

2015

# Bot Baseline

FRAUD IN DIGITAL ADVERTISING



# Table of Contents

## ANA & WHITE OPS REPORT

## 2015 BOT BASELINE

Foreword: Closing the Profit Windows of Bots .....	4
<b>I. Executive Summary .....</b>	<b>7</b>
<b>II. Detailed Findings .....</b>	<b>10</b>
1. Once Again, No Advertiser Was immune to Bots .....	10
2. Bot Impacts Ranged from \$250,000 to \$42 Million .....	10
3. The Majority of Returning Participants Did Not Improve .....	11
4. Bots Continue to Evade Detection and Create High Costs .....	13
5. Bot Operators Continue to Hide Bot Traffic Among Humans .....	18
6. Bots Get In When Targeted Audiences Do Not Meet Demand .....	22
7. Bots Use Complex Techniques to Increase Profits .....	26
8. Mobile: An Emerging Frontier .....	33
9. Being Aware and Involved Reduces Fraud Exposure .....	35
<b>III. Recommendations .....</b>	<b>33</b>
1. Action Plan for All Stakeholders .....	34
2. Action Plan for Buyers .....	35
3. Action Plan for Publishers, Platforms, and Exchanges .....	37
<b>IV. Appendix .....</b>	<b>38</b>
A. Methodology .....	39
B. Illustrative Terms and Conditions .....	40
C. Glossary .....	41

### About the ANA

The ANA (Association of National Advertisers) provides leadership that advances marketing excellence and shapes the future of the industry. Founded in 1910, the ANA's membership includes nearly 700 companies with 10,000 brands that collectively spend over \$250 billion in marketing and advertising. The ANA also includes the Business Marketing Association (BMA) and the Brand Activation Association (BAA), which operate as divisions of the ANA, and the Advertising Educational Foundation, which is an ANA subsidiary. The ANA advances the interests of marketers and promotes and protects the well-being of the marketing community.

### About White Ops, Inc.

White Ops is the leading provider of cyber-security services for the detection and prevention of sophisticated bot and malware fraud. Unlike traditional approaches that employ statistical analysis, simple blacklisting, or static signatures, White Ops effectively combats criminal activity by actually differentiating between robotic and human interaction within online advertising and publishing, enterprise business networks, e-commerce transactions, financial systems, and more, allowing organizations to remove and prevent fraudulent traffic and activity. By working with customers to cut off sources of bad Internet traffic, White Ops makes bot and malware fraud unprofitable and unsustainable for the cyber-criminals — an economic strategy that will eventually eradicate this type of fraud.

# Special Thanks to the Following ANA Member Company Participants

---





PlayStation.

PrestigeBrands



## Closing the Profit Windows of Bots

Botnets do not need to go undetected forever to be profitable. The smartest operators continuously infect new machines and monetize them differently to maximize yield. Even if a bot operator's programs get detected, the profits remain constant if the operator infects new computers faster than old infections are discovered. Bot operations, then, have a "profit window," a period of time from when a computer has been freshly infected until the bot is so widely detected that no one will pay for its impressions anymore.

Infections at the leading edge of the profit window, those that are "fresh," affect high-CPM advertising buys. Because most systems will not determine that the just-infected machines are now sending non-human traffic, high-CPM direct buys, programmatic private marketplace deals, and buys on top-tier platforms are all affected. Bots make their way into those deals from publishers which are buying expensive PPC (pay-per-click) traffic.

Infected machines that have existed for some time — the trailing edge of the profit window — are easier to catch, and fool fewer parties. Therefore, such bots have fewer buyers and only affect low-CPM buys. A different tier of publisher pays a lower price-per-click for that traffic, affecting buys on mid-tier programmatic platforms and lower CPM direct buys, "free" bonus, and incentive placements.

The bottom of the bot monetization barrel is the "platform of last resort," where buyers know they can go to buy cheap "tonnage" and long-tail publishers can make money with an audience paid for with the cheapest PPC traffic. Whichever high-volume inventory source is doing the worst job of purging bots off its platform in a quarter becomes the platform of last resort.

The platform migration of bot populations is not planned by the bot operators. Rather, it's a consequence of market forces. The best, most profitable traffic brokers adopt bot-blocking software to filter out all the bots that get caught, selling only the freshest infections to buyers paying a premium. Older bots get bought by the buyers who don't care and just want "tonnage." To close the profit window and stop funding bot traffic as much as possible, advertisers must take a stand against ad fraud by implementing the recommendations of this report and of groups such as the ANA and TAG.



**Michael Tiffany**  
Chief Executive Officer  
White Ops Inc.



The background features a complex network of black nodes of varying sizes connected by thin, light gray lines. Some nodes are significantly larger than others, creating a sense of hierarchy or importance. Dashed gray lines form large, overlapping circular shapes that intersect with the network. The overall aesthetic is clean, modern, and technical, suggesting a focus on data, technology, or global connectivity.

# I. Executive Summary

# Executive Summary

In 2014, White Ops and the Association of National Advertisers (ANA) partnered to release the Bot Baseline Study, considered by many to be the seminal report on advertising fraud. The 2014 study helped provide the industry with a better understanding of the impact of fraud on the online advertising ecosystem and provided a series of action steps to help stakeholders reduce fraud.

In 2015, White Ops and the ANA worked together again to repeat the study, this time with a larger group of participants: 49 advertisers versus 36 in 2014. These participants deployed White Ops detection tags on their digital advertising to measure bot fraud, or non-human traffic. Data was collected over 61 days from August 1 to September 30, 2015 (the same period as 2014). However, unlike 2014, the 2015 study was not publicly announced in advance. All participants received proprietary information on their buys. The aggregate data

is reported here.

Million average lost per participant	Billion estimated global losses in 2016
--	---

HIGH BOT IMPACT

## Major Findings

### BOT PROFITS INCREASED IN 2015

#### a. Financial Impact Averaged \$10 Million per Participant, with \$7.2 Billion Estimated Global Losses Expected in 2016

The annual financial impact of bot fraud ranged between \$250,000 and \$42 million for the 49 participating advertisers and averaged about \$10 million per participant. The 2014 Bot Baseline Study estimated that advertisers would lose approximately \$6.3 billion globally to bots in 2015. With the overall rate of fraud unchanged in our current study and estimating a 15 percent increase in global digital spending in 2016, losses due to bots could be approximately \$7.2 billion globally in 2016.

#### b. Bots Are Fooling Detection and Prevention Efforts

- Bots exploit users' cookies to appear as humans in general detection and prevention systems.
- Bots spoofed viewability, showing nearly the same viewable rates as humans. Bots fooled list-based prevention technologies in programmatic buys.
- Desktop bots impersonated mobile devices to consume mobile media.

#### c. Bots Prey on Higher-Value Media

Media with higher CPMs (cost per thousand impressions) was more vulnerable to bots, as these segments provide a stronger economic incentive for botnet operators to commit fraud. Display media with CPMs over \$10 had 39 percent higher bots than lower-CPM media. Video media with CPMs over \$15 had 173 percent higher bots than lower-CPM media.

#### d. More Focused Targeting Results in Increased Fraud

- The high demand/limited supply for targeting certain high-CPM market segments, such as high-income demographics or Hispanics, means rewards are greater for bot operators which can seemingly supply the needed audience impressions in those segments.
- Hispanic-targeted programmatic media had 70 percent greater bots than non-Hispanic.
- Hispanic-targeted direct buys had 20 percent greater bots than non-Hispanic.

**\$10**

**\$7.2**

**39%**

Higher bot rates in display media over \$10 CPM

**70%**

Higher bot rates in Hispanic-targeted programmatic media

## BOT FRAUD RATES OVERALL SHOWED NO CHANGE IN 2015

### a. Overall Fraud Levels Ranged from 3 Percent to 37 Percent

In 2015, advertisers had a range of bot percentages varying from 3 to 37 percent, compared to 2 to 22 percent in 2014. But the overall rate of fraud was basically unchanged. Only about one third of the advertisers which participated in both 2014 and 2015 experienced a decrease in their bot rates, suggesting that advertising fraud needs to continue to be a focus in 2016.

### b. Traffic Sourcing Remains Problematic

Sourcing traffic (any method by which publishers acquire more visitors through third parties) results in greater fraud. Sourced traffic had more than three

times the bot percentage than the study average.

### c. Fraud Varies by Buy Type

- Direct buys had lower fraud. Programmatic buys had greater fraud. The high bot rates in programmatic video were expected given that video CPMs are significantly higher than other types of online media.
- Programmatic display ads had 14 percent more bots than the study average.
- Programmatic video ads had 73 percent more bots than the study average.
- Direct video ads, where measurable, were 59 percent less likely to have bots than the study average.
- Direct display ads were 14 percent less likely to have bots than the study average.

quality (especially providers of the costliest media: video).

## COMBINED DEFENSES CAN DEFEAT AD FRAUD

### Action Steps to Reduce Fraud

Industry stakeholders can work to reduce ad fraud by combining the use of anti-fraud technologies with proactive policies and strategies. No single solution protects any single stakeholder. Only combined, unified defenses can effectively thwart the ad fraud attacks that are coming from all directions.

#### a. The “Sell Side,” Including Publishers, Networks and Exchanges, Must:

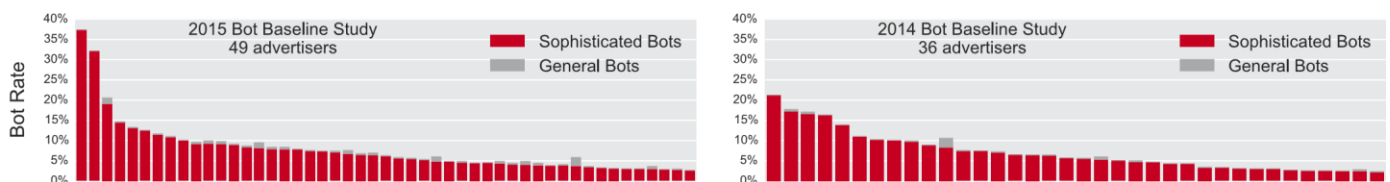
- Relentlessly monitor inventory for ad fraud. Cut off sources that supply bots.
- Consistently maintain transparency and allow buyers to monitor these media investments for

#### b. To Prevent Ad Fraud, Advertisers and Their Agencies Must:

- Be aware and involved.
- Understand the programmatic supply chain and request inventory transparency (especially programmatic video buys that tend to have higher CPM and higher fraud levels).
- Request transparency for sourced traffic.
- Include language on non-human traffic in terms and conditions.
- Use third-party monitoring to ensure compliance with anti-fraud policies.
- Require media quality measurement vendors to demonstrate effective anti-fraud technology and provide measurement transparency.
- Announce your anti-fraud policies to all external partners.
- Support the Trustworthy Accountability Group.

Figure 1: Bot Percentage for All Participants 2015 (left) and 2014 (right)

General bots are detectable using the industry spiders and bots list, while sophisticated bots require more complex techniques to detect.





The background features a complex network of black nodes of varying sizes connected by thin grey lines. Some nodes are significantly larger than others, creating a sense of hierarchy or importance. Dashed grey lines form large, overlapping circular shapes that frame the central text. The overall aesthetic is clean, modern, and technical.

## II. Detailed Findings

# Detailed Findings

## 1. Once Again, No Advertiser Was Immune to Bots

In 2015, the average advertiser's bot rate declined by only 0.2 percentage points compared to 2014. Overall, the 49 participants saw a wider range of sophisticated bot percentages in 2015 (3 to 37 percent) than in last year's study (2 to 22 percent). A quarter of the advertisers encountered bot rates of 9 percent or higher during the study period. The lowest bot level achieved by any advertiser across the study period was about 3 percent.

Much of the media purchased by the typical advertiser is clean, but when fraud does affect an advertiser, it tends to hit hard and in very concentrated areas. In 2014, 17 percent of advertisers were hit the hardest and were paying for 82 percent of the losses. In 2015, 30 percent of advertisers paid for 80 percent of the bots.

## 2. Bot Impacts Ranged from \$250,000 to \$42 million

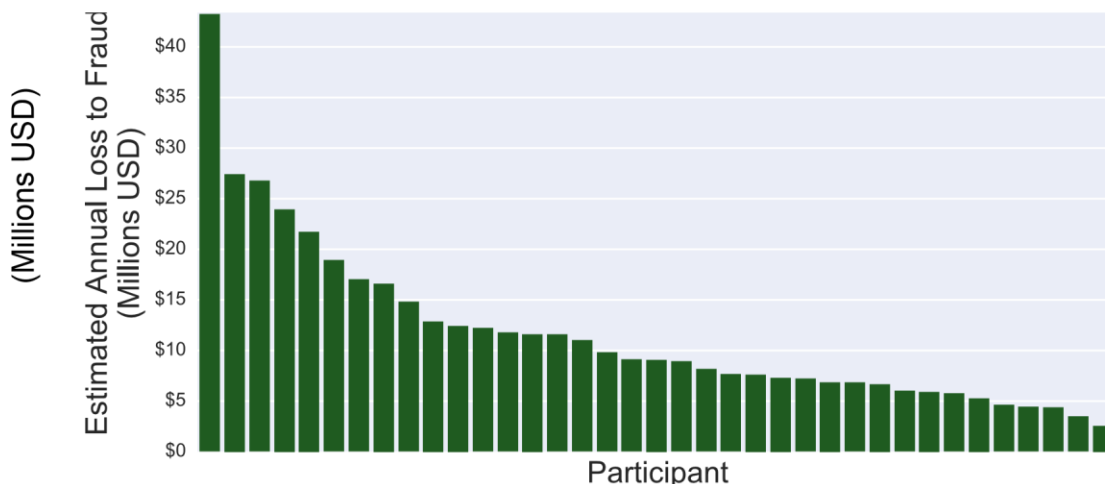
More than 10 percent of participants lost hundreds of thousands of dollars during the study due to "hotspots" — problematic ad campaigns that have high bot percentages. Some advertisers stand to lose tens of millions of dollars annually to the bot hotspots uncovered in this report if they do not change their strategies and buying policies.

**Figure 2: Annual Estimated Losses to Bot Fraud in 2015** Advertisers will lose millions to digital ad fraud in 2016.

The 10 participants with the highest digital ad spend would average \$20 million in estimated annual losses to bot fraud. The 29 participants with moderate estimated digital ad spend would average an estimated \$9 million lost in 2015, and the 10 participants with the lowest estimated annual spend would average \$2 million in estimated losses in 2015. The estimated average annual loss to bots among ANA 2015 study participants was \$10 million.

The participant with the lowest estimated bot impact also worked to reduce the actual cost of bot fraud by adding to insertion orders and contracts the requirement that it would not pay for bots. This participant deployed continuous monitoring technology to enforce its anti-fraud policies and contracts. The combined use of anti-fraud technology, policies, and strategies effectively eliminated the financial bot impact to that participant.

**“Sophisticated” invalid traffic, or bots, is the preferred term used by Media Rating Council (MRC) to describe the traffic produced by automated sources which is not detected by the common whitelists and blacklists used in the industry**

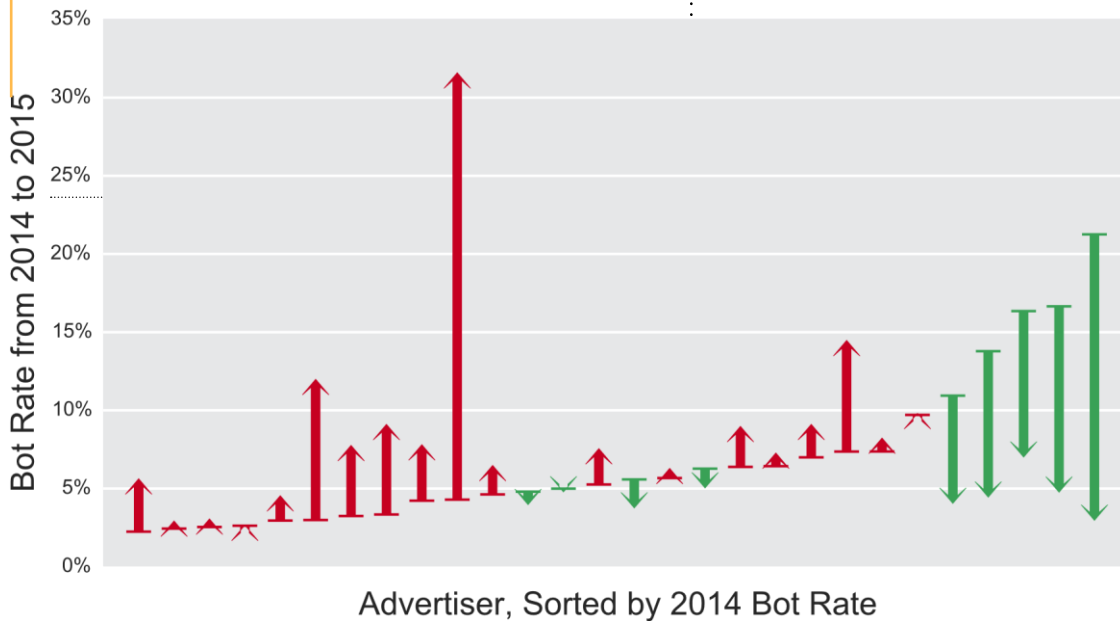


### 3. The Majority of Returning Participants Did Not Improve

In 2015, most returning 2014 study participants saw more bots. About one-third — only nine — of the 28 advertisers which participated in the study in 2014 and returned to be measured in 2015 saw a decrease in their overall fraud rates, suggesting that the problem is visible but has not been solved. Every company which experienced a fraud rate in 2014 of greater than 10 percent showed a decrease in the 2015 study, but all 11 companies with a 2014 fraud rate lower than 5 percent saw an increase in the amount of fraud they suffered during the recent study period.

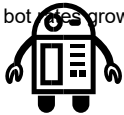
These results underscore that solving the puzzle of digital ad fraud is not a one-and-done project, but requires constant vigilance. Advertisers need to remain focused on fraud reduction to keep the most costly bots at bay.

**Solving the puzzle of digital ad fraud is not a one-and-done project, but requires constant vigilance**



**Figure 3: Most Returning Participants Saw More Bots**

Companies in green reduced their bot rates from the bar's high point to the low point, while companies in red saw their bot rates grow from the low point to the bar's high point.



## How Does Fraud Get into Media?

An ad buy is affected by bot fraud if a supplier between the advertiser and the web site showing ads is sourcing bots or is the victim of someone else who is trying to game the system by making the audience appear larger than it actually is.

AUDIENCE TARGETING MECHANISMS	Ads are served to bots that use stolen or spoofed	cookies or user IDs to exploit:

and programmatic buys  
on:

Ads are served to bots  
when publishers pay for  
visitors from a “botty”  
source, or if they  
partner  
with anyone doing so:

Ads are served through  
sourced traffic to bots on  
bogus sites in long-tail, run  
of network (RON),

**Bots consume ads at any or all of the following stages in the digital advertising supply chain:**

- Fraud can get in at the audience targeting stage, usually at the DMP (Data Management Platform) or DSP (Demand Side Platform) level. Additionally, retargeting in this stage can drive bots that clone real people’s cookies and fool audience modeling systems through all stages of the advertising process.
- Fraud can get in at the network or exchange level if a network or exchange has publishers sourcing traffic that includes bots.
- Fraud can get in at the publisher level if a publisher sources traffic to fulfill inventory requirements from companies that sell bots (note: this may be either knowingly or inadvertent). Publishers will also be vulnerable to fraud when they allow other sites to feature their content — known as audience extension — if the other sites source bot traffic.

## 4. Bots Continue to Evade Detection and Create High Costs

Bot operators continue to reap significant revenue from ad fraud. The most costly bots fool the detection systems of advertisers and suppliers by using freshly infected PCs and complex malware, while simpler bots use stale infections or more basic malware to gain profits from stakeholders who write off fraud as a business risk and do not focus on fighting fraud.

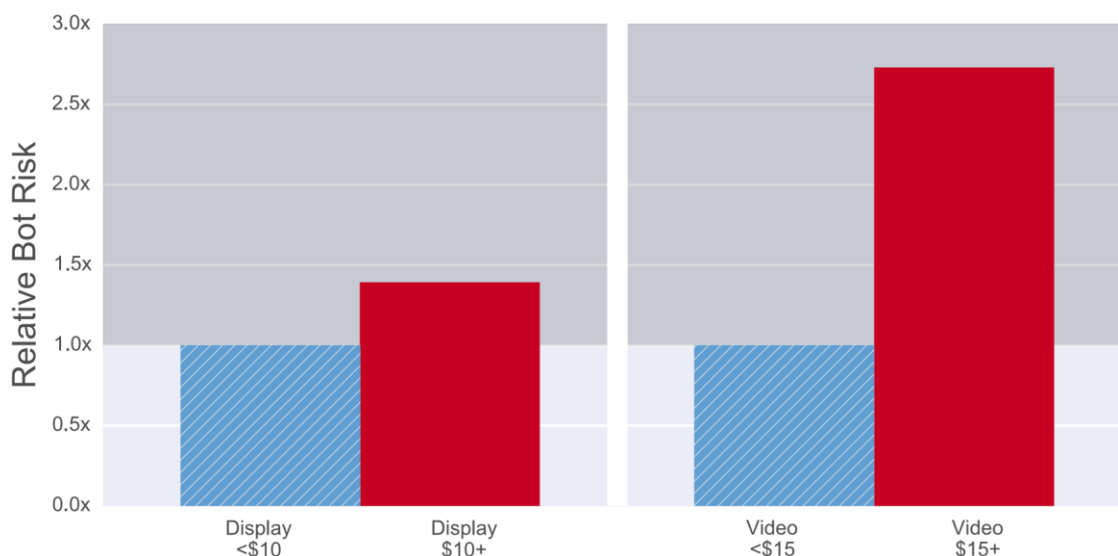
Advertisers and suppliers must defend against complex and simple bot populations using a combination of anti-fraud policies and bot detection technologies.

### a. More Valuable Ad Inventory Is More Susceptible to Bots

The display advertising campaign with advertising priced at or over \$10 per thousand impressions (CPM) had a median 1.39 times more bots than inventory priced below \$10 CPM.

The impact of higher CPMs is even more pronounced in video advertising. Video media campaigns with \$15-or-greater CPM had a median 2.73 times more bots than campaigns with less than \$15 CPM.

**Advertisers can reduce their actual bot cost by combining the use of anti-fraud technologies with policies that prevent payment for ad fraud**



**Figure 4:**  
Higher CPM  
Campaigns See  
More Bots



### b. Programmatic Ads Attracted the Most Bots — Direct Buys Were Cleaner

Buy type affected the bot rate significantly. In 2015, programmatic video advertising continued to attract more bots (as a proportion of overall traffic) than other types of advertising. With video CPMs remaining significantly higher than other types of online media and providing a stronger economic incentive to commit fraud, this was expected.

Advertisers which bought programmatic video had 73 percent higher bots than the study average, with a bot range in video programmatic buys of 1 to 70 percent. Only a small percentage of purchased direct buy video media was measurable due to publishers' transparency and measurement policies that did not permit full fraud assessment on the buys. However, on what was measurable, the range in bot percentages among participants was small.

Direct display media generally had lower risk of bots (0.86 fraud multiplier), with a wide range of bot percentages among participants. Direct video media, where measurable for the study, had the lowest fraud multiplier: 0.41.

## The fraud multiplier shows the relative vulnerability for bots compared to the study average of 1.00

Media Type and Buy Type	Bot Percentage Range	Fraud Multiplier
Direct Video	2–5%	0.41
Direct Display	2–40%	0.86
Programmatic Video	1–70%	1.73
Programmatic Display	2–30%	1.14

Table 1: Sophisticated Bot Ranges by Media Type and Buy Type for Study Participants

## Malware Will Increasingly Target the Advertising Ecosystem

There are many ways a user's computer comes under the control of a fraud operator — outright remote compromise via “drive-by downloads” exploiting a vulnerability, “bloatware” shipped with computers, black-box libraries unwittingly embedded into otherwise legitimate applications, or install wrappers that add remote-controlled services along with some functionality the user desires.

There are many ways to gain access; what's interesting is what's done with it. Advertising fraud has the curious status of almost seeming legitimate — you couldn't expect to get away with raiding a bank account or accessing someone else's Gmail account, but defrauding advertisers, even by using the host user's identifying cookies, doesn't seem nearly as criminal. While the ecosystem suffers, the end user sees very little impact from the fraud.

For the bot operator, however, the scheme is quite profitable. Many do not even operate their own infrastructure. So this sort of fraud has a surprising number of “legitimate” participants. We've found companies where not everyone at the company knew they were fraudulent operations.



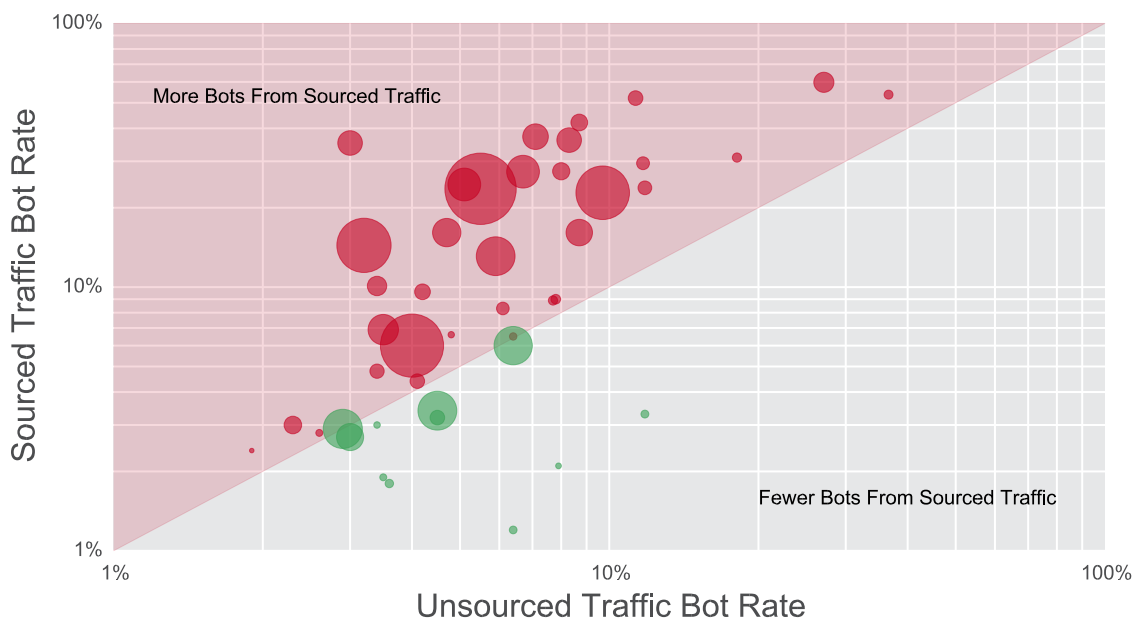
**Dan Kaminsky** Co-Founder and Chief Scientist, White Ops

### c. Sourced Traffic Continues to Move Bots into Media Buys

More than three-quarters of participants (red in chart below) had higher bot percentages in traffic bought from third-party sources compared to unsourced traffic.

Overall, sourced traffic was more than three times more likely to contain bots than unsourced traffic. Sourced traffic in 2015 showed a slight improvement over 2014, when sourced impressions were over four times more likely to come from bots.

**Sourced traffic was more than three times more likely to contain bots than unsourced traffic**



**Figure 5: Sourced Traffic Generates More Bots**

Sourced traffic was more than three times more likely to contain bots. Bubble area is proportional to traffic volume.

#### **d. Where Some Suppliers Improve, Others Show Higher Bot Rates**

While ad fraud due to bots has largely remained stable on average, across buyers, providers, exchanges, and ad tech platforms, fraud levels have changed on a granular level since the 2014 study. These changes are in response to not just anti-fraud technology, policies, and strategies, but also ongoing organic changes in the global ad fraud ecosystem.

Aggressive efforts to eliminate fraud at one major exchange had a substantial overall impact on the distribution of bots seen across this year's study data. In 2014, that exchange's traffic comprised 8.4 percent of all traffic and 24 percent of all sophisticated bots observed, with 31 percent of that exchange's own traffic consumed by sophisticated bots. In 2015, this exchange made a substantial effort to clean up its traffic. With approximately the same impression volume, the exchange supplied just 5.3 percent of the sophisticated bots across the 2015 study and lowered the sophisticated bot percentage of its own traffic to 6.5 percent.

Conversely, bot sources have been consistently observed not to simply shut down when blocked from their current pool of targets. There are large portions of the ecosystem that remain unprotected or less stringently protected, and bots target those exchanges, shifting to platforms and domains where their current methods still work. In other cases, bots move away from one target toward a more lucrative one when economic pressure is applied. Some bot operators, for example, have shifted the focus of their attacks from display to video, which pays a premium.

Because CPMs vary widely, the highest financial impact from bots does not necessarily come from the suppliers with the highest bot percentages. In 2014, the supply platform with the highest bot levels accounted for 24 percent of all the bot impressions in the study but only accounted for a small amount of the dollar losses. After purging the obvious bots from its supply this year, it accounted for only 6 percent of the bots in the 2015 study. However, because of the higher price point, even with the lower bot percentage, this platform accounted for approximately the same dollar losses due to bots as it did in 2014.

**Because CPMs vary widely, the highest financial impact from bots does not necessarily come from the suppliers with the highest bot percentages**

**Bots have been observed not to simply shut down, but to shift to other targets, when blocked by a stakeholder or when economic pressure is applied**

### e. List-Lookup-Based Programmatic Bot Prevention Did Not Work

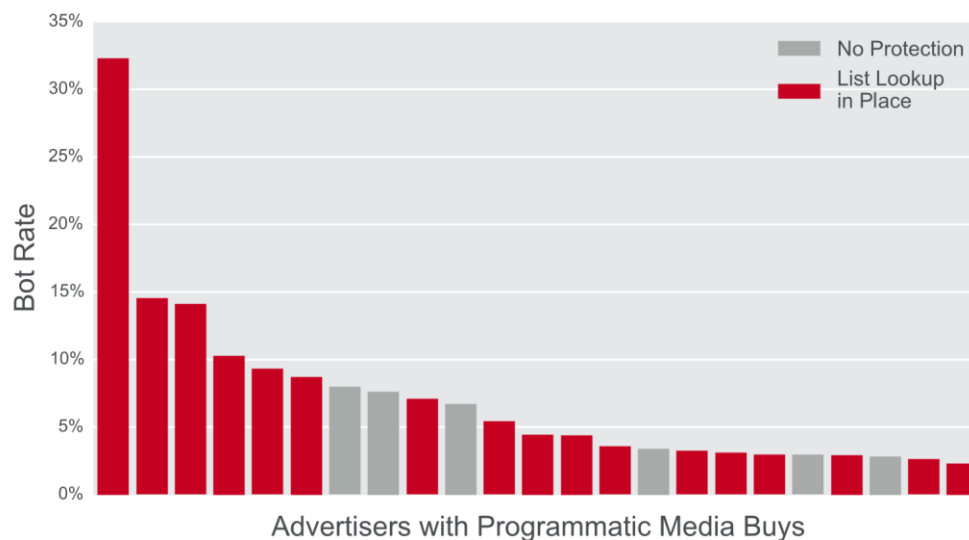
The increased bot exposure in higher-value inventory indicates that monitoring the highest value inventory is essential for lowering the financial impact from bot fraud in media investments. When protections such as fraud detection or prevention are put in place but not accompanied by proactive anti-fraud policies and anti-fraud responses to detection results, fraud can infiltrate media buys. The buyer's false sense of security can put it at higher risk.

Three in four programmatic media buyers participating in the 2015 Bot Baseline study were protected by fraud prevention solutions that relied on general fraud detection as defined by MRC (Media Rating Council). The general blocking solutions used list lookup in programmatic buys to prevent fraud rather than sophisticated detection that relied on more dynamic, security-based methods to identify fraud. Sophisticated bot prevention as defined by MRC uses bot impression behaviors to identify and

The programmatic buys protected by general/listlookup-based solutions did not show increased impression validity over unprotected programmatic buys. MRC itself does not recommend solely relying on general invalid traffic techniques.

Security-based fraud detection and blocking can be key tools in combating ad fraud, but maintaining accountability and transparency in all layers of the supply chain, including detection and prevention vendors, is required in order to effectively defend against fraud.

## The programmatic buys protected by general/list-lookup-based solutions did not show increased impression validity over unprotected programmatic buys



**Figure 6: Programmatic Bot Percentage Without Prevention and with List-Lookup-Based Blocking** List-lookup-based blocking did not protect programmatic buys.

prevent bots from consuming media; general bot prevention relies on a list-based approach to detect and block bots.

## 5. Bot Operators Continue to Hide Bot Traffic Among Humans

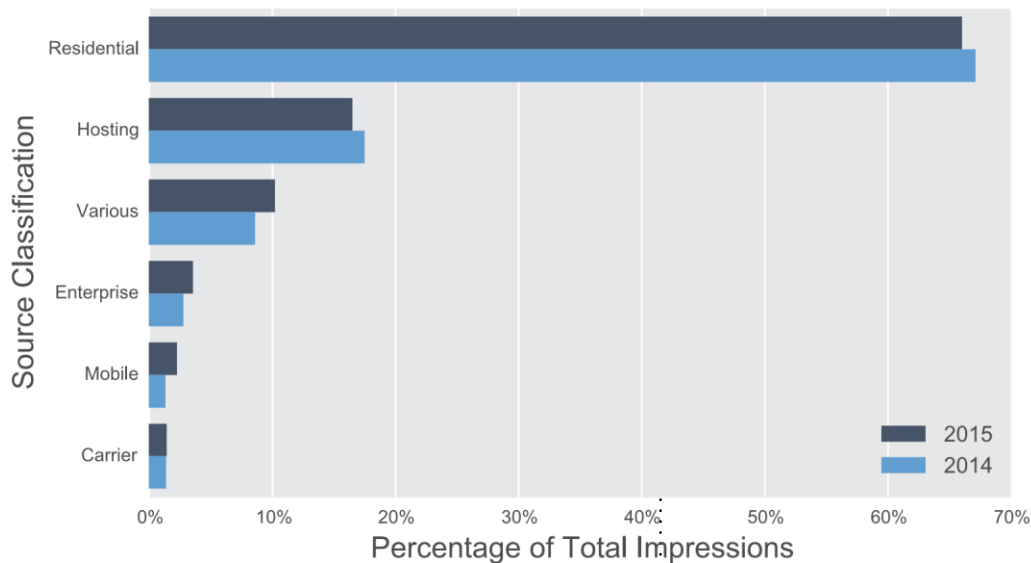
Among the sophisticated bot population not identifiable using the industry bots and spiders list, bots exhibit behaviors of varying complexity. More complex bots can mimic human browsing behaviors, while more basic bots are easily identifiable as bots by machine learning and statistical detection methods.

### a. The Majority of Bots Come from Residential Internet Addresses

Household computers accounted for the majority of bots seen by advertisers. Two-thirds of all sophisticated bot traffic came from residential Internet addresses. The use of residential IP addresses makes countermeasures based on blacklisting certain blocks of Internet addresses a difficult trade-off, as blacklisting removes valid human impressions with the blocked bot impressions.

**More complex bots can mimic human browsing behaviors, while more basic bots are easily identifiable as bots by machine learning and statistical detection methods**

NO CHANGE



**Figure 7:**  
IP Addresses  
of the Sources  
of Bots in 2014  
and 2015

Bot operators are using an increasing number of complex techniques to disguise their bots as humans. In addition to mimicking patterns such as time-based human behavior, the vast majority of the bots came from home networks, often using the existing browser cookies to appropriate real identities and appear as members of certain desirable demographics (see page 23, “Case Study: Advertiser’s Targeting Drove Bots to Its Own Site”).

The ability of bots to masquerade as legitimate human users is the by-product of a compromised browser. Whatever identity is associated with a browser is represented in all actions the browser takes; humandriven or not. This leads to bots adopting common targeting characteristics such as geography, age group, browsing histories, and any and all other demographics used to target ads.

The second most popular source of sophisticated bots were Internet addresses belonging to companies that host web servers and other systems, which accounted for 16 percent of sophisticated bot traffic.

The distribution of the main sources of bot traffic is almost identical to the 2014 Bot Baseline study, with slightly more bots coming from enterprise networks and mobile sources.



**b. The Bots on Infected Machines Are a Moving Target for Advertisers**

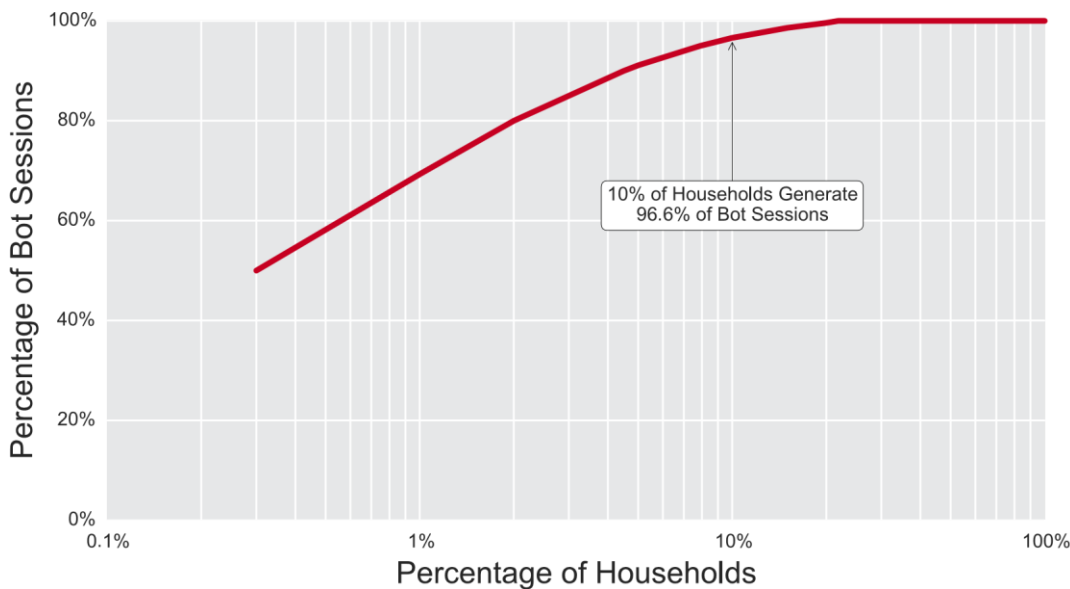
The fraud that is responsible for the majority of advertiser losses comes from the most freshly infected computers, since they have not yet been flagged as sources of bots in blacklists. These bots are unknown to the blocking mechanisms in general detection systems and cannot be blocked using machine learning.

At any one time, a relatively small number of households account for most of these successful bot impressions. Because the bots are running on real people’s computers, these same households are the source of real human visits as well. Almost 80 percent of successful bot traffic came from the 2 percent of households with the freshest malware infection. But that leading 2 percent changes constantly, as old or obvious bots get detected and new computers are infected to take their place. Over

the two-month period of this study, infections stretched across an incredible 10 percent of all the residential IP addresses seen.

This is why the bot problem continues to vex this industry. It is not enough to detect and even block bot traffic. If old infections are not discovered faster than new infections are made, all those efforts have zero impact on the profitability of the bots. To prevent the bots in this “profit window” (see page 4, “Closing the Profit Windows of Bots”) from raiding an ad buy, advertisers and suppliers must monitor for fraud using sophisticated detection methods and block new bot infections using sophisticated bot prevention technologies. Combined with proactive anti-fraud policies, sophisticated detection and prevention technologies can significantly reduce the threat from fresh bot infections that are in the profit window for botnet operators.

**Be careful how you block: Blacklisting removes valid human impressions with the blocked bot impressions**



**Figure 8:**  
The Proportion of Households That Account for Bot Percentage

### c. The Geographical Distribution of Bots Resembles Human Populations

Because sophisticated bots overwhelmingly come from malware-infected computers from residential IP addresses, the distribution of their sources concentrates in large metropolitan regions, resembling the distribution of people.

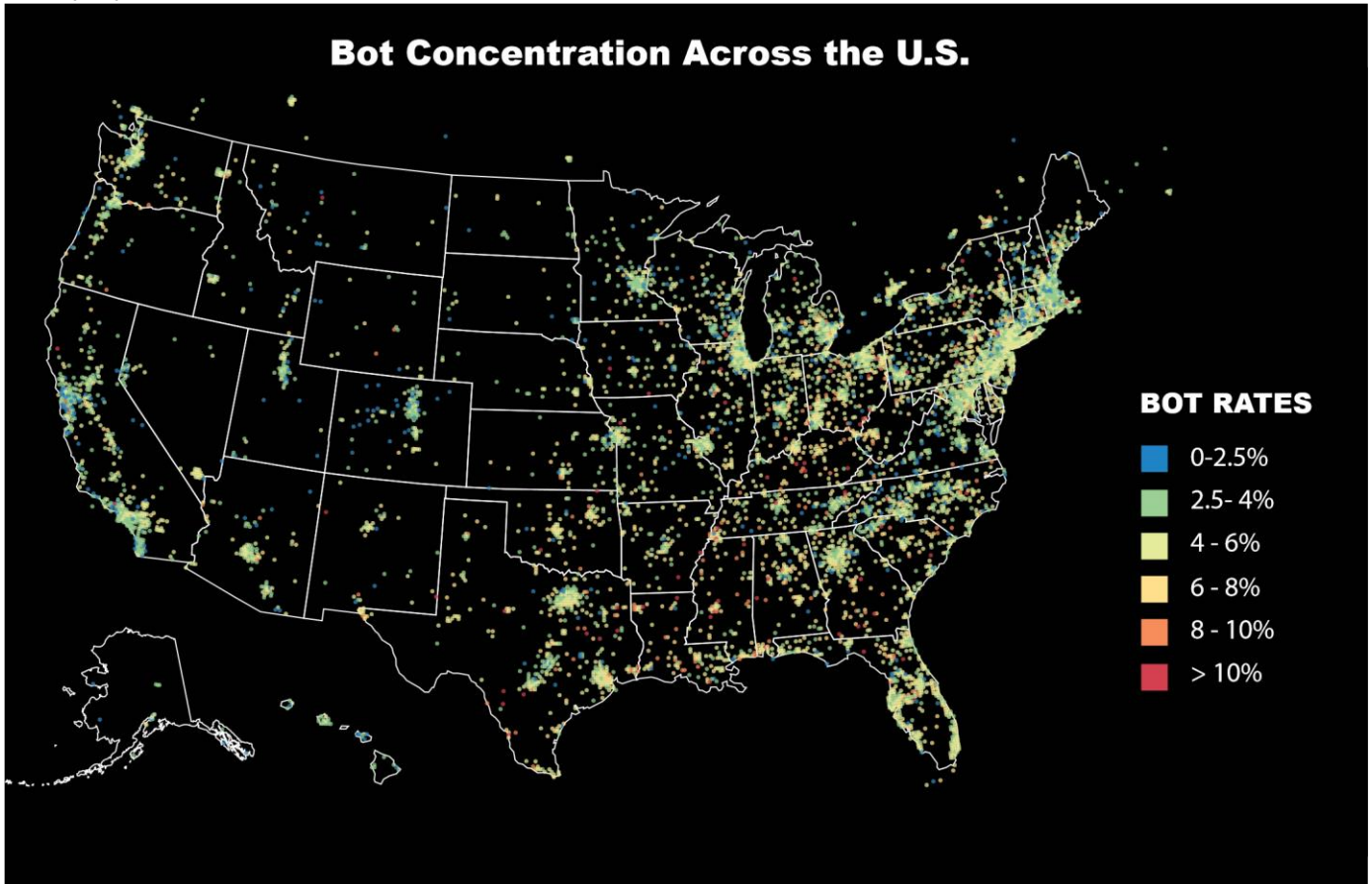
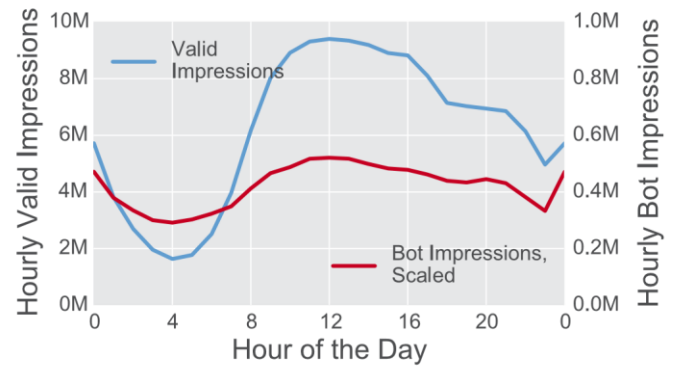


Figure 9: Map Showing Bot Concentrations Across the U.S.

NO CHANGE

#### d. Bot Operators Mimic Human Behavior by Focusing on Daytime Hours

Most people sleep at night and are awake during the day. Online advertisers know this and tend to target their online advertising to their target audiences' daytime hours. While bot operators likely know this as well and may ramp up activity during the day and decrease activity at night, the regular pattern of computer use — with most computers off at night — is likely responsible for bots mimicking a normal human's waking hours.



**Figure 10: Bot and Human Traffic Patterns Throughout the Day**

### e. Basic Bots Account for a Greater Proportion of Traffic at Night

Conversely, because human activity declines at night, some bots appear more active, comprising a larger portion of all traffic. The relative nighttime peak in fraud reached 15 percent of total traffic in the 2015 study, down from a peak of about 26 percent in the 2014 study. Both studies show a similar pattern of behavior, but the smaller proportion of bots in the nighttime hourly traffic could indicate that operators are doing a better job of shaping their traffic to resemble human traffic as their fraud operations become more complex.

### f. Bots Are Less Active on Weekends

Similarly, the operators of advanced bots mimic humans by preferring to send traffic to sites during weekdays, when there is a greater amount of legitimate human traffic as well. The decline of human activity on the weekend — with lower peaks during daytime hours — means that bots account for a greater proportion of traffic, but still tend to mimic human patterns of browsing in a complex manner.

**Bots that fill inventory for ad buys of specific demographics and locations exploit advertising orders for audiences which are typically in short supply**

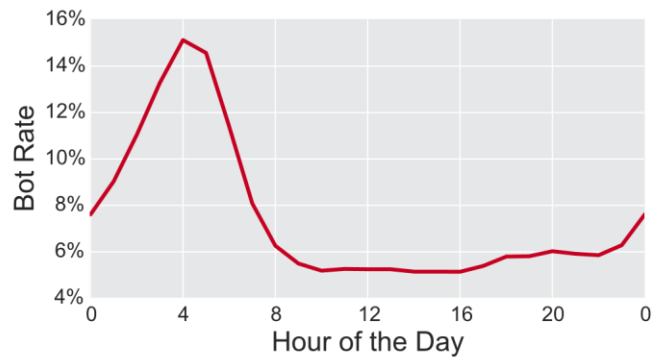


Figure 11: Percentage of Traffic Made Up by Bots, on Average for All Data

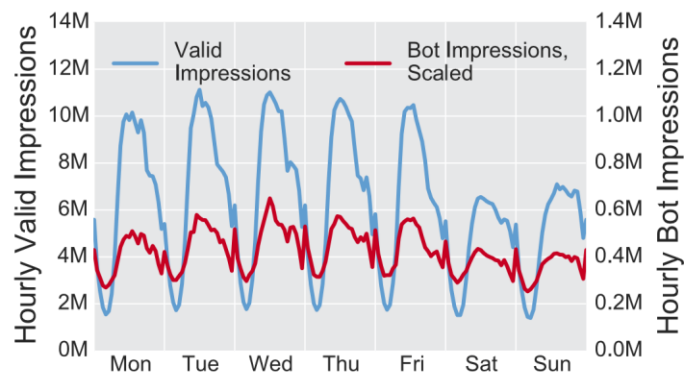


Figure 12: Weekend Distribution of Traffic by Impression Type

## 6. Bots Get in When Targeted Audiences Do Not Meet Demand

Marketers want to target specific demographics of consumers, whether high-income buyers of luxury goods, Hispanic home owners, or young couples living in California. Bots that fill inventory for ad buys of specific demographics and locations exploit advertising orders for audiences which are typically in short supply. These bots make a greater profit at the expense of advertisers seeking more targeted audiences.

The study saw much higher bot percentages in certain advertising campaigns based on demographic targeting or retargeting potential consumers. For example, in one campaign, retargeting previous web visitors resulted in 18.3 percent bots detected among nearly 38 million impressions.

### a. Bots Fill Hard-to-Reach Demographic Quotas

White Ops discovered a number of campaigns that were dominated by bots representing themselves as desirable demographics of limited supply. In one

campaign, for example, more than four million impressions provided by a single publisher appeared to be mostly young Asian visitors, but in fact were 96 percent bots.

This fraud is mainly caused by malware-infected home computers and laptops (see page 17, “The Majority of Bots Come from Residential Internet Addresses”) that are able to masquerade as human. The malware on the infected system was “driving” the same user’s browser, allowing it to use the cookies of the human owner of the machine and registering the demographic and targeting profiles seen in the 96-percent-bot segment as above.

If a particular nation, state, or city has an endemic infection of ad-fraud malware, that can have a strong impact on the resulting demographic profile — in this case, a young Asian audience.

### b. Hispanic Targeting Increases Bots

The impact of bots’ demographic targeting can be seen among campaigns that targeted Hispanic users. Sixteen study participants out of 49 reported Hispanic-targeted media to the study, totaling 300 million impressions.

The 50 top-volume domains targeted using Hispanic demographic data show that Hispanic-targeted campaigns are often more bot-infested than a

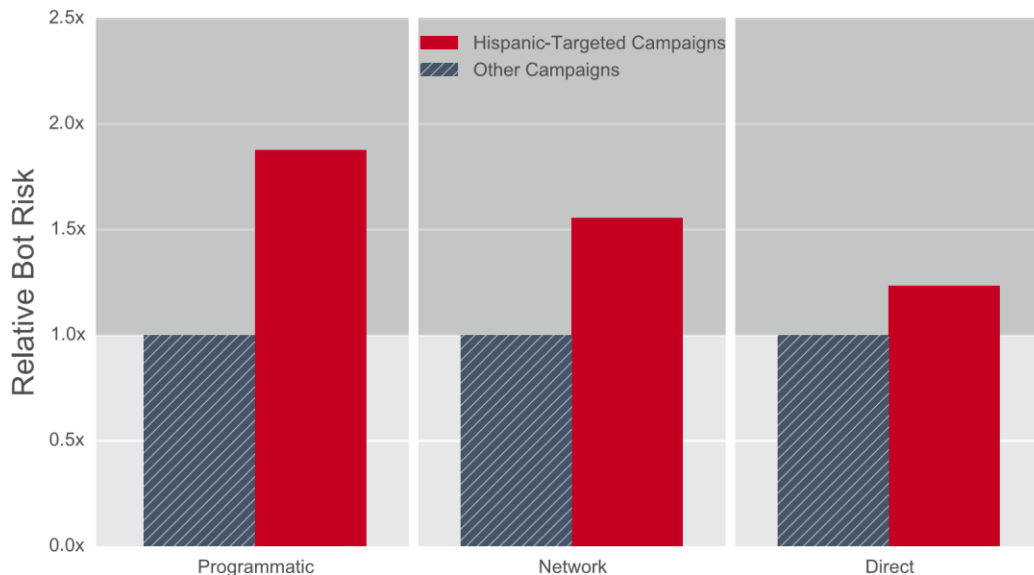


Figure 13:  
Hispanic-Targeted  
Campaign Bot Fraud  
Multipliers Compared  
to Non-Hispanic

nontargeted campaign served on the same domain. Across the highest-volume domains served by Hispanic-targeted campaigns, nearly all had higher bot rates, and many had bot rates near 100 percent.

Programmatic buys with Hispanic targeting were nearly twice as likely to encounter bot traffic than non-Hispanic-targeted programmatic media, with a fraud multiplier of 1.7. Hispanic-targeted network

buys had a fraud multiplier of 1.6 compared to the study average for network buys. Direct buys with Hispanic targeting had slightly increased bot percentages, with a fraud multiplier of 1.2.

## Programmatic buys with Hispanic targeting were nearly two times more likely to encounter bot traffic than non-Hispanic-targeted programmatic media



### c. Compromised Systems Allow Bots to Look Like Humans to Advertisers

Retargeting is a good way to advertise to interested users. By only purchasing ad space for users who have previously expressed some form of interest in the product, advertisers can filter out uninterested audiences. Recognizing that a user previously expressed interest in a site or product does not provide an effective method for reducing bot traffic in programmatic advertising. Bots are able to infiltrate retargeting segments and reap the higher CPMs advertisers pay to reach them.

In one investigation of a retargeting campaign, the bot rate was found to be 37 percent. The campaign's 10 million impressions delivered during the study period were spread across nearly 360,000 unique users and could be divided into three distinct categories: Human, Compromised, and Synthetic.

Almost 90 percent of the agents, representing about 57 percent of the advertising traffic, were entirely human. The remaining volume of advertising impressions was served by either compromised machines or entirely synthetic audiences. Synthetic audiences — agents with 100 percent bot traffic — were able to enter the campaign's targeting segment despite failing to exhibit true human behavior. This traffic came from a 4 percent subset of all agents and comprised 3 percent of all traffic.

The most bots came from the compromised audience, where agents are mixed human and bot traffic. While only making up a small number of agents — 5 percent — the compromised segment created 40 percent of all advertising traffic, and, with its 85 percent bot rate, made up about 92 percent of all bots seen by the campaign. These infected machines are able to drive a disproportionate amount of bot traffic,

as they are well disguised and they spearhead botnet profitability.

## Bots are able to infiltrate retargeting segments and reap the higher CPMs advertisers pay to reach them

Audience	Impression Volume*	Accounts for X-Percentage of User Agents	Sophisticated Bot Percentage
Human	57%	86%	0%
Compromised	40%	5%	85%
Synthetic	3%	4%	100%

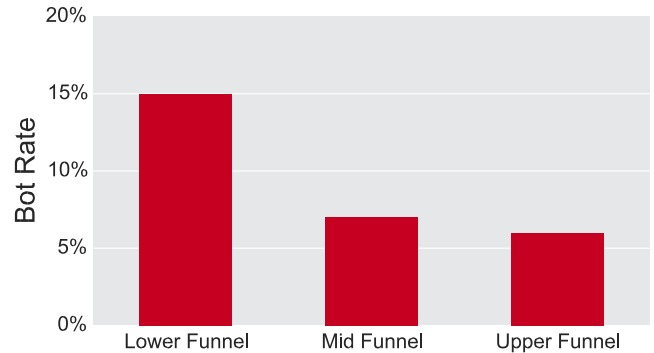
**Table 2: Makeup of Three Audiences in One Retargeting Campaign**

**d. Advertiser's Targeting Drove Bots to Its Own Site**

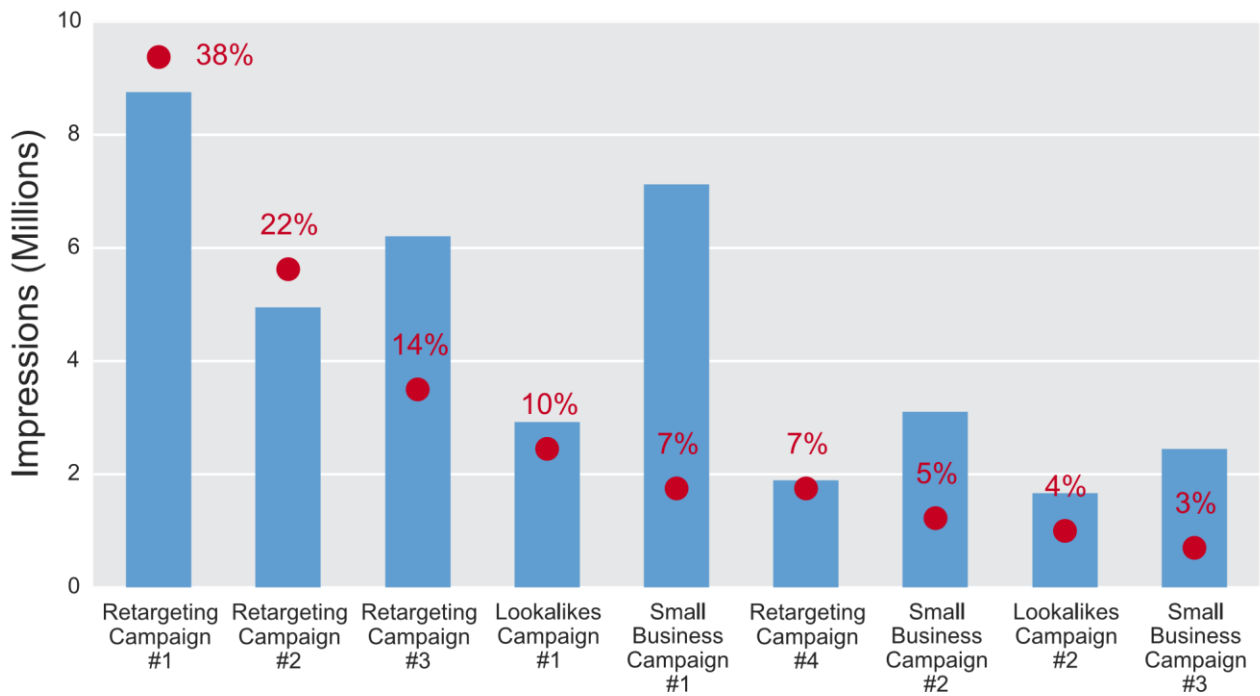
An advertiser retargeted visitors to its own e-commerce site, but the advertiser's targeting drove non-human traffic from awareness and branding campaigns to the advertiser's e-commerce site.

The bots were visiting the participant's e-commerce site to collect high-value impression cookies, including shopping cart abandonment and other interactive cookies. The advertiser's lower-funnel campaigns — those focused on closing a sale — saw 15 percent sophisticated bots on 40 million impressions in high-impact media (see chart at right).

This participant saw 38 percent sophisticated bots in its highest-volume campaign, which focused on retargeting, and winning back, potential customers. Other retargeting campaigns had 22 percent, 14 percent, and 7 percent sophisticated bots. Campaigns that did not retarget saw a range of 3 percent to 10 percent sophisticated bots (see chart below).



**Figure 14: Sophisticated Bot Percentages by Funnel in an E-Commerce Site**



**Figure 15: Bot Percentages and Impression Volumes in E-Commerce Campaigns** Red dots display the bot rate for the campaign.

## 7. Bots Use Complex Techniques to Increase Profits

Bot operators continue to develop their techniques and infrastructure to better profit from the advertising ecosystem, focusing on specific types of domains, aiming for higher advertising payouts in video, and creating better spoofing viewability to fool anti-fraud technology.

rendering capabilities, built into nearly all modern browsers, will naturally respond to JavaScript and Flash viewability measurements.

For measurable viewable impressions (impressions consumed by visitors whose browsers executed JavaScript measurement tags on a web site), the average sophisticated bot viewable rate was only 4 percentage points less than that of humans. The chart below shows how sophisticated bot traffic, designated by the red population, resembled the human population, designated in blue, both in measurable and viewable rates for nearly every advertiser in the study.

**The average viewable rate of sophisticated bot traffic is 43 percent, closely mimicking the average human viewable rate of 47 percent**

NO CHANGE

### a. Bot Operators Are Successfully Fooling Viewability Measures

Bots are successfully fooling viewability measures, according to data collected in the study. The average viewable rate of sophisticated bot traffic is 43 percent, closely mimicking the average human viewable rate of 47 percent.

The technical requirements for responding to a viewability measurement are basic JavaScript and Flash functionality. This comes standard with any desktop browser environment, and if a bot is simply controlling a browser session or scripting a browser rendering engine — as it most often is in the case of ad fraud — then the mere loading of a web page will exhibit viewability with no extra sophistication on the part of the bot. The presence of JavaScript and Flash

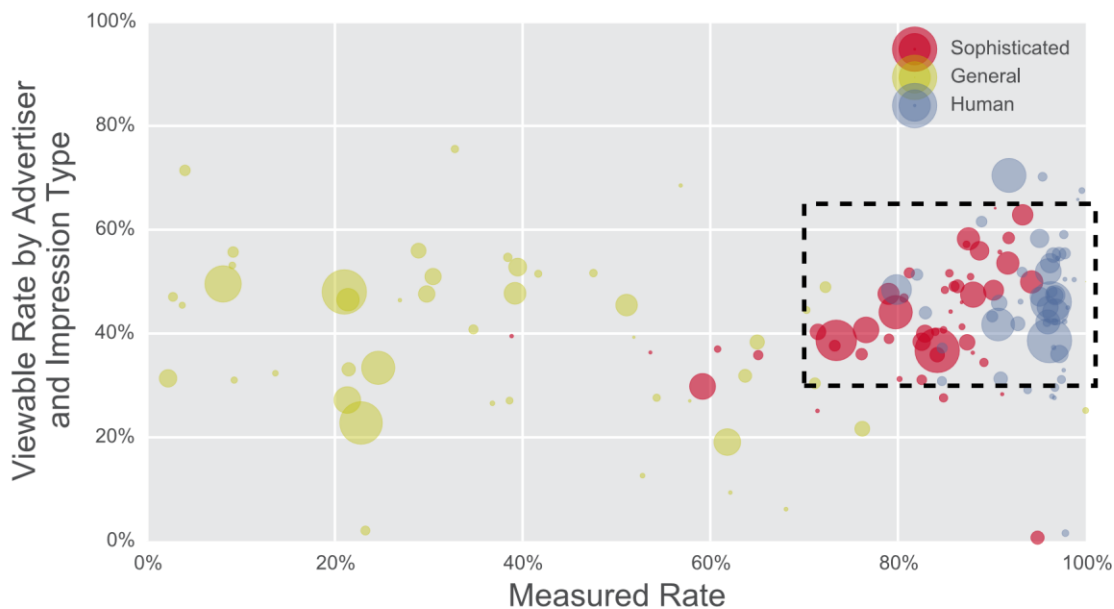


Figure 16:  
Viewability Rates  
and Impression Types

### b. Bot Traffic Favors Certain Domains, or Vice Versa

Domains focused on particular industries tended to have more bot traffic. Travel had the highest bot rate, with 17 percent of ad impressions identified as bots. Business, family, and finance sites were the next highest groupings. In the 2014 Bot Baseline study, bot traffic trended highest on financial sites (a 22 percent bot rate), family-focused domains (18 percent bots), and food-related domains (16 percent).

The shift in bots in 2015 from finance, family, and food to travel, business, and family domains likely does not reflect a change of focus on the part of bot operators. Rather, it likely reflects a change in buyer focus. As marketers change their targeting goals, bot traffic fills

in the gaps between what marketers want to reach and the real online audience.

**Travel domains had the highest bot rate, with 17 percent of ad impressions identified as bots**

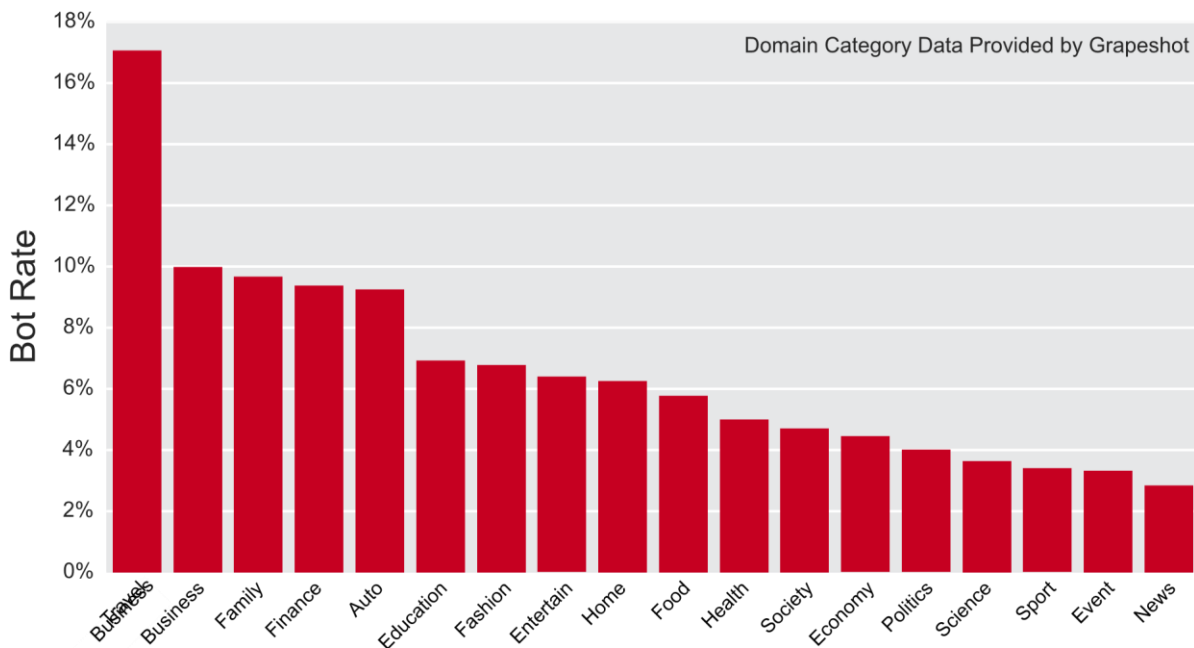


Figure 17: Bot Rate by Publisher Domain Category



**c. Advertisers Buying on the Same Web Pages**



### Can Have Dramatically Different Bot Rates


We observed in our data a top 2000 Alexa publisher experimenting with traffic sourcing for some of its subdomains. While most of the site attracted less than 2 percent bots, particular subdomains started showing much higher bot rates. We saw advertising from at least 21 study participants on these particular subdomains, and only four were able to achieve 10 percent or lower bot rates. The rest showed 45 to 90 percent bot rates. We also observed that different advertisers could buy on the same page over the same period of time and achieve dramatically different bot rates.

An advertiser cannot optimize against what it does not measure. Certainly, there are some predictors of high

bot rates — such as traffic sourcing, nighttime activity, hosting traffic (traffic originating from server IP addresses), very old browsers, programmatic buying, and non-premium publishers — but within every predictor there are examples of advertisers that achieve low bot rates and others that achieve high bot rates. Even within a particular premium publisher, on a particular page, over the same timeframe, advertisers can have markedly different bot rates. To fight ad fraud under these conditions, advertisers must be able to measure impression validity precisely and put policies and strategies in place to reduce fraud.

The good news is that proactive measurement and remediation by those hurt by fraud,






combined with evolving anti-fraud policies and strategies as recommended by organizations including TAG (Trustworthy Accountability Group), the IAB (Interactive Advertising Bureau), and the ANA (Association of National Advertisers), have already been shown to have a measurable impact.

The bot percentages of the five returning Bot Baseline participants with the worst impression validity in 2014 all improved dramatically — by an average of 11 percent. These participants' strategies and anti-fraud policies varied, but they had one thing in common: the strong intention to reduce the fraud in their media.

**An advertiser cannot  
optimize against what  
it does not measure**



## Complex Bot Behavior Does Not Require Brilliant Artificial Intelligence

Advertising bots can be quite simple. They just have to copy the behavior of the real people using the machines they've compromised. Bots built with slightly more advanced programming can mimic human behavior, fooling even more advanced data analysis. As we dug deep into the traditional defenses of the advertising ecosystem, trying to determine

why the size of the problem is so bad, we consistently encountered the mistaken belief that the malware we're up against must, with enough data, look "robotic." Unfortunately, only the most basic malware works that way.

The good guys haven't been asleep on the job, but they've been fighting the wrong fight. Catching complex bots that cost advertisers the most requires the identification of traffic patterns that look like humans, not just finding the basic bots that behave like robots.

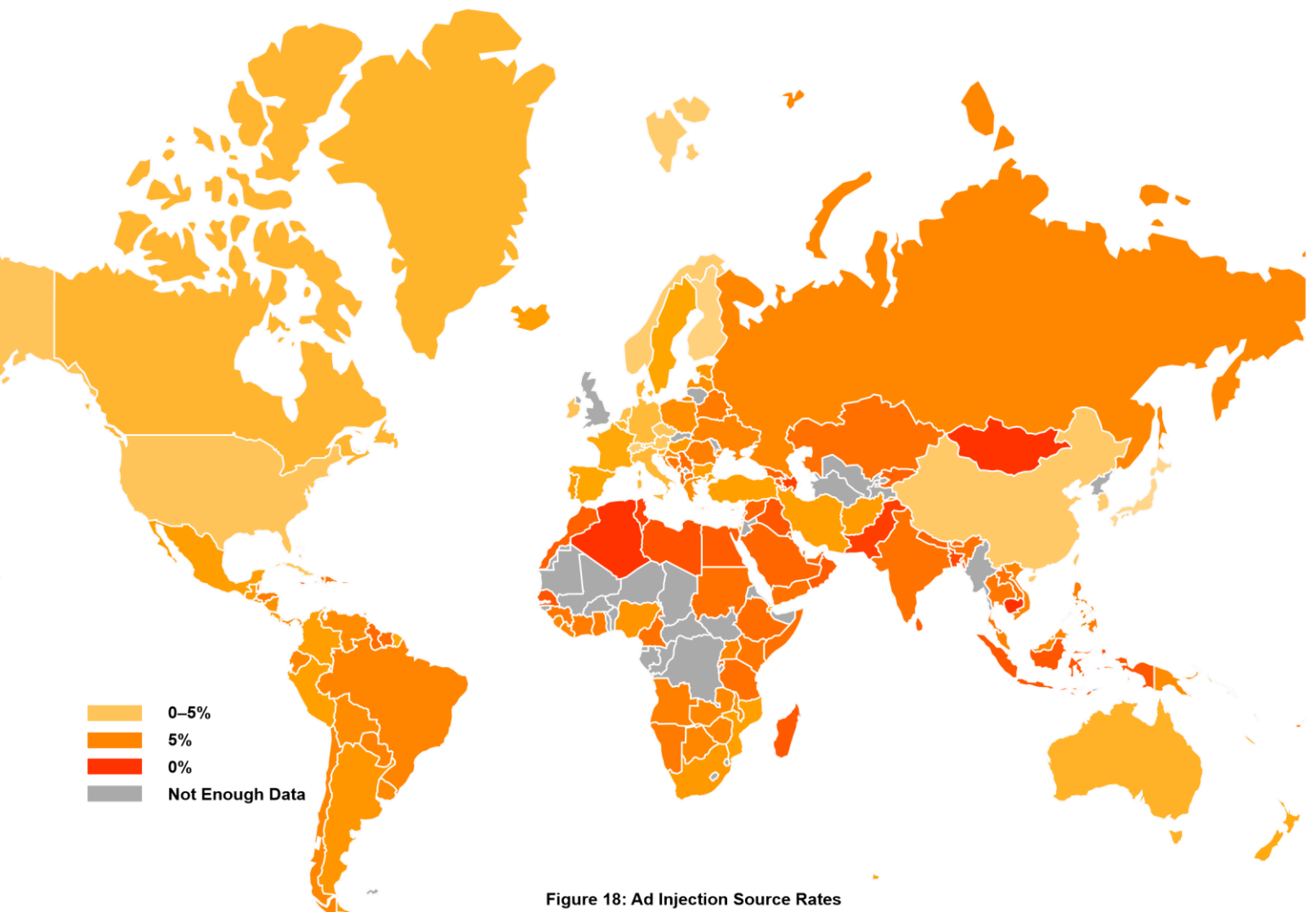
#### d. One Publisher Had 5.8 Percent Ad Injection

A special type of fraudulent traffic is generated by ad injection (the insertion of ads into an app, web page, or other online resource without the consent of the publisher or operator of that resource). Ad injection occurs when a third party intercepts web content on its way to the user, injects ads into the content without the consent of the content owner, and is paid for the ad impressions, undermining the online advertising ecosystem.

An analysis of a single premium publisher found that about 5.8 percent of visitors had some sort of ad injector present that displayed its own advertisements, resulting in an increase of 6 percent in impression volume to the site.

Over 250 different ad injection agents were present in these users, mostly from voluntary toolbar downloads and browser extensions, although the intention of these toolbars are not always obvious to the users who install them, especially in the case of software bundlers. The sources of the ad injection traffic to a particular publisher are shown below.

**Ad injection increased a single publisher's total impression volume by 6 percent**



## 8. Mobile: An Emerging Frontier

Malware is a major source of bots from desktops and laptops but has a very small infection footprint in the mobile ecosystem, particularly in the U.S. It continues to be difficult to propagate mobile malware at a wide enough scale for any significant level of mobile-driven bot traffic. Current infection rates of mobile devices are extremely low. Google typically reports that less than 1 percent of Android devices that connect to the Google Play store have any potentially unwanted software. In other markets, particularly emerging markets such as India and China with third-party Android app stores that often propagate fraudulent versions of legitimate apps, malware rates can be greater.

Because of the lack of information about mobile impressions, the study focused primarily on nonmobile visitors, with limited analysis of mobile traffic. However, the proliferation of mobile devices is clearly evident in the 2015 Bot Baseline study, with 38 percent of impressions originating from devices that report as mobile. Even with this high volume, bot populations in mobile are lower, as attackers have less of a malware footprint in mobile.

Despite the immature mobile fraud market, the threat models for mobile fraud are something to watch closely in 2016 as additional users migrate to this medium and ad pricing, volume, and economic opportunity begin to create more parity with the desktop counterparts.

We have identified three core threat vectors for mobile fraud:

### a. Desktop Botnets Impersonate Mobile Environments

This form of fraud includes mobile impersonation, where botnets or server clusters may:

- Impersonate mobile devices by manipulating the reported user agent string
- Spoof programmatic mobile ad requests that appear to come from mobile devices using

specific ad network or exchange software development kits

In the current study, White Ops observed that the majority of mobile fraud was not actually from mobile devices. A significant number of mobile-targeted advertisements were viewed by apparent desktop-driven bots impersonating mobile, despite having originated from major exchanges that were meant to be delivering mobile inventory.

Though mobile campaigns delivered only 6.7 percent of impressions to desktops, 85 percent of all bots on mobile-targeted inventory came from desktops. While mobile devices are vulnerable to malware and apps that make invisible ad calls in the background, desktop machines or servers offer greater processing capabilities with fewer power and connectivity constraints than mobile. This issue exists across large exchanges and mobile-only exchanges.

**Despite the immature mobile fraud market, the threat models for mobile fraud are something to watch closely in 2016**

### b. Desktop Users Viewing Mobile Advertisements Are Often Not Human

Because botnets are not currently a serious threat in the mobile ecosystem, malware running in the background on compromised mobile devices is not yet common. As the economic landscape shifts, and more ad spending moves to mobile devices, White Ops expects mobile-specific types of fraud — such as unethical developers displaying non-viewable ads within apps for profit — to gain prominence.

Looking at the macroeconomics behind the higher rate of bot fraud in video advertising, we see that the growth of spending has outpaced the growth in supply without an appropriate gain in price. If marketers as a group shifted spending to mobile just to escape fraud, the same thing would happen.

White Ops observed a clear trend toward fraud when mobile advertisements were viewed by what was detected to be desktop operating environments. While mobile sites, such as [m.whiteops.com](http://m.whiteops.com), can be easily viewed in a desktop browser, the study showed that nearly 18 percent of non-mobile devices that visit advertising inventory intended for mobile were nonhuman, sophisticated bots.

### c. Publisher (App) Fraud Is Uncontrolled on Mobile Devices

A third vector to monitor is that of unethical developers rendering hidden ads for profit. As it becomes more economically viable for criminals to run fraud models in the mobile ecosystem, publisher app fraud is an important area to keep an eye on in 2016. Improving mobile viewability standards and mobile fraud measurement may help reduce the impact of this type of fraud.

Stakeholders can reduce the bot impact from desktop systems going to mobile web pages by blocking desktop browsers that try to visit mobile pages

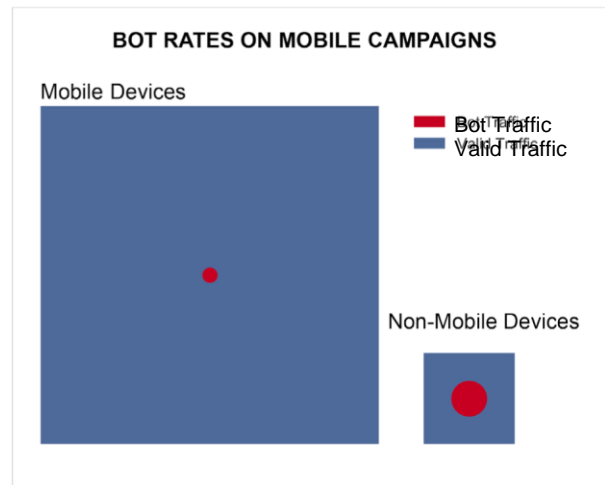


Figure 19: Mobile Bot Rates in Non-Mobile and Mobile Devices

Campaign Type	Device Type	Impressions	Bot Percentage	Bots
Desktop	Mobile	3.05 billion	0.41%	10.3 million
Mobile	Mobile	2.84 billion	0.20%	4.6 million
Mobile	Non-Mobile	207 million	17.59%	25.9 million

Table 3: Mobile Bot Rates by Device Type and Campaign Type

## 9. Being Aware and Involved Reduces Fraud Exposure

### a. Bots Shifted from Proactive to Less Involved Advertisers

Proactive policies and strategies to combat ad fraud can make a difference. White Ops identified two study participants which appeared very similar. The participants used the same agency and had similar technologies in place for fraud detection and bot prevention. Yet because of different policies and approaches to traffic sourcing, these two advertisers saw dramatically different results.

**Participant A** carefully selected its partners and required them to provide details of their traffic-sourcing policies. The advertiser also relied on programmatic buys for only 1 percent of its impressions. This practice yielded low sophisticated-bot impressions — between 1 and 10 percent across providers, averaging 3 percent — for the duration of the study.

**Participant B**, which operated in the same industry vertical, had an impression volume within the same range of 100 million to 300 million impressions but had

## Technologies that detect fraud are necessary, but not sufficient, to lower the bot rate; advertisers also need rigorous policies to reduce the impact of ad fraud in their paid media

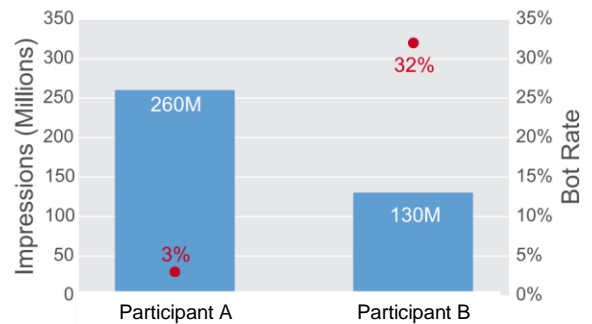


Figure 20: Anti-Fraud Policies and Strategies Reduce Bot Rates

### b. Major Exchanges and Platforms Have Reduced Their Bot Levels

Measured Trait	Participant	
	A	B
Average bot rate	3%	32%
Range of bot percentage in publishers with over 3,000 impressions	1–10%	1–62%
Fraud detection and fraud prevention in place	Yes	Yes
Anti-fraud buying patterns and policies in place	Yes	No
When surveyed, specified that the responsibility for combatting ad fraud lies with the advertiser	Yes	No

Table 4: Advertisers with Proactive Anti-Fraud Policies Had Much Lower Bot Rates

10 times the sophisticated bot rate. This participant’s media was mostly programmatic, with the sophisticated bot percentage among the participant’s publishers ranging from 1 to 62 percent, resulting in an average bot rate of 32 percent.

A study of nine of the highest-volume advertising technology platforms and exchanges that could be identified through HREF data showed that the highest-volume ad tech platform reduced its bot rates by 9 percentage points year



over year. The second-highest volume platform reduced bot percentages by 25 percent year over year, while for five platforms and exchanges, bot percentages stayed the same.

Some of the studied platforms and exchanges lacked transparency in their data, with incomplete loads accounting for 4 to 21 percent of traffic in 2015. The lack of transparency and inability of stakeholders to measure the impression validity could hide higher bot percentages in the unmeasured inventory.

Requiring transparency and fraud measurement capabilities from providers is critical to ensure that fraud levels remain low. It's recommended that buyers request transparency from publishers by building language into insertion orders that requires publishers to identify all third-party sources of traffic and to allow media validity measurement, including bot detection, on all media.

Platform providers that make clean inventory a priority can have less fraud than the direct display channel. One ad-tech provider of video advertising placed a huge emphasis on clean inventory in 2015, and it showed. This major video platform partnered with White Ops to reduce the bot impact in programmatic video media. For this platform, human impressions of video advertising designed to improve brand recognition and engagement rose 22 percent in

compared to campaigns that did not use the technology.

## The HREF data provided in web links gives information about the source and destination of an advertising impression

Requiring transparency and fraud measurement capabilities from providers is critical to ensure that fraud levels remain low

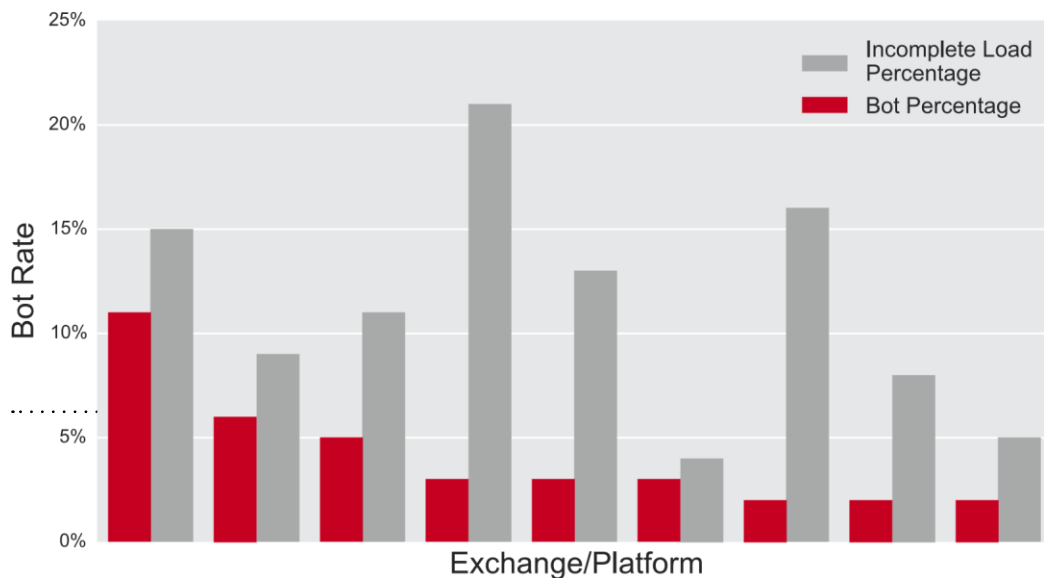


Figure 21: Bot Percentages and Incomplete Loads in Exchanges and Platforms campaigns using sophisticated bot prevention

**c. Survey: Awareness of Ad Fraud Has Improved**

In the course of the 2015 Bot Baseline study, White Ops surveyed study participants to discover the priorities and motivations of the participants as well as aspects of their media campaigns, such as target audience and media type, that cannot be discovered from impression data.

Out of the 42 advertisers which responded to the survey, the great majority use viewability as a metric of media quality. To reduce fraud, slightly more than half rely on a fraud detection solution, while slightly less use a fraud prevention solution. Others use anti-fraud policies or employ anti-fraud buying patterns.

Survey respondents stated that they would like to see improvements in transparency in regard to media buying, mobile impressions, and efforts by individual sites to protect advertisers against fraud. In addition, respondents requested more support for advertisers to fight fraud and insight into who should be preventing ad fraud in the supply chain.

Respondents agreed that the issue of digital ad fraud is important for the industry. Eighty-four percent of the surveyed advertisers considered the issue of digital ad fraud as either important or very important. Yet the advertisers surveyed had very different ideas of who should be responsible for combating fraud: about a quarter thought all parties should take responsibility, while more than a third place responsibility with the agency. Only 17 percent placed responsibility with the advertiser.

Anti-Fraud Solution	Self-Reported Usage Rate
Viewability	55%
Detection Vendor	55%

Prevention Vendor	43%
Anti-Fraud Policies	40%
Anti-Fraud Buying Patterns	29%

Table 5: Self-Reported Anti-Fraud Solutions in Place


**HOW CAN PROVIDERS BE MORE TRANSPARENT?**

- Allow third-party JavaScript-based tracking
- Reveal sources of traffic and their fraud levels
- Reveal programs such as audience extension
- Commit not to count fraud in billing

Who Should be Responsible for Combatting Fraud?	Participant Response Rate
The Publisher	21%
The Agency	36%
The Advertiser	17%
All Parties	26%

Table 6: Ad Fraud Accountability Survey Responses

**Survey respondents stated that they would like to see improvements in transparency in regard to media buying, mobile impressions, and efforts by individual sites to protect advertisers against fraud**

The background features a complex network of black nodes of varying sizes connected by thin, light gray lines. Some nodes are significantly larger than others, suggesting a hierarchy or central nodes. Dashed gray lines form large, overlapping circular shapes that intersect with the network. The overall aesthetic is clean, modern, and technical.

# III. Recommendations

# Recommendations

Stakeholders in the advertising ecosystem are taking action to reduce ad fraud, but the leading edge of fresh botnet infections are holding the size of the problem steady and causing the bulk of monetary losses to advertisers.

## **In 2015, Advertisers with the Lowest Impact from Bot Fraud:**

- Used legal language that removed the impact of fraud during the billing stage, placing legal language in contracts that stated the commitment not to pay for fraudulent impressions
- Selected media partners that proactively reduce fraud
- Leveraged the watchdog effect by announcing anti-fraud policies to partners and encouraging them to provide the highest-validity media
- Created open dialogues with providers about traffic sourcing and carefully selected the providers with a commitment to providing valid impressions
- Combined technology with anti-fraud policies and strategies to reduce fraud at all levels

In 2016, all stakeholders can work to reduce ad fraud by combining the use of anti-fraud technologies with proactive policies and strategies that reduce the impact of fraud across all stages.

---

## 1. Action Plan for All Stakeholders

### **a. Authorize and Approve Third-Party Traffic Validation Technology**

To effectively combat bots in their media buys, advertisers, publishers, and agencies must be able to deploy monitoring tools. This study was not deployed across all participants' placements, partly due to agency and publisher policies, which did not permit the monitoring software in certain placements. All participants in the advertising ecosystem need to be able to set policy and procedures to enable advertisers to deploy fraud detection technologies in their ad buys.

### **b. Require Clarity from Vendors on How They Combat Fraud**

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

### **c. Protect Against Fraud that Is in the Profit Window**

When possible, use sophisticated bot detection to shrink the profit window for ad fraud. Use sophisticated fraud detection solutions to reveal the hard-to-find fraud that is still fresh and profitable for the botnet operators because it is not yet listed in general detection databases.

### **d. Use Sophisticated Fraud Detection to Block Bots in Programmatic Media**

Protect programmatic media buys with sophisticated fraud detection as defined by MRC and avoid general blocking solutions that are not shown to significantly reduce fraud in programmatic buys.

### **e. Follow MRC Guidelines for Invalid Traffic Detection and Filtration**

MRC recently issued a strong set of [guidelines](#) for invalid traffic detection and filtration. The ANA recommends all digital measurement organizations adopt these guidelines and that sophisticated fraud detection vendors seek MRC accreditation for their detection procedures.

#### **f. Support the Trustworthy Accountability Group**

The IAB, 4A's, and the ANA announced in November 2014 the creation of the Trustworthy Accountability Group (TAG), a joint marketing-media industry program designed to eradicate digital advertising fraud, malware, ad-supported piracy, and other deficiencies in the digital communications supply chain. In the past year TAG has made significant strides in developing solutions to thwart fraud in the advertising supply chain while gaining strong support from its industry leaders. TAG has developed an [Anti-Fraud Working Group](#) with a mission to improve trust, transparency, and accountability by developing tools, standards, and technologies that enable the elimination of fraud. In May 2015 TAG unveiled its [Fraud Threat List](#), a shared database of domains that are known sources of non-human traffic. Shortly thereafter TAG launched the [Data Center IP list](#), which identifies sources of non-human traffic based upon IP addresses. Support of TAG's initiatives is a crucial step in creating a transparent and legitimate digital advertising ecosystem. Every company across the ecosystem should [register with TAG](#) in order to ensure they are doing business with trusted partners.

---

## **2. Action Plan for Buyers**

#### **a. Be Aware and Involved**

Advertisers must be aware of digital advertising fraud and take an active and vocal position in addressing the problem. Fraud hurts everyone in the digital communications supply chain, especially advertisers. Advertisers must therefore play an active role in generating positive change and should take responsibility for combating ad fraud.

#### **b. Request Transparency for Sourced Traffic**

Traffic sourcing correlates strongly to high bot percentages. It's recommended that buyers request transparency from publishers around traffic sourcing and build language into RFPs and IOs that requires publishers to identify all third-party sources of traffic. Furthermore, buyers should have the option of rejecting sourced traffic and running advertising only on a publisher's organic site traffic.

#### **c. Request Transparency for Audience Extension Practices**

Audience extension by publishers can introduce high bot percentages by extending content to providers that source traffic. It's recommended that buyers request transparency from publishers around audience extension and build language into RFPs and IOs that requires publishers to identify audience extension practices. Buyers should have the option of rejecting audience extension and running advertising only on a publisher's owned and operated site.

#### **d. Understand the Programmatic Supply Chain and Require Inventory Transparency**

The foundation of optimizing your media investment, including reducing bot fraud when using programmatic buys, is understanding the programmatic supply chain. Advertisers should ask about the role of each player in the process, know the partners of your partners, and then ask for inventory transparency to know where your programmatic advertising is running. You wouldn't "blindly" run your advertising in offline media such as television or print without knowing the specific networks or publications that carry your advertising. Why accept anything less in programmatic buying?

#### **e. Include Language on Non-Human Traffic in Terms and Conditions**

Insertion orders should include language that the company will only pay for non-bot impressions. Additional language should be added to your terms and conditions to address the issues discussed in this study. An illustration of one approach to the definition of fraudulent traffic and the safeguards that might be negotiated between advertisers and media companies is provided in the appendix (developed by Reed Smith, the ANA's outside legal counsel). You should consult with your own counsel to develop specific provisions that best serve your company's individual interests (see Appendix B: Illustrative Terms and Conditions, page 40).

#### **f. Use Third-Party Monitoring**

Monitor all traffic with a consistent tool. We recommend relentless monitoring to get the best value out of your ad investment. Use monitoring and bot detection to reveal the bots in retargeting campaigns, weed bots out of audience metrics, and protect higher-value inventory that may have increased fraud exposure. Protect against ad fraud to be sure that bots are not being pushed into your media from other proactive stakeholders. Monitor your top-100 volume sites to prevent making payments to cash-out sites.

#### **g. Use Frequently Updated Blacklists**

For blacklists to be effective, they need to be updated at least daily, must be very specific (micro-blacklisting), and must accompany other defenses.

#### **h. Announce Your Anti-Fraud Policy to All External Partners**

In combination with covert, continuous monitoring practices, the watchdog effect will change behavior, reduce fraud, and encourage others to join the fight.

#### **i. Equip Your Organization to Fight Ad Fraud: Budget for Security**

Across many industries, the typical cost of security amounts to an overhead of 1 to 3 percent. In the credit card ecosystem, that security spending has lowered the losses due to fraud to just \$0.08 per hundred dollars. Lowering bot fraud in advertising to those levels could potentially return many multiples of the security spending needed to achieve it.

#### **j. Involve Procurement**

Many ANA member companies have marketing procurement groups which should be a partner in the fight against bot fraud. The best marketing procurement organizations reduce waste and help improve marketing ROI by ensuring that every dollar is invested to deliver maximum growth and profitability. The fight against bot fraud can directly reduce waste and improve ROI, meeting procurement objectives.

#### **k. Demand the Data**

Ask suppliers for maximum visibility into bot levels in their inventory. Ask for third-party monitoring or certification of specific inventory to demonstrate that the inventory meets human impression requirements.



## 3. Action Plan for Publishers, Platforms, and Exchanges

### **a. Continuously Monitor Sourced Traffic**

Publishers should always monitor sourced traffic, know their sources, and maintain transparency about traffic sourcing. Publishers, platforms, and exchanges which are serious about reducing bot fraud should eliminate sources of traffic that are shown to have high bot percentages and monitor their vendors at all times.

### **b. Purge the Fraud; Increase Your Prices**

Clean up the fraud in your supply. Once you can demonstrate higher value from higher valid impression percentages, the value of your media will increase.

### **c. Protect Yourself from Content Theft and Ad Injection**

Use a service such as domain detection or bot detection to monitor for evidence of ad injection and for content scraping — from copying content from a site to in some cases monetizing the scraped content with ads on an unsanctioned site. A bot detection service can measure actual numbers of bots in high-bot traffic, allowing payment for the human audience while eliminating bots from the billing process.

### **d. Allow Third-Party Traffic Assessment Tools**

Publishers can enable advertisers to improve the granularity of their traffic performance by authorizing third-party tracker measurement and third-party monitoring for characteristics such as viewability, engagement, and bot detection.



# IV. Appendix

# A. Methodology

## 1. Study Data Sets

In 2014, White Ops and the ANA set out to gain a better understanding of the impact of fraud on the online advertising ecosystem. White Ops worked with 36 members of ANA to measure more than 5.5 billion ad impressions over 60 days. The results were illuminating. Bot fraud accounted for a substantial portion of the impressions paid for by advertisers, far more than many of those advertisers expected

In 2015, White Ops worked with the ANA to repeat the Bot Baseline Study with a larger group of advertisers to gain greater visibility into ad fraud due to bots. The study included:

- 49 advertisers from 10 industries: auto, beer/spirits, CPG, financial services, health care, hospitality/travel, insurance, restaurant, retail, and technology
- 28 returning participants and 21 new participants
- Data collected over 61 days from August 1 to September 30, 2015
- 10 billion total impressions examined across 1,300 campaigns

For the 61 days of the study, from August 1 to September 30, 2015, ANA participants deployed White Ops detection tags on their digital advertising. White Ops collected 19.3 billion impressions, of which half did not satisfy the study's conditions — either failing to completely load the JavaScript tags, referred to as unmeasurable traffic, or coming from mobile devices. In total, the study focused on 9.7 billion non-mobile, measurable impressions.

## 2. Data Collection

Where possible, the White Ops technology gathered information directly at the time of impression. No data or results were provided to study participants during the data collection period. Because of the lack of information about mobile impressions, the study focused on non-mobile visitors only, with limited analysis of mobile traffic. In addition, impressions were considered unmeasurable in cases where they did not execute any JavaScript. White Ops does not count bots detected by the industry spiders and bots list as “sophisticated bots.” Instead, these impressions are designated as “general bots,” and include legitimate automated search spiders as well as easily-detected malicious bots. This is the same methodology used in the 2014 Bot Baseline study. Viewability was measured per MRC guidelines using page geometry, browser optimization, and other methods. White Ops' capability to measure viewability was not yet accredited at the time of the measurement.

## 3. Reporting

Following the end of the data collection period, participants received comprehensive bot fraud reports on their studied media. Data aggregated in this report preserves anonymity for all study participants.

## 4. 2014 and 2015 Data Sets Are Not Fully Comparable

In 2015, White Ops encountered a significantly different study population compared to 2014, with the following differences:

- In 2014, a handful of large companies had high bot rates, contrasting with the more even distribution of participant bot rates in 2015.
- Media classifications in 2015 partially rely on participant surveys. Advertisers and their agencies tagged the studied ads to designate media types (such as display ads or video ads) and buy types (such as direct or programmatic) and to designate operational policies.
- In certain cases, detection code was blocked or evaded, resulting in incomplete loads.

## B. Illustrative Terms and Conditions

Consider adding specific language to your insertion order terms and conditions to address the issue of digital ad fraud. An illustration of one approach to the definition of fraudulent traffic and the safeguards that might be negotiated between advertisers and media companies appears below (developed by Reed Smith, the ANA's outside legal counsel). You should consult with your own counsel to develop specific provisions that best serve your company's individual interests.

### **Fraudulent Traffic**

(a) "Fraudulent Traffic" means the inclusion in reports, bills or other information and materials associated with this Agreement, of data that counts or uses in calculations, anything other than natural persons viewing actually displayed Ads in the normal course of using any device, including, without limitation, browsing through online, mobile or any other technology or platform. For the avoidance of ambiguity, Fraudulent Traffic includes, without limitation, the inclusion or counting of views: (i) by a natural person who has been engaged for the purpose of viewing such Ads, whether exclusively or in conjunction with any other activities of that person; (ii) by non-human visitors; (iii) combinations of displays directed or redirected by any combination of (i) and/or (ii); and (iv) that are not actually visible to the human eye, discernible to human senses or perceived by a human being.

(b) Media Company will establish, implement and use all commercially reasonable technology and methodologies to: (i) prevent Fraudulent Traffic; (ii) detect Fraudulent Traffic should it occur; and (iii) promptly take steps to prevent continuation and/or recurrence of occurrences thereof. Media Company will ensure, by agreement, instruction or any other legally enforceable means, that all third parties to which Ads are delivered, displayed or made available (including, without limitation, DSPs) have adopted and implemented technology and methodologies (and agreed in writing thereto) to ensure Media Company is in compliance with the foregoing obligations. Media Company agrees that Advertiser shall have no obligation hereunder, for compensation, liability or otherwise in respect of Fraudulent Traffic and shall not be billed or required to pay for Fraudulent Traffic. To the extent any payment attributable to Fraudulent Traffic is or may be paid by Advertiser, Media Company shall, within five (5) days, reimburse and refund such payment to Advertiser, together with reasonably adequate documentation to substantiate the accuracy of any such reimbursement or refund. Unless otherwise included in another audit provision hereunder, Advertiser or its designated auditors shall be entitled to audit the books and records (including, without limitation, log files) of Media Company for the purpose of determining compliance with these Terms.

(c) Media Company will (i) upon request by Advertiser or Agency, permit Advertiser and/or Agency to deploy fraud detection, traffic validation or other technologies on Ads to measure compliance with these Terms, (ii) disclose to Advertiser and Agency in writing (and update on an on-going basis) its practices for sourcing third-party traffic and audience extension, (iii) disclose to Advertiser and Agency in writing (and update on an on-going basis) its practices for reducing Fraudulent Traffic, (iv) provide third-party monitoring or certified reports of the Deliverables upon request.

## C. Glossary

### **Ad**

An online advertisement of any sort

### **Ad Fraud**

The inclusion in reports, bills, or other analytics of anything other than natural persons consuming ads in the normal course of using any device

### **Ad Injection**

The visible or hidden insertion of ads into an app, web page, or other online resource without the consent of the publisher or operator of that resource

### **Ad Inventory**

Available online advertising space; an aggregation of available online ad slots

### **Advertiser**

A company, brand, or individual which pays a third party to display or act as agent for the display of ads

### **Blacklisting**

Using lists of known bad IPs, domains, or other parameters to prevent the serving of ads to those parameters

### **Bot(s) (Non-Human Traffic or NHT)**

Automated entities capable of consuming any digital content, including text, video, images, audio, and other data. These agents may intentionally or unintentionally view ads, watch videos, listen to radio spots, fake viewability, and click on ads.

### **Bot Detection**

The detection and differentiation of bot traffic and bot impressions from human traffic and human impressions

### **Bot Prevention**

The prevention of bot traffic and bot impressions in inventory before the inventory is bought/sold

### **Bot Traffic**

Automated website or other online traffic and/or ad consumption driven by or resulting from bots

### **Botnet**

A group of infected computers that generate automated web events. The infrastructure used to create many types of bots

### **Broker**

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

### **Campaign**

A group of ads belonging to an advertiser that share a single idea and theme and which may be made up of different types of ads, and which may be run on multiple publishers, sites, or other channels and in multiple formats

### **Cash-Out Site**

A website, app, or other resource that is capable of delivering ads, and is operated by perpetrators of ad fraud for the purpose of extracting money from the online advertising ecosystem

### **Desktop Impressions (or Non-Mobile Impressions)**

Ad impressions coming from web pages browsed to by user agents tagged as desktops, laptops, and gaming consoles

### **Domain**

A unique name that identifies and can be used to access an Internet resource such as a web site

### **DSP (Demand-Side Platform)**

A platform that allows advertisers or their agencies to manage multiple exchange accounts and bid across those accounts

### **DMP (Data Management Platform)**

Software that aggregates first-party and third-party data in a centralized location and format for advertisers or their agencies

### **Exchange**

A technology platform that facilitates the buying and selling of ads and related data from multiple sources such as publishers and networks of publishers



**Funnel, Brand**

A concept that breaks down the impact of advertising on an audience into different phases and objectives. At the top of the funnel, advertisers focus on brand awareness and attitudes toward the brand. In the middle of the funnel, advertisers attempt to move potential customers from awareness to intent to buy, including convincing customers to prefer the advertiser's product. Finally, at the bottom of the funnel, advertisers focus on converting the advertisement into a sale.

**General Bots (or Known Bot)**

Bots that can be detected through the industry bots and spiders list and known browser list

**HREF Domain**

The domain where a particular ad impression, video play, page view, or other online event occurred

**Human Impression (or Valid Impression)**

An impression legitimately served to a real human not intentionally or unintentionally engaged in any form of ad fraud

**Impression**

A singular instance of the delivery of a particular online ad in a specific online inventory space. The basic economic unit of online advertising, generally as recorded by ad servers for the purposes of billing advertisers or their agencies

**Incomplete Load (or Non-Measurable)**

Cases where the JavaScript tag was not fully loaded due to factors such as page abandonment or site configuration

**IP, IP Address**

A unique numerical address corresponding to a particular device or set of devices connected to the Internet

**Mobile (or Mobile Impressions)**

Impressions coming from web pages browsed to by user agents using the mobile tag

**Monitoring**

Paying attention to ads and their formats and the publishers, sites, and channels on or in which they are displayed for the purpose of detecting differing levels of ad fraud, allowing for the optimization of spending to reduce ad fraud

**Placement**

A subset of ads under a specific campaign belonging to an advertiser that is related to a specific ad size and inventory slot

**Private Marketplace**

A seller-controlled auction-based buying environment that requires a passkey (usually a Deal ID) in order for the buyer to participate

**PPC (Pay-per-Click)**

A method of buying and selling ads in which the buyer pays the seller an agreed-upon amount of money per click that is generated

**Publisher**

The operator of a website or network of websites, and the producer or curator of content for those sites. A seller of online advertising inventory, and often a buyer of third-party traffic

**Retargeting (or Behavioral Retargeting)**

The process of delivering ads to particular users based on previous online activity

**Site or Web Site**

A set of related web pages, often served from a single domain

**Sophisticated Bot**

A bot not listed in the industry bots and spider list and known browser list

**Sophisticated Bot Percentage**

The percentage of total traffic for which sophisticated bots are responsible, compared to total traffic

**Traffic**

Visits to a particular site, page, or other online resource; impressions related to a particular ad

**Traffic Sourcing or Sourced Traffic**

Any method by which publishers acquire more visitors through third parties

**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, and a consumer of an advertiser's ads





