Digital Forensics in the Landscape of Cloud Environment: Challenges and Future

Adit Ishraq
Dept. of Computer and System Sciences (DSV)
Stockholm University
Stockholm, Sweden
adis6811@student.su.se

Jobayer Bin Showkat
Dept. of Computer and System Sciences (DSV)
Stockholm University
Stockholm, Sweden
josh7989@student.su.se

Khondaker Refai Arafat
Dept. of Computer and System Sciences (DSV)
Stockholm University
Stockholm, Sweden
khar1482@student.su.se

***Abstract*— Cloud digital forensics is an emerging and critical field as cloud computing continues to transform information technology. This paper explores the challenges and future trends in cloud forensics, focusing on public cloud environments. It categorizes these challenges according to the phases of the cloud forensic process: identification, preservation and collection, examination and analysis, and presentation. Key issues include the retrieval of log files, transient data, physical accessibility, vendor trust, data integrity, crime scene reconstruction, tool insufficiency, large data volumes, encryption, and testimonial complexity. The paper also discusses open problems and future trends, such as the impact of containerization, server-less computing, AI, advanced cryptographic techniques, IoT and block-chain technologies on forensic practices. The research aims to provide a comprehensive overview of the current state of cloud forensics, highlight ongoing challenges, and suggest directions for future research.**

Keywords*— *Cloud Forensics, Digital Evidence, Cloud Computing, Data Preservation, Data Integrity, Encryption, Multi-Tenancy, Chain of Custody, Server-less Computing, Block-chain, IoT.

INTRODUCTION

The rapid adoption of cloud computing has revolutionized data storage and processing, leading to significant changes in how digital forensics is conducted. Cloud digital forensics involves the application of forensic science techniques to cloud environments, addressing unique challenges posed by the cloud's distributed and dynamic nature. This paper examines these challenges and categorizes them according to the phases of the cloud forensic process. The identification phase involves retrieving information from log files and dealing with transient data and lack of physical accessibility [1, 2]. The preservation and collection phase focuses on maintaining data integrity and chain of custody [1, 3, 4]. The examination and analysis phase tackles issues related to forensic tool sufficiency, large data volumes, and encryption [5, 6, 7, 8, 9]. The presentation phase addresses the complexities of presenting technical evidence in court [1]. The paper also explores future trends and open problems in cloud forensics, emphasizing the need for ongoing research and the development of new methodologies to handle the evolving landscape of cloud computing.

BACKGROUND

The proliferation of cloud computing has transformed how data is stored, processed, and accessed. This shift has significant implications for digital forensics, which must adapt to the decentralized, multi-tenant, and often encrypted nature of cloud environments. Traditional forensic techniques are insufficient for these dynamic and complex settings, necessitating the development of specialized methodologies and tools. This research explores the various challenges faced by forensic investigators in cloud environments, categorizing these challenges according to different phases of the forensic process.

RESEARCH GAP

Despite advancements in digital forensics, significant gaps remain in addressing the specific challenges posed by

cloud environments. Existing forensic tools and methodologies are often inadequate for handling the dynamic, distributed nature of cloud data. There is a lack of standardized procedures for log retrieval, data preservation, and chain of custody in cloud contexts [1,2,4]. Additionally, the rapid evolution of cloud technologies, such as containerization and server-less computing, introduces new complexities that current forensic practices do not fully address [10].

Research Questions and Sub-Questions

RQ: How can digital forensic methodologies be adapted to effectively address the unique challenges posed by cloud computing environments?

Sub-questions:

a. What are the primary challenges in identifying and retrieving forensic evidence in cloud environments, particularly in multi-tenant and decentralized settings?

b. How can data integrity and chain of custody be ensured during the preservation and collection phase in cloud forensics?

c. What role do emerging technologies such as advanced cryptography, server-less computing, and block-chain, IoT, AI play in shaping future cloud forensic practices?

Methodology (Literature Review)

To address the research questions, a comprehensive literature review will be conducted. The literature review will include the following steps:

Search academic databases (e.g., IEEE Xplore, ACM Digital Library, Google Scholar) for articles, conference papers, and reports related to cloud digital forensics.

Organize the identified challenges according to the phases of the cloud forensic process: identification, preservation and collection, examination and analysis, and presentation. Summarize the key issues within each phase, referencing specific studies and expert opinions.

Review articles and reports that discuss emerging technologies and their impact on cloud forensics. Identify potential open problems and areas for future research, focusing on the integration of new technologies into forensic practices. Highlight ongoing challenges, suggest best practices, and propose directions for future research.

The literature review will provide a solid foundation for understanding the complexities of cloud digital forensics and will inform the development of strategies to address the identified challenges. This methodology ensures a thorough examination of existing knowledge and identifies gaps that need further exploration.

Aim of the Study

The aim of this study is to explore and develop effective adaptations of digital forensic methodologies to address the unique challenges posed by cloud computing environments. This includes identifying and mitigating issues in forensic evidence retrieval, ensuring data integrity and chain of custody, and incorporating emerging technologies such as advanced cryptography, server-less computing, and block-chain into cloud forensic practices. The study seeks to provide a comprehensive understanding of the complexities involved in cloud digital forensics and propose strategic solutions to enhance the efficacy and reliability of forensic investigations in cloud contexts.

Delimitation of the Study

This study focuses specifically on the forensic challenges and methodologies related to public cloud environments, as these present the most significant complexities due to multi-tenancy, decentralized infrastructure, and jurisdictional variability. Private cloud environments, while briefly mentioned, are not the primary focus. The study is also delimited to examining the forensic processes in IaaS, PaaS, and SaaS models, with an emphasis on the identification, preservation and collection, examination and analysis, and presentation phases. While emerging technologies such as AI, block-chain, and advanced cryptography are considered, the study does not delve into proprietary forensic tools or specific case studies but rather provides a broad overview and theoretical framework for addressing the outlined challenges.

Service Models (IaaS, SaaS, PaaS) of Cloud Computing

In cloud computing, there are three primary service models: Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). Each model offers a different level of abstraction and control

over the IT resources provided by the cloud provider, with varying responsibilities for the customer and the cloud provider as shown in Fig. 1

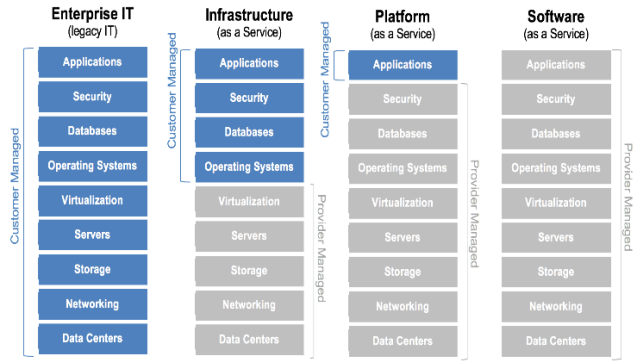

Fig. 1. Cloud-shared responsibility service model in IaaS, PaaS and SaaS (Source: https://mycloudblog7.wordpress.com/wp-content/uploads/2013/06/screen-shot-2015-06-09-at-2-13-05-pm1.png)

CLOUD SERVICES AND REGULATORY LANDSCAPE

Recently, organizations across all sectors increasingly rely on cloud service providers (CSPs) for IT infrastructure, data storage, and software solutions. Several regulatory bodies worldwide play crucial roles in the cloud services landscape:

A. European Union Agency for Cybersecurity (ENISA):

ENISA is enhancing cybersecurity across the European Union. It provides guidelines and best practices to address cybersecurity and regulatory issues related to cloud services.

B. General Data Protection Regulation (GDPR):

GDPR imposes stringent standards for the processing and protection of personal data, impacting cloud services by ensuring robust data privacy and security measures.

C. National Institute of Standards and Technology (NIST):

NIST provides a framework for cloud computing that addresses security, privacy, and interoperability to assist organizations.

D. International Organization for Standardization (ISO):

ISO/IEC 27017 develops standards for security controls for cloud services and ISO/IEC 27018 for protecting personal data in the cloud.

CLOUD DIGITAL FORENSICS TOOLS AND TECHNOLOGIES

The cloud digital forensics is bolstered by tools that facilitate investigations in cloud computing environments. Among them Magnet AXIOM Cloud, Mandiant CloudLens, Autopsy, X-Ways Forensics, Azure Security Center, AWS CloudTrail are prominent. These tools are used for cloud data collection and analysis, supporting AWS, Azure, and Google Cloud to recover, examine, and preserve cloud-based evidence.

Offline Digital Forensics Tools:

EnCase Forensic, AccessData Forensic Toolkit (FTK), DEFT (Digital Evidence and Forensics Toolkit), Digital Forensics Framework (DFF) are widely used software for acquiring, analyzing, and reporting digital evidence from various devices and file systems.

CLOUD FORENSIC CHALLENGES

This section provides an overview of the challenges encountered in cloud forensics, categorized according to the procedural phases of cloud forensic investigations.

A. Identification Phase

- *Retrieval of Information from Log Files:*

Importance: Log files are critical for forensic investigations as they provide detailed records of system activities and user interactions.

Challenges: In cloud environments, retrieving log files is highly complex due to several factors:

Cloud Haze: The abstract nature of cloud services obscures direct access to the underlying infrastructure.

Multi-Tenancy: Multiple clients share the same physical resources, complicating the isolation and retrieval of specific logs.

API Access Only: Clients typically access services via APIs, which do not offer comprehensive logging capabilities [11].

IaaS Model Restrictions: While logs are essential for understanding VM behavior in the Infrastructure as a Service (IaaS) model, cloud providers may impose restrictions on log storage, access, and sharing. These restrictions can limit the effectiveness of logs in forensic investigations.

Log Collection Services: Cloud service providers may neglect or even hide log collection services, leading to issues such as decentralized log storage, fluctuating data availability, difficulties in preserving logs, limited accessibility, and the non-existence of critical log data. Furthermore, logs might be in incompatible formats, making it challenging to combine and analyze data from different sources [12].

- *Transient Data:*

Importance: Transient data includes temporary information stored in VMs that can be crucial for forensic analysis.

Challenges: The dynamic behavior of VMs in IaaS platforms like Azure, Digital Ocean, and AWS presents significant challenges:

Volatile Data Preservation: Volatile data can be lost during shutdowns or restarts of VMs, making it crucial for forensic experts to understand these behaviors to preserve important data [5,6,7].

Service Structure Navigation: Forensic professionals must navigate the diverse and complex service structures of different IaaS providers to identify and secure transient data effectively.

- *Lack of Physical Accessibility:*

Importance: Traditional digital forensics relies on direct access to physical hardware to collect evidence.

Challenges: In cloud environments, direct physical access to hardware is impractical due to:

Global Deployment: Data is stored across globally distributed data centers, complicating localization.

Jurisdictional Issues: Physical hardware storing data cannot be easily seized because systems are spread across different legal jurisdictions.

Fixed Settings: Unlike on-premises systems where hardware can be controlled directly, cloud environments do not allow for such control, limiting the ability to collect physical evidence [13].

- *Identification at the Client Side:*

Importance: Evidence can be located on both the cloud provider's side and the client's side, especially in SaaS and PaaS environments.

Challenges: Identifying and capturing client-side data presents several difficulties:

Quick Data Capture: Investigators must act swiftly to capture data before it can be altered or destroyed by the perpetrator.

Jurisdictional Variability: Different legal jurisdictions complicate the identification and preservation of client-side data.

Proof Location: Evidence may be spread across multiple interfaces and locations, requiring comprehensive and coordinated efforts to secure it [5].

- *Vendor Dependency and Trust:*

Importance: Cloud Service Providers (CSPs) play a critical role in providing access to necessary data for forensic investigations.

Challenges: Dependency on CSPs introduces several issues:

Reluctance to Release Information: CSPs may be hesitant to provide data, especially in multi-tenant environments where releasing data could affect other clients.

Authenticity Concerns: There are concerns about the authenticity and integrity of the data provided by CSPs, especially when dealing with non-expert personnel who may not follow strict forensic protocols.

Impact on Forensic Findings: Relying on CSPs for evidence discovery in SaaS and PaaS models can affect the validity and reliability of forensic findings due to potential biases and incomplete data [11].

By addressing these challenges, cloud forensic investigations can improve their effectiveness and reliability, ensuring that critical evidence is accurately identified and preserved for legal proceedings.

B. Preservation and Collection Phase

- *Integrity and Stability in Multi-Tenancy and Privacy:*

Importance: Maintaining the integrity and stability of evidence is crucial in cloud forensics, especially in multi-tenant environments (IaaS, PaaS, SaaS) where multiple users share the same physical resources.

Challenges:

Data Retention: Ensuring data retention across different jurisdictions is complex, with varying laws affecting how data can be stored and accessed. This can hinder the preservation of evidence.

Evidence Reliability: The reliability of evidence can be compromised due to the shared nature of cloud resources. This may render the evidence inadmissible in court if it cannot be shown to be untampered.

Authenticity Issues: Investigators must often rely on third parties (CSPs) for data authentication, which can be

problematic. There is a need for increased trust and assurance in the data provided by these third parties [3].

Data Consistency: The dynamic nature of cloud environments means data can change frequently, making it difficult to ensure consistent and stable evidence collection [16].

- *In-House Staffing:*

Importance: Effective cloud forensic investigations require a multidisciplinary approach involving technical experts, legal consultants, and external specialists.

Challenges:

Expert Collaboration: Coordinating between various experts is necessary to navigate the complex technical and legal landscapes of cloud forensics.

New Technologies Expertise: Staff must be proficient in the latest technologies and forensic tools to effectively manage and analyze cloud-based data.

Resource Allocation: Ensuring that the organization has adequate staffing and resources dedicated to forensic investigations can be challenging, especially for smaller organizations [15].

- *Crime Scene Reconstruction in Criminal Investigations:*

Importance: Reconstructing the sequence of events leading up to a crime is a fundamental part of forensic investigations.

Challenges:

VM Termination: In cloud environments, VMs can be terminated after malicious activities, making it difficult or impossible to reconstruct the crime scene accurately.

Data Volatility: The transient nature of data within VMs adds complexity, as crucial evidence can be lost if not captured promptly.

Lack of Physical Access: Without direct access to physical hardware, traditional methods of crime scene reconstruction are not applicable, requiring innovative approaches to gather and piece together digital evidence.

- *Chain of Custody:*

Importance: Maintaining a clear and documented chain of custody is essential for ensuring the admissibility of evidence in court.

Challenges:

Multi-Jurisdictional Legislation: Different legal jurisdictions have varying requirements for evidence handling, complicating the chain of custody in cloud environments.

CSP Engagement: Cloud service providers play a critical role in the custody chain, and any lapses on their part can jeopardize the entire investigation.

Initial Failure Point: The initial potential failure point in the chain of custody is often at the CSP level, where improper handling or delays in providing data can occur [4].

- *Data Imaging:*

Importance: Creating an exact forensic image of a system or instance is vital for thorough investigation and analysis.

Challenges:

IaaS Environment: In IaaS models, capturing a disk image of a VM involves creating a precise copy in a specific file format (e.g., EWF). If the VM is restarted or shut down, data might still be recoverable, but if destroyed, it is permanently lost.

PaaS Environments: Data collection in PaaS environments relies heavily on the CSP. This dependency becomes problematic if the data is managed by third-party subcontractors, adding another layer of complexity.

Third-Party Management: When data is managed by a subcontractor, obtaining a reliable and comprehensive forensic image can be difficult, as the subcontractor's cooperation and compliance with forensic protocols are not always guaranteed [17].

- *Bandwidth Constraints:*

Importance: The increasing volume of data in cloud environments necessitates efficient data collection methods.

Challenges:

Data Volume: The sheer amount of data in cloud storage, often reaching petabytes, complicates the collection process. Researchers need to make forensic copies of VM instances, which is bandwidth-intensive.

Available Bandwidth: Limited bandwidth can slow down the imaging process, making it harder to obtain timely and complete copies of the necessary data.

Data Transfer: Transferring large data volumes within the constraints of available bandwidth requires efficient planning and possibly leveraging high-speed connections or incremental data collection techniques to manage the load effectively.

C. Examination and Analysis Phase

- *Insufficient Forensic Toolset:*

Importance: The accuracy and reliability of forensic tools are critical for cloud forensic investigations, as these tools are essential for analyzing digital evidence.

Challenges:

Lack of Comprehensive Vetting: Many commercial tools designed for cloud forensics have not undergone thorough testing to ensure their accuracy and reliability. This can lead to errors in investigations and potentially inadmissible evidence.

Enhancing Tool Reliability: By providing measurable assurance of tool accuracy, the CFTT (Computer Forensics Tool Testing) program helps enhance the credibility of forensic tools [17].

- *Large Data Volumes:*

Importance: The immense and ever-growing volume of data stored in cloud service provider (CSP) facilities presents significant challenges for forensic investigations.

Challenges:

Data Overload: The sheer amount of data, often in the petabytes, can make it difficult to identify relevant digital evidence. This complicates the processing and analysis stages of an investigation.

Data Reduction Methods: Research gaps exist in developing efficient data reduction methods, data mining techniques, and intelligence evaluation strategies. These methods are essential for filtering through vast amounts of data to find meaningful evidence.

Effective Information Filtering: Investigators must implement appropriate collection and filtering techniques to manage the large data volumes in cloud infrastructures effectively. This includes utilizing both open-source and proprietary tools to streamline the identification and extraction of relevant evidence [18].

- *Encryption:*

Importance: Encryption is widely used by cloud clients to protect data, posing a significant hurdle for forensic investigations.

Challenges:

Key Accessibility: Accessing encryption keys is crucial for decrypting data. Without these keys, forensic investigators may find it impossible to analyze the encrypted content.

Owner-Dependent Evidence: If only the data owner has the encryption key, obtaining this key can be challenging, especially if the owner is uncooperative or unavailable. This can undermine the integrity and completeness of the evidence.

CSP Encryption Practices: Many CSPs also employ their own encryption technologies, adding another layer of complexity. Investigators must navigate these encryption protocols, which may vary widely among different providers, to access and analyze the data [8].

- *Log Format Standardization:*

Importance: Standardized log formats are essential for efficient and effective analysis of data collected from various cloud service models.

Challenges:

Variety of Log Types: Cloud environments generate a wide range of log formats, depending on the service model (IaaS, PaaS, SaaS) and the specific CSP. This diversity makes it difficult to combine and analyze logs.

Cost and Complexity: Analyzing these diverse log formats is a costly and complex operation. Investigators need to standardize and integrate logs from multiple sources to reconstruct events and identify relevant evidence.

Combining Log Forms: Accessing and interpreting a large number of different log formats requires significant effort and expertise. This process is essential for creating a cohesive and comprehensive view of the events under investigation [15].

In summary, the examination and analysis phase of cloud forensics faces numerous challenges, including the insufficient vetting of forensic tools, handling large data volumes, navigating encryption protocols, and dealing with diverse log formats. Addressing these issues requires ongoing efforts to improve forensic tools, develop efficient data processing methods, and standardize log formats, ensuring that investigators can effectively analyze and present digital evidence.

D. Presentation Phase

- *Password or Key Retrieval:*

Importance: Accessing encrypted data is a critical challenge in cloud forensic investigations. Without cooperation from involved parties, decrypting data can be exceedingly difficult.

Challenges:

Advanced Tools: Tools like John the Ripper and Hashcat are essential for password retrieval. These tools use techniques such as brute force and dictionary attacks

to crack passwords, aiding investigators in accessing encrypted data [19].

Memory Dumps: Analyzing memory dumps can provide crucial information for retrieving encryption keys. When a system is running, encryption keys may be stored in volatile memory (RAM). By capturing and analyzing memory dumps, investigators can sometimes extract these keys, which are vital for decrypting protected data.

Enhancing Capabilities: The combination of advanced password-cracking tools and memory analysis significantly enhances investigators' capabilities to deal with encrypted data. However, the process is time-consuming and requires specialized knowledge and resources.

- *Testimonial Complexity:*

Importance: Presenting technical details in court can be challenging due to the complexity of digital forensics and the limited technical knowledge of juries.

Challenges:

Simplifying Technical Concepts: Investigators must transparently disclose their methods and procedures, ensuring they provide a clear, understandable explanation of cloud computing, digital forensics, and the investigation process. This includes detailing how evidence was preserved and recorded.

Expert Testimony: Expert witnesses must be able to convey complex technical information in a manner that is comprehensible to the jury. This requires not only deep technical knowledge but also the ability to communicate effectively with non-technical audiences.

Clarity and Precision: Every piece of evidence must be presented with great care, emphasizing clarity and precision. This ensures that jurors understand the relevance and reliability of the evidence, which is crucial for the legal process [20].

- *Documentation and Record Keeping:*

Importance: Proper documentation and record keeping are essential to maintaining the integrity of the evidence and the chain of custody throughout the investigation.

Challenges:

Chain of Custody: Researchers must ensure that all parties involved in the investigation adhere to established methodologies and standards to preserve the chain of custody. This is crucial for proving that the evidence has not been tampered with or altered in any way.

Comprehensive Documentation: Electronic documentation must encompass all stages of the investigation. This includes initial data collection, analysis, and the final presentation of findings. Each step must be meticulously documented to demonstrate the reliability and authenticity of the evidence.

Convincing the Jury: Investigators must be able to convince the jury that the proof obtained during the investigation has been properly documented and remains unaltered. This requires presenting a thorough and transparent record of the entire forensic process, highlighting the adherence to proper procedures and standards.

In conclusion, the presentation phase of cloud forensic investigations involves significant challenges related to password retrieval, testimonial complexity, and documentation. Investigators must use advanced tools and techniques to access encrypted data, simplify complex technical information for juries, and ensure meticulous documentation to maintain the integrity of the evidence. These efforts are crucial for successfully presenting and defending digital evidence in a court of law.

EMERGING TECHNOLOGIES AND THEIR CHALLANGES IN CLOUD FORENSIC

The landscape of cloud digital forensics is continually evolving, prompting researchers to explore future directions to enhance forensic practices in the cloud. As cloud computing technologies advance, adapting forensic methodologies to address emerging trends becomes increasingly essential.

A. Homomorphic Encryption:

It allows computations on encrypted data without decrypting it, preserving privacy. The complexity of these techniques poses significant difficulties for forensic analysis, as traditional methods of data inspection are ineffective [21].

B. Multiparty Computation:

It enables multiple parties to jointly compute a function over their inputs while keeping those inputs private. Forensics must adapt to decrypt and analyze data processed through such secure methods without compromising the privacy guarantees [22].

C. Federated Learning:

This is a machine learning approach where models are trained across multiple decentralized devices without centralizing data. This approach complicates the collection and analysis of training data, as the data remains distributed and privacy-preserved across multiple nodes [22].

D. Block-chain-Based Cloud Systems:

This is the use of block-chain technology to manage cloud data, ensuring transparency and security in digital transactions. The decentralized nature of block-chain complicates data retrieval and validation, making it harder to trace the provenance and integrity of digital evidence [23].

E. Containerization:

Containerization is a technology that packages applications and their dependencies into isolated units called containers, ensuring consistent operation across different computing environments. Containers are ephemeral, and data within them can be lost or altered when they are destroyed or moved [24].

F. Micro-services:

Micro-services is an architectural style that structures an application as a collection of small, loosely coupled services, each responsible for a specific functionality and communicating through APIs. Micro-services architecture spreads data and operations across many small, independent services, complicating data correlation and evidence gathering [24].

G. Server-less Computing:

Server-less computing is a cloud execution model where the cloud provider dynamically manages the infrastructure. Server-less functions are stateless and can execute in unpredictable locations, making it hard to track data and execution context [24].

ADVANCED TECHNOLOGIES TO SOLVING CHALLENGES IN CLOUD FORENSICS (FUTURE TRENDS)

A. Advanced Cryptographic Techniques:

Homomorphic Encryption allows computation on encrypted data without needing decryption. Investigators can perform necessary analysis directly on encrypted datasets, preserving data confidentiality and integrity. This technique ensures that sensitive data remains protected while still allowing forensic analysis, thus addressing privacy concerns in multi-tenant environments [26].

B. Multiparty Computation:

It enables collaborative computation without revealing individual inputs. Facilitates joint forensic investigations across multiple entities (e.g., different cloud service providers or clients) without compromising sensitive data. This technique ensures secure, private data analysis and is especially useful in cases involving multiple jurisdictions and stakeholders [26].

C. Internet of Things (IoT) in Enhanced Logging Mechanisms:

IoT devices can generate secure, tamper-proof logs. By implementing robust logging mechanisms, IoT devices can provide valuable forensic data. Secure logs help investigators trace the sequence of events leading up to an incident, even across a distributed network of IoT devices. This is crucial for maintaining a chain of custody and ensuring the reliability of evidence [28].

D. Standardized Protocols of IoT:

Developing and adopting standardized data collection and communication protocols. Standardization ensures that data from various IoT devices can be easily integrated and analyzed in forensic investigations. It helps in overcoming the challenge of heterogeneity in IoT ecosystems and facilitates the smooth aggregation of forensic data [27].

E. Block-chain and Distributed Ledger Technology (DLT)

Block-chain technology ensures that once data is written, it cannot be altered. The immutability of block-chain records provides a verifiable and tamper-proof log of transactions. This is crucial for maintaining the integrity of forensic evidence. Forensic investigators can rely on block-chain to validate the authenticity of digital evidence, ensuring that it has not been tampered with [27].

F. Decentralized Data Management:

Distributed ledgers allow data to be stored across multiple nodes in a decentralized manner. Decentralized storage enhances data availability and resilience, making it harder for malicious actors to compromise evidence. This also aids in cross-jurisdictional investigations, as data stored on a block-chain can be accessed and verified globally, without relying on a single point of control [27].

G. Block-chain in Transparent and Auditable Trails:

Every transaction on a block-chain is recorded with a timestamp and cryptographic proof. Block-chain provides a transparent and auditable trail of actions, which is invaluable for forensic investigations. Investigators can trace the origin and changes to digital assets, ensuring a clear chain of custody. This transparency helps in presenting credible evidence in legal proceedings [27].

H. Data Mining in Pattern Recognition:

Data mining techniques can uncover hidden patterns and correlations within large datasets. In cloud environments, data mining can help identify suspicious activities, anomalies, or trends indicative of cyber-attacks or unauthorized access. By analyzing vast amounts of data, data mining assists investigators in detecting potential security breaches and irregularities [13].

I. *Data Mining in Behavioral Analysis:*

Data mining enables the creation of behavioral profiles based on user interactions and system activities. By analyzing user behavior and system interactions, data mining can aid in the identification of malicious activities or insider threats. It allows forensic investigators to establish baseline behaviors and detect deviations that may indicate unauthorized or abnormal behavior [13].

J. *Machine Learning in Anomaly Detection:*

Machine learning algorithms can learn patterns from historical data and detect anomalies or outliers. In cloud environments, machine learning algorithms can automatically identify suspicious activities or deviations from normal behavior. By training on historical forensic data, machine learning models can detect novel threats or emerging attack patterns, enhancing the proactive detection of security incidents [21].

K. *Machine Learning in Predictive Analysis:*

Machine learning models can predict future security incidents based on historical data and current system states. Predictive analytics can help forecast potential security threats or vulnerabilities in cloud environments. By analyzing patterns in past incidents and system behavior, machine learning algorithms can provide early warnings of impending security risks, allowing organizations to take proactive measures to mitigate threats [21].

L. *Artificial Intelligence (AI) in Automated Incident Response:*

AI-powered systems can automatically respond to security incidents based on predefined rules or learned patterns. AI-driven incident response systems can rapidly detect and contain security breaches in cloud environments. By automating response actions such as isolating compromised systems or blocking suspicious network traffic, AI enhances the efficiency and effectiveness of incident response efforts, minimizing the impact of security incidents [25].

M. *Natural Language Processing (NLP):*

NLP techniques enable the analysis of unstructured textual data, such as log files, system reports, or communication transcripts. In cloud forensic investigations, NLP can assist in extracting relevant information from textual data sources, facilitating evidence collection and analysis. By parsing and understanding natural language text, NLP systems can identify keywords, entities, or sentiment that may be indicative of suspicious activities or security incidents [25].

N. *Integration of Data Sources:*

Integrating diverse data sources and formats to provide a comprehensive view of cloud environments. By combining data from multiple sources, including logs, network traffic, and system events, forensic investigators can gain deeper insights into security incidents and improve the accuracy of their analyses. Integrated data mining and machine learning techniques enable the correlation of disparate data points to uncover hidden relationships and patterns [25].

O. *Continuous Learning Models:*

Developing machine learning models that can adapt and evolve over time in response to changing threat landscapes. Continuous learning models enable forensic systems to stay up-to-date with emerging threats and evolving attack techniques. By continuously analyzing new data and incorporating feedback from security incidents, these models can improve their accuracy and effectiveness in detecting and mitigating security risks in cloud environments [25].

RESULTS AND DISCUSSION

In cloud environments, forensic investigations face several key challenges. Retrieving information from log files is difficult due to the abstract nature of cloud services, shared resources among multiple clients, and limited logging capabilities through APIs. Transient data, such as volatile information within virtual machines (VMs), can be lost during shutdowns or restarts, and forensic professionals must navigate various service structures to secure this data. The global distribution of data complicates localization and introduces jurisdictional issues, further limiting physical access. Identifying and capturing data quickly at the client side is critical to prevent alteration or destruction, but different legal jurisdictions add complexity. Investigators also rely heavily on cloud service providers (CSPs) for data, facing challenges related to CSPs' reluctance to release information and concerns about data authenticity. Ensuring data integrity and maintaining the chain of custody are difficult in multi-tenant environments, with varying data retention laws and shared resources potentially compromising evidence reliability. Effective investigations require collaboration among technical and legal experts and adequate resource allocation. Reconstructing crime scenes is challenging due to VM termination and transient data volatility. The chain of custody is further complicated by differing legal requirements across jurisdictions and CSPs' crucial role. Data imaging in cloud environments, particularly in IaaS and PaaS models, faces issues related to the precise capture of disk images and reliance on third-party subcontractors. Additionally, large data volumes and limited bandwidth can slow down the imaging process. Emerging technologies offer potential solutions: advanced

cryptographic techniques like homomorphic encryption and multiparty computation enhance data confidentiality, IoT devices can generate secure logs, block-chain ensures data immutability, and data mining and machine learning aid in pattern recognition and anomaly detection. AI-powered systems can automate incident response, and natural language processing (NLP) assists in analyzing unstructured textual data.

Conclusion

Cloud digital forensics is a critical and rapidly evolving field that faces numerous challenges due to the unique nature of cloud environments. Effective adaptation of forensic methodologies to these settings is essential to address issues related to data retrieval, integrity, and analysis. Emerging technologies such as AI, block-chain, and advanced cryptographic techniques present both opportunities and challenges, necessitating continuous innovation and research [10]. By addressing these challenges, forensic investigators can better preserve, analyze, and present digital evidence, ensuring the integrity and reliability of forensic processes in the cloud era. Collaborative efforts among researchers, practitioners, and policymakers are crucial to developing robust forensic frameworks and tools that can meet the demands of this dynamic landscape.

References

[1] Damshenas, M.; Dehghantanha, A.; Mahmoud, R.; bin Shamsuddin, S. Forensics investigation challenges in cloud computing environments. In Proceedings of the 2012 International Conference on Cyber Security, Cyber Warfare and Digital Forensic (CyberSec), Kuala Lumpur, Malaysia, 26–28 June 2012; pp. 190–194.

[2] Alobaidli, H.; Nasir, Q.; Iqbal, A.; Guimaraes, M. Challenges of cloud log forensics. In Proceedings of the SouthEast Conference, Atlanta, GA, USA, 2 December 2017; pp. 227–230.

[3] Aydin, M.; Jacob, J. A comparison of major issues for the development of forensics in cloud computing. In Proceedings of the 8th International Conference for Internet Technology and Secured Transactions (ICITST-2013), London, UK, 9–12 December 2013; pp. 77–82.

[4] Orton, I.; Alva, A.; Endicott-Popovsky, B. Legal process and requirements for cloud forensic investigations. In Cybercrime and Cloud Forensics: Applications for Investigation Processes; IGI Global: Hershey, PA, USA, 2013; pp. 186–229.

[5] Birk, D.;Wegener, C. Technical issues of forensic investigations in cloud computing environments. In Proceedings of the 2011 Sixth IEEE International Workshop on Systematic Approaches to Digital Forensic Engineering, Oakland, CA, USA, 26 May 2011; pp. 1–10.

[6] Zawoad, S.; Hasan, R. Cloud forensics: A meta-study of challenges, approaches, and open problems. arXiv 2013, arXiv:1302.6312.

[7] Zimmerman, S.; Glavach, D. Cyber forensics in the cloud. IA Newsl. 2011, 14, 4–7.

[8] Almulla, S.; Iraqi, Y.; Jones, A. Cloud forensics: A research perspective. In Proceedings of the 2013 9th International Conference on Innovations in Information Technology (IIT), Al Ain, United Arab Emirates, 17–19 March 2013; pp. 66–71.

[9] Sibiya, G.; Venter, H.S.; Fogwill, T. Digital forensic framework for a cloud environment. In IST-Africa 2012 Conference Proceedings; International Information Management Corporation (IIMC): Dublin, Ireland, 2012.

[10] Montasari, R.; Hill, R. Next-generation digital forensics: Challenges and future paradigms. In Proceedings of the 2019 IEEE 12th International conference on global security, safety and sustainability (ICGS3), London, UK, 16–18 January 2019; pp. 205–212.

[11] Zawoad, S.; Dutta, A.K.; Hasan, R. SecLaaS: Secure logging-as-a-service for cloud forensics. In Proceedings of the 8th ACM SIGSAC Symposium on Information, Computer and Communications Security, Hangzhou, China, 8–10 May 2013; pp. 219–230.

[12] Marty, R. Cloud application logging for forensics. In Proceedings of the 2011 ACM Symposium on Applied Computing, Taichung, Taiwan, 21–24 March 2011; pp. 178–184.

[13] Poisel, R.; Tjoa, S. Discussion on the challenges and opportunities of cloud forensics. In Proceedings of the Multidisciplinary Research and Practice for Information Systems: IFIP WG 8.4, 8.9/TC 5 International Cross-Domain Conference and Workshop on Availability, Reliability, and Security, CD-ARES 2012, Prague, Czech Republic, 20–24 August 2012; Proceedings 7; Springer: Berlin/Heidelberg, Germany, 2012; pp. 593–608.

[14] Adams, R. The emergence of cloud storage and the need for a new digital forensic process model. In Cybercrime and Cloud Forensics: Applications for Investigation Processes; IGI Global: Hershey, PA, USA, 2013; pp. 79–104.

[15] Ruan, K.; Carthy, J.; Kechadi, T.; Crosbie, M. Cloud forensics. In Proceedings of the Advances in Digital Forensics VII: 7th IFIP WG 11.9 International Conference on Digital Forensics, Orlando, FL, USA, 31 January–2 February 2011; Revised Selected Papers 7; Springer: Berlin/Heidelberg, Germany, 2011; pp. 35–46.

[16] Chen, G.; Du, Y.; Qin, P.; Du, J. Suggestions to digital forensics in Cloud computing ERA. In Proceedings of the 2012 3rd IEEE International Conference on Network Infrastructure and Digital Content, Beijing, China, 21–23 September 2012; pp. 540–544.

[17] Computer Forensics Tool Testing (CFTT). Available online: https://www.cftt.nist.gov/ (accessed on 30 October 2023).

[18] Quick, D.; Choo, K.K.R. Impacts of increasing volume of digital forensic data: A survey and future research challenges. Digit. Investig. 2014, 11, 273–294. [CrossRef]

[19] Kanta, A.; Coray, S.; Coisel, I.; Scanlon, M. How viable is password cracking in digital forensic investigation? Analyzing the guessability of over 3.9 billion real-world accounts. Forensic Sci. Int. Digit. Investig. 2021, 37, 301186. [CrossRef]

[20] Ruan, K. Cybercrime and Cloud Forensics: Applications for Investigation; IGI Global: Hershey, PA, USA, 2013.

[21] Basilakis, J. Cloud-Based Homomorphic Encryption for Privacy-Preserving Machine Learning in Clinical Decision Support; Western Sydney University: Sydney, Australia, 2020.

[22] Alexandru, A.B.; Pappas, G.J. Secure multi-party computation for cloud-based control. In Privacy in Dynamical Systems; Springer: Singapore, 2020; pp. 179–207.

[23] Aggarwal, B.K.; Gupta, A.; Goyal, D.; Gupta, P.; Bansal, B.; Barak, D.D. A review on investigating the role of block-chain in cybersecurity. Mater. Today Proc. 2022, 56, 3312–3316. [CrossRef]

[24] Jambunathan, B.; Yoganathan, K. Architecture decision on using microservices or serverless functions with containers. In Proceedings of the 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT), Coimbatore, India, 1–3 March 2018; pp. 1–7.

[25] Ahmed, S.F.; Shuravi, S.; Afrin, S.; Rafa, S.J.; Hoque, M.; Gandomi, A.H. The Power of Internet of Things (IoT): Connecting the Dots with Cloud, Edge, and Fog Computing. arXiv 2023, arXiv:2309.03420.

[26] Kanagavelu, R.; Wei, Q.; Li, Z.; Zhang, H.; Samsudin, J.; Yang, Y.; Goh, R.S.M.; Wang, S. CE-Fed: Communication efficient multi-party computation enabled federated learning. Array 2022, 15, 100207. [CrossRef]

[27] Natarajan, H.; Krause, S.; Gradstein, H. Distributed Ledger Technology and Blockchain; Technical Report; World Bank: Washington, DC, USA, 2017.

[28] Montasari, R.; Hill, R. Next-generation digital forensics: Challenges and future paradigms. In Proceedings of the 2019 IEEE 12th International conference on global security, safety and sustainability (ICGS3), London, UK, 16–18 January 2019; pp. 205–212.