Chasing Shadows: A deep dive into the latest obfuscation methods being used by ShadowPad

pwc.co.uk[/issues/cyber-security-services/research/chasing-shadows.html](https://www.pwc.co.uk/issues/cyber-security-services/research/chasing-shadows.html)

08 December, 2021

By Adam Prescott, Cyber Threat Intelligence Analyst, PwC

While monitoring for the backdoor known as ShadowPad, our threat intel practice discovered a bespoke packing mechanism – which we named ScatterBee – being used to obfuscate malicious 32-bit and 64-bit payloads for ShadowPad binaries. The obfuscation mechanism has been briefly touched on in open source;¹ however in this blog we detail how the technique works, ways to analyse binaries obfuscated in this manner, and how to find further samples obfuscated with this bespoke method. This content has previously been made available privately to clients via PwC's intelligence subscription service.

During our analysis, further malicious samples were uncovered which indicate that one or more users of ShadowPad have access to ScatterBee, and have highly likely delivered some of these malicious payloads via watering hole attacks on sites that are used to deliver Adobe Flash update files.

Most of the malicious ScatterBee files can be directly linked back to a China-based threat actor that we are currently tracking as Red Dev 10.

Analysis

Throughout the rest of this blog we will detail a series of obfuscation techniques that, when combined, we assess is the result of a packing mechanism we call ScatterBee. The ScatterBee packing mechanism consists of control flow obfuscation, string encoding, dynamic API resolutions, several anti-analysis techniques and shellcode decoding/decrypting.

For anyone wanting to replicate the analysis detailed in this blog, we have provided an accompanying [GitHub repository](https://github.com/PwCUK-CTO/ScatterBee_Analysis) containing scripts and a walkthrough.

The malicious DLL loader

Discovery of the ScatterBee obfuscation began with a file tagged by ESET on an online multi-antivirus scanner as "a variant of Win32/Shadowpad.L".

This DLL exports one seemingly benign function called "log", which just writes a given string to %TEMP%\log.txt, as well as exporting its entry point function.

The entry point function, which is automatically called when an executable loads this DLL, contains guardrails to make sure the executable loading log.dll has specific bytes at specific offsets, as seen in Figure 1.

Figure 1 - Checking bytes in calling executable

Searching for files with these bytes at these positions returns an MPRESS packed file that is likely a legitimate version of BDReinit.exe, a component of BitDefender. We have observed a similar guardrail technique in previous ShadowPad samples.²

Once the malicious DLL has verified it is being loaded by the target version of BDReinit.exe it will overwrite the parent executable's entry point with a call into its own code.

Figure 2 - Overwriting the entry point of the parent executable

This is a common technique used by various malware families originating from China-based threat actors – notably in PlugX loaders – to gain execution of the malicious DLL's code while running as the original and legitimate executable's process.

Once the parent executable has finished loading its required DLLs, it will then execute code from its entry point, which now points to code in the malicious DLL. This is where the first unique obfuscation technique employed by ScatterBee is found.

Figure 3 - Calls to an obfuscated jump routine

Each of the calls to loc_100095f1 in Figure 3 are used to calculate where the next instruction to be executed is located. The code in this function makes use of pairs of inverted conditional branches to identical locations to further obfuscate how the destination is calculated, as seen in Figure 4.

Figure 4 - Opposing conditional branches

The result of the obfuscated code is to take the return address (the memory location immediately after the call) that is on the stack, get the next four bytes from memory, add them to the return address and then jump to the calculated address.

Figure 5 - Offsets used to calculate destination addresses

In the first highlighted example in Figure 5 the current return address is 0x100128c0; adding the 32-bit value 0xffff81ed to this address results in a target address of 0x1000aaad. From this point on every single instruction in the malicious DLL is followed by an obfuscated jump to the next address, preventing disassemblers from being able to follow the control flow of the sample. As a first attempt at deobfuscating the malicious code we replaced the calls to the obfuscated address calculation function by jmp instructions which jump to the correct location. The results of this can be seen in Figure 6.³

Figure 6 - Fixed control flow

The resulting code has similar instructions to a standard function epilogue (push ebp; mov ebp, esp) but then has a strange comparison instruction comparing the stack register – esp – to 0xe1cf. This is the second technique that ScatterBee employs to obfuscate control flow. Throughout the malicious code, the stack is compared to various low values and then a conditional jump is placed after the check. This fools disassemblers into thinking the code could take the jump if the current stack register is a small value. In practice, it is impossible for the stack register to be a small value, as on x86 and x64 systems the stack is placed in high memory ranges. Further, the targets of the conditional jumps are often into the middle of existing instructions, or to code halfway through functions which prevents disassemblers and decompilers from correctly analysing the flow of execution.

Both of these techniques are likely applied as part of a custom compiler pass⁴ as they significantly modify the control flow of the binary, which is easiest to do before the final assembly instructions have been generated. It is uncommon for China-based actors to employ such extensive custom obfuscation techniques and indicates either a greater level of capability or a greater need to avoid detailed analysis once discovered than other China-based threat actors. Similar techniques have been seen used by financially motivated threat actors (e.g. DoppelPaymer ransomware binaries) who go to extreme lengths to avoid their malware being analysed.

There are several approaches that could aid in statically analysing code obfuscated in this way, however we have taken the route of rebuilding the malicious binary with the jump and stack obfuscations removed. In doing this, the resulting binary will be very close to what would be produced from compiling the original source code with a standard compiler.

Figure 7 - Deobfuscated code from Figure 6

The results of this deobfuscation can be seen in Figure $7.\overline{5}$ This demonstrates the benefit of this approach as in Figure 6 only the first three meaningful instructions were able to be displayed in an analysis tool, whereas in the deobfuscated binary a plain disassembly listing is evident, showing many more instructions while taking up less space.

We chose to leave the stack comparison instructions in the deobfuscated binary while removing the fake branches for two reasons; firstly they do not affect execution of the sample as the obfuscation technique ensures they are never placed between a valid comparison instruction and its resulting conditional jump; and secondly each numerical value used in the obfuscated comparison instruction occurs exactly once in the original obfuscated sample; this means that when analysing the deobfuscated sample an analyst can verify that the output of the deobfuscation tool is accurate by searching for the constant value used in the original binary and checking the expected instructions in both binaries match up.

With a rebuilt binary, decompilation tools were then able to successfully analyse the malicious binary. The differences in outputs are clearly demonstrated in Figure 8 and Figure 9 with the same code being attempted to be decompiled in both figures.

Figure 8 - Decompilation before deobfuscation

Figure 9 - Decompilation after deobfuscation

The next obfuscation technique employed by ScatterBee is to resolve API functions dynamically at runtime. This is achieved by decoding strings specifying the library and function names required, then searching the Process Environment Block (PEB) for the kernel32 functions LoadLibraryA and GetProcAddress and using them to retrieve a pointer to the needed function. The string encoding algorithm is used extensively by ScatterBee obfuscated binaries for API call obfuscation, data obfuscation and string obfuscation.⁶

The encoding algorithm is a stream cipher that takes a 32-bit value as a seed and for each byte in the encoded string:

- Multiplies the current seed by 17:
- Subtracts the 32-bit constant value 0x443246ba from the seed;
- Stores the result as the seed for the next iteration; and,
- Sums each byte of the resulting seed to give the final XOR byte to use with the current encoded byte.

This algorithm will generate a pseudo random sequence of bytes that will be different for each seed used. Different values have been observed being used as the subtraction value in the algorithm. Sometimes the algorithm terminates when it decodes a null character, while other implementations have it run over a fixed number of bytes.

Once these obfuscation methods have been dealt with, it is possible to analyse the functionality of this malicious DLL. It will look for a file in the same folder called log.dll.dat and read the contents. The first four bytes of the file are a littleendian integer to use as the seed value with the previously described encoding algorithm. In this instance, the value 0x107e666d is added to the seed during each iteration instead of having 0x443246ba subtracted.^Z

A buffer is created in memory for the decoded payload, using VirtualAlloc with a length 4,096 bytes greater than the length of the payload. The extra space is so that the malware can generate a random number less than 4,096 via a call to QueryPerformanceCounter, and then use the value as an offset into the buffer to write the payload. This will prevent some detection methods that rely on malicious payloads being written at the start of memory segments and also hinder analysts in determining the entry point of the payload when analysing the sample dynamically.

The malicious payload

The payload is position independent shellcode that uses the same ScatterBee obfuscation techniques as the loader. After deobfuscating the payload to rebuild analysable code there are numerous calls to addresses that are outside the payload's loaded memory (Figure 10).

Figure 10 - Calls using invalid memory locations

This is caused by a further obfuscation technique that is employed by ScatterBee shellcode to patch specific parts of the shellcode at run time. The logic for how the shellcode finds and applies the patches to its own memory is described below.

The first function that the shellcode calls searches through its own memory for a configuration data section by checking that there are six specific integer values consecutively in memory. It XORs every four bytes in the shellcode with 0xAD48FB1D, checking whether the following integer matches the result. Once a match is found it then checks that the next following integer, XORed with 0xE642D205, matches its subsequent integer value and that the integer after that, XORed with 0x868910EE, also matches its subsequent integer value. The valid data in this sample that signifies the start of the configuration information is shown in Figure 11.

Figure 11 – XOR bytes at start of config

The three integers that immediately follow these XOR bytes represent the size of the code section (0xC9000), data section (0x3000) and patch metadata section (0x5AD0) of the shellcode. It further checks the integrity of the payload by checking that the first byte of the shellcode is 0xE9, which corresponds to the initial jmp instruction used by the malware. This is designed to thwart a common malware analysis technique of loading a payload into memory with a breakpoint on the first instruction which has the effect of replacing the first byte (0xE9) with 0xCC.

Once the shellcode has passed these checks it uses the patch metadata section to overwrite data in its own memory. The metadata section is an array of pairs of four-byte integer values, the second integer value in each pair is used as the value to overwrite the four bytes in the shellcode at the offset specified by the first integer value.⁸

The same code from Figure 10 after the patching has been applied can be seen in Figure 12.

Figure 12 - Patched function calls

After we have removed the ScatterBee obfuscation layers from the shellcode, the final payload can be analysed in detail. In this instance, the payload matches what is described as ShadowPad.4 in open source.⁹

An example of configuration information in a 32-bit sample is shown in the following structure:

Each of the offsets in the arrays at 0x34 and 0x72 in the configuration structure point to an obfuscated string that is used by the ScatterBee encoded ShadowPad payloads to specify sample specific variables such as C2s and filenames to use. The obfuscated strings consist of one WORD to use as a decoding seed, a WORD specifying the length of the encoded string, and then the encoded data.

Examples of each of these decoded strings with a description of possible usage is shown in the table below. The first 19 entries correspond to the array starting at 0x34 and the final four entries correspond to the array starting at 0x72.

Table 2 - Configuration strings

64-bit variations

As alluded to previously, we also found 64-bit versions of ShadowPad obfuscated with ScatterBee. The following table shows the details of one of the 64-bit loaders.

Whereas all of the 32-bit loaders found so far have had the filename log.dll, we have found that the 64-bit ScatterBee loaders are named either mscoree.dll or secur32.dll. The functionality of these loaders is identical to the 32-bit variants, in that they search the current directory for a file with the same name as themselves with ".dat" appended (secur32.dll.dat or mscoree.dll.dat), then deobfuscate and load it into memory.

In 32-bit versions of ScatterBee loader files, there are a limited number of strings in plaintext in the .data section of the malicious binary, along with plaintext stack strings for kernel32, LoadLibraryA and GetProcAddress, whereas in the 64 bit samples there are no strings relating to the ScatterBee encoded sections. Some 64-bit ScatterBee files also employ a different encoding algorithm to the stream cipher in various places, and may hint that several different users of ScatterBee have added their own take on obfuscation to the tool.

The encoding algorithm predominantly seen in 64-bit ScatterBee samples uses a combination of MD5 and AES to decode data. The process is as follows:

- Takes 16 hardcoded bytes from the shellcode, along with the final four bytes of the encoded data and MD5 hash them;
- Uses CryptDeriveKey with the resulting hash as input to generate an AES-128 key; and,
- Calls CryptDecrypt on the encoded data (minus the final four bytes used to create the key).

The below table shows the details of an encoded, 64-bit ScatterBee payload.

The configuration data in these samples starts with a similar section of six DWORDs that are used as XOR markers, along with the sizes of the code, data, and patch sections of the payload. However the subsequent data is a series of obfuscated chunks.

Each chunk begins with a four-byte marker that contains the chunk ID in the high byte, and the length of the chunk in the lowest three bytes. For example, the DWORD 0x80000774 has a chunk type of 0x80 and a length of 0x774 bytes. The chunks are decoded by using either of the previously described algorithms - the stream cipher, or the MD5 and AES algorithm.

In the payloads that we have access to, these chunks contain various different IDs. The chunks with 0x80 as their chunk type contain similar data to the 32-bit configuration data, 10 although the encoded strings can use either the AES encoding algorithm or the stream cipher algorithm. Chunks with an ID of 0x02 contain 0x20 bytes of unknown data followed by a valid PE file. These PE files are ShadowPad modules that further enhance the capabilities of the running backdoor. In files seen by PwC, some are obfuscated with ScatterBee techniques and some are not. We have not seen any other IDs in chunks from samples that we have analysed, however, from the code in the ShadowPad backdoor it supports further chunks with IDs of 0x83, 0x84, 0x90, 0x91, 0x92 and 0xa0.

Further malicious files

Pivoting on the names of the DLL loaders and the code sequences used to calculate the obfuscated jumps uncovered 25 malicious DLLs obfuscated with ScatterBee and 10 further malicious payload files that use ScatterBee obfuscation/packing.

Pivoting on the stack comparison code also uncovered trojanised flash installers, a malicious loader and a ZIP archive (detailed in Table 3). All of these malicious files are part of an execution chain that executes variants of ShadowPad, and so far PwC has not found any files obfuscated with ScatterBee that do not deliver ShadowPad, likely indicating that ScatterBee is a core part of the build process of one or more ShadowPad users.

Table 3 - Early stage ScatterBee files

The trojanised installers both contain the same logic for executing their embedded payloads. The initial file is a 64-bit Windows executable that writes two files from its resource section to disk in the folder returned by GetTempPathA. The names and descriptions of the files dropped are as follows:

- A hack tool, named "Microsoft.Win32.TaskScheduler.dll", for carrying out operations relating to Windows Task Scheduler¹¹; and,
- A malicious second stage loader named "td.Principal.UserId =.exe".

The Task Scheduler hack tool is not executed by this or any other stage seen by PwC, and is highly likely an artefact of the build process that supports a persistence mechanism not used by this sample. However, the second stage loader ("td.Principal.UserId =.exe") is executed by the trojanised installer in a call to CreateProcessA. This second stage loader is a .NET executable responsible for dropping and executing a legitimate Adobe Flash installer and a DLL search order hijacked copy of Oleview.exe, as well as creating a task that runs as a LogonTrigger.

First, the malicious loader will attempt to disable all network adapters returned by a query of "SELECT * From Win32_NetworkAdapter". Then, it reads five resources from its own resources section and writes them to disk as the following:

Table 4 - Dropped files

Next it will create a new TaskDefinition (registered as "FlashUpdate") with an action that is triggered by LogonTrigger with the following details:

Table 5 - Persistence task details

With the persistence task registered, the .NET executable reenables the network adapters and creates three processes to execute the dropped .exe files.

The first three files in Table 4 are a DLL search order hijacking triplicate of files with similar functionality to the ScatterBee files described earlier in this report. When the legitimate Oleview.exe is executed by the .NET executable it will load the malicious IVIEWERS.dll, which will in turn load and execute the malicious ScatterBee obfuscated ShadowPad payload contained in IVIEWERS.dll.dat.

flashplayerax_install_cn_fc.exe is also executed by the .NET executable and is a legitimate Adobe Flash installer.

helper.exe is a binary written in Go, which acts as a HTTP server and serves up the response "Hello!" when any client connects to it. It is highly likely that this is another artefact left in the loader by accident, or that is still under development, to allow the threat actor to gain persistence on the victim machine via a secondary backdoor.

The file f7ef194f2dcc341ba03f76872cb7c0dfbae8f79118f99cf73dfccfb146c4e966, from Table 3, is a similar dropper to the first stage of the trojanised installers; however, in this case it simply drops the three OLEVIEW related files straight to disk and executes them.

These first stage droppers also have strings and logic embedded in them to support dropping and executing two further files that were not present in these samples - %TMP%\flsh.exe and %TMP%\schost.exe.

Among the additional DLLs discovered, there was a cluster of eight files that stand out from the rest.

Table 6 – Xiamen submitted DLLs

All of the files in Table 6 were submitted to an online multi-antivirus scanner from locations in Xiamen, China. All of the files have an exported DLL name of Dll.dll, and all apart from the last one were also submitted by the same submitter ID within the space of about 20 minutes. Each of these files are slightly different DLLs: some are MFC binaries, some are meant to be run as Service DLLs; however, all of them contain almost identical copies of ScatterBee packed shellcode to load a .dat file into memory. Only two of these samples have code that would enable the ScatterBee shellcode to run if loaded with an appropriate executable file:

- f768bd36e88ffa496e7b6c538f2259cbdab0317e88432a99050f550b4c9f2f12; and,
- c738af04c5b531abdb303a68cfb8994bb8db6e088bf99b45f85bdb863d3fb3e5.

All the other files either, will not run the packed shellcode, or would require another loader beyond just an executable importing their DLL.

The clustering of file submissions from the same location, the similarity of the files exported names, the presence of almost identical copies of ScatterBee packed shellcode, the mixture of functioning and none-functioning samples, and the submission name of ALTTEST.dll in many of these samples all add weight to the possibility that a developer or user of ScatterBee is based in Xiamen, and was testing and/or developing the ScatterBee packer during January 2021. Alternatively, there is a possibility that these submissions are from a researcher related to Positive Technologies, as their public blog on this malware family was published the day after these submissions to the online multi-antivirus scanner.

Targeting

Based on submissions to an online multi-antivirus scanner of the obfuscated payloads, it is highly likely that the threat actor using the ScatterBee obfuscated ShadowPad binaries has targeted:

A military organisation in Afghanistan;

- An aviation organisation in Hong Kong; and,
- A company with a security operations centre based in India.

There are also numerous submissions from users based in China, some of which may represent testing whether the current version of the malicious file is detected by antivirus vendors, and others that are likely organisations based in China that are being targeted by a ShadowPad user. This targeting is consistent with our historical tracking of ShadowPad victims, based on communications with known command and control servers.

Infrastructure

When extracting the ShadowPad payloads from the ScatterBee encoded payloads we found the following C2s in use in the configuration sections of the backdoors:

Table 7 - ScatterBee encoded ShadowPad C2s

The first four domains in Table 7 were already tracked by us as Red Dev 10, and have resolved to IP addresses that have previously shown up in our scans for ShadowPad C2s. Pivoting on these domains and IPs uncovers a highly connected set of infrastructure that includes the following domains, most of which also have numerous subdomains that have been observed used as C2 addresses in other variants of ShadowPad.

Table 8 - Red Dev 10 domains

Red Dev 10 has made a habit of using NameCheap and Namesilo when registering its domains, and this activity follows that pattern. In addition, the subdomains under several of these domains also follow a pattern of having between 8 and 12 random alphanumeric characters, which, combined with domains registered by NameCheap and Namesilo that

resolve to IP addresses assigned to The Constant Company, as well as being parked resolving to 127.0.0[.]1 when not in use, allows analysts to pivot and find more potentially malicious domains.

While investigating this cluster of infrastructure, several of the domains shared self-signed SSL certificates that were themed around Microsoft. This, together with the domain names chosen in Table 8, shows a general pattern of trying to spoof the legitimacy of infrastructure employed by these campaigns.

The remaining C2s from Table 7 are not easily linked together beyond being found in ScatterBee encoded ShadowPad samples, which leaves open the possibility that there may be multiple groups using the packer, or that for some operations that greater care is taken to compartmentalise the activity.

Putting together the use of ShadowPad (predominantly a tool used by China-based threat actors), C2 infrastructure that we have previously tracked as Red Dev 10, and the likely targeting of targets aligning to previous ShadowPad usage, we assess that most of this activity is highly likely Red Dev 10, with the possibility that a small subset of this activity could be an as yet unknown China-based threat actor.

Conclusion

PwC has been tracking ShadowPad since 2017 and has observed numerous evolutions of the technical capability. During this time, there has also been widespread reporting about its use in supply chain attacks. Despite this, multiple threat actors continue to use ShadowPad for long term compromise of sensitive organisations, including in the military and telecommunications sectors. This activity aligns extremely closely to the threat actor we track as Red Dev 10, which is a known ShadowPad user.

The ScatterBee obfuscation technique documented in this report is likely the latest attempt to minimise detection in victim networks. Whether this technique is exclusively used by one threat actor, or a general development of ShadowPad capability, remains to be seen.

More detailed information on each of the techniques used in this blog, along with mitigations, can be found on the following MITRE pages:

ewig Positive Technologies, https://www.ptsecurity.com/ww-en/analytics/pt-esc-threat-intelligence/higaisa-or-
Department of the state of the strategy with the state DA plugin that fixes these jumps
 $\frac{1}{2}$ and state the

Appendix A – Indicators of compromise

Related content

Cyber security: Research

[Explore our latest cyber security research, developed to share new knowledge on critical cyber security issues.](https://www.pwc.co.uk/issues/cyber-security-services/research.html)

Contact us

[Form](https://cloud.uk.info.pwc.com/cyber-security-sign-up-to-updates)

Hide

© 2015 - Sat May 28 13:50:16 UTC 2022 PwC. All rights reserved. PwC refers to the PwC network and/or one or more of its member firms, each of which is a separate legal entity. Please see www.pwc.com/structure for further details.