proofpoint

ET Intelligence Replist Overview

Introduction

Emerging Threats ET Intelligence Rep List is published in two separate lists. One file contains the IP listings and the other file contains the Domain Name listings. Each list is published hourly and is available for download from our server using an authorized license key issued for evaluation purposes or through purchase.

In total, the ET Intelligence Rep List (IP & Domain files) contains hundreds of thousands of entries in CSV format with downloadable updates every hour. About 30% of the list, on average, is turned over about every 3 to 4 weeks.

Entries include IP addresses and Domain Names, which have been observed by our own sensors operating in various large-scale networks globally, as well as a proprietary Malware Sandnet. There is no IP or Domain data that is derived from the open source community, or other sources. This data is based on real world observation & captured over a period of up to 3 years. While the requirement for direct observation can limit the sheer size of the numbers in the IP and domain lists, we feel that this burden of proof is paramount to accuracy. We strive to keep the false positive ratings as low as possible to avoid unwanted and unwarranted alerts.

Each entry in an ET Intelligence Rep List is categorized, according to the type of behavior detected, and given a Reputation Score ranging from 0 to 127. The rolling list contains roughly 1 million entries scored, BUT we only publish entries with a score greater than 20. This results in a published list containing 50 to 100k Domain Names and 350 to 500K IP addresses. Large fluctuations in the list are possible from day to day, resulting from the discovery and take down of large Botnet infrastructures.

About 80% of these listings are unique, as some of the listings are duplicated across two or more categories.

Note that the categories represent observed behavior, and should not serve as an absolute indicator of malicious intent. The goal is to produce indications of malware activity/behavior (categories) mapped to a level of certainty (reputation score) that those observed behaviors are what we believe them to be. This design will have an

impact on the information's use by the operator or system that consumes it.

We recommend using the following two specific use cases as best practice examples. The first use-case, involves using the Rep List for *reporting* purposes. In this example, a system such as a log aggregator or SIEM (Security Information and Event Management) tool parses events or logs, in order to match IP addresses or domains found in logs with those found on the reputations lists. For this use-case, a more liberal inclusion of categories will prove beneficial. In this case, the Rep List is leveraged for informational purposes and provides greater context for the analyst or analytics of reported events. In this use-case, more information is better, with the categories showing different behaviors being observed that form a complete and conclusive profile of modern malware detected. Be advised, there are some categories, that alone, may not indicate behavior that is certain to be malicious – Spyware as an example. We have, however, based on our vast experience with malware, constructed the categories such that, these behaviors, when observed together by a single actor typically represent the activities of malicious malware. In this use-case, more information provides context, and not disruption.

In our second use case, if the intent is to provide a mechanism for any type of network blocking, or *prevention*, then the following caution should be observed. As stated above, rep list categories are constructed based upon individual malware behaviors, that when observed together provide very accurate and granular context to the analyst or analytics. The categories ARE NOT constructed to be considered a final and binary block list (sometimes called black/white list). This is because there are categories representing certain behaviors that may not be malicious when observed independently.

Scoring

Reputation scores range from 0 to 127. This range was chosen so that it could be seamlessly ingested into the Suricata platform. With Suricata, IP reputation data (and very shortly domain reputation) can be integrated into the system on a per rule basis with specific directives, allowing the operator to block or alert based upon the type of information in each ET Intelligence rep list.

An IP Address or Domain is given a score within each category it is assigned. Reputation score is indicative of the confidence we have in assigning each category (or categories, if

there are more than one) to the IP Address or Domain. Reputation score, within a category, is driven by both volume of activity and type of activity. For example, some aspects of reputation are driven by ETPro signatures deployed worldwide, in this case the volume of hits and types of signatures both impact the assigned score within the category they serve. The scores do not necessarily reflect level of risk or maliciousness, only that there is more or less reason to believe that the assigned category is accurate and appropriate.

Categories

Categories represent behaviors, and as with the reputation an individual can have in life, actions speak louder than words. The more an individual tells the truth, the better chances they will have of gaining a reputation for honesty. While the more lies told, the greater chances of being known as dishonest. Scoring would be analogous to the number of lies told, and types of situations where one summoned the courage to tell the truth—here volume and type will have a big influence on the end result. The resulting reputation will either be honesty, or dishonesty—similar to our categories.

The combination of categories and scores form a context in which the informed analyst can pass judgement. Examples of judgement could be: high risk, low risk, neutral, good, and bad. The context of the situation, along with one's goals, and his/her values produce judgement. The ET Intelligence Reputation lists are designed to provide the operator with categories that inform judgement rather than assign it. Armed with the wealth of context provided, the analyst will have all of the necessary information needed to take actions supporting the goals and values of the organization—both quickly and efficiently in any situation.

Categorization Legend

Each ET Intelligence Rep List entry is given one of the following categories, and is displayed in the list with its corresponding category number:

1,CnC,Malware Command and Control Server
2,Bot,Known Infected Bot
3,Spam,Known Spam Source
4,Drop,Drop site for logs or stolen credentials
5,SpywareCnC,Spyware Reporting Server

6,OnlineGaming,Questionable Gaming Site
7,DriveBySrc,Driveby Source
9,ChatServer,POLICY Chat Server
10,TorNode,POLICY Tor Node
13,Compromised,Known compromised or Hostile
15,P2P,P2P Node
16,Proxy,Proxy Host
17,IPCheck,IP Check Services
19,Utility,Known Good Public Utility
20,DDoSTarget,Target of a DDoS
21,Scanner,Host Performing Scanning
23,Brute_Forcer,SSH or other brute forcer
24,FakeAV,Fake AV and AS Products
25,DynDNS,Domain or IP Related to a Dynamic DNS Entry or Request
26,Undesirable,Undesirable but not illegal
27,AbusedTLD,Abused or free TLD Related
28,SelfSignedSSL,Self Signed SSL or other suspicious encryption
29,Blackhole,Blackhole or Sinkhole systems
30,RemoteAccessService,GoToMyPC and similar remote access services
31,P2PCnC,Distributed CnC Nodes
33,Parking,Domain or SEO Parked
34,VPN,VPN Server
35,EXE_Source, Observed serving executables
37,Mobile_CnC,Known CnC for Mobile specific Family
38,Mobile_Spyware_CnC,Spyware CnC specific to mobile devices
39,Skype_SuperNode,Observed Skype Bootstrap or Supernode
40,Bitcoin_Related,Bitcoin Mining and related
41,DDoSAttacker,DDoS Source

Periodically we add and retire categories from the published lists. To date, we have kept data and reputation attribution mechanisms active, and associated with 40 different categories. Some categories, however, may be retired from publishing because it is more effective to globally white-list trusted sites instead of having them as a category. For others, we have removed categories that might be confusing or not relevant. In each case we may continue to track and maintain the category, but not publish it—this allows us to tweak attribution mechanisms, as well as create categories for R&D purposes.

proofpoint

ET Intelligence Rep List Sampling

Generally, Reputation Score over 50 is reliable. If it is over 100 consider it highly reliable. We don't publish until an item has over 20 points in most categories.

ET Replist File Formats

ET Intelligence provides several different files for your convenience and parsing needs for both Domain and IP objects. While the core information is the same for each file, the fields and format do vary.

Detailed Domain Replist Format (CSV)

Domain Name, Category, Score, First Seen, Last Seen, Ports

143.ns098.com,1,80,2013-01-09,2013-01-17,8080 7070
143.ns529.com,1,45,2013-01-19,2013-01-19,8080
14308.noip1.nl,1,82,2013-01-11,2013-01-16,8000 8003
14308.noip1.nl,35,109,2013-01-12,2013-01-16,443
14335.pqpq.net,1,57,2013-01-08,2013-01-10,80
14339.noip2.nl,1,122,2013-01-15,2013-01-23,8000 8003 9004
c56c30fa24ebee89dfc3d5c80a3077f1.info,1,107,2013-01-14,2013-01-20,
c56c30fa24ebee89dfc3d5c80a3077f1.org,1,107,2013-01-14,2013-01-20,
c5a.shuisumuli.com,1,127,2012-08-11,2013-01-24,53
c5cc591e2980433838dc9b28bccc5b17.co.cz,1,32,2012-12-31,2013-01-06,
c5cc591e2980433838dc9b28bccc5b17.cz.cc,1,32,2012-12-31,2013-01-06,
c5cc591e2980433838dc9b28bccc5b17.info,1,32,2012-12-31,2013-01-06,
ilo.brenz.pl,1,127,2012-02-28,2013-01-24,80 65520
ilo.brenz.pl,9,118,2012-04-11,2013-01-15,80
iloveyouyuyu.3322.org,1,87,2012-12-25,2013-01-17,9999
ilpns.biz,1,85,2013-01-21,2013-01-22,

Simple Domain List Format (CSV)

domain, category, score
leu.su,27,126
tazl.ru,35,74
jma1.biz,9,123
f5v9w.com,29,126
gbcno.com,28,127

Domain List Format (json)


```
{
        "ngrb0ts.co.cc" : {
            "Blackhole" : "73",
            "ChatServer" : "73"
   },
   "euromillions.sd.en.softonic.com" : {
     "SpywareCnC" : "72"
   },
   "laoboer.3322.org" : {
     "CnC" : "92"
   },
   "4273b.perfectchoice1.com" : {
     "Blackhole" : "53"
   },   "53d5e.perfectchoice1.com" : {
     "Blackhole" : "51"
   },
   "theworld-browser.sd.en.softonic.com" : {
     "SpywareCnC" : "77"
   },
   "rp.thebestdownload-manager.com" : {
     "SpywareCnC" : "37"
   },
```


Detailed IP List (CSV)

IP Address, Category, Score, First Seen, Last Seen, Ports
109.71.162.100,1,35,2013-01-15,2013-01-15,1935
109.92.91.247,2,47,2012-12-24,2012-12-25,14554 14803 23717 18398
109.92.94.137,2,72,2013-01-05,2013-01-05,17840 13697 17874
109.93.116.219,1,52,2013-01-06,2013-01-06,21537
109.95.160.155,1,72,2013-01-11,2013-01-11,49741
110.139.65.244,1,87,2013-01-13,2013-01-13,14068 10015

Simple IP List (CSV)

ip, category, score
1.1.1.22,29,120
1.1.1.112,29,120

1.9.98.94,31,77
2.30.1.24,15,42
36.8.96.7,15,100
46.5.0.30,15,57
46.5.16.1,15,40
5.108.1.0,15,40
5.18.62.3,15,47
5.34.8.97,15,52
75.85.7.5,21,50

Background & Legend for each Category:

1, CnC, Malware Command and Control Server
Observed or DGA predicted domains and IPs that are command and control for known Trojans. These listings are specifically criminal, differentiated from spyware and user tracking domains, which are classified in SpywareCnC.

2, Bot, Known Infected Bot
A host observed checking in to a command and control server, or exhibiting clear indications of unwanted and criminal code on the host.

3, Spam, Known Spam Source
We don't track all spam sources, but those observed sending spam or being rejected as blacklisted are included.

4, Drop, Drop site for logs or stolen credentials
Differentiated from CnC servers, but sometimes overlapping. Anywhere we see stolen data or credentials being pushed. Does not include droppers being served or other exe movement.

5, SpywareCnC, Spyware Reporting Server
Servers and domains observed being used to serve or track user activity. Specifically not clearly criminal, but we avoid plain ad-serving sites as much as is possible. Generally these are going to be toolbars, rogue gaming, free screensavers, etc.

6, OnlineGaming, Questionable Gaming Site
Gambling, flash games, and similar that installs a client and report or track user activity. Most of these do not cross the line of criminal, but are differentiated from plain

spyware activity.

7, DriveBySrc, Driveby Source
Kit redirectors, exploit serving, or injected/compromised sites that either have attempted to or will lead to a compromised browser. DriveBySrc is a category representing sites that have been utilized by various exploit kits such as Neosploit or Blackhole, where via some sort of HTML injection, a users browser is redirected to these sites that deliver an exploit by java or some other method.

9, ChatServer, POLICY Chat Server
Observed chat activity, including but not limited to IRC, Jabber, Google Talk, MSN, AIM, ICQ, Baidu, GaduGadu, etc. This is not an indication of hostile activity, only known chat activity. Can be cross-correlated with CnC to help mitigate legitimate IRC networks in use as CnC.

10, TorNode, POLICY Tor Node
Identification of Tor exit nodes and participants seen in the network.

13, Compromised, Known compromised or Hostile
A bit of a catchall category for hosts that are observed hostile including compromised web servers, brute forcers, or otherwise not easily classifiable activity.

15, P2P, P2P Node
Observed clients and sources of generally legitimate file sharing, including traditional bittorrents, limewire/kazaa, qvod, and others.

16, Proxy, Proxy Host
Observed proxy endpoint for http, stun, socks, etc.

17, IPCheck, IP Check Services
IP and geo check services. Generally public services which are very often abused by malware or dyndns activity.

19, Utility, Known Good Public Utility
Known good nets and services such as Google search frontends, Bing, etc…

20, DDoSTarget, Target of a DDoS
Observed DDoS targets by traffic, or observed commands to launch attacks to these nets.

21, Scanner, Host Performing Scanning
Web vulnerability scanning; open relay scanning, network and service recon, and often Nessus or other scanner activity.

23, Brute_Forcer, SSH or other brute forcer
All observed authentication brute forcing, including SSH, imap, VNC, etc.

24,FakeAV, Fake AV and AS Products
Fake antuspyware and av product sites being sold or distributed. Often overlaps with CnC.

25, DynDNS, Domain or IP Related to a Dynamic DNS Entry or Request
Host or domain observed using DynDNS.

26, Undesirable, Undesirable but not illegal
Some hack tool forums, metasploit updates, etc. Not illegal, but of interest on an otherwise controlled network.

27, Abused TLD, Abused or free TLD Related
Activity or DNS related to rogue TLD and GTLDs such as .tk, co.cc, and others. Not always hostile, but of interest. Reserved for rogue registrars…registrars that either can't, or choose not to police their domains, and TLDs that are free or with less accountability. The TLD .su is a perfect example. ICANN can't claim a domain back from a country (Soviet Union) that no longer exists. Malicious activity in this category isn't certain, but typically suspicious.

28, SelfSignedSSL, Self Signed SSL or other suspicious encryption. Self-signed or other invalid SSL certificates in use.

29, Blackhole, Blackhole or Sinkhole systems
Known sinkhole in use by a trusted organization. Will often overlap with CnC.

30, Remote Access Service, GoToMyPC and similar remote access services
Observed but often legitimate remote access services like Kaseya, Gotomypc, Citrix, and others.

31, P2PCnC, Distributed CnC Nodes
Zeus and other families that use P2P as a CnC mechanism. Separated category to handle the volume and transient nature of these hosts.

33, Parking, Domain or SEO Parked
Known parked domain or parking server.

34, VPN, VPN Server. VPN Protocols observed terminating at this address. Possible anonymizing service.

35, EXE_Source, Observed serving of an Executable. Not necessarily hostile, but will often coincide with CnC.

40, Bitcoin_Related is a category based upon observed activity related to P2P Bitcoin mining. The focus here is not a user making transactions with Bitcoins, but actual Bitcoin clients running in the Bitcoin P2P network. This can help identify the rogue use of computing resources in Bitcoin mining operations.

41, DDoSAttacker, Source of DDoS attack traffic.