FACTORS INFLUENCING DIGITAL FORENSIC INVESTIGATIONS: EMPIRICAL EVALUATION OF 12 YEARS OF DUBAI POLICE CASES

Ibtesam Al Awadhi, Janet C Read
University of Central Lancashire
School of Computing, Engineering and Physical Sciences. Preston, UK
{IAlawadhi, JCRead}@uclan.ac.uk

Andrew Marrington
Zayed University
College of Technological Innovation. Dubai, UAE
andrew.marrington@zu.ac.ae

Virginia N. L. Franqueira
University of Derby
College of Engineering and Technology. Derby, UK
v.franqueira@derby.ac.uk

ABSTRACT

In Digital Forensics, the number of person-hours spent on investigation is a key factor which needs to be kept to a minimum whilst also paying close attention to the authenticity of the evidence. The literature describes challenges behind increasing person-hours and identifies several factors which contribute to this phenomenon. This paper reviews these factors and demonstrates that they do not wholly account for increases in investigation time. Using real case records from the Dubai Police, an extensive study explains the contribution of other factors to the increase in person-hours. We conclude this work by emphasizing on several factors affecting the person-hours in contrast to what most of the literature in this area proposes.

Keywords: cyber forensics, digital forensics, empirical data, forensic investigation, Dubai police

1. INTRODUCTION

Year on year, digital forensic teams face the mounting challenge of diversification of storage devices and distribution of data across many storage areas. In single investigations, practitioners are now expected to search more storage than they were five years ago (Irons & Lallie, 2014). This growth and spread of data raised considerations on how to best manage the analysis of material with finite human resource and time constraints. Forensic tools can take some of the work from the human element but still there is a need to better understand how to allocate and manage person-hours so that investigations can be concluded within reasonable time and reliable findings. This research adds to the understanding in this field by studying real case records from the Dubai Police for the past 12 years. The growth in cases is measured and the main factors behind time spent in their investigation are identified. This research contributes for the understanding of the effects

which volume and heterogeneity of evidence items cause to person-hours spent by Digital Forensic (DF) practitioners.

2. LITERATURE REVIEW

Many research papers studied empirically the current status of DF investigation capabilities and identified challenges which affect different aspects of DF investigations. Gogolin (2010) conducted interviews with practitioners from 45 agencies in Michigan, USA. He identified the current status of experience and investigation capabilities of law enforcement. Dezfoli et al. (2013) conducted a statistical study to cover the trends of several aspects of DFs and security. The research suggests some factors which need to be considered by the digital forensic investigations in order to adapt to the new challenges in the field. Irons & Lallie (2014) demonstrated a yearly growth in the number of forensic investigations, the amount of data being investigated, and the amount of data being investigated per case using the annual data published by the FBI from 2007 to 2011. The authors concluded that digital crimes are increasing remarkably every year.

The literature suggests a need to improve the use of the available resources and move beyond the capabilities of the current forensic tools. Each DF process entails a number of challenges including: heterogeneous sources, data diversity, anti-forensics, volume of digital evidence, legal issues, and maintenance of efficiency levels of DF departments. Many practical solutions have been implemented by different DF departments to militate against those challenges. Examples include: features introduced into DF commercial tools, the use of random sampling (Roy, 2014), triage (James, 2014), enhanced previewing (Shaw & Browne, 2013), information visualization (Prefuse, 2013), distributed DFs (Roussev & Golden, 2004) and the use of data mining tools analysis (Nirkhi, Dharaskar, & Thakre, 2012).

3. THE RESEARCH STUDY

To date, research in DF has mainly focused on solutions to technical problems or the analysis of issues faced by practitioners, often not supported by empirical data. This paper reports on the analysis of factors associated with person-hours based on completed cases from the Dubai Police.

3.1 The Dubai Police DF Department

The Dubai Police DF Department is composed of 32 investigators. The department is structured in different sections: Computer, Network, Mobile, Programs & Databases, Photos & Videos Analysis and Voice Analysis. Each section follows the standard DF processes (i.e., acquisition, examination, analysis and reporting) with the goal to examine digital media in a forensically sound manner. The Computer section deals with evidence found in computers, embedded systems, and static memory. It deals with crimes like unauthorized access, intellectual property theft or misuse of information, illicit pornography possession, theft of services, forgery, invasion of privacy, denial of service, sabotage, extortion, embezzlement, espionage, terrorism, racketeering, money laundry, human trafficking, corruption, harassment and discrimination, organized crimes, suicide, threat, and blackmail. The Network section monitors and analyzes computer network traffic for the purpose of data gathering. Hence, this section differs from the other sections because it deals with volatile and dynamic information. The most common crimes investigated by this branch are network breaches, network piracy, unusual network activities, eavesdropping, botnets, targeted

attacks, obtaining information by unauthorized computer access, economic espionage and damage or destruction of property. The Mobile section is concerned with recover/extraction of data from devices like mobile devices, PDAs, GPS navigation devices, tablet computers. This section includes cases like mobile malware analysis, human trafficking, impersonation, defamation and slander, harassment and discrimination, threat/intimidation and theft. The Programs & Databases section covers the cases with databases and their related metadata and cached information. Some examples of the cases in this section are: database breaches, unlicensed commercial voice over IP activities, online piracy and fraud. The Photos & Videos Analysis section and Voice Analysis section analyze photos, videos and voice files related to different types of crimes, for example, revealing the identity of a thief.

3.2 Data Gathering

The records for this study were collected manually due to the fact that the source of the records were spread between the manual archives, databases and acquisition verification reports in different sections of the Dubai Police DF department. The original records in the databases were stored in the Arabic language so were sampled and translated into English and inserted into a new database specifically for this study. Thus, the process of gathering the data took a long time (almost ten weeks).

3.3 Sampling Methods

Records from January 2003 to February 2015 were initially collected. The data used for the study was selected from February 2003 to December 2014 to make sure that the examiners were not working on pending cases which started before 2003, and to be sure that all the selected cases were completed. There were three sources for the collected data. The first source of data was the case records database which held 8620 records. The DB contains information about examiners, cases, and evidence and data from four sections of the Digital Forensics Department. There were two factors for records selection. Only cases received by the Computer, the Network, the Mobile, and the Programs & Databases sections were selected. Outliers were then deleted from the database. The outliers were determined using Cook's distance analysis (Kim, 1996). Records with Cook's distance values above 2 or less than -2 were considered as abnormal records. In this way, 5097 records were selected for this study and 3523 were disregarded because they were either outliers or not relevant for the study such as classified cases were considered as not relevant because the database did not include all the required information. The remaining 5097 records were stored in a new database called “Complete Case Details”. The second source of data was documentation related to acquisition and verification, which consisted of 4398 reports. The final source of data was the inventory database with more than 600 records which included specifications of the devices.

3.4 Study Variables

The dependent variable for the analysis is person-hours per case. Independent variables are the number of cases, the case received date, the total volume per case, the total number of evidence items per case, the total number of examiners per case, the total number of evidence items per examiner at the same time the case type, the case request details, and the number of evidence types.

3.5 Limitations

This study aimed to cover all sections under the Dubai Police DF Department; however, cases from the Photos & Videos Analysis and Voice Analysis sections were not included. Those sections were previously under the Fingerprint Department and were only incorporated into the DF Department in 2013,

and their records prior to 2013 were not available.

This paper focuses on factors such as the increase in the number of cases and hard disk volume vs. person-hours, independent of the complexity of the cases concerned.

4. GENERAL DESCRIPTION OF DATA

Descriptive statistics show that the number of cases increased each year as illustrated in Figure 1. There were 51 cases in 2003 and more than 900 cases in 2013. In 2010 there was an extraordinary increase in the number of cases due to several high profile crimes in that year which led to the need to initiate more cases. Generally, the number of cases kept increasing throughout the past twelve years. The Computer and Mobile sections received the highest number of cases in across the years.

Across the years, the average time spent in investigation per case was between 100 to 200 hours. There was an exception in 2007, 2008 and 2009 where the average time was less than 100 hours. The average person-hours per case reached a peak of 198 hours in 2014. The averages of the total number of person-hours across the different DF sections show that the cases in the Network section took longer to investigate than other sections in most of the years except for the years 2003, 2006 and 2008. The cases in the Computer section required the second highest amount of time for investigation in all the years from 2004 to 2010. In 2003, the Computer section cases required the highest amount of time for investigation. Since 2011, the cases in the Computer section need less time to investigate than the cases in the Mobile section and Databases & Programs section. It is also noticeable that the time to investigate the cases in the Mobile section has been steadily increasing since 2012.

The volume average of evidence items per case also increased over the years. In years 2011 and 2014 there was a 20% increase For example, the average jumped from 171GB per case in 2010 to 900GB in 2011 and 1186GB in 2014. The Network section received the highest volume average of DF items compared to other sections in the years between 2003-2010 and 2013. In 2011, the Databases section received the highest volume of DF items. In years 2012 and 2014, the Computer section received the highest volume of DF items.

The average of the total number of evidence items per case was between 1 and 2 items among all the years (1.67 in 2003 and 2.09 in 2014) except 2011 where it reached a peak of 3.77. The Databases & Programs section received the highest average of the total number of evidence items in years 2003 - 2006. After that, the Network section remained in a peak from 2004 to 2014.

The average number of examiners working in a case remained between 1 and 2 over the years for all sections (2003-2014). However, the load of evidence items each examiner had at a particular time has fluctuated. The number of evidence items each examiner had at once was around (1.9, 2.92, 2.18, 1.69, 2.01, 2.58, 2.28, 4, 3, 3, 5.21, and 5.89) respectively for years 2003 to 2014. As we can see, the number of evidence items has kept increasing over the years reaching almost 6 items at once in 2014.

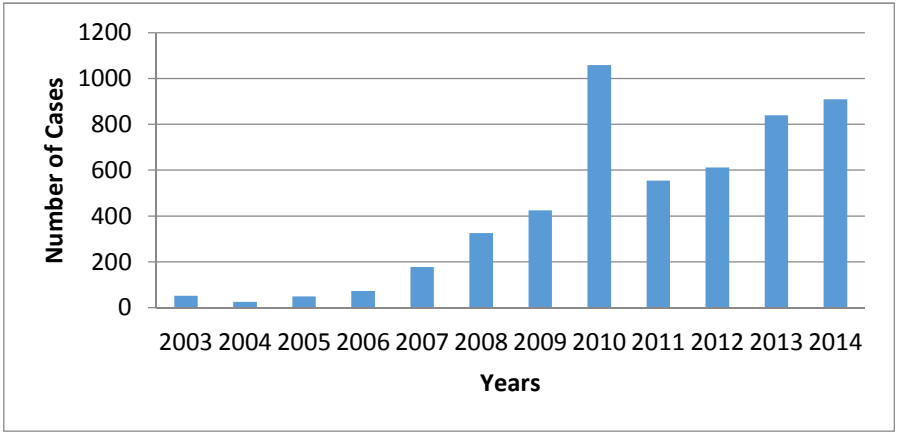

Figure 1. Number of cases from 2003 to 2014 investigated by the Dubai Police DF department

5. ANALYSIS

Data analysis for this study was carried out using a variety of statistical techniques. Data were analyzed using the computerized statistical analysis program SPSS (Version 20). Pearson's Correlation was used to measure the linear correlation between the variables.

5.1 Person-Hours vs. Years

As shown in Figure 1, the number of cases steadily increased over the years. Every year there was a 48% increase in the number of cases compared to the previous year. The Pearson Correlation analysis of the number of cases vs. year equals .884. This means that the total number of cases and the year are strongly correlated. Thus, the number of cases increases progressively every year.

It is important to determine if the total number of person-hours per case has also increased through the years. Simple linear regression was utilized as a key method of regression analysis to study the relationship between this bivariate data, as shown in Figure 2. The Pearson Correlation coefficient .117 reveals a weak relationship between year and person-hours per case. This suggests that person-hours per case did not significantly increase over the years. The majority of cases over the years required less than 200 hours of investigation process for all the evidence items in the case. The number of cases that took more than 200 hours increased slightly over the years and reached a peak by 2011 when it took 4368 hours (nearly 2 years) to investigate 64 evidence items in one case with a size of 28 Terabytes. Despite outlying cases like these, the majority of cases still take less than 200 person-hours to complete.

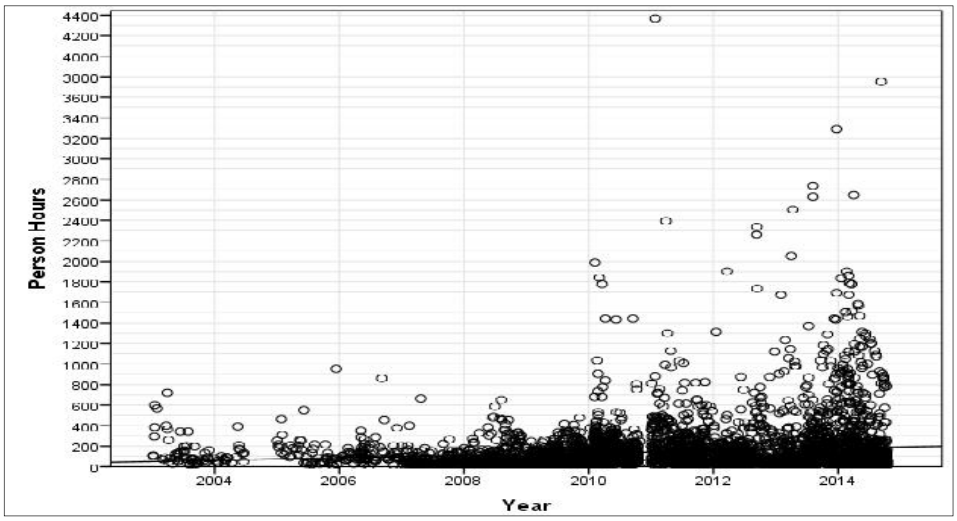

Figure 2. Relationship between person-hours and the year in the Dubai Police DF department

5.2 Person-Hours vs. Volume

The total volume per case also did not affect positively the time of investigation. The Pearson Correlation shows a weak relationship between volume and time spent on each case with a correlation coefficient of 0.388.

5.3 Person-Hours vs. Evidence Items

The Pearson Correlation coefficient between the number of evidence items involved in a case and person-hours spent investigating the case was 0.266. This represents a weak relationship between these bivariate data.

The total number of evidence types per case does not affect the time needed for investigation. The Pearson Correlation of 0.278 indicates a weak relationship between these variables.

Our analysis also showed no significant impact of the category of case on the time taken to conclude the investigation, although on the whole, fraud cases seemed to take more time than other case types.

Although the Pearson Correlation shows that the relationship between the total number of evidence items and year is weak, as discussed in section 5.1, an increase was observed in the number of evidence items over the years.

6. DISCUSSION OF FACTORS INFLUENCING PERSON-HOURS

This section highlights several factors that might influence the increase in person-hours with the caveat that the increase might be a result of the combination of those factors. Environmental factors that might influence the amount of person-hours include hardware specification of workstations used by investigators, availability of investigative software (e.g., specialist DF tools/versions), examiners' experience, complexity of the case and availability of case details. Sections 6.1-6.4 discuss experiments representing some of those assumptions.

6.1 Person-Hours vs. Volume Experiment

There are cases where the total volume of evidence items varied while person-hours were

similar. Several filters were applied on the database to select the desired collection of records. First of all, cases with a total storage volume equal to 4 GB and 1024 GB were selected. 1092 records are selected out from 5097 records in the database. The records were then grouped by the total number of person-hours; this resulted in 85 distinct groups. It was found that the majority of the cases with a total of 4 GB volume in a case were received between 2003 and 2011 and most of the cases with 1024 GB volume were received between 2012 and 2014. Thus, the circumstance of spending similar total numbers of person-hours in cases could be explained by factors that must have changed over time; for example, workstation specifications, DF tools and DF practitioners' experience. These tools were more primitive between 2003 and 2011 compared to 2012 to 2014. The cases received in recent years were investigated with better capabilities and sophisticated workstations, tools and experience.

6.2 Number of DF Practitioners vs. Volume Experiment

There are many cases where the number of DF practitioners varied but the total volume of evidence items per case did not. To understand this, cases with volume of 2048 GB were selected from the year 2013 and separately cases with a total volume of 20480 GB were selected from the year 2014. There were 47 such cases found in 2013 and 7 such cases found in 2014. By analyzing the records, it was identified that the cases with identical volume took less total hours with higher numbers of forensic practitioners than fewer forensic practitioners. The selected records were analyzed in terms of case priority where the total volume is fixed and the priority of the cases varied. This experiment shows that the cases with high priority are more often assigned to more practitioners than the cases with normal priority. Hence, it is likely that if two cases are received with the same specifications but different priorities, a higher number of forensic examiners will be assigned to the high priority cases compared to the normal priority cases.

6.3 Practitioners' Experience Experiment

In this experiment the examiners were divided in three groups depending on their experience: (novice) less than 3 years of experience, (proficient) 3-7 years of experience, and (expert) more 7 years of experience. Cases with similar volume (512 GB) with one examiner working per case were selected from one specific year (2013) resulting in the selection of 256 records. The Pearson Correlation equals -.233. This shows that the strength of association is small. However, the analysis of these records also showed that there was evidence that novice examiners spent more time than the other levels of examiners, and that expert examiners spent the least time in investigations.

6.4 Person-Hours vs. Case Details Experiment

It is well known among digital forensic practitioners that the amount of details that comes in the case request to describe what is required from the examiner to search for affects the person-hours. It is assumed that the cases with more details and specifications could be investigated faster than cases with generic or little information. This experiment sought to understand the effect of case details variable on person-hours. A case details value was incorporated into the dataset, which could be either 'specific' or 'general'. *Specific* indicated that the case had provided search keywords and/or details like asking for the existence of a specific type of file in the hard drive or the case request provided the forensic examiner with personal details of the suspects. *General*

indicated that the case had no request details like extracting all the personal information for the suspect from the hard drive without specifying the file type or kind of information looking for.

For this experiment 75 records were selected for cases meeting the following requirements: total volume of 512 GB, 1 evidence item, 1 examiner, fraud case, and received in 2014. From the analysis of these records it was found that the assumption of specific details' influence in investigation time is most likely true. If the case request comes with more specifications, the examiner could target the required evidence from the investigated device. However, if only generic information is provided, then the DF examiner will spend more time to extract all evidence that might relate to the case.

7. DISCUSSION SUMMARY

Using real data from an active DF department, this study evaluates the relationship between different factors that are thought to impact DF investigation person-hours. Unexpectedly, the number of person-hours is not found to have a strong relationship with the years, volume and evidence items. However, from the descriptive statistics and the analyses conducted, combinations of factors were found that have an effect on the person-hours spent conducting a DF investigation case.

First, there was a significant increase in the number of cases through the years, especially from 2010 to 2014. Whereas it was expected that there would be a strong relationship between the total time of investigation and years, the results indicate the opposite. Interestingly, the descriptive analysis at the beginning of the study indicated that the average number of person-hours per case increased over the years. Furthermore, in the year 2014 the mean total person-hours increased sharply to reach 198 hours per case. Therefore, we can conclude that the majority of the cases spent similar person-hours over the years. Thus, the Pearson Correlation between the time and year is not affected but there is an increase in the total person-hours spent in number of cases over the years.

It was expected that the increase in the total storage data volume per case would lead to an increase in the number of person-hours and vice versa. While the descriptive statistics show a dramatic increase especially in the last four years. In contrast, total volume per case does not affect the total time of investigation. This means that even cases with small volume might take as much time as cases with high volume due to several factors. Moreover, this study illustrates the relationship between the number of examiners and volume, the load of evidence items per examiner and total number of examiners and total number of evidence items per case. The relation between the number of examiners and the total volume per case is also not strong. However, the total number of cases each DF practitioner needs to examine at the same time has increased over the years. A strong relationship between the total number of examiners and the total number of evidence items per case is shown. This means that the case distribution among the DF practitioners relies on the number of items per case where more items leads to a higher number of examiners being assigned. However, it is more convenient to make the decision on variables, volume and number of evidence items, to be able to reduce the amount of time examiners spent on the cases with high volume. We can conclude that the pressure of cases, which leads to an increase in the total volume each examiner is asked to investigate in the same time, is one of the factors behind the delay of investigation.

It was assumed that both the number of evidence items per case and the total number of evidence types per case would not affect the total time of investigations. There is no noticeable difference between the values of those variables through the years. From this we can conclude that both total number of evidence items per case and total number of evidence types per case are not considered to be factors in the delay of investigations.

There are several observations noted in the analyses. The analysis reported in section 6.1 examined selected cases with similar person-hours but with two volume sizes. It found that there are several factors behind this circumstance like workstation specifications, DF tools version and DF practitioner's experience. Thus, absence of improvements to those factors might lead to the delay in investigation. In section 6.2, the selected cases had similar volumes and were from a specific period of time in order to check the effect of number of examiners per case. This study found that the cases with higher numbers of examiners spent less time than the cases with lower numbers of examiners. In section 6.3 the examiner's experience is tested. It was shown that there is no significant impact of experience on the total time of investigation. Section 6.4 examined how the amount of information, which comes with the case request, affected the person-hours. The study demonstrated that it is most likely to take less time if enough request details are provided for the DF practitioners assigned to the case.

This research uses empirical data from the DF Department of the Dubai Police. The data is relatively unique in the DF field since the amount of data used allows robust and accurate results.

8. CONCLUSION

This paper reported on the analysis of 12 years of archived cases investigated by the DF department of the Dubai Police between 2003 and 2014. The study showed that there is no single factor which affects the time of investigation. Thus, combinations of many factors correlated cause delays in investigation. The analyses which were conducted by selecting cases where they met certain specifications to find out the most effective factor of DF investigation delay illustrated several interesting results which will be studied further in the course of this research. It will be interesting to complete the study by deeply examining selected cases to check how much volume the examiners receive in real cases and measure the volume they actually examine out of the total volume received. Future work needs also to focus on complexity of analyzing evidence, and on recommendations to reduce the person-hours and, therefore, improve the efficiency of law enforcement DF departments.

ACKNOWLEDGEMENT

I would like to express my special thanks of gratitude to my sponsor the Dubai Police as well as my work colleagues who gave me the golden opportunity to complete this project and achieve my goals.

REFERENCES

Irons, A., & Lallie, H. S. (2014). Digital Forensics to Intelligent Forensics. *Future Internet, 6*(3), 584-596.

Gogolin, G. (2010). The Digital Crime Tsunami. *Digital Investigation, 7*(1–2), 3-8. doi: http://dx.doi.org/10.1016/j.diin.2010.07.001

Dezfoli, F. N., Dehghantanha, A., Mahmoud, R., Sani, N. F. B. M., & Daryabar, F. (2013). Digital Forensic Trends and Future. *International Journal of Cyber-Security and Digital Forensics (IJCSDF), 2*(2), 48-76.

Roy, M. B. (2014). *An analysis of the applicability of federal law regarding hash-based searches of digital media.* Monterey, California: Naval Postgraduate School.

James, J. I. (2014). Multi-Stakeholder Case Prioritization in Digital Investigations. *Journal of Digital Forensics, Security and Law, 9*(2), 59-72.

Shaw, A., & Browne, A. (2013). A practical and robust approach to coping with large volumes of data submitted for digital forensic examination. Digital Investigation, *10*(2), 116-128.

Prefuse. (2013). the prefuse visualization toolkit. from http://prefuse.org/

Roussev, V., Richard, G. Breaking the Performance Wall: The Case for Distributed Digital Forensics. In Proceedings of the 2004 Digital Forensics Research Workshop (DFRWS). Aug 2004, Baltimore, MD.

Nirkhi, S. M., Dharaskar, R., & Thakre, V. (2012). Data Mining: A Prospective Approach for Digital Forensics. *International Journal of Data Mining & Knowledge Management Process, 2*(6), 45.

Kim, C. (1996). Cook's distance in spline smoothing. *Statistics & probability letters, 31*(2), 139-144.

