The Seybold Report

ISSN 1533-9211

DOI 10.5281/zenodo.7099360

INTEGRATIVE FORENSICS FRAMEWORK APPROACH FOR INTERNET OF THINGS USING BLOCK CHAIN

K.VENKATAGURUNATHAM NAIDU[1] and Dr. RAJAVARMAN.V.N[2]

[1]Research scholar, Department of Computer Science and Engineering, Dr. MGR Educational and Research Institute, Chennai, Tamilnadu, India.Email: venkatspmvv@gmail.com
[2]Professor, Department of Computer Science and Engineering, Dr. MGR Educational and Research Institute, Chennai, Tamilnadu, India. Email: rajavarman.vn@drmgrdu.ac.in

Abstract:

IoT devices have been widely embraced in the recent decade, allowing for the collection of varied data from various contexts. Data storage presents difficulties since the data may be corrupted and the integrity of the data may be damaged without being discovered. Integrity and data provenance are necessary in these situations in order to discover the source of any occurrence and prove it in legal disputes. To overcome these difficulties, blockchain presents significant prospects since, due to its distributed nature, it can safeguard data integrity. However, there are also drawbacks associated with keeping large amounts of data on a public block chain, such as substantial transaction fees. In this research, we offer a very cost-effective and trustworthy digital forensics methodology that uses many low-cost block chain networks as interim storage before committing evidence to Ethereum. Merkle trees, which store hashes of recorded event data from IoT devices hierarchically, are used to decrease Ethereum costs. The identification of compromised devices, as well as the gathering and storage of evidence about alleged hostile conduct in IoT networks, have emerged as high priority issues. This study describes a block chain-based solution that deals with the collecting and storage of digital forensic data. The system makes use of a private forensic evidence database to store the acquired evidence, as well as a permissioned block chain to provide security services such as integrity, authentication, and non-repudiation, allowing the evidence to be utilized in a court of law.

Key Words: IoT device, Data Integrity, digital forensics, Ethereum, Block chain

1. INTRODUCTION

The evolution of communication technologies, sensing devices, and inexpensive computing devices has ushered in the internet of things (IoT) era, which allows for the collection and transmission of different ambient data to remote places [1]. In several industries, such as transportation, energy, healthcare, agriculture, and hospitality, IoT is becoming the de facto technology [2]. The data received from various IoT devices is utilised in these apps to do comprehensive analyses in order to make educated judgments and take action. However, in other cases, the data is vital for maintaining key infrastructure (such as power systems and transportation) and detecting and understanding breakdowns. In particular, if there are failures due to human mistake or malicious assaults, it is critical to be able to identify the source of the problems and punish those accountable. IoT data must thus be transmitted and stored securely in order for these apps to function.This needs techniques for storing IoT data in order to conduct digital forensics investigations. Because the data must be submitted as evidence in the event of a disagreement, it must be stored in a secure manner that cannot be deleted or manipulated. Because it can enable authenticity verification, data provenance, and data integrity, emerging Blockchain technology can be a great match for such applications

The Seybold Report
ISSN 1533-9211
DOI 10.5281/zenodo.7099360

[3]. The creation of a permissioned blockchain (i.e., a private blockchain network) that permits only specified entities to join and exchange information with some untrusted parties would be an excellent option. The stakeholders, on the other hand, would not be cooperative in this regard. Furthermore, the security of a private blockchain is determined by the number of users, and small numbers may pose a risk in terms of consensus. To reduce maintenance costs and boost confidence, it makes more sense to utilise a public blockchain. However, there is a cost difficulty with trans-actions in the case of public blockchain. This is especially true of major blockchain networks like Ethereum and Bitcoin. As the number of IoT devices writing to blockchain grows, the solution will become unscalable. While alternative less expensive ledger systems can be used instead of Ethereum, their dependability will be greatly reduced since these ledgers may not have enough nodes, allowing a 51 percent assault to be carried out with less effort. As a result, cost effective solutions for storing IoT data in public blockchains are required. In this paper, we propose a multi-chain approach for IoT data integrity verification, in which we use multiple relatively affordable blockchain networks such as EOS [4] and Stellar [5] (in comparison to Ethereum and Bitcoin) for temporarily storing the hash of the IoT data before they are permanently stored on Ethereum. The benefits of this strategy are twofold: First, it is cost-effective since we save incident information mostly in EOS and Stellar, which are significantly less expensive, and only a daily summary of all transactions is written to Ethereum. Second, the suggested architecture is more secure and resistant to a percent 51 consensus assault [6]. Before the summary is published to Ethereum, the attacker must either hack both blockchain networks in the first level on the same day, or modify data in both Ethereum and one of the blockchains in the first level,but it is very difficult. This paper additionally uses advanced intrusion detection and distributed ledger technology (DLT) solutions developed as part of the Cyber-Trust project (https://cyber-trust.eu/) to address challenges in the forensic evidence collection, preservation, and investigation process for IoT environments. In more detail, several techniques implanted at an entry point, such as anomaly detection,monitoring and profiling, enable supervision of the condition and conduct of Internet of things devices, effectively increasing the discovery of possible attacks and 0-day threats together with the gathering of evidence in the case of malicious conversations identified. The gathered data, together with the information required for correlation and subsequent examination of an attack's created events, is saved in the evidence database (evDB), which is maintained by the Internet service provider (ISP). The metadata is published on a blockchain that is maintained by the ISPs, allowing law enforcement agencies (LEA) to successfully trace back an attack to its source by keeping the chronological ordering of attack evidences on a worldwide scale. The suggested method, known as Cyber-Trust blockchain (CTB), allows in order to establish the chain of custody by recording and keeping the historical history of dealing digital evidence, entities involved as in investigation stage, including Law enforcement agencies, prosecutors, must access as well as need to handle digital data. To fulfil privacy needs, the CTB system uses HyperLedger Fabric and creates a permissioned blockchain.

The Seybold Report
ISSN 1533-9211
DOI 10.5281/zenodo.7099360

2. BACKGROUND

Below background topics supported us in proposing integrative forensic framework

2.1 Basics of Blockchain

Blockchain The blockchain is a digital ledger that stores encoded blocks of data and connects them to create an informational timeline with a single truthful source.The Files are disseminated, creating a asset immutable record , as opposed to being copied or relocated. An asset is decentralised, allowing for transparency and real-time public access. A visual record of revisions protects the document's integrity and increases confidence in asset. Almost any company may benefit from using block chain because of its built-in security features and public ledger.

Public Blockchain The term"public blockchain network" refers to a blockchain network that anybody may join at any time. In general, there are no limits on who can participate. Furthermore, anybody with access to the ledger may participate in the consensus process. One of the public blockchain platforms, for example, is Ethereum. Ethereum [7] is the world's most popular and reliable Smart Contract-oriented Blockchain network. It is a permissionless, public blockchain, which implies that anybody may access Ethereum's information and conduct transactions on their own. The Solidity programming language is used by Ethereum. To create an Ethereum virtual machine-compiled contract Each contract does have a gas cost which is calculated based on the contract's total workloads and memory use.

Private Blockchain Blockchain technology known as a private blockchain is one in which the network is managed by single entity.As a result, it is not open to the general public to participate.In practise, all private blockchain systems will have some kind of permission process in place to identify who is logging on to the site. These platforms are essentially created by private blockchain solutions for an organization's internal network connection infrastructure. for example private blockchain platforms are Steller,EOS. The first smart contract-oriented blockchain network, Stellar [5], seeks to provide customers with a scalable paymentgateway. Stellar's block/contract mining time is roughly 3-5 seconds, making it incredibly scalable. Thousands of transactions can be confirmed every second. EOS [4] is a wellknown and very effective Blockchain Network. It gets its name from the Ethereum Operating System. As a consensus system, EOS employs delegated proof of stake, which is both efficient and energy efficient. Deploying a smart contract to the EOS network is simple and cheap, but the contract developer should have enough EOS, CPU, and RAM to make the most of EOS bandwidth.

Merkle Tree In computer science, Merkle trees, sometimes referred to as Binary hash trees, are a typical kind of data structure. They are used to more efficiently and securely encrypt blockchain data in cryptocurrencies such as bitcoin.It's a mathematical data structure built up of hashes of different data blocks that aggregate up all of the transactions in a block. Along with data consistency and content verification, it also enables quick and secure content verification across huge datasets.

The Seybold Report
ISSN 1533-9211
DOI 10.5281/zenodo.7099360

Beacon Chain The Beacon Chain is the heart of the Ethereum 2.0 system chain, and it's in charge of maintaining the Casper Proof of Stake consensus mechanism for itself and all of the shard networks.

2.2 Detection of Intrusion in IoT

For identifying prospective threats in a network, intrusion detection systems commonly use signature-based and anomaly-based methodologies, with the latter relying on the monitoring of a network's devices for any aberrant behavioral patterns [8]. The framework proposed by Nguyen, et al. autonomously identifies anomalies in an IoT network [9] in order to detect compromised IoT devices. This is achieved by classifying devices according to their categories and creating normal profiles which are utilised for the identification of abnormalities using a self-learning framework. In order to protect smart home ecosystems against distributed network attacks by suspicious Internet of things, Siotome, a privacy-preserving architecture, proposed in [10]. The system has the capacity of detecting,monitoring ,analysing Internet - of - things potential threats as well as providing an efficient security framework by leveraging machine learning techniques to establish the best operational setups.

3. RELATED WORK

Various concepts for using blockchain for data integrity verification have been offered. [11], [12], and [13] are working on a general blockchain-based data provenance architecture for IoT produced data. They begin by identifying the concerns connected to security and trust, and then discuss how blockchain may be used to address these issues. In [14], the authors propose a system for vehicle accident scenarios that is designed to record data in blockchain when an accident occurs. To protect anonymity, they deploy a simpler public key infrastructure designed specifically for automotive networks. The blockchain data is utilised to resolve any disputes between the insurance, the owner, and the manufacturer. A similar strategy was utilised by Gipp et al. [15] on cellphones that are deployed as dashboard cameras in automobiles. When a smartphone detects an accident using its accelerometer sensor, it begins recording the scenario and, at the conclusion, calculates the hash to be uploaded to the public blockchain.It stores the aggregate of the hashes to reduce the cost to a minimal. The user can offer the original video together with the hash to establish that the video saved on the phone has not been tampered with. Block- DEF [16] proposes a blockchain based safe digital evidence structure. The objective is to keep evidence and information about evidence separate. They propose a lightweight blockchain design in order to minimise data bloat. They claim that it is a scalable platform for securing and tamper-proofing evidence. ProvChain [17] aims to offer data assurance for IoT sensors-collected data. Instead of storing the entire data, they generate the hash of the data and store it on the blockchain network. In two respects, our work varies from earlier studies: For starters, hash-based storage isn't the only option. To conserve even more space, we use Merkle trees. Second, and more importantly, we collaborate with several low-cost blockchain networks to improve dependability and security while keeping costs low.

ISSN 1533-9211

DOI 10.5281/zenodo.7099360

4. METHODOLOGY

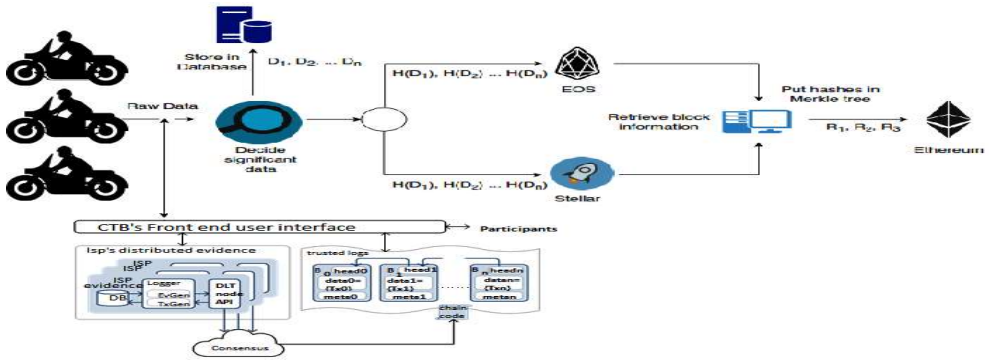
4.1. Motivation

In our scenario, we're looking for a cost-effective solution for a bike rental company that wants to retain sensor data from its bike in a method that allows it to guarantee data integrity in future retrievals. Not only would a secure integrity mechanism lower their insurance premiums, but it would also assist them in promptly addressing future consumer disputes. The most apparent way is for stakeholders to create a permissioned blockchain. When numerous untrustworthy parties want to communicate information, this sort of solution is appropriate. Hyperledger [18] is IBM's solution for this sort of scenario. Because insurance companies are often not helpful owing to management expenses, this alternative is ruled out in our scenario. However, because the rental firm still has to store data in an irreversible manner, a public blockchain might be a viable solution. Consequently, the second preference could be to publish the data right away to the Ethereum network, that is a very secured blockchain platform.However, given the enormous volume of transactions in IoT scenarios, publishing every single transaction on Ethereum would be prohibitively expensive. Ethereum might be useful in various situations, such as asset transfer via smart contract. When the ownership of an automobile is transferred, for example, the money transfer is accomplished. However, it is not a particularly cost-effective solution in circumstances where we require frequent transactions. Another method is to save data in a database and regularly write the calculated hash of the stored data to Ethereum (i.e., once a day). When compared to the prior technique, this will save money and maintain data integrity when it is written to blockchain. While this saves money, it does not ensure data immutability for as long as the data is stored in the data center database. As a result, while the cost is reduced, this solution has security difficulties. As a result, we choose a more cost-effective method based on several blockchain networks, as outlined below.

4.2. Proposed Integrated framework

The propsed architecture contains two parts.one is for low cost preservation of digital evidence and another one is to provide cyber trust block chain for detecting malicious activity and store evidence regarding such incidents.

Fig1:Architecture of Integrative Forensics framework



The Seybold Report
ISSN 1533-9211
DOI 10.5281/zenodo.7099360

Low Cost Forensic framework Our suggested architecture, uses public blockchain technology to solve the trust problem and create a trustless environment. Data storage on a public blockchain, on the other hand, is both costly and fraught with privacy problems. As a result, we turn to less expensive public blockchain options. Although Bitcoin and Ethereum are the most well-known and long-lasting blockchain systems, others, such as Stellar and EOS, have been around for a long time.Because of the small number of users and popularity, they may not be as trustworthy as Bitcoin and Ethereum, but the cost of utilising these platforms is substantially lower. We propose a multi-factor integrity (MFI) approach that combines various low-cost blockchain platforms with Ethereum since utilizing only one of these platforms may not be secure. The concept is comparable to employing a backup system in the event of a system breakdown. We want to make data more resilient in the event that one of the systems goes down or gets compromised. MFI makes it more difficult for a hostile actor to tamper with IoT data that is kept in the company's database invisibly. If one blockchain is hacked or broken, a hostile actor still has at least one more barrier to overcome in order to jeopardise the data's integrity. It's worth noting that all of these platforms are smart contract-based, allowing for seamless communication between them. We use hash functions and the Merkle tree to decrease the size of data that has to be published to public blockchains. Our cost-cutting strategy is as follows:

i) The IoT edge device from a bike uploads the hash of IoT data to the first level of the multi-chain in the first step system. As previously said, only useful data is selected based on predetermined occurrences or criteria. As long as there is useful data, the hash of all this information is sent to both the Stellar & EOS during the day.

ii) At the conclusion of each day, a synchronisation procedure begins, during which the rental company's data centre retrieves the confirmed transactions that have been sent to the first-level blockchains Stellar and EOS. The data centre then constructs a Merkle tree from verified transactions and determines its Merkle root for each one.

iii) Another integrity factor is represented by the Merkle root computed in the preceding step. As a result, it is sent to Ethereum, a more secure and dependable blockchain, and a duplicate is kept in a local database for use in forensic investigations. To avoid transaction costs, Ethereum is only utilised for the hash of all hashes within a day.

Table 1:Cost comparisons of different Block chains

Approach	Total Cost in dollors
Multi chain(Steller+EoS+Etherreum)	443
Ethereum Only(fun call)	13140
Ethereum Only (new contract)	69350

The Table 1 compares the costs of various techniques to ours. As can be seen, the Ethereum-only strategy is rather costly, costing roughly 70K dollors. The bike firm will not be interested in deploying it, despite the fact that it is exceedingly secure and dependable. The alternative Ethereum approach, which uses function calls, is significantly more inexpensive, costing roughly 13K dollors. This is due to the fact that the contract deployment fee is a one-

The Seybold Report
ISSN 1533-9211
DOI 10.5281/zenodo.7099360

time fee, and the hashes are always written to the contract. Nonetheless, this is still a lot more costly than 443 dollors. Our method saves a lot of money and may be highly appealing to a corporation to use. In future If we use Beacon Chain which is the heart of the Ethereum 2.0 system chain the processing time and Processing cost will reduce much more.

Cyber-Trusts blockchain

The major purpose of Cyber-Trust in any domain is to reliably detect compromised local networks and/or components from the network's other components, so that suitable remedial actions may be implemented. To simplify the gathering and subsequent correlation of forensic evidence from many independent sources, intrusion detection technologies are being used at both the device and network levels. To resist cyber-attacks and aid evidence gathering, crucial information from IoT devices is stored on the blockchain and may be accessed later when, for example, a verification of correct operation is required or elements of the system's software must be updated or patched reliably. This means that attributes like as a device's firmware, configuration files, and other data are registered into the Cyber-Trust blockchain at the start of the system's operation and, if necessary, validated against a history of previously legitimate states to guarantee that they have not been tampered with.

Structural Components: The (SGA) smart gateway agent is in charge of collecting network data, such as forensic evidence (all of those relating to the forensic evidence gathering process shown in Fig. 1), monitoring the network's health status, profiling the behaviour of IoT devices, and serving as the primary link with the fundamental parts of a system operating at ISP layer.The SGA executes device fingerprinting when a new device is registered in order to derive the device's behavioural patterns based on network flows provided the device is initially in a clean condition. Furthermore, the SGA employs a lightweight IDS that actively monitors the communication of linked devices to detect anomalous activity, transferring any suspicious traffic to the platform's back-end for deep packet inspection (DPI). In addition to the foregoing, the SGA employs the manufacturer's utilisation description (MUD) to provide device-specific network profiling in order to facilitate accurate feature-set extraction for anomaly identification. A smart device agent (SDA) is placed on more competent IoT devices, such as smart phones, that permits the direct capture of information (including evidence) from end-user IoT devices. The SDA is more restricted because it is primarily in charge of monitoring the device's usage, critical files, and security firmware integrity, patching status, and vulnerabilities. The back-end of the Cyber-Trust platform, and specifically the profiling service, is routinely synced with information on run-time processes and hardware resources utilised (PS). Evidence collection: When the SGA (resp. SDA) detects doubtful network activity and, respective device activity, the relevant evidence is gathered and submitted to the ISP to be kept in the evDB. In network assaults, the evidence consists of IP packets (among other data), but in device-level attacks, it may consist of the full device's picture. The entire procedure is intended to, at the very least, achieve the following objectives:

(i) Make sure the Integrity and Confidentiality of Digital evidence during Storage and Transfer;

The Seybold Report
ISSN 1533-9211
DOI 10.5281/zenodo.7099360

(ii) Guarantee that the digital evidence gathered from and intended to Secure Devices which have Formed a trustworthy relationship through an Authentication Procedure to verify the remote device's hardware/software configuration (such as the BIOS, MBR, and firmware); and

(iii) Compute a non-repudiated provable hypothesis.

Process for Verifying Integrity

When an occurrence occurs that leads to a disagreement about who is to blame, the proposed framework will be utilized to investigate and prove what occurred. Basically, an insurance company working on a claim or a law enforcement officer on the site of an accident must confirm the data's veracity Once the data's integrity has been established, the errant party may be identified without a doubt. In order to do so, the investigator/officer must first access the appropriate forensic data housed in the data centre, according to our architecture. S/he must next gather the submitted transactions that include the hashed data to the first level blockchains, as well as the relevant Merkle root values and Merkle pathways of those transactions. If the computed Merkle root matches the value recorded in Ethereum, the investigator/officer can be certain that the data centre has provided him/her with valid/tamperproof IoT hash data.Additionally, s/he is aware that the existence of a transaction upon that blockchain has been validated by a number of multi-chain miners, as well as the multi-chain system's long PoW or computation time safeguards the reliability of hash information.

5. CONCLUSION AND FUTURE WORK

We suggested a forensics structure in this study that comprises of two layers of numerous blockchain networks. In the event of a disagreement, the system's aim is to validate the validity and integrity of data acquired from various IoT devices. We worked together to establish a more secure, tamper resistant,and cost-effective solution by combining different blockchains. We conducted a cost study using real-world prices from three well-known blockchain networks. The results showed that our system significantly saves expenses and is thus appealing to businesses. Additional low-cost blockchain platforms, when they become available in the future, can be added to the system to strengthen its resilience to potential assaults.In addition we introduced Cyber-blockchain-based Trust's solution to recognise harmful behaviour and collect such activities evidences. Designing of more cost effective realistic IoT Forensic frame work can be taken as part of future work.

References

1. Ashton, K., 2009. That "internet of things" thing. RFID journal, 22(7), pp.97-114.
2. Bandyopadhyay, D. and Sen, J., 2011. Internet of things: Applications and challenges in technology and standardization. Wireless personal communications, 58(1), pp.49-69.
3. Li, S., Da Xu, L. and Zhao, S., 2015. The internet of things: a survey.Information Systems Frontiers, 17(2), pp.243-259

The Seybold Report
ISSN 1533-9211
DOI 10.5281/zenodo.7099360

4. EOS,https://eos.io/
5. Stellar, https://www.stellar.org/.
6. Baliga, A., 2017. Understanding blockchain consensus models. In Persistent
7. Wood, G., 2014. Ethereum: A secure decentralised generalised transaction ledger. Ethereum project yellow paper, 151(2014), pp.1-32.
8. C. J. Fung, O. Baysal, J. Zhang, I. Aib, and R. Boutaba, Trust management for host-based collaborative intrusion detection, in Managing Large-Scale Service Deployment, F. De Turck, W. Kellerer, and G. Kormentzas, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg,2008, pp. 109122.
9. T. D. Nguyen, S. Marchal, M. Miettinen, M. H. Dang, N. Asokan, and A. Sadeghi, Dot: A crowdsourced self-learning approach for detecting compromised iot devices, CoRR, vol. abs/1804.07474, 2018
10. H. Haddadi, V. Christophides, R. Teixeira, K. Cho, S. Suzuki, and A. Perrig, Siotome: An edge-isp collaborative architecture for iot security, in Proceedings of International Workshop on Security and Privacy for the Internet-of-Things (IoTSec) 2018. ETH Zu rich, 2018,1st International Workshop on Security and Privacy for the Internet-of- Things (IoTSec); Conference Location: Orlando, Florida, USA; Confer- ence Date: April17, 2018.
11. Danko, M., Mercan, S., Cebe, M. and Akkaya, K., 2019. Assuring the Integrity of Videos fromWireless-based IoT Devices using Blockchain.The Sixth National Workshop for REU Research in Networking and Systems, Monterey, CA, 2019.
12. Olufowobi, H., Engel, R., Baracaldo, N., Bathen, L.A.D., Tata, S. and Ludwig, H., 2016, October. Data provenance model for Internet of Things (IoT) systems. In International Conference on Service-Oriented Computing (pp. 85-91). Springer, Cham.
13. Polyzos, G.C. and Fotiou, N., 2017, August. Blockchain-assisted information distribution for the Internet of Things. In 2017 IEEE International Conference on Information Reuse and Integration (IRI) (pp. 75-78).IEEE.
14. Cebe, M., Erdin, E., Akkaya, K., Aksu, H. and Uluagac, S., 2018.Block4forensic: An integrated lightweight blockchain framework for forensics applications of connected vehicles. IEEE Communications Magazine, 56(10), pp.50-57
15. B. Gipp, J. Kosti, and C. Breitinger. Securing Video Integrity Using Decentralized Trusted Timestamping on the Bitcoin Blockchain. In MCIS, p. 51. 2016.
16. Tian, Z., Li, M., Qiu, M., Sun, Y. and Su, S., 2019. Block-DEF: a secure digital evidence framework using blockchain. Information Sciences, 491, pp.151-165.
17. Liang, X., Shetty, S., Tosh, D., Kamhoua, C., Kwiat, K. and Njilla, L.,2017, May. Provchain: A blockchain-based data provenance architecture in cloud environment with enhanced privacy and availability. In Proceed- ings of the 17th IEEE/ACM international
18. symposium on cluster, cloud and grid computing (pp. 468-477). IEEE Press.Cachin, C., 2016, July. Architecture of the hyperledger blockchain fabric.In Workshop on distributed cryptocurrencies and consensus ledgers (Vol.310, p. 4).