Cybereason Lab Analysis

Dissecting Domain Generation Algorithms

Eight Real World DGA Variants

cybereason

Even though attackers use various tools to compromise a network, there are core activities that form the foundation of each malicious operation.

ESTABLISHING COMMAND & CONTROL

One essential component is establishing command & control (C&C) communication between the attacker and hacked network. Detecting and blocking the attacker's C&C attempts is a useful approach for shutting down a variety of malicious operations.

A COMMON C&C METHOD

DGAs have quickly become the main method attackers use to remotely communicate with the sophisticated malicious tools they've created. Adversaries have stopped using hard-coded domain lists and IP addresses, which are useless once blocked.

DGAs by comparison are easy to implement, difficult to block, and may be impossible to predict in advance and can be quickly modified if the previously used algorithm becomes known. A DGA typically has three components:

- A time-sensitive "seed"
- A domain "body" generator that uses this seed
- A set of top-level domains (TLDs)

Often, the seed is simply the current date in some standard format. The domain body generator is the main part of a DGA, and can basically be anything—a random string of characters, concatenation of random words, a constant part followed by a changing suffix, and so on. The set of TLDs, however, must contain real-world values that determine under which Web entities the generated domains are registered.

TRADITIONAL METHODS FAIL TO DETECT AND BLOCK DGAS

Even when a certain DGA is known (for example, by reverse engineering a malware sample), it's still difficult—or even impossible—to effectively block it. First, there is the sheer number of possible domains that can be generated. Gameover Zeus, for example, generates 1,000 domains every day. This amounts to 365,000 domains that need to be generated in advance and blocked, which would strain on firewalls and other network-filtering solutions. *And that's just for one, single DGA for a year.*

While the amount of domains that need to be blocked is problematic and some registrars are very uncooperative with law enforcement agencies, the seed can be the real issue. The date can be predicted indefinitely, but it's not the only value that can constantly change. The DGA can use, for example, the daily trending hashtag on Twitter, the current exchange rate of the U.S. dollar to the Japanese yen, the temperature in Rio de Janeiro and basically any value that can be reliably obtained via the Internet by both the malware and its operator. Predicting these values in advance is *of course* impossible, and most filtering solutions do not support dynamic generation of domains to block.

Law enforcement and government agencies from across the world, including the FBI, have attempted to take control over these domains at the source by going after the registrars, as seen in Operation Tovar. But even government organizations have limits to their power.

In the case of Operation Tovar, the FBI, was unable to take over domains registered under the Russian TLD. And accessing the TLD name servers requires spending huge amounts of time and effort to obtain a warrant, which had to be renewed every six months.

Some researchers have tried to detect randomly-generated domains by their patterns, without knowing the algorithm in advance, and had some moderate success. The problem with this approach is two-fold. First, there is a strong chance for false positives, as many legitimate websites use load-balancing servers and other strange looking domain names, and the tiny ratio of DGA traffic compared to regular traffic makes false positives almost a certainty.

Secondly, DGA body generators can take many forms and aren't necessarily a long string of random characters (see the following examples, detected in Cybereason customer environments). These domains can't be detected using traditional security methods.

CYBEREASON DETECTS NEW DGA VARIANTS

The Cybereason platform uses a unique approach for DGA detection, looking for behaviors associated with DGAs instead of looking for DGA variants. Using this approach, Cybereason Labs has identified new DGA variants in customer environments. We hereby describe these variants and the outcome of our investigation.

EIGHT DGA VARIANTS DISSECTED BY CYBEREASON

Name	"Body"	Used TLDs
Unknown Russian	Seven random letters, prefixed by a subdomain. The subdomains seen include "flag" followed by a number and "pop"	.ru .com
Word-based	Two English words chosen and concatenated from a list of 384 (0x180) words. At least three variants were seen, each with a different words list	.net
Necurs	8-20 random letters	.ac (Ascension Isl) .bz (Belize) .cc (Cocos Islands) .cm (Cameroon) .co (Colombia) .cx (Christmas Isl) .de (Germany) .eu (European Union) .ga (Gabon) .im (Isle of Man) .in (India) .ir (Ireland) .jp (Japan) .ki (Kiribati) .kz (Kazakhstan) .la (Laos) .me (Montenegro) .mn (Mongolia) .ms (Montserrat) .mu (Mauritius) .mx (Mexico) .nf (Norfolk Isl) .nu (Niue) .pw (Palau) .ru (Russia) .sc (Seychelles) .sh (Saint Helena) .so (Somalia) .su (Soviet Union) .sx (Sint Maarten) .tj (Tajikistan) .to (Tonga) .tv (Tuvalu) .tw (Taiwan) .ug (Uganda) .us (USA) .org .pro .net .com .bit .biz .xxx
Dridex	Random English words concatenated together, sometimes offsetted or broken	.me (Montenegro) .mn (Mongolia)
Angler exploit-kit	11-19 random letters and digits, though letters are much more likely	.com
Unknown DWORD-based	A random DWORD value, in its textual hexadecimal representation	.com .net .info
Pykspa	5-11 random letters	.com, .net, .org, .info .cc (Cocos Islands)
Unknown Punycode-like	Long Punycode-like string (starts with "www.xn--") with a constant part and a random part which is six random digits	.com

1. UNKNOWN RUSSIAN DGA

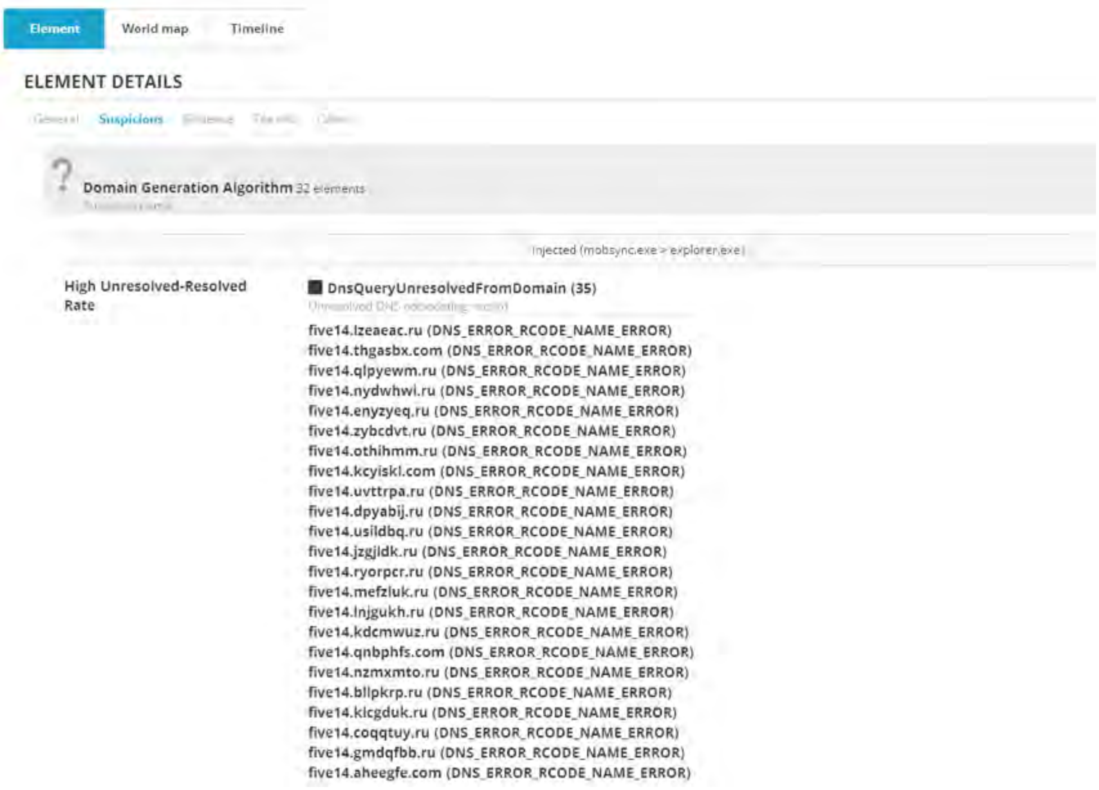
What is it?

A Russian malware using an unknown DGA.

Mechanism of Action

Each day, 35 domains are generated by randomly selecting seven letters, suffixing them with either the .ru or the .com top-level domains and prefixing them with the word “five” followed by a number. This is unusual, since most DGAs do not bother with any subdomain, and perhaps this is why it’s done in this case, to give some semblance of legitimacy. The malicious code usually injects itself into explorer.exe to evade detection.

The following is a screenshot of the detected DGA in a customer environment:



ELEMENT DETAILS

Domain Generation Algorithm 32 elements

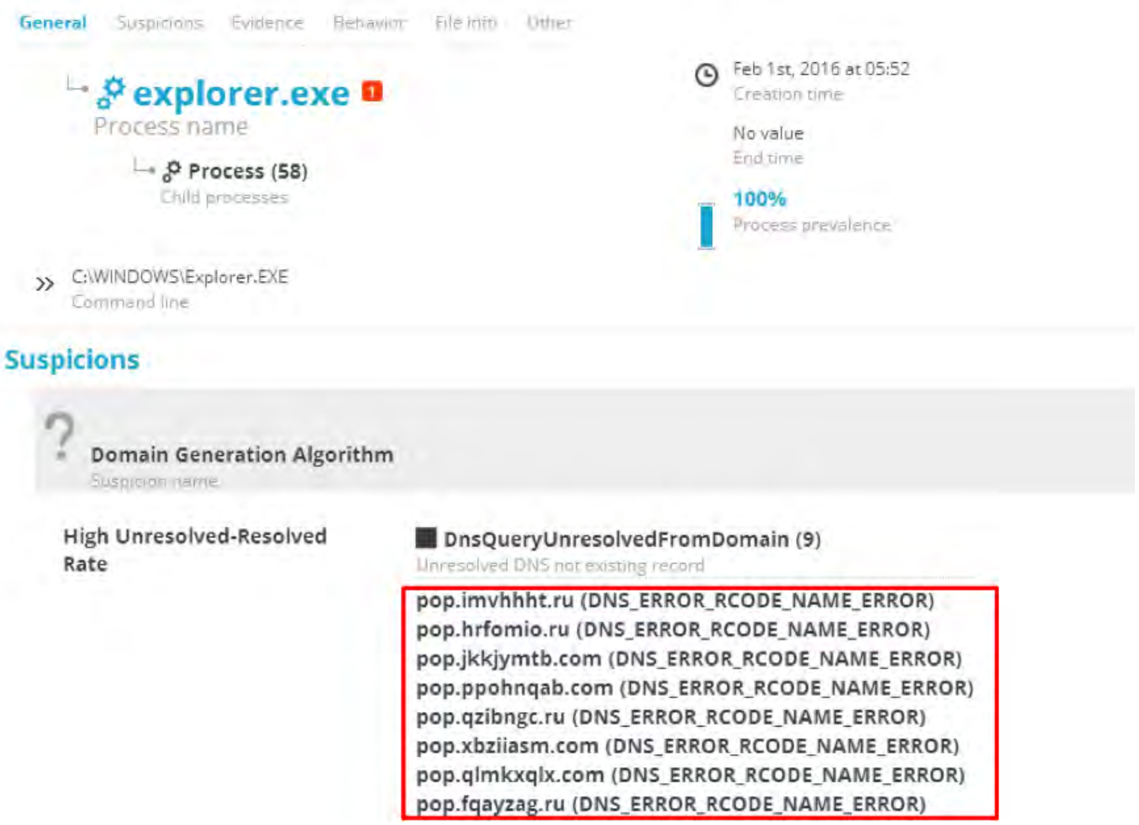
Injected (mobsync.exe > explorer.exe)

High Unresolved-Resolved Rate

■ DnsQueryUnresolvedFromDomain (35)

```
five14.lzeaeac.ru (DNS_ERROR_RCODE_NAME_ERROR)
five14.thgasbx.com (DNS_ERROR_RCODE_NAME_ERROR)
five14.qlpyewm.ru (DNS_ERROR_RCODE_NAME_ERROR)
five14.nydwhwi.ru (DNS_ERROR_RCODE_NAME_ERROR)
five14.enyzyeq.ru (DNS_ERROR_RCODE_NAME_ERROR)
five14.zybcdvt.ru (DNS_ERROR_RCODE_NAME_ERROR)
five14.othihmm.ru (DNS_ERROR_RCODE_NAME_ERROR)
five14.kcyiskl.com (DNS_ERROR_RCODE_NAME_ERROR)
five14.uvttrpa.ru (DNS_ERROR_RCODE_NAME_ERROR)
five14.dpyabij.ru (DNS_ERROR_RCODE_NAME_ERROR)
five14.usildbq.ru (DNS_ERROR_RCODE_NAME_ERROR)
five14.jzgjldk.ru (DNS_ERROR_RCODE_NAME_ERROR)
five14.ryorpcr.ru (DNS_ERROR_RCODE_NAME_ERROR)
five14.mefzluk.ru (DNS_ERROR_RCODE_NAME_ERROR)
five14.lnjgukh.ru (DNS_ERROR_RCODE_NAME_ERROR)
five14.kdcmwuz.ru (DNS_ERROR_RCODE_NAME_ERROR)
five14.qnbphfs.com (DNS_ERROR_RCODE_NAME_ERROR)
five14.nzmxmto.ru (DNS_ERROR_RCODE_NAME_ERROR)
five14.bllpkrp.ru (DNS_ERROR_RCODE_NAME_ERROR)
five14.klcgduk.ru (DNS_ERROR_RCODE_NAME_ERROR)
five14.coqqtuy.ru (DNS_ERROR_RCODE_NAME_ERROR)
five14.gmdqfbb.ru (DNS_ERROR_RCODE_NAME_ERROR)
five14.aheegfe.com (DNS_ERROR_RCODE_NAME_ERROR)
```

Another very similar variant was detected on another machine in the same organization, but in this case only nine domains are generated, and the prefix is the constant word “pop”: see the screenshot below.



explorer.exe 1
Process name

Process (58)
Child processes

Feb 1st, 2016 at 05:52
Creation time

No value
End time

100%
Process prevalence

C:\WINDOWS\Explorer.EXE
Command line

Suspicions

Domain Generation Algorithm
Suspicion name

High Unresolved-Resolved Rate

DnsQueryUnresolvedFromDomain (9)
Unresolved DNS not existing record

```
pop.imvhhht.ru (DNS_ERROR_RCODE_NAME_ERROR)
pop.hrfomio.ru (DNS_ERROR_RCODE_NAME_ERROR)
pop.jkkjymtb.com (DNS_ERROR_RCODE_NAME_ERROR)
pop.ppohnqab.com (DNS_ERROR_RCODE_NAME_ERROR)
pop.qzibngc.ru (DNS_ERROR_RCODE_NAME_ERROR)
pop.xbziiasm.com (DNS_ERROR_RCODE_NAME_ERROR)
pop.qlmkxqlx.com (DNS_ERROR_RCODE_NAME_ERROR)
pop.fqayzag.ru (DNS_ERROR_RCODE_NAME_ERROR)
```

2. WORD-BASED DGA

What is it?

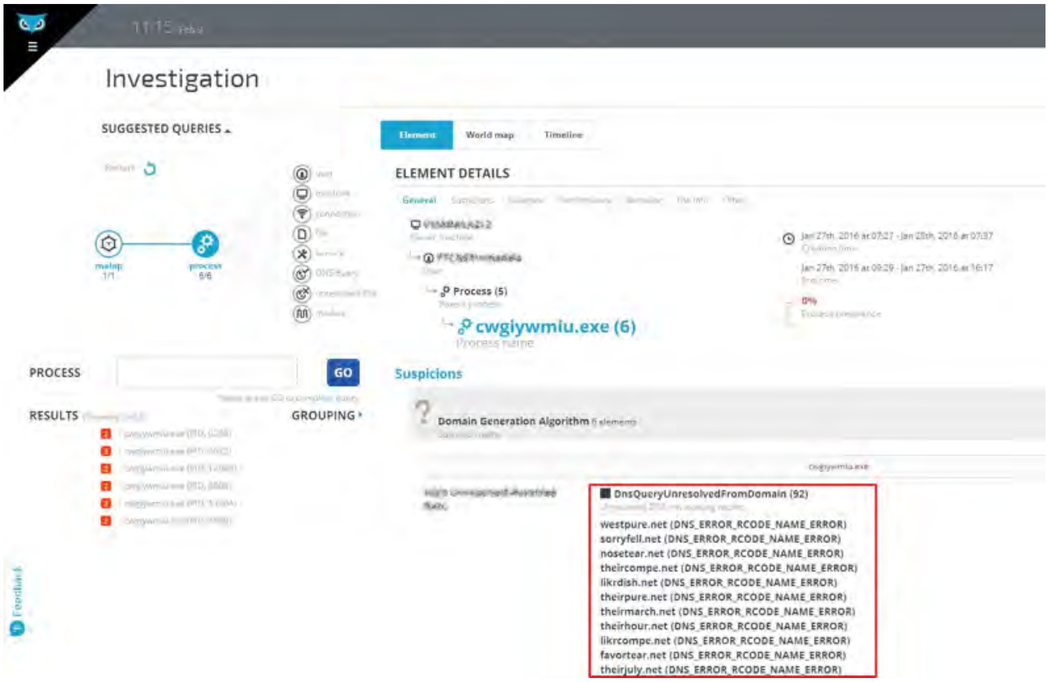
This seems to be the same DGA as an unnamed malware analyzed by Crowdstrike in 2013.

Mechanism of Action

Domains are generated by randomly choosing two English words from a hard-coded list and concatenating them together under the .net top-level domain.

With a list of 384 (0x180) words, this comes to approximately 150,000 possible combinations. The difficulty of detecting this simple algorithm is that the domains do not seem to be randomly generated, and the commonly used words may appear in many legitimate domain names.

However, it seems that this malware may use several different word lists. By simply replacing this list, the attackers can completely evade detection of the old algorithm. This includes words such as “july,” “table,” “city,” “favor,” “dish” and others.



Process (5)

cwgiywmiu.exe (6)

Process name

Jan 27th, 2016 at 07:27 - Jan 28th, 2016 at 07:37

Jan 27th, 2016 at 09:29 - Jan 27th, 2016 at 16:17

0%

Suspicions

Domain Generation Algorithm

DnsQueryUnresolvedFromDomain (92)

```
westpure.net (DNS_ERROR_RCODE_NAME_ERROR)
sorryfell.net (DNS_ERROR_RCODE_NAME_ERROR)
nosetear.net (DNS_ERROR_RCODE_NAME_ERROR)
theircompe.net (DNS_ERROR_RCODE_NAME_ERROR)
likrdish.net (DNS_ERROR_RCODE_NAME_ERROR)
theirpure.net (DNS_ERROR_RCODE_NAME_ERROR)
theirmarch.net (DNS_ERROR_RCODE_NAME_ERROR)
theirhour.net (DNS_ERROR_RCODE_NAME_ERROR)
likrcompe.net (DNS_ERROR_RCODE_NAME_ERROR)
favortear.net (DNS_ERROR_RCODE_NAME_ERROR)
theirjuly.net (DNS_ERROR_RCODE_NAME_ERROR)
```

3. NECURS DGA

What is it?

A nasty backdoor malware.

Mechanism of Action

Randomly-generated strings of eight to 20 characters in length suffixed with one of the many multiple exotic top-level-domains, such as .ga (Gabon), .im (Isle of Man) and .sc (Seychelles). This makes it harder for law enforcement agencies to take down these domains.

The following are sample domains detected by Cybereason:

QUJFVNN.TO

CRWKBMX.TW

FFJVGCIF.MN

JNHUTIIV.TV

YJENASPDAN.IN

AODXYTMXLB.COM

OLKQXMAEUIWYX.XXX

BPWENCSDVRJXJI.PRO

SNDXKVGEFQQCFCTJ.PW

FQOXIBDVBYCNSAPPXC.NU

DOOKMSWEMEXLTBSUAL.SU

OPCALVWELIIISUHXARKR.BIT

4. DRIDEX DGA

What is it?

Dridex is a strain of banking malware that leverages macros in Microsoft Office to infect systems.

What does it do?

Concatenates English words and parts of words chosen in random from a small list, suffixed by the .mn (Mongolia) and .me (Montenegro) top-level domains.

Unlike the malware described by Crowdstrike, in this variant the words are often broken, shifted and padded with random characters, significantly increasing the number of possible combinations and making detection much harder.

The following are sample domains detected by Cybereason:

CLIENTALALAXP.MN

CLIENTALNOTHING.ME

USERALCLICLIENT.ME

AGENTCLIENTCLIENT.ME

JSCJSCAXPCLIALLOW.ME

JSCCLIENTAGENTDISA.ME

DISAALALLOWDISALLOW.ME

ALLOWCLIENTAXPALAGENT.ME

CLIAGENTDISALLOWALLOW.ME

CLIALJSCNOTJCLIENTCLI.ME

5. ANGLER EXPLOIT-KIT DGA

What is it?

A widely used exploit kit.

Mechanism of Action

While the domains generated by this DGA were previously connected to the Angler exploit kit, they do not appear to be generated by the known DGA, and may be a new variant. The algorithms strings randomly chosen characters and digits under the .com top-level domain.

The following are sample domains detected by Cybereason:

V6PNSC80LL.COM

B9U5R3RJMPP.COM

YM5R99EX5Q8.COM

MBSIGLGFQIH2.COM

GSJZNQCOHIKO.COM

VEG2671WMX88.COM

DLNOYYVQSOZHH.COM

BFZFLQEJOHXMQ.COM

AJFSZWOMNHDFCYY.COM

EXAGQLXTMOPSFT8.COM

FWOGZPAGLGOVLIMY.COM

JVRRMMKYEJDEYLCQ.COM

LKLHJONIUDKKHCWO.COM

CADDBSGSCNYDZOH5F.COM

CEUNNFOHGWJYAUA9H.COM

NQZHTFHRMYMTVBQJE.COM

OVLREWGRHHVAJBOTX.COM

OTPWFJOKPOZOOMNK2O.COM

CNEISZDKHZEKQEUBUT.COM

EMUXMJDBTNWCQRFN0G.COM

OWASALWIGURWYVNNPV.COM

PMNYPARTDBVYHCZDJS.COM

6. UNKNOWN DWORD-BASED DGA

What is it?

Unknown malware injected inside svchost.exe.

Mechanism of Action

The DGA of this malware seems to generate a random DWORD (a 32-bit integer, with a maximum value of approximately 4 million) then converts it to its hexadecimal format and suffix the result with either the .com, .net or .info TLDs. This DGA has not been disclosed before online, making it unique. This appears to be the first time this DGA has been discussed, making it a new discovery. There aren't any references to this DGA online.

The following are sample domains detected by Cybereason:

04F645A5.COM

15AF64DD.INFO

2518F789.COM

2AF14345.INFO

39E076F7.INFO

3E0CA533.NET

428BF932.COM

4E32A34D.INFO

59D1FC99.NET

6CC69779.NET

78E05B8B.NET

7C7F4A6E.COM

974381F6.NET

9890D1FA.INFO

B06CB4A1.NET

C50A4E79.COM

D3270391.NET

D41FCED5.NET

DB0311C2.INFO

F7A1F33B.INFO

7. PYKSPA DGA

What is it?

A stealthy botnet that uses Skype.

Mechanism of Action

Randomly generated strings of characters of varying lengths suffixed with the **.com**, **.net**, **.org**, **.info** and **.cc** (Cocos Islands) top-level domains.

The following are sample domains detected by Cybereason:

CFAOBN.COM

QQQCLQFO.CC

HYEHGNR.NET

SWGDOM.INFO

FVGCWBMX.ORG

HGZGHCYJ.NET

USCNXQES.ORG

GVMVMEQD.NET

LEZBMAH.INFO

IJDVHZYQS.NET

JUKIULBI.INFO

ASOOGYCRE.NET

EYHKHBTPYG.NET

PDOYVFIGFG.NET

IXLMYGMNDWJ.CC

CIJFTOCHT.INFO

ATRAEAUZWUJ.ORG

YRWRWYZSQL.INFO

XGUGUSBBOK.INFO

RZXFYIIXJOE.INFO

8. UNKNOWN PUNYCODE-LIKE DGA

What is it?

Unknown malware that generates domains that look like Punycode—non-English domain names—but are, in fact, randomly-generated gibberish.

The following are sample domains detected by Cybereason:

WWW.XN--ZALGO003446-SJGB60AIGHL2I8JC3B0A2A97FTBLL0CZA.COM

WWW.XN--ZALGO012841-SJGB60AIGHL2I8JC3B0A2A97FTBLL0CZA.COM

WWW.XN--ZALGO029243-SJGB60AIGHL2I8JC3B0A2A97FTBLL0CZA.COM

WWW.XN--ZALGO075952-SJGB60AIGHL2I8JC3B0A2A97FTBLL0CZA.COM

SUMMARY

Instead of trying to fight each DGA variant separately, a nearly impossible task, Cybereason concentrates on the ripples it leaves. We detect the technique, not the variant. And since no legitimate process will ever use DGA, just detecting it incriminates the process as malicious.

This is a part of Cybereason's "Aikido approach": Using the opponent's strength against him/her. The more adversaries try to hide, the more suspicious they appear.

About the Author

Uri Sternfeld

Research Team Leader, Cybereason Labs

Uri is Team Leader of the Research Team at Cybereason Labs.

He has more than 15 years of experience in software design, cybersecurity and technology research. His areas of focus are cyber-forensics, reverse engineering and data mining automation.

cybereason

Cybereason was founded in 2012 by a team of ex-military cyber security experts to revolutionize detection and response to cyber attacks. The Cybereason Malop Hunting Engine identifies signature and non-signature based attacks using big data, behavioral analytics, and machine learning. The Incident Response console provides security teams with an at-your-fingertip-view of the complete attack story, including the attack timeline, root cause, adversarial activity and tools, inbound and outbound communication used by the hackers, as well as affected endpoints and users. This eliminates the need for manual investigation and radically reduces response time for security teams. The platform is available as an on premise solution or a cloud-based service. Cybereason is privately held and headquartered in Boston, MA with offices in Tel Aviv, Israel and Tokyo, Japan.