Research Article

Digital Forensics Challenges in Cyberspace: Overcoming Legitimacy and Privacy Issues Through Modularisation

Moses Ashawa1, Ali Mansour2, Jackie Riley1, Jude Osamor1, Nsikak Pius Owoh1

1Department of Cyber Security and Networks, Glasgow Caledonian University, Scotland, United Kingdom
2Department of Computing and Information Systems, University of Bedfordshire, England, United Kingdom
Email: Moses.Ashawa@gcu.ac.uk

Received: 31 October 2023; Revised: 18 December 2023; Accepted: 18 December 2023

Abstract: The significance of the cloud environment is growing in the current digital world. It provides several advantages, such as reduced expenses, the ability to adjust to different needs, adaptability and enhanced cooperation. The field of digital forensic investigations has encountered substantial difficulties in reconciling the requirement for efficient data analysis with the increasing apprehensions regarding privacy in recent times. As investigators analyse digital evidence to unearth crucial information, they must also traverse an intricate network of privacy rules and regulations. Given the increasing prevalence of remote work and the necessity for businesses to be adaptable and quick to react to shifting market circumstances, the cloud infrastructure has become a crucial asset for organisations of various scales. Although the cloud offers benefits such as scalability, flexibility and enhanced collaboration, it presents difficulties in digital forensic investigations regarding data protection, ownership and jurisdictional boundaries. These concerns are becoming increasingly significant as more data is kept in the cloud. In this paper, we present three major challenges that are faced during cloud-based forensics investigation. We analyse the extent to which different data formats increase complexity in forensics investigations in cyberspace. This paper analyses three core challenges facing digital forensics in the cloud environment: legitimacy, complexity and an increase in data volume, looking at the implications these have on data liable for legal issues in court. These challenges contribute to the backlog in digital forensics investigations due to a lack of modularisation of the procedures. To address these concerns, modularisation model is proposed to offer a way to integrate traditional processing functions while ensuring strict adherence to privacy protocols. To overcome these challenges, we propose modularisation as a strategy for improving the future of digital forensic research’s operational efficiency, overcoming the identified challenges faced during cloud-based investigations and demonstrating how organisations can mitigate potential risks associated with storing sensitive information in the cloud.

Keywords: legitimacy and privacy, dynamic allocation of resources, cloud-forensics, data increase, standardisation, modularisation

1. Introduction

The rapid growth of digital technology and its impact on society have given rise to numerous challenges faced by forensic examiners, especially in cloud-based forensic investigations. Cloud-based digital forensics is becoming...
increasingly important as more businesses move their operations and data to the cloud. However, this shift has also led to new challenges for digital forensics investigators, such as the difficulty of accessing and analysing data stored in a cloud environment, data volume increase, complexity and legitimacy. These challenges and many more have posed significant challenges for digital forensics investigations in terms of technical complexities and legal considerations, leading to issues related to data privacy, ownership, and jurisdictional boundaries. For example, in a case of a cyber-attack on a company’s cloud infrastructure, investigators may face difficulties obtaining data access due to the cloud provider’s security measures and privacy policies. These challenges constitute substantial difficulties in reconciling the requirement for efficient digital forensic analysis with privacy and legitimacy in recent times.

For instance, when analysing cloud-based digital evidence to extract crucial information, examiners must also navigate a complex network of privacy rules and regulations during investigation, such as obtaining proper legal authorization and ensuring compliance with data protection laws. Legitimacy issues arise when considering the reliability and admissibility of digital evidence obtained from cloud platforms. The dynamic nature of cloud computing and the potential for data alteration or tampering further complicate the forensic analysis process. As a result, digital forensic examiners face the ongoing challenge of striking a balance between the need for efficient analysis and respecting privacy rights while also ensuring the legitimacy of the evidence collected from large data volumes from different locations.

In cloud forensics, an investigation may involve multiple jurisdictions if the cloud provider is located in a different country than the victim organisation, leading to legal complexities and challenges in coordinating international cooperation. Apart from legal complexities, other challenges arise in finding and extracting legitimate and reliable evidence due to the sheer volume of data generated by cloud-based applications and services, which is relatively different from traditional digital forensic investigation. Digital forensics science examines digital artifacts in a systematic and coherent study of digital evidence to answer authentication, reconstruction, and classification questions to assist criminal investigations and prosecution. At its inception in late 1984, digital investigation was focused mainly on extracting and analysing digital evidence on computer devices and related components; however, as technology advances, more of our daily lives are conducted online, and as a result, over 90% of crimes nowadays are recognised as having digital elements [1]. As there are advanced forensic tools and skilled digital forensic practitioners, an increase in the number of sophisticated crimes connecting multiple electronic devices may not be successfully analysed with today’s digital investigation tools. The reason is due to some of the digital challenges, such as encryption, incompatibility of data formats, increases in data volume, data management issues, a fragile commercial market, complexity, and legitimacy. These issues and many more have constituted core challenges in digital forensics. Without a clearly stated plan to enhance the efficacy of digital forensic tools and processes and to enhance their relevance for the future, the tools and processes may be degraded and become obsolete in analysing emerging digital crimes without tempering privacy and evidence legitimacy.

To address these challenges, there is a need for a framework that will create a balance between classic and disaster digital forensic investigations. This framework should incorporate traditional forensic techniques while also accommodating the unique aspects of digital evidence in disaster situations. It should prioritise the preservation and collection of digital evidence, considering the potential challenges of disrupted infrastructure and limited resources. Additionally, the framework should emphasise the collaboration between various stakeholders, including law enforcement agencies, disaster response teams, and digital forensic experts, to ensure a comprehensive and efficient investigation process.

In this paper, we made the following contributions:

- First, the paper presents a comprehensive literature review, analysing previous studies on the next-generation digital forensic readiness by examining the key components and best practices of these frameworks.
- Secondly, we examine the core digital forensic challenges associated with cloud-based forensic investigation. Cloud-based forensic investigation presents several challenges due to the unique nature of cloud computing. One of the core challenges is the lack of physical access to the cloud infrastructure, making it difficult to collect evidence directly from the physical storage devices. The reliance on third-party providers for data storage and processing introduces issues related to jurisdiction, privacy, and data integrity, which can complicate the forensic process.
- Lastly, the paper presents a modularisation framework for cloud-based digital forensic investigation. Modularisation framework takes into consideration the unique challenges posed by cloud computing environments, such
as the distributed nature of data storage and the dynamic allocation of resources. Modularisation will improve the future of digital forensic research’s operational efficiency and overcome the identified challenges.

The remainder of the paper is structured as follows: Section 2 provides critical analysis of related works in traditional digital forensic investigation processes and frameworks. The section provides a critical analysis of related works in traditional digital forensic investigation processes and frameworks. This section highlights the gaps, strengths, and limitations in the existing previous studies. Section 3 presents the core cloud-based digital forensics challenges, such as the increase in data volume, complexity, legitimacy, and security, and looks into the challenges of preserving the integrity of digital evidence in a cloud environment in relation to privacy and legitimacy. Section 4 presents the proposed framework. In section 5, we discussed the scope and performance analysis of the proposed model. The conclusion and future work is presented in section 6 accordingly.

2. Related works

Different works in digital forensics exist that involve ontologies, creating file formats, thumbnails, and schemas. However, there remains limited work on digital forensics modularisation to project the future of digital forensic research on tools and processes. In bringing digital forensics into education, Garfinkel explained the need for a data standardisation set by introducing the concept of corpora as a strategic approach to furthering digital forensic research [2]. Many instructors use corpora for teaching work, though the approach has yet to be widely adopted. Over time, digital forensics processes have changed because of technological advancements. Many emerging devices are IoT-oriented, which has the benefits of remote access to servers with digital forensic tools and large volumes due to cloud-connected applications. To this end, many changes have been seen in the digital forensic process. Even though applied digital forensics process models are not standardised [3], there are international standards such as ISO/IEC 27037:2016 [4] that provide requirements [5] and principles [6] for handling the digital investigation process and evidence. IoT devices require extensions and model changes in digital forensic analysis; hence, there is a need to update ISO/IEC 27037:2016 standards to address cloud forensics challenges associated with these devices.

Ali and Kaur [7] designed a framework called the “Next-generation digital forensic readiness Bring Your Own Device (BYOD) framework”. Their framework focused on detecting and protecting the BYOD environment from advanced persistent attacks, which usually constitute threats and can damage critical infrastructure. Their work reviewed the requirements for detecting and preventing those attacks by conducting digital forensic investigations and model reviews to find digital evidence showing that some attacks originate from cloud service vulnerabilities. Threats associated with BYOD and the cloud represent challenges as many commercial and corporate networks become exposed to intrinsic threats through them. Even though some elements of modular computer forensics are included in some existing forensics frameworks, however, many of the existing frameworks failed to satisfy the requirements for parallelism, which is essential for digital forensics modularity [8].

Alex and Kishore developed a framework for collecting digital forensic artifacts outside the cloud service environment to extenuate cloud service providers’ dependency by introducing a forensic monitoring plane for inbound and outbound monitoring of cloud environment connections [9]. Bit-by-bit forensic imaging of encryption streams can be achieved and stored as logs by a forensic server, and the forensic investigator does not have to interact directly with the cloud service provider. During the investigation, bit-by-bit stream images can be analysed separately. However, the challenge in this model is identifying and selecting the stream that contains the sought-after evidence. Permutation by timeframe can be deployed to determine the stream or data frame that contains evidence; however, this can be time-consuming, and circumstances where two or more user streams separate evidence will lead to an extra task of evidence reconstruction.

Generally, when stream-based forensics are required in an investigation, it is easier to do it on the disc by reading from the beginning to the end of the disc than on a cloud. Evidence of reconstruction and building of relevant file systems and boundaries hierarchy becomes easier and provides data retention and surplus information [10]. However, evidence reconstruction has issues associated with reliability assurance in the procedure, tool, method, examiner, and data set domains. Issues associated with evidence reconstruction introduced accountability and transparency questions concerning the digital investigation, leading to a compliance problem with Association of Chief Police Officers (ACPO)
guidelines and other scientific requirements for forensic evidence, as well as fragmented digital forensic processes and procedures, especially when it involves cloud investigations, especially when depending on the server [11]. It is risky and unwise to completely depend on the forensic server and cloud service provider’s logs and audit trails without proactively collecting and preserving digital forensics data.

3. Preliminary studies

3.1 Core digital forensic challenges

The cloud computing model offers essential services with a broad spectrum of applications in science, including digital forensic science, to enhance the investigation and establishment of facts of interest concerning criminal, civil, and other regulatory laws. Cloud development services are useful in applying scientific and technological principles to reconstruct cloud-related events by identifying, acquiring, preserving, interpreting, and reporting potential evidence utilising proven methods. Cloud-based digital forensic investigations involve the application of a cloud model. Even though the cloud model is composed of essential characteristics such as on-demand self-service, where a customer can unilaterally access computing capabilities, including server time and network storage without needing human interaction, Services such as measured service, resource pooling, broad network access, and rapid elasticity are other essential characteristics that the cloud model offers that constitute challenges to digital forensics in different domains. There are several issues and challenges that digital forensics is faced with, including technical [12], legal [13], and resource [14].

This paper discusses some of the cloud-related challenges the authors considered the most significant challenges facing digital forensics worldwide. Though the challenges discussed have a broad cross-section ranging from anti-forensics, legal jurisdiction, training, architecture, technical standards, and practises, the authors focused their analysis on the challenges with a cloud-based environmental root cause. The following subsections analyse the core challenges facing digital forensics.

3.1.1 Increasing data volume

While humans may be tech-savvy, the increasing use of digital devices can lead to information overload and a lack of understanding of how to analyse and manage complex data sets properly. As a result, there is a need for efficient and effective data management and analysis tools to make sense of this data and derive insights that can inform decision-making and drive business growth. With the right tools and skills, humans can leverage technology to unlock the full potential of their data and gain a competitive edge in today’s digital age. The increasing use of digital devices has led to the generation of complex data sets, making efficient data management and analysis tools crucial for deriving insights and driving business growth. With the right tools and skills, humans can harness technology to unlock the full potential of their data and gain a competitive advantage in today’s digital age. Due to the rapid change in data storage, many digital devices have their storage media connected to the cloud. Cloud platforms for data storage have an ever-expanding capacity, increasing from terabytes to zebibytes (see Figure 1). Therefore, individuals and businesses need to adapt to this changing landscape and utilise cloud-based data management solutions to store and analyse their data effectively. By doing so, they can save on physical storage costs, access their data from anywhere, and collaborate seamlessly. The future of data management is in the cloud, and it is up to us to embrace this technology and leverage it to our advantage. This data is generated and stored by both individuals and organisations.

However, digital investigation and analysis connected to the cloud present a big challenge to digital forensics research because the larger the storage size, the more acquisition and analysis time is required. This challenge of increasing data volume causes a delay in clearing backlogs and bringing offenders to justice in a timely fashion. While cloud-based data storage may offer benefits in terms of accessibility and cost savings, it also presents challenges for digital forensics research due to the increased time required to acquire and analyse large volumes of data. Although many digital practitioners strongly support using automated processes to reduce case backlogs, automation jeopardises the quality of the evidence even though it may reduce the number of case backlogs. It is important to consider the potential risks and benefits of automating case management processes before implementing them. While automation may reduce case backlogs, it may also compromise the quality of evidence, which could have serious consequences.
in legal proceedings. It is crucial to strike a balance between efficiency and accuracy when implementing automated processes in case management. It is important to explore alternative solutions to reducing case backlogs that may not compromise the quality of evidence. For example, increasing staffing levels, improving training and development programmes, and implementing better communication and collaboration processes can all help reduce case backlogs without compromising the quality of evidence. Ultimately, it is essential to carefully weigh the potential risks and benefits of any automation solution before implementing it and continually monitor and evaluate its effectiveness to ensure that it achieves the desired outcomes.

While automation can be a useful tool for reducing case backlogs in case management, it is important to consider the potential risks and benefits before implementing it. It is crucial to strike a balance between efficiency and accuracy and explore alternative solutions that may not compromise the quality of evidence. It is also important to continually monitor and evaluate the effectiveness of any automation solution and investigate how other industries have successfully implemented automated processes without compromising quality. By following these principles, case management can benefit from automation while maintaining.

The challenge of an increase in data volume impacts not only law enforcement but also victims, witnesses, and suspects. Investigating cloud services varies per service model that is being provided to the client, such as public cloud [16], private cloud [17], and hybrid cloud [18]. For example, if a client is compromised and a request is generated to a cloud service provider (CSP), the overall success of the investigation relies heavily on the cooperation of the CSP if they are willing to provide both logical and physical access for the investigation. In many cases, relying on the laws and legal procedures for acquiring the desired permissions fails because of the border laws, and hence the investigation becomes unsuccessful. Sometimes, the delay or inability of the CSP to give permission or to provide logical and physical access can be due to the cloud being compromised at the CSP site [19]. This is usually due to a lack of logical and physical security policies. Other concerns, such as data volatility, issues of geographical locations, cloud forensic expenses, remote access security, the high cost of forensics, and the ineffectiveness of traditional forensics tools, among others, are issues associated with cloud forensics challenges and the increasing volume of data. A detailed counterexample is a case of Apple refusing to unlock an iPhone used by a suspect in a terrorist attack, citing privacy concerns and setting off a legal battle with the Federal Bureau of Investigation (FBI). This scenario demonstrates how, even with legal procedures and permissions, cooperation from technology companies can be difficult to obtain in cloud service investigations.

![Figure 1. Data size increasing to zeibibyte [15]](image-url)
3.1.2 Complexity

The movement of organisations’ and individuals’ data into the cloud has increased the complexity of digital investigations as data is not confined to a single host [20-21]. Devices and infrastructure connected to the cloud have their data shared between physical and virtual devices. During an investigation, there are insufficient techniques to maintain digital forensic capabilities for investigating interrelated cases. Figure 2 demonstrates that classic forensics uses less complex and sophisticated tools and methodologies to investigate crimes on single digital devices, while disaster forensics deals with investigations in the larger world with no confined data. By implication, the complexity of the investigation and the consequences of having data liable for legal issues in court increase. In addition, investigating crime in the larger world with many data connections to many hosts requires more expertise, complex tools, and time to reconstruct evidence completely and correctly in such scenarios. These constraints and many more impact the quality of the investigation and make it very complex.

The complexity challenge in digital forensics research has recently increased due to the rising volume of digital examinations and investigative device submissions. Many of these devices and applications have different data formats, some with end-to-end encryption, which requires the examiners to have deep expertise in digital forensics and a body of cyber security knowledge to examine, respectively successfully. Rapid growth in the Internet of Things (IoT), which in a few years may exceed human interactions, is another key issue adding complexity to digital forensics research. This implies that to sustain the abilities of digital forensics research, there must be new data extraction and analysis techniques to keep in line with the criminal justice system and avoid being cut off from artifact sources. Criminal defence includes cloaking techniques such as encryption, obfuscation, and information hiding.

Figure 2. Increased cloud storage increases the complexity associated with classic and disaster forensics

Despite cooperation between international jurisdictions, investigating cybercrime and collecting evidence are essential to building a firm case. To do this, security professionals need the best research tools. Digital forensics is the basis of real-life research often closely associated with cyber expansion. Today’s digital society is exposed to cyber criminals and fraud that pose financial losses and personal dangers; therefore, the new wave of forensic tools needs to be designed to support heterogeneous research that protects privacy and provides scalability. Additional complex challenges arise from anti-forensics techniques, whereby criminals deploy different techniques such as encryption, steganography, covert channels, tail obfuscation, and data wiping to evade detection and make forensic investigation more difficult. For example, in a recent cybercrime involving a major corporation, the criminals used encryption and steganography techniques to hide their activities and evade detection. However, with the help of advanced digital forensics tools and techniques, investigators were able to uncover the hidden data and build a strong case against the perpetrators, resulting in the successful prosecution and the recovery of stolen assets. However, it is important to note
that even with advanced digital forensics tools, there can still be challenges in uncovering hidden data. For instance, if criminals use strong encryption algorithms and take proper measures to secure their communication channels, it can be extremely difficult to recover the data and build a case against them.

3.1.3 Legitimacy and security

The challenge of legitimacy in digital forensics investigations in the cloud is a complex issue that requires careful consideration of various factors such as privacy, security, and jurisdiction. As more organisations and individuals move their data to the cloud, it is essential to develop robust and transparent policies and procedures for conducting digital forensics investigations that can withstand legal scrutiny and ensure the integrity of the evidence collected. For example, when a company’s confidential information has been leaked from its cloud storage, the digital forensics investigation must follow strict protocols to ensure the evidence collected is admissible in court. The forensic process may involve obtaining proper legal authorisation, preserving the chain of custody, and conducting the investigation in compliance with privacy laws and regulations. However, this is easier said than done, as evidenced by a case in which the FBI seized servers belonging to a web hosting company, resulting in the deletion of innocent customers’ data and disrupting their businesses. This incident highlights the need for clear guidelines on how digital forensics investigations should be conducted in cloud environments to prevent collateral damage.

As illustrated in Figure 3, many modern infrastructures are now virtualised, and organisations contract and delegate services to third parties across different parts of the world. Most countries working on delegated services or providing them do not have defined investigation guidelines or principles. For example, some organisations in the UK contract and delegate their services to other countries and continents, such as Africa, Asia, and America, among others. Digital investigation cases and evidence from such countries may be questionable in a court of law because of the Lack of standardised investigation guidelines. Also, even in countries with those principles, executing investigations legitimately without violating laws in bordered and borderless scenarios becomes challenging. As a result, legacy data handling in those scenarios is complex due to varying regulations and policies on data preservation and retention. According to the research [22], a lack of data localisation on a single host results in the failure of scientific validation of digital artefacts. For instance, a UK-based healthcare organisation may delegate its medical transcription services to a company in India. If there is a data breach or cybersecurity incident, investigating the incident may be challenging due to differences in investigation guidelines and laws between the two countries. Additionally, if the investigation involves legacy data, it may be difficult to preserve and retain the data in compliance with both UK and Indian regulations.

Many countries may have a low or no budget for legal aid, and winning cases in court mostly depends on the strength of the argument in the courtroom and not on the legitimacy of the evidence. In addition, there is unaffordable expertise in most of these countries, even when the case is to be investigated in the country where the crime was committed. As a result, many organisations will resort to offering the investigation case to the cheapest company or
individual, regardless of the quality of the investigation and the integrity of the evidence. Though the integrity and legitimacy of such evidence from the investigation outcome may not be proof, the defendant has no financial fitness to order or contract the services of a competent investigator, leaving the defendant at risk.

Another issue bordering on legitimacy is regulation and constant amendments in the General Data Protection and Regulations Act [23], which have increased civic consciousness of personal data privacy. Other factors, such as religion, can also constitute a problem in those bordered and borderless scenarios during an investigation, prohibiting certain forms of data, such as pornography and blasphemy, from being accessed. For example, in a country with low budgets for legal aid, a defendant accused of a crime may have to rely on a cheap investigator without the necessary expertise or integrity. This can result in using illegitimate evidence in court and increase the risk for the defendant. Additionally, religious constraints may limit access to certain data types during investigations, further complicating the process. For example, in a country with limited resources for legal aid, a defendant may not be able to afford a competent investigator and may be forced to rely on subpar evidence. Additionally, religious restrictions may limit the types of data accessed during an investigation, further complicating the search for legitimate evidence.

3.2 Privacy and security challenges

Cybercrime’s increase has culminated in substantial business model challenges such as Crime as a Service (CaaS) [24], which has become organised crime [25]. Crime as a Service provides different attack vectors that threat actors deploy to attack networks and digital infrastructures by providing various access points. These access points range from vulnerabilities in applications, tools, frameworks, and techniques to cloud services. We analyse some key cloud-based digital forensics issues, and the positive impact that overcoming these issues through standardisation and modularisation would have on current and future digital forensic research. For example, a company may fall victim to a CaaS attack where the threat actors use compromised cloud services to gain access to sensitive data. The digital forensics team would need to investigate the attack and trace it back to the source, but this process can be complicated due to the lack of standardisation among cloud service providers and the modularisation of their systems. Overcoming these issues could streamline the investigation process and improve prevention and response measures.

3.2.1 Use of fog metadata for authentication

By leveraging the metadata associated with a fog instance, users can authenticate themselves and gain access to the resources they need. This approach can provide an additional layer of security and streamline the authentication process for cloud-based applications and services. For example, a cloud-based application could require users to authenticate themselves using their fog instance metadata, such as their IP address and region. This would ensure that only authorised users with valid Fog instances can access the application, reducing the risk of unauthorised access and data breaches. However, this approach can also be vulnerable to attacks such as IP spoofing, where an attacker can mimic a valid fog instance and gain access to the application. Additionally, if the fog instance metadata is not properly secured, it can be easily accessed by unauthorised users and used for malicious purposes.

Data movement or relocation within the cloud environment has significant consequences for fog metadata, such as its creation time and the last time the data was modified or accessed. This significant metadata may change during its collection and migration process. When this happens, it results in a metadata authentication problem [26], and sometimes it could cause metadata damage [27]. This is a major challenge, especially during the e-discovery procedure for data or information governance. Even though other forensic experts suggest using other metadata, such as fog login files, for versioning fog metadata authentication, this does not provide wholesome integrity on a data file in the cloud, and as such, it is not suitable for this cloud challenge.

3.2.2 Data location and dependence on cloud providers

Choosing a reliable and secure cloud provider that can guarantee data privacy and protection is important. The location of data and dependence on cloud providers are important considerations for businesses and individuals who store and access data online. Additionally, understanding the location of data can help ensure compliance with local regulations and laws. For example, a multinational corporation with offices in different countries may need to store
sensitive company data in the cloud. They would need to choose a cloud provider with servers in each country where they operate to comply with local laws and regulations regarding data privacy.

Due to the multiple tenants in various data centres [28], the issue of transparency and distribution restrictions arises in the cloud regarding information or data retrieval. Access to resources suspected of being related to a crime involves an expanded scope, possibly including many geographical locations. Forensic parameters such as chain of custody, resource identification, and evidence identification are essential for accessing the system of unknown geolocations that can impact investigations. Overcoming this challenge will enhance the discovery and retrieval of data from multiple data centres. Knowing the geolocations or venues would make it easier to find digital evidence and identify all the resources needed to maintain the chain of custody if the challenge is overcome.

With the help of user credentials, acquisition can be done without necessarily involving the assistance of cloud providers. However, to comply with the legal procedures, it becomes necessary to cooperate with the providers, and in most cases, with limited return cooperation. Overcoming this challenge could lead to better-equipping cloud providers regarding human and physical resources. Thus, assisting forensic investigators during investigation for data retrieval to collect complete evidence by specifying it within the summons. Similarly to the challenge of imaging cloud evidence, having all digital evidence completely imaged in the cloud is practically impossible, even though fractional imaging can be done. However, admissibility is associated with this when the evidence, often partially imaged, is presented in court. Overcoming this challenge will streamline the cloud imaging process and enhance evidence acceptability in a court of law.

3.2.3 Synchronising network forensics timestamps

Proper timestamp synchronisation is essential for effective network forensics analysis and can greatly improve the accuracy and reliability of investigative findings. This process could be crucial for accurately reconstructing digital events and identifying potential security threats. By ensuring that all devices on the network use the same time source and are synchronised to the same time, investigators can accurately correlate events and establish activity timelines, which can be achieved using Network Time Protocols (NTP) or Precision Time Protocol (PTP). For instance, in a cyber attack investigation, multiple devices may need to be analysed to determine the sequence of events. By synchronising the timestamps across all devices, investigators can accurately piece together the attack’s timeline and identify the breach’s source.

Timestamp synchronisation is one of the main difficulties in network forensics investigations. This is because a cloud-based inquiry involves many physical equipment dispersed across several geographical locations between cloud structures. Cloud infrastructures have a large number of distant web clients at various stations. Timestamp synchronisation is crucial for the validity and admissibility of evidence [29]. Improving the link between artifacts and timeline size and overcoming network forensics timestamp synchronisation throughout the cloud ecosystem would boost investigation and evidence relevance. For example, in a cyber-attack investigation, the timestamps of various log files from different cloud servers need to be aligned to identify the exact sequence of events leading to the attack. Without accurate timestamp synchronisation, pinpointing the attack’s source and assigning responsibility would be difficult. In this case, advanced network forensics tools and techniques that address timestamp synchronisation could help streamline the investigation process, reduce errors, and expedite incident response. Therefore, it is important to address the challenge of timestamp synchronisation in network forensics investigations on cloud infrastructures to improve the accuracy and admissibility of evidence. This can be achieved by improving the link between artifacts and timeline size, ultimately enhancing the investigation process and the relevance of evidence.

3.2.4 Lack of standardisation

Lack of standardisation of scientific principles in cloud forensic investigations can lead to inconsistencies in the results obtained and hinder the overall effectiveness of the investigation process. This poses significant challenges for investigators. Without clear guidelines and best practices, investigators may struggle to collect and preserve evidence consistently and reliably, leading to errors, inconsistencies, and potentially compromised investigations. The scientific community must work together to establish a standardised approach to cloud digital forensics investigations, ensuring that evidence is collected and analysed systematically and transparently. This will help improve the reliability and
accuracy of investigations, ultimately leading to more effective and poses a significant challenge for investigators and forensic analysts. Without a standardised approach, it becomes difficult to ensure the accuracy and reliability of digital evidence, which can have serious consequences in legal proceedings. Therefore, the industry must develop and adopt standardised scientific principles for cloud digital forensics investigations to ensure the integrity of evidence and maintain public trust in the justice system. For example, in a case involving a cyber-attack on a cloud-based financial institution, investigators may need to collect evidence from multiple cloud service providers and various virtual machines. Collecting and preserving this evidence could be complicated without clear guidelines, leading to potential errors and compromised findings. A standardised approach would ensure that all evidence is collected consistently and reliably, improving the accuracy of the investigation and increasing the likelihood of successful prosecution. Therefore, the scientific community must establish standardised principles and guidelines for cloud forensic investigations to ensure reliable and consistent results, which will improve the quality of investigations and increase trust in cloud services and their security measures.

Cloud forensics lacks standardisation in the forensic testing and validation of the investigation tools, principles, techniques, and models [30]. The lack of cloud forensics standardisation hinders forensic evidence soundness and reliability due to the process and tools used for acquiring evidence. This shortage leads to issues of admissibility, competence, and trustworthiness of the cloud provider (CP) as an ineffective and immediate response when evidence occurs on the side of the CP due to missing terms in the service level agreement (SLA) or contract. As highlighted in [31], one of the security challenges in the cloud during the investigation is that CP may be more concerned with service restoration than preserving evidence when an incident occurs. In order to preserve their reputation in severe cases, many cloud providers would prefer not to report the incident or cooperate in the digital investigation process due to fear of losing their reputation with customers and the public. Further, due to the difference in SLA with their clients, there are missing terms in the contract or agreement stating the requirements that the CP maintain or produce appropriate artifacts within the stated time frame constraints. Overcoming these issues will enhance evidence-based soundness, method validity and repeatability, and admissibility in a court of law. Also, it would make digital forensics investigations uncomplicated and facile by guaranteeing that CPs handle a compromised cloud environment as a crime scene to preserve the genuineness, integrity, and admissibility of essential digital evidence. It would, in turn, clearly delineate the responsibilities of CP for timely forensic data provision and preservation.

Standardisation is a core principle for future operations processes to ensure future improvement in digital forensic research. Cloud-based digital forensics faces standards challenges, such as the absence of basic standard operating procedures, the Lack of cloud service providers’ interoperability, and the Lack of testing and validation processes and procedures. Lack of standardisation of hardware, software, techniques, and policies surrounding cloud forensics [32] across industries and several countries brought in the challenge of evidence dependability and soundness, which are acquired using those tools and policies. Several unsuccessful attempts have been made to create new formats and abstractions [33-34] when representing an individual’s whole life’s data in a few bytes of information informing signature pen tester metrics, user profiles, and filesystem metadata. It may be difficult and practically impossible to standardise all the digital forensic tools, models, and processes due to some form of diversity in the procedures. However, high-volume processes, tools, models, and procedures can be identified for standardisation. Thus, aiding forces to converge on specific and single models and methods. Standardisation will enhance the process and design validation for training and development to avoid duplication of processes nationally. Customary technology can be adopted more simply once digital forensic research processes are standardised.

The cost of standardising processes and models can be very high due to the required time, resources, and expertise [35]. Nevertheless, the future operation of digital investigation research is more important than the cost of standardising it now. Achieving this will provide flexibility in the tools, model-switching process, and components for new solutions.

3.2.5 Geographic borders

The challenge of geographic borders, jurisdiction, and a lack of network protocols in cloud forensic investigation is a complex issue that requires careful consideration. The global nature of cloud computing means that data can be stored in multiple locations, making it difficult to determine the appropriate jurisdiction for an investigation. Additionally, the lack of standard network protocols in cloud environments can make collecting and analysing data challenging. As such, there is a need for a coordinated approach to cloud forensic investigation that considers these challenges and works
to establish standardised protocols and procedures for investigating cloud-based crimes. Implementing the proposed technique will require collaboration between law enforcement agencies, cloud service providers, and regulatory bodies to ensure that investigations are conducted effectively and efficiently while respecting privacy and data protection laws. Addressing these challenges will ultimately ensure the security and integrity of cloud-based data and maintain trust in the digital economy. For example, in cybercrime, where the perpetrator uses a cloud service provider to store and transfer data across borders, identifying the appropriate jurisdiction for investigation can be difficult. More complication is experienced by the dearth of standard protocols in cloud environments, which may result in important digital evidence being overlooked or improperly handled during the investigation process. To overcome these challenges, law enforcement agencies and cloud providers may need to work together to establish protocols for cross-border investigations and develop specialised tools for forensic analysis in the cloud.

Understanding system boundaries in cloud computing is usually challenging because the cloud environment is composed of multi-tenanted environments that are complicated to define by digital investigators. The operation of many cloud services in multiple countries with different legal frameworks governing data protection and accessibility and no universal treaty impedes investigations in multiple cloud-tenanted environments. To overcome this challenge, the investigation range should be concentrated and focused within appropriate confines to ensure that all the forensics data collected is relevant to the investigation. This can empower law enforcement and investigators to obtain data relevant to the investigation effectively. In many cloud environments, transfer protocols such as TCP/IP v6 dumps and TCP segments are unavailable [36]. These transfer protocols are essential for forensic examinations to discover more evidence.

Overcoming this cloud forensic challenge could help forensic examiners have access to network traffic dumps to discover other relevant digital forensic evidence. For example, in a case involving intellectual property theft, investigators could focus their investigation on the cloud service provider and the specific user account used to access and transfer the stolen data. By obtaining relevant data from this specific account, such as login times and IP addresses, investigators could track down the perpetrator and build a strong case against them. Additionally, by using specialised tools to extract network traffic dumps from the cloud environment, investigators could uncover additional evidence related to the theft, such as communications between the perpetrator and accomplices or other suspicious activity on the network. Therefore, forensic examiners must adapt to the challenges of cloud computing and develop new techniques and tools to gather and analyse digital evidence in the cloud effectively.

4. Proposed model

Modularisation involves breaking the investigation process into smaller, more manageable modules to analyse and gather evidence from cloud-based platforms effectively. By doing so, investigators can focus on specific areas of interest and reduce the risk of missing important information. This approach allows for easier collaboration among team members, leading to a more efficient and thorough investigation. For example, a forensic examiner may modularise their investigation of a cloud-based email service by analysing the metadata of each email separately before moving on to the content of the messages. It ensures a more efficient and organised examination process. This approach saves time and resources and allows for a more thorough examination of each module, leading to more accurate and reliable results. By utilising modularisation, investigators can streamline their workflow and improve their investigative process.

There are a lot of compatibility and diversity issues in digital forensics, ranging from hardware, storage devices, and embedded devices to software developed for investigations. For example, several forensic software tools are developed in Python, C, and Java and designed to run on operating systems such as iOS, Windows, and Linux. Therefore, digital forensic tool developers face cross-language and cross-platform design, major barriers to forensic tool development. Although frameworks consisting of Apache webserver modules exist [37], no digital forensic information processing framework exists in that format with different functionalities for sector processing, object byte streaming, timestamp, structured data plugins, threat, and protection level correction [38].

Frameworks, especially those associated with the plugin, should be centred on a call-back prototype. Interacting with different devices with distinct filesystem layouts becomes challenging during a digital forensic investigation. To solve this challenge, modularisation is required to facilitate the single development of each data format, and those once-
developed data formats can then be used in various settings or scenarios on multiple digital devices and applications with different plugins. Modularisation will permit the use of the same plugin in a multi-threaded or single-threaded forensic implementation. Even though frameworks such as fiwalk.py and exit have limited call-backs, their application programming interfaces also need to be expanded to incorporate forensic reporting correlation. To achieve modularity, application programming interfaces should grant forensic module usage in interactive and batch forensic techniques. Figure 4 summarises the proposed modularisation technique, composed of the essential components.

Other existing frameworks, such as PyFlag [39-40], are generic and do not provide the scale and workflow automation needed for investigation construction. For example, in a cybercrime investigation, investigators may need to analyse large amounts of data from different sources and in various formats. Using PyFlag or other generic frameworks would require investigators to manually extract and transform the data before analysis, which could be time-consuming and error-prone. Another limitation of the existing frameworks is that many lack cross-platform integration and were algorithmic and non-specific [41] during their creation process, but they are deployed to learn many general digital forensic details. The modularisation approach enhances the future of digital forensic research amid constant and daily emerging technologies. It will provide a framework for deploying the same algorithms and making them specific on handheld devices and large cutting-edge clusters with multiple computational nodes. This will promote stable and balanced application programming interfaces that could be embraced by traditional forensic tools such as FTKImager [42] and EnCase [43], allowing composability and technological transition from the study test site to the user community.

Other modularity elements, such as structured query language, can also form the integration component of such a framework to enable search, results display, and storage of evidence details in the database by extraction and visibility tools, respectively. Thus, enabling several forensic investigators to use the tools simultaneously. A standardised query language design does not store all the essential information in the existing framework’s databases. In designing a structured query language to achieve this, the design architecture of a modular digital forensic framework must include all the significant...
information in the database since not all information held in the filesystems is detectable by all applications. Integrating a structured query language database in the framework will augment the conservation and maintenance of the structured query language database’s old fields for plugin usage, even if the fields are no longer valuable.

The plugin component of the modularisation would enhance the emission of structured information from the captured data for visualisation and reporting via an automated incident and event system. The call-back model will utilise different threats consisting of single and multiple threats connecting to multiple cloud servers. The other component of modularisation is the application programming interface (APIs), which would correlate reporting and forensics, encompassing batch forensics. The batch forensics sub-component would enhance the batch processing of evidence and aid in the appraisal of gaps in the modularisation model and their remediation. Similarly, it would improve the model by identifying the existing gaps with digital forensic investigators, digital forensic quality assurance engineers, or other groups of users. Analysing streams of captured data could achieve this by event occurrence. By so doing, the creation of abstraction could be enhanced by data manipulation for forensic data processing of discs and cloud streams.

The first logical step in having an effective modular digital forensics system that could solve identity management challenges, validation issues, scale factor problems, and autonomous operation associated with cloud forensics is that the implementation of such a system should include essential elements such as visibility, filter, and report components. As demonstrated in Figure 5, the suggested modular model comprises two major components: the traditional processing model component and the modular forensic system. Using the modular model, investigators can quickly identify the gaps in their investigation process and take appropriate measures to fill them, leading to a more robust and effective investigation. For instance, in an organisation investigating a cyberattack, the batch forensics sub-component can analyse large amounts of data collected from various sources to identify potential threats and suspicious activities. The APIs can provide an interface for integrating different digital forensic tools and techniques to process the data efficiently.

The traditional processing model component comprises the stream-based disk analysis model, where a device’s hard disc is analysed by byte stream, considering file hierarchy reconstruction. While this component is essential in analysing hard disc drives, it may be less significant when working with devices that have solid-state drives because of

![Figure 5. Proposed modularisation model with integrated traditional model processing functions](image-url)
the lack of moving heads in solid-state drive (SSD) architecture and the nature of its unique file system structure. The stochastic module of the traditional processing component is essential in selecting different sections of the drive sectors for random processing without necessarily observing multiple queues that may be caused by batch processing of data. This module is significant, as the tendency to omit data in small drive sectors may be possible. However, the random selection function of the stochastic model component will provide a layer of abstraction for identity management where small components of the data stored in smaller sectors of hard drives can be searched easily for identity resolutions.

More so, the module will aid in identifying different anonymous data sources, such as social media accounts and networks. The last component of the proposed modular digital forensic system is the prioritised model component for separate investigations of high priorities such as child pornography and terrorism. The prioritised investigations would be sequenced and run with parallel tracking and adequate feedback. The modular systems would then control all the traditional components for evidence visibility, filtering, and reporting, where all cloud and single-host system searches such as IP search, social media search, cloud service provider requests, intelligent search, and external data mining can be performed. For example, the stochastic module may not be as significant in a criminal investigation involving a suspect’s laptop with an SSD. However, the prioritised model component can be useful in identifying and investigating high-priority crimes such as terrorism or child pornography. The modular system can also perform cloud and single-host system searches for evidence filtering and reporting, aiding investigators in searching for anonymous data sources related to the case, such as social media accounts and networks.

5. Scope and performance analysis of the proposed model

The limitations and scope of the digital modularisation model with integrated traditional model processing functions to overcome the challenges of digital forensic challenges in cyberspace are:

5.1 The scalability of the digital modularisation model

In the context of large-scale and complex cybercrime investigations, the scalability of the digital modularisation model becomes a crucial aspect to consider. As the volume, variety, and velocity of data continue to increase exponentially, it is essential for the model to adapt and handle these challenges effectively. The modularisation approach provides a framework that allows for the seamless integration of diverse data sources and the efficient processing of vast amounts of information. However, in cases where criminals employ sophisticated encryption techniques, the modularisation approach may struggle to decipher and integrate this information, hindering the investigation process and limiting its scalability.

5.2 Privacy concerns in digital forensic investigations

In recent years, the field of digital forensic investigations has faced significant challenges in balancing the need for effective data analysis with the growing concerns surrounding privacy. As investigators delve into digital evidence to uncover critical information, they must also navigate a complex web of privacy laws and regulations. To address these concerns, modularisation model will offer a way to integrate traditional processing functions while ensuring strict adherence to privacy protocols. However, it is crucial to examine the limitations of this model to fully understand its effectiveness in preserving privacy during digital forensic investigations. One of the limitations of the proposed modularisation model is the potential for data fragmentation. By breaking down the processing functions into smaller modules, there is a risk of data being scattered across different systems and locations. This can make it difficult to maintain control over the privacy of the information being analyzed. Additionally, the complexity of integrating multiple modules from different vendors or sources can introduce vulnerabilities that could be exploited by malicious actors. Therefore, it is essential to thoroughly evaluate the security measures implemented within the modularisation model to ensure that privacy is not compromised.
5.3 Adapting to evolving cyber threats

Adapting to evolving cyber threats: Examine how the digital modularisation model addresses emerging challenges posed by new technologies, advanced malware, encryption techniques, and other evolving cyber threats. For example, in the digital modularisation model, instead of relying on a traditional monolithic system, organizations can break down their infrastructure into smaller components that can be easily updated and upgraded to counter new cyber threats. This allows for the implementation of more advanced encryption techniques and continuous monitoring systems to detect and respond to advanced malware in real-time. However, in instances where an organization’s smaller components are not effectively updated or upgraded, this will limit the framework. In such cases, these outdated components may become vulnerable to new cyber threats, as they lack the necessary security patches and advancements found in a monolithic system. This can result in a higher risk of successful cyber attacks and compromises to sensitive data within the organization.

6. Conclusion and future work

Digital device technological advancements increase data production, which in turn increases network bandwidth and data storage capacity. As a consequence, numerous organisations have migrated their data to the cloud in an effort to resolve the issue with data storage. Nevertheless, the utilisation of cloud services in forensic investigations presents obstacles on account of the wide range of software, hardware, filesystem configurations, intricacy, and augmented data volume. In cloud-based digital forensics investigations, digital forensic examiners encounter a number of these fundamental obstacles. The current state and prospective trajectory of digital forensics research are influenced by these obstacles. In order to tackle these obstacles, we offer suggestions that can be interpreted as methods to progress cloud-based digital forensic investigations. Legitimacy, data volume increase, and complexity challenges in digital forensics were identified as factors including low or non-existent funding for legal aid, constant amendments to the General Data Protection and Regulations Act, cloaking techniques, border laws, and emerging filesystem layout variants. In order to address these obstacles, we proposed a modularization framework for implementation. In order to surmount concerns related to privacy and legitimacy while conducting a digital forensic investigation in the cloud, a comprehensive understanding of the legal and ethical implications at play is essential. The modularization framework proposed will infuse traditional forensic processing model components into modular forensic that addresses other forensic issues such as randomization and sequential stream of data byte in stream-based investigation.

To address the challenges and issues identified and to enhance the future of digital forensic research, we suggest digital modularisation and analysis of its significance in overcoming the identified challenges, such as integrating a structured query language database in the framework to augment the conservation and maintenance of the structured query language database’s old fields for plugin usage even if the fields are no longer valuable. Another significance of having a digital forensics modular framework to overcome the core challenges is that such a framework would promote stable and balanced application programming interfaces that existing traditional forensic tools could embrace. The future direction of this work is to implement a test set containing different data formats and file system layouts, including multiple data domains, to measure the performance and assess the abstraction of network streams and their scalability in real-world digital modular scenarios.

Conflict of interest

The authors declare no competing financial interest.

References


