Quantifying Relevance of Mobile Digital Evidence as They Relate to Case Types: A Survey and a Guide for Best Practice

Shahzad Saleem
Stockholm University, Borgarfjordsgatan

Ibrahim Baggili
Stockholm University, Borgarfjordsgatan

Oliver Popov
Stockholm University, Borgarfjordsgatan

QUANTIFYING RELEVANCE OF MOBILE DIGITAL EVIDENCE AS THEY RELATE TO CASE TYPES: A SURVEY AND A GUIDE FOR BEST PRACTICE

Shahzad Saleem[1], Ibrahim Baggili[2] and Oliver Popov[1]

[1] Department of Computer and Systems Sciences
Stockholm University, Borgarfjordsgatan 12
SE-16407 Kista, Sweden
{shahzads,popov}@dsv.su.se
[2] Tagliatela College of Engineering
UNHcFREG (University of New Haven Cyber Forensics Research and Education Group)
Department of Computer & Electrical Engineering and Computer Science
University of New Haven, 300 Boston Post Road
203-932-7198
West Haven, CT 06516, USA

ABSTRACT

In this work, a survey was conducted to help quantify the relevance of nineteen types of evidence (such as SMS) to seven types of digital investigations associated with mobile devices (MD) (such as child pornography). 97 % of the respondents agreed that every type of digital evidence has a different level of relevance to further or solve a particular investigation. From 55 serious participants, a dataset of 5,772 responses regarding the relevance of nineteen types of digital evidence for all the seven types of digital investigations was obtained. The results showed that (i) SMS belongs to the most relevant type of digital evidence for all the seven types of investigations; (ii) MMS belongs to the most relevant type of digital evidence for all the types of digital investigations except espionage and eavesdropping where it is the second most relevant type of digital evidence; (iii) Phonebook and Contacts is the most relevant type of digital evidence for all types of digital investigations except child pornography; (iv) Audio Calls is the most relevant type of digital evidence for all types of digital investigations except credit card fraud and child pornography; and (v) Standalone Files are the least relevant type of digital evidence for most of the digital investigations. The size of the response dataset was fairly reasonable to analyze and then delineate by generalization, relevance based best practices for mobile device forensics, which can supplement any forensics process model, including digital triage. For the reliability of these best practices, the impact of responses from the participants with more than five years of experience was analyzed by using one hundred and thirty three (133) instances of One-Way ANOVA tests. The results of this research can help investigators concentrate on the relevant types of digital evidence when investigating a specific case, consequently saving time and effort.

1. INTRODUCTION

This research work aimed at investigating the level of relevance of certain types of mobile digital evidence (such as SMS) related to specific types of cases (such as credit card fraud). The outcome of this research aids in the reduction of the size of data to be processed during an investigation, or speeding up the processing of certain types of digital evidence that might hold a higher priority in specific types of cases. The outcome of this research (guide for best practices) can help the investigator to fulfil the requirements of the "Extended Abstract Digital Forensics Model with 2PasU" (Saleem, Popov, & Bagilli, 2014). Most importantly, this research can also supplement and enhance the performance of any forensics process model including digital forensics triage.

Evidence is defined as "a matter of fact the design of which is to persuade the existence or non-existence of another matter of fact" (Routledge, 2004). Although this definition is broad, there are typically two main characteristics that are important when choosing evidence to be submitted to a court of law; namely relevance and weight. Relevance and weight are established in part by the scientific methods involved in reconstructing the evidence.

Computers and Mobile Devices (MDs) are important scientific inventions of the modern era. Due to the ubiquity of computers and MDs, our digital footprint has become immense, leaving behind a sizable amount of digital evidence. It is believed that around 80 to 90 percent of legal cases involve digital evidence (Baggili, Mislan, & Rogers, 2007; Rogers, 2004; Science Daily, 2009). In contemporary times, MDs are penetrating society faster than personal computers. The total number of mobile subscribers in 2013 is 6.8 Billion which is almost equal to our total population of 7.1 Billion (International Telecommunication Union (ITU), 2013). This indicates deep penetration and wide acceptance of MDs in our society. To that extent, we limit the scope of this initial work to MDs. However, we note that our approach is generic and can be extended to include evidence from other types of digital devices.

In its nature, digital evidence is a huge, messy, slippery, abstract and a transformed interpretation of reality (Casey, 2011; Palmer, 2001; Ryan & Shpantzer, 2002). Due to its nature, finding relevant digital evidence, from the image of a MD is a complex research problem. It can be solved if we know the types of evidence that have better chances of having leads to further or solve an investigation. These initial findings can then guide us to reveal the remaining relevant evidence needed to solve the case at hand.

This research is focused on how to find relevant types of digital evidence for a specific case type. To answer this question we (i) classified digital evidence found in MDs into five classes and a total of nineteen sub-types, (ii) identified seven different types of digital investigations, and (iii) conducted a survey to capture and tag the factor of relevance for the classified types of digital evidence in our identified investigation types.

This research has not only identified and tagged the factor of relevance to digital evidence, but its outcome also helped in creating a guide for best practices in mobile device forensics. This guide is a prioritized list of nineteen types of digital evidence based on their relevance for solving or furthering a particular type of case.

2. RELATED WORK

In the context of a proverb "justice delayed is justice denied", we attempted to improve the overall performance of the digital forensics process with an aim to reach proper conclusions in a timely manner, thus ensuring

that justice is served in time. The solution presented in this work is generic, with an ability to expand horizontally by supplementing existing forensic processes including digital forensics triage and vertically by including more types of digital evidence and investigations. Digital forensics triage is a relatively new and growing research area. Its aim is to reduce backlogs in digital forensics laboratories by focusing the resources where they are needed the most (Mislan, Casey, & Kessler, 2010; Pearson & Watson, 2010a).

Digital forensics triage has evolved to overcome the complexities associated with the nature and large amount of data. The word "triage" originates from the French language. Earlier, it was specifically related to the medical field, referring to "the sorting of the patients according to their condition, urgency to the need of treatment and allocating the resources accordingly to maximize their chances of survivability" (Encyclopedia Britannica, 2013; Oxford University Press, 2013). However, in today's world of digital forensics, triage is the process of redirecting the resources to the relevant stages of the forensics process model to maximize the performance of the overall process, ensuring that proper conclusions are reached in a timely manner (Baggili, Marrington, & Jafar, 2014).

Conducting a full scale digital forensics examination is a daunting activity and usually returns forty to fifty percent (40% to 50%) negative results while also producing backlogs in the forensics laboratory (ADF Solutions Incharge, 2013). Modern digital forensics triage tools can help reduce these backlogs by onsite examination of the suspects' digital devices while trying to employ automation and minimal investigator intervention. These tools can be configured to focus on specific types of relevant digital evidence with certain properties. The problem, however, is twofold:

The configuration process is based on the prioritization of digital evidence for a specific case which in turn depends on personal preferences and heuristics.

The tool then extracts and reports all the evidence based on the initial configuration. The investigator interprets the results to identify at least the first few leads in order to uncover as much relevant and weighty digital evidence as possible.

Our research tagged a generalized factor of relevance to each type of digital evidence based on the consensus of surveyed experts in the domain of digital forensics. Hence, it can solve both the aforementioned problems by equipping the investigator with pre-raid knowledge required to configure triage tools and allocate the resources to the most relevant digital evidence; consequently improving the overall throughput of the process.

Despite being criticized, highly automated digital forensics, or push button forensics, is gaining popularity. Forensic tools are constantly evolving with interfaces for automation. The push button forensics approach can facilitate investigations while increasing the performance of the overall investigative process (James & Gladyshev, 2013).

Fabio et al. presented related work (Marturana, Me, Berte, & Tacconi, 2011) to help improve the performance of digital forensics triage. They suggested using a machine learning technique to predict whether a mobile device under investigation was a part of a crime or not.

James and Gladyshev automated the exclusion of irrelevant exhibits from "Child Exploitation Material (CEM)" (James & Gladyshev, 2013). In their research, a live Linux distribution called "Computer Aided Investigative Environment (C.A.IN.E.)" was used to automatically process CEM. On average this enhanced preview processing technique took approximately twenty-four hours to process one exhibit. Analysts then used the output reports to decide whether to include or exclude a particular exhibit for further full scale analysis. The results were encouraging with some false positives and no false negatives. They showed that the

technique reduced the total number of exhibits to be processed in order to reach the correct conclusion. The objective of our research is also data reduction in the entire process of digital forensics while still being able to reach correct conclusions.

In the context of digital forensics triage, the research presented in this paper, is actually a continuation of the Rogers et al. work that is to first rank and then deal with the pieces of evidence with respect to their volatility or relevance in solving or furthering the case at hand, thus saving time (Rogers, Mislan, Goldman, Wedge, & Debrota, 2006). Therefore, rather than redirecting resources to different sub-processes, the research actually prioritized different digital evidence on the basis of their relevance with respect to the ongoing investigation.

Our work is in line with the definition of triage as implemented in the medical field where patients are sorted to maximize their chances of survival. During this research, digital evidence is sorted based on their relevance to maximize the chances of the extraction of relevant digital evidence in a shorter time span. The results can be used (1) to generate a guide for best practices for mobile device forensics based on prioritized lists of digital evidence found on MDs; and (2) as an extension of a digital forensics tool, thereby improving the tool's ability to process relevant digital evidence in a timely manner.

2.1 Evidence: Relevance and Weight

The quality of evidence relies upon its integrity, relevance and weight, which we define below:

Relevance: is the relationship between evidence and the fact under consideration. If the evidence increases or decreases the probability of the fact being proved or disapproved then it is relevant. Finding relevant evidence is the aim of every investigator.

Weight: is the extent or degree to which digital evidence is relevant to the case at hand. The form of generalization used to show the relevance of evidence will affect its weight, e.g. the bolder the generalization the heavier the weight of the evidence (Routledge, 2004).

In short, a weighty connection between evidence and a case has to be established in the court to help in its resolution. However, the duty of a digital forensics expert during mobile device forensics is a bit different from the party presenting the evidence in the court. The forensics expert has to find the evidence on which the art of presentation, logic and common sense can be applied to establish the required connection between the evidence and the case at hand.

2.2 Extended Abstract Digital Forensics Process Model

On the basis of actors and their responsibilities, we divided the process of digital forensics into two stages:

1. Before the evidence is presented in the court. We call it the out of court (OC) stage.

2. While the evidence is being presented in the court. We call it the in court (IC) stage.

We fine-tuned Reith's "abstract digital forensics process model" (Reith, Carr, & Gunsch, 2002) based on the above classification and our knowledge of abstract digital forensic process models (National Institute of Justice, 2001; Palmer, 2001). All of the sub-processes, except presentation, have an aim to eventually find relevant and weighty digital evidence while preserving integrity, whereas in presentation, the evidence is presented in the court of law in such a way that its relevance and weight is legally established to the case at hand with an aim to eventually further or solve the case. Therefore, we included presentation in the IC stage and the other sub-processes in the OC stage.

Moreover, actors and their objectives are different in both the OC and IC stages. Our research work targets and improves the performance of OC stage sub-processes (from identification to analysis), which in turn synergistically improves the IC stage sub-process.

The presentation sub-process is further subdivided into two more sub-processes called reporting and presentation. In total there are eleven sub-processes after our extension, namely: (1) identification, (2) preparation, (3) approach strategy, (4) preservation, (5) collection, (6) examination, (7) analysis, (8) reporting, (9) presentation, (10) archiving, and (11) returning evidence. Explanations to the extensions in Reith's model are presented in the following section.

Reporting: constitutes an expert's conclusion along with the relevant and weighty digital evidence. It deals with summarizing the findings and providing the explanations and conclusions along with the appropriate digital evidence. It is written for the layman using abstract terminologies with suitable references.

Presentation: is an art to infer from digital evidence using logic and common sense. The output of the reporting sub-phase is fed into this phase as an input. The aim of this phase is to present the findings in such a way that relevance and weight of the digital evidence is established to the case at hand. In the mind of the trier of the fact, it will create an argument for the existence or non-existence of some other matter of fact thus helping in solving or furthering a specific case. Moreover, any questions related to the integrity of digital evidence are answered during this phase.

Archiving: deals with strictly securing digital evidence along with its chain of custody for any potential future usage.

As shown in Figure 1, there is a large amount of data at the beginning of a forensic process. As we move through the OC and IC stages, depicted by rectangle ABCD, the data is reduced into digital evidence. Entities that are involved in this reduction process learn, through experience, that some types of digital evidence are more relevant in furthering and or solving specific case types. In this research, we fed that experience back into the initial stages of the digital forensics process with a survey. We also created, by generalization, SAO Best Practices (SAOBP) for Mobile Device Forensics. This can help investigators concentrate only data subset A´B´ of the set of data AB to reach relevant and weighty evidence CD, which is eventually presented in court. Depending on the investigative context and settings (civilian law enforcement or military), SAOBP can be used at any or all the phases of a forensics process. Due to increasing storage potential of MDs, backlogs are usually generated during the examination and analysis phases and SAOBP can possibly improve the situation as described in Section 5.
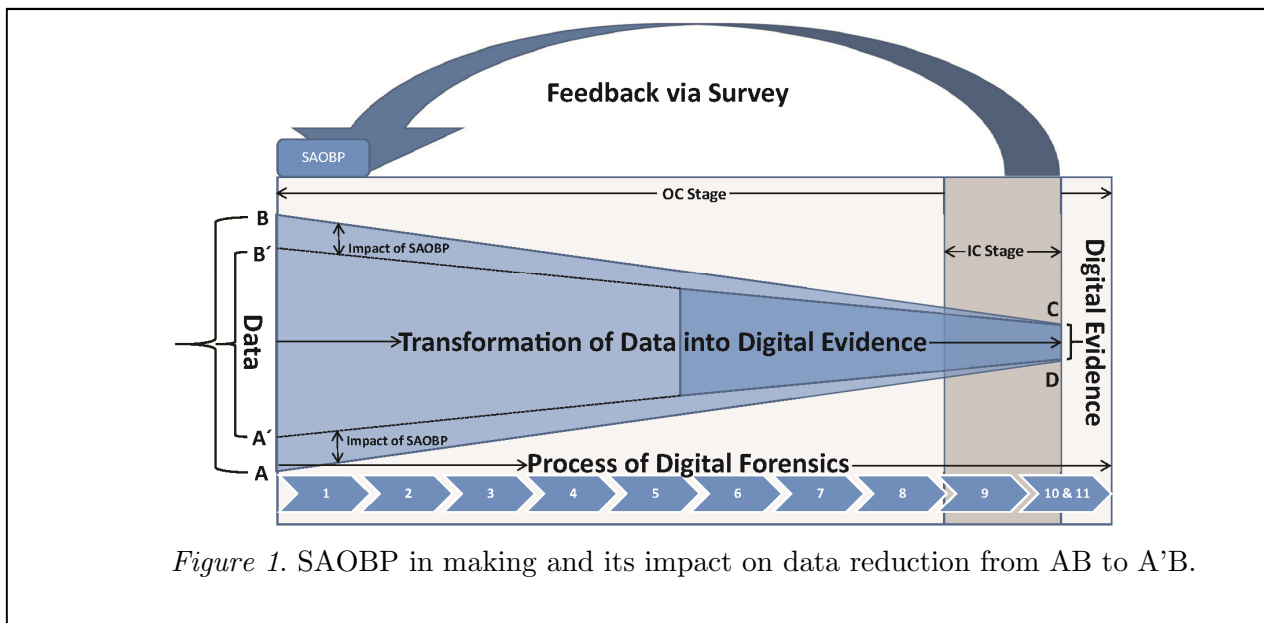
Figure 1. SAOBP in making and its impact on data reduction from AB to A'B.

,

3. SURVEY

We conducted a survey using a snowball sampling technique to benefit from the experience of experts (which is based on past cases or literature) and tagged the factor of relevance to all the types of digital evidence as they relate to case types. It is important to note that the research process followed during this work is general and applicable to every type of digital investigation, but we have limited ourselves to MDs. The limitation of this scope was due to two major reasons: i) MDs have become ubiquitous, and ii) the need to create a concise survey with a reasonable response rate.

3.1 Survey Design

A survey with generic and open-ended questions was not appropriate since we required a quantitative representation of the factor of relevance for digital evidence as they relate to specific case types. Moreover, questions with an abstract notion of digital evidence and digital investigation (investigations having some digital side associated with them) were not suitable to clearly tag the factor of relevance to every type of digital evidence in a specific case. Therefore, we identified and classified both digital evidence (Kubi, Saleem, & Popov, 2011) found

in MDs and digital investigations associated with MDs (Anobah, 2013). The classification of digital evidence associated with MDs is thoroughly discussed in previous publications (Kubi et al., 2011; Saleem, Popov, & Kubi, 2013), so we only discussed the classification of digital investigations in this article.

However, the types of evidence were still abstract and independent from the applications. For instance, SMS is SMS whether the native application or an installed application handles it, and an audio call is an audio call whether performed by a native application or any other user application. Based on discussions with the experts in industry and academia (Anobah, 2013), there appears to be seven major types of investigations associated with MDs namely:

Drug Trafficking: is a global black market of illegal drug trade. It includes cultivation, manufacturing, distribution, and sale of drugs prohibited by law. MDs play an important role in all the phases of this crime by providing the means to store, communicate, present and process relevant information.

Rape: is a sexual assault without the victim's consent. MDs play an important role in these types of cases by processing, storing and or communicating data associated with the crime.

Murder: is the act of unlawfully killing someone. Many digital artifacts from both the murderer and the victim can be associated with their MDs.

Credit Card Fraud: generally, the motive behind this kind of crime is to fraudulently acquire unauthorized funds from an account or to obtain goods either without payment or paying via the victim's account. MDs play important role in carrying out this type of criminal activity.

Harassment: is the term represents a range of offensive behaviors intended to disturb or upset someone. MD usage is common in this domain.

Espionage/Eavesdropping: Espionage or spying involves stealing or obtaining confidential or secret information without the permission or knowledge of its holder or owner. Eavesdropping is secretly listening to someone's private conversation without their consent. This type of crime has an associated digital side where the involvement of MDs is not out of question.

Child Pornography: is a kind of pornography portraying sexual activities involving a child. MDs play an important role in this crime.

We do acknowledge that there are other types of crimes strongly related to MDs such as human trafficking, however, at the time of the dissemination of the survey, only the abovementioned types of cases were used.

The first page of the survey outlined the objectives of the research activity and clearly described what one should expect during his/her participation. It also captured a mandatory and explicit consent of participation from each participant. The participants were clearly briefed that they can skip any or all the questions at any time during the survey.

Next, we briefly explained the concepts associated with this survey including the different types of digital evidence and the scale being used for tagging the factor of relevance.

We then captured each participant's demographic information, after which every participant was asked seven questions to tag the factor of relevance to all the types of mobile digital evidence for each of the seven types of digital investigations. To eliminate any bias, the participants were requested to answer each of the questions related to relevance in full (if possible). To complete the survey, an open ended question was also asked at the end, in order to capture any suggestions and comments to help our future work.

3.1.1 Scale

A ratio scale of eleven discrete points starting from zero to ten was used. It had the following properties:

The minimum value was zero (0) and the maximum was ten (10), where zero means irrelevant and ten means most relevant.

The difference between two points was constant, i.e. if one type of digital evidence (X) received 4 points and another type of digital evidence (Y) received 8 points then "Y" is twice as relevant as "X".

3.1.2 Sample Size

We used Equation 1 to calculate a reasonable sample size needed for our survey (iSixSigma, 2013).

$$n = \left(\frac{Z_{\frac{\alpha}{2}} \times \sigma}{E} \right)^2 \qquad \textbf{Equation 1}$$

Where:

$Z_{\frac{\alpha}{2}}$ = Z-Score of a normal distribution. It is 1.96 for a normal distribution with 95% confidence interval.

σ = Standard deviation. Population standard deviation was unknown, so we used the initial survey results to calculate the estimated value. Its value was equal to 2.82.

E = Standard error. In our study we were comfortable with 0.9 to 1.0 unit difference from the population mean μ.

n = Sample Size. It is equal to 47 if we use 0.9 unit difference from the population mean and 30 if we use 1. In reality, we obtained more than 50 participants for almost every question for our survey.

3.1.3 Data Cleansing

The basic assumption behind the survey was that every type of digital evidence had a different level of relevance in solving or furthering a specific type of investigation. The aim was to find that level of relevance for all the types of evidence in every type of investigation.

According to the classification, we had nineteen types of digital evidence and seven types of investigations. So, in total, we had to analyze one hundred and thirty three (133) cases. Most of the measurements of kurtoses and skewness were not significantly high and the distributions were discrete, therefore, D'Agostino Kurtosis, D'Agostino Skewness and Jarque-Bera goodness of fit tests were employed to measure the departure from normality. The tests indicated that at α = 0.01 we did not have enough evidence to reject the null hypothesis (data follows a normal distribution) for the majority of the cases i.e., eighty six (86), seventy eight (78) and eighty seven (87) cases out of one hundred and thirty three (133) with D'Agostino Kurtosis, D'Agostino Skewness and Jarque-Bera methods, respectively. So, for the entire set of data with 5772 responses, we concluded that an approximately normal distribution was being followed.

Furthermore, to gain a robust level of relevance, the data had to be cleaned. Data cleansing was conducted using the following two strategies:

1. Removing incomplete responses: Removing responses from participants who had incomplete responses regarding relevance was the strategy to identify and keep the responses from serious participants.

2. Eliminating the outliers: The respondents were bound to choose a rating between

zero and ten when determining the level of relevance. Moreover, the distance between any two adjacent levels of relevance was constant. Hence, respondents could not specify a drastically smaller or larger level of relevance.

Therefore, outliers in our case were identified by:

a. The value of the level of relevance

b. The size of the group of respondents who voted for a specific level of relevance

This survey used the concepts of a voting system to tag the appropriate level of relevance to each type of digital evidence with the help of consensus among the respondents. Therefore, majority groups were given preference with respect to the total number of their members.

Firstly, frequency (f) of votes for each level of relevance was calculated. Using these frequencies, the proportion of participants voting for each level of relevance was computed using Equation 2. Then the data was arranged in ascending order of proportion. After that, starting from the bottom, all the groups of voters were selected and their proportions were added together until $\sum p_i < 0.9$, ensuring the representation of at least ninety percent (90%) of the majority vote. In doing so, we eliminated all of the groups which did not have a sufficient number of members compared to the rest of the groups and were thus treated as outliers. This cleaned data was used for subsequent analysis. Additionally, central tendency was computed by using weighted averages for this set of cleaned data.

$$p_i = \frac{f_i}{\sum_{i=0}^{10} f_i} \qquad \textbf{Equation 2}$$

Where:
$\sum p_i = 1$ for i = 0, 1, 2, 3 … 10

Initially, we acquired a set of raw data with one hundred and twenty seven (n=127) respondents before the execution of data

cleansing techniques. After data cleansing following strategy 1, a set of fifty-five (n=55) respondents was obtained. The data was subsequently cleaned using strategy 2 and then used for data analysis.

The approach (Approach A) used in this research is not affected by the distance from the mean (Narasimhan, 1996). It only considers the groups with more members. Standard Deviation (SD) around the mean approach (Approach B) takes both the value of data points and their distance from the mean into account. According to the three sigma rule, 2 SD around the mean is a poorly conservative research criteria (Leys, Ley, Klein, Bernard, & Licata, 2013) and covers ninety five (95%) of the area under the distribution curve (Narasimhan, 1996). Reducing the sample by removing the outliers to a level which covers at least 90% of the area under the distribution curve was used (approximately 1.645 SD).

Figure 2. Histogram while removing outliers.
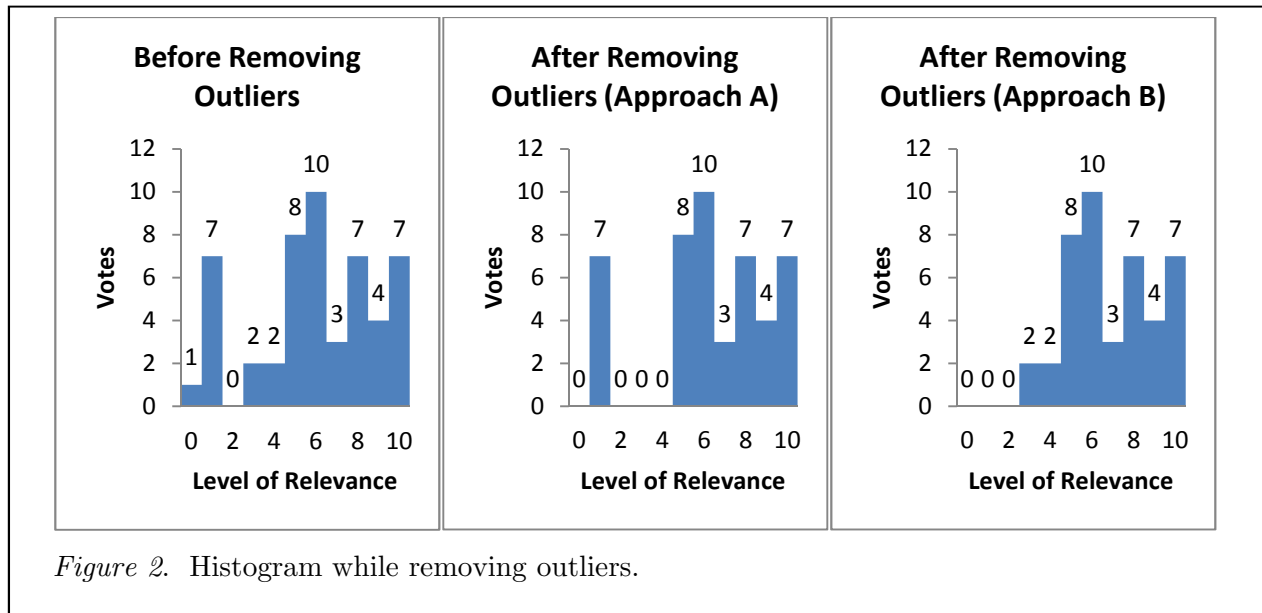
Histograms in Figure 2 indicate the difference between the two data cleansing approaches when used for the "Calendar Entry" type of digital evidence in a "Drug Trafficking" investigation. Relevance levels (RL) 0, 3 and 4 are considered outliers by Approach A, while RL 0 and 1 are considered outliers by Approach B. Continuing with Approach B, it can be seen that even though RL1 is tied at third with 7 votes, it is deemed an outlier (because of its distance from the mean) and thus, by using weighted average, Approach B produces an RL of 6.91 as an overall level of relevance for Calendar Entries in a drug trafficking investigation. On the other hand, Approach A produces an RL of 7.3 for the same type of digital evidence, which is more appropriate.

4. RESULTS AND DISCUSSION

Clarification, demographics, level of relevance of digital evidence and digital investigation and tagging the level of relevance to digital evidence were the four main parts of our survey. We discuss the results of each part of our survey in this section.

4.1 Clarification

This was the first page in our survey. Results indicated that a total of one hundred and ninety eight (198) individuals accessed the survey, out of them one hundred and twenty seven (127) gave their consent of participation and seventy seven (77) dropped out at this stage.

4.2 Demographics

This section of our survey was devoted to capture demographic details. The participants were from the continents of North America, Europe, Africa, Asia and Australia. Table 1 shows the summary of the demographic details.

1. Gender: Fifty five (55) serious participants answered this question. Fifty three (53) out of fifty five (55) were male (96.4 %) and only 2 were female (3.6 %).

2. Age: Most of the participants belonged to the age groups of 41-45 and above 50 years with a 25.5% and 23.6% respective representation. Age groups of 31-35, 46-50 and 36-40 had almost equal representation that is 16.4%, 16.4% and 14.5% respectively. This even distribution of age groups allowed us to obtain results from a mixed and random set of respondents.

3. Areas of Interest: On the basis of our literature review (Brinson, Robinson, & Rogers, 2006; Garfinkel, 2010; National Security Database, 2013; Palmer, 2001; Yadav, 2011), the areas of interest were classified into four categories, namely: (1) computer forensics, (2) network forensics, (3) small scale digital forensics and (4) forensics data analysis. The participants were able to choose any or all of them depending on their areas of interest. Once again we saw a healthy distribution of areas of interests among the participants

4. Experience: The experience of the majority of the respondents fell into the groups of 6-10, 3-5 and 0-2 years respectively. A total of fifty four (54) respondents answered this question and four (4) of them had more than fifteen

(15) years of experience. So, in terms of experience, we again received a wide distribution of participants in our pool. It ensured that the opinion of the experts at both the advanced and entry levels had an appropriate representation.

Table 1
Demographics

Demographics			
		Response Total	Response Percent
Gender	Male	53	96.4 %
	Female	2	3.6 %
	Total Respondents	55	
Age	Below 18	0	0 %
	18 - 25	0	0 %
	26 - 30	2	3.6 %
	31 - 35	9	16.4 %
	36 - 40	8	14.5 %
	41 - 45	14	25.5 %
	46 - 50	9	16.4 %
	Above 50	13	23.6 %
	Total Respondents	55	
Areas of Interest	Computer Forensics	52	94.5 %
	SSDD Forensics	30	54.5 %
	Network Forensics	16	29.1 %
	Forensic Data Analysis	35	63.6 %
	Others	6	10.9 %
	Total Respondents	55	
Experience (Years)	0 - 2	5	9.3 %
	3 - 5	16	29.6 %
	6 - 10	25	46.3 %
	11 – 15	4	7.4 %
	15 +	4	7.4 %
	Total Respondents	54	

A total of fifty five (55) respondents answered this question. Most of them, fifty two (52) participants or approximately ninety four percent (94.5%), were related to computer forensics. It was expected since computer forensics is the oldest branch of digital forensics. The number of participants related to small scale digital forensics and

forensics data analysis were equal to thirty (30) and thirty five (35), respectively.

4.3 Level of Relevance of Digital Evidence and Digital Investigations

This was a Yes/No question. We assumed that every type of digital evidence has a different level of relevance in solving or furthering a specific case type. This question captured the experts' view on this matter, and in the set of cleaned data, all fifty-five (55) or 100% of the participants confirmed our assumption.

However, if we discuss the set of raw data with incomplete answers and outliers included, then two (2) out of seventy seven (77) participants were identified with dissenting opinion. Since dissenting opinions are essential for initiating a scientific discourse, an additional attention was paid to these two participants.

Our analysis revealed that one of them dropped out of the survey after answering this question. She belonged to the age group of 36-40 years, had 6-10 years of experience in forensics data analysis, and used trusted third party evaluation results to select the appropriate tool for her forensic activities. She did not believe that every type of digital evidence has a different level of relevance in solving or furthering a specific case type.

The second participant with dissenting opinion was above 50 years of age. He had 6-10 years of forensics experience with all types of digital devices. He also dropped out of the survey after answering this question. The reason he cited was that he was unable to generalize because every case has unique aspects during an investigation.

He is partly right, but generalization is possible as confirmed by the consensus of 97% of the participants. Every type of a crime has its unique characteristics as well as some common ones. Based on these commonalities, we classified them into seven types, which also show that generalization in terms of the

level of relevance of digital evidence for each type of investigation is possible. This phenomenon eventually allows, in a general sense, to tag the factor of relevance to each type of digital evidence for a particular case. It can aid the investigator in finding the first few leads which, in turn, may speed up the process of revealing the remaining relevant digital evidence covering both the specific and general aspects of a particular crime.

4.4 Tagging the Level of Relevance to a Digital Evidence

There were a total of five thousand seven hundred and seventy two (5772) responses from fifty five (55) serious respondents. All the types of digital evidence were rearranged in descending order of their weighted average relevance for all types of investigations. Digital evidence were then grouped into five levels in such a way that the entire range of relevance levels is divided into five equal steps using Equation 3.

$$step = (RL_1 - RL_6)/5 \qquad \textbf{Equation 3}$$

Whereas:

$RL_1 =$ The maximum level of relevance acquired by any type of digital evidence for a specific type of investigation.

$RL_6 =$ The minimum level of relevance acquired by any type of digital evidence for a specific type of investigation.

By using the step size from Equation 3, five different levels of relevance in descending order were obtained by using Equation 4 and Equation 5.

$$RL_{i+1} = RL_i - step \qquad \textbf{Equation 4}$$

$$level_i = RL_i - RL_{i+1} \qquad \textbf{Equation 5}$$

Where:

$$i = 1, 2, 3, 4, 5$$

The group of digital evidence at level one (L1) comprises the range of relevance which is greater than 80% of the entire range. Similarly L2 consists of 60% to 80%, L3

consists of 40% to 60%, L4 consists of 20% to 40% and finally L5 is the range of relevance which is less than 20% of the entire range. For simplicity, we assumed and categorized the evidence as following:

1. L1 are called Grade "A" evidence.

2. L2 are called Grade "B" evidence.

3. L3 are called Grade "C" evidence.

4. L4 are called Grade "D" evidence.

5. L5 are called Grade "E" evidence.

The tables containing detailed information are presented in the Appendix. The following section discusses the first three most relevant levels of digital evidence (Grades A, B and C) based on the summary graph for all seven types of digital investigations.

1. **Relevance of Digital Evidence for Drug Trafficking**: SMS, Phonebook/Contacts, Audio Calls, MMS, Graphics/Pictures and Email Entries are considered Grade "A" digital evidence for investigating drug trafficking and range from an RL of 9.68 to an RL of 8.65 respectively. So, maximum attention should be given to these six types of digital evidence to at least find the first few leads. EMS, with an RL of 7.62, is a Grade "B" digital evidence, while Grade "C" digital evidence comprises of Video Files, Video Calls, Memo and Notes, Calendar Entries, URLs Visited and Tasks to Do Lists which range from an RL of 7.04 to an RL of 5.83 respectively (Figure 3).

Word Files, PDF and PowerPoint Files are Grade "E", the least relevant group, of digital evidence for a drug trafficking investigations and range from an RL of 4.35 to an RL of 3.11 respectively. PowerPoint files, with an RL of 3.11, is the least relevant type of digital evidence for a drug traffic investigation, while SMS, with an RL of 9.68, is the most relevant type of digital evidence.
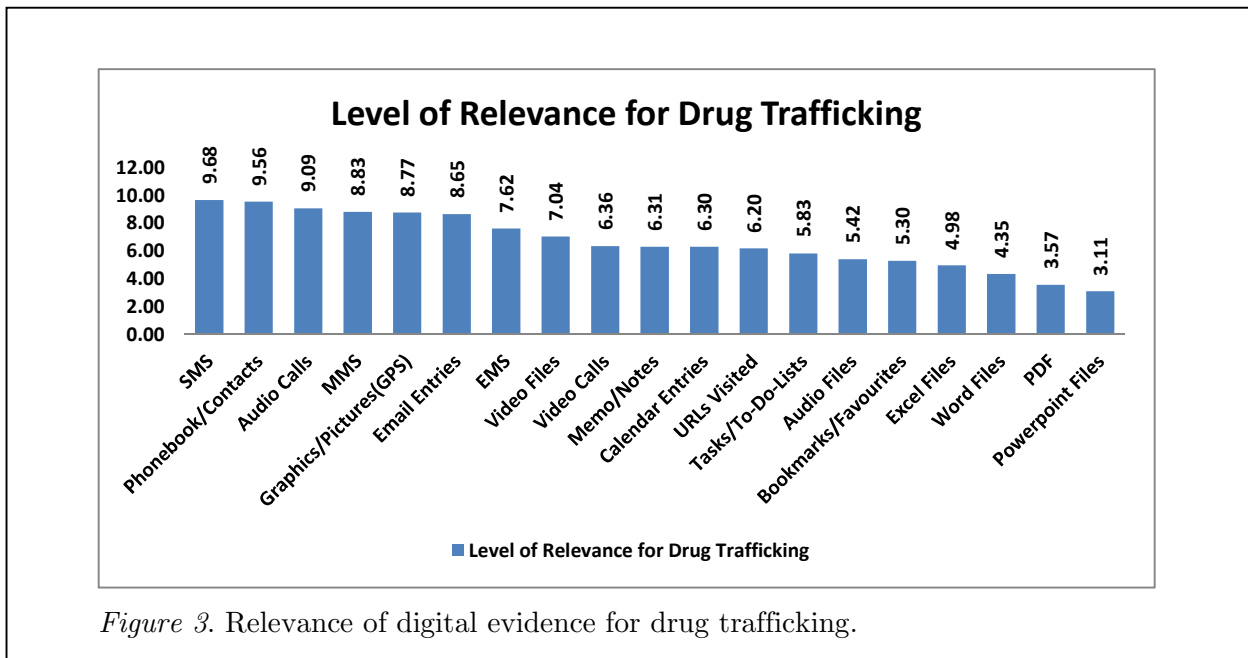
Figure 3. Relevance of digital evidence for drug trafficking.

2. **Relevance of Digital Evidence for Rape:** Figure 4 depicts that SMS, Graphics/Pictures (GPS), Phonebook/Contacts, MMS and Audio Files are Grade "A" types of digital evidence with a range of RL 9.33 to RL 8.77 respectively. They have the highest potential to further or solve investigations related to a rape investigation. Video Files, EMS, Email Entries and Video Calls are deemed Grade "B" digital evidence that range from an RL of 7.65 to an RL of 6.84. Lastly, Calendar Entries, Audio Files and URLs Visited are Grade "C" digital evidence spanning from an RL of 6.13 to an RL of 5.47.

SMS, with an RL of 9.33 is the most relevant type of digital evidence while PowerPoint Files, with an RL of 2.27 is the least relevant type of digital evidence in solving or furthering a rape investigation.
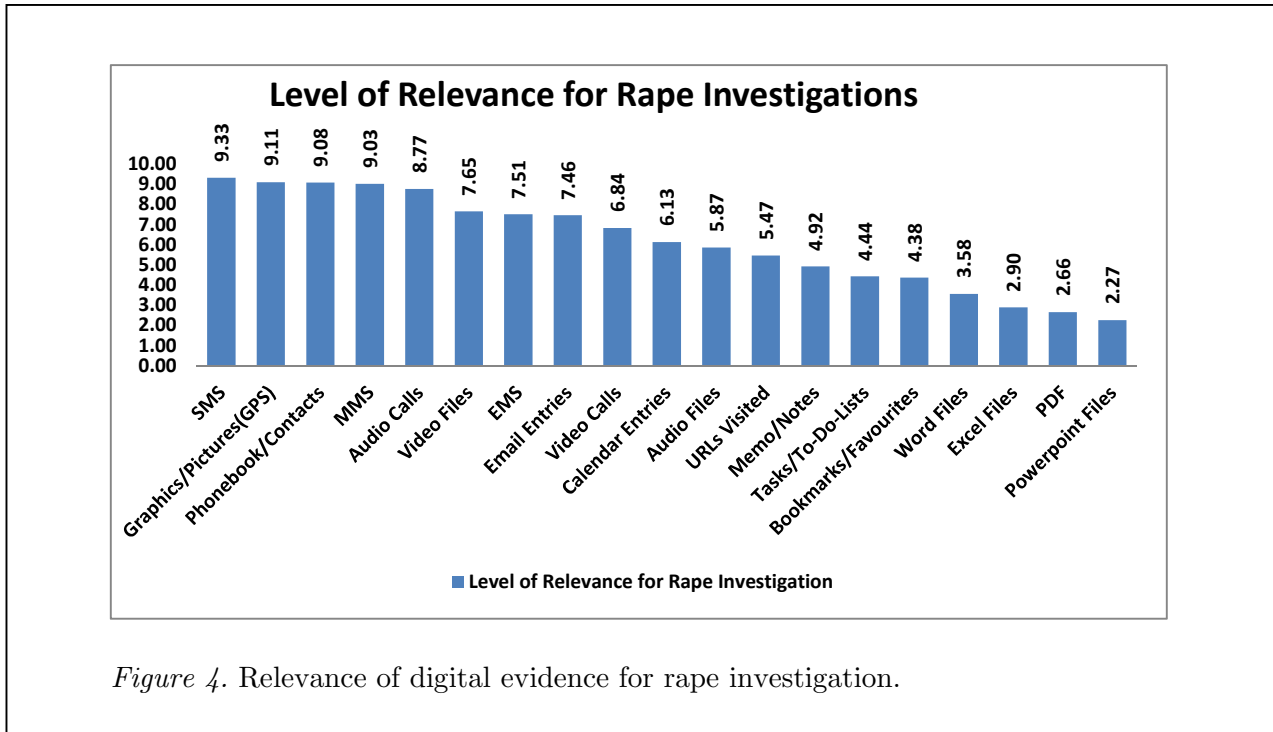
Figure 4. Relevance of digital evidence for rape investigation.

3. **Relevance of Digital Evidence for Murder:** Figure 5 shows that SMS, Phonebook/Contacts, Audio Calls, MMS and Email Entries are Grade "A" types of digital evidence for murder investigations. Maximum priority should be given to these types of digital evidence to help unearth, understand and extract clues in order to solve or further a murder investigation. Grade "B" is awarded to Graphics/Pictures (GPS), Calendar Entries, EMS and Memo/Notes, while URLs Visited, Video Files, Tasks/To-Do-Lists and Video Calls constitute Grade "C" digital evidence.

SMS, with an RL of 9.68, is the most relevant while PowerPoint Files, with an RL of 4.35, is the least relevant type of digital evidence to solve or further a murder investigation.
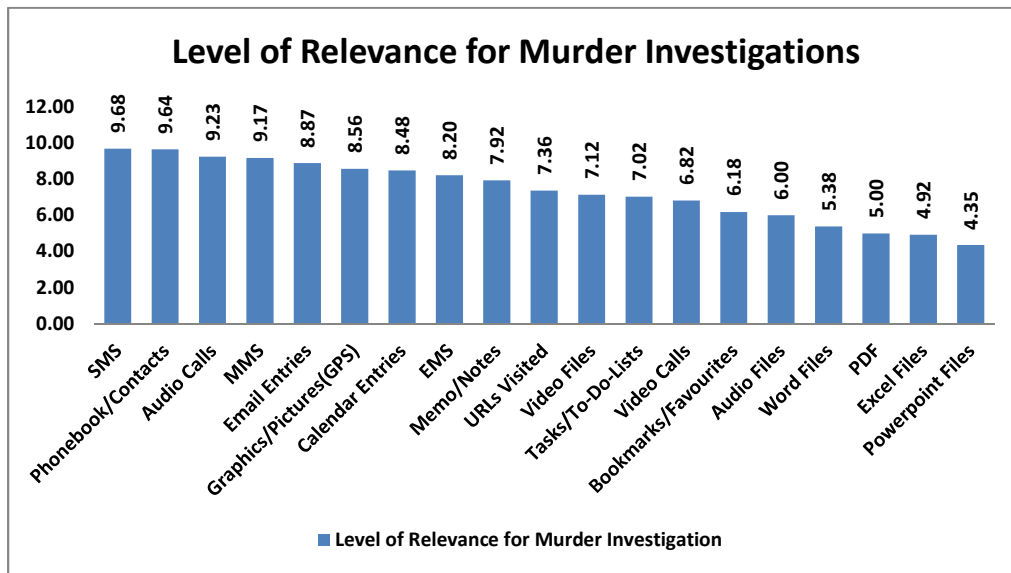
Figure 5. Relevance of digital evidence for murder investigation.

4. **Relevance of Digital Evidence for Credit Card Frauds:** Figure 6 shows that SMS, Email Entries, Phonebook/Contacts, URLs Visited and MMS are Grade "A" types of digital evidence for credit card fraud investigations. The highest priority should be given to Grade "A" digital evidence to solve or further an investigation of this type. Grade "B" evidence constitute Audio Calls, Bookmarks/Favourites, Excel Files and Graphics/Pictures (GPS) while Word Files, EMS, Memo/Notes, Calendar Entries, Tasks/To-Do-Lists are considered Grade "C" digital evidence.

SMS, with an RL of 8.84, is the most relevant while PowerPoint Files, with an RL of 5.05, is the least relevant type of digital evidence to solve or further credit card fraud investigations.
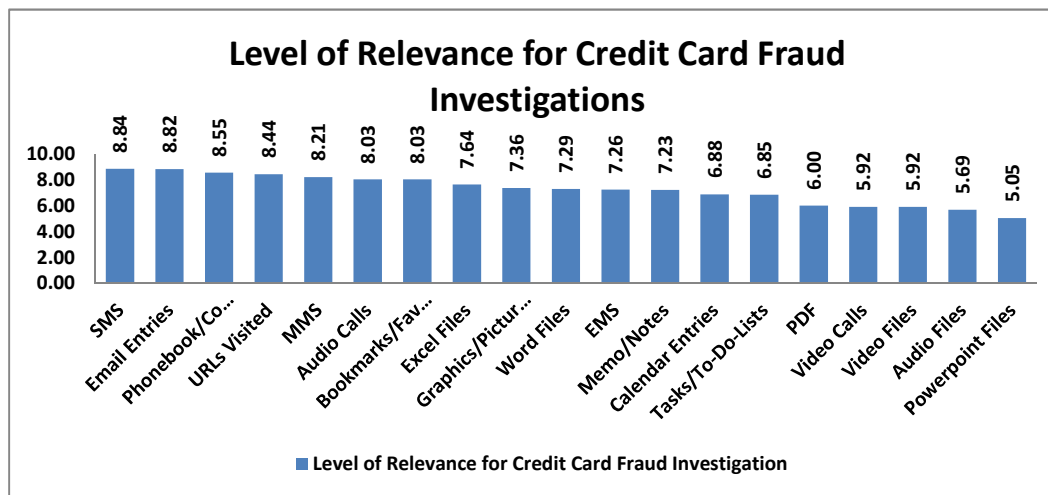
Figure 6. Relevance of digital evidence for credit card fraud investigation.

5. **Relevance of Digital Evidence for Harassments**: Figure 7 shows that SMS, MMS, Phonebook/Contacts, Audio Calls, Email Entries, Graphics/Pictures (GPS) and EMS are considered Grade "A" digital evidence when solving or furthering an investigation of harassment. Video Files, Video Calls, Audio Files and Calendar Entries are classified as Grade "B" digital evidence. Lastly, Memo/Notes, URLs Visited, Tasks/To-Do-Lists and Bookmarks/Favourites are Grade "C" types of digital evidence. SMS (RL of 9.84) is the most relevant while Excel Files (RL of 3.0) is the least relevant type of digital evidence in furthering or solving a case involving harassment.

Level of Relevance for Harassment Investigations

Values (left to right): SMS 9.84, MMS 9.59, Phonebook/Contacts 9.51, Audio Calls 9.50, Email Entries 9.38, Graphics/Pictures(GPS) 9.00, EMS 8.59, Video Files 8.00, Video Calls 7.97, Audio Files 7.41, Calendar Entries 7.11, Memo/Notes 6.85, URLs Visited 6.84, Tasks/To-Do-Lists 6.41, Bookmarks/Favourites 6.11, Word Files 5.11, Powerpoint Files 3.45, PDF 3.21, Excel Files 3.00

■ Level of Relevance for Harassment Investigation

Figure 7. Relevance of digital evidence for harassment investigation.

6. **Relevance of Digital Evidence for Espionage/Eavesdropping:** For espionage/eavesdropping investigations, Audio Calls, Phonebook/Contacts, SMS and Email Entries are deemed Grade "A" types of digital evidence spanning from an RL of 9.37 to an RL of 9.13. Graphics/Pictures (GPS), Audio Files, MMS and Video Files are Grade "B" digital evidence while URLs Visited is Grade "C" digital evidence with an RL of 8.39.

Audio Calls, with an RL of 9.37, is the most relevant type of digital evidence while PowerPoint Files, with an RL of 7.11, is the least relevant type of digital evidence in furthering or solving espionage/eavesdropping investigations.

Following are some of the unique properties associated with the level of relevance of digital evidence to further or solve an investigation related to espionage/eavesdropping (Table 2):

a. The range, meaning, the difference between the most and the least relevant types of digital evidence is very small. It is only 2.263 points.

b. The difference between Grade "A" and Grade "E", or the intergrade difference, is very small at only 1.576 points.

Table 2
Range and Difference between Grades A and E

	Range = (RL_{max} - RL_{min})	Grade A - Grade E
Espionage/Eavesdropping	2.263	1.576
Credit Card Fraud	3.789	2.518
Child Pornography	4.851	3.384
Murder	5.325	3.491
Drug Trafficking	6.574	4.298
Harassment	6.838	5.145
Rape	7.055	5.194

The range and intergrade differences for espionage and eavesdropping investigations are the smallest when compared to the other types of investigations. In terms of relevance, all nineteen types of digital evidence are not far apart. Thus, even with the five groups of evidence (from Grades A to E); almost equal importance must be given to all of the types of digital evidence while investigating this type of a case.
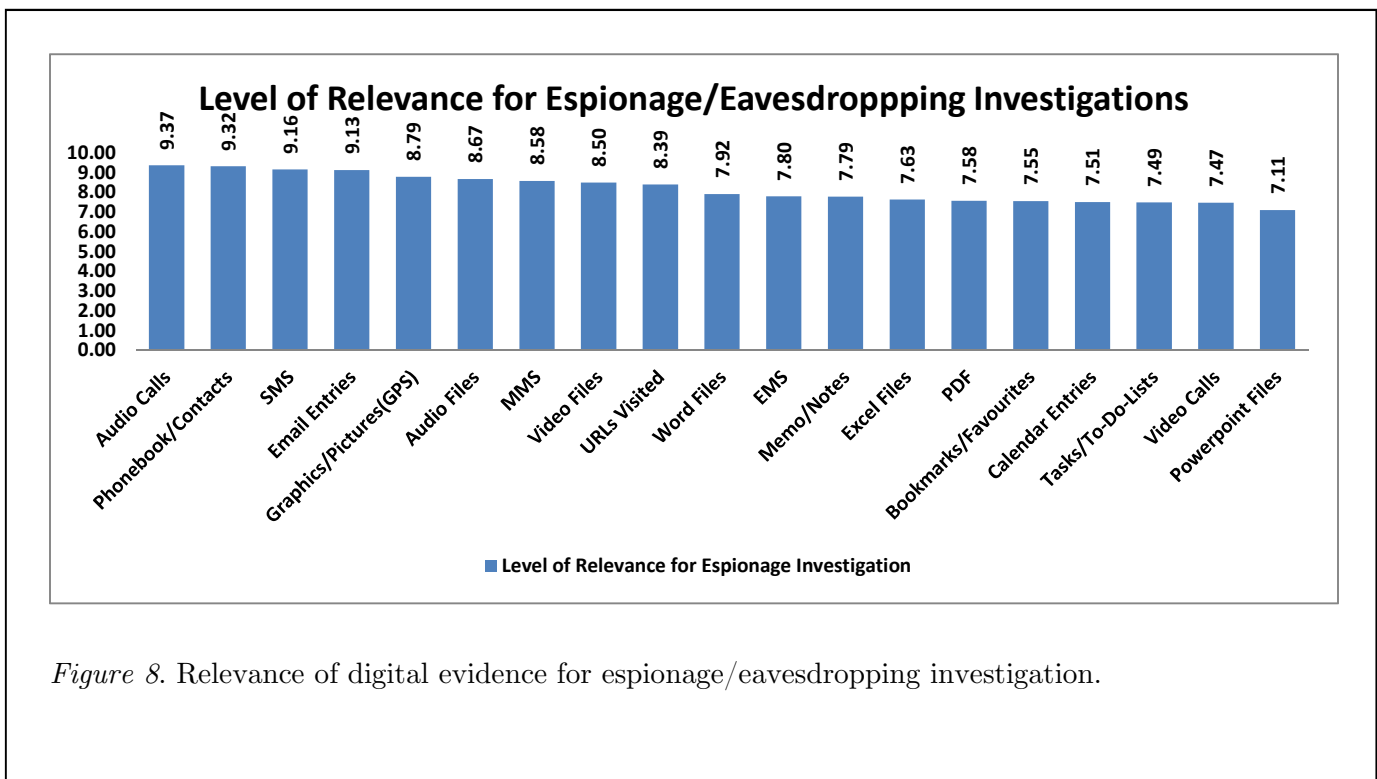
Figure 8. Relevance of digital evidence for espionage/eavesdropping investigation.

7. **Relevance of Digital Evidence for Child Pornography:** Graphics/Pictures (GPS), Video Files, URLs Visited, Bookmarks/Favourites, Email Entries, SMS and MMS are Grade "A" types of digital evidence. This is the most relevant group, so due attention must be given to them in order to further or solve the case of child pornography. The Grade "B" group of digital evidence consists of Phonebook/Contacts, EMS and Audio Calls, while Video Calls are in Grade "C".

Graphics/Pictures (GPS), as expected, is the most relevant type of digital evidence with an RL of 9.83 for investigating child pornography. On the other hand, PDF files, with an RL of 4.97, are the least relevant type of evidence in this type of investigations.
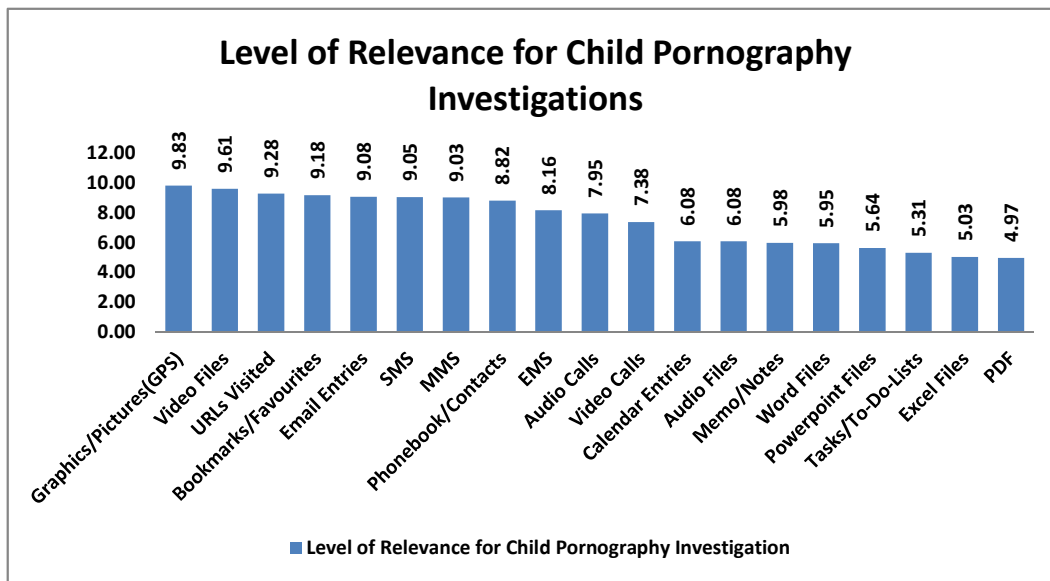
Figure 9. Relevance of digital evidence for child pornography investigation.

5. SUMMARY AND SAO BEST PRACTICES FOR MOBILE DEVICE FORENSICS

Following are some of the important points to be noted when Grade "A" digital evidence from all of the types of investigations are summarized and analyzed.

1. SMS is Grade "A" digital evidence in all the types of investigations. It stands at the top of the Grade "A" digital evidence for five out of the seven types of investigations. It stands at third position in the group of Grade "A" digital evidence for the investigation of Espionage/Eavesdropping and at sixth position for the investigation of Child Pornography.

2. Phonebook/Contacts are considered Grade "A" digital evidence for six out of the seven types of investigations. Child pornography is the only type of investigation where Phonebook/Contacts are included in Grade "B" digital evidence. However, in this case, it stands at the top of the group of Grade "B" digital evidence.

3. Audio Calls is Grade "A" digital evidence for five out of the seven types of investigations. Credit card fraud and child pornography are the two types of investigations in which audio calls are categorized as Grade "B" digital evidence.

4. MMS is Grade "A" digital evidence for six out of the seven types of investigations. Espionage/Eavesdropping is the only type of investigation in which it is labeled Grade "B".

5. Graphics/Pictures (GPS) is deemed Grade "A" digital evidence for the investigations of drug trafficking, rape, harassment and

child pornography. Therefore, it is one of the most relevant types of digital evidence for four out of seven types of investigations.

6. Email Entries are Grade "A" digital evidence for six out of the seven types of investigations. Rape is the only type of investigation in which it is considered Grade "B" digital evidence.

7. URLs Visited is the least common type of digital evidence categorized as Grade "A" digital evidence. It is included as Grade "A" digital evidence in just two types of investigations. They are credit card fraud and child pornography.

To see the cumulative importance of each type of evidence across all the types of investigations, a Borda count (Bowen, 2001) was computed. It ranked all the Grade "A" type of digital evidence for all the types of investigations. The summary of the results is presented in Figure 10.

Across all the types of investigations, SMS and Phonebook/Contacts are the most important types of digital evidence, followed by Audio Calls, MMS and Email Entries as the second most important group of digital evidence. Graphics/Pictures are the third most important type of digital evidence and are not very far (in terms of relevance) from the second group. URLs visited are the fourth most important type of digital evidence in this regard. Similarly this concept can be extended to all the other groups from Grade "B" to Grade "E" across all the types of investigations.

Preference and resource allocation should follow the grading system presented in this research. Grade "A" digital evidence should be given the highest level of preference and resources while Grade "E" digital evidence should be allocated the least amount of resources and preference. This will ensure proper resource allocation and thus a better

chance to further or solve the investigation in a shorter time span.

All of the research activity presented from Figure 3 to Figure 9 is summed up in Table 3 which transforms into SAO best practices for mobile device forensics with respect to the relevance of digital evidence.
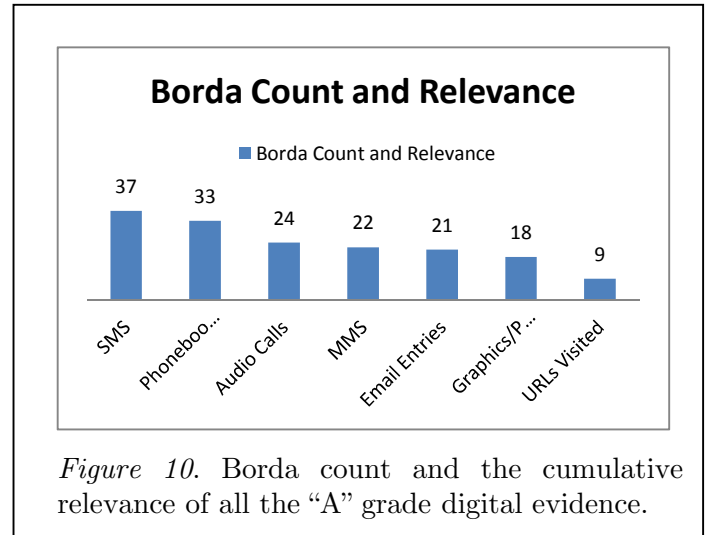
Figure 10. Borda count and the cumulative relevance of all the "A" grade digital evidence.

5.1 SAO Best Practices for Mobile Device Forensics

Table 3 contains SAO best practices for mobile device forensics based on the relevance of digital evidence for every type of investigation. These are the prioritized lists based on descending levels of relevance of digital evidence to solve or further an investigation. Forensics experts should follow these lists to allocate their resources for investigating a specific case.

Some of the important findings from Table 3 are:

1. SMS is Grade "A" digital evidence for all of the types of investigations.

2. MMS is Grade "A" digital evidence for all of investigations except EE, where it belongs to Grade "B" digital evidence.

3. Similarly PBC is Grade "A" digital evidence for all of the types of investigations except CP, where it is Grade "B" digital evidence.

4. ADC is Grade "A" digital evidence for five out of the seven types of investigations. It is Grade "B" digital evidence for investigating CCF, CP.

Standalone files mostly share the least relevant level for all of the types of investigations. Hence, they are Grade "E" digital evidence for most investigation types.

Table 3
SAO Best Practices for Mobile Device Forensics

	DT	RP	MRD	CCF	HMT	EE	CP
Grade A	SMS, PBC, ADC MMS, GP, EML	SMS, GP, PBC MMS, ADC	SMS, PBC, ADC MMS, EML	SMS, EML, PBC URLV, MMS	SMS, MMS, PBC ADC, EML, GP EMS	ADC, PBC, SMS EML	GP, VDF, URLV BKF, EML, SMS, MMS
Grade B	EMS	VDF, EMS, EML VDC	GP, CDR, EMS MN	ADC, BKF, EXL GP	VDF, VDC, ADF CDR	GP, ADF, MMS VDF	PBC, EMS, ADC
Grade C	VDF, VDC, MN CDR, URLV, TTD	CDR, ADF, URLV	URLV, VDF, TTD, VDC	DOC, EMS, MN CDR, TTD	MN, URLV, TTD BKF	URLV	VDC
Grade D	ADF, BKF, EXL	MN, TTD, BKF	BKF, ADF	PDF, VDC, VDF	DOC	DOC, EMS, MN EXL, PDF	CDR, ADF, MN DOC
Grade E	DOC, PDF, PPT	DOC, EXL, PDF PPT	DOC, PDF, EXL PPT	ADF, PPT	PPT, PDF, EXL	BKF, CDR, TTD VDC, PPT	PPT, TTD, EXL PDF

Table 4
SAO Best Practices for Mobile Device Forensics (5+ Years of Experience)

	DT	RP	MRD	CCF	HMT	EE	CP
Grade A	SMS, PBC, ADC, EML	PBC, SMS, ADC, GP, MMS	SMS, PBC, EML, MMS, ADC	EML, SMS, PBC, URLV, ADC	SMS, ADC, PBC, MMS, EML, EMS, GP	ADC, PBC, SMS, GP	GP, URLV, VDF, BKF, SMS, MMS, EML
Grade B	GP, MMS, EMS	EML, EMS, VDC, VDF, CRD	CRD, EMS, MN, GP, URLV	BKF, MMS, EXL, EMS	VDC, VDF, CRD, URLV	EML, VDF, MMS	PBC, EMS, ADC
Grade C	VDF, VDC, CRD, MN	ADF, MN, URLV, TTD	TTD, VDC, VDF, BKF	CRD, MN, DOC, GP, TTD	BKF, ADF, MN, TTD	ADF, EMS, DOC, URLV	VDC
Grade D	TTD, URLV, BKF, EXL, ADF, DOC	BKF	ADF, DOC, EXL	PDF, VDC	DOC	PDF, EXL	CDR, ADF, MN DOC
Grade E	PDF, PPT	DOC, EXL, PDF PPT	PDF, PPT	VDF, PPT, ADF	PPT, EXL, PDF	MN, BKF, PPT, CRD, TTD, VDC	PPT, TTD, PDF, EXL

Comparative analysis of SAO best practices (SAOBP) for mobile device forensics was also conducted to see if the experience has an effect on the results. As evident from

*Table 3*Table 3 SAOBP is an ordered prioritized list of digital evidence grouped into different grades depending on their level of relevance to a particular case. The responses of only the individual who had an experience of more than five years (super experienced respondents) were filtered to perform this comparative study. It showed that eighty five (85) out of 133 instances of digital evidence were reordered. Twenty five (25) out of these eighty five (85) types of digital evidence needed intergroup reordering i.e., a shift from one grade to another while rest of the sixty

(60) elements just required intragroup reordering, as is evident from Table 5.

Table 5
Frequency of Reordering

	DT	RP	MRD	CCF	HMT	EE	CP	Total
Intergrade Reordering	5	4	4	3	2	7	0	**25**
Borderline Elements	5	4	3	3	1	7	0	**23**
Total Reordering	11	12	15	13	14	13	7	**85**

Types of digital evidence with similar levels of relevance were grouped together and categorized with the same grades. So intragroup reordering does not have serious consequences while intergroup reordering changes the grade of digital evidence and hence it's associated required resources and thus deemed more serious. But it was further observed that twenty three (23) out of twenty five (25) types of digital evidence requiring intergroup reordering involved borderline entities only e.g., VDF in RP is the first element in the group with Grade "B" and hence a borderline item in its group. The difference of the level of relevance of the first element of Grade "B" and the last element of Grade "A" is very low. Therefore, most of the intergroup reordering with borderline elements will have minimal overall effect on the amount of required resources assigned in the light of SAOBP. Table 5 shows the details of all the instances of reordering.

Although the differences between Table 3 and Table 4 are small, they encouraged us to formally examine how, if any, the analysis of the variance is significant. A total of one hundred and thirty three (133) one way ANOVA tests were performed for every type of digital evidence and digital investigation in the SAOBP. The following testing procedure was adopted:

1. All the responses were arranged with respect to the experience of the respondents which resulted in five groups.

2. Null hypothesis (Ho: $\mu1 = \mu2 = \mu3 = \mu4 = \mu5$) was tested to check if there is no significance difference between the means of all the five groups at 0.05 level of significance.

3. Alternate hypothesis (H1: $\mu1 \neq \mu2 \neq \mu3 \neq \mu4 \neq \mu5$) was accepted in case we had enough evidence in the form of p-value < 0.05 and F > F critical which means at least one sample has a different mean.

4. Having no sufficient evidence to reject the null hypothesis means that, although the responses were segregated with respect to the experience, even then there was no significant difference between the choices of the level of relevance for a type of digital evidence.

Table 6 is a summary of ANOVA tests and it shows the absence of sufficient evidence to reject the null hypothesis for most of the cases i.e., 118/133 (88.72%). Generally, we can formally conclude that experience did not have a sufficient impact on the overall outcome and SOABP with a significance level of 0.05. Excel sheets with raw data containing all the responses, all the details of ANOVA and Normality Testing are publically assessable[1].

---

[1] http://cs2lab.dsv.su.se/datasets/ or http://www.unhcfreg.com (under datasets)

The following sections discuss SAOBP without considering the experience as a factor.

5.2 Properties of Table 3: SAO Best Practices for Mobile Device Forensics

Some important properties of Table 3 to interpret SAO best practices for mobile device forensics are:

1. It represents the relevance of different types of digital evidence graded from Grade "A" to Grade "E".

2. All the grades in a specific type of investigation should be interpreted independently from similar grades in another type of investigation. For instance getting a Grade "A" in the investigation of drug trafficking is not the same as getting a Grade "A" in a murder investigation. Therefore, "A" in drug trafficking might start from an RL of 9 while, in a murder investigation, it can start from an RL of 8.5.

3. In any specific type of investigation, all the grades were derived by dividing the total range of relevance into five equal parts as discussed in Section 4.4.

4. Each grade in a specific investigation can be shared by one or more types of digital evidence.

5. If a grade is shared by more than one type of digital evidence then each of them can carry different levels of relevance. This level is preserved by their respective positions within a group of evidence sharing the same grade. For example, if SMS and Audio Calls share Grade "A" in an investigation but SMS is more relevant than Audio Calls then, while representing/discussing the group, they will be written/discussed in an order to preserve their level of relevance. Therefore, by following the aforementioned rule, SMS will always be written/discussed before Audio Calls.

6. To make the SAO table presentable abbreviations instead of full names are used. The following is a list of all the abbreviations and their corresponding full names:

Drug Trafficking (DT), Murder (MRD), Rape (RP), Credit Card Fraud (CCF), Harassment (HMT), Espionage/Eavesdropping (EE), Child Pornography (CP), SMS, Phonebook/Contacts (PBC), Audio Calls (ADC), MMS, Graphics/Pictures (GP), Email Entries (EML) , EMS, Video Files (VDF), Video Calls (VDC), Memo/Notes (MN), Calendar Entries (CDR) , URLs Visited (URLV), Tasks/To-Do-Lists (TTD), Audio Files (ADF), Bookmarks/Favourites (BKF), Excel Files (EXL), Word Files (DOC), PDF, PowerPoint Files (PPT).

Table 6
Frequency of Reordering

	DT	RP	MRD	CCF	HMT	EE	CP
PBC	Ho	Ho	Ho	Ho	Ho	Ho	Ho
CDR	Ho	Ho	Ho	Ho	Ho	H1	Ho
MN	Ho	Ho	Ho	Ho	Ho	Ho	Ho
TTD	Ho	Ho	Ho	Ho	Ho	Ho	Ho
SMS	Ho	Ho	Ho	Ho	Ho	Ho	Ho
MMS	Ho	Ho	H1	H1	Ho	Ho	Ho
EMS	Ho	Ho	H1	Ho	Ho	Ho	Ho
ADC	Ho	Ho	Ho	Ho	Ho	Ho	Ho
VDC	H1	Ho	H1	Ho	H1	H1	Ho
EML	H1	Ho	H1	Ho	Ho	Ho	Ho
URLV	Ho	Ho	Ho	Ho	Ho	Ho	Ho
BKF	Ho	Ho	Ho	Ho	Ho	Ho	Ho
ADF	H1	Ho	Ho	Ho	Ho	Ho	Ho
VDF	Ho	Ho	Ho	Ho	Ho	Ho	Ho
GP	Ho	Ho	Ho	Ho	Ho	Ho	Ho
DOC	Ho	H1	Ho	H1	Ho	Ho	Ho
EXL	Ho	Ho	H1	H1	Ho	Ho	Ho
PPT	Ho	Ho	Ho	Ho	Ho	Ho	Ho
PDF	Ho	Ho	Ho	Ho	Ho	Ho	Ho
Ho Rejection	**3**	**1**	**5**	**3**	**1**	**2**	**0**

5.3 Benefits of SAO Best Practices for Mobile Device Forensics and Triage

Table 3 has prioritized lists of digital evidence in descending order of the level of relevance for all of the types of investigations. Investigators should obtain guidance from these lists to concentrate on the most relevant types of digital evidence in order of relevance shown by the table. This will help optimize the allocation of resources in order to achieve the best results; meaning the investigators will have better chances of finding relevant types of evidence with lesser effort and in a shorter time span, and in doing so, will improve the efficiency of the overall process. This fact also coincides with the advantages of another important field of forensics called "Digital Forensics Triage".

Pearson and Watson (2010b) argue that empowering the first responder with the tools and training to find the actionable intelligence immediately, in the form of digital evidence can potentially:

1. Give a positive feedback to the analysis phase.

2. Prevent the laboratory from processing non-yielding data.

3. Reduce the probability of the generation of backlogs in the laboratory.

Figure 11. Forensics processing pyramids (Pearson & Watson, 2010b).
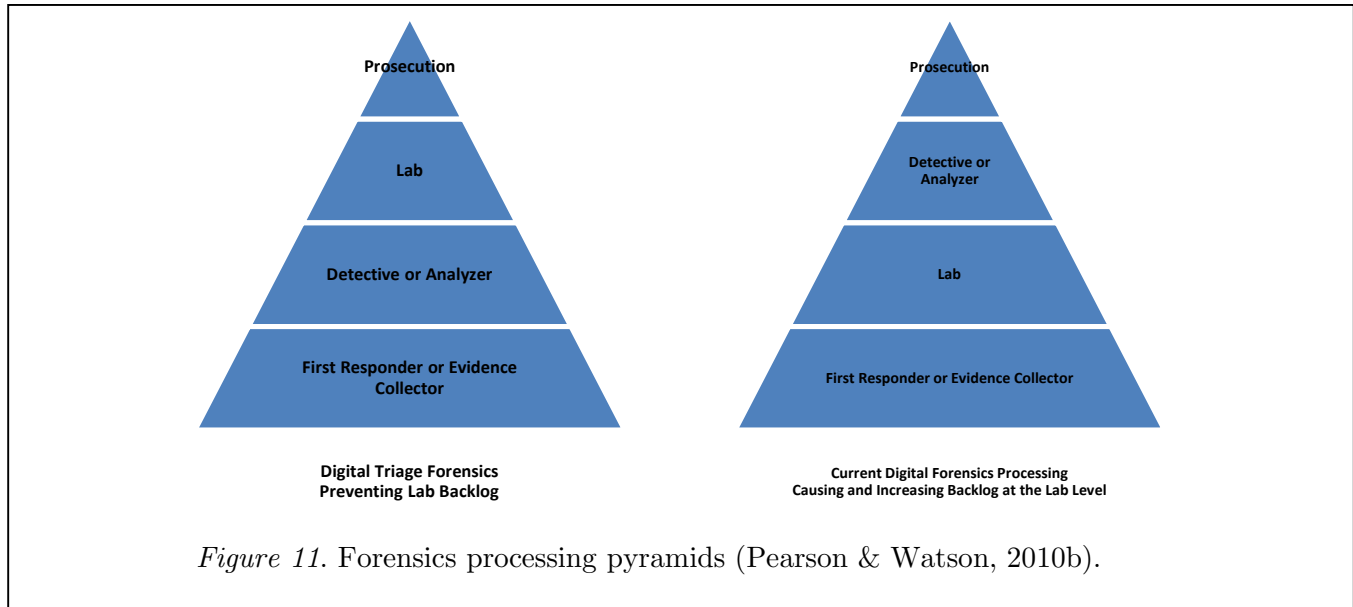
As depicted in Figure 11, Pearson and Watson (2010b) classify current digital forensics processing and digital forensics triage into four levels. In traditional digital forensics processing the first respondent collects and forwards all the media to the lab without any exploitation attempts. In the second level, the lab is responsible for all of the processing and becomes backlogged due to the sheer volume of data that must be analyzed. Consequently, the detective or analyzer at the third level then has to wait for the results from the lab due to a backlog.

These problems can be resolved by digital forensics triage. This would allow the first responder, at the first level, to use tools and his/her training to find actionable intelligence immediately. The detective or analyzer now sits on the second level where they acquire relevant digital evidence immediately, which saves the lab from having backlogs. The lab, now at the third level, finishes the process by exploiting the data found to have value by the first responder to work on a reduced set of data.

In Pearson's research (Pearson & Watson, 2010b) the resources are relocated to the sub-processes where they are needed the most i.e. resource prioritization is emphasized. However,

in the world of digital forensics, digital triage, as distilled in this research, is the process of first ranking and then dealing with the evidence with respect to its volatility or relevance in solving or furthering the case at hand; thus saving time (Rogers et al., 2006). Hence our SAOBP prioritize the relevant evidence for a specific case type. These best practices can fit into the current as well as the digital triage forensics process models.

In the current forensics process model, SAOBP can fit into the laboratory on the second level. The lab initiates its processing on the prioritized lists of the digital evidence from SAOBP to obtain the initial leads. These leads can become seeds for further search strategies to uncover all of the possible relevant digital evidence in a shorter period of time.

In the digital triage forensics process model, SAOBP can fit into the first level where the first responder can now have a prioritized list to guide him/her in finding at least the first few pieces of relevant digital evidence. Later on, these initial pieces of evidence will act as leads to uncover the remaining segments of the puzzle in a shorter period of time. Moreover, during digital forensics triage, conducting on-site examination and analysis can provide time sensitive leads

and thus huge psychological advantage to the investigators.

According to Black and Yeschke (2003), suspects are psychologically more vulnerable during the first few hours of their initial contact with the police. This vulnerability is even higher if the contact occurs on the scene or at the suspect's place of business/dwelling. This vulnerability causes the suspects to be more cooperative, which can eventually help in a timely advancement or solution to the case at hand. During these initial stages, the aim of the investigator is to find the digital evidence that can have positive effect on the suspect's willingness to cooperate (Rogers et al., 2006). Rogers et al. (2006) advocate the use of on-site triage to achieve this objective. Having pre-raid information about the relevant digital evidence (for a specific case) is important to produce better search strategies and save time. In this regard, they have presented three types of cases. Our research carries this concept even further by introducing more types of cases and digital evidence in an MD. Generalization, in our research, is based on a worldwide survey; hence, our results carry quantifiable weights.

Based on SAOBP, we have also devised tables (presented in the Appendix) carrying an ordered list of digital evidence with respect to their proportional relevance. These proportions are important to help allocate resources by using the equation below:

$$r_i = p_i * R_t \qquad \textbf{Equation 6}$$

Where:

i = 1, 2, 3 ... 19

r_i are the resources required for current (ith) digital evidence.

p_i is the proportional relevance of the current (ith) digital evidence.

R_t are the total resources at an expert's disposal.

Optimum resource allocation can be performed by using the mathematics in Equation 6 and our prioritized lists of digital evidence based on their relevance in solving or furthering a specific case. Optimum resource allocation has the potential to speed up the overall forensics process, which is a necessity for investigations, especially with the amount of data that has to be processed.

6. FUTURE WORK

As explained earlier, SAOBP can potentially improve the performance on all or any of the sub-processes of a digital forensics process depending on the settings of investigative context. In military settings, it can help the first responder to concentrate on a sub-set of the data while examination and thus give a positive feedback to the analysis phase to improve the overall performance. In law enforcement settings, SAOBP can improve the performance of examination sub-process. It can help the examiner to concentrate on the prioritized list of evidence to either reach to the proper conclusions or at least to the first few leads to reveal the remaining necessary pieces of digital evidence required to solve the case at hand.

The evolution in this field means the lists of the types of evidence, investigations and forensics process are expanding rapidly. The technique adopted to formulate SAOBP is generic and can potentially expanded horizontally to cover the evolving list of forensics processes and also vertically to include more types of digital evidence and investigations. To suffice the needs of evolution we have classified the types of digital evidence into seven abstract classes namely (i) PIM, (ii) Messaging, (iii) Call Logs, (iv) Email Entries, (v) Internet History, (vi) Standalone Files and (vii) Application Related Files. Each class has a list of application independent instances for example an SMS is an SMS whether sent by native application or by a third party application (Handcent, etc.) similarly a dialed number can be associated with the native application or any other third party application (Whats App, Viber, Skype, etc.). However the list of abstract instances and even the classes can expand to cover the new types of evidence. For example, in future we will add another application independent instance of

"text chat" performed by, for instance, Facebook, Skype, Viber, etc.

SAOBP is based on a survey, where experts were contacted to tag the factor of relevance in the light of their experience. The outcome can actually be treated as a profile where different types of evidence are ordered with respect to their relevance to the case types. In future, we want to build another similar type of profile based on the real case files. We will then compare these two profiles to see if we can find any reasonable discourse between them. Finally, these two profiles will work as corrective and predictive profiles for a decision support system, which will consequently fine tune the SAOBP.

We further hope to extend the study to include other branches of digital forensics as well. Additionally, we intend to augment the results obtained from our survey with the performance measures of mobile device forensic tools. This will help us in selecting the most appropriate tool for a specific type of a case both in terms of performance and relevance, especially in the cases where the investigator has to rely on just one tool.

REFERENCES

ADF Solutions Incharge. (2013). Triage computers to reduce forensic backlogs and lower costs. Retrieved September 13, 2013, from http://www.adfsolutions.com/products/triage-examiner

Anobah, M. (2013). *Testing Framework for Mobile Forensic Investigation Tools*. Stockholm University.

Baggili, I., Marrington, A., & Jafar, Y. (2014). Performance of a Logical, Five- Phase, Multithreaded, Bootable Triage Tool.pdf. In *Advances in Digital Forensics X* (pp. 279–295). Springer.

Baggili, I., Mislan, R., & Rogers, M. (2007). Mobile phone forensics tool testing: A database driven approach. *International Journal of Digital Evidence, 6*(2). Retrieved from http://www.utica.edu/academic/institutes/ecii/publications/articles/1C33DF76-D8D3-EFF5-47AE3681FD948D68.pdf

Black, I., & Yeschke, C. L. (2003). *The Art of Investigative Interviewing,* 2nd ed. Butterworth-Heinemann.

Bowen, L. (2001). The Borda Count Method. Retrieved September 12, 2013, from http://www.ctl.ua.edu/math103/voting/borda.htm#Determine1

Brinson, A., Robinson, A., & Rogers, M. (2006). A cyber forensics ontology: Creating a new approach to studying cyber forensics. *Digital Investigation, 3*, 37–43. doi:10.1016/j.diin.2006.06.008

Casey, E. (2011). *Digital evidence and computer crime: Forensic science, computers, and the Internet*, 3rd ed.

Encyclopedia Britannica. (2013). Marriam-Webster: Triage. Retrieved July 25, 2013, from http://www.merriam-webster.com/dictionary/triage

Garfinkel, S. L. (2010). Digital forensics research: The next 10 years. *Digital Investigation, 7*, S64–S73. doi:10.1016/j.diin.2010.05.009

International Telecommunication Union (ITU). (2013). ICT Facts and Figures. Retrieved September 23, 2013, from http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2013.pdf

iSixSigma. (2013). How to determine sample size. Retrieved May 03, 2013, from http://www.isixsigma.com/tools-templates/sampling-data/how-determine-sample-size-determining-sample-size/

James, J., & Gladyshev, P. (2013). Challenges with automation in digital forensic investigations. *arXiv Preprint arXiv:1303.4498*. Retrieved from http://arxiv.org/abs/1303.4498

James, J. I., & Gladyshev, P. (2013). A survey of digital forensic investigator decision

processes and measurement of decisions based on enhanced preview. *Digital Investigation*, *10*(2), 148–157. doi:10.1016/j.diin.2013.04.005

Kubi, A., Saleem, S., & Popov, O. (2011). Evaluation of some tools for extracting e-evidence from mobile devices. In *Application of Information and Communication Technologies*, 603–608. Baku: IEEE. doi:10.1109/ICAICT.2011.6110999

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental ...*, 4–6. Retrieved from http://www.sciencedirect.com/science/article/pii/S0022103113000668

Marturana, F., Me, G., Berte, R., & Tacconi, S. (2011). A Quantitative approach to triaging in mobile forensics. *2011 IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications*, 582–588. doi:10.1109/TrustCom.2011.75

Mislan, R. P., Casey, E., & Kessler, G. C. (2010). The growing need for on-scene triage of mobile devices. *Digital Investigation*, *6*(3-4), 112–124. doi:10.1016/j.diin.2010.03.001

Narasimhan, B. (1996). The normal distribution. Retrieved September 11, 2013, from http://www-stat.stanford.edu/~naras/jsm/NormalDensity/NormalDensity.html

National Institute of Justice. (2001). Electronic crime scene investigation: A guide for first responders. Retrieved February 15, 2012, from https://www.ncjrs.gov/txtfiles1/nij/187736.txt

National Security Database. (2013). Digital forensic analysis. Retrieved February 05, 2013, from http://www.nsd.org.in/digital-forensic-analysis/

Oxford University Press. (2013). Oxford Dictionaries: Triage. Retrieved July 25, 2013, from http://oxforddictionaries.com/definition/english/triage?q=triage

Palmer, G. (2001). A road map for digital forensic research. Digital Forensic *Research Workshop (DFRWS)*. Retrieved from http://www.dfrws.org/2001/dfrws-rm-final.pdf

Pearson, S., & Watson, R. (2010a). *Digital Triage Forensics: Processing the Digital Crime Scene.* (M. Harrington, Ed.) (1st ed.). Syngress.

Pearson, S., & Watson, R. (2010b). Introduction: Using the digital triage forensics model to collect and process cell phones and SIM cards. In *Digital Triage Forensics: Processing the Digital Crime Scene* (pp. ix–xi). Syngress. doi:10.1016/B978-1-59749-596-7.00012-7

Reith, M., Carr, C., & Gunsch, G. (2002). An examination of digital forensic models. *International Journal of Digital Evidence*, *1*(3), 1–12.

Rogers, M. K. (2004). DCSA: A Practical Approach to Digital Crime Scene Analysis. West Lafayette: Department of Computer Technology, Purdue University. Retrieved January 17, 2013, from http://www2.tech.purdue.edu/cit/Courses/cit556/readings/DCSA.pdf

Rogers, M. K., Mislan, R., Goldman, J., Wedge, T., & Debrota, S. (2006). Computer forensics field triage process model. In *Conference on Digital Forensics, Security and Law,* 27–40. Retrieved from http://www.digitalforensics-conference.org/CFFTPM/CDFSL-proceedings2006-CFFTPM.pdf

Routledge. (2004). Introduction. In *Cavendish: Evidence Lawcards,* 3rd ed., 1–8. Routledge-Cavendish.

Ryan, D. J., & Shpantzer, G. (2002). Legal aspects of digital forensics. In *Proceedings: Forensics Workshop.* Retrieved from

http://euro.ecom.cmu.edu/program/law/08
-732/Evidence/RyanShpantzer.pdf

Saleem, S., Popov, O., & Bagilli, I. (2014). Extended Abstract Digital Forensics Model with 2PasU. *Procedia Computer Science*, *35*, 812–821.

Saleem, S., Popov, O., & Kubi, A. (2013). Evaluating and comparing tools for mobile device forensics using quantitative analysis. *Digital Forensics and Cyber Crime: Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, *114*, 264–282. doi:10.1007/978-3-642-39891-9_17

APPENDIX I

Tables with the detailed information of our survey related to the relevance of digital evidence are presented in this appendix.

Prop. Rel. means Proportional Relevance which is normalized by using the following equation.

$$PR_i = \frac{RL_i}{\sum_{i=1}^{19} RL_i}$$ **Equation 7**

Table 1
Relevance of Digital Evidence for Drug Trafficking Investigation

Grade	Types of Evidence	0	1	2	3	4	5	6	7	8	9	10	Weighted Average	Prop. Rel.
Grade A	**SMS**									4	8	38	**9.68**	**0.076**
	Phonebook/Contacts									5	11	32	**9.56**	**0.075**
	Audio Calls								7	5	12	23	**9.09**	**0.072**
	MMS				2		3		3	5	7	27	**8.83**	**0.070**
	Graphics/Pictures(GPS)						3		5	10	9	21	**8.77**	**0.069**
	Email Entries						3		6	9	11	17	**8.65**	**0.068**
Grade B	**EMS**			5		2	3	3	4	7	7	16	**7.62**	**0.060**
Grade C	**Video Files**		4				5	6	10	8	5	8	**7.04**	**0.055**
	Video Calls	5		5		2	8		2	5	6	12	**6.36**	**0.050**
	Memo/Notes		7			4	5	6	10	3	5	8	**6.31**	**0.050**
	Calendar Entries		7				8	10	3	7	4	7	**6.30**	**0.050**
	URLs Visited			4		3	13	6	9	3	3	5	**6.20**	**0.049**
	Tasks/To-Do-Lists		5	7	3		5	5	9		6	8	**5.83**	**0.046**
Grade D	**Audio Files**	4	6		4		8	3	7	4	3	6	**5.42**	**0.043**
	Bookmarks/Favourites			4	8	4	13	5	6	3		4	**5.30**	**0.042**
	Excel Files		8	5	3	3	6	5	8	3		5	**4.98**	**0.039**
Grade E	**Word Files**	3	8	5	5	3	7	5	7			5	**4.35**	**0.034**
	PDF	6	8	9	6		7		7			4	**3.57**	**0.028**
	Powerpoint Files	10	6	10	6		6		6			3	**3.11**	**0.024**

Table 2
Relevance of Digital Evidence for Rape Investigation

Grade	Types of Evidence	0	1	2	3	4	5	6	7	8	9	10	Weighted Average	Prop. Rel.
Grade A	**SMS**								4	2	11	23	**9.33**	**0.079**
	Graphics/Pictures(GPS)									13	8	17	**9.11**	**0.078**
	Phonebook/Contacts								5	5	11	18	**9.08**	**0.077**
	MMS				2				4	2	9	23	**9.03**	**0.077**
	Audio Calls							2	6	4	14	13	**8.77**	**0.075**
Grade B	**Video Files**				5		4		4	6	8	10	**7.65**	**0.065**
	EMS				3	4	4	3	4	3	7	13	**7.51**	**0.064**
	Email Entries				3		8	2	5	5	5	11	**7.46**	**0.064**
	Video Calls	4	3				4	3	2	2	8	11	**6.84**	**0.058**
Grade C	**Calendar Entries**		4	3	2	2	7	2	5		7	7	**6.13**	**0.052**
	Audio Files	2	3		6	3	3	3	5	4		9	**5.87**	**0.050**
	URLs Visited	3	4	3			10	3	4	3	2	6	**5.47**	**0.047**
Grade D	**Memo/Notes**	4	4	3	5		9	3	3		2	7	**4.93**	**0.042**
	Tasks/To-Do-Lists	3	6	8		2	7	4		3		6	**4.44**	**0.038**
	Bookmarks/Favourites	3	6	3	3	2	11		4			5	**4.38**	**0.037**
Grade E	**Word Files**	8	8		4	4	7	3	2			4	**3.58**	**0.030**
	Excel Files	9	8	6	3	3	7					4	**2.90**	**0.025**
	PDF	9	9	2	4	4	8					2	**2.66**	**0.023**
	Powerpoint Files	10	8	4	4	3	6		2				**2.27**	**0.019**

Table 3
Relevance of Digital Evidence for Murder Investigation

Grade	Types of Evidence	0	1	2	3	4	5	6	7	8	9	10	Weighted Average	Prop. Rel.
Grade A	**SMS**										13	27	**9.68**	**0.069**
	Phonebook/Contacts									3	9	30	**9.64**	**0.069**
	Audio Calls						2			6	9	23	**9.23**	**0.066**
	MMS						3			6	8	25	**9.17**	**0.066**
	Email Entries							3	3	8	7	18	**8.87**	**0.063**
Grade B	**Graphics/Pictures(GPS)**					2	4		3	6	6	20	**8.56**	**0.061**
	Calendar Entries						3	3	4	7	8	15	**8.48**	**0.061**
	EMS					4	5		2	5	7	17	**8.20**	**0.059**
	Memo/Notes					3	5	4	3	4	5	15	**7.92**	**0.057**
Grade C	**URLs Visited**	4					3	6	3	6	3	14	**7.36**	**0.053**
	Video Files		2		3	3	6	2	2	7		15	**7.13**	**0.051**
	Tasks/To-Do-Lists		3	3		3	5	3	2		7	14	**7.03**	**0.050**
	Video Calls	4	4	3					3	3	9	13	**6.82**	**0.049**
Grade D	**Bookmarks/Favourites**	4			4	4	5	3	4	4		11	**6.18**	**0.044**
	Audio Files	5	3	3		2	5	4		4	2	13	**6.00**	**0.043**
Grade E	**Word Files**	3	8	3	3		4		7	3		11	**5.38**	**0.038**
	PDF	6	7	3			5		4	2	2	9	**5.00**	**0.036**
	Excel Files	6	7	4			7		4		1	11	**4.93**	**0.035**
	Powerpoint Files	8	8	4			6		2	3		9	**4.35**	**0.031**

Table 4
Relevance of Digital Evidence for Credit Card Fraud Investigation

Grade	Types of Evidence	0	1	2	3	4	5	6	7	8	9	10	Weighted Average	Prop. Rel.
Grade A	SMS						3		3	7	6	19	**8.84**	**0.064**
	Email Entries					2			4	8	5	19	**8.82**	**0.064**
	Phonebook/Contacts						5		4	8	8	17	**8.55**	**0.062**
	URLs Visited				2		2		6	6	7	16	**8.44**	**0.061**
	MMS						6	3	2	7	6	14	**8.21**	**0.059**
Grade B	Audio Calls		2				5		4	8	4	15	**8.03**	**0.058**
	Bookmarks/Favourites					2	7		4	6	6	14	**8.03**	**0.058**
	Excel Files		3		3			4	2	9	7	11	**7.64**	**0.055**
	Graphics/Pictures(GPS)		3	2			4	3	3	6	7	11	**7.36**	**0.053**
Grade C	Word Files		5				3	3	4	7	5	11	**7.29**	**0.053**
	EMS		2		2		6	5	5	3	4	12	**7.26**	**0.053**
	Memo/Notes		2		4		4	2	8	2	6	11	**7.23**	**0.052**
	Calendar Entries		2		3		8	5	7	3	2	11	**6.88**	**0.050**
	Tasks/To-Do-Lists			2	5		5	3	7	6	2	9	**6.85**	**0.050**
Grade D	PDF	3	5				8	3	5	6	2	7	**6.00**	**0.043**
	Video Calls	6			4		6		8	7	3	5	**5.92**	**0.043**
	Video Files	2	4	3			8	5	3	3		10	**5.92**	**0.043**
Grade E	Audio Files		5	4	2		8	4	5	2	2	7	**5.69**	**0.041**
	Powerpoint Files	4	7	4			5	2	2	6	2	6	**5.05**	**0.037**

Table 5
Relevance of Digital Evidence for Harassment Investigation

Grade	Types of Evidence	0	1	2	3	4	5	6	7	8	9	10	Weighted Average	Prop. Rel.
Grade A	SMS										6	31	**9.84**	**0.072**
	MMS									5	6	28	**9.59**	**0.070**
	Phonebook/Contacts									5	10	26	**9.51**	**0.069**
	Audio Calls								3		10	25	**9.50**	**0.069**
	Email Entries									6	12	21	**9.38**	**0.069**
	Graphics/Pictures(GPS)						3	2		4	8	22	**9.00**	**0.066**
	EMS				2		2	3		6	4	20	**8.59**	**0.063**
Grade B	Video Files				2		5	3	3	3	6	14	**8.00**	**0.058**
	Video Calls	2		2			4		2	3	7	17	**7.97**	**0.058**
	Audio Files			2	3		5	3	3	5	3	13	**7.41**	**0.054**
	Calendar Entries		5			2	3	4	3	3	4	13	**7.11**	**0.052**
Grade C	Memo/Notes		3		4	4	4		3	6	3	12	**6.85**	**0.050**
	URLs Visited				3	3	6	4	6	4	6	5	**6.84**	**0.050**
	Tasks/To-Do-Lists		3		5	3	5	2	3	4	3	9	**6.41**	**0.047**
	Bookmarks/Favourites		3	3	3		6	4	8		6	5	**6.11**	**0.045**
Grade D	Word Files	6	6		4			2	2	6	3	6	**5.11**	**0.037**
Grade E	Powerpoint Files	6	9	4	3	3	6	3		4		2	**3.45**	**0.025**
	PDF	7	10	3	4		6		2	4	2		**3.21**	**0.023**
	Excel Files	6	9	5	3	4	5		3	3			**3.00**	**0.022**

Table 6
Relevance of Digital Evidence for Espionage/Eavesdropping Investigation

Grade	Types of Evidence	0	1	2	3	4	5	6	7	8	9	10	Weighted Average	Prop. Rel.
Grade A	Audio Calls								2	3	12	21	9.37	0.060
	Phonebook/Contacts								2	5	10	21	9.32	0.060
	SMS								4	4	12	18	9.16	0.059
	Email Entries						2			7	12	18	9.13	0.059
Grade B	Graphics/Pictures(GPS)					2	3		2	3	7	21	8.79	0.056
	Audio Files					4	3	1	2	9	17		8.67	0.056
	MMS				2		2		4	5	8	17	8.58	0.055
	Video Files				2	2	2	3	5	8	16		8.50	0.055
Grade C	URLs Visited					4	2	5	4	10	13		8.39	0.054
Grade D	Word Files		3			3	3	4	3	5	16		7.92	0.051
	EMS			3	3		3			4	9	13	7.80	0.050
	Memo/Notes				4	2	4		4	3	6	15	7.79	0.050
	Excel Files		3		3		2	3	4	2	4	17	7.63	0.049
	PDF		4			3	2	2	3	3	5	16	7.58	0.049
Grade E	Bookmarks/Favourites		2				7	4	4	4	4	13	7.55	0.048
	Calendar Entries	3			4			2	4	5	4	15	7.51	0.048
	Tasks/To-Do-Lists				4	4	4		3	3	6	13	7.49	0.048
	Video Calls	3		2			5	2		3	11	12	7.47	0.048
	Powerpoint Files		4	4		2	2		3	3	5	15	7.11	0.046

Table 7
Relevance of Digital Evidence for Child Pornography Investigation

Grade	Types of Evidence	0	1	2	3	4	5	6	7	8	9	10	Weighted Average	Prop. Rel.
Grade A	Graphics/Pictures(GPS)										7	33	9.83	0.069
	Video Files									5	5	28	9.61	0.067
	URLs Visited							2		7	7	24	9.28	0.065
	Bookmarks/Favourites								3	7	8	20	9.18	0.065
	Email Entries								4	7	10	18	9.08	0.064
	SMS				2		2			5	3	27	9.05	0.064
	MMS				2		2			4	7	25	9.03	0.063
Grade B	Phonebook/Contacts				2			1	1	8	9	18	8.82	0.062
	EMS			3		3			4	6	4	18	8.16	0.057
	Audio Calls			2	4		2		2	7	8	15	7.95	0.056
Grade C	Video Calls	3			4		3	2	2	6	6	14	7.38	0.052
Grade D	Calendar Entries	3	4	4			7		5	3	3	11	6.08	0.043
	Audio Files	2	5	4		3	4		4	5		13	6.08	0.043
	Memo/Notes	3	3	5	3		3		4	5	6	8	5.98	0.042
	Word Files	5				6	6	3	4	5		8	5.95	0.042
Grade E	Powerpoint Files	4	4		3	10	3	2	3	2	8		5.64	0.040
	Tasks/To-Do-Lists	4	4	5	3		5		4	3	3	8	5.31	0.037
	Excel Files	5	4		3	3	8	3	3	4		6	5.03	0.035
	PDF	4	6	3	3		8		3	4		8	4.97	0.035