HASIL CEK_Sunardi , Ridho Surya Kusuma

by Sunardi , Ridho Surya Kusuma Digital Evidence Security System Design Using Bloc

Submission date: 25-May-2023 09:18AM (UTC+0700)
Submission ID: 2101271507
File name: Evidence_Security_System_Design_Using_Blockchain_Technology.pdf (1.23M)
Word count: 4856
Character count: 27385

International Journal of Safety and Security Engineering
Vol. 13, No. 1, February, 2023, pp. 159-165
Journal homepage: http://iieta.org/journals/ijsse

Digital Evidence Security System Design Using Blockchain Technology

Sunardi[1], Ridho Surya Kusuma[2*]

[1] Department of Electrical Engineering, Universitas Ahmad Dahlan, Yogyakarta 55164, Indonesia
[2] Department of Informatics, Universitas Siber Muhammadiyah, Yogyakarta 55253, Indonesia

Corresponding Author Email: ridhosuryakusuma@sibermu.ac.id

https://doi.org/10.18280/ijsse.130118

Received: 2 December 2022
Accepted: 8 January 2023

Keywords:
chain of custody, blockchain, forensic, k-means, digital evidence

ABSTRACT

Digital evidence plays an essential role in meeting the forensic need to uncover cybercrime and search for trace information of perpetrators. Digital evidence is vulnerable to system changes, human error, theft, deletion, and data manipulation, requiring security efforts to maintain authenticity. This study offers optimization of the chain of custody systems to maintain digital evidence integrity using authentication applications connected to the website server database. The design of the chain of custody system uses blockchain technology and K-means clustering algorithm. This research process consists of two stages. The first stage is the prototype of blockchain-based user access authentication applications. The second stage is the implementation of K-means clustering to determine the place of data storage according to its classification. The results of this study are the maximum security for blockchain-based chain of custody with the efficiency value of this application of 94.73% and the system load value of 0.223%. The total cost of deploying the application is 0.026702786 ETH. Based on this research can help to secure digital evidence information.

1. INTRODUCTION

The security of computer network infrastructure in companies, organizations, and governments has a unique role in maintaining the sustainability of information systems from cyber-attacks and minimizing losses [1]. The security of computer network infrastructure is crucial for protecting sensitive and confidential information, preventing unauthorized access, ensuring system availability and reliability, and mitigating the risk of cyber attacks.

One application of computer network security systems using sniffing or monitoring techniques allows administrators to monitor the data traffic and all activities and save the network monitoring results into a log file [2]. Logs are a source of information that can troubleshoot, record system breaches, attack activity, forensic needs, and investigations to find digital evidence according to standards [3].

Generally, investigators will secure and manage those results to a centralized local device after discovering digital evidence of a cyberattack because it is temporary [4]. However, digital evidence security systems are often overlooked and have the possibility of vulnerabilities so that attackers can exploit the system to modify and even delete data [5]. Some of the solutions in previous research, namely proposing a security framework using blockchain (Block-Def) with a loose clutch structure, multi-signature system, and mechanisms Consensus practical byzantine fault tolerance (PBFT) so that the storage of digital evidence information is stored separately, has validity, and integrity [5].

The study entitled digital forensic approaches for the Amazon Alexa ecosystem was proposes a new approach in cloud-native forensics with client-side forensics in support of practical investigations and using cloud-based digital evidence storage [6]. Based on this, the next challenge is obtaining, managing, and ensuring digital evidence storage [7]. This study aims to overcome the problem of security and management of digital evidence.

The problem in this study offers a systems approach to digital proof security with the combination of blockchain technology and machine learning. This research aims to secure digital evidence using blockchain technology and K-means clustering algorithms. The blockchain serves as a user authentication security system in managing and storing data [8]. Further application of one of the unsupervised machine learning technologies is the K-means clustering algorithm that serves as a digital proof routing system according to priority clusters [9, 10]. The result of this study is a prototype blockchain-based security application. The prototype efficiency value in this study was 94.73%, with ten users, 50 processes, and 90 records in the database. The results of calculating the system load are 0.223% with the number BCEntries 0.066592786ETH and TotalEntries 30 user requests.

Based on previous exposure, the development of digital evidence security systems is a challenge in digital forensic activities that include management, validity, integrity, originality, and storage scalability. The research seeks a new approach by building a prototype security system as a digital proof management solution based on blockchain technology and K-means clustering algorithms. The following is the structure of this research as a whole, which consists of an introductory section; material and methods contain information on the proposed approach; the result contains the application of the system based on the proposed approach; discussion describing the advantages, disadvantages, and implementation of the system; and conclusions explain the actual contribution of this research.

2. MATERIAL AND METHODS

This section describes the basic theory that supports research into realizing digital evidence security using blockchain and machine learning.

2.1 Blockchain

Blockchain is a distributed network between computers or nodes that allows all nodes to share responsibility for managing the network [11] so that every application or code that runs on the blockchain must have a smart contract [12]. A smart contract is the application of a program/protocol to carry out the functions needed in various blockchain technologies that allow processing transactions or interactions without the involvement of third parties [13, 14]. The process of running a smart contract requires resources on the network, namely user accounts, Ether (currency in Ethereum), and gas as a unit of value for ether payments so that applications can run on the Ethereum network [15].

2.2 K-means clustering

K-means clustering is an effective and efficient algorithm in grouping data [16, 17]. Clustering means the process of managing data into several groups or clusters. This algorithm is one of the branches of unsupervised machine learning that does not require training in its application [18] and group management based on the similarity of unique features in data [19]. This study uses the k-means algorithm to group digital evidence partitions based on size, time, and without user intervention.

The purpose of the K-means algorithm is to find clusters in the given input data. The process of determining the value of K using the Elbow technique with Eq. (1).

$$WSS = \sum_{i=1}^{m} (x_i - c_i)^2 \quad (1)$$

After obtaining the value of K through Eq. (1), the system will randomly assign several centroids and measure the distance of each data point from the centroid. The centroid is the arithmetic mean of an object's shape from all points in an object. Referring to Eq. (1), within the sum of squares (WSS) is the sum of the squared distances between each cluster member and its center. x_i is the data point and xc_i is the closest point to the centroid.

2.3 Proposed approach

This study proposes a security system in managing digital evidence using a combination of blockchain technology and machine learning. Blockchain acts as a user authentication system to access the digital evidence. At the same time, machine learning function is to grouping the digital evidence into a separate cluster. The following is a flowchart of the system flow in this study, as shown in Figure 1.

Figure 1 provides information on the flow of digital evidence security in this study which consists of six stages. The first stage is logging into the Ethereum blockchain platform. The second stage is logging in to the decentralized application (DAPP) [13]. The third stage is entering data descriptions and uploading digital evidence files. The fourth stage is managing files and categorizing them into specific clusters. The fifth stage determines the address and sends the file according to the cluster to the database. The sixth stage is recording and allocating each digital evidence in the repository system.

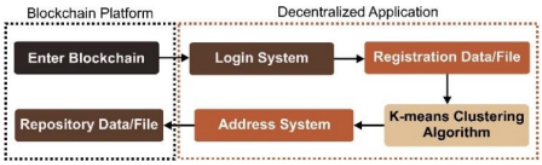

Figure 1. Flowchart of securing digital evidence in a blockchain network

3. RESULT

This section describes the results of designing a digital security application through an authentication system for blockchain-based user access rights and implementing the K-means algorithm to determine data according to the group.

3.1 Deploy DAPP on blockchain

This study uses a DAPP based on the Ethereum blockchain platform to secure digital evidence. Some supporting software, namely ganache, acts as a local environment for blockchain databases, and metamask stores ether balances or digital wallets for transaction activities—the following results of deploying the application to the Ethereum blockchain environment, as shown in Figure 2.

Figure 2. The result of deploying the system into the Ethereum blockchain platform

Figure 2 provides information on the total payment for smart contracts for an application that can run on the blockchain system of 0.0161408 ETH and displays the status of applications running at the URLs http://localhost:3000 and http://192.168.100.12:3000.

3.2 Login system

This study's login and authentication system are essential as the main gate of system security to manage and access applications. The development of this system uses solidity programming and runs on the Ethereum block with smart contracts. The following are the results of the login system in this study, as shown in Figure 3.

Figure 3 describes the flow of a login system using a blockchain which consists of six stages. The first stage is creating a simple user account by filling in the user ID, username, and password requirements. The second stage is to make a payment transaction in Ether in 0.003989 to input user

data into the application through a metamask account. The third stage presents a record of transaction information which includes account addresses and the amount of gas in ganache, ganache's role is as a blockchain database. The fourth stage is the login process using the previous user data. The fifth stage is verification between user data and metamask accounts. If the user and account data do not match, then the login will not be successful. The sixth stage displays a successful login process using a user with the name Zakila and the address of the metamask account that matches with the blockchain.

3.3 Registration data

This section sends and saves the information data digital evidence into the system by filling in some related information. This process becomes a special requirement for managing and determining digital evidence's cluster and identity, as shown in Figure 4.

Figure 4 shows a data registration form that functions as identity and information related to digital evidence before entering the blockchain system. The form consists of User ID, Full Name, Time and Date, Scope Network, Evidence Type, Size Data, and Digital Evidence. User ID and Full Name function to determine the identity of the user or investigator who accesses the system and enters data. Time and Date provides information on the time of securing digital evidence by users. Scope Network is a description of the origin or place of investigation and file acquisition. Evidence type serves to determine and classify file types. Label Size Data and Digital Evidence play an essential role in the authentication of file information. If the file size does not match, the user cannot process the data and submit it to the blockchain system. This study's registration data input scenario uses ten blocks with different accounts, and each account enters two digital evidence data. The following is the display of the registration data as shown in Table 1.

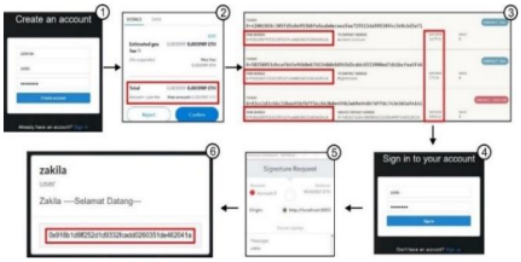

Figure 3. The flow of login system using blockchain

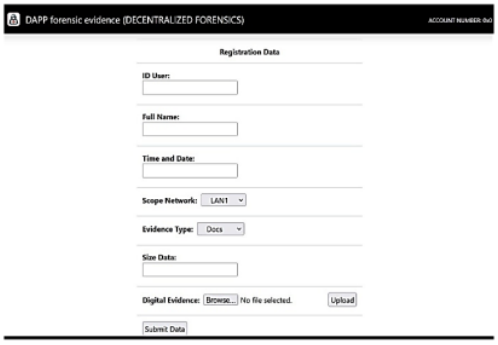

Figure 4. Display of digital evidence data registration form

Table 1. Digital Proof dataset on Ethereum block

Block	Address User	Date	Scope Type	Size Data1	Size Data2
0	0x918b1D6fF252d1D9332Fcadd0260351dE462041A	25/05/2021	LAN1PCAP	833 MB	791 MB
1	0x64174A8Fe32f29718daf568CdF4b3313c0E09376	26/05/2021	LAN2PCAP	392 MB	925 MB
2	0xaF8A092071D8eD48FF92432915C98b112F90c650	27/05/2021	LAN1PCAP	747 MB	445 MB
3	0x78e25C5a0245a7076a788b3B435d7E8247a19d41	28/05/2021	LAN4PCAP	405 MB	217 MB
4	0xc56B44717adAaA098fea71044883Bf3abcc00832	29/05/2021	LAN3PCAP	159 MB	624 MB
5	0xe847f0a3259497ae4bb0A1E25A4D488c04Dd04fE	30/05/2021	LAN1PCAP	481 MB	109 MB
6	0xdc47e6F51a23dc191DC3b76a52612a07Ec3968C3	31/05/2021	LAN4PCAP	553 MB	468 MB
7	0x081563e0b234b3cCF29008b8bff56AE674a5a691	01/06/2021	LAN3PCAP	393 MB	331 MB
8	0x05ff0C406Fd61b597e9f83D19A6E593c38cf37ba	02/06/2021	LAN1PCAP	321 MB	775 MB
9	0xE6f22b41A4077da945A6494899be9C9209c7900B	03/06/2021	LAN2PCAP	785 MB	193 MB

Table 1 provides information on registration data consisting of 10 blocks, Address User, Date, Scope, Type, Size_Data1, and Size_Data2. The data management resides in the Ethereum blockchain network environment and packages user information via User Address. Furthermore, this research will manage the data in the Size_Data1 and Size_Data2 columns of type PCAP into specific clusters. This grouping serves to protect and support investigators in the forensic data search process.

3.4 K-Means clustering algorithm

This section describes the concept of using the K-means algorithm in processing any data information that enters the system and determines the appropriate cluster. The following is the pseudo-code for Algorithm1 in Figure 5.

```
const k = 3
let preventCallBack = false
async function getData (){
    if(!preventCallBack){
        preventCallBack = true
        let clusterCenterLocs = ['ridho']
        let clusterCenterNewLocs = ['surya kusuma']
        while(JSON.stringify(clusterCenterLocs)!=
        JSON.stringify(clusterCenterNewLocs)){
        clusterCenterLocs = []
        for(let i=0; i<k; i++){
        clusterCenterLocs.push(chart.data.datasets[0].data[i])
        await labelDataPoints()
        await recenterClusterCenters()
for(let i=0; i<k; i++){
clusterCenterNewLocs.push(chart.data.datasets[0].data[i])}
function getCluster(){
    let clusters = []
    for(let i=0; i<k; i++){
        clusters.push([])}
if(k){getData();}
function recenterClusterCenters(){
    return new Promise(async(resolve,reject)=>{
        const clusters = getCluster()
        clusters.forEach((cluster,i)=>{
            let sumX = 0
            let sumY = 0
            cluster.forEach((dataPoint,j)=>{
            sumX = sumX + dataPoint.x
            sumY = sumY + dataPoint.y})
            if(cluster.length){
        chart.data.datasets[0].data[i] = {x:Number((sumX/cluster.length).
```

Figure 5. Algorithm 1 calculation k-means on Javascript

```
pragma solidity >= 0.7.0 < 0.9.0;
contract dataStorage{
    mapping(address => digitalEvidence) eviPcap;
    struct digitalEvidence{
        uint user_id;
        string fullName;
        uint dateTime;
        string scopeN;
        uint sizeData;}
    function addPcap(address _addr, uint user_id, string memory fullName,
    uint dateTime, string memory scopeN, uint sizeData) public{
    eviPcap[_addr] = digitalEvidence(user_id, fullName, dateTime, scopeN, sizeData);}
    function getAddress(address _addr) public view returns(uint){
    return eviPcap[_addr].sizeData;}}
```

Figure 6. Algorithm 2 cluster k-means on solidity

Figure 5 shows the application of pseudo-code algorithm1 uses Javascript programming to calculate the K-means formula to group data. The algorithm is for the process of calculating the K-means formula so that it can group data. The following is the pseudo-code for algorithm2 in Figure 6.

Figure 6 shows Pseudo-code algorithm 2 uses Solidity programming, which accommodates, manages, and classifies data structures on the Ethereum blockchain. The algorithm represents the classification process with a mapping structure; the mapping structure stores many of information such as user id, full Name, Date Time, scope N, and size Data using the address key. The get Address function acts as code to verify data and valid users. If not a legitimate user, then the output value is zero. Use the add PCAP function to insert digital evidence data information into the digital Evidence structure.

3.5 Address system

Blockchain uses an address system to secure or package information on transaction activities, accounts, and contracts of a program in the form of hashing. The public can see the address but cannot determine the precise information, so only certain parties, such as application developers and users, can understand it. This section presents some smart contract address information while deploying a digital proof security application on the Ethereum blockchain network. The following are the results of the application address recording, as shown in Table 2.

Table 2. Smart contract DAPP forensic evidence

Deploy	Smart Contract	Gas	Addr. Input	ETH
Initial Migration	0x82BE84c99C9bB980Fc1148e8eE9F2b535b62275A	201843	0x60806...60033	0.004033686
Autentication	0x7eB3b17a1E41668B032C22186e6FD724D353611b	605211	0x60806...60033	0.01210422
Digital Evidence	0x14A919590E83B987aF5f7A3273Db70076A794CD0	376817	0x60806...60033	0.01056488

Table 3. Log of data processing activities into blockchain

Block Address	Data Address	Gas Cost	Trans. Hash	Contract Address	Information
0x918b1D6...462041A	0xf2c298b...0000	147976	0x420018...cbd5ef1	0x7eB3b17...353611b	Create Account1
0x78e25C5...7a19d41	0xf2c298b...0000	132976	0xc78ab7...629b5ed	0x7eB3b17...353611b	Create Account2
0x64174A8...0E093760	xf2c298b...0000	132976	0x6b3e7c…e88935b	0x7eB3b17...353611b	Create Account3
0xc56B447...cc00832	0xf2c298b...0000	132976	0xa97e4a…ddc05db	0x7eB3b17...353611b	Create Account4
0xe847f0a...4Dd04fE	0xf2c298b...0000	132976	0x582500...fea6f46	0x7eB3b17...353611b	Create Account5
0x918b1D6...462041A	0xca7e31...0000	136861	0x697645...052e7850	x14A9195...A794CD0	Input Data1
0x78e25C5...7a19d41	0xca7e31...0000	136897	0x8e054b...0f19c0a	0x14A9195...A794CD0	Input Data2
0x64174A8...0E09376	0xca7e31...0000	136885	0xffa2bdf...78abd4d	0x14A9195...A794CD0	Input Data3
0xc56B447...cc00832	0xca7e31...0000	136885	0xd746bc...6491974	0x14A9195...A794CD0	Input Data4
0xe847f0a...4Dd04fE	0xca7e31...0000	136885	0x9e5e9f…d6197680	x14A9195...A794CD0	Input Data5

Table 2 contains information on the three main functions: Initial Migration, Authentication, and Digital Evidence. The Initial Migration function is an application identity that will enter the Ethereum blockchain network. Authentication has a role in adding and login system users. Digital Evidence's role is to add and store data to the blockchain. Each function has a unique address. The smart contract manages this unique address by determining the cost of gas, address input, and the amount of ETH to run a function.

All function processes refer to the same input address and the ETH section provides information on the total price per function deployed, namely 0.00403 ETH; 0.0121 ETHs; 0.0105 ETH; and the total cost of deploying the application is 0.026702786 ETH. The addresses in the research are generated and deployed using the Ganache software, so the address record results will differ from other studies. Following are the results of the data input scenario activity log in the application, as shown in Table 3.

Table 3 provides an overview of using DAPP forensic evidence, including creating user accounts and inputting digital evidence data. The information in Table 3 includes a Block address indicating the source of internal user identity. Data Address is the packaging of user information to create an account on the application. Gas Cost provides a breakdown of the cost of running the application. Transaction Hash is a particular record in the form of hashing that records the transaction process on the blockchain; Contract Address has a role in accommodating and being the purpose of the data input process; and information as a description of a process.

3.6 Repository data

This section provides a visualization of data processing results into each cluster based on file size information. The following shows the results of grouping data into three clusters using the K-means algorithm, as shown in Figure 7.

Figure 7 shows the results of the digital proof data input process, which includes PCAP, DOCS, Image, and Txt data into the blockchain system. The use of digital evidence data in this study is random based on a case study of network forensics. The system in this study manages and divides the data into three clusters. The following are the results of the first cluster of PCAP data, as shown in Table 4.

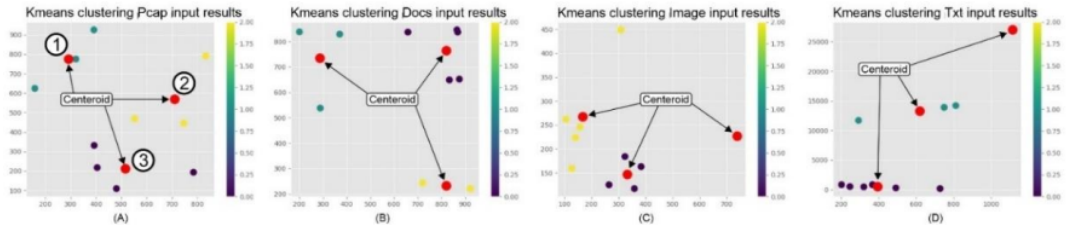

Figure 7. Visualization of digital evidence on data: (A) PCAP, (B) DOCS, (C) Image, and (D) Txt

Table 4. The first cluster of the Pcap digital proof dataset on the Ethereum block

Block	Address User	Type	Size Data1	Size Data2
0	0x918b1D6fF252d1D9332Fcadd0260351dE462041A	PCAP	833 MB	791 MB
2	0xaF8A092071D8eD48FF92432915C98b112F90c650	PCAP	747 MB	445 MB
6	0xdc47e6F51a23dc191DC3b76a52612a07Ec3968C3	PCAP	553 MB	468 MB

First Cluster of PCAP Digital Proof Dataset on Ethereum Block

Table 5. Second cluster of digital proof dataset Pcap on Ethereum block

Block	Address User	Type	Size Data1	Size Data2
1	0x64174A8Fe32f29718daf568CdF4b3313c0E09376	PCAP	392 MB	925 MB
4	0xc56B44717adAaA098fea71044883Bf3abcc00832	PCAP	159 MB	624 MB
8	0x05ff0C406Fd61b597e9f83D19A6E593c38cf37ba	PCAP	321 MB	775 MB

Second Cluster of PCAP Digital Proof Dataset on Ethereum Block

Table 6. The third cluster of the Pcap digital proof dataset on the Ethereum block

Block	Address User	Type	Size Data1	Size Data2
3	0x78e25C5a0245a7076a788b 3B435d7E8247a19d41	PCAP	405 MB	217 MB
5	0xe847f0a3259497ae4bb0A1E25A4D488c04Dd04fE	PCAP	481 MB	109 MB
7	0x081563e0b234b3cCF29008b8bff56AE674a5a691	PCAP	393 MB	331 MB
9	0xE6f22b41A4077da945A6494899be9C9209c7900B	PCAP	785 MB	193 MB

Third Cluster of PCAP Digital Proof Dataset on Ethereum Block

Table 4 provides data information in the first cluster, namely data transmission from blocks zero, two, and six. Each block has a unique address and has two data with different sizes. The address is the identity of the data owner or data source. The data size of each data becomes a K-means Clustering parameter in grouping and determining the centroid value. PCAP data sizes in this first cluster are 833 MB, 791 MB, 747 MB, 445 MB, 553 MB, and 468 MB. Based on this, the centroid values are X: 290.67 and Y: 774.67, following Figure 4 point 1 (A). The second cluster of PCAP data are shown in Table 5.

Table 5 displays information on data belonging to the second cluster, which consists of blocks one, four, and eight. PCAP data sizes in this second cluster are 392 MB, 925 MB, 159 MB, 624 MB, 321 MB, and 775 MB. Based on these sizes, the centroid values are X: 516 and Y: 212.5, following Figure 4 point 1 (A). The display of three PCAP data clusters is shown in Table 6.

Table 6 provides data information that belongs to the third cluster, which consists of blocks three, five, seven, and nine. PCAP data sizes in this third cluster are 405 MB, 217 MB, 481 MB, 109 MB, 393 MB, 331 MB, 785 MB, and 193 MB. Based on these sizes, the centroid values are X: 711 and Y: 568, which follow Figure 4 point 1 (A). One of the benefits of implementing the K-means clustering algorithm in this study is intending to facilitate the search system and managing digital evidence data records based on the cluster.

4. DISCUSSION

Data management in this study focuses on calculating K-means in Javascript rather than Solidity because the data type structure only consists of integers, strings, and Booleans. Solidity is not general-purpose programming; programming must be as simple as possible to avoid a complicated and lengthy calculation process. The more complicated and lengthier a process is, the higher the price of deploying smart contracts and gas for each order. The following describes the use of DAPP forensic evidence in this study, as shown in Figure 8.

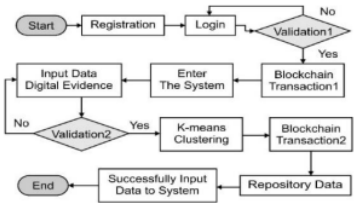

Figure 8. Flowchart design DAPP on this research

Figure 8 provides information on the flow of use of the Digital Forensic DAPP application, which consists of the

registration process, login, user validation, first blockchain transaction, enter the system, data input, data validation, K-means algorithm calculation, second blockchain transaction, data repository, and verification. Data entered the system successfully. Based on the flowchart, the following system efficiency values use Eq. (2) [20].

$$E = 100 - \left(\frac{U_{BC} * NoS}{\sqrt{ReDB}}\right) \quad (2)$$

Eq. (2) consists of E as a symbolic representation for efficiency, UBC, which is the number of internal users on the application system, NoS means the number of services available for users, and ReDB represents the number of data records that enter the blockchain system. Based on the formula's calculation, the system efficiency value in this study is 94.73%, with 10 users, 50 processes, and 90 records in the database. The following is a mathematical calculation to determine the system load percentage on this DAPP using Eq. (3).

$$Ldr = \frac{BCEntries}{TotalEntries} * 100\% \quad (3)$$

Eq. (3) consists of Ldr, which represents the system load on application performance, BC Entries is the number of Ether, and Total Entries is the total number of system entries created by the user. The result of the calculation of the load system is 0.223%, with the number of BC Entries 0.066592786ETH and Total Entries 30 user requests.

4.1 System advantage

The advantage of the system in research is that it attempts to combine blockchain network technology and unsupervised machine learning to secure digital evidence data—blockchain technology functions as an environment for securing application access by authorized users and digital evidence data information. Machine learning acts as a mathematical calculation in managing and grouping digital data based on size to make it easier for investigators to search for data according to forensic needs.

4.2 System disadvantage

This section discusses the system's shortcomings in this study, namely the application of user authentication features and digital proof data input using separate smart contracts, thereby increasing the cost of deploying applications. These shortcomings can be considered for further research to build both features into a unified whole to save the cost of deploying applications.

4.3 System implementation

This research has important implementation for developing security and protecting the authenticity of digital evidence from tampering and other threats. The results of this study make it easier for incident response teams or IT security to maintain the authenticity of digital evidence because blockchain technology allows digital evidence to be associated with unique and verified identities. Each digital token has a unique digital signature issued by the creator and recorded on the blockchain. Thus, the authenticity of evidence can be verified transparently and cannot be manipulated. This digital evidence security system is a maximum security innovation to protect the chain of custody based on blockchain technology and machine learning.

5. CONCLUSIONS

A digital evidence security system using blockchain technology can significantly enhance the security and reliability of digital evidence. This research focuses on developing digital evidence security into blockchain network technology through DAPP and unsupervised machine learning technology so that it can maintain the integrity and place digital evidence according to the cluster. This research includes the process of the valid user authentication system and the digital data input process with random data in the PCAP format. The efficiency value of using this application is 94.73% with ten users; a System load value of 0.223% with ten users and 30 requests; and the total cost of deploying the application is 0.026702786 ETH. Using the combination of the two technologies, investigators can obtain valid and reliable digital evidence, which can help increase trust and transparency in business and legal processes. The next direction for developing this application is the digital evidence search feature according to the cluster and forensic needs.

ACKNOWLEDGMENT

The authors would like to express appreciation and gratitude to Universitas Ahmad Dahlan and Universitas Siber Muhammadiyah for funding this research.

REFERENCES

[1] Toapanta, S.M.T., Ochoa, I.N.C., Sanchez, R.A.N., Mafla, L.E.G. (2019). Impact on administrative processes by cyberattacks in a public organization of Ecuador. In 2019 Third World Conference on Smart Trends in Systems Security and Sustainablity (WorldS4), London, UK, pp. 270-274. https://doi.org/10.1109/WorldS4.2019.8903967

[2] Umar, R., Riadi, I., Kusuma, R.S. (2021). Mitigating sodinokibi ransomware attack on cloud network using software-defined networking (SDN). International Journal of Safety and Security Engineering, 11(3): 239-246. https://doi.org/10.18280/ijsse.110304

[3] Kebande, V. R., Ray, I. (2016). A generic digital forensic investigation framework for internet of things (iot). In 2016 IEEE 4th International Conference on Future Internet of Things and Cloud (FiCloud) Vienna, Austria (pp. 356-362). IEEE. https://doi.org/10.1109/FiCloud.2016.57

[4] Al-Sharif, Z.A., Al-Saleh, M.I., Alawneh, L.M., Jararweh, Y.I., Gupta, B. (2020). Live forensics of software attacks on cyber–physical systems. Future Generation Computer Systems, 108: 1217-1229. https://doi.org/10.1016/j.future.2018.07.028

[5] Tian, Z., Li, M., Qiu, M., Sun, Y., Su, S. (2019). Block-DEF: A secure digital evidence framework using blockchain. Information Sciences, 491: 151-165. https://doi.org/10.1016/j.ins.2019.04.011

[6] Chung, H., Park, J., Lee, S. (2017). Digital forensic approaches for Amazon Alexa ecosystem. Digital investigation, 22: S15-S25. https://doi.org/10.1016/j.diin.2017.06.010

[7] Lillis, D., Becker, B., O'Sullivan, T., Scanlon, M. (2016). Current challenges and future research areas for digital forensic investigation. arXiv preprint arXiv:1604.03850. http://arxiv.org/abs/1604.03850

[8] Zheng, Y., Li, Y., Wang, Z., Deng, C., Luo, Y., Li, Y., Ding, J. (2019). Blockchain-based privacy protection unified identity authentication. In 2019 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC) Guilin, China, pp. 42-49. https://doi.org/10.1109/CyberC.2019.00017

[9] Arora, P., Varshney, S. (2016). Analysis of k-means and k-medoids algorithm for big data. Procedia Computer Science, 78: 507-512. https://doi.org/10.1016/j.procs.2016.02.095

[10] Yu, S.S., Chu, S.W., Wang, C.M., Chan, Y.K., Chang, T.C. (2018). Two improved k-means algorithms. Applied Soft Computing, 68: 747-755. https://doi.org/10.1016/j.asoc.2017.08.032

[11] Hasan, H.R., Salah, K., Jayaraman, R., et al. (2020). A blockchain-based approach for the creation of digital twins. IEEE Access, 8: 34113-34126. https://doi.org/10.1109/ACCESS.2020.2974810

[12] Alam, A., Rashid, S.Z.U., Salam, M.A., Islam, A. (2018). Towards blockchain-based e-voting system. In 2018 international conference on innovations in science, engineering and technology (ICISET), Chittagong, Bangladesh, pp. 351-354. https://doi.org/10.1109/ICISET.2018.8745613

[13] Yang, X., Chen, Y., Chen, X. (2019). Effective scheme against 51% attack on proof-of-work blockchain with history weighted information. In 2019 IEEE International Conference on Blockchain (Blockchain) Atlanta, GA, USA, pp. 261-265. https://doi.org/10.1109/Blockchain.2019.00041

[14] Shorman, S.M., Allaymounq, M., Hamid, O. (2019). Developing the E-commerce model a consumer to consumer using blockchain network technique. International Journal of Managing Information Technology (IJMIT), 11(2): 55-64. https://doi.org/10.5121/ijmit.2019.11204

[15] Bragagnolo, S., Rocha, H., Denker, M., Ducasse, S. (2018). Smart Inspect: solidity smart contract inspector. In 2018 International workshop on blockchain oriented software engineering (IWBOSE), Campobasso, Italy, pp. 9-18. https://doi.org/10.1109/IWBOSE.2018.8327566

[16] Abbas, S.A., Aslam, A., Rehman, A.U., Abbasi, W.A., Arif, S., Kazmi, S.Z.H. (2020). K-means and k-medoids: Cluster analysis on birth data collected in city Muzaffarabad, Kashmir. IEEE Access, 8: 151847-151855. https://doi.org/10.1109/ACCESS.2020.3014021.

[17] Govender, P., Sivakumar, V. (2020). Application of k-means and hierarchical clustering techniques for analysis of air pollution: A review (1980-2019). Atmospheric pollution research, 11(1): 40-56. https://doi.org/10.1016/j.apr.2019.09.009

[18] Sinaga, K.P., Yang, M.S. (2020). Unsupervised K-means clustering algorithm. IEEE Access, 8: 80716-80727. https://doi.org/10.1109/ACCESS.2020.2988796.

[19] Fard, M. M., Thonet, T., Gaussier, E. (2020). Deep k-means: Jointly clustering with k-means and learning representations. Pattern Recognition Letters, 138: 185-192. https://doi.org/10.1016/j.patrec.2020.07.028.

[20] Toapanta, S.M., Quimis, O.A.E., Gallegos, L.E.M., Arellano, M.R.M. (2020). Analysis for the evaluation and security management of a database in a public organization to mitigate cyber attacks. IEEE Access, 8: 169367-169384. https://doi.org/10.1109/ACCESS.2020.3022746

HASIL CEK_Sunardi , Ridho Surya Kusuma

ORIGINALITY REPORT

3%	3%	3%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	iieta.org Internet Source	1%
2	www.iieta.org Internet Source	1%
3	Segundo Moises Toapanta, Omar Alexander Escalante Quimis, Luis Enrique Mafla Gallegos, Ma Rocio Maciel Arellano. "Analysis for the Evaluation and Security Management of a Database in a Public Organization to Mitigate Cyber Attacks", IEEE Access, 2020 Publication	1%
4	www.mdpi.com Internet Source	1%
5	Rusydi Umar, Imam Riadi, Ridho Surya Kusuma. "Mitigating Sodinokibi Ransomware Attack on Cloud Network Using Software-Defined Networking (SDN)", International Journal of Safety and Security Engineering, 2021 Publication	1%

Exclude quotes	On	Exclude matches	< 1%
Exclude bibliography	On		