Building a Public RPZ Service to Protect the World's Consumers

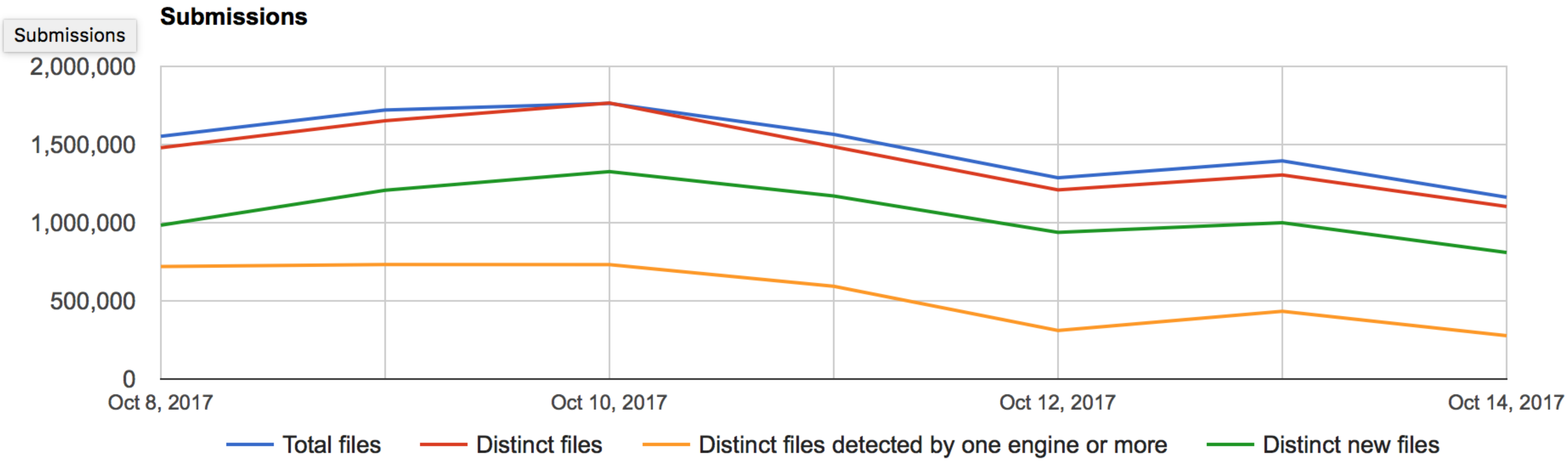
John Bambenek, Manager of Threat Systems
Fidelis Cybersecurity

The content of this presentation can be considered TLP:WHITE.
I will identify any specific data points I discuss that are more sensitive and shouldn't be disclosed as we go.

Introduction

- Manager of Threat Systems with Fidelis Cybersecurity
- Part-Time Faculty at University of Illinois in CS
- Handler at the SANS Internet Storm Center
- Provider of open-source intelligence feeds… DGAs! ☺
- Run several takedown oriented groups and surveil threats

The Problem Illustrated (from Virustotal)



The Reality

- There is a much smaller set of actual malware tools, EKs, and criminal service providers.
- Problem: Most cybercrime impacts consumer networks that are unprotected by security vendors, enterprise SOCs, etc.
- Consumers generally will not secure their devices, pay for security, or clean up malware infections except in rate cases.

Good News

- We have tons of open-source data, blocklists, and tracking systems out there publishing near-time data on threats.
  - My DGA feeds
  - Malware Domains
  - Abuse.ch trackers
  - Phishtank
  - Literally hundreds more (and that’s just Western ones)

Requirements to Protect Consumers

- Must have no or extremely low false positives
  - Compromised websites
  - Shared hosting
- Must not impact the user experience
- Must not increase cost
- Must be automated to deploy/update
- Ideally include some ability for security awareness of infected users

Two Partial Solutions

- Almost all malware requires DNS at some point in the flow of traffic.
- Every piece of malware wants to “talk somewhere”.
- Solution:
  - RPZ
  - Auto-generated Firewall Rules

RPZ Primer

- DNS Response Policy Zones are, in essence, a DNS firewall. On a resolver, you create zone files for things you want to protect constituents from.

- Instead of getting www.badsite.com, they get something else depending on what you define.

RPZ Indicator Types

- Things RPZ can detect on:

  - Hostnames and Domains
  - The resolved IP address
  - The nameserver hostnames used
  - The nameserver IP addresses used

RPZ actions

- You can return NXDOMAIN
- You can log the query but let it happen
- You can modify the query to point the constituent to a "walled garden".
  - Instead of going to www.badsite.com, they go to an IP you define.
  - You can use this to tell the victim they are infected, do some security awareness, and work on cleaning up infections.

RPZ Example

```
   $TTL 60
   @          IN    SOA  localhost. root.localhost.  (
                        282   ; serial
                        3H  ; refresh
                        1H  ; retry
                        1W  ; expiry
                        1H) ; minimum
            IN    NS    localhost.
oysjtyymfwbhfxv.com CNAME .
*.oysjtyymfwbhfxv.com CNAME .
ccvjoddsmsoheev.net CNAME .
*.ccvjoddsmsoheev.net CNAME .
paunsiqcihxtmgv.biz CNAME .
*.paunsiqcihxtmgv.biz CNAME .
```

RPZ Resolved IP address example

```
# 5.8.37.0/24 (Listed in Spamhaus DROP List SBL284078)
24.0.37.8.5.rpz-ip    CNAME .
# 36.93.0.0/16 (Listed in Spamhaus DROP List SBL310189)
16.0.0.93.36.rpz-ip   CNAME .
```

RPZ Nameserver examples

```
# Block ns1.bambenekconsulting.com

ns1.bambenekconsulting.com.rpz-nsdname CNAME .

# Block nameservers at 8.8.8.0/24

24.8.8.8.8.rpz-nsip   CNAME .
```

Sounds good… what’s the catch?

- RPZ is pretty straight-forward to set up… the problem is getting data.
- We have dozens of hostname/domain lists.
- We have hundreds of IP lists.
- Some are documented, many are not. Few have any real confidence indicators.
- What about false positives? Compromised sites? Shared infrastructure?

DGA Example

- Usually a complex math algorithm to create pseudo-random but predictable domain names.

- Now instead of a static list, adversary has a dynamic list of hundreds or thousands of domains and adversary only needs to have a couple registered at a time.
  - newfandultimati.cc,Domain used by tinba,2017-08-23 16:00,http://osint.bambenekconsulting.com/manual/tinba.txt
  - ybguvvvvcduv.trade,Domain used by tinba,2017-08-23 16:00,http://osint.bambenekconsulting.com/manual/tinba.txt

DGA Difficulties

- Word-list based DGAs:
  - windbearboxreceive.com, Domain used by matsnu DGA
  - winner-care-sir.com, Domain used by matsnu DGA
  - theirtheandaloneinto.com, Domain used by Rovnix DGA
  - thathistoryformertrial.com, Domain used by Rovnix DGA

DGA Difficulties

- DNS is under the ***complete control*** of the adversary. They can point any of their domains to anywhere they want.
- What if a domain pointed to these IPs and people ingest them into their firewall?

198.41.0.4
192.228.79.201
192.33.4.12
199.7.91.13
192.203.230.10
192.5.5.241
192.112.36.4
128.63.2.53
192.36.148.17
192.58.128.30
193.0.14.129
199.7.83.42
202.12.27.33

DGA Difficulties

- DNS for malicious domains is under the complete control of the adversary (until/unless we seize the domain).

- Using resolved IPs for RPZ or firewall rules without any filtering is giving the adversary control over your firewall or RPZ zones.

- As of now, there are no good nameserver/nameserver IP feeds (that I’m aware of).

Generating RPZ files

- For each source and each DNS record type, generating a zone is a matter of a for loop.
- Allows local locations to choose their own confidence for each file.
- Allows for different policies by zone.
  - For instance, different landing pages for phishing vs malware C2s.
- Possible to create global whitelists to prevent essential infrastructure for being blocked (i.e. root servers)

DGA Feeds

- My DGA feeds include all 4 indicators types (domain, IP, nameservers, nameserver IPs).
- Use domain names (unless wordlist or shorter than 7 characters).
  - For word-lists / short domains, log but don’t block.
- Don’t use IPs at this point.
- Not using nameserver details at this point (no good way to do it automatically yet)

Malware Configs

- Every malware has different configurable items.
- Not every configuration item is necessarily valuable for intelligence purposes. Some items may have default values.
  - Free-form text fields provide interesting data that may be useful for correlation.
  - Mutex can be useful for correlating binaries to the same actor.
- How to get to the identity of someone using Cobalt Strike to attack you?
- KEY POINT: Non-operational data is still useful for intelligence purposes.

Sample DarkComet Data

Key: CampaignID	Value: Guest16
Key: Domains	Value: 06059600929.ddns.net:1234
Key: FTPHost	Value:
Key: FTPKeyLogs	Value:
Key: FTPPassword	Value:
Key: FTPPort	Value:
Key: FTPRoot	Value:
Key: FTPSize	Value:
Key: FTPUserName	Value:
Key: FireWallBypass	Value: 0
Key: Gencode	Value: 3yHVnheK6eDm
Key: Mutex	Value: DC_MUTEX-W45NCJ6
Key: OfflineKeylogger	Value: 1
Key: Password	Value:
Key: Version	Value: #KCMDDC51#

Sample njRat config

Key: Campaign ID Value: 1111111111111111111

Key: Domain Value: apolo47.ddns.net

Key: Install Dir Value: UserProfile

Key: Install Flag Value: False

Key: Install Name Value: svchost.exe

Key: Network Separator Value: |'|'|

Key: Port Value: 1177

Key: Registry Value Value: 5d5e3c1b562e3a75dc95740a35744ad0

Key: version Value: 0.6.4

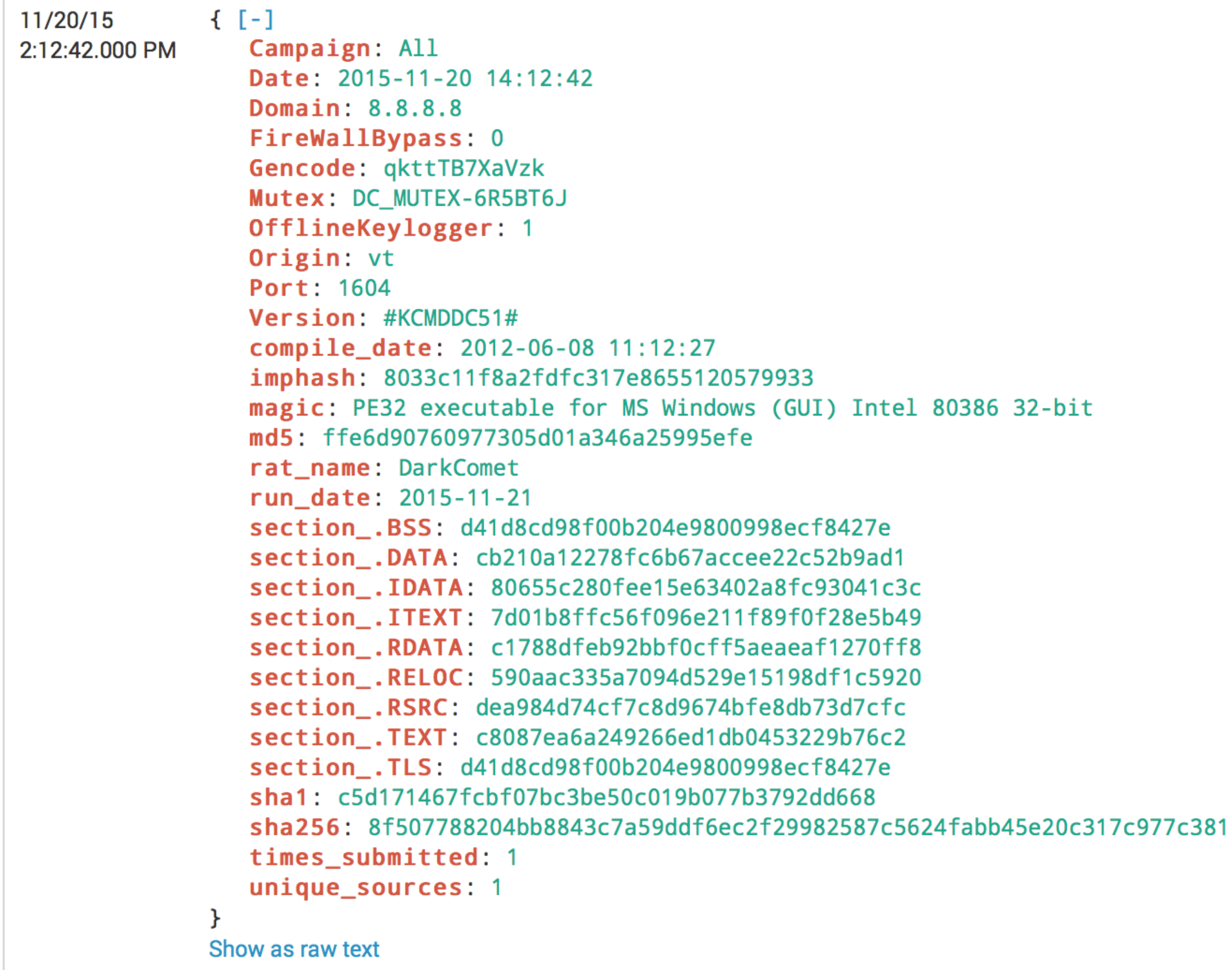
Bad Data?

- Malware builders can be used by adversaries for a variety of purposes.

- They know we mine configs for purposes of creating feeds and the like.

- There have been cases with “bad data” sent to VirusTotal and other places for the purposes of poisoning automated feed generation.

Bad Data?



```
11/20/15             { [-]
2:12:42.000 PM          Campaign: All
                        Date: 2015-11-20 14:12:42
                        Domain: 8.8.8.8
                        FireWallBypass: 0
                        Gencode: qkttTB7XaVzk
                        Mutex: DC_MUTEX-6R5BT6J
                        OfflineKeylogger: 1
                        Origin: vt
                        Port: 1604
                        Version: #KCMDDC51#
                        compile_date: 2012-06-08 11:12:27
                        imphash: 8033c11f8a2fdfc317e8655120579933
                        magic: PE32 executable for MS Windows (GUI) Intel 80386 32-bit
                        md5: ffe6d90760977305d01a346a25995efe
                        rat_name: DarkComet
                        run_date: 2015-11-21
                        section_.BSS: d41d8cd98f00b204e9800998ecf8427e
                        section_.DATA: cb210a12278fc6b67accee22c52b9ad1
                        section_.IDATA: 80655c280fee15e63402a8fc93041c3c
                        section_.ITEXT: 7d01b8ffc56f096e211f89f0f28e5b49
                        section_.RDATA: c1788dfeb92bbf0cff5aeaeaf1270ff8
                        section_.RELOC: 590aac335a7094d529e15198df1c5920
                        section_.RSRC: dea984d74cf7c8d9674bfe8db73d7cfc
                        section_.TEXT: c8087ea6a249266ed1db0453229b76c2
                        section_.TLS: d41d8cd98f00b204e9800998ecf8427e
                        sha1: c5d171467fcbf07bc3be50c019b077b3792dd668
                        sha256: 8f507788204bb8843c7a59ddf6ec2f29982587c5624fabb45e20c317c977c381
                        times_submitted: 1
                        unique_sources: 1
                     }
```

Scrubbing Malware Configs

- Phase 1:
  - Eliminate RFC 1918 IP addresses
  - Use only dynamic DNS hostnames
- Phase 2 (not yet operational):
  - For all non-RFC 1918 IP addresses and domain names, do syn() check on port to check if up.
  - Skip common ports.
  - If something is listening there, it's “bad enough”

What is an Exploit Kit?

- Set of tools (prominently web-based) that exploit vulnerabilities in software (browser, Adobe, Java, etc) to spread malware.
  - Relatively static list of exploits each kit uses and they vary.
  - Rarely (but sometimes) use 0-days.
- They operate as a criminal service and “sell infections” of whatever provided malware.
- Primary defense: patch your OS and applications.

Using a crawler

- Inefficient because it will request more than what you are looking for.

- Crawlers are also resource intensive the broader you are looking for behavior.

- It can, however, have a global footprint and be thorough.

Using a crawler

- Luckily, we don’t have to make our own crawler when Microsoft will give Bing crawler malicious URLs to MAPP/VIA members.

- About 26M malicious webpages daily were seen which Microsoft gives a 99% confidence interval too.
  - Much more than EKs.

Using Bing Malicious URLs

```
8/4/2016 4:58:27 PM	http://0000-programasnet.blogspot.com.ar/2011/03/my-defragmenter-my-defragmenter-es-un.html?action=backlinks&widgetId=Blog1&widgetType=Blog&responseType=js&postID=6994789541307753585	216.58.216.193	us	15169	MalwareNetwork

8/4/2016 4:51:46 PM	http://0000-programasnet.blogspot.com.ar/2011/03/pocopique-tv-programa-para-ver-tv.html?action=backlinks&widgetId=Blog1&widgetType=Blog&responseType=js&postID=7841830628282890204	216.58.192.129	us	15169	ES

8/4/2016 6:06:13 PM	http://0000-programasnet.blogspot.com.ar/2011/07/reparacion-de-impresoras.html	216.58.192.129	us	15169	ES

8/4/2016 6:26:04 PM	http://0000-programasnet.blogspot.com.ar/2011_02_24_archive.html	216.58.192.129	us	15169	MalwareNetwork

8/4/2016 4:34:23 PM	http://0000-programasnet.blogspot.com.es/2011/02/descarga-chat-para-facebook.html?action=backlinks&widgetId=Blog1&widgetType=Blog&responseType=js&postID=2134381520774268527	216.58.192.225	us	15169	MalwareNetwork
```

Populate Data from Exploit Kits

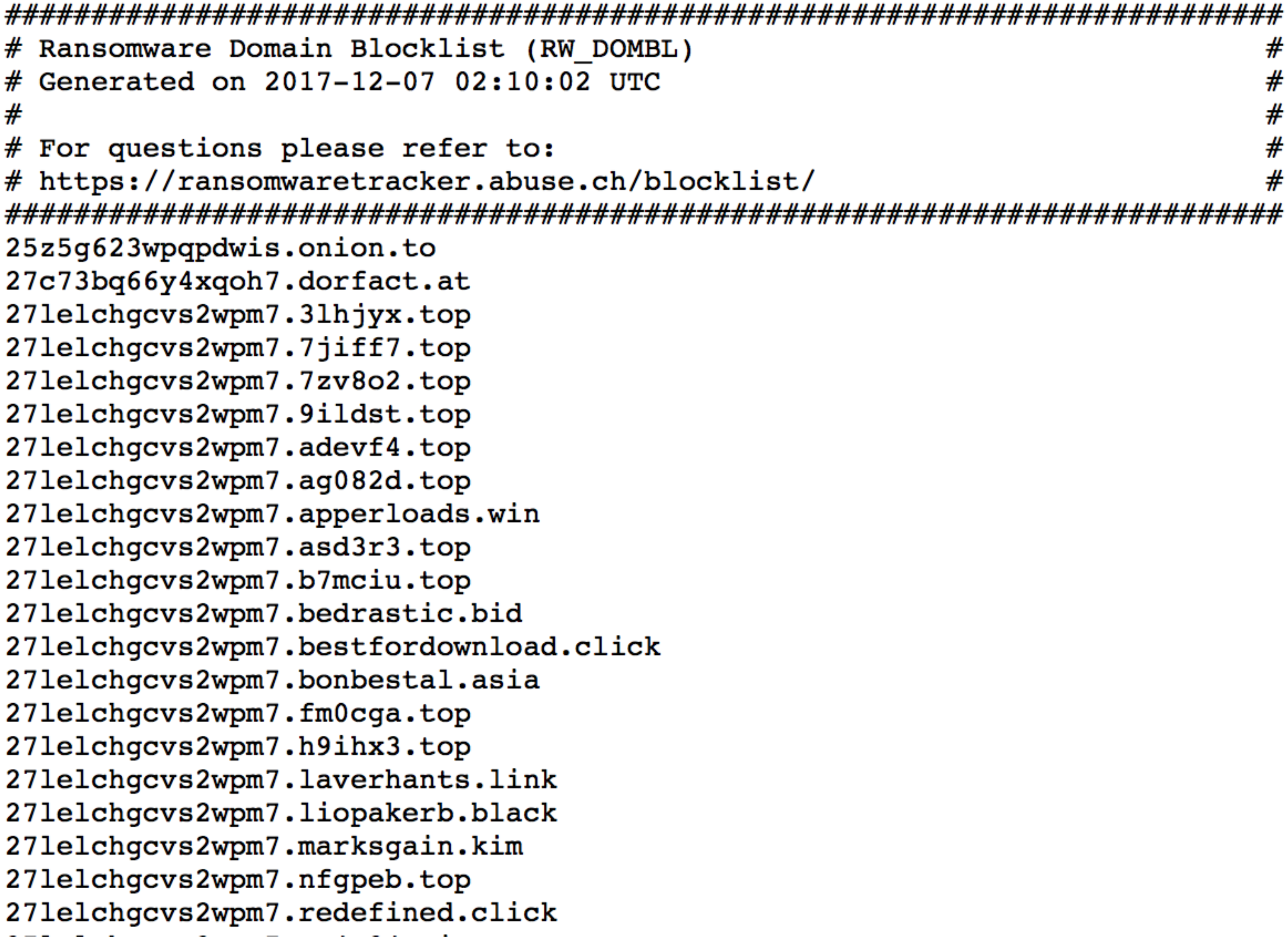
- EKs have a hierarchical structure but the deeper levels also need to be aware of the landing pages to prevent people artificially getting malware directly from the source.

- Some of these systems have vulnerabilities that will give you more info about the overall EK infrastructure, including landing pages and traffic delivery systems.

Protecting Consumers From Ransomware

- Abuse.ch has a ransomware tracker at https://ransomwaretracker.abuse.ch/ where you can download domain blacklists, IP blacklists, and URL blacklists.
- Some risk for false positives but it is well documented.
- Can even protect against tor2web based C2s.

Ransomware Tracker



```
##########################################################################
# Ransomware Domain Blocklist (RW_DOMBL)                                 #
# Generated on 2017-12-07 02:10:02 UTC                                   #
#                                                                        #
# For questions please refer to:                                         #
# https://ransomwaretracker.abuse.ch/blocklist/                          #
##########################################################################
25z5g623wpqpdwis.onion.to
27c73bq66y4xqoh7.dorfact.at
27lelchgcvs2wpm7.3lhjyx.top
27lelchgcvs2wpm7.7jiff7.top
27lelchgcvs2wpm7.7zv8o2.top
27lelchgcvs2wpm7.9ildst.top
27lelchgcvs2wpm7.adevf4.top
27lelchgcvs2wpm7.ag082d.top
27lelchgcvs2wpm7.apperloads.win
27lelchgcvs2wpm7.asd3r3.top
27lelchgcvs2wpm7.b7mciu.top
27lelchgcvs2wpm7.bedrastic.bid
27lelchgcvs2wpm7.bestfordownload.click
27lelchgcvs2wpm7.bonbestal.asia
27lelchgcvs2wpm7.fm0cga.top
27lelchgcvs2wpm7.h9ihx3.top
27lelchgcvs2wpm7.laverhants.link
27lelchgcvs2wpm7.liopakerb.black
27lelchgcvs2wpm7.marksgain.kim
27lelchgcvs2wpm7.nfgpeb.top
27lelchgcvs2wpm7.redefined.click
```

Bottom Line

- Of the hundreds of open-source feeds online today, I run only a couple.
- I am very familiar with a couple more.
- I have no global visibility, what not impactful to Western networks could be very impactful elsewhere (baidu, for instance, is flagged by many DGA detection algorithms).
- The others, I rely on what’s online (and there often isn’t much), how do I assess confidence?
- Instead of using “.” (NXDOMAIN), use CNAME rpz-passthru.
  - Results still logged so you can assess false positives.

Quad9s

- An open-resolver was set up using a similar approach, Quad9s (9.9.9.9).
- Set up with locations all over the world and free to use.
- Ideal for consumers (or consumer devices) looking to just point to something.
- What you gain in simplicity, you lose in telemetry.
- You can’t deliver targeted security awareness.

How to Deliver This?

- I have an RPZ server (rpz.bambenekconsulting.com) that can do DNS zone transfers (sign up form soon).
- I can deliver “master” zone files that aggregate all of this or hundreds of zone files so you can mix or match, but how do I communicate changes?
- Do you want me controlling your DNS policy?
- I’ve opted just in the last week to simply make a RPZ zone file generator framework instead so YOU can decide your own policy and confidence levels in data.

Future Work?

- Taking bulletproof and other criminal networks by ASN and blocking their entire IP space.
- Finding some way to find malicious nameservers in an ***automated*** way so I can stand behind them if they are blocked.
- Getting every consumer facing service provider to adopt some methodology like this.

To Sign up For RPZ Data

- I have a server providing RPZ zone transfers, you can sign up for that here:
  - URL: https://docs.google.com/forms/d/1rcLFEfSmo09lPQM8YT4VU3ixTwZ-1lK_0G5R3wk5oJY/viewform?edit_requested=true

- Coming soon:
  - Open-source RPZ zone-file generating tool (code published soon)
  - Github URL: https://github.com/bambenek/rpz-gen

Questions & Thank You!

John Bambenek / john.bambenek@fidelissecurity.com
Twitter: @bambenek

To access my DGA feeds go to:
http://osint.bambenekconsulting.com/feeds

To request access to Barncat Malware MISP go to:
https://www.fidelissecurity.com/resources/fidelis-barncat

Fidelis™ Cybersecurity