WAKU-RLN-RELAY: Privacy-Preserving Peer-to-Peer Economic Spam Protection

Sanaz Taheri-Boshrooyeh∗†, Oskar Thorén†‡, Barry Whitehat†, Wei Jie Koh§, Onur Kilic¶, and Kobi Gurkan∥
∗Vac Research and Development, †Status Research and Development, Singapore, ‡Unaffiliated,
§Independent, ¶Unaffiliated, ‖cLabs

Abstract—In this paper, we propose WAKU-RLN-RELAY as a spam-protected gossip-based routing protocol that can run in heterogeneous networks. It features a privacy-preserving peer-to-peer (p2p) economic spam protection mechanism. WAKU-RLN-RELAY addresses the performance and privacy issues of the state-of-the-art p2p spam prevention techniques including peer scoring utilized by libp2p, and proof-of-work used by e.g., Whisper, the p2p messaging layer of Ethereum. In WAKU-RLN-RELAY, spam protection works by limiting the messaging rate of each network participant. Rate violation is disincentivized since it results in financial punishment where the punishment is cryptographically guaranteed. Peers who identify spammers are also rewarded. To enforce the rate limit, we adopt the suggested framework of Semaphore and its extended version, however, we modify that framework to properly address the unique requirements of a network of p2p resource-restricted users. The current work dives into the end-to-end integration of Semaphore into WAKU-RLN-RELAY, the modifications required to make it suitable for resource-limited users, and the open problems and future research directions. We also provide a proof-of-concept open-source implementation of WAKU-RLN-RELAY, and its specifications together with a rough performance evaluation.

Index Terms—P2P, Spam Protection, Messaging, zkSNARKs, Zero-Knowledge, Anonymity, Routing, Pub/Sub, Gossipsub

I. INTRODUCTION

WAKU [1] is a family of peer-to-peer (p2p) protocols for anonymous and privacy-preserving communication. It is designed to be able to run in resource-restricted environments e.g., limited computational power, bandwidth, and storage space. Being p2p means that WAKU relies on no central server. Instead, peers collaboratively deliver messages in the network. WAKU-RELAY constitutes the transport layer of WAKU and aims at being privacy-preserving in which no one knows the owner and the receiver of a message except the two ends of the communication. Many of the design choices in this layer are centered around the anonymity requirement. WAKU-RELAY follows a publisher-subscriber messaging model and adopts a gossip-based routing protocol. It is a thin layer over the libp2p GossipSub [2] routing protocol. In a pubsub messaging model, peers congregate around topics they are interested in and can send messages to topics. Each message gets delivered to all peers subscribed to the topic. Each peer has a constant number of direct connections/neighbors. To publish a message, the author forwards its message to a subset of neighbors. The neighbors proceed similarly till the message gets propagated in the network of the subscribed peers. The adopted gossip-based routing allows concealing the message receiver hence greatly benefits receiver anonymity [3].

In addition to WAKU-RELAY, WAKU features other types of protocols for running in resource-restricted environments. Among which are the request/response protocols of 12/WAKU2-FILTER [1] which is a lightweight version of WAKU-RELAY for devices with limited bandwidth, and 13/WAKU2-STORE [1] by which resourceful peers can persist and offer historical messages to the querying nodes. Details of these protocols are out of the scope of this paper and can be found in the Waku RFCs [1].

As an open messaging network, WAKU-RELAY is prone to spam messages. A spammer usually indicates an entity that uses the messaging system to send an unsolicited message (spam) to large numbers of recipients. However, in WAKU-RELAY with an open and gossip-based structure, spam messages not only affect the recipients but also all the other peers involved in the routing process where they have to spend their resources e.g., computational power, bandwidth, and storage capacity on processing spam messages. As such, we define a spammer as an entity that uses the messaging system to publish a large number of messages in a short amount of time, in other words, has a high messaging rate. We define messages in this way as spam regardless of the intention of the spammer, the content of the message, and the number of intended recipients.

The state-of-the-art p2p spam protection techniques for messaging systems i.e., Proof of Work (PoW) [4] deployed by e.g., Whisper [5] and Peer scoring [2] method adopted by libp2p have their own shortcomings. The PoW technique imposes a high computational cost for messaging hence devices with limited resources won’t be able to participate and benefit from the messaging system. The peer scoring method is prone to censorship and is also subject to inexpensive attacks where the spammer can send bulk messages by deploying millions of bots. The spam protection methods deployed by centralized messaging platforms e.g., WhatsApp exhibit privacy issues where they usually ask users to disclose and commit to some piece of personally identifiable information e.g., phone number and email address at the registration time. In addition to this, the central provider is aware of messages owned and received by a particular user which is against privacy.
The economic-incentive spam protection mechanism of WAKU-RLN-RELAY aims at coping with the aforementioned issues where

1) It suits p2p systems and does not rely on any central entity.
2) It is efficient in terms of computation, storage, memory, and bandwidth usage, hence suitable for a network of heterogeneous peers with limited resources.
3) It respects users’ privacy unlike reputation-based and centralized methods. It also allows global identification and removal of spammers while not requiring personally identifiable information from the participants.
4) It deploys economic-incentives to contain spammers’ activity. Namely, there is a financial sacrifice for those who want to spam the system. Additionally, there is a financial reward to monitor the network for spam and remove the offending users.

At a high level, WAKU-RLN-RELAY guarantees no one can publish more than one message per epoch i.e., a time interval, without being financially charged. The spammer gets removed from the network and will not be able to publish further messages. The epoch length is a system parameter that should be configured to meet the desired messaging rate and application throughput. For example, while an epoch duration of 1 second i.e., a messaging rate of 1 per second might be acceptable for a chat application, might be too low for communication among Ethereum network validators.

II. PRELIMINARIES: RATE LIMITING NULLIFIER

A. Semaphore

Semaphore [6] is a zero-knowledge signaling framework on Ethereum. It allows a set of users to broadcast arbitrary signals (where the signal is any value like a string, vote, etc.) while proving they are among a group of authorized users without disclosing their identities. Use-cases of Semaphore are anonymous authentication and private voting. It also utilizes external nullifiers to prevent double-signaling. An external nullifier can be seen as a voting booth where each user can only cast one vote [7]. Casting a second vote for the same booth will be rejected. Similarly, each signal is bound to an external nullifier, and each group member is allowed to signal only once for that external nullifier. While nullifiers can limit the signal rate per user, there is no way to identify and remove users who violate this rate limit. In an attempt to address this shortcoming, an extended version of this method, using Shamir Secret Sharing (SSS) [8] has been proposed called Rate-Limiting Nullifier or RLN for short. In RLN [9], if a user attempts more than one signal for the same external nullifier, it reveals its identity/private key by which it has registered to the group. As such, the identified identity key can be removed from the group and the user won’t be able to signal anymore. The removed user will be also financially punished. The user initially deposits some funds when joining the group, and the fund will be rewarded to anyone who identifies double signaling of that user.

B. RLN: Semaphore with Shamir Secret Sharing

In this section, we present an overview of the extended version of Semaphore that utilizes Shamir secret sharing [8] to enable the identification of spammers.

Each member has a private key $sk$ and a public key $pk = H(sk)$ where $H$ is a Cryptographic hash function. These are also called the identity key and the identity commitment, respectively. The list of group members is stored in an Identity Commitment Tree which is a Merkle tree [10] whose leaves are members’ public keys. The root of the tree is denoted by $\tau$. The membership of a user in the identity commitment tree is proven by providing a branch of the tree that connects the root to that leaf corresponding to the user’s $pk$. This branch is called authentication path and we denote it by auth. The tree is stored on a smart contract deployed on the Ethereum blockchain. Each user additionally needs to deposit some funds in the contract at the time of registration.

Signaling is bound to a publicly known external nullifier denoted by $\emptyset$. When publishing signal $m$, a group member utilizes (2,n)-Shamir secret sharing to derive a share $(x, y)$ of its private identity key $sk$ where $x = H(m)$ and $y = sk + H(sk, \emptyset) * x$. $(x, y)$ will be published alongside with the signal $m$. The publishing user also calculates an internal nullifier $\phi$ as $\phi = H(H(sk, \emptyset))$ and publishes it together with the signal $m$. Finally, a user needs to prove in a zero-knowledge manner that

1) its secret key $sk$ belongs to the identity commitment tree namely i.e., provides proof of membership.
2) $(x, y)$ is a valid share of its identity key.
3) the internal nullifier $\phi$ is correctly calculated.
The above items are proven through zkSNARKs i.e., a zero-knowledge Succinct Non-interactive Argument of Knowledge proof system where the circuit represents the aforementioned constraints. The public inputs to the zero-knowledge proof system are the message hash, the external nullifier $\emptyset$, internal nullifier $\phi$, the share of identity secret key $(x, y)$ and the tree root $\tau$ whereas the private inputs provided by the message owner are the identity secret key $sk$, the index of $pk$ in the tree, and the authentication path $auth$. In WAKU-RLN-RELAY, we utilize Groth16 [11] for the proof system. The parameter generation can be done through a multi-party setup similar to [12], [13], [14], [15].

The user submits its signal $m$, identity key share $(x, y)$, and internal nullifiers $\phi$ to the contract to be stored and accessible by the rest of the group members. The contract maintains the state of all the group signals together with all the metadata that can be used to identify double signaling. If a user attempts two different signals for the same external nullifier $\emptyset$, their internal nullifiers will collide, which signifies a double signaling attempt. Furthermore, the two shares of the user’s identity key can be used to reconstruct its $sk$ and remove the user from the group. The $sk$ reconstruction stems from the fact that each line (polynomial of degree 1) can be uniquely reconstructed by using two distinct points on it. In the case of double signaling, the spammer reveals two distinct shares $(x, y)$ and $(x', y')$ on line $y = sk + H(sk, \emptyset) \cdot x$, which enables the reconstruction of the line and its evaluation at $x = 0$ which is $sk$. The user who reconstructs the $sk$ also gets rewarded by a portion of the slashed user’s stake. This can be done by communicating the recovered $sk$ to the contract.

III. WAKU-RLN-RELAY CONSTRUCTION

In this section, we describe the flow of the economic spam detection mechanism in WAKU-RLN-RELAY as also depicted in Figure 1.

A. Overview of WAKU-RLN-RELAY vs Semaphore

WAKU-RLN-RELAY adopts the extended variant of Semaphore i.e., RLN which provides the additional capability of identifying spammers. However, further adjustments are required to make it fit a p2p routing system. Following is the list of these adjustments that mainly affect the state of the contract.

1) In WAKU-RLN-RELAY the state of the contract keeps a simple ordered list of users’ identity commitments (instead of Merkle tree) and the Merkle tree is kept off-chain by individual peers. This is in contrast to the original setting of Semaphore where the contract holds the entire commitment tree. The reason for this design shift is to mitigate the significant computational cost/ gas consumption associated with member insertion and deletion which is logarithmic in the number of registered members. While the cost associated with insertion in Semaphore could be amortized by using batch insertion, this solution does not apply to the deletion since deletion affects random leaves of the tree which cannot be necessarily batched together. WAKU-RLN-RELAY mitigates this problem as insertion and deletion modify a single item of the list.

2) In WAKU-RLN-RELAY, users’ messages, and their metadata are not stored in the contract which is in contrast to Semaphore. In WAKU-RLN-RELAY messages are stored off-chain and are distributed through the WAKU-RELAY routing protocol. It has the benefits of 1) having higher message propagation speed and 1) being more economic (mesaging is for free) as opposed to the on-chain message store. In the on-chain message storage of Semaphore, published messages will not be visible until blocks containing those message transactions get mined. This results in an unnecessary and undesirable delay which is not acceptable for messaging systems with e.g., 1.1 million messages per second\(^1\). In an attempt to address this issue, WAKU-RLN-RELAY decouples the message propagation and storage from the contract state and provides a p2p and off-chain medium for message transportation i.e., WAKU-RELAY [1] and storage i.e., 13/WAKU2-STORE [1]. The off-chain storage, adopted by WAKU-RLN-RELAY, has another advantage of being more economic. That is it saves the financial cost (related to the gas consumption) associated with message insertion into the contract state. This cost for one-time messaging scenarios like voting systems can be tolerable, however, it is far from practical in a messaging application where there are millions of messages transmitted per second.

Due to these adjustments, sending messages in WAKU-RLN-RELAY is for free i.e., does not need gas consumption. Furthermore, message transmission is not affected by the underlying blockchain and its consensus layer. Such separation allows the utilization of various optimization techniques on the message transmission delay which would be otherwise impossible due to reliance on the blockchain.

B. Setup and Registration

A peer willing to publish a message is required to register. Registration is moderated through a smart contract deployed on the public Ethereum blockchain. The state of the contract contains the list of registered members’ public keys i.e., identity commitment keys. An overview of registration is illustrated in Figure 2.

For the registration, a peer creates its identity private key $sk$ and its commitment $pk = H(sk)$ and sends a transaction to the contract to register its identity commitment $pk$ in the group. The transaction also transfers the $v$ amount of Ether to the contract. This amount is deposited on the contract to prevent spam activity.

C. Maintaining the identity commitment tree

The construction and maintenance of the identity commitment tree, unlike the original proposal of Semaphore [6], is delegated to the peers. As we previously mentioned, the reason is that the cost associated with member deletion and

\(^{1}\)https://www.oberlo.ca/blog/whatsapp-statistics
F. Routing and Slashing

Peers follow the regular routing protocol of WAKU-RELAY and in addition check the metadata of each relayed message to identify and spot spam messages and slash spammers. The routing procedure is depicted in Figure 3.

Upon the receipt of a message \( (m, (x, y), \phi, \text{epoch}, \tau, \pi) \), the routing peer needs to decide whether to route it or not. The decision relies on the following factors:

1) If the \text{epoch} value attached to the message has more than \text{Thr} gap with the routing peer’s current epoch, the message is considered invalid and must be dropped. This is to prevent a newly registered peer from spamming the system by messaging for all the past epochs. \text{Thr} is a system parameter.

2) The message must contain valid proof \( \pi \) that gets verified by the routing peer.

3) The messaging rate is not violated.

If the preceding checks are passed successfully, then the message is relayed. In case of invalid proof, the message is dropped. If spamming is detected, the publishing peer gets slashed and the slasher gets rewarded.

To identify spam messages, each routing peer keeps a local record of the identity key share \((x, y)\) and the internal nullifier \(\phi\) of all of its valid incoming message bundles for the past \text{Thr} epochs. This list is called a nullifier map. The routing peer checks every new message against this list to spot spam messages. This list does not have to capture the entire history because any incoming message whose epoch is older than the last \text{Thr} epochs is discarded by default (as explained in item 1 of section III-F). The routing peer utilizes this list to locally identify spam messages and spammers as follows.

1) The routing peer first verifies the \( \pi \) and discards the message if not verified.

2) It checks for the presence of another past message with an identical internal nullifier i.e., \( \phi \). If nothing is found, then the message gets relayed, otherwise:

   a) If the identity share i.e., \((x', y')\) of the older message is different from the incoming message share i.e., \((x, y) \neq (x', y')\), then slashing takes place

   b) If \((x, y) = (x', y')\), then the message is a duplicate and should be discarded.

\textbf{Race condition:} There is a possibility of a race condition when slashing an identified spammer. The race happens when a peer submits the identified identity key of the spammer to the contract in plain format and at the same time, another peer observes and copies the same information to reclaim the spammer’s fund. This yields a race issue. To avoid this, the commit and reveal technique can be utilized, i.e., the routing peer sends a commitment to the spammer’s identity key (instead of the plain identity key) to the contract so that no one else can replicate and reclaim it. Later it opens the commitment and reveals its knowledge of the spammer’s identity key.

\textbf{Maximum Epoch Gap:} As mentioned earlier, the gap between the epoch observed by the routing peer and the
one attached to the incoming message should not exceed a threshold denoted by $\text{Thr}$. $\text{Thr}$ can be measured based on 1) the network transmission delay which is the maximum time that it takes for a message to be fully disseminated in the network 2) the clock asynchrony i.e., the maximum difference between the Unix epoch time perceived by the network peers which can be due to clock drifts. With a reasonable approximation of the preceding values, one can set $\text{Thr} = \lceil \text{NetworkDelay} + \text{ClockAsynchrony} \rceil$ where $T$ is the length of the epoch in seconds (the NetworkDelay and ClockAsynchrony are also in seconds). By this formulation, $\text{Thr}$ measures the maximum number of epochs that can elapse since a message gets routed from its origin to all the other peers in the network.

IV. SECURITY AND PERFORMANCE ANALYSIS

Security: WAKU-RLN-RELAY achieves global spam protection while preserving user anonymity, namely 1) Peers do not have to disclose any piece of personally identifiable information in any phase i.e., neither in the registration nor the messaging phase 2) Peers can prove that they have not exceeded the messaging rate in a zero-knowledge manner and without leaving any trace to their identity public keys. Our solution also protects network resources against spammers as the spam messages are dropped immediately and not propagated. Malicious participants that may attempt to send messages with invalid proofs to exhaust the resources of the network will also fail because the effect of their attack is 1) limited to their direct connections and will not impact the entire network, since such messages are not relayed further due to invalid proof 2) is easily addressable by leveraging peer scoring mechanism [2]. Sybil attack is also mitigated by making registration expensive. Economic incentives are guaranteed cryptographically via secret sharing.

Performance: The proof of concept implementation as well as specifications of WAKU-RLN-RELAY are available in [16] and [1], respectively. We utilize RLN library [17] for our implementation. Generating membership proof to a group size of $2^{32} \approx 4$ billion peers takes $\approx 0.5$s on an iPhone 8 [17]. Proof verification run time is constant and takes $\approx 30$ms. Each peer persists a 32B public and secret keys and a prover key with $\approx 3.89$MB in size [17]. Storage of a membership tree with depth 20 takes up 67MB from each peer (this can be optimized to 0.128KB using the proposal of [18]).
A. Future Work

WAKU-RLN-RELAY is currently a Proof of Concept (POC), and its design and development are in progress. Below, we shed light on some of our future work.

Evaluating Merkle tree computation overhead: We would like to evaluate the running time associated with the Merkle tree operations. Indeed, the need to locally store Merkle tree on each peer was one of the unknowns discovered during this POC and yet the concrete benchmarking result in this regard is not available.

Enhancing performance by off-chain solutions: Another possible improvement is to replace the membership contract with a distributed group management scheme e.g., through distributed hash tables. This is to address possible performance issues that the interaction with the public Ethereum blockchain may cause. For example, the registration transactions are subject to delay as they have to be mined before being visible in the state of the membership contract. This means peers have to wait for some time before being able to publish any message. The same issue exists for slashing and other smart contract-related functions. The use of state channels and optimistic Rollups are other ways to overcome transaction delays and achieve better performance.

Lowering the storage overhead per peer: Currently, peers need to maintain the entire tree locally which incurs a storage overhead linear in the size of the group. This overhead can be tackled using a hybrid architecture where peers with adequate storage capacity retain the tree and supply the necessary information to the resource-limited peers upon request. Alternatively, the storage efficient method of [18] can be utilized where peers maintain a partial view of the tree yet can construct and update the tree root and their authentication path based on the group updates. Keeping the partial view lowers the storage complexity to $O(\log(N))$ where $N$ is the size of the group. In this solution, the authentication path of the slashed keys can become available assuming that peers, at the registration time, 1) provide their authenticated path encrypted under their secret identity key and 2) prove the former in a zero-knowledge manner. The use of Verkle tree [19] or polynomial commitments [20] are other paths to follow.

Cost-effective member insertion and deletion: Depending on Ethereum gas costs, the cost associated with membership is 40$k$ gas which translates to more than 20 USD (at the time of writing of this paper). We aim at finding a more cost-effective approach. For example, by using batch insertion and deletion, the cost can be reduced to 20$k$ gas. Alternatively, layer two solutions or more cost-efficient blockchains with support for smart contracts can be utilized.

B. Open Problems

Below is the list of open problems for which no immediate solution is known hence demanding more long-term research.

Exceeding the messaging rate via multiple registrations: While the economic-incentive solution has an economic incentive to discourage spamming, there is still the possibility of expensive attack(s) for exceeding the messaging rate. An attacker pays for multiple e.g., $k$ registrations, and uses its aggregate quota for messaging i.e., $k$ messages per epoch. Such attacks can be mitigated by increasing the entry barrier via a higher membership fee. Following this argument, the high fee associated with the membership, which is discussed in section IV-A, can indeed be beneficial for spam prevention.

Escaping punishment by early withdrawal: A spammer can escape from getting slashed by withdrawing its fund from the contract before its spam activity gets caught. While this means the attacker burns its initial membership fund (the fee paid to register its key to the group), it allows saving the other part of the fund that can be otherwise rewarded to slashers.

ACKNOWLEDGMENT

We acknowledge Corey Petty (Corey@status.im) for his valuable comments on the paper.

REFERENCES

[1] “Specifications of waku protocols stack,” https://rfc.vac.dev/, accessed: 2022-03.
[2] D. Vyzovitis, Y. Napora, D. McCormick, D. Dias, and Y. Psaras, “Gossipsub: Attack-resistant message propagation in the filecoin and eth2.0 networks.” arXiv preprint arXiv:2007.02754, 2020.
[3] N. Unger, S. Dechand, J. Bonneau, S. Fahl, H. Perl, I. Goldberg, and M. Smith, “Sok: secure messaging,” in 2015 IEEE Symposium on Security and Privacy. IEEE, 2015, pp. 232–249.
[4] C. Dwork and M. Naor, “Pricing via processing or combating junk mail,” in Annual international cryptology conference. Springer, 1992, pp. 139–147.
[5] “Whisper specifications,” https://eips.ethereum.org/EIPS/eip-627 , accessed: 2022-03.
[6] K. Garkan, K. W. Jie, and B. Whelan, “Community proposal: Semaphore: Zero-knowledge signaling on etherum,” 2020. [Online]. Available: https://github.com/appliedzkp/semaphore/blob/master/spec/ Semaphore%20Spec.pdf
[7] https://semaphore.appliedzkp.org/, accessed: 2022-03.
[8] A. Shamir, “How to share a secret,” Communications of the ACM, vol. 22, no. 11, pp. 612–613, 1979.
[9] “Rln original proposal,” https://ethresear.ch/semaphore-rln-rate-limiting-nullifier-for-spam-prevention-in-anonymous-p2p-setting/5009, accessed: 2022-03.
[10] R. C. Merkle, “A digital signature based on a conventional encryption function,” in Conference on the theory and application of cryptographic techniques. Springer, 1987, pp. 369–378.
[11] J. Groth, “On the size of pairing-based non-interactive arguments,” in Annual international conference on the theory and applications of cryptographic techniques. Springer, 2016, pp. 305–326.
[12] S. Bowe, A. Gabizon, and I. Miers, “Scalable multi-party computation for zk-snark parameters in the random beacon model,” Cryptology ePrint Archive, 2017.
[13] “https://medium.com/coinmonks/announcing-the-perpetual-powers-of-tau-ceremony-to-benefit-all-zk-snark-projects-c3da86af8377.”
[14] “Implementation of multi-party computation (mpc) ceremony for zk-snark parameters,” https://github.com/kobigurk/phase2-bn254/tree/master/powersoftau, accessed: 2022-03.
[15] “First phase of a multi-party trusted setup ceremony based on zcash powers of tau ceremony,” https://github.com/weijieko/perpetualpowersoftau, accessed: 2022-03.
[16] “Reference implementation of waku in nim,” https://github.com/status-im/nim-waku, accessed: 2022-03.
[17] “Rln library implemented in rust,” https://github.com/kilic/rln, accessed: 2022-03.
[18] “Storage efficient merkle tree update,” https://github.com/vac2p2/research/blob/master/rln-research/merkle-tree-update.md, accessed: 2022-03.
[19] J. Kuszmaul, “Verkle trees,” Verkle Trees, pp. 1–12, 2019.
[20] A. Kate, G. M. Zaverucha, and I. Goldberg, “Polynomial commitments,” Tech. Rep. 2010.