OLLSCOIL TEICNEOLAÍOCHTA
BHAILE ÁTHA CLIATH
DUBLIN
TECHNOLOGICAL
UNIVERSITY DUBLIN

Digital Forensics

Week 1 – Intro to Module and File Carving

Date: Friday 19th September 2025

Course Code: TU856 TU857 TU858 (Full-Time)

TU Dublin – Grangegorman Campus

School of Computer Science

Seoladh Cláraithe / Registered Address
OT Baile Átha Cliath - Teach na Páirce Ghráinseach Ghormáin
191 An Cuarbhóthar Thuaidh, D07 EWV4, Éire

TU Dublin - Park House Grangegorman
191 North Circular Road, D07 EWV4, Ireland

OT Baile Átha Cliath
Gráinseach Ghormáin
D07 H6K8, Éire

TU Dublin
Grangegorman
D07 H6K8, Ireland

~ +353 1 220500
~ tudublin.ie

Overview

- Intro to the module
- Intro to Digital Forensics
- File Carving

Slides Created: 23rd July 2023
Slides updated: 28th August 2025

Welcome to Forensics

TU Dublin - School of Computer Science (Grangegorman Campus – Central Quad)

Introduction

About Jonathan:

Jonathan McCarthy

Lecturer – School of Computer Science

Contact Details:

If you have any questions or queries regarding the Forensics module please send me a mail:
jonathan.mccarthy@TUDublin.ie

Note: its @TUDublin.ie not @myTUDublin.ie

Module Timetable

Course schedule

- 2 hours lectures
  - Friday 10:00-11:00 (CQ-008 Large Lecture Room 1)
  - Friday 11:00-12:00 (CQ-010 Large Lecture Room 2)
- 2 hours labs
  - Friday 12:00-14:00
    - CQ-235 Specialist Computer Lab 6
    - CQ-227 Specialist Computer Lab 3
    - CQ-238 Specialist Computer Lab 8

Module Notes

- All course content will be available through BrightSpace.
- The module has been set to facilitate self enrollment.

- If you cannot access the content please email Jonathan asap!!
- **jonathan.mccarthy@TUDublin.ie**

Module Aim (from Course Document)

- Introduce the students to the principles Forensics
- Give the students a thorough understanding of the techniques involved in collecting, storing and maintaining admissible electronic evidence
- Provide them with an in-depth practical Forensics knowledge in real-life
- Expand the student’s ability to analyse computer systems and storage media to enable them to complete a comprehensive investigation of information stored and moved electronically.

Learning Outcomes

1. Define and explain the components of Forensics.
2. Discuss, relate and organise the fundamental concepts of Forensics.
3. Critically analyze different aspects of Forensics, Legal and Ethics.
4. Experiment and demonstrate ability to use various tools (Commercial and Open Source) in a computer forensic lab.
5. Collect admissible e-Evidence from a Computer System or storage media which can be presented in the court.
6. Correctly and completely document a forensic investigation.

Assessment Methods

- Written examination – 50%

- Continuous assessment – 50%
  - Forensics Assignment – 30%
  - Forensics Assignment – 20%

Late Submissions

Rules for late submissions:

Week 1: 6% for the first day, 3% for each day thereafter.

Week 2: 3% for each day thereafter.

Week 3: No submissions accepted, zero grade.

Note: All penalties are calculated per day started.

CA Schedule

- For our assignment submissions we will strictly adhere to the CA schedule.
- If anyone is experiencing personal difficulties we can cater for this using the EC process.

The Definition and Importance of Computer Forensics

- Computer forensics is the retrieval, analysis, and use of digital evidence in a civil or criminal investigation.
- Any medium that can store digital files is a potential source of evidence for a computer forensics investigator.
- Computer forensics is a science because of the accepted practices used for acquiring and examining the evidence and its admissibility in court.
- **Forensically sound** means that during the acquisition of digital evidence and throughout the investigative process the evidence must remain in its original state.
- Moreover, everyone who has been in contact with the evidence must be accounted for and documented in the **Chain of Custody** form.

The Importance of Computer Forensics

Crime Investigation

- Helps law enforcement agencies solve cybercrimes such as hacking, fraud, cyberbullying, terrorism, and identity theft.
- Provides digital evidence that can be legally admissible in court.

Corporate Security

- Detects insider threats, data leaks, and intellectual property theft.
- Investigates policy violations within organizations.

Incident Response

- Identifies how cyberattacks occurred, the scope of damage, and the techniques used by attackers.
- Helps prevent future incidents by analyzing weaknesses.

Data Recovery & Integrity

- Recovers deleted, hidden, or encrypted files.
- Maintains the chain of custody so evidence remains authentic and untampered.

Legal and Compliance Needs

- Many industries (banking, healthcare, government) require forensic practices for compliance.
- Ensures that organizations meet regulations like GDPR, HIPAA, etc.

National Security

- Tracks cyber terrorism, espionage, and cross-border cyber threats.
- Provides intelligence for defense and security agencies.

Terminology

- **U.S. Computer Emergency Response Team (US-CERT):**

"...define computer forensics as the discipline that combines elements of law and computer science to collect and analyze data from computer systems, networks, wireless communications, and storage devices in a way that is admissible as evidence in a court of law."

Source: https://www.cisa.gov/sites/default/files/publications/forensics.pdf

- **NIST Special Publication 800-86, "Guide to Integrating Forensic Techniques into Incident Response":**

"Digital forensics, also known as computer and network forensics... is considered the application of science to the identification, collection, examination, and analysis of data while preserving the integrity of the information and maintaining a strict chain of custody for the data."

Source: https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-86.pdf

Terminology

- One of the main difficulties in defining computer crime is that situations arise where a computer or network was not directly involved in a crime but still contains digital evidence related to the crime.
- As an extreme example, take a suspect who claims that she was using the Internet at the time of a crime. Although the computer played no role in the crime, it contains digital evidence relevant to the investigation.
- To accommodate this type of situation, the more general term *computer-related* is used to refer to any crime that involves computers and networks, including crimes that do not rely heavily on computers.
- Notably, some organizations, such as the U.S. Department of Justice and the Council of Europe, use the term *cybercrime* to refer to a wide range of crimes that involve computers and networks.

Terminology

- In an effort to be inclusive and most useful for practical application, the material in this course covers digital evidence as it applies to any crime and delves into specific computer crimes that are defined by laws in various countries.
- The term *digital investigation* is used throughout this course to encompass any and all investigations that involve digital evidence, including corporate, civil, criminal, and military.

Terminology

- The term *computer forensics* also means different things to different people.
- Computer forensics usually refers to the forensic examination of computer components and their contents such as hard drives, compact disks, and printers.
- However, the term is sometimes used more loosely to describe the forensic examination of all forms of digital evidence, including data traveling over networks (a.k.a. network forensics).
- To confuse matters, the term *computer forensics* has been adopted by the information security community to describe a wide range of activities that have more to do with protecting computer systems than gathering evidence.
- As the field has developed into several distinct sub disciplines, including malware forensics and mobile device forensics, the more general term *digital forensics* has become widely used to describe the field as a whole.

Agencies Involved in Computer Forensics Investigations

- Federal Bureau of Investigation
- U.S. Internal Revenue Service
- United States Secret Service
- Federal Law Enforcement Training Center
- National White Collar Crime Center
- INTERPOL
- High Tech Crime Investigation Association
- Computer Technology Investigators Network
- InfraGard
- An Garda Siochana
- New Scotland Yard

Case Studies

Digital forensic case studies (Notion Digital Forensics)

https://notiondigitalforensics.com.au/digital-forensics-case-studies

Case studies (NHS Counter Fraud Authority)

https://cfa.nhs.uk/about-nhscfa/digital-forensics-unit/cases

What is File Carving

- File carving is a process used in digital forensics.
- It looks at extracting data from a drive (storage device) without using the filesystem.
- Carving is extracting structured data from a larger set of raw data based on characteristics(patterns) in the raw data.
- Carving is a powerful process to a digital investigation by:
  - Identifying and recovering from raw, deleted or damaged file systems, memory, or swap space data
  - Recovering files not recognized by the OS or the filesystem

What is File Carving

- File Carving is a digital forensics technique used to recover files from unallocated disk space, memory dumps, or corrupted storage when the file system structures (like FAT, NTFS, or ext tables) are missing or damaged.

- File carving is the process of reassembling files from raw data fragments on a disk or memory image, based only on file content (signatures, headers/footers, patterns) rather than metadata like filenames, timestamps, or directory paths.

Why Carve?

- A deleted file may no longer appear in the file system, but its raw data is still present on the disk.

- A forensic tool (e.g., Scalpel, Foremost, Autopsy, X-Ways) can carve that image out by recognizing the JPEG header and footer in raw binary data.

- Applications
- Recovering deleted files (photos, videos, documents).
- Data breach investigations (finding hidden or wiped evidence).
- Memory forensics (retrieving executable fragments from RAM dumps).
- Disaster recovery (extracting files from damaged storage devices).

Limitations of Carving

- Works best for non-fragmented files.
- Carved files may lack metadata (filename, creation date).
- Highly fragmented files (like large videos) can be partially unrecoverable without context.

File Header

- A header is a sequence of bytes (in hexadecimal or ASCII) at the beginning of a file.
- It usually includes the magic number, and it can identify the file type and sometimes contains metadata (like format version, size, encoding, etc.).
- Example (in hex):
  - JPEG: starts with FFD8 (FF D8 FF)
  - PDF: starts with %PDF (25 50 44 46).
  - ZIP: starts with 50 4B 03 04.
- Think of the header as the file's "opening tag".

Magic Number
A magic number is a specific sequence of bytes at the very beginning of a file.
Its main purpose: identify the file type.

File Footer

- A footer (also called a trailer) is a sequence of bytes that appears at the end of a file.
- It indicates the file’s conclusion.

- Example (in hex):
  - JPEG: ends with FFD9
  - PDF: ends with %%EOF.
  - Some file types (like MP3) don’t have a fixed footer — carving those requires structure analysis.

- Think of the footer as the file’s “closing tag”.

Headers / Magic Numbers

File Type	Magic Number (Hex)	ASCII Equivalent
JPEG	FF D8 FF E0 (or FF D8 FF E1)	—
PNG	89 50 4E 47 0D 0A 1A 0A	‰PNG....
PDF	25 50 44 46	%PDF
ZIP	50 4B 03 04	PK..
GIF	47 49 46 38	GIF8
EXE (Windows)	4D 5A	MZ

Identifying a file on a disk

- Each file will have a header and footer distinguishing the beginning and end of the file.
- Each file type will have a unique header/footer
- The header can also be referred to as a magic number.
- On the disk we are trying to identify specific types of file headers and/or footers and carve out blocks between these two markers
- Not all files have standard footers, these can be difficult to work with.
- Depending on the filesystem the data can be in different locations, we will see this in more detail in week 3.

Header to Footer Carving

- Most file types have standard headers and footers
- To recover data an analyst can carve out everything between the JPEG header and footer to recover the image file.
- A hex editor can be used to work with the disk / image.

File Structure Based Carving

- This uses the internal layout of a file
- This looks at the header, footer and the file metadata
- Popular carvers use this technique (eg. Photorec and Scapel)

Content-based carving

- This is used to recover file data that has not been recovered from file structure carving.
- This uses machine learning and statistic-based algorithms to look for statistical patterns or signatures indicating language or file content.
- All possible data clusters are gathered that appear to be related to the image. These are used to try extract and find meaningful data/information.

Potential Issues

- No entry in the File table for a given file
- File just exists in unallocated space.
- File still exists in unallocated space but some sectors have been reused
- Files may be incomplete
  - Start, end, middle sectors may have been reused
- Majority of file carving programs will only recover files that are contiguous
- Files on SSDs may be fragmented
  - out-of-order and missing sectors

Scalpel

- Scalpel is a file carving and indexing application that runs on Linux and Windows. The first version of Scalpel, released in 2005, was based on Foremost 0.69. There have been a number of internal releases since the last public release, 1.60, primarily to support our own research.

- As of 6/27/2013 Scalpel has been released under the Apache 2.0 License and the source is available at The Sleuth Kit github repository.

- Source: https://github.com/sleuthkit/scalpel

PhotoRec

- PhotoRec is file data recovery software designed to recover lost files including video, documents and archives from hard disks (Mechanical Hard drives, Solid State Drives...), CD-ROMs, and lost pictures (thus the Photo Recovery name) from digital camera memory. PhotoRec ignores the file system and goes after the underlying data, so it will still work even if your media's file system has been severely damaged or reformatted.

- Autopsy uses PhotoRec as part of its ingest modules. We will see this hext week.

- Source: https://www.cgsecurity.org/wiki/PhotoRec

In-Class Demo of File Carving

Questions