Beyond Blind Defense:
Gaining Insights from Proactive App Sec
46% of Web Traffic is Bots

Source: Distil Networks Bad Bot Report 2016
Why Trust? Verification Required.
Bad Bots Cause the Majority of Website Problems

19% of Traffic Causes the Following Problems

- Web Scraping
- Security Breach
- Spam
- Online Fraud
- Skewed Analytics
- Downtime
- Digital Ad Fraud
- API Abuse
# Web App Security Requires Complementary Solutions

## Core Competency

<table>
<thead>
<tr>
<th>DDoS Mitigation</th>
<th>Firewall</th>
<th>WAF</th>
<th>Distil Bot Protection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volumetric attacks on infrastructure</td>
<td>Network layer attacks</td>
<td>Application coding exploits</td>
<td>Automated abuse, misuse, and attacks (scraping, fraud, account takeover, etc.)</td>
</tr>
</tbody>
</table>

## Techniques

<table>
<thead>
<tr>
<th>DDoS Mitigation</th>
<th>Firewall</th>
<th>WAF</th>
<th>Distil Bot Protection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scrubbing centers, Large pipes</td>
<td>Access Control Lists (ACLs), Rules-Based</td>
<td>App layer understanding, ACLs, Rules-Based</td>
<td>Real-time Analysis, Fingerprinting, Honeypotting, Machine learning, Behavioral modeling</td>
</tr>
</tbody>
</table>
Majority of Bots are Advanced Persistent Bots (APBs)

APBs have one or more of the following abilities:

**Advanced**
- Mimic human behavior
- Load JavaScript
- Load external resources
- Support cookies
- Browser automation (Selenium, PhantomJS)

**Persistent**
- Dynamic IP rotation
- Distribute attacks across IP addresses
- Hide behind anonymous and peer-to-peer proxies

2016 Distil Bad Bot Report
That Majority of Bad Bots Now Use Multiple IP Addresses

Bots which dynamically rotate IP addresses, or distribute attacks are significantly harder to detect and mitigate.
Loading Assets & Bots Mimicking Humans

% of bots able to load external assets (e.g. JavaScript)

- Able to Load: 53%
- Not able to load: 47%

These bots will skew marketing tools such as (Google Analytics, A/B testing, conversion tracking, etc.)

% of bots able to mimic human behavior

- Bots Behaving Like Bots: 61%
- Bots Mimicking Human Behavior: 39%

These bots will fly under the radar of most security tools
The Antidote to Blind Defense

Identify
Hi-Def Fingerprint

Distil
Real-time Detection

Respond
Control with Certainty
Hi-Def Fingerprinting Eliminates Blind Defense

- IP Address
- Header & User Agent Information
- Cookie Browser
- 200+ Attributes of data (Navigator, WebGL, Plugins, Audio, Video, etc.)
- Tamper proofing layer
Delivering a Clear Picture of Your Web Traffic

Low Resolution Fingerprint
“Unactionable”

Hi-Def Fingerprint
“Actionable”
The Benefits of a Hi-Def Fingerprint

Pre-access Identification: Inspect traffic before website accessed

Hi-Def Fingerprint: Analyze over 200 device attributes

Verified Accuracy: False Positive Report

Respond With Confidence: Make better decisions

Shareable Public Fingerprint: Available for Known Violators Database & SIEM and other security products
Detect and Distil Traffic

Pre-Access Detection

Pre-Access Challenge

Progressive Detection

Performed Prior to Access to Website

Access to Website Granted
No Longer Blind Defense
Complete Visibility into False Positives

17 million CAPTCHAs served
78 solved
False Positive Rate = 0.00000458

<table>
<thead>
<tr>
<th>Date</th>
<th>Total Traffic</th>
<th>Total Served</th>
<th>Attempts</th>
<th>Solved</th>
<th>Failed</th>
<th>Success Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aug 7th, 2016</td>
<td>23,418,461</td>
<td>17,023,198</td>
<td>1,699</td>
<td>78</td>
<td>1,621</td>
<td></td>
</tr>
</tbody>
</table>

[Graph showing daily CAPTCHA requests]
Controlled Responses Based on Traffic Type

- Monitor
- Challenge
- Un-block Form
- Whitelist
- Hard Block

Specific countries
Aggregators

Exceeds rate limits
Behavioral anomalies

Known Violator Datacenters
Known Violator Database

Partners
Performance tools

Blacklist
Competitors
The Only Easy and Accurate Way to Protect Web Applications from Bad Bots, API Abuse, and Fraud.
Two Months of Free Service + Traffic Analysis

Offer Ends: October 30, 2016

www.distilnetworks.com/trial/
Some questions - any that you prefer to answer that we should include?

1. How can you identify traffic as malicious on your website, before it has done anything?

2. An attacker can use a browser automation tool like selenium. That would be hard to detect and it is not a real user, correct?

3. Why would you care if a bot is scraping your site? It’s public information, this is not a security problem.

4. Is a vulnerability scanner identified? Does it have a fingerprint if it doesn’t use a browser?

5. What types of companies use this bot protection?

6. You have to have blind trust on your website, everyone is anonymous on the web. Nobody wants to block customers from visiting our site. There must be false positives that you see, how do you handle that?
Detect and Distil Traffic

Pre-Access Detection  Pre-Access Challenges  Progressive Detection

Access to Website Granted
Detect and Distil Traffic

Pre-Access Detection
- Known Violators Database
- Known Violators DataCenter
- Known Violators Tools
- IP Blacklists
- Geo-Fencing Check
- Bad user agents
- Header Analysis Check

Pre-Access Challenges
- Javascript validation
- Browser validation
- Proof of work puzzle

Progressive Detection
- Device rate limits
- Deceptive honeycode
- Machine learning behavioral analysis

Access to Website Granted
How the Distil Bot Detection Solution Works

1. Fingerprints each incoming connection and compares it to our Known Violators Database
2. If it’s a new fingerprint, validates the browser to determine if it’s a Bot or Not
3. Based on your preferences, automatically tags, challenges, or blocks the bot
Flexible Deployment Options

- Install on virtualized or bare metal appliance(s)
- High availability configurations with failover monitoring
- Heartbeat up to Distil Cloud
- Deploys in days

Physical or Virtual Appliances

Content Delivery Network

- Automatically compresses and optimizes content for faster delivery
- 17 global datacenters automatically fail over when a primary location goes offline
- Automatically increases infrastructure and bandwidth to accommodate spikes
- Deploys in hours
Advanced Bot Detection Increases Accuracy

**Browser Validation**
Detects all known browser automation tools, such as Selenium and Phantom JS

Protects against browser spoofing by validating each incoming request as self reported

**Behavioral Modeling and Machine Learning**
Machine-learning algorithms pinpoint behavioral anomalies specific to your site’s unique traffic patterns

Self optimizing algorithms improve bot detection and mitigation without manual configuration
Device Fingerprinting
Fingerprints stick to the bot even if it attempts to reconnect from random IP addresses or hide behind an anonymous proxy or peer-to-peer network

Tracks distributed attacks that would normally fly under the radar

Without Impacting Users Sharing the Same IP
Avoids blocking residential users or organizations that might share the same NAT as the bot or botnet
Known Violators Database
Real-time updates from the world’s largest Known Violators Database, which is based on the collective intelligence of all Distil-protected sites

Distil customers are automatically protected against new threats discovered anywhere on the network.
Main use cases: Apps calling APIs, Partner APIs
Bots move to API if unable to access main site
Most solutions track API usage by IP

This makes them blind to a couple of key use cases

- **Server sourced** API clients are hosted by cloud providers that can cycle IP’s at will
- **Mobile application sourced** clients are behind Wireless provider proxy networks (many devices share an IP)
- **Web browser sourced** clients can be behind a consumer ISP NAT - shared IP for many browsers
Our flexible ID system can track API usage on a device level. For example:

- **Server sourced** API clients tracked by session or vendor ID’s
- **Mobile application sourced** clients tracked by device ID (Apple IDFA, or Google Advertiser ID)
- **Web browser sourced** clients tracked by session ID or if Distil Web Security is present: Distil UID’s
API Security Controls

- Configurable ACL
- Country and Organization fencing
- Progressive Thresholds and Graduated Rate Limiting
- Multi-Token Support
- Custom Rules
- Anonymous Proxy Blocking
- Referrer Blocking
- Blocking by Datacenter
Dedicated Analyst Team

**Proactive Analysis**
Human analysts proactively investigate and report on website anomalies.

**Priority Investigation**
Receive priority support on incidents, investigations, and support tickets.

**Account Review**
Explore new features, refresh your training, and influence our roadmap.

**Custom Reporting**
Tailored reports pairing your security data with business objectives.

**Engineering Assistance**
Leverage our team to plan and implement engineering projects.
The World’s Most Accurate Bot Detection System

**Inline Fingerprinting**
Fingerprints stick to the bot even if it attempts to reconnect from random IP addresses or hide behind an anonymous proxy.

**Known Violators Database**
Real-time updates from the world’s largest Known Violators Database, which is based on the collective intelligence of all Distil-protected sites.

**Browser Validation**
The first solution to disallow browser spoofing by validating each incoming request as self-reported and detects all known browser automation tools.

**Behavioral Modeling and Machine Learning**
Machine-learning algorithms pinpoint behavioral anomalies specific to your site’s unique traffic patterns.
How Companies Benefit from Distil

- Increase insight & control over human, good bot & bad bot traffic
- Block 99.9% of malicious bots without impacting legitimate users
- Protect data from web scrapers, unauthorized aggregators & hackers
- Slash the high tax bots place on internal teams & web infrastructure