A Hierarchical, Objectives-Based Framework for the Digital Investigations Process

Nicole Lang Beebe
University of Texas at San Antonio
nbeebe@utsa.edu

Jan Guynes Clark
University of Texas at San Antonio
jgclark@utsa.edu

ABSTRACT

Although digital forensics investigations can vary drastically in their level of complexity, each investigative process must follow a rigorous path. Previously proposed frameworks are predominantly single-tier, higher order process models, focusing on the abstract, rather than more concrete principles of the investigation. We contend that these frameworks, although useful in explaining overarching concepts, require additional detail in order to be useful to investigators and researchers. We therefore propose a multi-tier, hierarchical framework to guide digital investigations. Our framework includes objectives-based sub-phases that are applicable to various layers of abstraction.

Keywords

Digital investigative process, digital forensics, computer forensics, analysis, framework

INTRODUCTION

The traditional physical forensic science discipline developed along side its underlying physical and biological sciences over the course of several decades (Palmer 2002). In contrast, the digital forensic science discipline is fledgling and significantly lags behind its relatively well developed underlying computer science (Palmer 2001; Palmer 2002; Carrier and Spafford 2003). Because of this, a more pervasive, well-established digital investigative process framework is still evolving. Such a framework will provide a common starting place from which established theory (e.g. computer science theory and forensic science theory) can be scientifically applied to the digital forensic science discipline. The framework will also enable new theory development and the identification of research and development requirements. The resultant scientific rigor that will be applied using the framework as its foundation will transform current digital forensics practices into digital forensic *science*, as defined by academics and practitioners at the first Digital Forensics Workshop in 2001:

> Digital Forensic Science – "The use of scientifically derived and proven methods toward the preservation, collection, validation, identification, analysis, interpretation, documentation, and presentation of digital evidence derived from digital sources for the purpose of facilitation or furthering the reconstruction of events found to be criminal, or helping to anticipate unauthorized actions shown to be disruptive to planned operations." (Palmer, 2001: 16)

A review of the prevailing digital investigative process models presented to date (Palmer 2001; DoJ 2001; Reith et al. 2002; Carrier and Spafford 2003; Mandia et al. 2003; Mohay et al. 2003; Nelson et al. 2004) discloses a predominant focus on single-tier, higher order process models. While this is a natural starting point, the complexity level associated with the digital investigative process necessitates a multi-tier, hierarchical framework. We purport that the digital investigative framework is completed by the addition of lower order objectives-based sub-phases for each higher order phase. A simple analogy is useful here.

Arguably, flying an airplane can be a function of a higher order framework consisting of three phases: take-off, fly, and land. Few pilots could accomplish this task without obtaining more detail regarding each of the phases. Similarly, more detailed information is needed within each phase of a digital investigation framework in order to improve usability for practitioners and researchers alike.

Our framework is based upon phases, sub-phases, principles, and objectives. Phases and sub-phases are distinct, discrete steps in the process that are usually a function of time and suggest a necessarily sequential approach. Principles, on the other hand, are overarching procedures, guidelines, and/or methodological approaches that overlap some or all of the phases and sub-phases. Unlike phases, principles are not distinct, discrete steps in the process; instead, they represent goals and objectives sought throughout the process. Documentation is an example of a principle.

Our proposed phases and sub-phases are objectives-based. The higher order process phases presented by researchers and practitioners to date are largely objectives-based, but this characteristic has not been a stated requirement as such. This is an important point at this juncture, because it is vital, yet not necessarily the natural inclination for the development of sub-phases. Its vitality is a function of usability. Practitioners find objectives-based steps more useful than task-based steps, because the uniqueness of each situation and "digital crime scene" (Carrier and Spafford, 2003) necessitates a non-checklist approach. A different subset of steps is likely taken in each situation. The decision of which steps to take in any given situation is more easily made when the steps are outlined in an objectives-based fashion, as opposed to a task based fashion. As an example, it is easier for a practitioner to decide whether a step described as "Determine whether unauthorized software has been installed" is relevant to the investigation than the task described as "Examine the Registry." Many tasks apply to more than one step and can easily be matrixed to steps, which allows practitioners to select relevant objectives-based steps and then accomplish recommended tasks for each step. There are also benefits of this approach related to the furtherance of research and development activities, which will be discussed later in this paper.

Another important characteristic of our framework is that the phases and sub-phases are applicable to various layers of abstraction (Carrier 2003). Abstraction layers are used to analyze and translate data into more manageable formats, either via translation from a lower level representation to a higher, more human readable level representation (e.g. translation from binary to ASCII), or via data reduction mechanisms to lessen the amount of data to be analyzed by humans (e.g. intrusion detection system). The abstraction layer corresponds to the translation rule that will operate on the input. This is of particular importance with regard to the Data Analysis Phase. As we define digital crime scenes, we will encounter data in various layers of abstraction. As such, it is important that the digital investigative process framework be robust enough to apply to various layers of abstraction. Separate frameworks for each layer of abstraction or each type (i.e. media analysis vs. network analysis) would be too cumbersome.

Our proposed framework consists of higher ordered (first tier) phases, more definitive sub-phases (second tier), objectives, and framework principles. In the following section, we will describe the first tier phases and process principles for the framework. Although all phases should consist of sub-phases, for the purpose of this paper, we will focus on the sub-phases of the Data Analysis Phase. The rationale for selecting the Data Analysis Phase is its importance, complexity, and the relative absence of detailed attention paid to it to date. Additionally, it provides the best opportunity to demonstrate the necessity for a hierarchical, objectives-based digital investigations process framework. The third section then discusses the benefits and draw-backs of the proposed framework.

PROPOSED DIGITAL INVESTIGATIVE PROCESS

The present effort proposes a digital investigative process framework that parsimoniously encapsulates all phases and activities outlined in prevailing models presented to date (Palmer 2001; DoJ 2001; Reith et al. 2002; Carrier and Spafford 2003; Mandia et al. 2003; Mohay et al. 2003; Nelson et al. 2004). The first tier framework is presented in Figure 1 below, followed by a discussion of each phase, including phase definition and example activities that characterize each phase. Table 1 then maps phases and steps from previously presented models to the proposed model's first tier phases.

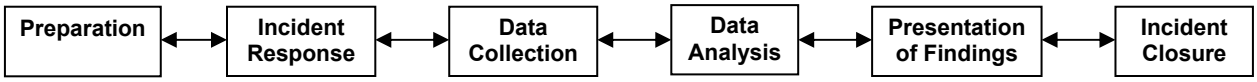

Figure 1.
Single Tier Digital Investigations Process Framework

First-Tier Phases

Preparation Phase

Simply put, a digital investigation requires digital evidence, which is not entirely existent by default and is frequently damaged or destroyed during standard containment, eradication, and recovery activities. Thoughtful preparation maximizes the availability and quality of digital evidence when needed, while minimizing the associated organizational, financial burden—a trade-off that equates to an organization's decision regarding its "forensic readiness" posture (Rowlingson 2004). The Preparation Phase keeps this goal in mind and includes those steps taken by companies to maximize digital evidence availability in support of deterrence, detection, response, investigation, and prosecution related to computer security incidents. Preparation activities include, but are not limited to:

- Assess risk considering vulnerabilities, threats, loss/exposure, etc.;
- Develop information retention plan (both pre/post-incident);
- Develop an incident response plan, including policies, procedures, personnel assignments, and technical requirements definition;
- Develop technical capabilities (e.g. response toolkits);
- Train personnel;
- Prepare host and network devices;
- Develop evidence preservation and handling procedures;
- Document results of activities; and
- Develop legal activities coordination plan (both pre/post-incident).

Incident Response Phase

The Incident Response Phase consists of the detection and initial, pre-investigation response to a suspected security incident. The purpose of this phase is to detect, validate, assess, and determine a response strategy for the suspected security incident. Incident response activities include, but are not limited to:

- Detect or suspect unauthorized activity;
- Report detected or suspected unauthorized activity to the proper authority;

- Validate the incident;
- Assess damage/impact via interviews of technical/business personnel, review of pertinent logs, review of network topology, etc.;
- Develop a strategy regarding containment, eradication, recovery, and investigation, considering business, technical, political, and legal factors/goals;
- Coordinate, as applicable, managerial, human, legal, and law enforcement resources; and
- Formulate the initial investigative plan for data collection and analysis.

Data Collection Phase

Some data and information will be initially collected during the Incident Response Phase, which is necessary to validate the incident and determine the impact and nature of the incident. The formal Data Collection Phase, however, occurs subsequent to the Incident Response Phase once a decision has been made that a digital investigation will ensue, regardless of its scope or anticipated legal or administrative actions. The purpose of the Data Collection Phase, therefore, is to collect digital evidence in support of the response strategy and investigative plan. Data collection activities include, but are not limited to:

- Complete "live response" data collection, which likely began during the Incident Response Phase;
- Obtain network-based evidence from applicable sources (e.g. intrusion detection systems, routers, firewalls, log servers, etc.);
- Obtain host-based evidence from applicable sources (e.g. volatile data, system date/time information, hard drives or forensic duplicates thereof, etc.);
- Obtain removable media evidence from applicable sources (e.g. backup tapes, floppy disks, CD-ROMs, flash memory devices);
- Install activity monitoring capability (e.g. network monitors, system monitors, surveillance cameras);
- Ensure integrity and authenticity of the digital evidence (e.g. write protection, hashes, etc.); and
- Package, transport, and store the digital evidence.

Data Analysis Phase

The Data Analysis Phase is arguably the most complex and time consuming phase in the digital investigations process. The purpose of the Data Analysis Phase is confirmatory analysis (to confirm or refute allegations of suspicious activity) and/or event reconstruction (answer "who, what, where, when, why, and how" type questions). Data collected during the Data Collection Phase is surveyed, extracted, and reconstructed during the Data Analysis Phase. Data analysis activities include, but are not limited to:

- Transform the voluminous amount of data collected during the Data Collection Phase into a more manageable size and form for analysis;
- Conduct initial data survey to recognize obvious pieces of digital evidence and assess the skill level of the suspect(s);
- Employ data extraction techniques (e.g. keyword searches, extraction of unallocated space and file slack, file timeline/mapping, hidden data discovery/extraction, etc.); and
- Examine, analyze, and event reconstruct the data to answer critical investigative questions.

Presentation of Findings Phase

The purpose of the Presentation of Findings Phase is to *communicate* relevant findings to a variety of audiences, including management, technical personnel, legal personnel, and law enforcement. The authors selected "presentation of findings" over "reporting" to name this phase because it suggests careful consideration about how to best communicate information to various audiences. A technical report, which is the natural inclination of digital forensic analysts, tends to document relevant information, but does so in a manner that not necessarily helpful for those who then act upon the information provided. The Presentation of Findings may be written, oral, or presented in both formats. The presentation(s) are intended to provide both succinct and detailed confirmatory and event reconstruction information regarding the data examined in the Data Analysis Phase.

Incident Closure Phase

The purpose of the Incident Closure Phase is four-fold and includes the following steps:

- Conduct a critical review of the entire process and investigation to identify and apply lessons learned;
- Make and act upon decision(s) that result from the findings presentation phase;
- Dispose the evidence (e.g. return to owner, destroy, cleans and re-use (forfeiture)—all as applicable and legally permissible); and
- Collect and preserve all information related to the incident.

Comparison of Previous Models

The proposed model incorporates all phases and activities proposed by previous researchers. Each model can be easily mapped to the proposed model. Table 1 maps phases and steps from previously presented models to the proposed model's first tier phases.

Table 1.
Mapping of Previous Models to Proposed Model (First Tier)

	Prep.	Incident Response	Data Collect.	Data Analysis	Findings Present.	Incident Closure
Palmer, 2001 (DFRWS model)						
Identification		✓				
Preservation			✓			
Collection			✓			
Examination				✓		
Analysis				✓		
Presentation					✓	
Decision						✓
Department of Justice, 2001						
Collection			✓			
Examination				✓		
Analysis				✓		
Reporting					✓	
Reith et al, 2002 (Abstract Model)						
Identification		✓				
Preparation [for the current investigation]		✓				
Approach strategy		✓				
Preservation			✓			
Collection			✓			
Examination				✓		
Analysis				✓		
Presentation					✓	
Returning evidence						✓
Mandia et al, 2003						
Pre-incident preparation	✓					
Detection of incidents		✓				
Initial response		✓				
Formulate response strategy		✓				
Data collection			✓			
Data analysis				✓		
Reporting					✓	
Carrier and Spafford, 2003						
Readiness	✓					
Deployment		✓				
Digital crime scene investigation			✓	✓	✓	
Preservation			✓			
Survey				✓		
Documentation	✓	✓	✓	✓	✓	✓
Search and collection			✓	✓		
Reconstruction				✓		
Presentation					✓	
Review						✓
Nelson et al, 2004						
Initial assessment		✓				
Approach strategy (“design”)		✓				
Resource determination		✓				
Copy evidence			✓			
Risk identification & mitigation		✓				
Test approach strategy (“design”)			✓	✓		
Data analysis and recovery				✓		
Data investigation				✓		
Report					✓	
Critique						✓

Digital Investigation Principles

Certain principles apply to all phases of the digital investigations process, and should therefore not be cordoned off as distinct, discrete phases or steps. Doing so diminishes their impact on the overall process. Such principles represent overarching goals, which necessitate different actions be taken during each phase of the digital investigations process to achieve those goals. Two key digital investigation principles are *evidence preservation* and *documentation*.

Evidence Preservation

The goals of the evidence preservation principle are (1) to maximize evidence availability and quality, and (2) maintain the integrity of the evidence during the digital investigation process. During the Preparation Phase, steps should be taken to ensure the availability and quality of digital evidence when needed. During the Incident Response Phase, steps should be taken to ensure initial validation and assessment ("live response") activities preserve evidence. During the Data Collection Phase, which is where this principle is generally relegated to, steps should be taken to ensure data is collected in a forensically sound manner. Some example activities include creating forensic duplicates, employing write protection technology, calculating checksums and hashes, and employing environmental protections. During the Data Analysis Phase, forensic working copies are created as needed. Additionally, the analyst must be cognizant of which steps and processes modify working copies (e.g. file access times) and performs steps methodically from least invasive to most invasive and/or continually returns to use of clean copies. During the Presentation of Findings Phase, the evidence preservation principle suggests that findings should be communicated in a manner that facilitates future corroboration. Finally, during the Incident Closure Phase, proper evidence disposition measures are employed and information related to the entire process is retained.

Documentation

The goal of the documentation principle is to permanently (or semi-permanently) record all information relevant to and/or generated during the digital investigative process to support decision making and the legal, administrative, etc. processing of those decisions. Information that should be documented includes, but is not limited to the following (Palmer 2001; DoJ 2001; Reith et al. 2002; Carrier and Spafford 2003; Mandia et al. 2003; Nelson et al. 2004):

- **Preparation Phase:** Risk assessment and management information and decisions; policies; procedures; "known good" system information and hashes; training program requirements and progress; and legal coordination.
- **Incident Response Phase:** All information related to detection of the incident and any information pertaining to "who, what, where, when, why, and how" type questions; witness statements; and damage information, including direct costs and personnel time.
- **Data Collection Phase:** Information pertaining to the "state" of systems when data is collected (physical connections, processes running, network interface mode, date/time, open ports, etc.); evidence identification mechanisms (marking); and chain of custody information.
- **Data Analysis Phase:** Digital forensics tools used; processes, actions, and approaches taken during the analysis; and all findings, including things later deemed irrelevant.
- **Findings presentation phase:** Communication of findings from both technical and non-technical view points; and a formal record of relevant assumptions, observations, and conclusions. The findings documentation should be: timely; professional; follow the "ABC's of writing" (accuracy, brevity, and clarity); and be restricted to only what is known, not supposed.

- **Incident Closure Phase:** Permanently (or semi-permanently) retain all related documentation created during the digital investigations process.

Second-Tier Phases

Several authors who proposed single tier digital investigations process frameworks have suggested that additional sub-phase development is needed to provide adequate detail needed in the overall framework (Palmer 2001; Reith et al. 2002; Carrier and Spafford 2003; Mandia et al. 2003). As presented in the Introduction of this paper, the complexity level associated with the digital investigative process necessitates a multi-tier, hierarchical framework with objectives-based sub-phases that are applicable to various layers of abstraction. The second tier sub-phases should be inclusive of all possible types of crime and digital evidence and consist of tasks subordinate to specific objectives of interest. While the objectives-based sub-phases (OBSP) will remain largely consistent from situation to situation, the specific tasks that populate the sub-phases will vary from situation to situation, depending on the objectives sought in any given situation. Additionally, some tasks and sub-tasks may be applicable to more than one objective. As a result, the tasks can be matrixed to the set of digital forensic objectives as deemed appropriate. This enables the digital forensic examiner to quickly determine which objectives and specific tasks are applicable to the incident and approach strategy at hand. Similarly, researchers and tool developers can identify 'holes' wherein tools and methodologies do not exist at the desired level for objective achievement. The sub-phases of our model are depicted in Figure 2.

Each first-tier phase in the digital investigations process requires sub-phase development at this time for the desired framework to be complete. The overall hierarchy is depicted in Figure 2 below, where OBSP stands for Objectives-Based Sub-Phase:

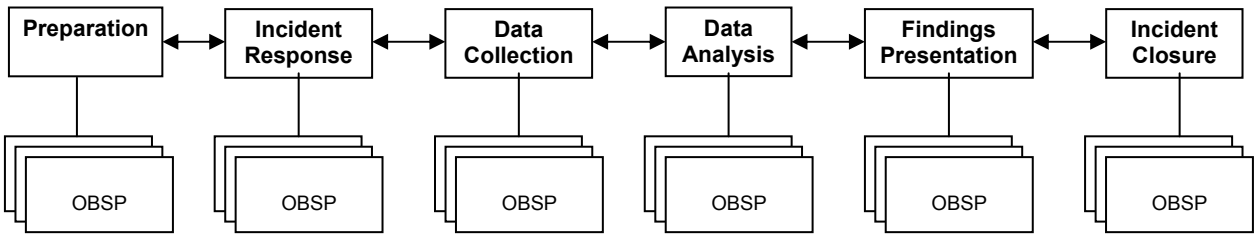

Figure 2.
Two-Tier Digital Investigations Process Framework

For the purpose of this paper, we are providing the detail of the Sub-Phase Framework only as it applies to the Data Analysis Phase. This phase was selected because of its importance, complexity, and the relative absence of detailed attention paid to it to date. Additionally, it provides the best opportunity to demonstrate the necessity for a hierarchical, objectives-based digital investigations process framework.

Review of Previously Proposed Analytical Phases

As previously presented, the purpose of the Data Analysis Phase is confirmatory analysis (to confirm or refute allegations of suspicious activity) and/or event reconstruction (answer "who, what, where, when, why, and how" type questions). Data collected during the Data Collection Phase is surveyed, extracted, and reconstructed during the Data Analysis Phase.

Digital investigations processes presented by prior researches have differed regarding the exact definition and, more importantly, the boundaries of the Data Analysis Phase. Several models dedicate separate phases for 'examination' and 'analysis' (Palmer 2001; DoJ 2001; Reith et al. 2002). In doing so, the

examination phase is primarily characterized by search and extraction activities, whereas the analysis phase is primarily characterized by subsequent activities that generate useful information from the extracted data. Nelson et al (2004) present a similar two phased approach to examination and analysis, except they apply a different label to their phases.

Mandia et al (2003) describe a single data analysis related phase, but subdivide it into "Data Preparation" and "Data Analysis" sub-phases. In this case, Data Preparation includes: file list creation, deleted data recovery, unallocated space recovery, statistical data collection, partition table and file system identification, file signature analysis, and known system file identification. The Data Analysis sub-phase includes: email/attachment extraction, installed application review, string searches, software analysis, logical file review, Internet browser history review, live system data collection review, network based evidence review, identify and decrypt encrypted files, and specialized analyses.

Carrier and Spafford (2003) describe three first-tier phases that relate to the overall data analysis function. These include the 'Survey Phase,' the 'Search and Collection Phase,' and the 'Reconstruction Phase.' The Survey Phase "…finds obvious pieces of digital evidence for the given class of crime…[and]…show[s] the investigator the skill level of the suspect and what analysis techniques the investigation will require" (Carrier and Spafford 2003: 11). The Search and Collection Phase consists of data extraction and processing (e.g. keyword searches, unallocated space analysis, file activity timeline analysis, code reverse engineering, encryption analysis, and log reviews). The Reconstruction Phase "…put[s] the pieces of the digital puzzle together" (Carrier and Spafford 2003: 12) and answers investigative "who, what, where, when, why, and how" questions.

Mohay et al (2003) describe two first-tier phases that relate to the overall data analysis function. These include: (1) 'the live system processing and data collection,' and (2) 'the analysis of secured data.' Their process is written more from a network forensics perspective, thus 'live system processing and data collection' includes volatile data acquisition, copying system files, logical volume imaging, and obtaining system date/time information. Likewise, 'analysis of secured data' includes logical analysis of the media structure, collecting operating system configuration information, file system mapping information collection and analysis, file signature analysis, identifying file content and type anomalies, evaluating program functionality, text string and key word searching, evaluating virtual memory, and evaluating ambient data. Overarching analysis techniques include system usage analysis, Internet usage analysis, time-line analysis, link analysis, and password recovery and cryptanalysis.

Proposed Data Analysis Sub-Phases

We purport that the analytical phases be bounded by one Data Analysis phase accompanied by an iterative set of sub-phases. To the maximum extent possible, first-tier phases should be sequential and non-iterative. This is not to suggest that they cannot be applied iteratively when needed, nor does it suggest that lessons learned from each phase are not fed back into the overall process for the purpose of continuous improvement. Instead, it merely suggests that when properly prepared and executed, first-tier phases are predominantly non-iterative in nature. Conversely, digital investigations experience suggests typical data analysis activities are very much iterative in nature. The typical digital forensic analysis process is characterized by decisions to search for certain data artifacts, subsequent data extraction and analysis, and then decisions to search for new, different, and/or additional data artifacts. These subsequent decisions are due to a variety reasons, including, for example, approach strategy refinement based on information obtained and discovery of unanticipated data artifacts or evidence. This is the reason Data Analysis sub-phases must be iterative in nature.

The proposed Data Analysis sub-phases include: Survey, Extract, and Examine (referred to as the "SEE Data Analytical Approach"). As is the case with general land surveying, the primary purpose of the Survey Sub-Phase is mapping. In a land survey, surveyors are tasked with providing a detailed, precise description of a land area including its topography, elevation, boundaries, geographical coordinates, conformity with standards, and object/spatial relationships. Analogously, in a digital data analysis survey, the analyst is tasked with providing detailed, precise descriptions of various aspects of a digital object's "landscape." Some examples include mappings of file systems, logical disk partitioning, disk geometry, "landmark" locations, and irregularities. This mapping data provides and enables the following: familiarity with the digital object under analysis, indications of suspect skill level, and location of obvious and potential evidence.

The Survey Sub-Phase mapping data then facilitates data extraction activities. The purpose of the Extract Sub-Phase is to extract data from the digital object according to stated objectives, using the mapping data from the Survey Sub-Phase. Techniques such as keyword searches, proprietary format data deconstruction, mining for hidden data, filtering, pattern matching, and file signature analysis are examples of Extract Sub-Phase activities.

The purpose of the third and final sub-phase, the Examine Sub-Phase, is to examine the extracted data to achieve confirmatory and/or event reconstruction goals. During this phase conclusions are made regarding the presence or absence of digital evidence (confirmatory analysis) and/or the answers to "who, what, where, when, why, and how" questions (event reconstruction analysis). Analytical techniques such as log reviews, image and text viewing, chronological and correlation assessment, and decryption are examples of Examine Sub-Phase activities.

The SEE Data Analytical Approach can be applied iteratively any number of times, with regard to any data analysis objectives, and targeting any type of evidence at any layer of abstraction. This approach is visualized in Figure 3.

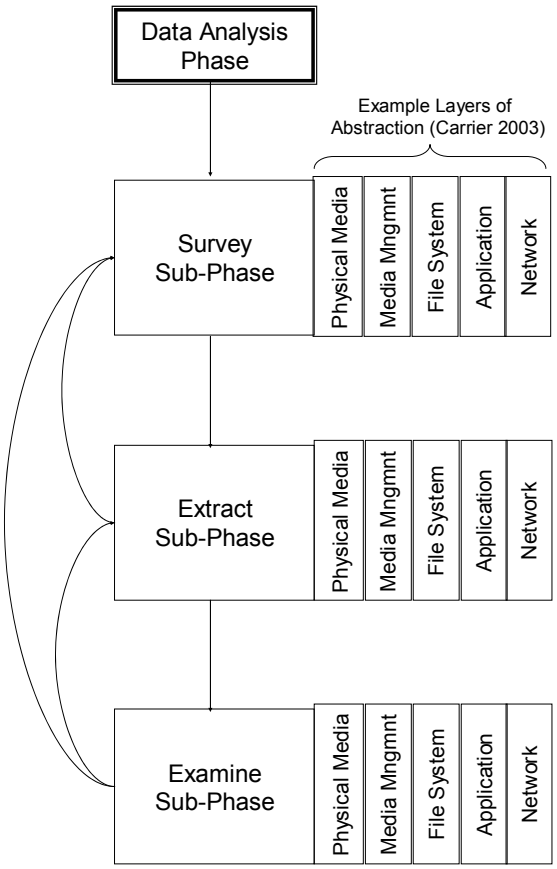

Figure 3.
SEE Data Analytical Approach

Objectives-Based Tasks for the Proposed Data Analysis Sub-Phases

The SEE Data Analytical Approach provides a phased approach to digital forensic analysis activities. The specific activities performed will vary according to the response and investigative goals of each incident. To help examiners develop an analytical approach strategy for any given incident, we propose an initial set of analytical objectives with subordinate task hierarchies. This way, forensic examiners can decide which analytical objectives are applicable to their given situation and follow the subordinate task roadmap pertaining to those objectives. The proposed set of objectives is arguably incomplete and is proposed to stimulate additional discussion and development within the digital forensic science and practitioner community. It is also assumed that this objectives-based task hierarchy will remain dynamic and change as technology changes. Figure 4 provides an initial set of data analysis objectives. Figure 5 then provides an initial task hierarchy based upon these objectives (DoJ 2001; Vacca 2002; Kruse and Heiser 2002; Mandia et al. 2003; Casey 2003a, b; Mohay et al. 2003; Nelson et al. 2004).

SEE Data Analytical Objectives

1) Reduce the amount of data to analyze
2) Assess the skill level of the suspect(s)
3) Recover deleted files
4) Find relevant hidden data
5) Determine chronology of file activity
6) Recover relevant ASCII data
7) Recover relevant non-ASCII data
8) Ascertain Internet (non-email) activity history
9) Recover relevant email and attachments
10) Recover relevant 'personal organizer' data (e.g. calendar, address books, etc.)
11) Recover printed documents
12) Identify relevant software applications and configurations
13) Find evidence of unauthorized system modification (e.g. trojan applications)
14) Reconstruct network based events

Figure 4.
SEE Data Analytical Objectives

SEE Data Analytical Task Hierarchy

1) Reduce the amount of data to analyze
    a) Signature analysis and filter out 'known goods' from the analysis data set
    b) Chronological ordering and focus
2) Assess the skill level of the suspect(s)
    a) Look for evidence of data hiding and wiping utilities
    b) Look for evidence of activity hiding
3) Recover deleted files
    a) Identify and recover deleted files via file system information
    b) Identify and recover deleted files via deleted file repositories (e.g. "Recycler")
    c) Identify and recover temporary files
    d) Rebuild deleted partitions
4) Find relevant hidden data
    a) Apply steganography detection measures
    b) Search for file signatures of known encryption schemes
    c) File signature analysis (file signature and extension mismatches)
    d) Search for data streams
    e) Search non-data storage areas for data
    f) Search for unusual hidden files and directories
    g) Recover passwords/phrases

(Continued on next page.)

SEE Data Analytical Task Hierarchy (cont.)

5) Determine chronology of file activity
    a) Analyze file mapping data
    b) Metadata analysis
6) Recover relevant ASCII data
    a) Conduct ASCII string ("keyword") searches at applicable levels of analysis (logical level, physical level, unallocated space, free space, file slack, swap, etc.)
    b) Apply automated filtering and/or indexing techniques
    c) Apply manual filtering and/or analysis techniques
    d) Export relevant data
7) Recover relevant non-ASCII data
    a) Signature based search and data extraction
        i) Repair damaged headers
        ii) Reconstruct file fragments
        iii) Utilize physical level "carving" (AKA "salvaging") techniques
    b) Apply automated filtering techniques
    c) Apply manual filtering and/or analysis techniques
    d) Export relevant data
8) Ascertain Internet (non-email) activity history
    a) Extract Internet browsing data
        i) Chronological history analysis
        ii) Search cache for URLs, Boolean searches, images, etc.
        iii) Cookie metadata analysis
        iv) "Favorites" analysis
    b) Extract "chat"/instant messaging activity (logs, name lists, cache)
    c) Extract newsgroup data (logs, membership lists, cache)
9) Recover relevant email and attachments
    a) Deconstruct email files stored on servers and workstations
    b) Recover deleted email an attachments
    c) Employ keyword searches
    d) Employ time-line based searches
    e) Employ participant (to, from, etc.) based searches
    f) Search for web-based email cached on media
        i) Search for base 64 encoded text
        ii) Search for email addresses
    g) Review temporary files for previously downloaded attachments
    h) Analyze headers of relevant email recovered
    i) Examine email server logs
10) Recover relevant 'personal organizer' data (e.g. calendar, address books, etc.)
    a) Deconstruct 'personal organizer' files
    b) Apply automated filtering and/or indexing techniques
    c) Apply manual filtering and/or analysis techniques
    d) Export relevant data
11) Recover printed documents
    a) Recover print spool documents
    b) Analyze print spool metadata (e.g. date/time, user, etc.)
    c) Correlate printed documents with those recovered physically and logically

(Continued on next page.)

SEE Data Analytical Task Hierarchy (cont.)

12) Identify relevant software applications and configurations
    a) Generate listing of installed and uninstalled applications
    b) Extract application configuration/initialization data
    c) Analyze extracted data
13) Find evidence of unauthorized system modification (e.g. trojan applications)
    a) Hash analysis
    b) Log file analysis
    c) Identify unauthorized user accounts, groups, and permissions
    d) Identify rogue processes and services
    e) Examine 'jobs' scheduled
    f) Analyze startup files
    g) Analyze configuration files
    h) Identify unauthorized network trusts
    i) Analyze unauthorized applications, tools, processes, etc.
    j) Identify electronic eavesdropping mechanisms
14) Reconstruct network based events
    a) Extract network device logs (e.g. intrusion detection system, firewall, router)
    b) Extract victim system logs
    c) Correlate activities between device logs
    d) Reconstruct chronological account of activity
    e) Determine incident impact (scope of compromise, business impact, damage)
    f) Examine network traffic collected via network monitoring tools
        i) Install network monitors
        ii) Define and apply filters
        iii) Reassemble and examine sessions
    g) Identify initiating Internet protocol (IP) address, source port and service, date and time, and *modus operandi*

Figure 5.
SEE Data Analytical Objectives-Based Task Hierarchy

MODEL DISCUSSION AND CONCLUSIONS

Benefits of the Proposed Model

Carrier and Spafford (2003) presented a five-point requirement set by which proposed digital investigative process models may be judged. These requirements are summarized as follows:

- Basis in existing physical crime scene investigation theory;
- Practicality—matching steps taken in actual investigations;
- Technology neutrality to ensure the process isn't constrained by current products and procedures;
- Specificity to facilitate technology requirement development; and
- Applicability to all possible user communities.

The first-tier framework, which consists of preparation, incident response, evidence collection, analysis, and incident closure phases, as well as the SEE Data Analytical Approach, which consists of Survey, Extract, and Examine Sub-Phases, both leverage lessons learned over time from the physical crime scene investigation process. Principles such as data preservation and documentation permeate all phases of the investigative process. Confirmatory analysis goals are differentiated from event reconstruction analysis goals. Finally, Investigative approach strategy development is a function of investigative objective selection.

The proposed digital investigative framework offers unique benefits over previously proposed frameworks in the areas of practicality and specificity. Previously proposed frameworks lack the level of detail needed to achieve practicality for investigators. This inadequate level of detail also hinders requirements development and gap analysis activities by researchers and tool developers. While the present effort is admittedly incomplete in that it only presents an initial sub-phase structure for the Data Analysis Phase, it proposes a multi-tier, hierarchical, objectives-based framework that provides the needed levels of practicality and specificity.

The proposed framework also fulfils the remaining two requirements—technology neutrality and wide user community applicability. The former requirement, technology neutrality, is relatively easy to achieve. The lack of neutrality in some models introduced to date is predominantly a function of the lack of concerted effort to apply scientific principles and develop a digital investigations process for the field. The latter requirement, wide user community applicability, is evident in the proposed first-tier framework. The challenge is in developing sub-phases that meet the needs of all potential user communities. We argue that the SEE Data Analytical Approach and objectives-based data analysis task hierarchy are applicable to all user communities, and will be much more so when the framework is enhanced via community input.

In addition to fulfilling the five requirements outlined above, a hierarchical, objectives-based framework has several other benefits. First, specific digital forensic activities will materialize for framework users by simply determining which data analysis objectives are applicable to specific incidents. For example, a typical network intrusion incident approach strategy will normally include objectives 1-7 and 12-14, omitting objectives 8-11 from the analysis approach strategy. Conversely, in a criminal child pornography investigation, objectives 1-13 will likely be of interest, whereas objective 14 probably will not be. Once the hierarchical task framework is sufficiently populated, specific tasks will be clearly delineated for each data analysis objective.

The proposed framework is also amenable to the development of helpful matrices. Such matrices can easily be developed that correlate tasks to objectives, tools to tasks, and capabilities to tools. This will provide practitioners, researchers, developers, and legal community members with an efficient means to:

- Clearly delineate a data analysis approach strategy and keep analytical objectives at the forefront;
- Identify analytical steps to take to achieve data analysis objectives;
- Establish analytical process standards and guidelines;
- Determine which tools can be used for various data analysis tasks;
- Identify and track which tools have been scientifically tested;
- Track margin of error associated with tools; and
- Determine where research and develop is needed.

Limitations of the Proposed Model

As stated previously, the proposed set of objectives is arguably incomplete and is proposed to stimulate additional discussion and development within the digital forensic science and practitioner community. This initial effort was predominantly focused on traditional, main-stream computer and network forensics in which hard drives are the primary digital device analyzed. Additional work is needed to ensure the proposed model is applicable across various layers of abstraction (Carrier 2003) and to other digital devices, such as personal data assistants (PDAs), digital cameras, telephones, and removable data storage devices. Collaborative inputs from both academic and practitioner communities are needed to increase robustness and rigor of the proposed framework.

It is also conceivable that platform (i.e. operating system) specific renditions of the task hierarchy may be needed. In the case of analyzing hard drives, a significant area of emphasis for digital forensics, tools, techniques, and procedures vary widely between file systems and operating systems. This again supports the argument for creating an objectives-based digital investigations process framework. Still, once objectives are identified, task hierarchies might be enhanced via additional levels of specificity unique to file systems, operating systems, etc.

Finally, the proposed framework would benefit from case studies—hypothetical and/or actual—with a methodical approach to identifying objectives and task hierarchies. Doing so would optimize objective and task development efforts. Additionally, it would facilitate scenario development to help users, researchers, and tool developers understand how to leverage and apply the proposed framework.

REFERENCES

Carrier, Brian "Defining Digital Forensic Examination and Analysis Tools Using Abstraction Layers." *International Journal of Digital Evidence* (1:4), Winter 2003, pp 1-12.

Carrier, Brian, and Spafford, Eugene H. "Getting Physical with the Digital Investigation Process," *International Journal of Digital Evidence* (2:2), Fall 2003, pp 1-20.

Casey, Eoghan *Digital Evidence and Computer Crime - Forensic Science, Computers and the Internet* Academic Press, Cambridge, 2003a, p. 265.

Casey, Eoghan (ed.) *Handbook of Computer Crime Investigation*. Academic Press, London, 2003b.

DoJ "Electronic Crime Scene Investigation - A Guide for First Responders," U.S. Department of Justice, pp. 1-82.

Kruse, Warren G., and Heiser, Jay G. *Computer Forensics - Incident Response Essentials* Lucent Technologies, Indianapolis, 2002, p. 398.

Mandia, Kevin, Prosise, Chris, and Pepe, Matt *Incident Response & Computer Forensics*, (Second ed.) McGraw-Hill/Osborne, Emeryville, 2003, p. 507.

Mohay, George, Anderson, Alison, Collie, Byron, Vel, Olivier de, and McKemmish, Rodney *Computer and Intrusion Forensics* Artech House, Boston, 2003, p. 395.

Nelson, Bill, Phillips, Amelia, Enfinger, Frank, and Steuart, Chris *Guide to Computer Forensics and Investigations* Thomson Learning Inc. - Course Technology, Canada, 2004, p. 689.

Palmer, Gary L. "A Road Map for Digital Forensics Research - Report from the First Digital Forensics Research Workshop (DFRWS) (Technical Report DTR-T001-01 Final)," Air Force Research Laboratory, Rome Research Site, Utica, pp. 1-48.

Reith, Mark, Carr, Clint, and Gunsch, Gregg "An Examination of Digital Forensic Models," *International Journal of Digital Evidence* (1:3), Fall 2002, pp 1-12.

Rowlingson, Robert "A Ten Step Process for Forensic Readiness," *International Journal of Digital Evidence* (2:3), Winter 2004, pp 1-28.

Vacca, John R. *Computer Forensics - Computer Crime Scene Investigation* Charles River Media, Inc., Hingham, 2002, p. 731.

ABOUT THE AUTHORS

Nicole Lang Beebe (nbeebe@utsa.edu) is a Research Assistant at the University of Texas at San Antonio, where she is working on her PhD in Information Technology. Previously, she was a Senior Network Security Engineer with the Science Applications International Corporation (SAIC), where she conducted commercial digital forensics investigations and information/network security vulnerability assessments for government and commercial customers. She has been a federally credentialed computer crime investigator for the Air Force Office of Special Investigations (AFOSI) since 1998 (she served on active duty through 2001 and as a Reservist since that time). She is a Certified Information Systems Security Professional (CISSP) and holds degrees in electrical engineering and criminal justice.

Jan Guynes Clark (jgclark@utsa.edu) is a Professor at the University of Texas at San Antonio, which is a National Security Agency (NSA) designated Center of Academic Excellence. Dr. Clark is a Certified Information Systems Security Professional (CISSP), has a Ph.D. in Information Systems, and numerous publications on a variety of information systems topics.