Department of Computer Science
University of Pretoria

Computer Security II
COS 721

Study Guide
Version 2015.1.0.0

2015
Chapter 1

Overview

In 2015 the module will explore the nature of and scientific basis (or bases) for
digital evidence. We will differentiate between scientific evidence, other evidence
(including eyewitness testimony) and expert opinion. However, our focus will be
on the first category: scientific (digital) evidence.

We consider scientific evidence to be conclusions that can be justified based on
science. To illustrate, consider a non digital forensic discipline: Suppose science
tells us that the probability of two humans having the same genetic markers is 1/n,
for some n. Suppose further that the world population is currently \( p=7\ 000\ 000 \)
000. Finally, assume that we found a DNA trace at the scene of a crime and we
have obtained a DNA sample from a suspect. Can we say that the suspect was
definitely at the scene of the crime if \( n=p \)? What if \( n=2p \)? Or 100 p? Or 1 000
000 p? At what point is the error rate so low that we can conclude that the two
samples match?

One of the immediate questions is whether we are able reach any scientifically
justified conclusions in the digital realm. If we can, the next question has to be:
how certain are we about our conclusions? Note that no satisfactory answer to the
latter question exists (in general) at this time. This may be reason to doubt that
we are indeed able to reach justified scientific conclusions in the digital realm in
general.

In the bigger forensic environment this course focusses solely on examination
(or analysis) of evidence, and ignores (almost) all other phases of the (digital)
forensic process.

To appreciate the module you have to be able to explore and engage with a
range of topics. Questions about the nature of science are inherently philosophi-
cal. We can (and should) learn from fields that reach scientific conclusions about
the physical or biological world and offer such conclusions as evidence. Hence we
need to think about DNA evidence, pathological evidence, toolmark evidence and
questioned documents, to name but a few - even though most of us do not have any
specialist knowledge about these fields. We should extrapolate from what they can (and cannot) do, and apply the lessons to our field. We should study the existing (albeit unsatisfactory) models that attempt to provide a scientific basis for digital evidence. And, finally, we should attempt to establish one or more scientific bases for digital evidence - and critique whatever we propose.

1.1 Prerequisites/Course assumptions

1. All (or almost all) current theories of digital evidence use (introductory) automata theory as a basis. We assume that students taking the module are familiar with basic automata theory.

2. We will venture into algorithmics and assume that students are familiar with basic computability and complexity theory. This, for those unfamiliar with the topic, requires some mathematics skills.

3. Error rates normally take the form of Type I and Type II errors; students should be familiar with these notions, as well as the basic probability theory that informs them.

4. Obviously we have to discuss the nature of science to assess attempts to be scientific. This will entail some philosophical discussion. Prior philosophy knowledge is not assumed; however, students should not be scared by the prospect of some philosophising.

5. In order to learn from other fields we need to learn (and talk) about topics outside our fields of expertise. Some fields may be rather different from our field (and sometimes may include blood, gore and tears).

The “ideal” student for this course is the student who is excited about the prospects listed above and participate in a journey to find answers where few exist, and good answers are even scarcer. A desire to explore, a wide general interest and not minding changing the world are all desirable attributes of such an ”ideal” student for this course.
Chapter 2

The course

2.1 Module design

The module will consist of three phases that will follow, as far as possible, on each other as follows:

- **Preparation**, in which we will learn what the requirements of scientific evidence are, how other disciplines meet those requirements and where scientific (or seemingly scientific) evidence has failed us in the past;

- **Examination models**, in which we will study (and critique) existing models that aim to examine digital artefacts and derive scientifically justified evidence from such artefacts; and

- **Quo vadis?**, in which we will consider the way forward for scientific digital evidence.

Each of these phases will consist of a number of themes, that have already been alluded to in the description of the phases above. Where the phases will (as noted) in general occur sequentially, the themes within each phase will not necessarily be presented in order in each phase. In fact, more than one theme may be addressed by a single lecture and later lectures may return to a theme covered during a previous lecture. Below the themes of each phase are introduced. Note that the plan is to spend about seven lectures for the first phase, and then three lectures for each of the two later phases.

2.1.1 Preparation

The major themes of the preparation phase are as follows:
Science: For evidence to be considered scientific, it has to meet the requirements of science. What are those requirements? Naturally, the domain of science is contested and time will only allow us to consider the dominant theories of science (such as the logical positivism of the Vienna Circle, Popper and Kuhn).

Law: The law restricts what may be proffered as evidence, and then further restricts what may be seen as scientific evidence. We will be interested in the latter category, noting that the law not only depends on the legal system and jurisdiction, but also often enough on the details of a specific case. However, this does not prevent one from exploring specific standards that have been set, such as the US Frye and Daubert standards (and their local ‘equivalents’, or absence thereof).

Gold-standards: Some (non-digital) disciplines are seen as examples of applied science that other disciplines have not yet reached. Why are those disciplines deemed successful? What are the ‘recipes’ of their success? Can we abstract ideas from such disciplines and apply them to digital evidence?

Failures: What was deemed to be scientifically certain evidence did fail in the past with dire consequences for individuals involved. Why was incorrect evidence offered in such cases? Was the science at fault? Was what was deemed scientific not scientific after all? Was it mere operator error, where a technician pushed a ‘wrong’ button? Or did the individual scientist make a professional mistake by interpreting appropriate science incorrectly? Again, the few examples we will be able to study will hopefully contain valuable lessons to consider when thinking about digital evidence.

Statistics: Often questions are asked about error rates of specific methods. Probability theory provides the basis of such error rates and, on a more positive note, certainty and confidence intervals. One needs to comprehend the language of statistics to think about certainty in may of the themes listed above, as well as in the digital domain. The module assumes a basic knowledge of probability theory. This theme will therefore not be presented during lectures, but a class test will be used to ensure that all participants have (or have acquired) a sufficient understanding of the language of science.

Computing / the Digital: It goes without saying that the subsequent phases of this module depends heavily on a comprehension of computing. Given that this course is offered as an honours module in computer science, it is assumed that students are familiar with the foundational concepts of computing, without revisiting the concepts during lectures. Students’ grasp of
these foundational concepts will be assessed in two class tests: one on basic automata theory and one on computability.

2.1.2 Examination models

A few models have been proposed as bases for scientific digital evidence. Some of these models will be discussed during lectures. About three lectures will be devoted to such models, with one or two models to be discussed during each lecture.

Some common themes will be observed across the various models:

The role played by automata theory;

The ubiquitous problem presented by error rates (or, conversely, certainty); and

Notions of science used as foundation for these models.

2.1.3 Quo vadis?

This theme will explore the road ahead: How does digital evidence meet the demands of science (as observed by the philosophers of science and as prescribed by law)? How do we incorporate the lessons from our peer disciples (if possible, at all)? How do we avoid the pitfalls evidenced by examples of improper ‘science’ and even the concerns we note about existing models to examine digital evidence?

The answers are not clear yet, but the following themes clearly have to be covered by our discussion:

Certainty

Science of technology

The use of technique (‘craftsmanship’, ‘craft’, or ‘art’; Modern Greek: ‘texni’; Afrikaans: ‘tegniek’, sometimes also translated in English as technology, but then not intended to refer to machines or automation) in the examination of digital evidence.
Chapter 3

Assessment

The semester mark will be calculated as follows:

- Class test 1 (on probability theory): 15%
- Class test 2 (on automata theory): 15%
- Class test 3 (on computability theory): 15%
- Class test 4 (on existing theories of digital evidence): 25%
- Two written assignments, 15% each

Students will have the option to submit an assignment on the topic of up to two class tests as an alternative to writing the test. (Those who miss a class test for any reason — such as sickness — will be expected to submit the corresponding assignment; however, the maximum of two substitute assignments still applies.)

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. No subminimum is specified for the examination.

3.1 Assignments

A number of assignments will be posted on the website during phase 1 of the module. Students are expected to select any two. The assignments are designed to explore insights gained about the themes covered during phase 1.
Chapter 4

On integrity

4.1 Plagiarism Policy

This department considers plagiarism as a serious offense. 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 — tel 012-420-2052
Consultation hours: See http://mo.co.za/consult

5.2 Interaction with the lecturer

Students are welcome to come and discuss any aspects of the course without appointment during the consultation hours linked to above. Students are also welcome to make an appointment (in person or telephonically) to discuss work at
other times. (The consultation period is also a good time to reach the lecturer telephonically — at other times he may be hiding somewhere to get some work done.) The deluge of email has unfortunately rendered email an ineffective tool for this purpose, and time often does not permit this lecturer to reply to email.
Chapter 6

Study Material

The reading list is available on the course site. It will be updated throughout the semester.
Chapter 7

Lecture schedule

Will be available on the course site soon.