Forensic Science International: Digital Investigation 32 (2020) 300923

Contents lists available at ScienceDirect

Forensic Science International: Digital Investigation

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

DFRWS 2020 EU – Proceedings of the Seventh Annual DFRWS Europe

Towards Sound Forensic Arguments: Structured Argumentation Applied to Digital Forensics Practice

Virginia N.L. Franqueira [a, *], Graeme Horsman [b, **]

[a] School of Computing, University of Kent, Canterbury, CT2 7NF, UK
[b] School of Health & Life Sciences, Teesside University, Middlesbrough, TS1 3BA, UK

ARTICLE INFO

Article history:

Keywords:
Structured argumentation
Toulmin
Investigation process
Digital forensics
Logical reconstruction

ABSTRACT

Digital forensic practitioners are increasingly facing examinations which are both complex in nature and structure. Throughout this process, during the examination and analysis phases, the practitioner is constantly drawing logical inferences which will be reflected in the reporting of results. Therefore, it is important to expose how all the elements of an investigation fit together to allow review and scrutiny, and to support associated parties to understand the components within it. This paper proposes the use of 'Structured Argumentation' as a valuable and flexible ingredient of the practitioners' thinking toolbox. It explores this approach using three case examples which allow discussion of the benefits and application of structured argumentation to real world contexts. We argue that, despite requiring a short learning curve, structured argumentation is a practical method which promotes accessibility of findings facilitating communication between technical and legal parties, peer review, logical reconstruction, jury interpretation, and error detection.

1. Introduction

Forensic investigators are frequently required to gather large amounts of data from a diversity of seized devices, online forums and/or cloud storage for the investigation of criminal cases. This exponentially growing volume, increasing variety and complexity of data to be analysed, known as a "big data problem" (Noel and Peterson, 2014), imposes numerous challenges to investigators (Quick and Choo, 2014). This data typically contains numerous pieces of evidence of different types collected using a variety of individual forensic tools and techniques such as hard drive evidence, mobile device evidence, social media evidence, physical evidence, and evidence from interviews. Tools provide reconstruction capabilities limited to their specific domain, e.g., timeline analysis of files metadata, or relational analysis of who communicated with whom. It is left to the human investigator to reason about how evidence of different types are logically connected and how they fit together in the case's "big picture" to be able to produce a summary narrative of findings and of conclusions. However, humans' ability to make sense of a large volume of evidence is restricted, and unlikely to keep pace with trends of increasing data volume for single cases (Quick and Choo, 2014), not only impacting on forensics practitioners but also on the ability of other stakeholders to logically understand and process findings. Therefore, there is a need for methods and tools to better equip digital forensic investigators with logical analytical capabilities.

Alongside the above mentioned phenomenon, there has been a push both in the domain of Forensic Science and of Digital Forensics (DF) to increase rigor, standardization and transparency in practices and reporting (see for example the Codes of Conduct documentation set by the Forensic Science Regulator in the UK (Forensic Science Regulato, 2014)). Such a move is seen as a way of improving the reliability of results produced from any given forensic examination (Casey, 2018). Given the reliance placed upon forensic evidence in many criminal cases, methods which ensure both the robustness of this work and the ease in which it can be interpreted must be seen as beneficial. The European Network of Forensic Science Institutes (ENFSI) published, in 2015, guidelines for evaluative reporting in forensic science aiming to ensure "that the reports capture both the value and the limitations of the findings expressed in a manner understandable to a wide range of users including the Police, lawyers and juries" [5, Page 3]. Aligned with this initiative, the Royal Statistical Society (RSS) published a series

* Corresponding author.
** Corresponding author.
E-mail addresses: v.franqueira@kent.ac.uk (V.N.L. Franqueira), g.horsman@tees.ac.uk (G. Horsman).

https://doi.org/10.1016/j.fsidi.2020.300923

of 4 guidelines related to communication and interpretation of statistical evidence (Royal Statistical Society, 2019). The guide number 3 aims at enriching the "thinking toolbox" of forensic practitioners presenting two graphical methods for inferential reasoning (Roberts and Aitken, 2014) (refer to Section 2). The DF community is arguably at a cross roads in relation to how it proceeds in the reporting of case data. Despite calls to move towards a more forensic science-orientated model for results reporting in terms of providing a "quantifier" in relation to the evidential weight of any given results and methods for offering the concise and accurate description of case-based events, in reality this is a difficult task to achieve (Kwan and Lai, 2008; Council, 2009). The nature of digital evidence means that, in many cases, whilst factual content can be established, there are often areas of uncertainty in regards to the digital actions of an user which are in some cases difficult to quantify (Nordgaard and Rasmusson, 2012). As a result it remains important that DF practitioners understand where any areas of concern may exist within the results of their case and that they are able to structure the elements which their investigation comprises of, in a way which allows them and other parties related to a given case to be able to understand all of the components within it, and how these have been addressed within the examination.

1.1. Problem statement

Digital forensics examinations are complex in nature and structure. Often, multiple elements are collected, examined and interpreted by the practitioner where these results are brought together to address an investigation hypothesis. Dependant on the number of evidential elements of any case, it can become difficult to logically organise all key facts of a given case to allow full, transparent scrutiny and evaluation of the investigatory process both by the practitioner themselves, peers who may undertake review of the work, and those involved with the wider investigation of the case (such as law enforcement, legal professionals, defence council, and jury) (Horsman, 2019).

1.2. Contribution of the paper

The contribution of this paper is threefold. (1) It proposes the use of Toulmin's structured layout for arguments as a practical mechanism for logical reconstruction, where inferences between digital and/or non-digital evidence can be exposed and refined for forensic argumentation. (2) It illustrates Toulmin's model using three case examples that permit exploring its applicability in real world contexts, motivating discussion about potential benefits and limitations. (3) The examples also show the flexibility of structured argumentation for use at different levels of abstraction and for different purposes – e.g., to help support case conclusions and summaries in a readable and accessible way, to help with hypotheses elaboration and falsification, and to focus further investigative work.

2. Inferential reasoning for forensic proof

"Forensic scientists are constantly drawing inferences at all stages of their work" [7, Page 23]. For example, they are the basis to link evidence to proof, determine new hypotheses, decide on next course of actions and draw conclusions.

There are different types of logical inferences which, collectively, represent a "thinking toolbox" (Paul Roberts and Jackson, 2015). They are: deduction, induction and abduction.

Deduction is helpful to reach (i.e., *deduce*) a conclusion anchored on premises known (or widely accepted) to be true. A key aspect of deduction is that a conclusion invariably follows if the premises are sound. This means that, even when new premises emerge or additional premises are added to the pool of premises, the same conclusion must necessarily hold with certainty (Paul Roberts and Jackson, 2015). This form of logic only occasionally happens in forensic practice and in combination with other types of reasoning. For instance, premises resulting from mathematical theorems may allow a specific conclusion to always be drawn – such inferences are considered as "logically robust" [7, Page 42].

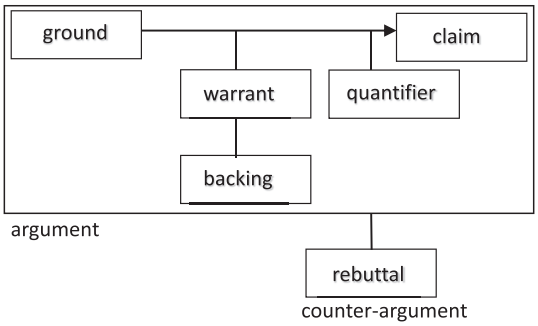

Fig. 1. The layout of an argument (adapted from Toulmin [16, Page 104]).

Induction is the mostly used type of inference in forensics practice [7, Page 43]. It allows reasoning under uncertainty and reaching conclusions which are *probabilistic* (under a certain degree of confidence), *defensible* (if new evidence or information emerges), and *ampliative* (establishing generalisable new knowledge or theories based on case specific knowledge (Pollitt et al., 2018)). Therefore, conclusions are never 100% doubt-free. A baseline must be set to determine the level of uncertainty considered acceptable. For instance, under civil law, the baseline of acceptance required by the standard of proof is expressed as "under the balance of probabilities" (i.e., *more likely than not* (Roberts and Aitken, 2014)) while, under criminal law, the baseline of acceptance is much higher and expressed as "beyond reasonable doubt".

Abduction is a variant of induction, and yet another type of inference useful for forensics practitioners. It allows the formation of hypotheses and their validation by questioning "how" or "what if", a practice known as *falsification* [14, Page 54].

Blending all the above types of reasoning, there are also what is called **common-sense inferences** (Roberts and Aitken, 2014). They often remain implicit, unexpressed, and take the format of generalisations, presumptions, assumptions, story filling-gaps, or beliefs (Twining, 2006). These are "rarely subjected to systematic critical examination" [12, Page 540].

The next section elaborates on a structure useful for inferential reasoning of these different types, including common-sense inferences.

3. Toulmin' structured argumentation

Stephen Toulmin (1958) proposed a layout for arguments which exposes the logical process behind them, and allows reasoning about their validity. The layout is composed of 6 elements, as illustrated in Fig. 1.

A **claim** (C) is what is under evaluation, i.e., is what one wants to establish as true or false. It can be, e.g., a conclusion, a decision, an expert opinion, a hypothesis, or a case statement.

A **ground**[1] (G) can be an evidence collected, a fact, a piece of

[1] Toulmin (1958) calls this element "Data". To avoid confusion, we adopt the terminology used by Haley et al. (2008).

information, data produced, a scientific finding, a legal precedent or an observation which gives support to a claim. A claim not supported by one or more grounds does not constitute an argument [16, Page 106].

A **warrant** (W) is an inferential leap which connects a ground to a claim, i.e., it is a bridge-statement. Such connection can be of different types, e.g., a cause/effect relationship, an empirical generalisation, an interview statement, and any form of "common sense" glue statement regarded as true, as discussed in Section 2.

A **backing** (B) adds credibility or authority to a warrant. It may be, e.g., laws, statistics, test results, regulations, standards and accepted best practices.

A **quantifier** provides a degree of certainty or confidence attached to a claim without considering rebuttals. It can be expressed in terms of qualitative measurements (like "probably", "frequently", "sometimes", "usually") or in terms of quantitative measurements (like probabilities).

A **rebuttal** is a counter-argument which diminishes confidence in a claim, therefore, affecting the initial quantifier (i.e., the confidence) attached to that claim. It can take the form of an exception, a reservation, a new fact, additional evidence and novel information. It is, by its very nature, an *argument* which can be either detailed using elements of an argument layout (Fig. 1) or summarized in the format of a statement. A rebuttal can "attack" a ground, a warrant and, occasionally, a backing.

The layout of an argument by Toulmin is rather static, and has been adapted by Newman et al. (Newman and Marshall, 1991) to incorporate recursiveness. This concept enables the notion of "threads of argument" or "dialog" where counter-arguments (i.e., rebuttals) can be attacked recursively as well, therefore, restoring (to a certain extent) the confidence on the original claim.

Toulmin recognised arguments of the type $G \rightarrow W \rightarrow C$ or $G \rightarrow B \rightarrow C$ [16, Page 124]. In practical argumentation applied to Computing (e.g. (Kelly, 1998; Huhn and Zechner, 2010; Burgemeestre et al., 2010; Cyra and Górski, 2007; Potts and Bruns, 1988; Graydon and Knight, 2008; Goodenough et al., 2013; Haley et al., 2008; Franqueira et al., 2011),), these 6 elements have been used as building blocks with flexibility. For instance, arguments may be abstracted from quantifiers (as we consider in this paper), a claim may be supported by several grounds, and not every single ground may be linked to the claim through a warrant or backing. Such flexibility supports well the "thinking toolbox", allowing representation of the different types of inferences discussed in Section 2.

3.1. Related work

Structured argumentation has been applied to different purposes where the common goal was to expose knowledge and assumptions, which led to a conclusion, to systematic critical scrutiny. This practice helps to build confidence on a target audience that the conclusion reached is justifiably true. For example, it has been applied to build safety cases (Kelly, 1998) and dependability cases (Huhn and Zechner, 2010), to demonstrate compliance to laws and regulations (Burgemeestre et al., 2010; Cyra and Górski, 2007), to trace and justify software design decisions (Potts and Bruns, 1988), to establish confidence in software development (Graydon and Knight, 2008; Goodenough et al., 2013), to show security requirements satisfaction (Haley et al., 2008), and to expose threads of risks/mitigations for risk assessment (Franqueira et al., 2011; Yu et al., 2015).

In fields indirectly related to forensics, Toulmin's argumentation has also been extensively used. For example, it has been applied to help decision making aiming at transparent accountability in cases of child protection (Duffy, 2011). It has been used as an instrument for validation of claims about offenders' profile (Laurence Alison et al., 2003); in this study, results indicated that 80% of the 4000 claims analysed were unsubstantiated and 31% were falsifiable. For a survey of developments in argumentation, refer to Bench-Capon and Dunne (2007).

In the field of Digital Forensics, Toulmin's argumentation scheme has been scarcely applied. Boddington (2012) used it to expose a claim in a child abuse imagery case, and validate it. The validation considered admissibility of the evidence, plausibility of the evidence and corroboration among evidence; these consisted the basis for rebuttals. Pasquale et al. (2013) applied it in the context of forensic readiness for incident response.

Analytical methods and models to validate digital evidence based on other argumentation schemes, and to evaluate uncertainty of forensic inferences have also been proposed in the literature. For instance, they have utilised Bayesian Networks (e.g. (Biedermann and Vuille, 2016; Kwan and Lai, 2008),), Probability Theory (e.g. (Overill et al., 2013),) and Complexity Theory (e.g. (Overill and Silomon, 2011),). The Royal Statistical Society also recommends the use of Bayesian Networks or the Wigmore Chart Method to model inferential relations for forensic proof (Roberts and Aitken, 2014). Those mathematical approaches can be used to complement Toulmin's argumentation model as the *quantifier* element (Fig. 1). The model, however, makes it much clearer with regards to what exactly is under evaluation. Moreover, the fact that Toulmin's arguments are "readable" is another benefit which makes it suitable for all stakeholders and digital forensic practitioners. In terms of visual representation of argument schemes, Wigmore charting (i.e., modified Wigmorean analysis (MWA) (Anderson, 2007)) produces a diagram accompanied by a legend-like key list. Evidences of different types may be connected via directed lines and a set of 8 symbols. For example, "vertical lines indicate tends to support; horizontal lines indicate tends to negate or weaken or tends to corroborate" [36, Page 101]. Roberts and Aitken [7, Page 93] note that a number of related charts may end up nested. Although MWA, likely the Toulmin scheme, exposes the logical structure of inferences, unlike Toulmin's model, its complexity and applicability for practical use for DF remains to be established.

In comparison to developments and proposals in traditional forensic science evidence evaluation techniques (see (Iyer and Lund, 2017; Berger and Slooten, 2016)), the DF community has yet to see significant in-roads in the creation and acceptance of methods for determining uncertainty in any given digital evidence. Whilst as noted above, some approaches have been proposed, there are arguably none which are widely adopted.

The next section illustrates Toulmin's scheme for practical argumentation using three example cases.

4. Case studies

In order to illustrate the benefits of structured argumentation, the following three examples are offered.

4.1. Case 1

Case one is an advanced-fee fraud case judged around 5 years ago. The report of the case has been made available to one of the authors by law enforcement at the time.

The case involved a gang of criminals which operated for at least 6 years. Its participants were based in several countries, but targeted elderly people in the UK and the USA. The investigation was a big operation led by UK agencies where the defendant, leader of the gang (hereinafter called suspect 'X'), resided and was arrested from. The investigation involved hundreds of interviews with victims, a

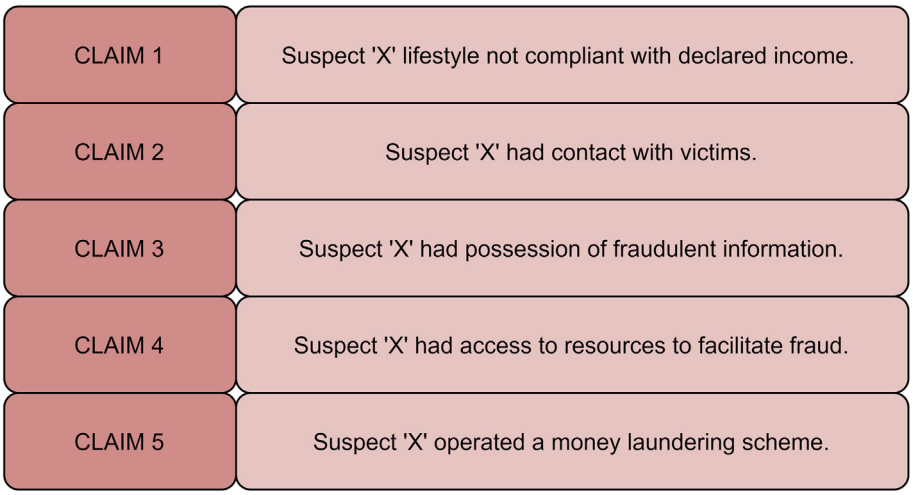

Fig. 2. Claims for case 1 – an advanced-fee fraud case.

number of search and seize operations, and a high volume of money transactions. Several mobile phones, loose SIM cards, laptops, USB sticks, and a bulk of paperwork containing PII (personal identifiable information) and material related to fraud were collected from the suspect's address at the time of arrest.

This case was selected to illustrate structured argumentation applied to a large case. The initial allegation which triggered the arrest of suspect 'X' was another type of crime; only during the investigation, digital forensic examiners realised that they faced an advanced-fee fraud case. Fig. 2 shows several claims which are typical for this type of crime. They are the basis for a logical reconstruction and validation of evidence.

Fig. 3 provides a refinement of Claim 2. It starts with the fact (captured in Ground 1) that suspect 'X' had in his possession, at time of arrest, a number of mobile phones (n mobiles) and a number of loose SIM cards (m cards). There is another fact based on evidence collected, and captured in Ground 2: all the seized mobiles and SIM cards were linked as well since they had contact entries matching each other.

There are three warrants (Warrant 1, 2 and 3) which link the claim to Ground 1, i.e., they link victims to the seized mobiles and SIM cards. These are evidence which show information on the devices related to known victims, and Western Union (WU) reference numbers recovered from these devices associated with known victims and money transferred by them (according to paperwork handed over by known victims to the Police). Warrant 2 is supported by statistics (Backing 1) showing that WU and Money Gram are preferred method by criminals for fraudulent activities. This is the case because money can be transferred without any association with a bank account, therefore, transactions leave no trail apart from their reference numbers (MTCN (Money Transfer Control Number) in case of WU transfers). Therefore, the main argument for claim 2 is contained within an outer box with the claim, grounds, warrants and backing.

Fig. 3 also shows Rebuttal 1 as a counter-argument expelled out by suspect 'X' in interview after arrest. The suspect affirmed that he owned only one of the mobile phones seized and received it as a gift one year before (therefore, implying having no connection with the other devices and SIM cards seized). However, Rebuttal 2 brings forward the (counter-) counter-argument that all of the phones and SIM cards contained evidence linking suspect 'X' with family members prior to one year before being seized. Rebuttal 3 links all phones and SIM cards based on the uncovered fact that they contained photos with associated gang agents. These rebuttals are backed by testing which showed no signs of photo tampering and of them been downloaded. Rebuttal 4 adds another counter-argument to Rebuttal 1 – suspect 'X', at the time of arrest, when search and seize were taking place, confirmed ownership of the seized phones and SIM cards. Therefore, rebuttals 2–4 restore confidence in the initial argument (i.e., Claim 2).

The first case highlights the ability of using structured argumentation for logical reconstruction of a large case, well beyond the role fulfilled by case management tools. It allows the breakdown of the case into a set of claims and the organisation of evidence of different types, acquired at various points in time during the investigation, in a logical but still readable way. The rebuttals are not spelled out in full arguments here; they are summarised in statements although one is supported by a backing. Nevertheless, together, Rebuttals 2–4 might be enough to counter Rebuttal 1, and restore confidence on the initial claim.

4.2. Case 2

Case two (set out in Fig. 4) provides the structured argumentation based on the facts of the Mitesh Patel investigation reported via the media.[2] To stress, some additional details of this case have been assumed or added via creative licence in order to demonstrate the application of structured argumentation. This case was chosen given it provides a vehicle for discussion when a mix of digital evidence types are present within an investigation, where all are crucial to determining the activities of a suspect and each supplement the overarching investigation.

Fig. 4 commences with the defining of a primary Claim, that suspect 'X' murdered victim 'Y', where through an examination of available sources of digital evidence, this claim may be capable of being addressed. There are three supporting facts which underpin this investigation, notably that 'Y' is dead, 'Y's' body was located at their jointly owned dwelling and that both 'X' and 'Y' were married. Such facts exposes their underpinning relationship and are confirmed via the course of available documented records and Police interviews. From the outset it is warranted that 'X' was in the dwelling at the point in which 'Y' was still alive on the day of the incident. At this stage, a suspect rebuttal is faced whereby 'X'

[2] https://www.gazettelive.co.uk/news/teesside-news/murder-accused-mitesh-patel-cheated-15409286https://www.bbc.co.uk/news/uk-england-tees-46348189.

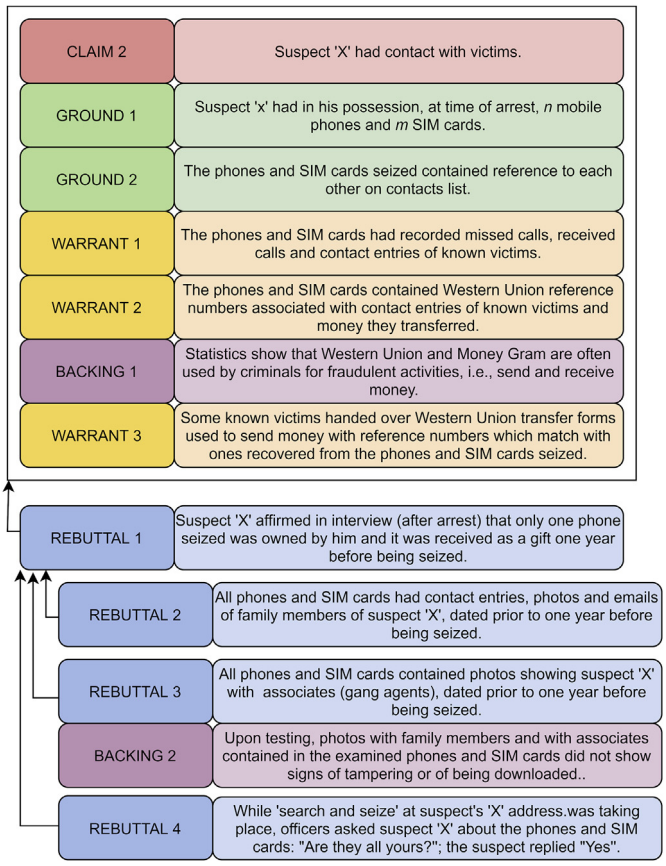

Fig. 3. Refinement of claim 2 for case 1.

affirms to have been out of the dwelling at the time of 'Y's' death, came home to find 'Y' dead and went outside the dwelling to contact the Police. This initial argument (contained in an outer box) provides the underpinning structure from which subsequent investigatory processes are both based, and which they must address. Any rebuttal raised is a question towards the validity of any assumed facts of the case which legal processes may seek to rely upon. As a result, rebuttals must be addressed both in terms of counter-argument or by acknowledging the potential validity of the original rebuttal itself.

An unacknowledged rebuttal provides a potential threat to the reliability of any set of results/hypotheses derived from the investigation. As demonstrated here, Rebuttal 1 is countered via the use of CCTV evidence. Facts regarding the implementation of CCTV are established in order to dispute 'X's' rebuttal regarding their actions. Backing in the form of testing and validation procedures should be offered to provide support for warrants (whenever possible), increasing their evidential weight. As part of the course for all investigations, multiple rebuttals may be raised, each in need of being addressed. In turn, multiple counter-arguments may exist to address a single rebuttal.

Case 2 highlights two further challenges faced by the practitioner when describing their evidence. If we consider Rebuttals 2, 3 and 4, each are addressed with Backings 1–4. Yet, arguments could be made to neither 'out-weigh' the other in terms of argumentation. Cases for either can be made, despite arguably that robust testing should overcome arguments of data inaccuracy. Nevertheless, a possibility of the Rebuttal's truth still remains, regardless of its strength. Herein lies the need for a "quantifier"; a measure of certainty. As previously stated (in Sections 1 and 3.1), this

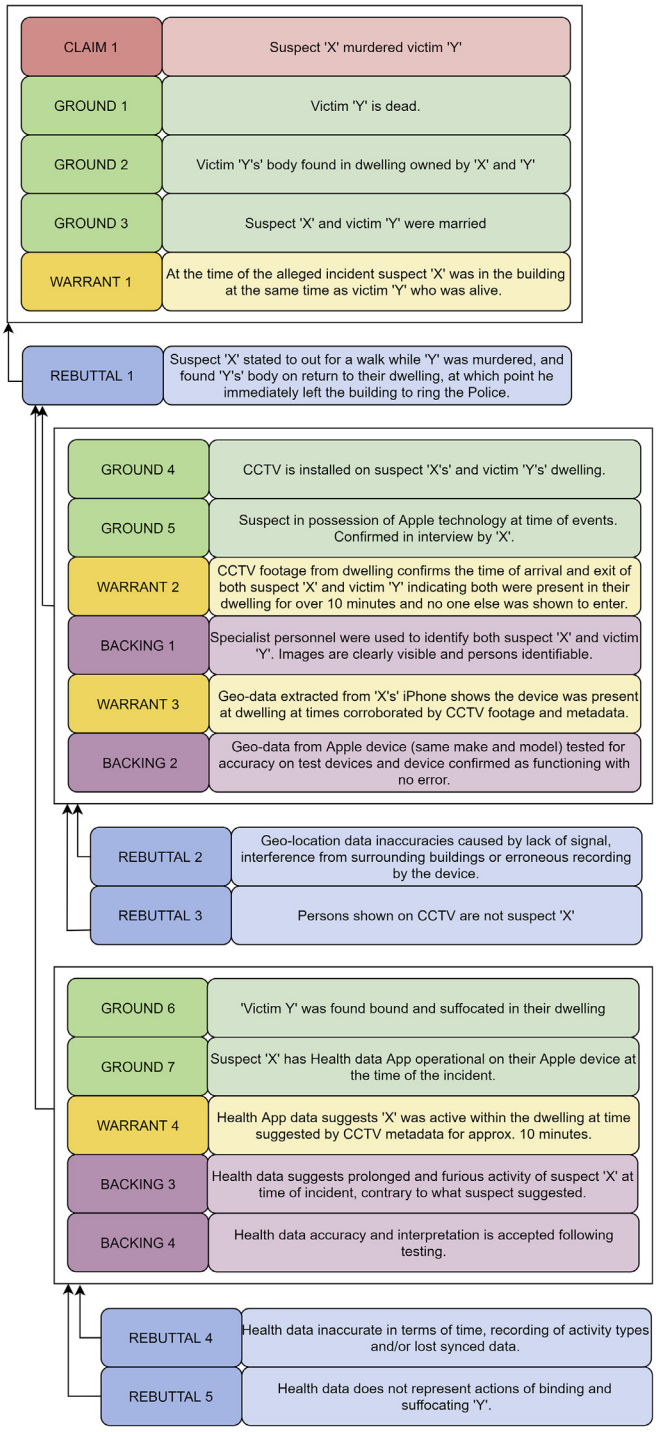

Fig. 4. Structured argumentation applied to case 2.

measurement is currently one of the key research problems for the field of DF to date, one which is arguably still to be addressed.

The second challenge relates to Rebuttal 5, a valid rebuttal without refute backed by scientific evidence. Rebuttal 5 is arguably controversial, where counter claims are capable of being made in various forms. Yet, the point which is being made is valid, there is no way to confirm that the health data presented represents an act of murder. This is a matter of subjective interpretation, supported by all the presented facts of the case.

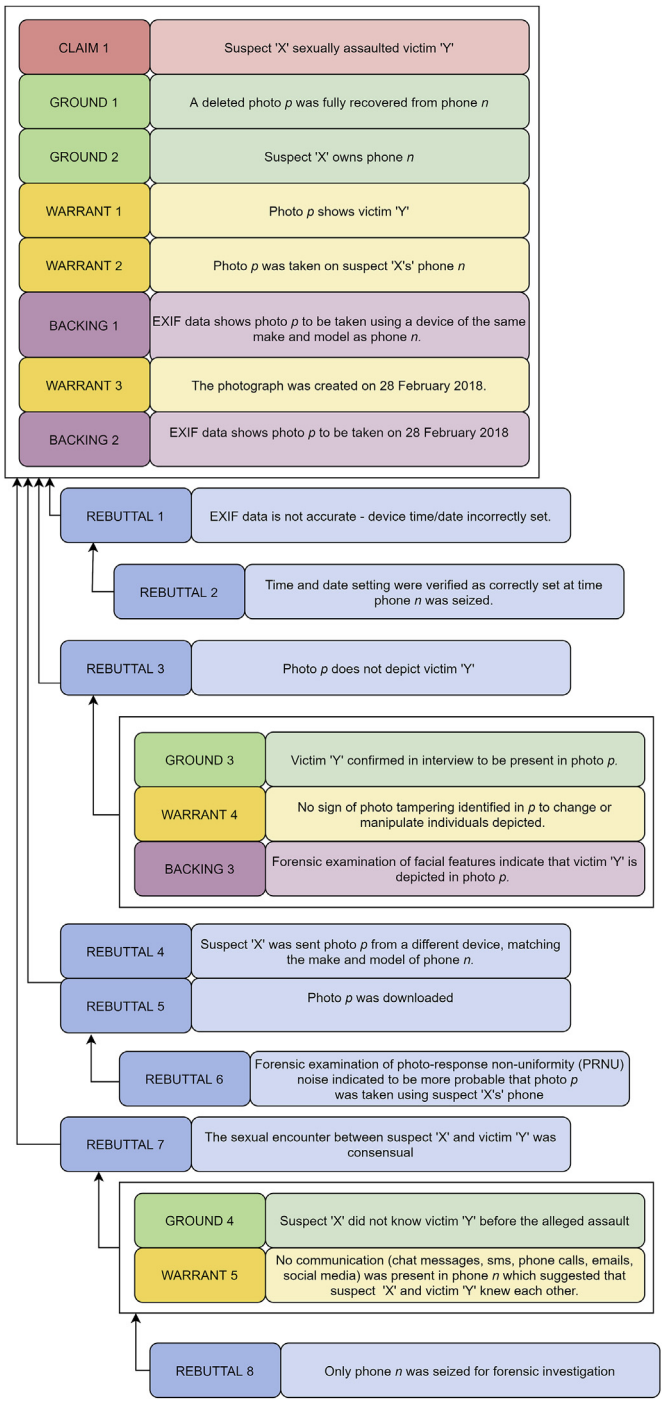

Fig. 5. Structured argumentation applied to case 3 (adapted from (Casey, 2018)).

4.3. Case 3

Case three is adapted from Casey's presentation in (Casey, 2018). Fig. 5 provides a breakdown of a fictitious sexual assault scenario as introduced by Casey involving the discovery of a picture on a suspects phone of a victim who they have allegedly sexually assaulted. This case demonstrates the development of a primary argument, followed with the raising of eight rebuttals.

The founding claim lies with an alleged sexual assault committed by suspect 'X' against victim 'Y'. The grounds of this claim lie with the discovering of a deleted image file (photo *p*) recovered from phone *n* and that 'X' owns '*n*'. Warrants (which link the claim to Grounds 1 and 2) include both the contents of *p* showing 'Y' and that metadata suggests *p* was taken on *n*, at an appropriate time and date. Both warrants are supported through Backings 1 and 2. This initial argument is based on accepted testing and interpretation of EXIF data retained within.jpg pictures (of which *p* is a type).

Rebuttal 1 questions the procedural elements of the investigations which can be counter-rebutted in Rebuttal 2 providing correct device handling, seizure and processing has occurred. Rebuttal 3 focuses on the content of *p*. A counter-argument for Rebuttal 3 is elaborated as a full argument where Ground 3 is based on the evidence that 'Y' acknowledged their presence in the photo. Warrant 4 accompanying forensic analysis (Backing 3) confirms the authenticity of the photo and 'Y's' presence.

Rebuttals 4 and 5, are also counter-arguments for the original claim, and provide a challenge to the practitioner given *p* was recovered from unallocated space where associated file system metadata is no longer available. Rebuttals 4 and 5 are viable and difficult to factually refute, albeit PRNU data is offered in Rebuttal 6.

Finally, Rebuttal 7 (also counter-argument for the original claim) involves consent. For some activities, consent is not a defence, however it is assumed for the purpose of this hypothetical case that it could be. Here, surrounding circumstances of the case can support a refute of this, including those noted in Ground 4 and Warrant 5 – evidence of 'X' and 'Y's' surrounding interactions and relationship.

This case highlights several arguments "attacking" the original argument, included in the outer box with the claim. The argumentation here uses some simple statement-like rebuttals but also some full counter-arguments with ground, warrant and sometimes backing as well. It documents, and tries to anticipate defense counter-arguments, and exposes how the evidence logically follows.

5. Discussion

This section elaborates on perceived benefits and limitations of *structured argumentation* applied to digital forensics practice (Sections 5.1 and 5.2). It is worth noting, as already mentioned in Section 3, that this paper abstracts from the element "quantifier" of an argument.

5.1. Potential benefits of structured argumentation

Following the proposals made in this work for the use of structured argumentation as a means of displaying the results of a digital investigation, the following assumed benefits of this method are offered.

5.1.1. Decipher-ability

One of the challenges surrounding cases involving digital evidence types is the presentation of often complex and diverse forms of information. Presenting the importance and value of evidential content along with how such information fits within the overall investigation scenario is a goal for the practitioner, but also a difficult task. Using structured argumentation, it is argued that case information can be represented in both a readable and accessible way. This is important as case information is not only in need of interpretation by the DF practitioner but also by legal and law enforcement entities who require investigative findings to be presented in a manner which is digestible to those who do not have the same level of forensic knowledge as the investigating practitioner. It is proposed that structured argumentation as a technique can

offer support.

5.1.2. Logical reconstruction

Structured argumentation is designed to support the logical representation of a scenario. A structured argumentation representation of a case offers a clear understanding of what a practitioner has done, why any process work has been done and the meaning of this work within the greater context of a given scenario. Those viewing such a representation of a case are in a position to see the justification for all investigatory processes undertaken by a practitioner and how they have dealt with rebuttals and counter-arguments to warrants made. In this way, this proposal also offers another approach to reconstruction, one of "logical reconstruction" rather than temporal, relational or functional (Casey, 2011).

5.1.3. Peer review

A structured argumentation approach facilitates peer review at both technical level and legal level. In doing so, the gap between the legal and forensics/technical disciplines is bridged where both sides are able to understand, input to, and critically evaluate the work that has been undertaken. As a result, structured argumentation could also be viewed as a tool for quality management, by allowing increased accessibility to investigatory procedures for the purposes of review.

5.1.4. Jury interpretation

Structured argumentation may also support juries in their interpretation of evidence, allowing them to identify where there is a strong argument and weak areas which may be counter-able. As noted above, this method of case presentation arguably provides a more accessible way into the intricacies of criminal investigation processes.

5.1.5. Error detection

If we consider that structured representations allow the greater ability to peer review by all those entities involved in an investigation, by this very nature there is an increased ability to detect case errors. Any method which increases the transparent scrutiny of both investigation processes and results will arguably provide for a more robust evaluation of findings and a greater chance of detecting any apparent weaknesses in the work undertaken.

5.1.6. Flexibility

The last benefit worth mentioning is the flexibility provided by structured argumentation applied to DF. The case examples discussed in Section 4 show such flexibility to a certain extent. In case 1, it has been used as a mechanism for logical reconstruction for a number of claims as manageable building blocks for the forensic specialist to summarise and support transparent conclusions to be considered by the Court. In cases 2 and 3, it has been used for hypothesis elaboration driving, e.g., focus of further investigative work, preemptively considering likely arguments from the defence council or vice-versa. Furthermore, different purposes may use structured argumentation at different levels of abstraction. On the one hand, it may be used *during the process of investigation* as a working tool to help the forensic specialist to understand how recovered digital evidence fit together at atomic level, or to help preparing for challenging interviews with suspects where details are paramount. On the other hand, it may be used *after the process of investigation* as a tool to help document the results of the investigation at a higher level of abstraction suitable for discussion and appreciation by other stakeholders.

5.2. Potential limitations of structured argumentation

This section discusses three potential limitations of the structured argumentation proposed: quality, risk, and overhead.

5.2.1. Quality of argumentation

Quality is one aspect often discussed in the literature in regards to Toulmin's structured argumentation are convincingness, soundness, and completeness.

Convincingness relates to whether the argumentation is compelling enough to assure an intended audience that the conclusion reached is reasonable (Haley et al., 2008).

Soundness relates to whether the argumentation fulfills the argumentation schema and whether it is based on "true premises" (Graydon and Knight, 2008).

Completeness relates to whether nothing has been omitted that could lead to a different outcome about a claim (Shum and Hammond, 1994).

A known problem in argumentation is the subjectivity involved in identifying arguments and counter-arguments (affecting soundness), and the difficulty in determining completeness. Proposals to reduce these problems have relied upon the help of (1) pre-defined critical questions (Walton, 1996), (2) what-if scenarios (Baroni et al., 2009), (3) expert assurance checks (Graydon and Knight, 2008), (4) guidelines (Lipson and Weinstock, 2019) and (5) how/why questions (Haley et al., 2008). Although we can certainly adapt some of these approaches (e.g. (3) and (4)), attributes more tailored to the domain of digital forensics investigations should be considered.

Boddington (2012) proposed the following attributes to validate argumentation in digital forensics: *admissibility*, *plausibility* and *corroboration*. Both former attributes relate to whether the evidence used in arguments were obtained lawfully, are relevant to the case, are accurate, consistent or unambiguous (often captured by means of backings), while the latter attribute prompts for checks about linkage among evidence (captured in grounds, warrants, backings and rebuttals in relation to a claim) and consistency along the argumentation threads.

5.2.2. Risks involved in or exposed by argumentation

One risk practitioners may face is to be trapped into the dangers of exploring every single theoretical possibility in terms of argument and counter-argument and their refinement with too much detail. This approach would lead to the phenomenon of "combinatorial explosion" (Shum and Hammond, 1994). This means that a certain level of imagination and creativity [7, Page 46], together with experience in using the argumentation scheme, is required to set the correct level of granularity.

Another risk, not really involved in argumentation but exposed by argumentation, is the possibility of unacknowledged rebuttals. In such circumstances, either the investigators may be required to check if any further examination could be done, or they might raise an informed discussion with the Crown Prosecution Service (CPS) regarding when, and when not, to proceed with a prosecution attempt. What is important, however, is to avoid results identified in the study reported by Alison et al. (Laurence Alison et al., 2003) regarding offenders profile where 80% of the 4000 claims analysed were unsubstantiated and 31% were falsifiable.

5.2.3. Overhead of argumentation

Structured argumentation may be seen as an overhead if benefits are not recognised as outweighing drawbacks, i.e., without buy-in from digital forensics practitioners. Undoubtedly, there is a learning curve involved to understand the basic rules and to gain practice with the approach proposed. However, since the approach

does not require any specialised background (theoretical or mathematical) and draws from inferences that forensic practitioners already make subconsciously during their work, a short training should suffice.

Structured argumentation may also be viewed as time consuming (in the already stressed environment of forensic laboratories) and effort draining. It can, however, become a very practical tool to support practitioners all the way through their investigations embedded in case management.

6. Conclusion

This paper proposed the use of *structured argumentation* based on Toulmin's argument layout with 6 elements (5 were used in this paper) for digital forensics practice. It illustrated the approach with three examples based on real cases or on the literature. Despite the need for further empirical evaluation, the method indicated several relevant benefits aligned with the push for a more science-oriented model for digital forensics investigations, which is gaining momentum among the community.

Structured argumentation helps to organise knowledge logically although preserving readability. Therefore, on the one hand, it becomes a way to exercise “logical reconstruction” in contrast to temporal, relational and functional types of reconstruction. On the other hand, it uses a simple structure along with open text making the approach accessible to all parties involved along a case (such as juries and Crown Prosecution Service), helping to bridge the communication gap between technical and legal parties. Another potential benefit relates to a gain in transparency and accountability derived from the fact that structured argumentation exposes knowledge thus facilitating peer review and error detection. By making the logical connections among evidence of different types clearly laid out, it is possible to find inconsistencies, unsupported arguments or counter-arguments, informing decisions and further investigative work.

We plan to develop further the argument element we abstracted from in this paper, namely “quantifier”, and work on the aspect of validation of arguments. The *quantification of evidence* in terms of weight/reliability is a much debated area, which requires careful evaluation and development due to its potential impact on the delivery and presentation of digital evidence. Future developments will focus on evaluating the concept of a quantifier and how this will function in practice.

References

Anderson, T.J., 2007. Visualization tools and argument schemes: a question of standpoint, Law. Probab. Risk 6 (1–4), 97–107.

Baroni, P., Cerutti, F., Giacomin, M., Guida, G., 2009. An argumentation-based approach to modeling decision support contexts with what-if capabilities. In: AAAI Fall Symposium. Technical Report SS-09-06. AAAI Press, pp. 2–7.

Bench-Capon, T., Dunne, P.E., 2007. Argumentation in artificial intelligence. Artif. Intell. 171, 619–641.

Berger, C., Slooten, K., 2016. The LR does not exist. Sci. Justice 56, 388–391.

Biedermann, A., Vuille, J., 2016. Digital evidence, 'absence' of data and ambiguous Patterns of Reasoning. Digit. Invest. 16, S86–S95.

Boddington, R., 2012. A case study of the challenges of cyber forensics analysis of digital evidence in a child pornography trial. In: Conference on Digital Forensics, Security and Law. ADFSL, pp. 155–172.

Burgemeestre, B., Hulstijn, J., Tan, Y.-H., 2010. Value-based argumentation for justifying compliance. In: DEON'2010. Springer, pp. 214–228.

Casey, E., 2011. Digital Evidence and Computer Crime, third ed. Elsevier Press.

Casey, E., 2018. Clearly conveying digital forensic results. Digit. Invest. 24, 1–3.

Council, N.R., 2009. Strengthening Forensic Science in the United States: A Path Forward.

Cyra, L., Górski, J., 2007. Supporting compliance with safety standards by trust case templates. In: Proc. of the ESREL'07 (European Safety and Reliability) Conference: Risk, Reliability and Societal Safety, vol. 2. Taylor & Francis Ltd, pp. 1367–1374.

Duffy, J., 2011. Explicit argumentation as a supervisory tool for decision making in child protection cases involving human rights issues. Practice: Social Work in Action 23, 31–44.

European Network of Forensic Science Institutes, 2015. ENFSI Guideline for Evaluative Reporting in Forensic Science [Online] http://enfsi.eu/wp-content/uploads/2016/09/m1_guideline.pdf. (Accessed 18 September 2019).

Forensic Science Regulator, Codes of Practice and Conduct, 2014 [Online] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/351220/2014.08.28_FSR-C-107_Digital_forensics.pdf. (Accessed 30 September 2019).

Franqueira, V.N.L., Tun, T.T., Yu, Y., Wieringa, R., Nuseibeh, B., 2011. Risk and argument: a risk-based argumentation method for practical security. In: RE'11 Proceedings. IEEE Press, pp. 239–248.

Goodenough, J.B., Weinstock, C.B., Klein, A.Z., 2013. Eliminative induction: a basis for arguing system confidence. In: Proceedings of the 2013 International Conference on Software Engineering. IEEE Press, pp. 1161–1164.

Graydon, P., Knight, J., July 2008. Success arguments: establishing confidence in software development. Tech. Rep. CS-2008-10, University of Virginia.

Haley, C., Laney, R., Moffett, J., Nuseibeh, B., 2008. Security requirements engineering: a framework for representation and analysis. IEEE Trans. Software Eng. 34 (1), 133–153.

Horsman, G., 2019. Formalising investigative decision making in digital forensics: proposing the digital evidence reporting and decision support (DERDS) framework. Digit. Invest. 28, 146–151.

Huhn, M., Zechner, A., 2010. Arguing for software quality in an IEC 62304 compliant development process. In: ISoLA'10: Proc. of the 4th International Conference on Leveraging Applications of Formal Methods, Verification, and Validation - Volume Part II. Springer-Verlag Press, pp. 296–311.

Iyer, H., Lund, S., 2017. Likelihood ratio as weight of forensic evidence: a closer look. J. Res. Nat. Inst. Stand. Technol. 122, 1–32.

Kelly, T.P., 1998. Arguing Safety - A Systematic Approach to Safety Case Management. Ph.D. thesis. University of York.

Kwan, C.K.L.F., Lai, M.P., 2008. Reasoning about evidence using bayesian Networks. In: IFIP International Conference on Digital Forensics. Springer, pp. 275–289.

Laurence Alison, O.E., Smith, Matthew D., Rainbow, L., 2003. Toulmin's philosophy of argument and its relevance to offender profiling. Psychol. Crime Law 9, 173–183.

Lipson, H., Weinstock, C.. Evidence of assurance: laying the foundation for a credible security case [Online]. https://www.us-cert.gov/bsi/articles/knowledge/assurance-cases/evidence-assurance-laying-foundation-credible-security-case. (Accessed 24 September 2019) (May 2008).

Newman, S.E., Marshall, C.C., 1991. Pushing Toulmin too far: learning from an argument representation scheme. Tech. Rep. SSL-92–45, Xerox PARC.

Noel, G.E., Peterson, G.L., 2014. Applicability of latent dirichlet allocation to multi-disk search. Digit. Invest. 11 (1), 43–56.

Nordgaard, A., Rasmusson, B., 2012. The likelihood ratio as value of evidence – more than a question of numbers, law. Probab. Risk 11, 303–315.

Overill, R.E., Silomon, J.A., 2011. Complexity based forensic analysis of the trojan horse defence. In: Proceedings of the Sixth International Conference on Availability, Reliability and Security (ARES'11). IEEE Press, pp. 764–768.

Pasquale, L., Yu, Y., Salehie, M., Cavallaro, L., Tun, T.T., Nuseibeh, B., 2013. Requirements-driven adaptive digital forensics. In: 21st IEEE International Requirements Engineering Conference (RE'13). IEEE, pp. 340–341.

Paul Roberts, C.A., Jackson, G., 2015. From admissibility to interpretation: new guidance on expert evidence. Crim. Law Justice Wkly. 179, 538–542.

Pollitt, M., Casey, E., Jaquet-Chiffelle, D.-O., Gladyshev, P., 2018. A Framework for Harmonizing Forensic Science Practices and Digital/Multimedia Evidence, oSAC Technical Series 0002R1 [Online]. https://www.nist.gov/system/files/documents/2018/01/10/osac_ts_0002.pdf. (Accessed 12 June 2019).

Potts, C., Bruns, G., 1988. Recording the reasons for design decisions. In: ICSE'88. IEEE Press, Los Alamitos, CA, USA, pp. 418–427.

Quick, D., Choo, K.-K.R., 2014. Impacts of increasing volume of digital forensic data: a survey and future research challenges. Digit. Invest. 11, 273–294.

Richard E Overill, K.-P.C., Silomon, Jantje A.M., Tse, R., 2013. Quantification of digital forensic hypotheses using probabilistic theory. In: Proceedings of the 8th International Workshop on Systematic Approaches to Digital Forensics Engineering (SADFE'13). IEEE Press, pp. 71–75.

Roberts, P., Aitken, C., 2014. The logic of forensic proof: inferential reasoning in criminal evidence and forensic science [Online] http://www.rss.org.uk/Images/PDF/influencing-change/rss-inferential-reasoning-criminal-evidence-forensic-science.pdf. . (Accessed 18 September 2019).

Royal Statistical Society, Practitioner guides, [Online] https://www.rss.org.uk/RSS/Influencing_Change/Current_projects_sub/Statistics_and_the_law_sub/Practitioner_guides.aspx, accessed: 18/09/2019. (n.d.) .

Shum, S.B., Hammond, N., 1994. Argumentation-based design rationale: what use at what cost? Int. J. Hum. Comput. Stud. 40 (4), 603–652.

Toulmin, S.E., 1958. The Uses of Argument, first ed. Cambridge University Press.

Twining, W., 2006. Narrative and generalizations in argumentation about questions of fact. In: Rethinking Evidence: Exploratory Essays, second ed. Cambridge University Press, pp. 332–343.

Walton, D.N., 1996. Argumentation Schemes for Presumptive Reasoning. Lawrence Erlbaum Associates, Mahwah NJ, USA.

Yu, Y., Franqueira, V.N.L., Tun, T.T., Wieringa, R.J., Nuseibeh, B., 2015. Automated analysis of security requirements through risk-based argumentation. J. Syst. Software 106, 102–116.