

DOCUMENT ANALYSIS

COMP 597 – Computer Forensics

Deleted Data



- What are the different types of deleted files?

Deleted Data

- ❑ Deleted files
- ❑ Orphaned files
- ❑ Unallocated files
- ❑ Overwritten files

Files



- We have two types of metadata:
 - ▣ File system level
 - ▣ Application level

Files are the Key



Where do we find a specific file type?

1. Aren't always what they appear to be
2. Can be hidden in strange places
 1. Can be embedded in other files
 2. Stored in metadata
 3. Steganography
 4. Windows Registry
3. Metadata can hold clues

File Identification



- File extensions
 - ▣ Control behavior of files
 - ▣ Generally identify the type of file

File Metadata



- File header information
- System Attributes
- Embedded Data

File Identification

- File headers
 - ▣ Used by application to identify type of file
 - ▣ Not so easily changed (but can be)
 - ▣ Provides the starting point for data carving utilities

File Identification

Table 15.4 Headers and Footers of Common File Types^a

File Type	Header	Footer
JPEG	Usually FF D8 FF E0 or FF D8 FF E1 and sometimes FF D8 FF E3	FF D9
GIF	47 49 46 38 37 61 or 47 49 46 38 39 61	00 3B
Microsoft Office	D0 CF 11 E0 A1 B1 1A E1	N/A

^aAdditional common file signatures are tabulated at http://www.garykessler.net/library/file_sigs.html.

From: Digital Evidence and Computer Crime 3rd edition, Eoghan Casey

File Identification



- File Signatures Table
- File Signatures Database

File Identification - Exercise







How do we determine the type of a file in Windows OS with no file extension?

File Identification

3 Results Found For 504B030414000600

	<u>Extension</u>	<u>Signature</u>	<u>Description</u>
☆	<u>DOCX</u>	<u>50 4B 03 04 14 00 06 00</u> ASCII PK●●●●●●	MS Office 2007 documents Sizet: 8 Bytes Offset: 0 Bytes
☆	<u>PPTX</u>	<u>50 4B 03 04 14 00 06 00</u> ASCII PK●●●●●●	MS Office 2007 documents Sizet: 8 Bytes Offset: 0 Bytes
☆	<u>XLSX</u>	<u>50 4B 03 04 14 00 06 00</u> ASCII PK●●●●●●	MS Office 2007 documents Sizet: 8 Bytes Offset: 0 Bytes

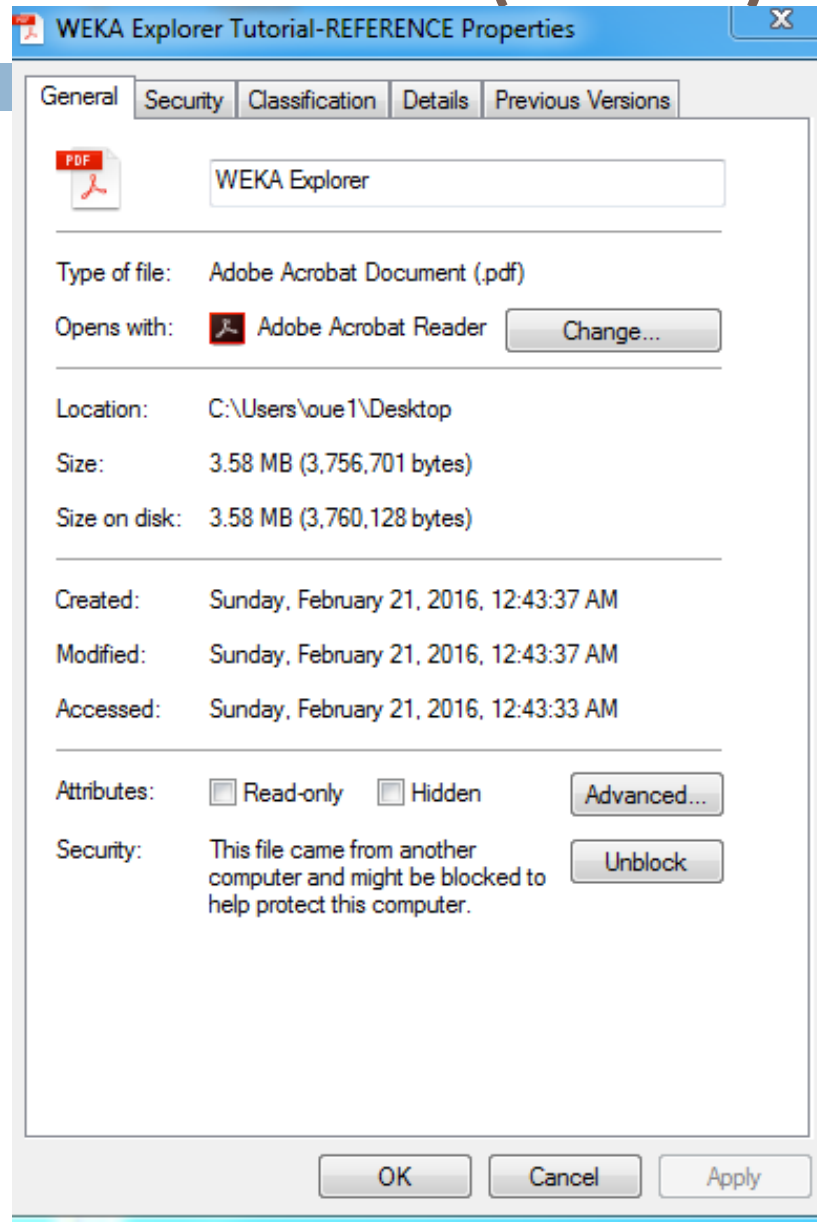
File Identification - Metadata

Name	Type	Compressed size
 _rels	File folder	
 docProps	File folder	
 ppt	File folder	
 [Content_Types]	XML Document	1 KB

System Attributes (MAC)

- Modify: the time when the contents are last altered
- Access: modified when the file is views/opens/copys
- Create: the first time the file is saved in the file system

System Attributes (MAC)



System Attributes (MAC)

Exercise:

- Suppose we are copying a.doc from folder A to folder B
- What will be the changes in the MAC attributes

File Tunneling



- Happens when a file is deleted and a new file with the same name is created shortly afterward

Embedded Data



Applications that create files generate metadata:

- ▣ User Name
- ▣ User Initials
- ▣ Organization Name
- ▣ Computer Name
- ▣ Revision Log
- ▣ Version Log
- ▣ Storage Location



Hiding Information

Hiding Information



- Using Hidden Text
- Embedded Data

Hidden Text

Microsoft Word:

Things that I ate:

Monday:

- *One cup of vegetable soup*
- *Rice, vegetables & yogurt*
- *Broiled fish fillet*

Embedded Data



- Can we hide information in embedded data?

Embedded Data - EXIF



- ❑ Origin Information
- ❑ Image Information
- ❑ Camera Information
- ❑ Additional Information



Data Recovery

File Carving



- What is File Carving?

File Carving

- Locating files in a raw data stream such as:
 - ▣ unallocated clusters on a hard drive
 - ▣ physical memory dumps
 - ▣ raw network traffic

File Carving

- ❑ hachoir-subfile
- ❑ Foremost
- ❑ Scalpel
- ❑ PhotoRec

File Carving



- What are the limitations of file carving?

References



- Digital Archaeology
- Digital Forensics with Open Source Tools