Implementation of Digital Forensics Investigations Using a Goal-Driven Approach for a Questioned Contract

Clive Blackwell
Oxford Brookes University
Oxford, UK
CBlackwell@brookes.ac.uk

Shareeful Islam
University of East London
London, UK
shareeful@uel.ac.uk

Benjamin Aziz
Portsmouth University
Portsmouth, UK
benjamin.aziz@port.ac.uk

Abstract

This paper introduces a new systematic process for describing digital investigations that focuses on forensic goals and anti-forensic obstacles and their operationalisation in terms of human and software actions. The main contribution of the paper is to demonstrate how this process can be used to capture the various forensic and anti-forensic aspects of a real world case study involving document forgery.

Keywords: Digital forensics investigations, KAOS, goals, obstacles, questioned documents

1. Introduction

There is a need to extend the typical digital forensic investigation process to handle increasingly complex cases involving large quantities of data from multiple computers and other devices such as mobile phones and portable storage, and many ways that computers can be involved in crime and leave evidence of wrongdoing. The forensic investigation process must cope with many difficulties inherent in evidence collection and analysis from both intentional and deliberate causes that may cause evidence to be incorrect, incomplete, inconsistent or unreliable.

Among the several existing Digital Forensics Investigations (DFI) processes in the literature, most of the work emphasizes on collecting evidence for the investigation or directly starting with the crime. Existing digital forensic processes generally focus on the different
investigation stages such as collection, preservation, examination, analysis and presentation [1, 2, 3, 4]. In addition, many investigations are bottom-up focusing on collecting and analysing data by a complete search of all supplied media based on keywords or regular expressions. However, it may be inefficient to examine all the supplied media as it can lead to lengthy backlogs. In addition, it may be ineffective as it miss vital evidence because searching for low-level patterns may miss evidence, as Casey shows with several examples where the case cannot be proven by these low-level techniques alone [5].

An appropriate systematic process is missing related to the analyses of the crime scene DFI requirements. Moreover, a forensic investigation should also address issues related to the anti-forensics, particularly when time, cost and resources are critical constraints in the investigation. Our work contributes to this direction by adopting a goal-driven methodology in specifying the requirements of a DFI. More specifically, we initiate the DFI process by systematically identifying the main goals for the investigation and analysing the obstacles that could obstruct these goals. The proposed approach integrates the anti-forensics dimension within the digital forensic investigation process at the level of requirements that overcome the deliberate obstacles.

In this way, our methodology supports existing forensic processes by offering a systematic investigation strategy to manage evidence so that it supports the achievement of the investigative goals and overcomes technical and legal impediments in a planned way.

Many formal methodologies exist for requirements engineering and analysis, including i*Tropos [6] and KAOS [7]. In this paper, we focus on KAOS in line with existing works [8,9]. According to Leigland and Krings [10], such adopting of a formal and systematic approach has several benefits, which can be classified as: procedural by reducing the
amount of data and aiding their management; technical by allowing digital forensic investigations to adapt to the technological changes underlying them; social in that the capabilities of the perpetrators are captured within the social as well as technical dimension; and finally legal in that it allows the expression of the legal requirements of the forensics investigation.

We demonstrate the applicability of our approach by considering a recent case involving alleged document forgery and questionable claims made by Paul Ceglia against Mark Zuckerberg of Facebook [11], where we construct a systematic goal tree analysis of the requirements underlying the DFI in this case. The analysis helps outline the main obstacles to the various claims and evidence that the case investigation revealed, and further proposes how the requirements underlying such claims and evidence are operational by means of investigator activities together with forensic system and software operations.

2. Related Work

There are several works that focus on the forensic investigation process and techniques relating to anti-forensics. Here we provide a brief overview of the approaches that are relevant to our work. Kahvedić and Kechadi [12] present a Digital Investigation Ontology as an abstraction of concepts and their relationship for the representation, reuse and analysis of Digital Investigation knowledge. The ontology model is based on four dimensions: Crime Case, Evidence Location, Information, and Forensic Resource. The approach models the knowledge within the windows registry using keys and values. Reith et al. [13] propose an abstract digital forensics model that consists of nine different components from identification, preparation, analysis, presentation and returning evidence. The model supports future digital technologies for non-technical observers. Hunton [14] uses utility
theory for cybercrime execution and analysis models. The work shows that law enforcement officers could make important uses of cybercrime execution and analysis models when investigating crimes by analysing the evidence regardless of the level of complexity of the committed crime. Carrier & Spafford [15] consider a digital device as a digital crime scene and uses process model for the forensic investigation. The process consists of five categories or phases: Readiness phase, deployment phase, physical/digital crime scene investigation phase and presentation phase. Huber et al. [16] emphasise on the necessity of approaches for crime analysis of online social networks and Cloud-based service types. The approach shows techniques to gather digital evidence from online social networking sites. Harris [17] presents techniques for destroying, hiding and eliminating evidence resources as part of anti-forensic activities. Recommendations (mostly relating to investigators) such as educational level, real-world experience, and willingness to think in new directions, are emphasized for handling anti-forensic issues. Dahuar and Mohammad [18] identify forensic challenges such as time, cost, vulnerabilities of forensic software, victim privacy and the nature of the digital evidence as being the main challenges of an anti-forensic process. Several mentioned works focus on the systematic forensic investigation process mainly with emphasis on collecting and analysing the evidence. Our investigation process differs from these as we initiate the process with identification of goals for the investigation and analysing the anti-forensic issues that could obstruct any stage of the investigation process. Therefore, the process presented in this paper combines both forensic and anti-forensic issues within an investigative framework.

3. The Proposed Process

We propose a systematic process, as illustrated in Figure 1, for understanding investigative processes starting with the crime context analysis and ending with the appropriate actions
for analysing the evidence. It combines the anti-forensic issues during the forensic investigation process so that possible obstructions of the investigation can be identified, analysed and overcome. The process consists of four activities that define major areas of concern for the forensic investigation. The individual activities include the steps concerning the creation of artefacts such as goals, obstacles, evidence and forensic actions relating to the incident. These artefacts are incrementally combined to produce the incident report containing both textual and graphical representations. The process defines roles that take the responsibility for a specific set of artefacts and perform a set of activities within the process in order to produce or modify the artefacts. The activities are performed sequentially, and, if necessary, a number of iterations are performed for individual activities until they are completed adequately.

Figure 1. A Systematic Process for Forensic Investigations
**Activity 1: Understanding Incident Context**

This is the first activity of the proposed process that initializes the forensic investigation process, where the main focus is on understanding the background of the incident context. In particular, a brief overview of the incident includes pre-incident preparation, choosing the investigation team, determining the investigation strategy, discovering the complexity and severity of the incident, and establishing the boundary of the forensic process. Once the incident context is identified, the forensic team makes a plan how the investigation will be performed. This includes choosing a strategy to isolate, secure and preserve the state of the physical and digital evidence. The plan should consider the investigation constraints such as the size of media involved, time and budgetary restrictions, and the availability of resources such as tools, equipment and skills.

**Activity 2: Identify and analyse goals**

Once the incident context is defined, the next activity is to identify and model the goals for the forensic investigation. Commonly forensic investigations have primary goals such as achieving a successful investigation and collecting and preserving the evidence for court. The explicit determination of goals aids the justification and delimitation of the scope of the investigation process. The goals may also include suggesting investigative leads and abandoning leads that are likely to be fruitless, as well as proving cases by discovering and overcoming their potential weaknesses.

The next step of this activity is to analyse the identified goals so that the higher-level goals are refined into sub-goals. In particular, this step considers how various phases of the investigative process link the sub-goals with the main goal and supports the analysis of the incident. For instance, collecting evidence as a goal can refine into gathering evidence from
different systems, devices and the Internet, possibly in unusual locations. In the Ceglia case, there was much information found on his equipment indicating other locations where important evidence might be located, such as in undisclosed email accounts and on third party systems like those belonging to his lawyers.

Sub-goals may be linked by either AND or OR refinement relations to construct the goal model. AND refinement specifies all sub-goals that must be satisfied in order for the parent goal to be satisfied, while OR refinement specifies that any one of the sub-goals is sufficient for the satisfaction of the parent goal [7].

**Activity 3: Identify and analyse obstacles**

Obstacles are the causes that reduce the ability to achieve the goals. Therefore, obstacle identification and analysis refers to what could go wrong during a forensic investigation specifically in relation to evidence collection, preservation and analysis within the forensic process. For a successful forensic investigation, we need to identify all plausible obstacles to determining the facts of the incident. It helps to determine the obstacles in advance, as this facilitates the determination of a course of action to overcome them. We determine the anti-forensic methods and tools to identify the obstacles that directly or indirectly reduce the reliability of the digital evidence. This activity might need specific technology-dependent tools to handle the anti-forensic issues.

This step assesses the potential damage incurred by obstacles for the overall investigation; in particular, the difficulties of finding evidence, exhausting the anticipated forensics investigator’s time and resources, misleading by manipulating essential metadata like hashes and timestamps, and storing data anonymously on the Internet rather than locally.
Generally, forensic investigation evidence should be admissible (i.e., must be able to be used in court), authentic (i.e., original and unchanged), reliable (i.e., correct and accurate), complete (i.e., all relevant evidence is available and correct), and believable (i.e., easy to understand and credible to a jury). An obstacle can also oppose the integrity, completeness, reproducibility, timeliness and believability of both a forensic activity and outputs produced by the activity. Perpetrators can also interfere with the forensic tools such as Encase, FTK and WinHex. Obstacle analysis focuses on understanding what type of obstruction is done by the anti-forensic actions. Therefore, obstruction of any of these properties is an obstacle for the digital forensics investigation.

**Activity 4: Identify forensic actions**

The final activity of the process is to identify the appropriate forensic actions that must be applied based on the critically of the incident. These actions operationalise the goal satisfaction to determine a suitable response strategy to resolve the incident. For choosing appropriate actions, it is necessary to understand the severity of the risk due to the occurrence of the incident and obstacles due to the anti-forensic activities. Risk can have various dimensions depending on individual, organisational or public domain. These dimensions could include financial loss, loss of reputation or privacy, intellectual property theft and others. It is also necessary to consider the legal constraints related to incident notifications (to the regulatory authority) and the quality of the documentation of the investigation’s goals and requirements, before choosing the actions.

For example, the Ceglia investigation examines the authenticity of the contract and the supporting contextual evidence such as relevant emails. An obstacle may exist in that the original evidence of document authenticity is not provided, but the supporting evidence
seems to suggest that such authenticity is plausible at first sight. Hence, the forensic actions would focus on the use of low-level tools to find anomalies in metadata and timestamps.

The selected forensic actions should be implemented for the successful completion of the investigation. The victim organisation should have an active response stance posture on this occasion. This step monitors the effectiveness of the implemented control action.

4. **Ceglia versus Zuckerberg and Facebook**

We use a real world forensic case involving Paul Ceglia who filed a complaint seeking a share of Facebook to demonstrate the applicability of our approach. This section demonstrates the systematic application of the four activities discussed above to aid the forensic investigation with a posterior analysis of the Ceglia case. We indicate later how the same activities could be applied to support a new investigation.

Paul Ceglia is an entrepreneur who engaged Mark Zuckerberg to perform some work on his project called StreetFax around the time Zuckerberg founded Facebook in 2003. Ceglia paid Zuckerberg $1,000 for work on StreetFax and claims he paid $1,000 to fund Zuckerberg’s "face book" project. He produced the ‘Work for Hire’ contract that is apparently signed by himself and Zuckerberg covering the two projects [19]. According to Ceglia, the agreement stated that Ceglia would get 50% of the "face book" project in exchange for funding initial development. Zuckerberg clearly did discuss Facebook with Ceglia, which was supported by multiple email exchanges between the pair.

Based on the Complaint, the court ordered Mr. Ceglia to produce relevant electronic assets such as an electronic copy of the contract, copies of the purported e-mails, and computer and electronic media under Ceglia’s control. The court also issued an Electronic Asset Inspection Protocol for inspecting the collected electronic evidence, involving mainly the
digital forensic evidence, requesting that the investigators check the authenticity and availability of the evidence and provide a report to the court.

**Activity 1: Understand Incident Context**

We examine the evidence for the agreement by investigating the validity of the ‘Work for Hire’ contract and contemporaneous emails provided by Ceglia that support his version of events, along with the subsequent discovered evidence. The main digital forensic analysis is provided by the Stroz Friedberg expert report [11] for Zuckerberg that was made publicly available after its submission into court. In addition, there were several expert reports on the physical evidence, especially those of Gerald LaPorte [20] and Frank Romano [21].

This first activity of the process focuses on understanding the issues relating to the investigation. The main scope for the investigator is to confirm the authenticity of submitted claims by Ceglia relating to the ‘Work for Hire’ contract and purported e-mails, including checking the timestamps and formats of the collected evidence. In addition, the evidence should be forensically sound to support the electronic asset inspection protocol, and, in particular, it should identify if any of the evidence is a forgery.

A crucial first issue is to acquire all of Ceglia’s computer equipment and any other devices used for his dealings with Zuckerberg, such as his parent’s computer that was found to contain the original contract, and to discover and preserve evidence from his online activities including the use of multiple email accounts. The complexity of the investigation mainly arises from the huge quantity of electronic data from different geographical locations and the need to preserve and check all the possible evidence. The forensic evidence was obtained from three hard drives, 174 floppy disks, and 1087 CDs. Relevant evidence could be present in image files, e-mail communication, and from draft and deleted
documents. Appropriate skills and tools existed for the investigation, and we do not consider issues like investigation team management, time and budget here.

**Figure 2. Ceglia Case Overview**

**Activity 2: Identify and analyse goals**

The overall goal in the Ceglia case is to prove the ‘Work for Hire’ contract a forgery, which will cause his claim for part ownership of Facebook to fail as it is the only supplied evidence capable of proving his version of events. The main goal can be refined into sub-goals related to the production and analysis of all relevant computer and electronic media including the purported contract and e-mail. All electronic evidence should be forensically sound.

An initial generic goal tree for document forgery developed from previous similar cases can help determine an initial approach that focuses attention on the likely evidence and its
potential locations. There are three branches of the goal tree for demonstrating the invalidity of the ‘Work for Hire’ document, whereas a further branch attempts to show the case should fail on legal technical grounds because of withheld or spoiled evidence.

In theory, it is sufficient to prove forgery in one way only, and so the goal is an OR refinement of these four possibilities as shown in figure 2. However, we should consider proving forgery in multiple ways to make the case resilient to unanticipated new evidence and legal challenges. We decomposed all four branches of the goal tree, but chose to explain the most convincing branch that makes the fewest assumptions by directly attempting to show the contract a forgery, as then Ceglia’s case must fail because the purported contract was the only convincing evidence.

**Activity 3: Identify and analyse obstacles**

Several obstacles impede the goal of the investigation to show the ‘Work for Hire’ contract a forgery. Obstacles to the direct proof of forgery are the unavailability of the original documents, and instead the use of copies to support the contract. We show how the obstacles were overcome within the analysed branch of the goal tree in figure 3, which forms one quarter of the overall tree. The obstacles slope the opposite way to goals, and are coloured red and appear darker than goals when viewed black and white. We show further goals to overcome many of the obstacles as children of the obstacle nodes, but any obstacle without any child goal node has not been defeated. The evidence is convincing in this case; however, in other cases of alleged document forgery, the discovered obstacles to direct proof may be considerable, so that the other branches giving weaker substantiation may be investigated instead, as shown in figure 2.
Figure 3. Decomposition of part of goal tree for Ceglia vs Zuckerberg to prove the ‘Work for Hire’ contract a forgery
The two main pieces of evidence supplied by Ceglia are the alleged ‘Work for Hire’ contract and supporting emails. There is the apparent authenticity of the contract based on its content, and the supporting email that appears to give a consistent account supporting Ceglia’s version of events. An important obstacle to proving forgery is that the original contract and supporting email are not available. Investigation therefore has to rely on secondary evidence from deleted and draft contract files, and e-mail cut and pasted into a Word document. However, the purported emails have formatting differences in the email headers that are inconsistent and indicate that content of e-mails are manually typed and edited. There is evidence of possible spoliation, in particular multiple reinstallations of the Windows operating system and that relevant files were deleted and overwritten. Therefore, the obstacles for this case are mainly unverified and incomplete evidence with the primary evidence being unavailable as there was no exact copies of the Work for Hire Document found on the investigated media.

Activity 4: Identify and execute forensic actions

The forensic actions need to achieve sufficient goals and negate necessary obstacles to achieve the primary goal of showing the ‘Work for Hire’ contract a forgery. Most of the nodes are OR branches, so there only has to be one successful path to the root from a leaf node holding sufficient evidence, and there are always alternatives to undefeated obstacles. As mentioned before though, it is safer to prove the case in multiple ways. We have decomposed each branch of the primary proof goal completely to demonstrate the case in four different ways, as we now indicate briefly and show in figure 2.

In the first branch, there was no independent evidence for the ‘Work for Hire’ contract, save the eyewitness that witnessed a contract signature, but the StreetFax contract had better
provenance and so it is more likely that it was signed. The third branch contains convincing evidence for the authenticity of the StreetFax contract, which as we indicated before shows the ‘Work for Hire’ contract a forgery, as there was only one contract between the two parties. Broom [22] in his expert report for Ceglia gave an alternative hypothesis that Zuckerberg or his agents could have forged the StreetFax contract. However, this is convincingly refuted by the discovery of the StreetFax contract independently in Ceglia’s email and on a server belonging to Ceglia’s lawyer from 2004 six years before the start of the case [11 pp 19-21]. In the fourth branch, we check for evidence of spoliation and withholding of evidence. The evidence includes deletion of relevant files such as the StreetFax contract and draft ‘Work for Hire’ documents, and deletion of email and deactivation of e-mail accounts in an apparent attempt to avoid discovery. System evidence relating to spoliation was found with the multiple reinstallations of the operation system that overwrote the data on the hard disk, but this could have an innocent explanation.

We now discuss the second branch showing the evidence for forgery of the ‘Work for Hire’ contract shown in figure 3. Although the content appears plausible, the metadata provides the evidence for forgery. Several actions lead to convincing evidence including:

- **Checking for inconsistency in e-mails**: The emails give a plausible account and support Ceglia, but the headers are inconsistent demonstrating that they have been changed [11 pp 29-31]. Generally, when an e-mail is created header information is automatically generated. Therefore, inconsistency in the e-mail headers in the copies supplied in a Word document by Ceglia means that the e-mail was not cut and pasted appropriately from the original authentic source, but may have been
fabricated with some extra information that was not in the original e-mail. This does not appear to have an innocent reason and gives strong evidence of forgery.

- **Verification of timestamps and file size**: Timestamps are saved when a file is created, accessed and modified, as well as the time of sending e-mail. This requirement supports checking the authenticity of the contract, as the document must be created before sending as an attachment by e-mail, and the size of the attachment must be same as the copy preserved in any media. In addition, the timestamps of the draft contracts show several inconsistencies as discussed at length in the Stroz Friedberg report [11 pp 33-36], and shown at the bottom of figure 3. The anomalous file timestamps do not appear to have an innocent explanation and together with evidence of backdating the system clock seem conclusive. Similarly, the purported emails contain a Date line that contains the date and time the email was sent followed by an inconsistent time zone for the time of year the email were written, but the anomaly for daylight saving hours could have a benign explanation.

- **Checking system logs for clock anomalies**: Typically, modern operating systems automatically adjust the system clock. However if the user changes the system clock and the difference of the clock with the network clock is more than 15 hours then an error message is recorded in the system event log. Together with the same 94-day inconsistency in the ordering of restore points, they show the system clock has been altered. Finding such independent evidence of time anomalies is suggestive that the system clock was deliberately altered.

- **Confirm matching of contract files**: This action verifies that the contract in two different media is the same because they have the same hash value. The comparison
of the files can be executed by generating hash values for the TIFF image files found in the attached e-mail and the other discovered on media supplied by Ceglia.

- **Obtain the up-to-date original ‘Work for Hire’ contract**: The ‘Work for hire’ contract is crucial for Ceglia’s case, and it is therefore necessary to obtain the original document and its absence from the Ceglia media is notable. Note also that there were many different inconsistent drafts discovered on the Ceglia media, which signify doubt about the authenticity of the document.

- In addition, the physical tests demonstrate beyond reasonable doubt that the ‘Work for Hire’ contract was created using a fake page 2 attached to the legitimate page 1 from the StreetFax contract. This was demonstrated in multiple ways by several different experts especially in the reports by Gerald LaPorte [??] and Frank Romano [??], as shown in the right-hand branch in figure 3, but we do not discuss this further as it is outside the digital forensics boundary.

The artefacts produced from the previous activities are incrementally combined to produce the forensic investigation report. The report should also include the status of the implemented forensic actions and their effectiveness. Although, the Stroz Friedberg expert report for Zuckerberg [11] was comprehensive and highlighted all the relevant points, a more systematic exposition of the overall argument would have given a clearer narrative.

**Discussion**

There are many useful points of a systematic goal tree analysis using KAOS that can be incorporated into forensic investigations illustrated by the Ceglia case study including:
• Reuse of knowledge about previous similar cases, shown by the common upper branches of the goal tree

• Formulation and execution of an investigation strategy, where there can be advance planning to overcome known obstacles, such as having to analyse copies of the contract and email rather than the originals

• Helping formulate and analyse alternative hypotheses, such as whether the anomalies in the time zones in email headers were indicative of fraud or could have alternative innocent explanations

• Clarifying the reliance on assumptions. The opposing parties agreed that there was only one contract signed by Zuckerberg, which is an assumption needed to prove the ‘Work for Hire’ contract a forgery by showing the StreetFax contract is authentic

• Helping explain the overall argument for the case by combining all the claims in each branch into a coherent, comprehensive and consistent narrative

One limitation is the absence of detailed analysis of timelines and timestamps that is crucial to most investigations. The goal tree decomposition may suggest possible avenues of investigation by creating requirements to discover anomalous temporal metadata, but they would be broad and possibly difficult for an analyst to perform. We plan to investigate how the goal tree analysis may inform and integrate with a timeline tool.

5. Conclusion and Further Work

This paper presented a new systematic process for DFIs, which consists of four main activities for understanding the context of incidents, the identification and analysis of the goals needed in the DFI, the identification and analysis of any obstacles to the DFI process, and the identification and execution of the required actions and operations that must be applied (by the investigators or their software) in order to satisfy the main investigative
goals. We used the process to model the real world case study of Ceglia versus Zuckerberg and Facebook involving alleged contract forgery.

There are several further directions for expanding on the work presented here. These would include the definition of a framework for the extraction of common patterns for describing goal-driven DFIs, their obstacles and their operationalisation. Document forgery as in the Ceglia case study would be an excellent domain to investigate. A potential limitation of the paper is that we have used an existing case study, which is overcome largely by our comprehensive modelling of the entire case, and would be further aided by a more general model for document forgery that could be applied to new cases.

Additionally, we plan also to utilise formal languages and formal verification tools to provide more rigour in specifying a forensic investigation and in providing proof-of-evidence that certain events or relevant information are of a high quality of assurance to prove that our claims hold or not.

References

International Workshop on Systematic Approaches to Digital Forensic Engineering (SADFE), IEEE, 89-93, 2010.


