Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 08, August-2021

Criminal Digital Forensic Investigation Application based on Blockchain

Charan T S
PG Student
Department of MCA
PES College of Engineering Mandya,
Karnataka, India

K M Sowmyashree
Assistant Professor
Department of MCA
PES College of Engineering Mandya,
Karnataka, India

***Abstract*—In recent years, blockchain has generated a lot of buzz. Although studies have been conducted on blockchain technology, it still requires more research before it can be fully implemented. By adopting blockchain, the system's decentralized nature will help to increase transparency and make the system safer for everyone. Blockchain is a digital ledger of transactions that is duplicated and spread over the whole network of computer systems on the blockchain, and it is a system of preserving information in a way that makes it difficult or impossible to edit or hack. Blockchain platforms haven't progressed far enough to be compatible with existing Digital Forensic (DF) tools, methodologies, and procedures. Collecting, identifying, evaluating, analysing, preserving, and presenting evidence is a difficult task in existing forensics. That's where the blockchain's decentralised nature comes into play, with forensic data being stored in the private network using the blockchain's nodes in a peer-to-peer network. It's also well-suited to the Department of Defense's requirements for evidence integrity and provenance across jurisdictional boundaries. Digital forensics will also be more secure, and the investigation will be more transparent to those on the other side of the jurisdictional border, due to blockchain technology.**

Keywords— Digital forensic, Blockchain, SHA algorithm, AES encryption, Forensic evidence chain.

I. INTRODUCTION

Blockchain is the newest and potentially most affordable means to access cloud storage, as many small companies contribute to cloud storage through processing power and storage space. The crypto-hash, timestamps and transaction data of the preceding block are contained in each block. Blockchain is a growing series of documents called blocks that are connected together via encryption. The data of users is broken up into small chunks by blockchain cloud storage solutions. Then build another security layer and scatter around the network. This is feasible using blockchain characteristics like hashing, private / public encryption and transaction data (ledgers). Another advantage is that because the node only store the needed information that is required, does not store too much excessive information. The users are only given a portion of the data, so all sensitive information is kept safe and secure. Redundancy of data and load balancing measures are employed to ensure fast access and high availability. Digital forensics is a process of in-depth analysis of digital devices and data within a legal context, such as a criminal investigation or civil enquiry [8]. Crime scene refers to the field where criminals commit crimes, generate and leave crime traces, and evidences within a certain range of time and space. Crime scene analysis is the starting point of crime investigation, thus it is of vital importance to the success of cases solving [5].

In the digital forensics investigation practices, there are over hundreds of digital forensics investigation procedures developed all over the world. Each organization tends to develop its own procedures. Some focused on the technology aspects in data acquisition, some focused on data analysis portion of the investigation [3]. Our DF will assist in overcoming many challenges, such as hacking, and will improve the security of evidence collected in the course of a criminal investigation. The evidence is stored in a secure manner, ensuring that no third person have access to the evidence. This technology will increase the system's security by twofold.

II. LITERATURE SURVEY

This paper conducted preliminary forensic research on the blockchain-based forensic investigation framework by considering the diversity of devices, evidence items, data formats, and more in the complicated IoT environment [2]. The author of this paper stated that cybercriminals steal information from IoT devices and other sources, and that in order to prevent this, they used blockchain technology to create a framework for storing data and preventing it through various methodologies. They only mentioned preventing cybercriminals in the IoT, whereas we propose a framework that will aid in the investigation of crimes such as murder.

New tools are necessary to improve IoT forensics, especially because antiforensic techniques will continue to become increasingly sophisticated [6]. Because the number of cybercriminals is increasing every day, the authors claim that new forensic tools that can handle IoT data are needed. So we're working on a framework that's not for IoT devices, but for assisting in criminal investigations and adding another layer of security. This framework is its coverage. Through this framework, different standards and procedures could be linked together in a more holistic way. Digital forensics investigation is no longer viewed from pure technical aspects. Business, system and legal aspects are incorporated [3]. The authors have created a comprehensive digital forensic framework that is both business and legal in nature. As a result, I created a completely crime-oriented framework that will aid in the investigation of crimes.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

III. PROPOSED SYSTEM

This criminal digital forensics will give double security for digital evidences by utilizing hash functions such as SHA-256 and the Advanced Encryption Standard, also known by its original name Rijndael, to generate a hash and successfully form a verifiable evidence chain. Details about evidence identification, preservation, analysis, and presentation will be documented in block chains. The DFIF will assist in investigating and updating the investigation status more accurately. This project will assist in determining how we can provide or improve the security of the investigative framework.

By incorporating forensic investigation and additional information, Blockchain Technology can overcome the tampering, hacking and make the investigation more accurate and secure and make data acquisition and validation more accurate and informative. The forensic investigation uses Blockchain to build a close-loop system that provides significant forensic analysis benefits in a timely and cost-effective manner. We are also giving double encryption by using the strong algorithms like SHA-256 and AES algorithm and store everything in the system which uses blockchain technology which will make even secure the storing system, because blockchain have a functionality like immutability. Every information will be stored in the form of blockchain's block and everything is secured and these will make hard for anyone who try to manipulate the data, how much time the data is being altered or tampered in the application the original data can be easily recovered.

A. *System Architecture*

In this architecture, the application manager or admin will add important actors to the application, such as police officers with their respective police station name, higher officers with their roles, forensic staff, and medical staff, as well as their details. The user IDs and passwords will be generated programmatically, and those IDs and passwords will be sent to the actors via email, which they will receive after the application manager has added them to the application. For mail purposes, we will use the SMTP protocol, and they will be able to change their passwords if necessary, which will be saved in the database.

A police officer will file a First Information Report (FIR) and update the information in the application if a crime (murder) occurs within their jurisdictional border. The forensic team will gather all of the information and use it to create a report based on their findings. The medical staff will review the information gathered by the forensic team, and if anything unusual occurs, such as a murder, they will generate a report on the death hour, blood splatter, and fingerprints. This information will be saved, and a report will be generated in the application. The higher officers will monitor the crime investigation and can also see the medical and forensic report.
.

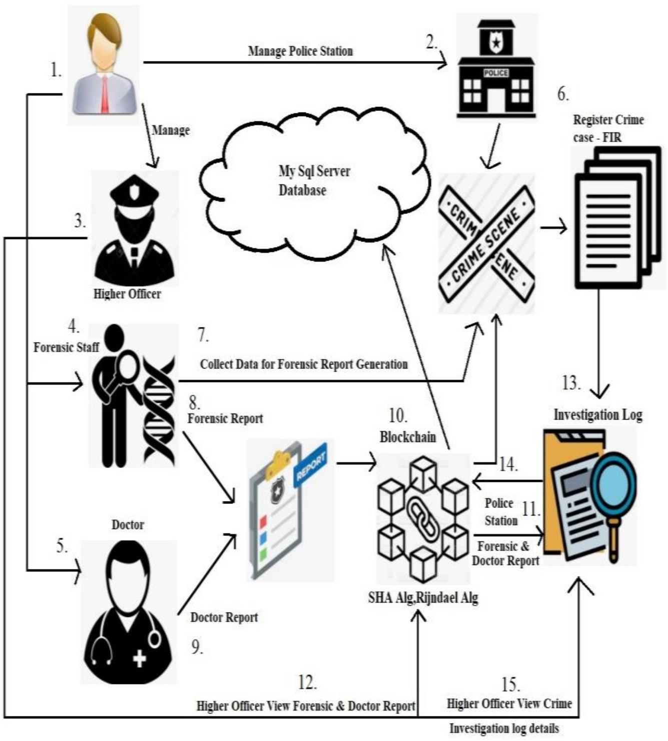
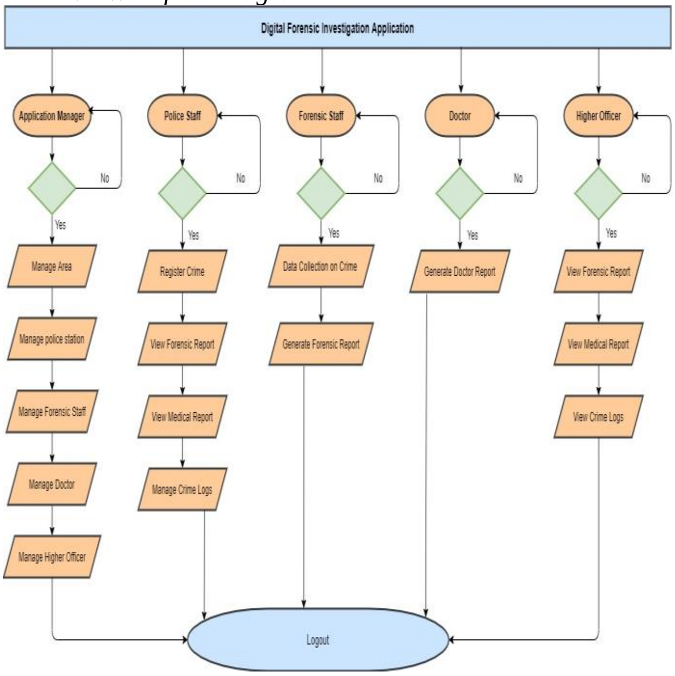


Fig 1. System Architecture

I. IMPLEMENTATION

A. *Workflow Diagram*



The FIR information from the police station's, the forensic team's report, and the doctor's report will all be updated in the application, and a report will be generated. In the programme, these data will be stored in hash format. The data will be converted to a hash using the SHA-256 algorithm, and the hash will be passed to the AES encryption algorithm, as well as a few other entities. The AES encryption algorithm will run for ten rounds before generating a 128-bit key. This key will be stored in the form of a blockchain database.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

B. AES Algorithm:

AES is symmetric key cryptographic algorithm published by MST. The algorithm was proposed by Rijndael. It is also known as Rijndael encryption algorithm. AES is a DES replacement. The block cypher technique is used by AES. Plain text and cypher text must be the same size. The size of an input key is the same as plain text.

AES supports data lengths of 128, 192, and 25 bits, as well as three different key lengths of 128, 192, and 256 bits. AES is made up of several rounds of processing different key bits, such as, Encryption is performed 10 times with a 128-bit key. The 192-bit key performs 12 encryption rounds. The 256-hits key encrypts data for 14 rounds.
AES algorithm Steps:

1. Byte Substitution (Sub Bytes): The 16 input bytes are substituted by looking up a fixed table(S-box). The result is in a matrix of four rows and four columns
2. Shift rows: Each of the four rows of the matrix is shifted to the left.
3. Mix Columns: Each column of four bytes is now transformed using a special mathematical function. This function takes as input the four bytes of one column and outputs four completely new bytes, which replace the original column. The result is another new matrix consisting of 16 new bytes. It should be noted that this step is not performed in the last round.
4. Add round key: The 16 bytes of the matrix are now considered as 128 bits and are XORed to the 128 bits of the round key. If this is the last round then the output is the ciphertext.

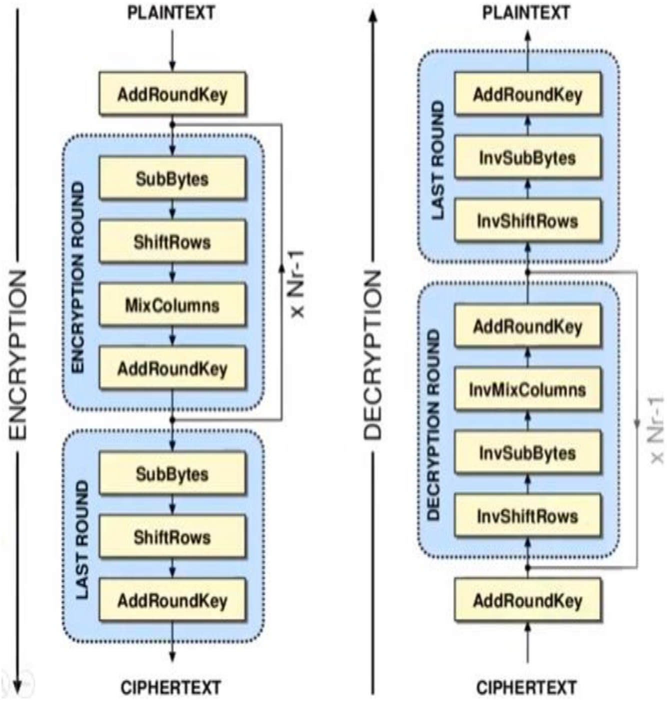


Fig 2. AES algorithm encryption and decryption

C. SHA-256 Algorithm

Working of SHA-256

- Arbitrary length message string 'M'.
- Convert it to binary.
- Pre-processing Stage:
  a. Padding a message M
  b. Parsing the Message M
  c. Setting initial hash values H_0……H_7
- Hash Computation
  a. Prepare the Message Schedule
  b. Initialize a, b, c, d, e, f, g and h
  c. Compute for the intermediate hash values
- Append Hash values $H_{0,}$ $H_1,$ $H_{2,}$ $H_{3,}$ $H_{4,}$ $H_{5,}$ $H_{6,}$ H_7
- 256-bits Message digest.

Steps in Encryption

1. Shift right: sha256 works on bits once and zeros all data can move represents in bits and it operates on chunks of 32 bits which we refer to as a word stated a word of 2 bits. Shift right it shifts the bits over a number of position that’s a 32 right a 32 bit shift.
2. Right rotation or rotation right: similar to aright shift instead of shift in the bits off until it never come back again, it moves them over onto the left-hand side to create this rotational bitwise operation or circular right shift.
3. Exclusive-or: it produces one if one of the two bits you’ve been put into it. It gives balanced representation of the bits.
4. Padding: Before stating an hash function we need to do padding because hash function likes to hash data in chunks of 512 bits at a time, So this is only 24 bits of data we need to pad it out with zeros basically to make it 512 bits. We need to append 1 at the end of the binary string 010000010100001001000011 → 1

Step 1: Let take a message string that is small letters M = "abc"
Step2: This string of small letters abc is converted to binary shown below:
a – 01100001
b = 01100010
c = 01100011
The new message M represented in bits is shown with 1 which is the length of the message is 24 bits.
M = 01100001 01100010 01100011
Step 3: Now the pre-processing stage starts with padding the message and before the hash computation begins. This process ensures that the padded message is multiple of 512 bits in preparation to parsing the message to 512 bits.

- Given the message size of 1 of 24 bits, a bit '1' is appended at the end of the message M = 01100001 0110001001100011 → 1
- The length of the padded message is now 25.
- Next is to append zeroes to get the total number of 0 bits. We use the equation:
  1+1+k — 448 mod 512
- This step ensures that the message is a multiple of 512 bits.
  As 1 = 24 bits
  k = 448 - (24+1) = 423 zero bits.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

K is 423 bits so that's why 423 bits are appended to the message lengths

- The length of the padded message is now 448.
  01100001 01100010 01100011 10000000 00000000
  00000000 00000000 00000000 00000000 00000000
  00000000 00000000 00000000 00000000 00000000
  00000000 00000000 00000000 00000000 00000000
  00000000 00000000 00000000 00000000 00000000
  00000000 00000000 00000000 00000000 00000000
  00000000 00000000 00000000 00000000 00000000
  00000000 00000000 00000000 00000000 00000000
  00000000 00000000 00000000 00000000 00000000
  00000000 00000000 00000000 00000000 00000000
  00000000 00000000 00000000 00000000 00000000
  00000000
- The remaining 54 bits represent the value of the original length of the message which is 24.
- I - 24 bits
  24 in binary is 11000 so we appending the this 24 to the end of the message of
  I- 00000000 00000000 00000000 00000000 00000000
  00000000 00000000 00011000
- The 64 bit representation of integer 24 in binary and 59 zeroes that is 11000 and is append to the padded message
  01100001 01100010 01100011 10000000 00000000
  00000000 00000000 00000000 00000000 00000000
  00000000 00000000 00000000 00000000 00000000
  00000000 00000000 00000000 00000000 00000000
  00000000 00000000 00000000 00000000 00000000
  00000000 00000000 00000000 00000000 00000000
  00000000 00000000 00000000 00000000 00000000
  00000000 00000000 00000000 00000000 00000000
  00000000 00000000 00000000 00000000 00000000
  00000000 00000000 00000000 00000000 00000000
  00000000 00000000 00000000 00000000 00000000
  00000000 00000000 00000000 00000000 00000000
  00000000 00000000 00000000 00000000 00000000
  00000000 00000000 00000000 00011000
  - The length of the padded message is now 512 bits.
  - M = 01100001 (8 bits) 01100010 (8 bits) 01100011 (8 bits) 1 (1 bit) 0 ...0000... (423 bits) 011000 (64 bits)
  - Converting this binary to hexadecimal: M = 61 62 63 80 00....00 18

Step 4: The padded message in hex representation is now parsed into blocks named M_1 to M_n In this example the padded message consists of one block. Since the 512 bit of an input block will represent a 16-31 bit words, the first 32 bits of the message block M_1 denoted by M_1 sub 0 to M_1 sub 15

Parsed Message $M^{(1)}$: = 61 62 63 80 00..00 18
$M_0^{(1)}$ = 61626380
$M_1^{(1)}$ = 00000000
.
.
.
$M_{14}^{(1)}$ = 00000000
$M_{15}^{(1)}$= 00000018

Step 5: The hash values H_o to H_7 are then initialized. The initial hash values for SHA-256 in hex representation arc as follows.
H_0 = 6a09e667
H_1 = bb67ae85
H_2 = 3c6ef372
H_3 = a54ff53a
H_4 = 510e527f
H_5 = 9b05688c
H_6 = 1f83d9ab
H_7 = 5be0cd19

Step 6: After the Pre-Processing stage is the absolute hash computation, the words of the padded message $M^{(1)}$ are then assigned to the words of the Message Schedule WT.

W_t is derived using the following equation:

$$W_t = ROTL(\sigma_1(W_{t-2}) + W_{t-7} + \sigma_0(W_{t-15}) + W_{t-16})$$

In this Example there is only one Message block , thus the message schedule is only from Wo to W15-

W_1 = 61626380	W_9 = 00000000
W_2 = 00000000	W_{10} = 00000000
W_3 = 00000000	W_{11} = 00000000
W_4 = 00000000	W_{12} = 00000000
W_5 = 00000000	W_{13} = 00000000
W_6 = 00000000	W_{14} = 00000000
W_7 = 00000000	W_{15} = 00000018

Step 7: The working part a to h is then initialized and is equivalent to the values Ho to H_7
a = H_0 = 6a09e667
b = H_1 = bb67ae85
c = H_2 = 3c6ef372
d = H_3 = a54ff53a
e = H_4 = 510e527f
f = H_5 = 9b05688c
g = H_6 = 1f83d9ab
h = H_7 = 5be0cd19

Step 8: To compute the final hash, we must run 64 iterations of the equation below:
Using our 64 words (W_t) from the table above and the initial hash values (a to h).

```
$For\ t = 0\ to\ 63:$
{
$T_1 = h + \sum_1^{\{256\}}(e) + Ch(e,f,g) + K_t^{\{256\}} + W_t$
$T_2 = \sum_0^{\{256\}}(a) + Maj(a,b,c)$
$h = g$
$g = f$
$f = e$
$e = d + T_1$
$d = c$
$c = b$
$b = a$
$a = T_1 + T_2$
}
```

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

$$Ch(e,f,g) = (e \wedge f) \oplus (\neg e \wedge g)$$

$$Maj(a,b,c) = (a \wedge b) \oplus (a \wedge c) \oplus (b \wedge c)$$

$$\sum\nolimits_1^{\{256\}}(e) = ROTR^{6}(e) \oplus ROTR^{11}(e) \oplus ROTR^{25}(e)$$

$$\sum\nolimits_1^{\{256\}}(a) = ROTR^{2}(a) \oplus ROTR^{13}(a) \oplus ROTR^{22}(a)$$

Step 9: After 64 rounds, a new set of 1 to h variables were derived at t = 64 To compute for the intermediate hash values the final equation used. Finally, we can convert the intermediate hash values into hexadecimal form, giving us our parts for the final message digest:

Step 10: Append the final hash values

$$H_0^{(N)} \parallel H_1^{(N)} \parallel H_2^{(N)} \parallel H_3^{(N)} \parallel H_4^{(N)} \parallel H_5^{(N)} \parallel H_6^{(N)} \parallel H_7^{(N)}$$

Step 11: 256-bit message digest the message digest is "ba7816bf8f01cfea414140de5dae2223b00361a396177a9cb410ff61f20015ad".

CONCLUSION

Because forensic reports are an important part of crime investigation and evidence collection, we are securing them with the Rijndael algorithm, SHA algorithm, and Blockchain technology. The details of the crime forensic report are sent to the police department in a secure and authenticated manner, which will aid future criminal investigations. The goal of this study is to create a blockchain-based DF investigation framework environment that can enable forensic investigation with high authenticity, immutability, and traceability between evidential entitlements and examiners. The application is tamper proof, so that forensic report is highly secured using blockchain technology. We've used more secure algorithms in this application to make it more secure. The application will become even more secure and powerful as a result of the blockchain technology.

REFERENCES

[1] Zibin Zheng, Shaoan Xie, Hongning Dai, Xiangping Chen,and Huaimin Wang "An Overview of Blockchain Technology:Architecture, Consensus, and Future Trends" (2017): 978-1-5386-1996-4

[2] Shancang Li, Tao Qin, and Geyong Min Blockchain-Based "Digital Forensics Investigation Framework in the Internet of Things and Social Systems." Computational Social Systems (2019): 2373-7476

[3] RICCI S.C.IEONG "FORZA – Digital Forensics Investigation Framework That Incorporate Legal Issues." Digital investigation(2006).

[4] Giannis Tziakouris "Cryptocurrencies— A Forensic Challenge or Opportunity for Law Enforcement? An INTERPOL Perspective." Security & Privacy (2018): 1558-4046

[5] Shuai Wang , Xiao Wang , Peijun Ye , Yong Yuan , Shuo Liu , and Fei-Yue Wang "Parallel Crime Scene Analysis Based on ACP Approach." Computational Social Systems (2018): 2373-7476

[6] Luca Caviglione, Steffen Wendzel, Wojciech Mazurczyk "The Future of Digital Forensics: Challenges and the Road Ahead." Security & Privacy (2017): 1558-4046

[7] Deqing Zou, Jian Zhao, Weiming Li, Yueming Wu , Weizhong Qiang , Hai Jin , Ye Wu , Yifei Yang "A Multi-Granularity Forensics and Analysis Method on Privacy Leakage in Cloud Environment" Internet of Things Journal (April 2019): 2372-2541

[8] Darren Quick, Kim-Kwang Raymond Choo "IoT Device Forensics and Data Reduction." 2169-3536

[9] Shady Mohamed Soliman, Baher Magdy; Mohamed A. Abd El Ghany "Efficient implementation of the AES algorithm for security applications" (2016) 2164-1706

[10] Wanzhong Sun, Hongpeng Guo, Huilei He, Zibin Dai "Design and Optimized Implementation of the SHA-2(256, 384, 512) Hash Algorithms." (2007): 2162-755X.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)