A Novel Approach for Digital Evidence Management using Blockchain

Sagar Rao
Department of Computer Engineering
Don Bosco Institute of Technology
Mumbai 400070, India
raosagar215@gmail.com

Shalomi Fernandes
Department of Computer Engineering
Don Bosco Institute of Technology
Mumbai 400070, India
shalomi1998@gmail.com

Samruddhi Raorane
Department of Computer Engineering
Don Bosco Institute of Technology
Mumbai 400070, India
samruddhi.raorane20@gmail.com

Shafaque Syed
Department of Computer Engineering
Don Bosco Institute of Technology
Mumbai 400070, India
shafaque.dbit@dbclmumbai.org

***Abstract*— In the proceedings of the court of law, it is of vital importance that the digital evidences are produced untampered during the entire Chain of Custody (Coc) process. Recent advancements in computer technology have enabled hackers to compromise with the integrity of the evidences. In order to secure these evidences from such external agents we have proposed a Blockchain model. Unlike the traditional security systems, Blockchain Technology uses cryptographic hashing in which each block has a cryptographic hash of its previous block i.e., A growing list of records that are called Blocks. These blocks are secure and tamperproof. The idea of this work is to handle the digital evidence from the time it is retrieved to the point it is presented as an evidence in the court of law. It ensures safeguarding of the evidence during the entire Chain of Custody process, which is fundamental to the idea of digital forensics. The aim of Digital Forensics is to maintain evidences in their most original form while performing a structured investigation by the collection, identification and validation of the digital information for the purpose of reconstructing events from the past. The proposed system ensures integrity, traceability, authenticity and security of the evidences.**

Keywords—Digital Forensics, Digital Evidence, Chain of Custody,Blockchain

I. INTRODUCTION

Digital Forensics (DF) is defined as the method that includes identifying, investigating, preserving, analyzing, validating and presenting digital evidence in a way that is traditionally permitted by the law in any legal proceedings i.e. the Court of Law. Digital evidence is gaining more and more importance because it is used to prove facts or to convict personnel engaged in cyber-crimes. The aim of digital forensics is to ensure that digital evidence is permissible in the court of law. Subsequently, it is extremely important to ensure the integrity of digital evidence during its whole lifecycle in any forensic investigation. Digital evidence is difficult to handle and maintain as compared to physical evidence because it has characteristics like easy to transmit, fragile in nature, vulnerable to tampering and removal, defenseless against altering and evacuation and time sensitiveness.

The main perspective in digital forensics is Chain of Custody (CoC). [1] Digital Chain of Custody is a procedure to handle digital evidence in investigations. It is used to maintain and document digital evidence in chronological order right from when the incident took place until it is presented to the court. In this process, it goes through different levels of hierarchy right from the first evidence collector to the higher authority court. CoC is a very important part of the investigation process. CoC records all the minute details related to the evidence. It records the five W's who, when, why, where and how with the evidences in each stage of investigation. In order to get the evidence to be accepted by the court, CoC must be maintained properly and should ensure that it should be tamper-proof. It is very important to preserve the integrity of the evidence.[2]

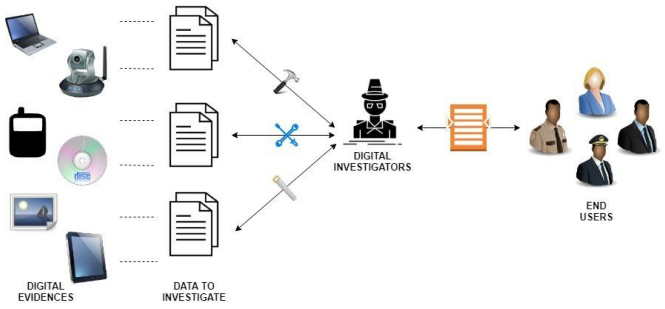


Fig. 1. Digital Forensics Process

The idea of this project is to handle the digital evidence right from the time it is gathered until it is presented as proof in the court. During this process, several people handle it and there is a possibility that the evidence may get tampered. Hence to ensure tamper resistance, it is important that the integrity, authenticity, and security is maintained. During the literature survey, we realised that a Blockchain ensures these features best and the details of the same have been discussed in the following section.[3]

II. Related Work

Satoshi Nakamoto is the name of an individual or group of individuals who have been credited for the innovation of the Blockchain Technology in the year 2008. We take inspiration from Satoshi Nakomoto's pioneering work for securing Electronic cash transactions using blockchain. We intend to implement a similar framework for securing the Chain of Custody for Digital Evidence Management.[4]

In the work of Chung and Park [5] there is talk about certain strategies for digital forensics related to the IVA Alexa's ecosystem. Recently, due to a rapid range of advancements of Internet of Things (IoT) devices in our day to day lives, numerous people are changing from standard home devices to connecting with diverse IoT consumer products. In such kind of situations, a lot of data is being manufactured in real-time responding to user's behaviors. Until now, there has been a low number of studies detailed on the Amazon Echo and its ecosystem with regards to digital forensics. This work brought ahead a new unified approach combining both cloud-native as well as client-centred forensics for Amazon's Alexa ecosystem as a factor of making an effort to prepare for IoT.

The paper Simson L. Garfinkel [6] gives a framework of present research on forensic bearings and proclaims that, to accomplish the society needs to accept systematic and extensible approaches for information portrayal and forensic proceedings. Today Digital Forensics is an essential gadget for solving crimes related to computers, as well as for gathering digital evidence. This work points out that we are in a "Golden Age of Digital Forensics," which is concluding rapidly.

Coronel, Cedillo, and Camacho, [7] explains how digital evidence can be handled at the client's side pertaining to techniques of identification, collection, recognizable proof, assortment, analysis, etc. Likewise, an audit of how principles are being utilized in cyber forensics concentrated on the client-side.

Breitinger, Gradeja and Baggili, [8] discusses about two main objectives. The first objective is to give a synopsis of all the available datasets that can be utilized by the investigators and how to discover them. The second objective is emphasizing on sharing the datasets to allow researchers to recreate the results and better the standard results.

The phases involved in crime scene investigation are: preserving the scene; surveying for evidence; documenting the evidence and scene; searching for the evidence; reconstructing the scene. All these phases are important, but preserving and documenting evidence is more important. The people who can work on the digital evidences are: The First Responders, The Forensic investigators, The Police officers, The Court experts, The Victims and the Suspects. Each of them can affect the evidences. This violates integrity. [9]

According to Lee, Kim, Lee and Lim [2], integrity can be guaranteed using hashing.
There are three methods suggested in this.
i) MDC publication method.
ii) Message Authentication Code (MAC) Authentication method.
iii) Public Authentication System with the Public Key Infrastructure (PKI).

As per Flores and Jhumka [10], CoC is monitored according to the 4 principles:
1: The insider is not at the liberty of making any alterations to the original evidence.
2: Whenever the original data is required to be produced, a sound justification must be provided for implementing such an action.
3: The audit trail produced of all of the events should produce the same result regardless to the nature of the arbitrator that it is subjected to.
4: The person responsible for investigation must ensure that all of these principles have been obeyed.

Prayudi and Sn [1] gives a summary of the extent to which the issues and challenges are faced in the digital chain of custody. It also discusses ways of handling CoC. Some aspects of handling digital chain of custody are forensic format, storage, security assurance of storage and access control of storage.

Selamat, Sahib, Hassan, Abdollah, and Abidin, [11] demonstrates the traceability data. It describes about the traces and the sources associated with investigation process. Tracing is a method of discovering the origin or cause of certain scenario. Tracing discovers the traces left in evidences. The traces are collected and preserved. These traces are then inspected and analyzed. This paper proposes a Traceability model.

In [12] a Blockchain based mobile edge computing model is created for data sharing and in-home therapy management. The model framework uses a Tor network for a distributed system. Blockchain is used to store the therapeutic data of the patients. The patient data should be very confidential, thus considering this the support of Blockchain is taken. [13] The therapy framework uses blockchain so that the patient is the true owner of the data. Blockchain preserves the privacy and ownership of the data. The blockchain stores the therapy metadata but the actual data is such as multimedia files are stores in a distributed or centralised database.

In today's technological phase, there are a lot of inventions. These inventions can be protected by law and are known

to be intellectual property. Protecting intellectual property is very important. Blockchain is used to store intellectual property, which is proposed in [14]. The time stamping feature of Blockchain is very useful in this case. Blockchain is used for intellectual property confirmation as the creator of the work is the person who accesses the intellectual property document at first and thus the ownership of the creation can be proved by tracing back in the particular timestamp in the Blockchain. Blockchain is also used for intellectual property transactions where the transactions are between the person using the intellectual property and the owner of that property.

Roman Beck, Michel Avital, Matti Rossi and Jason Bennett Thatcher discuss how a reliable distributed storage system is very important for relationships between businesses and organizations in their work [15]. The economy depends on the trust between individuals and organizations for creating and storing important records. The Blockchain technology can be used in various industries and businesses such as shipping industries for tracking the shipping containers, pharmaceutical industries for tracking the medicines, healthcare industries, etc. Blockchain would provide such industries a secure record storage and a proper ownership of the asset by each business.

With the fourth industrial evolution, the importance and use of Blockchain has increased drastically. It is being used in various fields ranging from finance to commerce. The research in this paper [16] was focused on the use of this technology in the field of education. Currently, Blockchain technology is being used in academia mainly for keeping track of students achievements and certificates/degrees. Data like students' interest, learning experience, etc along with the academia details are stored in the Blockchain. This helps to tackle degree fraud as it is a growing issue in the traditional system.

The Blockchain system is going to be very complex in nature and not easily maintainable. The immutability feature can cause issues as even legitimate users cannot edit their data. Furthermore, any technological failure can cause significant issues in the education system.

III. Existing System

Evidence is the foundation of every cybercrime. In crime scene investigations, two of the most vital factor is collecting the digital evidences and preserving the digital evidence. Every Electronic evidence has a physical form and they can be identified with naked eyes. Examples of Electronic evidences can be computer, mobile phone, camera, CD, hard disk, etc., whereas digital evidences are evidences that are recuperated or extracted from electronic evidences. They can be a file, an email, a short message, an image, a video, a log or a text.

Physical as well as digital crime scene investigations both are composed of a small number of basic building blocks such as preservation of the crime scene; surveying for evidence; documentation of the evidence and scene; searching for evidence; reconstruction of the scene; and presenting a theory. Physical and digital evidence both together create an entire theory about the crime scene. [6]

The management and handling of digital evidence is a tedious task. This is why a system environment that supports the implementation of managing and handling digital chain of custody is needed by the institution of law-enforcement for supporting the managing and handling of cybercrime investigations. The main issue in the chain of custody is related to the documentation of the evidence.

The flaws in the current system are:
(i) Paper documents: may last for many years but may also fail in few minutes due to extreme heat, fire, heaters, radiators, shredding.
(ii) Tape, CD's, disks: may last a couple of years but may fail in few minutes due to excessive heat or due to electromagnetic, a strong magnet and high impulse vibrations can be used for overwriting.
(iii) Storage of media data is difficult.
(iv) Tampering of Evidence thus not providing integrity and confidentiality.

IV. Proposed System

The proposed architecture has five main components namely Participants, Front-End, Flask Application, Core Modules, Blockchain Network, and Evidence Store as shown in Figure 2 and 3.

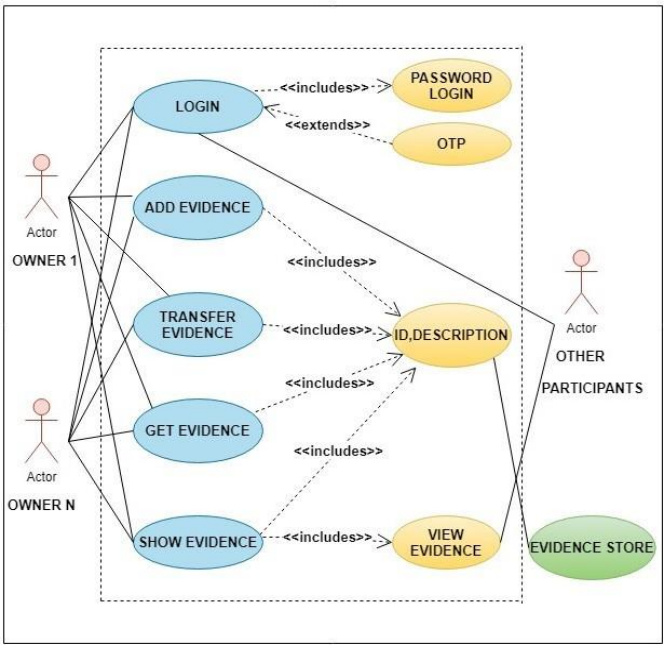


Fig. 2. Use Case-How participants in the network will access the Modules

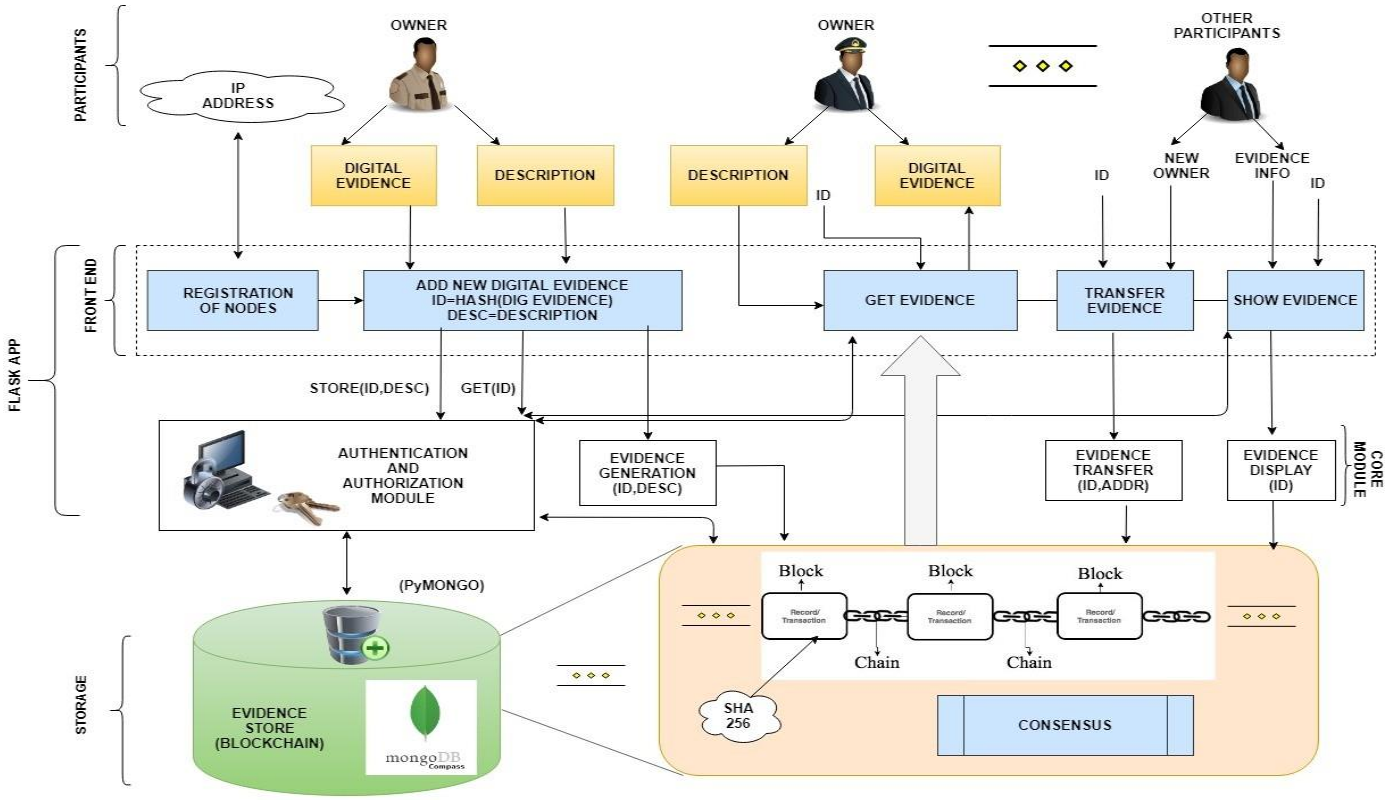


Fig. 3. Block Diagram of the working of the entire Blockchain Network

1. Participants:

Participants are the people in the network who store the blockchain information. In the blockchain model, participants are the Forensic Investigators whose work is to gather as much information about the digital evidence as possible and record it in the Blockchain. Participants after gathering the evidence information store the evidences in the Blockchain.

2. Front-End:

The blockchain model's front-end is developed with the help of python and flask application. All the modules are added in the front end.

3. Flask Application: Flask is a web framework of python and thereby flask helps in building a web application. The flask application contains the authentication and authorization modules, the core modules and the front end.

4. Core Modules:

The core modules are the main functions that the participants use while using the blockchain network. The participants use these core modules to change the state of the blockchain.

5. Blockchain Network:

It comprises of a Peer-to-Peer network where all the participants are under a consensus agreement and are connected to each other. This is a decentralised network where there is no central authority.

6. Evidence Store:

The whole chain of custody which is stored as a Blockchain is stored in the MongoDB database. Each Block has an index, a timestamp, the information about the evidence, proof, and the hash of the previous block. For connecting the front-end to the database PyMongo is used. Whenever the chain is updated, the same change is reflected in the database.

In proposed Forensic-Chain model transactions either record the details about the evidence or the evidence transfer event on the network.

A. Implementation Details:

This work has been split up into multiple modules and they are as follows:

1. Registration of Nodes: This module deals with the registering of nodes or users to the Blockchain network so that each node is visible to every other node in the Blockchain network.

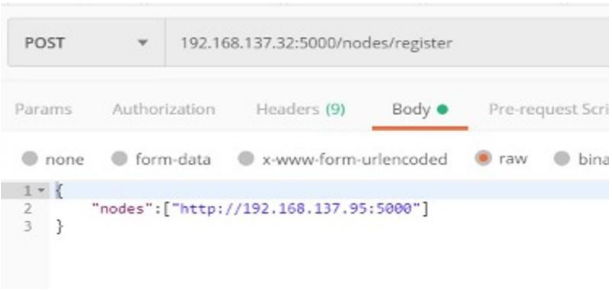


Fig. 4. Registration of Nodes

2. Creation of the Evidence:

This module takes all the information related to the evidence from the participants and adds the evidence to the Blockchain. Each evidence has a unique identification, i.e, ID. The ID of the evidence is created using the hash of the evidence. After the participant adds all the evidence information, mining is performed to create the block. Mining is the process of block creation.

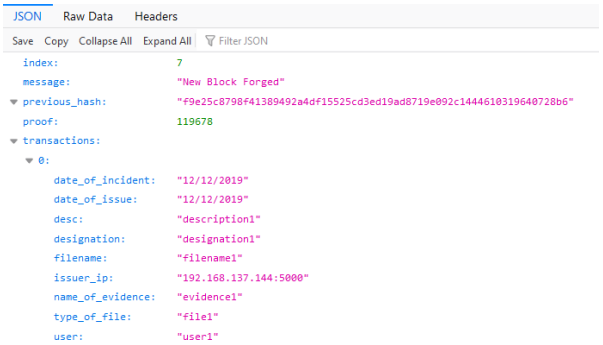


Fig. 5. Creation of the Evidences

3. Display of the Evidence:

Evidence Display module takes an evidence name or ID as an input and returns the evidence information from the Blockchain. The only check this function does is to ensure evidence already exists.

Fig. 6. Display of the Evidences

4. Transfer of the Evidence:

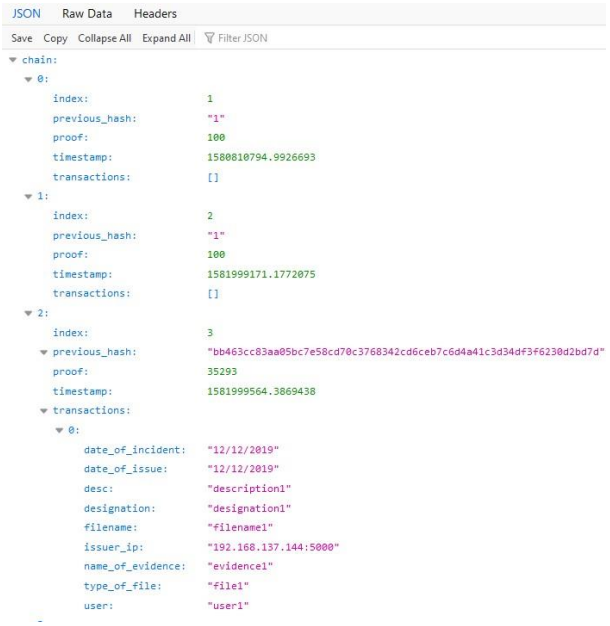
Evidence Transfer module is used for the synchronization of the Blockchain across the Blockchain network. When a participant adds a new block in the Blockchain, the new block must be shown and reflected among all the participants.

Thus this module is used so that the Blockchain data across the network is the same and there is no dissimilarity in the Blockchain across the participants in the network.

Fig. 7. Transfer of the Evidences

5. Get Evidence CoC(Chain of Custody):

Returns the whole chain of evidence from the first evidence to the last evidence.



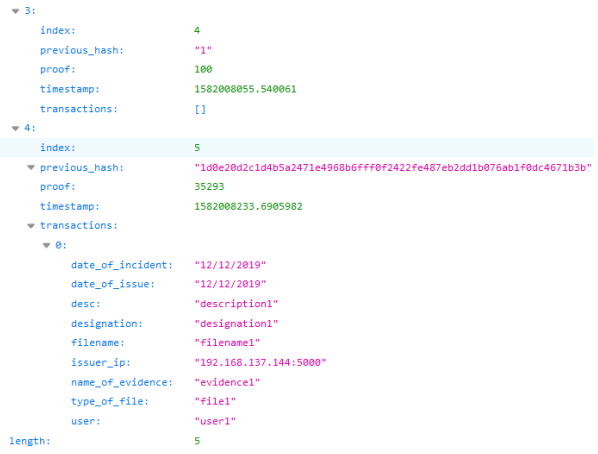


Fig. 8. Entire Chain of Custody

B. Benefits of proposed system:

CHAVI- A Novel Approach for Digital Evidence Management has great potential and has benefits to forensics based applications. It helps in preserving integrity, transparency, security, authenticity and auditability of digital evidence.

Few benefits are summarized below:

i) Collecting, preserving and validating digital evidence can be built easily using Forensic chains.

ii) Blockchain driven Forensic-Chain eliminates the requirement of a trusted and reputed third party for validating certain claims or the transfer of evidence and achieving consensus.

iii) Blockchain by its pattern applies integrity, transparency, authenticity, security, and auditability thus making it possibly one of the best choices for maintenance and tracing back of the forensic chain of custody.

iv) The blockchain will be private as the data should only be shared among trusted participants

V. Further Work

Further work includes building a practical, larger blockchain to use for processing and storing whole evidence files. The blockchain will be a private one as the data should only be shared among trusted participants. This will help the investigators to handle the digital evidence safely and thus guarantees the integrity of the digital evidence. All the pieces of evidence will have a guarantee of being secure and tamperproof.

VI. Conclusion

In this paper, a new system for digital evidence management is proposed which utilizes Blockchain technology. Blockchain by its pattern applies the concepts of integrity, transparency, security, authenticity, and auditability. This makes it one of the best choices for the maintenance and tracing back of the forensic chain of custody. The digital evidence will be handled by this system that has been proposed from the time it is retrieved until it is presented as evidence in the court. It will guarantee integrity, traceability, authenticity and security of the evidence. It will also help guarantee that this digital evidence gets admissible in the Court of Law.

References

[1] Yudi Prayudi and Azhari Sn. Digital chain of custody: State of the art. *International Journal of Computer Applications*, 114:975–8887, April 2015.

[2] Seokhee Lee, Hyunsang Kim, Sangjin Lee, and Jongin Lim. Digital evidence collection process in integrity and memory information gathering. In *First International Workshop on Systematic Approaches to Digital Forensic Engineering (SADFE'05)*, pages 236–247, Nov 2005.

[3] Auqib Hamid Lone and Roohie Naaz Mir. Forensic-chain: Blockchain based digital forensics chain of custody with poc in hyperledger composer. *Digital Investigation*, 28:44–55, March 2019.

[4] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. *Cryptography Mailing list at https://metzdowd.com*, 03 2009.

[5] Sangjin Lee Hyunji Chung, Jungheum Park. Digital forensic approaches for amazon alexa ecosystem. *Digital Investigation*, 22:S15–S25, August 2017.

[6] Simson L. Garfinkel. Digital forensics research: The next 10 years. *Digital Investigation*, 7:S64–S73, August 2010.

[7] Karina Campos Bryan Coronel, Priscila Cedillo and Jessica Camacho. A systematic literature review in cyber forensics: Current trends from the client perspective. In *IEEE Third Ecuador Technical Chapters Meeting (ETCM) 2018*, pages 1–6. IEEE, 2018.

[8] Frank Breitinger Cinthya Grajeda and Ibrahim Baggili. Availability of datasets for digital forensics and what is missing. *Digital Investigation*, 22:S94–S105, August 2017.

[9] J. Cosic and M. Baca. Do we have full control over integrity in digital evidence life cycle? In *Proceedings of the ITI 2010, 32nd International Conference on Information Technology Interfaces*, pages 429–434, June 2010.

[10] D. A. Flores and A. Jhumka. Implementing chain of custody requirements in database audit records for forensic purposes. In *2017 IEEE Trustcom/BigDataSE/ICESS*, pages 675–682, August 2017.

[11] S. R. Selamat, R. Yusof, S. Sahib, N. H. Hassan, M. F. Abdollah, and Z. Z. Abidin. Traceability in digital forensic investigation process. In *2011 IEEE Conference on Open Systems*, pages 101–106, September 2011.

[12] Abdur Rahman, M. Shamim Hossain, George Loukas, Elham Hassanain, Syed Rahman, Mohammed Alhamid, and Mohsen Guizani. Blockchain-based mobile edge computing framework for secure therapy applications. *IEEE Access*, PP:1–1, November 2018.

[13] Asaph Azaria, Ariel Ekblaw, Thiago Vieira, and Andrew Lippman. Medrec: Using blockchain for medical data access and permission management. pages 25–30, 08 2016.

[14] Junyao Wang, Shenling Wang, Guo Junqi, Yanchang Du, Shaochi Cheng, and Xiangyang Li. A summary of research on blockchain in the field of intellectual property. *Procedia Computer Science*, 147:191–197, January 2019.

[15] Roman Beck, Michel Avital, Matti Rossi, and Jason Thatcher. Blockchain technology in business and information systems research. *Business Information Systems Engineering*, 59, November 2017.

[16] Guang Chen, Bing Xu, Manli Lu, and Nian-Shing Chen. Exploring blockchain technology and its potential applications for education. *Smart Learning Environments*, 5(1), March 2018.