Behavioural Evidence Analysis: A Paradigm Shift in Digital Forensics

Barkha Shree, Manav Rachna University, Faridabad, India
Parneeta Dhaliwal, Manav Rachna University, Faridabad, India

ABSTRACT

Recent developments in digital forensics (DF) have emphasized that along with inspection of digital evidence, the study of behavioural clues based on behavioural evidence analysis (BEA) is vital for accurate and complete criminal investigation. This paper reviews the existing BEA approaches and process models and concludes the lack of standardisation in the BEA process. The research comprehends that existing BEA methodologies are restricted to specific characteristics of the forensic domain in question. To address these limitations, the paper proposes a standardised approach detailing the step-by-step implementation of BEA in the DF process. The proposed model presents a homogenous technique that can be practically applied to real-life cases. This standard BEA framework classifies digital evidence into categories to decipher associated offender characteristics. Unlike existing models, this new approach collects evidence from diverse sources and leaves no aspect unattended while probing criminal behavioural cues, thus facilitating its applicability across varied forensic domains.

KEYWORDS

BEA, Behavioural Analysis, Behavioural Evidence Analysis, Criminal Profiling, Digital Evidence, Digital Forensics, Forensic Analysis

INTRODUCTION

The role of technology is ever-increasing in today’s world. The inception of technology has provided some unique opportunities (Sedera & Cooper, 2019) towards creating a modern society but it can also be abused and misused by individuals. The plethora of tools and technologies (Dawson & Omar, 2015) available today that allow criminal acts to occur anywhere in the world serve as an incentive for criminals. The offenses carried out by criminals facilitated by a computer (McMurdie, 2016), computer networks or any other type of information communications technology are known as Digital crimes. It also involves classic crimes (such as murder, blackmailing, kidnapping, defamation), that misuse technological adeptness and accessibility to information. Formerly, criminal investigations revolved around the analysis of physical evidence from the crime scene. Whereas today, in the digital era, the evidence to assay is in an electronic or digital form (Macdermott, Baker, & Shi, 2018) and is called Digital Evidence. Digital evidence is any information related to the crime that is stored or transmitted in digital form. It may comprise of computer-generated log files, browsing history, or metadata and may be accrued from computer systems, smart digital devices, or network traffic.

The mechanism required to obtain and analyze digital evidence to solve criminal cases is called Digital Forensics (Årnes, 2017; Sammons, 2015). It is a new and fast-growing form of investigative
practice wherein the forensic specialists use modern forensic software tools to capture and examine digital evidence. Experts in digital forensic investigation are facing unseen challenges due to the new, advanced technologies, used in digital devices, and adopted by criminals alike. The evidence in many cases is not sufficient to narrow down suspects. The difficulties can be overcome through a hybrid digital forensic task process (Rahman & N. A. Khan, 2016) that incorporates other dimensions like behavioural clues of the offender into the traditional DF process.

The discovery and examination of behavioural clues of the offender and the victim from digital evidence is known as Behavioural Evidence Analysis (BEA) (Turvey, 2016; Turvey & Esparza, 2016). BEA is beneficial for creating a profile of the criminals based on their offense behaviour and establishing a strong evidence base for similar crimes in future. This way it can help in focusing on an investigation with speed and in the right direction, and in predicting the offender’s behaviour and motivation (Turvey, 2016). These results can further facilitate narrowing down a suspect’s behaviour by mapping it to a criminal profile thus assisting in elimination of suspects.

The existing literature only offers a customary description of the techniques of BEA, and its usefulness in investigating digital crimes. Despite the recognized efficacy of BEA, there exists no standard technique for the practical incorporation of BEA in solving real life criminal cases. Moreover, the models designed previously are constrained by the specific characteristic of the forensic domain being dealt with, limiting their use to particular categories of crime. This paper attempts to address these discovered gaps in the literature in several ways.

The paper proposes an approach that stipulates a well-defined, unambiguous, and extensive technique for “how” BEA can practically be implemented within the DF process. This new framework advances the current BEA process methods and is a standardized, homogenous approach to BEA. The model applies all aspects of BEA sequentially ensuring that the basic procedure of behavioural analysis is intact during investigation. Its prototype is designed to prudently include any unmonitored or often neglected evidence that may be left out in other methodologies. The system presented is holistic in nature, ensuring that no factor is ignored during the identification of behavioural clues of the offender. This extends the applicability of the model across varied forensic domains. The proposed approach modularly sorts the evidence extracted from digital devices into categories and derives different behavioural characteristic of the offender for each category. These cues are later integrated to devise a complete pen picture of the perpetrator of crime. This helps in narrowing down the pool of suspects and/or identification of the probable assaulter. Owing to the proficiency of the new framework to deal with all kinds of evidence, unlike previously designed systems, it can be used for all kinds of crimes, digital or conventional.

The research paper is structured as follows: Firstly, it describes background knowledge; Secondly, it reviews related work in previously developed DF models that incorporate aspects of BEA. Thirdly, it illustrates a comparative study of existing BEA approaches, identifying their advantages and limitations. Following that the paper proposes a standardized and integrated approach to BEA. Then given are the challenges and future directions before lastly presenting the conclusion.

**BACKGROUND KNOWLEDGE**

The expansion of technology in modern society transforms people's lifestyles (Siddiqui, 2016) thus having a significant impact on human behaviour (Holt, Bossler, & Seigfried-Spellar, 2017; Holt et al., 2018b). However, with the overwhelming abundance of technology, tools, and their simplicity of use (Bieser & Hilty, 2018), it becomes mandatory to delineate the positive and negative impacts of digital transformation on society. The anonymity offered by the internet has led to an upsurge in the number of crimes involving computers and cybernetics (MacDermott, Baker, Buck, Iqbal, & Shi, 2020). The impact of globalization and cyber developments (Tang, Alazab, & Luo, 2017) is such that the digital crimes perpetrated at both local and global level is increasing (Ronchi & Politecnico, 2019).
Digital crimes encompass several criminal acts like hacking (DeTardo-Bora & Bora, 2016), damage to computer resources, disruption of networks (Swarna Priya et al., 2020), malware attacks (Mamoun Alazab & Broadhurst, 2017; Moutaz Alazab, Alazab, Shalaginov, Mesleh, & Awajan, 2020), money related frauds, information theft, identity theft, cyber vandalism (Chowdhury, 2016), denial of service, copyright violation (ADYGEZALOVA, ALLALYEV, KISELEVA, & GRIGORIEVA, 2018), piracy (Hoy, 2017) and child pornography (Desai & Narayankar, 2015). The motivation behind these crimes may involve subjects that are personal, commercial, political, anarchical, or other, however, the rudimentary attack and intrusion techniques generally remain identical. The incessantly increasing complexity of digital crimes exhibit both the enhanced complexity of our digital technology and services mix, in addition to ongoing advancements in security (Iwendi et al., 2020; MacDermott et al., 2020; Taleby Ahvanooey, Li, Zhu, Alazab, & Zhang, 2020). This may pose new challenges to gather evidence in case of digitally committed crimes and their successful resolution.

The proliferation of digital crimes presents an essential investigative challenge vis-à-vis the review of evidence residing on the mass of devices (Mamoun Alazab, Venkatraman, & Watters, n.d.) that can be exploited for perpetrating a crime. The process involves the identification, acquisition, and analysis of digital evidence from the “devices of interest” at a crime scene (Årnes, 2017). The comprehensive probing of digital evidence can establish the occurrence of any online or offline criminal activity. Thus, there is a need for a forensic investigation process that examines digital information in the form of digital evidence and identifies actors within an exchange, determines intent and behaviours, and ascertain a timeline of events (Al Mutawa, Bryce, Franqueira, Marrington, & Read, 2019). As the technological complexity and storage capabilities of digital devices is increasing (Rahman & N. A. Khan, 2016), the forensic digital investigations involving such systems would include a more intricate digital evidence procurement and analysis (Holt et al., 2018a; MacDermott et al., 2020).

Digital Forensics Framework

In digital data investigation, relevant evidence is extracted from electronic devices because of digital forensic processes. The framework for the DF process includes the stages of evidence collection/acquisition, examination, analysis and lastly report generation (Hassan & Hassan, 2019). During the collection/acquisition phase (Varol & Ülgen Sönmez, n.d.), all devices of interest toward uncovering digital evidence are captured from the crime scene and data preservation is undertaken to retain the authenticity of evidence. The examination phase (Holt et al., 2018a) of digital forensics investigation, is concerned with the recovery or extraction of digital data from electronic devices. The analysis phase (Varol & Sönmez, 2017) employs state-of-the-art technology, tools and sophisticated software to yield substantial results in identifying the perpetrator of the crime and setting out the timeline of occurrences. The report thus generated must be accurate, precise and its integrity must be established before presenting it in the court of law. Digital forensics practices (Sammons, 2015) must also be assured by quality assurance measures for investigatory standards. Figure 1 demonstrates the traditional digital forensics process.

Cognitive and human factors in forensic science have been given much attention recently (Sunde & Dror, 2019). Forensic processes (Sammons, 2015) must be collaborated with the cognitive and human aspects to enrich it, so it proves to be a more efficient and comprehensive evaluation that considers other facets like human behavioural and psychological elements. Thus, apart from the evidence found at the crime scene, the behavioural patterns of the victim and the perpetrator of the crime prove beneficial in solving the case and prosecution of the criminal. The behaviour of the criminal analyzed from the crime scene can be used to draw a profile of the offender’s mannerisms. This helps in easy identification of the potential suspect or determination of his likely characteristics.

Criminal Profiling

Criminal profiling (Balogun & Zuva, 2019; Garcia, 2018) includes analysis of offender behaviour by interpreting digital or physical evidence and understanding the crime scene along with its underlying
circumstances. It is a practice whereby the likely characteristics of a criminal are predicted based on the behaviours exhibited while carrying out a crime. While analyzing personality traits from a criminal profile, the proclivity of an offender can be assessed. Suspicious activities may be found out by reviewing an offender’s behavioural characteristics in previously committed crimes (Turvey, 2016). It can be said that an offender’s profile forms the nucleus of malicious behaviour constituents.

Criminal profiling aims to create a profile which can depict behavioural aspects of a criminal under the same circumstances to commit similar crimes (Almond, McManus, & Curtis, 2019). Criminal profiling process creates an offender profile using three general forms of illative reasoning, namely, deductive, inductive, and abductive. Deductive profiling (Turvey, 2012) follows a case-based approach. It crafts a profile of the characteristics of the possible offender by examining behavioural evidence from the criminal case in question. Inductive profiling (Warikoo, 2014) analyzes data from criminal records on criminals convicted in different types of cases (e.g., murder, rape) to create a pool of suspects. Abductive profiling (Fahsing & Ask, 2018) uses the available evidence to create and evaluate hypotheses, make predictions, deliver explanations, or infer conclusions. This draws a similarity to the investigative approach used by Sherlock Holmes.

While constructing the offender profiles, various strategies may be adopted which later determine the specific area of use of these profiles to study behaviour of criminals. These strategies (Bartol & Bartol, n.d.) may be geography based, psychology based, suspect based, or crime scene based as mentioned in Table 1.

**Behavoural Evidence Analysis and Its Role in Digital Crime Investigation**

Behavioural Evidence Analysis (Turvey, 2016; Turvey & Esparza, 2016) considers a deductive, case-based approach to forensic criminal investigations. It inspects particular criminal cases to detect specific behavioural or personal traits of the suspect. BEA involves four types of analysis (Turvey, 2016) comprising Equivocal forensic analysis (Al Mutawa, Bryce, Franqueira, & Marrington, 2015), forensic victimology (Petherick, 2019; Petherick & Ferguson, 2015), identification of crime scene characteristics (Almond et al., 2017; Wang et al., 2018) and identification of offender traits (Turvey, 2012).

The first step includes analyzing the physical evidence to determine offense behaviour. This process goes through the scientific assessment of case details which includes deep study and evaluation of digital evidence while retaining logical reasoning and critical thinking. The second step of forensic victimology comprehends the role of the victim in the crime. It deals with finding out the specific traits of victim’s lifestyle and mannerisms that led to his selection, assessing the risk exposure of victim to the crime. The third step involves crime scene examination to reveal the choices undertaken by the offender in relation to the offense (signature ways and modus operandi). It implicates finding hidden aspects of crime scene which can uncover further evidence to define behavioural attributes of the suspects more clearly. In the last stage, behavioural traits of the suspect are specified based on
offender characteristics revealed in preceding stages. The discovered behavioural characteristics of the offender are to be used in conjunction with the traditional DF investigation process to provide a more coherent reconstruction of the crime. Figure 2 depicts the incorporation of BEA technique into the traditional digital forensics process.

BEA incorporated into the DF process enhances the quality (Al Mutawa et al., 2019) of the results yielded by the traditional DF process. The traditional DF technique considers only the digital evidence extracted from the devices of interest. At times, these evidences may not be complete or sufficient and reconstruction is not possible. In such scenarios, BEA used with DF plays a vital role in aiding the investigator to determine the course of action. The behavioural clues supplement the digital evidences, facilitate the narrowing down of suspects and thus speed up the investigation. The benefits of using BEA in digital forensics is depicted in Figure 3.

There are multiple theories that are put to work in the identification of different behavioural characteristics during all stages of the BEA process. They may be based on evolution (Nesse, 2015), psychodynamics (Balfour & Tasca, 2015; Battistelli & Farneti, 2018; Selater, Piper, Brown, & Day Selater, 2019; Willmott, Boduszek, & Robinson, 2018), behaviour (Walinga, 2019), biology (Beaver & Walsh, 2018; Farahany, 2015; Tiitonen et al., 2015; Veroude et al., 2016) or sociocultural differences (Azuh, Fayomi, & Ajayi, 2015; Maunder & Crafter, 2018; Schneider, Rollitz, Voracek, & Hennig-Fast, 2016; Zittoun, 2020). These major theories, along with the approach, implications and limitations are listed in Table 2.

### RELATED WORK

This section presents the survey of the existing work on behavioural analysis, discusses the various Digital Forensics Investigation Process Models (DFIPMs) developed so far and gives a brief understanding of the existing Behavioural Digital Forensics methods that incorporate human factors in a DF investigation.
The consideration of behavioural physiognomies while probing crimes has been emphasized in several studies.

Bryant (Bryant, 2016) pointed out that criminology can prove effective in investigating digital crimes. The application can aid investigators by giving pointers and leads and access to a wealth of information.

Kaati et al. (Kaati, Shrestha, & Sardella, 2016) argued that the psychological state of the criminal and his tendency of committing an offense can be deciphered from their spoken or written language for example, threat notes or verbal linguistics.

Tonkin et al. (Tonkin et al., 2017) used a dataset of serial sexual assaults from five countries comprising 3,364 cases. He employed various statistical procedures to examine offender behaviour. The results revealed that patterns of offender behaviour could be established vis-à-vis the modus operandi, extent of involved violence or victim restraint techniques. This also successfully linked sexual crimes carried out by the same offender.

Beauregard et al. (Beauregard, Busina, & Healey, 2017) conducted studies on 624 convicted sexual criminals and their associated case files. Based on the probability of confession in an interrogation, five profiles each for the offender and the victim were identified. It has helped in designing interrogation strategies for every profile to obtain criminal confessions.
Rogers (M. Rogers, 2003) in his research described the efficacy of behavioural analysis in the digital crime investigation process. He argued that both conventional and digital crimes essentially follow similar procedures for investigation: inspecting events of crime, offender motivations, and modus operandi.

Lowe (Eagle & 2002, n.d.) emphasized on the study of criminal behaviour in cases of crimes against persons, property, gang crimes and so on. This research focused on examining the role of criminal profiling in the operational phase of an investigation.

Bennett and Hess (Bennett, Learning, & 2007, n.d.) analyzed a variety of crimes like sex offenses, crimes against children, robbery, drug abuse, computer crimes and so on. They provided details on developing the suspect’s traits based on the modus operandi information, psychological profiling, geographic profiling, and racial profiling.

Digital Forensics Investigation Process Models (DFIPMS)

The inexorable evolution of technology has increased the number and complexity of cybercrimes (Lillis, Becker, O’Sullivan, & Scanlon, 2016). Due to this, numerous challenges have arisen in the digital forensics sphere (Vincze, 2016). These may be related to difficulties in law enforcement, technical complications, lack of expertise among investigating personnel and absence of standardized DF procedures (Karie & Venter, 2015). This suggests a need for the development of inventive techniques (Ch, Gadekallu, Abidi, & Al-Ahmari, 2020) and methodologies over conventional ones for probing digital offenses. This section discusses the various existing DFIPMs designed over the years.

An Abstract DF Model (Reith, Carr, & Gunsch, 2002) was developed that was not dependent on specific technologies. It had nine phases, namely, identification, preparation, approach strategy,
preservation, collection, examination, analysis, presentation, and returning evidence. A drawback of this model was that it was too general for practical use and its testing and validation was difficult.

Carrier and Spafford (Carrier & Spafford, 2003) designed an Integrated Digital Investigation Model that had an equal number of phases for both physical and digital crime scenes. It had five groups further expanding into 17 phases. The groups included readiness, deployment, physical crime scene investigation, digital crime scene investigation, and review. The practical applicability of this model was not established in real cases.

Enhanced Digital Investigation Model (Baryamureeba & Tushabe, n.d.) was a modification of the Integrated Digital Investigation model that split the investigation into physical and primary (digital) crime scenes. This system had five steps, namely, readiness, deployment, trace back, dynamite, and review. This model was later found to be useful only in investigation of digital frauds (Mir, Shoaib, & Shahzad Sarfraz, n.d.).

Beebe and Clark (Beebe & Clark, 2005) presented a multi-tiered framework called Hierarchical Objective-based Model. The first tier consisted of six stages, (1) preparation, (2) incident response, (3) data collection, (4) data analysis, (5) presentation of findings, and (6) incident closure. The second tier included objectives-based subphases and task hierarchies. This approach was tested on fictional cases and provided no detailed explanation of levels other than analysis stage.

Perumal (Perumal, 2009) claimed that no DFIPM developed previously focused on the full scope of investigation and only stressed on specific aspects. He presented a model that comprised seven steps. The stages included planning, identification, reconnaissance, analysis, results, proof and

<table>
<thead>
<tr>
<th>Theory</th>
<th>Approach</th>
<th>Implications</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evolutionary theory</td>
<td>Deals with selective pressures influencing behaviour</td>
<td>Mental instincts and behaviours are adaptive</td>
<td>Based on assumptions rather than evidence</td>
</tr>
<tr>
<td>(Nesse, 2015)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychodynamic theory</td>
<td>States that components of adult personality are based on childhood experiences</td>
<td>Unconscious mind affects behaviour</td>
<td>Ignores role of socialization and free will</td>
</tr>
<tr>
<td>(Balfour &amp; Tasca, 2015; Battistelli &amp; Farneti, 2018; Sclater et al., 2019; Willmott et al., 2018)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Behaviourist theory</td>
<td>Outlines the impact of external stimuli on behaviour</td>
<td>Circumstances cause an individual to learn behaviour</td>
<td>Undervalues intricacy of human behaviour and role of biological factors</td>
</tr>
<tr>
<td>(Walinga, 2019)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitive theory</td>
<td>Describes logical thinking and judgement as a result of state of mind of an individual</td>
<td>Individual’s interpretation of social situations determines behaviour</td>
<td>Free will is not recognized</td>
</tr>
<tr>
<td>(Robinson-Riegler &amp; Robinson-Riegler, 2016)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Humanistic theory</td>
<td>Lays emphasis on individual choice and free will of human nature</td>
<td>Inner feelings and self-image impact human behaviour</td>
<td>Does not identify human experience</td>
</tr>
<tr>
<td>(Bland &amp; Derobertis, n.d.; Winston, 2015)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biological theory</td>
<td>Holds biological factors responsible for influencing personality of an individual</td>
<td>Genetic determinants affect behaviour</td>
<td>Environmental influence and consciousness are not accounted for</td>
</tr>
<tr>
<td>(Beaver &amp; Walsh, 2018; Farahany, 2015; Tihoen et al., 2015; Veroude et al., 2016)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sociocultural theory</td>
<td>Believes that all groups that an individual belongs to (gender, racial, religion etc.) influence future and decision making</td>
<td>Cultural rules of social interactions shape behaviour</td>
<td>Disregards genetics, cognition, subconscious</td>
</tr>
<tr>
<td>(Azuh et al., 2015; Maunder &amp; Crafter, 2018; Schneider et al., 2016; Zittoun, 2020)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
defense, and diffusion of information. However, this model did not stand up to its claims and only focused on acquisition of different kinds of digital evidence. Another drawback was the feasibility of its implementation to real life scenarios.

Cohen (Cohen, 2010) developed a DF model that concentrated primarily on the examination stage of investigation (analysis, interpretation, attribution, reconstruction). This process model addressed the legal challenges relating to digital evidence, but its practical realization was not demonstrated.

Agarwal et al. (Agarwal & Gupta, n.d.) designed the Systematic Digital Forensic Investigation Model for handling static and volatile digital evidence. This scheme had 11 phases, viz. preparation, securing the scene, survey and recognition, documentation of the scene, communication shielding, evidence collection, preservation, examination, analysis, presentation, and result. The applicability of this model was restricted to digital fraud and cybercrimes.

Ademu et al. (Ademu, Imafidon, & Preston, 2011) proposed an iterative multitiered model for DF investigation where each phase included various subphases. The four tiers were identified as preparation, interaction, reconstruction, and presentation. This model inherited its structure from previous DF models but did not portray a detail of the phases and subphases. Its real-world applicability has also not been tested.

Kohn et al. (Kohn, Eloff, & Eloff, 2013) crafted an integrated model that consisted of six phases and many subphases. In this, documentation was recognized as an incessant process to be executed during all phases of investigation. The phases constituted preparation, incident, incident response, physical investigation, digital forensics investigation, and presentation. The utility of this model was limited to organizational crimes.

Montasari et al. (Montasari, Peltola, & Evans, 2015) proposed an eight-stage model involving readiness, identification, incident response, collection, examination, analysis, presentation, and incident closure. The authors did not provide a detailed description of each phase and hence this model was of very little use to DF practitioners.

Karie et al. (Karie, Kebande, & Venter, 2019) presented a model that identified five stages, namely, (1) initialization, (2) data sources identification, (3) application of deep learning (DL) methods, (4) forensic reporting/presentation, and (5) decision making and case closure. The 5-stage structure focuses on evidence which is the most important and diverse part of DF. An important advantage of using this methodology is that machine learning (Gadekallu et al., 2020) can be implemented in each of its stages. However, the implementation of this design varies depending on the type of evidence.

**Behavioral DF Models**

The DFIPMs discussed in the last section disregard the social, motivational or behavioural scopes of criminal actions that play a major role in recognizing probable evidence during the investigation. There was no DFIPM developed until 2014 that accredited the efficacy of involving behavioural analysis within the DF investigations.

However, the need of integrating other disciplines within the DF investigation process has been realized by researchers lately. There is very limited study in incorporating BEA in an investigation of digital crimes. There exists even lesser literature that describes the utility of behavioural analysis to construe digital evidence in these crimes. The models designed so far have here been classified into three types based on the number of stages, viz. 3-stage, 4-stage, and 6-stage models. Figure 4 presents a diagrammatic representation of the structures of these models.

**The 3-Stage Model**

Silde and Angelopoulou (Silde & Angelopoulou, 2014) proposed a Cyberstalker profiling methodology by integrating BEA into a traditional DF investigation framework. This model involved three steps: (1) discovery/accusation, (2) examination, and (3) analysis. Discovery steps is the survey phase where assessment of the devices of interest is done. The examination stage consists of evidence search and
acquisition, storage, recovery and preservation. Finally, the analysis comprises organization, in-depth review, and reconstruction of evidence.

Each phase includes several investigative processes (e.g., search and collection, recovery, harvesting) and profiling stages (e.g., equivocal forensics analysis, victimology). Additionally, the input (e.g., offender skill level, modus operandi) and output (e.g., evidence location, antiforensics) is defined within each stage. This is more of a theoretical model without practical applicability in different scenarios. It requires the complete details about victim or offender behaviour and activities for proper implementation.

The 4-Stage Model

Mutawa et al. (Al Mutawa et al., 2019) proposed a model that included four phases: (1) review, (2) recognition and collection, (3) examination and analysis, and (4) interpretation and reporting.

The initial stage of review deals with available evidence and observes potential offender motivations, behaviour, and crime scene characteristics. It encompasses contextualization and classification of case as well as prioritization of evidence. The recognition step identifies the authorship of the evidence files stored in digital devices. The third phase involves examining and analyzing the acquired data to produce answers to all investigation questions. The final step attempts to reconstruct the crime based on results of analysis in previous stages. This would then be used to build a comprehensive report.

Here, BEA is involved in the review phase as well as in the examination and analysis phase. Prioritization of evidence in the review stage reduces the time and effort involved. Examination and analysis stages help to identify suspects related to the offense. The 4-stage model examines offender motivations through hypotheses building. However, it is influenced by the availability of significant amounts of digital evidence and case information. It is not useful for digital crimes which involve limited or no human interaction. It proves to be ineffective in crimes where automated tools are used to commit the crime.

The 6-Stage Model

This model by Rogers (M. K. Rogers, 2016) had six phases: (1) case classification, (2) context analysis, (3) data collection, (4) statistical analysis, (5) timeline analysis/visualization, and (6) decision/opinion.

The first phase focuses on identifying the type of case under investigation for example, pornography, identity theft, extortion, homicide, etc. Next step is to understand the context of the case. This allows investigators to focus on the most likely locations of evidence on the suspect’s computer or social media accounts. Data collection is the acquisition of evidence and its storage in a format that can be studied for patterns, linkages, and timeline analysis. The statistical analysis stage conducts a frequency analysis on the web history of the suspect to reveal patterns related to his/her online behaviours. The visualization of a timeline can be used to determine a temporal pattern of the computer system’s usage. In the final step, the analyst will need to craft an opinion based on the results from each of phases and determine the correct finding.

The 6-stage methodology recognizes patterns to create an online behavioural profile of the suspect. It associates a timestamp to evidence, to discover patterns or linkages that were previously missed when the evidence is looked at in an isolated manner. On the negative end, the framework does not mention explicit guidelines for DF practitioners on how to implement it.

Despite the limitations identified in the above models, they offered a useful insight into the development of the model proposed later in this research.
APPROACHES TO BEHAVIOURAL EVIDENCE ANALYSIS IN DF PROCESS

A variety of approaches and techniques have been proposed for BEA in DF investigations. They are based on the human psychology, digital signatures of an individual, or solely on the behavioural traits extracted from the digital evidence.

The Social Media Evidence (SME) analysis (Arshad, Jantan, & Omolara, 2019) approach to BEA collects offender’s behavioural clues from the Open-Source Networks. Information from Open-Source Networks fall into four categories namely, user, activity, network, and content. The metadata, including date and time stamps, is maintained by social media sites for each of these categories. This method provides abundant information about human behaviour and relationships thus enabling to find motive and opportunity involved in crime. However, it is not a straightforward process due to legal and technical issues.

The DL technique of BEA analysis (Karie et al., 2019) involves the incorporation of DL into the analysis of forensic evidence. It identifies the four stages of the DF process and integrates Deep Learning Cyber Forensics (DCLF) as another stage into it. DCLF is an evidence-based approach that contains four stages: evidence collection, storage, preservation, and analysis. This framework suggests the implementation of ML techniques into each stage discussed above and enlists precautions to be
undertaken at every step. However, the implementation of this mechanism highly varies with the type of evidence in question. This technique discussed a generic model using diagrams, suggested use of clustering and classification algorithms in the process and measured the performance in terms of reduced human errors.

The Cognitive Bias analysis approach (Sunde & Dror, 2019) discusses three kinds of cognitive bias with respect to the DF practitioner examining the evidence. These biases are anchoring bias, availability bias and confirmation bias. During decision making, anchoring bias ensues when persons use an initial piece of information to make consequent judgments. The availability bias occurs when people evaluate the probability of an event, or frequency of it happening by the ease with which examples and instances come easily to mind. Confirmation bias is the tendency of individuals to favor information that confirms their previously held views or beliefs. However, this technique limits itself to only three cognitive biases and does not address other existing biases. It also disregards the factors that influence cognitive bias such as training or motivation of the DF practitioner involved. For its design, this approach used data and related information from the case’s evidence pool. The performance was calculated in terms of bias mitigation through peer review and hypotheses testing.

The Criminal Profile analysis (Almond et al., 2019) technique of BEA studies criminal characteristics such as demographics (age and gender), socioeconomic upbringing, psychological traits, trends or technology of crime, social relationships. This method is useful for cases where usual police investigative strategy may not be relevant. A disadvantage of this technique is that it is not an exact science as it is based on forms of prediction and speculation and hence cannot solely be used in an investigation. This approach used the data from SCAS UK database of 651 male stranger rape offenses against females who are 16 years or over. It utilized χ² analysis and logistic regression to determine association between offender behaviour and nationality.

The Behavioural Forensics analysis (Al Mutawa et al., 2019) methodology of BEA focuses on the inclusion of behavioural assessment during the forensic examination of physical or digital evidence. It involves the stages of review, recognition and collection, examination and analysis, interpretation, and reporting. This technique provides a multidisciplinary approach to examination of offender motivations through hypotheses building. However, it is influenced by the availability of significant amounts of digital evidence and case information. The method is ineffective in crimes where automated tools are used to commit the crime and is not useful for digital crimes which involve limited or no human interaction. This approach was devised by the application BEA to 35 real cases of cyberstalking and dissemination of indecent images of children and evaluated using five real digital crimes cases from Dubai Police archives. The performance was measured based on effective focusing of investigation, direction for location of further relevant evidence and interpretation of offender/victim behaviour.

The Digital Thinking Style analysis (Adeyemi, Razak, Salleh, & Venter, 2017) approach to BEA is based on the existence of a unique pattern of human thinking style on the Internet. It believes that the thinking style of persons on the internet can be classified into different categories with each style having a set of features. ML techniques can then be applied easily to each class to analyze it. This method also finds application in e-profile creation, e-learning in access control systems, and for comprehending search engine results. The highest accuracy obtained in the model is still below perfection and there is a smaller number of subclasses extracted for each category of thinking style. This methodology was sketched by considering server-side web data of 43 respondents collected for 10 months. It employed thinking styles measuring instruments and supervised machine learning techniques for finding conclusions. The results were then measured for accuracy, recall, precision based on root mean square error, kappa statistics and F-measure.

The Web Browser analysis (Akbal, Güneş, & Akbal, 2016) approach of BEA investigates a crime by utilizing the suspect’s internet browsing information for gathering evidence. This may include URLs visited by user, search words, search history, download history or cache data. In this, the usage of different web browsers in the same period is examined. A drawback of this technique is that the suspect could have deleted browsers data which may be difficult to recover. Also, the time and data
format used by the web browser must be considered by the analyst. This approach was developed by analyzing evidence from different versions of browsers (Internet Explorer, Google chrome, Firefox, Safari, Opera etc.). The utility of this tool was grounded on how browsers save data, what information can be recovered and how different operating systems store records.

A comparative analysis of these approaches with the benefits and limitations is illustrated below in Table 3.

**PROPOSED APPROACH**

As observed from the literature, no standard approach exists for application of BEA in the criminal investigation process. Due to the absence of a generic BEA procedure, the adoption of behavioural data into forensic examination becomes a tough task. All BEA approaches as listed in Table 3 make use of different behavioural features or patterns, restricting their use only to a particular class of forensic domain. While handling criminal investigations, multiple methodologies may be required to deal with the evidence in question. Thus, it becomes difficult to remember each factor, and one or

<table>
<thead>
<tr>
<th>Approach</th>
<th>Methodology</th>
<th>Benefits</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Media Evidence analysis (Arshad et al., 2019)</td>
<td>Focuses on data and metadata from SM sites</td>
<td>• New, dynamic, with abundant information</td>
<td>• Legal, technical issues</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Metadata aids investigation (timestamps, location tags, device info)</td>
<td>• Massive data</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Contextual view of data and metadata</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Inconsistency, diversity in SM platforms</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• SME not self-authenticating</td>
</tr>
<tr>
<td>Deep Learning techniques in analysis (Karle et al., 2019)</td>
<td>Focuses on evidence (collection, storage/preservation, analysis)</td>
<td>ML can be implemented in each stage</td>
<td>Implementation varies with type of evidence</td>
</tr>
<tr>
<td>Cognitive bias analysis (Sund &amp; Dror, 2019)</td>
<td>Focuses on anchoring bias, availability bias, confirmation bias of DF practitioner</td>
<td>Identifies causes and solutions to cognitive bias</td>
<td>Other factors influence cognitive bias (training, motivation etc.)</td>
</tr>
<tr>
<td>Criminal profile analysis (Almond et al., 2019)</td>
<td>Focuses on offender’s demographics (age and gender), socioeconomic upbringing, psychological traits, technology, social relationships</td>
<td>Narrows down suspects</td>
<td>• Not generic</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Prediction based</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Used with other techniques</td>
</tr>
<tr>
<td>Behavioural forensics analysis (Al Mutawa et al., 2019)</td>
<td>Focuses on BEA (review, recognition and collection, examination and analysis, interpretation and reporting)</td>
<td>• Multidisciplinary approach</td>
<td>• Ineffective for crimes with limited or no human interaction</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Hypotheses building finds offender motivation</td>
<td>• Theoretical model</td>
</tr>
<tr>
<td>Digital thinking style Analysis (Adeyemi et al., 2017)</td>
<td>Based on pattern of human thinking style on Internet</td>
<td>• Creates classes of thinking styles</td>
<td>• Accuracy below perfection</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• ML algorithms easily applicable</td>
<td>• Less subclasses extracted for each class</td>
</tr>
<tr>
<td>Web browser analysis (Akbal et al., 2016)</td>
<td>Analyses use of different web browsers in same period</td>
<td>Confirms occurrence of crime</td>
<td>• Difficult to recover deleted browser records.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Data varies with web browser's type</td>
</tr>
</tbody>
</table>
more important techniques may be left out during analysis. This highlights the need for an integrated system to BEA that will address the limitations of the earlier methods.

Figure 5 shows the proposed design for a standardized model of BEA implementation. It portrays the application of the notion of an integrated BEA approach. This model aims to give a standard process for embedding behavioural analysis into forensic investigations. The proposed approach functions as a homogenous, clearly defined model that covers all necessary factors of behavioural analysis of digital evidence. It assimilates all aspects of BEA into a proper framework that employs them sequentially during the BEA process and ensures that the investigation considers all behavioural cues associated with the offender to yield effective results.

The proposed model is grounded on four basic stages of a forensic model, namely, identification, collection, analysis and reporting. BEA is amalgamated in the third stage that is, analysis stage of the model. In the first stage, devices captured from the crime scene are filtered which results in the identification of the devices of interest. These are devices (computers, laptops, mobile phones, tablets, any other electronic gadget) from which the evidence needs to be extracted. The second stage of the model involves the recovery of digital evidence from these devices. All the digital evidence extracted from this step are fed into the third stage of model that is, analysis. Analysis is the key stage of the proposed model. The input to this stage is the digital evidence extracted in the previous step.

Fundamentally, this phase is accountable for establishing the impact of evidence and drawing conclusions based on the analysis of the evidence. In the proposed model, not only the digital evidence is analyzed but also behaviour is investigated from the identified digital evidence. Three types of analysis are done to investigate behaviour from the DE. These three analyses can be carried out irrespective of the order in which they are performed. First type of analysis is the equivocal analysis in which we infer the “demographic profile” of the suspect(s). In the demographic profile the DE used as parameters are victim statement, witness statement and crime scene photographs. Second type of analysis performed to deduce the behaviour of the suspect is the victimology. In the proposed model, through victimology we try to infer the “risk assessment” and “victim-offender relation.” For risk assessment and victim-offender relation, DE considered are victim’s background, occupation, hobbies, criminal record and drug or alcohol abuse. The third type of analysis done as a part of BEA is identification of crime scene characteristics. In crime scene characteristics, we seek to infer the offender’s “technical expertise” and “social profile.” DE considered for this are the weapon or time of attack, method of gaining entry, type of location, modus operandi, SME, browsing records (URLs with location and time), and digital thinking patterns.

Now, both the “as-is” digital evidence and behavioural evidence are fused to identify the suspect or narrow down the pool of suspects. The “as-is” digital evidence are those which are captured from the crime scene and can be used directly in determining an offender’s characteristics. Some of the “as-is” DE are police records, autopsy reports, degree of harm caused, and forensic evidence like residual files, data, e-mails, photographs, contact numbers, transactions, and so on. The final stage of reporting involves the decision-making process by interpretation of the results and reporting to the requesting party.

Consider an example case involving digital forensics investigation, where a former school volleyball coach pleaded guilty to kidnaping and sexually assaulting two 13-year-old girls he met online. When the two girls went missing, the first-place investigators looked was for the digital clues in their iPods and smartphones. The girls were soon found in the basement of a 23-year-old man, who was then charged with felony criminal sexual conduct, kidnaping and solicitation of a child.

In the investigation of the above case using the proposed approach, the first step involved the identification of the devices where digital clues relating to the victim or offender could be found. These “devices of interest” were identified as the iPods and smartphones belonging to the two victims. The second stage consisted of extraction of all the digital evidences like emails, messages, log-in details to websites or any residing files from these devices. Here, the log-in details or profiles of the two girls on the online dating website played an important role to identify their cyberspace activities.
Under analysis stage, “as-is” digital evidence included e-mails, text messages and data recovered from the victims’ smartphones that were exchanged with the offender. It also included the sexually explicit photographs present on the seized devices indicative of the type of conversation between the girls and the assaulter.

The “crime scene characteristics” included internet as the method of gaining access to the victim, via a social media site that provided avenues for people to “talk to strangers.” In terms of his social profile, this was suggestive that the perpetrator could be either a naive criminal or a serial offender. Next, the social media evidence extracted from the victims’ profiles on the website exposed online sexual chats (Amuchi, Al-Nemrat, Alazab, & Layton, 2012) between the assaulter and the victims. It could be concluded that the offender possessed technical expertise. The “equivocal forensics” consisted of the witness’ statements wherein the parents told the investigator how the girls went out of the home and never returned. It was also revealed that the two girls were not happy at their homes. The parents of the victims provided certain information about the car (number plate details, color) in which the girls were seen last leaving their houses accompanied by a man. Witness statements of parents and neighbors were recorded which described the demographic details of the possible offender for example, facial features of the person driving the car were Mongoloid, who appeared
to be a young man in his twenties. Lastly, the “victimology” phase determined the risk assessment of the victim to fall prey to the crime and the victim-offender relationship. Both the victims were teenagers at a vulnerable age and with easy access to the internet. The unmonitored internet usage mannerisms, on risky dating sites that allowed talking with strangers, put the girls at a very high-risk level. Evidences showed that the relationship between the victim and offender was mainly sexual and that the offender was unknown to the victim before the social media interaction between them.

As a result of the digital forensics investigation and considering the behavioural traits of the offender, the criminal was identified to be a 23-year-old former school coach who lured the girls to leave their homes and kept them in the basement of his house where he sexually assaulted both. The final stage of reporting involved making a detailed report of the case and preparing a charge sheet to be presented in court for the persecution of the criminal. The perpetrator in this case was sentenced to roughly an 11-year prison term.

In the discussed case, the proposed model leads to a rough pen-picture of the assaulter. This it does by deciphering the behavioural mannerisms of the offender from available evidences. The analysis of evidence using this model results in the demographic profile, social profile, technical expertise, of the offender and also his relationship with the victim. Using a flavor of BEA with DF, the method narrows down the offender search, enabling his identification and further prosecution. This was not possible through DF process alone. Rather than sticking to just one forensic domain, the framework covers all classes of evidences like social media footprints, offender demographic or psychological traits, victim habits, digital thinking patterns, web records etc. This holistic approach makes it consistently applicable to almost all kinds of criminal cases. Hence, the proposed design is a step in the direction of standardization of the BEA process and its inclusion in the traditional DF investigations. Figure 6 shows a tabular comparison of the proposed model with the existing studies and how it overcomes the limitations of the latter.

CHALLENGES AND FUTURE DIRECTIONS

Future research directions may include showcasing the extensive utility of the proposed approach by implementing it over more real-life criminal cases. A comprehensive comparative analysis of the existing approaches with the new design must be done by implementing both on the same cases across different crime domains like murder, rape etc.

As the digital investigations pose fresh challenges, further research is required on how BEA can be made faster by using ML (Koul, Becchio, & Cavallo, 2018; Sculati, 2015) and DL technologies. It can be noteworthy to ascertain how their use may be extended to the proposed BEA model. A major challenge is to direct the BEA investigation process in case of insufficient evidence and reconstruction of evidences in these scenarios. It is required to understand what contribution can be harnessed from the latest technologies like Artificial Intelligence (Rigano, n.d.; Russell, Dewey, & Tegmark, 2015) under such circumstances.

CONCLUSION

The exponential growth of technology is impacting all walks of human life. The increasing complexities and capabilities of digital technology has triggered the need for monitoring its use to prevent any abuse. Thus, the ways and methods to deal with digitally committed crimes must also advance to match the new face of digital proficiencies. Hence, there is a need to augment the existing digital forensics processes. The assimilation of Behavioural Evidence Analysis with the DF process boosts the potential of solving criminal cases in this digital era.

The existing approaches to BEA suggest diverse procedures for implementation of BEA thus lacking the required standardization in the process. However, the proposed BEA model aims at drawing a standard system that can be used throughout all criminal investigations. The previously
developed prototypes were largely theoretical, lacking practical utility whereas the new methodology has applicability in real life cases. Though the efficacy of BEA has been elaborated in the reviewed literature but it did not provide a clear procedure to implement it within the DF process. The presented structure overcomes this limitation and gives a detailed step-by-step method of incorporating BEA in forensic investigations.

The earlier techniques focused on one class of evidence or specific characteristics of the forensic domain in question (social media evidence or offender traits or human thinking patterns etc.) and using one technique implied neglecting other available evidences. Conversely, the new framework considers all types of evidences and influencing factors thereby extending the use of BEA technique over several forensic domains across all genres of digital crimes. Previously, a combination of approaches was required to encompass all evidences leading to delay in solution. On the contrary, the given prototype is a time saver design as all factors to be analysed are considered in the same system simultaneously. The proposed approach visibly overcomes the limitations of the previously developed BEA models. It is applicable to not only the current digital crimes, but can also prove useful for new unseen crimes in the future. Some challenges that need to be handled and scope for future work have also been discussed in the paper for pursuing research in this domain.

**Figure 6. Tabular comparison of proposed model with existing studies**

<table>
<thead>
<tr>
<th>Existing approaches</th>
<th>Proposed model</th>
</tr>
</thead>
<tbody>
<tr>
<td>• No standard approach for application of BEA in DF process. Different literatures suggest diverse procedures leading to inconsistency in use of BEA</td>
<td>• Provides a standard approach that can be used consistently in all criminal DF investigations</td>
</tr>
<tr>
<td>• Mainly theoretical models lacking practical applicability</td>
<td>• Practically applicable to real life criminal cases</td>
</tr>
<tr>
<td>• No clear procedure of “how” to apply BEA stepwise in the forensics process</td>
<td>• Well-defined, unambiguous, and extensive technique lucidly describing each step in application of BEA</td>
</tr>
<tr>
<td>• Focussed on one particular class of evidence or on specific characteristic of the forensic domain in question</td>
<td>• Considers all types of evidences (SM evidence, web records, human thinking patterns etc), spanning a variety of forensic domains</td>
</tr>
<tr>
<td>• Integration of approaches required to encompass all evidences leading to delay in solution</td>
<td>• All-in-one, time saver approach as all factors/evidences are considered in the same system simultaneously</td>
</tr>
</tbody>
</table>

---

36
REFERENCES


Barkhashree received her MSc in Computer Science from the University of Delhi in 2015. She has worked as Assistant Professor, Computer Science at the University of Delhi. She is currently a research scholar at Manav Rachna University, Faridabad Haryana.

Parneeta Dhaliwal (PhD) completed her M.Tech and PhD from Netaji Subhash Institute of Technology, University of Delhi. She is currently working as Associate Professor at Manav Rachna University, Haryana and is guiding many PhD students. Her areas of interest are Machine Learning, Blockchain, Big Data, and Digital Forensics. She has published several papers in International journals of high repute and has worked on many research projects sponsored by the Government of India.