COMP 5350 / 6350 – Project #2

The second project for COMP 5350 / 6350 will focus on automating file recovery based on file signatures and analysis of a Windows 10 registry.

Schedule:

Project #2 Assigned: 30 October
Project #2 Due: 4 December

Students Project Requirements:

Each team will be provided with a disk image collected during a forensics investigation.

- ✓ Project2.dd

Automated File Recovery

In Project #1 our focus was on understanding file system structures and recovering user generated files. In this project instead of using a step-by-step process based on file system boundaries, we will instead recover files by making use of file signatures. The objective of Project #2 is to develop a Python script that will take a disk image as an input, locate file signatures, properly recover user generated files without corruption, and generate a SHA-256 hash for each file recovered.

The disk image provided will contain numerous file types including:

- ✓ MPG
- ✓ PDF
- ✓ BMP
- ✓ GIF
- ✓ ZIP
- ✓ JPG
- ✓ DOCX
- ✓ AVI
- ✓ PNG

The following resource will assist with determining file signatures for each file type:
https://www.garykessler.net/library/file_sigs.html

The following program is **an example** of what the kind of information that will be found after the program takes in a disk image. You may configure the output however you would like, but filename, start and end offset, and SHA-256 results must be provided.

Example Output:

./FileRecovery.py Project2.dd

The disk image contains 8 files

File1.mpg, Start Offset: 0x100000, End Offset: 0x200000
SHA-256: 9f86d081884c7d659a2feaa0c55ad015a3bf4f1b2b0b822cd15d6c15b0f00a08

File2.pdf, Start Offset: 0x100000, End Offset: 0x200000
SHA-256: 9f86d081884c7d659a2feaa0c55ad015a3bf4f1b2b0b822cd15d6c15b0f00a08

File3.gif, Start Offset: 0x100000, End Offset: 0x200000
SHA-256: 9f86d081884c7d659a2feaa0c55ad015a3bf4f1b2b0b822cd15d6c15b0f00a08

File4.mpg, Start Offset: 0x100000, End Offset: 0x200000
SHA-256: 9f86d081884c7d659a2feaa0c55ad015a3bf4f1b2b0b822cd15d6c15b0f00a08

File5.pdf, Start Offset: 0x100000, End Offset: 0x200000
SHA-256: 9f86d081884c7d659a2feaa0c55ad015a3bf4f1b2b0b822cd15d6c15b0f00a08

File6.png, Start Offset: 0x100000, End Offset: 0x200000
SHA-256: 9f86d081884c7d659a2feaa0c55ad015a3bf4f1b2b0b822cd15d6c15b0f00a08

File7.pdf, Start Offset: 0x100000, End Offset: 0x200000
SHA-256: 9f86d081884c7d659a2feaa0c55ad015a3bf4f1b2b0b822cd15d6c15b0f00a08

File8.docx, Start Offset: 0x100000, End Offset: 0x200000
SHA-256: 9f86d081884c7d659a2feaa0c55ad015a3bf4f1b2b0b822cd15d6c15b0f00a08

Recovered files are located in ~/RecoveredFiles

Final Report:

Each team will provide a final report that answers the questions from the grading rubric. The format of the final report will include the following sections:

1) Executive summary
2) Problem description
3) Description of analysis techniques utilized
4) Tables and screenshots
5) Conclusions and Recommendations

A single page report will not adequately answer all questions so be prepared to have an in-depth analysis and description of the methods you used to answer the questions. **In the final report ensure you document code utilized from any other sources and describe how the code works!**

Grading Rubric:

The grading rubric that will be used to grade each disk image will be based on the following criteria:

Activity	%	Pts
Are the correct starting and ending offsets specified for each file?	10%	50
Are the correct number of files recovered?	10%	50
Is the file recovery process documented in the code?	50%	250
Are the files correctly recovered?	30%	100
Total	100%	500

Project Grading:

Letter grades will be assigned based on a 10-point scale:

90 – 100 = A
80 - 89.9 = B
70 - 79.9 = C
60 - 69.9 = D
< 60 = F