imperva

REPORT | 2020

Bad Bot Report

Bad Bots Strike Back

Contents

About the Bad Bot Report

Imperva's *2020 Bad Bot Report* investigates the daily attacks that sneak past sensors and wreak havoc on websites.

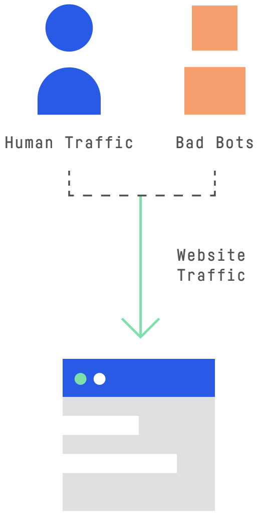
This is the 7th Annual Bad Bot Report. It's based on 2019 data collected from Imperva's global network and includes hundreds of billions of bad bot requests anonymized over thousands of domains. Our goal is to offer guidance about the nature and impact of automated threats to those of you on the frontlines 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 of which manipulate lower-level network protocols.

Bad bots interact with applications in the same way a legitimate user would, making them harder to prevent. They enable high-speed abuse, misuse, and attacks on your websites, mobile apps, and APIs. They allow bot operators, 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, brute-force login, digital ad fraud, spam, transaction fraud, and more.

The Rebranding of Bad Bots



Bad bots have long been the scourge of the internet. They lurk amidst real human traffic. Many businesses misunderstand the negative impacts of unfettered automated traffic. But others know that bad bots are not benign and have a very focused motivation—to make money.

We've come a long way from the early days of basic ticket scalping bots. Today, on sites like ticketbots.net, you can purchase a sophisticated range of customized spinners, drop checkers, ticket downloaders, and pdf generators to purchase tickets to any event on any platform globally.

To jump to the front of the line to buy limited-edition sneakers, it is easy to purchase any of the hypebots or sneaker bots available on websites like aiobots.com, hypebots.com, or anothernikebot.com.

But the next evolution of bad bot development is already underway. Bad bots are trying to improve their image and appear legitimate. This new wave of bot operators are building businesses on their ability to scrape proprietary data from websites, package the data, and provide competitive data feeds to any company willing to purchase—all neatly packaged as "business intelligence" services.

Bad Bots as a Service

This rebranding of "bad bots as a service" demonstrates itself in many ways. First, through the adoption of professional looking websites offering business intelligence services called pricing intelligence, alternative data for finance, or competitive insights. Typically, these businesses offer data products focused on specific industries. Second, there is increased pressure to purchase scraped data within your industry. No business wants to lose in the market place because the competition has access to data that is available to purchase. Finally, there is the growth of job postings looking for people to fill positions with titles like Web Data Extraction Specialist or Data Scraping Specialist. In this environment, it is difficult to see the bot problem disappearing any time soon.

The Legality of Web Scraping

In the most significant legal ruling in the ongoing HiQ vs Linkedin case[1], the Ninth Circuit appellate court ruled in favor of allowing bots to scrape publicly available content. Linkedin is attempting to prevent the automated scraping of profiles by aggregator HiQ and is appealing the ruling. The litigation may yet end up in the United States Supreme Court[2].

[1] Data Scraping Survives! (At Least for Now) Key Takeaways from 9th Circuit Ruling on the HIQ vs. Linkedin Case – The National Law Review – September 30, 2019

[2] LinkedIn Files Petition to the Supreme Court in HiQ Web Scraping Case – The National Law Review – March 12, 2020

The Rise of Mega Credential Stuffing Attacks

Beyond content and price scraping, the biggest bad bot problem is credential stuffing and credential cracking. Every website with a login is subject to these attacks and a new phenomenon is emerging—the rise of mega credential stuffing attacks targeting one company.

A recent attack that Imperva mitigated lasted 60 hours and included 44 million login attempts. In general, the availability of billions of breached credentials has helped fuel the rise in credential stuffing, but such large scale attacks can cause significant infrastructure strain leading to slowdowns or downtime. These large application layer credential stuffing attacks might be as damaging as volumetric DDoS attacks to any organization unprepared to handle such a high volume of bad bot requests.

A recent attack that Imperva mitigated lasted 60 hours and included 44 million login attempts

Breaches and Loyalty Programs

With the continuing proliferation of customer loyalty programs online, credential stuffing attacks are increasingly lucrative. Inside every loyalty program account is some digital currency available to redeem or transfer to another account. The availability of credentials from data breaches combined with the growth of online loyalty programs is providing the perfect platform for bot operators to attack every e-commerce business. The rise of problems associated with account takeover is unfortunately inevitable.

Nonprofits Suffering from Unscrupulous Bot Operators

Stolen credit card numbers are a problem for more than the card owner. Credit card enumeration is run against website payment processors to determine if the credit card is valid—and non-profits suffer more than most. Bot operators enumerate through credit card numbers and make small donations to nonprofit organizations. If the donation is successful using a card number, the bot operator knows that the credit card is valid and can be used elsewhere to commit further fraud. But the problem for the nonprofit doesn't end there. The card owner will see the fraud on their account and complain to the credit card company. The credit card company now has to deal with chargebacks involving the non-profit. While the unsuspecting non-profit cannot afford to spend time and incur fees refunding such fraudulent donations.

Every Industry Has a Different Bot Problem

While the goal of each bad bot operator might be different depending on their industry, bots are the tool of choice and are vital to their success. There is an ecosystem within many industries that rely on bots for survival. Without their use, many such operators would struggle to compete. In many cases, deploying bad bots is an essential business practice.

Every industry has its own bad bot problem and ecosystem of bot operators. Some of these include:

- AIRLINES – There is an ecosystem of online travel agents, aggregators, and competitors that use bots to scrape content—including flight information, pricing, and seat availability—while criminals attempt to fraudulently access user accounts that contain loyalty program awards and credit card information[3].

- **E-COMMERCE** – Competitors use bad bots to aggressively scrape pricing and inventory information. Grinchbots and Sneakerbots create denial of inventory problems for customers seeking limited edition items. Criminals use bad bots to commit fraud by stealing gift card balances and to access user accounts and credit card information[4].

- **EVENT TICKETING** – Brokers, scalpers, hospitality agencies, and corporations use bad bots to check for ticket availability and to purchase available seats to resell on secondary markets. Criminals access user accounts to steal tickets and credit card information[5].

Bad bots will continue to strike back.

Bad Bots Increase Infrastructure Costs

For any business whose website, mobile app, or API is the unfortunate target of malicious bots, they have to deal with more problems. Not only does it have to deal with the competitive pricing pressure resulting from the scraping bots, but it has to maintain infrastructure uptime and redundancy so that real customers aren't inconvenienced. In addition, they also suffer from skewed decision-making metrics because their web traffic has been polluted by bad bots.

Social Media Bad Bots in Elections

Influencer bots are a tool used to spread propaganda. The role of influencer bots on social media will take center stage as the United States presidential election gets closer. Automated traffic launched by bot operators who remotely manage a vast number of aggregated social media accounts will aim to influence and change votes.

Bad Bots Strike Back

The bot problem is real for every website and mobile app. Businesses have tried to protect themselves by adding bot protection capabilities to their solutions. But bot operators are expanding their operations and evolving into legitimate businesses. With increased financial resources, bot operators are also developing new methods to evade common bot detection techniques that ensure the arms race continues.

Only recently have business leaders become savvy to what bad bots do. Many are incredulous about the scams being perpetrated. One thing is certain, with the rebranding of bot operations into business intelligence companies, the hiring of professional data extraction experts, and investment in new techniques to evade detection, bad bots will continue to strike back.

[3] How Bots Affect Airlines – Imperva Threat Research

[4] How Bots Affect E-commerce – Imperva Threat Research

[5] How Bots Affect Ticketing – Imperva Threat Research

Understanding What Bad Bots Do

BAD BOT PROBLEM	HOW IT HURTS THE BUSINESS	SIGNS YOU HAVE A PROBLEM	INDUSTRIES TARGETED
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	All businesses that show prices • E-commerce • 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 • Finance • Ticketing
Account Takeover (a.k.a., Credential Stuffing, Credential Cracking)	• Stolen credentials tested on your site. If successful, the ramifications 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 chargebacks	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 • Dating sites • Communities Promotion abuse • Gambling
Credit Card Fraud (a.k.a., Carding, Card Cracking)	• Criminals testing credit cards numbers to identify missing data (exp. date, CVV). • Damages the fraud score of the business • Increases customer service costs to process fraudulent chargebacks	• Rise in credit card fraud • Increase in customer support calls • Increased chargebacks processed	Any site with a payment processor • E-commerce • Nonprofit/Charities • Airlines • Travel • Ticketing • Financial • Gambling

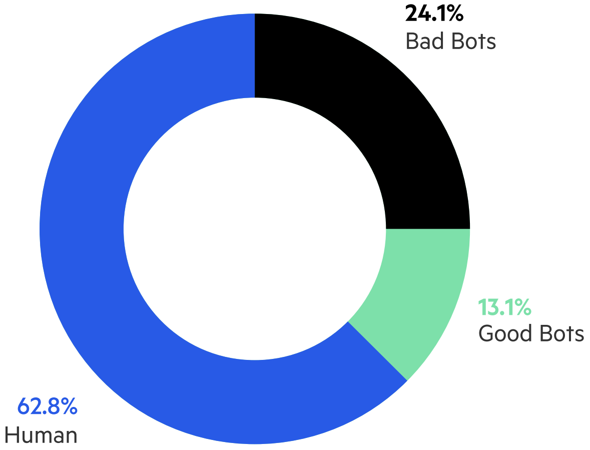
BAD BOT PROBLEM	HOW IT HURTS THE BUSINESS	SIGNS YOU HAVE A PROBLEM	INDUSTRIES TARGETED
Denial of Service	• Slows the website performance causing brownouts or downtime • Lost revenue from unavailability of website • Damaged customer reputation	• Abnormal and unexplained spikes in traffic on particular 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 the gift card balance page. • Increase in customer service calls about lost balances	E-commerce
Denial of Inventory	• Bots hold items in shopping carts, preventing access by valid customers • Damaged customer reputation because unscrupulous middle men hold all inventory until resold elsewhere	• Increase in abandoned items held in shopping carts • Decrease in conversion rate • Increase in customer service calls about lack of availability of inventory	Scarce or time-sensitive items • Airlines • Tickets • E-commerce (Sneakers)

Executive Summary of Findings

Bad Bot Traffic Rises to Highest Ever

In 2019, bad bot traffic rose to its highest ever percentage of 24.1 percent of all traffic. 37.2 percent of all internet traffic wasn’t human. Human traffic increased by 1.1 percent to 62.8 percent of all traffic.

Bad Bot v Good Bot v Human Traffic 2019



Bad Bots Amount all Website Traffic in 2019	**24.1%**
Percentage change in bad bot traffic from previous year	▲ 18.1%

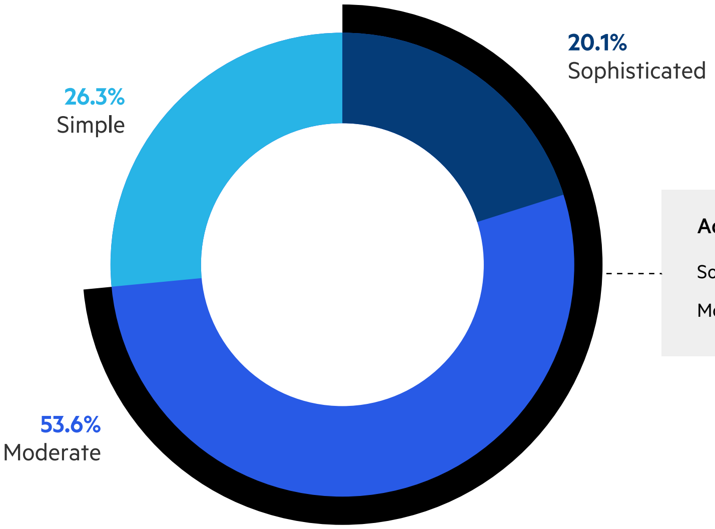
Good Bots Traffic Percentage in 2019	**13.1%**
Percentage change in good traffic from previous year	▼ 25.1%

Human Website Traffic Percentage in 2019	**62.8%**
Percentage change in human traffic from previous year	▲ 1.1%

Bad Bot Sophistication Levels Remain Consistent for the Third Year

Advanced persistent bots (APBs) continue to plague websites and often avoid detection. APBs cycle through random IP addresses, enter through anonymous proxies, change their identities, and mimic human behavior.

Bad Bot Sophistication Levels 2019



Advanced Persistent Bots	**73.7%**
Sophisticated Bad Bots	20.1%
Moderate Bad Bots	53.6%

The Bot Problem Affects Every Industry

Every business has a unique bot problem. Some bad bot problems run across all industries while others are industry-specific. Websites with login screens are hit by bot-driven account takeover attacks. Content and price scraping is rampant and is undertaken by bots.

TOP 5 INDUSTRIES BAD BOT TRAFFIC %		
1	Financial	47.7%
2	Education	45.7%
3	IT & Services	45.1%
4	Marketplaces	39.8%
5	Government	37.5%

TOP 5 INDUSTRIES SOPHISTICATED BAD BOT TRAFFIC %		
1	Marketplaces	28.5%
2	Real Estate	24.3%
3	Ticketing	22.5%
4	IT & Services	22.1%
5	Nonprofits	20.4%

More than Half of Bad Bots Claim to Be Google Chrome

Bad bots continue to follow the trends in browser popularity, impersonating the Chrome browser 55.4 percent of the time. The use of data centers reduced again in 2019 with 70 percent of bad bot traffic emanating from them—down from 73.6 percent in 2018.

Bad bots report as either Chrome, Firefox, Internet Explorer, Safari	79.4%
Bad bots hiding in data centers	70.0%
Bad bots using Amazon ISP	11.6%

Bad Bots Are All Over the World

With most bad bot traffic emanating from data centers, the U.S. remains the “bad bot superpower” with 45.9 percent of bad bot traffic coming from the country. For the third year in a row, the most blocked attacks originate in Russia (21.1 percent). Bots deployed from Amazon reduced significantly to 11.6 percent.

U.S. remains the “bad bot superpower”

TOP 5 BAD BOT TRAFFIC BY COUNTRY		
1	United States	45.9%
2	Netherlands	8.0%
3	Canada	6.3%
4	China	4.8%
5	Germany	4.1%

TOP 5 MOST BLOCKED COUNTRY		
1	Russian	21.1%
2	China	19.0%
3	Romania	8.6%
4	Turkey	8.5%
5	Vietnam	6.6%

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. Credential stuffing to perform account takeover is a prominent tactic of bad bots.

The Open Web Application Security Project (OWASP) provides a list of the different bad bot types in its *Automated Threat Handbook*.[6]

How Do Good and Bad Bots Differ?

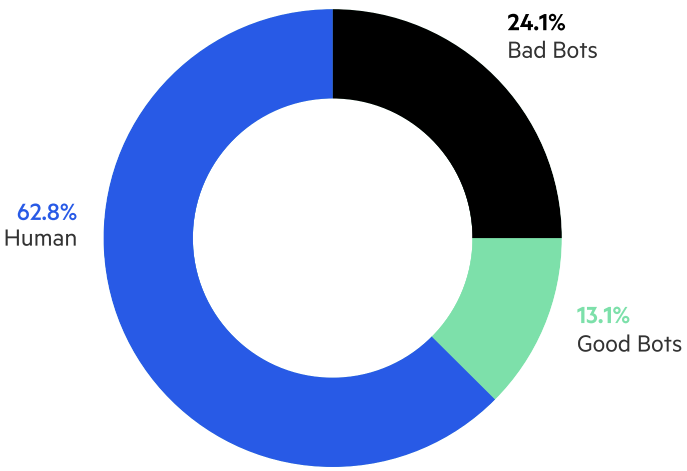
In simplistic terms, good bots ensure that online businesses and their products can be found by prospective customers. Examples include search engine crawlers such as GoogleBot and Bingbot that, through their indexing, help people match their queries with the most relevant sets of websites.

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. If your website analytics are polluted with bots, any decisions based on the origin of that traffic is potentially flawed. Being able to intelligently separate traffic generated by legitimate human users, good bots, and bad bots is essential for making informed business decisions.

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 v Good Bot v Human Traffic 2019



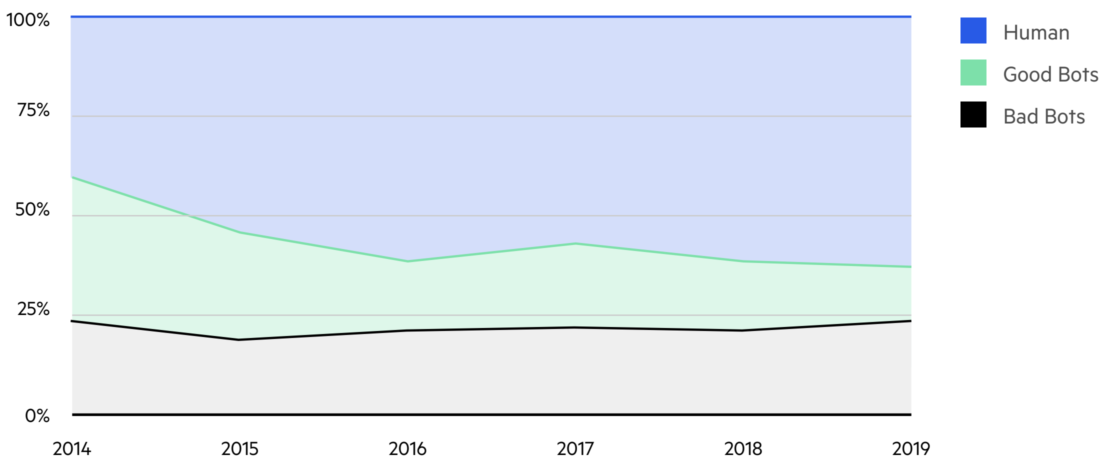
[6] Automated Threat Handbook – OWASP – February 2018

Bad bot traffic percentage is the highest ever

In 2019, bad bots accounted for 24.1 percent of all website traffic—an 18.1 percent increase over the prior year. This was the highest percentage of bad bot traffic seen since The Bad Bot Report's inception.

Good bots decreased by 25.1 percent compared with the prior year, accounting for 13.1 percent of all traffic. This past year the proportion of human traffic increased by 1.1 percent, totalling 62.8 percent of all internet traffic.

Bad Bot v Good Bot v Human Traffic 2014 - 2019



Bad Bot v Good Bot v Human Traffic 2014-2019

	2014	2015	2016	2017	2018	2019
Bad Bots	22.8%	18.6%	19.9%	21.8%	20.4%	24.1%
Good Bots	36.3%	27.0%	18.8%	20.4%	17.5%	13.1%
Humans	40.9%	54.4%	61.3%	57.8%	62.1%	62.8%

The bad bot traffic percentage is the highest ever, eclipsing the previous peak in 2014—and accounts for almost 1 in 4 web requests.

The good news is that the percentage of human users is up for the third year in a row. But it is worth reiterating that human traffic comprises only 62.8 percent of all internet traffic. When the goal is to attract real humans to your website, these numbers show that the bot problem remains a major problem.

Bad Bot Sophistication Levels

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

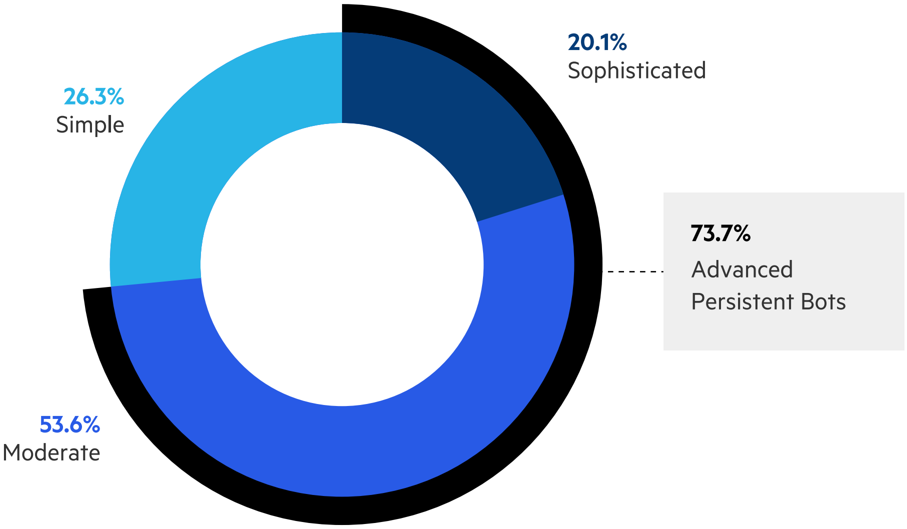
SIMPLE – Connecting from a single, ISP-assigned IP address, this type connects to sites using automated scripts, not browsers, and doesn't self-report (masquerade) as being a browser.

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 are a combination of moderate and sophisticated bad bots. 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. They use a mix of technologies and methods to evade detection while maintaining persistence on target sites.

Bad Bot Sophistication Levels 2019



For the third year in a row, the sophistication levels are very consistent.

Simple bots, which are easiest to detect, accounted for 26.3 percent of bad bot traffic. Meanwhile, the majority of non-human traffic (53.6 percent) came from those classified as moderate. And sophisticated bad bots, the most difficult to detect, comprised of 20.1 percent of automated traffic last year.

Advanced persistent bots (APBs) accounted for 73.7 percent of all 2018 bad bot traffic—slightly higher than the prior year. Because they can cycle through IP addresses and switch user agents, simple IP blacklisting is wholly ineffective.

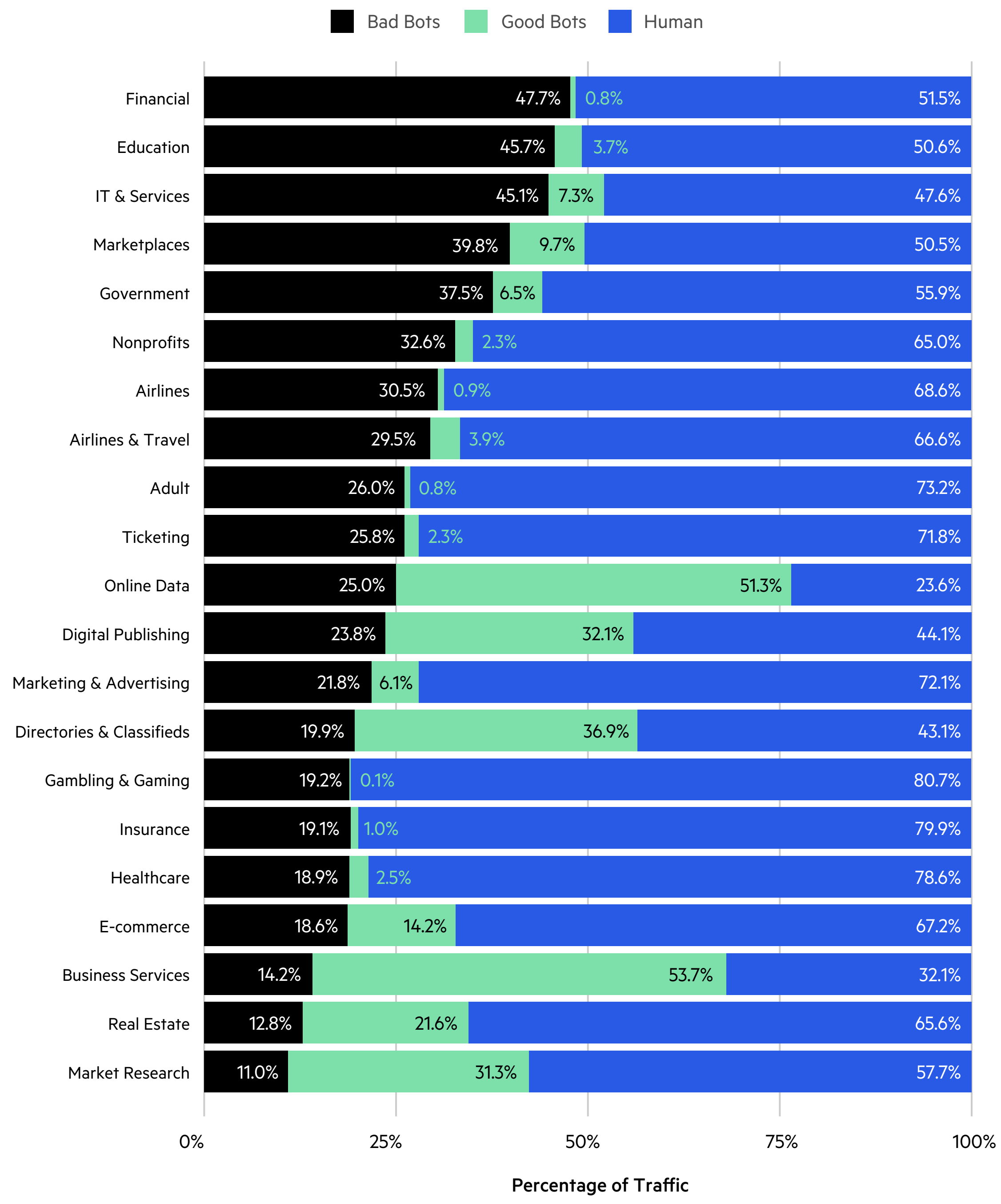
APBs, sometimes known as low and slow bots, carry out significant attacks using fewer requests and can even delay requests, all the while staying below request rate limits. This method reduces the 'noise' generated by many bad bot campaigns.

Bad Bots by Industry

By examining traffic from various industries, a deeper insight into the bot problem is possible.

As more organizations add bot protection to their security profile, a larger data set is gathered across more industries. For the 2019 Bad Bot Report, data was collected from 20 industries. For this report, the number of industries expanded to 21 by adding nonprofit organizations.

Bad Bot v Good Bot v Human Traffic 2014-2019



Note: Minimum required to include an industry segment = 75 million requests.

FINANCIAL SERVICES COMPANIES, for the second year, have the highest percentage of bad bots with 47.7 percent. Such companies typically suffer from bad bots attempting to access user accounts using credential stuffing.

EDUCATION had 45.7 percent bad bot traffic. Bots are deployed by malicious operators looking for research papers, class availability, and to access user accounts.

MARKETPLACES are another industry that suffers from a high percentage of bad bots, comprising 39.8 percent of traffic. This is similar to the bots on e-commerce sites that scrape prices and content and attack account logins.

GOVERNMENT, with 37.5 percent of bad bots, is interested in protecting business registration listings from scraping bots, and in stopping election bots from interfering with voter registration accounts.

NONPROFIT ORGANIZATIONS have 32.7 percent bad bot traffic. Bots using the donation pages to test stolen credit card numbers are a nuisance and a financial burden that many nonprofits cannot afford to endure.

AIRLINES have a very complex problem with 30.5 percent of their traffic comprising bad bots. 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. In addition, airlines suffer from account takeover issues as bad bot operators attempt to get into user accounts and empty them of accumulated air-mile balances.

TICKETING, one of the first industries ever targeted by bad bots, has 25.8 percent automated traffic. Scalping bots, seat inventory checkers, and credential stuffing bots that access user accounts are most prevalent on these sites.

GAMBLING AND GAMING COMPANIES, with 19.2 percent bad bot traffic, suffer from aggregators that relentlessly scrape for ever-changing betting lines. Account takeovers are also a major problem because each account contains money or loyalty points that, once compromised, can easily be transferred to another user and emptied.

E-COMMERCE sees a wide range of bad bot attacks. These include price scraping, content scraping, account takeovers, credit card fraud, and gift card abuse. Having one of the largest datasets, e-commerce has 18.6 percent of the bad bot traffic.

Financial Services has the highest percentage of bad bots with

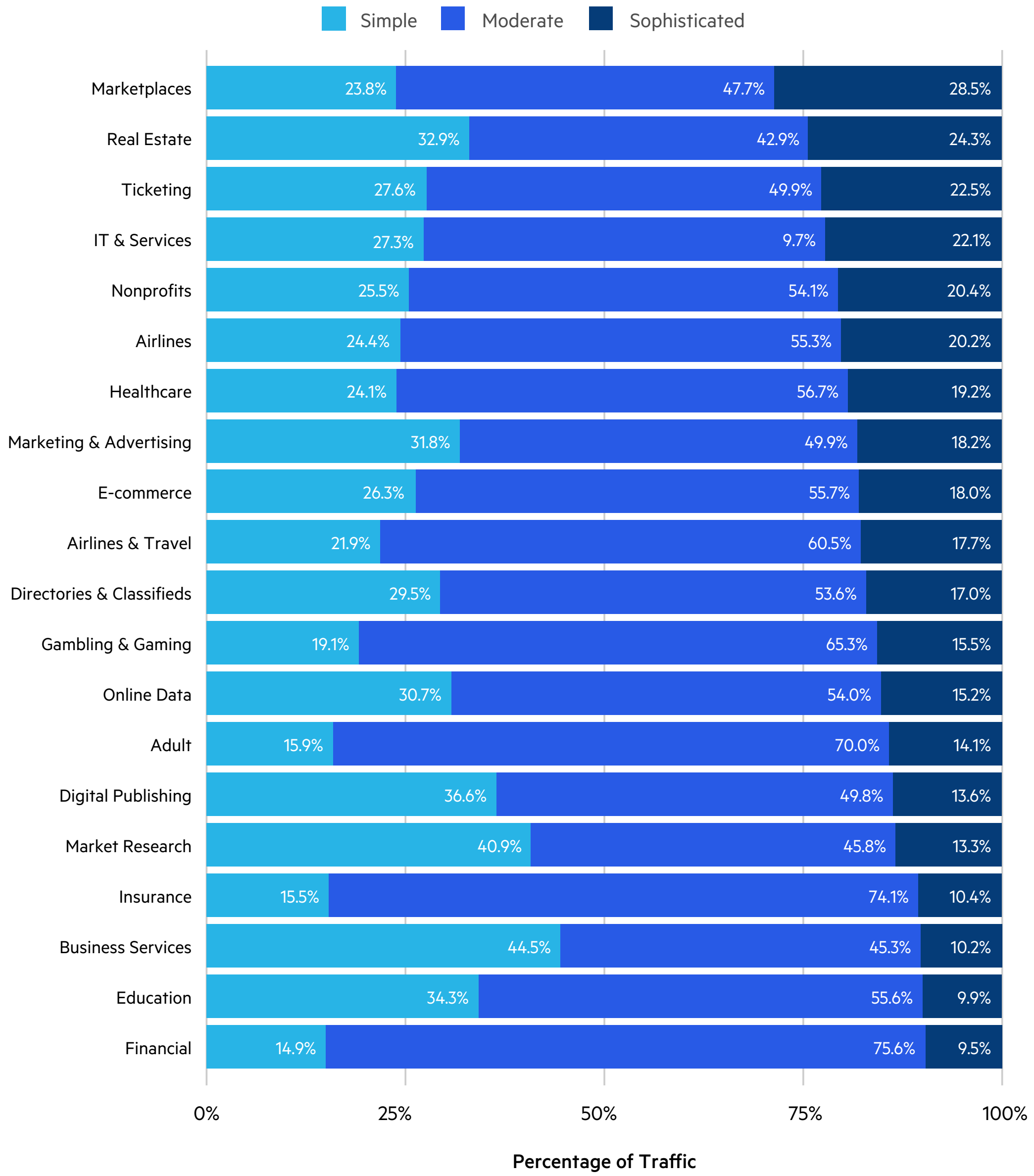
47.7%

Bad Bot Sophistication by Industry

Comparing bad bot sophistication levels by industry reveals a very different picture. Marketplaces, real estate, ticketing, IT & Services, nonprofits and airlines see the highest proportion of sophisticated bots.

It's important to understand that the volume of bots doesn't necessarily align with the sophistication of bot attacks. For example, a sophisticated bot may make fewer requests to achieve its goal.

Bad Bot Sophistication for 2019 – By Industry



Bad bots continuously target all of 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, such as checking credit card numbers. Another may be scraped for pricing content, while a third may be victimized by bad bots checking gift card balances.

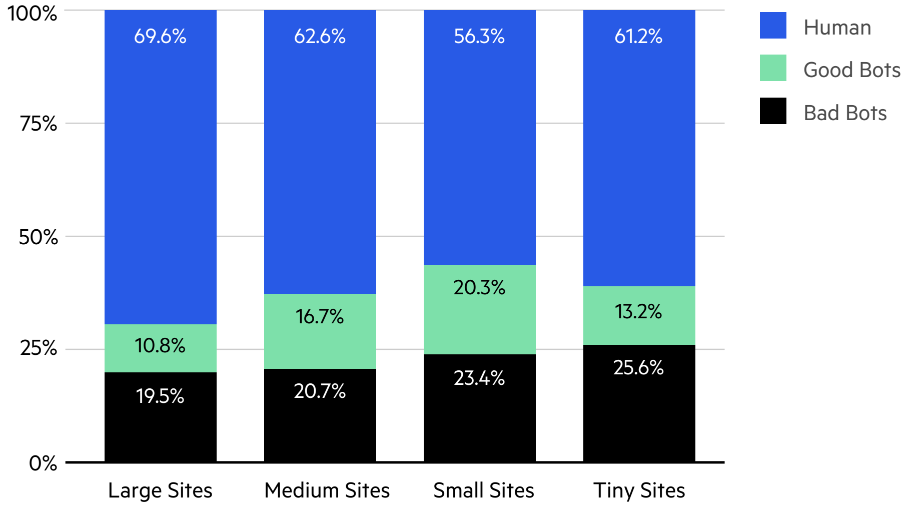
Every bot problem is unique; factors to consider include the nature of the business, its website content, and the goal of the adversary.

Bad Bot Traffic by Website Size

In this report, Imperva defines website size according to its Alexa index[7], 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's Google.com. We used Alexa rankings to categorize sizes as follows:

- Large = Alexa 1 – 10,000
- Medium = Alexa 10,001 – 50,000
- Small = Alexa 50,001 – 150,000
- Tiny = Alexa 150,000+

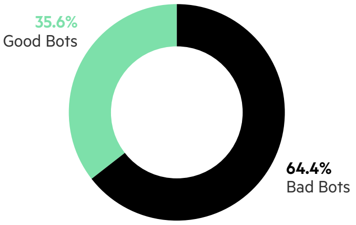
Bad Bot v Good Bot v Human to All Sizes Sites 2019



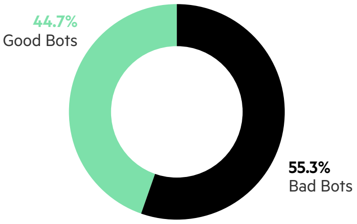
Bad bot volume is up for every website size. Tiny sites have the highest proportion of bad bot traffic at 25.6 percent.

The following four charts show the bad to good bot traffic ratio for large, medium, small, and tiny sites. The highest ratio of bad bots (65.9 percent) to good bots (34.1 percent) is on large sites. For the first time, there are more bad bots compared to good bots on each size of website.

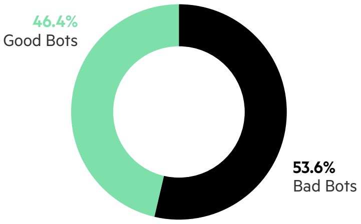
Bad Bot v Good Bot Ratio on Large Sites 2019



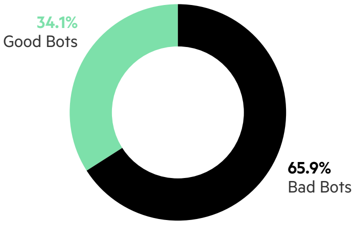
Bad Bot v Good Bot Ratio on Medium Sites 2019



Bad Bot v Good Bot Ratio on Small Sites 2019



Bad Bot v Good Bot Ratio on Tiny Sites 2019



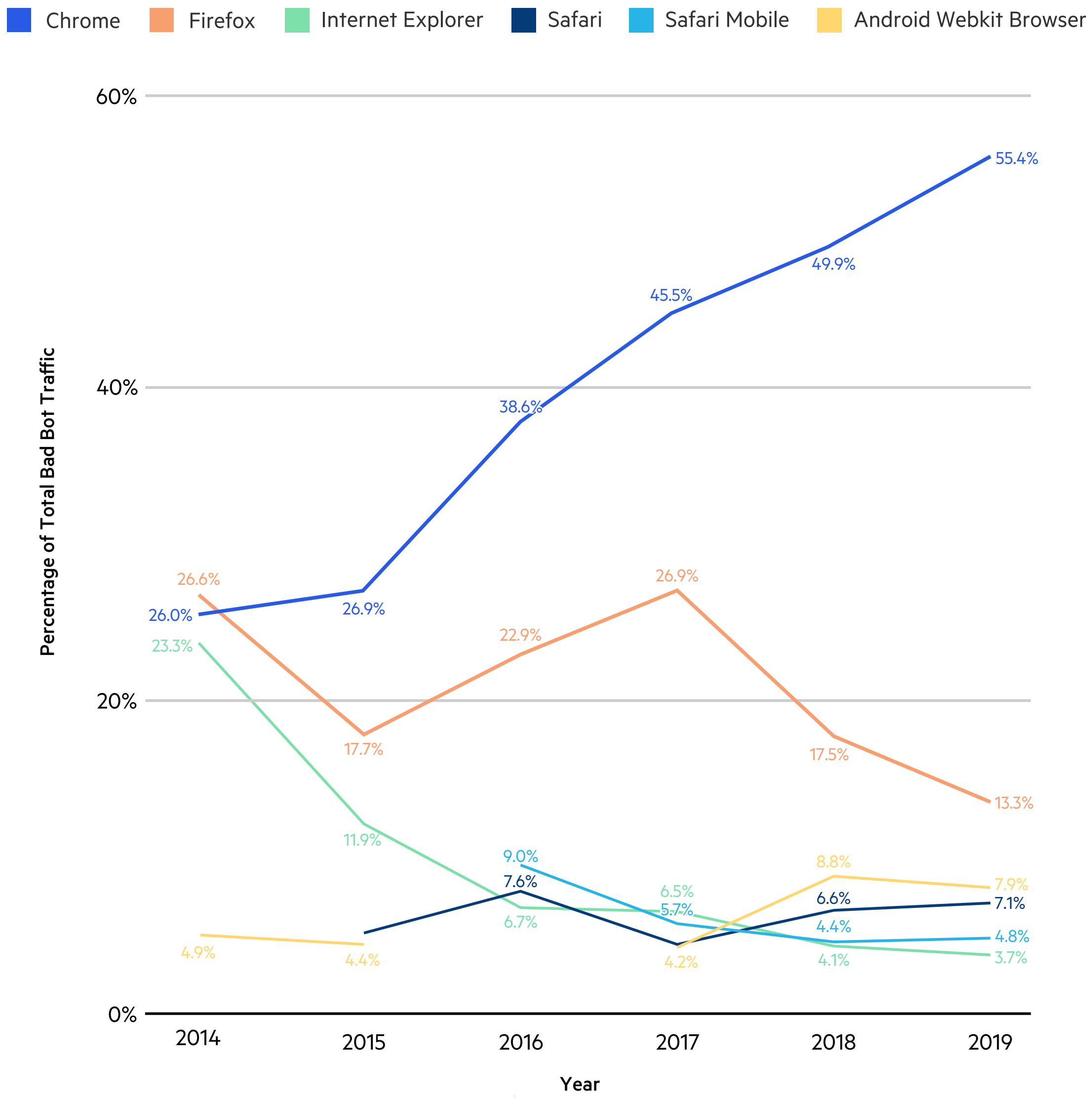
[7] alexa.com

Bad Bot Identity: Impersonating Chrome

Bad bots must disguise their identity to avoid detection. They do so by reporting their user agent as a web browser or mobile device. While the majority of bad bots claim to be the most popular browsers, during 2019 bad bots claimed a total of 517 different identities (user agents), only six less than in 2018.

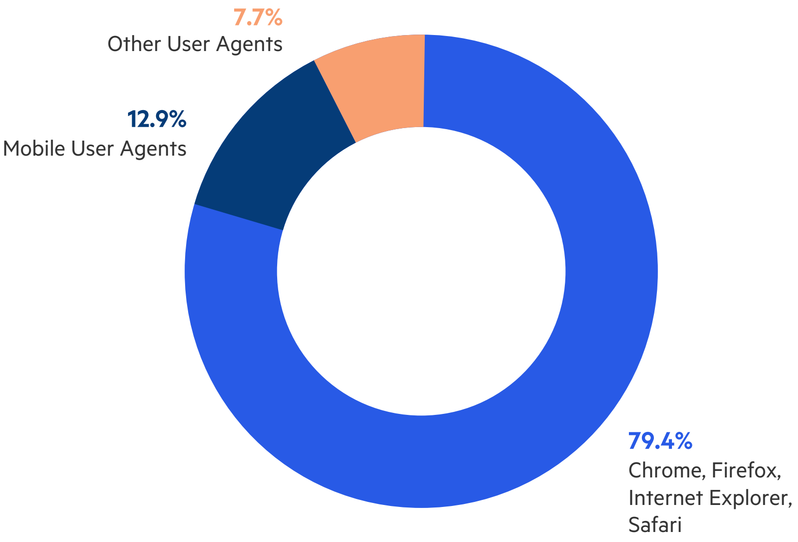
In 2019, Chrome continued the trend of being the most popular fake identity used by bad bots, with over half (55.4 percent) of them making this claim. Firefox dropped for the second year in a row to 13.3 percent but is still the second most popular claimed identity. Android Webkit Browser dropped in popularity and was claimed by 7.9 percent, and is the only mobile browser in the top three.

Top Self Reporting Browser by Bad Bots 2014–2019

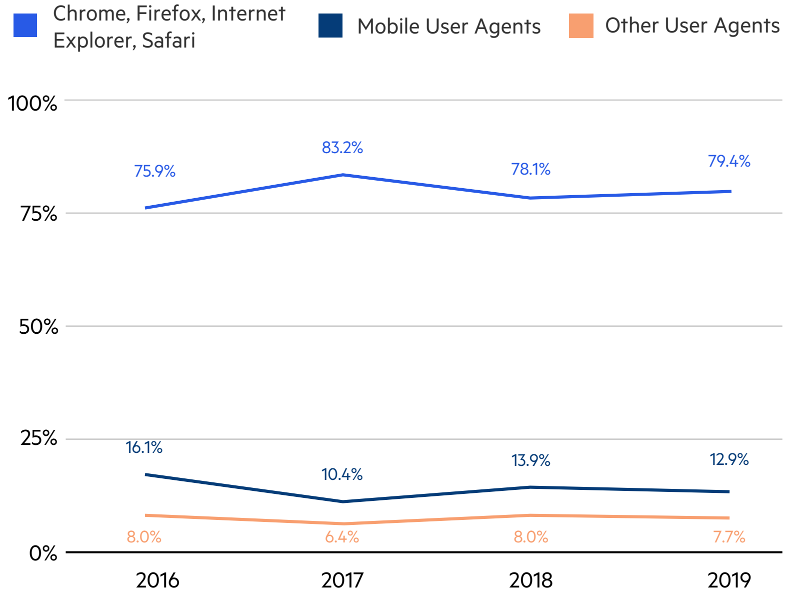


An increasing majority of bad bots (79.4 percent) are self-reporting as either Chrome, Firefox, Safari, or Internet Explorer, slightly higher than the previous year. Mobile browsers such as Safari Mobile, Android, and Opera decreased to 12.9 percent from 13.9 percent last year. The remaining 7.7 percent reported themselves as other user agents, such as Googlebot and Bingbot.

Bad Bot Reported User Agent Types 2019



Bad Bot Reported User Agent Types 2016–2019



Bad Bots Are Still Growing Old

A small number of bad bots are not trying too hard to hide. Examining the age of the browsers claimed by bad bots reveals that a small amount is using ones that were released over 20 years ago. For old browsers, the top ten are in the same order as last year but we can see that in 2019 the percentage of traffic claiming these browser versions is decreasing for each compared to the previous 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 2018-2019

		2018	2019
Year Released	Browser	Bad Bot Market Share %	Bad Bot Market Share %
1999	Internet Explorer 5	0.044%	0.042%
2000	Internet Explorer 5.5	0.009%	0.003%
2001	Internet Explorer 6	0.699%	0.343%
2002	Netscape 7	0.051%	0.005%
2004	Firefox 1	0.111%	0.041%
2005	Netscape 8	0.002%	0.001%
2006	Internet Explorer 7	0.823%	0.479%
2006	Firefox 2	0.135%	0.064%
2007	Netscape 9	0.002%	0.001%
2008	Firefox 3	0.116%	0.002%

Why Use Out-of-Date Browsers?

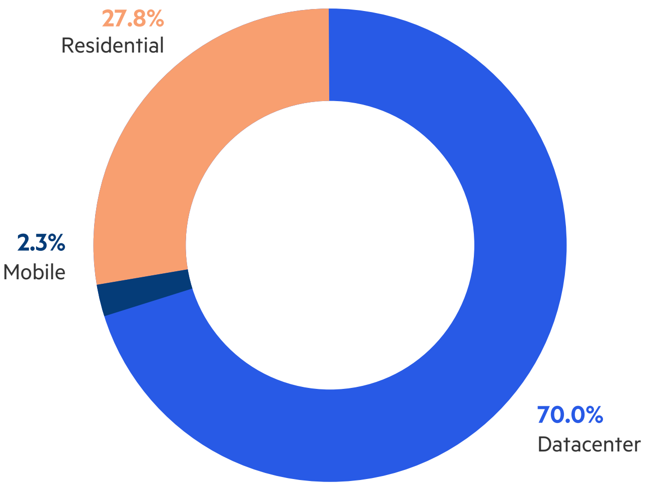
Perhaps some bad bots were written many years ago and remain on the prowl. Some may have targeted systems that only accept specific browser versions. Others may be out of control programs, bouncing around the internet in endless loops, still causing collateral damage.

Bad Bots Going Residential

Data centers are still the source of the majority of bad bots at 70 percent. But this number is less than last year's 73.6 percent. The global availability of low-cost cloud computing is what accounts for this dominance of data center use. However, bad bot traffic from residential ISPs increased for the third year in a row from 22.7 percent to 27.8 percent in 2019.

Bad bot traffic from mobile ISPs decreased to 2.3 percent this year from 3.6 percent in 2018. This indicates that bots only use mobile ISPs when the cheaper residential or data center options are not effective.

Bad Bot Traffic by ISP Type 2019



Amazon Bad Bot Market Share Drops

Bad bots were launched from 2,080 ISPs during 2019.

- While Amazon is the leading ISP for originating bad bot traffic, the proportion has dropped significantly to 11.6 percent in 2019 from 18.0 percent the previous year.
- DataWeb Global Group has moved from seventh in 2018 to second this year with 5.8 percent of bad bot traffic.
- OVH has increased its percentage to 3.7 percent and moved from fourth position last year to third in 2019.

Mobile ISPs: A Specialized Weapon

Data center traffic comprises the majority of bad bot traffic. But mobile ISPs also play an important role when bot operators find their data center traffic is blocked. Mobile ISP bad bot traffic is still a small percentage and usage overall dropped from the previous year.

TOP 10 BAD BOT ORIGINATING ISPS 2019

Rank	ISP	% of Traffic
1	Amazon.com	11.6%
2	DataWeb Global Group B.V.	5.8%
3	OVH Hosting	3.7%
4	China Telecom	2.4%
5	Cogent Communications	2.4%
6	Host1Plus	1.9%
7	Digital Ocean	1.8%
8	Apple	1.5%
9	Hetzner Online GmbH	1.3%
10	Google Cloud	1.2%

TOP 10 MOBILE ISPS

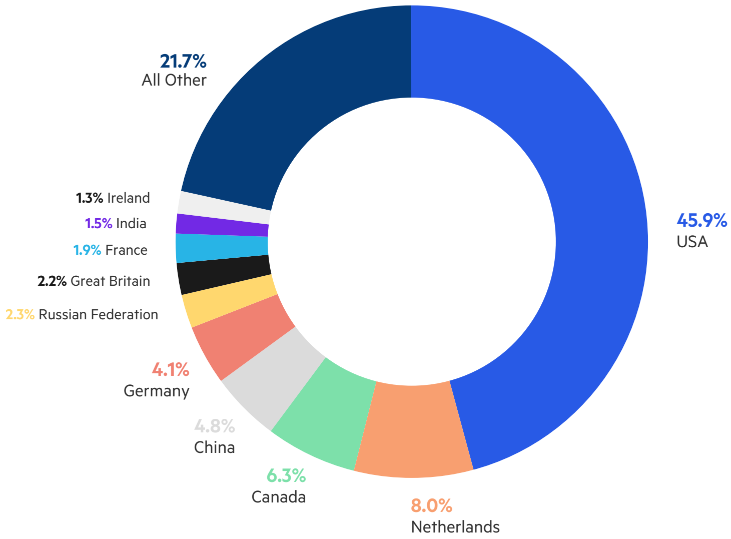
Rank	ISP	% of Traffic
1	Virgin Media	0.34%
2	Telefonica de Espana	0.32%
3	AT&T Wireless	0.30%
4	China Telecom Zhejiang	0.29%
5	Verizon Wireless	0.28%
6	T-Mobile USA	0.28%
7	China Telecom Jiangsu	0.27%
8	China Telecom Guangdong	0.20%
9	Orange Espana	0.18%
10	Vodafone Spain	0.12%

Where Bad Bots Originate

- For the sixth year running, the United States topped the list of bad bot originating countries. While it remains the only bad bot superpower from which 45.9 percent of all bad bot traffic originates, it has dropped since 2018 when it was the source of 53.4 percent of bad bot traffic.
- The Netherlands is in second place with 8.0 percent of all bad bot traffic—up from 5.7 percent the prior year.
- Canada has moved from fifth to third on the list responsible for 6.3 percent of bad bot traffic.
- China's volume of traffic has increased to 4.8 percent.
- India (1.5 percent) and Ireland (1.3 percent) join the top 10 countries where bad bots originate.

For the sixth year running, the United States topped the list of bad bot originating countries.

US Bad Bot v Rest of the World 2019



Russia and China: The Most Blocked Countries

Russia is the most blocked country for the third year running. China has moved from 4th place to 2nd as significantly more companies are blocking traffic from China compared to the previous year. Romania, Turkey, Vietnam, and Germany also appear in the top 10 most blocked countries this year.

TOP 14 MOST BLOCKED COUNTRIES

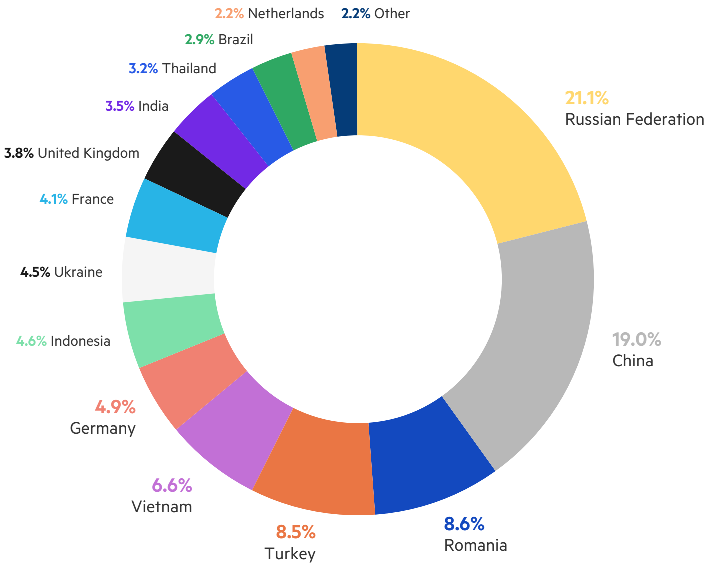
1	Russian Federation	8	Ukraine
2	China	9	France
3	Romania	10	United Kingdom
4	Turkey	11	India
5	Vietnam	12	Thailand
6	Germany	13	Brazil
7	Indonesia	14	Netherlands

Why Block Countries?

Many companies use geofencing blacklists to choke off large portions 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 who have suffered attacks from countries that haven't traditionally generated good traffic may block all traffic from that country as a sensible protection measure.

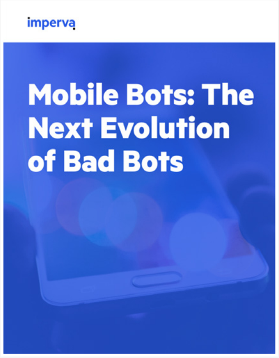
- Russia is still the most blocked country by 21.1 percent.
- China is now the second most blocked country, accounting for 19 percent of country-specific block requests, up from 11.2 percent in 2018.
- Romania is now the third most-blocked country at 8.6 percent. Turkey is a very close 4th place at 8.5 percent.

Most Blocked Countries 2019



Imperva Threat Research Lab

Threat Research


imperva

Mobile Bots: The Next Evolution of Bad Bots

MOBILE BOTS: THE NEXT EVOLUTION OF BAD BOTS

Key Finding

5.8% of mobile devices on cellular networks are used in bad bot attacks.

imperva

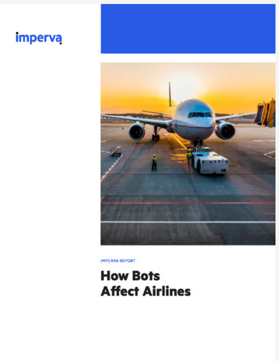
The Anatomy of Account Takeover Attacks

THE ANATOMY OF ACCOUNT TAKEOVER ATTACKS

Key Finding

Average of 2–3 account takeover attacks per month.

Industry Research


imperva

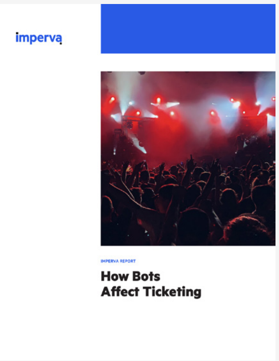
IMPERVA REPORT

How Bots Affect Airlines

HOW BOTS AFFECT AIRLINES

Key Finding

51 airlines with bad bot traffic higher than 50%.


imperva

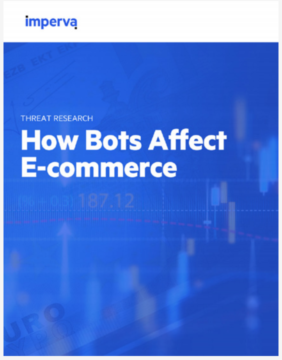
IMPERVA REPORT

How Bots Affect Ticketing

HOW BOTS AFFECT TICKETING

Key Finding

Bad bot traffic is 39.9% across 180 ticketing domains.


imperva

THREAT RESEARCH

How Bots Affect E-commerce

HOW BOTS AFFECT E-COMMERCE

Key Finding

17.7% of traffic on e-commerce sites is from bad bots.

Recommendations

Bots are on your website every day, and attack characteristics become more advanced and nuanced over time. How should businesses go about protecting themselves? Unfortunately, every site is targeted for different reasons, and usually by different methods, so there is no one-size-fits-all bot 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 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	< 52	< 60
Chrome version	< 57	< 65
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	GIGENET	OVH HOSTING	CHOOPA, LLC

3. BLOCK ALL ACCESS POINTS

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? They 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 PROTECTION 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 who use 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.

Industry analysts agree, which is why Gartner has added bot defense as a core requirement for WAF and CDN vendors. Your defenses need to evolve as fast as the threats, and to do that you need dedicated support from a team of experts.

About Imperva Application Security

Imperva Application Security mitigates risk for your business with full-function defense-in-depth, providing protection wherever you choose to deploy - in the cloud, on-premises, or via a hybrid model.

Imperva offers advanced analytics to quickly identify the threats that matter:

- Web Application Firewall (WAF) solutions which block the most critical web application security risks
- DDoS protection with a 3-second mitigation SLA
- API Security that integrates with leading API management vendors
- Advanced Bot Protection for defense against all OWASP automated threats
- Runtime Application Self-Protection (RASP) for security by default against known and zero-day vulnerabilities
- Developer-friendly Content Delivery Network (CDN) for the utmost performance.

Through FlexProtect, our unique licensing model, you can deploy Imperva Application Security how and when you need it. FlexProtect helps protect your applications wherever they live — in the cloud, on-premises or in a hybrid configuration.