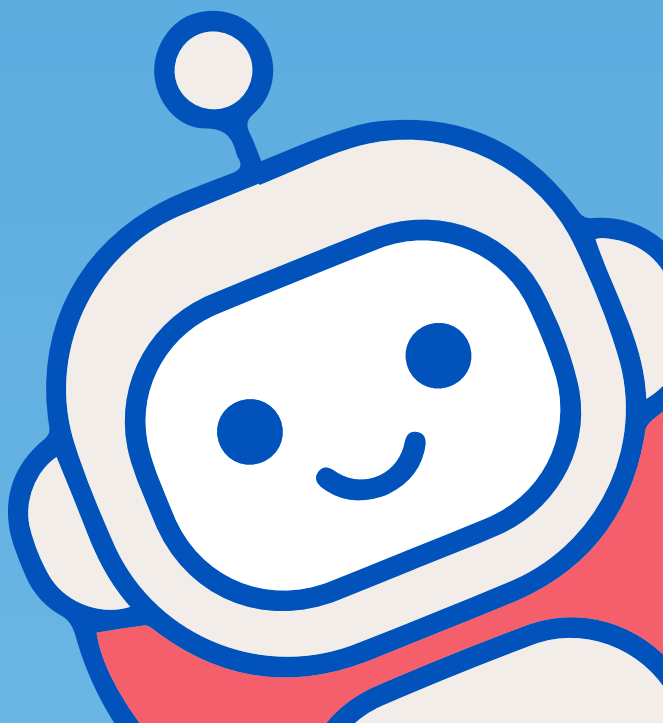
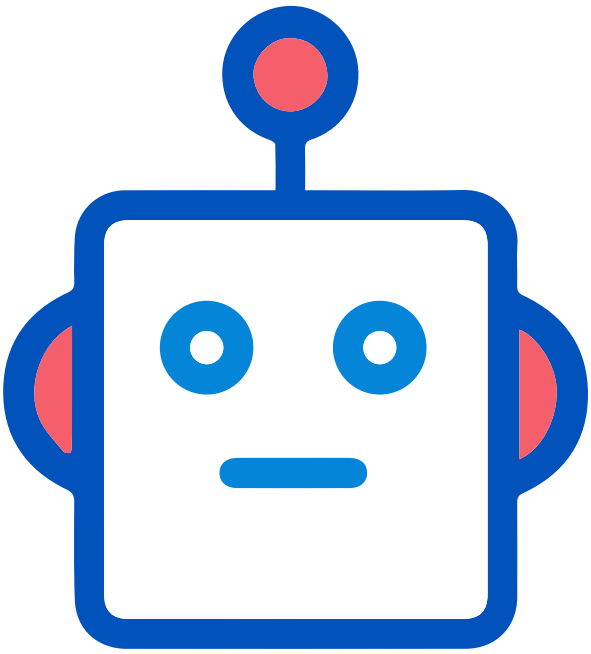
WHITEPAPER

Need for Managed **Bot Mitigation Solutions**



Bots Are Everywhere.



A bot—short for “internet bot” or “robot” is a software or application that has been created to operate as an agent of a computer program or user to simulate human activities. Bots are designed expressly for automating simple, repetitive, and mundane tasks.

Bots are everywhere in today’s technology. The chatbot that handles customer services is an example of the bot. Google’s search engine spider which crawls the website is also a bot. You have likely encountered bots on Spotify, Bank of America, Amazon, and many other services and product sites.

Bots can be broadly classified into two major types based on the intent/purpose of the Bots

Good Bots:

As mentioned before bots are created to do repetitive tasks, when the tasks that the good bots do are of legitimate nature then those bots are called good bots. For example, you might create a bot to check the availability of site over internet at periodic interval and report if it is not available. For this bot may regularly try to access a site (launch a http request) from various locations and trigger alarm if the response is other than 200 OK. This bot is solving a legitimate use case of monitoring availability of the site, would come under the category monitoring bots and would be classified as a good bot. There are various other categories of bots which are classified as good bots like search bots, chat bots, shopping bots etc.

Bad Bots:

Now if the tasks done by the bot is done for malicious purpose, then those bots are classified as bad bots. For example, an attacker might create a bot to flood an application with requests, so that it run out resources bringing down the application. For this bot will launch http requests from various locations. Number of requests will be high enough to bring down the application. Such bots fall under DDOS bot category as they are used to carry out DDOS attacks and is classified as bad bots. There are various other categories of bots that are classified as bad bots like scraper bots, spam bots, phishing bots etc.

Challenges of Bot Management:

Any Bot management solution has two major classification challenges.

1. Differentiating humans vs bots
2. Differentiating good bot vs bad bots.

Differentiating humans vs bots

The fundamental challenge of bot management is the nature of requests that the bot generates. Unlike exploits, where you could differentiate a request with an intent of exploiting vulnerabilities based on nature of the request, bots by virtue of it being a tool to do repetitive tasks, end up generating requests which will not differ much from requests made by humans. Nature of the requests won't change much, what will change is the intent of requests and this intent needs to be derived from the behavior of the request/requests. Certain characteristics of the request can be used to derive if the request is originating from a bot or not, but these characteristics are not well defined and would keep changing making it hard to differentiate a bot from human.

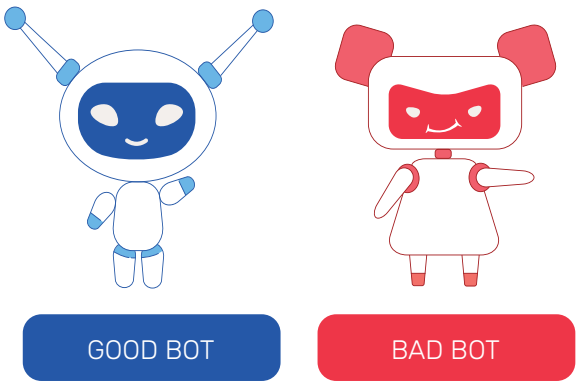
Differentiating good bots vs bad bots

Another major complexity when it comes to bot management is that it is hard to differentiate between good bots vs bad bots. If you see the example quoted before, monitoring bots and DDOS attack bots technically do the same thing. Both bots send http requests over internet, but the intent of these bots differ, intent can only be identified by the outcome that is achieved. Based on the intent, the way these bots are employed will differ and any bot management solution should be able to track the bot behavior and differentiate them as good vs bad bot.

For example, to differentiate between a bot that is employed for monitoring vs ddos, a bot management solution would have to track the number of requests coming from the same type of bots over time. Monitoring bots would periodically check the site, but volume of requests would not be very high, whereas bots that is trying to bring down the application would generate high volume, sustained or spiky, requests . So, in this case differentiation between good vs bad bot can be done by tracking the volume of requests

This is a very simple example but generally bots are more sophisticated, and it becomes very complex very fast. For example, think about a scraper vs search bot. Both bots would crawl your page and gather information. This information in case of search bot would be used for making your website more accessible whereas a scraper would use this information for malicious purpose like sharing pricing information to competition, creating an alternate site with similar content etc. Good bot mitigation solutions should be able to work through these complexities and accurately identify a bot and its intent.

According to research, nearly 40% of internet traffic are bots with nearly 25% of all requests coming from bad bots. Generally, bots are considered to be a nuisance but in reality, it is a common tool employed by threat actors to carry out complex attacks. Bots are sophisticated in nature and easily go undetected by conventional mitigation solutions.



What about Machine Learning (ML) & Artificial intelligence (AI)?

Technology is fast evolving, ML and AI has come a long way and do a much better job than humans at tasks such as image classification, translation, play and win complex games such as chess. Unfortunately, AI and ML is not a holy grail to solve all problems. In spite of having capabilities of mimicking human intelligence, AI cannot replace human intelligence when it comes to ways of understanding problem and finding solutions. But when it is about reducing errors and faults in the operational tasks and when it is about finding anomalies and irregularities, AI is way ahead of the human efficiency and capability. Apart from adding a robust security layer AI is super-efficient in evaluating the mistakes and all the errors that human intelligence is prone to commit. ML on other hand uses past data to derive patterns and help predicting future use cases. Effectiveness of ML depends on availability of right data sets and for any self-learning algorithms the data sets should be accurate. Unfortunately these self-learning algorithms are prone to manipulation from attackers who also have access to AI/ ML tools, algorithms and employ techniques like data poisoning, learning attacks in manipulating the algorithms as in real life it is impossible to keep feeding filtered out accurate data-sets to these algorithms to learn.

In short, it will be naïve to think that that a solution only using AI , ML algorithms can be effectively able to identify unknown bots , categorize them based on intent and block them. Bot management is complex, and a good bot management solution should employ all tools at its disposal. Of course, AI and ML have role to play, an important role to say the least as they are very effective in identify anomalies/ irregularities in through of data which humans cannot do. But what most solutions don't consider is that they are not as effective as humans in interpreting these anomalies to effectively identify bot and its intent. No technology is as effective as human intelligence in this aspect.

The Need of Managed Bot Mitigation Solution

Let's take a simple example of distributed DDOS happening from a bot where the attacker is launching 1000s of bots from all over the world accessing a single page. To identify that bot is malicious you will need to track the behaviour of requests from these bots. Now all the requests from these bots will be making a single request to a page (homepage in this case) and nothing more. This is an anomaly if requests from an entity is tracked. So this can be flagged, which an AI/ML tool can do effectively. But now can it be blocked out right? maybe not, what if there is product campaign to download an article. In this case also all the entity will only do one request to the download page . Behaviour wise it is same, 1000s of IPs are accessing a single page but intent is different and differentiating between them is not easy , unless a human is notified who can take a call. This is again a simple example to explain the concept, of course in this use case, a better algorithm may effectively identify and block bot without false positive, but in real life this becomes very complex for any algorithm to effectively take mitigation call all the time. There will be less complex cases where AI/ML tool can take a call with high confidence, but in many cases a notification to humans who take an intelligent call is more effective. AI/ML used to augment human intelligence and action is the most effective solution and best use of such technologies.

AppTrana – Bot Mitigation (Under the hood)

AppTrana is the only solution that provides a comprehensive Managed Bot Protection solution that actually works. It uses perfect amalgamation of human intelligence and AI/ML to provide an effective solution to the customers. As mentioned before AI/ML technologies are very effective in analyzing huge set of data and highlighting anomalies. Various models are built to identify these anomalies,

Examples of some models are

Correlated Risk-Scoring: All the newly added modules (Allow Good Bots/ Block Good Bot Pretenders, Tor IP, User Agent Based Detection, Suspicious Countries, IP Reputation, and Data Center IP) add their respective risk scores for an identity (generally an IP address), the risk score continuously gets adjusted and correlated to block the identity when it goes above a threshold

Good Bot Pretender Detection: There are legitimate bots that are beneficial, like search bots. One does not want to block such bots, but attackers take advantage of this and try to impersonate good bots. AppTrana has special checks to identify such bots.

Heuristics: Bots have unique characteristics, and they leave a pattern behind. AppTrana uses various heuristics derived from these patterns to detect bots. The model is constantly updated using both human inputs and self-learning techniques.

Integrity Checks: Various Integrity checks including browser integrity, HTTP request, IP reputations etc. are employed to identify if the bots are legitimate or not.

Fingerprinting & JS Detections: JS (JavaScript) detection is one of the common techniques employed to differentiate between a bot and human traffic. If the traffic is suspicious, JS challenges are thrown to identify bots. Cookies are injected to fingerprint each device/bot allowing AppTrana to write more complex checks based on behavior of devices/bots accessing your site.

Workflow Validation: Bots can be identified through workflow validation rules. Workflows can be defined based on valid, normal user behavior. Bots generally do not follow these workflows and can be easily identified with such well-defined workflow rules written by security experts working closely with customers.

Behavior Anomaly detections: Behavior of requests are tracked and abnormalities vs. regular behavior are immediately identified using various models and blocked. These models are self-learning and adjust to normal application variance automatically.

These models are built such a way that if there is a high degree of confidence that the requests are coming from malicious bots then immediate mitigation actions are taken automatically. In some cases the confidence score may not be high enough and all the models can determine is it is an anomaly. In such cases humans are notified who look at the notifications and take further actions as needed. These notifications come to our security experts who take a call on what needs to be done. Most times this happens without any intervention required from the customer. Certain times, additional inputs may be needed from customers to understand the application behavior. In such cases our security experts will reach out to customer and work with them to take further actions.

Enhanced Bot Protection

By leveraging our big data architecture, we have now built a correlation around our existing bot policies. For each request, various modules (Allow Good Bots/ Block Good Bot Pretenders, Tor IP, User Agent Based Detection, Suspicious Countries, IP Reputation, and Data Center IP) of the bot protection feature would inspect these requests simultaneously and collectively decide if that request was made by a bot or a human. And if made by a bot, whether it is a bad bot or not.

For every request, each module does these checks individually and if the check passes, it adds a risk score for the identity from which the request is generated (generally, the identity here is an IP address). Hence, for every identity making a request to the website protected by AppTrana, a risk score is added. For all the identities, the risk score starts with zero, and then, based on the behavior of such requests, various bot modules add risk scores to determine if the identity is a malicious bot or not.

AppTrana WAAP with these enhancements for Bot Protection, provides:

1. Behavioral & real-time analysis of bot traffic
2. Correlated Risk Scoring for bot detection & blocking based on Tor IP, IP Reputation, Suspicious Countries, Data Center IP risk scores
3. Better bot protection through custom controls
4. Advanced analytics and real-time visibility into bot mitigation

Combination of automated and manual techniques is very effective and provides comprehensive coverage against all type of automated threats as identified by OWASP.

OWASP Automated Threats Coverage Details

Threats	Characteristics	Coverage Details
Carding	Multiple payment authorization attempts used to verify the validity of bulk stolen payment card data	Protected
Token Cracking	Mass enumeration of coupon numbers, voucher codes, discount tokens, etc	Protected
Fingerprinting	Elicit information from the web, application and database servers about the supporting software and framework types and versions	Protected
Scalping	Obtain limited-availability and/or preferred goods/services by unfair methods	Protected
Expediting	Perform actions to hasten the progress of usually slow, tedious or time-consuming actions on behalf of a person	Protected
Credential Cracking	Identify valid login credentials by trying different values for usernames and/or passwords	Protected
Credential Stuffing	Mass log in attempts used to verify the validity of stolen username/ password pairs	Protected
Captcha Bypass	Solve anti-automation tests	Protected
Card Cracking	Identify missing expiry dates and security codes for stolen payment card data by trying different values	Protected
Scraping	Collect application content and/or other data for use elsewhere	Protected
Crashing Out	Buy goods or obtain cash utilizing validated stolen payment card or other user account data	Protected
Sniping	Last minute bid or offer, for goods or services	Protected
Vulnerability scanning	Crawl and fuzz application to identify weaknesses and possible vulnerabilities	Protected
Denial of Service	Target resources of the application and database servers, or individual user accounts, to achieve denial of service (DoS)	Protected
Skewing	Repeated link clicks, page requests or form submissions intended to alter some metric	Protected
Spamming	Malicious and/or more benign information addition, that appears in public or private content, databases or user messages	Protected
Foot printing	Probe and explore application to identify its constituents and properties	Protected
Account Creation	Create multiple accounts for subsequent misuse	Protected
Account Aggregation	Use by an intermediary application to collect together accounts and interact on their behalf	Protected

BOT protection is complex and you need a solution like AppTrana that provides comprehensive managed protection. Check out the solution now.

ABOUT INDUSFACE

Indusface is a leading application security SaaS company which secures critical Web, Mobile & API applications of 2000+ global customers using it's award winning fully managed platform that integrates web application scanner, web application firewall, DDoS, BOT Mitigation, CDN and threat intelligence engine. The company has been funded by Tata Capital Growth Fund, is Great Place to Work certified, mentioned by analysts like Gartner, Forrester etc in their reports and has been the recipient of many awards such as the Economic Times Top 25, NASSCOM DSCI Top Security Company, Deloitte Asia Top 100 and several other such prestigious recognitions.

Bengaluru | Vadodara | Mumbai | New Delhi | San Francisco

Contact: marketing@indusface.com