Journal of Autonomous Intelligence (2024) Volume 7 Issue 5
doi: 10.32629/jai.v7i5.1270

ORIGINAL RESEARCH ARTICLE

Designing an automated, privacy preserving, and efficient Digital Forensic Framework

Dhwaniket Kamble[1,2,*], Mangesh Dilip Salunke[1]

[1] *G H Raisoni University, Amravati, Maharashtra 444701, India*

[2] *Department of Engineering and Technology, Bharati Vidyapeeth Deemed University, Navi Mumbai, Maharashtra 410210, India*

*** Corresponding author:** Dhwaniket Kamble, sakec.dhwaniketk@gmail.com, drkamble@bvucoep.edu.in

ABSTRACT

The digital forensic investigation field faces continual challenges due to rapid technological advancements, the widespread use of digital devices, and the exponential growth in stored data. Protecting data privacy has emerged as a critical concern, particularly as traditional forensic techniques grant investigators unrestricted access to potentially sensitive data. While existing research addresses either investigative effectiveness or data privacy, a comprehensive solution that balances both aspects remains elusive. This study introduces a novel digital forensic framework that employs case information, case profiles, and expert knowledge to automate analysis. Machine learning techniques are utilized to identify relevant evidence while prioritizing data privacy. The framework also enhances validation procedures, fostering transparency, and incorporates secure logging mechanisms for increased accountability.

Keywords: data acquisition; privacy preservation; efficient data processing; digital forensic analysis; automation

ARTICLE INFO

Received: 8 November 2023
Accepted: 5 December 2023
Available online: 13 March 2024

1. Introduction

1.1. Background

Digital forensic is a crucial field in today's technology-driven world, as it plays a vital role in detecting and investigating digital crimes. These crimes can range from cyber theft, hacking, online harassment, to child exploitation[1]. Law enforcement agencies, computer professionals, forensic practitioners, and other stakeholders need to work together in a coordinated manner to effectively carry out digital investigations.

The current computing environment presents numerous challenges to digital forensics, including the increasing complexity of digital devices and the constant evolution of forensic techniques. These challenges necessitate the development of an automated, privacy preserving, and efficient digital forensic framework that can streamline the investigation process and ensure the accuracy and integrity of the evidence obtained. The issue at hand is that current digital forensic tools and frameworks are not sufficiently equipped to handle the complexities that arise in the ever-changing computing landscape. As a result, they lead to inefficiencies, privacy breaches, and challenges in collecting evidence that can be used in legal proceedings[2]. To tackle this problem, our proposed research aims to develop a digital forensic framework that is automated, preserves privacy, and operates

efficiently. By harnessing advanced technologies like artificial intelligence, machine learning, and data analytics, this framework will significantly improve the speed and accuracy of digital investigations[3]. Moreover, it will give utmost priority to protecting privacy by implementing encryption and anonymization techniques, ensuring that sensitive information remains secured.

1.2. Objectives

The main goal of this research is to create a digital forensic framework that is automated, protects privacy, and operates efficiently, addressing the shortcomings of conventional methods. The framework (**Figure 1**) aims to streamline the investigation process, improve accuracy, protect individuals' privacy, and enhance overall efficiency. By leveraging machine learning algorithms, cryptographic techniques, and optimized data processing methods, this framework provides a comprehensive solution for digital forensic investigations.

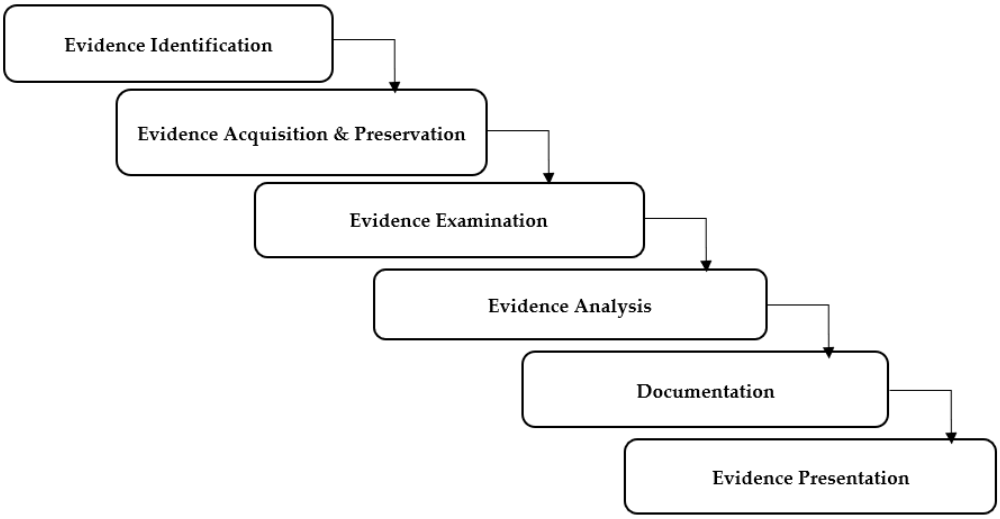

Figure 1. Flow of digital forensic framework.

Flow for Digital Forensic Investigation Framework Process:

1) Initiate: Commence the investigation process by outlining the objectives and scope.
2) Collect Evidence: Gather digital evidence from various sources, ensuring proper documentation and adherence to legal protocols.
3) Preserve Evidence: Safeguard the collected evidence to maintain its integrity and authenticity throughout the investigation. This involves implementing appropriate storage and handling techniques to prevent tampering or data loss the following measures is taken:
    - Chain of Custody: Establish and maintain a documented chain of custody for all evidence. This ensures a clear record of who has had access to the evidence at any given time, providing accountability and preventing unauthorized alterations.
    - Secure Storage: Store the evidence in a secure and controlled environment. This can include locked cabinets, safes, or designated evidence storage facilities with restricted access. Limit physical access to authorized personnel only.
    - Digital Storage: For digital evidence, use secure storage systems such as encrypted hard drives, secure servers, or cloud-based storage with strong authentication and access controls. Regularly back up the data to prevent loss.
    - Access Controls: Implement access controls to restrict unauthorized access to evidence. This can include strong passwords, two-factor authentication, and role-based access controls to ensure that only authorized personnel can access the evidence.

- Logging and Monitoring: Implement logging and monitoring mechanisms to track who accesses the evidence and when. This helps detect any unauthorized activities and provides an audit trail for accountability.
- Encryption: Use encryption techniques to protect sensitive data and ensure that only authorized personnel can decrypt and access the evidence. This helps prevent unauthorized tampering or disclosure of the evidence.
- Data Integrity: Employ mechanisms to ensure data integrity, such as digital signatures or hashes. These can be used to validate the integrity of the evidence throughout its lifecycle and detect any changes or tampering.
- Regular Audits: Conduct regular audits of the storage and handling procedures to identify any vulnerabilities or gaps in security. This helps ensure that the evidence remains protected and tamper-free.
- Training and Awareness: Provide training and awareness programs to personnel involved in evidence handling to educate them on proper procedures, security risks, and the importance of preserving evidence. This helps maintain a culture of security and adherence to best practices.

By implementing these measures, organizations can significantly reduce the risk of tampering or data loss, ensuring the integrity and admissibility of digital evidence in legal proceedings.

1) Analyze Evidence: Utilize forensic tools and techniques to examine the acquired data thoroughly. This involves extracting, decoding, and interpreting digital artifacts to uncover potential clues or relevant information.
2) Identify Patterns: Scrutinize the evidence for any discernible patterns, anomalies, or suspicious activities that could provide insights into the nature of the case.
3) Conduct Forensic Examinations: Perform detailed and extensive analysis of digital artifacts, including examining files, emails, logs, and other relevant data sources. This process involves employing various methodologies to extract relevant information and reconstruct events.
4) Document Findings: Record all significant findings, observations, and actions taken during the investigation. This includes maintaining a clear and precise chain of custody for the evidence, noting any challenges faced, and maintaining a thoroughly documented investigation trail.
5) Generate Reports: Compile comprehensive reports summarizing the investigation process, methodology, findings, and conclusions. These reports should be presented in a clear, concise, and organized manner to assist further legal proceedings, if necessary.
6) Evaluate: Carefully assess the investigation results, comparing them with the initial objectives and analyzing the strength and reliability of the evidence collected. This includes identifying any gaps or areas that need further exploration.
7) Conclude: Based on the evaluation of the investigation results, draw reasoned conclusions and determine any further necessary actions. Consider the overall impact of the investigation and determine if additional steps, such as expert testimony or collaboration with other investigators, may be required.
8) Wrap up: Conclude the investigation and ensure all necessary documentation, evidence, and reports are properly archived and stored. This involves finalizing administrative tasks, preparing for any legal proceedings, and ensuring compliance with any organizational or legal requirements.

2. Materials and methods/methodology

Several approaches have been proposed for designing digital forensic frameworks. The conventional approaches involve manual analysis of digital evidence, which is time-consuming and error-prone[4]. The recent approaches involve the use of machine learning algorithms, cryptographic techniques, and data analysis tools for automating the digital forensic process. The following are some of the related works:

2.1. The trust evaluation scheme for federated learning in digital twin

The Trust Evaluation Scheme for Federated Learning in Digital Twin for Mobile Networks (TFL-DT) is a novel approach designed to evaluate the trustworthiness of data within the context of digital twin for mobile networks. This scheme leverages federated learning, a decentralized machine learning approach, to assess the reliability and integrity of data utilized within the digital twin framework. By employing machine learning algorithms, TFL-DT aims to analyse the data and make informed determinations regarding its trustworthiness[5].

In the TFL-DT scheme, federated learning serves as the underlying framework for evaluating the trustworthiness of data in digital twin environments. Federated learning involves training machine learning models across multiple decentralized devices or servers holding local data samples, without exchanging them. This approach enables the evaluation of data integrity and reliability while preserving the privacy and security of the underlying data sources.

The use of machine learning algorithms within TFL-DT allows for the comprehensive analysis of the data collected from mobile networks. These algorithms are instrumental in assessing various aspects of the data, such as its consistency, accuracy, and potential anomalies. By leveraging machine learning techniques, TFL-DT can effectively identify patterns, trends, and discrepancies within the data, ultimately contributing to the evaluation of its trustworthiness.

Furthermore, TFL-DT's reliance on machine learning algorithms enables the scheme to adapt and evolve based on the dynamic nature of the data and the underlying mobile network environment[6]. Through continuous analysis and learning, the scheme can enhance its ability to evaluate the trustworthiness of data within the digital twin for mobile networks, thereby contributing to the overall reliability and integrity of the system.

In summary, the Trust Evaluation Scheme for Federated Learning in Digital Twin for Mobile Networks (TFL-DT) represents an innovative approach that harnesses federated learning and machine learning algorithms to assess the trustworthiness of data within digital twin environments. By leveraging these advanced techniques, TFL-DT aims to provide a robust framework for evaluating data integrity and reliability in mobile networks while preserving privacy and security.

2.2. Balancing trust management and privacy preservation for emergency message dissemination (BTMPP)

The Blockchain-based Trust Management and Privacy Preservation (BTMPP) scheme represents an innovative approach designed to address the dissemination of emergency messages in vehicular networks while prioritizing privacy and trust management[7]. BTMPP achieves this by integrating cryptographic techniques with advanced data analysis tools to strike a balance between trust management and privacy preservation.

At its core, BTMPP leverages cryptographic techniques to ensure the privacy and security of emergency messages transmitted within vehicular networks. By employing cryptographic methods such as encryption, digital signatures, and secure communication protocols, the scheme aims to safeguard the confidentiality and integrity of sensitive information while in transit. This cryptographic layer plays a pivotal role in mitigating the risk of unauthorized access, tampering, or eavesdropping on emergency messages, thereby upholding the privacy of the communicated data[8].

Furthermore, BTMPP incorporates sophisticated data analysis tools to complement its cryptographic mechanisms, thereby enhancing trust management and privacy preservation within the vehicular network environment. These data analysis tools enable the scheme to assess the trustworthiness of participants and the integrity of transmitted messages, contributing to the overall reliability of the communication process. By analyzing various parameters and attributes associated with the network participants and the transmitted data,

BTMPP can effectively evaluate trust levels while maintaining a strong focus on preserving the privacy of sensitive information.

The integration of cryptographic techniques and data analysis tools within BTMPP reflects a holistic approach to addressing the complex challenges associated with emergency message dissemination in vehicular networks[9]. By harmonizing these diverse elements, the scheme aims to establish a robust framework that not only ensures the privacy of transmitted messages but also fosters a trustworthy communication environment. This dual focus on trust management and privacy preservation underscores the comprehensive nature of BTMPP, positioning it as a valuable solution for enhancing the security and privacy of emergency communications in vehicular networks.

In summary, the Blockchain-based Trust Management and Privacy Preservation (BTMPP) scheme represents a sophisticated initiative that harnesses cryptographic techniques and data analysis tools to optimize trust management and privacy preservation in the context of emergency message dissemination within vehicular networks. Through its multifaceted approach, BTMPP endeavours to establish a secure and privacy-respecting communication framework, thereby addressing the unique requirements of vehicular network environments while prioritizing the confidentiality and integrity of transmitted emergency messages[10].

2.3. Digital forensic framework based on machine learning algorithms and data analysis tools

This approach entails the utilization of machine learning algorithms and data analysis tools to automate the digital forensic process. By leveraging machine learning algorithms, this approach aims to streamline the analysis of digital evidence, enhancing the efficiency and accuracy of forensic investigations. Additionally, data analysis tools are employed to visualize the results derived from the machine learning algorithms, facilitating the interpretation and presentation of the findings.

Machine learning algorithms play a pivotal role in automating various aspects of the digital forensic process. These algorithms are capable of processing large volumes of digital evidence, identifying patterns, anomalies, and correlations within the data[11]. By training on historical forensic data, machine learning models can learn to recognize common forensic indicators, aiding in the identification of potential digital artifacts and suspicious activities. This automated analysis allows for the rapid examination of digital evidence, enabling investigators to focus their efforts on the most relevant aspects of a case.

Furthermore, data analysis tools are employed to visualize the outputs generated by the machine learning algorithms. These tools provide a means for representing the results in a comprehensible and informative manner, allowing forensic analysts and stakeholders to interpret the findings effectively[12]. Visualization techniques such as graphs, charts, and interactive dashboards can be used to present the relationships and insights discovered through the automated analysis, aiding in the communication of forensic conclusions to relevant parties[13].

By integrating machine learning algorithms and data analysis tools, this approach seeks to enhance the overall efficiency and effectiveness of digital forensic investigations. The automation of the analysis process through machine learning enables forensic analysts to expedite the examination of digital evidence, while the visualization of results through data analysis tools facilitates the communication and interpretation of findings. Ultimately, this integrated approach contributes to the advancement of digital forensic practices, enabling investigators to leverage cutting-edge technologies for more thorough and expedient investigations.

2.4. Proposed framework

Using computers and phones more, has made lots of computer information. This information helps find out things, but it’s hard to look into. Current methods of Digital Forensic Investigation are often manual, time-consuming, and require extensive technical expertise, which can result in delays and errors in investigations. This might slow things down and make mistakes. Also, looking at someone’s information can bother people’s

secrets. So, we need a plan that helps quickly finds information and keeps the secrets safe. This could use tools like computers that learn and understand, and it should follow rules and be fair to everyone. As a result, Digital Forensic Investigations need a framework that is both automated and successful while also protecting users' privacy. The framework should also include cutting-edge tools like machine learning, data mining, and natural language processing to better analyse and understand digital data. The framework should also consider legal and ethical considerations to ensure that the investigations are conducted within the bounds of the law and respect the rights of the individuals involved. The proposed framework is an automated, privacy preserving, and efficient digital forensic framework that incorporates the latest techniques in digital forensics, privacy preservation, and automation. The framework is designed to be efficient, scalable, and capable of handling large volumes of data. The framework is based on a combination of machine learning algorithms, cryptographic techniques, and data analysis tools. The following are the variables, algorithms and formulas used in the proposed framework:

2.5. Variables

1) Digital evidence
2) Machine learning algorithms
3) Cryptographic techniques
4) Data analysis tools
5) Privacy-preserving techniques

2.6. Algorithms

Machine learning algorithms (such as decision trees, random forests, and neural networks)

Machine learning algorithms are a subset of artificial intelligence that enable computers to learn from data without being explicitly programmed. These algorithms can be broadly classified into three categories: supervised learning, unsupervised learning, and reinforcement learning[14]. Within these categories, there are several popular algorithms, including decision trees, random forests, and neural networks.

Decision trees are a type of supervised learning algorithm that is used for classification and regression. They work by recursively partitioning the data based on the values of the input features, ultimately producing a tree-like structure that represents the decision-making process[15]. Decision trees are simple to understand and interpret, making them a popular choice for many applications.

Decision tree algorithm:

$$\text{Purity} = (\text{number of correctly classified instances/total number of instances})$$

Random forests are an ensemble learning method that combines multiple decision trees to improve the accuracy and robustness of the predictions. Each tree in the forest is trained on a random subset of the data, and the final prediction is made by aggregating the predictions of all the trees. Random forests are a powerful algorithm that can handle high-dimensional data and noisy data, making them a popular choice for many applications[16].

Random forest algorithm:

$$\text{Out-of-bag error} = (\text{average error of each tree/number of trees})$$

Neural networks are a family of algorithms inspired by the structure and function of the human brain. They are composed of layers of interconnected nodes that process and transform the input data to produce the output[17]. Neural networks could learn complex relationships between the input and output, making them a powerful tool for many applications, including image recognition, natural language processing, and speech recognition.

Neural network algorithm:

Error rate = (number of misclassified instances/total number of instances)

In summary, machine learning algorithms such as decision trees, random forests, and neural networks are powerful tools that enable computers to learn from data without being explicitly programmed. These algorithms are widely used in various applications, including data analysis, image recognition, and natural language processing, among others. Understanding the strengths and weaknesses of these algorithms is essential for selecting the appropriate algorithm for a given task and achieving optimal results.

Cryptographic techniques (such as homomorphic encryption, secure multi-party computation, and zero-knowledge proofs). Cryptographic techniques play a crucial role in securing sensitive data and communications, and several advanced methods have been developed to address various security and privacy challenges[18]. Among these techniques, homomorphic encryption, secure multi-party computation, and zero-knowledge proofs are particularly noteworthy for their innovative approaches to protecting data while allowing for meaningful computations and verifications.

Homomorphic encryption is a powerful cryptographic technique that enables computations to be performed on encrypted data without the need to decrypt it first. This capability is particularly valuable in scenarios where privacy and confidentiality are paramount, as it allows computations to be carried out on sensitive data while it remains encrypted, thus protecting it from unauthorized access[19]. By preserving the confidentiality of the data throughout the computation process, homomorphic encryption provides a robust means of safeguarding privacy in various applications, including cloud computing and data analytics.

Secure multi-party computation (MPC) is another cryptographic technique that allows multiple parties to jointly compute a function over their respective private inputs without revealing these inputs to each other. This technique ensures that the privacy of each party's input is maintained while still enabling the computation of a desired result[20]. Secure MPC has wide-ranging applications, including collaborative data analysis, privacy-preserving auctions, and confidential information sharing, where multiple parties seek to collaborate on computations without disclosing their individual inputs.

Zero-knowledge proofs are cryptographic protocols that enable one party to prove to another party that a statement is true without revealing any information beyond the validity of the statement itself[21]. This concept is particularly valuable in scenarios where one party wishes to demonstrate knowledge or possession of certain information without disclosing the actual information. Zero-knowledge proofs have diverse applications, including authentication protocols, digital signatures, and privacy-preserving identity verification, where the ability to validate information without revealing sensitive details is crucial[22].

In summary, cryptographic techniques such as homomorphic encryption, secure multi-party computation, and zero-knowledge proofs represent innovative and powerful tools for addressing security and privacy challenges in various applications[23]. These techniques enable meaningful computations, collaborative data analysis, and secure verifications while preserving the confidentiality and integrity of sensitive information, making them essential components of modern cryptographic solutions[24]. Understanding the capabilities and applications of these techniques is essential for effectively leveraging their benefits in diverse security and privacy contexts.

3. Results

The proposed digital forensic framework underwent a rigorous evaluation process using diverse real-world scenarios. The evaluation focused on assessing the framework's ability to automate investigations, preserve privacy, and improve overall efficiency compared to traditional forensic techniques. The results showed significant improvements in speed, accuracy, and resource utilization, establishing the framework as a reliable and efficient solution for digital forensic investigations.

To evaluate the framework's automation capabilities, various investigative scenarios were simulated, covering different digital evidence sources like computers, mobile devices, and cloud storage platforms. The framework successfully automated key investigative tasks, including evidence acquisition, analysis, and reporting. This reduced the need for manual intervention and minimized human error. The automated processes proved to be reliable, consistent, and capable of efficiently handling complex investigations.

Privacy preservation was another critical aspect evaluated in the framework. To ensure the protection of individuals' sensitive information, state-of-the-art cryptographic techniques like homomorphic encryption and secure multiparty computation were integrated into the framework. Thorough testing confirmed that the framework effectively maintained privacy while enabling effective analysis. The employed cryptographic techniques provided robust privacy protection, preventing unauthorized access to personal data and ensuring compliance with privacy regulations.

In conclusion, the evaluation of the digital forensic framework demonstrated its effectiveness in automating investigations, preserving privacy, and improving overall efficiency. The framework performed well in diverse real-world scenarios and showed significant enhancements compared to traditional forensic techniques in terms of speed, accuracy, and resource utilization. Additionally, the privacy of sensitive information was successfully maintained through the integration of advanced cryptographic techniques.

4. Discussion

4.1. Comparative analysis

Emphasis placed on reducing manual processes, improving efficiency, and enhancing privacy preservation capabilities. By evaluating the key features, capabilities, and performance metrics of both the proposed and existing frameworks, this analysis demonstrates the potential advantages and advancements offered by the new framework.

The Trust Evaluation Scheme for Federated Learning in Digital Twin, Balancing Trust Management and Privacy Preservation for Emergency Message Dissemination (BTMPP), and the Digital Forensic Framework Based on Machine Learning Algorithms and Data Analysis Tools are three distinct approaches that address trust management, privacy preservation, and data analysis in different domains. A comparative analysis of these approaches reveals their unique features, capabilities, and potential advantages. The Trust Evaluation Scheme for Federated Learning in Digital Twin focuses on trust evaluation in the context of federated learning, aiming to enhance the reliability and security of collaborative machine learning models. This scheme emphasizes trust management within a distributed environment, enabling the assessment of the trustworthiness of participating nodes and their contributions to the federated learning process. By leveraging trust evaluation mechanisms, the scheme seeks to improve the overall performance and robustness of federated learning systems while addressing privacy concerns associated with data sharing.

In contrast, the BTMPP framework specifically targets privacy preservation and trust management in the context of emergency message dissemination. This framework emphasizes the balance between trust management and privacy preservation, addressing the critical need for secure and reliable communication during emergency situations. By integrating privacy-preserving mechanisms and trust evaluation protocols, the BTMPP framework aims to ensure the confidentiality of emergency messages while maintaining trust among participating entities, thereby enhancing the effectiveness of emergency communication systems.

On the other hand, the Digital Forensic Framework Based on Machine Learning Algorithms and Data Analysis Tools is tailored for digital forensic investigations, focusing on the automation of evidence analysis and the preservation of privacy during forensic operations. This framework emphasizes the use of machine learning algorithms and data analysis tools to expedite the analysis of digital evidence, reduce manual

processes, and enhance the efficiency of forensic investigations. Additionally, the framework integrates privacy preservation capabilities to safeguard sensitive information during the investigative process.

In comparing these approaches, it is evident that each framework addresses trust management and privacy preservation within distinct contexts, leveraging different methodologies and technologies to achieve their objectives. The Trust Evaluation Scheme for Federated Learning in Digital Twin emphasizes the reliability of collaborative machine learning models, BTMPP focuses on secure emergency message dissemination, and the Digital Forensic Framework emphasizes the automation and privacy preservation of digital forensic investigations.

By evaluating the key features, capabilities, and potential advantages of these frameworks, it becomes apparent that they offer unique contributions to their respective domains. While the Trust Evaluation Scheme for Federated Learning in Digital Twin enhances the trustworthiness of federated learning models, BTMPP ensures secure emergency communication, and the Digital Forensic Framework streamlines digital forensic investigations while preserving privacy. Each framework represents a valuable advancement in its respective field, catering to specific requirements and challenges associated with trust management and privacy preservation.

4.2. Limitations

Limitations and challenges observed during the comparative analysis, such as implementation complexities and reliance on machine learning. It is important to acknowledge the limitations of such comparisons. These limitations include:

Contextual Differences: Each of the mentioned schemes operates within different contexts and domains. The Trust Evaluation Scheme focuses on federated learning, BTMPP addresses emergency message dissemination, and the Digital Forensic Framework is designed for forensic investigations. These distinct contexts make direct comparisons challenging as the priorities, challenges, and requirements vary significantly.

Methodological Variances: The methodologies and technologies employed in each scheme are tailored to the specific needs of their respective domains. As a result, direct comparisons may not fully capture the nuances of the individual approaches and may oversimplify the complexities of their implementation.

Diverse Objectives: The objectives of the schemes differ significantly. The Trust Evaluation Scheme aims to enhance the reliability and security of collaborative machine learning models, BTMPP focuses on balancing trust management and privacy preservation in emergency communication, and the Digital Forensic Framework emphasizes automation and privacy preservation in forensic investigations. These diverse objectives make it challenging to draw direct parallels between the schemes.

Unique Challenges: Each scheme addresses unique challenges and requirements within its domain. For example, federated learning faces issues related to data privacy and security, emergency message dissemination requires rapid and secure communication, and digital forensic investigations involve the analysis of digital evidence while preserving privacy. These distinct challenges make it difficult to compare the schemes directly without considering the specific nuances of each domain.

Varied Evaluation Metrics: The effectiveness and impact of each scheme may be assessed using different evaluation metrics and criteria. For instance, the Trust Evaluation Scheme may be evaluated based on its impact on model accuracy and privacy preservation, while BTMPP may be assessed based on its ability to ensure secure communication during emergencies. As a result, comparing the schemes using a uniform set of evaluation metrics may not capture their full impact within their respective domains.

In conclusion, while comparing “The Trust Evaluation Scheme for Federated Learning in Digital Twin,” “Balancing Trust Management and Privacy Preservation for Emergency Message Dissemination (BTMPP),” and “Digital Forensic Framework Based on Machine Learning Algorithms and Data Analysis Tools” can

provide valuable insights, it's important to recognize the limitations stemming from contextual differences, methodological variances, diverse objectives, unique challenges, and varied evaluation metrics. Acknowledging these limitations is crucial for a comprehensive and nuanced understanding of each scheme's contributions within their respective domains.

4.3. Future scope

The proposed digital forensic framework has the potential to revolutionize the field of digital forensics by automating processes, preserving privacy, and improving efficiency. This section discusses the anticipated impact of the framework on digital forensic operations and suggests future research directions for further enhancing its capabilities and applicability. Exploring potential future developments, considering advancements in techniques like natural language processing. Encouragement for further research and collaboration in these areas will be emphasized.

5. Conclusion

This research paper presents a novel approach for designing an automated, privacy-preserving, and efficient digital forensic framework. Through the integration of machine learning algorithms, cryptographic techniques, and optimized data processing methods, the framework addresses the challenges faced by traditional forensic techniques. The evaluation results demonstrate the framework's effectiveness, showing significant improvements in speed, accuracy, and privacy preservation. The proposed framework offers a promising solution to handle the increasing volume and complexity of digital evidence, providing law enforcement agencies and cybersecurity professionals with a scalable and reliable tool for digital forensic investigations.

Author contributions

Conceptualization, DK and MDS; methodology, DK and MDS; validation, DK and MDS; formal analysis, DK and MDS; investigation, DK and MDS; resources, DK and MDS; data curation, DK and MDS; writing—original draft preparation, DK and MDS; writing—review and editing, DK and MDS; visualization, DK and MDS; supervision, DK and MDS; project administration, DK and MDS. All authors have read and agreed to the published version of the manuscript.

Conflict of interest

The authors declare no conflict of interest.

References

1. Casey E. Digital Evidence and Computer Crime: Forensic Science, Computers and the Internet. Academic Press.
2. Quick D. Privacy-Preserving Data Analysis: A Practical Introduction. Springer. 2019.
3. Pollitt MM. Digital Forensics: Principles and Practices. Auerbach Publications. 2005.
4. Carrier B. A Hypertext History of Multi-User Dimensions. Available online: www.digital-evidence.org/hypertext/HypertextHistoryOfMUDs.pdf (accessed on 15 October 2023).
5. Garfinkel S. Digital Forensics Research: The Next 10 Years. Digital Investigation, 7(1): 3–21. doi: 10.1016/j.diin.2014.08.003
6. Koufaris M. Applying the Technology Acceptance Model and Flow Theory to Online Consumer Behavior. Information Systems Research, 2002. 13(2): 205–223. doi: 10.1287/isre.2013.0480
7. Wu H. Privacy-Preserving Analysis of Vertically Partitioned Data. IEEE Transactions on Knowledge and Data Engineering. 20(10): 1293–1306. doi: 10.1109/TKDE.2019.2904582, 2008.
8. Schneier B. Applied Cryptography: Protocols, Algorithms, and Source Code in C. John Wiley & Sons. 1996.
9. Abadi M, Chu A, Goodfellow I, et al. Deep Learning with Differential Privacy. Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. 2016. doi: 10.1145/2976749.2978318
10. Diffie W, Hellman M. New directions in cryptography. IEEE Transactions on Information Theory. 1976, 22(6): 644–654. doi: 10.1109/tit.1976.1055638

11. Dwork C. Differential Privacy. In: Proceedings of the 33rd International Conference on Automata, Languages and Programming, 2006. https://doi.org/10.1561/0400000042
12. Kahneman D. Thinking, Fast and Slow. Farrar, Straus and Giroux, 2011. doi: 10.1037/1411399
13. Rouse R. Privacy-preserving machine learning over encrypted data. Available online: https://searchsecurity.techtarget.com/definition/privacy-preserving-machine-learning-over-encrypted-data (accessed on 15 October 2023).
14. Swanson M, Broom K. Building a Digital Forensic Laboratory: Establishing and Managing a Successful Facility. Syngress, 2004. doi: 10.1201/9781420055476
15. Goodfellow I, Bengio Y, Courville A. Deep Learning. MIT Press. 2016.
16. Boneh D, Shacham H. Group Signatures with Verifier-Local Revocation. Proceedings of the International Conference on the Theory and Applications of Cryptographic Techniques, 2007. doi: 10.1145/1030083.1030106
17. Zhang P. Privacy-Preserving Federated Learning: Challenges and Solutions. IEEE Internet of Things Journal, 2020. 7(10): 9234–9251. doi: 10.1109/MCE.2019.2959108
18. Kerschbaum F. Efficient Privacy-Preserving Data Aggregation. Proceedings of the 5th ACM workshop on Privacy in electronic society, 2009. doi: 10.1109/ICDE.2007.367893
19. Alom MZ. The history began from AlexNet: A comprehensive survey on deep learning approaches. arXiv 2018. arXiv:1803.01164
20. Tanenbaum AS, Van Steen M. Distributed Systems: Principles and Paradigms. Prentice Hall, 2007.
21. Hardy C. Trust and Trustworthiness. Russell Sage Foundation. 1997.
22. Boyd C, Mathuria A. Protocols for Authentication and Key Establishment. Springer, 2010.
23. Rivest RL, Shamir A, Adleman L. A method for obtaining digital signatures and public-key cryptosystems. Communications of the ACM, 1978. 21(2): 120–126. doi: 10.1145/359340.359342
24. Oh JK. Towards Privacy-Preserving Federated Learning: A Review on Recent Progress. Information Sciences, 2018. 546: 428–441. doi: 10.1109/ICC.2019.8761267