Reverse Engineering Dridex and Automating IOC Extraction

S appgate.com/blog/reverse-engineering-dridex-and-automating-ioc-extraction



Dridex^[1] is a major banking trojan that appeared somewhere around 2011, continually evolving ever since.

The APT (Advanced Persistence Threat) known as TA505^[2] is associated to Dridex, as well as with other infamous malware such as TrickBot and Locky ransomware.

Once installed, Dridex can download additional files to provide more functionality to the trojan. Simply put, there are four main components:

Downloader	Infected Microsoft Office documents that downloads or drops Dridex's Loader
Loader	Responsible for downloading and installing the main module.
Bot	This is the main module, which contains most of Dridex's functionalities.
Modules	Additional files that provides more functionalities, downloaded by the Loader.

In the last few days, our team detected recent Dridex samples through our live hunting process and, after analyzing them, we've decided to publish an analysis of the main features implemented by this threat that makes detection and analysis difficult.

Therefore, in this post, we will focus on the second layer (Loader), showing the technical details regarding:

- 1. Unpacking;
- 2. Dynamic API Calls;
- 3. Encrypted Strings;
- 4. C2 Network Communication.

In addition, we are publishing a tool that statically extracts IOCs from the latest Dridex binaries, as we believe that this can enable organizations to reduce analysis time spent during incidents or prevent the malware family altogether.

1. Unpacking

Unpacking is an important step because Dridex doesn't write the unpacked payload to disk, instead, the custom packer loads it directly to memory and there it stays. First, to analyze and extract the IOCs, we must get the unpacked sample, and this is not a difficult task.

You can easily execute the packed sample and then run the amazing PE-sieve[3] tool, from <u>hasherezade</u>, which will extract the payload from memory. However, if you are curious like us and wants to understand how it works, the first step in the unpacking process is the

allocation of an encrypted content into memory:

100040CF	. 897D B8	MOV DWORD PTR SS:[EBP-48],EDI
100040D2	. E8 73FBFFFF	CALL whoami.10003C4A
100040D7	. 8B45 BC	MOV EAX, DWORD PTR SS:[EBP-44]
100040DA	. 0345 EC	ADD EAX, DWORD PTR SS: [EBP-14]
10003C4A=whoami.10003C4A		
Address Hex dump		ASCII
00240000 AF 9D 89 53 9B CC F2	E6 AB B6 64 96	BB 43 17 21 ⁻ ‰S>Ìòæ«¶d-»C-[!
00240010 2F 1E D3 75 2E 9B E0	C4 DC 7D 76 53	2E 00 49 FD ∕Óu.>àÄÜ}vSIý
00240020 E5 95 D2 54 C0 8F 08	B7 A7 2D 94 FC	04 A0 8F F4 å•ÒTÀ• ◘ •§-"üJ ô
00240030 0C 40 35 77 11 B6 E4	C9 9D 2A A4 37	A2 65 DF BF .@5w ∢ ¶äÉ*¤7¢eß¿
00240040 3A ED FE 62 F9 2E 34	FC D3 76 3C 84	4B D3 53 A7 :íþbù.4üÓv<"KÓS§
00240050 1A 41 72 8F 08 EA AA	FØ DC C5 F2 4F	D4 8F C7 2C →Ar•DêªðÜÅòOÔÇ,

Encrypted DLL in Memory.

This data is an encrypted small DLL that is responsible for unpacking the Dridex loader. Also, we noticed that the magic bytes (MZ) for the MS-DOS header was not present, probably to avoid automatic filetype identifications.

Address	Hex	c di	ump														ASCII
00340000	A 4	6C	ØA	7F	EØ	AA	00	00	04	00	00	00	FF	FF	00	00	¤l.∏àªJÿÿ
00340010	B8	00	00	00	00	00	00	00	40	00	00	00	00	00	00	00	,@
00340020	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	
00340030	00	00	00	00	00	00	00	00	00	00	00	00	90	71	85	EC	qì
00340040	ØE	1F	BA	ØE	00	B4	Ø 9	CD	21	B 8	01	4C	CD	21	54	68	∄º∄.´.Í!, LÍ!Th
00340050	69	73	20	70	72	6F	67	72	61	6D	20	63	61	6E	6E	6F	is program canno
00340060	74	20	62	65	20	72	75	6E	20	69	6E	20	44	4F	53	20	t be run in DOS
00340070	6D	6F	64	65	2E	ØD	ØD	ØA	24	00	00	00	00	00	00	00	mode\$

Decrypted DLL in Memory.

After the decryption process, the main executable transfers the execution to the allocated DLL, which unpacks the Dridex Loader and then replaces the main executable code with the payload's code.

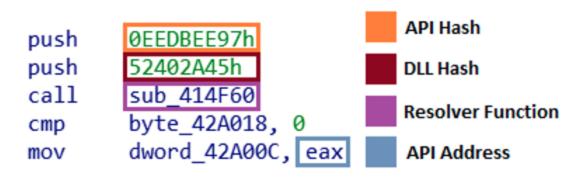
	D	eci	ryp	te	d D	DLL													
00341157									E8 44070000					CALL 003418A0					
0034115C								31	C Ø				Х	OR	EAX	,EAX			
0034115E								8B	4D	E8			M	OV	ECX	, DWOR	D PTR SS:[EBP-18]		
00344464										24					CDV				
Address	Hex dump																ASCII		
00360000	4D	5A	D8	66	80	1D	00	00	04	00	00	00	FF	FF	00	00	MZØf€」ÿÿ		
00360010	B8	00	00	00	00	00	00	00	40	00	00	00	00	00	00	00	,@		
00360020	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00			
00360030	00	00	00	00	00	00	00	00	00	00	00	00	EØ	00	00	00	à		
00360040	ØE	1F	BA	ØE	00	B4	<u>09</u>	CD	21	B 8	01	4 C	CD	21	54	68	µºµ.´.Í!, LÍ!Th		
00360050	69	73	20	70	72	6F	67	72	61	6D	20	63	61	6E	6E	6F	is program canno		
00360060	74	20	62	65	20	72	75	6E	20	69	6E	20	44	4F	53	20	t be run in DOS		
00360070	6D	6F	64	65	2E	ØD	ØD	ØA	24	00	00	00	00	00	00	00	mode\$		
							Dı	ide	ex	Pa	ylo	ad							

Dridex Payload in Memory.

2. Dynamic API Calls

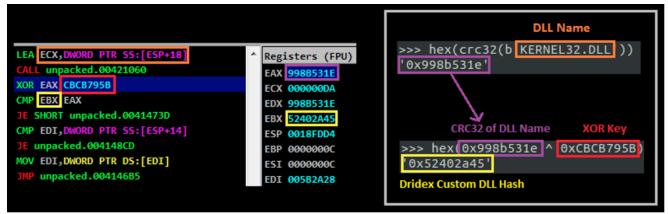
Dridex doesn't have an import table like regular PE files. Instead, all the API calls are dynamically resolved by the malware using a custom technique to avoid detection by APIs and to make reverse engineering more difficult.

To resolve an API, it calls a "resolver" function passing two parameters, both custom hashes. The first one is a hash for the DLL name that contains the API, and the second one is the hash of the API name that Dridex needs to resolve.



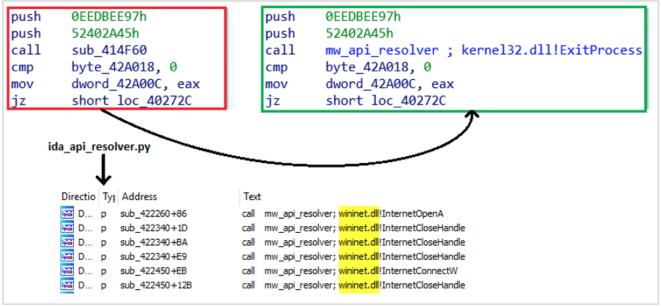
The way the API is resolved isn't trivial, but in summary, if the API wasn't already resolved by the malware, it parses the DLL linked list, located in the process' PEB (Process Environment

Block), generating a CRC32 hash for each DLL and API name and compares that with the ones that was passed into the function. Aside from the CRC32 hash, the value is also "xored" with a custom key, making the values unique on each Dridex sample.



Dridex Searching for DLL Using Custom Hash.

Therefore, in the example above, the first hash stands for "kernel32.dll". Using this same logic, <u>we created a small IDA script</u> which resolves API calls automatically and inserts a comment where the function is called, so we can easily search where in the code certain DLLs or APIs are being used.



Block with Encrypted Strings.

Each chunk contains one or more strings that are encrypted with RC4, where the key is the first 40 bytes (in little-endian format) of each block.

Address	Hex dump	Decryption Key = First 40 bytes
004254B0 004254C0	0E EA 9C AA 00 50 1F 87 5E C3 A0 EA 4A 6D A1 BA E8 4A D4 71 EC B2 92 51 AA CB FE 1B 0A 29 0D 4E 1D 6D 88 F3 74 B9 53 29 1D DA F2 82 97 34 6A F9 B8 F1 19 48 CD SF EE EA 26 8D AA E5 CA FD 92 5D	key = data[:40][::-1] Encrypted Data = Remaining Bytes
004254F0	F6 B8 DE 24 12 C6 00 9A 98 FD 15 2C 80 02 C2 C5 C4 7B 6C 3A 74 BD 64 D7 43 65 34 43 D4 B0 45 CB CE E3 9B D7 FD BF C7 5C D6 CF 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 <	enc_data = data[40:]
>>> AR	d Data Using RC4 C4.new(key).decrypt(enc_data) WARE/TrendMicro/Vizor\x00\\VizorUniclie	entLibrary.dll\x00ProductPath\x00\x00'

Decrypting Dridex Strings.

Aside from the encrypted strings, we also found plain text strings in this recent Loader that were not present in older versions, which can be used for identification (Yara rule):

- bot
- vnc
- socks
- uacme
- list
- cve-2015-0057 (mod5)
- TrendMicro (mod9)
- NetChecker (mod10)
- rep: DIILoaded (dmod5)
- rep: DIIStarted (dmod6)
- rep: NetGood (dmod7)
- rep: NetFail (dmod8)
- rep: NetPart (dmod9)
- rep: StartedInLo (dmod10)
- rep: StartedInHi (dmod11)

4. C2 Network Communication

The loader is responsible for the initial C2 communication and for downloading additional files, such as the bot and the modules. Analyzing the function that does such communication, we could see that the command sent to the C2 is passed as a parameter, and it's the CRC32 hash of the command string.

mov lea	ecx, [ebp+var		mov lea	edx, 18F8C844h ; ecx, [ebp+var_10]	"list" r	equest
call	mw_c2_request	edx, 69BE7CEEh		<pre>mw_c2_request request</pre>		
	lea call	ecx, [ebp+var_E4 mw_c2_request	.]			

Function for C2 Communication.

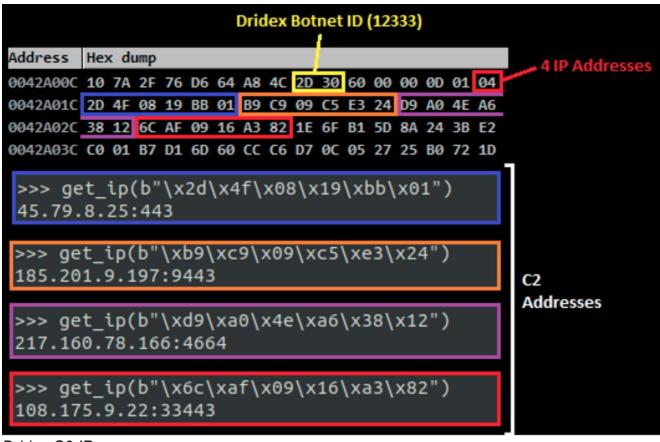
Curiously though, that despite using the hashes of the commands in the communication, this specific sample has the same commands in the strings, as mentioned above, such as the "list" that requests a list of IPs to connect or the "bot" that is the command to download the next stage.

Before doing the request, the botnet ID and the C2 addresses are parsed.

00405090	> 66:8B46 04	MOV AX, WORD PTR DS:[ESI+4] ESI = IP
00405094	. ØFB7C8	MOVZX ECX,AX ESI + 4 = Port
00405097	. 8B06	MOV EAX, DWORD PTR DS:[ESI]
00405099	. 894424 14	MOV DWORD PTR SS:[ESP+14],EAX
0040509D	. 8D4424 0C	LEA EAX, DWORD PTR SS:[ESP+C]
004050A1	. 66:894C24 18	MOV WORD PTR SS:[ESP+18],CX
004050A6	. 8D4C24 14	LEA ECX,DWORD PTR SS:[ESP+14]
004050AA	. 50	PUSH EAX
004050AB	. C74424 20 0000	MOV_DWORD_PTR_SS:[ESP+20],0
004050B3	. E8 18BF0100	CALL unpacked.00420FD0 Transform bytes to text
004050B8	. 8B0D D4A34200	MOV ECX,DWORD PTR DS:[42A3D4]
004050BE	. 8BD0	MOV EDX,EAX
004050C0	. 83C1 04	ADD ECX,4
004050C3	. FF31	PUSH DWORD PTR DS:[ECX]
004050C5	. FF32	PUSH DWORD PTR DS:[EDX]
004050C7	. E8 948F0000	CALL unpacked.0040E060
004050CC	. 8D4C24 0C	LEA ECX,DWORD PTR SS:[ESP+C]
004050D0	. E8 2B7A0000	CALL unpacked.0040CB00
004050D5	. 0FB605 1BA0420	MOVZX EAX, BYTE PTR DS: [42A01B] Pointer to next address
004050DC	. 8D76 06	LEA ESI, DWORD PTR DS:[ESI+6] 4 bytes (IP) + 2 bytes (Port)
004050DF	. 47	INC EDI
004050E0	. 38F8	CMP EDI,EAX
004050E2	.^7C AC	JL SHORT unpacked.00405090
004050E4	> A1 D4A34200	MOV EAX, DWORD PTR DS: [42A3D4]
004050E9	> 5F	POP EDI
Stack DS:[0	L 0018FB3C]=008936F0	0, (ASCII "185.201.9.197:9443")
static ps.[c		C2 Address

Dridex Parsing Command & Control IPs

This information is stored in the PE ".data" section, represented in bytes, along with the Dridex botnet ID.



Dridex C2 IPs.

Once these addresses are parsed, Dridex then sends a POST request to one of the C2 addresses and, bypassing the SSL, we could see that the data is encrypted. If the server doesn't respond as expected, Dridex continues to send the same content to the other IPs.

```
1 POST / HTTP/1.1
 2 Cache-Control: no-cache
 3 Host: 45.79.8.25
 4 Content-Length: 1935
 5 Connection: close
 7 ôX qÖ6(~òðj"!ol Ô9ÉîNvQDÉðò(P%A±6¦Ý×$\$\úÈÀE'jï´0"xqú"ãÁÔ¶Ú|Û¥f>\
 8 XÒÚeH³UHTQþUÜmÇÃ, ìÅAiú07. ŐäÿG⅔¦%£+¬íYwbäWþ²#&øÓFÅ*;Ó⅔
 9 = \pm = |z - \dot{A}d \times D
10 V¢%iï¿åhËßtaOã¥
11 îįĺÇI±²æ8woa
12 ×GÌóÌtÜ~zÍc&) ¶hElTD"_ÙÌ8ùÙKAe*)i_ÑrØ*RHI]çO×N,ÏøþhÏZFÇ6
13 #è{c§è'±b¶(½]'÷ôL}ñr&
14 ô¬zÖh" ²ÆJyæñ¢àIV#"Öà2½"wzýakL»I±Ü, ²Zôjg`4e`£ÉÝÓVP¢2âIí3°îÙ¡AÚ#á
   ×o@áÇõAÇ?C¢üqZY#fÙ$Ò¥jÅcÈÓÇ{£¢éypÌ4Úð+~Ï@{$¤{m¬kÒøLQZ:.º»°«Ñ·¾éø
15 :
16 exer >seôu?. ~íAjwx 8H a²Àêxs /T.*h býæh«Ő2óÁÊe´Ê58?ïØô;õac(ÅÛáI)
17 éyZő\⁻C²tiEiîRlx#Ôz-tþấú¥"8àÉîkìFɬxL©X>éUvAïí,§æYUŰbà Ö=ÚØ\ÌÒÅMÓ§
18 eùvÉhŐCH6ßß]l7wuòICd^<xÄZã'W$ü'D+/5¶ñ±8Á~_TýYWő7DÑpæ';p.GCE'Ò&4«</p>
   `ó7yÍXùÚ3ÊÆ.äÀ3qî^°Ãëiì7ĭwC|Ep,ܺ∼l-czøRÂcó3ëÙr#ôµ3m#(zÂT¦dä2ÀñÂ
19 [,{cèõ=RòfÖeß Zö¦r<80üÀB*Páí _ `äaðDæS&jh#OKpe±÷ hãqÍOùûn øËðx6l
   Å$þ/Å@õÞ#XDPËÒ\iDÎ'n§PÿAP±?CÝyÒBã§NÄ<ÃöêzÎiZla\W*,åÍ,qÍl(Ô@DièR×
20 F-5L黵§&'jöå?hÒSmm)ÚiÆ!u§÷Laù8/ð-çOEøU%ÆåàÑÜ yIjy¿åê$ijê
```

Dridex POST Request to C2.

After analyzing the function, we found that the malware uses the first 4 bytes as a checksum for the encrypted bytes, with CRC32 hash.



Encrypted Network Data Checksum

If the checksum matches, the data can be decrypted correctly. The algorithm used by Dridex is RC4 and the encryption/decryption key is stored among Dridex decrypted strings.

Decryption Key Stored Among Encrypted Strings
>>> key = b'xrAuVcgsoW0BBPhAH5w5aQ1Q2UuZQidMhZYugaYvCPvgttsD9jQkM'
>>> ARC4.new(key).decrypt(post_data[4:]) RC4 Decryption
<pre>b'0WIN-U2809HMVU7F_3e2135e6969588478c7f6c3967284c546b8a9cb0d4ee755f615d82f4f124b54a .00 beta (x64) (17.00 beta);Google Chrome (85.0.4183.102);Google Update Helper (1.3 osoft Visual C++ 2008 Redistributable - x64 9.0.30729.6161 (9.0.30729.6161);Microsc (9.0.30729.6161);Process Hacker 2.39 (r124) (2.39.0.124);Python 2.6 (2.6.150);Pythc uest Tools 0.141 (0.141);VMware Tools (10.1.6.5214329);Starting path: \x00\x00\x058 n\\AppData\\Roaming\nCommonProgramFiles(x86)=C:\\Program Files (x86)\\Common Files\ s\nCommonProgramW6432=C:\\Program Files\\Common Files\nCOMPUTERNAME=WIN-U2809HMVU7F CK=N0\nHOMEDRIVE=C:\\HINDEPATH=\\Users\\admin\nLOCALAPPDATA=C:\\Users\\admin\\AppDat CESSORS=2\nOS=Windows_NT\nPath=C:\\Windows\\system32;C:\\Windows;C:\\Windows\\Syste \\\nPATHEXT=.COM;.EXE;.BAT;.CMD;.VBS;.VBE;.JS;.JSE;.WSF;.WSH;.MSC\nPROCESSOR_ARCHIT TIFIER=Intel64 Family 6 Model 94 Stepping 3, GenuineIntel\nPROCESSOR_LEVEL=6\nPROCE Files(x86)=C:\\Program Files (x86)\nProgramFiles=C:\\Program Files (x86)\nProgramW6 2\\WindowsPowerShell\\v1.0\\Modules\\\nPUBLIC=C:\\Users\\Public\nSESSIONNAME=Consol s\\admin\\AppData\\Local\\Temp\nTMP=C:\\Users\\admin\\AppData\\Local\\Temp\nUSERDOW rs\\admin\nwindir=C:\\Windows\nwindows tracing flags=3\nwindows tracing logfile=C:\</pre>
Decrypted Data

Decrypted Network Data.

The image below illustrates some of the main fields used in the first POST request made by this Dridex sample.

30	57	49	4E	2D	55	32	38	4F	39	48	4D	56	55	37	46
5F	33	65	32	31	33	35	65	36	39	36	39	35	38	38	34
37	38	63	37	66	36	63	33	39	36	37	32	38	34	63	35
34	36	62	38	61	39	63	62	30	64	34	65	65	37	35	35
66	36	31	35	64	38	32	66	34	66	31	32	34	62	35	34
61	30	2D	1D	B1	El	11	44	C8	F8	18	40	00	00	01	EF
37	2D	5A	69	70	20	31	37	2E	30	30	20	62	65	74	61
20	28	78	36	34	29	20	28	31	37	2E	30	30	20	62	65
74	61	29	3B	47	6F	6F	67	6C	65	20	43	68	72	6F	6D
65	20	28	38	35	2E	30	2E	34	31	38	33	2E	31	30	32
29	3B	47	6F	6F	67	6C	65	20	55	70	64	61	74	65	20
48	65	6C	70	65	72	20	28	31	2E	33	2E	33	35	2E	34

OWIN-U2809HMVU7F _3e2135e69695884 78c7f6c3967284c5 46b8a9cb0d4ee755 f615d82f4f124b54 a0-.±á.DĚø.@...ï 7-Zip 17.00 beta (x64) (17.00 be ta);Google Chrom e (85.0.4183.102);Google Update Helper (1.3.35.4



Data Sent by Dridex to C2.

5. Automating IOC Extraction

Along with this blog post, we are <u>releasing a python script that automates the IOC extraction</u> <u>from the Dridex</u> loader. These are the main features implemented by our "Dridex Analysis Toolkit":

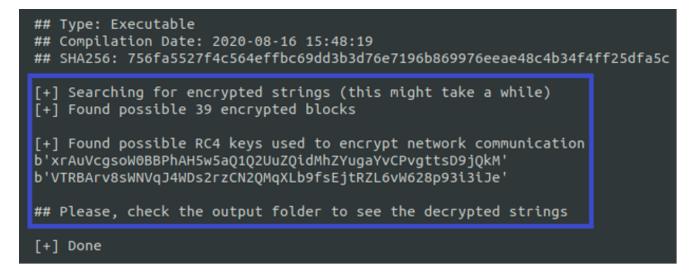
- Extract Botnet ID;
- Extract C2 IP Addresses;
- Decrypt Strings;
- Decrypt Network Communication.

Furthermore, since most Dridex payloads comes from memory dumps, the script also tries to unmap the PE file to disk, so it can get the right offsets to parse, example:

```
:/tmp/dridex$ python dridex_analysis_toolkit.py -f ./unpacked_mem_dump.bin -v -c
    ## Dridex Analysis Toolkit
[+] Output will be saved at: ./unpacked_mem_dump.bin_output
[+] Binary info
## Type: Executable
## Compilation Date: 2020-08-16 15:48:19
## SHA256: 3c86a3c9ef5f62c816bbcd5094022e4192e310be2af9ba768afa97856bc3803a
[-] File seems to be mapped, trying to fix...
[+] Section '.text' moved to 0x400
[+] Section '.rdata' moved to 0x23200
[+] Section '.data' moved to 0x29200
[+] Section '.reloc' moved to 0x29600
[+] Saving unmapped file to: ./unpacked_mem_dump.bin_unmapped.bin
[+] Done
[+] Found C2 parsing function, pattern from 2020
'bot id' at section '.data'
'c2_table' at section '.data'
## Botnet ID:
12333
## C2 Addresses
45.79.8.25:443
185.201.9.197:9443
217.160.78.166:4664
108.175.9.22:33443
[+] Done
```

Extracting C2 Addresses from Dridex Payload Automatically.

By using the "-s" option, the script searches and decrypts the payload strings and prints any possible RC4 keys:



RC4 Keys Found by the Script.

You can then use these keys to decrypt any network communication by using the "-n" option:



Decrypted Network Data.

The script also writes the output data into a folder:



Conclusion

In this post we show the main technical characteristics that makes Dridex a difficult malware to detect and analyze. By publishing this analysis and the automation script, our intention is to help analysts understand how key parts of Dridex work and to help organizations detect and extract Dridex IOCs as early as possible, so that appropriate actions are taken faster.

IOCs

Packed Dridex Loader

d506f18f771ec417c27a6528c17f08ee9d180d40a0a9c6b6ef93b7a39304b96a

Unpacked Dridex Loader

756fa5527f4c564effbc69dd3b3d76e7196b869976eeae48c4b34f4ff25dfa5c

C2 Addresses

45.79.8[.]25:443

185.201.9[.]197:9443

217.160.78[.]166:4664

108.175.9[.]22:33443

Botnet ID

12333

RC4 Keys

xrAuVcgsoW0BBPhAH5w5aQ1Q2UuZQidMhZYugaYvCPvgttsD9jQkM

VTRBArv8sWNVqJ4WDs2rzCN2QMqXLb9fsEjtRZL6vW628p93i3iJe

- [1] https://malpedia.caad.fkie.fra...
- [2] https://malpedia.caad.fkie.fra...
- [3] https://github.com/hasherezade...