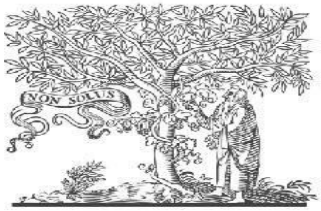
International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

COPY RIGHT



ELSEVIER SSRN

2021 IJIEMR. Personal use of this material is permitted. Permission from IJIEMR must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. No Reprint should be done to this paper, all copy right is authenticated to Paper Authors

IJIEMR Transactions, online available on 26th Nov 2021. Link :http://www.ijiemr.org/downloads.php?vol=Volume-10&issue=Issue 11

10.48047/IJIEMR/V10/ISSUE 11/83

Title *IMPLEMENTING SECURE AND EFFICIENT DISTRIBUTED NETWORK PROVENANCE FOR THE ON/OFF-BLOCKCHAIN PERFORMANCE*
Volume 10, ISSUE 11, Pages: 513-520

Paper Authors **PREETI GUPTA, DR. Dr. Rajeev yadav**

USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per UGC Guidelines We Are Providing A Electronic Bar Code

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

IMPLEMENTING SECURE AND EFFICIENT DISTRIBUTED NETWORK PROVENANCE FOR THE ON/OFF-BLOCKCHAIN PERFORMANCE

CANDIDATE NAME- PREETI GUPTA
DESIGNATION- RESEARCH SCHOLAR Glocal School of Computer science engineering,
THE GLOCAL UNIVERSITY, SAHARANPUR, UTTAR PRADESH
GUIDE NAME- DR. Dr. Rajeev yadav
DESIGNATION- PROFESSOR Glocal School of Computer science engineering,
THE GLOCAL UNIVERSITY, SAHARANPUR UTTAR PRADESH
PUBLICATION YEAR @ 2021

ABSTRACT
Utilizing blockchain technology, Secure and Efficient Distributed Network Provenance for IoT provides a reliable and dependable answer to the pressing problems of data integrity and provenance in IoT networks. In this piece, we discuss the need for SEDNP in the IoT and suggest a blockchain-based architecture to implement it. We present a unified provenance query model and design a provenance digest strategy that 1) allows compact (constant size) on-blockchain digests of provenance data and a multilevel index regardless of provenance data volume and 2) verifies the on-blockchain digests to guarantee the correctness and integrity of provenance query results. Additionally, we integrate a verifiable computation (VC) framework with a blockchain testing network and run extensive tests with this setup. In order to prove that SEDNP may be put to practical use, experimental data are offered as performance benchmarks.
Keywords: Provenance, Architecture, Efficiency, Algorithm, Benchmarks

I. INTRODUCTION

The Internet of Things (IoT) has revolutionized the way we interact with technology, permeating various sectors of our daily lives, ranging from smart homes and wearable devices to industrial automation and healthcare applications. The exponential growth of IoT devices and their ubiquitous presence has led to a massive influx of data being generated, exchanged, and processed within these networks. However, the increasing scale and complexity of IoT ecosystems have exposed them to a plethora of security and data integrity challenges. One of the most pressing issues faced by IoT networks is ensuring the authenticity and provenance of data exchanged between interconnected devices. Traditional centralized solutions often struggle to handle the massive data volumes and may introduce single points of failure, making them susceptible to malicious attacks and data manipulation.

IoT networks are characterized by their interconnectedness, where multiple devices communicate and share data with each other in real-time. However, in such a dynamic and heterogeneous environment, it becomes increasingly challenging to track the origin, history, and authenticity of data. Provenance, in this context, refers to the ability to establish the lineage and history of data, ensuring that it can be traced back to its original source, and any

modifications or transformations it underwent along its journey. Ensuring the integrity of provenance is vital in various IoT applications, such as supply chain management, healthcare systems, autonomous vehicles, and critical infrastructure monitoring, where the accuracy and reliability of data are paramount.

Blockchain technology, best known for its association with cryptocurrencies like Bitcoin, has garnered significant attention in recent years for its potential to revolutionize various industries beyond finance. The core features of blockchain, including decentralization, immutability, and transparency, make it an ideal candidate for solving the provenance challenges in IoT networks. By deploying a blockchain-based approach, the proposed system seeks to eliminate single points of failure, enhance data reliability, and provide an auditable and tamper-resistant ledger of events across the IoT ecosystem.

II. REVIEW OF LITERATURE

Xu, Guangquan et al., (2021) The IoV, or Internet of Vehicles, is a subset of the IoT. There are primarily two security issues with the IoV: (1) the IoV's central server may be underpowered to handle the centralized authentication of the exponentially growing number of linked cars, and (2) the IoV may not be resilient enough to single-node assaults. In order to address these issues, this research introduces SG-PBFT, a distributed blockchain-based PBFT consensus method for the Internet of Vehicles. The distributed architecture may lessen the load on the primary server and safeguard against assaults on individual nodes. By using a score grouping method, the SG-PBFT consensus algorithm is able to increase the efficiency of the conventional PBFT consensus process. We have shown via experimentation that our approach may significantly increase consensus efficiency and defend against single-node assaults. In particular, our approach's consensus time is only around 27% of what is needed by the state-of-the-art consensus algorithm (PBFT) when the number of consensus nodes exceeds 1000. Other situations requiring high consensus efficiency may benefit from our suggested SG-PBFT's adaptability.

Hu, Rui et al., (2020) As a prototypical example of cyberization, Internet of Things (IoT) facilitates the linking of real-world objects with the web to provide high-tech, data-driven benefits to business operations and individual lives alike. The variety, complexity, and dynamic nature of the Internet of Things, although exciting, also provide significant obstacles for IoT applications. In particular, identifying the origins of specified data makes it open to insertion attacks raised by a variety of parties at any stage of data transport or processing. Data provenance was developed as a solution to these concerns; it documents where data came from and the steps that were taken to generate and analyze it. Although various similar studies have been offered, a systematic analysis of data provenance in IoT is currently lacking in the literature. Following an examination of the characteristics of IoT data and applications, this article first proposes a set of guidelines for the design of provenance in IoT. We then use the criteria to conduct a comprehensive evaluation of the current schemes for IoT data provenance and to

analyze the merits and shortcomings of each. Finally, we highlight a number of research questions that need to be answered.

Singh, Saurabh et al., (2019) The development of the smart home has been bolstered by the rising need for a human-independent, pleasant living. In order to meet the needs of its inhabitants, the average highly intelligent house is equipped with many Internet of Things devices that generate processes and enormous amounts of data. With this growing need comes a slew of worries about the scalability, efficiency, and security of a smart home system. All of these problems are a pain to keep track of, and the current research don't provide enough detail to find solutions. Taking into account the current conundrum posed by the need for both security and efficiency, this essay provides a safe and efficient smart home design that combines blockchain and cloud computing technologies. Blockchain technology's distributed nature allows it to process acquired smart home user data and create a transactional copy of that data. Our suggested methodology analyzes network data and identifies the association between traffic aspects using multivariate correlation analysis to keep smart home networks safe. Using a variety of metrics, including throughput, we determined that blockchain is a viable security option for the incipient Internet of Things network and so justified its inclusion in our proposed design.

Siegel, Josh et al., (2017) Resource consumption and concerns about privacy and security slow the expansion of the Internet of Things (IoT). To accommodate future growth, a solution should be developed that simultaneously addresses safety, efficiency, privacy, and scalability. We offer a strategy based on human context and cognition that makes use of cloud computing to make IoT possible on low-resource gadgets. We introduce a framework that makes use of process knowledge to deliver anonymity and isolation via remote data fusion and abstraction-based security. We describe five components of the architecture and discuss the fundamental ideas behind the "Data Proxy" and the "Cognitive Layer." While the Cognitive Layer use these models to keep tabs on the system's development and to simulate the effects of orders before they are actually carried out, the Data Proxy utilizes them to digitally replicate things with minimum input data. With the help of the Data Proxy, sensors in a system may be sampled efficiently to reach a desired "Quality of Data" (QoD) goal. A vehicle tracking application is used to illustrate the efficiency gains made possible by this design. Finally, we think about the potential for this architecture to help remove technical, financial, and emotional hurdles to wider IoT adoption in the future.

Zhou, Wenchao et al., (2011) This paper presents secure network provenance (SNP), a unique approach that allows networked systems to provide explanations to their operators for the states in which they find themselves, such as the presence of a suspicious routing table entry on a certain router or the origin of a specific cache item. By helping operators locate broken or misbehaving nodes and evaluate the harm those nodes may have caused to the rest of the system, SNP enables

network forensics. The tamper-evident qualities of SNP make it possible for operators to discover when compromised nodes lie or falsely implicate correct nodes, making it ideal for use in adversarial scenarios. We also detail the development of SNooPy, a flexible SNP platform. We use SNooPy on three real-world examples: the Quagga BGP daemon, a declarative version of Chord, and Hadoop MapReduce, to show how versatile and useful it is. Our findings suggest that SNooPy can effectively explain state in an adversarial situation, is easy to implement, and has reasonable costs.

III. RESEARCH METHODOLOGY

We use a 2.30 GHz Intel Core laptop with 8 GB of RAM to run our VC framework. In specifically, we use the R1CS programming language to realize the Pinocchio library's Python interface and the libsnark library's C interface. Please take note that all function inputs will now have the "NIZK" enabled. C code for the function F is then translated into a QAP that may be programmed. The acquired QAP is used to build zk-SNARK over the alt-bn128 curve in the libff crypto library of libsnark.

We use the proof-of-authority consensus technique to build an Ethereum testing network with full functionality. The testing network consists of two authority nodes and many user nodes, all of which are hosted on the same laptop. The transactions are validated by the authority nodes, while the users' nodes manage the network. With the help of solidity, we model a provenance smart contract that saves provenance digests from several network administrators.

The SEDNP runs on the public Ethereum blockchain, which is managed by Ethereum miners, in production environments. To exchange blockchain addresses, network admins need just produce and make public unique addresses. Together, they agree on a provenance contract for storing and reading cryptographic provenance authenticators and negotiate the public specifications of the VC scheme.

IV. DATA ANALYSIS AND INTERPRETATION

First, we will examine the efficacy of F in Algorithm. The digest scheme is next analyzed in detail. We conclude with an analysis of the costs and benefits of the multilevel index technique and a demonstration of SEDNP's optimized on-chain storage overhead.

Evaluation of Function F

The index query function F over the index I combines a linear search over the second-level index IS with a keyword-based lookup over the first-level index IF. We begin with some reference points for linear queries over IS. After that, we provide our extensive assessments of the global query function F.

Where m is the number of provenance tuples and n is the number of dimensions in the query and index, Q is the vector containing the query and IS is the matrix containing the index. The total number of dimensions, n, is equal to the sum of the dimensions of the range values, n1, and the keyword values, n2, or n = n1+n2. In the experiment, we made the final n2 dimensions of Q and the subindex of vi, denoted by Ivi, "integers." We begin with a study of Q and each index Ivi in IS, contrasting their respective range values.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Then, we add the comparison result to RI as a similarity score, based on the last n2 dimensions of Ivi and Q. The time it takes to prepare a common reference string (CRS), generate a proof, and check that proof is how we quantify the computational cost.

Figures 1 and 2 show that the CRS setup time and prover cost both increase linearly with m and n. Input size has no effect on the verifier's time required to complete the process (0.027 s). Since comparison operations are substantially more costly than algebraic operations in a circuit, a greater n1 raises the QAP complexity and the setup/prover/verifier cost. In only a few seconds after an initial CRS setup, the costs of a prover and verifier become manageable.

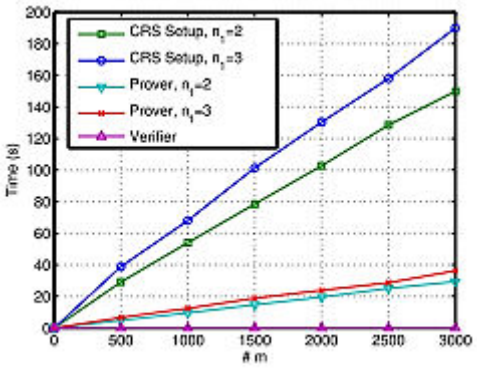


Figure 1: Processing time for CRS setup and prover cost m, n = 100.

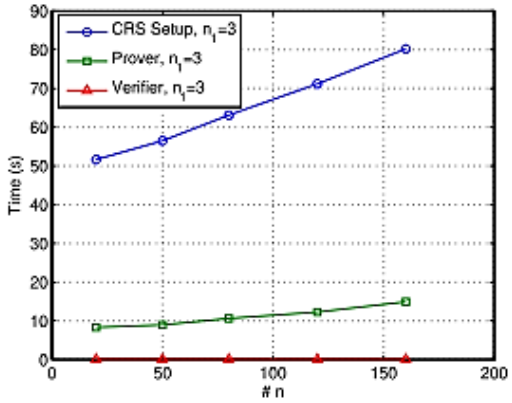


Figure 2: Processing time for CRS setup and prover cost n, m = 1000.

Storage expenses for the IS search are summed up in Table 1. In the experiment, Q and I are treated as hidden input such that the multiexponentiation components they are related with are not separated out. Libff provides efficient representations of group elements. The complexity of the created QAP is represented by the number of QAP variables. The storage cost exhibits the same linear growth with m, n. When compared to vk, ek is 10^3 times higher in size, whereas π_I size is unaltered.

Table 1: Storage Cost For I_S Search versus m, n = 100

# M (10^2)	ek (10^6 bits)		vk (10^3 bits)		# QAP Variables		π_I (bits)
	n_1 =2	n_1 =3	n_1 =2	n_1 =3	n_1 =2	n_1 =3	
5	305	386	162		167101	200101	2294
10	610	773	322		334101	400101	2294
15	915	1168	481		501101	600101	2294
20	1221	1547	641		668101	800101	2294

Evaluation of Digest Scheme

The index digest's storage and processing costs are summed together in Table 2. The cost of storing a group member denoted by |G|. Exponentiation in G_1/G_2 is denoted by the symbols E1/E2. Paring is represented by the letter "P." It is important to note that using the batch verification method, the ten may be grouped into four pairs. since n increases, the storage and computing cost for the prover goes up linearly with m, n, whereas the cost for the verifier

stays the same since n = n2 + n. The prover calculation cost is still manageable, though, since exponentiation operations in the alt-bn128 curve only take a few nanoseconds to complete. Take note that, in comparison to the libsnark implementation of F, the proof size and verification cost of the digest scheme are somewhat larger since it necessitates extra components mF, mS I with two more tests of their proper spans.

Table 2: Index Digest Cost Versus m, n

Component	**DK_E**	**DK_V**	**π_D**
Size	4mn' $\lvert\mathbb{G}_1\rvert$	3 $\lvert\mathbb{G}_2\rvert$	4 $\lvert\mathbb{G}_1\rvert$
Operation	Key generation	Prover	Verifier
Complexity	3mn'E_1 + 3E_2	2mn' E_1	4P

On/Off-Chain Tradeoff

An n2∗ m-dimensional IF and a mn-dimensional IS are used to instantiate the multilevel provenance index. The presence or absence of a keyword in the Ivi subindex is represented by the vector IF. To conduct a linear query over all of the subindexes Ivi that include a particular keyword k∗, we first identify all of the subindexes Ivi that contain k∗. The following considerations informed the design process.

- The cost of CRS production exceeds that of the prover and verifier phases. Therefore, CRS is produced once, and then instantiated with various Q and I for use in answering provenance queries..
- The magnitude of the reversed keyword subindex Using various search terms may alter Iki. In order for the C implementation of the Pinocchio VC framework to operate, the input size of the F must be specified so that the subscript for an array may be calculated. As a result, Iki is implemented as a constant vector of size n2.

Compact digest DL and DI minimizes the on-blockchain storage and processing costs. Off-blockchain processing and storage cost is growing, but improvement is still possible. This makes sense given the prohibitive nature of storing and computing on-chain compared to off-chain. One solution is to keep track of each Iki subindex as a separate digest on the distributed ledger. This doubles the amount of space needed on the blockchain if the index IF has n dimensions. We experiment with the reduction factor pr = |Ik∗|/m, another performance metric, to provide a more nuanced study of the tradeoffs between on- and off-blockchain performance. We use n = 100, m = 1000, and n1 = 2 to create the provenance index and modify pr to illustrate the performance disparity. Figures 3 and 4 show that the search space is decreased by a factor of pr, therefore a smaller pr results in cheaper off-chain storage and calculation costs. System designers are still given the freedom to optimize the on-chain vs off-chain tradeoff for certain provenance scenarios when using SEDNP.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

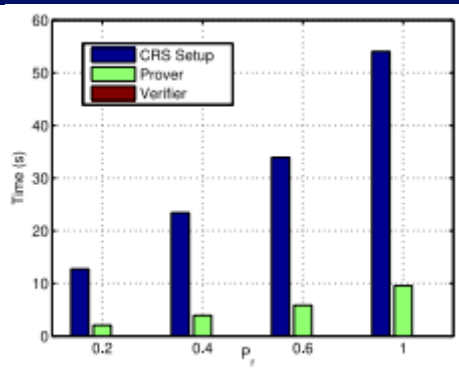


Figure 3: On/off-chain computation tradeoff

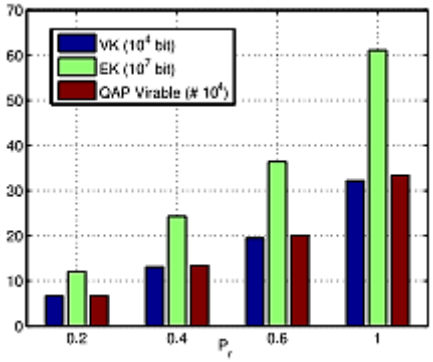


Figure 4: On/off-chain storage tradeoff

V. CONCLUSION

By exchanging the prohibitive on-blockchain storage and processing cost for manageable off-blockchain overheads, the proposed SEDNP has made the blockchain-based provenance architecture for the IoT feasible. This ensures the validity and integrity of query results while also minimizing the on-blockchain storage and processing cost for network provenance, regardless of the quantity of the provenance data. This study lays the groundwork for a more reliable and trustworthy Internet of Things via the use of a secure and efficient distributed network provenance mechanism. The suggested method uses blockchain technology to improve the trustworthiness of IoT networks by ensuring that data is consistently recorded and accessible. This study lays the groundwork for future developments in the Internet of Things, propelling breakthroughs in the safe administration of IoT data provenance.

REFERENCES: -

1. Xu, Guangquan & Liu, Yihua & Xing, Jun & Luo, Tao & Gu, Yonghao & Liu, Shaoying & Zheng, Xi & Vasilakos, Athanasios. (2021). SG-PBFT: a Secure and Highly Efficient Blockchain PBFT Consensus Algorithm for Internet of Vehicles.
2. Hu, Rui & Ding, Wenxiu & Yang, Laurence. (2020). A survey on data provenance in IoT. World Wide Web. 23. 10.1007/s11280-019-00746-1.
3. Liu, Dongxiao & Ni, Jianbing & Huang, Cheng & Lin, Xiaodong & Shen, Xuemin. (2020). Secure and Efficient Distributed Network Provenance for IoT: A Blockchain-Based Approach. IEEE Internet of

Things Journal. PP. 1-1. 10.1109/JIOT.2020.2988481.

4. Singh, Saurabh & Ra, In-ho & Meng, Weizhi & Kaur, Maninder & Cho, Gi. (2019). SH-BlockCC: A secure and efficient Internet of things smart home architecture based on cloud computing and blockchain technology. International Journal of Distributed Sensor Networks. 15. 155014771984415. 10.1177/1550147719844159.
5. Siddiqui, Muhammad Shoaib & Rahman, Atiqur & Nadeem Al Hassan, Adnan. (2019). Secure Data Provenance in IoT Network using Bloom Filters. Procedia Computer Science. 163. 190-197. 10.1016/j.procs.2019.12.100.
6. Elkhodr, Mahmoud & Mufti, Zuhaib. (2019). On the Challenges of Data Provenance in The Internet of Things. International Journal of Wireless & Mobile Networks. 11. 43-52. 10.5121/ijwmn.2019.11304.
7. Siegel, Josh & Kumar, Sumeet & Sarma, Sanjay. (2017). The Future Internet of Things: Secure, Efficient, and Model-Based. IEEE Internet of Things Journal. PP. 1-1. 10.1109/JIOT.2017.2755620.
8. Suhail, Sabah & Hong, Choon & Ahmad, Zuhaib & Zafar, Faheem & Khan, Abid. (2016). Introducing Secure Provenance in IoT: Requirements and Challenges. 39-46. 10.1109/SIoT.2016.011.
9. Zhou, Wenchao & Fei, Qiong & Narayan, Arjun & Haeberlen, Andreas & Loo, Boon & Sherr, Micah. (2011). Secure Network Provenance. 295-310. 10.1145/2043556.2043584.