Privacy-aware Distributed Ledger for Product Traceability in Supply Chain Environments

Sidra Aslam[1,2] - Michael Mrissa[1,2]*

[1]InnoRenew CoE, Livade 6, 6310 Izola, Slovenia
**{firstname.surname}@innorenew.eu*

[2]University of Primorska, Faculty of Mathematics, Natural Sciences and Information Technology, Glagoljaška ulica 8, 6000 Koper, Slovenia

Abstract

Wood supply chain stakeholders need traceability of individual products as well as protection from disclosure of their global activity (e.g. operation volumes), thus making data privacy a major concern. However, typical solutions to manage privacy-sensitive data are centralized and rely on third parties. Therefore, they suffer from single point of failure, trust and performance issues. Distributed ledger technology ensures data replication, immutability and availability, however, privacy-sensitive data remains publicly available. In this paper, we propose a framework design that combines distributed ledger technology with ring signature, mix networks and distributed hash table to manage data privacy. We illustrate the applicability of our solution with a product traceability scenario for the wood supply chain.

Keywords: distributed ledger, wood supply chain, privacy, security

Introduction

Supply chain management (SCM) has gained massive attention for both industry and academia [6]. Through the SCM process, all the stakeholders (producers, transporters, suppliers, customers, etc.) communicate with each other to trace products and increase overall quality. However, they generally have challenging requirements, as they need traceability and at the same time they need protection against information disclosure about their activity. Typical centralized solutions for product traceability are subject to the *single point of failure* problem that makes them vulnerable to attacks [6, 15, 16] and creates bottleneck that hinders scalability. Therefore, privacy-aware, decentralized solutions for product traceability are highly needed.

In this paper, we propose a framework based on Distributed Ledger Technology (DLT) to manage privacy-sensitive information for the Wood Supply Chain (WSC). DLT - and its most famous implementation, blockchain - mitigates the above issues by providing immutable decentralized storage to ensure data transparency and availability over the network [15]. DLT usage relates to a large number of application domains such as smart home, internet of things, supply chain, and finance, etc. It eliminates the need of a third party to establish trust as it records data in a data structure that is replicated on distributed nodes and where data fragments are related to each other with cryptographic techniques (hash functions) to guarantee integrity and immutability. Using DLT requires carefully looking at privacy concerns because all data and privacy-sensitive information over the network is publicly accessible, as proven by the amount of existing research papers [13, 4] and notes[1].

This paper is structured as follows: Section 2 highlights the need for a privacy solution to protect data along the wood production chain. Section 3 presents our contribution to support privacy aware distributed information management along the wood production chain. Section 4 reviews most relevant work in the area and shows how wood production chain currently lacks distributed privacy-aware solutions. The paper is concluded in Section 5.

Motivating Scenario and Research Problem

In this section, we introduce a simple furniture production scenario that motivates the need for DLT and highlights our research challenges. As presented in Fig. 1, we have identified 6 actors that participate to a WSC. 1) Wood cutting company identifies and cuts specific trees. 2) Transport company drives wood logs to storage warehouse. 3) Storage warehouse company sorts, processes and stores logs. 4) Furniture assembly company cuts logs and assembles furniture. 5) Furniture shop company stores and exposes assembled furniture for sale. 6) Customer buys wooden furniture and verifies product origin.

This scenario highlights the need for traceability, trust and anonymity in the supply chain. Indeed, all the actors of the chain want to provide full product traceability to the customers. A typical solution to this issue consists in inserting RFID chips in the trees and logs produced. RFID chips are detected at every stage of the WSC to provide stakeholders with required traceability. At this stage, typical third-party centralized solutions to store RFID data form a Single Point Of Failure (SPOF), thus motivating research towards decentralized solutions. In particular, blockchain is a

[1] https://pdfs.semanticscholar.org/549e/7f042fe0aa979d95348f0e04939b2b451f18.pdf

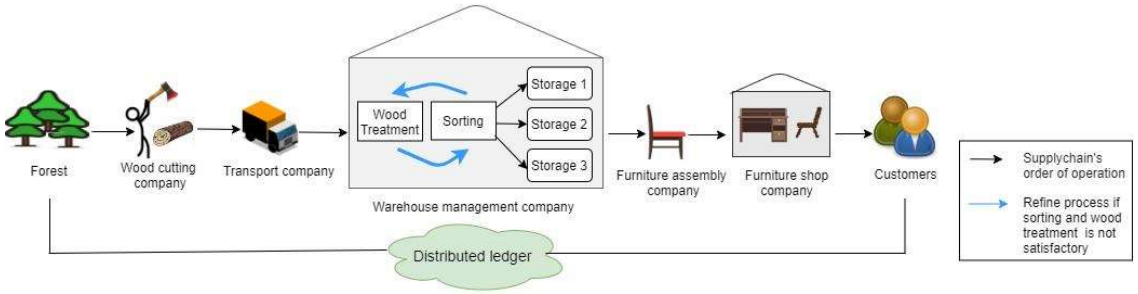

Figure 1: Workflow of wood supply chain system

type of DLT where transactions between users are stored in blocks and replicated on all participating nodes, thus avoiding SPOF. Typically, users generate public and private keys and use public keys as addresses to identify themselves in transactions. Miners use techniques such as proof of work or proof of stake [11] to add and validate blocks. In order to prevent modification of recorded transactions, blocks are linked to each others. Each block contains a hash of the previous block (except the first block in a chain called "genesis") [9].

However, our scenario also highlights that, although the actors of the WSC are willing to provide individual customers with traceability, they also do not wish to have their production information publicly accessible. There is a need to develop solutions that guarantee that it is not possible to access all product data and, for instance, draw statistics from it. Therefore, based on this scenario, our solution must provide 1) confidentiality to protect data from unauthorized access, 2) integrity to verify the originality of content, 3) availability to make sure that data should be always accessible upon request, 4) anonymity to protect the relationship between user identity and data content (unauthorized users must be unable to link data with their owner) and 5) scalability (the proposed solution must manage large number of users and transaction data with reasonable performance).

Such requirements highlight our motivation to combine blockchain with other solutions to provide anonymity and unlinkability of data on decentralized storage and at the same time enable traceability for customers over specific data items. In the following, we formulate a list of scientific locks (SL) that rise from this scenario and the requirements of the actors described above:

- SL1: Unlinkability between data and user's identity. In our context, protecting privacy means SL1a) protecting access to users' identities available in transactions, and SL1b) preventing disclosure of users' identities that could be found out by exploiting temporal dependencies between transactions. The difficulty is to find a solution that integrates smoothly with blockchain and does not affect its original operation.

- SL2: Management of data updates. In our motivating scenario, actors need to update information at each point of the chain (for example product location). The difficulty is to find a way to circumvent original blockchain design to allow data updates.

- SL3: Data security and fine-grained user access. Data stored on DL must be protected from unauthorized access with encryption mechanisms and users' actions on data must be

managed through access control models. The difficulty is to provide decentralized solutions that integrate smoothly with blockchain.

- SL4: Acceptable usability and response time. System design need to be usable with low response time. However, high response time is a major issue for blockchain. Answering previous scientific locks will increase the system complexity, the difficulty is to integrate solutions while maintaining low computational cost.

To answer these scientific locks, we present in the following a solution that integrates relevant technologies in a single framework. We show how the framework integrates those different aspects to answer the aforementioned scientific locks and we provide insight on the relevance of our framework with the help of our motivating scenario.

Conceptual Framework

In this paper, we propose a privacy-aware decentralized information management framework to support WSCs (or any distributed information system). This framework combines blockchain technology with ring signature, mix network and distributed hash table as shown in Fig. 2.

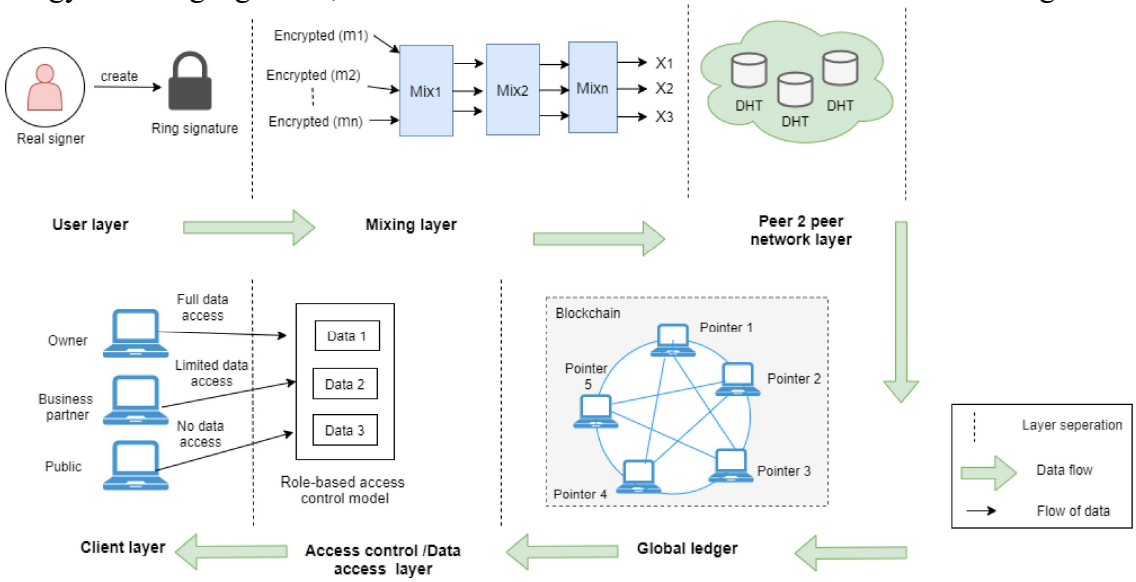

Figure 2: Overview of our conceptual framework

First, our framework answers SL1a using ring signature, that allows a group of several actors equipped with public/private keys (called a ring) to jointly sign a message. Messages going out of a ring can be verified to prove they have been emitted within the ring and at the same time guarantee anonymity of the emitter's identity.

Second, our framework makes use of mix network to prevent user' identity disclosure through temporal dependency obfuscation, thus answering SL1b. Indeed, mix networks randomize temporal relationships between messages, thus making it impossible to observe time dependency between messages and draw conclusion on the identity of message emitter.

Third, we address SL2 by jointly using blockchain and DHT. Due to its design, DHT addresses performance and data update issues related to blockchain. Therefore, we use DHT to store actual data, and blockchain to store pointers to the data in the DHT. This solution offers both the benefits of blockchain immutability and DHT performance and possibility for update. Indeed, data can be modified in the DHT but pointers cannot be deleted or modified because of blockchain immutability.

SL3 is resolved using role-based access control. In order to preserve the decentralized nature of our framework we adopt Decentralized Public Key Infrastructure (DPKI) [1] to enable role-based access control.

Fourth, we address SL4 by using DHT to store and manage data. Combining DHT with blockchain greatly improves performance as demonstrated in [10].

In our WSC scenario, actors such as the wood cutting company will store data through ring signature, then mix network, so that we will not be able to identify who put the information into the DHT and pointers to it in the blockchain. The rules for role-based access control will be recorded so that different levels of access are granted to different actors.

In our example, general access is not possible, however, business partners can access limited data (such as knowledge of what kind of tree species was cut) and data owner (company employees) have full access. Location data can be updated as the wood follows its progress in the supply chain.

Related Work

In this section, we review the most relevant existing research on privacy management for blockchain. In [12], the authors present a blockchain-based distributed platform for anonymized trading of datasets. The work is decentralized except a third party to deal with privacy policy management. An dynamic identity management system based on the bitcoin blockchain is proposed in [2] that also relies on third party. It ensures users' data confidentiality and allows a user to control their identities over public networks. In [5] identity management in blockchain is addressed using zero knowledge protocol and smart contracts. However, the work cannot handle malicious verifiers and thus other cryptographic schemes need to be combined with it to achieve better security.

In [8] privacy issues in blockchain-based IoT are explored. Ring signature is used to provide anonymity on the healthcare blockchain. However, DoS and modification attack is still possible because an attacker can make services unavailable for a user and inject falsified transactions in a network. In 2015, [17] presented a personal data management system using blockchain technology that combines blockchain and off-blockchain storage to provide privacy and allow users to control their data. However, it is not efficient to manage and query large volume of data.

In 2018, [7] described identity management on the blockchain by using Sovrin, uPort, and ShoCard. It allows users to control their transactions and enhance the data's transparency due to storage of data on distributed nodes. The authors in [3] propose blockchain-based medical data management system that allows patients to access their medical data across different providers and patient databases. This system is helpful to manage authentication, confidentiality, and data sharing. In paper [15] authors discussed blockchain technology in a supply chain system. Smart contracts are used to secure transactions between parties and remove the middle party. The major drawback of the proposed scheme is scalability. A decentralized supply chain system using smart contract is presented in [16]. The proposed system is useful to trace goods and recipe ingredients through supply chain system. Moreover, it is unable to trace damaged or loss goods during supply chain process. In [14], authors provided a survey of cryptographic techniques to resolve immutability

issue of blockchain. Similarly, they discussed advantages and limitations of these techniques when used in either public or private blockchain.

Table 1 summarizes advantages and limitations of existing work and relates to our identified scientific locks. We summarize which aspects are addressed in Table 2.

Table 1: Summary of related work analysis

Ref.No	Challenge Addressed	Approach	Advantages	Limitations
[12]	Anonymized dataset (SL1b)	Blockchain based distribution scheme	Decentralization; Data owner can trace their data	Need third party to manage privacy policy; Privacy risk remain
[2]	Ensure identity in bitcoin blockchain (SL1a)	Zero-knowledge proofs	Dynamic update of identities	Need third party
[5]	Anonymity of user data over blockchain (SL1b)	Zero-knowledge protocol and Smart contract	Provide data privacy; Eliminate need of third party	Require more security schemes; Unable to work when verifier is malicious
[8]	Secure management of healthcare data on blockchain (SL1a) (SL1b) (SL3)	Ring signature, Smart contract, and Cryptographic techniques	Decrease blockchain bandwidth and computational power	DOS attack is possible; Scalability
[17]	Decentralized user private data management (SL3) (SL4)	Fine-grained access control, Encryption, and Decryption schemes	Remove third party; User has control over their personal data	Framework cannot manage large volume of data
[7]	Digital Identity management on the blockchain (SL1a)	Sovrin, Uport and ShoCard	Completely decentralized transparency; User can control data transactions	Lack of usability to manage cryptographic keys; Unable to cope with data transparency
[3]	Blockchain based medical data management system (SL3)	Smart contract, Proof of work, and Syncing algorithm	Patients can access their medical record; Authentication, Avoid single point of failure	Cannot maintain large volume of data; centralized security

Conclusion

In this paper, we illustrate the need for privacy-aware product traceability in wood supply chains with a scenario that shows the challenges of this specific application domain, before identifying related scientific locks. We then propose a framework that relies on a combination of ring signature, mix network, blockchain and DHT to provide adapted measures that allow traceability for authorized users without disclosing all the business information of the different actors.

Further work includes exploring different options for decentralized authentication, assessing the solution performance in real or simulated environment, and deploying a proof-of-concept prototype in collaboration with an existing company.

Table 2: Comparative analysis of our proposed framework

Ref. no/Year	Privacy features		Security properties	Traceability
	Identity protection	Transaction data protection		
[1, 2017]	Yes	No	Confidentiality	No
[2, 2016]	Yes	Yes	Confidentiality	No
[5, 2019]	Yes	No	Confidentiality	No
[7, 2018]	Yes	Yes	Integrity	No
[8, 2018]	Yes	No	Confidentiality; Integrity; Availability	No
[9, 2019]	Yes	Yes	Confidentiality; Integrity;	Yes
[14, 2019]	No	No	No	Yes
[15, 2018]	No	No	No	Yes
[16, 2015]	Yes	Yes	Availability	No
Our framework	Yes	Yes	Confidentiality; Integrity; Availability	Yes

References

[1] Karl Aberer, Anwitaman Datta, and Manfred Hauswirth. A decentralized public key infrastructure for customer-to-customer e-commerce. *International Journal of Business Process Integration and Management*, 1(ARTICLE):26–33, 2005.

[2] Daniel Augot, Hervé Chabanne, Olivier Clémot, and William George. Transforming face-toface identity proofing into anonymous digital identity using the bitcoin blockchain. In *2017 15th Annual Conference on Privacy, Security and Trust (PST)*, pages 25–2509. IEEE, 2017.

[3] Asaph Azaria, Ariel Ekblaw, Thiago Vieira, and Andrew Lippman. Medrec: Using blockchain for medical data access and permission management. In *2016 2nd International Conference on Open and Big Data (OBD)*, pages 25–30. IEEE, 2016.

[4] Alex Biryukov, Dmitry Khovratovich, and Ivan Pustogarov. Deanonymisation of clients in bitcoin p2p network. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, pages 15–29, 2014.

[5] Yogita Borse, Anushka Chawathe, Deepti Patole, and Purnima Ahirao. Anonymity: A secure identity management using smart contracts. *Available at SSRN 3352370*, 2019.

[6] Fang Dong, Pengcheng Zhou, Zijian Liu, Dian Shen, Zhuqing Xu, and Junzhou Luo. Towards a fast and secure design for enterprise-oriented cloud storage systems. *Concurrency and Computation: Practice and Experience*, 29(19):e4177, 2017.

[7] Paul Dunphy and Fabien AP Petitcolas. A first look at identity management schemes on the blockchain. *IEEE Security & Privacy*, 16(4):20–29, 2018.

[8] Ashutosh Dhar Dwivedi, Gautam Srivastava, Shalini Dhar, and Rajani Singh. A decentralized privacy-preserving healthcare blockchain for iot. *Sensors*, 19(2):326, 2019.

[9] Kristoffer Francisco and David Swanson. The supply chain has no clothes: Technology adoption of blockchain for supply chain transparency. *Logistics*, 2(1):2, 2018.

[10] Yahya Hassanzadeh-Nazarabadi, Alptekin Küpçü, and Öznur Özkasap. Lightchain: A dhtbased blockchain for resource constrained environments. *CoRR*, abs/1904.00375, 2019.

[11] Sungmin Kim and Joongheon Kim. Poster: Mining with proof-of-probability in blockchain. In *Proceedings of the 2018 on Asia Conference on Computer and Communications Security*, pages 841–843, 2018.

[12] Shinsaku Kiyomoto, Mohammad Shahriar Rahman, and Anirban Basu. On blockchain-based anonymized dataset distribution platform. In *2017 IEEE 15th International Conference on Software Engineering Research, Management and Applications (SERA)*, pages 85–92. IEEE, 2017.

[13] Malte Moser. Anonymity of bitcoin transactions, 2013.

[14] Eugenia Politou, Fran Casino, Efthymios Alepis, and Constantinos Patsakis. Blockchain mutability: Challenges and proposed solutions. *IEEE Transactions on Emerging Topics in Computing*, 2019.

[15] Sara Saberi, Mahtab Kouhizadeh, Joseph Sarkis, and Lejia Shen. Blockchain technology and its relationships to sustainable supply chain management. *International Journal of Production Research*, 57(7):2117–2135, 2019.

[16] Martin Westerkamp, Friedhelm Victor, and Axel Küpper. Blockchain-based supply chain traceability: Token recipes model manufacturing processes. In *2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, pages 1595–1602. IEEE, 2018.

[17] Guy Zyskind, Oz Nathan, et al. Decentralizing privacy: Using blockchain to protect personal data. In *2015 IEEE Security and Privacy Workshops*, pages 180–184. IEEE, 2015.