2018-2019

Bot Baseline

Fraud in Digital Advertising

White Ops

ANA

May 2019

Table of Contents

A Special Thank You to the Following ANA Member Company Participants

Amica
Auto Home Life

ABInBev

BANK OF AMERICA

Benjamin Moore

BISSELL

Boehringer Ingelheim

BUSH'S BEST

Chick-fil-A

CHOICE HOTELS

Coca-Cola

COLGATE-PALMOLIVE COMPANY

COMCAST

Denny's

E*TRADE

Eastern Bank
JOIN US FOR GOOD

EA

Ford

GM

HEINEKEN

HERSHEY'S

HISCOX
encourage courage

SMUCKER'S
THE J. M. SMUCKER COMPANY

JPMORGAN CHASE & CO.

Keurig Dr Pepper

Kimberly-Clark

LAND O'LAKES, INC.

LA QUINTA
BY WYNDHAM

LEGO LEGOLAND
FLORIDA RESORT

L'ORÉAL
USA

MARS WRIGLEY
confectionery

MillerCoors
A MOLSON Coors COMPANY

Nestlé

Penguin Random House

PEPSICO

REGIONS

RUST-OLEUM

State Farm

TARGET

TIAA

Unilever

UNITEDHEALTH GROUP

USAA

usbank

U.S. Cellular

VITA COCO

A Letter from the Authors

This is our fourth Bot Baseline report, and in some important ways, things are better than they have ever been. Illegitimate traffic sourcing is declining. Less sophisticated cybercriminals have abandoned their fraud schemes. Industry coalitions have formed to combat more sophisticated threats. Advertisers are spending more on channels with robust fraud protection measures. We are seeing our founding faith turn into validated fact: the battle against fraud is winnable.

However, the battle is not won yet. Ad fraud persists in new, creative forms. Every year, cybercrime becomes more advanced, evolving in response to the selection pressure of fraud detection practices. We saw this clearly with 3ve, the fraud operations that implemented sophisticated tag evasion as well as IP hijacking — where groups of IP addresses were taken over — to bypass security measures.

Though transparency issues have always plagued digital advertising, they have become more pressing than ever. In this report, we reveal the importance — and pervasiveness — of the limitations in complete third-party auditability of ad impressions.

Now is the time for advertisers to push for the ability to hold all ad impressions to the same high standard of validatability. Imagine if every CAPTCHA on the internet was the same. It wouldn't work very well. Validation with only a pixel is like serving a CAPTCHA that never changes. Only a dynamic challenge can be used to catch a dynamic adversary.[1]

Despite these challenges, we are optimistic about the future. Ads.txt, and the reduction in spoofing it achieved, has proven the power of industry-wide cooperation. Federal indictments such as the ones handed down for the 3ve and Methbot operations demonstrate that cybercriminals can be held accountable, and face real consequences, for their actions.

As an industry, we now have the momentum, if we persevere, to fix the fraud problem in a deeply meaningful way. We believe that marketers can help bring about a future where ad fraud is not profitable enough to be worth the risk.

Michael Tiffany

Co-Founder and President, White Ops

[1] A CAPTCHA is a program or system intended to distinguish human entities from those that are automated; typically this is a method to prevent spam and automated extraction of data from websites.

About the Study

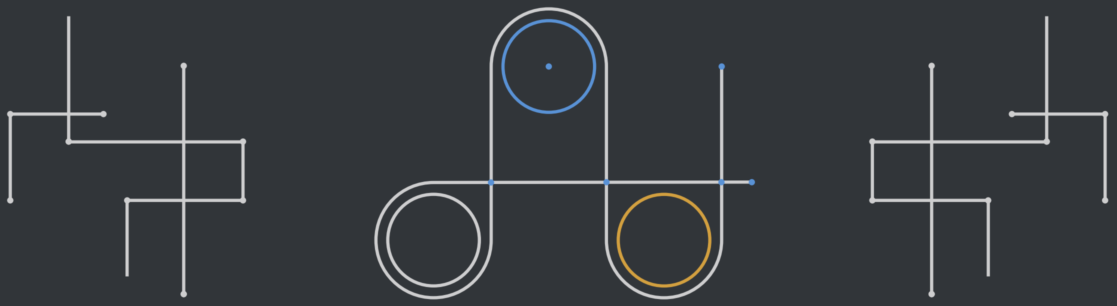
For the fourth time, White Ops and the ANA have partnered to measure bot fraud in the digital advertising ecosystem. Previous studies measured bot fraud in the digital advertising ecosystem in August/September 2014, August/September 2015, and November/December 2016.

In the latest study, 50 ANA member companies participated. White Ops worked with brand advertisers and their agencies to analyze digital advertising activity data between August 1, 2018 and September 30, 2018.

Measurements of fraud found in the global marketplace are derived from White Ops' ANA study participant data.

In this year's Bot Baseline, we share:

- **Baseline measurements of Sophisticated Invalid Traffic (SIVT),[2] excluding SIVT that was thwarted or otherwise not paid for.**
- **Initial metrics of the measurability and auditability of all the digital media paid for by study participants, illustrating the uphill battle toward full transparency.**
- **Practices related to the detection and prevention of digital ad fraud.**
- **Recommendations on best practices that will help advertisers protect budgets by ensuring their advertising dollars are not stolen by fraudsters.**



[2] General Invalid Traffic (GIVT) includes known non-human or fraudulent sources that can be identified with industry lists like the IAB Bots and Spiders List or parameter-based detection techniques. Sophisticated Invalid Traffic (SIVT) includes invalid traffic that is purpose-built to evade detection.

The Size and Scope of the Study

Of the 50 ANA member participants in the current study, 26 contributed data to previous Bot Baseline reports. Eleven have participated in all four studies, nine participated in three studies, and six participated in two studies; the remaining 24 participated for the first time.

Our study examines advertising by brand marketers. It does not include search or paid social media campaign data.

50
Participants

2,400
Campaigns

130K
Placements

606K
Domains

27B
Impressions

Major Findings

For the First Time, More Fraud Will Be Stopped This Year than Will Succeed

- Today, fraud attempts amount to 20 to 35 percent of all ad impressions throughout the year, but the fraud that gets through and gets paid for now is now much smaller.
- We project losses to fraud to reach **$5.8 billion globally**[3] in 2019. In our prior study, we projected losses of **$6.5 billion for 2017**. That 11 percent decline in two years is particularly impressive considering that digital ad spending increased by 25.4 percent between 2017 and 2019. A detailed breakdown by device and media type is on the next page.
- For the first time, the majority of fraud attempts are getting stymied before they are paid for, by DSPs and SSPs filtering fraudulent bid requests, by clawbacks, or by other preventative measures. Absent those measures, losses to fraud would have grown to at least $14 billion annually.
- Fraud is an evolving threat. Fraud rates have been growing in new formats, and continued vigilance is needed. But there is no denying the structural gains the industry has made in this fight.

[3] Loss estimates are based strictly on digital spending in the categories included in the study: video, display, and other CPM formats for desktop and mobile devices.

Projected Fraud Losses in 2019 by Category[4]

Desktop Study Sample Size: 13.6 Billion Impressions

8% Display

down from 9% in 2017

14% Video[5]

down from 22% in 2017

12% Other

Rich Media, Takeovers, etc.

Mobile Study Sample Size: 13.5 Billion Impressions

3% Display

8% In-App Video

14% Web Video

7% Other

[4] Fraud rates exclude social media and search.

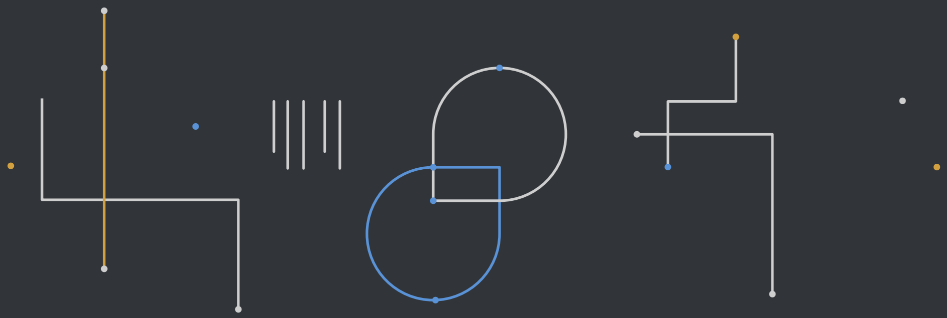
[5] As in prior years, video still shows more fraud.

How the Industry Has Made Gains Against Fraud

- Anti-fraud measures have made traffic sourcing more difficult and expensive.
- Thanks to traffic sourcing transparency efforts led by the Trustworthy Accountability Group (TAG), traffic vendors have gone further underground, reducing "retail" bot buying on the open web.
- Ads.txt has reduced domain spoofing.[6] Additionally, TAG requires publishers to have completed ads.txt files if they want to be Certified Against Fraud.
- Far more dollars are now being spent through programmatic platforms with built-in fraud prevention measures.
- Arrests, like those of the alleged masterminds behind the 3ve and Methbot operations, have brought real consequences to botnets operating overseas.

The Best Buyers Are Doing Better than Ever

- In our first study in 2014, fraud affected everyone, those with straightforward media plans and very sophisticated ones alike.
- This year, every buyer in the study knew about fraud risk; 90 percent had MRC-accredited fraud verification measures in place to deal with this risk.
- Performance was loosely correlated with study participation; long-time participants performed better than they did in previous years.
- For the top quintile of buyers, fraud was nearly nonexistent. However, discrepancies between ad server impressions numbers and verification impressions numbers were present for even the most sophisticated buyers.



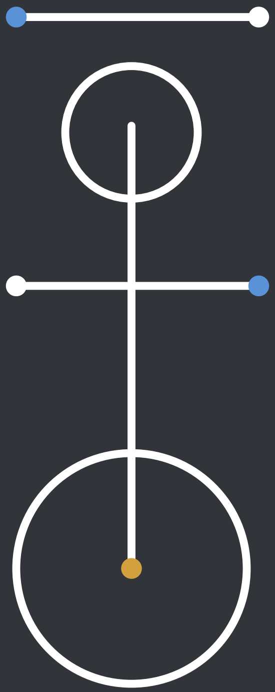
[6] Ads.txt, created by the IAB, stands for Authorized Digital Sellers. The mission of ads.txt is to increase transparency in the programmatic advertising ecosystem. It is a simple, flexible, and secure method that publishers and distributors can use to publicly declare the companies they authorize to sell their digital inventory.

It's Still Too Hard to Know What You're Buying

- The ability to hold all ad spending to the same high level of validatability should be one of marketers' top concerns in 2019. The time has come for marketers to stop tolerating — and stop paying for — outdated media formats like VAST 2 that cannot support the highest levels of third-party validation.
- By selling media under conditions that do not support high third-party validation, even trustworthy media companies are essentially part of the problem, since the vast sea of good but low-transparency inventory provides cover for the fake inventory sold under the same formats.
- Fraud detection and prevention are only as good as their implementation. Low Invalid Traffic (IVT)[7] measurements do not provide a full picture when half of a media plan is highly validatable with accredited third-party, dynamic JavaScript verification, but the rest is covered by a mix of half-measures like 1x1 pixels, list-based protections, and limited, static integrations.

Case Study: Trading Desk Video

One CPG participant purchased 44 million video impressions from an agency trading desk. Of the 44 million impressions, only 48 percent were validatable at the highest standard, with a dynamic fraud test served via JavaScript.



[7] Invalid Traffic is often referred to as Non-Human Traffic (NHT). It is the total of General Invalid Traffic (GIVT) and Sophisticated Invalid Traffic (SIVT).

The Good News

Fraud Volumes Are Growing in Newer Frontiers, Such as Mobile and OTT, But Reductions in Desktop Fraud Prove This Fight Is Winnable

The rates for desktop ad fraud were the lowest in the history of the Bot Baseline study.

Strides have been made in desktop display, where buying undetectable bot traffic is becoming too expensive to be profitable. Fraud in desktop display was 11 percent in 2014, 9 percent in 2016–17, and 8 percent in 2018–19.

This is how we win: if the industry makes cybercrime unprofitable, cybercriminals will lose the incentive to invest in ad fraud.

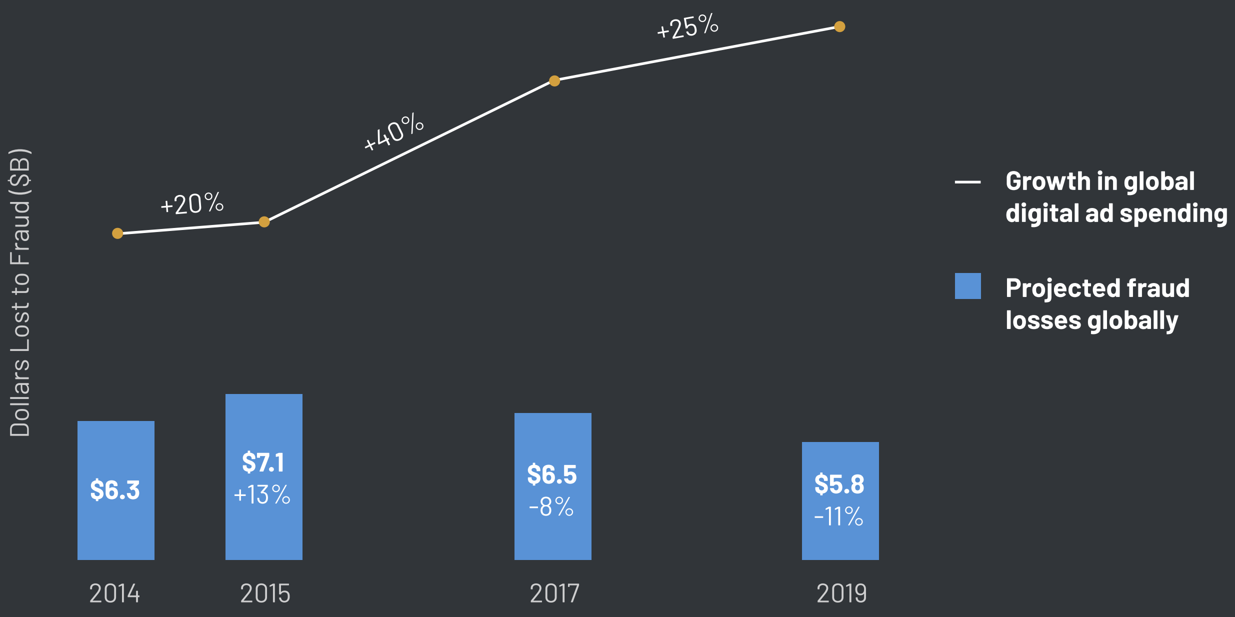
Case Study: Nonprofit

5.8 million display impressions were purchased from a century-old nonprofit organization, with just 22 percent validatable at the highest level.

Structural Improvements to the Fight Against Ad Fraud

In the five years since our first study, ad spending in the categories most vulnerable to fraud — video and display advertising across desktop and mobile devices — has more than doubled,[8] but losses to ad fraud have not. Increased awareness of the problem among marketers raised the priority of dealing with it, and leaders throughout the industry have risen to the challenge. First, this slowed the growth of ad fraud. Now, for the first time, we are projecting losses this year to be smaller than in the first year of our study.

Dollars lost to ad fraud year over year



[8] PWC Global Entertainment & Media Outlook 2018-22.

There are five reasons that ad fraud has dropped in the last year:

1. Advertisers are directing more of their money through buying channels with dedicated, independent fraud prevention measures, an option that wasn't available at the time of our first study.

2. Buying bot traffic has become harder. Many bot traffic vendors have been driven out of business or gone underground, reducing the once-prevalent availability of bot traffic sold on the open web to anyone with a credit card.

3. Buying bot traffic has become more expensive. The market price for sophisticated bot traffic has risen, limiting the arbitrage opportunity for buying visitors and showing them enough ad units to make a profit.

4. Initiatives like ads.txt have helped to reduce desktop spoofing, with 78 percent of the top volume domains in the study using ads.txt files to prevent their inventory from being successfully spoofed by the time of our study.

5. Consequences are changing the incentives for perpetrating fraud. The business of committing ad fraud has become significantly riskier, as shown by a number of arrests in the last year — something that nobody would have dreamed would happen just a few years ago.

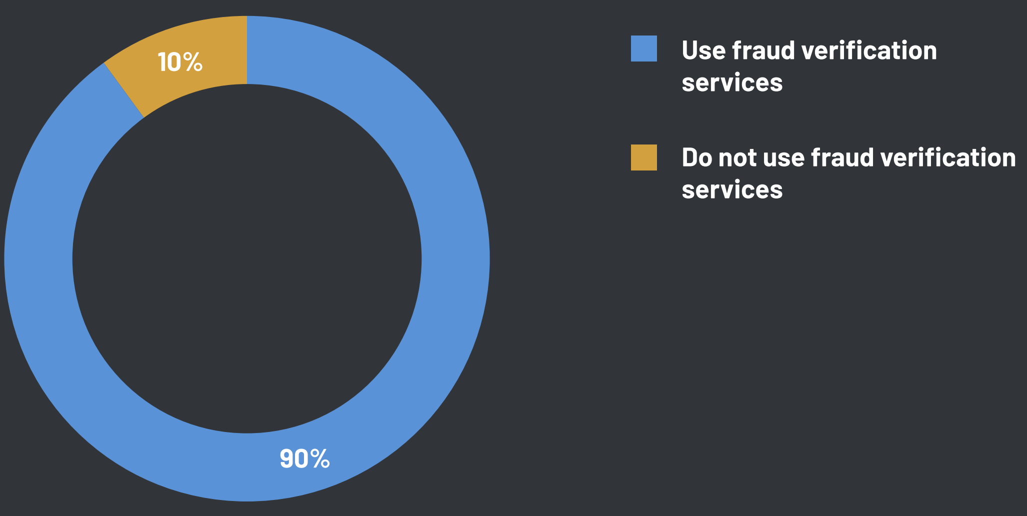
   Cybersecurity companies, industry experts, and governments have come together to dismantle fraud infrastructure and take down botnets and their operators.

More Money Is Now Being Spent Through Ad Tech Platforms with Built-In Fraud Prevention Measures

Since 2017, advertisers have moved more money into channels with built-in fraud prevention measures. In years past, it was commonplace for ad platforms to rely on their own homegrown safety and security measures; they were all in-house, with limited coordination between platforms, varying widely in resources and staff. Ultimately, they were grading their own homework.

Today, in contrast, there are third-party security firms treating fraud-fighting as a dedicated function. When security is separated from audience and media intelligence, collaboration across platforms becomes substantially easier to scale. These measures help marketers avoid bidding on invalid traffic by default, without extra effort.

Breakout of Bot Baseline participants who employ fraud verification measures:



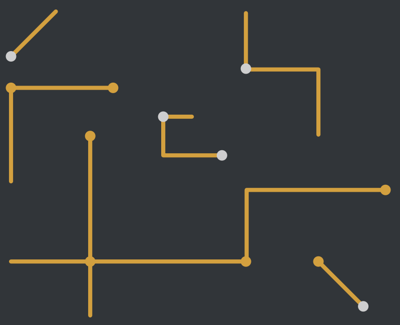
Traffic Vendors Go Underground

It used to be easy to buy high-quality bot traffic. In past years, we even discovered traffic vendors selling traffic by verification "filter," tuned to look good to different measurement vendors. Traffic aggregators kept track of the bot varieties flagged as suspicious or low quality by certain advertising measurement or verification vendors, ensuring they could only sell traffic that would go undetected by buyers.

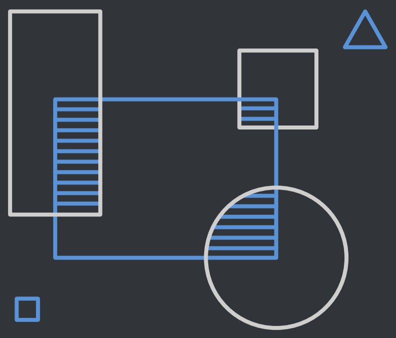
Notorious threat actor group KovCoreG exploited this to monetize the Kovter malware, infecting millions of machines in the aggregate while making the majority of their earnings off the freshest million infections at any given time. Traffic from freshly infected computers fetched the highest price. Older infections were sold for "tonnage" on low-CPM ad campaigns with little or no ad verification. This pattern, adopted by most of the top-tier fraud operators, created the appearance of fraud protection without tipping over the apple cart of widespread traffic sourcing.

But this year, we are pleased to report the beginning of a turnaround. It is still possible to buy sophisticated, realistic bot traffic, but it has become far more expensive and harder to source.

Two forces have upended this model:



As with 3ve, major botnets have been dismantled. When a botnet gets dismantled, its traffic, no matter the quality, cannot be sold anymore.



The extraordinary effort led by TAG to shine a light on traffic sourcing has severely reduced demand for low-quality traffic vendors that are caught selling bot traffic.

Retail vending of bot traffic on the open web has substantially diminished.

Realistic bot traffic used to be accessible to anyone with a credit card. Now, more of the sophisticated bot-buying has been forced underground, to invitation-only forums or transactions made over chat apps. That requires a more sophisticated buyer. You can still buy sophisticated, nearly undetectable bot traffic, but it is harder, which diminishes its scale.

Over the four Bot Baseline studies, the volume of paid traffic acquisition has increased by a small percentage of total ad traffic; however, bot traffic has decreased substantially. This is not just good news for advertisers. It is also good for all the legitimate audience aggregators who can send real people to publishers paying handsomely to reach a wider audience.

Case Study: Publicly Traded Company

A CPG participant purchased 102 million impressions from a publicly traded company; of the 102 million impressions, just 34 percent (35 million) were validatable at the highest standard.

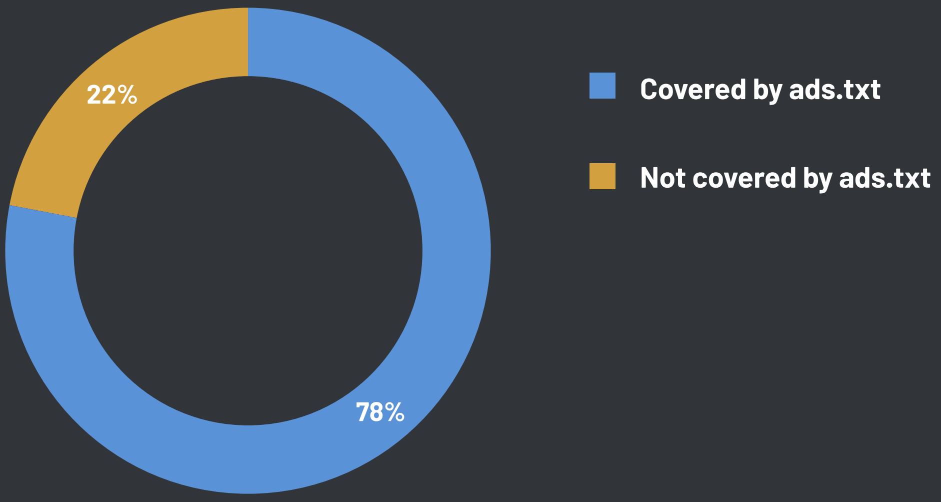
Cybersecurity Measures Like Ads.txt Have Reduced Desktop Spoofing

Ads.txt has had a particularly positive impact on desktop video. This shows that with widespread implementation of creative security solutions, the industry can push back against ad fraud.

Ads.txt was introduced in 2017 by the IAB in an attempt to curb spoofing, the practice of forging programmatic bid requests to appear to come from well-known domains. It's the RTB equivalent of selling knockoff designer handbags. With ads.txt, a simple text file hosted by a publisher's webserver lists all companies authorized to sell its ad inventory, while programmatic platforms include a file that confirms the publisher inventory they are permitted to sell. This simple but elegant solution gives buyers more clarity on the companies allowed to sell a publisher's inventory.

During our study, 78 percent of the overall study volume was covered by ads.txt, with 14 percent of all domains having published a valid ads.txt file by the time of the study. This shows that there is a long way to go to bring the long tail of domains into the ads.txt program. But even at the current state of adoption, a majority of ad impressions on the web are now being covered.

Percentage of overall study volume covered by ads.txt

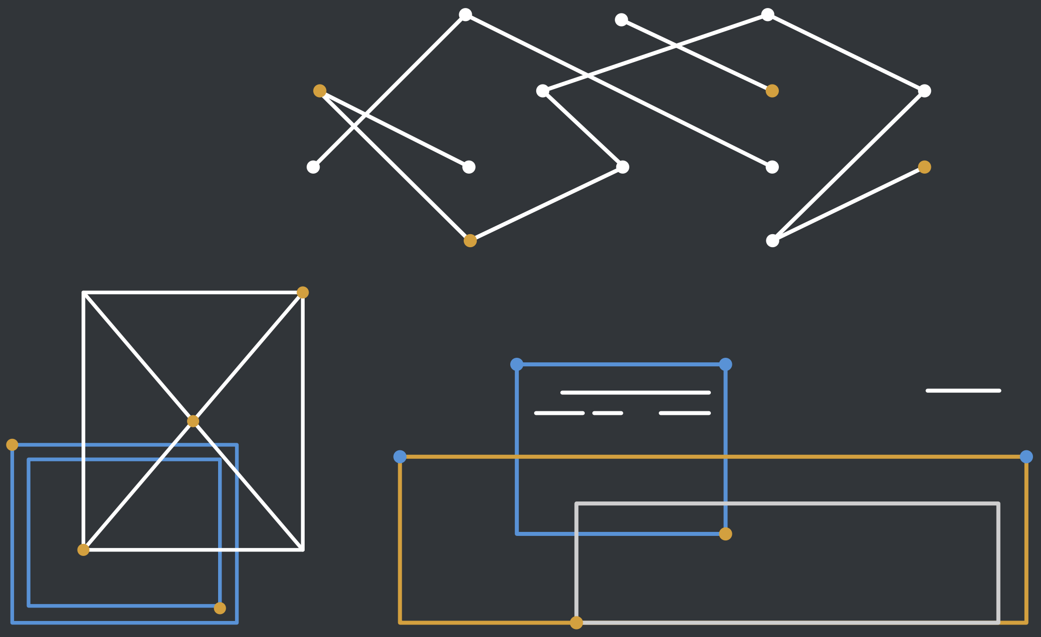


The Formation of Industry Coalitions

There is mounting evidence that fraud elimination works best when we as an industry work together. We saw this with the implementation of ads.txt, which has significantly contributed to the decline of desktop spoofing. Now, many larger platforms have partnered with cybersecurity expertise outside their own organizations to stop invalid traffic prior to auction.

The more publishers and platforms implement ads.txt and other cybersecurity measures, the greater the chance that cybercriminals will be deterred. This is precisely why the ANA, the 4A's, and the IAB are co-founders of the Trustworthy Accountability Group (TAG), a cross-industry nonprofit focused on fighting criminal activity like advertising fraud.

Cooperation proved essential in the takedown of 3ve, one of the world's most advanced ad fraud operations. The wide-scale infection footprint of 3ve covered millions of internet-connected devices, with over a million under control at any one time. It could not be eliminated by any single entity. Taking it down and indicting the alleged perpetrators required industry collaboration, dubbed Operation Eversion, spanning advertising platforms, security experts, and government agencies.



The Bad News

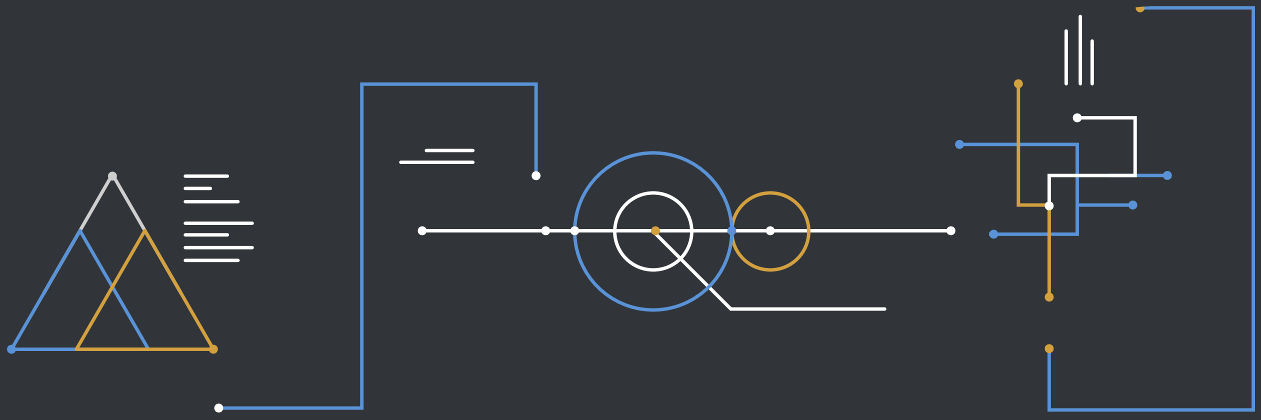
The decrease in ad fraud is a reason for celebration. Less fraud means less revenue lost to cybercriminals. Increased implementation of cybersecurity defenses, raising the cost and risk of the crime, has helped to dissuade some would-be fraudsters from pursuing this line of cybercrime altogether.

However, the digital ad industry still has a long way to go before ad fraud is completely eliminated. The most immediate — and fixable — issue the advertising industry faces is the incredibly uneven auditability of ad campaigns across formats and environments. A shockingly large percentage of video advertising, for instance, is sold strictly in the outdated VAST 2 format, with a very low level of third-party validatability, while other publishers are bending over backwards to provide full, dynamic, audited JavaScript execution from accredited third parties.

In effect, different publishers are being held to different standards of validatability, with very little marketer awareness.

But transparency is far from the only challenge. As we saw with the botnet 3ve, the bot operators that remain in business have substantially increased their scale and sophistication since our first study in 2014. In 2018, we witnessed highly advanced tag evasion and selective tag execution ("monkey-patching") techniques to prevent fraud detection code from running correctly.

Moreover, as cybercriminals see reduced profit opportunities in desktop ad fraud, they are moving to other formats. Mobile fraud is increasing in some formats. There are also rising threats in OTT and server-side ad insertion (SSAI), as well as murky waters surrounding incentivized traffic. We believe that all of these issues will become more prevalent in the coming year.



Less than Half of All Impressions Are Fully, Transparently Validatable

The next stage in eliminating fraud is to raise every platform and format to the same high standard of validatability. White Ops observed more than 50 billion third-party ad server impressions over the course of the study period. Of those impressions, fewer than half met the highest standard of third-party validatability.

In many cases, impressions could not support a dynamic fraud detection tag via JavaScript or an equivalently dynamic challenge via SDK or other integration. Those impressions are vulnerable to the most sophisticated fraud operations, which have proven that they can defeat static detection. In other cases, placements included walled gardens with limited third-party validatability by design. In yet more cases, the process is still so onerous or error-prone to get fully dynamic third-party tagging working that it just did not happen because of time constraints. In the aggregate, each of these gaps adds up.

Among Bot Baseline participants, we encountered the lowest standards in mobile video.

Device Type	Buy Type	Media Type	Validatable at the Highest Level
Desktop	Direct	Display	51.8%
Desktop	Direct	Video	36.8%
Desktop	Programmatic	Display	50.8%
Desktop	Programmatic	Video	60.3%
Mobile	Direct	Video	28.8%
Mobile	Direct	Display	50.7%
Mobile	Programmatic	Display	46.1%
Mobile	Programmatic	Video	33.6%

Most Video Ads Still Don't Support VAST 4, the Gold Standard in Transparency

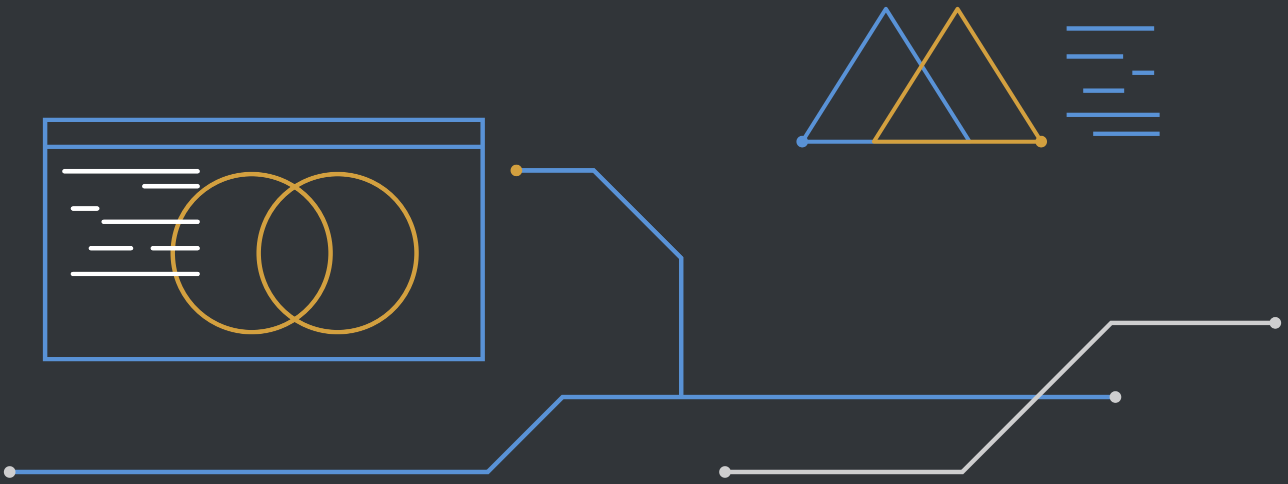
Back in 2008, the IAB introduced Video Ad Serving Template (VAST) to give advertisers a standardized video ad template to work from. There have been several incarnations of VAST; the most recent version is VAST 4, which was updated to version 4.1 in November 2018.

VAST 4 is the most comprehensive template yet, as it finally supports vendor verification and viewability. Before, vendors were forced to request that clients run VPAID tags, which allow deployment of JavaScript, supporting measurement of viewability or fraud; however, this was not an ideal solution.

1. This was not VPAID's intended use (it was originally created to track interactions).
2. This "workaround" only works for one vendor's tags. Those marketers who want to leverage viewability and ad fraud solutions were forced to choose one single vendor tag.
3. VPAID runs in browsers, which makes it ineffective in app-heavy mobile inventory.
4. With the introduction of VAST 4, VPAID has been deprecated by major platforms.

The good news:

VAST 4 is here. We recommend that marketers start requesting support for the most recent version of VAST 4 and implementation for their video ad tags, and push for compatibility with all publisher and ad tech platforms as well as video players.



The Quality of Real-Time Independent Monitoring Is Still Inconsistent

Real-time third-party monitoring and auditing still have much room for improvement. The gold standard in third-party monitoring and auditability is technology that can serve a dynamic challenge to bots via JavaScript that changes, an SDK that supports dynamic challenges, or some other integration measure that presents bots with a challenge the bot author cannot anticipate. This is the technology that White Ops used to measure traffic during this study. That gold standard faces several hurdles, however.

Outdated Technology

Many large publishers are still using outdated technology that permits only low-fidelity validation, like 1x1 pixels. This is particularly prevalent with video ads. When ad servers and video players only support VAST 2 videos — which do not allow JavaScript — advertisers have very little insight into the amount of fraud on those ad impressions.

Unfortunately, by sticking with VAST 2, even publishers that are themselves trustworthy are effectively providing cover to fraudsters who are using those low-fidelity formats to hide their fraud.

Technical Limitations

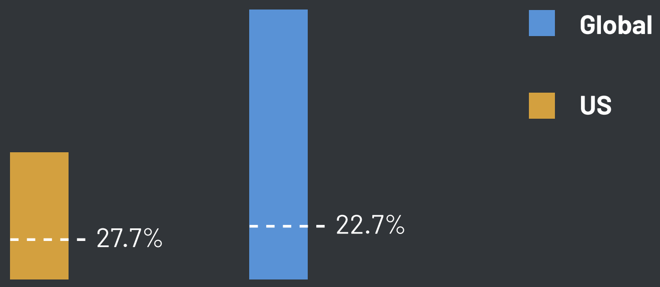
Over the years, there have been technical limitations in third-party ad servers (both buy- and sell-side), publisher and ad tech platforms, and/or video player technology. The time has come for brand advertisers to push their partners to adopt newer technology.

Walled Gardens

Some large digital advertising platforms are referred to as “walled gardens,” as they exert more access control over their platforms than publishers on the open web. Since advertisers rely on many different ad tech services and tools to monitor and understand their campaigns, less access means less visibility and independent validatability.

Walled gardens also typically hold more user data than publishers on the open web. Care must be taken to give marketers the ability to hold every publisher to the same high standard of validatability while respecting and caring for user privacy. We are all consumers too. Let’s build a world we want to live in.

Percentage of total digital ad spending on walled gardens



Case Study: Publisher

Participants collectively bought 34 million display impressions from a popular publisher website, of which only eight million were validatable at the highest standard.

Human Error

Human error can occur during implementation of the JavaScript tag, the most common of which is adding a JavaScript tag as a 1x1 tag, which does not let the code fire properly. It is important to check implementations post-launch, and to frequently review reporting for major discrepancies.

Intentional Evasion

Malware can selectively execute — a technique sometimes called monkey-patching — to avoid running code in the intended way. Enterprise-scale botnets block the solutions they cannot beat. It is important to note that evasion of measurement is one of the clearest indicators of fraudulent behavior. After ruling out an error, a consistently high discrepancy between fraud detection tags and impression counts should be considered invalid. Providers that silently drop fraud detection tags are often trying to hide something.

In Practice, High-Fidelity Validation Is Surprisingly Uneven

Verification and analytics companies sometimes issue 1x1 tags for placements that do not allow JavaScript, but these tags are fundamentally low-fidelity. They generally do not have the ability to employ much more than the obvious fraud detection techniques: frequency counts, cohort and co-visitation analysis, and checks against industry lists, like those managed by TAG and the IAB.

The Frontiers in Fraud

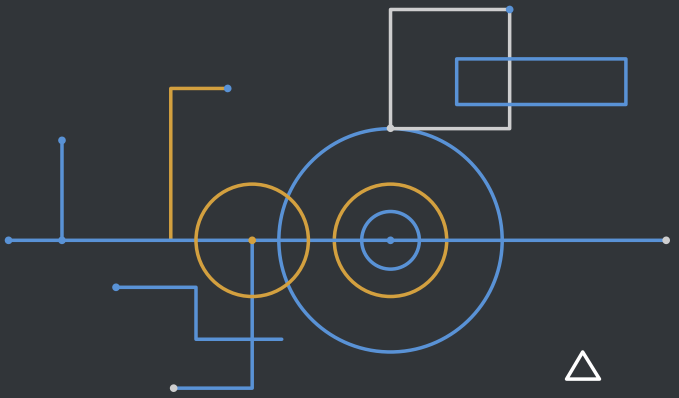
Fraud on the Mobile Web: Easy to Do, Hard to Profit

Mobile display ads harbored surprisingly little SIVT. This is not because mobile is inherently safer, or that fraud is harder in mobile. It is actually a consequence of the economics of fraud and the number of ad units per page. The cost of bot traffic has risen, while mobile display CPMs and ad units per page are relatively low. As a result, it is simply hard to buy bot traffic that is sophisticated enough not to get caught at a price that is low enough to turn a profit when the total page CPM for a typical mobile session without video is so low.

In mobile web video, CPMs are higher, making it easier to turn a profit even for publishers paying a high price-per-bot-visitor. As a result, fraud rates there are higher.

In-App Fraud: Growing and Innovative

In 2019, the average time spent per day with mobile devices by U.S. adults will surpass that of TV.[9] The most common threats we are seeing today are app spoofing[10] and hidden ads.[11] Marketers must remain vigilant and push to ensure *all* mobile media can be validated at the same high level as desktop media. To curb spoofing, suppliers should now adopt app-ads.txt. Advertisers should work only with those platforms and publishers that have adopted both.



[9] Source: eMarketer, April 2018.
[10] App spoofing is when an app sends a fake app bundle ID, to disguise itself as a different (often premium) app.
[11] Hidden ads are non-viewable ads, typically deployed with purpose-built deception code to fake viewability.

Server-Side Ad Insertion (SSAI)

Server-Side Ad Insertion is a way to stitch video and ad content together on the server side, instead of the client side, so that content can be served as one continuous stream. When SSAI came out, it gave the industry a way to be “more like TV” and provided a mechanism to prevent ad blockers from detecting ads.

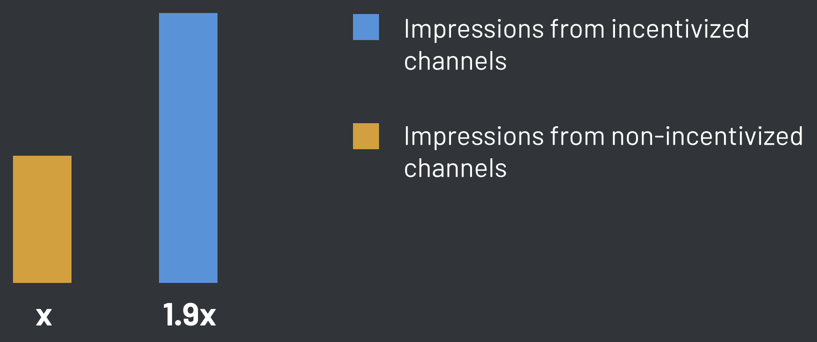
Along the way, SSAI breaks some assumptions built into many impression-tracking and ad verification services. This can blind some verification techniques, increasing the risk of false negatives. Conversely, SSAI makes many normal behaviors appear automated, which can increase the risk of false positives. Evolution in both technology and industry standards is needed to reach a uniformly high level of independent validatability.

Connected TV (CTV)

We expect CTV to be one of the fastest-growing markets for advertisers, so brand advertisers must protect themselves. Marketers will spend a projected $20 billion on CTV over the next two to three years.

Specific threats that we are seeing include OTT device impersonation, SSAI spoofing, app spoofing, device farms, hidden ads, and fraudulently incentivized ads. We recommend that marketers work with platforms protected by fraud verification companies. We also encourage CTV platforms to adopt protocols or guidelines such as the app-ads.txt. We believe that these steps are critical to ensure that all supply is authenticated.

Bottiness in incentivized channel impressions



Recommendations

Steps to Take to Reduce Fraud

Require clarity from vendors on how they combat fraud

Always ask your vendor how it measures for invalid traffic — whether it matches against a list (using general detection methods) or uses sophisticated IVT detection method(s) as defined by MRC. When possible, use solutions that are proven to reduce fraud in targeted media and buy types.

Work with vendors who have implemented ads.txt

Ads.txt was created by the IAB Tech Lab to help publishers create lists of authorized media sellers. Over the course of this study, 78 percent of traffic came from domains authorized by ads.txt. The average rate of SIVT across domains not authorized by ads.txt was 2.2 times higher than domains authorized by ads.txt.

As more companies implement the protocol, fraudsters will have even less leeway to commit crimes. IAB Tech Lab has ads.txt aggregations compiled from periodic internet-wide crawls of ads.txt files. In this study, the list was cross-referenced with the domains encountered during the BB4 study to identify which domains had ads.txt implemented.

Use an MRC-accredited anti-fraud vendor

We recommend you use an anti-fraud vendor to filter your ad campaigns to detect and remove sophisticated invalid traffic and fraud; MRC has audited and accredited several of these vendors for compliance with industry standards which helps ensure coverage of significant threat models. A list of MRC-accredited anti-fraud vendors can be found at: http://www.mediaratingcouncil.org/Digital%20Landscape.pdf. See the section entitled "Sophisticated Invalid Traffic Detection/Filtration."

Reject the temptation to buy for "tonnage"

Tonnage is space on high-risk and less-valuable placements, which are filled with undisclosed incentivized inventory and adware. While in general we stand against incentivized traffic, we believe the industry should enlist ad tech platforms and publishers to begin proactively declaring traffic as incentivized with an incentivized flag.

Include language on non-human traffic in your terms and conditions

As in previous years, we recommend that insertion orders include strong language surrounding the types of traffic your company will agree to pay for. One participant includes the phrase, “Final numbers to be actualized based on third-party reported non-SIVT impressions.”

Insist that your company will only pay for non-SIVT or non-IVT impressions. Consult your legal counsel to develop provisions that best serve your company’s specific needs and interests.

Join an industry coalition working to fight ad fraud

As we saw in the battle against 3ve, as well as the success of ads.txt, there is evidence that fraud solutions work best when stakeholders work together. White Ops and the ANA support industry organizations such as TAG and IAB, and we encourage the industry as a whole to do the same.

Implement a comprehensive fraud detection solution, and work only with vendors who have done the same

A combination of in-house defenses and third-party cybersecurity software that detects fraud before any purchase has been made is the best layered protection against cybercriminal activity.

Sourced traffic management: buy from providers that are complying with TAG’s disclosure guidelines

Marketers must take responsibility for more active stewardship of their media investments. Media is often the largest marketing expenditure at most companies, and marketers should be responsible for assuming greater internal stewardship of their media investments.

This includes taking an active role in having practices and guidelines in place to reduce the risk of digital advertising fraud and to ensure transparency through verifiable auditing standards. Marketers who ignore this or completely outsource fraud reduction/prevention do so at their and their shareholders’ risk.

Steps to Take to Increase Measurability

Insist upon transparency from partners, and regularly check for discrepancies

Partnerships should be based on transparency. Set fraud goals early in your campaign and have open discussions with partners that don't meet those goals. Proactively ask about traffic sourcing practices and hire TAG-certified vendors where possible. Additionally, marketers should be aware when ads are incentivized so they can make informed decisions on whether they want to spend in that channel.

Continue to advocate for JavaScript and VAST 4 support, compatibility, and implementation

Advocate for robust third-party SIVT measurement of all your supply and publisher partners. This includes asking for or requiring JavaScript execution, which will allow for better data collection. We recommend that marketers start requesting VAST 4 support and implementation for their video ad tags and push for compatibility with all publisher and ad tech platforms as well as video players.

Fraud in walled gardens must be measurable

This study did not measure bot fraud in the walled gardens, and our projections of dollars lost to fraud do not include spending on walled gardens. Today, marketers are seeking enhanced trust and transparency from their partners, elevating this issue in importance. It is imperative that marketers can hold every buying channel to the same high standard of validatability. The ANA and White Ops call on all publishers, including walled gardens, to allow third-party fraud detection. Visibility into fraud will lead to corrective action.

Ad Fraud Glossary: An Exhaustive List

1x1 (Also referred to as 1x1 Pixel, 1x1 Tag): An image measuring one pixel by one pixel, often used for tracking online events.

Ad Injection: Visible or hidden insertion of ads into an app, web page, or other online resources without the consent of the publisher or operator.

Ad: An online advertisement impression of any sort.

Advertiser: A company, brand, or individual who pays a third party to display ads.

Adware: Software, sometimes automatically installed on user devices, that displays ads to users while the software is running.

App Spoofing: When an app sends a fake app bundle ID to disguise itself as another (sometimes premium) app.

Audience Extension: This term is used in two different ways: 1. A technique used by some advertisers to reach a bigger but still relevant audience by using a known audience segment's common characteristics to target people outside the segment who share the same characteristics. This is often called lookalike modeling. 2. A technique used by some publishers to extend the volume of an advertiser's campaign by fulfilling some of the campaign by retargeting the publisher's audience while they visit other sites.

Audience Segmentation: The process of dividing people into subgroups based upon a defined set of criteria such as demographics, geographic location, etc.

Auto-Play: When a sound, video, or any other type of media plays, generally as part of an ad, without any user interaction, often when the user loads a web page or other resource.

Automated Browsing: A program or automated script that requests web content (including digital ads) without user involvement and without declaring itself as a *crawler* (which see).

Bot (aka Non-Human Traffic or NHT): Automated software agents capable of interacting with digital content, including apps, text, video, images, audio, APIs, and other data. These agents may intentionally or unintentionally interact with apps and webpages, view ads, watch videos, listen to radio spots, fake viewability, and click on ads.

Bot Detection: The detection and differentiation of bot traffic, bot impressions, and bot interactions.

Bot Farm: A group of computers generating bot traffic that is not part of a botnet of infected computers. Bot farms can be run in a data center or elsewhere.

Bot Fraud: Ad fraud perpetrated with bots.

Bot Impression: An ad impression served to a bot.

Bot Percentage: The percentage of traffic made up by bots, or the percentage of impressions served to bots. Also called *bottiness*.

Bot Traffic: Automated website or other online traffic and/or ad impressions driven by or resulting from bots.

Bottiness: See: *Bot Percentage*.

Botnet: A group of infected devices — computers, IoT devices, potentially anything connected to the internet — that generate bot traffic. The owners of these devices are usually unaware that their devices are infected.

Browser Extension: A plug-in that extends the functionality of a web browser.

Campaign: A planned series of ads and their formats together with the publishers, sites, and other channels the ads are run on or purchased through.

Cash-Out Site: A website, app, or other resource capable of delivering ads, operated by the perpetrators of ad fraud for the purpose of exfiltrating money from the online advertising ecosystem. Also called a *ghost site*.

Clawback: The recovery of ad spend from a partner, for instance due to ad fraud in a particular campaign or placement.

Click Farm: A type of ad fraud in which a large group of human workers (in one or multiple locations) is paid or otherwise incentivized to view or click on ads on behalf of a third party that economically benefits from it.

Connected TV (CTV): A television that is connected to the internet.

Crawler (aka Spider): An internet bot that systematically browses the World Wide Web, typically for the purpose of indexing (see also: *Known Crawler*).

Data Center: A hosting facility for servers. Since most consumer internet activity — like browsing, watching videos, and playing games — is done with consumer devices, not servers, ad traffic from data centers is suspicious. However, it is not inherently invalid. Consumer traffic can be routed through data centers when using VPNs or cloud desktops. Ad traffic may appear to come from data centers when web serving accelerators or server-side ad stitching is involved. Cooperative efforts like the TAG Data Center IP list (see also: *Trustworthy Accountability Group*) are used to help identify data center traffic that is never valid.

Device Farm: A large group of mobile devices used to inflate mobile interaction numbers, like clicks or app installs. Device farms are the mobile equivalent of *Bot Farms* (which see). They can be controlled via automation or, like a *Click Farm* (which see), by repetitive human work.

Domain Blacklisting: Using lists of known bad domains to prevent ads from being served to those domains.

Domain: Common shorthand for domain name. A unique name that identifies an internet resource such as a website. The ANA's domain is ana.net.

Demand-Side Platform (DSP): A platform that allows advertisers or their agencies to automate the purchase of media across multiple supply sources through real-time bidding.

Event: In the context of online advertising, a page view, ad impression, or video play.

Engagement: A metric of varying specificity that provides a qualitative evaluation of a user's interaction with a given ad, web page, or app.

Exchange: A technology platform that facilitates the buying and selling of ads or ad data in real time ("RTB") from multiple sources such as publishers and networks of publishers.

False Representation: An ad request for inventory that is different from the actual inventory being supplied, including ad requests where the actual ad is rendered to a different website or application, device, or other target (such as geography).

General Invalid Traffic (GIVT): Known non-human or fraudulent sources that can be identified with industry lists like the IAB/ABC International Spiders and Bots List or with parameter-based detection techniques. Formerly known as *Known Bots*.

Ghost Site: See: *Cash-Out Site*.

Hidden Ads: Ads made intentionally non-viewable, typically deployed with purpose-built deception code to fake viewability.

Human Impression: An impression legitimately served to a real human who is not intentionally or unintentionally engaged in any form of ad fraud.

Impression: An instance of the delivery of an online ad.

Incentivized Behavior: The use of an explicit incentive, for instance a financial reward, to drive users to interact with one or more ads for the sole purpose of receiving the incentive.

Incentivized Ad: An impression served to a human who is paid or otherwise incentivized to interact with the ad.

Intended Domain: The domain an ad is expected to be served on.

Invalid Traffic (IVT): The sum of all non-human or fraudulent traffic (SIVT + GIVT).

IP Address or IP: A unique numerical address corresponding to a device or set of devices connected to the internet.

IP Blacklisting: Using lists of known bad IPs to prevent the serving of ads to those IPs.

IP Geolocation: Determining the approximate physical location of a device connected to the internet at a given point in time by using information associated with or deduced from that device's IP address.

Known Crawler: A program or automated script that requests content and declares itself as non-human. That declaration can be made through a variety of identification mechanisms, the most common of which is in the user-agent string that every browser sends to a server when making a request. These crawlers are usually included in the IAB Bots and Spiders List (see also: *Internet Advertising Bureau*).

Lookalike Targeting: See: *Audience Extension*.

Long Tail: Websites with relatively low traffic that may offer value to advertisers due to their appeal to specific or niche audiences of users.

Make-Good: Credit given to an advertiser or its agency to compensate for an error in the composition, placement, or delivery of an ad.

Man-in-the-Browser Attack: An internet attack that infects a user's online interactions by taking advantage of vulnerabilities in browser or app security to modify ads, web pages, or online transactions. This occurs without the knowledge or consent of the user or the resources with which the user intended to interact.

Manipulated Behavior: A browser, application, or other program that triggers an ad interaction without a user's consent, such as an unintended click, an unexpected conversion, or false attribution for installation of a mobile app.

Micro-Blacklist: A blacklist that is updated and expires frequently to enhance its effectiveness against advanced and adaptive threats.

Misleading User Interface: A web page, application, or other visual element modified to falsely include one or more ads.

Native Advertising: Ads and ad placements that match the look, feel, and function of the publication or app in which they appear.

Network: A group of sites that are owned and operated by a single entity or that agree to sell some of their inventory together.

Over the Top (OTT): A reference to content providers that distribute streaming media directly to viewers over the internet.

Page View: A single request to load a single page of a website.

Phantom Layer: Websites operated specifically for laundering ad fraud by obscuring the source of inventory and impressions entering the online advertising ecosystem.

Pop-Under: Windows that appear or open under the user's current browser window so that they become visible only when the original window is closed.

Pop-Up: Windows that appear or open above or on top of the user's current browser window.

Potentially Unwanted Application or Program (PUA, PUP): Software that may be perceived as unwanted by the user because it may compromise privacy or weaken the computer's security.

Proxy: A server that sits between a client application, such as a web browser, and a server. It intercepts all requests to the real server to see if it can fulfill the requests itself. If not, it forwards the request to the real server.

Proxy Traffic: Traffic that has been routed through one or more servers. Though proxied traffic is not necessarily evidence of IVT, botnets are sometimes used as proxy exit points to make traffic from one part of the world appear, to come from the devices that make up the botnet.

Publisher: The operator of a website or network of websites, and the producer or curator of content for those sites.

Reach: The total number of different users exposed, at least once, to an ad or campaign during a given period.

Real-Time Bidding (RTB): Commonly referred to as programmatic advertising, real-time bidding is used to buy and sell advertising, per impression, in an auction between automated ad buying platforms that takes place in the milliseconds between when a browser or app requests an ad and an ad is displayed.

Retargeting or Behavioral Targeting: The process of delivering ads to users based on their previous online activity.

Run of Network (RON): An ad or campaign displayed on a large collection of websites without the ability to choose target-specific sites, placements, or domains.

Server-Side Ad Insertion (SSAI): A way to stitch video and ad content together on the server side, instead of the client side, so that content can be served as one continuous stream.

Site or Website: A set of related web pages, usually served from a single domain.

Sophisticated Invalid Traffic (SIVT): Invalid traffic that is purpose-built to evade detection.

Sourced Traffic: Any method by which publishers acquire more visitors through third parties.

Spider: See: *Crawler, Known Crawler*.

Supply-Side Platform (SSP): A technology platform that enables publishers to sell their ad inventory in automated fashion, via *RTB* (which see), to a wide number of potential purchasers.

Traffic: Visits to a site, page, or other online resource.

Traffic Broker: Third-party arbitrageurs that buy traffic from suppliers and sell to publishers, often media agencies, retargeting platforms, or traffic extension platforms.

Trustworthy Accountability Group (TAG): A joint media and marketing industry program that was created with a focus on four core areas: eliminating fraudulent digital advertising traffic, combating malware, fighting ad-supported internet piracy to promote brand integrity, and promoting brand safety through greater transparency. TAG was created by the American Association of Advertising Agencies (4A's), the Association of National Advertisers (ANA), and the Interactive Advertising Bureau (IAB), and works collaboratively with companies throughout the digital ad supply chain.

User: A person who uses a computer or other device or network service. In the context of online advertising, a visitor to a publisher's site or user of an app.

Video Ad Serving Template (VAST): A universal specification developed by the IAB for serving video ads.

Video Player Ad-Serving Interface Definition (VPAID): A mechanism established to help allow third-party measurement of video ads.

Virtual Private Network (VPN): A technology that creates an encrypted and presumably more secure connection over a less secure network like the Internet. It was developed as a way to allow remote users and branch offices to securely access corporate applications and other resources. (see also: *VPN Traffic*).

VPN Traffic: Traffic that is viewed over a *Virtual Private Network (VPN)*. Without careful differentiation, this traffic may be flagged as invalid when the VPNs exit point is housed in a data center.

Walled Gardens: Some large digital advertising platforms are referred to as "walled gardens" as they exert more access control over their platforms than publishers on the open web do.

About the Study Partners

About the ANA

The ANA (Association of National Advertisers) makes a difference for individuals, brands, and the industry by driving growth, advancing the interests of marketers, and promoting and protecting the well-being of the marketing community. Founded in 1910, the ANA provides leadership that advances marketing excellence and shapes the future of the industry. The ANA's membership includes more than 1,850 companies and organizations with 20,000 brands that engage almost 100,000 industry professionals and collectively spend or support more than $400 billion in marketing and advertising annually. The membership is comprised of more than 1,100 client-side marketers and more than 750 marketing solutions provider members, which include leading marketing data science and technology suppliers, ad agencies, law firms, consultants, and vendors. Further enriching the ecosystem is the work of the nonprofit ANA Educational Foundation (AEF), which has the mission of enhancing the understanding of advertising and marketing within the academic and marketing communities.

About White Ops

White Ops is a cybersecurity company that protects the Internet from malicious bot activity. Globally, software-as-as-service from White Ops determines the validity of nearly 100 billion transactions per day on behalf of over 200 customers. Our proactive adaptation, Internet-scale, and multi-layered methodology have made us the platform of choice for some of the largest and most forward-thinking platforms and brands. For more information, visit www.whiteops.com.

Bot Baseline 2018-2019

www.ana.net | info.ana.net | www.whiteops.com | info@whiteops.com