A Case-Based Reasoning Method for Locating Evidence During Digital Forensic Device Triage

Graeme Horsman, Christopher Laing, Paul Vickers*

Northumbria University, Department of Computer Science and Digital Technologies, Pandon Building, Camden Street, Newcastle-upon-Tyne, NE2 1XE, U.K.
Tel. +44(0)191 243-7614

Abstract

The role of triage in digital forensics is disputed, with some practitioners questioning its reliability for identifying evidential data. Although successfully implemented in the field of medicine, triage has not established itself to the same degree in digital forensics. This article presents a novel approach to triage for digital forensics. Case-Based Reasoning Forensic Triager (CBR-FT) is a method for collecting and reusing past digital forensic investigation information in order to highlight likely evidential areas on a suspect operating system, thereby helping an investigator to decide where to search for evidence. The CBR-FT framework is discussed and the results of twenty test triage examinations are presented. CBR-FT has been shown to be a more effective method of triage when compared to a practitioner using a leading commercial application.

Keywords: Digital forensics, triage, case based reasoning, Bayesian reasoning, knowledge reuse

1. Introduction

Triage is a technique used in many disciplines, most notably in the field of medicine as a way of prioritising injured or ill patients for treatment [25]. It can be viewed as a way of organising a workload to allow for the efficient allocation of available resources [10]. More recently it has found its way into the Cybersecurity lexicon where it is used to categorise threats [35] allowing an organisation to determine during incident response which events should be dealt with first based on their severity and available resources [11]. When applied to digital forensics (DF), the meaning of triage differs depending on the context in which it is applied but, as Casey [15] suggests, its goal is to speed up an investigation by attempting to identify evidential exhibits and files quicker.[1]

Triage can mean the prioritisation of physical exhibits for investigation (for which we coin the term *high-level triage*) but it can also signify an interrogation of data held on a target digital device (which we call *device triage*, or DT). In addition to known facts about a case, decisions made during high-level triage are also commonly based on a suspected offence type or the physical location of an exhibit at the scene of a crime.

DT involves the identification of evidence on a suspect system from amongst non-case relevant data, while allocating as few resources as possible [32]. DT is employed to speed up a DF investigation, attempting to cut down the time it takes to identify evidence and is the focus of this article. Cybercrime and the use of technology to commit crime are on the increase [17]. Bem *et al.* [9] suggest that this is leading to increased caseloads which, in turn, are causing difficulties in the field of DF.

High tech crime units are experiencing investigation backlogs [42] and DF practitioners are facing increasing pressure to effectively manage their workloads and process their investigations more efficiently. This has led to DF software developers championing their DT tools as a way of increasing investigation efficiency [1, 2, 19]. Pollitt [33] argues that these tools have fallen short of the requirements needed to deal with current DF cases. DF practitioners have yet to consistently use DT for investigating digital media.

The United Kingdom's Association of Chief Police Officers (ACPO) good practice guide for computer-based electronic evidence [3] provides guiding principles for DF investigations. However, ACPO [4] have been cautious to recommend DT owing to the perception that it carries an increased risk of missing evidential files. We argue that DT has the potential to reduce mounting case backlogs but to do this DT techniques must be improved.

Apprehension over the use of DT may be due, in part, to a limitation of many current DT applications, namely the use of pre-coded and fixed scripts. Until a vendor releases a software update these scripts can remain unchanged for months. This means a DF practitioner is limited to using the same evidence gathering script in their DT investigations, even though the way in which a particular offence is committed may have changed. In such a scenario the chance of missed evidence is increased, and reluctance to conduct DT is understandable. Such scripts are frequently derived from an estimate of which data types would

*Corresponding author
Email addresses: graeme.horsman@northumbria.ac.uk (Graeme Horsman), christopher.laing@northumbria.ac.uk (Christopher Laing), paul.vickers@northumbria.ac.uk (Paul Vickers)

[1] Abbreviations used: CBR-FT = Case-Based Reasoning Forensic Triager; DF = Digital Forensics; DT = Device Triage; ERR = Evidence Relevance Rating; PRF = Primary Relevance Figure

be likely to reside on a system for a given offence. Consequently, this approach opens up the DT process to criticism and, arguably, to an increased risk of investigation errors.

DT can be applied at both the scene of a crime (pre-seizure) and within the confines of the forensic laboratory (post-seizure) [10], each with different purposes and consequences. Pre-seizure DT aims to eliminate devices from an investigation and carries the greatest risk of missed evidence as devices of low priority may be omitted from an investigation due to time and resource constraints. Any evidence missed during pre-seizure DT could have serious ramifications, as evidence may be left with a guilty suspect. Even though post-seizure DT occurs in a secure laboratory environment and all the potential evidence is available (providing pre-seizure DT has not occurred first), it is possible that an inadequate DT could still prevent an exhibit from proceeding to a full examination, and thus, failure to find the evidence. Current DT applications have been criticised for lacking the investigative experience needed to extract relevant data from a system [5].

In order to improve DT a greater emphasis must be placed on achieving higher precision in the identification of relevant evidence. One possible way to achieve this involves the reuse of knowledge from past DF investigations in order to establish where evidence is commonly found. Patterns of suspect activity can help to isolate key areas of a system to be targeted during DT. Reusing investigation data in this way would allow DF practitioners to extract data from a system based on the probability of evidence being present in particular locations. This could transform the current approach of guessing where evidence may be located into an informed decision based on where evidence has been regularly found in the past.

This paper presents our Case-Based Reasoning Forensic Triager (CBR-FT) which uses patterns of behaviour to detect evidential activity in DT. Through the use of past DF case results we demonstrate the tool's ability to target evidential files. We offer CBR-FT as a method of implementing knowledge reuse for the benefit of DT. It also goes some way to answering Pollitt's call for triage to be treated "as a formal process that can be measured for efficiency and efficacy" [33].

Current approaches to triage are discussed in Section 2. Sections 3–7 introduce the new CBR-FT framework and its functionality. The results of twenty triage examinations are presented in relation to the offence of fraud (Section 8). The precision and recall of CBR-FT during DT are discussed and compared to EnCase Portable [21], a commercial DF DT application.

2. Approaches to Triage

Research often focuses on the development of basic frameworks that highlight general approaches to high-level triage [38]. There is little published research in the area of DT and even when techniques are presented their functionality often remains insufficiently tested as they are rarely used in actual DF DT investigations [28]. Commercial applications for DT do exist (e.g., [1, 2, 19]) but share the same fundamental weaknesses. Using predefined scripts, an investigator executes the application and data is automatically collected for review. These scripts are often coded to search for and retrieve specific evidential artefact types. This can lead to very large quantities of data being recovered which the DF practitioner must then interpret.

In addition, commercial approaches to DT frequently use hash set analysis and keyword searching to identify evidential files without the need for a practitioner to view file content. Keyword searching is a process of looking for relevant alphanumeric strings which may be contained within evidential files on a system [7]. Hash analysis involves the comparison of two file hash values in order to establish a match [39]. A file hash value is generated by an algorithm which produces a character string which is unique to a given binary file. The MD5, SHA-1, and SHA-256 hash functions are all commonly used in DF analysis. Except in very exceptional circumstances [41], two files with matching hash values will contain exactly the same binary data.

Using a hash set (a collection of hash values from known evidential files) to perform a hash analysis of a target system has the potential to lead to the identification of evidential files. This process is often used in the identification of indecent child images [16] through the use of the Child Exploitation and Online Protection (CEOP) hash sets. Hashing can be used not only to identify relevant files but also to filter out known non-evidential files (e.g., standard operating system files) which could reduce the overall time needed to carry out an examination. However, as discussed below, both hashing and keyword searching approaches can limit the effectiveness of DT because they are too restrictive, leading to a failure to identify digital evidence.

2.1. Limitations of Hash Analysis for DT

During hash analysis, should any aspect of a target file be altered, (e.g., altering one pixel in a picture) the file's hash value would change even though the target file is essentially the same, thereby rendering hash analysis ineffective. In addition, hash sets must contain hash values of files known to be evidential. What constitutes evidential value in one case may not in another, especially in crimes such as fraud. This is unlike offences involving indecent images, for example, where a single file can be evidential regardless of the system on which it is found. For example, a hash value from one particular corporate financial file is unlikely to be of value when dealing with an investigation from another company. In reality, such files would not exist in both scenarios as these files maintain company specific data and make hashing ineffective. Creating a hash set of files from company A would therefore be unlikely to highlight files found in company B.

A limitation of hashing is that it is defeated when a copy of a file is altered slightly (e.g., by cropping a photograph). Kornblums [27] piecewise hashing approach uses a "context triggered rolling hash" to highlight known files which have been modified or amended slightly but relies on *a priori* knowledge of the modified files. Perceptual hashing is one way to combat these limitations as it makes judgements about the human perceptual similarity of files rather than by comparing their binary representations. Perceptual hashing offers some flexibility as the user can identify files which maintain a certain level of similarity as opposed to an exact match [26]. However, perceptual hashing

maintains processing overheads which would increase the overall length of the DT process, thereby negatively affecting the efficiency of the investigation.

Therefore, hashing techniques are only helpful in limited DT scenarios. Hashing (both normal and perceptual) works well for certain file types, particularly picture files (as they often remain unedited by a user) and therefore lends itself to particular offence types concerning these types of files. However, there are many other crime types (e.g., fraud) which do not involve the types of file for which hashing is well suited and require, instead, semantic analysis of file content.

2.2. Limitations of Keyword Searching for DT

Keyword searching also raises issues for DT. First, keyword searching can take a considerable amount of time which works against the goal of triage which is to prioritise cases as quickly as possible. Second, defining keyword dictionaries can prove problematic because, in many DT investigations, the surrounding circumstances of the case may not yet be fully known, making key terms difficult to identify. Third, the key terms identified are subjective, being based upon the experience of the investigator, leading to varying degrees of success. It must be noted, however, that techniques to automatically generate key term dictionaries do exist [36] including ontological structures which create and maintain domain related keyword knowledge bases [20]. Although such techniques have the potential to be manipulated and applied in a DT investigation, the difficulty remains in automatically generating a key term database on a subject (the case under investigation) about which little is known at the time.

Finally, compound or compressed files may possess internal structures which cannot be easily identified through a simple binary keyword search. An example includes the latest .docx files used in Microsoft Word which now maintain an internal structure similar to those found in .zip files. Although techniques are available to mount and display some of these file types, doing so complicates the investigation causing it to take longer. The final concern arises if evidential files exist in a format containing no searchable text (e.g., a scan of a document).

2.3. Linguistic Analysis

Linguistic analysis techniques can prove useful for identifying relevant textual documents [13]. However, in complex situations such as fraud, it would need to be known in advance what specific language would be used to indicate that a crime has been committed. Although linguistic techniques may easily identify chat orientated towards a specific evidential topic (e.g., grooming), it may not prove so easy in crimes such as fraud. Distinguishing an evidential email from a non-evidential email may not be as straight forward as analysing the text content of both emails alone. What determines evidential value could be the direction in which the email is sent (email metadata) and this often requires capturing the file along with further analysis provided by the practitioner. That said, it is noted that linguistic analysis techniques might have a future role to play in DT. As it lies beyond the boundary of the present work there is scope to explore this area in future work to develop techniques to function in conjunction with CBR-FT.

2.4. Relevant Evidence and Superfluous Data

A key concern when undertaking any DT investigation is failing to identify all relevant evidence. However, this concern has to be balanced against time spent during a DT investigation collecting non-evidential data. Thus, an ideal DT system will exhibit maximum precision (the proportion of files retrieved that are relevant) whilst maintaining a suitable level of recall (retrieving relevant evidential files from a target system) to provide adequate assistance during decision making for case prioritisation. Precision is an especially important measure. Because a practitioner must review all recovered data to establish whether any evidential files exist it is important to keep the volume of data of no evidential value to a minimum. Often, evidential files make up a relatively small proportion of the overall volume of recovered data [22]. Sifting non-relevant data to extract evidence requires an examiner to use their expertise or knowledge of where evidence is situated on a system.

The CBR-FT framework aims to limit the gathering of non-relevant data by targeting only known evidential areas of a system based on past DF investigation experiences. The reason for this choice is the hypothesis that when people save and store files on their system they will tend to consistently use particular locations. The evaluation (see Section 8) supports this view as CBR-FT gave better results in evidence gathering tasks than the file-type-based method used by EnCase.

Because CBR-FT does not rely on hash or keyword matching it is harder for a suspect to circumvent DT by adjusting file content to evade a hash match. Instead, CBR-FT is concerned with extracting data from specific locations on a system reducing reliance on maintaining keyword or hash sets which are often incomplete or outdated. CBR-FT focuses on where evidence has been located in previous cases (regardless of file type or content) and targets these areas. As discussed in Section 8, this approach to DT yields higher precision and recall.

3. What is CBR-FT and How Can It Be Applied?

We have previously suggested a framework for the reuse of DF investigative knowledge for the auditing of DF examinations [23]. This article builds upon the fundamental ideas previously expressed and presents a fully implemented framework adapted for use in DT.

3.1. CBR-FT Structure and Functionality

The CBR-FT framework provides a mechanism for gathering the results of past DF investigations and utilising this knowledge to predict the existence of evidence on a system for a given offence type. The framework maintains a structure for gathering and storing DF investigation data and employs *user-contributory case-based reasoning* (UCCBR) [24].

UCCBR lets multiple DF practitioners record and submit the results of their investigations as 'cases' to CBR-FT's knowledge base. This is achieved by the creation of a case shell, generated by a script driving the practitioner's proprietary forensic software (in this case, EnCase [12]). The script regulates the format of the practitioner's case submission ensuring that data has the

correct structure while allowing case specific data to be made anonymous [23].

The case shell allows for a simple and efficient submission process which makes minimal demands on the practitioner's time. Submissions are automated in an attempt to remove the need for an expert to validate the content of each case. The script validates the format of data in the case and only allows a practitioner to enter details of files highlighted in their examination results (referred to as bookmarks in EnCase). Each submitted case contains the locations (file paths) of evidential files found on the suspect system and an investigator-assigned evidence relevance rating reflecting the relevance to the case of the evidence found there.[2]

Each location on a system that contains evidence related to an investigation is given an *evidence relevance rating* (ERR) by the investigating practitioner. The ERR represents the investigator's assessment of the relevance to the case of the evidence found in that location. The ERR is a value between 0.1 (low relevance) and 0.9 (very relevant) in increments of 0.1. For example, for a particular case, evidence found at location `C:\Folder` might be assigned an ERR of 0.8 to denote that the data found there was highly relevant, that is, the evidence would be strongly relied upon in determining the outcome of the investigation. Research has shown that more fine-grained scales do not provide optimal opinion information during a rating exercise [18, 34].

3.2. Bayesian Inference

The ERR is a measure of the perceived relevance to a case of evidence for a given location. Bayes' theorem uses prior knowledge of probability distributions to make posterior inferences about related events [40]. Therefore, it is well suited to the DT scenario in which uncertainty exists about which areas of a suspect system to interrogate for evidence, and knowledge of the relevance of evidence in prior cases can be used to prioritise locations to search in new cases.

CBR-FT uses the ERRs as a prior probability distribution in a Bayesian model to determine the priority of particular locations for searching during DT. The Bayesian model used by CBR-FT is given in Equation 1. An explanation of the terms used in Equation 1 is given in Table 1.

$$P(L|E) = \frac{P(E|L)\ P(L)}{P(E|L)\ P(L) + P(E|\neg L)\ P(\neg L)} \quad (1)$$

The term being sought is $P(L|E)$, the probability that a given location, L, contains evidence, E, that is relevant to the case in question. The known prior probabilities are $P(L)$ (the probability that location L is reported by an investigator) and $P(E|L)$ (the probability that evidence retrieved from L is relevant to the case). $P(L)$ is derived from the number of cases in the knowledge base that contain evidence at the the particular location. $P(E|L)$ is the average ERR of the given location in the knowledge base. Thus, we know how relevant to past cases investigators judged evidence in a given location to be.

From this we may calculate $P(L|E)$ which we call the *primary relevance figure* (PRF) for each location. The PRF is the inferred probability that a given location is likely to contain evidence that is relevant to a DT case. As new cases are added to the knowledge base the PRF (i.e., $P(E|L)$) will change to reflect the new case knowledge. Once PRFs have been calculated for each location in the knowledge base a leaderboard can be created to show areas of likely high and low importance on a given system for a suspected offence type.

$P(E|\neg L)$ represents the likelihood that a practitioner would miss relevant evidence at a given location. This value is given by the investigator and is based on their knowledge and experience of interrogating the locations during investigations. For the example discussed in the next section, the $P(E|\neg L)$ values were assigned by an independent practitioner after looking at all the different locations contained within the case knowledge base.

Table 1: Definition of Terms in Equation (1).

Component	Explanation
$P(L)$	Probability that the location, L, is reported by an examiner. This is derived from the number of cases in the knowledge base that contain evidence at a particular location.
$P(\neg L)$	Probability that the location is not reported by an examiner.
$P(E\|L)$	Probability of evidence, E, at the location, L, being important to case given the location is reported by the examiner. This is a probability distribution generated from the evidence relevance ratings given by a practitioner in each case.
$P(E\|\neg L)$	Probability of evidence at the location being important to a case given the location is not reported by an examiner. This is a value based on the subjective judgment of the DF investigator.

3.2.1. Example

An example of the generation of a PRF for the `MyDocuments` location is now presented. This location has a $P(E|L)$ distribution function value of 0.82, generated from the investigator-assigned ERR values which denote the relevance to past cases of evidence found in that location. $P(E|\neg L)$ has value of 0.1 reflecting the low probability that relevant evidence exists at the location but that an examiner has failed to report it. This value was assigned by an independent practitioner.

The probability, $P(L)$, of the `MyDocuments` location being reported by an investigator is 0.53 (as it appears in 25 of the 47 cases in the knowledge base), therefore $P(\neg L)$ (the probability of the `MyDocuments` location not being reported) is 0.47. It is then a simple matter to use Equation (1) to calculate the probability, $P(L|E)$, of the `MyDocuments` location being reported given that the evidence at that location is relevant, which is 0.90.

In this way CBR-FT is able to reflect that a particular piece of evidence may be important to a particular case, but that in

[2] This article focuses on Microsoft Windows operating systems as 99.2% of cases received by our supporting DF organisation were running versions of Windows.

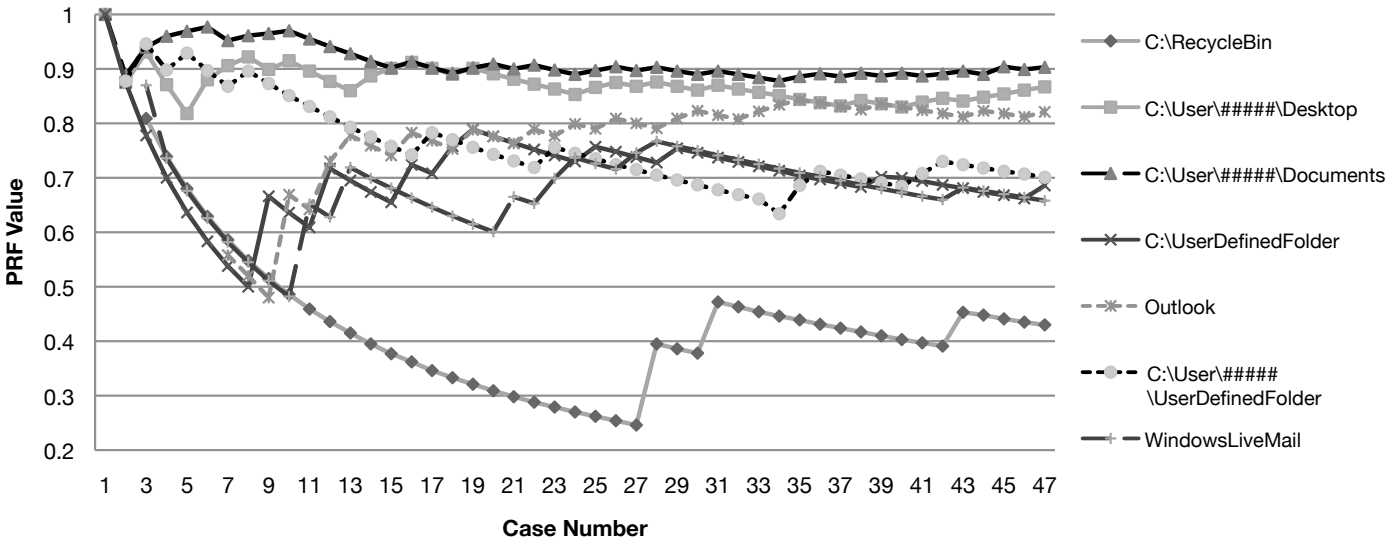


Figure 1: Demonstrates the PRF of each evidential location. As more cases enter the CBR-FT knowledge base, the PRF of each location fluctuates to correspond with any findings in a DF investigation. Each observation on the seven plots shows the change caused by the addition of a new case to the knowledge base.

other cases evidence in this location may be of little importance or that no evidence exists in this location at all. Triage carries the inherent risk that important evidence will be missed when prioritising cases to investigate. With the PRFs, which are are representative of the way in which a particular offence has been committed in the past, CBR-FT mitigates this risk by indicating the locations most likely to contain relevant evidence for the particular case type. Of course, for the PRF values to be reliable (a true reflection of the relevance of a particular location in all offences of a particular type), the CBR-FT knowledge base must contain enough cases to accurately reflect its target population.

4. Knowledge Base Construction

A knowledge base should aim to reach a point of saturation, that is, a size such that its set of sampled data reliably reflects the population from which it was drawn [30, 31]. Determining the saturation point depends on the size of the sample population, and when a population is fixed or known a saturation size value can be determined [29]. However, this research is concerned with the investigation of criminal activity, and as the ways in which digital crime is committed adapt with technological developments, then the location of evidence may also change. Consequently, the CBR-FT knowledge base will not have a fixed population; new cases will be continually added, and true saturation is unlikely to be reached.

As a precise saturation point cannot be identified, the reliability of the knowledge base is determined by the stability of each location's PRF. Fig. 1 shows how the PRF figures can vary greatly during the initial stages of knowledge base growth. In the initial stages the PRFs are less reliable but the variations decrease as the knowledge base increases in size. However, the complexity of DF investigations and the evolving methods used to commit offences mean it is unrealistic to expect each evidential location to achieve a static PRF. Therefore, the goal is to reach a point where the PRF of each location undergoes minimal change with the addition of new cases, that is, while not saturated *per se* the knowledge base has become stable.

In the knowledge base used in this research the PRFs appear have reached a level of stability as the addition of further cases is not likely to dramatically change the PRF outcomes (see Fig. 1). For example, when the knowledge base was being constructed, at the point at which it contained only nine cases the average change in PRF when the tenth case was added was 0.0457. However, the average change in PRF from case 46 to 47 was only 0.00857.

5. Using the Proposed CBR-FT for Digital Triage

CBR-FT conducts DT in two stages which are now discussed.

5.1. Stage 1: Extraction of Data from High Priority Locations

In Stage 1, CBR-FT extracts data from system locations with the highest PRFs. These values are shown in Table 2 and represent the highest probability of locations holding relevant evidence to the investigation. In this study we have focused on fraud offences, and have constructed a fraud knowledge base with associated PRFs (Fig. 2).[3]

In Stage 1, only locations with a PRF above 0.5 have been used (i.e., locations more likely than not to yield relevant evidence). The reason for this is to reflect the fundamental principle in DT that resources are often scarce and as few of them as possible should be allocated when trying to identify evidential data. Searching all system locations for evidence is inefficient, especially as many locations rarely contain relevant evidence. Table 2 shows the highest ranking locations and their associated PRFs.

[3]Note, different types of offence would have different characteristics which would directly affect the locations that hold relevant evidence. In turn, this would mean that a location's PRF would likely differ between offence types, hence the need for offence type-specific knowledge bases.

Table 2: Locations used in Stage 1

Location	PRF
User\Documents	0.903
User\Desktop	0.867
Local\Microsoft\Outlook	0.822
User\UserDefinedFolder	0.701
Volume\UserDefinedFolder	0.658
AppData\Local\Microsoft\WindowsLiveMail	0.658
User\Documents\UserDefined	0.585

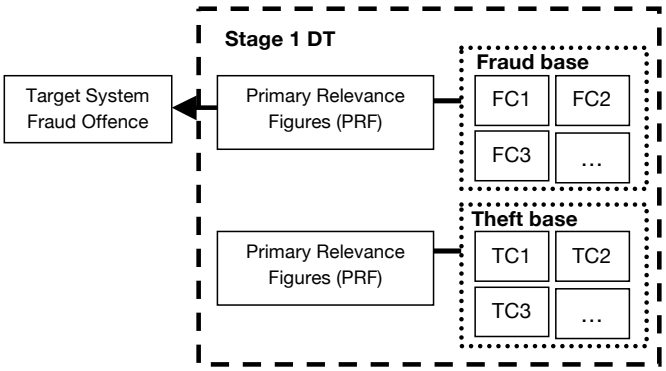


Figure 2: Structure of the CBR-FT framework for Stage 1. Distinct offence types have their own knowledge base. Here, a fraud base contains a number of cases ($FC_1 \ldots FC_n$) and a theft base contains cases $TC_1 \ldots TC_n$.

Should the results prove negative (i.e., no evidence is present at any of the seven locations in Table 2), the DF practitioner is faced with two ways to progress. First, the suspect system can be removed from the case or lowered in priority for further investigation. Second, the examiner has the choice of configuring CBR-FT to search locations with lower PRFs (≤ 0.5). Since each evidence location's PRF reflects the probability of relevant evidence existing in that location, lower ranked areas will have less chance of yielding evidential data whilst increasing the volume of non-relevant data returned (and the consequent effect on precision). Ultimately, it is for the examiner to determine the allocation of the PRF boundary and risk missing evidence. However, if the initial sweep of locations with PRFs > 0.5 yields no relevant evidence, configuring CBR-FT to collect data in locations rated ≤ 0.5 is a way of confirming or falsifying the initial results.

5.2. Stage 2: Data Extraction Based on Case Characteristics

When Stage 1 is complete and all data recovered from the seven locations has been reviewed, an investigator has the option to proceed to Stage 2. If any of the locations were found to contain evidential data they can be viewed as a potentially incomplete profile of suspect activity on the system. Stage 2 takes this incomplete profile and focuses on identifying similar patterns of suspect activity in individual cases stored in the CBR-FT knowledge base (Fig. 3).

For example, if Stage 1 recovers evidential data at the locations `User\Documents` and `User\Desktop`, Stage 2 involves an analysis of cases contained in the knowledge base where evidence had previously been reported at these locations. Each case in the knowledge base is examined for matching activity. If any cases are highlighted as containing evidence in the same locations, any additional reported locations (which were not examined during Stage 1) are collected for Stage 2. These locations will have a lower PRF than locations used in Stage 1 yet could be relevant as both cases share characteristics. This stage is designed to highlight any potential anomalous behaviour seen in past cases where evidence has been found in locations which may not be regularly reported.

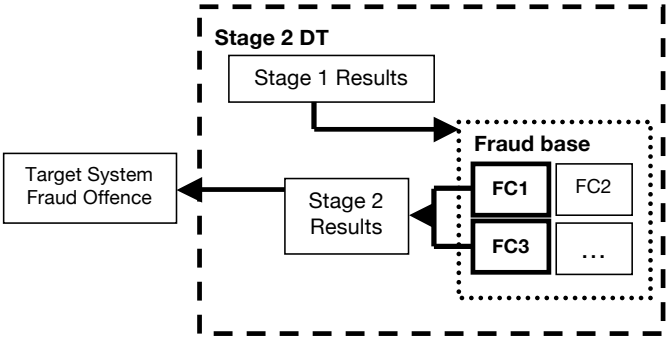


Figure 3: Structure of the CBR-FT framework for Stage 2. The cases FC_1 and FC_3 are those cases in the knowledge base with locations matching those found in the current investigation during Stage 1. These locations are analysed to identify a further set of lower-priority locations of possible evidence.

To prevent an influx of additional locations, matching cases are prioritised according to their match percentage. A match percentage is generated by identifying how many locations found during Stage 1 are present in the knowledge base. By default, Stage 2 results are generated from cases with a 100% match, therefore in the above example, cases previously reporting evidence at both `User\Documents` and `User\Desktop` will be used.

The CBR-FT system is based on the hypothesis that similar criminal activities share comparable traits. Modus operandi (MO) is a synopsis of suspect behaviour and can be used to identify an offence [43]. Each investigation undertaken by a practitioner reveals a suspect's MO which includes the locations in which a suspect has left evidence [37]. This information can be exploited to link past evidential activity to a current investigation [6]. The start of a DF DT investigation presents the greatest volume of data to be interpreted by an investigator and posses a significant challenge [14]. Yet decreasing this set of data whilst maintaining a high degree of recall would allow a practitioner to carry out this task more efficiently. This can be achieved by an analysis of past cases to target frequently seen MOs and this concept is used by CBR-FT during DT.

6. Evidential Area Identification

CBR-FT is concerned with system locations rather than actual files. By highlighting specific areas of a system, CBR-FT can decrease the amount of non-relevant data retrieved during DT. These specific areas are known as file paths and they must be

generalised. A generalised path contains no personal or system specific data, and can be applied to future DT cases. Each system contains files and folders with names that are often unique to a suspect. These file and folder names may differ in other DF cases but their position in the folder hierarchy remains the same. As CBR-FT is concerned with the location of data, CBR-FT transforms each file path into a universal file path to ensure its appropriateness for use in the knowledge base (Fig. 4).

Universal Path 1: `Volume\Users\######\Desktop\`
Actual Path 1: `C:\Users\Administrator\Desktop\`

Universal Path 2: `Volume\UserDefinedFolder\`
Actual Path 2: `C:\Company Docs\`

Figure 4: Example correspondences between universal file paths and original evidential file paths.

Fig. 4 shows how both the user profile name and volume label are generalised. All personal and system specific data has been substituted to create a generalised path string. For example, should the initial case have involved a suspect user profile of John, clearly a path match would not occur during DT unless the system being triaged also contained a profile named John. Issues would also arise where multiple profiles existed on the system. By standardising the path, '`######`' can be applied to all target systems irrespective of the profile name. The concept is similar when using `Volume` as a standard volume label identifier. By using `UserDefinedFolder`, unique folder names can also be standardised.

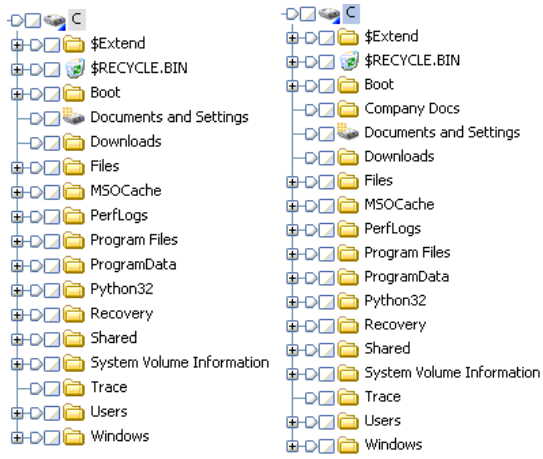


Figure 5: The standard structure skeleton (SSS) is demonstrated (left) in graphical format, highlighting an anomalous folder in the suspect system (right) file system.

One drawback of the universal file path is that looking at a volume's root folder has the potential for recovering large quantities of non-relevant data owing to the presence of system folders containing standard operating system files (e.g., `C:\Windows`). CBR-FT implements a *standard structure skeleton* (SSS) to identify which folders are made by a user and therefore available for DT (Fig. 5). The SSS acts as a system structure hash and is a record of folders that have been generated automatically by a system and which often contain only system files. By maintaining the SSS it is possible to mark these standard locations so they are not searched. Any folder not in the SSS is flagged as a `UserDefinedFolder` (e.g., `Company Docs` in Fig. 4) and files in that location would be recovered (assuming the PRF is above the specified threshold). Of course, if any file paths in the SSS (including standard system areas) are identified as evidence locations in any of the cases in the CBR-FT's knowledge base, then these areas are not restricted from DT. The SSS can be amended to include additional locations to be omitted during DT.

7. Evaluation

We have compared the use of CBR-FT against a leading commercial DT application, EnCase Portable [21], as the two systems take different approaches to DT. EnCase allows for the recovery of data based on file type and offers a set of preset file categories:

1. Picture files
2. Documents files
3. Emails

EnCase recovers data according to its category regardless of where it resides on a system. This offers the advantage that should a suspect attempt to hide evidence by placing it in an unusual location, EnCase can still recover it providing it is of the correct file type. However, there are drawbacks to this approach.

First, the onus is on EnCase to maintain a database documenting all possible file types which may be relevant. Because EnCase only searches for three categories of evidence, its scope for evidence recovery is limited. If evidence exists in a form outside these three categories EnCase will fail to recover it. This could result in a case failing to proceed to a full investigation or being given a lower priority.

Second, because EnCase gathers all files of the same file type from a system regardless of their importance to the investigation, the volume of data recovered could be unnecessarily large with much redundancy, thus increasing the time needed to complete a DT investigation.

CBR-FT's focus on using known locations removes the reliance on establishing a known set of evidential file types. This offers the potential to correctly pinpoint areas of evidence on a suspect system which, in turn, might increase investigation efficiency by reducing the volume of non-relevant data recovered (increasing the precision). EnCase uses a fixed approach for all investigations, and this may not be suitable in all circumstances. CBR-FT can adapt, so as the knowledge base expands, CBR-FT's decision-making will change, reflecting the way in which crimes are currently being committed.

The system was evaluated by taking twenty DF cases that had previously undergone a full DF investigation by an independent practitioner prior to this research. These twenty cases were all concerned with crimes of fraud and were not contained in the CBR-FT. It is against the evidence found during the 47 prior investigations recorded in the knowledge base that the CBR-FT

system is evaluated. The test cases used were selected on the basis of their availability, that is, they were the twenty cases that were made available by our contributing DF organisation.

For each case both EnCase and CBR-FT were used to run a triage activity. Both tools were run without direct investigator involvement, CBR-FT by comparing each case against its knowledge base and EnCase by executing its automatic DT data gathering scripts. Note, fraud cases rarely contain evidence in picture or image files. As an investigator would not be likely to search for image files in this type of case it was decided not to let EnCase retrieve files of this type as doing so would artificially increase the redundancy of the results and, hence, give unrealistically low precision.

8. Results

This section presents the results of DT examinations of the twenty separate cases of the same offence type (fraud) carried out using EnCase and CBR-FT. System performance is evaluated in in terms of recall and precision.

8.1. Stage 1 Results — Recall

Except in rare situations where the number of files involved is low, a recall rate for relevant evidence of 100% is an unrealistic expectation. Recall is the proportion of relevant evidence on a target system that is recovered during DT. A DT process with a high level of recall will be seen as more reliable and, ultimately, more useful to a practitioner as a greater quantity of relevant data will be recovered. Recall rates for this study are calculated by comparing the amount of evidence recovered during DT with the evidence retrieved during the prior independent investigations. We cannot say with certainty that the previous investigator recovered 100% of the relevant evidence that existed on the target systems, but all twenty cases had previously been submitted to the DF practitioner's client and were successful in proving or disproving a suspected offence. Therefore, we take the practitioner's results as the *de facto* 100% target. Table 3 shows the recall and precision rates for EnCase and CBR-FT for the twenty DT cases

It can be seen that CBR-FTs Stage 1 approach recovers more evidence than EnCase's automated collection scripts in 8 of the 20 cases and that both techniques have equal recall in 10 cases. As EnCase extracts data by type rather than location, where EnCase's recall is low, it is due to evidence existing in a format which is not targeted for collection and which is, therefore, missed. In 15 cases, CBR-FT recovered fewer files in total than EnCase (see Table 3) but matched or beat EnCase's recall performance by retrieving an equal or greater number of evidential files.

There were two cases (16 and 17) in which EnCase outperformed CBR-FT for recall. In these cases relevant evidence existed in a location on the system which was not targeted by CBR-FT because of its low PRF. However, this evidence existed in a format known to EnCase and so was collected, giving EnCase the greater recall. Of course, if the investigator were to proceed to CBR-FT's Stage 2 and the evidence were subsequently found then when the case is added to the knowledge base this location's PRF would increase, making it more likely to be searched in subsequent investigations. In the ten cases showing equal performance evidence existed in a format known to EnCase or a location searched by CBR-FT which caused both methods to recover the same quantity of evidence.

8.2. Stage 1 Results — Precision

The precision of a DT process is the proportion of retrieved data that is relevant, i.e., the number of relevant files retrieved divided by the total number of files retrieved. Table 3 shows the numbers of files retrieved by each method and the number of evidential files that existed in that case. Because all retrieved files will have to be reviewed by the investigator to determine whether it contains information of evidential value, precision is the more important measure for DT. A lower precision score means a greater volume of non-relevant data is collected which, in turn, makes the DT process longer and less efficient.

It is accepted that precision rates will generally be low due to the large quantities of files residing on suspect systems and, as evidential data commonly makes up only a small subset of the overall number of files on a target system, there is a high chance of non-relevant files being collected. Therefore, a goal of this research is to improve the precision of DT.

Table 3 shows the precision of both EnCase and CBR-FT in each of the twenty test cases. The precision rates attained by the two methods differ markedly with CBR-FT, in some cases, outperforming EnCase by an order or magnitude. A complementary view is given in Fig. 6 which shows the number of non-relevant files retrieved per evidential file retrieved.

The precision of each DT technique is sensitive to the different ways in which evidence is stored on a suspect system. EnCase's precision is sensitive to file type. If a suspect maintains many document files on their system, yet only a small proportion are evidential, the precision will drop as all files of this type are collected. On the other hand, CBR-FT's precision is sensitive to file location. If the suspect chooses to store large quantities of non-evidential data in the same location as evidential data, CBR-FTs precision will decrease.

CBR-FT's precision score matched or beat that of EnCase in 17 of the 20 cases. CBR-FT's precision was worse in 3 cases (7, 8, 17) because the suspects in these cases had stored a lot of non-relevant data in the evidential locations. What this means for the investigator is that while CBR-FT retrieved evidential files at a rate comparable with EnCase (a leading DF tool) the higher precision means that less time would need to be spent sifting the retrieved files to find the evidence. In cases where the investigator thinks the amount of retrieved evidence is too low and that more might be found, CBR-FT could be operated in its Stage 2 mode to widen the net. This was done and the results are presented below.

8.3. Stage 2 Results

Because EnCase does not offer a two-stage DT approach the Stage 2 results only apply to CBR-FT. This stage offers an optional approach to the examiner should they suspect that further evidence may exist and wish to try and increase the

Table 3: The recall and precision rates of EnCase and CBR-FT for the twenty test cases. For both systems the total number of files retrieved and the number of evidential files retrieved are given from which recall and precision are calculated. The second column shows the total number of evidential files available for retrieval for each case (i.e., the number of files found by the independent practitioner and which were sufficient for forensic needs). Figures in **bold text** indicate where one tool outperformed the other.

		EnCase				CBR-FT			
Case	Available Evidence Files	Files Retrieved	Evidence Files Retrieved	Recall	Precision	Files Retrieved	Evidence Files Retrieved	Recall	Precision
1	1	8706	1	100%	0.000114%	3621	1	100%	**0.000276%**
2	3	0	0	0%	0%	0	0	0%	0%
3	6	2758	0	0%	0%	65	0	0%	0%
4	6	9246	5	83%	0.054%	4759	5	83%	**0.105%**
5	7	0	0	0%	0%	0	0	0%	0%
6	9	11083	9	100%	0.00081%	7640	9	100%	**0.001178%**
7	10	13962	10	100%	0.072%	14065	10	100%	0.071%
8	14	95553	14	100%	**0.0001465%**	195896	14	100%	0.00007%
9	16	9182	16	100%	0.174%	2801	16	100%	**0.571%**
10	24	10010	7	29%	0.070%	3508	20	**83%**	**0.570%**
11	35	1774	0	0%	0%	7660	35	**100%**	**0.004569%**
12	49	8354	46	94%	0.551%	587	46	94%	**7.836%**
13	83	7403	44	53%	0.594%	3177	83	**100%**	**2.613%**
14	86	31123	57	66%	0.0018%	10046	81	**94%**	**0.00806%**
15	114	8257	30	26%	0.363%	683	102	**89%**	**14.934%**
16	121	11878	82	**68%**	**0.690%**	2544	36	30%	1.415%
17	262	15597	262	**100%**	**0.016798%**	5896	0	0%	0%
18	393	20447	1	0.3%	0.0000489%	1233	2	**0.5%**	**0.001622%**
19	776	33509	332	43%	0.010%	4751	364	**47%**	**0.077%**
20	1301	35491	1078	83%	0.030%	32535	1090	**84%**	**0.034%**

Table 4: Evidence locations highlighted in Stage 2.

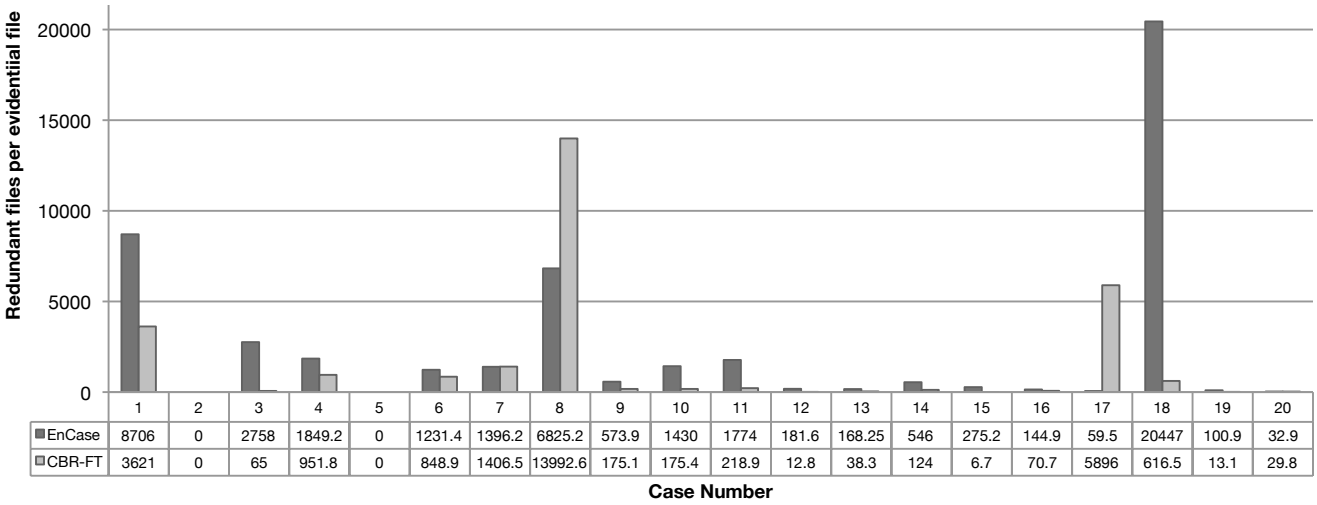
Location	PRF
`C:\User\#####\AppData\Local\Microsoft\TemporaryInternetFiles`	0.406
`C:\Unallocated`	0.335
`C:\SystemVolumeInformation`	0.329
`C:\User\#####\Pictures\SamplePictures`	0.329
`C:\User\#####\AppData\Roaming\Microsoft\Windows\Recent`	0.128
`C:\User\#####\AppData\Local\Microsoft\Windows\Explorer`	0.018

precision of their DT examination. Test cases 1, 6, 7, 8, 9, 11 and 13 are excluded from Stage 2 DT as 100% of the evidence was recovered during Stage 1 (see Table 3). Cases 2, 3, 5 and 17 are deemed anomalous in comparison to the other cases we have presented and are discussed in section 8.3.1.

Test case 16 has been used to demonstrate the Stage 2 process. Stage 1 confirmed evidence was present at locations `C:\User\#####\Documents` and `Outlook`. Stage 2 involves the identification of individual cases in the knowledge base which have previously recorded evidence in the same locations which were found during the current investigation undergoing stage 1 DT.

As Stage 1 identified evidence at two locations (`Outlook` and `C:\User\#####\Documents`) only cases in the knowledge base containing both these locations are used. Analysis of the CBR-FT knowledge base revealed eight of the cases within it had previously documented evidence at both these locations (KB cases 7, 9, 16 19, 22, 26, 44 and 47 in Fig. 1). This is seen as a 100% match as all the evidence locations confirmed during Stage 1 were confirmed in cases 7, 9, 16 19, 22, 26, 44 and 47. In addition, these cases contained six other evidence locations (Table 4) which had not been interrogated during Stage 1. Carrying out Stage 2 triage on these locations for test case 16 revealed that evidence also existed in the `SamplePictures` location and this subsequently raised the recall of CBR-FT during DT in this test case to 62% from 30%.

It is also possible to search the CBR-FT knowledge base for cases with a partial match. For case 16 this would involve looking for cases in the knowledge base that contained either one of the locations found during Stage 1. This may increase the chances of relevant evidence locations being returned during Stage 2, but could also increase the volume of non-relevant



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
EnCase	8706	0	2758	1849.2	0	1231.4	1396.2	6825.2	573.9	1430	1774	181.6	168.25	546	275.2	144.9	59.5	20447	100.9	32.9
CBR-FT	3621	0	65	951.8	0	848.9	1406.5	13992.6	175.1	175.4	218.9	12.8	38.3	124	6.7	70.7	5896	616.5	13.1	29.8

Figure 6: The redundancy per evidential file retrieved for the twenty test cases (redundancy = no. files retrieved/no. evidential files retrieved).

locations returned and this approach was not favoured during testing. The PRFs assigned to each location also assist the examiner during Stage 2 when determining which additional locations to collect files from.

Running stage 2 DT on the remaining four test cases where case matches were based on a 100% match in locations, cases 4, 12, 14, and 15 saw no improvement in recall. The remaining evidence in these cases was found in areas of the hard drive not identified in the knowledge base. In case 10, the recall of CBR-FT rose to 100% as the remaining evidence stored in the recycle bin was collected as this location was identified. For cases 18, 19 and 20 no cases in the knowledge base existed with a 100% match. Cases were then identified based on a 75% match. Using locations from these matching cases improved the recall seen in test case 19 to 68% and test case 18 to 0.8%. No improvement was seen in test case 20.

8.3.1. Anomalies

We have highlighted cases 2, 3, 5, and 17 as anomalous in comparison to the rest of the cases and these merit further discussion. In case 17, CBR-FT's recall is 0. This is due to all existing evidence being stored in the recycle bin. As nothing was highlighted during Stage 1, it is impossible to carry out Stage 2 DT. However, we note in section 5.1 that in this scenario the examiner has the option of expanding upon the locations chosen during Stage 1 (see Table 2). The recycle bin location maintains a PRF of 0.430 and is ranked 8th in our system, just outside the top seven. If the examiner opted to capture data from this location as we advise, the recall in case 17 would rise to 100%, matching EnCase and with a precision of 0.044%.

In case 3, all evidence was found in the unallocated clusters of the suspect system, leading to both techniques recovering nothing and again leading to it being impossible to carry out Stage 2. The unallocated areas of the system are ranked 11th in our system with a PRF of 0.335. If, as with case 17, the examiner chooses to expand upon locations in Stage 1, the potential to locate this evidence is increased.

Cases 2 and 5 are both anomalous as the target system appeared to be an empty drive, with evidential files found in unallocated clusters. This explains the lack of files recovered by both tools. As described above, unallocated clusters are ranked 11th in CBR-FT and, therefore, should the examiner choose to follow the Stage 1 directions, this area of the drive would be highlighted and the potential for file recovery exists. However, it is noted that due to the time penalties of analysing unallocated clusters, this may be unsuitable for DT. Yet, in providing the PRF of the unallocated clusters location and given the negative results found during Stage 1, the information provided by CBR-FT should aid the examiner in determining the priority of the case.

9. Conclusions and Further Work

Workload pressures on digital forensics investigators and the increasing criminal use of computer technology suggest that improved device triage techniques are required. This article has demonstrated a new approach to device triage, CBR-FT. It has been shown that CBR-FTs precision rate for the recovery of evidential data during device triage is an improvement over a leading existing commercial tool and that its recall (in 15 of the 20 cases) was as good or better than the commercial tool. In 17 cases precision was higher than that of EnCase, that is, CBR-FT recovered more evidence and less non-relevant data than EnCase. While this article has only used fraud crimes for test cases, CBR-FT could be applied to other crime types and this is an area for further work. In addition, the research only focused on the Windows operating system, but the principles of CBR-FT could also be applied to other operating systems.

9.1. Knowledge Base Time Sliding View

A limitation of the CBR-FT knowledge base is that it represents the most currently likely locations within a system that contain evidence for a given offence type. However, the key word here is *currently*; given the complexity of DF investigations and the range of methods used to carry out an offence,

common evidential locations will likely vary over time, and hence we introduce the concept of a *time sliding view* of criminal cyber activity. At present this concept remains untested yet merits brief discussion.

DF investigations are complex and contain many forms of evidence that can contribute to an offence. Some offences can become dependent on technology and as technology changes, then so will the location and its associated evidence. Let us assume that for a given time t the knowledge base is effectively saturated for t and therefore suitable for use in DT for cases where the offence is committed at this time. However, at a later time t' an offence is committed; the knowledge base is only saturated for t and might not be effective for DT at t' as the offence could present different characteristics. By adding new cases to the knowledge base as they arise, the PRFs for each location will be updated and saturation can be maintained. By adding another case at t', the knowledge base may no longer represent t; it may now be representative of t' and effective for the DT of future offences.

Given that cases will be constantly added to the knowledge base it can never become fully saturated, only reaching a level of representation for a particular time; as the knowledge base evolves, then so will the PRFs. Consequently, the CBR-FT system may be able to reflect changes in the way offences are committed. The adaptability of the PRFs produced from a time sliding knowledge base is, therefore, a potential advantage over current commercial systems, as it may have the ability to account for the changes over time in the way offences are committed. This would mean that where other techniques have a static script based approach to DT, CBR-FT could change the locations it searches based on the way in which a crime type evolves, derived from additional cases entering the knowledge. This could, in turn, increase the reliability of CBR-FT for DT. As noted, this feature is at present untested and validation will be carried out as more cases enter the CBR-FT system.

9.2. Prior Knowledge

In the next step (having acquired separate knowledge bases of different offence types), the intention is to identify if an offence has been committed using a target system without any prior knowledge. It is suggested that user activity on a target system may share similar characteristics to cases stored in the CBR-FT knowledge base, and this will allow for the creation of profiles of suspect activity. This technique would be useful in a suspect crime scene device triage situations, for example when border control agencies are attempting to carry out spot check examinations of digital media. Further work would also consider the possibility of incorporating notions of uncertainty using interval probability theory, when determining the probability of evidence being present. This would allow us to define the dependability of any located evidence on a system in an attempt further improve the CBR-FT's precision and recall abilities.

9.3. Clustering

Beebe *et al.*'s [8] method of clustering improves information retrieval performance for document searching. As part of future work we anticipate the incorporation of a similar technique could increase CBR-FT's precision performance. We have demonstrated the ability to extract evidence from a case whilst maintaining good recall, and improved precision rates. Large quantities of non-relevant data can be excluded from an investigation, significantly decreasing the amount of data a practitioner needs to review. Yet, implementing Beebe *et al.*'s technique over this smaller set of data could improve our precision further, thereby allowing evidential files to be individually targeted based on their related semantic content and is a consideration for future work.

9.4. Expanding the Size of the Knowledge Base.

The number of cases housed in the current knowledge base is 47. The expansion of the current knowledge base through the addition of future cases will enable the reliability of the framework to be monitored. As the PRFs of locations are affected when more data is added, future case additions would highlight when certain locations become more prominent in DT investigations. As noted above (section 9.2), this would permit a more thorough exploration of the time sliding view more thoroughly as well as the exploration of the effects that a larger knowledge base may have on future test cases. A limitation of this suggestion is the availability of additional cases for submission to CBRFT's knowledge base.

9.5. Additional Knowledge Bases

An additional limitation of this research is that it focused only on offences of fraud. However, CBRFT works by identifying shared characteristics amongst cases of a given type. Therefore, it is expected that the framework would also yield positive results in the triage of different crime types where evidential areas could be targeted in the manner demonstrated with fraud cases above. As Fig. 2 shows, all that is needed is further knowledge bases of cases belonging to different crime types. This will be tested through the collection of cases for an additional knowledge base documenting a different offence type and constitutes further work to be carried out.

Acknowledgements

We would like to thank Dr Pete Philipson of Northumbria University's Department of Information Sciences and Mathematics for his advice on Bayesian inference. Funding for this research came from Northumbria University's Graduate Tutor scheme.

[1] AccessData, AD Triage: Release Notes, `http://www.accessdata.com/support/product-downloads`, 2013. Last checked: February 2013.

[2] ADF Solutions, Inc., Triage computers to reduce forensic backlogs and lower costs, `http://www.adfsolutions.com/products/triage-examiner`, 2012. Last checked: February 2013.

[3] Association of Chief Police Officers, Good practice guide for computer-based electronic evidence, `http://www.7safe.com/electronic_evidence/ACPO_guidelines_computer_evidence.pdf`, 2007. Last checked: February 2013.

[4] Association of Chief Police Officers, Good practice and advice guide for managers of e-crime investigations, http://www.acpo.police.uk/documents/crime/2011/201103CRIECI14.pdf, 2011. Last checked: February 2013.

[5] D. Ayers, A second generation computer forensic analysis system, Digital Investigation 6 (2009) S34–S42.

[6] K. Baumgartner, S. Ferrari, G. Palermo, Constructing Bayesian networks for criminal profiling from limited data, Knowledge-Based Systems 21 (2008) 563–572.

[7] J. Beckett, J. Slay, Digital forensics: Validation and verification in a dynamicwork environment, in: HICSS '07 Proceedings of the 40th Annual Hawaii International Conference on System Sciences, IEEE Computer Society, 2007, pp. 266a–.

[8] N.L. Beebe, J.G. Clark, G.B. Dietrich, M.S. Ko, D. Ko, Post-retrieval search hit clustering to improve information retrieval effectiveness: Two digital forensics case studies, Decision Support Systems 51 (2011) 732–744.

[9] D. Bem, F. Feld, E. Huebner, O. Bem, Computer forensics - past, present and future, Journal of Information Science and Technology 5 (2008) 43–59.

[10] M.C. Bottrill, L.N. Joseph, J. Cawardine, M. Bode, C. Cook, E.T. Game, H. Grantham, S. Kark, S. Linke, E. McDonald-Madden, R.L. Pressey, S. Walker, K.A. Wilson, H.P. Possingham, Is conservation triage just smart decision making?, Trends in Ecology and Evolution 23 (2008) 649–654.

[11] W.K. Brotby, Information Security Governance: A Practical Development and Implementation Approach, John Wiley & Sons, Hoboken, N.J., 2009.

[12] S. Bunting, EnCase Computer Forensics — The Official EnCE: EnCase Certified Examiner Study Guide, John Wiley & Sons, 3rd edition, 2012.

[13] J.K. Burgoon, J.P. Blair, T. Qin, J.F.N. Jr, Detecting deception through linguistic analysis, in: H. Chen, R. Miranda, D.D. Zeng, C. Demchak, J. Schroeder, T. Madhusudan (Eds.), Intelligence and Security Informatics, number 2665 in Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2003, pp. 91–101.

[14] B.D. Carrier, H. Spafford, Eugene, Automated digital evidence target definition using outlier analysis and existing evidence, in: Refereed Proceedings of the 5th Annual Digital Forensic Research Workshop, DFRWS 2005, Astor Crowne Plaza, New Orleans, Louisiana, USA.

[15] E. Casey, Triage in digital forensics, Digital Investigation 10 (2013) 85–86.

[16] Child Exploitation and Online Protection Centre, CEOP relationship management strategy, http://www.ceop.police.uk/documents/ceopdocs/relationship_management_strategy.pdf, 2011.

[17] W. Chung, H. Chen, W. Chang, S. Chou, Fighting cybercrime: A review and the taiwan experience, Decision Support Systems 41 (2006) 669–682.

[18] V. Cicchetti, Domenic, D. Shoinralter, P.J. Tyrer, The effect of number of rating scale categories on levels of interrater reliability: A monte carlo investigation, Applied Psychological Measurement 9 (1985) 31–36.

[19] Evidence Talks, Announcing the arrival of SPEKTOR 2.8, http://www.evidencetalks.com/index.php?option=com_content&view=article&id=135:sfi-27&catid=83&Itemid=513, 2012. Last checked: February 2013.

[20] M. Fernández, I. Cantador, V. López, D. Vallet, P. Castells, E. Motta, Semantically enhanced information retrieval: an ontology-based approach, Web Semantics: Science, Services and Agents on the World Wide Web 9 (2011) 434—452.

[21] Guidance Software, EnCase Portable, http://www.guidancesoftware.com/encase-portable.htm, no year. Last checked: February 2013.

[22] I. Hong, H. Yu, S. Lee, K. Lee, A new triage model conforming to the needs of selective search and seizure of electronic evidence, Digital Investigation in press (2013).

[23] G. Horsman, C. Laing, P. Vickers, A case based reasoning framework for improving the trustworthiness of digital forensic investigations, in: 11th IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom-2012), IEEE, Liverpool, UK, pp. 682–689.

[24] G. Horsman, C. Laing, P. Vickers, User-contributory case based reasoning for digital forensic investigations, in: L. O'Conner (Ed.), EST 2012: Third International Conference on Emerging Security Technologies, IEEE Computer Society, Lisbon, Portugal, 2012.

[25] K.V. Iserson, J.C. Moksop, Triage in medicine, Part I: Concept, history, and types, Annals of Emergency Medicine 49 (2007) 275–287.

[26] T. Kalker, J. Haitsma, J.C. Oostveen, Issues with digital watermarking and perceptual hashing, in: ITCom 2001: International Symposium on the Convergence of IT and Communications, pp. 189—197.

[27] J. Kornblum, Identifying almost identical files using context triggered piecewise hashing, Digital Investigation 3 (2006) 91—97.

[28] M. Losavio, D.W. Keeling, A.S. Elmaghraby, A distributed triage model for digital forensic services to state and local law enforcement, in: Systematic Approaches to Digital Forensic Engineering, 2009. SADFE '09. Fourth International IEEE Workshop on, pp. 36–37.

[29] R.C. MacCallum, K.F. Widaman, S. Zhang, S. Hong, Sample size in factor analysis, Psychological Methods 4 (1999) 84–89.

[30] M. Mason, Sample size and saturation in PhD studies using qualitative interviews, Forum: Qualitative Social Research 11 (2010).

[31] J.M. Morse, Determining sample size, Qualitative Health Research 10 (2000) 3–5.

[32] S. Pearson, R. Watson, Digital Triage Forensics: Processing the Digital Crime Scene, Syngress, 2010.

[33] M.M. Pollitt, Triage: A practical solution or admission of failure, Digital Investigation In press (2013).

[34] C.C. Preston, A.M. Colman, Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences, Acta Pyschologica 104 (2000) 1–15.

[35] L.P. Rees, J.K. Deane, T.R. Rakes, W.H. Baker, Decision support for cybersecurity risk planning, Decision Support Systems 51 (2011) 493–505.

[36] E. Riloff, Automatically constructing a dictionary for information extraction tasks, in: AAAI-93: Proceedings of the Eleventh National Conference on Artificial Intelligence, July 11–16, MIT Press, Washington, DC, 1993, pp. 811–816.

[37] M. Rogers, The role of criminal profiling in the computer forensics process, Computers & Security 22 (2003) 292–298.

[38] M.K. Rogers, J. Goldman, R. Mislan, T. Wedge, S. Debrota, Computer forensics field triage process model, Journal of Digital Forensics, Security and Law 1 (2006) 19–38.

[39] V. Roussev, Y. Chen, T. Bourg, G.G. Richard, III, md5bloom: Forensic filesystem hashing revisited, Digital Investigation 3, Supplement (2006) 82–90.

[40] J.B. Tenenbaum, T.L. Griffiths, C. Kemp, Theory-based bayesian models for inductive learning and reasoning, Trends in Cognitive Sciences 10 (2007) 209–318.

[41] X. Wang, D. Feng, X. Lai, H. Yu, Collisions for hash functions MD4, MD5, HAVAL-128 and RIPEMD, Cryptology ePrint Archive, Report 2004/199, 2004. Last checked: February 2013.

[42] Wiltshire Police, Wiltshire Police corporate risk register, http://www.wiltshire-pcc.gov.uk/Document-Library/WPA-committee-papers/Police-Authority/July-2012/Agenda-Item-7-Force-Risk-Register.pdf, 2012. Last checked: February 2013.

[43] K. Yokota, S. Watanabe, Computer-based retrieval of suspects using similarity of *Modus Operandi*, International Journal of Police Science & Management 4 (2002) 5–15.