5

The transforming grid of digital forensics to intelligent forensics – relook into the applicability of artificial intelligence in current investigation techniques

Parvathi S Shaji

Department of Law, University of Kerala, India
parvathishaji11@gmail.com

Abstract. In this era of stepping into the 21st century, were each and every moment of human is getting converted to genetic information, transactions in all spheres are going paperless, e-commerce is becoming common man's necessity, digital devices acquiring its forms in multifarious profiles becoming part and parcel of human existence itself. The data is attaining the status of a valuable asset. On the other phase of the story, the concern of security and trust is elements under potential risk. The cyber criminals are getting brilliantly equipped with all the latest technologies beyond the imagination of a law enforcement agency or an investigator. Indeed, this is adding to the fear of any individual who is unknowingly becoming a prey for just can be a simple reason of purchasing a life essential commodity from an online site. The paper tries to examine paucity that exists in the traditional digital forensics' application techniques. With the interminable growth in the rate of cybercrime and at with a sophisticated intricacy of involvement of technology coupled in the nature of the crime committed trans-boundary, the law enforcement agencies are left strangled to conduct the digital forensics or investigation process precisely and, in a time, -bound manner. Indeed, the apparent inability of existing technologies and method adopted is acting as a laidback escape for cyber criminals. The paper enumerates on the crucial requirement for switching over to the application of artificial intelligence. Artificial intelligence which is composed of specialised intelligent agents that act exclusively based on the expert's knowledge of the technical domain. The prime goal line is with respect to analysing and correlating the data contained in the evidence of a specific case at hand and thereby with the utilisation of its expertise, presenting the most relevant evidence to the respective investigator. The element of accuracy and prompt results again enhances the discipline of digital forensics. The paper analysis on the different ambits, both legal and technological aspects involved in the transformation of the discipline of digital forensics to intelligent forensics. The major elements involved in the application of artificial intelligence techniques is through a development of multiagent system and case-based reasoning. The paper attempts to illustrate on the myriad concepts like processing and handling of digital evidence, utility of intelligent toolkit, network and cloud forensics, social network analysis, privacy related concerns for the acquisition of data from the virtual regime. The question as to whether the techniques with respect to artificial intelligence will be able to reduce the gap between the technology adopted by investigative law enforcement agencies and the ones used by the perpetrators and chase along even tapping up with the unpredictable criminal mindsets. The paper seeks to react whether the transformation address the challenges of the more; larger and more complex domains in which cybercrimes are taking place. The paper also tries to answer whether the application of artificial intelligent in the digital investigation technique can sort the challenges at myriad levels faced by the law enforcement agency or an investigator while handling the digital attack and by being technical equipped for any kind of harm caused by a criminal perpetrator.

Keywords: Digital Forensics, Artificial Intelligence, Cyber Crime, Cyber Security, Intelligent Forensics.

1 Introduction

As it is a matter of reality, that new machineries are emerging all the time in this innovative world of technological wonders. These techniques indeed are placed on record in the different methods such as data-on-demand ability of cloud technologies, the convenience of mobile platforms and other variant forms of digital gadgets. With the advancement of technological innovation in rapidity that's considerably incompatible with the technology at hand of digital investigators posing a serious challenging phase in whole discipline of digital investigation. The criminal perpetrators on the other side is increasingly using the newest advanced technology within the committing of crimes that too having novel and distinct characteristics. Also factors such as increasing magnitude of storage, multitude of data evidence sources and continual increases in computational power. Consequently, these are contributing to the rise within the backlog of digital mediums being left to be digitally investigated. Adding thereto due to transboundary concerns faced by the investigating authority and issues in reference to the location and acquisition of the digital evidence. The range of data sources again reaches its peak when an investigation involves social media resulting in storage concerns. The current traditional investigative technologies also step aback once they encounter with secure technologies such as with the advent of encryption, covering full disk encryption, secure network communication, secure processors and anonymous routing potentially resulting in making the situation more complex for the investigating officer to charter it down. With these series of issues, the necessity for the incorporation of the applicability of the artificial intelligence in digital forensics technique is a matter to be apprehended and analysed in switching over to newer investigative technology for adapting to the newest advances exhibited in the criminal use of technology. The transformation of traditional digital forensic technique to intelligent forensics is the need of the hour since the culprits are always peeping behind under the disguise of their technical intelligence in this virtual cyber space. The paper tries to examine the multilevel applicability of artificial intelligence in the domain of digital investigation techniques at the various process involved in the investigation procedure.

2 Understanding the Discipline of Digital Forensics

In the mid 1960's Donn Parker noticed the phenomenon that when people entered the computer centre, they left their ethics at the door (Terrel Bynum, 2001). On a simple note computer forensics has emerged out of the need to unravel, document and enable prosecution of computer crime. Further in the 1970s and 1980 relatively personal computers became common and individuals and businesses began to use them on a regular basis. Thus, subsequently law enforcement agencies noticed the emergence of a new class of crime i.e., computer related crime. The emergence of computer forensics was largely in response to a demand for service from the law. By the 1990s Law Enforcement Agencies (Hereinafter refereed as LEA's). in every technologically advanced country were aware of computer crime, and had a system in place to investigate and to prosecute such activities. Many research centres and scientific groups were also formed, and therefore the software industry began to offer various specialized tools to help in investigating computer crimes (Michael G.N, 2000).

For early investigators involved in computer related crimes it became immediately obvious that if their response and findings were to be of any use as court evidence they had to comply with the same rules as any other conventional investigations. The primary thing every investigator has to be aware of is Lockard's exchange Principle:

"Anyone or anything entering a crime scene takes something of the scene with them, or leaves something of themselves behind with they depart" (Richard, 2001)

Thus, it became clear that when investigating computer related crime, an equivalent basic rules applied as in during a non-computer related crime scene investigation. The investigation process includes phases of physical scene preservation, survey and reconstruction using collected evidence, all of which is formally documented (Ewa Huebner; Derek, Bem 2007). The first computer forensics training course appeared around 1989 at University of North Texas and the first International law Enforcement Conference on Computer Evidence was hosted in 1993 in Australia. With of these developments at hand, computer forensics became a unique discipline of science, and in many areas, it requires a special approach, different tools, as well as specialized education and training. The first period in computer forensics history is characterized by handling with relatively small capacity devices and a comparatively bit amount of information. Thus, paving way for the emergence of a novel discipline.

Technology is a double edged sword which will be utilised in economic sustainability, to aid in the arrest of cyber criminals etc., and there are various tools which will assist LEAs in investigating cyber-crime cases and in cyber-crime evidence collection, drafting and creating hard evidence, however an equivalent technology could also employed by cyber criminals to commit offences worse still the forensic tools could also be employed by these cyber criminals to hide their tracks for instance a criminal may use the disk wipers to clean the hard disks rendering forensic tools immobilized to recover evidence.(Virginiah S & Mohammad T, 2012).

National Institute of Standards and Technology (NIST) defines digital forensics as an applied science for "the identification, collection, examination, and analysis of data while preserving the integrity of the information and maintaining a strict chain of custody for the data" (Brian C, 2003).

There are major investigate contingents that arise the necessity for forensic techniques and tools. The following institutional frameworks play a significant role as far this discipline is concerned:

1. Law Enforcement – focuses on gathering evidence
2. Organizations, Business or e- commerce - for use in keeping the business on track using reasonably effective techniques and ensuring safe online purchasing.
3. Academia-ensures accuracy of result driven from precise, repeatable methods.
4. Prosecution - elaboration of the analysis during a court of law.
5. Judiciary- scrutinizing the findings against judicial standards.

Further understanding the other aspects, computer forensics is primarily concerned with the proper acquisition, preservation and analysis of digital evidence, typically after an unauthorized access or use has taken place. In broad terms, a forensics life cycle involves the following phases (Nina Godbole & Sunita Belapure, 2011).

Preparing for the evidence & identifying the evidence -When there exists an enormous amount of potential evidence available for a legal matter and it is also possible that the vast majority of the evidence may never get identified. In cases where there is in place a single computer or in case of networked pattern of systems, in the former case every sequence of events within a single computer leads to the interactions with files they produce and manage, and also with regard to log files and audit trails of various sorts and in case of latter it extends to all networked devices, potentially all over the world. Thus, definitely it's a matter of tedious task to prepare and identify the evidence.

Collecting and recording digital evidence-Digital evidence can be collected from many sources[1]. One of the most vital aspect is that special care must be taken when handling computer evidence as most digital evidence is easily changed, and once changed it is usually impossible to detect that a change has taken place unless other measures have been taken. Since such a kind of concern exist, the investigator calculates a cryptographic hash of an evidence file and to record that hash elsewhere, usually in an investigator's notebook, so that one can establish at a later point in time that the evidence has not been modified as the hash was calculated.

Storing and transporting digital evidence-In storage, digital media must be properly maintained for the period of time required for the purposes of trial. Depending on the particular media, this may involve any number of requirements ranging from temperature and humidity controls to the need to supply additional power or to read media. Storage must be adequately secure to assure proper chain of custody, and typically, for evidence areas containing large volumes of evidence, paperwork associated with all actions related to the evidence areas containing large volumes of evidence, paperwork associated with all actions related to the evidence must be kept to assure that evidence does not go anywhere without being properly traced. Evidence is often copied and sent electronically, on compact disks or on other media, from place to place. Original copies are normally kept in secure location to act because the original evidence that is introduced into the legal proceedings. Therefore, adequate care must be taken in transportation to prevent spoliation also.[2]

Examining or investigating digital evidence-As a general rule one should not examine digital evidence unless one has the legal authority to do so. Considering the aim of a digital evidence examination, "imaging of electronic media"[3] becomes necessary. During imaging process of electronic media, a write protection device or application is generally used to ensure that no information is introduced onto the evidentiary media during the forensic process. At crucial points throughout the analysis, the media is

[1] There include two kinds of sources: Obvious sources which includes computers, cell phones, digital cameras, hard drives, CD-ROM, USB memory devices and so on. On the hand Non-obvious sources include setting of digital thermometers, black boxes inside automobiles, RFID tags and webpages.

[2] For instance, in a hot car, digital media tends to lose bits.

[3] The process of creating an exact duplicate of the original evidentiary media is often called Imaging. Computer Forensics software packages make this possible by converting an entire hard drive into a single searchable file- this file is called an image. Using a stand- alone hard drive duplicator or software imaging tools such as DCFLdd, IXimager or Guymager, the entire hard drive is completely duplicated. This is usually done at the sector level, making a bit stream copy of every part of the user- accessible areas of the hard drive which can be physically store data, rather than duplicating the file system. Thereby the original drive is then removing to secure storage to prevent tampering.

verified again, referred to as “hashing”, in order to make sure that the evidence is still in its original state. (Nina Godbole & Sunita Belapure, 2011, P 346)

Analysis, interpretation & attribution: Analysis, interpretation and attribution of evidence are the foremost difficult aspects encountered in most forensics’ analysis. Within the digital forensics’ arena, there are usually exists only a finite number of possible event sequences that could have produced evidence. However, the actual number of possible sequences could also be almost unfathomably large. In essence, almost any execution of an instruction by the computing environment containing or generating the evidence may have an impact on the evidence. Basically, all digital evidence must be analysed to determine the type of data that is stored upon it.

Reporting- Once the analysis is complete, a report is generated. The report could also be in a written form or an oral testimony or it may be a combination of both. Finally, evidence, analysis, interpretation and attributions must in the end be presented in the form of expert reports, depositions and testimony. (Josaih Dykstra & Alan T. Sherman, 2012). The following are the broad elements of the report:

- Identifying of the reporting agency;
- Case identifier or submission number;
- Case investigator;
- Identity of the submitter;
- Date of receipt;
- Date of report;
- Descriptive list of items submitted for examination, including serial number, make and model;
- Identity and signature of the examiner
- Brief description of steps taken during examination, such as string searches, graphics image searches and recovery erased files
- Results or conclusions

Testifying-This phase involves presentation and cross examination of expert witnesses. Depending on the jurisdiction and legal frameworks in which a cybercrime is registered, certain standards may apply with reference to the issues of expert witnesses. Digital forensics evidence is generally introduced by expert witnesses except in cases where non- experts can bring clarity to non-scientific issues.

Thus, the chain of evidence and accuracy of digital evidence is extremely important in cyber forensics investigation. Therefore, experienced human investigators can often analyse crime trends precisely, but since the incidence and complexity of crime increase, human errors occur, analysis time increases and criminals have longer time to destroy evidence and escape arrest. By increasing efficiency and reducing errors, crime data mining techniques can facilitate police and enable investigators to allocate their time to other valuable tasks.

3 Understanding the Different Variants of Forensics - Cloud Forensics and Network Forensics

Cloud computing forensic science is the application of scientific principles, technological practices and derived and proven methods to reconstruct past cloud computing events through identification, collection, organization and reporting of digital evidence. In each step there are tools and techniques available. Traditional methods and tools of forensics cannot cope up with the cloud forensics due to the very fact that the retrieval of the information, the main lead of any case, is diversely located and hence difficult to succeed in. (Arjit Paul; Mayuri Kiran Anvekar; K. Chandra Sekaran, 2012). Cloud computing is predicated on extensive network access, and network forensics handles forensic investigation privately and public networks. However, cloud forensics also includes investigating file systems, process, cash, and registry history. Every data is vital for the investigation. So, within the collection phase, goal is to gather as much as data which has relevancy to the investigation (Prasad Purnayae, 2015).

The vital areas of concern with reference to investigation procedure in cloud platform is that the complexities a LEA face in the data acquisition procedure , the transboundary jurisdictional issues, concerns relating to the ownership of the cloud storage and geographic location and the varied problems in the data acquisition from different cloud system deployment models. Another major trouble maker is that of identifying and then subsequently imaging the data source. For instance, in a public cloud storage infrastructure which may possess a dozen of server or data sources located at different geographic locations against which the data may be dynamically routed and stored (Raun, Keyn, Joe Karby, Tahar Kechadi & Mark Crosbie, 2011).

The investigator or the concerned LEA has to recognize the precise locations of the data before being able to image the data, thus in itself a forensic challenge. For imaging large sets of data necessitates a novel approach to the technology and aiding mechanism for the investigators. With respect to timelining which forms a prime part of the investigation process, but the uncertainties that circumference the location of data make it more difficult to timeline. Since the file metadata does not store information relating to its movement and an officer find it quite difficult to handle the movement history of data over any given period.

On the other hand, network forensics is taken into account as a sub-branch of digital forensics concerning the monitoring and analysis of computer network traffic for the needs of data gathering, legal evidence or intrusion detection. Network forensics is additionally the process of gathering and examining raw data of network and systematically tracking and monitoring traffic of network to make sure of how an attack takes place. It aids in identifying unauthorised access to computer system and networks (Abhishek Srivastav, Imran Ali, 2014).

"Until recently, it was sufficient to look at individual computers as isolated objects containing digital evidence. Computing was disk centred collecting a computer and several disks would assure collection of all relevant digital evidence. Today, however, computing has become network-centred as more people rely on e-mail, e-commerce, and other network resources. It is no longer adequate to think about computers in isolation as many of them are connected together using various network technologies. Digital investigators/examiners must become skilled at following the cyber trail to seek out related digital evidence on the public Internet, private networks, and other commercial systems. An understanding of the technology involved will enable digital investigators to recognise, collect, preserve, examine, and analyse evidence associated with crimes involving networks."

Under the network forensics, the OSCAR methodology is relied upon for performing the investigation. The series of process are as follows:

Obtain Information- The collection of prime and critical information such as general information about the incident itself and the environment where it took place in, such as the date and time when an incident was discovered, persons and systems involved, what has initially happened, what actions have been taken since then, who is in charge, etc . The goals of the investigation should be well planned and prioritized.

Strategize- The second vital process involves the proper planning to be carried out in connection with the investigation procedure. There should be proper plan of action for prioritizing the acquisition process taking into concern the according to the instable nature of the sources, how potential value it can be for the investigation and the effort needed to get them.

Collect evidence- On the basis of the prior plans for the acquisition of the evidence intertwined with each identified source. Three entities should be points should be taken on matter.

1. Documentation: Any activity on the part of the investigating officer should be properly and systematically tagged and time lined. Any system accessed should be logged and the log must be stored safely following the same guidelines as the evidence itself. The log should include time, source of the evidence, acquisition method and the involved investigator.
2. Store/Transport: The chain of custody consisting of elements like showing the seizure, custody, control, transfer, analysis, and disposition of evidence, physical or electronic.
3. Analyse: An investigator resorts to number of variant methodologies and tools on the course of analysing process. Forensics researcher Brian Carrier described an "intuitive procedure" during which obvious evidence is first identified and then "exhaustive searches are conducted to start filling in the holes." (Carrier, 2006). The method opted by the respective investigative officer for analysis will depend on the case and what leads are already present.

4. Report: This will deal with conveying the results of the investigations to the client. It must be understandable by non-technical persons like managers, judges, etc. It must be factual and defensible in detail.

Indeed, cloud forensics and network forensics are considered as the sub sets of newer versions of digital forensics. The network forensics is now a novel emerging concept with respect to a network security while the cloud forensics applies majorly to issues covering cloud computing and allied applications.

4 Artificial Intelligence – Introductory Analysis

The task of defining artificial intelligence (hereinafter referred as AI) is left as a difficult one as there doesn't exist a clear definition of the same. As there follows a long series of questions to the categoric definitions as being laid down. By defining AI in terms of "creating a computer process that acts intelligently" but again left with the query what defines intelligence or "creating a computer process that can mimic human behaviour" leaving behind a challenging inquiry on do humans always act intelligently, what happens if a computer can normally perform better than a human. Another definitions refer to "rational behaviour" or engaging in a task that are hard for a computer can do. Considering variant elements impact on defining artificial intelligence, AI can be pragmatically as creating a computer process that acts in a manner that an ordinary person would deem intelligent. (Alastair Irons; Harjinder Singh Lallie,2014)

AI can be considered as an area of computer science that emphasizes the creation of intelligent machines that work and react like humans. Few of the interesting activities AI are designed to include speech recognition, learning, planning, problem solving, ability to manipulate and move objects etc. (Ahmad Habeeb, 2017). Advances in the field of machine learning is the matter of the hour. The line between mathematics and philosophy is blurry when we address artificial intelligence. The prime goals of AI include the creation of expert system and implementing human intelligence in machine. It is indeed multidisciplinary in nature includes the field of science, biology, psychology, linguistics, mathematics, and engineering.

Recently AI has been gaining more attention in different fields of science, technology and development fields. AI technology carries a variant feature when compared to a robot as an AI is being programmed to adapt and make decisions based on environmental factors surrounding it. For instance, these decisions take an innovative stand which ranges from a smart refrigerator refilling the ice in a freezer to a driverless car riding like how a human switch over in taking decisions instantly (Stuart J. Russell; Peter Norvig,2002).

5 Analyzing the application of AI in the discipline of digital forensics

5.1 Analyzing the Technical Components Involved in the Interplay

Idea of Representation of Knowledge and the Reasoning Process: The entity of inter-adaptability.

The concept intertwined with the representation of knowledge forms the vital part of most of the AI systems. The series of considerations include the knowledge representation with regard to representing the reason and formally structuring the same. The representation can be about the properties of objects in the domain and how these facts can be processed or even with respect to the application of these process. Recently, there persisted the realization that reasoning over multiple sources of knowledge is considered vital. This resulted in the creation of ontologies for domains4 that can be shared amongst applications and systems. The technologies such as XML, RDF 5are being utilised. Conceivably, here that AI has the potential to have the foremost effect on digital forensics, in providing expertise to assist the standardisation of the representation of data and information in the digital forensic domain. When paucity comes in with the quality of the above procedure, results in causing hindrance within the information exchange for even the most basic programming phase of a digital forensics procedure like the exchange of image information between forensics imaging tools, which ultimately pulls back the discipline of digital forensics in comparison to other scientific domains where there persist continuous effort in the production of standard domain ontology (Philip Turner, 2005).

The worth that follow the discipline of digital forensics by the creation of standardised international domain ontology is a remarkable one. For example, in a trans-boundary multi-jurisdictional case, it would provide a formal framework for the channelizing the digital evidence, also provide other benefits enabling the creation of a large, reusable case repository (D. A. Duce, F. R. Mitchell and P. Turner, 2007). This can be utilised in testing the performance of experts including a human or AI system. There also persist the utility in a standardised ontology with respect to reusing the collection

[4] A domain ontology (or domain-specific ontology) represents concepts which belong to a part of the world, such as biology or politics. Each domain ontology typically models domain-specific definitions of terms

[5] Extensible Markup Language (XML) is a markup language that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable. RDF stands for Resource Description Framework. RDF is a framework for describing resources on the web.

of background knowledge[6] and timeline specification which in turn aid the AI techniques.

The challenging concern involved in the application of AI in the forensic investigation is the AI technique or algorithm to explain the reasoning process.[7]

1. Among the symbolic reasoning, the most common type is the expert system. An expert system follows a predefined rule base. An expert system any point, the expert system has to provide an explanation of the reasoning for the conclusions obtained. Thereby, enabling an outside entity to critically analyse the reasoning process and to highlight any errors there might be with the reasoning used.
2. A Case Based Reasoners, another type of symbolic AI. The CBRs are based on well understood notions from psychology on how domain experts rely heavily on their past experiences, and when faced with a problem, will attempt to match the problem to atleast one they have experienced. Thus, the primary principle comes into picture only when all the possible similar cases in their experience is exhausted. In a CBR system, large set of case collection is obtained and a metric system is adopted to match the current case at hand. If fails to find the perfect match, but if a match is found that is deemed to be close enough, then the system may attempt to adapt the action of the matched cases to the instant case using a process refereed to as repair rules.
3. Pattern Recognition is yet another category for identifying specific types or clusters of data in an investigation. The software attempt to identify the parts of a picture, recognising a pattern in an e-mail message which indicates spam, or a pattern in disk image or a sound file. Most of the software relies on statistics or probabilistic reasoning or both. But the more complex and precise forms of image recognition that might be used to locate certain types of picture, rely on an understanding of how the human perpetual system works. (Dr. Faye Mitchell, 2010).

[6] Background knowledge is the term given to knowledge about a domain that is often common sense, and Often extremely large (e.g. If I throw a ball in the air it will normally come down; this windows file is normally found in this position in the directory tree). AI systems can be set up to use this knowledge to help their reasoning processes

[7] AI techniques are often divided into two categories: symbolic (those that reason with discrete entities in a knowledge base) and sub symbolic (those where the knowledge is spread around the representation structure).

6 Data Mining and Knowledge Discovery in Databases (DM/KDD)

The whole process of data mining and knowledge discovery in databases involves a multitude technique involving amalgamation of AI's like statistical analysis and probabilistic technique combined in order to analyse large collection of data. Technically, this process is a form of Exploratory Data Analysis (EDA), for instance the user provide command to the system for highlighting files with characteristics Q, and the system utilises Data Visualisation to highlight and recognizing potential relationship to the user. This find its merits in the digital forensics as the human perceptual system has the ability to distinguish patterns in extremely complex data. Data Mining and Knowledge Discovery in Databases also exhibits a concept termed as interestingness measure which aids us to decide whether there are any meaningful patterns in the set data. Thus, the Data Mining and Knowledge Discovery in Databases has to be relied on by the investigator during the initial phase of assessment. One of the demerits that persist is that the chance of missing relevant pieces of information as the reasoning process do not normally use background knowledge or complex reasoning.

7 Process of Adaptation through Machine Learning

The branch of AI that deals with the ability of the software to adapt is called Machine Learning (hereinafter referred as ML). When it comes to the application of Machine Learning Technique to digital forensics, the ML techniques can be classified under two variants, one is with respect to use of ML as a method of trying to refine the knowledge source to keep it updated referred to as refiners and other one is using ML to gather the initial knowledge called as the learners.

For instance, it will be possible for a human perception to tell by taste whether or not a whisky was a malt whisky or not, but not be able to predict exactly about what made it taste like a malt whisky. In such circumstances, it is possible for an AI system to learn about what the concept is by using a learning system. Such systems normally rely on the use of training sets which contain pre-classified examples which, along with the algorithm, form the basis of the learning system. The success or failure depends on the credibility and suitability of the learning algorithm and the quality of the data set used (Brian Carrier,2003).

- Social Network Analysis and Application of AI

The Social Network Analysis utilises graph theory 8and other related graphical techniques to allow for the analysis of networks (Mithas, 2012). The utility of SNA has been established in variant areas. They include discovering of hidden group, i.e., a group of individuals planning an activity over a communication medium without announcing their intentions, another one includes aiding the investigator in discovering organizational structure, also with respect to demonstrating how networks of people changes during an emerging situation(Diesner, J.; Frantz, T.L.; Carley, K.M; 2005). Further the SNA technique allow the investigator to work out on the density of communications, the strength of connections between the people and the factor of influencing power of a person in a network (Baumes, J.; Goldberg, M.; Hayvanovych, M.; Magdon-Ismail, M.; Wallace, W.; Zaki, M., 2006). The real science behind the making of such a technique is based on graph based mathematical analysis allowing the investigator to identify patterns in group behaviour and in particular identifying the key parts of the network. As with respect to the technical version, many variants of open tools are being utilised such as NetworkX, Pajek and Gephi, also industrial solutions such as i2 analyser.

- Investigation Toolkit

7.1 Multi-Agent Digital Investigation Toolkit (MADIK)

A multi-agent digital investigation toolkit is a multiagent system to assist the digital forensics expert during the examination process. The system comprises of a group of Intelligent System Agents that perform different analysis on the digital evidence related to a case on a distributed manner. In this toolkit, each ISA contains a set of rules and a knowledge base, both based on the experience of the expert involved in the specific case at hand. Since the fact that the examination of digital evidence in crime investigations share resemblances, MADIK uses case- based reasoning technique to determine which agents are better employed in which kind of investigation. This successively end in allowing the agents to reason about the evidences in a way that is more capable to the specific case in question. For instance, if we would like cite the sets in dowry abuse case. The ISA will initially use the hash sets related to dowry abuse cases, thus giving the examiner a quicker feedback on the existence of such files in a piece of evidence. Outlooking the technical aspect, the MADIK was implemented using the Java Agent Development Framework (JADE), fully developed with the Java language. JADE was used since it simplifies the implementation of multiagent systems, over a distributed platform (Mark d'Inverno; Michael Luck; Michael M. Luck, 2004).

Currently, the MADIK uses six kinds of specialised intelligent agents, they are as follows:

[8] A visual representation of data, in the form of graphs, helps us gain actionable insights and make better data driven decisions based on them.

1. Hash Set Agent: It calculates the MDS hash from a file and does the task of comparing it with its knowledge base, which contains sets of files and classify as ignorable or important.
2. File Path Agent: It tend to preserve its knowledge base a set of collection of folders which are commonly used by several application which may be of interest to the investigation like P2P(peer-to-peer) sharing, VoIP and instant messaging applications.
3. File Signature Agent: It scrutinizes the initial 8 bytes of the file headers to determine if they match the file extension. If someone alters the file extension in order to hide the true purpose of the file, this will be detected by this agent.
4. Timeline Agent: It inspects the entities such as date of creation, access and modification to determine events like system and software installation, backups, web browser usage and other related activities which will be having trail connection with the instant case of investigation at hand.
5. Windows Registry Agent: It studies the existing files which has connection with the windows registry and extracts valuable information such as system installation date, time zone configuration, removable media information and others.
6. Keyword Agent: It hunts for keywords and uses regular expression to extract information from files such as credit card numbers, URLs or e-mail addresses.

The MADIK which has absorbed the case-based approach provides a way to improve in analysing and correlate the findings in a meritorious manner when compared to the current system of acquisition and extraction of data. It also provides ample opportunity for the investigative agents to improvise the results over time by learning from previous cases (Andrew Case; Andrew Cristina; Lodovico Marziale; Golden G. Richard; Vassil Roussev, 2008).

7.2 AUDIT: Automated Disk Investigation Toolkit

AUDIT is engaged with the task of integrating and configuring the tools automatically for both general and specific investigations. For instance, with reference to searching the disk for evidence in graphic files, emails, documents and hidden locations. Also detailed search for items such as credit card and social security numbers can also be done. The toolkit comprises of three entities: a database of investigative tasks and tools; a knowledge domain with constructs defining rules and facts; and a core engine or an expert system.

Within the database component, two tables that maintain information regarding the tools that will be utilised by the AUDIT and the investigative tasks that an average investigator generally performs. The knowledge base contains facts and rules, some of which are predefined and embedded into the system and others that are created during the investigation. Facts and rules can be added, deleted and modified as required. The core engine controls the running execution of the system using the database component, the knowledge base and therefore the user input. The expert engine reads tool specifications and investigative tasks from the database and creates new rules and facts as needed. It also links the investigative tasks and therefore the tools with respect to the knowledge domain and user input and feedback. The AI part of AUDIT is mainly the embedded expert system and knowledge domain that is represented in it. In AUDIT, we used the open source expert system tool CLIPS which provides an entire platform to make rule and or object based expert systems and is additionally used to represent an expert's technical knowledge (Tye Stallard ; Karl Levitt,2003).

Analysing the technicality in AUDIT. Knowledge is represented via rules and facts. A rule in CLIPS consists of two parts: IF and THEN commands. In the IF portion of the rule, facts are listed that determine whether the rule is to be applied or not. A collection of facts is called a pattern and pattern matching is done by CLIPS to decide if the THEN portion is activated. In this case the rule is said to be active, else it is passive. If the facts hold (pattern matches), then actions in the THEN portion will be executed by the CLIPS inference engine. Multiple rules may be active at any time and the ordering of execution can depend on the salience value in the IF portion. The IF portion of the rule has a different characteristic than an IF statement in conventional programs. It works as WHENEVER, because facts can be changed anytime during the program execution. The inference engine executes actions of all active rules. Most of the actual rules used in AUDIT are more complex. In this rule, the user is asked to provide his/her technical expertise and need of help for investigation. Based on the answer received from the user some certain facts will be added to the facts list by using the assert command of CLIPS. The IF portion of the rule consists of the two lines before the symbol and the THEN portion of the rule is after that. This rule will be activated when we have no information about the user's expertise (Rainer Poisel and Simon Tjoa, 2011).

8 Analysing the Transformation of Traditional Digital Forensics into Intelligent Digital Forensics – Inevitable Revamping

Intelligence play a prime role in criminal investigations and is indeed the application of the artificial intelligence to digital forensics takes on a number of components of various stages of the investigation process involved starting with the gathering of digital evidence, the preservation of digital evidence, the analysis of digital evidence and the presentation of the evidence. The skill and expertise element of an investigating officer in each of these stages plays a vital role. Human perceptions and involvement are always folded by myriad of technical difficulties especially in the case of digital forensics. Here comes the crucial role played by the application of artificial intelligence in the process of digital forensics through useful set of tools and primely dealing exclusively on the speed and volume concerns of digital investigation cases. The course of action enables a speedy tracking of the required data sets and eliminating dormant files and static system files from digital investigations mainly by the application of hash algorithms.

The term digital intelligence covers a number of meanings. According to Mithas, who advocates that business managers can gain a significant advantage by having the intelligence to understand, analyse and use digital technology so as to provide competitive benefit and advantage, something that he refers to as digital intelligence. (Mithas S, 2010).

Stanhope's view however is somewhat different and he proposes that digital intelligence is:

The capture, management, and analysis of data to provide a holistic view of the digital customer experience that drives the measurement, optimization, and execution of marketing tactics and business strategies (Ribaux, O.; Baylon, A.; Roux, C.; Delémont, O.; Lock, E.; Zingg, C.; Margot, P, 2010).

Intelligent forensics exhibits an inter-disciplinary approach, which utilises technological advances and applies resources in a more intelligent way to solve an investigation. Intelligence forensics encompasses a range of tools and techniques from artificial intelligence, computational modelling and social network analysis in order to focus digital investigations and thereby increasing the efficiency. It can be applied both proactively i.e., before a case occurs and reactively i.e., post the occurrence of an incident.

Digital forensic intelligence are often drawn from intelligence led activities, also through routine investigations quite often, the intelligence drawn thereof stores in databases. There exist a variety of examples of such intelligence databases within the forensic science domain, for instance, the UK National DNA Database (NDNAD), the UK National Fingerprint Database (IDENTI) and the USA Integrated Automated Fingerprint Identification System (IAFIS).

With regard to the analysing the switchover of traditional digital forensics to intelligent forensics, the major elements of challenge can be categorised under two entities: legal and computational. Legal encounters include transgression with reference to the jurisdictional concern. On the other hand, computational challenges comprise of abnormal states of the computing machine, for instance, sector containing data in an abnormal part of disc or abnormally formatted data packets, data out of normal bounds or issues concerning personal relational data which point to unusual relationships.

As a persisting solution, the knowledge-based systems can be instituted to capture legal expert's understanding of the principles of the law and be able to signal unusual behaviour. A neutral network are often synced to categorize appropriate behaviour and are even able to model the behaviour of different users so that it would be possible to signal use patterns for the currently logged in user. Data mining and machine techniques can be used to discover patterns of behaviour and flag exceptions. Along with big data analytics and high-performance computing platforms, it is possible to develop systems, which continuously learn and improve system performance in order to keep up with changing trends in the computer forensics arena. Such techniques could be used to automate aspects of the identification, gathering, preservation and analysis of evidence both post hoc and proactively.

9 Conclusions

The viability in the utility factor of the application of artificial intelligence in digital forensics is the need of the hour taking into concern the environment of cybercrime with respect to its changing and growing scale. While relooking into the different forensics' procedure ranging from identifying, collecting, recovering, analysing and documenting there necessitates a more structures and efficient inclusion of technical tools and equipment which need to be merged in the discipline of digital forensics. For extensively combatting with the existing and future challenges allied with cyber-crime, there exhibits the need to enhance the use of the resources available and move out of the capabilities and constraints of the tools and techniques presently utilised by the current forensic arena. As technology is making leaps and bounds in the recent time frame and will continue to exponentially demonstrate the progress beyond our imagination. Whether with an email containing a virus attacking a random computer to serious crime hacking the national security surveillance of a jurisdiction is a matter of threat at myriads of spheres. The limitation of human perceptions and involvement and elimination of human error and switching over to machine detecting anomalies post and pre -phase involved the criminal activities. Indeed, the improvement in the acquisition and presentation of evidence will undergo a transformation considering the application of artificial intelligence as a smart applicability in our digital forensics. Thus, the technical challenges can be deployed at a greater extend reducing the delay and time lapse in the arena of digital investigation. Thus, paving a higher demand for the utilization of technical experts and demanding the applicability of artificial intelligence is in demand for increasing the efficiency and reliability of the digital forensics investigative techniques and the process involved.

References

1. Abhishek Srivastav, Imran Ali, Network Forensics an Emerging Approach to A Network sis, International *Journal of Computer Science and Engineering Technology,* India: Transstellar Journal Publications and Research Consultancy Private Limited, Feb 2014 ,5 (2), 118- 123. ISSN 2249-7943.

2. Alastair Irons and Harjinder Singh Lallie, Digital Forensics to Intelligent Forensics, *Future Internet*, Switzerland: Multidisciplinary Digital Publishing Institute (6) 585-59612 Sept 2014. ISSN 1999-5903.
3. Arjit Paul, Mayuri Kiran Anvekar & K. Chandra Sekaran, Cyber Forensics in Cloud Computing. *Computer Engineering and Intelligent Systems* [online].US: IISTE, 3(2), 29-36 .[viewed date April 28, 2020] .Available from:https://iiste.org/Journals/index .php/CEIS/article/ view/982/902.
4. Baumes, J.; Goldberg, M.; Hayvanovych, M.; Magdon-Ismail, M.; Wallace, W.; Zaki, M., Finding Hidden Group Structure in a Stream of Communications. Berlin: Springer, 2006. ISBN 978-3-540-34478-0.
5. Brian Carrier, Defining Digital Forensic examination and Analysis Tools using Abstraction layers. *International Journal of Digital Evidence*. Leibinz: Michael Ley, 2003, 1(4), 1-12. ISSN 1742-2876.
6. Bynum, Terrell, Computer and Information Ethics. In Edward N. Zalta. The Stanford Encyclopaedia of Philosophy [online]. Edition.: Stanford University, Apr 18 2018. ISBN 1095-5054. [viewed on 22 March 2020]. Available from <https://plato.stanford. edu/archives/sum 2018/entries/ethics-computer/>.
7. Carrier, B.D. Basic Digital Forensic Investigation Concepts. *International Journal of Digital Evidence* [online]., Leibinz: -Michael Ley, August 7,2012,1-10 [viewed date October 7th, 2018] Available from <http://www.digitalevidence.org/di_basics.html>.
8. D. A. Duce, F. R. Mitchell and P. Turner, The Use of Artificial Intelligence in Digital Forensics: An Introduction, *Digital Evidence and Electronic Signature Law Review*, Geneva: Pario Communications Limited Publication,2007 (7) 35-41. ISSN 1756-4611
9. Dr. Faye Mitchell, The Use of Artificial Intelligence in Digital Forensics: An Introduction, *Digital Evidence and Electronic Signature Law Review*. Geneva: Pario Communication Limited 2010 (7) 35-41. ISSN 1756-4611.
10. Ewa, Huebner; Derek, Bem, Computer Forensics Analysis in a Virtual Environment Future, I*nternational Journal of Digital Evidence*. New York, Utica, January 2007, (6),1-13. ISSN 1742-2876
11. Mark d'Inverno and Michael Luck. Understanding Agent Systems. 2nd ed. Berlin, Germany: Springer Series in Agent Technology ,2004. ISBN 3-540-40700-6.
12. Mark d'Inverno; Michael Luck and Michael M.Luck, Understanding Agent Systems. 2nd edn. Berlin, Heidelberg, New York Springer Science & Business Media, 2004. ISBN 3-540-40700-6.
13. Michael, G. Noblett, Mark M. Politt and Lawrence, A. Presley. Recovering and Examining Computer Forensic Evidence. *Forensic Science Communications*, Amsterdam: C. Cattaneo, C. Jackowski, 2000, 2(4), 10-21. ISSN 0379-0738.
14. Mithas, S. Digital Intelligence: What every Smart Manager Must Have for Success in an Information Age; 3rd ed. North Potomac, MD, USA, FinerPlanet: Nov 17, 2015, 34-55 ISBN: 0984989633
15. Nina Godbole & Sunita Belapure, Cyber Security: Understanding Cyber Crimes, Computer Forensics and Legal Perspectives. India: Wiley, 2011, 342.ISBN 978-8126521791.

16. Philip Turner, Unification of digital evidence from disparate sources (Digital Evidence Bags), *Digital Investigation*. Amsterdam, Elsevier Publishers 2005, 2(3), 223-228. ISSN 1742-2876.
17. Prasad Purnayae Prasad Purnayae, Cloud Forensics: Volatile Data Preservation, 4 *International Journal of Computer Science Engineering*, 2015,4(2), 41-43. ISSN. 2319-7323
18. Rainer Poisel and Simon Tjoa, Forensic Investigations of Multimedia Data: A review of the State of Art (Stuttgart, Germany, May 10-12, 2011), Sixth Conference on IT Security Incident Management. ISBN 978-1-4577-0979-1.
19. Raun, Keyn, Joe Karby, Tahar Kechadi & Mark Crosbie, Cloud forensics. In: Peterson, Gilbert & Sujeet Shenoi. *Advances in Digital Forensics.* 1st ed. U.S: Springer, Aug 2018. ISBN 978-3-319-99277-8.
20. Ribaux, O.; Walsh, S.J.; Margot, P. The contribution of forensic science to crime analysis and investigation: Forensic intelligence, *Forensic Sci. Int.* Elsevier, 2006, (3) 171–181.ISSN 0379-4410
21. Richard Saferstein. Forensic Science Handbook .2 ed. New Jearsey: Pearson, 2001. ISBN 978-0130910585.
22. Sherman Josaih Dykstra & Alan T. Sherman, acquiring forensic from Infrastructure-as-a-service cloud computing: Exploring and evaluating tools, trust, and techniques, *Digital Investigation*,2012,9, 590- 598. ISSN. 1742-2876
23. Stuart J. Russell and Peter Norvig. Artificial Intelligence: A Modern Approach. 2nd edition, Prentice-Hall, USA,2002. ISBN 0-13-790395-2.
24. Tye Stallard, Kart Levitt, Automated Analysis for Digital Forensic Science: Semantic Integrity Checking (Orlando, Dec 4-8,2017) Proceedings of the 19th Annual Computer Security Applications Conference. ISBN 978-1-4503-5345-8.
25. Virginiah, Sekgwathe; Mohammad Talib, Cyber Forensics: Computer Security and Incident Response, International Journal of New Computer Architectures and their Applications. Hong Kong, Society of Digital Information and Wireless Communication, January 2012, (2), 127-137.ISSN 2220-9085