2010

Towards an Automated Digital Data Forensic Model with specific reference to Investigation Processes

Johan Scholtz
Auckland University of Technology

Ajit Narayanan
Auckland University of Technology

DOI: 10.4225/75/57b2b3b340ce2
Originally published in the Proceedings of the 8th Australian Digital Forensics Conference, Edith Cowan University, Perth Western Australia, November 30th 2010

Towards an Automated Digital Data Forensic Model with specific reference to Investigation Processes

Johan Scholtz and Professor Ajit Narayanan
Auckland University of Technology
Auckland, New Zealand
phdforensics@gmail.com
fzw0309@aut.ac.nz
ajit.narayanan@aut.ac.nz

Abstract

Existing digital forensics frameworks do not provide clear guidelines for conducting digital forensics investigation. However, had a framework existed, investigations based on known procedures and processes would follow strict prescribed standardisation. This should direct investigations following a set method for comparisons; ensuring future investigation is following one standard.

Digital forensics lack confirmed and tested methods; this became obvious when we consider varied interpretations of the same case by participants using different investigation methods. Previous research covered several approaches to setting a forensics framework, which are mere adaptations of previous models. We found that only a few models present a framework that defines or delivers qualified likeness between the different disciplines. From this, possible pattern analysis from different disciplines is possible (Kohn, 2007).

This underlines the need to standardise processes, to ensure proven and consistent results. Digital Forensics Science needs a new approach, defining and standardising investigation processes by affirming an investigation framework. Present research does not enough cover how existing forensic frameworks are used as guideline while conduct investigations. As a result, wide general interpretations are possible instead of following a set standard. Investigation processes and in particular how data confirmation is conducted during and after investigation becomes questionable as well. This also challenges data consistency and the legality of investigation processes when a non-standard framework is used without forming a sound theory based on proven models.

Keywords

Digital data forensics, automated investigation, forensic data bank, digital standardization.

INTRODUCTION

We question whether standardization of a digital forensics framework, would present a model whereby investigation could be automated. This automated digital forensics framework, may then prescribe specific processes, built on standardisation, guiding investigators to use the same model for confirming all cases. Collecting global forensic data to set up a corpora, unique to digital forensics is only possible if technicians moved towards an academic approach for investigation analysis.

We test feasibility of existing frameworks that would strengthen investigation processes. This was difficult to confirm from the collected data, because only a few participants had used a framework before, during and after their investigations. This opens the argument that a standardized framework would deliver consistent results. We also determined whether existing digital forensics frameworks are sufficient to conduct investigations covering aspects of interdisciplinary dependencies. If such dependencies exist, digital forensics might link certain crimes across disciplines' and find similarity in prediction of such cases. We questioned if existing digital forensics frameworks met forensic investigators' expectations, indicating completeness and if creating a standardised automated template is possible.

We also paid particular attention to resolving the obvious tension between investigation groups that appears to have a negative influence on the forensics discipline's coherence, since each group upheld a standing over the other's procedures. The mismatch between the expectations of forensic technicians and academics with different interpretations of results came to light. This highlights the problems for setting up a standardised platform. Participants' responses showed that regulation of certification and training, should present a standardized framework.

Investigation Processes

Secrecy and privacy of real digital forensics investigations influenced participant involvement. Participants were invited to contribute with 3 years and more investigation experience. Based on participants' feedback, we tried to discover whether investigators are consistent in producing repeatable results which peer investigators can confirm if they use the same investigation procedures.

A major challenge exist for digital forensics, for instance;"are automated investigation methods possible?" If an automated investigation method was followed it would inherently support a standardised framework. Investigations would then be carried out faster based on standardisation that allows recognition of both forensic technician and academic perspectives. Digital forensics investigators with an agenda of "just-get-the-job-done" might argue the digital forensics discipline does not need to follow a standardised platform. Reliability of actual investigation processes stresses the importance of proof and legality of data origin. Setting up a standardised control for digital data forensics would potentially show a firm commitment to specific management controls, which could lead to a fully automated process. Despite the need for standardising and confirmation of data consistency, this sequence is missing from participants' interpretations. *We suggest resolving this by creating a global corpus, with strict privacy guidelines in place while gathering data. To our surprise this was not considered favourably by the participants*. As a result, this study finds that participants reject the idea for a global data bank. One reason reflected from the participants' shows that privacy and secrecy of the trade as a protected field, is dividing participants between technical and academic investigators.

From the participants' feedback (Figure 1) we note that standardisation borders on impossibility, since there are too many different cases and forming a generic automated platform would be nearly impossible. Two distinct groups emerged; about 25% of the participants said they were interested in a new model, while the rest pointed out the present order of investigation was acceptable and required no changes.

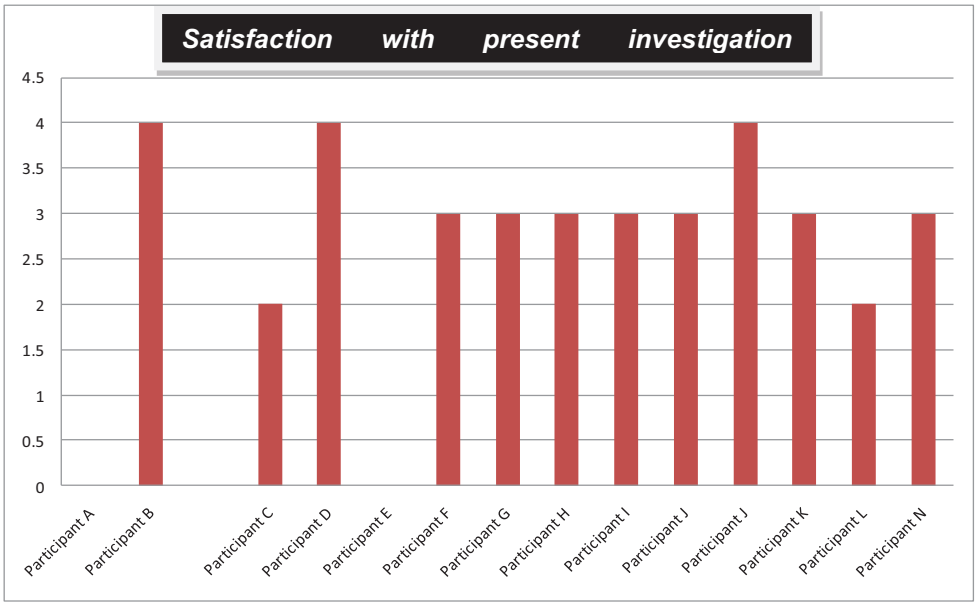

Figure 1 – Satisfaction with present investigation processes. A higher score suggest satisfaction with the existing Investigation Processes

EXISTING INVESTIGATION FRAMEWORK

Comparing existing frameworks, first we note similarity between a few concepts. We use the "traditional forensic framework" – a generic framework as described by (Carrier, 2004). Note the similarity of the concepts in (Table 1), adapted from Aanya-Isijola (2009).

Kruse and Heiser (2001)	*DOJ*	*Lee*	*Casey (2002)*	*DFRW, Palmer (2001)*	*Reith(2002)*	*Ciardhuain (2004)*
Acquire evidence	Collection	Recognition	Recognition	Identification	Identification	Awareness
Authenticate Evidence	Examination	Identification	Preservation	Preservation	Preparation	Authorisation
Analyse Data	Analysis	Individualisation	Classification	Collection	Approach	Planning
	Report	Reconstruction	Reconstruction	Examination	Strategy	Notification
				Analysis	Preservation	Search and Identify Evidence
				Presentation	Collection	Collection
				Decision	Examination	Transportation
					Analysis	Storage
					Presentation	Examination
					Returning Evidence	Hypothesis
						Presentation
						Proof/Defense
						Dissemination

Table1. Frameworks comparison

From a traditional forensics framework and comparing different investigation models, we try finding a broad framework that lends itself to automated procedures. While listing participants' arguments for and against automated investigations, we note divided interest in creating automated procedures. The *present participant group did not use any automated investigation processes and corpora creation making this a separate issue for future research.* Although these processes are separate issues, setting up a digital forensic corpus should be a starting point. We suggest that data queried from this data base, would allow possible automated investigation processes.

Differences between two distinct groups emerged - uncertified investigators in contrast to certified investigators. We might add another group here – academic investigators. The latter group focuses on interpretation and prediction as well as on setting a range of standardised types of questions. *Academic investigators adopt an investigative or comparative analysis, whereas the uncertified investigators mainly complete their investigations at result level, without deeper analysis. No reference is made to predictive analysis from this group.*

Underlying tension became clear between investigators with an academic approach and forensic technician investigators that are only interested in solving cases. No firm agreement for setting standardised investigation

and training methods has been reached yet among the professionals in this field. This underlying tension might even be resolved with an alternative practical approach to either retrain technicians or retrain academics trying to find a midway in framework and applied similarity, providing the retraining adheres to a standardised process of investigation. Currently these two opposites are extremely divided. (Hom-anek, 2009). Only a few participants make use of a set framework, suggesting that only a small group follows a (standardised) method, while conducting investigations. This further emphases' the need for standardisation of investigation procedures and training. Future automated processes, for instance *promoting automated investigation rather than automated software methodologies* should be the order of the day.

AN OVERVIEW OF EXISTING DIGITAL FORENSICS

Previous research covered several approaches to setting a forensics framework, which are adaptations of previous models. We found that only a few models present a framework that defines or delivers qualified similarity between the different disciplines. From this possible pattern analysis from different disciplines is possible (Kohn, 2007).

(Figure 2) presents small differences between participants' views on the importance of using a traditional framework model. Reporting is considered by participants being similar to preparation and awareness as we determined from other frameworks. Participants consider Data Comparison as the last activity, or the least important in the investigation process. This supports our suggestion that digital forensics investigators do not conduct data comparisons, mainly because no corpora exist to compare it against, or they do not see the need in finding similarity in cases.

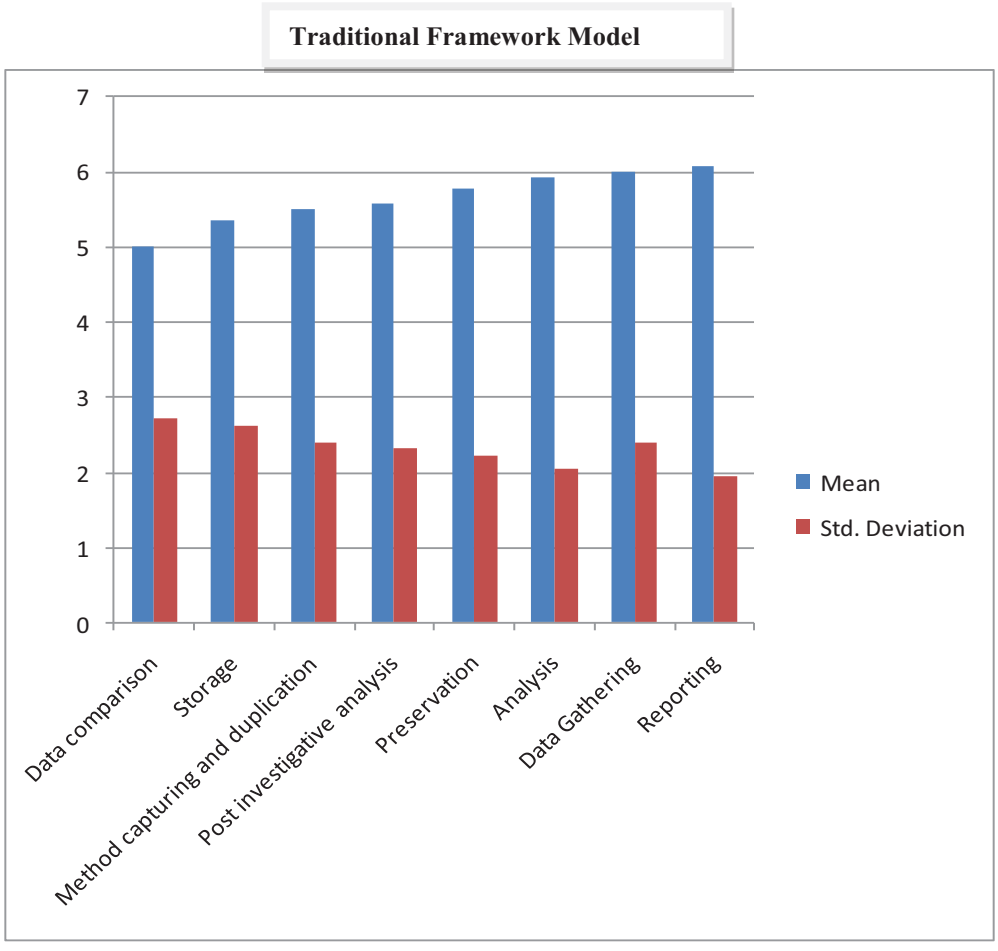

Figure 2 - Importance of using a traditional framework model. A higher score suggests importance.

Traditional data comparisons to test findings have never been conducted because it never has been proved. We noted a good link between the higher scale of the traditional framework processes, for instance, Reporting followed by Data Gathering and Analysis.

When we look at Data Comparison as the least important, this fits results from participants, suggesting that if participants are not contributing to a database; they have nothing to compare their results against. Clearly data comparison is not important to all investigators. Noteworthy is the importance of Post-Investigative analysis. This response is contradictory to the actual benefit of collecting data, since only one participant uses the data effectively after collection for analysis.

Although, this fits well with 50% of participants who point out they collect previous data from a post-investigative prospective, they are not using it efficiently.

We also recorded whether investigators use frameworks and if changes to their frameworks, or early theories would have some impact on the way they conduct investigations. It is assumed that investigators should at least follow a method for archiving their investigation processes while they are performing their investigations. Earlier research covered several approaches to setting a forensics framework, which are adaptations of previous models. We found that only a few models present a framework that defines or delivers qualified similarity between the different disciplines. Pattern analysis from different disciplines should be possible from these models.

We examined several frameworks and expected to find a model that would make data extraction from corpora possible and to group data for clear investigation analysis (Selamat, 2008). This would allow analysis of expected clustered data. Binding features of alternative procedures to investigations and presenting a new model of standardisation was expected. Therefore we recommend creating a framework that handles automated predictions.

A few points of interest showed in the research:

- Existing forensic investigative frameworks do not allow automated fast tracking of digital data investigations. Finding associated data clusters of specific digital forensics activity is still being researched. (Olivier, 2009). Oliver also suggests using a multidimensional model compared to one-dimensional file systems, making alternative relational structures possible.
- It is the author's view that automated search would enable other disciplines to find touch points to the case scenario associated to their own investigation procedures. For instance, creating a multi-disciplinary relational structure that includes networking and real-time mobile disciplines might create links to other processes at a digital forensics sub-level. From this perspective, digital forensics researchers can create specific scripts capturing mismatched or irregular pattern analysis of digital data that allows a customised automated investigation framework.
- We explored if investigators conduct re-assessable and verified investigations, based on their early theory. Investigation processes are questioned if a standardised framework was not in place through detailed procedures and processes. The importance of creating forensics corpora on a reproducible forensics platform with accessible real data was also mentioned as alternative solution.
- The feasibility of a framework that could potentially lead to develop new procedures and processes, using an automated forensic investigation was explored. It is envisaged that a workable automated template would support a scientific investigative process by setting up a firm scientific research platform.

INVESTIGATION GUIDELINES

When a general formula of a standardised method or framework exists, these models would assist the non-forensic expert to conduct investigations according to specific guidelines. We see this issue emerging from research by Garfinkel (2009) who states that forensic science is not yet a true science because the research community has not so far adopted understanding and rigour of reproducible test results. This is clear from a wider understanding when perspectives on re-creating forensic corpora are considered. Although a few frameworks try to prove a model to which certain procedures could be linked, most procedures still do not deal with the core issue – the gap between technical aspects of digital forensics and the judicial process. This is an unqualified problem as the argument between technical specialists and legal practitioners' boils down to whom knows best (Broucek, 2006).

The underlying ideas of reconnaissance, reliability and relevance are beyond the ability of the existing framework the most important to set up a link between factual information and judicial review. (Leong, 2006).

We note that at least one generic framework exists that, defines the phases of investigations. This is a basic framework and does not allow for deeper investigation processes.

Establishing and Organising Forensics Capacity

Digital data forensics has been on a fast track by setting up a defined framework and is still a young science as Ahmad (2006) shows, compared to other sciences. A few frameworks currently exist that address basic investigation processes, although digital forensics is a *relative young* discipline they do not show enough adaptability to the ever changing digital forensics science. This raises further questions when proposing a framework that works according to strict and consistent procedures and processes. It has to be noted that "relative young discipline" when referring to digital forensics is an obsolete term and although used by most researchers, it has been in use for about 15 years now.

We also note this from a study by Brinson (2006) that refers to the infancy of cyber forensics. The lack of awareness about various related disciplines hamper the forensic investigation process because no framework exists so far that caters for investigation processes and combines academic development of specialist skills but still covers other disciplines as well.

Several frameworks have set the ground rules for digital forensics investigations; Carrier (2004) focuses more on an event-based digital framework, whereas Beebe (2005) proposes a framework that is objective-based. Both these frameworks in essence still refer to the traditional model initially proposed by Palmer (2001). In contrast, we note that Baryamureeba (2004) suggests another approach that separates the investigations at the primary and secondary crime scene while describing the phases as iterative rather than linear. As we found out from participants, most companies develop their own procedures for examining gathered data. These basic processes of acquisition, identification, evaluation and admission as evidence form the foundation of the forensic framework. However, we need to analyse these procedures in depth to show a more defined structure. This is because researchers need to distinguish between investigations frameworks that define the physical investigation steps compared to a more abstract framework that defines abstract entities as described in (Carrier, 2003).

For example, "Abstraction layers occur in multiple levels. The file system itself is a layer of abstraction for the stream of bytes from the disk media. Within the file system are extra layers of abstraction and the result is a smaller stream of bytes that represents a file, which is then applied to an application level of abstraction and it is processed further."(Carrier, 2003).

Given the complexities of abstraction, there is room for error when the meaning of phase, activities, components, processes, stages, steps and classes are misinterpreted. A more recent framework suggested by Freiling (2007) focuses on analysis as a means to improve the investigation. It involves pre-incident preparation, pre-analysis, analysis and post-analysis.

As with most other framework, these are mere guidelines aiding the investigators in assessing data and compiling these into unique groupings. We state this because no study yet has delivered a framework that prescribes digital gathering procedures for the different interdisciplinary procedures.

The importance of investigation varies as participants' responses show in (Figure 3). From participants' feedback, we showed the lowest ranking in this graph suggests the order of importance or actions taken by participants during investigation. In other words, most investigators started with collection as the first action and only later did they consider duplication, theory and prediction.

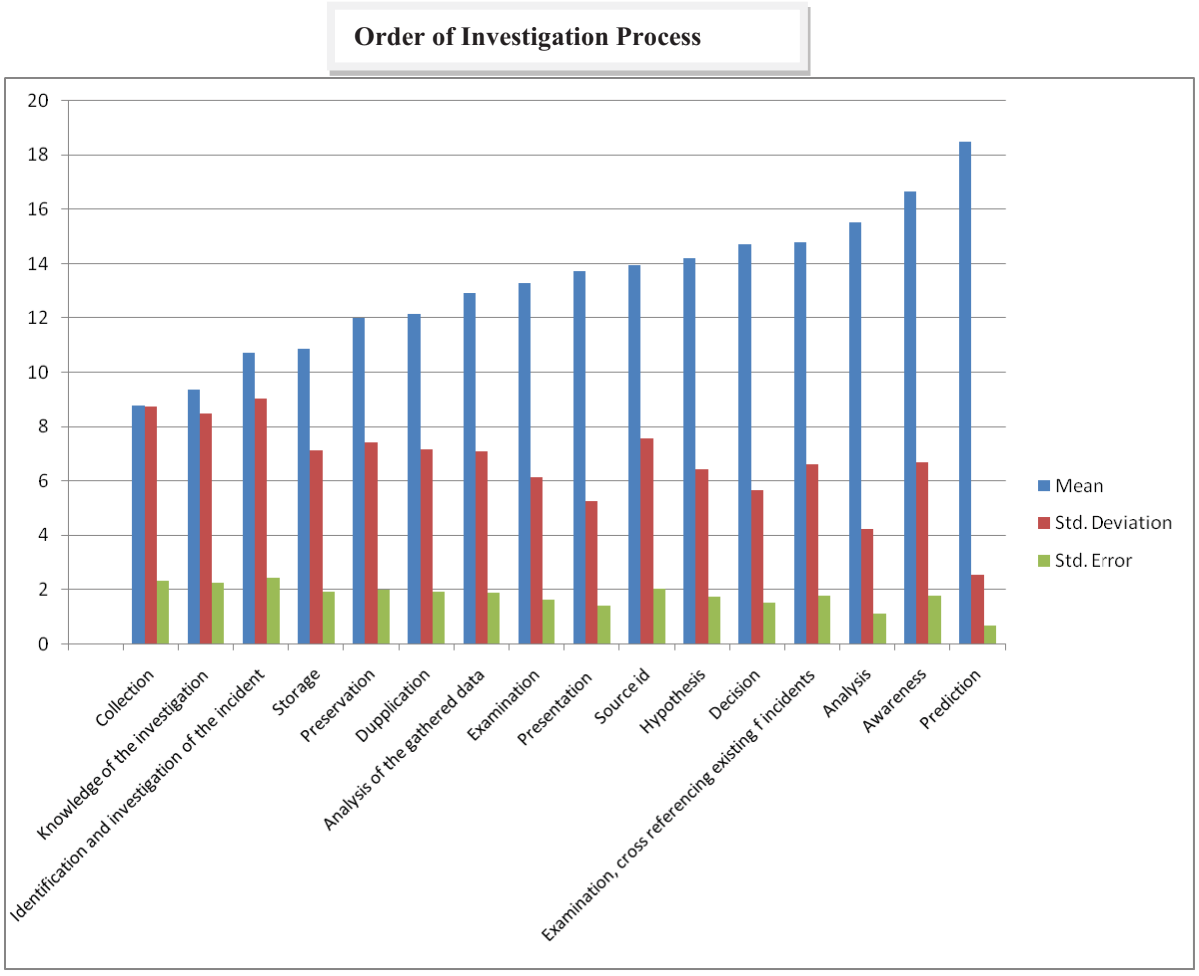

Figure 3. A lower score is more important or occurs first in the investigation process

Figure 3, also supports the idea that investigators do not consider prediction as an alternative at this stage. Since reasonable prediction is only possible after data analysis from a data bank, we might not see this happening soon since the digital forensic discipline does not yet have corpora to test against. This also affects automated prediction as a vast number of data is required to analyse pattern instances. From the gathered data we determined that some participants are more concerned about collecting data than to predict data at the other end of the scale.

Further, Radack (2009) shows a basic forensic investigation process consisting of collection, examination, analysis, and reporting. These processes are better understood from in a wider procedural basis of:

- Preparation: case briefings, engagement terms, interrogatories, spoliation, prevention, disclosure and discovery planning, discovery questions;
- Record: Drive imaging, indexing, profiling, search plans, cost estimates, risk analysis;
- Investigate: Triage images, data recovery, keyword searches, and hidden data review, communicate, iterate;
- Report: Oral vs. Written, relevant document production, search statistic reports, chain of custody reporting, case log reporting;
- Testify: Testimony preparation, presentation preparation, testimony.

Re-classification and training

(Beebe, 2009) states that digital forensics is lacking a common body of knowledge. When we consider the digital forensics industry is ever changing to new technology, it is easy to see that an existing body of knowledge is ever changing. This might lead to miss-interpretations or confusion of complex forensic concepts and finally leads to difficulty in getting credited results, this in turn influence time and resources in solving cases. Some investigators maintain a "get the job done approach" thereby putting defendants at risk, since proper case preparation and investigation might not be their focus. These investigators often maintain "I am a professional forensic investigator" attitude; this might not necessarily reflect their skill or experience. Setting a certification regulated pre-requisite linked to proper training would ensure investigators are in fact as good as they say they are.

Therefore, if investigations are not based on a sound theoretical basis and do not use a theory that draws up a framework describing the processes followed, we assume investigators would not find a connection point in solving the case or proving where these principles originated from, thus failing to present academic reasoning or comments in court.

If there is no certification, then no awareness exists of academic interpretation and prediction – thus no standardised platform is present. We now could expect un-classified findings from unskilled investigators that might not follow a standardised framework.

Digital Forensic Investigators should prove an inquisitive approach to solving cases, backed by theoretical knowledge about finding and associating hidden data. This should at least be the expectation, since defendants should get the best protection allowing the benefit of doubt. Developing a Digital Forensics Body of Knowledge (DFBK) has been slow for several reasons; including the lack of experienced participants and the lack of collaboration among digital forensics professionals. (Hom-anek, 2009)

We suggest having high-level, peer-reviews in a selected journal that is published for a selected audience, by it identifying and guiding training and certification needs, since technological progress demands constant retraining and updating to lessen redundant information. Given the responses we received and further assumptions of the digital forensics trade in the field, we cannot confirm that all aspects of responsible investigation processes are followed by the participants. Based on our findings to date, we cannot support whether investigators with limited academic knowledge should be able to defend a case in court – irrespective of field experience. If such an investigator makes it to the court, the process of arguing a rigid standardisation procedure loses accreditation.

Hannan (2003) also voiced a similar view. Setting a high standard of certification and academic research should bind this to an implementation model thus following specific recommended guidelines. This would allow seamless updating of specific procedures when challenged by newly developed processes. Also, this will ensure all investigators are bound to controls and will set a standardised working discipline.

Once a standard for investigation is set, creating a global data bank should be possible. Further to this, if relational data structures are settled, bulk data generation is possible that allows for sifting of data through predicting and identifying future happenings. We also foresee using semantic search functions, interlinked with predictive models whereby pattern recognition and re-occurrences of similar crimes are indentified.

As this paper shows, future research in this particular field is required; however, getting enough data might hamper its progress.

Frequency of Preferred Investigation Software

Participants showed a preference for Encase, FTK manager or EncaseFTK to conduct their investigations (Figure 4.) We were surprised to see Ubuntu on the lower end of the scale, since this is a free Linux based operating system; we expected this to have a higher ranking based on specific application functionality. In addition, various software packages were explored in this research. It seems that only a few core packages were chosen for the sake of simplicity in conducting investigations. High costs of getting off-the-shelve software and yearly maintenance fees make this very costly for novice investigators.

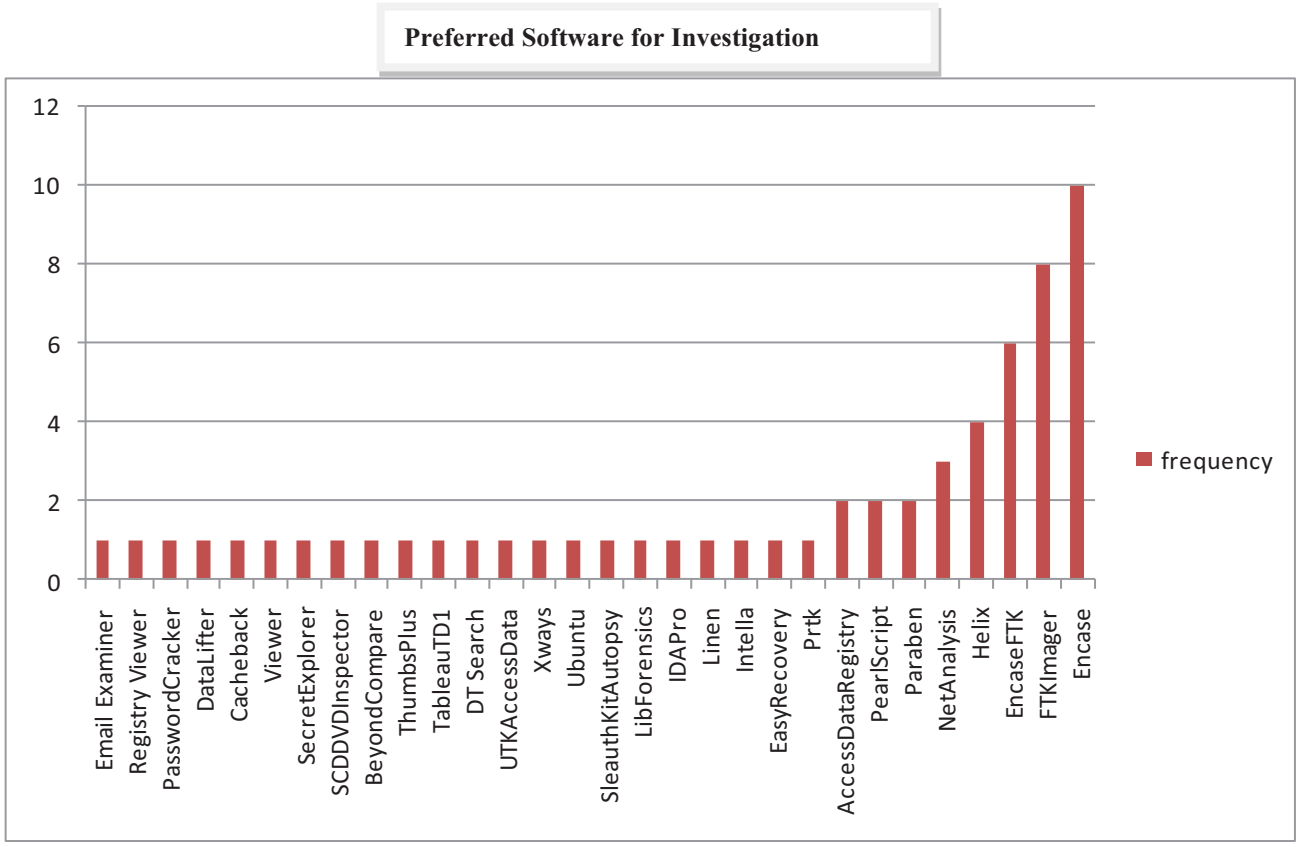

Figure 4 – Software Preferred by Survey Participants – the higher score is more important

Automated software seems to guide investigation intensity instead of automated procedures. As the field study suggests, these investigators are also known as "button investigators" since they do not have proper training and insights of underlying investigation processes. It becomes questionable if these investigations reflect the investigators' real experience and skills or we only get software-based solutions without interpretation? It seems that a few core investigation software packages have been written and most investigators follow this trend even though new investigations requires alternative approaches, geared towards demanding research of file fragment tracking.

To the contrary, it seems that only a few investigators follow an academic approach by writing specific scripts, solving challenging cases.

If the present prescribed and recognised methods, as suggested by industry, are the only model to work with, investigators might present incomplete solutions without recognising a standard requirement. Extensive training and re-skilling might correct this shortfall.

RECOMMENDATIONS

Creating a Forensic Corpus

Recent research (Kahvedzic, 2009) proposed a model to describe an investigation at different levels of detail. This suggests that independent vocabulary can be used to describe the researching process in more detail. In similar manner, we could use this notion to present a data structure that defines specific groupings of similar concepts and their qualities, by it ensuring representation of variables in a relational data structure. This database should be scalable ensuring new entities are related to the existing structure. Global forensic researchers would input data accordingly into this database. We envisage a noticeable benefit to the forensic community if members contribute and share their resources. This would provide true data sets that would help in setting up a platform in digital forensics whereby an automated framework might devise. Predictions are inherently impossible since every digital crime is different.

Digital Forensic Automation (DFA) is not yet possible because of the diverse data bank structure and the lack in member contributions. Data banks will have to be created with a vast number of tables and covering many, although not all possible scenario variables. According to (Garfinkel, 2007) such a database does not exist. If we had enough data to build a structured taxonomy in forensic modelling which defines how we would conduct investigations, and produce groups of similar clusters, we might succeed in getting a higher accuracy level. This would also lead to a platform whereby associations among predicted data are more defined from a procedural point of view.

(Rurbin, 2005) proposed a framework that displays the benefits of computer intelligence technologies. It uses automatic evidence extraction and provides a basis to build more knowledge through reusability. It is the author's view that such a corpus should be precisely modelled allowing sub-level classification and expansion. This might result in a Real Digital Forensics Data Corpus (RDFDC) that would be useful in data analysis.

(Beebe, 2009) suggests using an Intelligent Analytical Approach where artificial intelligence and other intelligent search would enable successful retrieval by making use of algorithms. *This supports my view that higher emphasis should be placed on semantic rather than literal searching techniques that should substitute traditional literal searches. This allows for a structured, but still adaptive, relational data structure by improving data indexing.* This would eventually present a match based on "fuzzy hashing" which requires a complete paradigm shift from traditional forensics investigation approaches. This means we should step away from the overwhelming traditional search patterns and move to prediction of similar cases. We suggest using predictive Markov models, analysing data for predictive similarity in events and consider a fuzzy reclassification of data models. Using a Fuzzy logic approach in data classification and clustering, presenting a new approach into re-classification that is not bound to factual rigour, but rather focuses on occurrences and predictability.

We now need to discover how these top structures would look like and how they can be build in such a way that quick changes can be made within the framework, whilst still preserving the basic flow of the investigation. This would allow development within collective phases and improvements as sub-phases are built on comparisons that lead to data mining options. Research by (Kohn, Eloff, Olivier, 2008) is considered here based on assumptions and suggestions with an end result of building a fuzzy analysis of likely outcomes of the investigation.

Enhanced Automated Investigation Framework

We would like to create a (RDFDC) from global forensic contributors. However, we foresee difficulty creating a databank that allows automated procedures. Creating a new predictive automated model, as presented from the research findings has fairly low-level approval from participants. Nevertheless, in equal comparison, creating a model based on the existing frameworks did not look promising. This is because a vast number of variables play a role and it is difficult to discover the likelihood of similar events in a digital crime.

One of the drawbacks had been the small corpora of research in forensic data banks. This is because of the non-existence of a proper data bank to confirm case findings. Setting up a relational database should strengthen and reflect the reliability of these predictions.

A view response from participants shows a shift to new methods of investigation and database creation. For instance, better classification could lead to better predictions based on many case examples. Given that many inconsistencies and preferences exist in various data gathering methods, questions about the feasibility of an automated digital forensics method is raised. We note a strong response for and against automated procedures. A participant reflected on forensic automation as follows:" *Forensic automation is already becoming a problem by giving untrained examiners a false sense of security when in reality; they are not conducting an examination at all. When used properly some automation is good. However, it is not to the point where any time soon, an automated tool can conduct a thorough enough examination to be trustworthy."*

Most concerns are that no investigation is ever the same as before and automated procedures might miss unique or complex associations which only an investigator can interpret. This is arguable; since human investigators might not always associate similar cases from the databank with present investigations, when similarity of the cases exists. However some participants presented a clear preference for automation, suggesting the importance of having a basis of early investigation processes. This notion is similar to the author's perspective that a tiered approach from first level investigation allows a more detailed assessment of the reasons unique to the particular investigation scenario.

The term automated forensics and automated tools are quite often misunderstood – as the participant above comments. We also note reference to software that reflects a click and drag scenario, which does not promote an academic research platform. We suggests automated forensics should prescribe the processes of investigation when the investigator makes use of previous data, based on predictive analysis from a data bank, which contains previous data and make use of forensic software to a lesser extent.

We suggest changing the mindset from "automated software" to "automated analysis" whereby investigators could sift through the first level of classification and control sub-levels of the investigation, with running of specific code scripts – suitable for level comparison and prediction. The question now is "what" to do with the data and related issues like sub classification and inter-relational dependencies. After the classification takes place the investigator move to the level of "how" to do it. The full spectrum of the investigation takes the form of Methods (What) vs. Procedures (How).

Automated responses would only be successful if a big enough (RDFDC) had been set up. This would enable creating a data set that supports forensic examiners' to search for the best combination of words or relevant case selection identifiers. Researchers are not aiming for automation because they do not have enough large corpora of forensically interesting data to develop reliable automated algorithms and tools. Instead, much research in both the academic and corporate worlds has underlined developing interactive visualisation tools. Since they are designed to be performed by a trained individual, tool failures can be more readily tolerated. Questions about "forensics expert's" ability come to mind.

Investigator preparation

Aspiring investigators are drawn to short courses, often a day or a week long that would lead to potential misinterpretations of complex digital forensic cases. Existing certification training programmes, for instance the Global Information Assurance Certification (GIAC) and the SANS Institute are forerunners in certification pathways. To become a GIAC Certified Forensic Analyst (GCFA), you are only required to pass one proctored exam (150 questions, with 4-hour time limit) and achieve 69.3% (104 of 150 questions). The SANS Computer Forensic Investigations and Incident Response certification, covers a 6 days training session. (SANS, 2010) and (GIAC, 2010)

Motivation for this is supported by recent research by Hom-anek (2009) who reports their findings on how information security training forms a basis for digital forensics qualification. Digital forensics professionals are regarded as having functional roles when performing their skills and training. Hom-anek also notes that only 42% of digital forensic investigators have competency that matches their job roles. This figure confirms that training is a major shortfall in most investigators' skills package.

We should consider all choices when conducting investigations, where earlier exposure to similar cases and level of expertise is important. Enforcing a grading level that allows only experienced investigators to conduct investigations at a specific level presents a degree of professionalism in the digital forensics discipline. This means that representation in court should only be allowed for those that are skilled enough. Theoretical expertise sets the standard that reflects a comprehensive understanding of the specifics of any particular investigation case.

Points of interest

- Research to date provides evidence of forensic frameworks that only provide guidelines for major forensic instances.
- The field study shows regulation of training and certification might provide a basis for standardising academic requirements for this discipline. (GIAZ, 2010)
- Extensive research based on a data base structure is required to enable predictions based on existing data. Proposed forensic scenario based on an initial generic platform would form the first stages of the research.
- Further development of a dynamic framework would enable sub-level associations/clusters. With hidden Markov and fuzzy logic implementation this would allow smaller data sets to be used with more certainty, and would also still allow for predictive assumptions.

CONCLUSION

We found that even with the number of frameworks discussed in this paper and the possibility of having only one framework that include all others, we still would not convince investigators to use a generic framework, since all investigations are different. The challenge is to set up a framework that produces a "gliding scale" of possibilities; this scale could then be used with a data base that contains the same entities and variables of the most consistent occurrences of similar types of crimes, by matching possible case results with predicted results. This will open new a direction for predicted analysis of digital forensics to regulate a standardised new approach to investigation processes.

This paper addressed the difficulty in settling a framework the digital forensics industry follows. As our research show participants hardly use a standardised framework during investigations. As discussed in this study, monitoring, logging and preservation of case data need to form a data bank, by setting up corpora for confirmed methods of fuzzy approaches and predications. It is questionable whether forensic investigators would follow a standardised procedure at all—considering they have been following their own customised methods so far. This presents a problem for standardisation and eventually automation. On training and skill improvement, we noticed that a few participants regarded own experience higher than formal qualifications.

SUGGESTED FUTURE RESEARCH

We propose further study in forensic profiling, particularly establishing a basis of interaction between automation and profiling, thus creating a stepping stone for early time-saving when a typical investigation is conducted. (Rogers, 2003) reference a research by (Pethenck, 2002) showing how criminal profiling might be achieved if a broader guideline is used as it was earlier suggested by the FBI. If we use this FBIs typology which was criticised for not having enough empirical testing, we might come up with a higher socio-criminal identity, by classifying potential criminals according to characteristics of typical groupings. We might find similar patterns emerging when users in the computing field are studied.

(Elsaesser, 2001) as referenced by (Stallard, 2003) presented an approach to create an automated hypothesis of computer attacks. This earlier research has potential for further development. This framework then simulates the computer attack and assumes matches to a target configuration using recognition techniques through searching for unique supporting data or patterns of the investigation. By using this approach we could also broaden the use of data extracts from data bases, finding the relation between fields and tables thereby getting patterns in similarity if constrained items are sifted using redundancy validation. This process of linking characteristics of a specific crime through the tiered level descriptors would allow a gradual disqualification of redundant data/characteristics and would lead to a sub-level classification or grouping of the crimes. This would be an eventual tool for aiding forensic investigations.

We support more recent research from Beebe, on a realignment stepping away from the overwhelming traditional search patterns and moving to prediction of (similar) cases.

We now move focus from standardisation in digital forensics to the cloud forensics environment. We expect finding and presenting choices to the user to manage their data with more control. From a user's perspective, everybody wants to have more permanent control over their data; this includes support of deletion on all servers, movement of data between servers and countries as well as better encryption. Allowing more freedom for the user to manage their files and being able to make a choice between servers. Also confirmation of permanent file deletion should give users encrypted privacy of choice.

From a digital forensics perspective, cloud investigators should have a sure way to track users and fragmented files over a series of servers, and being able to compile a directory of such data. Prediction of server locality based on the likelihood of distribution and storage is also investigated

REFERENCES

Aanya-Isijola, A. (2009) *Models of Digital Forensic Investigation*. School of Computing & Technology. University of East London.

Ahmad, A. (2006). *The Forensic Chain-of-Evidence Model: Improving the process of Evidence Collection in Incident Handling Procedures.* Department of Information Systems. Parkville, VIC, University of Melbourne.

Baryamureeba, V., Tushabe, F. (2004). *The Enhanced Digital Investigation Process Model*. Institute of Computer Science. Kampala, Uganda, Makerere University. http://www.dfrws.org/2004/day1/tushabe_EIDIP.pdf

Beebe, N., L. (2009). *Digital Forensics Research: The Good, the Bad, and the Unaddressed.* 5th Annual, IFIP, WG 11.9. The University of Texas at San Antonio. http://faculty.business.utsa.edu/nbeebe

Beebe, N. N., and Clark, J.G. (2005). A Hierarchical, Objectives-Based Framework for the Digital Investigations Process. Department of Information Systems and Technology Management. San Antonio. The University of Texas at San Antonio.

Brinson, A., Robinson, A., Rogers, M. (2006). *A cyber forensic ontology: Creating a new approach to studying cyber forensics*. Department of Computer & Information Technology. West Lafayette, Purdue University.

Broucek, V. T., P. (2006). Winning the Battles, Losing the War? Rethinking Methodology for Forensic Computing Research. School of Information Systems. University of Tasmania.

Carrier, B. (2003). "Defining Digital Forensic Examination and Analysis Tools Using Abstraction Layers." *International Journal of Digital Evidence 1(4).*

Carrier, B. D., and Spafford, E.H. (2004). *An Event-Based Digital Forensic Investigation Framework*. Center for Education and Research in Information Assurance and Security - Cerias. West Lafayette, IN 47907 USA, Purdue University.

Elsaesser, C., Tanner, M. (2001). *Automated Diagnosis for computer forensics*. The Mitre Corporation.

Freiling, F. C., Schwittay, B. (2007). A common Process Model for Incident Response and Computer Forensics. *Proceedings of Conference on IT Incident Management and IT forensics, Germany.*

Garfinkel, S. L. (2007). "*Forensic Corpora: A Challenge for Forensic Research."* http://www.simson.net/ref/2007/Forensic_corpora.pdf

Garfinkel, S., Farrell, P., Roussev, V., Dinolt, G. (2009). "Bringing science to digital forensics with standardized forensic corpora." *Science Direct 6(S2-SII).* www.dfrws.org/2009/proceedings/p2_garfinkel.pdf

GIAZ. (2010) GCFA Certification Bulletin. http://www.giac.org/certbulletin/gcfa.php

Hannan, M., Turner, P. (2003). *Australian Forensic Computing Investigation Teams: Research on Competence*. 7th Pacific Asia Conference on I.S. 1-13 Adelaide, South Australia.

Hom-anek, P., Apiwathanokul, C., Nachin, N., Pamomchaisirikit, S. Sripeamlap, T. (2009) Career Opportunities and Development for Asia Information Security Professional with the IT. Security Essential Body of Knowledge (EBK). http://www.tisa.or.th/downloads/6.TISA_TISET_Final_Presentation.pdf

Kahvedzic, D., Kechadi, T. (2009). "DIALOG: A framework for modelling, analysis and reuse of digital forensic knowledge." *Science Direct 6(S23-S33).*

Kent, K., Chevalier, S., Grance, T., Dang, H. (2006). *Guide to Integrating Forensic Techniques into Incident Response.* Gaithersburg, MD 20899-8930 Computer Security Division, Information Technology Laboratory, National Institute of Standards and Technology. http://csrc.nist.gov/publications/nistpubs/800-86/SP800-86.pdf

Kohn, M., Eloff, JHP., Olivier, MS. (2007). *Framework for a Digital Forensic Investigation.* Department of Computer Science, Information and Computer Security Architectures Research Group (ICSA). Pretoria, University of Pretoria.

Leong, R., S.C (2006). "FORZA – Digital forensics investigation framework that incorporate legal issues." *Digital Investigation 3S: S29-36.*

Olivier, M., S. (2009). *On Metadata context in Database Forensics*. ICSA Research Group, Computer Science, Pretoria, South Africa, University of Pretoria.

Palmer, G. (2001). DFRWS, Report from the first digital forensic research workshop. A roadmap for digital forensic research. The MITRE Corporation. Report DTR-T001-01 http://dfrws.org

Petherick, W. (2002). "*Criminal profiling: How it got started and how it is used*." http://www.crimelibrary.com/criminology/criminalprofiling2

Radack, S. (2009) *Forensic Techniques: Helping Organisations Improve Their Responses To Information Security Incidents.* Computer Security Division, Information Technology Laboratory, National Institute of Standards and Technology. Retrieved on 6 February 2010 from http://csrc.nist.gov/publications/nistpubs/800-64-Rev2/SP800-64-Revision2.pdf

Rogers, M. (2003). *The role of criminal profiling in the computer forensics process*. Centre for Education and Research in Information Assurance and Security (CERIAS). Purdue University. www2.tech.purdue/cit/course/cit556/readings/profile-rogers.pdf

Ruibin, G., Gaertner, M. (2005). "Case-Relevance Information Investigation: Binding Computer Intelligence to Current Computer Forensic Framework." *International Journal of Digital Evidence 4(1).*

SANS (2010) Focus: Learning how to discover new artefacts using application forensics. Retrieved on 04 June 2010 from http://www.sans.org/selfstudy/description.php?cid=13822

Selamat, S., R., Yusof, R., Sahib, S. (2008). Mapping Process of Digital Forensic Investigation Framework. Faculty of Information Technology and Communication. Melaka, Malaysia, Universiti Teknikal Malaysia. http://paper.ijcsns.org/07_book/200810/20081025.pdf

Stallard, T., Levitt, K. (2003). *Automated Analysis for Digital Forensic Science: Semantic Integrity Checking*. Department of Computer Science. Davis, University of California. http://www.asac.org/2003/papers/89.pdf