distil networks

2018 BAD BOT REPORT

The Year Bad Bots Went Mainstream

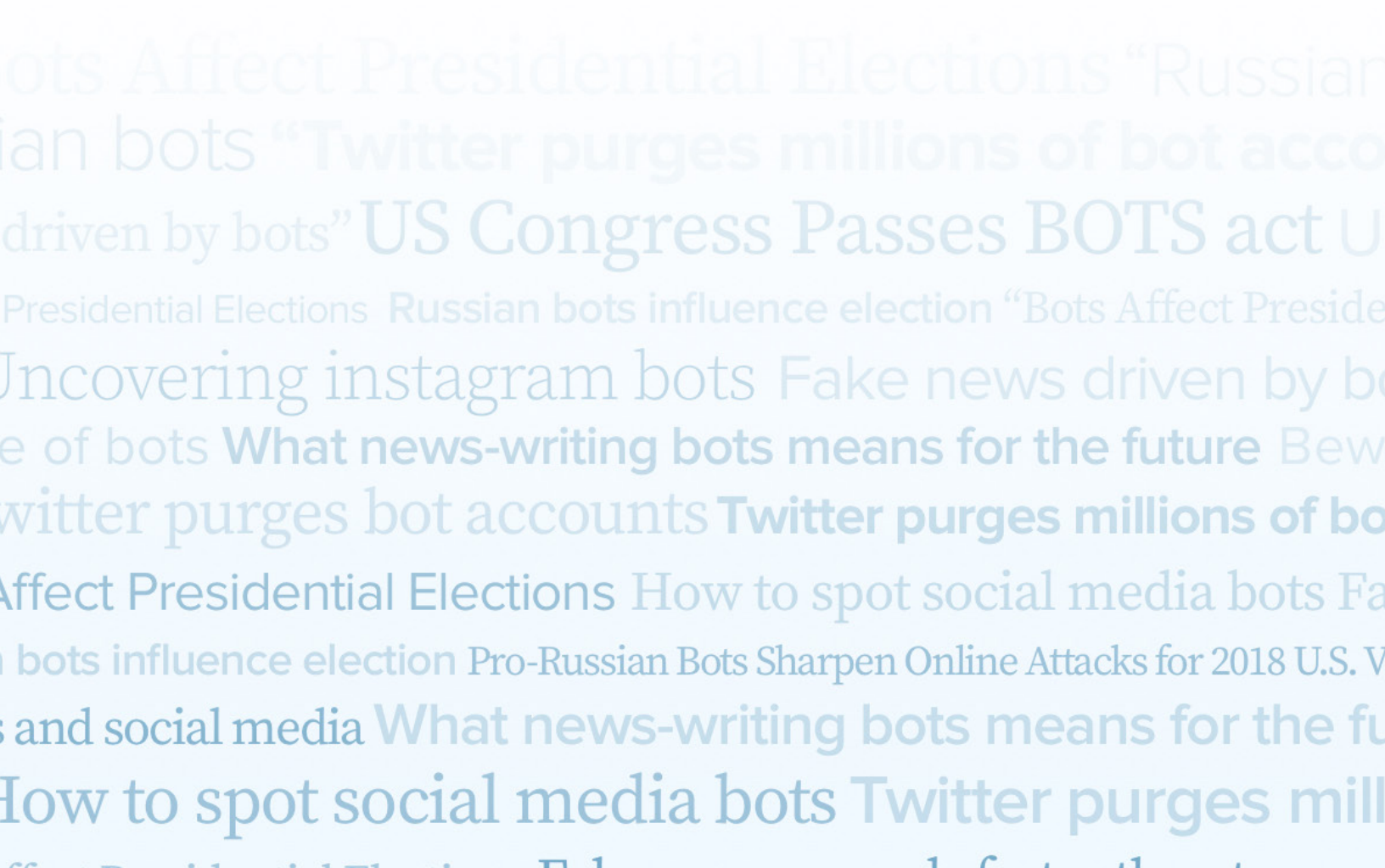


Table of Contents

About the Bad Bot Report

The 2018 Bad Bot Report investigates the daily attacks sneaking past sensors and wreaking havoc on websites. It's based on 2017 data collected from Distil Networks' global network, and includes hundreds of billions of bad bot requests, anonymized over thousands of domains. Our goal is to offer guidance about the nature and impact of automated threats to those of you on the frontlines of website security.

What makes this report unique is its focus on bad bot activity at the application layer (layer 7 of the OSI model). Automated application layer attacks differ from volumetric DDoS attacks, the latter manipulating lower level network protocols.

Bad bots interact with applications in the same way a legitimate user would, making them harder to detect. Bots enable high-speed abuse, misuse, and attacks on your websites and APIs. They enable attackers, unsavory competitors, and fraudsters to perform a wide array of malicious activities.

Such activities include web scraping, competitive data mining, personal and financial data harvesting, account takeover, digital ad fraud, spam, transaction fraud, and more.

“Bots the Matter” with the Internet?

This was the year that bots went mainstream. Today, it would be difficult for anyone to claim ignorance of the term bots after such a tumultuous twelve months. No longer are bots the preserve of cyber security experts. Instead, even the FBI is investigating their use into influencing the results of the last US presidential election.

A few of the most popular social media brands are suffering an existential crisis. Some were even hauled before the US Congress, because bots that exploited their platforms were used to amplify and spread fake news. Twitter founder Jack Dorsey eloquently described the problem companies like his face.

“We have witnessed abuse, harassment, troll armies, manipulation through bots and human-coordination, misinformation campaigns, and increasingly divisive echo chambers. We aren’t proud of how people have taken advantage of our service, or our inability to address it fast enough.”

Making matters worse, the New York Times conducted an impressive investigation into a social media growth service called Devumi, which allegedly sold fake followers to unsuspecting clients. Just days after the article was released, Twitter responded by purging millions of fake accounts that were bots.

But while the political world attempts to understand the impact of bad bots on democracy, much of their wider impact on the economy is misunderstood and grossly underestimated.

Legislation pertaining to bot behavior is being created the world over, and particularly within concert and sporting event ticketing. The US Congress passed the Better Online Ticket Sales (BOTS) Act of 2016, banning the use of software that circumvents security on ticket seller websites. Similarly, the UK proposed amending the Digital Economy Act to prevent bulk ticket purchasing by bots. Canadian provinces are also banning scalper bots.

But while such legislation is in place, it remains unclear if any of it has teeth, or that court cases can be effective in dealing with offenders. LinkedIn brought a lawsuit against hiQ Labs to prevent its proprietary content from being scraped and is using existing legal tools to protect its business—or at least to seek greater clarity as to what separates illegal hacking from legal data harvesting.

The difficulty with most bad bot behavior is in identifying its origin. Furthermore, bringing legal recourse against bot operators is both costly and time-consuming; if bot operators are in another country, the laws offer no guarantee of success.

Given that, preventing the bad bot problem using technology offers a viable alternative to help you protect your business. In many cases, it's cheaper and more effective than the slow moving legal process.

The reality is that bad bots account for more than one-fifth of all internet traffic. And most of it isn't targeting political elections, exploiting social media platforms, or buying concert tickets. Rather, the majority of bad bots are doing something else—damaging the economy and hurting businesses all over the world. Even yours.

So what are bad bots doing?

What Bad Bots Do

Left unaddressed, bad bots cause very real business problems that could harm the success—or even the continuance—of your organization. Examining the problems doesn't require deep knowledge of the technology behind attacks or the techniques used to prevent them. Instead it requires a solid understanding of your business.

FACT 1: Every business with an online presence is regularly bombarded by bad bots on its website, APIs, or mobile apps.

FACT 2: Unchecked bad bots cost businesses money every day. Different from the problem of data breaches, which are somewhat rare, automation abuse happens 24 × 7 × 365 because bad bots never sleep.

FACT 3: Bad bots are on your website for a purpose. Understanding what that purpose is helps you address the problem.

Many businesses ignore bots as a benign nuisance, simply because they don't understand the havoc they cause. Others try and handle the problem themselves, spending hours playing IP whack-a-mole, only to then suffer from ruined weekends dealing with alerts and inexplicable website problems. Here is what to consider in understanding some of the most common bad bot problems.

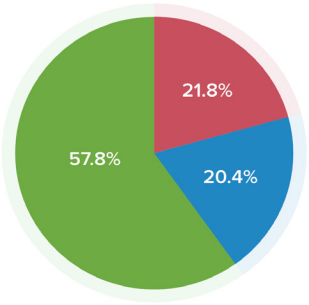
Bad Bot Problem	How it Hurts the Business	Signs You Have a Problem	Primary Targets
Price Scraping	Competitors scrape your prices to beat you in the marketplace. You lose business because your competitor wins the SEO search on price. Lifetime value of customers worsens.	Declining conversion rates. Your SEO rankings drop. Unexplained website slowdowns and downtime, usually caused by aggressive scrapers. Whenever you lower the price of an item, your main competitor lowers theirs within hours.	All businesses that show prices. - Ecommerce - Gambling - Airlines - Travel
Content Scraping	Proprietary content is your business. When others steal your content they are a parasite on your efforts. Duplicate content damages your SEO rankings.	Your content appears on other sites. Unexplained website slowdowns and downtime, usually caused by aggressive scrapers.	Similar to price scraping, but in addition: - Job boards - Classifieds - Marketplaces - Digital Publishing - Real Estate - Online Directories

Bad Bot Problem	How it Hurts the Business	Signs You Have a Problem	Primary Targets
Account Takeover (a.k.a., Credential Stuffing, Credential Cracking)	Stolen credentials are run against your website in order to test if they work. If successful the ramifications are are account lockouts, financial fraud, and increased customer complaints affecting customer loyalty and future revenues.	Increase in failed logins. Increase in customer account lockouts and customer service tickets. Increase in fraud (Lost loyalty points, stolen credit cards, unauthorized purchases). Increase in charge backs.	Any business with a login page requiring username and password.
Account Creation (a.k.a., Account Aggregation)	Free accounts used to spam messages or amplify propaganda. Exploit any new account promotion credits (money, points, free plays).	Abnormal increases in new account creation. Increased comment spam. Drop in conversion rates of new accounts to paying customer.	Messaging platforms - Social media *(Influencer/ propaganda bots)* - Dating sites - Communities Promotion Abuse - Gambling
Credit Card Fraud (a.k.a., Carding, Card Cracking, Cashing Out)	Criminals testing credit card numbers to identify missing data (exp. date, CVV). Damages the fraud score of the business. Increases customer service costs processing chargebacks.	Rise in credit card fraud. Increase in customer support calls. Increased chargebacks processed.	Any site with a payment processor - Ecommerce - Non-profit/ - Charities - Airlines - Travel - Tickets - Financial services - Gambling
Denial of Service	Slows the website performance causing slowdowns or downtime. Lost revenue from unavailability of website. Damaged customer reputation.	Abnormal, and unexplained spikes in traffic on particular resources (login, signup, product pages, etc). Increase in customer service complaints.	All Industries
Gift Card Balance Checking	Steal money from gift card accounts that contain a balance. Poor customer reputation and loss of future sales.	Spike in requests to gift card balance. Increase in customer service calls about lost balances.	Ecommerce
Denial of Inventory	Bots hold items in shopping carts, preventing access by “real” customers. Damaged customer reputation because unscrupulous middlemen hold all inventory until resold elsewhere.	Increase in abandoned items held in shopping carts. Decrease in conversion rates. Increase in customer service calls about lack of availability of inventory.	Scarce or time sensitive items. - Airlines - Tickets - Collectibles

Executive Summary of Findings

Bad Bots Are Up From Last Year

In 2017, 42.2% of all internet traffic wasn't human, and there were significant year-over-year increases in both bad bot (+9.5%) and good bot (+8.8%) traffic.

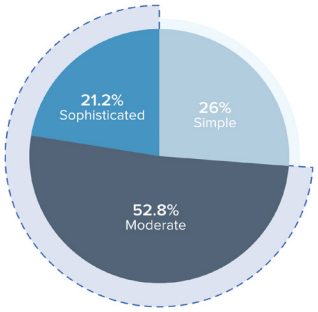


57.8% | Humans
21.8% | Bad Bots
20.4% | Good Bots

Bad Bot traffic percentage in 2017	21.8%
Growth in bad bot traffic in 2017	+9.5%
Good bot traffic percentage in 2017	20.4%
Growth in good bot traffic from previous year	+8.8%
Human website traffic percentage in 2017	57.8%
Decrease in human traffic from previous year	-5.8%

Bad Bot Sophistication Levels Remain Constant

Advanced Persistent Bots (APBs) continue to plague websites. APBs cycle through random IP addresses, enter through anonymous proxies, change their identities, and mimic human behavior.



74% | Advanced Persistent Bots (APBs)
21.2% | Sophisticated
52.8% | Moderate

Advanced persistent bots *(Moderate + Sophisticated bad bots percentage)*	74.0%
Sophisticated bad bots	21.2%
Moderate bad bots	52.8%

Every Industry Has a Bad Bot Problem

Some bad bot problems run across all industries while others are industry-specific. Websites with login screens are getting hit by bot-driven account takeover attacks two to three times per month. Meanwhile, nefarious competitors use bots to undercut prices on ecommerce sites, hoard seats on airline flights, and scalp the best concert tickets.

Top 5 Industries
Bad Bot Traffic %

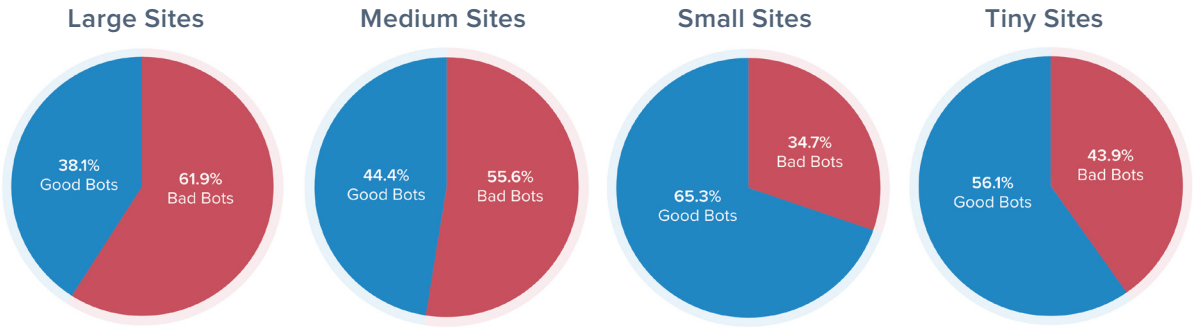
1	Gambling	53.1%
2	Airlines	43.9%
3	Finance	24.7%
4	Healthcare	24.4%
5	Tickets	23.0%

Top 5 Industries
Sophisticated Bad Bot Traffic %

1	Ecommerce	22.9%
2	Healthcare	22.3%
3	Airlines	19.7%
4	Travel	19.1%
5	Tickets	19.1%

Large and Medium Sized Websites Are More Enticing Targets For Bad Bots

For the second year in a row large sites were hit the hardest—21.7% of their traffic came from bad bots. Medium sites had the biggest year over year increase in bad bot traffic—36.9% more bad bot traffic than 2016.



How Bad Bots Hide

Bad bots continue to follow the trends in browser popularity, impersonating Chrome and Firefox 72.4% of the time. The weaponization of data centers accelerated in 2017 with 82.7% of bad bot traffic emanating from data centers—a 37% increase over 2016.

Bad bots report as either Chrome, Firefox, Internet Explorer, Safari	**83.0%**
Bad bots hiding in data centers	**82.7%**
Bad bots using Amazon ISP	**10.6%**

Bad Bots Are All Over The World

With most bad bot traffic emanating from data centers, it's no surprise that the US remains the bad bot superpower. But Russia became the most blocked country. France moved up to third as the country hosting the most bad bot traffic, and is the second most blocked country. This is due to cheap hosting availability from French company OVH. In 2017, OVH overtook Amazon to become the source of the most bad bot traffic in the world.

Top 5
Bad Bot Traffic By Country

1	United States	**45.2%**
2	China	**10.5%**
3	France	**9.9%**
4	Canada	**3.7%**
5	Germany	**3.3%**

Top 5
Most Blocked Country

1	Russia	**20.7%**
2	France	**20.4%**
3	Taiwan	**12.2%**
4	United States	**11.6%**
5	Ukraine	**9.2%**

SECTION 1

The Bad Bot Landscape

What is a Bad Bot?

Bad bots scrape data from sites without permission in order to reuse it (e.g., pricing, inventory levels) and gain a competitive edge. The truly nefarious ones undertake criminal activities, such as fraud and outright theft.

The Open Web Application Security Project (OWASP) provides a list of the different bad bot types in its *Automated Threat Handbook*[1]. This year a new attack type was added—Denial of Inventory.

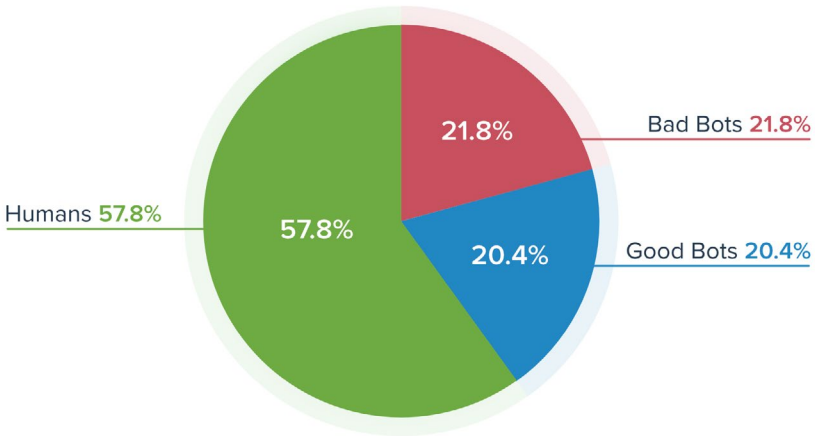
How Do Good and Bad Bots Differ?

In simplistic terms, good bots ensure that online businesses and their products can be found by prospective customers. Examples include search engine crawlers, like GoogleBot and Bingbot, that index websites to help people match their queries with the most relevant sets of websites.

Even Good Bots Can Be Bad News

Good bots can skew web analytics reports, making some pages appear more popular than they actually are. For example, if you advertise on your website, good bots can generate an impression, but that ad click never converts in the sales funnel; this results in lower performance for advertisers. Being able to intelligently separate traffic generated by legitimate human users, good bots, and bad bots is essential for making informed business decisions.

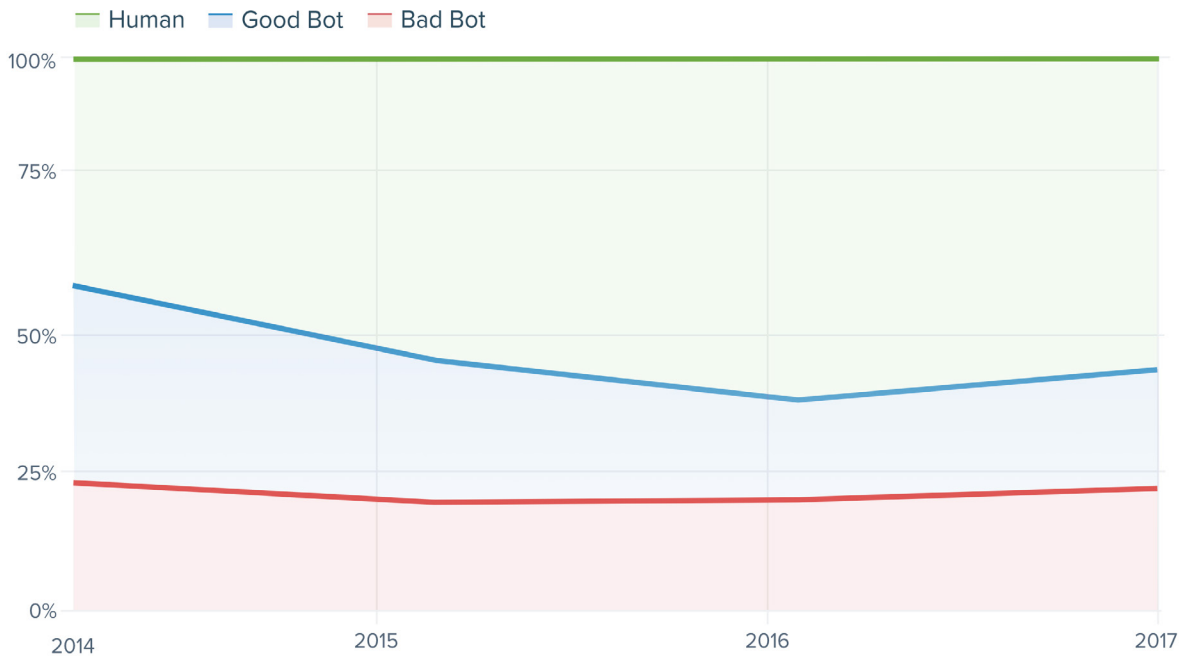
Bad Bot vs Good Bot vs Human Traffic, 2017



[1] www.owasp.org/images/3/33/Automated-threat-handbook.pdf

In 2017, bad bots accounted for 21.8% of all website traffic—a 9.46% increase over the prior year. Good bots increased by 8.76%, accounting for 20.4% of all traffic. In this year when bots went mainstream, the proportion of human traffic decreased by 5.76%, totaling 57.8% of all internet traffic.

Trend: Bad Bot vs Good Bot vs Human Traffic 2014-2017



Traffic Type	2014	2015	2016	2017
Human	40.9%	54.4%	61.3%	57.8%
Good Bot	36.3%	27.0%	18.8%	20.4%
Bad Bot	22.8%	18.6%	19.9%	21.8%

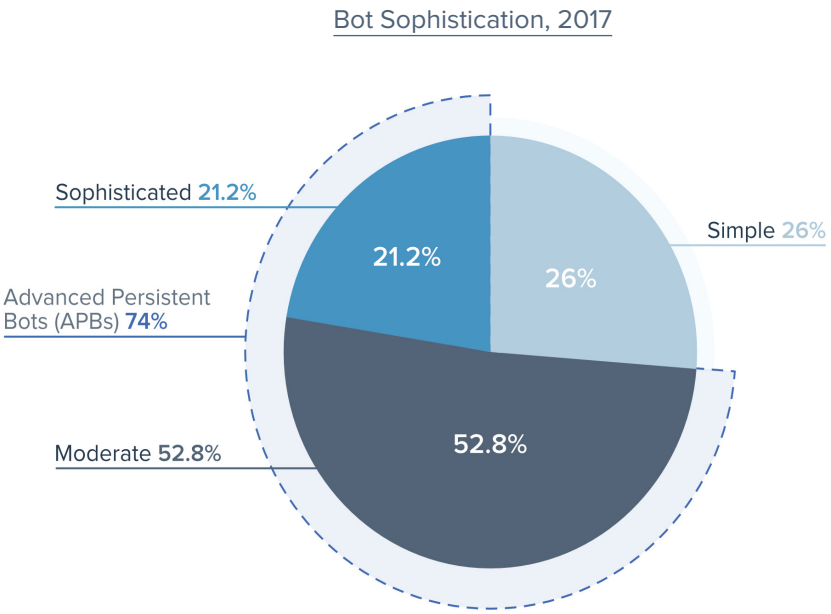
The bad bot traffic percentage has increased for two years running; 2017 was close to the peak seen in 2014.

Trying to attract human visitors is why any website exists. That human traffic comprises only about half of all internet traffic is a stark reminder of the bad bot problem.

Bad Bot Sophistication Levels

Distil Networks created the following industry standard system that classifies the sophistication level of the following four bad bot types:

- SIMPLE – Connecting from a single, ISP-assigned IP address, this type connects to sites using automated scripts, not browsers, and doesn't self-report (masquerade) as being a browser.
- MODERATE – Being more complex, this type uses "headless browser" software that simulates browser technology—including the ability to execute JavaScript.
- SOPHISTICATED – Producing mouse movements and clicks that fool even sophisticated detection methods, these bad bots mimic human behavior and are the most evasive. They use browser automation software, or malware installed within real browsers, to connect to sites.
- ADVANCED PERSISTENT BOTS (APBS) – APBs combine moderate and sophisticated technologies and methods to evade detection while maintaining persistency on targeted sites. They tend to cycle through random IP addresses, enter through anonymous proxies and peer-to-peer networks, and are able to change their user agents.



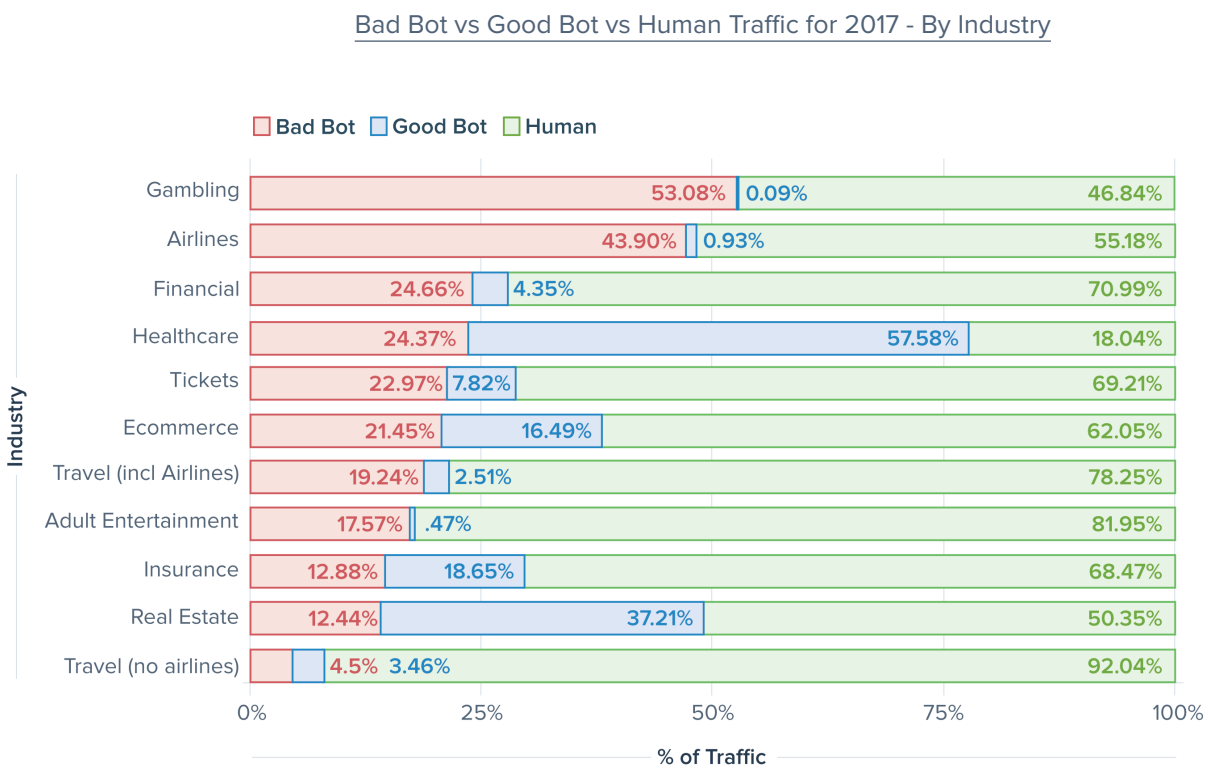
Being the easiest to detect and block, simple bots accounted for 26% of bad bot traffic. Meanwhile, the majority of non-human traffic (52.8%) came from those classified as moderate. And comprising 21.2% of such traffic last year, sophisticated bad bots once again proved to be the most difficult to detect.

Advanced Persistent Bots (APBs) accounted for 74% of all 2017 bad bot traffic—almost matching the prior year. Because they can cycle through IP addresses and switch user agents, simple IP blacklisting is wholly ineffective.

Sometimes known as “low and slow,” APBs carry out significant assaults using fewer requests and can even delay requests, all the while staying below request rate limits. This method reduces the “noise” generated by many bad bot campaigns.

Bad Bots By Industry

Looking at traffic from various industries, a deeper insight into the problem is evident. However, a higher bad bot percentage doesn’t equate to these industries doing less in their effort to address the problem. Rather, it reveals that they're targeted because of the content they produce or the data they store.



Gambling companies and airlines have the highest proportion of bad bot traffic at 53.08% and 43.90% respectively. Let's examine a few of the highlights.

Bad Bots Gamble 53.08% of the Time

Aggregators relentlessly scrape online gambling companies for the ever-changing betting lines they offer. Such aggressive activity causes denial of service problems and sends customers elsewhere.

Account takeovers are also a major problem. Each account contains money or loyalty points that, once compromised, can easily be transferred to another user and emptied.

Bad Bots Fly on Airlines (43.90%)

Breaking out the percentages for **Travel (no airlines), Travel (incl airlines)** and **Airlines,** the latter is faced with a much higher volume. Airline prices are scraped not only by direct competitors, but also by third-party players in the expansive travel ecosystem.

Unauthorized online travel agencies (OTAs), competitors, price aggregators, and metasearch sites use sophisticated scraping bots to abuse the business logic of booking engines. Querying for any ticket they can sell, they skew look-to-book ratios, increase GDS transaction costs, and are responsible for site slowdowns and downtime—causing customer dissatisfaction during disruptions. They dynamically package seat inventory with other products, stealing direct and ancillary revenue. And, they insert their own email addresses into reservations, thereby taking control of remarketing opportunities.

In addition, airlines suffer from account takeover issues as bad bot operators attempt to get into user accounts and empty them of air miles balances

Bad Bots Buy on Ecommerce Sites (21.45%)

As an industry vertical, Ecommerce sees the full gamut of bad bot attacks. These include price and content scraping, account takeovers, credit card fraud, and gift card abuse. With 21.45% of the nefarious traffic, the Ecommerce profile is closest to the overall bad bot total of 21.8%.

The Problem with Industry Data

Having a large number of bad bots doesn't necessarily mean a problem is worse. Traffic nuances exist between each industry, as well as within each company.

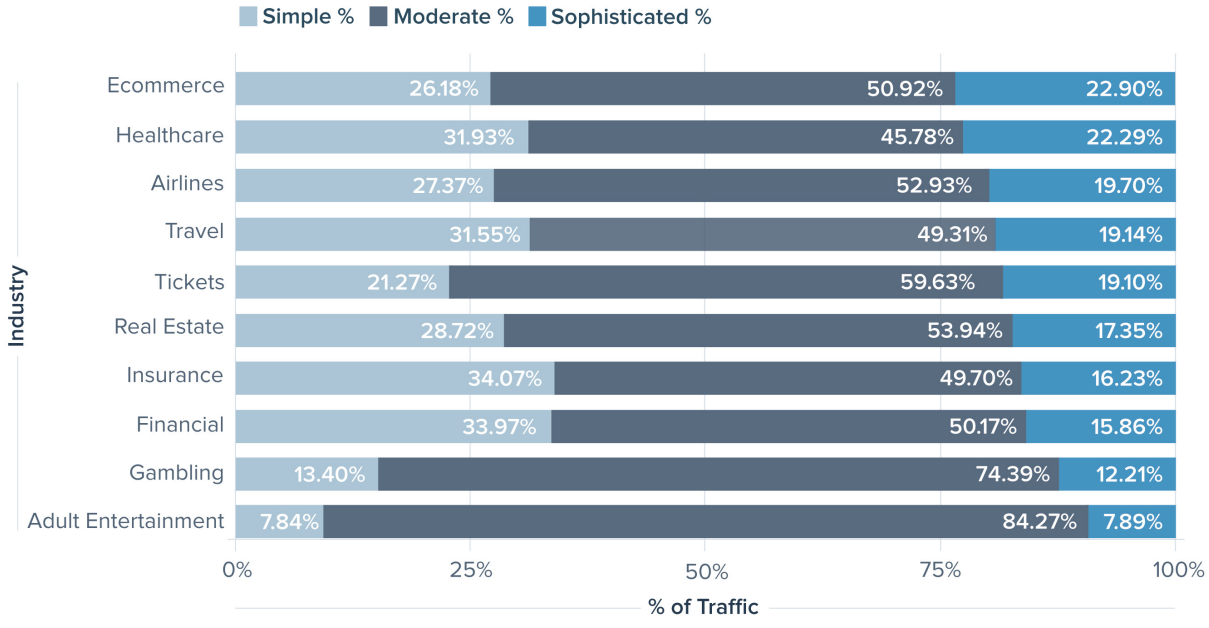
In only examining the chart above, it would seem that the **Ticketing** industry has a smaller problem than four other industries (Gambling, Airlines, Financial and Healthcare) with 22.97% of bad bots. But this is an inaccurate conclusion, as two other factors must also be considered—bad bot sophistication and the number of human requests. For the ticketing industry, what bad bots lack in volume they make up for by being the most evasive, persistent, and sophisticated yet seen. Human traffic volume on such sites is also higher than others, making the problem appear proportionally smaller.

That said, their effect is so strong that all major legislation against bad bots is focused on the ticketing industry. Ticketing has arguably the worst bad bot problem, even though the volume appears smaller.

Bad Bot Sophistication by Industry

Comparing bad bot sophistication levels by industry reveals a very different picture. Ecommerce, healthcare, and airlines see the highest proportion of the sophisticated type, which is inline with the overall data.

Bad Bot Sophistication by Industry, 2017



> The bad bot problem affects every industry. But every company has a unique bad bot problem.

Bad bots continuously target all these industries daily, with defenses requiring constant optimization. Every industry is attacked to check the viability of stolen credentials. Some are hit by sophisticated bots that repeatedly perform a specific task, like checking credit cards numbers. Another may be scraped for pricing content, while a third may be victimized by bad bots checking gift card balances.

Every bot problem is unique and depends on factors such as the nature of the business, its website content and the goal of the adversary. For example, when comparing bad bots to human proportions, gambling companies appear to have a highly significant bad bot problem. But in looking at the proportion of sophisticated bad bots, its problem seems less severe than others. Both of these statements may be correct when isolated, but the conclusion is the same. The bad bot problem affects every industry. But every company has a unique bad bot problem.

Bad Bot Traffic By Website Size 2017

We define website size according to its *Alexa index*[2], whereby sites are ranked by the amount of traffic received. An Alexa score of 1 means it's the most popular internet site—as of this writing that site is Google.com. We used Alexa rankings to categorize site sizes as follows:

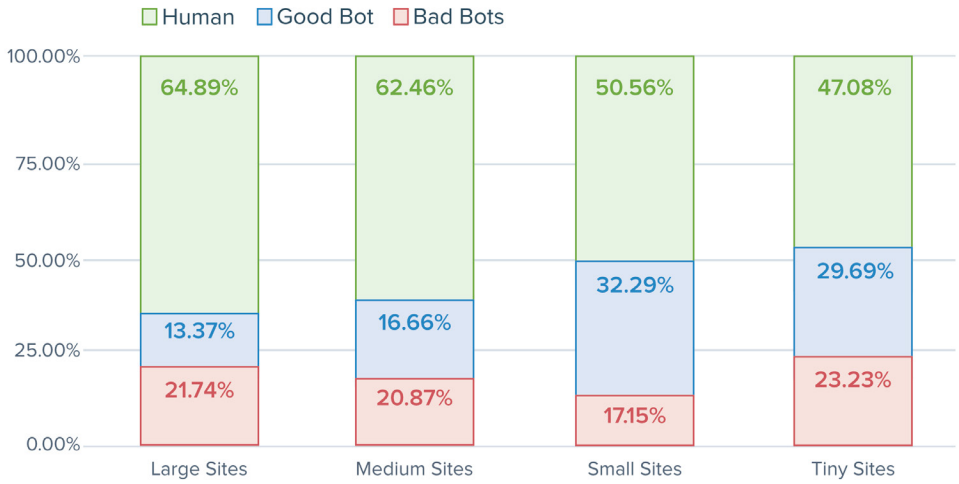
- LARGE = Alexa 1 – 10,000
- MEDIUM = Alexa 10,001 – 50,000
- SMALL = Alexa 50,001 – 150,000
- TINY = Alexa 150,000+

Although bad bots were everywhere in 2017, for the second year in a row large sites were hit the hardest—21.74% of their traffic came from them. This is only slightly less than that of the previous year (21.83%).

Bad bots increased proportionally for every other size of site. Medium sites grew the most (36.97%) over the previous year, followed by small sites which grew 20.75%, while tiny sites grew by 12.98%.

[2] alexa.org

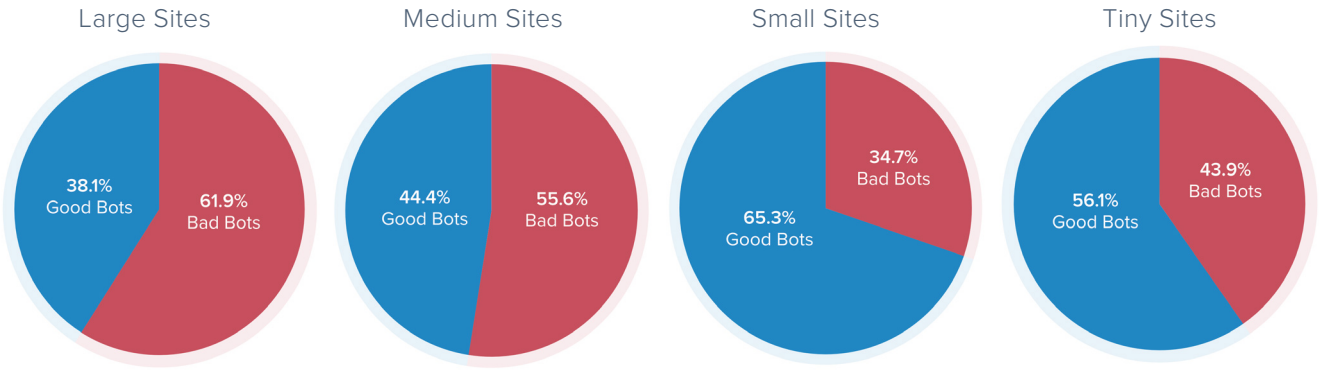
Good Bots, Bad Bots, and Human Traffic to All Sized Sites, 2017



The explanation for the lower proportion of human traffic on smaller sites has to do with how search engines work. Good bots like Googlebot and Bingbot crawl the web more or less equally, regardless of site size. However, larger sites are generally ranked higher in search engine results. Because humans rarely look past the first few search results, small and tiny sites don't get the same level of SEO traffic uplift as do large and medium sites. Each website size has a higher bad to good bot ratio than in 2016. Similar to last year, large and medium sites are more enticing targets for bad bots.

The following four charts show the bad to good bot traffic ratio for large, medium, small, and tiny sites. On large sites, it was 61.9% bad to 38.1% good—a slightly larger spread over 2016. On the other end of the scale, small sites only had a 34.7% bad bot percentage. Tiny sites saw the biggest change from last year, increasing from 28.6% to 43.9% in 2017.

Ratio of Bad Bots to Good Bots on Large, Medium, Small and Tiny Sites, 2017



Bad Bot Identities

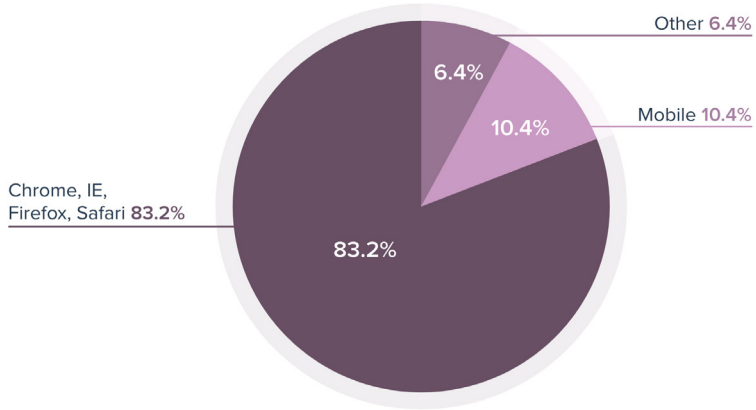
Bad bots must disguise their identity to avoid detection. They do so by reporting their user agent as a web browser or mobile device. In 2017, Chrome continued as the most popular fake identity with almost half (45.48%) of bad bots making this claim, followed by Firefox (26.94%). Safari Mobile was claimed by 5.72% and is the only mobile browser in the top five.

Top Self-Reported Browsers, 2015-2017

	2015		2016		2017	
Rank	User Agent	% of Total	User Agent	% of Total	User Agent	% of Total
1	Chrome	26.90%	Chrome	38.61%	Chrome	45.48%
2	Firefox	17.67%	Firefox	22.94%	Firefox	26.94%
3	Internet Explorer	11.94%	Safari Mobile	8.95%	Internet Explorer	6.47%
4	Safari	5.01%	Safari OSX	7.64%	Safari Mobile	5.72%
5	Android Webkit Browser	4.37%	Internet Explorer	6.70%	Safari OSX	4.33%

In a continuing trend, the majority of bad bots (83.2%) self-reported as either Chrome, Firefox, Safari, or Internet Explorer. Being a smaller proportion than last year, 10.4% asserted they were mobile browsers, such as Safari Mobile, Android, and Opera. The remaining 6.4% reported themselves as other user agents, such as Googlebot and Bingbot.

Bad Bot Reported User Agent Types, 2017



Mobile Browser Bad Bot Requests

An analysis of all requests self-reporting as mobile devices shows that 8.3% were bad bots.

Old Bad Bots Don't Die

Although the majority of bad bots detected in 2017 were programmed to report as being major browsers, not all were up-to-date. For old browsers, the top ten are in the same order as last year. Released in 1999, Internet Explorer 5 was again the oldest.

Clearly, the easiest way to prevent bad bots from hitting your website is to block out-of-date user agents from gaining access.

The 10 Oldest Self-Reported Browsers by Bad Bots, 2017

Year Released	Browser	% of Bad Bot Market Share
1999	Internet Explorer 5	0.037%
2000	Internet Explorer 5.5	0.016%
2001	Internet Explorer 6	0.771%
2002	Netscape 7	0.033%
2004	Firefox 1	0.142%
2005	Netscape 8	0.001%
2006	Internet Explorer 7	1.095%
2006	Firefox 2	0.117%
2007	Netscape 9	0.002%
2008	Firefox 3	0.184%

Why Use Out-of-Date Browsers?

Perhaps some bad bots were written many years ago and remain on the prowl. Some may have targeted systems that only accept specific browser versions. Others may be out-of-control programs, bouncing around the internet in endless loops, still causing collateral damage.

Bad Bots Hide in Data Centers

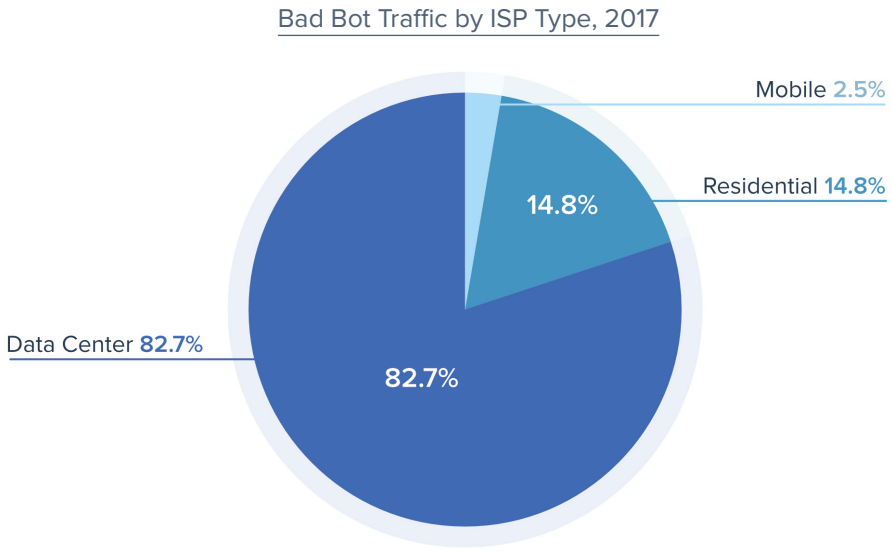
The weaponization of the data center escalates. Over four-fifths (82.7%) of bad bot traffic emanated from them in 2017, compared with 60.1% in 2016. The availability and low cost of cloud computing is the reason for this dominance of data center use.

The economics and success of using low-cost, cloud data centers probably explains why there was a drop in the amount of traffic from residential ISPs, falling from 30.5% to 14.8% in 2017.

Bad bot traffic from mobile ISPs also dramatically decreased— from 9.4% the year prior to 2.2% in 2017. This contradicts the trend where more of the most sophisticated bad bots use mobile ISPs to hide in plain sight. It indicates that mobile ISPs remain far more expensive to use for bad bot traffic compared with data center ISPs, so are only used in specific, targeted cases rather than for large scale attacks.

Why Mobile Botnets Are More Expensive

Renting a mobile botnet is more expensive because it costs more to build. It requires many of the following—installing malware on mobile devices, hacking cell towers, distributing mobile apps through app stores, building mobile device farms, and programming mobile emulators.



Amazon Loses Bad Bot Market Share

Amazon might be happy to see a decline in its bad bot market share, having dropped to second place. However, it still generated 10.62% of all bad bot traffic.

The French company OVH Hosting jumped to number one, responsible for 11.56% of all bad bot traffic. This is 194% growth over its 3.94% contribution last year.

And for the first time, Aliyun (A.K.A Alibaba) made third place, generating 5.64% of bad bot traffic.

Top 10 Bad Bot Originating ISPs, 2017

Rank	ISP	% of Bad Bot Traffic
1	OVH Hosting	11.56%
2	Amazon	10.62%
3	Aliyun Computing Co.	5.64%
4	DigitalOcean	3.20%
5	Microsoft Corporation	2.86%
6	Comcast Cable	2.77%
7	Time Warner Cable	1.83%
8	Choopa, LLC	1.32%
9	HiNet	1.18%
10	Google Cloud	1.10%

Mobile ISPs: The Special Forces Weapon

Data center traffic comprises the majority of bad bot traffic. But mobile ISPs play an important role when bot herders find their data center traffic is blocked. For example, during the *GiftGhostBot*[3] investigation, the bots moved from cheap data center ISPs to more expensive mobile ISPs so as to evade mitigation. Access to a mobile, malware-driven botnet is more expensive than renting one in a data center. The conclusion is that if attack economics make it worth their while, bot operators will spend more to achieve their goals.

Top 10 Mobile ISPs

Rank	Mobile ISP	% of Bad Bot Traffic
1	China Telecom Guangdong	0.70%
2	AT&T Wireless	0.60%
3	T-Mobile USA	0.48%
4	China Mobile	0.38%
5	Verizon Wireless	0.33%
6	KPN	0.27%
7	Orange Espana	0.27%
8	China Telecom Zhejiang	0.26%
9	Vodafone Spain	0.22%
10	Orange	0.21%

[3] Giftghostbot
https://resources.distilnetworks.com/all-blog-posts/giftghostbot-attacks-ecommerce-gift-card-systems

Advanced Mobile Tactics are Still in Their Infancy

Much theoretical talk has been made about bots using advanced techniques like mobile emulators and mobile device farms. Distil's analysis indicates that while these techniques are used by bot operators, they're still not a significant portion of the traffic.

Mobile Emulator Use is Low

Emulators are used on PCs to test and emulate mobile apps. Some bad bot operators use mobile emulators to launch attacks, but such frequency of use is fairly low. Most (93%) prefer to falsify smartphone request formats rather than go the extra length of using a mobile emulator app.

1.2%

Percentage of Devices Targeting Mobile Apps Identify as Mobile Emulators

0.5%

Percentage of Bad Bot Requests Generated by Mobile Emulators

0.5%

Percentage of Devices Targeting Mobile Apps Identified as Suspected Mobile Device Farm

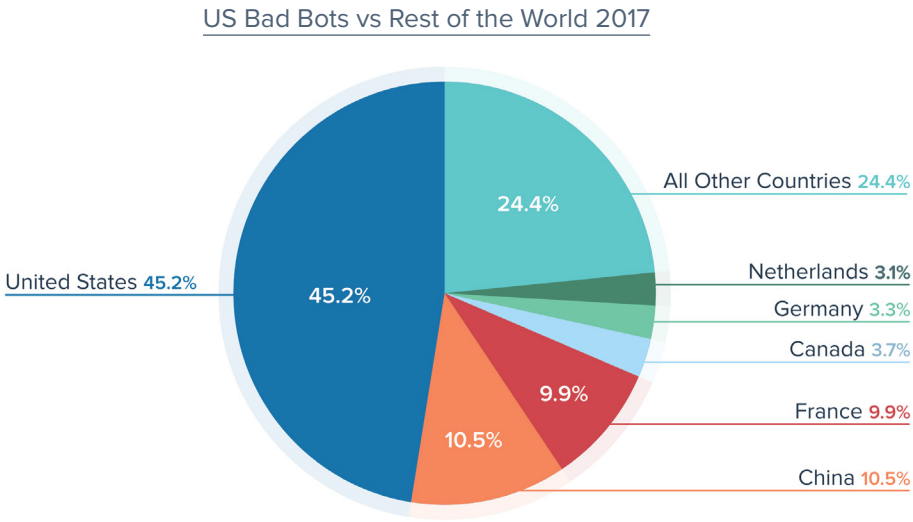
Mobile Device Farms are Used by Few Bot Operators

Mobile device farms like *Amazon AWS Device Farm or Google Firebase* are another technique available to test mobile apps. However, only 0.5% were suspected as being from mobile device farms running legitimate versions of native mobile apps, generating automated requests.

USA: The Bad Bot Superpower

For the fourth year running, the United States topped the list of bad bot originating countries. It remains the true bad bot superpower, but its position is slipping to less than half (45.2%) of bad bot traffic—down from 55.4% in 2016.

China moved into second place, reflecting the rise of Alibaba, and is responsible for 10.5% of bad bot traffic. France (primarily because of the popularity of OVH) has risen three positions from last year, generating 9.9% of such traffic.



Although this is where bad bot traffic originates, the operators aren't necessarily in the same locations. Such cyber criminals could be located anywhere in the world. But they seem to have a preference for US data center use, or at least spreading malware on US devices. That is because if you want to attack US businesses using bad bots, it's advantageous if they're programmed to originate from an American IP address—in that way they're less likely to be blocked.

Top Bad Bot Originating Countries, 2014-2017

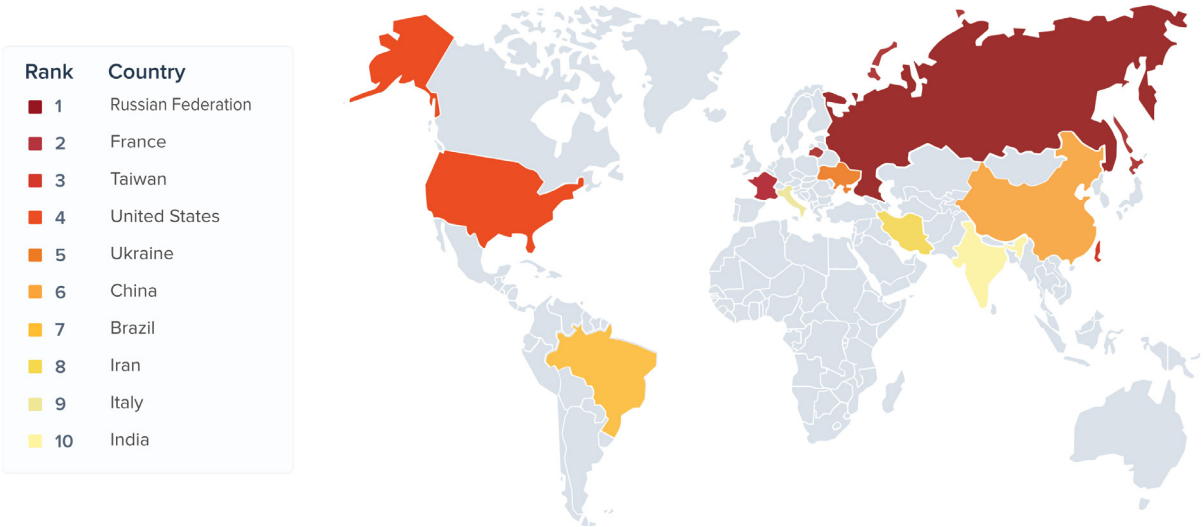
Rank	2014	2015	2016	2017	Change
1	United States	United States	United States	United States	--
2	Germany	India	The Netherlands	China	+1
3	Canada	Israel	China	France	+3
4	Italy	Germany	Germany	Canada	+1
5	France	France	Canada	Germany	-1
6	The Netherlands	United Kingdom	France	The Netherlands	-4
7	China	China	India	United Kingdom	+1
8	Russia	Canada	United Kingdom	Russia	+1
9	United Kingdom	Russia	Russia	Spain	+13
10	India	The Netherlands	South Korea	Ukraine	+18

Of the top ten countries of origin, the Netherlands has moved down four positions to number six.

Spain and the Ukraine moved into the top ten for the first time, while South Korea and India were bumped out.

Russia: The Most Blocked Country

Russia is the most blocked country by Distil Networks customers. Last year's number one—China—dropped to sixth.

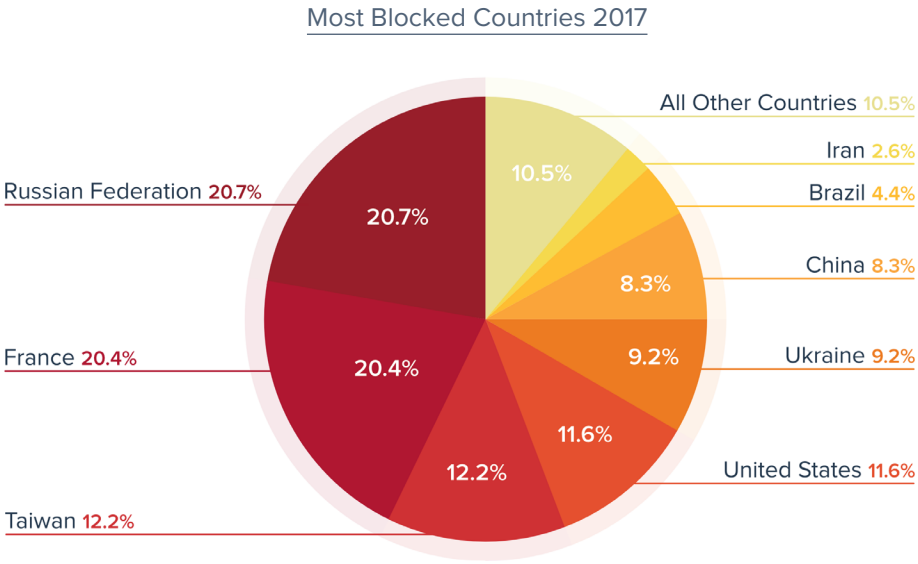


Why Block Countries?

Many companies use geofencing blacklists to choke off large swaths of unwanted traffic. In some cases, it simply doesn't make sense that foreign visitors would use a given site, so blocking chunks of foreign IP addresses is good hygiene. In other situations, customers have suffered attacks from countries that haven't traditionally generated good traffic, so have taken sensible protection measures.

For the first time the United States is on the most-blocked list—explained by international business growth on the Distil platform. And although the United States is by far the most dangerous country, only 11.6% of blocks include the country.

Russia, France, and Taiwan accounted for 53.24% of country-specific block requests. Blocking Russia may be a reaction to the influence of bots on the US presidential election.



Frequency of Account Takeover Attacks

Every time there is a breach and credentials are made readily available, any business with a login page should get ready for a rise in volumetric credential stuffing attacks.

Here, bot operators make two assumptions. The first is that people reuse their credentials on many websites. The second is that newly stolen credentials are more likely to still be active. This is why businesses should anticipate bad bots running those credentials against their website after every breach.

With the increase in breaches and the billions of available stolen credentials, the rise of account and credential fraud activity by bot operators has been startling.

The typical website sees account takeover attacks happen on average 2-3 times per month. But immediately following a breach, the increase in the number of account takeover attacks is 3 times the norm.

3x

Increase in Account Takeover Attacks Following a Breach

2-3

Number of Account Takeover Attacks Seen On Average Website Per Month

SECTION 2

Deep Dive: GiftGhostBot Explained

Identifying the Attack

Distil Networks analysts noticed that companies with gift card processing capabilities had increased malicious bot activity on their websites. In several instances, over half the traffic was on the gift card page alone, indicating a very targeted attack. Some sites saw millions of requests per hour on gift card pages—up to ten times their normal traffic.

GiftGhostBot was identified on nearly 1,000 customer websites, all protected by Distil Networks.

GiftGhostBot: An Advanced Persistent Bot

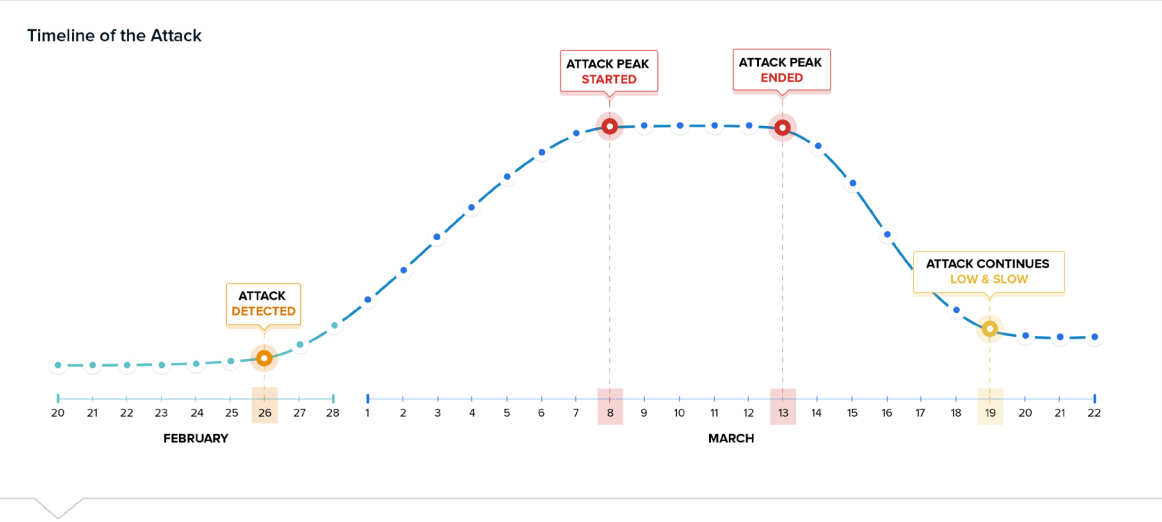
GiftGhostBot has multiple features that confirm it is an Advanced Persistent Bot (APB). First, it lies about its identity by rotating user agent strings. Second, it's heavily distributed across global hosting providers and data centers. Next, it's technically sophisticated in executing JavaScript, which mimics a normal browser. And it's persistent; if blocked using one method, it adapts—using a different attack technique upon its return.

Forensic Analysis

After forensic analysis of GiftGhostBot attributes, Distil identified five main attack profiles. The first three (Profiles 1, 2, 3) show how the attack began and its sophistication. Collectively, GiftGhostBot cycled through 740+ user agents. Once we began blocking the malicious behavior of those profiles, the bot demonstrated its persistence. It vanished, only to morph into two very different profiles (Profiles 4 & 5) upon its return—this time identifying itself as IPhone and Android user agents. Proving the assault was well-funded, its perpetrator didn't hesitate to increase the attack cost—even though each mobile ISP request costs at least five times more.

In examining the numbers of a typical retail customer, we saw some interesting attack characteristics. Using Distil's Hi-Def device fingerprinting technology, we detected on average 6,400 unique fingerprints per hour. Because the device fingerprint provides more accuracy than an IP address and user agent you see the average number of user agents detected were higher at 6,500 per hour, and that IP addresses were detected at an average rate of 29,000 per hour. All of these numbers indicate that the bot was widely distributing itself in its attempt to hide. Its average requests per hour numbered 1,716,000, reaching a peak of 4 million on some sites.

The Anatomy of GiftGhostBot



Profile of the Bot

PROFILE 1
Linuxx86_64
30 USER AGENTS

PROFILE 2
MacIntel
500 USER AGENTS

DETECTED
Once detected, morphed into new profiles.

PROFILE 3
Win32
210 USER AGENTS

DETECTED
Once detected, morphed into new profiles.

PROFILE 4
IPhone
30 IPHONE USER AGENTS

PROFILE 5
Linuxarmv71
54 ANDROID USER AGENTS

Scale of the Attack on a Typical Retailer

DEVICES DETECTED
6,400 Average per hour

USER AGENTS DETECTED
6,500 Average per hour

HIGH
15,000 Average per hour

IP ADDRESSES DETECTED
192.233.186.227
20,900 Average per hour

HIGH
25,000 Average per hour

REQUESTS DETECTED
1,716,000 Average per hour

HIGH
4M Average per hour

Recommendations

Bots are on your website every day, and attack characteristics become more advanced and very nuanced. How should businesses go about protecting themselves? Every site is targeted for different reasons, and usually by different methods, so there is no one-size-fits-all bot defense solution. But there are some proactive steps you can take to start addressing the problem.

Recommendations for Detecting Bad Bot Activity

1. **BLOCK OR CAPTCHA OUTDATED USER AGENTS/BROWSERS:**
   The default configurations for many tools and scripts contain user-agent string lists that are largely outdated. This step won't stop the more advanced attackers, but it might catch and discourage some. The risk in blocking outdated user agents/browsers is very low; most modern browsers force auto-updates on users, making it more difficult to surf the web using an outdated version.

   We recommend you block or CAPTCHA the following browser versions:

	BLOCK End of Life More than 3 years	**CAPTCHA** End of Life More than 2 years
Firefox version	<38	<45
Chrome version	<41	<49
Internet Explorer version	<10	10
Safari version	<9	9

2. **BLOCK KNOWN HOSTING PROVIDERS AND PROXY SERVICES:**
   Even if the most advanced attackers move to other, more difficult-to-block networks, many less sophisticated perpetrators use easily accessible hosting and proxy services. Disallowing access from these sources might discourage attackers from coming after your site, API, and mobile apps.

   Block these data centers:

Digital Ocean	OVH SAS	OVH Hosting
Amazon.com	DigitalOcean	Choopa, LLC
GigeNET		

3. PROTECT EVERY BAD BOT ACCESS POINT

Be sure to protect exposed APIs and mobile apps—not just your website—and share blocking information between systems wherever possible. Protecting your website does little good if backdoor paths remain open.

4. CAREFULLY EVALUATE TRAFFIC SOURCES

Monitor traffic sources carefully. Do any have high bounce rates? Do you see lower conversion rates from certain traffic sources? These can be signs of bot traffic.

5. INVESTIGATE TRAFFIC SPIKES

Traffic spikes appear to be a great win for your business. But can you find a clear, specific source for the spike? One that is unexplained can be a sign of bad bot activity.

6. MONITOR FOR FAILED LOGIN ATTEMPTS

Define your failed login attempt baseline, then monitor for anomalies or spikes. Set up alerts so you're automatically notified if any occur. Advanced "low and slow" attacks don't trigger user or session-level alerts, so be sure to set global thresholds.

7. MONITOR INCREASES IN FAILED VALIDATION OF GIFT CARD NUMBERS

An increase in failures, or even traffic, to gift card validation pages can be a signal that bots such as GiftGhostBot are attempting to steal gift card balances.

8. PAY CLOSE ATTENTION TO PUBLIC DATA BREACHES

Newly stolen credentials are more likely to still be active. When large breaches occur anywhere, expect bad bots to run those credentials against your site with increased frequency.

9. EVALUATE A BOT MITIGATION SOLUTION

The bot problem is an arms race. Bad actors are working hard every day to attack websites across the globe. The tools used constantly evolve, traffic patterns and sources shift, and advanced bots can even mimic human behavior. Hackers using bots to target your site are distributed around the world, and their incentives are high. In early bot attack days you could protect your site with a few tweaks; this report shows that those days are long gone. Today it's almost impossible to keep up with all of the threats on your own.

About Distil Networks

Distil Networks, the global leader in bot mitigation, protects websites, mobile apps, and APIs from automated threats. Fraudsters, hackers, and competitors use bots to commit online fraud, break into customer accounts, and gain an unfair competitive advantage.

As the sheer volume, sophistication, and business damage of these attacks grow, bots put a costly strain on IT staff and resources. Only Distil's unique, more holistic approach provides the vigilant service, superior technology, and industry expertise needed for full visibility and control over this abusive traffic.

The Distil team pioneered bot mitigation in 2011, and has been leading the way ever since. With Distil, there is finally a defense against automated attacks that is as adaptable and vigilant as the threat itself.

For more information on Distil, visit https://www.distilnetworks.com/block-bot-detection/ or follow @DISTIL on Twitter.

Confidentiality Statement