Five Major Bot Threats to Holiday E-commerce and How To Stop Them

As revenue for online and mobile e-commerce companies continues to grow, it attracts *both* more shoppers *and* more bad actors. This is why the 2019 holiday shopping season is expected to set records not only for online sales but also for losses due to malicious bot activity.

The Dark Web is flooded with stolen user credentials that serve as fuel for launching new bot attacks. The result: malicious bots now make up more than 50 percent of the traffic to retail sites, while attractive "returns on their investments" have driven bot creators to pursue ever-increasing levels of sophistication. Recent generations of bots are capable of imitating human behavior and perpetrating clever, large-scale fraud schemes.

Taken together, user apathy about password security, easy access to stolen credentials, escalating bot activity and the increasing sophistication of attackers present an alarming scenario for e-commerce over the holiday shopping season.

Read on to learn more about:

- The expected impact of bot attacks against retail websites and mobile apps
- The five major types of attacks to watch out for during the 2019-2020 shopping season
- How to detect and block malicious bots *before* they wreak havoc on your all-important holiday peak periods and damage your brand reputation

50%
of traffic to retail sites overall comes from bots

ATO attacks have resulted in approximately
$10 billion
in losses over the past two years

Malicious bots responsible for
$6.5 billion loss

E-commerce Losses from Malicious Bots Already in the Billions... and Heading Higher

The size of the bad bot problem is already significant.

- Account takeover (ATO) attacks, where bots use stolen credentials to hijack customer accounts, have resulted in approximately $10 billion in losses over the past two years
- Malicious bots are largely responsible for the $6.5 billion to $19 billion in losses expected in 2019 from digital advertising fraud[1]
- More than 50 percent of the traffic to retail sites overall comes from bots and they represent between 40 percent and 80 percent of retail login attempts

The situation is likely to get even worse, as bots continue to proliferate. For several years now, bot-generated traffic has surpassed human-generated traffic on the Internet. These bots include not only bad bots but also the ones that are considered beneficial, such as search engine crawlers. The more alarming bots that make the largest portion of the automated traffic are the malicious bots, that will not only expand in number but also evolve in sophistication.

There are several reasons to expect that merchants will face the most sophisticated bot-based attacks yet during the 2019 holiday shopping season.

1. Online and mobile sales growth attracts attackers.

As online and mobile commerce expands, cybercriminals follow the money and scale up their activity.

Up from $2.86 trillion in 2018[2], retail e-commerce sales worldwide are expected to hit $3.53 trillion in 2019, as they head for an estimated $6.54 trillion in 2022[3]. Citing data from 451 Research's Global Unified Commerce Forecast, this excerpt from a Forbes article[4] is particularly relevant:

"Consumers are increasingly turning to online and mobile channels to make purchases that they traditionally would have made at the cash wrap in years past. This deflection of spend has been fueled, in part, by the rise of online marketplaces and the on-demand economy against the backdrop of new purchase experiences like click-and-collect and mobile order-ahead. *This year, one out of every ten dollars spent globally will occur in a digital channel. By 2022, more than 17% of B2C sales around the world will occur online.*"

> *The more alarming bots that make the largest portion of the automated traffic are the malicious bots, that will not only expand in number but also evolve in sophistication.*

The same article goes on to indicate that the number of mobile e-commerce transactions—those conducted with a smartphone or tablet—will exceed the number of traditional e-commerce transactions—those conducted using a desktop or laptop—globally, for the first time in 2019.

Based on past trends, it is also reasonable to expect a disproportionate percentage of these online transactions—and, therefore, associated bot attacks—to occur during the highly-active holiday shopping season.

2. More stolen credentials at cybercriminals' and bots' disposal.

As the volume of viable credentials increases, so too does the volume of bad bots looking to exploit them.

2018 continued an all-too-familiar pattern, as numerous high-profile data breaches were revealed, including Marriott Starwood hotels with 500 million records exposed, Google+ with 52.5 million records exposed, and Panera with 37 million records exposed. Then there's the first half of 2019, with over 3800 breaches exposing a mind-boggling 4.1 billion records[5]. Overall, it is estimated that there are more than 1.4 billion stolen email/password combinations available on the dark web for purchase.

The result is a truly worrisome situation where the login credentials of most consumers have already been stolen and resold, and due to poor password practices, they are likely to work on multiple websites and accounts. Moreover, it's pretty clear that cybercriminals have already geared up to take advantage of the situation. The evidence: according to the 2019 Verizon Data Breach Investigations Report, 29 percent of reported breaches involved the use of stolen credentials, with web apps being the most popular vector of attack.

3. Retailers' outmoded defenses cannot stop increasingly sophisticated bots.

As one line of defense is put in place, cybercriminals will invariably advance bot technology to overcome it.

Cybercriminal syndicates, which increasingly deploy the most sophisticated bots available, are engaged in an arms race with retailers' IT security teams. Today's advanced bots impersonate real users and legitimate system behaviors either by

injecting a malicious extension into the user's browser or by simply executing the browser in a hidden window. The bots, or their operators, then piggyback other attacks on the valid identities and systems of real users.

Because these bots often have valid user credentials, do not make many requests from any single IP address, and mimic key aspects of human behavior, they won't trigger volumetric or IP reputation alarms. Typically, they're also able to evade commonly deployed signature-based defenses and web application firewalls.

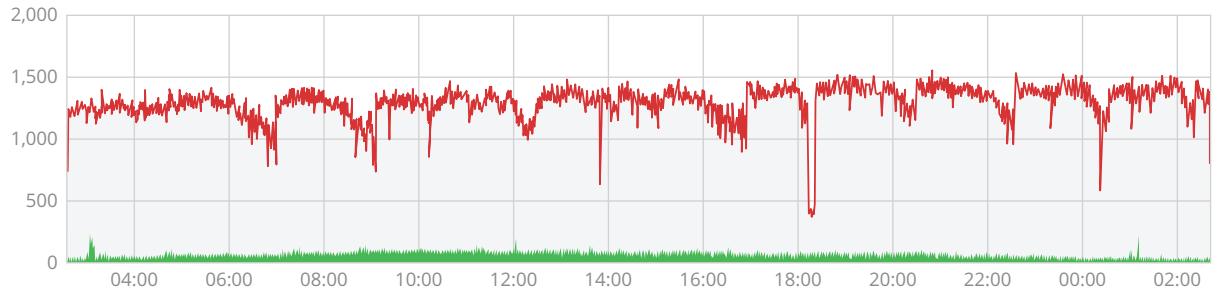
Five Types of Bot Attacks Against E-commerce Sites and Mobile Apps

The already costly bad bot situation is likely to get worse as the 2019 holiday shopping season unfolds. Gaining an understanding of the best way to prevent this from happening requires delving a bit deeper into the mechanics of malicious bot activity. Accordingly, this section reveals the top five types of bot attacks currently plaguing online retailers.

Account Takeover: A Bot Takes Over a Real User's Account

E-commerce criminals frequently use ATO attacks as the first step in a wave of attacks and fraud utilizing the login credentials of unsuspecting customers. This type of attack costs retailers billions of dollars of losses yearly.

The graph below shows a recent ATO attack that peaked at ~1500 attempts per second. When looking at the number of login attempts during an ATO attack, comparing the bot traffic in red to the legitimate one in green, it can clearly be seen that during the attack, over 90 percent of the login attempts were malicious. This type of attack typically originates from thousands of IP addresses and has an average success rate of 8 percent, which begs the question: how is that possible?



In ATO attacks, automated bots compromise thousands of accounts by exploiting curated lists of username-password login credentials. Cybercriminals generally mask their attacks by routing credential stuffing attacks through network proxies, or by rotating IP addresses to attempt many username-password combinations. This isn't as futile as it may appear. The average American user relies on just six passwords across 130+ accounts[6], so hackers can often reuse compromised credentials on other e-commerce sites where the user has shopped. Once attackers hit upon a valid combination, they can place false orders and commit other illegal acts.

Carding: Bots Use Stolen Credit Cards to Ring Up Charges

For markets where there is a high percentage of credit card usage, like the United States, carding—a form of credit card fraud—remains a significant issue. Here's how it works.

Carders use bots to test lists of recently stolen credit or debit card information, obtained from other hackers or the Dark Web, against merchant sites. The carders then use the proven credit card data to directly retrieve funds from associated accounts or to purchase gift cards which can easily be converted into high-value goods, such as cell phones, televisions, and computers. These goods are then resold—often via websites offering a degree of anonymity—for a nifty profit.

It is interesting to note that many carding attacks are very similar to ATO attacks in terms of how they work. The big difference is that while ATO attacks focus on defeating the login page and process for a merchant's site using lists of stolen usernames and passwords, carding targets the checkout page and process with stolen cards details.

In a variation of carding, attackers go after gift cards directly, instead of credit cards. Bots again test the cards, in this case by trying and finding PINs that work for stolen or fraudulently created gift card accounts. The end result is the same. Merchants lose real money when a stolen or fake card is redeemed online or at a store because the retailer must reimburse the customer. And customer satisfaction invariably suffers, too. No customer likes to learn that their account has been compromised and their funds have been stolen. Using external card verification services for every transaction can solve the problem. But, the verification costs can be prohibitive for retailers and the increased processing time can increase frustration for legitimate users.

Checkout Abuse: Bots Scalp or Hoard Products that are in Tight Supply

With a *scalping* attack, bots are used to rapidly purchase all the inventory of a highdemand product. The scalpers know the products can be resold for a quick profit on the secondary market. Prime examples are hot toys and limited release apparel and footwear items.

Five Types of Bot Attacks Against E-commerce Sites and Mobile Apps

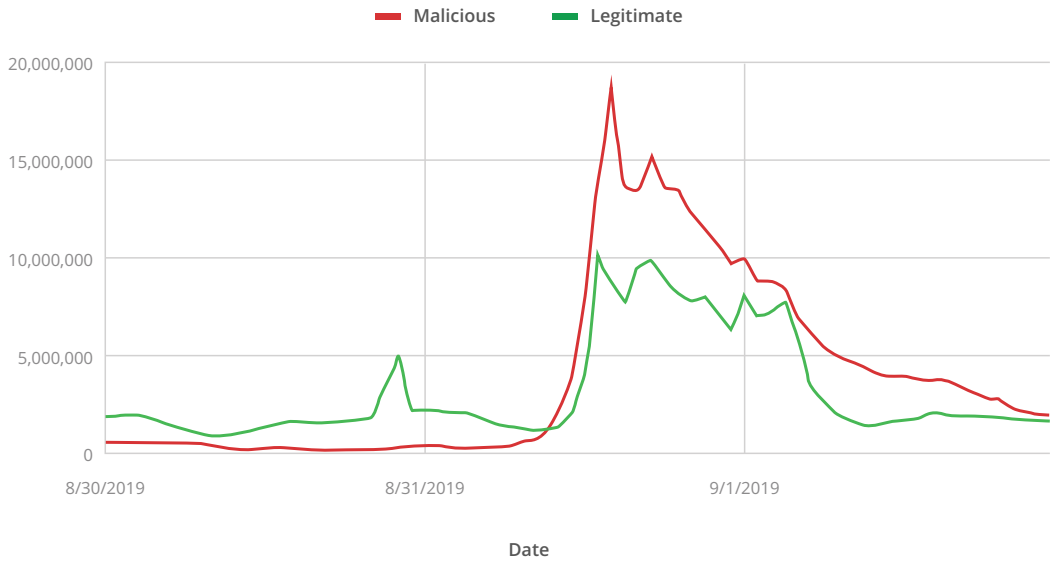
1. Account Takeover
2. Carding
3. Checkout Abuse
4. Web Scraping
5. Mobile Apps

These types of attacks are typically more sporadic, as they're associated with the release dates of new products. Just prior to release, bots prepare to pounce by constantly checking the product URLs to see if they have gone live. As much as 90 percent of website traffic may be generated by bots waiting for the sale to begin. Once the item goes on sale, the bots will continue to buy it until it's deleted, or until they purchase the quantity they want.

For certain high-demand products, like the coveted limited-edition of shoes, there is a whole industry dedicated to making sure that the potential buyers only have one option to get the shoes: pay higher than the suggested retail price on resale websites. The of-the-shelf sneaker bots are programmed to automatically purchase the latest shoes from an online retailer before anyone has a chance to click "add to cart". These bots are offered by professional organizations with 24-hour customer service and online help resources. Once the sneaker bot users get the shoe, they will resell the shoe on a third-party website for a price that is significantly higher than the original one.

This kind of bot activity burdens a retailer's infrastructure—often to the point of crashing massive sites. More importantly, this activity prevents the retailer from selling to its real, human customers. The retailer's brand, credibility, and reputation are damaged.

The graph below shows a scalping attack for limited-edition sneakers. In this case, a peak bot-request rate of nearly two times the rate for legitimate users kept many customers from getting the new footwear they wanted.



When attackers use automated fake shoppers to load up shopping carts with items that they never intend to purchase, this practice—known as *hoarding*—penalizes the retailer severely. This is also known as an application-layer distributed denial of service (DDOS) attack. Hoarder bots may put the same item back into a shopping cart thousands of times over the course of a few days, preventing human shoppers from buying the products. It also makes it very hard for merchants to understand what they have in stock, resulting in potential losses when they order more stock based on apparently depleted inventory.

Attackers may offer the hoarded product for sale on third-party sites, knowing they have all units of a retailer's inventory completely locked up in shopping carts. Shoppers cannot buy the product from the retailer advertising it for sale, so they may pay more to get it from a resale site. If the attackers can arbitrage a higher price selling to a frustrated consumer, they then make the purchase at the retailer's price from the shopping carts it controls. They first take the consumer's money, then buy from the original retailer and have the retailer ship directly to the attackers' customer, perhaps taking advantage of a retailer's free shipping. The merchant tends to book very few actual sales, losing significant revenue and is confused by seeing "sold out" notifications on the hoarded inventory.

Web Scraping: Bots Spy on Your Prices and Steal Your Content

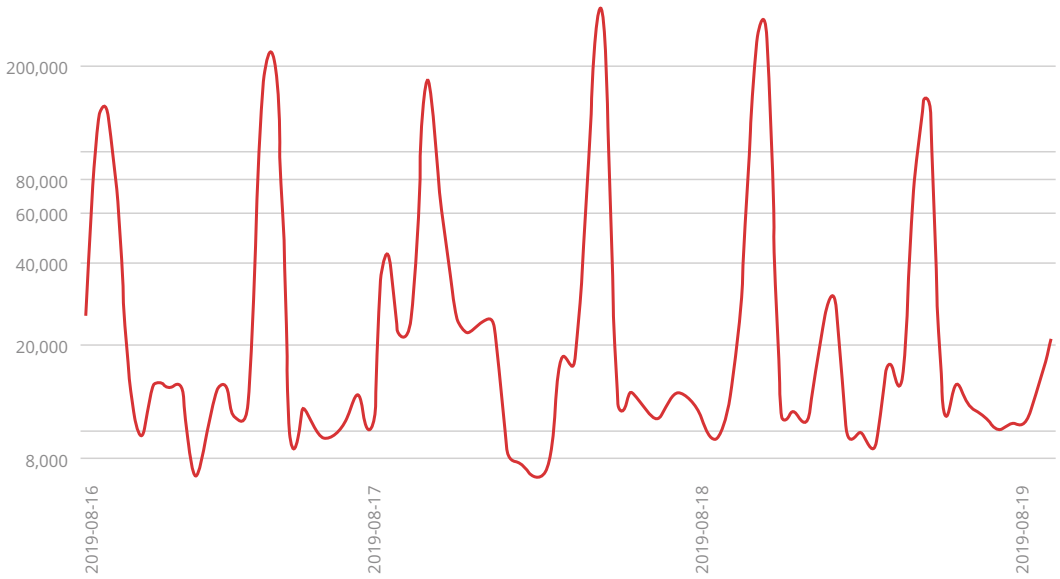
Price scraping by bots enables competitors, or their hired intelligence-gathering scrapers, to spy on your pricing tactics and keep up with your pricing moves. While not technically illegal, this practice harms retailers by giving their competitors easy access to their pricing strategies, category management, inventory levels, and even SEO and keyword tactics, which could severely hurt their competitive edge.

To stop price scraping, many companies hide prices for hot products on their sites. The price appears only after the item is selected for purchase. In some cases, the final price is only arrived at once the shopper or bot fills out a user profile. However, attackers have responded with bots that mimic real shoppers by putting the products into shopping carts. At that point, the price becomes visible and the bot scrapes it.

Up to 20 percent of traffic to carts at most retailers—every day, all day—is from price scraping bots that enter fake user information to get at the most accurate final prices. The extra traffic also skews retailers' analytics, as these scraping bots never buy anything. Conversion rates become distorted as bot activity makes them artificially low.

A real bonus for scrapers is getting a low-inventory message such as "Only two left in stock." When competitors learn you have limited inventory, which can be the signal to raise prices. During the holidays, this intel can be a meaningful competitive advantage.

The graph below shows a typical scraping pattern—in this case with over 31,000 different IP addresses and over a thousand different devices and browsers targeting more than 117,000 different paths on the website—which is very atypical for real users and a clear indication of malicious activity.



Content Scraping, in comparison, is where web scrapers are used to steal product descriptions, reviews, and inventory data. The theft and republishing of product reviews make a competitor look more established and reputable, and degrades the value of exclusive, copyrighted content on your website. Product reviews help competitors gauge product popularity.

Unlike price scraping, content scraping is illegal if another site publishes your content. As any retailer knows, good content is expensive to generate. Retailers pay writers to generate copy that describes products effectively and uses optimized keywords. With the help of web scraping bots, criminals steal this investment in content and benefit from using it on their own sites. This can also affect the SEO of the victim's website when search engines detect pages with duplicate content. When malicious sites that republish the stolen content rank highly, not only does it naturally hurt the ranking of the retailer's site that originally posted the content, but also the search engines may penalize the site, mistakenly assuming that it stole the descriptions, and hurt its ranking even further.

Mobile Apps: Another Major Front in The Bot Wars

According to Forbes, 2019 will mark the first time the number of m-commerce transactions—those conducted from a smartphone or tablet—will exceed the number of traditional e-commerce transactions—those conducted from a desktop or laptop—globally[4].

Ever-alert to growth opportunities, cybercriminals have followed shoppers from the desktop to the mobile device. They now employ bots to attack the mobile APIs that retailers use to power their native mobile applications. In doing so, bots are following consumers who spend more time shopping via native smartphone apps and less time on web browsers.

Moving onto mobile creates an entirely new set of problems for online retailers because people behave differently on their mobile devices. Traditional techniques used to determine whether the user is a human or a bot—such as IP addresses affiliated with home broadband accounts—no longer apply.

Mobile Attacks: Criminals Use Three Techniques

There are three primary techniques for mobile app attacks, all of which attempt to impersonate application behavior in some fashion.

- Attackers can call an application's APIs directly from any IP connection—without having to use the actual app or even a mobile device.
- Attackers can use the genuine application or a hacked version, running on thousands of instances of a mobile device emulator.
- Attackers can hack a device or more likely, an application on a device and then take over the application to launch the attack.

PerimeterX has tested leading mobile applications and found that, from a technical standpoint, they do very little to protect themselves[7]. Criminals have also figured this out. An all-too-common scenario in our experience is that, as merchants take steps to better defend their websites, threat actors simply shift their attention and efforts to their mobile assets. In addition, mobile attacks are, in general, more distributed—utilizing fewer requests per IP address.

How To Stop Malicious Bots

As bad bots become more automated and sophisticated, it becomes imperative to outsmart them before they gain access. The majority of bot-detection solutions—from static signature-based to more advanced behavior-based options—can only detect and mitigate after the user or bot begins to interact with the web page. Predicting prior to web page access prevents harmful bots from gaining any access.

To be clear, behavior-based methods are required to accurately distinguish bad versus good users, by tracking the user's interaction with the application, collecting device fingerprinting info and using additional client-side indicators. Such methods can use hundreds of different data signals—everything from network and device-level bread crumbs such as IP address, device make or version and user agent, to the typical paths used to navigate a site or application, to the details of how users, both in general and individually do different things including mouse movements and clicks, keyboard patterns and the rate of response between pages.

Indeed, more personal and granular behavioral traits, such as the way someone types, known as keystroke dynamics, and uses a mouse, and benchmarking these activities against a baseline, can be used to quickly distinguish human from a bot. For example, humans might hover over different elements on a page, until they reach a box where they can type. The actual typing, in terms of keystrokes, key down and key up events, and the timing of each is unique to each person. The fact is, human interactions are very distinct from the behavior of automated attacks. This isn't detected by looking at the behavior or path of URLs accessed, but more specifically at all aspects of what the user or bot is doing with the browser and application.

The challenge with relying on behavioral detection methods alone, however, is that in order to collect sufficient behavioral data, the user first has to interact with the application and access a page or data. But, if the user is a malicious bot, then the battle is already lost.

To address this challenge, it's essential to take the behavioral analysis a step further—into predictive detection. With this technique, detailed fingerprints are extracted in real-time to accurately predict which requests are issued by bad bots and attack tools, and which are legitimate users.

Machine learning enables PerimeterX Bot Defender to establish a nominal or common range for human interactions and allows the platform to create and auto-tune prediction rules. PerimeterX continually provides a large volume of real-time sensor data to a massively scaled detection cloud that predicts bots before they gain access. This not only yields unmatched accuracy in predicting and detecting the diverse flood of automated threats in today's environment but also enables PerimeterX Bot Defender to find ever more nuanced signs of attacks, even as they continue to evolve in sophistication.

Sophisticated Bot Attacks Require Investment in Sophisticated Defenses Before the Holiday Shopping Season Arrives

It is almost a certainty that every retailer will experience multiple, simultaneous attack types throughout the calendar year. Heading into the holiday shopping season, however, we expect professional criminal groups to attack on a larger scale, using increasingly advanced strategies and bots that take advantage of curated databases of stolen credentials.

Behavior analysis is the basis of newer security technology that can detect even the most advanced, stealthy bots and attacks. In particular, behavior analysis with predictive detection is what retailers should consider adopting. Predictive capabilities are an essential component needed to observe each user's activity and predict—before the user has access to web pages, account information, pricing, or gift cards—whether it's a legitimate human visitor or a malicious bot.

These defensive capabilities are needed year-round, but particularly during the allimportant holiday shopping season. Since this technology can be deployed on major retail websites very quickly, retailers are encouraged to investigate PerimeterX solutions before the peak shopping days arrive.

[1] www.emarketer.com/content/digital-ad-fraud-2019
[2] www.digitalcommerce360.com/article/global-ecommerce-sales/
[3] www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales/
[4] www.forbes.com/sites/jordanmckee/2018/09/11/global-digital-commerce-sales-to-near-6-trillion-by-2022
[5] www.forbes.com/sites/daveywinder/2019/08/20/data-breaches-expose-41-billion-records-in-first-six-monthsof-2019/#25ada33fbd54
[6] blog.dashlane.com/infographic-online-overload-its-worse-than-you-thought/
[7] www.perimeterx.com/blog/account-takeover-mobile-apps/

About PerimeterX

PerimeterX is the leading provider of application security solutions that keep your business safe in the complex digital world. Delivered as a service, the company's Bot Defender, Code Defender and Page Defender solutions detect risks to your web applications and proactively manage them, freeing you to focus on growth and innovation. Bringing together an elite engineering team, security research to continually update its solutions with current intelligence, and best-in-class customer enablement and support, the world's largest and most reputable websites and mobile applications count on PerimeterX to safeguard their consumers' digital experience.

perimeterx