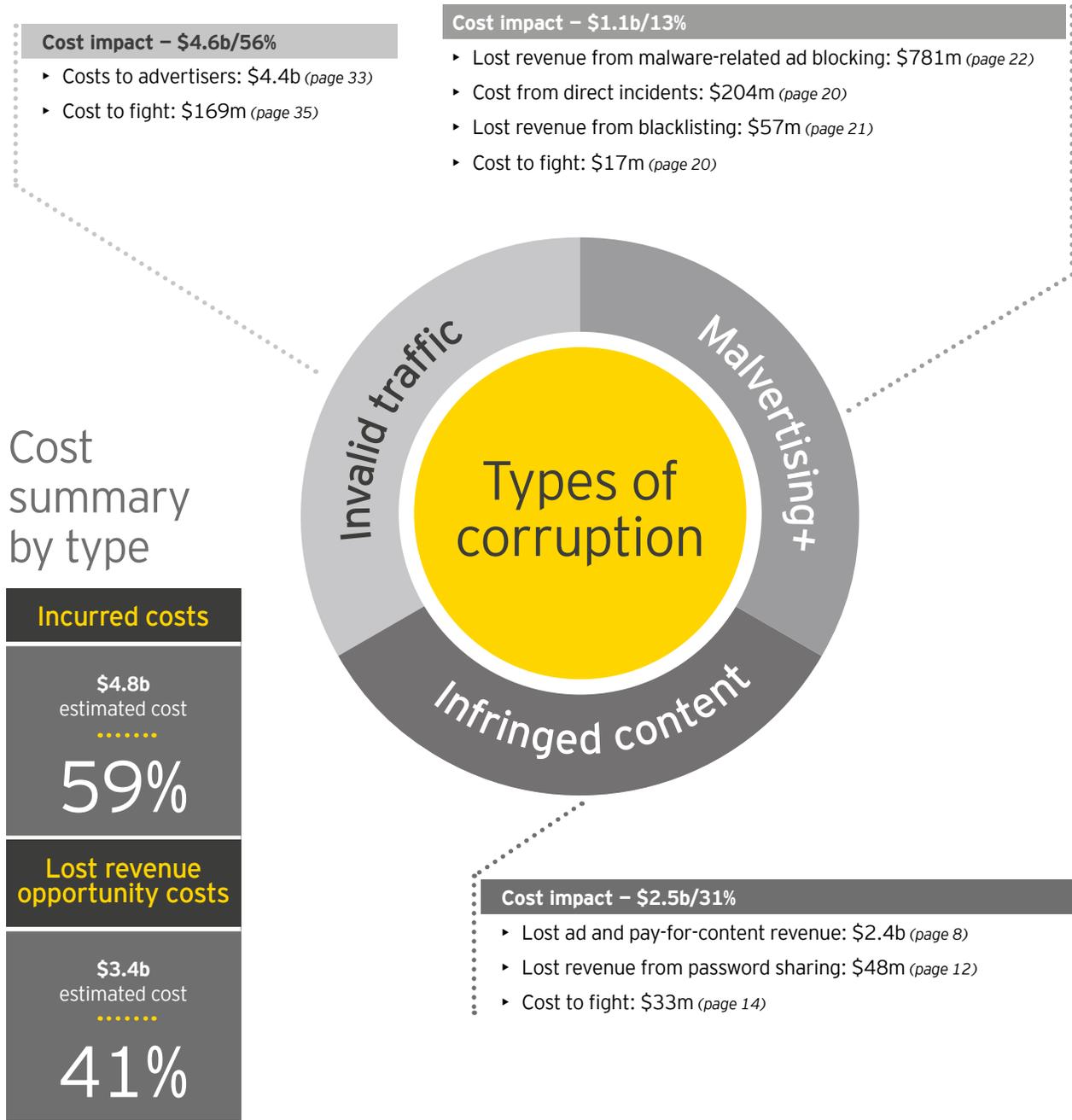
- ▶ **Study of studies** – We assessed a number of studies that other organizations have conducted in relevant supply chain areas over the last several years. We have referenced key reports to estimate certain costs.
- ▶ **Voice of the industry** – We selected and contacted 90 supply chain companies, including both publishers and ad tech companies, to complete a comprehensive 13-page questionnaire that included qualitative and quantitative areas. Of the 90 contacted, 30 companies completed the questionnaire.
- ▶ **Data analytics** – For several specific areas, we obtained data directly from third-party measurement and analytic organizations to estimate certain costs.

EY conducted this study independently on behalf of the IAB. EY did not audit the information provided to us and provides no opinion or other forms of assurance with respect to the report's findings.

Finally, we wish to say thank you and express our gratitude to the following IAB sponsors of this study:

Premier	Supporting	Participating
AppNexus	PubMatic	OpenX
MediaMath	Xaxis	PulsePoint
Rocket Fuel Inc.	YuMe	Videology

# Cost impact summary – \$8.2b\*



Note: The page numbers above contain a detailed explanation of our estimation approaches.

\* All amounts are in US dollars.

# Key findings

- ▶ Each studied category has an estimated cost impact above \$1 billion. Individually, they represent significant costs to the industry that should not be ignored. **However, as each category can be interrelated, they need to be considered collectively and equally when being addressed by the industry.** An excellent example is a consumer who visits an infringed content site containing malware that infects the consumer's browser with a robot that is later used to drive invalid traffic. If the industry can eliminate the profits earned by serving ads next to infringed content, it can reduce the amount of money available to drive illegal activities in the supply chain. It also has the opportunity to disrupt the corruption life cycle related to invalid traffic. To help the industry reclaim some of the \$8.2 billion in costs, EY believes that an improvement in some fundamental practices, such as knowing your business partners and investigating new relationships using address information, tax IDs and background checks, is critical.
- ▶ At \$4.4 billion, costs to advertisers from invalid traffic represent the most significant portion of incurred costs. In terms of distribution, 70% of the costs relate to performance-based pricing models, such as cost-per-click (CPC) and 30% relate to cost per month (CPM) based pricing models' costs. Related to consumer consumption, currently 72% of the costs are from desktop and 28% are from mobile. We also noted a range of rates (e.g., CPM-based mobile video has a 12.1% invalid traffic rate while CPM-based display desktop has a rate of 6.6%). **As the digital advertising industry continues to be dynamic related to pricing models, consumer consumption by delivery platforms and pricing by ad units, assessing the invalid traffic costs to advertisers, should holistically consider the rapid changes to business and fraud approaches.**
- ▶ At \$2.4 billion, infringed content represents the most significant portion of lost revenue opportunity costs. One key feature that drives consumers to infringed content is the desire and ability to access recently distributed content at no direct cost in the convenience of their homes. It is hard to say what the impact would be to distribution channels if access were eliminated. Would consumers turn to ad-supported or pay-for-content channels? How many would actually become paying customers? There's no conclusive way of knowing. However, our approach suggests a potential advertising revenue increase of \$456 million and a potential pay-for-content revenue increase of \$2 billion for the industry. The \$2 billion represents approximately 21 million US consumers who would be willing to spend \$8 a month on what is currently classified as infringed content. **Unless the industry collectively takes significant steps, there is a likelihood that the number of infringed content consumers will continue to increase.** Improving technology and bandwidth that make it easier for consumers to obtain content, aids to protect the anonymity of users, and an increasing culture of moral acceptance by consumers are all contributing factors. At the same time, it is becoming increasingly difficult for consumers to determine whether content is truly infringed. And even if they can tell the difference, they have a diminishing fear of legal repercussions.
- ▶ The remaining areas representing 16% of the total are estimated at \$1.4 billion. These areas include the cost to fight illegal activities, lost revenue from password sharing, lost revenue from search engine blacklisting when a website is impacted by malware and lost revenue from malware-related ad blocking.

“The industry needs to deal with the problem effectively and the fraud needs to be put to its death.”

Bob Liodice, Association of National Advertisers President and CEO, interviewed by Beet.tv, 26 February 2015.



# Infringed content

# Infringed content landscape

## Ad injection

This is a toolbar or adware that alters the site HTML prior to the browser rendering a served impression.

**Primary revenue:** Advertising  
**Major content:** Display content



## P2P community

A peer-to-peer (P2P) community allows users to browse for files on websites linking to content hosted by other connected computers or servers via a peer-to-peer distribution system.

**Primary revenue:** Advertising  
**Ancillary revenue:** Donations  
**Major content:** Music, movies, software, games, text and TV programs



## Storefront community

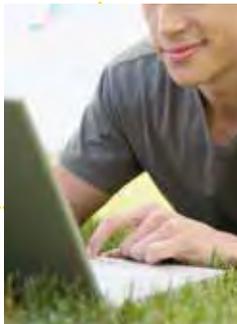
In a storefront community, users can purchase and download digital media from the site's own servers.

**Primary revenue:** Transactions  
**Ancillary revenue:** Advertising  
**Major content:** Music, movies, software, games, text and TV programs

## Subscription community

Subscription communities allow users to browse for files on websites linking to content hosted by other connected computers or servers via a P2P distribution system.

**Primary revenue:** Advertising  
**Ancillary revenue:** Donations  
**Major content:** Music, movies, software, games, text and TV programs



## Freemium community

Freemium communities give users access to P2P links or direct downloads of curated digital media content for free. It also enables them to pay or contribute content to the site for additional content access and/or quality.

**Primary revenue:** Subscriptions  
**Ancillary revenue:** Advertising and donations  
**Major content:** Music, movies, software, games, text and TV programs

## Embedded streaming

Embedded streaming offers a hosting site where users can upload and directly stream video content.

**Primary revenue:** Advertising  
**Ancillary revenue:** Donations and subscriptions  
**Major content:** Music, movies and TV programs



## Live TV streaming

Live TV streaming provides links to direct streams of live free-to-air and pay-per-view TV (including sporting events).

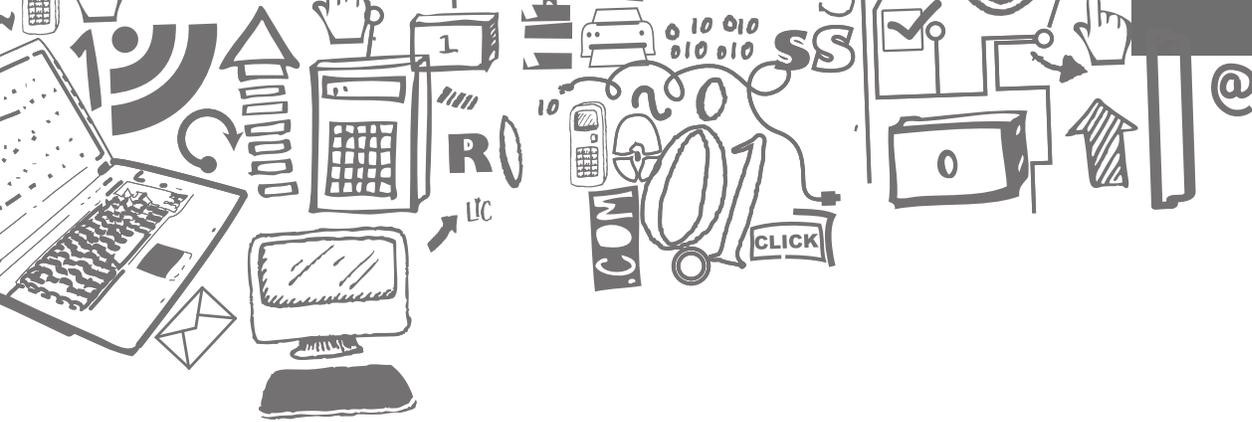
**Primary revenue:** Advertising  
**Ancillary revenue:** Donations  
**Major content:** Live TV



## VPN and proxy piracy

Virtual private network (VPN) and proxy piracy enables users to access content illegally by bypassing geolocation licensing restrictions. There is likely no direct revenue to criminals. However, it could impact geo-targeting and measurement.

**Major content:** Movies and TV programs



## Comprehensive description

Online digital piracy is the illegal practice of using the internet (via mobile, PC or other device) to access infringed content via websites and peer-to-peer networks. Content may include videos, live events, music, video games, text, software and applications. From a business perspective, copyright infringement operators generate revenue through advertising, subscriptions, donations and transactions.

### Broad digital infringed content categories include:<sup>1</sup>

- ▶ **Ad injection.** This is a toolbar or adware that alters the site HTML prior to the browser rendering a served impression without permission or compensation to the website or content owner.
- ▶ **Embedded streaming.** Embedded streaming offers a hosting site where users can upload and directly stream video content. Generally, these sites are financially supported by digital advertising, subscriptions and donations.
- ▶ **Freemium community.** Freemium communities give users access to P2P links or direct downloads of curated digital media content for free. They also enable users to pay or contribute content to the site for additional content access and/or quality.
- ▶ **Live TV streaming.** Live TV streaming provides links to direct streams of live free-to-air and pay-per-view TV, including sporting events. These sites are largely ad-supported, although some also accept donations from users for financial support.

- ▶ **P2P community.** This allows users to browse for files on websites that link to content hosted by other connected computers or servers via a peer-to-peer distribution system. Users can generally download the desired content files for free as the communities are largely ad-supported, although some also accept donations from users for financial support. The industry should pay special attention to newer infringed content distribution platforms that combine P2P and streaming characteristics and allow users to access video and music content using a clean and legitimate-looking application. These platforms differ from traditional P2P platforms in that they stream as components are delivered by other P2P participants rather than assembling a chosen file first and then storing it on a user's PC hard drive.
- ▶ **Storefront community.** In a storefront community, users can purchase and download digital media from the site's own servers. These sites are generally ad-supported or fee-based.
- ▶ **Subscription community.** Subscription communities allow users to browse for files on websites linking to content hosted by other connected computers or servers via a P2P distribution system. They provide links or direct downloads of curated digital media content (with the exception of free-to-air and pay TV) typically for a subscription fee. Alternatively, these communities may be ad-supported.
- ▶ **VPN and proxy piracy.** VPN and proxy piracy enables users to access content illegally by giving global users access to certain US-based digital content illegally (e.g., video content from OTT services) by paying the subscription fee and then bypassing geolocation restrictions. This corruption area is exacerbated by complicated video licensing agreements in which a legal OTT service will have a different content library depending on the country.

Additionally, there is the issue of password sharing between family members or individuals who are resource pooling to expand their content access. VPN and proxy piracy can sometimes be combined with password sharing to create a different level of corruption (e.g., a user living in another country uses a US VPN to access an OTT service by obtaining a password from someone who lives in the US).

1. *The six business models for copyright infringement – A data-driven study of websites considered to be infringing copyright*, a Google and PRS for Music commissioned report with research conducted by BAE Systems Detica, 27 June 2012.

# Infringed content

## Key drivers of infringed content

Factors that impact the growth of infringed content consumption can be divided into two groups:

- ▶ Infringed content site factors:
  - ▶ Profitable with a low cost of entry
  - ▶ Well-organized business models that new entrants can easily replicate
  - ▶ An expanding digital universe that brings more potential consumers who may use the infringed content market for some of their own content
  - ▶ Improving bandwidth that makes it easier for certain infringed content approaches
  - ▶ A growing acceptance of infringed content use by consumers
  - ▶ An increasing number of support-oriented companies that provide tools to users
- ▶ Infringed content user factors:
  - ▶ Increasing demand for content without waiting for a release or the next episode
  - ▶ Desire for lower content costs
  - ▶ Desire to access content remotely
  - ▶ Improving technology and bandwidth making it easier to obtain infringed content
  - ▶ Resistance to paying for content with advertisements
  - ▶ Easy availability of tools that protect the anonymity of infringed content users
  - ▶ An increasing culture of moral acceptance built on years of receiving a high degree of free content in other areas
  - ▶ Difficulty identifying whether content is truly infringed
  - ▶ Lack of fear of legal repercussions

## Industry initiatives to combat infringed content

To fight back, some current industry initiatives include:<sup>2</sup>

- ▶ Participate in the Association of National Advertisers (ANA) and American Association of Advertising Agencies' (4A's) *Statement of Best Practices to Address Online Piracy and Counterfeiting*. These leading practices recommend that marketers and their agencies include the following conditions in media placement contracts and insertion orders with ad networks and other intermediaries involved in their US-originated digital advertising campaigns on both domestic and foreign internet sites:
  - ▶ All such intermediaries shall use commercially reasonable measures to prevent ads from being placed on those sites dedicated to the infringement of the intellectual property rights of others because they have no significant, or only limited, use or purpose other than engaging in, enabling or facilitating such infringement.
  - ▶ All such intermediaries should implement commercially reasonable processes for removing or excluding such sites from their services and for expeditiously terminating noncompliant ad placements in response to reasonable and sufficiently detailed complaints or notices from rights holders and advertisers.
  - ▶ All such intermediaries should refund or credit the advertiser for the fees, costs and/or value associated with noncompliant ad placements or provide alternative remediation.
- ▶ Participate in best practices for ad networks to address piracy and counterfeiting, which recommends ad networks:
  - ▶ Maintain policies that prohibit websites dedicated to selling counterfeit goods or engaging in copyright piracy from participating in the ad network's advertising programs.
  - ▶ Maintain and post the best practices guidelines on the ad network's website.
  - ▶ Include in ad network policies language indicating that websites should not engage in violations of law.

2. *Statement of Best Practices to Address Online Piracy and Counterfeiting*, The Association of National Advertisers (ANA) and the American Association of Advertising Agencies (4A's), 3 May 2012.



# Infringed content

## Cost impact to industry

### Infringed content segmentation

There are four main types of digital infringed content sites: 1) direct download (DDL) sites; 2) linking sites; 3) P2P sites; and 4) video streaming host sites.

In terms of estimating the cost impact, we obtained usage data (e.g., number of downloads, unique visitors or unique IPs) from three different sources and then applied certain rate data (e.g., CPMs or monthly pay-for-content costs) under the two principal revenue models (i.e., ad revenue model and a pay-for-content model). For the monthly pay-for-content cost, we used \$8. This represents the additional revenue obtained if infringed content was 100% eliminated. The additional revenue could come from monthly streaming service, direct downloads of music or videos from an online store, purchased video on demand, an additional cable box adapter or a ticket to the movie theater.

Often, consumers are attracted to infringed content distribution channels because of the immediate access to recently distributed media, such as a new movie, song or television series. We used multiple sources and approaches to triangulate the cost impact range to the industry. Our goal was to estimate the potential revenue that could be earned if this content usage data was consumed at legal sites as opposed to infringed sites. For example, DDL sites are generally subscription-based with minimal advertisements. As a result, we treated the DDL content as a legitimate channel and assumed some banner and video ad impression activity.

**Estimation approach 1:** We utilized the following May 2015 usage metrics obtained directly from an analytics company that measures websites for purposes of identifying content infringement related to movies, television and music:

DDL	Linking sites	P2P	Video streaming
16,371,716	36,020,713	18,111,399	12,454,597

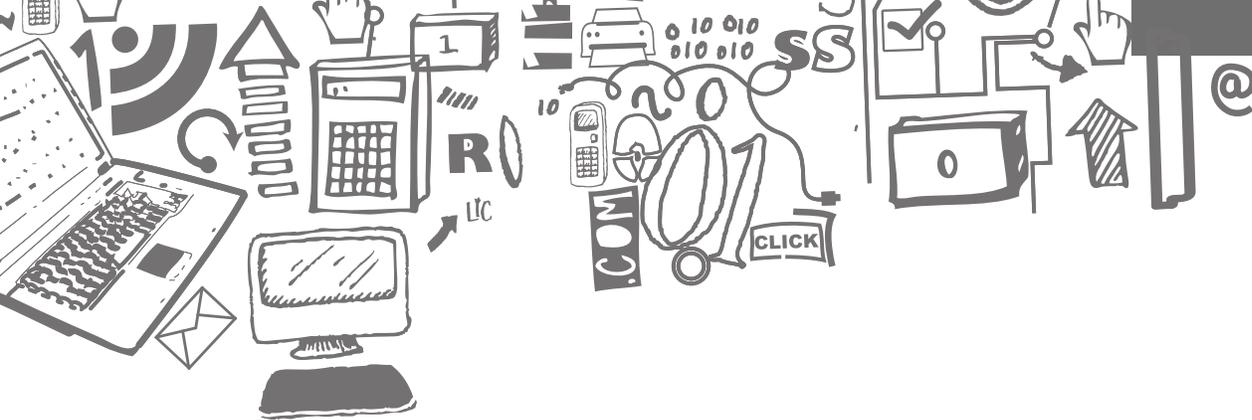
The analytics company calculated the usage (surrogate for a monthly unique visitor reach) metrics above by multiplying a monthly global reach estimate (per million users across all sites in each category) by the estimated number of global internet users of 3,188,000,000 by the percentage of US users divided by 100.

We multiplied the usage data above by an estimated monthly cost of \$8 by 12 months (replicating the annual revenue from a pay-for-content revenue model) to calculate the following:

DDL	\$1,571,684,736
Linking	\$3,457,988,448
P2P	\$1,738,694,304
Video streaming	\$1,195,641,312
Total estimated revenue	\$7,964,008,800

The analytics company also provided the following May 2015 estimated visits. These were defined as an entry to a web domain from a different web domain or from the beginning of an empty browsing session which expires after 30 minutes of inactivity:

DDL	Linking sites	P2P	Video streaming
229,659,429	826,319,617	189,906,569	117,802,924



We calculated the banner ad revenue below assuming an \$11.35 CPM<sup>4</sup> for 12 months (to annualize the May 2015 data obtained) for one viewed impression per visit (a conservative assumption). The video/audio ad impression revenue estimated below assumes: a) a \$21.28 CPM<sup>5</sup> for each visitor (uses the usage metric above) to a content hosting site; b) an average usage of four times a week for 12 months; c) three ad units viewed per half hour; and d) an average of a full hour of consumption based on the content (e.g., movies, television programming).

	Banner ad revenue on sites	Video/audio ad revenue on sites	Total ad revenue
DDL	\$31,279,614	\$100,336,354	\$131,615,968
Linking	\$112,544,732	\$220,757,983	\$333,302,715
P2P	\$25,865,275	\$110,998,244	\$136,863,519
Video streaming	\$16,044,758	\$76,329,741	\$92,374,499
<b>Total estimated ad revenue</b>	<b>\$185,734,379</b>	<b>\$508,422,322</b>	<b>\$694,156,701</b>

**Estimation approach 2:** We used the following monthly average unique visitor data based on Q3 2014 from a publicly available study performed by MediaLink on behalf of the Digital Citizens Alliance. The study, *Good Money Still Going Bad: Digital Thieves and the Hijacking of the Online Ad Business*, issued in April 2015, estimated that the top infringing sites (the top 596 infringing sites were measured) in 2014 earned \$209 million in advertising revenue.<sup>6</sup>

MediaLink performed an initial study in 2013 for the same organization. It estimated that the major sites hosting infringing content earned an estimated \$227 million in

advertising revenue that year.<sup>7</sup> The goal of the MediaLink studies was to estimate the revenue and profit earned at the top infringing content study. For this study, we used the volume metrics and then assumed commensurate CPMs were earned for the content involved (i.e., normal CPMs for quality content as opposed to low-dollar CPMs, which the infringing content sites typically earn).

DDL	Linking sites	P2P	Video streaming
14,909,600	9,819,500	21,311,700	9,324,000

We multiplied the monthly unique visitor data above to an estimated monthly cost of \$8 by 12 months (replicating the annual revenue from a pay-for-content revenue model) to calculate the following:

DDL	\$1,431,321,600
Linking	\$942,672,000
P2P	\$2,045,923,200
Video streaming	\$895,104,000
<b>Total estimated revenue</b>	<b>\$5,315,020,800</b>

The study also provided the following estimated monthly average page views for sites:

DDL	Linking sites	P2P	Streaming
191,600,000	182,800,000	383,800,000	190,100,000

The banner ad revenue below was calculated assuming an \$11.35 CPM<sup>8</sup> for 12 months (to annualize the monthly average Q3 2014 data obtained) for one viewed impression per visit (a conservative assumption). The video/audio ad impression

4. IAB internet advertising revenue report: 2014 full year results – April 2015, IAB, [http://www.iab.com/wp-content/uploads/2015/05/IAB\\_Internet\\_Advertising\\_Revenue\\_FY\\_2014.pdf](http://www.iab.com/wp-content/uploads/2015/05/IAB_Internet_Advertising_Revenue_FY_2014.pdf), accessed November 2015.

5. Ibid.

6. *Good Money Still Going Bad: Digital Thieves and the Hijacking of the Online Ad Business*, Digital Citizens Alliance, <https://media.gractions.com/314A5A5A9ABBBBC5E3BD824CF47C46EF4B9D3A76/298a8ec6-ceb0-4543-bb0a-edc80b63f511.pdf>, accessed November 2015.

7. Ibid.

8. IAB internet advertising revenue report: 2014 full year results – April 2015, IAB, [http://www.iab.com/wp-content/uploads/2015/05/IAB\\_Internet\\_Advertising\\_Revenue\\_FY\\_2014.pdf](http://www.iab.com/wp-content/uploads/2015/05/IAB_Internet_Advertising_Revenue_FY_2014.pdf), accessed November 2015.

# Infringed content

revenue estimated below assumes (a) a \$21.28 CPM<sup>9</sup> for each visitor (uses the usage metric above) to a content hosting site; (b) the usage is an average of four times a week for 12 months; (c) there are three ad units viewed per half hour; and (d) an average of a full hour of consumption based on the content.

	Banner ad revenue on sites	Video/audio ad revenue on sites	Total ad revenue
DDL	\$26,095,920	\$91,375,571	\$117,471,491
Linking	\$24,897,360	\$60,180,180	\$85,077,540
P2P	\$52,273,560	\$130,611,737	\$182,885,297
Video streaming	\$25,891,620	\$57,143,439	\$83,035,059
<b>Total estimated ad revenue</b>	<b>\$129,158,460</b>	<b>\$339,310,927</b>	<b>\$468,469,387</b>

**Estimation approach 3:** We used the following annual usage data from a P2P measurement company for 2014:

	Movies	Music	TV
P2P downloads	805,700,000	308,900,000	522,600,000
Unique IPs	55,000,000	38,000,000	29,900,000

We divided the 2014 unique IP estimate by a factor of 5.18 (estimate of IP addresses used in the US compared to the US population) and multiplied this factor by a cost of \$8 by 12 months (replicating the annual revenue from a pay-for-content revenue model) to calculate the following:<sup>10,11</sup>

Music revenue	\$704,914,286
Video revenue	\$1,018,656,371
<b>Total estimated revenue</b>	<b>\$1,723,570,657</b>

We calculated the banner ad revenue below assuming an \$11.35 CPM<sup>12</sup> for one viewed impression related to each download (a conservative assumption). The video/audio ad impression revenue estimated below assumes a \$21.28 CPM<sup>13</sup> be applied to the P2P downloads considering a likely number of spots for the media type (e.g., three ad units for a 30-minute TV show, 12 ad units for a 2-hour movie and six ad units for an hour of audio play). For example, the \$33,362,912 estimated video ad impression revenue for TV was calculated by dividing 522,600,000 downloads by 1,000 and then multiplying it by a CPM of \$21.28 per ad unit and then multiplying it by three spots per hour.

	Banner ad revenue on sites	Video/audio ad revenue on sites	Total ad revenue
Movies	\$9,144,320	\$205,735,125	\$214,879,446
Music	\$3,506,163	\$39,442,012	\$42,948,174
TV programming	\$5,931,533	\$33,362,912	\$39,294,445
<b>Total estimated ad revenue</b>	<b>\$18,582,016</b>	<b>\$278,540,049</b>	<b>\$297,122,065</b>

9. Ibid.

10. *Regional Internet Registries Number of IP Addresses Per Country*, BGP Expert, [www.bgpexpert.com/addressespercountry.php](http://www.bgpexpert.com/addressespercountry.php), accessed September 2015.

11. *Internet Usage and 2015 Population in North America*, Internet World Stats, [www.internetworldstats.com/stats14.htm](http://www.internetworldstats.com/stats14.htm), accessed September 2015.

12. *IAB internet advertising revenue report: 2014 full year results – April 2015*, IAB, [http://www.iab.com/wp-content/uploads/2015/05/IAB\\_Internet\\_Advertising\\_Revenue\\_FY\\_2014.pdf](http://www.iab.com/wp-content/uploads/2015/05/IAB_Internet_Advertising_Revenue_FY_2014.pdf), accessed November 2015.

13. Ibid.



# Infringed content

## VPN piracy and password sharing

Consumers are able to illegally access digital content through password sharing. In some cases, this action is compounded when consumers bypass their actual geolocation by using a virtual VPN located in another geolocation. A negative side effect of VPN usage is the accuracy impact to some passive digital measurement approaches, as well as country-based digital ad targeting. Absent a change to complex content agreements, the corruption impact is likely to grow as servers become more accessible, bandwidth strength increases and global internet access penetration increases.

To estimate the cost impact, we used publicly available studies or certain estimates quoted publicly. According to research issued by GlobalWebIndex in the first quarter of 2015 (32 countries were measured):<sup>14</sup>

- ▶ 51% of users cited access to better entertainment content as the number one reason for VPN usage. Many of the other reasons related to anonymity and accessing restricted sites; however, the 7th overall reason at 22% was to access restricted download sites such as torrent sites (which are generally used to obtain infringed content).
- ▶ The highest percentage of users of VPN/proxy servers at 35% live in Latin America. EY considers these estimates relevant to the infringed content assessment in the US media market because of the growing number of people migrating from Latin America to the US. It is possible that some family members remain behind and can access content remotely using a VPN and a shared password.
- ▶ There are approximately 28 million VPN server users in the US. This puts the US in a tie with Brazil for third place in terms of VPN server users. Only China at 157 million and India at 45 million have higher numbers of users. EY considers these estimates relevant to infringed content assessment in the US media market because these individuals tend to use these servers to access torrent sites to obtain infringed content.

- ▶ VPN server users skew younger (27% of ages 16 to 24 and 36% of ages 25 to 34 vs. 11% of ages 55 to 64), male (31% of males vs. 21% of females) and upper income (38% of top quarter of income group vs. 27% of the bottom quarter of income group).

According to a recent research report from Parks Associates, the practice of password sharing will cost the subscription video-on-demand (SVOD) industry more than \$500 million worldwide in 2015. Six percent of US broadband subscribers indicated they access a subscription OTT video service paid for by someone outside their home.<sup>15</sup>

To estimate the cost impact of password sharing to the SVOD industry in the US, we considered the following:

- ▶ Consumer price elasticity is not known (e.g., quantity demand decreases as price increases, and it is not known how many consumers would become a paying customer if they are currently accessing content for free).
- ▶ OTT services likely already consider password sharing when establishing their pricing strategy (e.g., monthly fee can increase based on the number of concurrent streams).

As a result, we conservatively applied a 9.58% factor (this factor represents the approximate percentage of people connected to the internet who live in the US) to the \$500 million from the Parks Associates global estimate to calculate an estimated rounded cost impact of \$48 million for the US only.<sup>16</sup>

EY was unable to obtain an estimate on the impact of VPN and proxy pirates.

14. Jason Mander and Felim McGrath, "VPNs and Proxy Servers," *GlobalWebIndex*, <http://www.globalwebindex.net/>, accessed November 2015.

15. *OTT Password Sharing Will Impact Pay-TV Network Revenue, Too*, Parks Associates, [www.parksassociates.com/blog/article/ott-password-sharing-will-impact-pay-tv-network-revenue](http://www.parksassociates.com/blog/article/ott-password-sharing-will-impact-pay-tv-network-revenue), accessed September 2015.

16. *Internet Users by Country (2014)*, Internet Live Stats, <http://www.internetlivestats.com/internet-users-by-country/>, accessed September 2015.



# Infringed content

Furthermore, the study results for the ad tech respondents indicated that their attitudes and initiatives were moving in the right direction to combat infringed content:

- ▶ 100% indicated that their organization includes commitments in their contracts not to purchase inventory on sites with infringed content.
- ▶ 99% said that their organization performs measures aimed at excluding sites with infringed content in response to reasonable and sufficiently detailed complaints from copyright holders and advertisers.
- ▶ 81% noted that their organization performs measures specifically aimed at removing or excluding sites with infringed content from platforms that use fraud or deception to avoid the requirements set by the advertiser or agency.
- ▶ 79% indicated that their organization's leadership is against advertisements being served to sites with infringed content.
- ▶ 79% said that they have witnessed their staff discussing concerns regarding advertisements appearing on sites with infringed content.
- ▶ 42% suggested that their organization has designated an individual or role responsible for mitigating risk of ads appearing on sites with infringed content.

“Internet usage continues to grow at a rapid pace, and with it, so does internet-based infringement.”

David Price, NetNames Director of Piracy Analysis, “Sizing the piracy universe,” *NetNames*, September 2013.





# Malvertising+

# Malvertising+ threat landscape

## Malvertising+ Delivery methods

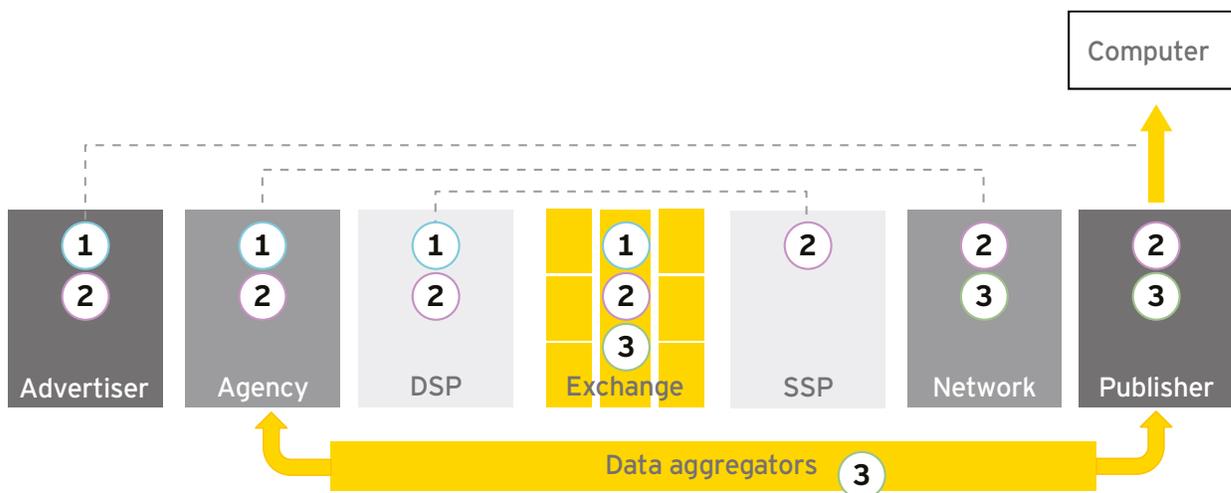
- ▶ **Deceptive download:** tricked into download
- ▶ **Drive-by download:** unintended software download
- ▶ **Link hijacking:** redirection to unintended site
- ▶ **Watering hole:** targeted drive-by download

## Uses and purpose

- ▶ **Ad bot creation:** uses infected machine for ad fraud
- ▶ **Nuisanceware:** adds unwanted features
- ▶ **Ransomware:** alters system until payment is made
- ▶ **Scareware:** scares user to pay for unneeded "fix"
- ▶ **Spyware:** collects consumer activity without consent
- ▶ **Virus/infection:** has ability to use consumers' device

## How an attacker views the landscape

- 1 Creating fraudulent content (fraudulent ads, fraudulent agencies, drive-by download ads, deceptive download ads, bad scripts, spyware, ransomware, scareware, viruses)
- 2 Altering good content (code injection, link hijacking, repository compromise)
- 3 Content mining (cookie hijacking, watering hole targeting)





# Malvertising+

- ▶ A **virus** infects the user's device and takes over part or all of its functions for malicious purposes (e.g., relay spam, harm computers, steal data and monitor activity).
- ▶ **Ad bot** creation uses the infected machine as a bot for impression and click fraud.

## **The following are ways malvertising+ can be established:**

- ▶ Creating fake advertisers or advertising agencies that pretend to represent legitimate clients in a buy.
- ▶ Gaining access to a library of affiliate marketing ad creatives and hijacking them to deliver malware.
- ▶ Compromising third-party scripts delivered with the ad or page content that are intended for measurement or related purposes.
- ▶ Attaching malware to a selected download that appears legitimate to the consumer.

For several of the methods listed above, nefarious malware attackers use malware delivery kits (available for relatively inexpensive price via the web) and advertising to infect a consumer's device.

## **Malvertising+ is able to exist and prosper in an ecosystem for many reasons, including:**

- ▶ Not prioritizing security within the creative quality assurance function, or having insufficient tools and resources to fight it.
- ▶ A single weak link in the digital advertising ecosystem that can be compromised to inject malware.
- ▶ Traditional PC defenses like antivirus and other tools are unable to determine in time whether a compromised third-party script or advertisement, such as a Flash-powered banner ad (which is not defined as malicious itself), is simply serving ad content or something more sinister.
- ▶ Attackers who use tactics to slip past the filtering systems. This may include: 1) enabling the malicious trigger after a delay of several days following the approval of the ad; 2) only serving the bad ad or script to every nth consumer; 3) targeting to, or away from, specific consumers based on identifier information such as IP address, operating system, browser and other parameters; 4) leveraging programmatic

or real-time bidding systems to further target consumers with specific operating systems, browser versions, Flash versions, geographic locations, or IP addresses that may indicate residential, university or corporate users, and potentially the institution or business the consumer is within; 5) launching attacks on weekends or holidays when it is likely ad operations personnel are away from the office or will take longer to respond to malware attacks; and 6) embedding malware in HTML headers, and steganographically embedding malicious code fragments in image and SWF files that are linked together to form an attack string at run time.

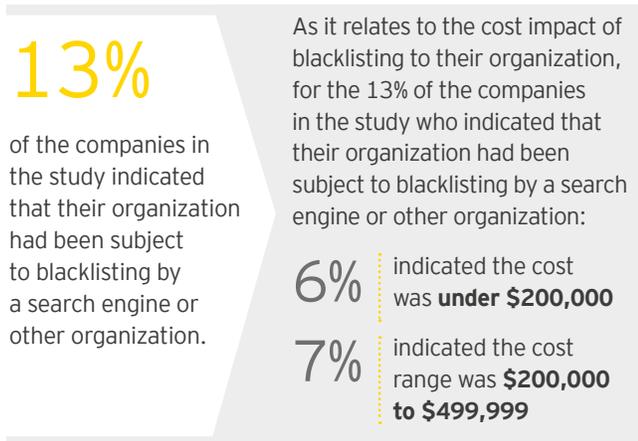
## **To fight back, some key preventive measures include:**

- ▶ Using ad-serving tools and controls that can scan the creative to detect and disable injected or unintended code (i.e., malware) before allowing ads to launch. This would entail scanning Flash or JavaScript files, either manually or by using sites that provide malware scanning tools. Companies should run these analyses on systems outside of their system to prevent infection of their internal systems and to prevent the identification of the environment as a test environment in which the malware should remain hidden and dormant.
- ▶ Evaluating business partners, including advertisers, agencies and third parties with whom companies work (background checks, credit checks, etc.), to determine if they are reputable and legitimate companies.
- ▶ Assessing third-party tech partners' diligence regarding their evaluation of business partners, internal security framework, and quality assurance over ad content and scripts received from partners.
- ▶ Identifying and closing holes on sites or internal systems.
- ▶ Reporting business partners involved in the ad-serving transaction that handle the ad content or provide third-party scripts in support of the transaction to the advertiser and agency.
- ▶ Finding a way for the good actors in the industry to share information to help reduce the level of malvertising+.



# Malvertising+

As it relates to blacklisting, our voice of the industry study noted the following:



“Visits to mainstream websites can expose consumers to hundreds of unknown or potentially dangerous third parties.”

“Online Advertising and Hidden Hazards to Consumer Security and Data Privacy,” Permanent Subcommittee on Investigations Majority and Minority Staff Report, United States Senate, 15 May 2014, <https://www.hsgac.senate.gov/media/permanent-subcommittee-on-investigations-releases-report-online-advertising-and-hidden-hazards-to-consumer-security-and-data-privacy>-accessed November 2015.

For purposes of estimating the cost impact of blacklisting, our study was interested in legitimate websites whose businesses were impacted by a malware security incident. During 2014, a US-based nonprofit anti-malware organization received 29,000 requests from websites (the direct request increases the likelihood that these represented legitimate sites) impacted by blacklisting requesting the organization to review the website and delist the site from these blacklists. The vast majority was cleaned within two days without assistance, which could represent cleanup or the malware only existed for a short period of time; however, approximately 2,000 requested a manual inspection by the organization, which is a strong indicator that they were not free of malware. To assign a cost related to blacklisting, we considered that the majority of these sites represented small businesses (further supported by our voice of the industry study, where only 13% of the IAB members had been impacted), and according to the IDC, the average annual revenue of a small business with a website, when adjusted for inflation, is \$6.35 million or \$17,386 per day.<sup>20</sup> We conservatively selected a 50% negative impact or \$8,693 per day for blacklisting over an average of two days (this period was used because most incidents were addressed within two days; however, some took longer), which results in an estimated cost of \$504,194,000 related to the total impact of malware blacklisting. Because the organization estimated 10% or less of the cases were due to malvertising, we calculated a range impact of zero to \$50,419,400 with a midpoint of \$25,209,700 (used in our estimate below to be conservative).

We also weighted and projected our voice of the industry data to calculate the impact to the larger organizations within the digital advertising ecosystem and estimated a cost of \$31,325,000. As a result, the total rounded cost impact is estimated at \$57,000,000.

20. *Small Business at a Glance*, Entrepreneur.com, <http://www.entrepreneur.com/page/216022>, accessed September 2015; EY analysis.



# Malvertising+

As it relates to ad blocking, our voice of the industry study noted the following:

**49%** of the companies in the study indicated that they measure the level of ad blocking at their websites or via their platforms for the ad technology companies.

For those measuring the ad blocking

**87%**

indicated the level was less than 10%

**2%**

indicated the level was 10% to 20%

**11%**

indicated the level was 20% or greater

As it relates to the cost impact of ad blocking to their organization

**72%**

indicated the cost was under \$200,000

**8%**

indicated a range of \$200,000 to \$499,000

**12%**

indicated a range of \$500,000 to \$999,000

**8%**

indicated a range of \$1,000,000 or more

For purposes of estimating the cost impact of ad blocking related to malvertising+, we calculated an estimated cost of \$781,000,000 as follows:

- ▶ Ad revenue generated per person not blocking was \$209.09 based on 2014 digital ad spend of \$49.5 billion divided by 236,739,760 (279,834,232 US digital population multiplied by 84.6% of the US internet population estimated to be not blocking ads).<sup>22</sup>
- ▶ Missed ad revenue was estimated at \$9,025,447,009 based on \$209.09 multiplied by 43,165,369 ad blockers (279,834,232 multiplied by 15.4% of the US population estimated to be ad blocking).
- ▶ As 17% of the PageFair respondents attributed the reason for ad blocking to privacy (directly related to security and malware), we calculated an estimate of \$1,534,325,991.
- ▶ We also weighted and projected our voice of the industry data to estimate an overall ad-blocking cost of \$157,675,000 and then applied the 17% factor from the PageFair study to estimate a cost of \$26,804,750 (ad blocking associated with malvertising+).
- ▶ The \$781,000,000 estimated rounded cost was based on the midpoint between \$26,804,750 and \$1,534,325,991.

“Ad blocking is beginning to have a material impact on publisher revenues.”

Mike Zaneis, CEO Trustworthy Accountability Group, “Publishers and adblockers are in a battle for online advertising,” *FT.com*, 29 March 2015, <http://www.ft.com/intl/cms/s/2/c84a647e-d3af-11e4-99bd-00144feab7de.html#axzz3rmaUjreu>, accessed November 2015.

22. IAB internet advertising revenue report: 2014 full year results – April 2015, IAB, [http://www.iab.com/wp-content/uploads/2015/05/IAB\\_Internet\\_Advertising\\_Revenue\\_FY\\_2014.pdf](http://www.iab.com/wp-content/uploads/2015/05/IAB_Internet_Advertising_Revenue_FY_2014.pdf), accessed November 2015.







# Invalid traffic

# Invalid traffic landscape

Invalid traffic can enter the ecosystem in several ways, and for several purposes, including:

Ad traffic is typically designed to deliver the right ad at the right time to the right user. Fraudulent invalid traffic generates ad-related actions to extract the maximum amount of money from the digital advertising ecosystem, regardless of the presence of an audience. Legitimate invalid traffic generates actions in the normal course of internet maintenance by non-human actors: search engine spiders, brand safety bots and competitive intelligence gathering tools.

## Audience extension

Audience extension increases inventory by selling the inventory of third parties as if it belongs to the site, incentivizing the content partner to increase traffic (and thereby revenue), which may ultimately include a downstream partner sending invalid traffic.

## Traffic sourcing

Traffic sourcing increases inventory through payments to third parties to drive traffic to the site, which may ultimately include a downstream partner sending invalid traffic.

## Cookie enrichment

This approach generates invalid activity on valid and reputable sites to build a cookie profile of increased value within targeted buying systems, and then visits fake websites to achieve higher CPMs for the ads delivered to the site.

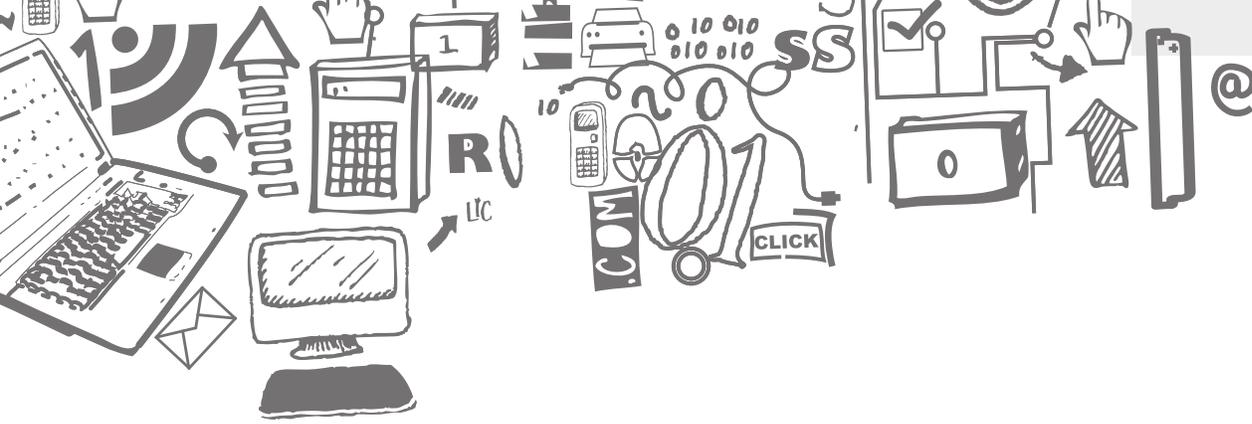
## Click fraud

Click fraud generates invalid click activity to illegitimately increase cost-per-click (CPC) revenue earned (network click fraud) or drive competitor marketing costs (competitor click fraud), which is more commonly present within the search-based advertising ecosystem.

## Illegitimate websites

Illegitimate ad-supported websites generate ad impressions using invalid traffic to collect revenues from advertisers.

The above methods affect searches, displays, videos, audio, mobile (web and in-application) and social.



## Comprehensive description

Invalid traffic (IVT) induces systems to generate ad-related actions for purposes other than support of the delivery of the right ad at the right time to the right user. This includes actions occurring across the ecosystem, which impact the search, display, video, mobile, audio and social areas. IVT may take the form of legitimate activity, as well as activity generated by bad actors for fraudulent purposes.

- ▶ **Fraudulent IVT** activity typically extracts the maximum amount of money from the digital advertising ecosystem, regardless of the presence of an audience.
- ▶ **Legitimate IVT** tends to generate actions during the normal course of internet maintenance by non-human actors, including actions executed by search engine spiders, brand safety bots and competitive intelligence gathering tools.

The Media Rating Council (MRC) further defines IVT in terms of the methods by which IVT may be detected:<sup>23</sup>

- ▶ **General IVT** is traffic identified through routine means of filtration. Key examples include data center traffic; bots and spiders or other crawlers masquerading as legitimate users; non-browser user-agent headers; hidden/stacked/covered or otherwise never-viewable ad serving, pre-fetch or browser pre-rendering traffic; and invalid proxy traffic.
- ▶ **Sophisticated IVT** is more difficult to detect and requires advanced analytics, multipoint corroboration/coordination or significant human intervention, etc., to analyze and identify. Key examples include: hijacked devices, hijacked tags, adware, malware, incentivized browsing, misappropriated content (if applicable), falsified viewable impression decisions and cookie stuffing.

IVT does not in any way represent legitimate traffic. As such, it is difficult to identify and prevent its monetization. Current studies vary widely in dimensioning the true impact of IVT. However, the general consensus is that IVT has a material cost impact. Impacts may include: depressed inventory CPMs and a reluctance to invest and allocate digital media spend; damaged reputation to organizations susceptible to exposure to fraudulent IVT; and the overall cost to fight.

With the rise of automation and ever-increasing complexity of the digital supply chain, the prevalence of IVT is expected to persist. Fraudulent IVT in this environment is exacerbated in ad transactions involving unknown sources, such as publishers purchasing low-cost traffic or open ad exchanges.

In general, IVT has the potential to have a direct monetary impact to buy-side organizations. Fraudulent IVT's impact may be the result of fraudulent publisher sites selling inventory to advertisers against known robotic traffic directed to the inventory. Alternatively, bad actors may operate fraudulent publisher sites in addition to perpetrating illegitimate cookie enrichment. Through cookies, bots are directed toward reputable sites to build cookie profiles that mimic traits of desirable consumers for ad targeting. The bad actor then sells inventory on the fraudulent site against these enriched cookies at a higher CPM. In the latter scenario, the publisher's reputation may be impacted as the intermediate steps of the cookie enrichment process involve the presence of IVT across premium or otherwise reputable publisher content sites.

“The digital advertising industry must stop having unprotected sex.”

Randall Rothenberg, CEO Interactive Advertising Bureau, “IAB Head: ‘The Digital Advertising Industry Must Stop Having Unprotected Sex’” *Businessinsider.com*, <http://www.businessinsider.com/iab-randall-rothenberg-supply-chain-2014-2>, accessed November 2015.

23. *Invalid Traffic Detection and Filtration Guidelines Addendum, Draft Version 5.0 – Public Comment Version*, Media Rating Council, June 30, 2015.

# Invalid traffic

Actions taken by publishers to maximize ad revenue may also inadvertently (if not blatantly) support and encourage the proliferation of fraudulent IVT within the digital supply chain. Although sell-side organizations may not be immediately monetarily impacted by fraudulent IVT, the reputational repercussion may ultimately result in a shift in ad spend away from publishers with practices that may facilitate fraudulent IVT. One such example includes traffic sourcing, whereby publishers sell more inventory than currently available. They subsequently seek out third-party publishers to purchase additional traffic to drive the audience toward sold inventory to fulfill the ad buy. In these situations, the third party may likewise seek additional third parties to fulfill the audience demands of the first-party publisher. In these situations, third-party sources may resort to using bot traffic to generate the necessary volume to meet inventory demands. The initial intent of the first-party publisher may not have been to perpetrate fraud in these situations. However, the environment of the ad buy transaction and third-party relationships increases the difficulty of maintaining transparency and accountability related to the quality of the audience fulfilling the ad buy.

A similar example regarding publisher-driven (potentially) fraudulent IVT relates to the practice of audience extension. In these situations, a publisher may represent to sell inventory under the publisher's ownership, but ultimately fulfill the ad buy through inventory placed on other sites owned by the publisher, affiliate sites or third parties. Although many of these transactions are conducted through legitimate means when the site placement of the sold inventory is transparent to the advertiser, lack of transparency in these transactions may lead to the serving of ads outside of the audience target of the media plan.



## Types of IVT

The following are additional examples of the specific sources of IVT present within the digital supply chain.

Impression/click/search impact (CPM, CPV and CPC impact)	
Non-human or illegitimate traffic sources	
<b>Hijacked device</b>	A user's device (browser, phone, app or other system) is modified to request HTML or make ad requests that are not under the control of a user and made without the user's consent.
<b>Crawler masquerading as a legitimate user</b>	A browser, server or app makes page-load requests automatically without declaring itself as a robot. Instead, the robot declares itself as a valid regular browser or app user agent where there is no real human user. In addition, robots can be programmed to mimic human behavior to develop a highly desirable profile that will incentivize a targeted ad campaign to serve an ad to that robot.
<b>Data-center traffic</b>	Traffic originates from servers in data centers, rather than residential or corporate networks, where the ad is not rendered in a user's device (there is no real human user).
<b>Adware traffic/ad injection</b>	A device where a user is present and additional HTML or ad requests are made by the adware independently of the content being requested by the user. Adware may also contain a function to inject an ad from the software onto a webpage as the user browses, rather than the ad being delivered by the publisher of the webpage.
<b>Proxy traffic</b>	Traffic is routed through an intermediary proxy device or network where the ad is rendered in a user's device where there is a real human user.
<b>Non-browser user-agent header</b>	A device declares a user-agent header not normally associated with human activity.
<b>Browser pre-rendering</b>	A device makes HTML or ad requests ahead of specific human-initiated navigation to the requested resources, for example, the process by which the Safari browser creates thumbnails for its new tab page.
Tag hijacking	
<b>Ad tag hijacking</b>	Ad tags are taken from a publisher's site and onto another site without the publisher's knowledge.
<b>Creative hijacking</b>	Creative tags are taken from a legitimately served ad so they can be rendered at a later time, without the consent of the advertiser or the contracted service provider.
Site/ad/audience attributes	
<b>Auto-refresh</b>	A page or ad unit may be enabled to request a new rendered asset more than once and at periodic intervals.
<b>Incentivized browsing</b>	A human user may be offered payment or benefits to view or interact with ads.
<b>Hidden ads</b>	Ads are placed in such a manner that they cannot ever be viewable (e.g., stacked ads, ads clipped by iframes, zero opacity ads).
<b>Misappropriated content</b>	Sites may contain copyrighted content or links to copyrighted content without the rights to monetize such content.
<b>Illegitimate sites</b>	Websites are built primarily to collect advertising revenue and offer little to no content to human audiences. These sites are often part of a network where each individual site collects a small amount of revenue to avoid suspicion.
<b>Falsely represented/domain spoofing/laundered impressions</b>	HTML or ad requests attempt to represent another site or device or other attribute, other than the actual placement. Additionally, a publisher's content management system (CMS) may be compromised when a fake page is created using a legitimate publisher's domain and markup code.
Affiliate/lead/conversion impact (CPA and CPL impact)	
Ad creative/other:	
<b>Cookie stuffing</b>	A client is provided with cookies from other domains as if the user had visited those.

# Invalid traffic

## Current response to address IVT

In response to IVT, industry participants have historically focused on standardization and developing technology that can help identify IVT within the ecosystem.

### Industry standards

Within the current digital supply chain ecosystem, commonly accepted practices to address the presence of IVT include adherence to filtration guidelines established by the industry. The MRC is expected to formally release the *Invalid Traffic Detection and Filtration Guidelines Addendum* in October 2015. The addendum establishes minimum requirements to identify and remove invalid traffic from advertising transactions.

Specifically, the addendum establishes two categories of invalid traffic. The first, "General Invalid Traffic," consists of traffic identified through routine means of filtration executed through application of lists or with other standardized parameter checks. Key examples include: known data-center traffic, bots and spiders or other crawlers masquerading as legitimate users; activity-based filtration using campaign or application data and transaction parameters from campaign or application data; non-browser user-agent headers or other forms of unknown browsers; and pre-fetch or browser pre-rendered traffic.

The second category, "Sophisticated Invalid Traffic," consists of more difficult to detect situations that require advanced analytics, multipoint corroboration/coordination or significant human intervention, etc., to analyze and identify. Key examples include: hijacked devices; hijacked sessions within hijacked devices; hijacked ad tags; hijacked creative; hidden/stacked/covered or otherwise intentionally obfuscated ad serving; invalid proxy traffic; adware; malware; incentivized manipulation of measurements; misappropriated content; falsified viewable impression decisions; falsely represented sites or impressions; cookie stuffing, recycling or harvesting; manipulation or falsification of location data or related attributes; and differentiating human and IVT traffic when originating from the same or similar source in certain closely intermingled circumstances.

The addendum also calls for organizations to maintain a business partner qualification process. The goal is to determine that upstream and downstream partners are legitimate entities, and that they themselves have similar processes to vet partners, and identify and remove invalid traffic from the transactions.

In addition to industry standards serving as guidelines to participants within the digital supply chain to detect and address IVT, all supply chain participants (publishers, ad exchanges, agencies) have a shared responsibility in this effort. Agencies should be aware of the legitimacy of the publishers to whom ads are being served and scale repatriation when impressions are identified as the result of IVT. Publishers should be aware of the risks posed to the value of their inventory and avoid practices that may incent IVT. Ad exchanges should work to detect and avoid, including IVT within sales transactions.

### Third-party vendors

To support transparency and accountability, and the need of buy-side organizations for additional intelligence regarding the activities of participants within the digital supply chain, third-party vendors have developed and marketed verification and fraud detection technologies. These technologies can validate ad delivery according to media plan, whether the ad content was ultimately viewable within a user's browser, and in certain cases support the detection of fraudulent activity. Through the availability of this data, participants within the digital supply chain gain additional tools and resources to police the ecosystem and spotlight the presence of IVT beyond what limited capabilities may have been available to services adhering to industry-standard filtration methodologies.

Verification and fraud services in particular allow advertisers to measure the risk relating to the placement of inventory to which ads are ultimately delivered. Such services identify the nature of the environments in which the advertisements are served. Using the information, verification services can typically confirm whether the ad was delivered on plan (i.e., delivered to the sites, devices, geographies or target audience), whether the environment of the publisher site may impact the prominence of the advertising (i.e., ad clutter, presence of competitor ads) or whether the content of the



# Invalid traffic

Long-tail publisher sites had a higher concentration of IVT, in comparison to premium and highly trafficked publisher sites.

- ▶ TAG will develop and enhance anti-fraud standards and protocols for all types of entities in the supply chain.
- ▶ TAG will develop tools both to identify fraudulent activity, and to better identify reputable companies in the supply chain that are not associated with fraudulent conduct.<sup>24</sup>

## Media Rating Council initiatives

The MRC has also coordinated with industry participants and trade organizations to modernize and strengthen existing industry standards to filter and disclose IVT for measurement purposes. The main tenets included within this effort focus not only on the modernization of existing guidelines to reflect the current online environment, but also on the standards to require processes to assess new IVT risks as they develop. Tenets will also consider the processes needed within organizations to understand and address the risks that other participants within the digital supply chain may introduce into the ad transaction. Lastly, the MRC's goals as they relate to this effort focus on reducing discrepancies that result from using myriad filtration methodologies across organizations and requiring responsible disclosure of the filtration methodologies an organization uses.

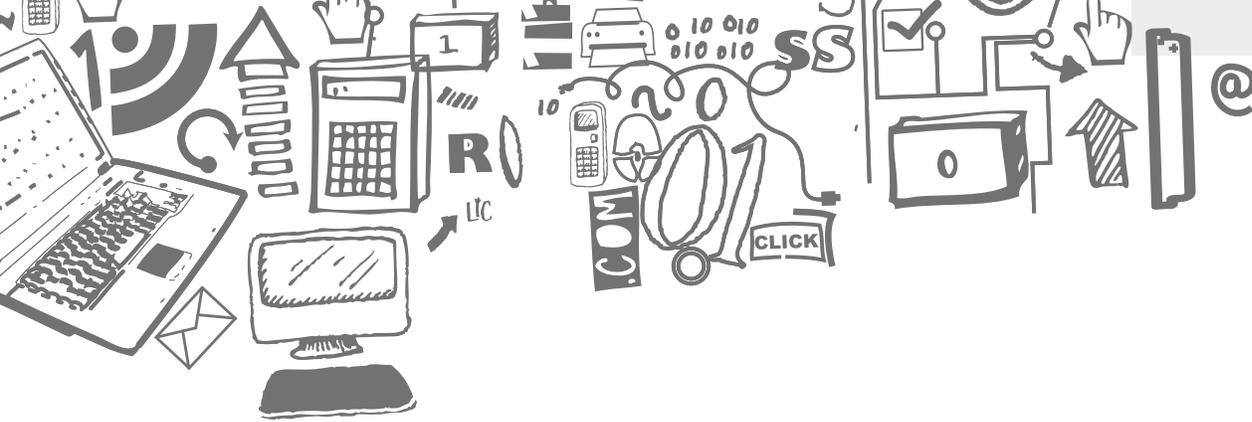
## Cost impact of IVT

### Cost from fraudulent traffic

To estimate the cost from fraudulent traffic, we used ad revenue data published within the IAB Internet Advertising Revenue report.<sup>25</sup> This helps us to define the size of US ad revenue generated across display, video and search ad formats delivered to desktop web and mobile platforms. To triangulate the impact of IVT on the digital ad ecosystem, we obtained multi-dimensional quantitative data representative of impressions transacted across each vendor's platform throughout 2015 from representative third-party measurement analytics organizations and ad serving/exchange vendors. After evaluating the data provided and adjusting for bias associated with the digital supply chain universe represented by each vendor data provider, we leveraged public research published by various IVT, fraud and analytics vendors as part of our study of studies research to develop a holistic estimation of IVT.

24. *Eliminate Fraudulent Traffic*, Trustworthy Accountability Group, <https://www.tagtoday.net/traffic/>, accessed November 2015.

25. *IAB internet advertising revenue report: 2014 full year results – April 2015*, IAB, [http://www.iab.com/wp-content/uploads/2015/05/IAB\\_Internet\\_Advertising\\_Revenue\\_FY\\_2014.pdf](http://www.iab.com/wp-content/uploads/2015/05/IAB_Internet_Advertising_Revenue_FY_2014.pdf), accessed November 2015.



Our analysis provided the following estimated percentage (invalid traffic rate applied to the 2014 revenue) and cost impact of IVT across pricing models and ad formats delivered to desktop and mobile platforms:

	Desktop		Mobile		Total	
	Percentage	Revenue	Percentage	Revenue	Percentage	Revenue
CPM-based: display	6.6%	\$500,000,000	9.8%	\$350,000,000	7.6%	\$850,000,000
CPM-based: video	11.1%	\$310,000,000	12.1%	\$160,000,000	11.4%	\$470,000,000
Performance-based	10.0%	\$2,340,000,000	10.0%	\$740,000,000	10.0%	\$3,080,000,000
<b>Total estimated IVT cost</b>	<b>9.3%</b>	<b>\$3,150,000,000</b>	<b>10.2%</b>	<b>\$1,250,000,000</b>	<b>9.6%</b>	<b>\$4,400,000,000</b>



# Invalid traffic

## Cost to fight

We weighted and projected our voice of the industry data to calculate the internal cost to fight impact for organizations within the digital advertising ecosystem and estimated a rounded cost of \$169,000,000. This cost was based on an average of 91 hours per week spent identifying, processing and analyzing invalid traffic. To project to a full year, we used a fully loaded wage hourly rate of \$62 for supervisory-level IT security practitioners in US-based organizations derived from Ponemon Institute's 2014 IT security spending tracking study.<sup>26</sup> There are also several third-party vendors that are available for hire to assist in identifying and eliminating invalid traffic for advertisers. EY was unable to estimate a cost for this initial study but will attempt to estimate a cost in any future studies.

**EY summary:** Through our study, we identified the estimated cost impact of IVT on the digital advertising supply chain to be \$4,600,000,000. This includes the costs from fraudulent traffic (\$4,400,000,000) and costs to fight associated with identifying and addressing IVT (\$169,000,000).

The majority, if not all, participants in the ecosystem, are aware that invalid traffic exists. These participants make decisions while considering these issues. If invalid traffic were to be significantly reduced or eliminated, the supply and demand relationship would change. There would be both a reduction of available inventory, and over time as confidence on the buy side improved, an increase in demand for the available inventory. While eliminating invalid traffic would not likely produce immediate material increases in CPMs, the change in the supply and demand relationship would increase CPMs over time.

Within the production impression data analyzed across our vendor participants, our research identified trends regarding the concentration of IVT consistent with many recent industry studies. These trends include:

- ▶ The inventory represented within our analyses primarily consisted of display content (>95% of impressions analyzed). However, we noted that video ad impressions contained higher concentrations of IVT in comparison to display impressions (11.4% in video versus 7.6% in display).
- ▶ IVT continues to increase in prevalence within the mobile ad ecosystem. The cost impact of IVT in mobile may continue to rise, although the levels of IVT within mobile advertising inventory may decline slightly as the levels of human traffic rise.
- ▶ The inventory represented within our analyses primarily consisted of ad network and ad-exchange-traded inventory (>75% of impressions analyzed). When assessed at a domain or sub-domain level, rates of detected IVT tended to cluster at either the low or high end of the continuum. That is, there were a number of domains and sub-domains noted with relatively low rates of IVT, as well as a number of domains and sub-domains noted with relatively high rates of IVT. Interestingly, there were relatively few domains and sub-domains noted with moderate rates of detected IVT.
- ▶ Within our analyses, we noted that IVTs were distributed similarly across the ad network and ad-exchange-traded inventory in comparison to direct publisher buys. As a reference point, however, we saw only a slight increase in the prevalence of IVT as a percentage of ad network and exchange-traded inventory, relative to direct publisher buys.
- ▶ Long-tail publisher sites had a higher concentration of IVT (greater than 4:1), in comparison to premium and highly trafficked publisher sites.

26. *The Cost of Malware Containment*, Ponemon Institute, sponsored by Damballa, January 2015.





