Proof of Idea

Gajendra Jung Katuwal
gajendra.katuwal@oso.network
Open Science Organization

v 0.0

September 29 2018

Abstract

Proof of Idea (PoI) is an algorithm to maintain a list of *ideas* of desired attribute in a decentralized fashion. Here, idea means an abstraction of any intellectual contribution (e.g. publication, code, data, etc.). In this document, we present a basic version of PoI and demonstrate how it can be used to create a network of validated ideas G_I. And we present how G_I can be the base idea layer of an open and community-owned decentralized knowledge dissemination platform that incentivizes the generation of quality ideas. PoI can potentially inverse the current pay-to-publish publishing model in research to publish-to-pay.

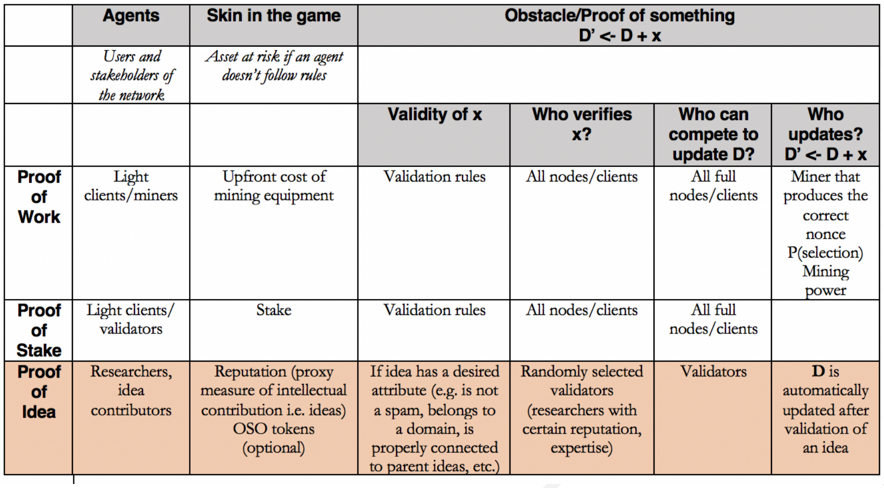
We have decided to make this paper and project public at this early stage in hope that we will receive valuable feedback for improvement.

[1]

1 Introduction

The research dissemination is conceptually a process to curate a list of scientific ideas (publications) with desired attributes. Loosely speaking, it is a quality control (QC) process. The QC process consists of two fundamental steps: 1) Make sure if an idea is not a spam (e.g. complete junk to spam the network) 2) Assign suitable attributes such as research domain, novelty, usefulness, readabiliy, etc. to it. The preprint servers such as https://arxiv.org/ and https://www.biorxiv.org/ serve the first step and the journals and conference proceedings serve the second step. For example, https://arxiv.org/ is a list of scientific ideas which are generally preprints and are available in pdf format. Within https://arxiv.org/, there can be sub-lists of ideas according to the subject domains. Similarly, https://www.biorxiv.org/ is another list of ideas which are generally preprints and are related to life science research. So, https://arxiv.org/ and https://www.biorxiv.org/ are the lists of new ideas which have been vetted to ensure that they are not junk but have not been properly validated by the experts. Similarly, Journal of Machine Learning Research http://www.jmlr.org/ is a curated list of certain attributes – "supposedly" novel and influential ideas in machine learning. The vetting process which ensures that only the ideas with desired attribute are appended to the list is called *peer-review.* During peer-review, peers (researchers conducting similar research) evaluate the idea (publication) separately and then come to a consensus to admit or reject the idea from being appended to the list. The peer-review process is usually cyclic (there is a feedback loop between the author and reviewers) and is facilitated by an editor.

[1] **Acknowledgments.** My conversations with Kirk Gosik and Noman Land were very helpful and have significantly shaped this document. Special Thanks to Noman for developing a PoC of the proof of idea in Solidity.



	Agents	Skin in the game	Obstacle/Proof of something D’ <- D + x			
	Users and stakeholders of the network	*Asset at risk if an agent doesn’t follow rules*				
			Validity of x	Who verifies x?	Who can compete to update D?	Who updates? D’ <- D + x
Proof of Work	Light clients/miners	Upfront cost of mining equipment	Validation rules	All nodes/clients	All full nodes/clients	Miner that produces the correct nonce P(selection) Mining power
Proof of Stake	Light clients/ validators	Stake	Validation rules	All nodes/clients	All full nodes/clients	
Proof of Idea	Researchers, idea contributors	Reputation (proxy measure of intellectual contribution i.e. ideas) OSO tokens (optional)	If idea has a desired attribute (e.g. is not a spam, belongs to a domain, is properly connected to parent ideas, etc.)	Randomly selected validators (researchers with certain reputation, expertise)	Validators	**D** is automatically updated after validation of an idea

Figure 1: **Comparison of Proof of Idea with popular consensus algorithms.**

Here at Open Science Organization (OSO) [1, 2], we believe that this knowledge dissemination process can be significantly improved if it happens in a web3-based open-source platform. The desirable features of the supposed platform are:

- Open source
- Decentralized (web3 based) and community-managed
- Intellectual contribution (idea) is the sole asset or resources required to collect the asset of the platform
- No pre-mined assets
- Reward of assets is proportional to the quality and importance of the intellectual contribution. This topic is important but highly controversial. The best way to move forward would be to design an efficient system where market can figure this out by itself.

Proof of Idea

Proof of Idea (PoI) was developed to achieve the above-mentioned desirable properties. PoI is an algorithm to curate *ideas* of desired attribute in a decentralized fashion. In PoI, a proof of an intellectual contribution (idea) is required to update the network (e.g. adding new idea, claiming assets in the network). **Note:** *I have made the initial design as simple as possible to help us get started. With this naive model, there will be several attack vectors which we will have to state explicitly and apply measures to discourage them.* See Table 1 for comparison of PoI with proof of work and proof of stake.

Figure 3 shows the conceptual layers of the envisioned knowledge dissemination platform based on PoI. Each blue layer uses some form of PoI and the sub-networks (similar to sub-reddits) allow to establish local rules such as particular version of PoI algorithm, review process, reward sharing, etc., within the main network. Layer 1 acts as the base idea layer containing the list of validated ideas. PoI in layer 1 basically checks if the submitted idea is a spam or

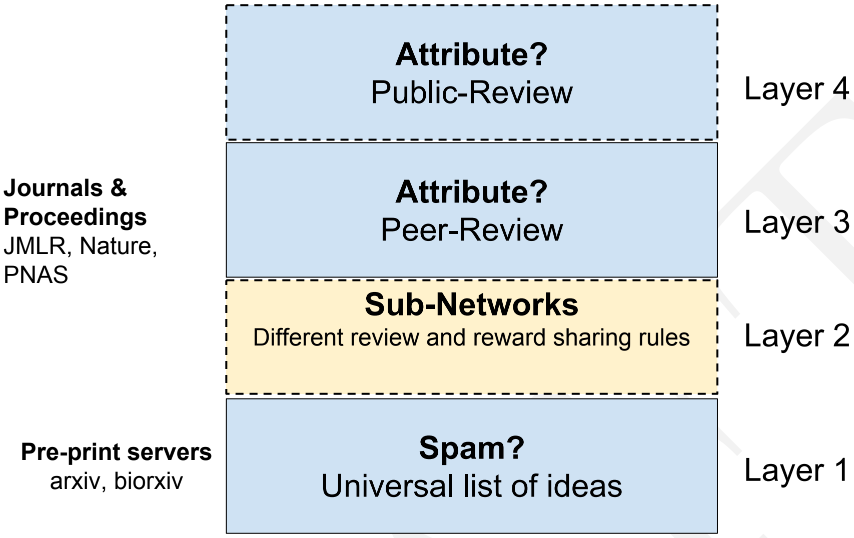


Figure 2: **Conceptual Layers of the Knowledge dissemination platform based on Proof of Idea (PoI).**

not (similar to verifying if a Bitcoin [3] transaction is valid). PoI in layer 2 facilitates the peer-review and PoI in layer 3 encourages perpetual validation of ideas (publications) i.e. review after the initial peer-review.

2 Algorithm Description

Below I will explain the PoI of layer 1 which basically acts as a spma filter and ensures the new idea is properly connected to its parent ideas; see Figure 3.

2.1 List of verified users

Let U be the list of verified users. Each user is represented by a unique public-private key pair where private key is owned by the user. U will be like a token curated registry (TCR). Mapping between the public key and real world identity can be performed by mutual verification among researchers. This is a separate work, so I will not discuss further in this document.

2.2 Submission of an idea

An author submits an idea I and associated domain D of the idea. (e.g. blockchain, machine learning, sociology, etc.).

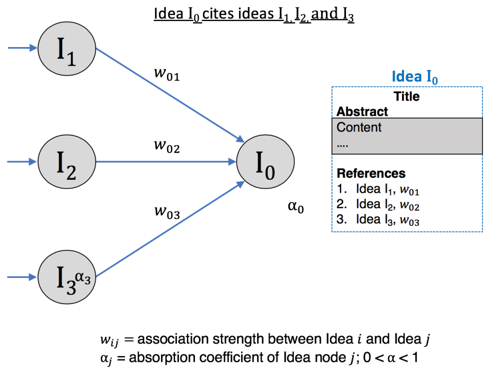

Idea I_0 cites ideas I_1, I_2, and I_3
I_1
I_2
I_3
α_3
I_0
α_0
w_{01}
w_{02}
w_{03}
Idea I_0
Title
Abstract
Content
....
References
1. Idea I_1, w_{01}
2. Idea I_2, w_{02}
3. Idea I_3, w_{03}
w_{ij} = association strength between Idea i and Idea j
α_j = absorption coefficient of Idea node j; $0 < \alpha < 1$

Figure 3: **Idea Flow.** Validators check if the is association strengths w_{ij} among ideas are fair. The reward is shared to the parent ideas (e.g. cited publications) proportional to the weights.

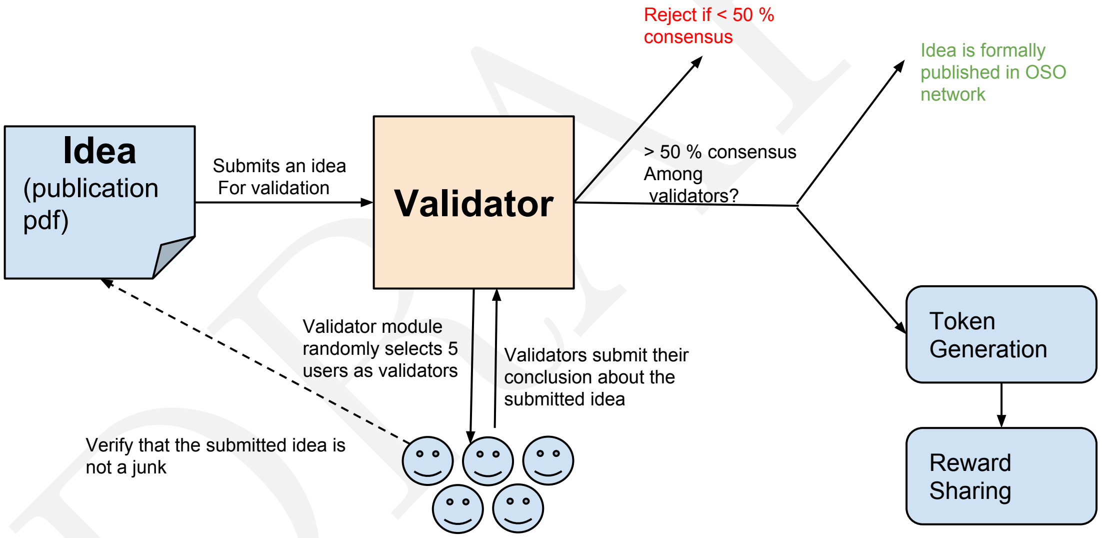

Reject if < 50 % consensus
Idea is formally published in OSO network
Idea (publication pdf)
Submits an idea For validation
Validator
> 50 % consensus Among validators?
Token Generation
Reward Sharing
Validator module randomly selects 5 users as validators
Validators submit their conclusion about the submitted idea
Verify that the submitted idea is not a junk

Figure 4: **Proof of Idea in layer 1: Curating a list of valid (not-spam) ideas.** Layer1 acts as a spam filter and ensures the proper connection of the idea with its parent ideas (cited publications). If the submitted idea gathers approval from the majority of the validators, it is formally appended to the idea network G_I and which triggers the token generation and reward sharing steps. Validators (along with other players in the process) are rewarded for their work (intellectual contribution) during the reward sharing.

2.3 Validation of the authenticity of the idea

Human users (e.g. researchers) will act as validators to check the authenticity of the submitted idea. For an idea, a subset of users from the list of verified users U are selected as validators. Selection of the validators is itself a separate topic of research. In initial implementation, a random subset of users are selected as validators. N_V (e.g. 7) validators are randomly selected from U and the selected validators will receive notification for validation. The validators have following tasks:

- check if the idea (publication is a garbage document to spam the network). This is a similar mechanism to that is used in biorxiv where someone checks the submitted paper before it gets published.
- check if the idea belongs to the domain D
- check if the relationship of the idea (connecting weights w) with its parent ideas (e.g. cited publications) is fair. The weights w determine the amount of value flow back. See [2].

If the consensus among the validators surpasses the threshold for consenus (some percentage, e.g. 50 %), the idea gets published in the network and there will be associated token generation and reward sharing event. For example, if 4 out of 7 validators agree that the submitted idea is not a spam and belongs to domain D, the idea I gets validated i.e. it gets formally published in the network. Here (in layer 1), 'published' means the idea object get recognized as a node in the idea network (represented by a digraph G_I) and its edges are determined by the dependency of the idea on its parent ideas (e.g. connecting weights w with parent ideas)

2.4 Token Generation

If an idea gets validated and is successfully submitted to 0S0 network, the idea gets rewarded with some cryptographic asset (proportional to the intellectual contribution) in the 0S0 network. This is opposite to the traditional model of scientific publishing where researchers have to pay to publish. **In 0S0 networks, researchers are paid to publish.** i.e. they are rewarded with assets in the network.

Let:

- **Fundamental unit of intellectual contribution:** 1 OSO token represent the smallest asset representing some form of intellectual contribution in the 0S0 network. OSO will be the utility token of the network and will be a medium of exchange to facilitate the trade of intellectual contributions (peer-review, bug bounty, etc.)
- **Total asset of the network:** Total number of OSO tokens that will ever be minted/generated (T_{OSO}) = 1 Trillion = 10^{12} = 1000,000,000,000
- **Idea reward rate** $R_I = 10^{-10}$ The idea reward rate is flexible and can be changed dynamically to adjust with the market price of OSO tokens. Note: Initially, the quality of all the validated ideas equal which is naive. We have to address this when we add a 'review layer on top of this initial design.

With this **publish-to-pay** design, when an idea gets validated and is successfully submitted to 0S0 network, certain number of OSO tokens are automatically generated. Idea reward rate controls the number of tokens generated by an idea. No.of tokens generated by an idea = Idea reward rate (R_I) * No. of tokens that can be potentially generated.

For genesis idea, No. of tokens that can be potentially generated = $T_{OSO} = 10^{12}$. So, the genesis idea will generate R_I*$T_{OSO} = 10^{-10} * 10^{12} = 100$ OSO tokens.

For second idea, No. of tokens that can be potentially generated = $T_{OSO} - 100 = 10^12 - 100 = 999,999,999,900$ So, the second idea will generate $R_I * 999,999,999,900 = 10^{-10} * 999,999,999,900 = 99.99999999$ OSO tokens.

And so on.

2.5 Reward (Token) Sharing

The newly generated OSO tokens are distributed to the players of the platform as follows:

- **Reward of Idea (50%)** i.e.(50%) of the generated tokens goes to authors. Authors can use these tokens to pay for the validation (review) cost to publish in any publishing channel of their choice within any sub-network of the 0S0 network.
- **Cost of Storage/Development/Maintenance (30%)** (30%) of the generated tokens goes to 0S0 fund. 0S0 fund is used for the development of the 0S0 platform. It will be used to reward the contributors and host the ideas in a p2p storage network such as IPFS [4] or its own network.
- **Cost of Idea (15%)** (15%) of the generated tokens are distributed to the cited publications.
- **Cost of Validation (5%)** (5%) of the generated tokens distributed to validators.

3 Attack Vectors

3.1 Unnecessary large file sizes to overload the network

There will be an upper bound to the size of the idea (paper here) to avoid unnecessary load to the network. 10 MB?

3.2 Haphazard creation of ideas

- Upper limit on how many ideas can be created in a time period (at least initially). This can be updated later.
- Upper limit on how many ideas can be submitted by a user

3.3 Sybil attacks

Upper limit on how many ids can be created in a time period i.e. how fast U grows.

4 Technical Implementation

Current implementation is Etheruem [5] based and is very basice. We are using https://metamask.io/ web-browser extension to interact with Ethereum Robsten testnet. There are two smart contracts:

4.1 Idea network (registry) G_I

https://gist.github.com/noman-land/64db658bbfa15c1d487ea1544c8e2d88

4.2 Validators

https://gist.github.com/noman-land/ce4c940bca841c6c2057df4fabd98d94

4.3 Token Ownership and Flow

Conceptually token ownership can be represented by a list L of mappings of idea wallets and tokens inside those wallets (address of idea, no. of OSO tokens)

Token flow means the change in the balances of the wallet ideas (second field of L) which equivalents to updating the list L. Specifically, a token flow is an event that can be represented by a list ΔL of transactions where each transaction updates the token balance of the corresponding wallet.

4.4 Challenges:

Theoretically, token flow can extend all the way to the genesis idea. In practice, token flow (value flow back) will extend up to a parent idea if the value flow back to it is greater than additional computational (hence economical) cost. Even with this restriction, the order of computation (i.e. no. of transactions or the length of ΔL) can be too large for current blockchain solutions. What makes it worse is that as the network grows, L and ΔL become larger making the token flow costlier.

4.5 Potential Solutions for Token Flow

4.5.1 Periodic updates of the list L

- update weekly ? Use of timestamp can be tricky because of triggering problem.
- update based on block number of the blockchain.
- update after every certain no. of ideas. Validator contract has a counter which counts the number of ideas that were successfully validated and formally submitted to the network. Use that counter value to trigger the token generation and reward sharing contract i.e. to update the list L

4.5.2 Look how brave payments handle the solution?

batch processing?

4.5.3 Use zk-snarks

https://ethresear.ch/t/on-chain-scaling-to-potentially-500-tx-sec-through-mass-tx-validation/3477

4.5.4 Use Bulk API of HUMAN protocol

https://medium.com/human-protocol/transfer-your-tokens-9-600x-more-efficiently-on-ethereum-using-the-bulk-api-fbc2f10669ed

4.5.5 Plasma chain

still thinking ...

5 How are we using Proof of Idea inside OSO community?

OSO members will be the initial list of the users (genesis users). Each member of the OSO community will generate a public/private key pair and will submit their public keys to be a user of the 0S0 platform. This will create a list of public keys that are controlled by genesis users. These public keys will be used in more user friendly identity solutions in later versions. Initially, an idea will be scientific publication, for example, 4 page pdf paper. There has to be an upper bound to the size of the idea (paper here) to avoid unnecessary load to the network.

5.1 Registration of new user

Option 1: A random subset of existing users can verify the new users and let their public key appended into the list of public keys LI Option 2: Anyone can join the network at their will sine we already know each other inside our community.

6 TO DO:

How to handle version change?

References

[1] G. J. Katuwal, "Open science organization- a decentralized autonomous organization for better scientific ecosystem," 2017. https://github.com/open-science-org/wiki/blob/master/OSO_white_paper.pdf.

[2] K. G. A. K. S. P. Katuwal, Gajendra J, "Open science organization- an idea platform," 2018. https://github.com/open-science-org/wiki/blob/master/OSO_Idea_Platform_whitepaper.pdf.

[3] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," 2009.

[4] J. Benet, "IPFS - Content Addressed, Versioned, P2P File System (DRAFT 3)," no. Draft 3.

[5] G. Wood, "Ethereum: a secure decentralised generalised transaction ledger," *Ethereum Project Yellow Paper*, pp. 1–32, 2014.