SY110
Digital Data: Files

Major Brian Hawkins, USMC

U.S. Naval Academy

Fall AY 2018

Outline

Files

Knowing that all data on an information system is stored as 0's and 1's, what is a file?

Then what's the difference?

- How programs interpret these bytes
- Rules for how files are interpreted called "file format"
  - Plain text (ex. .txt)
  - JPEG (ex. .jpg or .jpeg)
  - MP3 (ex. .mp3)
  - PDF (ex. .pdf)
- Different files are operated on by different programs – Word Documents (.doc or .docx) open with Microsoft Word, but .mp3s are opened by iTunes, Windows Media Player, etc.
- These programs understand the rules for interpreting the file formats they handle – Word knows how to open .doc, but not .mp3

File Extensions

File extensions

File names usually end with a "." and several characters; this is the "file extension". Some common file extensions include .jpg, .docx, .xlsx, .pdf, .mp3.

File extensions, like .jpg, .mp3, .docx, are conventions that the Operating System (Windows, Mac OSX, Linux) uses to determine what type of file it is, and what program to use to open it.

What if we change a file extension?

Let's create a text file, change its file extension, and open it.

File Headers

What are they?

How all files of a particular type begin – a standardized block that is the same for all files of the same type

Examples

- PDF: 0x25 50 44 46
- JPG: 0x52 49 46 46
- DOC: 0x0d cf 11 e0 a1 b1 1a e1 00 00

Why have them?

- It's really easy to modify the file extension
  - Validates that the file extension is correct, or, if a program can read multiple types of files, indicates what type to read it as
- Also contain file metadata – length, timestamps, etc

"GIFAR" Files

Attackers can use file formatting rules against unwitting victims!

GIFAR attack

Essentially a .jpg image with a Java archive (.jar) file inside – after the image has been loaded, and the .jpg footer indicates the end of the image, the .jar file is read by the Java Virtual Machine. The Java-specific code can be written for malicious purposes, and the victim of the attack is none the wiser!

Questions

Questions?