Practical 2 Report

Joshua Main-Smith joshuamainsmith@ufl.edu Practical 2 - Malware Reverse Engineering

2021-03-18

Contents

Executive Summary	2
Static Analysis	2
Hashes and Antivirus Check	2
Packing	3
Compilation Date and Subsystem	4
Program Imports	4
Suspicious Strings	5
PE Header Sections	6
Anti-Disassembly Techniques	6
Dynamic Analysis	6
Post-Execution Behavioral Analysis	6
Network Communication	7
Registry Keys Created or Modified	8
Files Created or Modified	9
Processes Started	10
Anti-Debugging and Anti-Virtual-Machine Techniques	10
Indicators of Compromise	12
Sources	14

Executive Summary

The binary exhibits behavior that is consistent with Trojan malware, with the majority of antivirus engines classifying it as such (VT). The binary has been identified as a variant of the AVE_MARIA family, which can be read further with Yoroi's analysis. The malware takes advantage of CVE-2017-11882, which is a Microsoft memory corruption vulnerability that allows an attacker to execute arbitrary code and has a high ranking in terms of severity. The goal of the attacker appears to be stolen system information, stolen credentials, and backdoor installation (for further susceptibility/infection). Further information can be read on Malwarebytes blog.

The sample malware analyzed here had various strings and imports that were deemed suspicious by PEStudio. Particularly interesting was the PE Header section, which contained a blacklisted section (.00cfg) and .tls, which is a section typical in anti-debugging techniques. There were several anti-debugging and anti-vm techniques found while debugging, which essentially consisted of recording timestamps of the victim machine then throwing an exception when execution appeared to be too slow.

The malware attempted to contact an IP address on 195.140.214.82:6703, which can be recognized due to the unusual port number associated with the IP address. Also, it deleted a file that holds meta information on file downloading activity, suggesting the malware was attempting to obscure its behavior by hiding the file it was attempting to download.

Finally, the malware opened several registry keys and manipulated the values of some that would allow the application to make several simultaneous connections to a host.

Static Analysis

Hashes and Antivirus Check

The following are the hashes generated from the VirusTotal report.

MD5 971a3320179e0494fdb70b138ada2446 SHA-1 b04bba3b8be297b6178e73d10f0380897e76464c SHA-256 9633d0564a2b8f1b4c6e718ae7ab48be921d435236a403cf5e7ddfbfd4283382 Vhash 016086551d551d1515156045z200677z90baz1bfz Authentihash 70e203d43f38c128fed15c70ec8479eb5989ab63f535bf190d9c1d54847e9b45 f396b39dbfa473ab2b7180d955fdc740 Imphash Rich PE header hash 94adf1b937a03026d0489a22843d03cc SSDEEP 12288:hkhSL4pH7FYiliicuueTh9yeJWrpDz29Wa+QB1t6gMvlTpa6NYjHhtkaJN:h72Z/8VWrpn2ZF1Ea1jBH TLSH T1E3654A07A7628111FCFE12F774FFB2B8552CABA2279892C352C5A5ECA7055F06D30E52 File type Win32 EXE Magic PE32 executable for MS Windows (console) Intel 80386 32-bit TrID Win32 Executable MS Visual C++ (generic) (47.3%) Win64 Executable (generic) (15.9%) TrID Win16 NE executable (generic) (10.6%) TrID Win32 Dynamic Link Library (generic) (9.9%) TrID Win32 Executable (generic) (6.8%) 1.38 MB (1446912 bytes) PEiD packer Microsoft Visual C++ 8.0 [Debug]

There were a total of 52 engines that detected this binary as being malicious, the majority classifying it as a trojan.

Packing

The binary appears to not be packed due to a few indicators discussed below. To start, the entropy of all the sections are around six and below.



This is consistent with binaries that are not packed as binaries that are packed will generally have sections with an entropy above seven. Further, in the section header one can see that there is minimal deviation between the raw size and virtual size for each section is minimal.

Name	Virtual Size	Virtual Address	Raw Size	Raw Address	Reloc Address	Linenumbers	Relocations N	Linenumbers	Characteristics
Byte[8]	Dword	Dword	Dword	Dword	Dword	Dword	Word	Word	Dword
.text	000B5071	00001000	000B5200	00000400	00000000	00000000	0000	0000	60000020
.rdata	000297DC	000B7000	00029800	000B5600	00000000	00000000	0000	0000	40000040
.data	0007DB5C	000E1000	00079C00	000DEE00	00000000	00000000	0000	0000	C0000040
.idata	0000150E	0015F000	00001600	00158A00	00000000	00000000	0000	0000	40000040
.tls	00000309	00161000	00000400	0015A000	00000000	00000000	0000	0000	C0000040
.00cfg	00000104	00162000	00000200	0015A400	00000000	00000000	0000	0000	40000040
.rsrc	0000043C	00163000	00000600	0015A600	00000000	00000000	0000	0000	40000040
.reloc	000067BE	00164000	00006800	0015AC00	00000000	00000000	0000	0000	42000040

Packed binaries tend to have much larger virtual sizes when compared to their raw size sectional counterpart. Finally, the string count numbers over four thousand and the imports number over 150.



Packed malicious binaries tend to have much fewer string and import counts as to obscure behavior.

Compilation Date and Subsystem

The apparent compilation date and subsystem of the program is December 09, 2020 and is a console application.

file-type	executable
cpu	32-bit
subsystem	console
compiler-stamp	0x5FD1540F (Wed Dec 09 16:47:43 2020)
debugger-stamp	0x5FD1540F (Wed Dec 09 16:47:43 2020)

Program Imports

There are a number of import modules, as listed below.

Module Name	Imports	OFTs	TimeDateStamp	ForwarderChain	Name RVA	FTs (IAT)
szAnsi	(nFunctions)	Dword	Dword	Dword	Dword	Dword
KERNEL32.dll	103	0015F4E4	00000000	00000000	0015FA82	0015F070
USER32.dll	27	0015F75C	00000000	00000000	0015FC4A	0015F2E8
GDI32.dll	2	0015F4B0	00000000	00000000	0015FC78	0015F03C
ADVAPI32.dll	4	0015F474	00000000	00000000	0015FCC4	0015F000
ole32.dll	9	0015F808	00000000	00000000	0015FD92	0015F394
OLEAUT32.dll	11	0015F6FC	00000000	00000000	0015FD9C	0015F288

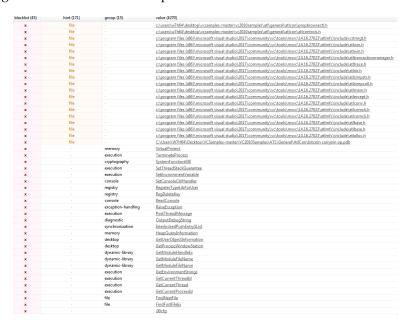
Kernel32.dll is a common DLL that is primarily involved in core functionality in relation to memory, file, and hardware manipulation. *User32.dll* contains user-interface items, like buttons, scrolling, response to user interactions, etc. *Gdi32.dll* is primarily used to display and manipulate graphics. *Advapi32.dll* is used primarily to provide access to core Windows components like the Service Manager and Registry. *Ole32.dll* is used to initialize and uninitialize the COM library and the *Oleaut32.dll* is a subset of COM that is used for OLE technologies and IPC (Wiki).



There were a few import functions that were blacklisted on PEStudio, as shown above. Several of these originated from the Kernel32.dll and appears to mostly be associated with memory manipulation, file handling, and execution. Advapi32.dll and Oleaut32.dll have API calls with the registry grouping.

Suspicious Strings

There were a number of blacklisted strings found by PEStudio, several of those being the ASCII value of the imported functions discussed earlier.



The rest appear to be directory paths to files that are either created or used. All but three of the blacklisted path strings navigate to a header file in Microsoft Visual Studio 2017. The last three navigate to a location under a user named *W7H64*, with the last being a database file (commonly used with MSVS).

PE Header Sections

A screenshot of the various sections for the binary has been included under the *Packing* section. The *text* section primarily contains executable code, *rdata* section has globally accessible read-only data, *data* section has globally accessible data, *.idata* contains import function information, *rsrc* section contains resources necessary for execution, *reloc* section has relocation information of library files, *.tls* contains TLS, a Windows storage class and is uncommon with most programs (it has code that is executed before the entry point, common in anti-debugging techniques), and *.00cfg*, which is commonly employed with MSVS and most likely contains call guard check information.

Anti-Disassembly Techniques

The binary attempts to obfuscate its contents using a few common techniques. One technique (which is common in the industry as well) is to remove the debugging information from the binary. This can be seen when attempting to view the debug symbols, as shown below.



The OEP as shown in Ghidra is also different than when executed in a debugger, suggesting that the binary changes the OEP of the program upon execution. This can make it more difficult to reverse alongside debugging when comparing addresses.

Dynamic Analysis

Post-Execution Behavioral Analysis

Upon execution, the first thing one would initially notice is the command line program starting then terminating, as can be seen below.



Having process monitor open would show that Practical2 running as a process in the background. Hybrid analysis noted that there was a shell code injection into the running process, as indicated below.



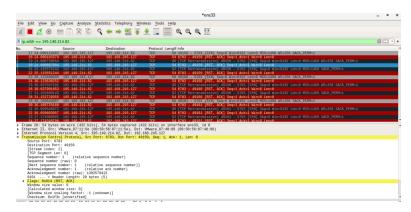
Further, there were several attempts to contact an IP address using an unusual port number (discussed below).

Network Communication

When ran in Virus Total, the binary attempted to contact the IP address 195.140.214.82 on port 6703.



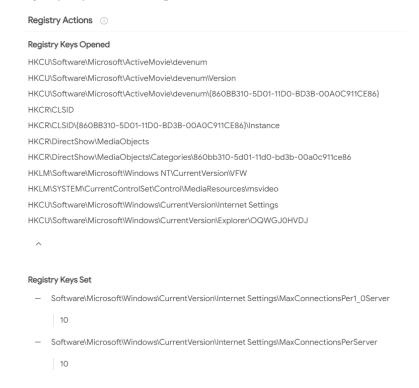
Wireshark was able to confirm this by attempting to contact the host at the above mentioned IP. The ports used on the host machine were iterated incrementally by one, starting at port 49159.



Something interesting to note is that the port used to contact 195.140.214.82 is atypical with most known services and may be an indicator of malicious behavior.

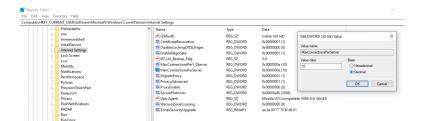
Registry Keys Created or Modified

Some registry keys that were opened or modified are shown below.



After running the binary in a virtual machine, the registry keys *MaxConnectionsPerServer* and *MaxConnectionsPer1_0Server* were set to the decimal value

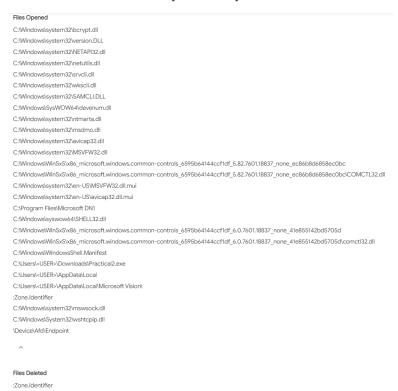
10.



This allows an application (or multiple applications) to make up to ten simultaneous network connections for a given process.

Files Created or Modified

There were several files touched by the binary, most of which were DLL files.



One file the Zone.identifier, was deleted from disk. This is a file that contains meta-information about the file, such as network file downloading activity. The binary deleting this file is an attempt to obscure its network behavior by hiding evidence of file downloads.

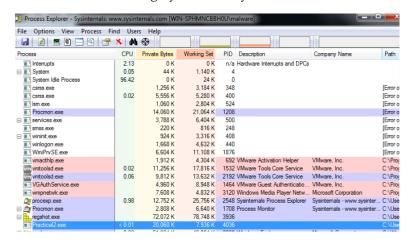
Processes Started

There was only one process started by the binary, which was its own process.

Analysed 1 process in total (System Resource Monitor).

☐ Practical2.exe (PID: 1944) **△** 52/70

This was also observed during dynamic analysis.



Anti-Debugging and Anti-Virtual-Machine Techniques

As mentioned in the static analysis section, there was a .tls section found in the binary. This is usually used in malware to execute code prior to a debugger pausing execution. The presence of .tls lends suspicious to the binary using anti-debugging techniques as this section is rarely present in legitimate applications. This is given further credibility with the presence of debussing strings like <code>IsDebuggerPresent</code> and <code>OutputDebugStringW</code>.

There was a point during debugging where the program got stuck in a loop where the program continually raised exceptions. This is one tactic used in anti-debugging techniques by misusing exception handling by recording timestamps to see how long it takes for the program to execute.

If it takes longer than it should, then there's an assumption that the program is being debugged and raises an exception. This is done by calling *QueryPerformanceCounter*, which is a high resolution API call that measures the time interval between calls in microseconds. An example of this call is shown below.

A further technique this binary uses in measuring timestamps is its use in RDTSC. The machines timestamp is read multiple times to look for any irregularities.

An anti-vm technique that the binary uses is with CPUID which, when executed on real hardware, will return 0. When executed on a VM, CPUID will return 1.

```
| OliBFAD | C745 F4 00000000 | mov dword ptr ss:[ebp=3],0 | mov dword ptr ss:[ebp=3],0 | mov dword ptr ss:[ebp=3],1 | mov dword ptr ds:[12F99CC] | mov dword ptr ds:[12F99CC], ex | mov dword
```

Indicators of Compromise

An obvious indicator would be a binary attempting to contact the above mentioned IP address, 195.140.214.82 on port 6703. This is especially the case considering the port number is unusual.

Two Yara rules I developed were the blacklisted strings from PEStudio. In order for the rule to be satisfied, all of the strings need to be present. Requiring the AND logic is preferred as to limit false positive rates.

```
Yara Rule
rule P2_strings
meta:
        description = "Blacklisted strings found in P2"
strings:
        $a = "InterlockedPushEntrySList"
        $b = "RegisterTypeLibForUser"
        $c = "RegDeleteKey"
        $d = "VirtualProtect"
        $e = "HeapQueryInformation"
        $f = "FindFirstFileEx"
        $q = "FindNextFile"
        $h = "SetThreadStackGuarantee"
        $i = "GetCurrentThreadId"
        $j = "PostThreadMessage"
        $k = "GetCurrentProcessId"
        $1 = "TerminateProcess"
        $m = "GetCurrentThread"
        $n = "GetEnvironmentStrings"
        $0 = "SetEnvironmentVariable"
        $p = "RaiseException"
```

```
$q = "GetModuleFileName"
$r = "GetModuleFileName"
$s = "GetModuleHandleEx"
$t = "OutputDebugString"
$u = "GetProcessWindowStation"
$v = "GetUserObjectInformation"
$w = "SystemFunction036"
$x = "SetConsoleCtrlHandler"
$y = "ReadConsole"
$z = "IsDebuggerPresent"
condition:
    all of them
}
```

The other Yara rule presented are two sections from the PE file, .tls and .00cfg. The latter was blacklisted in PEStudio while the former is typically only present in malicious software. Either can be present for the rule to take hold.

The last Yara rule includes a screenshot of Florian Roth's rule that was detected in Joe's Sandbox. The rule can be found here and includes both Codoso_Gh0st_1 and Codoso_Gh0st_2.

Roth's rule uses various parameters when deciding if it passes, matching on four strings between S1-S7, or four between C1-C8, etc.

Lastly, using the bytes that were mentioned earlier from the shell code injected into the parent running process would be a good indicator of compromise to use. To create a Yara rule for this, one would need to detect the shell code in memory during execution.

Sources

Hybrid Analysis VirusTotal Intezer Analyze ZeroBox Joe Sandbox Yara Repository