Department of Computer Science
University of Pretoria

Computer Security II
COS 721

Draft Study Guide
Version 2024.0.1

2024

Contents

1 Overview **2**
- 1.1 Rationale . . . 2
- 1.2 Prerequisites/Course assumptions . . . 4

2 The course **5**
- 2.1 Module design . . . 5
- 2.2 Prescribed material . . . 5
- 2.3 Test and assignments . . . 5

3 Assessment **7**

4 On integrity **8**
- 4.1 Plagiarism Policy . . . 8
- 4.2 On research integrity . . . 8

5 Instructors **9**
- 5.1 Contact details . . . 9
  - 5.1.1 Course coordinator . . . 9
  - 5.1.2 Interaction with the lecturer . . . 9

Chapter 1

Overview

1.1 Rationale

Most forensic disciplines are associated with a small set of questions that each discipline proclaims to be able to answer. The field of *questioned documents* (QD, also known as *forensic document examination*, FDE) speaks to the authenticity, origin or other attribute of a document. A typical question may deal with aspects related to the creation of the document. In bygone days, this may have entailed asking whether it was typed on an Acme (brand) typewriter, An Acme Model 123 typewriter or even whether it was typed on a specific *individual* typewriter. In a more modern context similar questions may be asked about whether a (physical copy of) a document was printed using a specific brand of printer, model of a printer or even a specific individual printer.

In a similar manner, ballistics deals with questions such as the trajectory of a round fired, the brand of weapon that was used to fire the round based on rifling marks on the round, and the specific weapon that fired the round based on other striation marks on the round and/or toolmarks on the casing. In a forensic context, DNA is used to associate DNA material with a donor; in the ideal case this association is with a unique individual, but in some cases the association may be with some (possibly unidentified) family member of an individual.

The DNA example can be used to emphasise two well-known, but often overlooked aspects of forensic science. Firstly, 'matching' DNA with an individual may support the individual's alibi, and therefore provide exculpatory evidence. Too often only evidence associated with 'guilt' is deemed valuable. Secondly, DNA is often used for paternity testing, which typically provides evidence in civil (rather than criminal) matters. The use of forensic science is not limited to criminal matters, and, in general, the practice of forensic science should not be deemed to primarily be part of some branch(es) of law enforcement. In our specific con-

text, a contract is a very typical example of a document that becomes contested — and disputed contracts are often a civil matter once it reaches a court. Forensic science provides evidence, based on science, that is useful to determine legal matters — be they civil or criminal.

Inman and Rudin discuss the "origin of evidence" in their eponymous paper [2]. They provide a more detailed account of the origin of evidence in a book on the same topic [1]. The evidence they refer to is forensic evidence. They propose a paradigm where "identification, classification or individualization, association, and reconstruction describe the practice of forensic science starting with the recognition of an item as evidence" [2, p.11]. However, their specific point of departure is divisible matter — and hence are referring to some physical basis for forensic science. There are two salient points to note, given this point of departure:

1. Many (or perhaps most) of the branches of forensic science may be grouped into a class of 'divisible matter' forensic science (or 'physical' forensic science); and

2. Digital forensics is a clear exception; amongst others the Inman and Rudin's tenet "that matter must divide before it can be transferred, is necessary to complete the paradigm" does not hold in the digital world; in the digital world perfect copies can be made without division and the questions of *original* and *copy* are often blurred.

Note that 'physical forensics' (just like digital forensics) does dot imply a small set of specific questions, but just points to the "origins of evidence". This, at least, raises the question whether digital forensics should not be similarly seen as an umbrella term, that covers more specific disciplines. If correct, the next question would be to ask which digital sub-disciplines are worth exploring? Arguably some would be very similar (in principle, but not necessarily in practice) to their physical counterparts. Some may have no physical counterpart. Some physical disciplines may have no digital counterpart.

A related question is whether the five categories identified by Inman and Rudin apply in a digital context. (The sixth principle — that transfer depends on division — does not hold in the digital context.)

Let us formulate the following 'hypotheses':

1. Some branches of physical forensic sciences have viable digital counterparts; and
2. The Inman/Rudin principles are meaningful in such a digital context.

Also note that I formulated a potential principle of 'programmed execution' since the videos used in the module have been produced. This principle is suggested as a possible alternative for what the divisible matter principle and intended for application in the digital sphere.

This module intends to explore these two hypotheses by

1. Attempting to define a notion of *Questioned Digital Documents* (QDD) / *Forensic Digital Document Examination*; and

2. Discussing various aspects of it to determine the extent to which the Inman/Rudin principles help to delineate and focus such a notion.

Clearly, a successful demonstration of the viability of QDD/FDDE supports hypothesis 1. To support hypothesis 2 it is necessary to show that the Inman/Rudin principles help to focus the discipline, without excluding any meaningful (forensic) evidence about questioned digital documents. Stated differently, do identification, classification, individualisation, association and/or reconstruction accurately predict what one can expect from a QDD discipline?

1.2 Prerequisites/Course assumptions

In this module we will attempt to *discover* patterns in digital documents, and use those for classification and other forensic questions. This means you have to be comfortable using tools that allow you to inspect files on a low level. More details are provided below. At the other extreme, it is necessary to provide concise and exact descriptions of any patterns we use. One of the common tools we will user for this is grammars. On the simpler side, grammars may be used in the form of regular expressions (and even be useful in the exploration of files at a low level, using a tool such as `grep`). However, knowledge of context free and context-sensitive grammars will also be required. Familiarity with the notions of tokenisation or lexical analysis, as well as with parsing (syntactical analysis) from compiler construction is highly recommended.

As noted, for this module we also assume that you are comfortable to 'tinker' with files. More specifically, we will use typical *nix tools, such as `tr`, `sed` and `grep`, to mention just a few. On a lower layer, tools such as `hexdump` / `hd` will be useful. Tools that extract metadata for specific file types (such as `exiftool` and `pdftk`) will be used in the contexts where they are meaningful.

A simple practice question: Change a file consisting of multiple lines into a single line, so that text processing tools can be used on the resulting string. Then split the file into its original lines again (that may now have somewhat different content). This is just an example of processing that may be useful — such processing will be used, but not be explained in the course.

Chapter 2

The course

2.1 Module design

Lectures will be presented in a face-to-face mode aligned with the class and test time tables published on `https://www.cs.up.ac.za/honours/`. The course will be presented in a typical face-to-face style, with regular indications of issues students should explore between lectures. These issues will typically be discussed at the start of the subsequent lecture.

2.2 Prescribed material

No textbook is prescribed. Material to be used will be accessible via the Internet. A reading list will be provided on the course site (and expanded during the semester).

Note that videos from a previous iteration of this module are available on the YouTube channel *NetworkProf*. Those videos may help you to understand some of the notions discussed in the current iteration. However, significant new insights have been gained since those videos were recorded and they are now deemed to be outdated.

2.3 Test and assignments

Your semester mark will be calculated as follows:

Semester test	30%
Assignment 1	30%
Assignment 2	40%
Semester mark	100%

The plan is to schedule the semester test in the period of around 4 to 11 September 2024. It will be a take-home test. It will cover the work completed by the time of the test as well as material with with students should be familiar (including grammars and string-oriented *nix commands).

In each of two assignments you will be expected to experiment with techniques discussed in the module, and report you results and/or reflect on the effectiveness of certain techniques discussed during the module.

The assignments will be short (approx 4 pages each) written assignments that you have complete on your own. The topics will be announced in due course.

For most assignments you will have an open choice of the type of document / file you want to explore. The intention is that you will apply the logic discussed in the module to document types not discussed during the lecture and report your findings. This forms part of the testing of the hypotheses mentioned at the start of this study guide.

Chapter 3

Assessment

A semester mark of 40% is required to be admitted to the examination.

The semester mark will count 60% towards the final mark, while the examination will count 40% towards the final mark. A subminimum of 40% for the examination and a final mark of 50% are required to pass the module.

Chapter 4

On integrity

4.1 Plagiarism Policy

This department considers plagiarism as a serious offence. Disciplinary action will be taken against student who commit plagiarism. For a formal definition of plagiarism, the student is referred to
`http://www.ais.up.ac.za/plagiarism/index.htm`
(From the UP Main page follow the Library link and then the Plagiarism link.)

4.2 On research integrity

See the Singapore Statement on Research Integrity at
`http://www.singaporestatement.org/`.

Chapter 5

Instructors

5.1 Contact details

5.1.1 Course coordinator

Prof MS Olivier — molivier@cs.up.ac.za

5.1.2 Interaction with the lecturer

The lecturer prefers in-person discussions about the work during consultation hours — see https://mo.co.za/consult. Discussion after formal lectures will also be possible. Note that a telephonic conversation during consulting hours (tel 012-420-2052) is a good option (although in-person visitors will get preference). Arranging a meeting outside formal consultation hours is a further possibility. Note that asynchronous communication presents a number of challenges — especially to make an appointment or to explain or discuss facets of the course content.

A class representative will be appointed and email correspondence with the lecturer should ideally occur via the class representative.

Bibliography

[1] Keith Inman and Norah Rudin. *Principles and Practice of Criminalistics: The Profession of Forensic Science*. CRC, 2000.

[2] Keith Inman and Norah Rudin. The origin of evidence. *Forensic Science International*, 126(1):11–16, 2002.