Chapter 5
Digital Forensic Data and Open Source Intelligence (DFINT+OSINT)

This chapter focuses on the externally sourced data aspect of the framework, and explores a process of data mining to extract entity information and a process of fusion with external source data to improve the knowledge discovery potential and intelligence from digital forensic data holdings.

The DRbSI and Quick Analysis process (Vol. 1 Chap. 4 and Vol. 2 Chap. 2) have potential applications for criminal intelligence purposes. In research testing, the ability to use DRbSI subsets and the Quick Analysis process with Bulk Extractor software (Garfinkel 2013) was explored to gain a rapid understanding of digital forensic data holdings. In this chapter, a process of value-adding to extracted data by drawing on OSINT resources is explored. This allows practitioners and intelligence personnel to add value to the data and information; thus, enabling a better understanding of the criminal environment.

The contributions of this chapter are:

- a process of semi-automated scanning of multiple digital forensic data subsets from a variety of devices, including computers, hard drives, mobile phones, portable storage, cloud storage, and Internet-of-Things (IoT) data, with a view to extract entity information; and
- value-add to the entity information by drawing on the resources of OSINT with a view to expanding cross-device and cross-case analysis, leading to improved overall knowledge relating to disparate cases.

In the next section, the background and related work for digital forensic intelligence (DFINT) and open source intelligence is outlined. The process of using DRbSI subsets in conjunction with semi-automated analysis to enable entity extraction and open source information searching is then discussed. Following this, the process of DFINT+OSINT analysis is applied to M57 test data to enable an understanding of its application, and then explore the application of the methodology to real-world data. The final sections discuss the research findings.

Material presented in this chapter is based on the following publication:
Quick, D. and K.-K.R. Choo, Digital Forensic Intelligence: Data Subsets and Open Source Intelligence (DFINT+OSINT): a Timely and Cohesive Mix, Future Generation Computer Systems, ISSN 0167-739X, http://dx.doi.org/10.1016/j.future.2016.12.032.

D. Quick and K.-K. R. Choo, *Big Digital Forensic Data*, SpringerBriefs on Cyber Security Systems and Networks, https://doi.org/10.1007/978-981-13-0263-3_5

5.1 The Role of Intelligence

The role of intelligence is to provide decision makers with independent and impartial information which is timely, accurate, relevant, verifiable, answers a question, and enables proactive decision-making (Gibson 2004). Criminal intelligence analysis is a term used to describe how information and intelligence can be used in the investigation of crime and persons involved or suspected of being involved in crime (UNODC 2011).

5.1.1 Intelligence-Led Policing (ILP)

The process of intelligence analysis is well entrenched within enforcement and other agencies, and includes a concept of "Intelligence led policing" (ILP), which is defined as;

> *the application of criminal intelligence analysis as an objective decision-making tool in order to facilitate crime reduction and prevention through effective policing strategies and external partnership projects drawn from an evidential base.* (Ratcliffe 2008).

ILP focuses on four elements, namely:

- targeting offenders,
- management of crime hotspots,
- linked crimes and incidents, and
- preventative measures.

Organised crime groups and terrorist organisations often consist of members with antecedents for the predilection of crime, and members often associate with persons with similar criminal background, but also draw on new recruits with limited criminal background. Organised crime was previously associated with the Cosa Nostra, but is nowadays quite different (UNODC 2011). Organised criminal groups now can have well-developed organizational structures, mainly established for obtaining power and/or wealth. Such groups include outlaw motorcycle gangs, Russian organized crime, Asian organized crime, African organized crime, drug cartels, and street gangs, such as; Asian, Korean, Hispanic, black, and white supremacy groups (UNODC 2011). It is reported that the complexity of these groups is increasing, with fluid structure-less networks, and increasing cooperation between different organized crime groups and networks (Choo 2008; Choo and Smith 2008; UNODC 2011).

Organised crime involvement now extends to; trafficking in human beings, drug trafficking, extortion, fraud, murder, and high-technology crime, facilitated with the growth in Internet resources, opening new opportunities for profit. The National Organised Crime Response Plan 2015–2018 of the Australian Government estimates that organised crime costs Australia $15 billion per annum in transnational crime, money laundering, identity crime, and the growth in technology facilitates this (Government 2015). The United Nations Office on Drugs and Crime (UNODC)

has reported an "escalation of high-technology crime is a challenging and relatively new arena for law enforcement" (UNODC 2011, p. 8), and that organised crime groups are more sophisticated and dynamic than ever before (UNODC 2011). The UNODC clearly states that "the challenge for law enforcement is to be prepared for this increasing sophistication in order to reduce the impact of criminal activities on our communities" (UNODC 2011, p. 8). With the increasing volume, variety, velocity, and veracity of digital forensic data, there is an opportunity to focus and learn from the information contained within this data, with an appropriate method to process the data in a timely manner.

The aim of a criminal intelligence analyst is to gather information in relation to criminals and criminal enterprises, and prevent or disrupt crime and criminal activity. The focus should, therefore, be to develop useful sources of information, including details of associates and relationships between individuals and their role in a criminal enterprise (UNODC 2011). The pervasive nature of ICT has resulted in a vast amount of information on devices, and data transiting from devices via the Internet to be stored in cloud storage environments. The role of digital forensics is to identify, collect, and preserve digital data which may provide assistance in legal enquiries. This is not limited to evidence alone, and there is a role for digital forensics in criminal intelligence analysis, supported with open source and external source intelligence and information.

5.1.2 Open Source Intelligence

OSINT has been utilised by many agencies and contributes to strategic, operational, tactical, and technical intelligence needs (Gibson 2004). OSINT is potentially a cost effective and rapid source of information, and the information and intelligence derived can potentially be shared (Best 2008). OSINT involves information extraction from publicly available sources (e.g. social networking sites). The Internet is now a major source of information, with estimates that data volume will grow from 4.4 zettabytes (ZB) in 2013 to 44 ZB by 2020, doubling in size every two years (IDC 2014a). The digitalised society we are witnessing has led to the globalisation of commerce, and opportunities for the globalisation of crime. This is enabled by the ability to easily travel across borders and transit international distances with relative ease.

This growth in data requires software which can provide for rapid content discovery, search, and retrieval. The Internet itself is not a source, but the means to access information sources, which can include media monitoring services, specialist information sources, 'grey literature' such as academic papers, and satellite imagery (Gibson 2004). Care must be taken with OSINT as information in the public domain is not necessarily verified and may be biased or inaccurate (UNODC 2011). Identification of sources of information is an ongoing process, as different means of communication go through a cycle of popularity. Search engines such as Google, which enable searching web sources by crawling and indexing hosted content, are

slow, but effective. Web mining tools that focus on specific sources can provide alerts to keyword terms when changes are made, or matches are located.

OSINT can be fast, flexible, dynamic, communicable, shareable, partner forming, can encompass rapid evaluation or in-depth analysis at strategic, operational, tactical, and technical levels, and identify and mitigate risk, i.e. from 'horizon scanning' to sophisticated targeting (Gibson 2004). One challenge with OSINT is the data source language and the need to translate the language, which may necessitate the use of a translation service, training for analysts, or the use of translation software (UNODC 2011).

Today's unparalleled access to global satellite data, coupled with Google street view pictures of a large percentage of populated environments, has provided for vast amounts of data and information. As this information is not classified, there are fewer barriers to information sharing, although the intelligence derived from OSINT may need to be classified. It is quite clear that data available via the Internet is potentially and currently the greatest data source ever available, and with the rapid rate that data is doubling, this is anticipated to only increase in future. The vast amount of data and information can assist with providing actionable intelligence to decision makers.

It is apparent that in today's connected world, with a vast volume of data a mere click of a mouse away, that criminals and organised crime groups will utilise new and sophisticated methods of communication. Hence, agencies tasked to investigate these groups have a need to gather information about activities undertaken in the digital realm, or with a digital footprint. Furthermore, with the growing number and storage capacity of mobile phones, computers, and portable devices increasing dramatically, there is a vast pool of information in relation to criminal activity.

The FBI has applied the process of criminal profiling to the digital offending realm, developing cyber-criminal profiles (Rogers 2003). However, this may not necessarily take into account the information available from victims' computer and devices, or the potential to develop psychological profiles of a wider range of crimes using the digital intelligence available from computers and devices, which when analysed can reveal a lot of information about the user.

A process of forensic feature extraction and cross drive analysis was first proposed in (Garfinkel 2006) and has been refined over the years to develop software to assist with processing and extracting entity information from digital media storage. The next step of this is to build on the extracted entity information, and again, an automated process is necessary, given the sheer volume of data and entities resulting from even small hard drive storage. Research in relation to the development of a method to value-add to the entity and information extracted is necessary, as is a method which is applicable in the current environment involving real-world cases and real-world data volume.

Beebe (2009) emphasised the need for research in crime network identification. The importance of applying intelligence analysis techniques to digital investigations was also discussed in Chap. 3 and this chapter continues this research focus by applying link analysis methods to digital forensic data and expanding on this with open source and external source information.

In Weiser et al. (2006) a repository of information is proposed to store information relating to digital forensic cases, to build a knowledge base of cases including a case tracking system that stores forensic discoveries related to a case, an expert system record of best practice, and certified tools index. This aspect of digital forensic intelligence appears to relate to a higher level focus to build a knowledge base, rather than a focus on the use of intelligence analysis techniques to assist with the process of digital forensic analysis.

In this research, the focus is on the merging of digital forensic analysis and intelligence analysis techniques to add value to the process of both digital forensics and intelligence analysis. A national repository knowledge base for digital forensic practitioners is an admirable goal, and with the appropriate security and controls, could be expanded to include intelligence extracted from digital forensic data, with the adoption of a suitable method of data reduction, such as that proposed in Vol. 1 Chap. 4.

5.1.3 Digital Forensic Intelligence + OSINT

With the vast quantity of data available from digital forensic analysis, there is a wealth of information which can be potentially improved with open source information to enable a better understanding of events or persons, and improve decision making opportunities. One issue affecting rapid and timely analysis of large volumes of digital forensic data is the growth in media volume and the associated increase in the time to undertake searches and review information. The proposed data reduction method in Vol. 1 Chap. 4 has enabled faster collection and processing of digital forensic data. This process of rapid extraction and analysis enables the processing of large volumes of digital forensic data in a timely manner.

In the next section, the proposed framework is outlined, with a goal of enabling the inclusion of open source information with digital forensic analysis to encompass building on the data that is extracted through digital forensic analysis, such as computer hard drives, mobile phones, tablets, media storage devices, IoT devices, and cloud stored data. Using the Data Reduction by Selective Imaging (DRbSI) data reduction process (Vol. 1 Chap. 4) (Quick and Choo 2016), and semi-automated entity information extraction (using Bulk Extractor software), a process of fusion with open source information is explored. The aim is to expand knowledge and intelligence from large volumes of data by a process of;

- digital forensic data reduction,
- semi-automated entity extraction, and
- external source OSINT searches.

In the proposed process for digital forensic and open source analysis, the information from a wide scope of devices and information storage is enhanced with open source intelligence. This is to enable those involved in an investigation or intelligence probe to make decisions with as much relevant knowledge as is available, in a timely

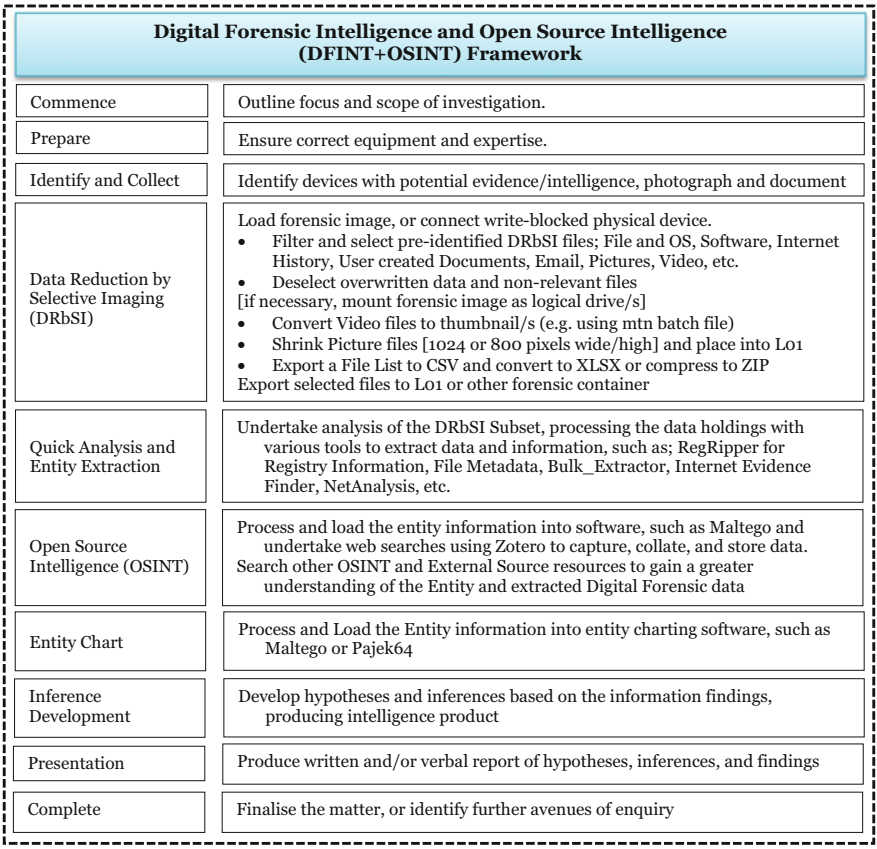

Fig. 5.1 Digital forensic intelligence and OSINT framework (Step 7)

manner using the proposed digital forensic intelligence and open source information framework.

5.2 DFINT+OSINT Method

The proposed framework for the process of digital forensic intelligence and open source intelligence (DFINT+OSINT) is based on the digital forensic intelligence analysis cycle (Chap. 2) and the digital forensic data reduction methodology (Vol. 1 Chap. 4). The proposed DFINT+OSINT framework (Fig. 5.1) is outlined in this section.

When working with open source data for intelligence purposes, consideration should be made regarding the following points for each source: authority, accuracy, objectivity, currency, and coverage (Gibson 2004). The use of a rating system for

intelligence and sources is also appropriate, such as the 4 × 4 system or 6 × 6 system (admiralty scale) (UNODC 2011).

As with any investigation or intelligence probe, timely and accurate notes of the steps undertaken should be maintained, and any information and evidence secured according to agency directions or industry best-practice (ACPO 2006; NIJ 2004, 2008) Legislation and legal authority must be checked and appropriate prior to commencing, and ensure compliance at all stages.

When working online, practitioners are advised to ensure identity protection measures are appropriate, and network security is paramount when dealing with source access via open Internet connections. Covert operations need to ensure the appropriate approvals and security is in place prior to commencing, as it potentially only takes one miss-step to jeopardise an operation or personnel involved, which in some cases may be life-threatening consequences.

Digital forensic analysis follows a well-established framework, namely: identification, preservation, analysis, presentation (McKemmish 1999). Intelligence analysis involves a similar process of collection, collation, analysis, and presentation. The Digital Forensic Intelligence Analysis Cycle (DFIAC) is a merger of the methodologies of digital forensics and intelligence analysis to form a process of Commence, Prepare, Evaluate and Identify, Collect, Preserve, Collate, Analyse, Inference Development, Presentation, Completion or Further Tasks Identified (Chap. 4).

Using the DFIAC to frame the process, a process of DFINT+OSINT is outlined as a sub-cycle of a wider Investigation or Intelligence probe, as follows:

Commence (Scope/Tasking): the focus, aims, and scope of analysis are outlined to enable preparation and guidance for the overall examination.

Prepare: gather the anticipated equipment required and expertise, including confirmation of legal authority, network security, and covert considerations.

Identify and Collect: identify devices with potential evidence and/or intelligence, and undertake appropriate physical examination and documentation, according to agency policy and procedures.

Data Reduction by Selective Imaging (DRbSI): collect a digital forensic subset of data from the identified device or media. This involves utilising the process outlined in Vol. 1 Chap. 4, i.e. connect the identified media, then run a filter for pre-identified files and data. If necessary, reduce the dimension of pictures, and thumbnail video data. Export a list of all files on the media, and export all this into a logical forensic container (e.g. L01, AD1, or CTR).

Quick Analysis and Entity Extraction: using the process of Quick Analysis and Entity Extraction as outlined in Chap. 2. The entity extraction process involves using software to process the various data types retained within the DRbSI subset, and merge the output into a single source of information. The use of software, such as Bulk Extractor, assists with extracting a range of key entity information.

OSINT: identify sources of data and potential evidence and intelligence, such as: data which may be on social media websites, global or local media reports, 'grey' literature, satellite images, Google Street view information, or other open source media. Once data or information source is identified, conduct searches for known relevant entities; names, addresses, email addresses, phone numbers, vehicle details,

and other information pointers. If relevant data is located, then it should be forensically handled where possible, i.e. printed to PDF or use screen capture (Microsoft Windows Snipping Tool) or screen capture software. Further in-depth analysis of websites may be necessary, and may locate additional information not normally presented publicly, i.e. within HTML source code for websites. The preservation process can align with a data collection and reduction process, such as that outlined in Vol. 1 Chap. 4 to collate a subset of relevant data. The use of software, such as Maltego, may assist with bulk data analysis and retrieval.

Entity Chart: the identified and collected OSINT data is then included with other data subsets and data extracted from computers, mobile phones, Internet of Things (IoT) devices, including data from cloud stored providers, for examination for evidence or intelligence, with a focus of that of the scope of the task. If during the process of analysis, additional sources of data are identified (e.g. other social media websites, media reports, or cloud stored data), the process forks to the 'Preparation' stage to collect the new data, whilst the analysis progresses.

Inference Development: with the knowledge gained during the process of collation and analysis, ideas are formed in relation to the questions of who, how, what, when, why, and where. The gained knowledge is used to form inferences about the investigation or intelligence probe to answer questions or outline findings. The intelligence and source data should be rated using a rating system, such as the 4×4 system or 6×6 systems (UNODC 2011).

Presentation: the findings of the overall process of analysis are formed into a report (written and/or verbal) which is communicated to the requesting persons involved in the investigation, legal process or probe.

Complete: the matter is finalised, and data archived according to agency practice and procedures. If further tasks are identified, the process continues in the cycle until complete. As part of the completion process, feedback should be sought from the persons involved to ensure the scope of the task has been met, and also practitioners should provide feedback to those involved in the task to ensure they are aware of the findings, and/or their role in the overall process.

5.3 Results: Digital Intelligence and OSINT from M57 Test Data

The growth in the volume and number of devices encountered in investigations has resulted in a need to collect and analyse growing volumes of digital forensic data in a variety of formats. The test data from the M57 corpus (Garfinkel et al. 2009) as previously used for digital forensic data reduction research, and successfully reduced the volume of data to 0.206% using the DRbSI process (Vol. 1 Chap. 4). A process of Quick Analysis was then outlined in this Volume to distil relevant information from the digital forensic data subsets. Subsequent to this, the digital forensic data from the test data (M57) computers, portable storage, mobile phones, and tablet devices, and

source data comprising approximately 498 GB was merged. This was successfully reduced to 4.25 GB of DRbSI subsets and logical container files encompassing potentially relevant information.

A bulk data analysis process was applied to the data subsets, using semi-automated entity extraction from the forensic subsets using Bulk Extractor 1.5.5 software (Garfinkel 2013). This scanned the DRbSI data subsets, and the output encompassed 2.02 GB, comprising 23,496 email features, and 22,962 picture files, in approximately 30 min. Mobile phone extracts from 41 mobile devices, comprising 207 MB, were merged into a single source file for analysis.

The data output from the Quick Analysis process which included parsing the Windows Registry files, Internet Evidence Finder, NetAnalysis software, and information extracted from a variety of data sources within the DRbSI subsets, was merged with the output from Bulk Extractor, along with the previously merged mobile phone extracted data, resulting in a very large file of extracted entity information with associated source and relationship links. This was loaded into Pajek64 software.

This process has encompassed the first five stages of the DFINT+OSINT framework (Fig. 5.1), and now moves to the next stage, by using the extracted single-source entity information with Maltego CE to explore the process of expanding knowledge of the persons and entities contained within the data by locating available OSINT relative to the M57 test data.

A guide to OSINT analysis from (Toddington_International 2016) is summarised as:

- Choose an effective search tool,
- Use extended search capabilities (GREP),
- Search deep web resources, such as; databases, electoral roles, telephone, business databases,
- Review LinkedIn, Facebook, Twitter, YouTube, Flikr, Instagram, PhotoBucket, web blogs, Tripod, online sales sites, such as; eBay, Gumtree, Craigslist, Whirlpool,
- Run WHOIS searches for domain names,
- IP addresses (extracted from Registry and Internet History), genealogy sites, maps, traceroute, wayback machine, review source HTML, trace emails,
- Search for names, usernames, account names, email, phone numbers, addresses, family members, friends, associates, image EXIF data, GEO DATA,
- Use Zotero to collect, collate, and store data as you go, and
- Use snipping tool or print to pdf.

Also, notes should be made of the searches conducted, including; keyword search terms used, which websites were examined, any email references, personal information, associates, online images, chat, social network, further avenues of enquiry, time and date, and importantly, legal authority to undertake the analysis.

The extracted entity information from the test data was loaded into Maltego CE, displayed in Fig. 5.2. This chart shows the interlinked nature of the disparate data sources and entities extracted from the M57 corpus DRbSI subsets.

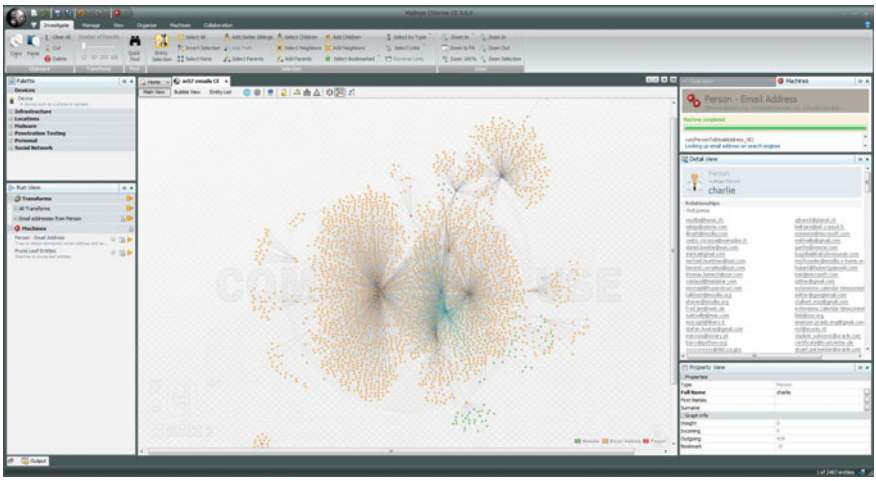

Fig. 5.2 Test Data loaded into Maltego CE

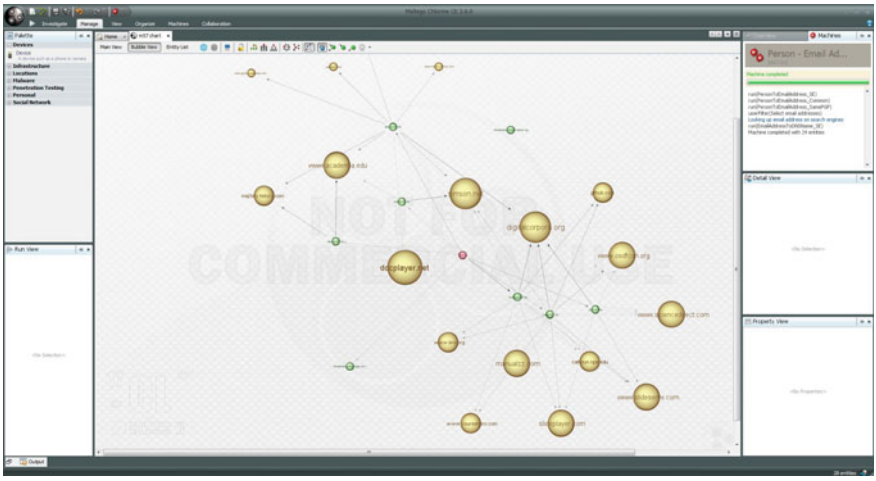

Fig. 5.3 OSINT URL references displayed in Maltego CE

Maltego Transform searches of the entities within the link chart were undertaken (Fig. 5.3). This resulted in a large amount of data matches, expanding knowledge in relation to the entities in the test data, and included URL locations with references to the email addresses and entities contained within the M57 corpus.

A selection of the URL matches located with OSINT is listed in Table 5.1. These show the type of external source data located which relates to entities contained within the extracted data from the DRbSI subsets. In a rapid and timely manner it was possible to add-value to the information in relation to the M57 data, expanding the knowledge-base with information available from open sources. Using this

Table 5.1 OSINT URL references located with Maltego CE

Name	URL
Jean	http://maj3sty.tistory.com/1034
Jean	http://maj3sty.tistory.com/category/%5B+%5D%20Forensic?page=8
Jean	http://www.doc88.com/p-1836912450568.html
Charlie	http://simson.net/ref/2012/2012-08-08%20bulk_extractor%20Tutorial.pdf
Alison	http://www.tuicool.com/articles/eiYNzuU
Jo	http://www.osdfcon.org/presentations/2015/McCarrin-Allen_osdfcon.pdf
Pat	http://digitalcorpora.org/downloads/bulk_extractor/BEUsersManual.pdf

process, digital forensic intelligence in conjunction with open source information has expanded knowledge. Whilst this increased information and knowledge-gain is of benefit in research, more importantly, this type of information and intelligence building can greatly assist in real world investigations.

5.4 Applying DFINT+OSINT to Real World Data

From the research outlined earlier, the entity information extracted from the M57 test data is similar to that which is extracted from real world data. The volume in real world data holdings is often much larger, which can result in longer search times. However, with more data there is better information potential.

There is also an opportunity to develop a method of refining the data and entities to that which relates to a case, and exclude or filter out generic data that exists in many operating systems and software installations, such as Microsoft Windows URL and email links for help and assistance. There is potentially a large volume of entities which can be excluded as these are unlikely to be related to an investigation. Undertaking a bulk extraction of a newly installed operating system and software and using this as a source of 'known-good' entities, which can then be used to remove these 'known-good' entities from real world data, would enable a further reduction in the number of entities for OSINT research. The use of the National Software Reference Library (NSRL) hash databases is also of potential benefit to further reduce the volume of entities extracted.

With real world data, data from IoT devices such as fitness bands which record GPS locations of the wearer may be relevant, and can include instances where it is uploaded to cloud storage. Security systems such as smart door locks which record biometric information as a person enters or exits a smart home could also be relevant, as is wireless internet connected doorbell systems with video recording capability

which record movement of persons near the device. Data from these systems can be extracted and merged with the digital forensic data for analysis, and provides for even more information for investigators and analysts to examine. Hence, a process of data reduction, quick analysis, and external source intelligence can be beneficial to those involved in an investigation to understand the context of the information available.

By adding value to the information contained within digital forensic data, there is an opportunity to explore cross-case intelligence analysis, which in real-world cases may highlight cross-case linkages which were previously unknown to investigators. Indeed, the first author has experience where case linkages were unknown to disparate investigations due to the different focus of the various investigations, i.e. drug importation investigations and local service area property crime offending, and when the cross-case linkages have been brought to the attention of the separated investigation teams, this has enabled a better understanding of the volume of disparate (but actually connected) offending.

By implementing a method of building a knowledge base of cases, such as that proposed by Weiser et al. (2006), there is an opportunity to assist with disparate cross-case linkages being discovered early enough in an investigation to ensure appropriate resourcing of investigations, with a potential for a more timely resolution of cases. This must be balanced with appropriate security of the information, and legal authority to access and review the data.

In the previous research, the metadata contained within archived cases from South Australia Police Electronic Evidence Section data backup archives was examined (Volume 1 and 2). The content of case data was not examined, and only reviewed the times and data from processing of limited archived meta-data.

One archived case examined comprised eighteen computers, laptops, portable storage, mobile phones, and tablet devices, totalling 2.7 TB of source data. The volume of data was reduced according to the DRbSI process outlined in Vol. 1 Chap. 4, resulting in 46.1 GB of DRbSI subsets. Full imaging took approximately 42 h, and the DRbSI process took less than 4 h. To test the use of a semi-automated analysis process, the subsets were batch processed with Bulk Extractor 1.5.5 software, processing in 1 h 19 min. In comparison, processing the original source data took 43 h and 11 min. As per the process undertaken with the M57 data, it would be possible to merge the output entity information for further analysis with Maltego and adding value to the information with OSINT data analysis (this process was not undertaken with this data).

In addition, the DRbSI subsets from 544 archived devices were loaded into NUIX 6.2.3, EnCase 6.19.7, and EnCase 7.10.5. Again, the data was not viewed, rather, the times for processing was noted. EnCase 6.19.7 took approximately 3 min to load and open the 544 L01 files. File signature analysis was run, and took 2 h and 8 min. Over 10 million files were presented for analysis, including 907,015 documents, 52,742 emails, 2,221,521 picture files, and 2,333 container files. Within this data was potentially relevant entity information which could be improved with OSINT data.

The subsets when loaded into NUIX 6.2.3, provided for metadata analysis, which included strategic intelligence analysis to identify the types of phones presented for analysis, identifying that mainly iPhone and Samsung mobile phones were present. This highlighted that research into the device storage of these devices is warranted as they appear to be quite popular in relation to other devices. Further device specific information was not viewed. In addition, a focus on the JPEG EXIF metadata highlighted that Panasonic, Nikon, and Canon camera identifiers were present (further analysis was not undertaken on this data).

This research demonstrated a capability to process DRbSI subset L01 files from real-world devices ranging from USB storage to multi-terabyte hard drives, and it was possible to load and process these with EnCase 6.19.7, EnCase 7.10.5 and NUIX 6.2.3, and a potential to conduct further analysis of the data using Bulk Extractor across the subsets which could be further enhanced with Maltego to locate any open source information relating to the entities. Further analysis of the data was not undertaken due to privacy considerations.

5.5 Discussion

As outlined, the literature review highlighted a need for further research in the use of intelligence analysis techniques with digital forensic data. A corpus of test data was collected to undertake experiments, and from this it was possible to determine a method to undertake in-depth analysis to value-add to the entity information present in digital forensic case data. This is encompassed in the proposed DFINT+OSINT framework. Aspects of this framework were applied to real world data, which indicated the potential application to real world cases. In this manner, intelligence from DFINT+OSINT data has been enhanced.

Criminal intelligence is described as a national asset, which should be; "collected once and used often for the benefit of many and therefore adds value to the decision-making process" (Australia 2013, p. 62). This principle is also applicable to digital forensic data, and the analysis of disparate case data can have benefits to society in solving and progressing cases in a timely manner. As IoT devices, computers, portable storage, mobile phones, and tablet devices become more pervasive throughout society, there will be a growing need for forensic analysis of these devices. With the growing volume of disparate data, there is a need to be able to undertake analysis on growing volumes of structured and unstructured data. The method outlined in the previous sections as applied to test data (M57) and real world data, demonstrated an ability to undertake analysis of a large volume of disparate data, and locate potential evidence and intelligence.

In experiments, it was possible to scan data-reduced subsets in a semi-automated manner, and then merge the output to enable the examination of a large volume of data in a timely manner for linkages across devices and cases. As more and more devices are seized and presented for digital forensic analysis, there will be a larger source of data for criminal intelligence analysis, potentially locating evidence and intelligence

to enable investigators and decision makers a better understanding of events from large volumes of data. Strategic and management level information can be drawn from digital forensic data; operational knowledge can be located and provided to investigators and managers, including information relating to crime trends. Tactical, target specific information can also be located and communicated in a timely manner.

By enhancing the information contained within digital forensic data by undertaking semi-automated analysis of entity information with open source data sources and information, there is an opportunity to fast-track investigations, and locate disparate linkages, which may otherwise remain unknown.

5.6 Summary

This chapter explored the aspect of external source data, which addresses Step 7 of the framework (Fig. 1.1). The process enables a fusion of open source (via the Internet) and closed and confidential source (digital forensic subsets) data sources. The use of semi-automated data mining software and external source collection software enabled the rapid processing of disparate case data to improve the knowledge discovery and intelligence potential of digital forensic data holdings.

References

All URLs were last accessed (and correct) on 5 November 2016

ACPO (2006). *Good practice guidelines for computer based evidence v4.0*, Association of Chief Police Officers. Retrieved March 5, 2014, from www.7safe.com/electronic_evidence.

Australia Co (2013). *Parliamentary joint committee on law enforcement inquiry into the gathering and use of criminal intelligence*, Canberra.

Best, C. (2008). Open source intelligence. *Mining massive data sets for security: Advances in data mining, search, social networks and text mining, and their applications to security, 19,* 331–344.

Beebe, N. (2009). Digital forensic research: The good, the bad and the unaddressed. In *Advances in digital forensics* (pp. 17–36). Springer.

Choo, K.-K. R. (2008). Organised crime groups in cyberspace: a typology. *Trends in organized crime, 11*(3), 270–295.

Choo, K.-K. R., & Smith, R. G. (2008). Criminal exploitation of online systems by organised crime groups. *Asian Journal of Criminology, 3*(1), 37–59.

Garfinkel, S. (2006). Forensic feature extraction and cross-drive analysis. *Digital Investigation*, *3*, Supplement, no. 0, 71–81.

Garfinkel, S., Farrell, P., Roussev, V., & Dinolt, G. (2009). Bringing science to digital forensics with standardized forensic corpora. In *DFRWS 2009, Montreal, Canada*. Retrieved September 9, 2009, from http://digitalcorpora.org/corpora/disk-images.

Garfinkel, S. (2013). Digital media triage with bulk data analysis and bulk_extractor. *Computers and Security, 32,* 56–72.

Gibson, S. (2004). Open source intelligence: An intelligence lifeline. *The RUSI Journal, 149*(1), 16–22.

Australian Government (2015). *National organised crime response plan 2015–2018*, Australia. https://www.ag.gov.au/CrimeAndCorruption/OrganisedCrime/Documents/NationalOrganisedCrimeResponsePlan2015-18.pdf.

IDC (2014a). *The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things*, EMC Corporation. Retrieved June 1, 2014, from http://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm.

McKemmish, R. (1999). *What is forensic computing?*

NIJ (2004) *Forensic examination of digital evidence: A guide for law enforcement*. http://nij.gov/nij/pubs-sum/199408.htm.

NIJ (2008). Electronic crime scene investigation: A guide for first responders, Second Edition. http://www.nij.gov/pubs-sum/219941.htm.

Quick, D., & Choo, K.-K. R. (2016). Big forensic data reduction: digital forensic images and electronic evidence. *Cluster Computing, 19*(2), 723–740.

Ratcliffe, J. (2008). Intelligence-led policing. *Trends and Issues in Crime and Criminal Justice*. Australian Institute of Criminology.

Rogers, M. (2003). The role of criminal profiling in the computer forensics process. *Computers and Security, 22*(4), 292–298.

Toddington_International (2016). *Online investigator's checklist*, Toddington International Inc. Retrieved July 7, 2016, from https://1x7meb3bmahktmrx39tuiync-wpengine.netdna-ssl.com/wp-content/uploads/TII_Online-Investigators-Checklist_v2-1.pdf.

UNODC. (2011). *United nations office on drugs and crime—Criminal intelligence manual for analysts*. New York, Vienna, Austria: United Nations.

Weiser, M., Biros, D. P., & Mosier, G. (2006). Development of a national repository of digital forensic intelligence. In *Proceedings of the conference on digital forensics, security and law*.