Malicious PDF File Analysis - No. 15

1. Report the number of objects in the file.

```
remnux@remnux:~/Downloads$ pdfid.py DF_Hack.pdf
Error opening file DF_Hack.pdf
[Errno 2] No such file or directory: 'DF_Hack.pdf'
remnux@remnux:~/Downloads$ pdfid.py DF_Hack1.pdf
PDFiD 0.2.8 DF_Hack1.pdf
 PDF Header: %PDF-1.0
 obj                    12
 endobj                 12
 stream                  2
 endstream               2
 xref                    2
 trailer                 2
 startxref               2
 /Page                   2
 /Encrypt                0
 /ObjStm                 0
 /JS                     1
 /JavaScript             1
 /AA                     1
 /OpenAction             1
 /AcroForm               0
 /JBIG2Decode            0
 /RichMedia              0
 /Launch                 1
 /EmbeddedFile           0
 /XFA                    0
 /URI                    0
 /Colors > 2^24          0

remnux@remnux:~/Downloads$
```

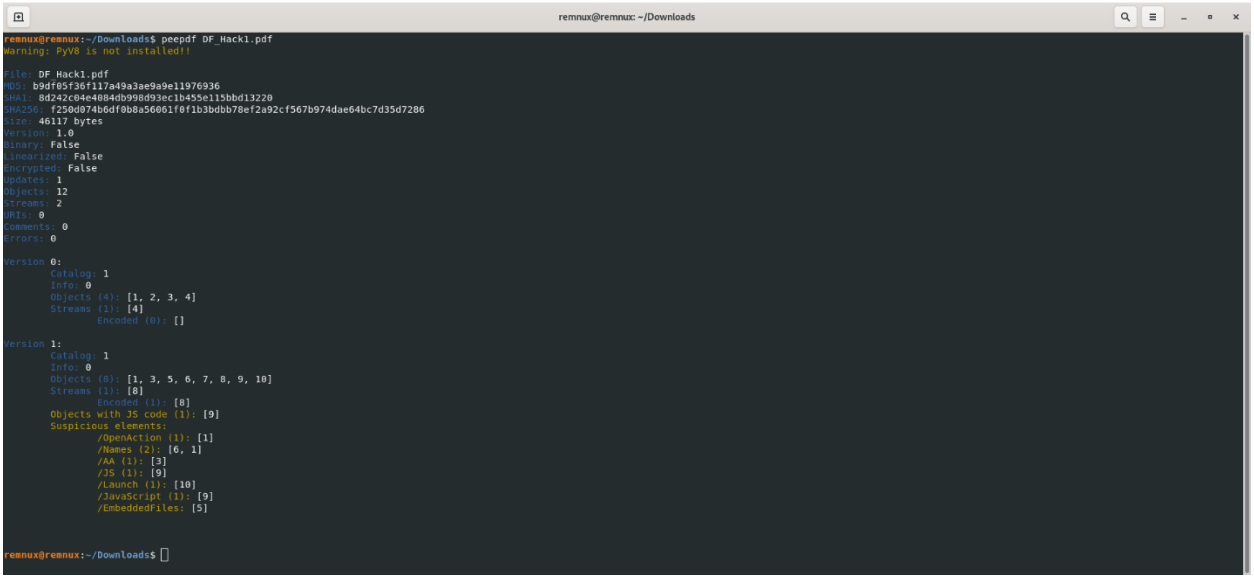
There are 12 objects in the pdf file.
Command : pdfid.py filename.pdf

2. Determine whether the file is compressed or not.

They have used Filter/FlateDecode for compression. Thus, the file is compressed.
We checked this in text editor.

3. Determine whether the file is obfuscated or not.

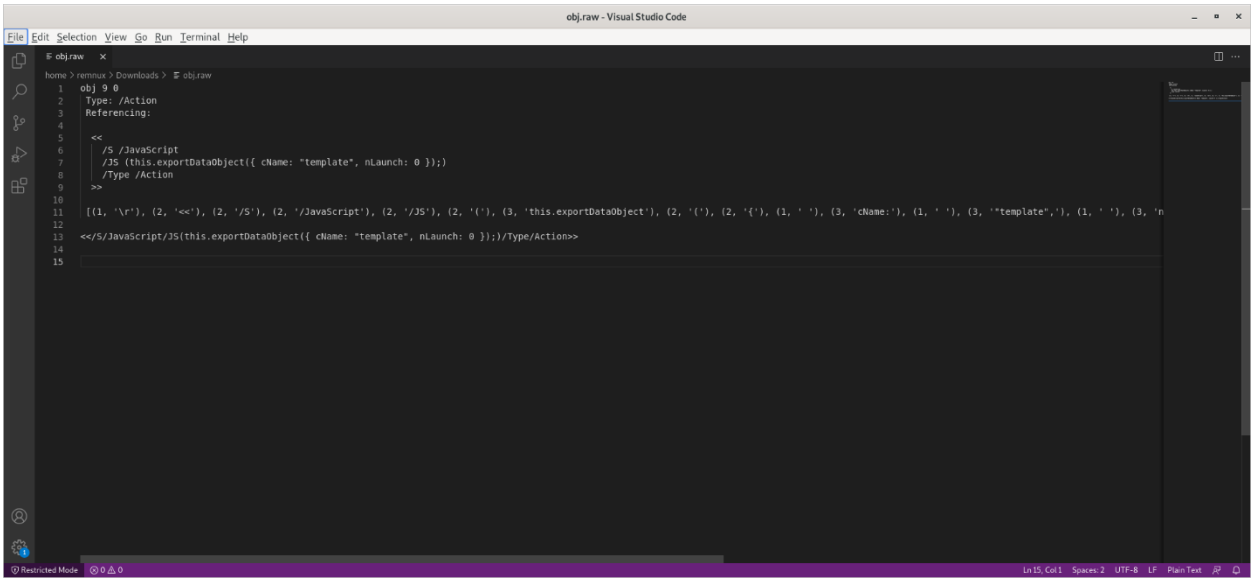
This pdf is not obfuscated. We can see one JavaScript and is not obfuscated. It just launches template.pdf.



The Encoded is object 8 but, not the object 9 which contains JavaScript. Hence the JavaScript is not obfuscated.
Command: peepdf filename.pdf

4. Find and Extract JavaScript.

We found JavaScript in object 9.

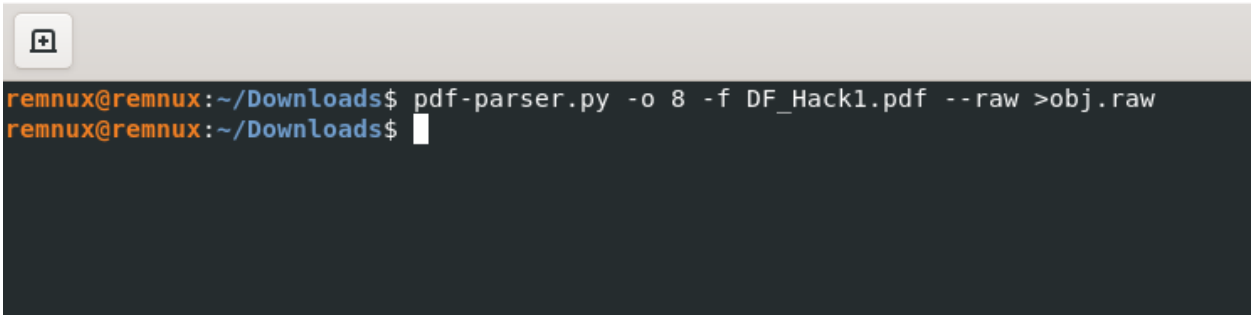


We found JavaScript in obj 9.

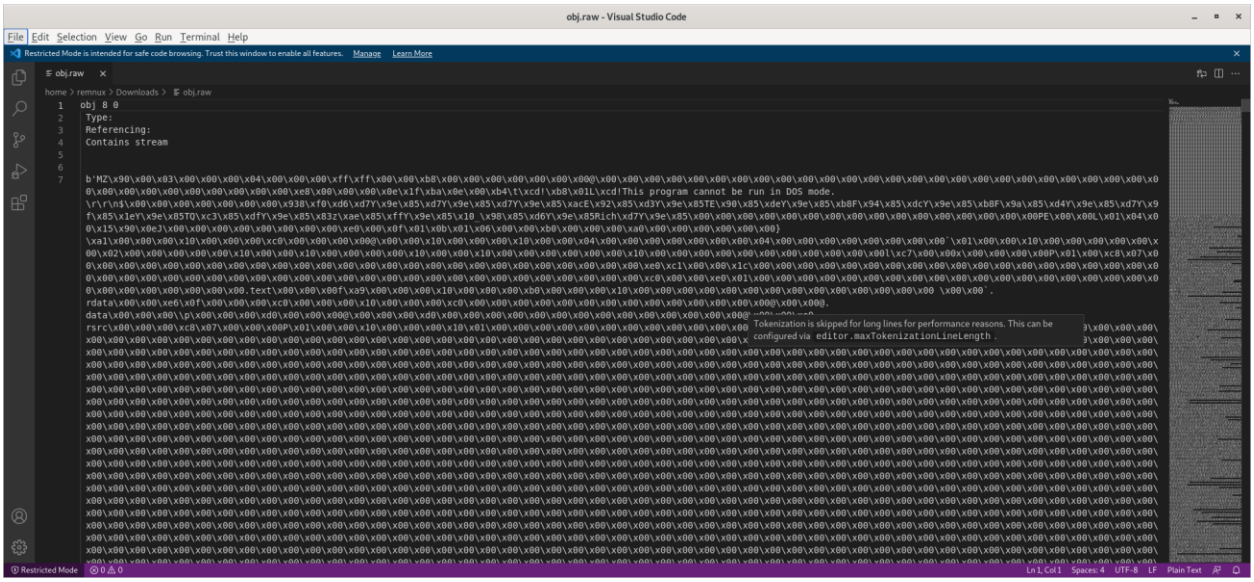
5. De-obfuscate JavaScript.

Since, the JavaScript is not obfuscated so, we cannot De-obfuscate JavaScript. But we have used filters to remove the 'FlateDecode' compression for the stream which is encoded.

6. Extract the shell code.

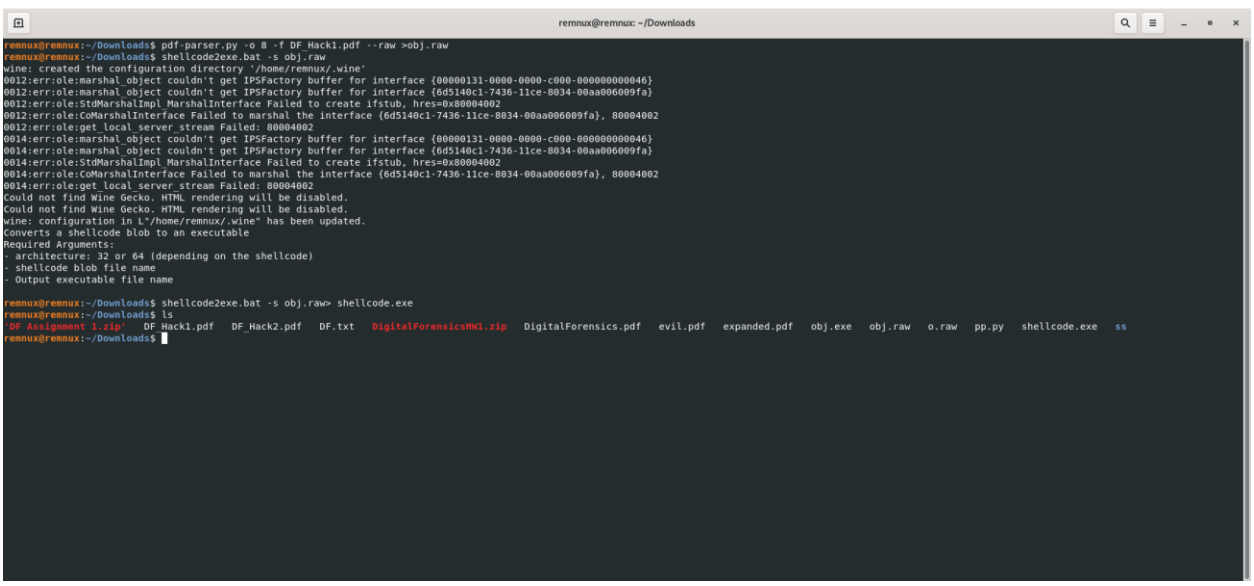


We found a code in the stream but, its not JavaScript.
Command: pdf-parser.py -o objNo -f filename.pdf –raw > outputfile



Extracted shell code.

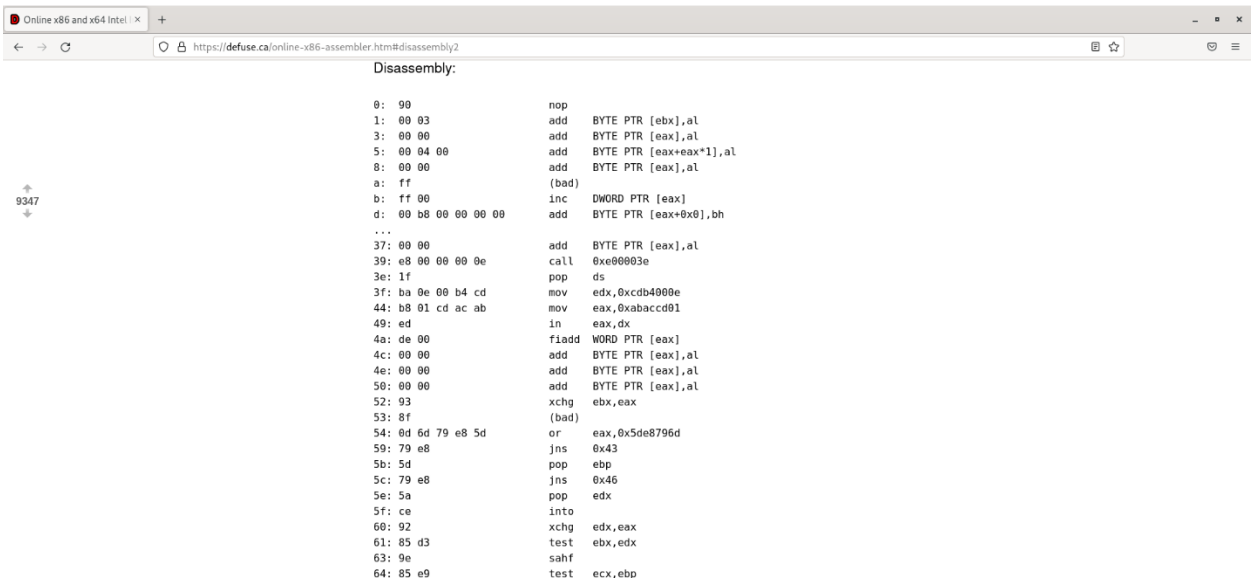
7. Create a shell code executable



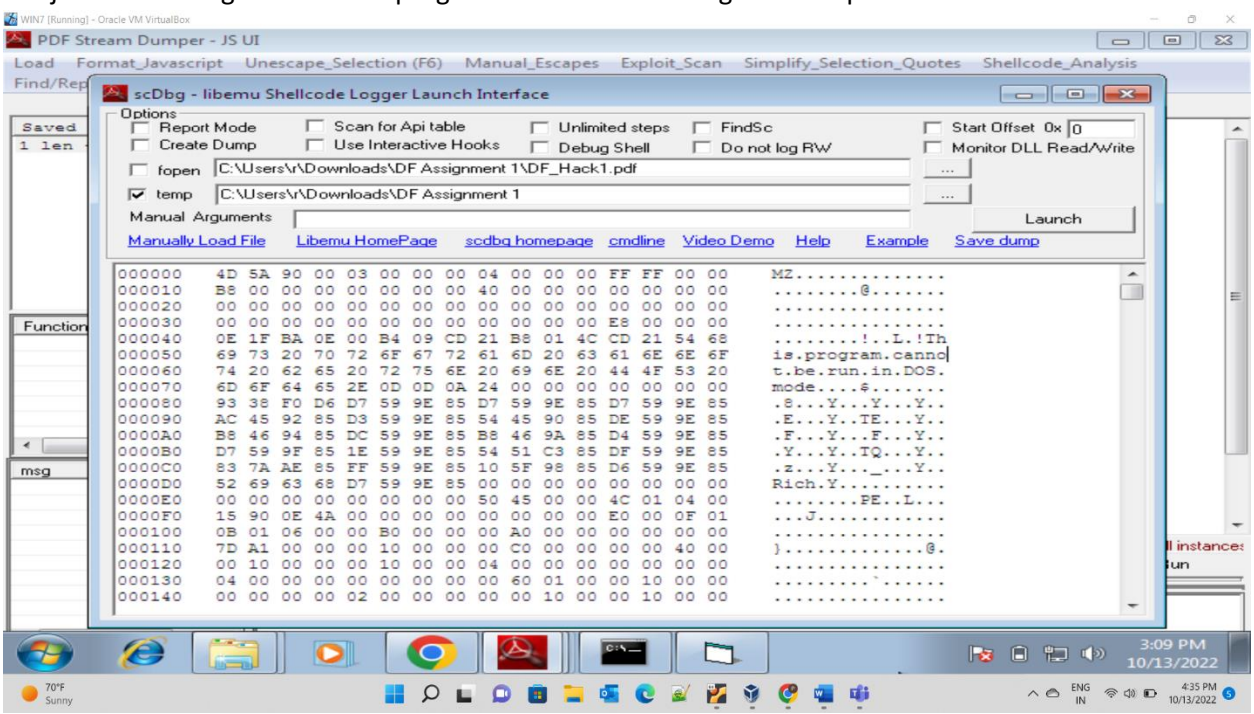
We have created shellcode.exe from the file obj.raw.
Command : shellcode2exe.bat -s extractedshellcodefile > outputfile.exe

8. Analyze shell code and determine what is does or even execute it using sctest or spider monkey.

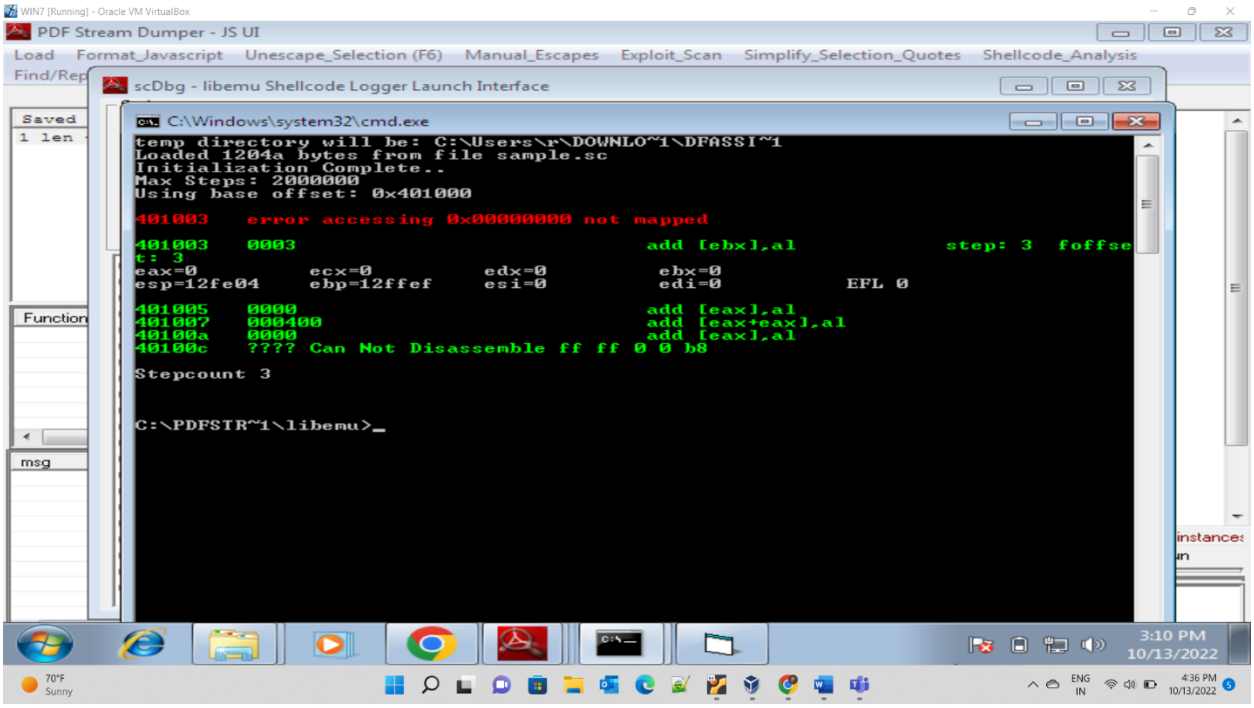
We cannot analyze the shellcode because it does not contain any obfuscated JavaScript and found assembly code which is encoded in the stream. It is in the format of ‘HEX’ and can convert to assembly code by using ‘x86’ format.



We just tried Scdbg to check the program in the stream using PDF dump streamer.



We executed it



The results tell nothing.