Extracting Metada From Files

pentestlab.blog/category/information-gathering/page/2

February 20, 2013

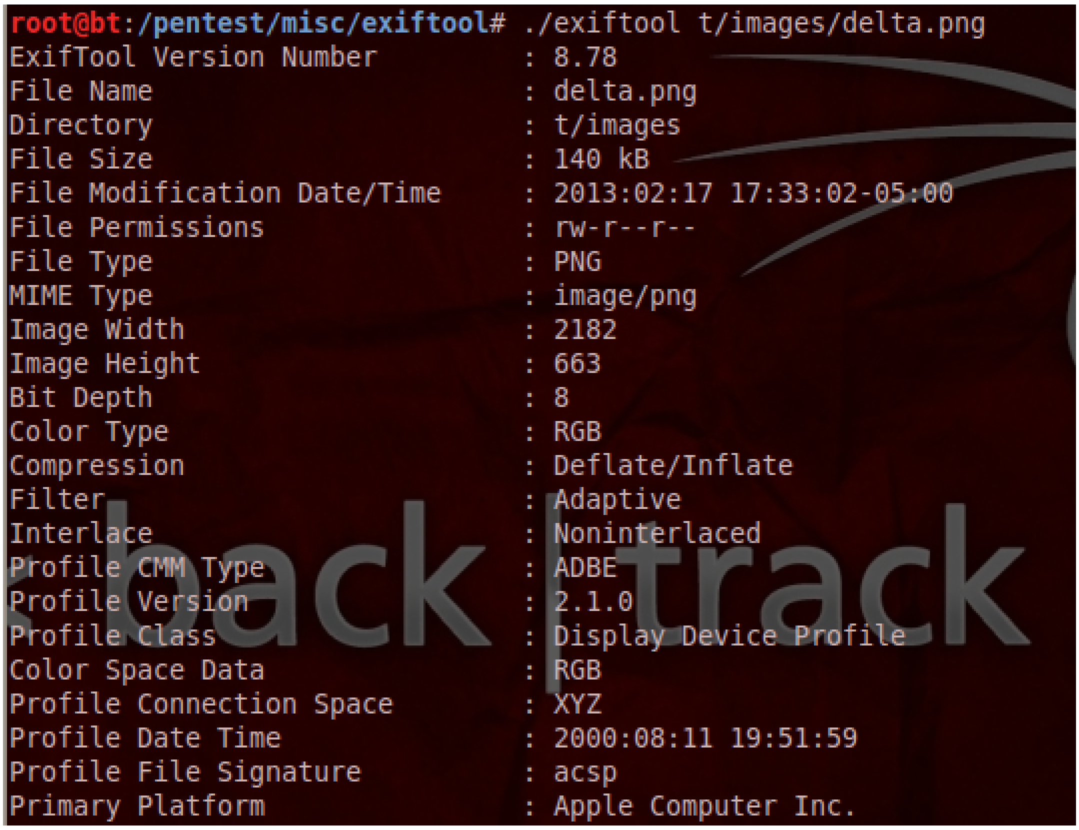
Penetration testers must be able to think outside of the box and to use whatever method is necessary in order to discover information about their targets.Malicious attackers will not stop in the conventional tactics and this should apply and to the penetration tester.Many organizations are uploading in their websites word documents and excel files without been aware that they expose sensitive information.This information is hidden in the metadata of the files.Also in application assessments (web or mobile) it is a good practice except of the common vulnerabilities to check and the metadata in order to see if this information can be used in a malicious way.In this article we will examine some of the tools that we can use for metadata extraction and what kind of information can unveil.

Exiftool

One of the tools that can extract Metadata information is the exiftool.This tool is found in Backtrack distribution and can extract information from various file types like DOC,XLS,PPT,PNG and JPEG.Typically the information that we would look for are:

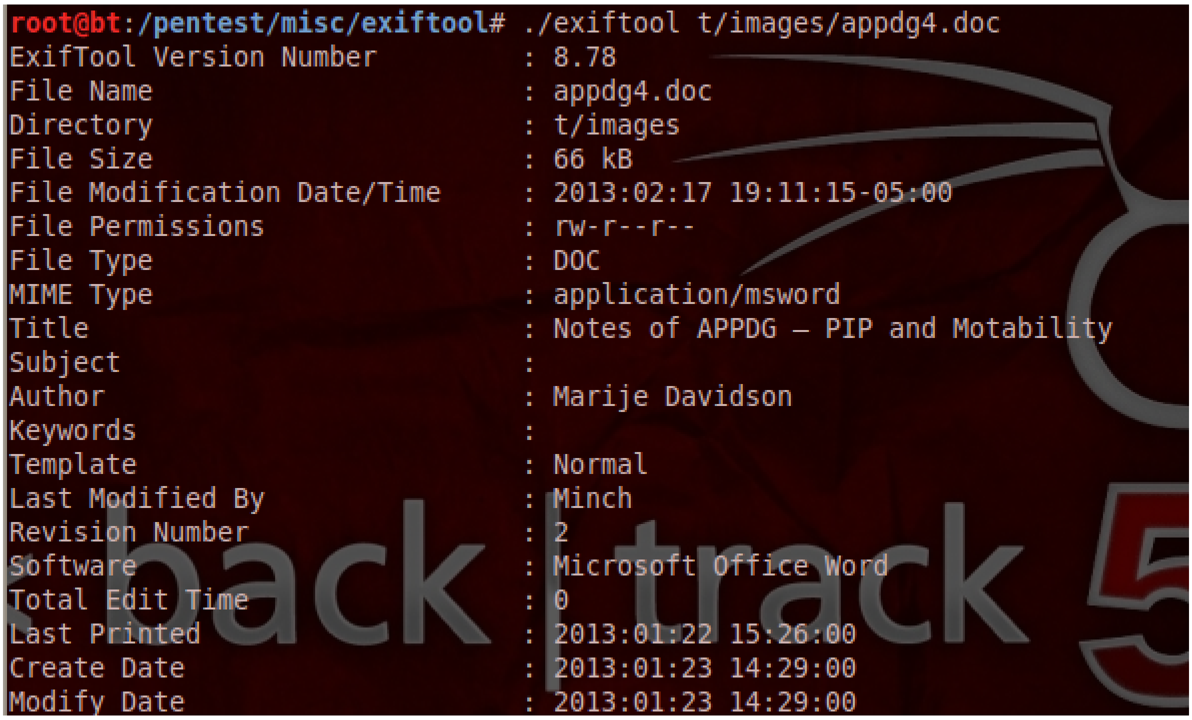
- Title
- Subject
- Author
- Comments
- Software
- Company
- Manager
- Hyperlinks
- Current User

Below is the information that we have obtained from an image and the metadata from a doc file.



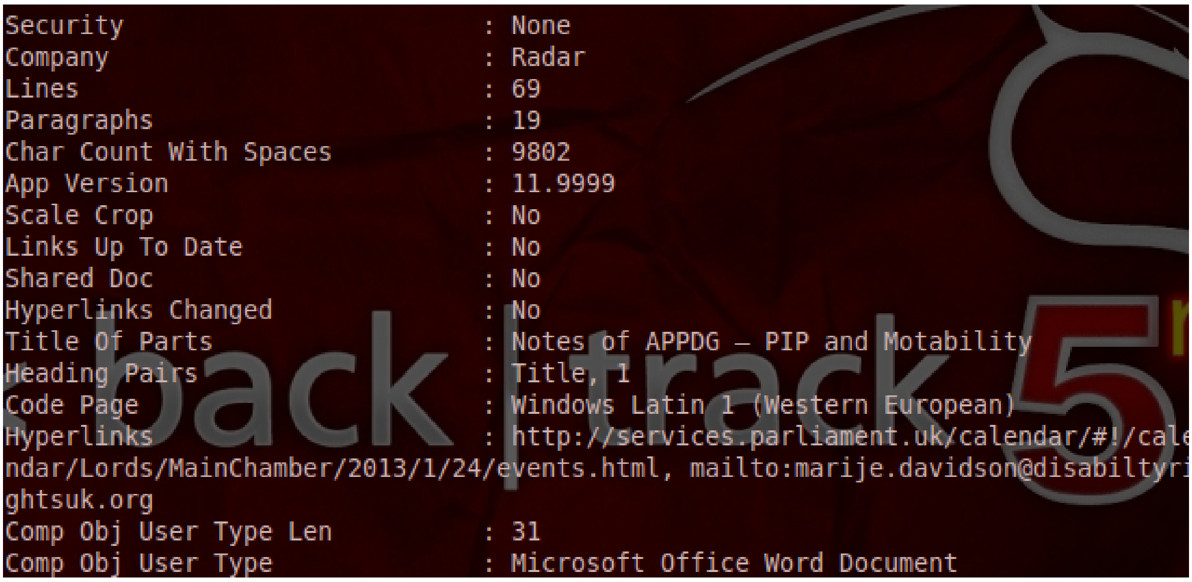
```
root@bt:/pentest/misc/exiftool# ./exiftool t/images/delta.png
ExifTool Version Number         : 8.78
File Name                       : delta.png
Directory                       : t/images
File Size                       : 140 kB
File Modification Date/Time     : 2013:02:17 17:33:02-05:00
File Permissions                : rw-r--r--
File Type                       : PNG
MIME Type                       : image/png
Image Width                     : 2182
Image Height                    : 663
Bit Depth                       : 8
Color Type                      : RGB
Compression                     : Deflate/Inflate
Filter                          : Adaptive
Interlace                       : Noninterlaced
Profile CMM Type                : ADBE
Profile Version                 : 2.1.0
Profile Class                   : Display Device Profile
Color Space Data                : RGB
Profile Connection Space        : XYZ
Profile Date Time               : 2000:08:11 19:51:59
Profile File Signature          : acsp
Primary Platform                : Apple Computer Inc.
```

Extracting metadata of an image – exiftool



```
root@bt:/pentest/misc/exiftool# ./exiftool t/images/appdg4.doc
ExifTool Version Number         : 8.78
File Name                       : appdg4.doc
Directory                       : t/images
File Size                       : 66 kB
File Modification Date/Time     : 2013:02:17 19:11:15-05:00
File Permissions                : rw-r--r--
File Type                       : DOC
MIME Type                       : application/msword
Title                           : Notes of APPDG – PIP and Motability
Subject                         :
Author                          : Marije Davidson
Keywords                        :
Template                        : Normal
Last Modified By                : Minch
Revision Number                 : 2
Software                        : Microsoft Office Word
Total Edit Time                 : 0
Last Printed                    : 2013:01:22 15:26:00
Create Date                     : 2013:01:23 14:29:00
Modify Date                     : 2013:01:23 14:29:00
```

Metadata of a doc file

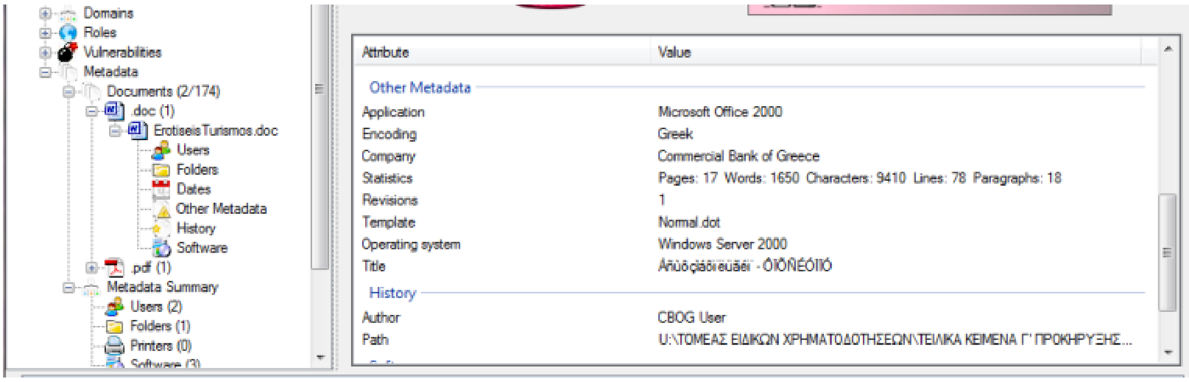


```
Security                          : None
Company                           : Radar
Lines                             : 69
Paragraphs                        : 19
Char Count With Spaces            : 9802
App Version                       : 11.9999
Scale Crop                        : No
Links Up To Date                  : No
Shared Doc                        : No
Hyperlinks Changed                : No
Title Of Parts                    : Notes of APPDG – PIP and Motability
Heading Pairs                     : Title, 1
Code Page                         : Windows Latin 1 (Western European)
Hyperlinks                        : http://services.parliament.uk/calendar/#!/cale
ndar/Lords/MainChamber/2013/1/24/events.html, mailto:marije.davidson@disabiltyri
ghtsuk.org
Comp Obj User Type Len            : 31
Comp Obj User Type                : Microsoft Office Word Document
```

Metadata of a doc file 2

FOCA

FOCA is another great tool for analyzing metadata in documents.It is a GUI based tool which make the process a lot of easier.The only thing that we have to do is to specify the domain that we want to search for files and the file type (doc,xls,pdf) and FOCA will perform the job for us very easily.Below you can see a screenshot of the metadata that we have extracted from a doc file.As you can see we have obtained a username an internal path and the operating system that the file has created.



Attribute	Value
Other Metadata	
Application	Microsoft Office 2000
Encoding	Greek
Company	Commercial Bank of Greece
Statistics	Pages: 17 Words: 1650 Characters: 9410 Lines: 78 Paragraphs: 18
Revisions	1
Template	Normal.dot
Operating system	Windows Server 2000
Title	Áñùôçìáôïëüãéï - ÔÏÕÑÉÓÌÏÓ
History	
Author	CBOG User
Path	U:\ΤΟΜΕΑΣ ΕΙΔΙΚΩΝ ΧΡΗΜΑΤΟΔΟΤΗΣΕΩΝ\ΤΕΛΙΚΑ ΚΕΙΜΕΝΑ Γ' ΠΡΟΚΗΡΥΞΗΣ...

FOCA – Metadata

Conclusion

As we have seen in this article metadata can unveil important information which can be used in conjunction with other attacks.Companies should be aware about this exposure of information that exist in their documents and before they upload something on public domain must use the appropriate tools first in order to remove the metadata from their files and to mitigate the risk.