LEDGER
ledgerjournal.org

ISSN 2379-5980 (online)
associated article DOI 10.5195/LEDGER.2021.239

A Decentralized Identity-Based Blockchain Solution for Privacy-Preserving Licensing of Individual-Controlled Data to Prevent Unauthorized Secondary Data Usage: Open Review

Meng Kang,[*] Victoria L. Lemieux[†]

Reviewers: Reviewer A, Reviewer B

Abstract. The final version of the paper "A Decentralized Identity-Based Blockchain Solution for Privacy-Preserving Licensing of Individual-Controlled Data to Prevent Unauthorized Secondary Data Usage" can be found in Ledger Vol. 6 (2021) 126-151, DOI 10.5195/LEDGER.2021.239. There were two reviewers involved in the review process, neither of whom has requested to waive their anonymity at present, and are thus listed as Reviewers A and B. After initial review by Reviewers A and B, the submission was returned to the authors with feedback for revision (1A). The authors responded (1B) and resubmitted their work. After subsequent evaluation by Reviewer B (2A), revisions made were deemed sufficient to address any concerns, thus ending the peer review process. Author responses have been bulleted for reader clarity.

1A. Review

Reviewer A

Does this paper represent a novel contribution to cryptocurrency or blockchain scholarship?

Yes

[*] M. Kang (meng.kang@ubc.ca) is a master's degree student at the School of Engineering at the University of British Columbia, Kelowna, BC, Canada.
[†] V. L. Lemieux (v.lemieux@ubc.ca) is an Associate Professor of Archival Science at the School of Information at the University of British Columbia, Vancouver, BC, Canada and Founder of and Co-Lead of Blockchain@UBC.

If you answered "yes" to the previous question, in one sentence, describe in your own words the novel contribution made by this paper:

A novel way of thinking about IP protection and data sharing through the use of Blockchain and SSI

Is the research framed within its scholarly context and does the paper cite appropriate prior works?

Yes

Please assess the article's level of academic rigor.

Excellent (terms are well defined, proofs/derivations are included for theoretical work, statistical tests are included for empirical studies, etc.)

Please assess the article's quality of presentation.

Excellent (the motivation for the work is clear, the prose is fluid and correct grammar is used, the main ideas are communicated concisely, and highly-technical details are relegated to appendixes).

How does the quality of this paper compare to other papers in this field?

Top 5%

Please provide your free-form review for the author in this section.

I rarely have an opportunity to read such a well written and interesting paper. The ideas are solid and well presented. I am not sure that DRMs will ever work, but this piece is as close to convincing me as it can get. Seriously interesting read!

Reviewer B

Does this paper represent a novel contribution to cryptocurrency or blockchain scholarship?

Yes

If you answered "yes" to the previous question, in one sentence, describe in your own words the novel contribution made by this paper:

The paper is one of the firsts addressing the important field of combining SSI and Confidential Computing. Since SSI benefits a lot from ledger technology sufficient connection to the topic of the journal is given.

ISSN 2379-5980 (online)
associated article DOI
10.5195/LEDGER.2021.239

Is the research framed within its scholarly context and does the paper cite appropriate prior works?

Yes

Please assess the article's level of academic rigor.

Good (not excellent but a long way from poor)

Please assess the article's quality of presentation.

Good (not excellent but a long way from poor)

How does the quality of this paper compare to other papers in this field?

Top 50%

Please provide your free-form review for the author in this section.

The paper “A Decentralized Identity-Based Blockchain Solution for Privacy-Preserving Licensing of Individual-Owned Data to Prevent Unauthorized Secondary Data Usage” is structured fairly well. The idea of the paper is to combine SSI with Confidential Computing (despite this category term is not used in the paper) for usage control purposes. This idea is significant and at the frontier of exactly the research, which is necessary for digital sovereignty. Also a connection to ledger technology is present through the SSI component. Despite this inspiring and important idea of the paper, substantial revisions of the paper are required before publication is recommended.

The novelty claim of the paper is in the field of the overarching system architecture to enable real usage control. However, the fundamental challenge, which is claimed to be solved, is not well formulated and also the solution to it is hidden below the technicalities of the chosen demo-implementation.

To understand the value of the paper the reader needs already to have both an overview in the field of identity management, confidential computing and advanced usage control as well as a deep technical understanding of the implementing technologies. This is a small circle of people only, and definitely a much too small fraction of the readers of the Ledger Journal only.

The reader is not able to gain the overview and insight from the paper. No overview of alternative architectures, nor a genesis or derivation of the chosen architecture is given. The term Confidenital Computing is not even mentioned, i.e. the category of computing techniques, where privileged access to data being processed is technically prevented, is not treated as a category. Merely homomorphic encryption is chosen and it appears to the reader that other techniques available to provide confidential computing are either not known to the authors or they remain unmentioned for unclear reasons (processor TEEs, multi-server TEEs, combinations of them, etc.).

The practical performance limitations of fully homomorphic encryption are not mentioned to the reader, nor it is discussed, whether these potentially could be solved in future (e.g. through dedicated processor hardware).

ISSN 2379-5980 (online)
associated article DOI
10.5195/LEDGER.2021.239

Revisions should be made to improve the requirements analysis and the architecture derivation. Also a comparison of the chosen setup with potential alternatives should be elaborated.

In the following some minor comments:

In Abstract Line 7 and Line 70 and many other places in the paper:
“owner of data”: Since the European GDPR and its understanding of privacy is referenced, it should be considered, that the concept of “ownership” generally does not apply to personal data. The right to use them cannot be sold and cannot be purchased. Unlike with the concept of ownership the granted right to use personal data can be withdrawn by the data subject anytime disregarding any trading. Please use in the privacy context a more precise terminology.
Line 11: platform economics can exist – when privacy preserving technology is used – independent from data abuses
Line 47 with FHE and other techniques of confidential computing the cloud usage is not limited to encrypted “storing” but also to privacy preserving “processing”
Line 113: Another unfortunate use of “owner” -> “holder”
Line 204/205: Difficulty with efficiency of FHE not made transparent to reader
Line 374: Should this be one sentence? “For ease of presenting the design. The architecture is divided into two parts.” -> For ease of presenting the design, the architecture is divided into two parts.
Line 378: Figure 7 shows the architecture of data sharing and storage
Line 434: developed
sorry for not having listed all typos

1B. Author Responses

Reviewer B

To understand the value of the paper the reader needs already to have both an overview in the **field of identity management**, **confidential computing** and **advanced usage control** as well as a **deep technical understanding of the implementing technologies**. This is a small circle of people only, and definitely a much too small fraction of the readers of the Ledger Journal only.

- We have added additional introductory information about all of these fields to further contextualize our proposed solution architecture.

-Add to SSI section by contextualizing it within the field of identity management and suggesting that it is the last evaluation of identity management.

- We have contextualized SSI within the field of identity management as the latest evolution of that field.

ISSN 2379-5980 (online)
associated article DOI
10.5195/LEDGER.2021.239

-Add confidential computing to the background literature, and contextualize homomorphic encryption as one aspect of confidential computing.

- We have added discussion of confidential computing to the background literature and contextualized homomorphic encryption as one aspect of confidential computing in section 2.3.

The reader is not able to gain the overview and insight from the paper. No overview of alternative architectures, nor a genesis or derivation of the chosen architecture is given.

- We have added additional information to highlight that we derived our solution architecture from a previous on PREM-DRM (see Gaber, T., Ahmed, A., Mostafa, A. "PrivDRM: a privacy-preserving secure Digital Right Management System." In *Proceedings of the Evaluation and Assessment in Software Engineering* 481–486 (2020)). We argue that no existing architectures combine all three areas that we do in this paper (i.e, SSI, confidential computing and advanced usage control) to achieve our objectives. We acknowledge that it is possible that there could be alternate architectures to the one we have chosen. Though we have not gone into to great depth about alternative architectures to keep the paper to a reasonable length, we have taken care to explain our design choices.

The term **Confidenital Computing** is not even mentioned, i.e. the category of computing techniques, where privileged access to data being processed is technically prevented, is not treated as a category. Merely homomorphic encryption is chosen and it appears to the reader that other techniques available to provide confidential computing are either not known to the authors or they remain unmentioned for unclear reasons (processor TEEs, multi-server TEEs, combinations of them, etc.).

- We have added confidential computing to the keywords defining out paper
- We have added a discussion of confidential computing to the introduction and section 2.3.

The practical performance limitations of fully homomorphic encryption are not mentioned to the reader, nor it is discussed, whether these potentially could be solved in future (e.g. through dedicated processor hardware).

- We have added mention of the practical limitations of FHE and also discussed the possibility of using dedicated processor hardware (i.e., TEE). We argue that if the confidential computing is done through e.g., an SGX hardware device, since it has to go to a third party's (e.g., Intel's) servers for remote authentication, it requires that all the parties involved trust the third party. Therefore TEE-based privacy computing is not suitable for some scenarios with higher privacy requirements. We reserve fuller exploration of hardware based solutions for future work.

Revisions should be made to improve the requirements analysis and the architecture derivation. Also a comparison of the chosen setup with potential alternatives should be elaborated.

ISSN 2379-5980 (online)
associated article DOI
10.5195/LEDGER.2021.239

- To address this comment, we have added a requirements section 4.1. Due to the length of the paper, we have chosen not to include extensive discussion of potential alternative setups.

2A. Second Round Review

Reviewer B

Did you review an earlier version of this submission? (If "no," please contact the editor.)

Yes

Has the submission been sufficiently revised to address your previous concerns?

No

If you answered "no" to the previous question, please provide more detailed feedback here.

The main concerns have been addressed.

Except that the use of the term "data owner" was not corrected. The use of the term is widespread amongst legally untaught technical scientists. However, the paper explicitly refers to the GDPR, in which term owner or ownership does not appear at all for the reason that personal data cannot be “owned”. The correct term is “data subject”, who is the holder of rights concerning the personal data of him or her. It is a question to the editor, whether such lack of precision is acceptable in a technical paper with reference to legal requirements.

Regarding processing efficiency of FHE: The average reader of Ledger, does not know that FHE still needs several orders of magnitude higher computing performance than other confidential computing implementations. The authors should mention this to not mislead the reader regarding the proposed solution. They could do that for example in Line 247, by clarifying what they mean with “realistic efficiency”.

Do you have any new concerns specific to this revision?

Yes

If you answered "yes" to the previous question, please provide more detailed feedback here.

Through the introduction of the subsection 4.1_Requirements, also following inconsistency has been introduced:

In Line 408 “Data Integrity” is demanded. However, in the brief discussion of the implementation options for confidential computing in the Lines 198-230, FHE is positioned as clearly superior over TEE from a security point of view. However, FHE has no feature to

ISSN 2379-5980 (online)
associated article DOI
10.5195/LEDGER.2021.239

guarantee the confidentiality and integrity of processing code and therefore cannot guarantee data integrity during processing. The reader should not be misled by a one-sided positioning of FHE. Probably a combination of available techniques is needed to reach the ultra-high security requirements defined in this paper. The weaknesses of FHE should not be concealed to the reader, who most of them are not familiar with FHE.

‘

ISSN 2379-5980 (online)
associated article DOI
10.5195/LEDGER.2021.239