Fare Scraping and Excess Transactions: The Real Cost to the Travel Industry

NETACEA

INTRODUCTION TO LOOK-TO-BOOK RATIOS AND FARE SCRAPING

The look-to-book ratio is the number of requests made per booking on an online travel site. Requests can be made by humans or bots, and the lower the look-to-book ratio, the better for the company, meaning conversions are high from genuine customers browsing the website. However, in some cases look-to-book ratios can exceed several thousands to one due to scraper bot activity.

In travel, web scraper bots are mainly used to collect fare and availability information by rival companies and aggregator sites are used for price comparison. Travel sites are frequently affected by aggregation services that use scraper bots to discover and publicize the availability of products or services such as flights, hotels or car rentals.

Threat actors advertise the scraped information at lower price points on a secondary site, motivated by the financial rewards of charging commissions, stealing personal data or generating advertising revenue. Due to the dynamic nature of travel pricing, this is fast becoming a top threat for the industry, exacerbated by increased competition driven by the pandemic.

In a 2021 survey by Netacea, 96% of travel companies said their website had been attacked by bots over the previous 12 months.

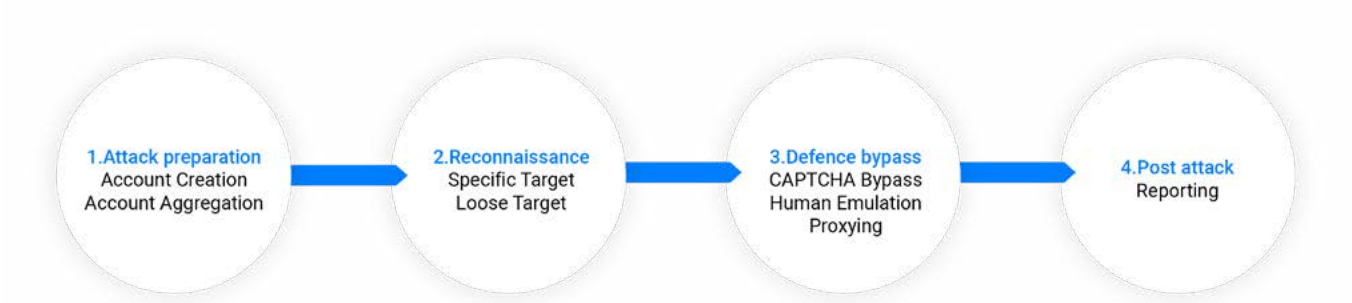
Netacea's Threat Research team has observed travel sites with 90% scraper bot traffic, and whilst this activity can be benign or even good, if uncontrolled it can impact top line revenue, bottom line profits and customer experience.

How do excess web requests lead to high look-to-book ratios?

Look-to-book ratios are regularly inflated by scraper bot traffic. Scraper bots can make excess web requests to an online travel agent or travel booking site which, in turn, negatively impacts your look-to-book ratio. This can be used to your competitors' advantage and is often used to gather the data needed for more sophisticated or damaging attacks such as spinning or denial of inventory. Preventing malicious scraper bots can cut out these attacks early as the attackers do not have the data they need to progress, leveraging you above your competitors by keeping business and technical costs down.

EXAMINING SCRAPER BOTS THROUGH THE BLADE FRAMEWORK

The BLADE Framework captures the various stages of a scraper bot attack, from attack preparation and reconnaissance to defense bypass and post-attack action.[1] Using this framework, businesses can isolate the individual attack stages used by the adversary to scrape travel websites, allowing for improved risk assessment, strengthened threat detection and mitigation capabilities, and a better-informed incident response process. To be able to effectively mitigate scraper bot attacks, businesses must be aware of the different stages that make up the attack process.

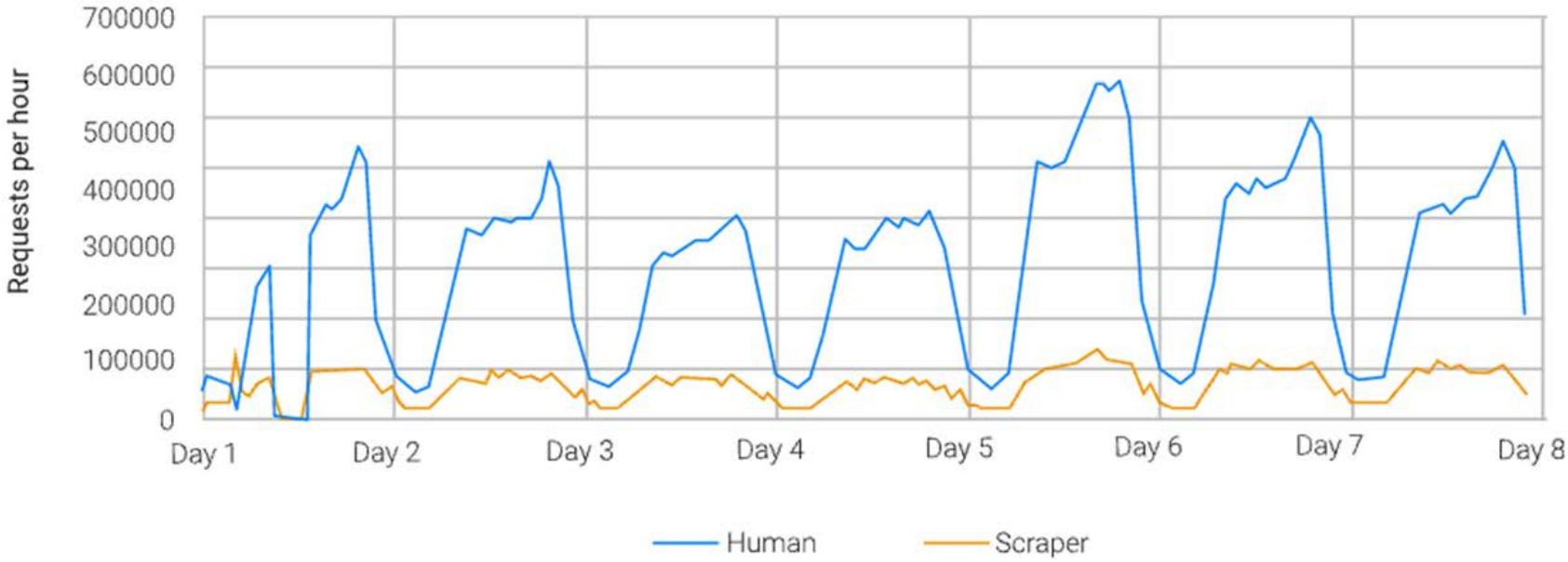


Scraper bot attacks usually begin with the attacker manually or automatically creating many user accounts using a real person's details, and aggregating information from multiple different user accounts. The adversary either specifies a target organization or identifies potential targets from a website. On travel websites, however, not all scraping requires a customer account; a simple scraper can gather information from multiple sites without logins to find the cheapest price for a flight, for example. The next stage is obfuscating the origin of the bot activity, seeking to bypass IP and geo-blocking by using a server application or appliance that acts as an intermediary for requests from clients to the web app. Bots imitate human behavior in order to circumvent behavior analytics-based defense measures such as CAPTCHA. Information gathered during the attack is then provided to the adversary and/or another interested party. This information may be used to inform future attacks or decision-making.

Sources:
1. BLADE Framework

Fig. 1 shows the time series of human and scraper requests on a website over a week period. On travel websites, scraper bots mimic human behaviour to appear as genuine users at peak times throughout the day, sometimes switching IP addresses to give the impression that the traffic originates from a human.

Fig. 1



WHAT DAMAGE ARE HIGH LOOK-TO-BOOK RATIOS CAUSING TO TRAVEL COMPANIES?

Excess traffic caused by aggressive scraping negatively impacts airlines and travel companies both on their bottom line and on their technical performance. Price scraping on travel websites has the potential to not only damage website sales, but also user experience, marketing analytics and brand reputation.

Business costs

- Additional costs (up to millions per year) to third-party services, charged based on traffic volumes such as Metasearch engines and PSS/GDS excess transactions.
- Extra costs for SIEM and anti-fraud solutions
- Loss of pricing visibility leading to disadvantage in competitive pricing
- Loss of control of customer journey
- Reduced conversions and misleading analytics from inaccurate number of website viewers interested in a certain product or booking, used as a basis for making business decisions
- Loss of ancillary revenues such as hotels, insurance, car hire, often more profitable than the original reservation

Technical costs

- Excessive infrastructure costs (up to 50%) used to serve bots which add nothing to your profitability
- IT teams stretched to deal with bots away from daily tasks
- Slowed website performance leading to negative effect on user experience
- Costly downtime in extreme cases
- Attackers gathering data used later by more sophisticated bots in spinner or denial of inventory attacks

What wider impact is this having on the travel industry?

Excess transactions leading to a high look-to-book ratio can be costly for airlines and online travel companies; the cost of pricing systems being queried can be both substantial and uncontrollable.

The online travel agency (OTA) and Metasearch engine landscape has changed substantially over the last decade. Whilst the Covid-19 pandemic froze bookings on OTAs, both the number of OTAs and subsequent traffic to these OTAs has grown unlike anything else in the industry.

Particularly in Europe, which typically has one OTA per country, each OTA is trying to beat the competition based on price and getting to the pole position of Metasearch engines.

However, many airlines – that is traditional airlines rather than emerging, low-cost airlines, which are extremely online focused – lack a strategy for how to work with OTAs in an efficient way. It is important for airlines to understand who is undercutting fares and how to control their distribution channels, to gain a holistic view of their entire distribution and a plan for both direct sales and sales on OTAs and Metasearch engines.

Many airlines talk about how the shift towards continuous or dynamic pricing – allowing airlines to provide infinite price points and adapt to supply and demand on a more granular level – is going to change the entire landscape further. Whilst it will be more difficult for scrapers to consistently check fares, attack methodology will also adapt to find ways around new technologies. Adversary groups will either adapt their attack pattern or move to a new attack entirely. As AI benefits the travel industry's pricing policies, it also benefits attackers who can use it to find ways around new defense measures.

HOW TO PREVENT SCRAPING, EXCESS TRANSACTIONS AND HIGH LOOK-TO-BOOK RATIOS

The travel industry is one of the most severely affected by bad bots and has been since the advent of online travel. As bots grow in sophistication and volume, it is crucial for travel websites to accurately detect and control bad bots without affecting good bots necessary for the steady running of your website, and genuine users' experience.

Choosing the right bot management solution is a significant decision for any business. Choosing an agentless solution offers the following benefits:

Fast and accurate protection of all endpoints, without the need for JavaScript and SDKs

Low-maintenance solution with no need for frequent updates

Invisible and secure layer of bot detection which ensures protection of your customers' privacy

GET STARTED WITH NETACEA BOT MANAGEMENT

Netacea Bot Management takes a consultative, agentless approach, paired with our server-side implementation and innovative Intent Analytics® technology to seamlessly integrate with your business and deliver accurate, intelligent and effective bot mitigation.

Try Netacea Bot Management for your business

Get in touch to arrange your free trial of Netacea Bot Management at hello@netacea.com.