The Return of the English Premier League:
Arb Betting in Action

Introduction to betting fraud

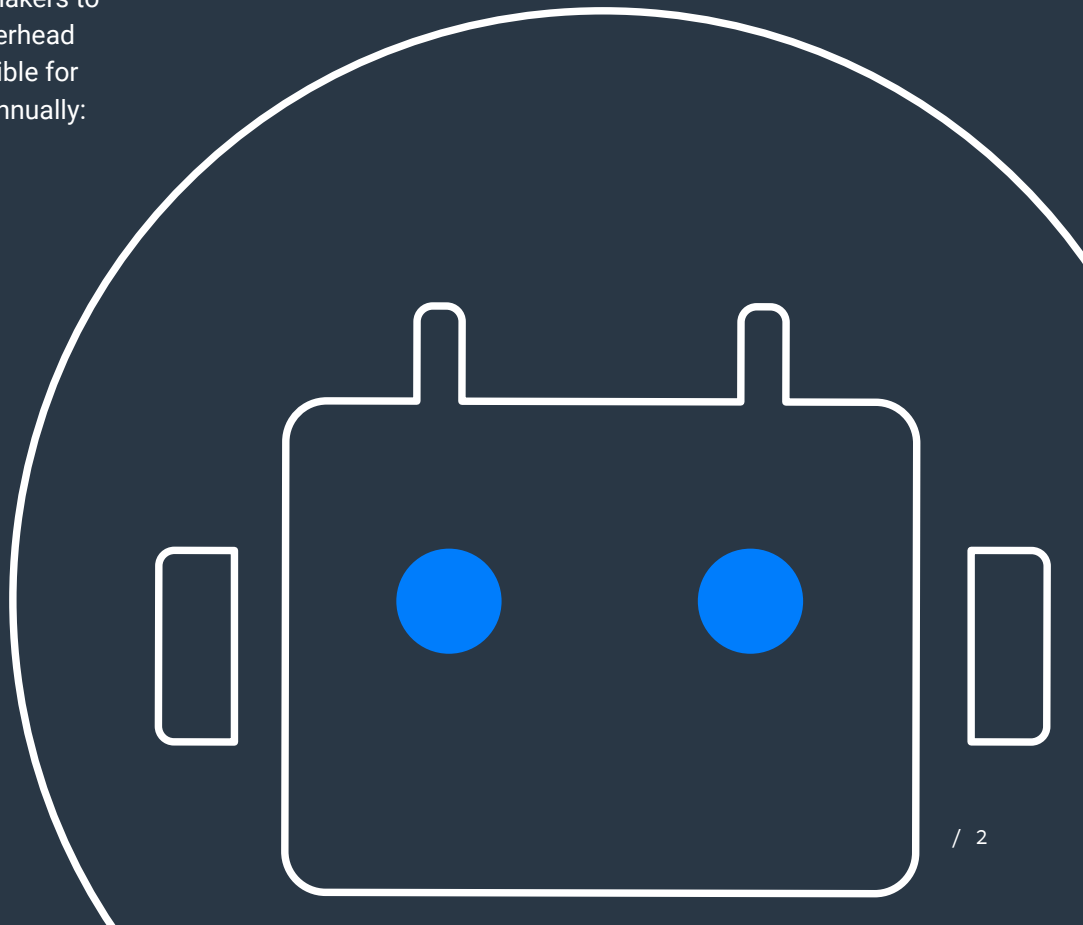
Globally, the sports gambling industry is estimated to be worth over $200billion. With more and more states in the USA passing legislation for sports betting, this figure is likely to grow massively in coming years.

Where there is industry growth, fraudsters are sure to follow. Unfortunately, there are several routes to profit for bad actors targeting bookmakers and their customers.

As customers of both sports betting sites and online casinos are likely to have financial assets stored in their accounts, either in credits or winnings, such sites are prime targets for phishing, credential cracking and account takeover attacks.

Another obvious attack vector is bonus abuse, where attackers will automate the creation of new accounts to take advantage of welcome offers on a large scale.

Perhaps the hardest attack for bookmakers to combat is one that creates a huge overhead in infrastructure costs and is responsible for millions in lost operational expense annually: Arbitrage betting scraper bots.



What is arbitrage betting?

Arbitrage betting, or arb betting, is a tactic used to guarantee a profit when making a series of opposing bets on the same event. It works by exploiting imbalances in the odds across multiple bookmakers, placing bets on opposing outcomes to guarantee that the bets never lose money, and potentially win. Arb betting is not illegal, however it has become an industry-wide threat to gambling businesses.

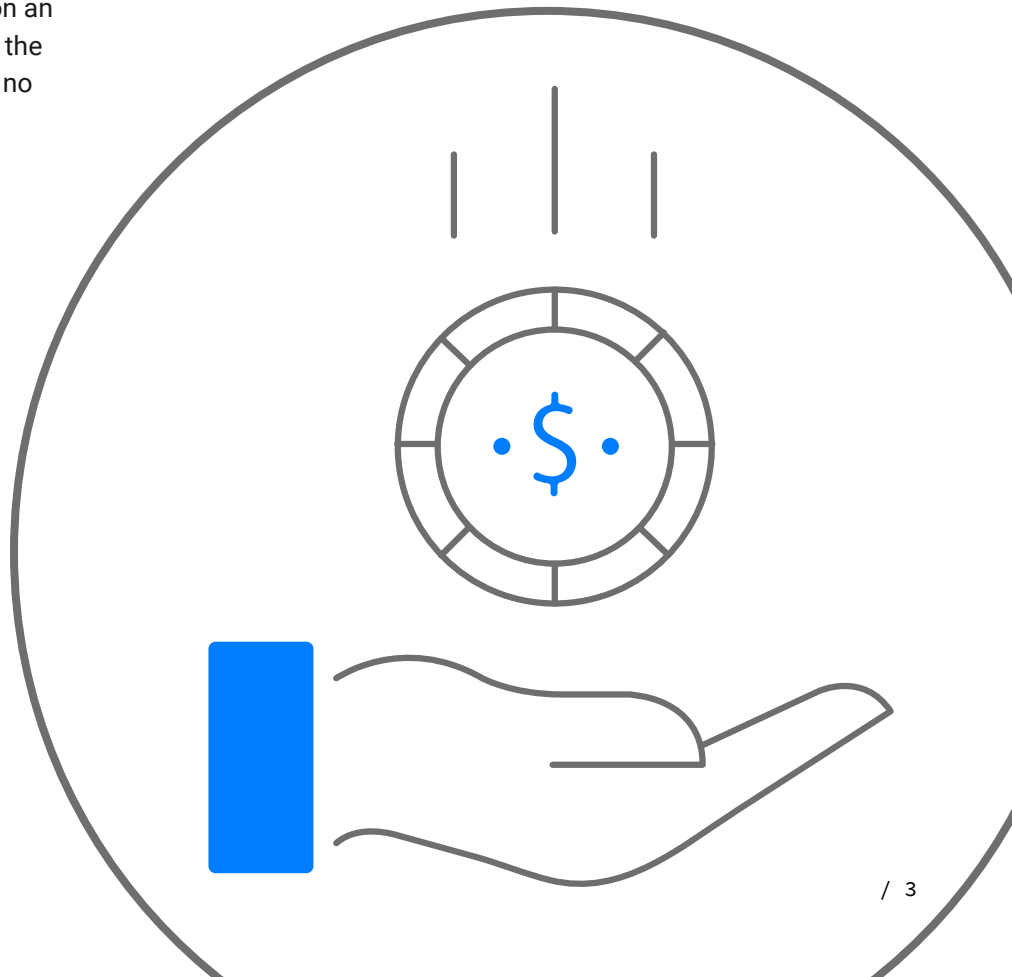
First, the arb bettor places a bet on an event with a bookmaker or betting site, before laying that bet on the same event using a betting exchange – betting on the opposite outcome. Essentially the arb bettor is selling the bet for more than they bought it for, so that no matter the outcome of the bet, they will turn a profit.

For example, the arb bettor will place a bet on a football team to win a game, whilst betting on an exchange that the opposing team will win. If the odds are imbalanced, they will make a profit no matter who wins the game.

Using scraper bots to achieve arb betting

Web scraping is a common automated technique across industries; common good or benign reasons for web scraping include search engine indexing, content aggregation and pricing intelligence. However, web scraping can have a damaging impact on a business if used with bad intent, for example to facilitate arb betting in the betting industry.

Arb bets exploit odds imbalances, which are often hard to come by and time-consuming to find. Arb bettors can improve their efficiency by using web scraper bots to automatically scrape the odds data they need from various betting websites, before simultaneously placing bets on all possible outcomes of an event at odds that guarantee a profit. This automated process means adversaries can generate arb bets quickly and efficiently.



Understanding arb betting through the BLADE™ framework

Bot attacks are often made up of several distinct actions or stages (for more information see the BLADE™ – or Business Logic Attack Definition – framework), and arb betting is no different.

We have used the BLADE framework to identify the arbitrage betting kill chain, which describes the potential stages this specific attack type follows in sequence (see Fig.1). This gives us a clear view of how the attack is made possible and what we might do to disrupt it at the most effective point in the chain.

Arb betting bots are, at their core, facilitated by scraper bots. **Resource development** is typically the first step of most business logic attacks. The infrastructure to run the bots must be acquired, especially because arb scraping is a resource-intensive activity.

Next is **reconnaissance**, where the target betting sites are selected and the relevant URLs to scrape content from identified. Arb bots must "scrape" (or extract) content from multiple bookmaker websites and betting exchanges, specifically the up-to-the-minute odds for available bets, so the structure of these pages and the information on them must be configured in the **attack preparation** stage. With **defense bypass** measures like IP rotation, human emulation and proxying in place, the **attack execution** stage can commence. As odds can change rapidly, arb betting is a time-sensitive operation. As a result, these scraper bots attack their targets aggressively to ensure the information is always accurate.

The scraped odds are then fed into an algorithm to determine where arb bets can be made for guaranteed profit, which is presented to users of specially built arb betting tools in the **post-attack** stage.

Some arb betting bots are even capable of automating the entire process of placing and laying the bets, using similar functionality for this stage as a retail or ticket scalper bot.

The cost of serving requests for content to these aggressive scrapers can run into millions of dollars annually. At peak times, scrapers make up most traffic to betting sites, often slowing down the site's performance for legitimate customers. This can cause conversions and pageviews to drop, damaging revenue even further.

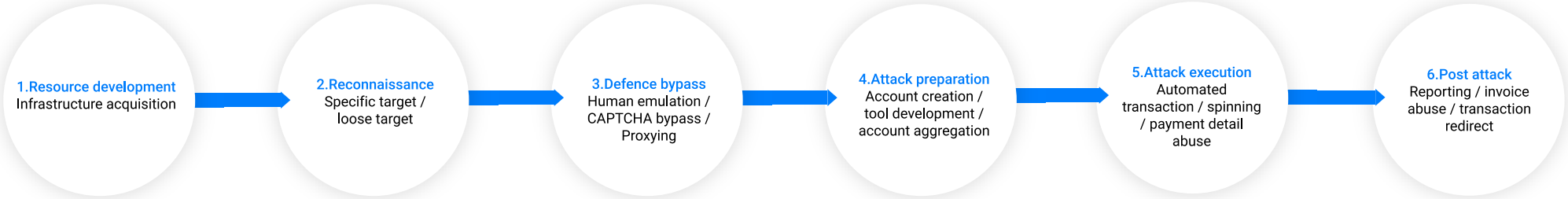


Fig. 1 – BLADE Arb Betting Kill Chain

Bookmaker faces high levels of bot traffic during Premier League opening weekend

While popular sporting events were put on hold over the course of the Covid-19 pandemic, the return of the world-famous English Premier League with full stadiums and match schedules in August 2021 was a milestone in the recommencing of the world's largest stadium events post pandemic.

On the opening weekend of the 2021/22 season, with demand for bookmakers higher than ever, Netacea investigated arb betting activity by monitoring odds scrapers on a popular bookmaker's website.

In these charts we can see clear spikes in web activity corresponding with most of the games themselves (Friday evening and Saturday afternoon).

Fig. 2 depicts the most targeted path on the bookmaker's website during the 48-hour period over which eight Premier League football matches were played across England. The most targeted path searched for 'football' rather than specific matches, suggesting arb betting bots at work looking not for specific football matches but instead looking to scrape information to place bets across the board. This is consistent with the reconnaissance phase of the attack where the bot operator will target specific URLs to gain the information they require in the most efficient manner.

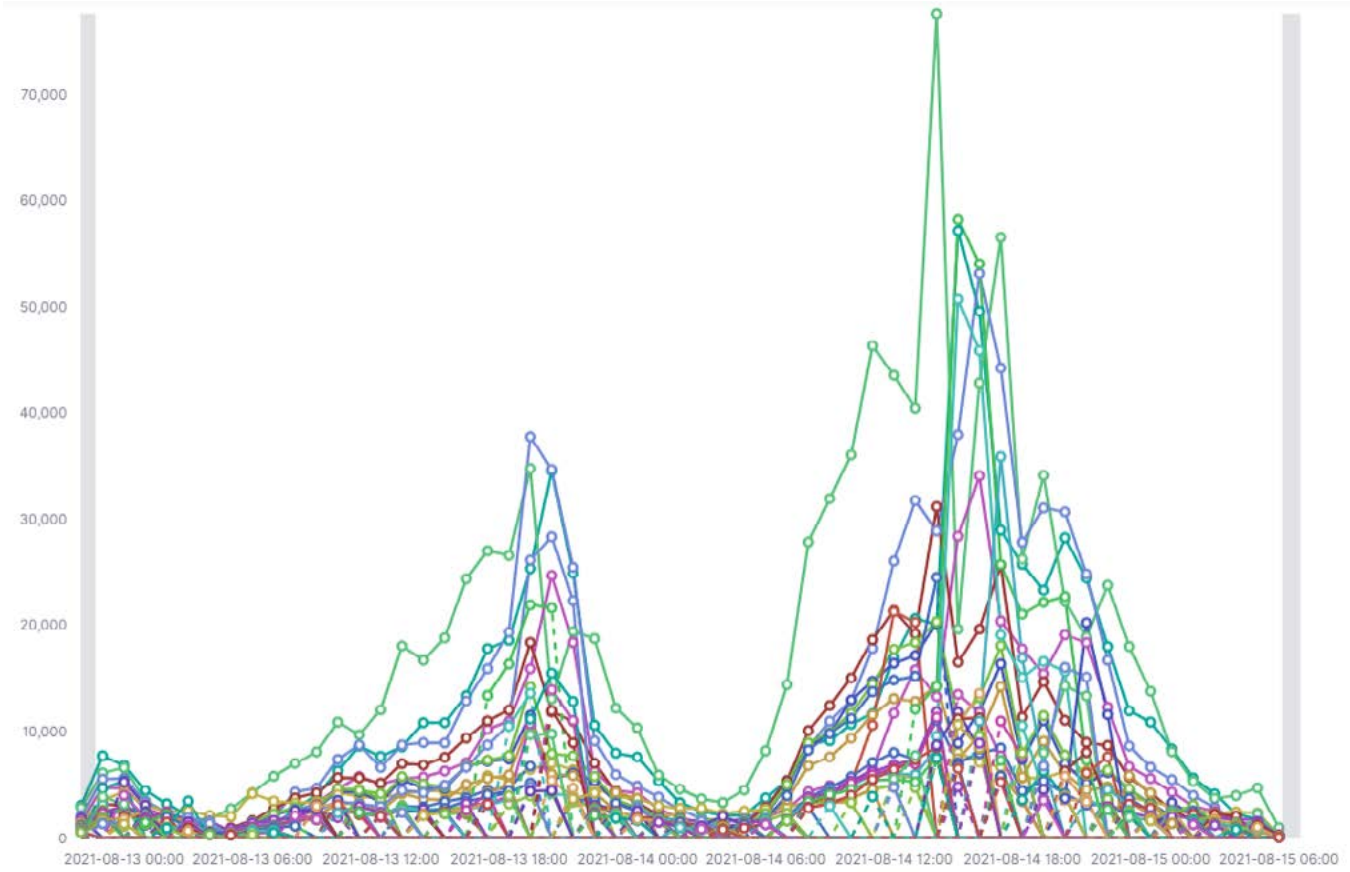


Fig. 2 – Traffic on bookmaker's 'football' website paths

In Fig. 3 we can look at the specific spike in traffic when it hit its peak – around the time in the afternoon of five Premier League matches at 3pm BST – and see the good bot vs. bad bot vs. human (other) activity during this spike. We can see evidence of defense bypass techniques, as bad bots (arb betting bots) attempted to emulate human activity by working in time with actual human traffic to hit the site at the times of match kickoffs, to avoid arousing any suspicion. The blue spike later in the afternoon is likely to represent humans cashing in their bets, which fails to correspond with the red spike as arb bettors tend to spike before the event rather than afterwards, since they scrape odds pre match.

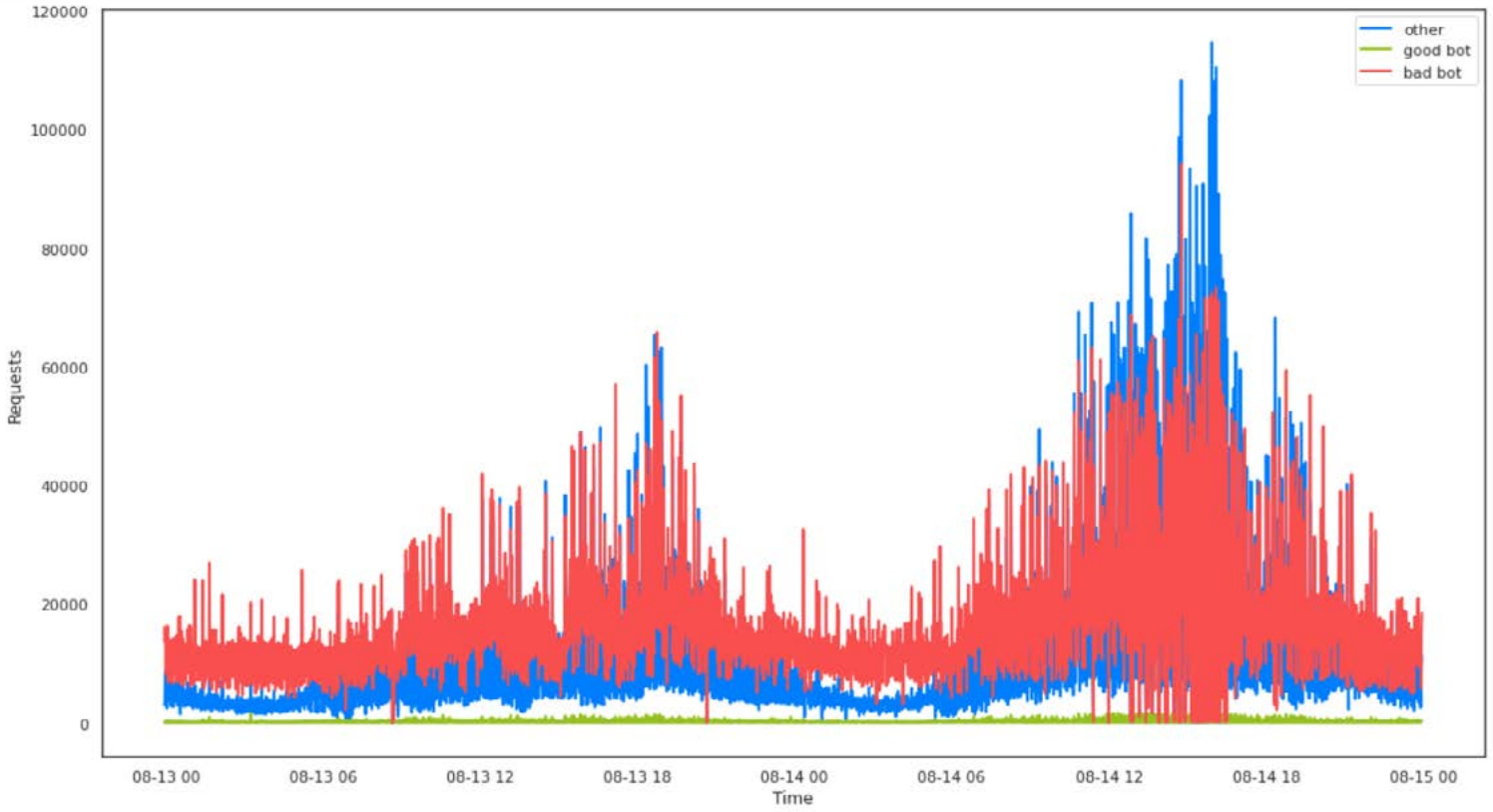


Fig. 3 – Good vs bad bot traffic on 'football' website paths

Traffic originated predominantly from three countries:

- UK
- Ukraine
- Mexico

After the home of the Premier League, it's unsurprising that Ukraine and Mexico, both big footballing nations, contributed the most traffic to the bookmaker's website paths during the opening weekend of top-division football. What is significant is the large numbers of countries contributing small amounts of traffic to the website. This could be the result of adversaries using proxy networks – gained either legitimately or illegitimately – to facilitate the scraping bots. Alternatively, this could also represent data centers around the world being used for the same purpose. Ultimately, making hundreds of millions of connections across one IP address would arouse suspicion; making hundreds of millions of connections via hundreds of IP addresses is far more likely to pass by volumetric-based security controls for the website. Thus, adversaries can carry out the bot activity without drawing attention to any non-human activity.

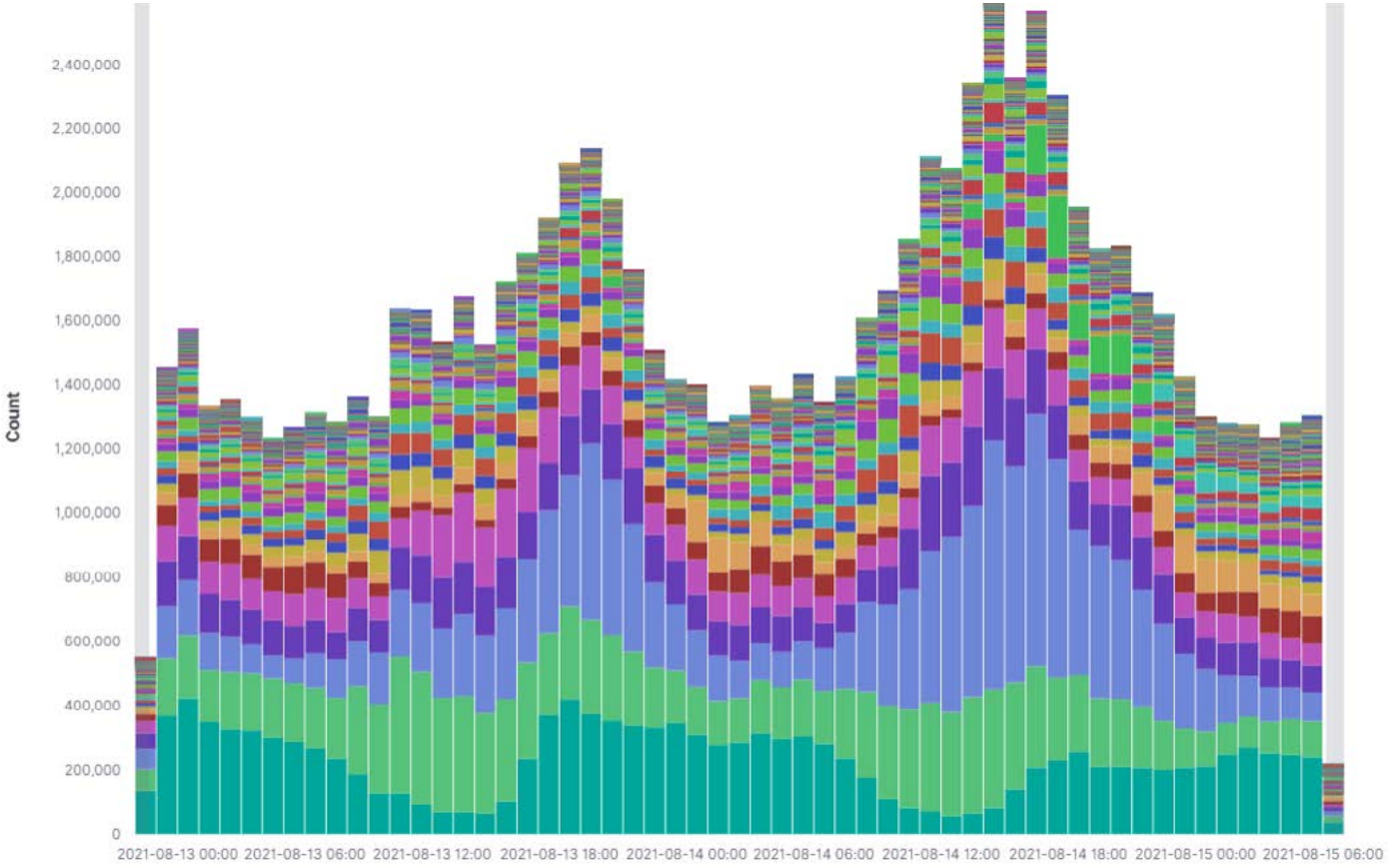


Fig. 4 – Origin countries for 'football' website paths traffic

How can bookmakers protect themselves?

Although bookmakers routinely ban and block users who are gaining an unfair advantage or going against their terms of service, arb bettors are often hard to distinguish, and the task is manual. Banning such users also does not stop the costly scraping activity carried out by automated arb betting bots.

As arb betting opportunities are time-bound and will quickly become invalid as bookmakers shift the odds for each event, scrapers must aggressively collect data to be effective. By slowing down or blocking this scraping activity, arb betting tools become ineffective to their users, and the bookmaker saves substantial money by not serving these requests.

Netacea has worked extensively with one of the biggest bookmakers in the world to identify and block scraper bots linked to arb betting activity. Our Intent Analytics™ engine uses advanced machine learning techniques to detect scrapers and categorize them based on the scraping activity, for instance, the information they are collecting and the patterns emerging in the collection methods. We can then challenge or block these bots, mitigating the damage they cause our client with minimal intervention needed.

Protect your betting business with Netacea Bot Management

To find out how much bots are costing your gaming and gambling business, visit netacea.com/impact-of-bots-calculator.

Or speak to the Netacea team to arrange a free demo at hello@netacea.com.