December 20, 2016

The Methbot Operation

whiteops

The Methbot Operation

White Ops has exposed the largest and most profitable ad fraud operation to strike digital advertising to date. Russian cybercriminals are siphoning millions of advertising dollars per day away from U.S. media companies and the biggest U.S. brand name advertisers in the single most profitable bot operation discovered to date. Dubbed "Methbot" because of references to "meth" in its code, this operation produces massive volumes of fraudulent video advertising impressions by commandeering critical parts of Internet infrastructure and targeting the premium video advertising space.

Using an army of automated web browsers run from fraudulently acquired IP addresses, the Methbot operation is "watching" as many as 300 million video ads per day on falsified websites designed to look like premium publisher inventory. More than 6,000 premium domains were targeted and spoofed, enabling the operation to attract millions in real advertising dollars.

The following report illustrates the sophistication and rapid evolution of the Methbot operation and its damaging effect on the advertising ecosystem on both the demand and supply sides. This analysis is possibly only a fraction of Methbot's true impact. Because White Ops is only able to analyze data directly observed by White Ops, the total ongoing monetary losses within the greater advertising ecosystem may be larger. At this point the Methbot operation has become so embedded in the layers of the advertising ecosystem, the only way to shut it down is to make the details public to help affected parties take action. Therefore, White Ops is releasing results from our research with that objective in mind.

Information available for download

- IP addresses known to belong to Methbot for advertisers and their agencies and platforms to block.
 This is the fastest way to shut down the operation's ability to monetize.
- Falsified domain list and full URL list to show the magnitude of impact this operation had on the publishing industry. These publishers were impersonated and deprived of revenue opportunities because of this operation.

A Snapshot of the Methbot Operation

Volume and Estimated Financial Impact

- \$3 to \$5 million in counterfeit inventory per day
- CPMs ranged from \$3.27 to \$36.72 with the average being \$13.04
- 200 400 million video ad impressions generated per day on fabricated inventory
- 250,267 distinct URLs spoofed to falsely represent inventory
- 6,111 premium domains targeted and spoofed
- High value marketplaces targeted including PMPs

Operational Infrastructure

- 852,992 dedicated IPs, many falsely registered as US ISPs
- 800 1,200 dedicated servers operating from data centers in the United States and the Netherlands

Advanced Techniques to Avoid Detection

- Faked clicks, mouse movements, and social network login information to masquerade as engaged human consumers
- Manipulation of geolocation information associated with the IP addresses
 under their control
- Special case countermeasures against code from over a dozen different ad tech companies
- Fully custom http library and browser engine with Flash support, all running under Node.js

Table of Contents

06	Introduction
08	Discovery and History
09	The Advertising Ecosystem (The Basics)
10	How Methbot Infiltrates the Advertising Marketplace
11	Advances Bot Behaviors
12	A New Type of Bot Operation
14	Financial Impact
16	Looking Ahead - A Call Towards Transparency
17	Methbot - A Technical Analysis
19	Bot Characteristics
21	Key Behaviors
22	Countermeasures by Methbot to Avoid
24	Dynamic Code Patching
25	Viewability and Behavioral Spoofing Behaviors
27	Human Input Simulation
28	Forged IP Registration: Data Centers
29	About White Ops

Introduction

In September, 2015 the White Ops security research team noticed a small amount of automated web traffic featuring a unique bot signature which was quarantined and placed into monitoring. This signature, internally called "C3" showed little activity affecting White Ops clients until October of the following year when the bot morphed into Methbot and began to scale and adapt aggressively.

The measured impact to the advertising ecosystem is unprecedented. By fabricating as much as \$5 million in video advertising inventory per day, Methbot far exceeds the financial damages done by previously discovered botnets. ZeroAccess is thought to have collected as much as \$900,000 per day¹, the Chameleon Botnet up to \$200,000 per day², and HummingBad up to \$10,000 per day³.

BOT OPERATION	TYPE	FOCUS	ESTIMATED LOSSES PER DAY
Methbot	bot farm	Programmatic video advertising	\$3,000,000
ZeroAccess	malware	Ad fraud and bitcoin mining	\$900,000
Chameleon	malware	Ad fraud	\$200,000
Avalanche ⁴	malware	Identity theft, access control	\$39,139
Ponmocup ⁵	malware	Theft	\$27,778
Metuji and Mariposa ⁶	malware	Identity theft, access control	unknown

To avoid detection, the group developed and cultivated an array of infrastructure dedicated to the Methbot ad fraud operation. Instead of the more traditional malware botnet structures, which involve attacks on existing IP addresses and piggybacking on residential computers, Methbot operators farm out their operations across a distributed network based on a custom browser engine running out of data centers on IP addresses acquired with forged registration data.

Using these forged IP registrations has allowed the Methbot operation to evade typical datacenter detection methodology. This marks an innovation that transcends beyond traditional botnets, allowing Methbot to scale beyond anything the industry has seen before and placing it in a new class of bot fraud.

 Jackson Higgins, Kelly. "ZeroAccess Botnet Surges." Dark Reading. October 30, 2012. Accessed December 12, 2016. http://www.darkreading.com/risk/ zeroaccess-botnet-surges/d/d-id/1138615.

2. Spider.io. "Chameleon Botnet." Spider.io. February 28, 2003. Accessed December 12, 2016. http://www.spider.io/blog/2013/03/ chameleon-botnet/. ARS Technica. "Legal raids in five countries seize botnet servers, sinkhole 800,000+ domains." Sean Gallagher. December 1, 2016. Accessed December 16, 2016. http://arstechnica.com/security/2016/12/legal-raids-in-fivecountries-seize-botnet-servers-sinkhole-800000-domains/

5. Trend Micro. "The state of botnets in late 2015 and early 2016." December 17, 2015. Accessed December 16, 2016. Noah Gamer. http://blog. trendmicro.com/the-state-of-botnets-in-late-2015-and-early-2016/

 We Live Security. "Nine bad botnets and the damage they did. Karl Thomas. February 25, 2015. Accessed December 16, 2016. http://www.welivesecurity.com/2015/02/25/nine-bad-botnets-damage/

^{3.} Polkovnichenko, Andrey, and Oren Koriat. "HummingBad: A Persistent Mobile Chain Attack." Check Point Blog. February 05, 2016. Accessed December 8, 2016. http://blog.checkpoint.com/2016/02/04/hummingbad-a-persistent-mobile-chain-attack/.

Discovery and History

In September 2016, White Ops detected a mutation in a previously low volume bot signature which had been flagged as "C3" since September 2015. The security research team continued to track the evolution of C3 as it innovated and grew into what would become known as "Methbot."

On October 5, 2016, Methbot began to scale aggressively, reaching as many as 137 million impressions per day by the end of the week. The operation continued to expand rapidly. By mid-October, the White Ops MediaGuard Prevention Service was detecting three to five billion bid requests per day from Methbot spread across multiple ad platforms. By the end of the month, the bot farm had spread to affect 32 distinct clients upon which White Ops had detected or blocked activity.

Following the initial ramp in October, Methbot continued to produce massive amounts of impression volumes while continuing to adapt its codebase daily in an effort to elude fraud detection and viewability vendors and avoid discovery in order to continue the operation.





The Advertising Ecosystem (The Basics)

The Advertising Ecosystem

As the digital advertising landscape has come to rely more heavily on technology to transact advertising and media plans, technology platforms built to facilitate trading between advertisers (demand side) and publishers (supply side) have become vital components. This has created a traditional marketplace where demand side and supply side meet in order to transact and do business.

The Advertisers

Agencies, on behalf of their advertisers, plan and execute media plans aimed at delivering the highest quality targeted audiences. These plans may be delivered using technology called Demand Side Platforms (DSPs) which connect programmatic exchanges and publishers to find audiences from the supply side. Often, advertisers will layer in demographic targeting data or look for premium inventory inside of private marketplaces (PMPs) on exchanges to deliver high value against their media plan.

The Publishers

The Supply Side Platform (SSP) monetizes content for publishers. Often it will gather inventory from publishers and networks to provide a wealth of audiences to the DSP and other bidders inside of the exchange. To meet specific advertiser needs, SSPs and publishers may package up premium segments of inventory and audiences to offer up inside of a PMP. When the auction is finalized, the highest bid is accepted, the creative is loaded and sent to appear to the consumer on the content they are consuming.



The Adverting Ecosystem (The Basics)

The Methbot Profit Machine

How Methbot Infiltrates the Advertising Marketplace

Since both human audiences and premium publisher inventory are in high demand, Methbot focuses on manufacturing both of these as its product. By supplying faked audiences and hijacking the brand power of prestigious publishers through faked domains and falsified inventory, Methbot is able to siphon away millions in real advertising dollars.



BOT FARM

How Methbot Profits:

- Impersonate established sites and fabricate inventory
- Runs on a custom built dedicated desktop browser
- · Fakes mouse movements and social network logins

Advanced Bot Behaviors

Appearing Human

Advertisers often rely on data stored on a user's machine in cookies to target advertising against demographic information, browser histories, past purchases, and many other data points. Methbot operators use this industry approach to their advantage and stuff crafted cookies into fake web sessions by leveraging a common open source library which allows them to maintain persistent identities containing information known to be seen electronically as valuable to advertisers. In this way they take advantage of the higher CPMs advertisers are willing to spend on more precisely targeted audiences.

Methbot operators also forge tried-andtrue industry measures of humanity. Cursor movements and clicks are faked and multiple viewability measures are faked to further mimic observed trends in human behavior. Additionally, sophisticated techniques are employed to provide an even more convincing picture of humanity. Methbot forges fake social network login information to make it appear as if a user is logged in when an impression occurs.

Geolocation Provider Dataset Manipulation

Programmatic advertising typically uses geolocation providers to ensure ads are delivered to desired geographic regions, often at premium prices. Manipulating this data so it appears to have originated from more "premium" regions can increase the value of the fraudulent inventory.

Counter-detection Behaviors

Methbot uses several techniques to evade detection logic and spoof verification measurements in order to ensure high value for the operation's fabricated inventory.

Methbot is a New Type of Bot Operation

Datacenter-based ad fraud operations are usually easy to detect. Therefore, most advanced ad fraud operations have traditionally relied on malware that infects residential computers and sits on the same IP addresses as human users so the botnet can operate fake browser sessions in the background and generate fraudulent advertising activity. However, this approach has been a limiting factor because of the work needed to continually infect new home computers — especially while existing infections are being discovered and cleaned by anti-malware vendors.

Knowing this scenario, Methbot operators invested significant time, research, development, and resources to build infrastructure designed to remove these limitations and provide them with unlimited scale. They used dedicated servers to run proxies in order to hide the single origin source of their operation. Using falsified documents, the perpetrators were able to obtain or lease 852,992 real IP addresses, putting them to work generating fraudulent ad calls that appeared to come from legitimate residential Internet providers such as Verizon, Comcast, Spectrum, and others. The value of these IP addresses alone is over \$4 million today, according to figures posted by IPv4 Market Group.



Traditional Botnets Rely on Malware-Infected Residential Computers

The Methbot Bot Farm



Dedicated Infrastructure Provides Unlimited Scale

- 800 1,200 dedicated Methbot servers
- Distributed system to leverage parallel, reliable, and redundant operations
- Browser impersonation to resolve against fabricated premium domains
- Forged browser aspects including objects like screen information, plugin list, built in functions, and supported events

Financial Impact

White Ops consulted with AD/FIN, a programmatic media intelligence company, for representative cost data on the Methbot URL list. The analysis produced through this partnership showed that Methbot generated ad impressions sell for anywhere from **\$3.27 CPM** to **\$36.72 CPM**. The average CPM for URLs manufactured by Methbot was \$13.04.

The financial repercussions of Methbot continue to reverberate through the industry. Since early October 2016 White Ops estimates it has been running at a daily rate of 200 million to 300 million impressions per day. AD/FIN's CPM data places a value of this daily activity between \$3 million and \$5 million dollars per day.

Calculated Range for Methbot

	LOW	BEST ESTIMATE	HIGH
# Impressions/day	200,000,000	300,000,000	400,000,000
СРМ	\$13.00	\$13.00	\$13.00
TOTAL	\$2,600,000	\$3,900,000	\$5,200,000

White Ops continues to monitor and detect this operation on behalf of its customers and client platforms. However, given the size and reach of this enterprise, other companies outside our purview may also have been affected. Our goal is to shut down Methbot and stop the monetization power of this enterprise. Therefore we are releasing our research so that advertisers, agencies, platforms, and publishers, can arm themselves with this data and bring a stop to Methbot.

WE ARE RELEASING THE FOLLOWING FOR DOWNLOAD AT WWW.WHITEOPS.COM/METHBOT

IP addresses known to belong to Methbot for advertisers, agencies, and technology providers to block so they can prevent ads from appearing on Methbot inventory.

Falsified domain list and full URL list to show the magnitude of impact this operation had on the publishing industry. These publishers were impersonated and deprived of revenue opportunities because of this operation.

Looking Ahead - A Call To Transparency

The current complexity, interconnectivity, and resulting anonymity of the advertising ecosystem enabled the Methbot operators to exploit the entire marketplace. An impression may pass through many hands before it lands on a page and the ad is served. Tracing that complete path back through the various marketplaces proves difficult due to walled gardens, reselling, competing interests, and limitations on human capital to devote to this initiative.

The industry can respond by aligning supply and demand interests to protect the entire ecosystem from these types of specific threats. Close relationships between publishers and their advertisers may help circumvent much of this obfuscation and increase transparency, which will make it more difficult for even advanced operations like Methbot to take advantage of the system. Fraud is fluid and adapts to pressures by mutating, changing its codebase, or, when blocked, moving to target other unprotected sources of revenue. A combination of human best practices and technological vigilance by verification companies can help the industry close ranks against these threats and increase certainty through transparency for everyone across the advertising spectrum.

White Ops has partnered with The Trustworthy Accountability Group (TAG) tagtoday.net industry associations to facilitate the delivery and propagation of the data necessary to help bring the Methbot operation to a halt. We hope the public release of this research will result in a rapid end to this enterprise as White Ops remains on constant watch for new threats on the horizon.

Contact: threatintel@whiteops.com

Methbot A Technical Analysis

Methbot -A Technical Analysis

What follows is a more technical analysis of the specific tools tactics and procedures (TTPs) employed by Methbot operators which resulted in the substantial financial losses to the advertising ecosystem as illustrated in the preceding pages.

Bot Characteristics

Methbot uses custom software running on server-based infrastructure with dedicated IP space. White Ops detection technology was able to use a JavaScript language feature called "reflection" to gather extensive, detailed information about its inner workings. The bot runs under Node.js, and uses several open source libraries to add other features. It operates primarily on a large scale multi-data center distributed system to leverage parallel, reliable, and redundant operations.

Some open source libraries and tools used in the bot include:

- tough-cookie for preserving session data between executions
- cheerio for parsing HTML
- JWPlayer for running ad tags and requesting video advertisements
- Node.JS

Methbot is able to camouflage itself as any of the major desktop browsers by spoofing their user agent strings. Google's Chrome is the browser identity of which White Ops detected the highest volume, including minor versions 53 and 54. Firefox 47, Internet Explorer 11 and Safari 9.1 and 9.2 are also represented. Methbot operators also spoofed operating system including Windows 10 — and some older versions — and several versions of Mac OS X from 10.6 to 10.12.1.

In addition to browser impersonation, Methbot uses a variety of methods to fool anti-fraud technologies, embedding the appropriate contextual responses to further the illusion of a browser with a human user.

Distributed Hardware

Methbot nodes are physical servers located in data centers in Dallas, TX and Amsterdam. Each server runs multiple instances of the Methbot browser component and a proxy. As of December 2016, we estimate that there are 800 to 1,200 servers in use.

Proxy Network

Contrary to most ad fraud operations, Methbot acquires IP diversity through proxies operating on their servers. Normally this single-source approach is not effective due to IP metadata providers identifying such traffic as associated with datacenters rather than end-user systems and the advertising industry recognizing this and blacklisting compromised IP addresses belonging to datacenters.

Methbot operators have avoided this problem by gaining direct control of large contiguous IP allocations and falsifying

registration details so that they appear to be residential ISPs in the United States such as Comcast, Cox, AT&T, Verizon, Centurylink and others. Some of the earlier records found show completely fabricated entities, such as "HomeChicago Int", or use names similar to well known companies such as "AmOL wireless Net"", and "Verison Home Provider LTD". Based on the traffic observed, White Ops has attributed 852,992 IP addresses under IPv4 allocations controlled by Methbot. For comparison, Facebook only uses about half this number.

Key Behaviors

Video advertising on premium web sites fetches some of the highest prices in digital advertising. Methbot hijacks the brand power of premium publishers by spoofing URLs in the call for a video ad in order to attract advertising dollars in the following way:



Counterfeit page: Methbot selects a domain or URL from a list of premium publishers, and fabricates counterfeit pages. The page contains nothing more than what is needed to support an ad, and the publisher's server is never contacted.

Offer inventory: Using the industry standard VAST protocol, Methbot requests a video ad from a network, using one of Methbot's identifiers so they will get credit for it.

Produce fake views and clicks: The video ad is loaded through a proxy and "played" within the simulated browser. Any specified anti-fraud and viewability verification code is also loaded and fed false signals in order to make the activity seem legitimate.

To date White Ops has observed 250,267 distinct URLs across 6,111 distinct domains that were generated by Methbot in the act of impersonating a user visiting a web page.

Sample of Methbot generated URLs

*for the complete list, please refer to the accompanying documents

- http://ibtimes.co.uk/video
- http://vogue.com/video
- http://economist.com/video
- http://espn.com/video
- http://www.cbssports.com/CBS_Air_Force_Falcons_Fall_Gear
- http://fortune.com/2016/09/28/department-stores-closings/
- http://foxnews.com/video

COUNTERMEASURES BY METHBOT TO AVOID FRAUD-DETECTION SERVICES

Methbot uses several techniques to fool anti-fraud and viewability companies.

Fraud detection firms commonly measure browser environments and look for inconsistencies in the collected data. For example, the user agent string may claim to emulated aspects of a browser include be Firefox, but the browser may behave in an inconsistent manner. By faking browser behavior, Methbot's goal is to maintain the ability to operate and collect revenue by providing fraud detection companies with enough data that their traffic will not be flagged as fraud.

Methbot uses several techniques to evade the logic of viewability measurement and fraud detection companies. These objects like screen information, plugin list, built-in functions and supported events.

Screen Object Construction

```
function Screen(browser) {
var oss = browser.os[0];
this.availWidth = browser.screenW;
this.availHeight = browser.screenH - ((oss==='W')? 40: 27);
this.width = browser.screenW;
this.height = browser.screenH;
this.colorDepth = 24;
this.pixelDepth = 24;
this.availLeft = 0;
this.availTop = (oss==='W')? 0: 23;
this.orientation = {angle:0, type:"landscape-primary"}
```



Events are artificially created and dispatched

Methbot defines

var event = win.document.createEvent('UIEvents'); event.initEvent('beforeunload', false, true); win.dispatchEvent(event); event = win.document.createEvent('UIEvents'); event.initEvent('unload', false, false); win.dispatchEvent(event);

unction __MethSetSetters() {

Object.defineProperties(this, {

DYNAMIC CODE PATCHING

Where such emulation is too much trouble, Methbot dynamically modifies third party scripts to return "known good" values instead, avoiding the need to implement complicated APIs. This technique was used against various fraud detection measures, viewability code from multiple vendors, and login-status functions from various social networks.

Attempting to defeat function analysis code by patching it



The White Ops security research team found traces of analysis code where Methbot developers dissected the logic of the most widely adopted fraud detection vendors on the web. It's apparent that they spent some time reverse-engineering these capabilities, manually running portions of measurement code inside legitimate browsers to learn what its output looks like, and then porting the logic to spoof those values in Methbot execution context.

VIEWABILITY SPOOFING BEHAVIORS

Viewability is a technical measurement used to verify if a digital ad was in view on a screen and is often used as a billable metric for advertisers to pay for ads. In addition to code specifically designed to defeat viewability measurements used by specific vendors, White Ops found routines for spoofing industry-standard measurements. In particular, flash VPAID events are expected and handled.

For example, depending on the event received several spoofing actions are taken in the following functions:



if (typ === 'firstQuartile' && rand < 0.017) this.__MethFlashClick();</pre>

Finally, clicks are simulated in a randomly generated fashion to achieve a realistic rate.

•••••••



When execution is not interrupted - and if linked images were created in the context, as VAST companion banners would be - human events are simulated and clicks in those images are faked.

HUMAN INPUT SIMULATION

Several calls were found to a function called humanEvents, all triggered via a randomized timer.

Click emulation, to fool click validation logic looking for input events.

WillMethWilteboc(),
<pre>var event = doc.createEvent("HTMLEvents");</pre>
<pre>event.initEvent("load", false, false);</pre>
<pre>next.dispatchEvent(event);</pre>
browser.humanEvents();
<pre>winMethScripts = winMethTempScripts.concat(winMethScripts);</pre>
<pre>winMethTempScripts = [];</pre>
<pre>winMethNextScript();</pre>

Right after the DOM is initialized, the onload event is artificially created and dispatched.

The White Ops security research team found code indicating that Methbot is able to appear to be logged into popular social networks, pushing the perceived value of its traffic even higher.

FORGED IP REGISTRATION - DATA CENTERS IMPERSONATING RESIDENTIAL ACTIVITY

While it's common to see malware-infected computers, as part of a botnet, used for IP diversity, this is the first time data centers have been observed impersonating residential internet connections. This makes the scale of this operation virtually unlimited, with none of the typical durability issues of maintaining a constant base of infected user machines.

The proxies used by Methbot allow its traffic to originate from any of 852,992 IP addresses. It avoids typical data center detection by falsifying registration details of those IPs so that they appear to belong to residential ISPs in the United States. White Ops discovered forged listings for organizations such as Comcast, Cox, AT&T, Verizon, Centurylink and others. In addition to real companies, completely fabricated entities such as "HomeChicago Int", "AmOL wireless Net"", and "Verison Home Provider LTD" were discovered.

Beyond simply avoiding blacklists, this type of forgery allows Methbot operators to sell to advertisers willing to pay a premium to reach US consumers.

One Example:

% This is the	AfriNIC Whois server.
% Information	related to '196.62.32.0 - 196.62.63.255'
% No abuse con	tact registered for 196.62.32.0 - 196.62.63.255
inetnum:	196.62.32.0 - 196.62.63.255
netname:	TIME-WARNER
	Time Warner Cable Inc.
country:	US
admin-c:	1P9-AFKINIC
tech-c:	IP9-AFRINIC
status:	ASSIGNED PA
mnt-by:	IP-ADMIN
mnt-lower:	IP-ADMIN
<pre>mnt-domains:</pre>	
mnt-routes:	
	adw0rd.yandex.ru@gmail.com 20151014
source:	AFRINIC 196.62.0.0 - 196.62.255.255
parent:	190.02.0.0 - 190.02.255.255
person:	IP Admin
address	TP Admin
phone:	+2482534202
e-mail:	adw0rd.yandex.ru@gmail.com
nic-hdl:	IP9-AFRINIC
changed:	adw0rd.yandex.ru@gmail.com 20151014
source:	AFRINIC

- Falsely registered to Time Warner Cable Inc, etc
- US listed as the country
- Suspicious contact
 email address
- Seychelles
 phone number

About White Ops

White Ops is a global leader in advertising fraud protection and offers verification and optimization solutions to the advertising industry. Combining data science with advanced solutions designed to detect and prevent fraudulent advertising activity, our company's mission is to stop the spread of advertising fraud through our human verification techniques. We work collaboratively with industry groups who are dedicated to preventing malicious activity in the advertising space and promoting transparency for the industry as a whole. White Ops is headquartered in New York City with satellite nodes operating in countries around the world.

To learn more please visit www.whiteops.com

whiteops

The Methbot Operation

CONTACT US threatintel@whiteops.com