Digital Investigation 24 (2018) 1–3

Contents lists available at ScienceDirect

Digital Investigation

journal homepage: www.elsevier.com/locate/diin

Editorial

Clearly conveying digital forensic results

Concerns about the scientific validity and reliability of forensic results are motivating organizations such as ENFSI[1] and OSAC[2] to formalize how evidence is evaluated and presented to decision-makers.

Most forensic practitioners know they should not express expert opinion as fact, and should not present evidence from only one viewpoint. A fundamental aspect of scientific practice, called falsification, involves seeking evidence that disproves the theory under consideration. Focusing on a single hypothesis could be an indication of bias, or a failure to consider alternative possibilities sufficiently. In addition, a strength-of-hypothesis approach (offering a probability of the hypothesis) creates a logical fallacy because it transposes the conditional aspect of the assertion.[3] However, many digital forensic practitioners are confused by the emerging expectation that they should evaluate forensic findings and express conclusions in terms of probability of the evidence. It does not help that these expectations are not always clearly articulated.

In the trial of Johnny Oquendo for the murder of Noel Alkaramla, investigators presented geolocation information extracted from a mobile device, but the judge expected a more scientific treatment of the evidence. The judge stated that the prosecution "*failed to meet their burden of demonstrating that the science underlying Google location services has gained general acceptance in the relevant scientific community.*" Although the judge raises questions about the reliability of the digital traces, his statement seems more concerned with the way the geolocation information was evaluated and presented in court. In actuality, such geolocation information can provide very strong evidence if treated properly.

Similar challenges have confronted cell site analysis, causing critics to call it "junk science." Cell site analysis uses information in Call Data Records (CDR) from telecommunication network operators, combined with geographic information and radio frequency propagation survey details, to assess whether a mobile device was, or was not, in or near an area at a given time. In an effort to improve the scientific validity and reliability of this form of analysis, the UK Forensic Science Regulator produced an appendix in the Codes of Practice and Conduct specifically addressing "Digital Forensics – Cell site analysis" (FSR-C-135 – published 13 June 2016) which includes these requirements:

> "(b) The consideration of one or more alternative hypotheses"
>
> "(c) The terminology used in reports shall be clearly defined and imply no bias.
>
> Phrases in reports such as 'in the vicinity of' may only be used if qualified; phrases such a 'consistent with' should not be used in reports unless all other scenarios the findings would be consistent with are given."
>
> "(d) Cell site analysis may be used to propose investigative avenues (i.e. to help form a hypothesis). If a hypothesis has been produced through a different process, cell site evidence should only be used to test whether that hypothesis is supported by the evidence; it should never be used to test whether the hypothesis supports the allegations or scenarios being put forward in the case independently of the evidence. Care should be taken not to transpose the conditional aspects of any assertion."

These requirements reflect a growing expectation that forensic practitioners treat digital traces in a manner that is becoming widely accepted in forensic science: evaluating and expressing the relative probabilities of the forensic findings given at least two mutually exclusive hypotheses. However, even if the intention is apparent, there is a lack of clarity about how to accomplish this objective.[4]

To foster clarity and understanding, it is necessary to concentrate on clearly defined core forensic concepts and processes. To this end, OSAC Technical Publication 0002 (https://www.nist.gov/document/osacts0002pdf), "A Framework for Harmonizing Forensic Science Practices and Digital/Multimedia Evidence" defines core forensic concepts and processes in the context of digital and multimedia evidence.

Who makes the decisions in court?

It is important for forensic practitioners to be mindful when they are delivering their expert assessment, and not just facts, particularly when presenting digital evidence in a courtroom context. As an example, consider a sexual assault case in which a deleted photograph of the victim was recovered from a mobile phone used by the defendant. Some digital investigators might present forensic findings in terms of the following assertions:

- The deleted photograph was fully recoverable.
- The photograph was taken using the defendant's mobile phone.
- The photograph was created on 28 February 2018.
- The person in the photograph is the victim.

[1] European Network of Forensic Science Institutes.
[2] Organization of Scientific Area Committees for Forensic Science.
[3] The transposed conditional is a logical fallacy that confuses the probability of the evidence given a hypothesis with the probability of the hypothesis given the evidence.
[4] The above requirement to "test whether the hypothesis is supported by the evidence" focuses on the hypothesis and could easily be misinterpreted as "test the strength of the hypothesis," which is an inappropriate approach. A clearer statement of the intended requirement is to "test the probability of the evidence given one hypothesis versus a given alternative hypothesis; never test the probability of the hypothesis given the evidence."

https://doi.org/10.1016/j.diin.2018.03.001

However, there are several problems within such statements. Firstly, they do not give due consideration to alternative (opposing) claims. Secondly, they present forensic findings as facts, which conceals the inherent decision process. Thirdly, they express conclusions in terms of the claim (hypothesis), rather than in terms of the probability of the evidence given the claim.

The responsibility of digital forensic practitioners is to focus on the digital traces, not to prove or disprove a specific claim. Even when digital traces lead to seemingly incontrovertible conclusions, it is important for forensic practitioners to keep in mind that subjectivity is involved in the evaluation of forensic findings. Ultimately, the judge or jury is responsible for combining the probability of evidence with all other information relevant to the case, and reaching a verdict.

Evaluating alternatives

When forensic examiners concentrate on proving or disproving a specific claim, there can be a risk of confirmatory bias. To mitigate this risk, an increasing number of best practice guidelines are instructing forensic practitioners to evaluate the probability of evidence given one claim versus a given alternative claim.

Continuing the case example above, it is necessary to consider whether the deleted file was recovered incorrectly, whether the photograph was downloaded rather than created on the mobile device, whether the device clock was backdated to 28 February when the photograph was created, and whether the individual in the photograph could be some other person.

When forensic tools automatically recover deleted files containing incriminating information, forensic practitioners must decide whether or not the recovered data are the actual, original contents of the deleted files. This process is necessary because most deleted file recovery operations involve an estimation of what data was allocated to the deleted file.[5] This process involves *authentication* and *evaluation*, as defined in OSAC Technical Publication 0002.

To avoid misinterpretation and missed digital traces, it is sometimes necessary for forensic practitioners to consider the broader alternatives. In sexual assault cases, for instance, forensic practitioners have a duty to disclose information potentially favorable to the defendant, such as digital traces of a consensual sexual relationship between the victim and defendant.

Focusing on evidence

Particularly in the U.S. and Europe, there is an increasing expectation that forensic practitioners clearly express the probability of the evidence given one claim versus an opposing claim. Note the difference between this approach and the earlier statements in the above sexual assault case example, based on language in the ENFSI Guideline for Evaluative Reporting in Forensic Science:

- In my opinion, the results observed in my forensic examination of the file system data structures and recovered content are exceedingly more probable if the deleted photograph was fully recoverable, than if the recovered content was from an unknown file.
- In my opinion, the results observed in my forensic examination of photo-response non-uniformity (PRNU) noise are far more probable if the photograph was taken using the subject mobile device, than if the photograph was taken using an unknown device.
- In my opinion, the results observed in my forensic examination of date-time stamps on the mobile device and embedded within the photograph are much more probable if the photograph was created on 28 February 2018, than if the photograph was created on an unknown date.
- In my opinion, the results observed in my forensic examination of facial features are exceedingly more probable if the person depicted in the photograph is the victim, than if the photograph is of an unknown person.

These statements convey an expert assessment of digital traces in specific domains of forensic expertise, and express a ratio of the probability of evidence given one claim versus the opposing claim. As with any forensic testimony, such statements require an explanation of the forensic examination of digital traces, and any contextual information that was taken into consideration.

Clearly conveying forensic results

One of the most hotly debated issues in forensics science is how to convey forensic results to decisions-makers most effectively.

Many forensic practitioners use categorical conclusion scales including multiple levels, such as ‘definitely’ and ‘probably not’. A risk of such scales is that they encourage forensic practitioners to evaluate one claim in isolation, and they often take an inappropriate strength-of-hypothesis approach, rather than probability of evidence.

In 2002, I developed a Certainty Scale, including decision criteria, to help forensic practitioners express their level of certainty in observed evidence in a way that non-technical decision-makers can easily comprehend. This Certainty Scale emphasizes an evaluation of evidence approach, and does not require (but does not exclude) the use of probabilities. A risk of this Certainty Scale is that forensic practitioners could use it inappropriately to express their level of certainty using a strength-of-hypothesis approach.

The Bayesian approach has gained traction in certain forensic disciplines, producing a ratio of the probability of evidence given alternative opposing claims, called a likelihood ratio (LR). A risk of this LR is that the resulting number (or verbal equivalent) could be difficult for a non-technical decision-maker to interpret, especially when dealing with probabilities.

Although measuring the probability of digital traces poses challenges, forensic practitioners already have approaches to studying specific situations. As noted in OSAC Technical Publication 0002:

> “While the foundations of digital/multimedia evidence are largely in computer science, computer engineering, image science, video and television engineering, and data science, the underlying digital traces are, in large part, created by actions of operating systems, programs, and hardware that are under constant development. As a result, it will not always be possible to test in advance the performance of such systems under every possible combination of variables that may arise in casework. However, it may be possible, to test the performance of a particular system under a particular set of variables in order to address questions arising in a specific case. For instance, digital documents created using a new version of word processing software can exhibit digital traces that were not previously known. The observed traces can be understood by conducting experiments; studying the software under controlled conditions. In this manner, generalized knowledge of digital/multimedia evidence is established and can be used

[5] The NIST Computer Forensic Tool Testing specification for deleted file recovery defines estimated content as “*A tool Estimates Content if it attempts to recover the content of a deleted file, beyond what is explicitly identified in the residual metadata*.”

by any forensic scientists to obtain reproducible, widely accepted results."

Even when probabilities are calculated based on a model and study, the outcomes are dependent on underlying assumptions, the model, and the study. In reality, probability reflects the beliefs of those who developed the model and conducted the study. Therefore, any assertion about probabilities of evidence must be viewed critically as a belief (albeit well-founded) rather than a fact of nature. In the end, there is human judgement involved in the process and, therefore, there is some degree of subjectivity in the results.

Playing by the same rules

Reasonable minds can differ, but all can agree on the role of forensic expert testimony: to help decision-makers understand forensic findings and their value in specific cases. Forensic expert testimony should not advocate for a specific outcome.

The benefits of evaluating forensic findings and expressing conclusions outlined above depends on both parties following the same rules. When the stakes are high (e.g., money, politics, prison) some experts can become biased, presenting digital evidence in a way that is most favorable to their client. Given this reality, a forensic expert following an inappropriate strength-of-hypothesis approach might seem more certain in their conclusions than a person following an appropriate approach of expressing the probability of evidence for one claim versus an opposing claim. If steps are not taken to counteract such situations, forensic practitioners could be discouraged from using an appropriate evaluation approach. To strengthen criminal justice, policy- and decision-makers must take a stand against forensic practitioners acting as advocates, and insist upon evaluation and expression of forensic findings in terms of the relative probabilities of evidence given at least two alternative claims.

Forensic practitioners can treat the emerging expectation to evaluate and express forensic results in a more formalized manner as an opportunity to better understand digital traces and forensic science, and to give decision-makers increased confidence in digital forensic expert testimony. OSAC Technical Publication 0002 provides a framework and definitions to help digital forensic practitioners realize this opportunity.

Eoghan Casey
Editor in Chief
Digital Investigation