Exercise 9

05/10/2021

File signature analysis

File signatures are data used to identify or verify the content of a file. Such signatures are also known as magic numbers. Almost all file types contain a *file signature* at the beginning of a file and some contain particular data patterns at the end of the file. These patterns at the beginning of a file and the end of a file may be called as *headers* and *footers* respectively.

File signature analysis is done primarily to check files are what they claim to be. Changing the extension of a file does not change its contents. For example, suppose we have a genuine jpg file called file.jpg. Renaming it as file.txt will not change its contents. You may check this using a hex editor. So we can easily detect a jpg file impersonating as a txt file by doing file signature analysis.

A signature analysis will compare a file's header or signature to its file extension. A file header identifies the type of file and is located at the beginning of the file's data area. The Windows operating system uses a file's extension to associate the file with the proper application. UNIX and Linux operating systems also use a file's header information to associate file types to specific applications.

Download **at least** two files with each of the following extensions from the Internet and keep them in a folder: jpg, png, bmp, gif, pdf

Use a hexadecimal editor such as Winhex (see https://www.x-ways.net/winhex/) or some other hexadecimal editor (see https://en.wikipedia.org/wiki/Comparison_of_hex_editors) to look at the hexadecimal contents of the file in order to find headers and footers. Check whether headers and footers are the same for the same file type.

See the following sites for more information about how file signatures look like.

https://en.wikipedia.org/wiki/List_of_file_signatures

https://www.garykessler.net/library/file_sigs.html

Include screenshots in your submission.