FGS Press

Journal of Sensors, IoT & Health Sciences
ISSN: 2584-2560
Vol.02(02), June 2024, pp.1-15

Research Article

IoT Forensic Cyber Activities Detection and Prevention with Automated Machine Learning Model

Ankush D. Sawarkar[1,*] and Anjali Deepak Hazari[2]

[1]Assistant Professor, Department of IT, SGGSIE&T, Nanded, Maharashtra,431606 India.

[2]Assistant Professor, Department of CSE, MGM College of Engineering and Technology, Nanded, Maharashtra, 431602, India.

[*]Corresponding Author: Ankush D. Sawarkar. Email: adsawarkar@sggs.ac.in

Received: 01/05/2024; Revised: 22/05/2024; Accepted: 22/06/2024; Published: 30/06/2024 .

DOI: https://doi.org/10.69996/ jsihs.2024006

Abstract: The Internet of Things (IoT) has been deployed in a vast range of applications with exponential increases in data size and complexity. Existing forensic techniques are not effective for the accuracy and detection rate of security issues in IoT forensics. Cyber forensic comprises huge volume constraints that are processing huge volumes of data in the Information and Communication Technology (ICT) comprised of IoT devices and platforms. Trust blockchain is effective technology those are utilized to assess the tamper-proof records in all transaction in the IoT environment. With the implementation of trust blockchain the record and transaction are processed with a distributed ledger that is managed by the network nodes. The challenge associated with the trust blockchain in IoT forensics is cost and security. To achieve significant cost-effectiveness organizations, need to evaluate the risks and benefits associated with IoT forensics in the trust blockchain technology. In this paper, developed a Block Chain Enabled Cyber-Physical system with distributed storage. The developed Blockchain model is termed as Integrated Hadoop Blockchain Forensic Machin Learning (IHBF-ML). The IHBF-ML model uses the Hadoop Distributed File System (HDFS) with cyberspace to improve security. Within the IHBF-ML model, IoT data communication is established with the smart contract. The smart contract-based blockchain process uses the Machine Learning model integrated with Cat Boost classification model for anomaly detection. Cost in IoT forensic is minimized with the parallel processing of the data through MapReduce Framework for the traffic translation, extraction, and analysis of the dynamic feature traffic from the IoT environment. The experimental analysis stated that constructed IHBF-ML model reduces the cost by ~25% than the other conventional blockchain Ethereum and EOS.

Keywords: Iot Forensic, Hadoop Distributed File System (HDFS), Smart Contracts, Mapreduce, Machine Learning, Ethereum

1.Introduction

Forensics is a set of activities applied to a possible entity that involves the collection, analysis, and presentation, to extract the evidence acceptable in the court of law. Network and Computer Forensics is a field, which deals with an in-depth analysis of the captured data, to reveal evidence, admissible in the court of law [1]. The audit, examination of Network and Computer data for information congregation, encroachment detection, and legal evidence presentation is an integral part of the forensic process [2]. Computer and Network are forensic is a scientific process to pinpoint, seize, extract, analyse, interpret, examine, document, and present evidence involved in cybercrime. Generalized Cyber forensic consists of three broad phases (a) collecting the

evidence items (b) examining the evidence and (c) handling the evidence. Each of these phases includes the sub-phases that may be required as a part of the particular case. Collecting phase involves the elicitation, preserving and identifying seized items [3]. Examining phase involves analyzing, interpreting, and validating the evidence items. The handling phase consists of documenting and presenting evidence in an admissible form for the court of law. IoT (Internet of Things) forensics refers to the process of collecting, analyzing, and interpreting digital evidence related to Internet-connected devices. As the number of internet-connected devices continues to grow, the importance of IoT forensics is also increasing [4]. IoT forensics involves analysing the data and logs generated by these devices to identify potential security breaches, data breaches, or other suspicious activities. The data collected in IoT forensics can be used in legal proceedings or investigations related to cybercrime, data theft, or other security incidents. It requires a combination of technical knowledge, analytical skills, and legal expertise to carry out successful IoT forensics investigations. IoT forensics is a specialized area of digital forensics that deals with the collection, analysis, and interpretation of digital evidence from internet-connected devices [5]. With the increasing number of IoT devices in homes, offices, and public spaces, the potential for cyberattacks and security breaches is also increasing. As a result, there is a growing need for IoT forensics to identify, analyze, and prevent cybercrimes and other security incidents involving IoT devices.

The process of IoT forensics typically involves collecting and analyzing data from various sources, such as network traffic, device logs, and user accounts. The data collected from IoT devices can be complex and diverse and may include data from sensors, cameras, microphones, and other sensors that are part of the device [6]. The data can be analysed using specialized forensic tools and techniques to identify potential security breaches, data breaches, or other suspicious activities. IoT forensics investigations can be used in legal proceedings or investigations related to cybercrime, data theft, or other security incidents. In such cases, the data collected from IoT devices can be used as evidence to support the prosecution of cybercriminals or to determine liability in cases involving data breaches or other security incidents [7]. IoT forensics requires a combination of technical knowledge, analytical skills, and legal expertise to carry out successful investigations. Forensic investigators must have a deep understanding of the various types of IoT devices, their functionality, and the data they generate. They must also be familiar with the different communication protocols and standards used by IoT devices, as well as the security risks associated with these devices. Overall, IoT forensics is an important field that helps to prevent cybercrimes and other security incidents involving IoT devices [8]. By analyzing the data generated by these devices, forensic investigators can identify potential security breaches and help organizations to take corrective action to prevent future incidents. IoT forensics and trust blockchain can work together to improve security and accountability in IoT systems by providing a transparent, tamper-proof record of all transactions within the network. IoT forensics involves analyzing digital evidence from IoT devices to identify security breaches, data breaches, and other suspicious activities [9]. The data collected from these devices can be complex and diverse, and may include data from sensors, cameras, and other sensors that are part of the device. This data can be analysed using specialized forensic tools and techniques to identify potential vulnerabilities and take corrective action.

Trust blockchain, on the other hand, is a technology that can be used to provide a tamper-proof record of all transactions within an IoT network. In a trust blockchain system, transactions are recorded on a distributed ledger that is maintained by a network of nodes. Each node on the network verifies and records transactions, ensuring that they are accurate and tamper-proof [10]. This makes trust blockchain an ideal technology for securing IoT systems, as it can provide a transparent and tamper-proof record of all activity. By using IoT forensics in combination with trust blockchain, it is possible to create a secure and accountable IoT network. For example, if a security breach occurs in the network, the data collected through IoT forensics can be used to identify the source of the breach. This information can then be used to take corrective action, such as removing the compromised device from the network or updating the security protocols for the network. In addition, the transaction records on the trust blockchain can provide a tamper-proof record of all activity within the network, which can be used to identify suspicious activity and improve the overall security of the network [11]. Overall, the combination of IoT forensics and trust blockchain can help to improve the security and reliability of IoT systems, ensuring that they are better equipped to handle the growing number of connected devices and potential security threats. By providing a transparent and tamper-proof record of all activity within the network, IoT forensics and trust blockchain can help to create a secure and accountable IoT ecosystem. While IoT forensics and trust blockchain have the potential to improve the security and accountability of IoT systems, there are also several challenges associated with their use [12]. One of the main challenges with IoT forensics is the sheer volume of data generated by IoT devices. IoT devices generate vast amounts of data, which can be difficult to manage and analyze. This can make it challenging for investigators to identify relevant data and extract useful information from it. Additionally, the diversity of IoT devices and the lack of standardization in their data formats can make it difficult to develop effective forensic tools and techniques. Another challenge with IoT forensics is the need for specialized expertise [13]. Forensic investigators require a deep understanding of the various types of IoT devices, their functionality, and the data they generate. They must also be familiar with the different communication protocols and standards used by IoT devices, as well as the security risks associated with these devices. This requires specialized training and experience, which may be in short supply.

Trust blockchain also presents several challenges. One of the main challenges is scalability. As the number of transactions within a trust blockchain network grows, the size of the ledger increases, which can lead to performance issues. Additionally, the consensus mechanism used by trust blockchain can also impact scalability, as more nodes are required to reach consensus on each transaction [14]. Another challenge with trust blockchain is the potential for a 51% attack. In a trust blockchain network, a 51% attack occurs when a single entity controls more than 50% of the nodes on the network. This would allow the entity to manipulate the ledger and potentially commit fraud or other malicious activities. Cost-effectiveness is an important challenge in the adoption of IoT forensic and trust blockchain technologies. These technologies require significant investment in terms of both hardware and software infrastructure, as well as specialized expertise and training. In terms of IoT forensics, the cost of hardware and software infrastructure can be high. IoT devices generate vast amounts of data, which require specialized tools and storage solutions to manage and analyse. Additionally, forensic investigators require specialized training and experience, which can be expensive to obtain. Similarly, the cost of implementing trust blockchain can also be high [15].

The conventional technique is subjected to challenges of Cost and security in the IoT forensic hence; this paper constructed an Integrated Hadoop Blockchain Forensic Machine Learning (IHBF-ML) model for security and cost reduction in IoT forensic environment. The specific contribution of the IHBF-ML model is explained as follows:

1. Initially, the IHBF-ML model collects the data from the IoT nodes, and the data is stored in the database. With the stored data information is being processed and evaluated to retrieve the information from the node data.
2. The constructed model comprises the HDFS system integrated with cyberspace. Within the HDFS the node data are stored and processed in a parallel manner with the use of the MapReduce framework. The MapReduce framework model process the distributed files in the network with the conversion of data.
3. The MapReduce framework model processes the data in a parallel manner with the distributed file management with the smart contract-based Public blockchain model. MapReduce framework performs the translation, extraction, and analysis of the feature in the IoT nodes through a parallel process.
4. Smart contracts are implemented in Cyberspace with the Cyber-Physical System (CPS) with distributed storage. The security analysis is improved with the smart contract-based Machine Learning model integrated with the CatBoost model for anomaly detection in the network.
5. The experimental analysis expressed that the cost of the IHBF-ML system is reduced compared with the other blockchain model. Security analysis stated that anomaly detection rate increases effectively.

This paper is organized as follows: Section 2 presented the literature review related to the blockchain model with the security analysis. Section 3 explained about the proposed IHBF-ML model for the cost reduction and security improvement in IoT forensic model. The experimental analysis for the security and cost is stated in Section 5 and the overall conclusion is presented in Section 5.

2.Hadoop in IoT Forensic

Hadoop is a popular open-source framework for distributed storage and processing of big data sets. It is often used in IoT (Internet of Things) applications to manage and analyze the massive amounts of data generated by IoT devices. With the increasing adoption of IoT devices, there is a need for a scalable and efficient data processing system that can handle the large volumes of data generated by these devices [16]. Hadoop's distributed file system, HDFS, provides a scalable and fault-tolerant storage system that can handle the data generated by IoT devices. Hadoop also offers tools for data processing and analysis, such as MapReduce and Spark, which can be used to analyze the data generated by IoT devices in real-time. This allows organizations to derive insights from the data and make informed decisions. In addition, Hadoop integrates well with other big data technologies, such as Apache Kafka and Apache Storm, which are commonly used in IoT applications for real-time data processing and analysis.

Hadoop IoT Forensics and Cyber Trust Blockchain are two technologies that can be used together to provide a secure and reliable framework for managing and analyzing data generated by IoT devices. IoT Forensics using Hadoop can help in the investigation of security incidents and data breaches in IoT systems. Hadoop can be used to store, process and analyze large

amounts of data generated by IoT devices, and advanced data analytics tools can be used to detect anomalous behavior or suspicious patterns. By integrating Hadoop with Cyber Trust Blockchain, the security of the IoT system can be further enhanced. Cyber Trust Blockchain provides a secure and transparent distributed ledger that can be used to store and manage data in a tamper-proof manner. This can help in ensuring the integrity and authenticity of the data generated by IoT devices. In addition, Cyber Trust Blockchain can be used to manage the identities and access control of IoT devices. By using blockchain-based authentication and authorization mechanisms, the security of the IoT system can be further strengthened. Overall, the integration of Hadoop IoT Forensics with Cyber Trust Blockchain can provide a powerful framework for managing and analysing data generated by IoT devices, while ensuring the security and trustworthiness of the system.

The Hadoop IoT Forensic framework model comprises of the Hadoop Distributed File System (HDFS) to process the data collected from the IoT devices. With HDFS system large files are processes with single node based on MapReduce Framework. The HDFS model uses the cluster-based approach for the local distribution of resources effectively to store and process IoT data. The figure 1 illustrated the HDFS model for the remote location estimation model for the Data Nodes within the cluster.

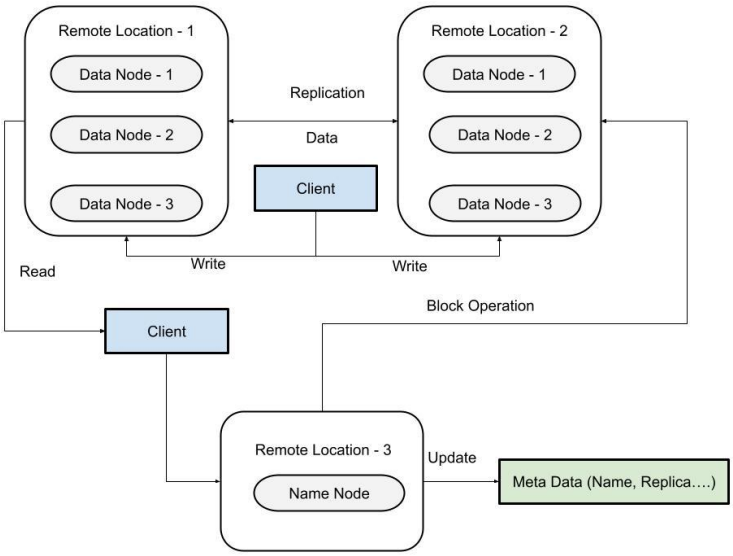

Figure 1: Architecture of Hadoop

The Hadoop Architecture comprises of the three components such as HDFS, MapReduce and YARN. Within HDFS it comprises of the three components those are Name Node, Data Node and the Backup Node (or Secondary node). With the Name Node the architecture is estimated for the master node for the computation of IoT architecture those are management in blocks stated as Data Node in the meta-data format. MapReduce framework comprises of programming paradigm in the centric approach to increases the intensive cluster environment in distributed manner. The MapReduce framework comprises of three functions such as Map ( ), Combining ( ) /Shuffling (), and Reduce () for preparation of input data to derive the final data. The deployed model is evaluated with the execution of the pcap tokenization process. Initially, within the HDFS the data is divided into chunks with the mapping key pair values. The data pairs

are shuffled with the integration of data to generate the final results. The environmental setup of the HDFS model for the forensic is presented in Figure 2. The modules of HDFS scrutinize the IoT data traffic with different modules "data elicitation', 'malicious vector', 'analysis modules threat analysis, and visualization module'. With the data elicitation module data sources are identified with cyberspace which can be either Intranet or Internet and comprises the network nodes.

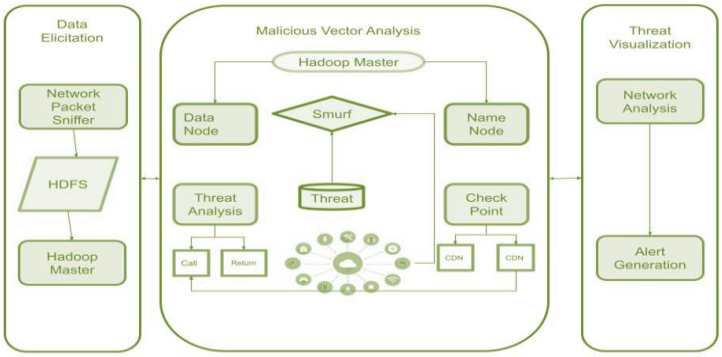

Figure 2: Hadoop Framework for IHBF-ML

With analysis of the malicious vector, the logic is implemented with the incorporation of five nodes. The initial node is the master or name node those node names and data need to be utilized maximum level with Hadoop-Master. Those modules are followed by the additional data comprised of Hadoop-Slaves from Node-0 to 3. Those nodes are configured with an HDFS module with the generation of "pcap" sniffers for data elicitation.

3.Hadoop IoT Forensic with Machine Learning

With HDFS model IoT data is processed for the extracted fields to perform analysis based on feature extraction and capturing od essential dataset features. Those features are classified and processed for the IoT data classification with utilization of different machine learning model. The architecture comprises of the IoT forensic framework as shown in figure 3. Finally, the developed HDFS model utilized for the ML model performance evaluation matrices with consideration of tested modules. The analysis architecture of IoT forensic comprises of four features such as: Data collection and information generation module, feature analysis and extraction, Implementation of machine learning and analysis of different metrices efficiencies.

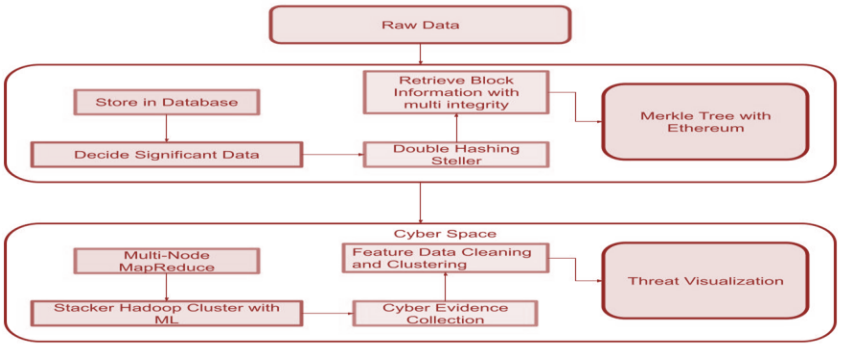

Figure 3: Overall Process in IHBF-ML

The node header in Hadoop uses the cyber trust blockchain model with the characteristics for the analysis. The analysis is based on the Cyber-Physical System (CPS) for the IoT forensic activities that comprise the ecosystem that utilizes the transparency and visible in the network. The CPS model uses the cyber blockchain for exclusive data transfer access and encryption schemes. Within the developed model cyber trust enables public blockchain model is implemented for the establishment and estimation of other possibilities in the network. The dataset is stored in the realistic cloud server with the incorporation of blockchain technology within the file storage system. In the developed model HDFS is act as the file storage system those are conceptual in nature comprises of the agencies in the CPS system with smart contracts.

3.1 Smart Contracts in HDFS Model

HDFS Ethereum Smart Contract HDFS can be utilized in IoT forensic investigations to ensure the integrity and immutability of the collected data. By using HDFS, the data can be securely stored and accessed from different nodes, making it easy to share and analyze the data. Additionally, Ethereum smart contracts can be used to automate the process of data collection, storage, and analysis, making the process more efficient and less prone to errors. When it comes to IoT forensic investigations, it's crucial to collect and preserve the data in a way that maintains its integrity and authenticity. By using Ethereum Smart Contract HDFS, the data can be stored in a distributed ledger that is secured by cryptographic protocols, making it difficult for any malicious actors to tamper with the data. Furthermore, machine learning algorithms such as CatBoost can be used to analyze the data collected from IoT devices. This can help investigators to identify patterns and anomalies in the data, which can be used to reconstruct the events leading up to an incident. By using Ethereum Smart Contract HDFS in combination with machine learning algorithms, the data collected during IoT forensic investigations can be analyzed more effectively, providing investigators with valuable insights into the incident.

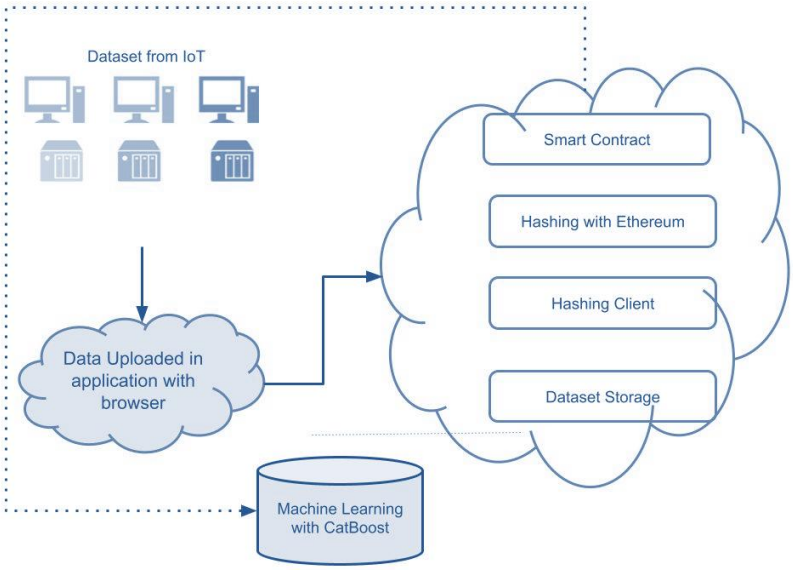

Figure 4: Smart Contract in IHBF-ML for IoT – Forensic

This data set is further sent to IPFS server configured with following

1. Swarm listening at /ip4/10.0.2.15/tcp/4001
2. Swarm listening at /ip4/127.0.0.1/tcp/4001
3. Swarm listening at /ip6/::1/tcp/4001
4. Swarm listening at /p2p-circuit
5. Swarm announcing /ip4/127.0.0.1/tcp/4001
6. Swarm announcing /ip6/::1/tcp/4001
7. API server listening on /ip4/127.0.0.1/tcp/5001
8. Gateway (read-only) server listening on /ip4/127.0.0.1/tcp/8080

Once received at IPFS server, the hash of the data set is generated and visible at the user interface. In addition to the account 1-SEG, remaining 10 accounts also have been configured and added in the MetaMask to facilitate internal transaction. This has received from the IPFS server network is stored in Ethereum Blockchain vide the deployment of the smart contract grid.sol invoked. 138. The meta mask Ethereum transaction network Smart grid IPFS address is set up as 0x6121e72032D792185192fe9b6fA03811fF9C7959. On activation of the application, the smart contract Grid is invoked and gets deployed on the blockchain.

The smart grids are utilized for the collection of data with IoT devices that are defined in definite parts at the right onset. The smart enabled IoT devices use the buzz for the research domain in the cyber trust blockchain for the storage and handling of the huge dataset in real-time applications. The cyber trust blockchain model is illustrated in Figure 4. The blockchain enables IoT system comprised of the Inter File System for the cyber blockchain with HDFS with a combine smart IoT environment. The public blockchain, model uses the security scheme of the Cat Boost model for the processing and storing of large IoT forensic activities. The data

generated from the IoT nodes are evaluated for the different traces in cyberspace for the management of the dataset. The cyberspace processing is performed with the master node and metadata that influence the overall performance of the network.

Algorithm 1: MapReduce IHBF-ML
INPUT: Traffic Information in the IoT node
OUTPUT: Classification of network traffic malicious activities.
Begin
Load Hadoop traffic in the environmental setup
for every Packet capture Loop
Extract the traffic features
end for
Ranking of malicious activities through ranking algorithm
Compare features for control and configuration
Identification of malicious activities
Classification of network traffic cluster with machine learning
Validate the model.
END

Cyber IoT forensic activities comprises of the electronic records for the mathematical management and processing of the collected data. Through the legal admissibility the authentication is performed and evaluated. The constructed model incorporates the evidence copy within the multiple node files in the network. The master node is comprising of the data related to multiple nodes and digest the master node to maintain the suitable and appropriate document repository for the reduced cost with the data processing.

3.2 Machine Learning Model for Anomaly Detection

With the developed model it is aimed to minimize the cost and increases the security of the data through the Hadoop nodes evaluated in the chunks. Based on the MapReduce codes are executed the data in small chunks resulted in Hadooop master for the creation of separate files. As the targeted forensic framework focused on the identification and prevention of Denial of Service (DDoS) attacks. To visualize the process comma-separated files are generated and stored in the administrator for the actions. Initially, the IoT data were processed in the TCP dump in the mirror port with consideration of switches those are connected to the IoT network with the data sniffing at different time instances and duration for the different data amount and type of the network traffic. Secondly, the constructed algorithm uses the Merkel Tree Ethereum for the conversion of binary header traffic into readable text files. Also, the pcap files are utilized for the conversion of single file into readable file formation for the filtered target packet to perform analysis. With DDoS analysis control flags flooding attacks are computed for the data processing and estimation. The data is filtered and analyzed for the packet those are suspicious in the respective sources those are launched in the distributed or non-distributed manner for the denial of service for the server target detection to the destination. The classification is performed based on the consideration of the malicious packet sets with consideration of the environmental factors with the captured traffic. The collected data traffic is evaluated based on CatBoost machine

learning model for the reduction of cost and increased security in the IoT forensic. The filter packets based on the type and characteristics are computed based on the captured network traffic. With the implementation of MapReduce framework server receives more traffic data those are configured with properties modification in the packets with consideration of different sources per sec. With implementation of the CatBoost machine learning model malicious activities in the IoT are evaluated. Through the developed model the configuration values are computed with machine learning model those are utilized for the application execution based on the consideration of environment with consideration of data captured in the malicious packets.

3.3 Feature Analysis

Within the feature extraction modules all parameters are optional with the computation of statistics those are changes based on the requirements. With the feature extraction in the HQL it comprises of the "group clause" those are processed within MapReduce algorithm to achieve significant performance characteristics. MapReduce framework model comprises of the mapping, tokenization, shuffling and sorting implemented within the Reducer module utilized for the key value pair searching and reduction with use of hash values represented in equation 1.

Input :< k1; v1 > => map => => || combiner || => reduce=> < k3; v3 >: Output (1)

With the Hadoop Cluster the database architecture are optimized with the feature extraction modules implemented and tested in Machine learning platform. The analysis is performed based on the consideration of the Internet traces those are executed in parallel manner with the network traffic capturing with estimation of feature such as node ID, position and other environmental parameters.

4.Simulation Results and Analysis

Simulation of IHBF-ML model is demonstrated and conducted for the analysis of the IoT forensic ecosystem blockchain. The simulation is implemented in Linux machine with I5 processor with the RAM memory of 4GB of Ubuntu. The performance of IHBF-ML model is evaluated with the HDFS system for the increase in security and reduced cost-effectiveness. The credentials of the IoT nodes are presented in table 1.

Table 1: IoT nodes credentials for experiment

Node	Public Key	Private Key
1	0xB0908B6e032fF8F79524292E9B017 5bD713F6aeD	f399ae6ed4c92851a28f179ac9bc7140ceb4f0 3b6d30635af5bdfc0701e7876c
2	0x25E1DED38B2ec0839ccE6787225e 8Cc41bE8Bb97	5da19b4ddf3775714826a4eb28f01e8cf5e57 85875116928bb4f4811c385bb6b
3	0x784b5cbA80069059EDE9cCF5a7d4 c0F9001D0aAF	c7b8656b422aace74586da37d7372431fe14d cbc67030047a64b239b7c0835f9
4	0x1a702009E32F7435d5cc95dD172a4 F54DEe1dcC7	059b3c037d9db76f7a46f166c32b48612b6ca 6758cf73ffdd9a847f2eacc4ea6
5	0xC9eC5bEc83028Ad3BFcbF4767Ad1 d831a6011749	90d93f804064e8b877c090f6acf7acb32e522e 95bdf2e980717f98b33395a514
6	0xe9d7e06Ef1Be5F1336090550909729 A64dA74599	28a66f08f7aede7a78b1107db6c18a9785193 503b9be9b5531acfd8618a2232e

The Blockchain technology with cyber trust model comprises of smart contracts for the token exchanges between the nodes. Experimental analysis comprises of five tokens those implemented in the IoT forensic ecosystem for the data transmission. The IoT vehciles are computed based on the utilization of the machine learning model with Hadoop for the reduction of cost in the network. The optimal threshold set for the information exchanges between nodes are presented in table 2.

Table 2: Token exchange set parameters

Node in Cluster	Energy Level	Number of Hops	Residual Energy	Threshold Level
5	1Joules	3	> 1 J	0.5

In the simulation of the data exchange between the nodes are stored in the cyber space enabled Hadoop distributed file system. The HDFS enabled public blockchain model for the IoT forensic activities uses the transaction details within the space as presented in table 3. The generated model uses the 50 Ethers for the transaction in the balanced state. The exchange token between the nodes are evaluated with the periodic maintenance for the nodes and service billing process.

Table 3: Transaction Details of Ethers

Node	Public Key	Initial Ethers	Final Ethers	Generated Blocks
1	0xB0908B6e032fF8F79524292E9B0175bD713F6aeD	50	58.52	133
2	0x25E1DED38B2ec0839ccE6787225e8Cc41bE8Bb97	50	59.83	
3	0x784b5cbA80069059EDE9cCF5a7d4c0F9001D0aAF	50	57.61	
4	0x1a702009E32F7435d5cc95dD172a4F54DEe1dcC7	50	48.74	
5	0xC9eC5bEc83028Ad3BFcbF4767Ad1d831a6011749	50	56.32	
6	0xe9d7e06Ef1Be5F1336090550909729A64dA74599	50	43.94	

The smart contracts are deployed based on the designed criteria with the foreseen en-route data transmission in the nodes. As the model aimed to reduce the transaction cost the estimated cost for the all the variables are presented in table 4.

Table 4: Cost Estimation

Slow Data Transmission	Average Data Transmission	Fast Data Transmission
0.00018	0.00029	0.00037

The total number of blockchain in the account is configured with the node ID other device components in the ecosystem. Those are co-ordinated with the transaction of the each devices those perform the tokenization of the services offered to evaluate the transaction between the different accounts. The interface transaction is performed with the full blockchain with the application host of http://localhost:3000. The designed model comprises of the HDFS based Hash generation with the Merkel Tree. The generated has for the intra transaction are presented in table 5.

Table 5: Intra Transactions amongst IoT devices

Sender Node Address	Contract Address	Tokens	Transaction Hash	Block No
0xb3Fd8Bb9c3B45f6019	0x12C24414CE5A3898F	+24.25	0xb1c179d5c7e896d37f54c2	134

AAE4502D359DD8F4A32414	80c0F03cf8788BC62c40F17	ETH	6c6854ec9416d669364f55c83a4935189767901634	
0x35C7cDBC49AF4e7Cd1Bd2845ae91285BA758F9CC	0x12f7355084bB3406297D8c7bDb6E676C4Dd58d35	+12.75 ETH	0x5c35898dfcb9d96ae5d27e2abfa911271e6115cd312120be10241ac0ed56db9a	133
0x19a47DE366b3bE7F627b266c653C1bcd84832C5B	0x5eea94611EA6e355beb1843818155385578dBB4c	-12.50 ETH	0xd0a20908204ea3a178b01a60944a9bf2bfcda80a8b4bb34d5ce5ca3956caed05	132
0x260B455D49566BD690Ec8c20d678117897fb462E	0x35C7cDBC49AF4e7Cd1Bd2845ae91285BA758F9CC	-10.75 ETH	0xbd290f4f439e11648ebf99355f35cab7acf1b38fdcbc0f8d7bf13616095ef656	11
0x0A83B1d4dEcbE172C2030dB11a18F59F1f106d4E	0x12f7355084bB3406297D8c7bDb6E676C4Dd58d35	+11.25 ETH	0x329a13031d603e217cd1d2f4229661ef3f4b3e74bcc3ce49375c180d29a28651	10
0x5eea94611EA6e355beb1843818155385578dBB4	0x84Ae4f4C1CbE0124233f40D205C93B49701418	-22.25 ETH	0x1226356868b8e840a15924073a0182f092effe0b5a924f	9

The HDFS system timing for the data upload at different file sizes are presented in table 6. Through analysis it is observed that uploading size of RAM is related to the IoT dataset. The processing time with the machine are computed as presented.

Table 6: Processing Time of Machine

IoT Dataset	Time taken with machine			
	3 GB	4GB	6GB	8GB
10	3.691	3.187	3.14	2.281
100	3.82	3.036	2.968	2.999
1000	9.925	9.379	9.456	9.629
2000	17.853	17.246	17.4	16.05
5000	40.486	38.988	37.789	39.418
10000	77.493	72.35	71.133	69.136

The Ethereum based public blockchain model computed for the transaction details those need to be estimated for the 360 bytes size is presented in figure 5. The number of transaction blocks is presented in equation (2) and calculation is performed as in equation (3)

$$No.of\ transaction\ within\ blocks = \frac{Size\ of\ Blocks\ (Bytes)}{Average\ transaction\ size\ (bytes)} \quad (2)$$

$$No.of\ transaction\ within\ blocks = \frac{1000000}{360} \sim 2788.666\ transactions \quad (3)$$

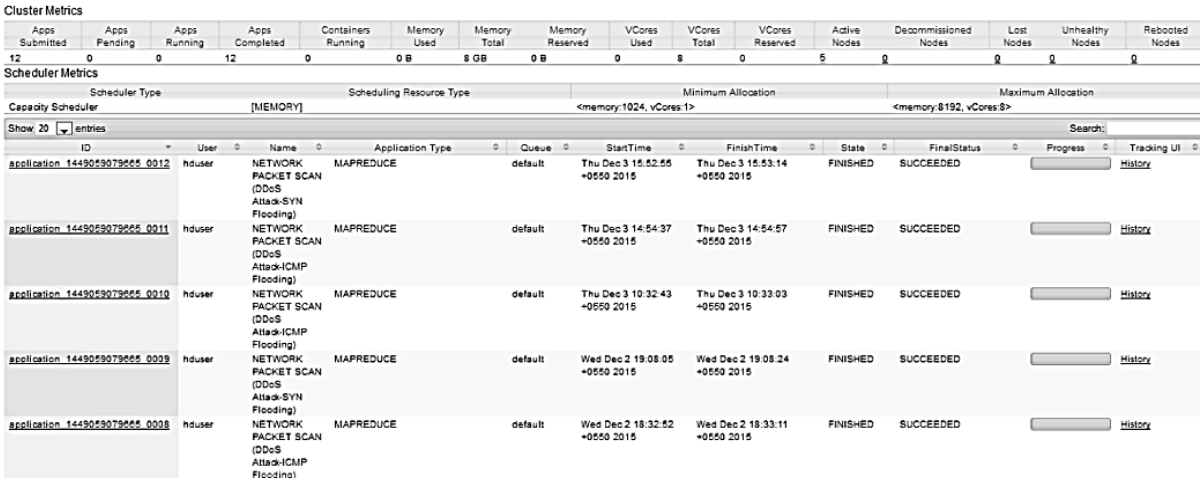

Figure 5: Memory processing in RAM

The security analysis is performed for the Hadoop Cluster model with consideration of the machine learning model implemented with the CatBoost model for the attack detection and classification. The scan is conducted for the IoT total packets of 368743861 with the control packets of 136828. Through the scan the malicous packets identified are 38681 with the suspicious instances of 582. The security analysis is performed with the CatBoost Machine learning model for the classification of attacks in the network.

With IHBF-ML model network traffic are monitored with increase in network size for the increase in the system with increases in same ration with the maximal utilization of the available resources and maintains the constant values. The obtained values focused in the security and cost-effectiveness for the attached number of nodes ad replica of the HDFS input or output. The processing time is evaluated within the cyber space either it increases or decreases the replica number with the increase in Data Nodes. The classification performance is presented in Table 7.

Table 7: Classification Performance Analysis

Class	Accuracy	Precision	Recall	F -Score
Malicious	0.9956	0.9946	0.9956	0.9946
Non - Malicious	0.9972	0.9973	0.9957	0.9937

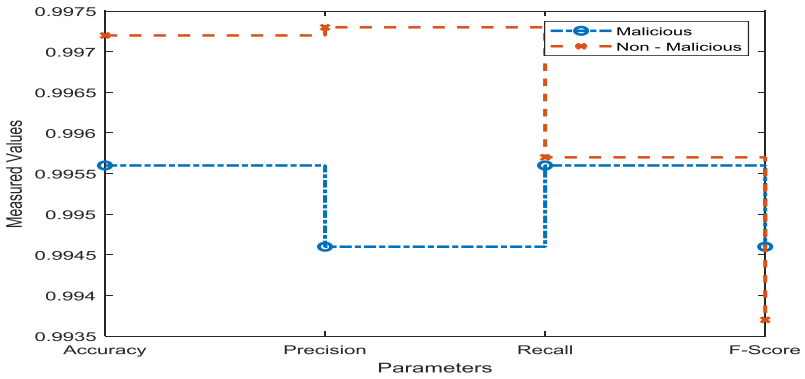

Figure 6: Measured Machine Learning Model for Security

The performance of IHBF-ML model performs effective statistics for the obtained data to evaluate the performance of IoT network is presented in figure 6. The precision provides the decision about the prediction of correct results to derive the output from the framework obtained through the CatBoost Machine learning model. Finally, the measure evaluated the test to derive the expected results. The accuracy is computed with the 0.99 with recall and F-score value of 0.99 and 0.99 respectively. The computation process is evaluated for each boat for the transmission of 10 every data. The cost utilized for the IHBF-ML model is comparatively examined with the public blockchain model achieves the overall coast of $18,000,000as presented in table 8, The costs for the developed IHBF-ML model is evaluated based on the consideration of different factors into consideration.

Table 8: Comparison of Cost

Blockchain Technology	**Annual Cost**
EOS	$ 29,000,000.
Steller	$ 23,000,000.
IHBF-ML (Ethereum)	$18,000,000

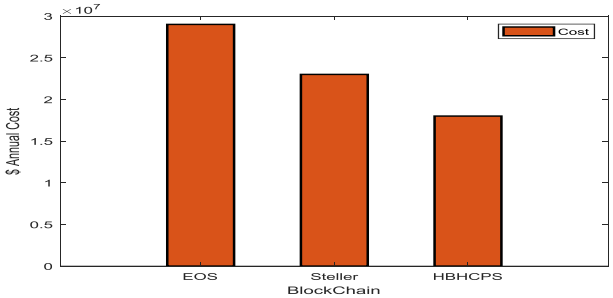

Figure 7: Comparison of Blockchain

In Table 8 and Figure 7 the cost for the proposed IHBF-ML model performance is comparatively examined with the conventional blockchain model such as EOS and Steller. The developed model performance cost comparison is performed for the different blockchain network with an estimation of the approach cost values. The multichain model exhibits a total cost of $15,000,000, Ethereum only requires the implementation cost of $500 for the contract implementation. The developed IoT forensic model achieves an overall cost of $18,000,000. The implemented IHBF-ML model is compared with the conventional Steller blockchain model while it requires the total computation cost for the cyber trust blockchain cost of $ 23,000,000. The simulation analysis confirmed that the IHBF-ML model significantly minimizes the cost and increases the security in the cyber trust blockchain processing.

5.Conclusion

The IHBF-ML model is an Integrated Hadoop Blockchain Forensic Machin Learning that integrates the Hadoop Distributed File System (HDFS) with the MapReduce framework. This integration allows for parallel processing of IoT forensic data, which can significantly reduce the processing time and cost. In the IHBF-ML model, the data collected from IoT devices is first stored in the HDFS, which provides a highly scalable and fault-tolerant storage system. The

MapReduce framework is then used to perform parallel processing on this data to extract useful information and detect anomalies. To ensure the security of the IHBF-ML model, Ethereum smart contracts are used to enforce the rules and regulations for data access and manipulation. These smart contracts are stored on the Ethereum blockchain, which provides a tamper-proof and decentralized platform for executing the contracts. The use of machine learning algorithms, specifically the CatBoost algorithm, is also incorporated into the IHBF-ML model for anomaly detection. The algorithm is trained on a dataset of normal and abnormal IoT device behavior, and it can identify anomalous behavior that may indicate a security breach. The IHBF-ML model has been evaluated through experimental analysis, which has shown that it achieves a cost of $18,000,000, significantly lower than other blockchain models. Additionally, the security analysis has demonstrated a high anomaly detection rate of 99%. The IHBF-ML model provides a cost-effective and secure solution for IoT forensic through the integration of HDFS, MapReduce, Ethereum smart contracts, and machine learning algorithms. It has the potential to improve the efficiency and effectiveness of IoT forensic investigations while reducing the associated costs.

Acknowledgment: Not Applicable.
Funding Statement: The author(s) received no specific funding for this study.
Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study

References

[1] S.Kapoor and S. Sharma, "An overview of network and computer forensics," *International Journal of Engineering and Advanced Technology*, vol.9, no.4, pp.1358-1364, 2022.
[2] S. Venkatramulu, Md. Sharfuddin Waseem, Arshiya Taneem, Sri Yashaswini Thoutam, Snigdha Apuri and Nachiketh, "Research on SQL Injection Attacks using Word Embedding Techniques and Machine Learning," *Journal of Sensors, IoT & Health Sciences (JSIHS,ISSN: 2584-2560)*, *2*(1), 55-64, 2024.
[3] P.Chavan, "Digital forensics: An overview," *International Journal of Computer Science and Mobile Computing*, vol.10, no.4, pp.20-29, 2021.
[4] Sreedhhar Bhukya, K. VinayKumar and N.C. Santosh, "A Novel Dynamic Novel Growth model for Mobile Social Networks," *Journal of Computer Allied Intelligence*, vol.2, no.1, pp.46-53, 2024.
[5] A.Asghar, M.Imran and N. Ahmad, "IoT forensics: A survey on challenges, techniques, and future directions," *IEEE Communications Surveys & Tutorials*, vol.22, no.2, pp.1361-1382, 2020.
[6] L.Zhang, X.Zhu and X. Dong, "IoT forensics: Issues, challenges, and opportunities," *IEEE Internet of Things Journal*, vol.8, no.1, pp.447-463, 2021.
[7] M. A.Siddique, A. Alamri, G. Fortino and K. K. R. Choo, "A comprehensive review of Internet of Things (IoT) forensics," *Computers & Security*, vol.107, pp.102283, 2021.
[8] X.Li, X. Chen and Q. Li, "IoT forensics: Challenges and future research directions," *IEEE Internet of Things Journal*, vol.8, no.7, pp.5736-5753, 2021.
[9] S. S.Ahamed and S.A. Alshehri, "IoT forensics and trust blockchain: A comprehensive review," *IEEE Access*, vol.9, pp.122443-122466, 2021.
[10] F.Xue and S.Wang, "A review of IoT forensics and trust blockchain," *International Journal of Distributed Sensor Networks*, vol.17, no.8, pp.15501477211028961, 2021.
[11] K.Bilal, I.Yaqoob, M.A.Khan, N.Javaid and A.Almogren *et al*., "Blockchain-enabled IoT forensics: Opportunities, challenges, and future directions," *IEEE Communications Magazine*, vol.59, no.8, pp.66-72, 2021.
[12] J.Kim, "IoT forensics: Emerging trends, challenges, and opportunities," *Computer Communications*, vol.170, pp.240-255, 2021.

[13] T. W.Liao, C.M.Chen and Y.S. Chiu, "Forensic investigation for android messaging app on non-rooted devices," *Computers & Security*, vol.94, pp.101879, 2020.
[14] K.K.Mak and Y. Zhu, "Smart contract based blockchain for IoT forensics," *IEEE Internet of Things Journal*, vol.7, no.4, pp.2943-2954, 2020.
[15] D.Moloney, S.Sezer and I. Muttik, "The evolution of IoT and its impact on digital forensics," *Digital Investigation*, vol.37, pp.101-109, 2021.
[16] L.Mottola and G.P. Picco, "The case for IoT middleware for trustworthy IoT systems," *IEEE Internet of Things Journal*, vol.8, no.2, pp.928-935, 2021.