

ClaimChain: Improving the Security and Privacy of In-band Key Distribution for Messaging

Bogdan Kulynych
EPFL SPRING Lab
bogdan.kulynych@epfl.ch

Wouter Lueks
EPFL SPRING Lab
wouter.lueks@epfl.ch

Marios Isaakidis
University College London
m.isaakidis@cs.ucl.ac.uk

George Danezis
University College London
g.danezis@ucl.ac.uk

Carmela Troncoso
EPFL SPRING Lab
carmela.troncoso@epfl.ch

ABSTRACT

The social demand for email end-to-end encryption is barely supported by mainstream service providers. Autocrypt is a new community-driven open specification for e-mail encryption that attempts to respond to this demand. In Autocrypt the encryption keys are attached directly to messages, and thus the encryption can be implemented by email clients without any collaboration of the providers. The decentralized nature of this in-band key distribution, however, makes it prone to man-in-the-middle attacks and can leak the social graph of users. To address this problem we introduce ClaimChain, a cryptographic construction for privacy-preserving authentication of public keys. Users store claims about their identities and keys, as well as their beliefs about others, in ClaimChains. These chains form authenticated decentralized repositories that enable users to prove the authenticity of both their keys and the keys of their contacts. ClaimChains are encrypted, and therefore protect the stored information, such as keys and contact identities, from prying eyes. At the same time, ClaimChain implements mechanisms to provide strong non-equivocation properties, discouraging malicious actors from distributing conflicting or inauthentic claims. We implemented ClaimChain and we show that it offers reasonable performance, low overhead, and authenticity guarantees.

KEYWORDS

E-mail encryption; Decentralization; Key distribution; Privacy

1 INTRODUCTION

Following the Snowden revelations it became clear that, given the dependence of citizens, governments, and corporations on electronic communications, there is a strong need for highly secure end-to-end encrypted communications. That is, the content of communications must not be accessed by third parties, so as to shield them from mass surveillance systems, domestic or foreign. Yet, so far we have only seen feeble and largely unsuccessful attempts by mainstream service providers such as Gmail and Yahoo to support fully encrypted e-mail [1, 8, 13]. Only a few minor e-mail providers embrace end-to-end encryption.¹²

To fill this void, a recently launched community-driven initiative, Autocrypt, is developing a new open specification for e-mail encryption. The goal is to facilitate key handling by mail user agents so that encryption can be deployed without the need for

collaboration of the service providers. The Autocrypt approach is supported³ by key e-mail clients such as Thunderbird+Enigmail, K-9 mail, and Mailpile, as well as a new messaging application for Android, DeltaChat.

Similarly to in-band PGP [24], Autocrypt embeds the encryption keys into the e-mail messages, but uses pre-defined headers instead of attachments. Furthermore, in the spirit of the PGP Web of Trust, these headers also contain cross-references to the keys of other users. The cross-references implicitly endorse the binding between these keys and the corresponding user identities.

This decentralized approach alleviates the privacy problem of centralized certification authorities, such as SKS Keyservers⁴ for PGP keys, which can observe users' key look-ups, and thus can infer their communication patterns. However, since no one has a global view of all the bindings in Autocrypt's decentralized approach, malicious users or providers can supply different user-to-key bindings to different recipients, effectively opening the doors to man-in-the-middle attacks.

This attack, whereby Alice can show to Carol and Donald different versions of Bob's key to manipulate encryption in her advantage, is commonly known as *equivocation*. A solution to render equivocation detectable and accountable could be to use CONIKS [15]. However, CONIKS' transparency logs are maintained by providers, and thus the scheme is not compatible with the Autocrypt principle of not requiring provider collaboration.

In this paper we present *ClaimChain*, a cryptographic construction that alleviates the authenticity and privacy problems of in-band key distribution in the setting of Autocrypt.⁵ Similarly to CONIKS, ClaimChains consist of chained blocks. Instead of a global log, however, each user has their own ClaimChain that contains all the information necessary to represent her claims about her own keys, and her beliefs about other users' keys, i.e., her cross-references. Chaining of blocks enables tracking and authenticating the evolution of beliefs and keys. Cross-references enable users to combine their contacts' beliefs—represented by their ClaimChains—to establish evidence about the binding between identities and keys.

To address the privacy issues of the Web of Trust cross-reference sharing model, ClaimChains use cryptographic access tokens to provide fine-grained control on who is allowed to read which claims. Moreover, ClaimChains' claim encoding schemes make it hard to

³<https://github.com/autocrypt/autocrypt/blob/master/doc/install.rst>

⁴<https://sks-keyservers.net>

⁵ClaimChains are currently being tested by the Autocrypt team (<https://py-autocrypt.readthedocs.io/en/latest/>)

¹<https://mailfence.com>

²<https://protonmail.com>

infer how users’ beliefs change over time. Finally, ClaimChains include mechanisms to reliably and efficiently prevent equivocation within a block, and a mechanism to detect equivocation across blocks. In doing so, ClaimChains ensure non-equivocation while minimizing the leakage of users’ friendship networks.

We prove the security and privacy properties provided by ClaimChains. We also provide an implementation of ClaimChains and show that it scales to accommodate the needs of large groups at an acceptable overhead cost. We also simulate the usage of ClaimChains for in-band key distribution using the Enron e-mail dataset⁶ in a privacy-preserving way: ClaimChain owners reveal cross-references only if that does not leak more information about their social graph than is already revealed by the e-mails themselves. We quantify the degree to which in-band key distribution can protect communication. We show that ClaimChains improve privacy, without diminishing considerably the ability to encrypt e-mail, and enable the detection of incorrect key information. Our main contributions are:

- We introduce ClaimChain, a cryptographic construction based on authenticated data structures. ClaimChains store claims about keys, thereby supporting key authentication in decentralized environments in a secure and privacy-preserving way.
- We define the properties that a decentralized and privacy-preserving key distribution system should offer and we formally model them for ClaimChain.
- We show that owners cannot equivocate about contact keys within blocks. Moreover, we provide a novel mechanism that enables ClaimChain owners to prove that they have not equivocated across different blocks about a particular contact key. Unlike other transparency-backed solutions [15], auditing the consistency of contact keys is possible without revealing their actual values in the ClaimChain history.
- We provide an implementation of ClaimChain and show that its computation and bandwidth requirements are reasonable and within reach of modern computers and networks.
- We evaluate the effectiveness of decentralized key distribution on a real e-mail dataset. We show that selective privacy-preserving distribution can be almost as effective as broadcasting all known keys, although total encryption is difficult to achieve.

2 PROBLEM STATEMENT AND GOALS

We assume a messaging system in which users embed their cryptographic keys in-band, i.e., into the messages themselves or into the message headers, as in Autocrypt. These keys are used to provide message confidentiality using opportunistic encryption [5]. That is, the communication is encrypted when users know each others’ keys, but falls back to plaintext when they do not.

Sending keys as part of message headers results in two problems. First, in terms of privacy, adding such headers reveals users’ social ties. Second, in terms of security, man-in-the-middle attackers can modify the header contents, since they are not authenticated. Moreover, malicious users can equivocate about others’ keys.

Design goals. We assume that all actors in the system, users and providers, may act maliciously. Our goal is to design a data structure that can store the binding between keys and identities, and is

suitable for integrating with in-band key distribution. The purpose of the structure is to support key validation, i.e., help users establish the authenticity of user-key bindings, as long as some users in the system are honest. Furthermore, it must protect users’ privacy without relying on centralized parties.

More concretely, we aim at providing the following properties. First, the structure must guarantee the *integrity* and *authenticity* of identity-key bindings, i.e., it should not be possible to replace or inject bindings without being detected. Second, we want to preserve the *privacy of cross-referenced information* and the *privacy of the social graph*. These properties ensure that only authorized users can access the key material in the structure and the identities of the bindings being distributed. Third, the structure must prevent users from *equivocating* other users with respect to the identity-key bindings that they share. That is, a user Owen should *not* be able to show to Alice and Bob different versions of a Charlie’s key, even if he withdraws Alice’s access to see Charlie’s keys. In the latter case if Alice ever regains access, she must be able to detect Owen’s misbehavior. Finally, our construction should not entail significant computational or communication overhead for the end users and providers to enable adoption at large scale.

Non-goals. In-band key distribution cannot ensure full availability of public encryption keys. The keys of one or more recipients may not be available to a sender at a time of sending a message, and thus, because of the opportunistic encryption operation, the message would be sent in the clear. We are therefore not concerned with ensuring 100% availability of keys. Instead, our goal is to secure the keys that *are* distributed without harming privacy. If guaranteeing encrypted communication is absolutely necessary, parties must exchange keys in a reliable way, e.g. through a centralized service or an out-of-band mechanism.

Furthermore, throughout this paper we consider that users have only one identity, and use one and only one structure to store key bindings of their contacts. If a user wishes to have different identities, she must create one structure per identity.

3 CLAIMCHAIN DESIGN

In this section we introduce ClaimChain, a structure to store key bindings in a secure and privacy-friendly manner.

3.1 Cryptographic preliminaries and notation

We denote sampling uniformly at random from a set X as $x \leftarrow \$ X$, and the assignment of an evaluation of a function $f(x)$ to y as $y \leftarrow f(x)$, regardless of whether f is probabilistic or deterministic. We denote concatenation of strings by $\|$.

Let λ be the security parameter. ClaimChain relies on the following standard cryptographic primitives. Let $\text{Enc}(k, m) \mapsto c$ and $\text{Dec}(k, c)$ denote an IND-CPA secure symmetric authenticated encryption scheme. ClaimChain uses an existentially unforgeable signature given by the algorithms $\text{Sig.KeyGen}(1^\lambda)$ returning the keypair $(\text{sk}_{\text{sig}}, \text{pk}_{\text{sig}})$, $\text{Sign}(\text{sk}_{\text{sig}}, m)$ returning a signature σ , and the verification function $\text{Sig.Verify}(\text{pk}_{\text{sig}}, \sigma, m) \mapsto \{\top, \perp\}$. We write $\text{DH.KeyGen}(1^\lambda)$ for the generation of a Diffie-Hellman (DH) keypair $(\text{sk}_{\text{DH}}, \text{pk}_{\text{DH}})$ using which we can non-interactively compute

⁶<http://www.cs.cmu.edu/~enron/>

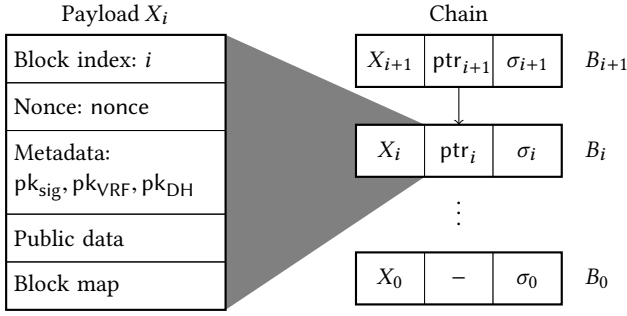


Figure 1: ClaimChain block structure

the shared DH key $s \in \{0, 1\}^*$ using $\text{SharedSecret}(\text{sk}_{\text{DH}}, \text{pk}_{\text{DH}}^R)$. Finally, let H be a cryptographic hash function from which we derive a family of hash functions $H_i : \{0, 1\}^* \rightarrow \{0, 1\}^{2\lambda}$, $i > 0$.

All schemes use a cyclic group \mathbb{G} of prime order q generated by g . We write \mathbb{Z}_q for the integers modulo q . Moreover, we assume the existence of a cryptographic hash function $H_{\mathbb{G}} : \{0, 1\}^* \rightarrow \mathbb{G}$ that hashes strings to group elements, and a hash function $H_q : \{0, 1\}^* \rightarrow \mathbb{Z}_q$ that hashes strings to the elements of \mathbb{Z}_q .

ClaimChains also require an information-theoretically hiding commitment scheme $\text{Commit}(r, m)$ that commits to values $m \in \mathbb{Z}_q$ given a randomizer $r \in \mathbb{Z}_q$. We instantiate this scheme using Pedersen’s commitment scheme [19]. Let $g_1, g_2 \in \mathbb{G}$ be random generators such that the discrete logarithms of g_1 and g_2 with respect to each other are unknown. Then, $\text{Commit}(r, m) = g_1^r g_2^m$.

ClaimChains use standard zero-knowledge proofs of knowledge, and in particular Schnorr’s proof of knowledge of discrete logarithms [21], to prove correctness of claims. We use the Fiat-Shamir heuristic [6] to derive non-interactive signature proofs of knowledge. For example, we write:

$$\text{SPK} \{ (r, m) : C = g_1^r g_2^m \} (t)$$

to denote the non-interactive signature proof of knowledge on a random string t for which the prover knows the commitment opening (r, m) . To focus on the semantics of the proof, we write

$$\text{SPK} \{ (r, m) : C = \text{Commit}(r, m) \} (t)$$

instead, to denote the same proof.

Finally, ClaimChains use a verifiable random function (VRF) [7, 16], given by the algorithms VRF.KeyGen and VRF.Eval . The function $\text{VRF.KeyGen}(1^\lambda)$ returns a keypair $(\text{sk}_{\text{VRF}}, \text{pk}_{\text{VRF}}) = (\text{sk}_{\text{VRF}}, g^{\text{sk}_{\text{VRF}}})$. Then, $h = \text{VRF.Eval}(\text{sk}_{\text{VRF}}, m) = H_{\mathbb{G}}(m)^{\text{sk}_{\text{VRF}}}$ is the VRF of the value m . Users prove that h was correctly computed by constructing the proof

$$\text{SPK} \left\{ (\text{sk}_{\text{VRF}}) : \text{pk}_{\text{VRF}} = g^{\text{sk}_{\text{VRF}}} \wedge h = \text{VRF.Eval}(\text{sk}_{\text{VRF}}, m) \right\} ()$$

The properties of VRF hashes are similar to that of cryptographic hashes: uniqueness of h for a given message and private key, collision resistance, and pseudorandomness (assuming no access to the corresponding proof) [18].

3.2 Overview

We consider that each user has a state made of information about herself and her beliefs about other users’ states. At a given point in time a users’ state is represented as a set of statements, called *claims*. Claims can be of two kinds. The first type of claim refers to a user’s own state. In particular, these may be statements on the user’s encryption keys, identity information (screen name, real name, or e-mail), or other cryptographic material such as verification keys to support digital signatures. The second type of claims, we call them cross-references, refer to other users’ states. A claim owner creates a cross-reference to endorse the referenced user’s state as being authoritative, i.e., a cross-reference indicates the owner’s belief that the self key material found in those users’ state is correct. A user’s state evolves over time as she rotates her keys and observes the evolution of others’ states. She stores snapshots of her state in a cryptographic data structure called a *ClaimChain*.

The core element of a ClaimChain is a block. A block includes all claims that the owner endorses at the time when she creates the block, i.e., a block is a snapshot of the owner’s state. Blocks form a chain. A block contains a payload X , a pointer to the previous block ptr , and a digital signature σ_i on the payload and the pointer. See Figure 1. The payload of the previous block X_{i-1} contains the verification key $\text{pk}_{\text{sig}}^{(i-1)}$ for the private key $\text{sk}_{\text{sig}}^{(i-1)}$ that signs σ_i .

We now describe each of the block components in detail.

The payload X_i has the following content (see Figure 1, left):

- *Block index*. The block’s position in the chain. The index of the genesis block is 0.
- *Nonce*. A fresh cryptographic nonce used to ‘salt’ all cryptographic operations within the block. It ensures that the information across blocks is not linkable.
- *Metadata*. The current signature verification key of the owner pk_{sig} , that is used to authenticate the next block of the ClaimChain; the current key pk_{VRF} to compute a verifiable random function used to support non-equivocation; and a Diffie-Hellman key pk_{DH} used to provide claim privacy.
- *Public data*. Application-specific data the owner wishes to make publicly visible. For in-band key distribution we set this to the owner’s self-claim on her current public encryption key.
- *Block map*. A high-integrity key-value map storing the claims, as well as access tokens that express access-control rights. This map has two core properties: i) a key can only be resolved to a single value, and ii) it enables the generation and verification of efficient proofs of inclusion of claims or access tokens. We implement the map using *unique-resolution key-value Merkle trees*, explained in more detail in Section 3.3. For our use case the map only contains cross-references.

The signature $\sigma_i = \text{Sign}(\text{sk}_{\text{sig}}^{(i-1)}, (X_i, \text{ptr}_i))$ authenticates the current block. A block B_i must have a valid signature under the verification key indicated in the payload of the previous block B_{i-1} . The genesis block of a ClaimChain is ‘self-signed’. The corresponding initial public signing key is included in the initial payload. Each block in the chain contains enough information to authenticate past blocks as being part of the chain, validate the next block, and, by transitivity, all future blocks as being valid updates. Therefore,

a user with access to a block of a chain that she believes is authoritative, can both audit past states of the chain, and authenticate the validity of newer blocks.

3.3 Low-level operations

We now describe how we implement claims and access tokens, and how they are combined into the block map.

Claims. We model claims as a tuple composed of a *label* l and a *body* m . The label is a well-known identifier associated with the identity of the user to whom the claim refers. The body is the state of that user at the time when the claim is generated, represented as the latest block of this user’s ClaimChain. For instance, a claim (‘bob@gmail.com’, B) represents the ClaimChain owner’s belief that the current state of the user associated with this Gmail account is represented by the block B .

For privacy reasons, claims in a ClaimChain are encrypted. Thus, they cannot be found directly by other users. To enable efficient search for concrete claims within a ClaimChain block, we introduce a *lookup key*, or *index*, i for each claim.

We illustrate the encoding of claims in procedure ENCCLAIM, see Figure 2. Consider a claim (l, m) that is to be included in a block. We first compute the unique VRF of its label and derive the claim’s lookup key i [lines 2–3]. Note that the computation of h includes a per-block nonce to ensure that the lookup keys for a given claim label look different across blocks, and therefore no patterns can be inferred from their appearance.

Recall that VRF hashes are unique. We use them to derive lookup keys to ensure that, given a label, all users retrieve the same claim, effectively supporting non-equivocation within a block. This use of VRFs is inspired by CONIKS [15]. We also include additional cryptographic elements in our encoded claims in order to obtain a stronger non-equivocation property than CONIKS. Specifically, we guarantee that equivocation is detectable *across blocks* without the need for key owners to intervene. The need for a detection mechanism stems from the fact that ClaimChain owners can give and withdraw access to claims at will. Thus, they can try to equivocate others by giving them access to different information in different blocks. To make this misbehaviour detectable we provide ClaimChain owners with the ability to prove statements about claims that other users cannot see. This way, if a proof cannot be completed, equivocation is revealed (see Section 4.1 for more details).

To prove statements on claim contents without revealing them, we commit to the claim body m [line 4]. Moreover, we construct a non-interactive proof π on k_π proving that the VRF h is correct and that the commitment com commits to m [lines 5–6]. When decoding a claim, users verify the proof π . The proof verification key k_π ensures that only authorized users can verify this proof.

Once π is computed, we encrypt m and π with a random key k [line 7]. Finally, the claim encoding consists of this ciphertext and the commitment com : $c = \text{Enc}(k, \pi \parallel m) \parallel \text{com}$.

The binding property of the commitment com and the validation provided by the proof π also ensure that all users with access to an encoded claim c must recover the same claim body m . This makes this encoding scheme an instance of committing, or non-deniable, encryption [9]. Hence, a malicious owner can not equivocate by supplying two different claim encryption keys to different users.

The procedure DECCLAIM, see Figure 2, describes the decoding of a claim. It takes as input the encryption key k , the VRF hash h , and the proof verification key k_π from the owner (see below). Then, users can decrypt the ciphertext using k [lines 2–3]; and verify the claim proof π , which includes the verification of the correctness of the VRF hash h and of the commitment com [lines 4–6].

Our claim encoding scheme offers four distinct security advantages. First, the use of the VRF ensures that lookup keys can only be produced by the owner of the chain, which as we describe below supports access control. Second, the lookup key is unique for a given label, and thus can be used to support non-equivocation for claims within a block. Third, the lookup key i and claim encoding c leak no information about the claim label or body. Fourth, it supports zero-knowledge proofs about claim contents, which enables the detection of equivocation across blocks.

Access capabilities. ClaimChain owners create cryptographic access tokens called *capabilities* to ensure that only authorized users can access specific claims. A single capability grants one authorized user access to one claim. We call the authorized users *readers*.

An encoded capability is an encryption of all the values needed to obtain a claim lookup key and decode the corresponding claim: the encryption key k , the VRF hash h , and the proof verification key k_π . We encrypt these using a key derived from a shared secret s between the chain owner and the reader. Similarly to claims, encoded capabilities have an associated lookup key i_{cap} , and a body cap .

The procedure ENCCAP, see Figure 2, describes how to encode capabilities. First, it computes the shared Diffie-Hellman secret s using the owner’s private DH key sk_{DH} and reader’s public DH key pk_{DH}^R [line 2]. The latter is available in the metadata of the reader’s ClaimChain. We use the secret s to derive both the capability lookup key i_{cap} [line 3], and the capability encryption key k_{cap} [line 4]. Then we encrypt the values h , k , and k_π using the key k_{cap} to obtain the capability encoding [line 5]: $\text{cap} = \text{Enc}(k_{\text{cap}}, h \parallel k \parallel k_\pi)$.

Chain owners store the encoded claim c under the lookup key i in the block map. Similarly, they store the encoded capability cap under the lookup key i_{cap} . To find a capability corresponding to a claim with label l in a ClaimChain block, a reader first computes the lookup key i_{cap} for label l using the shared secret with the ClaimChain owner. If the corresponding capability cap is in the block, she decodes it using DECCAP, see Figure 2. First, the reader derives the shared secret s [line 2], and computes the capability encryption key k_{cap} using the claim label l [line 3]. She can then decrypt cap using k_{cap} [line 4], obtaining the label’s VRF hash h , the encryption key k , and the proof verification key k_π . With this information the reader can compute the claim’s lookup key $i = H_1(h)$, find the claim, and decode it using DECCLAIM.

Block map. Encoded claims and capabilities are stored in the block map. We implement the block map using a unique-resolution key-value Merkle tree. Unlike a standard Merkle tree that implements an authenticated set data structure, a key-value tree is an instance of an authenticated dictionary [4]. It can be efficiently queried for a value that corresponds to a given lookup key. Our construction is similar to that of a binary search tree: the intermediate nodes contain pivots that define whether the querier should follow the left child or a right child; the leaf nodes contain the values. The construction allows queriers to be sure that retrieved values are

<pre> 1: procedure ENCClAIM(sk_{VRF}, l, m, nonce) 2: h ← VRF.Eval(sk_{VRF}, l nonce) 3: i ← H₁(h) 4: r ← \$Z_q, com ← Commit(r, H_q(m)) 5: k_π ← \$ {0, 1}^λ 6: π ← SPK{(sk_{VRF}, r) : pk_{VRF} = g^{sk_{VRF}} ∧ h = VRF.Eval(sk_{VRF}, l nonce) ∧ com = Commit(r, H_q(m))}(k_π) 7: k ← \$ {0, 1}^λ, c ← Enc(k, π m) com 8: return r, h, k, k_π, (i, c) 1: procedure ENCCAP(sk_{DH}, pk_{DH}^R, l, h, k, k_π, nonce) 2: s ← SharedSecret(sk_{DH}, pk_{DH}^R) 3: i_{cap} ← H₃(s l nonce) 4: k_{cap} ← H₄(s l nonce) 5: cap ← Enc(k_{cap}, h k k_π) 6: return (i_{cap}, cap) </pre>	<pre> 1: procedure DECCLAIM(pk_{VRF}^O, h, l, k, k_π, c, nonce) 2: c̄ com ← c 3: π m ← Dec(k, c̄) 4: ▷ Note the verification of π requires pk_{VRF}^O, h, l, k_π, com, m, and nonce. 5: if π is not a valid proof then 6: return ⊥ 7: return m 1: procedure DECCAP(sk_{DH}, pk_{DH}^O, l, cap, nonce) 2: s ← SharedSecret(sk_{DH}, pk_{DH}^O) 3: k_{cap} ← H₄(s l nonce) 4: h k k_π ← Dec(k_{cap}, cap) 5: i ← H₁(h) 6: return i, h, k, k_π </pre>
---	--

Figure 2: Low-level ClaimChain operations

unique, i.e., there cannot exist any other leaf nodes that correspond to the queried lookup key. We call this the *unique resolution property*. We formally define the property in Experiment 1 and prove it in Theorem 1 (both in Appendix A). We refer to Appendix A for further details on the construction.

The unique-resolution property guarantees that for a given lookup key i , respectively i_{cap} , there can only be one claim c , respectively capability cap . The uniqueness of the VRF value h , the property of the tree, and the commitment in the claim encoding, ensures that a ClaimChain owner can not equivocate within a block.

We note that ClaimChain blocks only need to include the root hash of the Merkle tree, not the whole tree. This is because our Merkle tree construction allows to produce an inclusion proof for items: a path from the root to the leaf node which contains the item. Thus, providing others with this paths is enough to convince them that the items are in the tree defined by the root in the block.

3.4 High-level operations

So far we have described how users can encode and decode claims. We now outline how these claims can be included in a ClaimChain and read from it. At a glance, owners create blocks with a set of encoded claims and corresponding encoded capabilities, and use them to extend their ClaimChains. Any user can validate the authenticity and integrity of the chain. Moreover, readers can retrieve the claims they are authorized to read. Section 4 illustrates how these operations can be used in the context of in-band key distribution.

Content-addressable store. ClaimChain owners store their blocks and trees in mutable *content-addressable stores*. These are key-value stores where the key must be the hash of the corresponding value. They are a good fit for ClaimChains because i) it is easy to verify their integrity by checking that all keys are the hashes of the respective objects they map to; and ii) an incomplete store cannot lead to an erroneous decision on the authenticity, inclusion or exclusion of any block or tree node. The store supports two operations:

- PUT(v). Record the value v in the store.

- GET(h). Return v such that $h = H(v)$, if present in the store.

Extending a chain. Whenever an owner decides to add new claims to her ClaimChain she uses the procedure EXTENDCHAIN in Figure 3. This procedure takes as input the public application data, a set of claims (l_j, m_j) to add to the block, an access control set acs consisting of the authorized reader-label pairs for these claims, the cryptographic keys necessary to create the block (keypairs for signatures, DH key exchange, and VRF), as well as the previous signing key sk'_{sig} included in the previous block, the pointer ptr to that block, and, finally, the user’s store.

To create a block the user first generates a random nonce that is used for all encoding operations [step 1]. She then encodes all the claims and capabilities [steps 2–3]. The set S of encoded values and their respective lookup keys are used to construct a Merkle tree with root hash MTR, as described in Algorithm 1 in Appendix A [steps 4–5]. She then constructs the block payload X using the nonce, the block metadata containing the public keys $(\text{pk}_{\text{DH}}, \text{pk}_{\text{sig}}, \text{pk}_{\text{VRF}})$, the public application data, and the root MTR of the Merkle tree. She signs the payload X and the pointer ptr to the previous block using the previous signing key sk'_{sig} (see Figure 1) [step 6]. Finally, she puts the obtained block, $B = (X, \text{ptr}, \sigma)$, into the content-addressable store [step 8].

Chain validation. Readers must always validate that new blocks correctly extend the chain that they have previously seen. To do so, users run the procedure VALIDATEBLOCKS, see Figure 4. The input to this procedure is a list of blocks B_i , where B_0 is the last validated block, and B_1 through B_t are the new blocks to be validated. For each new block B_i the reader first checks if the block includes all elements: the payload, signature, and the pointer [step 1]. Next, she retrieves the public key pk_{sig} from the preceding block B_{i-1} and verifies the signature in the block B_i [steps 2–3]. This verifies the authenticity of the chain. Finally, she verifies that the pointer in the block B_i is a hash of the preceding block B_{i-1} , which verifies the integrity of the chain [step 4].

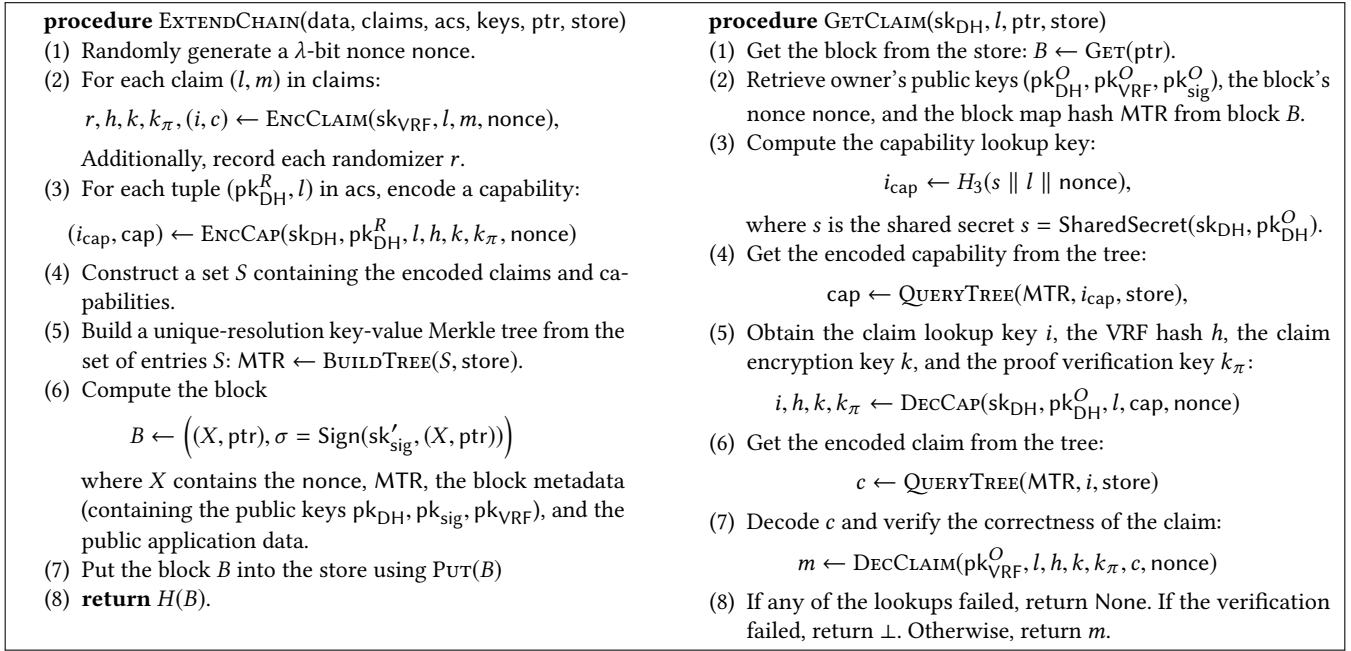


Figure 3: Extending and querying ClaimChains

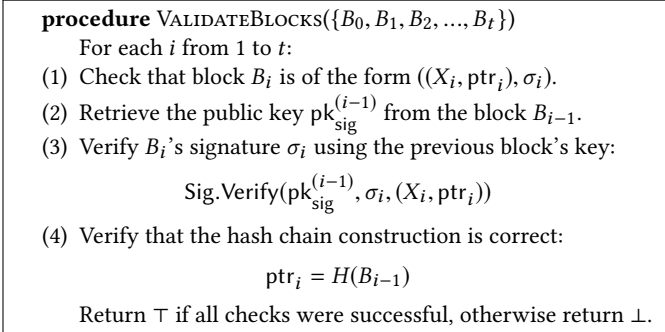


Figure 4: Block validation

Retrieval of the claim by label. After having validated the ClaimChain of an owner, the reader can query it to retrieve claims of interest using procedure GETCLAIM in Figure 3. This procedure takes as input the reader's private Diffie-Hellman key sk_{DH} , the claim label l , a pointer to the latest block ptr and the owner's store. The reader retrieves the block, and parses it to get the block's nonce, the owner's public keys, and the block map hash [steps 1–2]. She then derives the capability lookup key using the DH secret shared with the owner [step 3], queries the block map to retrieve the corresponding capability [step 4]. We refer to Algorithm 2 in Appendix A for the details of the QUERYTREE algorithm. Next, she runs the decoding procedure to obtain the claim lookup key i , the VRF hash h , the claim encryption key k , and the proof verification key k_π [steps 5]. She then obtains the claim encoding c by querying the tree with the claim's lookup key i [step 6]. Finally, the reader decodes and verifies the encrypted claim using h, k, k_π [step 7].

3.5 Security and privacy properties

We now sketch why the ClaimChain design fulfills the security and privacy objectives established in Section 2.

We note that *authenticity* and *integrity* are guaranteed through the usage of signature and hash chains respectively. Signatures guarantee that the information stored in a ClaimChain has been added by the owner of the chain. The usage of cryptographic hash functions for constructing the pointers between blocks guarantees that tampering with the ClaimChain content will be detected.

Privacy. ClaimChains provide *privacy of content* and *privacy of the social graph*. We capture these through the following properties:

- *Capability-reader unlinkability.* The adversary cannot determine for which honest user a capability has been created.
- *Claim privacy.* The adversary cannot learn anything about the labels and bodies of claims for which it does not have the corresponding capabilities.

Informally, these properties are provided by ClaimChains because the adversary can neither derive the capability lookup key, nor learn the contents of the encoded capability without the knowledge of the shared secret used to encrypt the them. This implies that an adversary without this key cannot read capabilities nor learn to whom they are destined (capability-reader unlinkability). Since the adversary cannot read the capability, it also does not learn the VRF hash h required to compute the claim lookup key, nor the claim encryption key k . Moreover, the pseudorandomness of the VRF hash h ensures that the adversary cannot compute h without the cooperation of the chain owner. Thus, the adversary cannot check whether a particular claim is included in the block.

Following a similar reasoning, the adversary cannot learn the content of a claim from its lookup key. Furthermore, the encoded

claim c does not reveal anything about the claim, except its length. Therefore, claim privacy holds, as long as all claims are of the same length, or padded to the same length.

We formalize these properties in Experiments 2 and 3, and prove them in Theorems 3 and 4 in Appendix B.

Non-equivocation. Our construction also *prevents equivocation*. Specifically, it guarantees the following two properties:

- *Intra-block non-equivocation.* Within a given block, a ClaimChain owner cannot include two different bodies encrypted to different readers, having the same claim label.
- *Detectable inter-block equivocation.* For any subset of ClaimChain blocks the owner can produce a proof that, for a given label l , all claims in these blocks belong to some set of allowed claims M without revealing the claims themselves.

The latter property ensures that a user cannot selectively withdraw access rights between blocks to equivocate users. We detail this attack and the proof that mitigates it in Section 4.1.

The intra-block non-equivocation relies on three properties of the ClaimChain construction. First, the uniqueness of the VRF hash h ensures that for a given label all readers will compute the same claim lookup key. Second, the unique-resolution property of our Merkle tree ensures that for a given lookup key all readers obtain the same claim encoding. Third, the claim commitment ensures that all readers will decrypt the same claim body.

We formalize both properties in Experiments 4 and 5, and prove them in Theorems 5 and 6 respectively in Appendix B.2.

4 USING CLAIMCHAINS TO SECURE IN-BAND KEY DISTRIBUTION

Recall from Section 2 that the goal of the ClaimChain data structure is to improve the security and privacy of in-band key distribution. In this section we describe how this can be achieved.

Building a ClaimChain. To use a ClaimChain, a user has to build blocks, containing her claims. When to update the ClaimChain depends on the owner’s preferences. For example, a user can update her chain whenever she rotates her own encryption public key, or when she needs to distribute new cross-references that are not present in her ClaimChain yet.

To update her chain, an owner runs the `EXTENDCHAIN` procedure (Figure 3). For this purpose, she encodes a set of claims representing all her current views of other users as cross-references in the following way. For each contact, she makes a cross-claim (l, m) , where l is the contacts’ e-mail, and m is the contact’s latest block.

Then, the owner must decide which of these claims she intends to make available to which of her contacts. This choice determines the access control set acs . The access control policy is governed by the user’s privacy preferences. Defining these preferences is beyond the scope of this work.

Recall that to implement access control the owner uses shared DH secrets with each of the readers. Thus, the owner needs to complete a round-trip of messages with a contact before she can give this contact access to her claims.

Finally, the owner puts her own public encryption key into the public application data section of the block. For our use case of in-band key distribution we assume that all keys are constant size. Hence blocks, and therefore claims, are constant size too. This

ensures claim privacy even though the encryption scheme leaks the length of the plaintext.

Distributing ClaimChains. To fulfill their purpose, ClaimChains must be made available to other users. For this, a user includes a content-addressable store containing blocks from her ClaimChain, and a subset of the Merkle tree nodes from her latest ClaimChain block, in every message she sends. The user keeps a record of which blocks they have sent to whom. To select the blocks to be sent, the sender checks her record, and includes all her ClaimChain blocks that the recipients of the current e-mail have not received yet.

The subset of the Merkle tree is selected to ensure that all information in the ClaimChain relevant to her message can be authenticated. More concretely, the sender produces resolution paths on the tree (see the `GETINCPATH` procedure in Algorithm 2 in Appendix A for the details) for each relevant claim and capability.

Receiving messages and validating ClaimChains. Upon receiving a message with a store containing ClaimChain data, a user first validates the received chain, running the `VALIDATEBLOCKS` procedure (Figure 4) to check if the new blocks extend a chain that has been seen previously. If the validation succeeds, the owner checks the *consistency* of the cross-references in the newly received part of the chain, i.e., whether all the cross-references to Charlie point to the blocks on a single chain. This partially prevents malicious chain owners from cross-referencing fake chains. See Section 4.1 for an example of such an attack, and the details of a consistency check procedure in case the receiver does not have access to the claim in some of the received blocks. If both checks succeed, she stores all the received blocks and tree nodes into her *gossip storage*. This enables her to query the sender’s ClaimChain later. The gossip storage contains all the block and tree nodes the user has received over time.

Message encryption. Following the opportunistic encryption paradigm, before sending a message, the sender checks if she has learned the public keys of all of the recipients through the ClaimChains she has received over time. If she cannot find all keys, she sends the message in plaintext.

To find the encryption keys she proceeds as follows. For every recipient with e-mail address l , and every ClaimChain with head ptr in her gossip storage `gossip_store`, she runs `GETCLAIM(skDH, l, ptr, gossip_store)`, see Figure 3, to find out whether it includes cross-references to this recipient. For every hit, she parses the corresponding claim and adds the cross-referenced ClaimChain block of the recipient to a *social evidence set* for this recipient. She then identifies the most recent block of the recipient’s ClaimChain (out of those present in her evidence set), i.e., the one that forms the longest hash chain, and uses the encryption public key in that block to encrypt the message. As a result of this process, the sender may discover new blocks of the recipient’s chain. She can then include the updated views as cross-references next time the chain is extended.

Resolving conflicts. This key resolution process may reveal conflicting views. For example, the blocks in the evidence set could point to two or more distinct chains. Another possibility is there could be a ‘fork’: two valid blocks with the same block index that extend a common parent block. In either case, ClaimChains conflicts are detectable and generate cryptographically non-repudiable

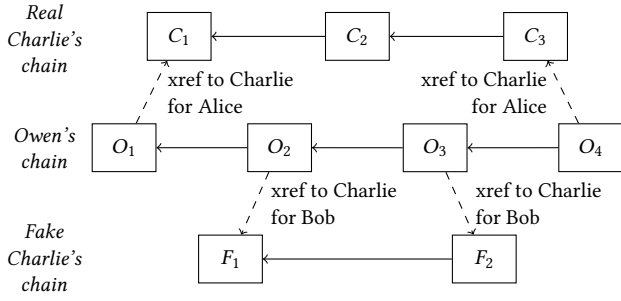


Figure 5: Inter-block equivocation

evidence. The design of mechanisms for sharing such evidence and deciding how to act on it is out of scope of this work. In Section 5.2, however, we empirically measure the number of distinct views that a sender would on average have about a recipient, to quantify if resolving conflicts is possible at all in a decentralized setting.

4.1 Detecting inter-block equivocation

ClaimChain’s intra-block non-equivocation property ensures that all readers of cross-references to Charlie’s chain see the same cross-reference in each block. However, chain owners may try to present different views to different users in different blocks by abusing the access-control mechanism. Thereby, the chain owner can equivocate between blocks.

Consider the following example, illustrated in Figure 5, in which the chain owner Owen shows Bob a fake cross-reference to Charlie’s chain, while showing the correct cross-reference to Alice. To do so, he never lets Alice and Bob see claims about Charlie’s chain in the same block. In block 1, he gives access to Alice, but not to Bob, while in blocks 2 and 3, he gives access to Bob, but not to Alice. Finally, in block 4, Owen again gives access to Alice but not Bob. If Owen has claims about Charlie’s true chain in blocks 1 and 4—the ones that Alice can read—and false claims about Charlie’s chain in blocks 2 and 3—the ones Bob can read—he is effectively launching an equivocation attack.

A trivial solution to prevent this attack would be to, upon suspicion, allow Alice and Bob to inquire about claims related to Charlie in the blocks where they do not have access. However, this can leak information about if and when the chain owner learned about Charlie’s updates. To be able to withdraw the access while preventing the described attack in a privacy-preserving way, ClaimChain enables the chain owner to prove, in zero knowledge, that she did not equivocate in the blocks where the cross-references were not accessible by the reader.

Consider again our example in Figure 5. When Alice regains access to Charlie’s references in block 4, she can use a detection mechanism to detect Owen’s equivocation attempt. In other words, she can determine that in the intermediate blocks 2 and 3, where she did not have access to the cross-references about Charlie, Owen referenced a different chain than the one she sees. Bob would also detect the equivocation if he regains read access.

To enable detection, upon giving the access to Alice in block 4 again, Owen constructs a non-equivocation proof as follows.

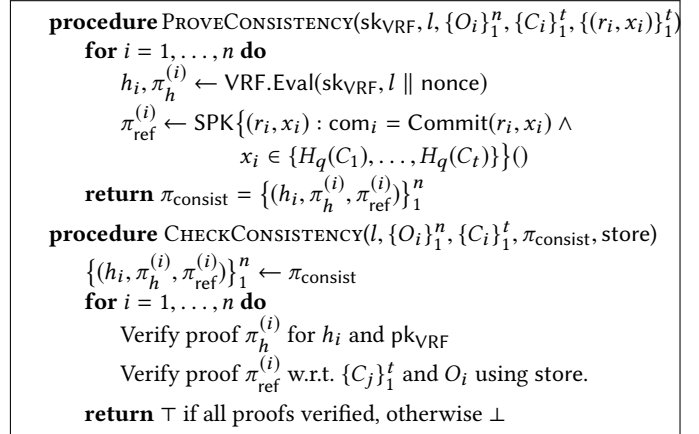


Figure 6: Proving and verifying that blocks O_i cross-reference the label l to the correct chain C_i .

- (1) Owen recomputes the VRF hashes $h_i = VRF.Eval(sk_{VRF}, l \parallel nonce_i)$ for all intermediate blocks, and computes proofs of correctness $\pi_h^{(i)}$:

$$\pi_h^{(i)} = SPK\{(sk_{VRF}) : pk_{VRF} = g^{sk_{VRF}} \wedge h_i = VRF.Eval(sk_{VRF}, l \parallel nonce_i)\}()$$

Alice can use the VRF hashes to locate the cross-reference to Charlie in the intermediate blocks of Owen. The proofs $\pi_h^{(i)}$ confirm that she found the correct claims for Charlie’s label l .

- (2) Owen proves in zero-knowledge that com_i commits to one of the intermediate blocks C_1, \dots, C_t on Charlie’s chain:

$$\pi_{ref}^{(i)} = SPK\{(r_i, x_i) : com_i = Commit(r_i, x_i) \wedge x_i \in \{H_q(C_1), \dots, H_q(C_t)\}\}()$$

Owen compiles all the tuples $(h_i, \pi_h^{(i)}, \pi_{ref}^{(i)})$ and sends them to Alice. Alice uses these tuples to check that each of the intermediate blocks belong the same chain of Charlie that she saw before. If Owen indeed equivocated as in the example, Alice can detect this, since the proof verification would have failed. A detailed description of this procedure is given in Figure 6.

5 EVALUATION

Experimental setup. We implemented a prototype of ClaimChains in Python.⁷ This implementation uses the petlib library [20] for elliptic curve cryptography operations, which internally relies on the OpenSSL C library. For the implementation of hash chains and unique-resolution Merkle trees we use the hippiehub⁸ library, which is written in pure Python. Our implementation uses AES128 in GCM mode for symmetric encryption; ECDSA, ECDH, and other elliptic curve operations with a SECG curve over a 256 bit prime field (“secp256k1”); and SHA256 as the base hash function.

All the lookup keys on the claim map are truncated to 8 bytes, which makes collisions unlikely for up to 2^{32} entries in the map.

⁷<https://github.com/claimchain/claimchain-core>

⁸<https://github.com/gdanezis/rousseau-chain>

Table 1: ClaimChain basic operations timing

	mean (ms)	std. (ms)
Label capab. lookup key computation	0.30	0.01
Label capab. decoding	0.33	0.01
Label capab. encoding	0.33	0.02
Claim encoding [π computation]	2.44 [2.38]	0.05 [0.05]
Claim decoding [π verification]	3.03 [2.96]	0.05 [0.05]

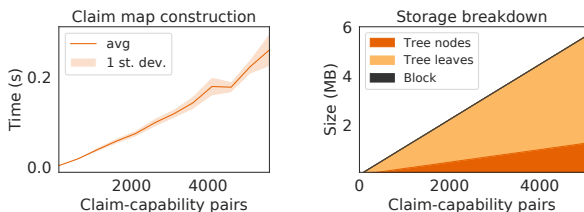


Figure 7: Total storage size and claim map construction time

The size of the per-block nonce is set to 16 bytes, and it is generated using the standard Linux `urandom` device.

Our experiments are also publicly available and reproducible.⁹ We extensively use Jupyter notebooks [11] and GNU parallel [22]. We run the experiments on an Intel Core i7-7700 CPU @ 3.60GHz machine using CPython 3.5.2.

5.1 ClaimChain operations performance

We now evaluate the performance of ClaimChains in terms of computation time and storage.

Timing. We first measure the computation time for encoding and decoding claims and capabilities as described in Section 3.3. We encode and decode 1000 claims and corresponding capabilities for random readers (i.e., encoded for a random DH public key). Each claim has a 32-byte random label and 512-byte random content. This reflects a realistic e-mail setting: 32-byte labels can accommodate e-mail addresses or their hash; and 512 bytes approximates the approx. 500-bytes block size in our experiments below.

Table 1 reports our measurements. The time for encoding, decoding, and computing lookup keys for capabilities is under 0.33 ms. The time to encode and decode claims is around 3 ms, consisting mostly of the proof computation and verification time.

The most computationally expensive operation that ClaimChain owners perform is constructing the block map when a new block is created. The map is constructed using the `BUILDTREE` procedure (see Algorithm 1 in Appendix A). We measure the time to create a block map of n claims with one capability each, i.e., readable by only one reader. We range n from 100 to 5,000. For each case we construct a unique-resolution key-value Merkle tree with the encoded entries. Figure 7 (left) shows the average time required to build the tree across 20 experiments. Even for 5,000 claim-capability pairs the operation takes under 0.3 seconds. In reality, we expect users to have much fewer entries per block (in our simulation using the Enron dataset this number rarely exceeds 1,000).

⁹<https://github.com/claimchain/claimchain-simulations>

Recall that along with the block, users send paths that prove the inclusion of relevant claims and capabilities in the block map tree. These are computed using the `GETPATH` procedure (see Algorithm 2 in Appendix A). We measure the time to compute and verify a proof for a single entry, as well as the proof size in terms of number of tree nodes and bytes. We use the same setting as in the previous experiment. Unsurprisingly, the computation and verification time, and the proof size scale logarithmically with the number of items in the map. For 5,000 items, computation and verification take on average about 150 milliseconds, and the proof consists of on average 20 tree nodes and takes about 1.5 KB.

Storage. We measure the size of a ClaimChain block, a block map tree, and values stored in the leaves of the tree (encrypted claims and capabilities). The size of the block map depends on the number of entries in the map and the size of claims. Figure 7 (right) shows the size breakdown depending on the number of items in the map. Note that the block itself only includes the root of the tree. Thus, the block size is constant (about 500 bytes), and can only grow if security parameters change (size of cryptographic public keys, or hash length increases), or additional data about the owner is added.

Inter-block equivocation detection. The cost of proving consistency is dominated by the proof $\pi_{\text{ref}}^{(i)}$. Using a straightforward instantiation with ‘or’ proofs, the prover and verifier must compute approximately $5t$ exponentiations to construct and verify $\pi_{\text{ref}}^{(i)}$, where t is the number of possible cross-referenced blocks. Therefore, a full consistency proof requires approximately $5nt$ exponentiations, where n is the number of intermediate blocks on the owner’s chain.

5.2 ClaimChain for in-band key distribution

In this section we study the use of ClaimChains for supporting in-band key distribution. We make use of the Enron dataset [10, 14] as a realistic test load to drive our experiments. It contains 500,000 e-mails of 147 Enron employees (230,000 after removing duplicates and non-readable e-mails).

To simulate the use of ClaimChains, we loop through the e-mails in this dataset chronologically, updating ClaimChains of senders and receivers after each sent e-mail. For the experiments we run simulations of 10,000 consecutive e-mails from the full log of e-mails. We consider that at the beginning of a simulation the senders’ ClaimChains are empty. During simulation they embed ClaimChain data in their messages as described in Section 4, and that upon receiving a message, users store all ClaimChain data locally.

We consider a scenario, called *private setting*, in which senders selectively share the cross-references using ClaimChains. The rights to access claims are granted incrementally via ‘introductions’. Every time there is an e-mail with more than one recipient, all recipients earn a capability to access the cross-references about the correspondents in the sender’s ClaimChain. Such capabilities are persistent over time. A user updates her chain when any of the recipients should receive a capability to a claim, and either the capability or the claim are not already in the ClaimChain.

As a baseline, we compare it to the *public setting*, another scenario in which users gossip all the information available to them to their correspondents. They do not use ClaimChains since there is no need for access control. This scenario is close to the operation of the PGP Web of Trust if the users always were attaching their

public key along with signatures on the keys of all their friends (cross-references). Since users share all the information they have available with everyone, the effectiveness of this setting reflects the fundamental limit for decentralized in-band key distribution.

To show how key propagation properties vary with the underlying social graph, we consider two groups of users from the dataset. First, only the 147 Enron employees, which represents a dense, well-connected, small social network with frequent messaging. Second, all users in the dataset, which is a sparsely connected network with many sporadic messages.

Resilience to conflicts in views. We first study how ClaimChains protect against attackers that aim at misleading users by reporting fake information about others. For each e-mail we record the amount of views that the sender has collected over time about each recipient. We call this quantity *social evidence diversity*. When the evidence diversity is 1 for a given recipient, the sender only has one view—that of her own. When it is 10, that means the sender knows of 9 other people that have cross-referenced the recipient’s chain. Intuitively, the higher is the diversity, the more users have to be corrupted by an adversary in order to convince the sender of a non-truthful ClaimChain state of the recipient.

Fig. 8 illustrates the results in one batch of 10,000 messages. Results in other batches are similar. Unsurprisingly, in the public setting the amount of evidence is much higher than in the private setting, since much more information is exchanged. We also see that including all users reduces the mean diversity, since the social graph is sparse and many users do not have the opportunity to gather enough information about their correspondents.

Social evidence can also differ because over time views get outdated. In this case, the availability of the latest ClaimChain state of a user in our decentralized in-band setting is fundamentally limited by users’ communication behavior. This is not a critical issue, since the construction of ClaimChains enables to differentiate ‘forks’ resulting from an attack from simple chain updates. For example, in our private setting, social evidence includes views of the same chain at different time instants (on average in 1% of cases among all users, and in 5% within the Enron employees). In all these cases, ClaimChain enables to establish which view is the most up-to-date in a given social evidence set.

Storage and bandwidth costs. To evaluate the overhead imposed by the use of ClaimChains for each user we record the size of the ClaimChain data being sent with each e-mail, and the required storage. We separately measure *self-storage*—the space taken by users’ own ClaimChain blocks and tree nodes, and *gossip storage*—the space taken by information received from other users.

The Figure 9 shows that these costs rise over time as chains grow. We observe a large variation in growth caused by the variation in users’ behaviour within the dataset. In the extreme, after 10,000 sent e-mails the required bandwidth per message is under 30 kB, the total size of the self-storage is under 50 kB, and the size of the gossip storage is under 2 MB. Note that we only report here the results in the private setting, since it employs ClaimChains.

Effectiveness of in-band key distribution. Recall that the end goal of the public-key distribution is to enable end-to-end encrypted communication. Thus, we measure the effectiveness of the distribution as the percentage of encrypted e-mails, i.e., the fraction of

times when a sender has received at least one ClaimChain block of *all* the recipients, directly or through gossiping, and could find the key to encrypt the message to them. We simulate 10,000 e-mails starting at arbitrary points in the e-mail log and record whether senders have enough information to encrypt the e-mail. Figure 10 shows how this proportion changes over time.

On the left we show the results within the set of Enron employees. At the beginning of the trace users learn many keys and increasingly more e-mails get encrypted. After some time the discovery rate decreases and there is a large variation in the proportion of encrypted e-mails. For this particular run, in the public setting, the overall proportion is 66%, decreasing to 57% in the private setting. We study the variance running the private setting simulations ten times at different points in the log. We observe that by the 2,000-th, 4,000-th, 6,000-th, and 8,000-th e-mail, on average 38% (± 15), 50% (± 11), 55% (± 12), 57% (± 12) of e-mails are encrypted¹⁰. The overall percentage of encrypted e-mails across the ten runs is 59% (± 9). Note that these rates do not consider possible key gossiping that could have happened through users outside of the company.

On the right we consider all users. Compared to the previous measurements, the proportion of encrypted messages is significantly lower, overall average being 26% in the public setting. This is because many senders and recipients are outside of the group of Enron employees and thus exchange fewer, or sporadic, e-mails. In the private setting the proportion decreases by only 4 p.p to 22%. Across all ten runs, the overall proportion is 23% (± 7). Note that these numbers are a lower bound on the number of encrypted e-mails. Since we do not observe the inbox of non-Enron users, we cannot establish the effectiveness of gossiping. Thus, the propagation may be better than shown in our measurements.

Takeaways. As expected, promiscuous public gossiping is more effective at propagating the key information than privacy-preserving sharing. Nonetheless, its advantage is relatively small. Sacrificing users’ privacy does not provide a significant increase in the proportion of encrypted e-mails. This suggests that selective revealing of cross-references, enabled by ClaimChain cryptographic mechanisms, can offer a better trade-off between privacy and utility than the traditional Web of Trust-like sharing model.

On the other hand, the gain in privacy comes at a cost in resistance to active attacks. Even though gossiping in the private setting does not significantly decrease the proportion of encrypted e-mails, it does significantly deteriorate the resilience to malicious users. The evidence diversity in the private setting can be up to 10 times lower than in the public setting. This means that on average fewer users need to be compromised to propagate inauthentic keys.

The public setting in our simulations represents an upper bound for key propagation as in this setting users share all the information available to them. Our results corroborate that, independently of the use of ClaimChains to secure cryptographic material, in-band key distribution is unlikely to achieve full coverage, and furthermore the coverage is largely unpredictable. In a decentralized setting, key propagation cannot be more effective unless additional communication channels are employed.

¹⁰By $\pm x$ we denote 95% Student t-distribution confidence interval in percentage points

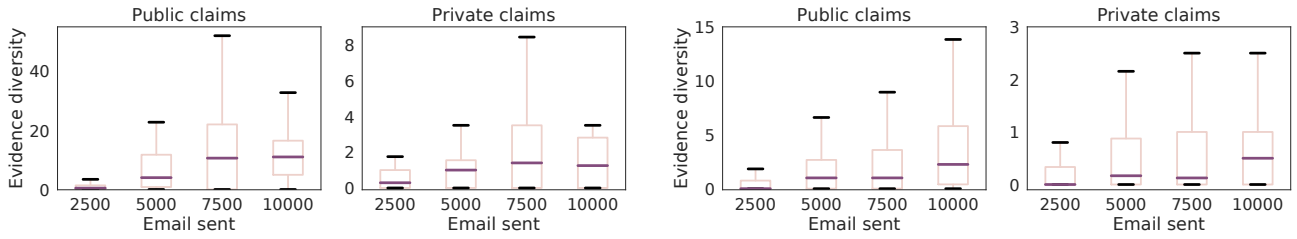


Figure 8: Evidence diversity (within Enron, left, and all users, right).

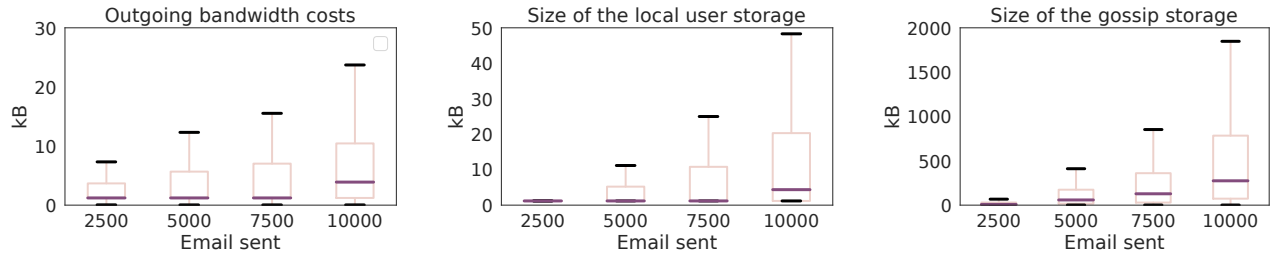


Figure 9: Storage and bandwidth measurements

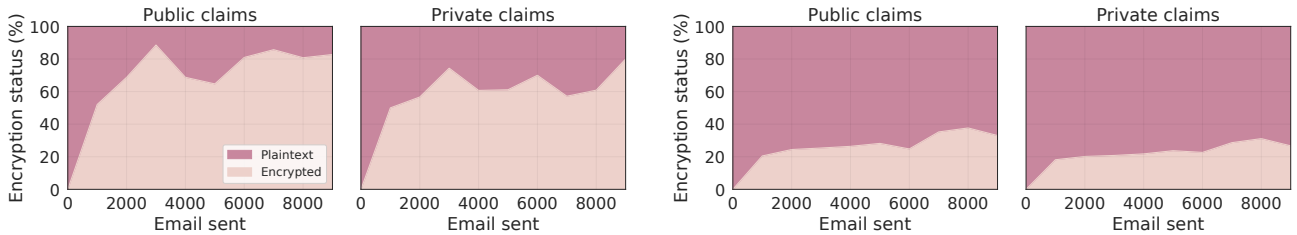


Figure 10: Encryption status of e-mails (within Enron, left, and all users, right). Proportions are computed over groups of 1,000 consecutive e-mails.

6 COMPARISON WITH EXISTING SYSTEMS

In this section we compare ClaimChain to other key distribution systems targeting e-mail communications. Designs that facilitate secure key distribution in other contexts, like Certificate Transparency [12] for HTTPS connections, are out of scope.

We consider four approaches that represent existing deployed and academic solutions. First, we consider the PGP in-band decentralized approach, where users attach their public keys in outgoing e-mails. This corresponds to the current implementation of Autocrypt where there is no gossiping of contacts' keys. Second, we compare ClaimChain to systems that employ highly available centralized key servers taking the SKS Keyservers, a pool of synchronized servers that store PGP keys, as reference. Third, we consider solutions that, like ClaimChain, use high-integrity data structures (hash chains and Merkle trees) so as to hold providers accountable for the bindings they serve. In this space we consider CONIKS [15],

a federated variant, and Keybase¹¹, a centralized design. Finally, we consider approaches such as Namecoin¹² that bind identities to cryptographic wallet addresses on Proof-of-Work blockchains [17].

We compare these systems in terms of functionality, computational and network costs. The result is summarized in Table 2 where we use the following notation: n —number of users, s —number of sent messages, r —number of received messages, b —maximum total number of contacts of any user after s sent and r received messages.

PGP key distribution. OpenPGP [2] is a format for sharing PGP key material. It enables users to vouch for each others' keys, constructing a Web of Trust. In-band PGP key distribution, whereby users' keys are embedded in messages, is vulnerable to attacks by malicious e-mail providers and network adversaries who, first, gain complete visibility of the users' Web of Trust, and second, can

¹¹<https://keybase.io>

¹²<https://namecoin.org>

Table 2: Comparison of key distribution systems from an end-user perspective. *Social graph visibility*: who learns the user’s social graph. *Active attack detection*: whether active attacks by malicious providers, users, and network adversaries can be detected. *Total key availability*: guarantee that recipients’ current encryption keys are always available to senders.

	In-band PGP	SKS Keyserv.	CONIKS	Keybase	Namecoin	ClaimChain
	E-mail provider	Public	Provider	Public	Public	Authorized readers
Social graph visibility						
Active attack detection	✗	✗	✓ [‡]	✓ [‡]	✓ [‡]	✓
Total key availability	✗	✓	✓	✓	✓	✗
Sending bandwidth, $O(\cdot)$	$s \cdot b$	$s \cdot b^2$	$s \cdot b \cdot \log(n)$ [†]	$s \cdot b \cdot (b + \log(n))$ [†]	$s \cdot b \cdot (b + \log(n))$ [†]	$s \cdot b^2 \cdot \log(b)$
Receiving bandwidth, $O(\cdot)$	$r \cdot b$	$r \cdot b$	$r \cdot \log(n)$ [†]	$r \cdot (b + \log(n))$ [†]	$r \cdot (b + \log(n))$ [†]	$r \cdot b^2 \cdot \log(b)$
Local storage, $O(\cdot)$	b^2	b^2	$b + \log(n)$	b^2	b^2	$r \cdot b^2 \cdot \log(b)$

[†] Without costs for auditing the transparency log / verifying blockchain history [‡] Requires global consensus on system’s state

replace the attached keys and in that way compromise the confidentiality of users’ communications. The Autocrypt implementation protects from servers launching man-in-the-middle attacks, but still reveals the contacts. Moreover, malicious users can equivocate by sharing different versions of others’ keys with different readers. ClaimChain can be seen as an alternative format to OpenPGP where cryptographic mechanisms hide contacts’ information. Furthermore, it ensures that all readers retrieve the same cross-reference for a contact, and makes past equivocation attempts detectable.

For in-band PGP key distribution users need to attach their keys and cross-references in their e-mails, hence requiring $O(s \cdot b)$ and $O(r \cdot b)$ outgoing and incoming bandwidth respectively. Locally, they need to store the keys of all their friends ($O(b)$), and their friends’ cross-references ($O(b)$ for each friend), resulting in $O(b^2)$ cost. As shown in Section 5, key propagation efficiency in this setting strongly depends on the social graph and on the users’ behavior, and is not constant.

Centralized PGP PKI providers allow to achieve 100% key availability, but introduce security and privacy concerns. The most widely deployed implementation, the SKS Keyservers, does not defend against malicious providers that serve fake keys, or that exploit key lookup requests to learn users’ relationships. In addition, they accept unauthenticated plaintext HTTP requests, and thus network adversaries can also perform these attacks.

Centralized PKIs obviate the need for attaching key material in outgoing e-mails. However, before sending a message, a user needs to obtain the latest key of each of the recipients (which contains their cross-references, $O(b)$), requiring $O(s \cdot b^2)$ bandwidth. Upon receiving a message, users look up the sender’s key (containing $O(b)$ cross-references), thus also resulting in $O(r \cdot b)$ network cost. As in the previous setting, users store the keys of their friends, including their cross-references, requiring a storage of $O(b^2)$.

Accountable key repositories. Recent approaches to key distribution rely on transparency logs [15] to prevent active attacks from malicious providers and network adversaries.

In CONIKS [15], providers maintain cryptographically signed hash chains that can be audited for serving the correct key bindings for their users. Key lookup responses in CONIKS include a Merkle tree proof of inclusion (size $\log(n)$). Hence, considering that users make a lookup request for each of the recipients when sending, and for the sender when receiving, the bandwidth cost is $O(s \cdot b \cdot \log(n))$ for sending and $O(r \cdot \log(n))$ for receiving. Users store the keys of the friends, and the current inclusion proof for their own key, which requires storage space bound by $O(b + \log(n))$. We do not

consider bandwidth costs related to verifying the history of the CONIKS provider or the consistency of users’ keys.

Keybase maintains a global auditable hash chain that contains commits to users’ individual *sigchains* through a global Merkle tree. These sigchains are self-signed objects that evolve over time and include information about owner’s keys, devices, online profiles, and friends. The global chain is occasionally cross-referenced into the Bitcoin’s blockchain guaranteeing that it cannot be tampered with. Keybase users can create cross-references to their contacts by adding a snapshot of their contacts’ state into their own sigchain. These cross-references are public, thus reveal user relationships.

The latest state of a Keybase user includes a Merkle proof in the Keybase tree and the cross-references of her friends, which is bound by $O(b + \log(n))$. Hence, the bandwidth cost for sending messages, assuming that senders retrieve the receivers’ latest keys from Keybase, is $O(s \cdot b \cdot (b + \log(n)))$. In the same way, the cost of receiving e-mails, including looking up the latest version of the sender key, is $O(r \cdot (b + \log(n)))$. These estimations do not consider the cost for validating the Keybase history, the consistency of user history, or the Keybase cross-references on the Bitcoin blockchain. Users store locally the keys of their friends and their respective cross-references, thus requiring $O(b^2)$ storage.

Unlike ClaimChains, neither CONIKS nor Keybase provide mechanisms to enable social verification while protecting the social graph of users. Furthermore, CONIKS, a federated design, and Keybase, a centralized design, both put providers in a privileged position to observe their users’ communications patterns. ClaimChains are designed to work in a decentralized setting that does not rely on the existence of a single entity with a global view of the system.

PKIs inspired by the Bitcoin blockchain. Some PKI systems leverage permissionless decentralized cryptographic ledgers. Besides high availability and resistance to tampering attempts, ledgers also provide a global namespace and mechanisms for achieving global consensus. For instance, Namecoin uses a proof-of-work blockchain [17] to implement key discovery for secure messaging apps. In Namecoin key material is encoded in the OpenPGP format, thus revealing the users’ social graph. Moreover, the public transactions reveal information such as how social graphs evolve, or how pseudo-identities are linked to the same owner. ClaimChains protect this information by means of cryptographic access controls. Furthermore, blockchain-based systems involve a fee for obtaining and managing key material. Instead of maintaining a global state, in ClaimChains each user controls a personal chain. Thus, there

is no need for mining blocks and thus there are no costs for key management operations.

From the perspective of end-users, the bandwidth and storage costs in Namecoin are the same as in Keybase. In a similar way, we do not consider costs involved in maintaining consensus in a global proof-of-work blockchain, or in verifying its history.

In-band key distribution with ClaimChains. Finally, we outline the costs involved with using ClaimChains. Besides her own information, owners include a claim for each of her cross-references and capabilities, along with the corresponding proofs of inclusion in the block map. Therefore, the size of a ClaimChain block is bound by $O(b^2 \cdot \log(b))$. Consequently, given that all messages attach a ClaimChain block, the total sending and receiving bandwidth cost is $O(s \cdot b^2 \cdot \log(n))$ and $O(r \cdot b^2 \cdot \log(n))$, respectively. Users are expected to store all ClaimChain blocks they receive, requiring a total storage of $O(r \cdot b^2 \cdot \log(n))$.

7 CONCLUDING REMARKS

In-band key distribution, as proposed by Autocrypt, is a promising direction towards achieving e-mail encryption without the collaboration from service providers. However, it suffers from security and privacy problems. To address these issues we introduced ClaimChains, a construction that can be sent in-band with e-mails to provide high-integrity evidence of key-identity bindings. Its cryptographic access control enables users to selectively reveal their contacts, preserving their privacy, while preventing equivocation attacks in which different users are shown different bindings.

We demonstrate that key propagation, and thus the ability to encrypt messages, is not affected much when using the privacy features of ClaimChains. However, users do obtain less evidence about other users' bindings, increasing the chances that wrong keys go unnoticed. On the negative side, our study shows that the coverage achieved by in-band key distribution is partial at best. In our realistic simulations we could achieve a maximum of 66% of e-mail encrypted, even within a well-connected social network.

However, we note that the design of ClaimChains is not tied to decentralized storage and distribution. Their strong security and privacy properties permit to host the content-addressable storage in semi-trusted providers without relying on them to return correct values. Such deployment of ClaimChains would greatly improve availability of ClaimChain data. But, to obtain perfect privacy, such scheme requires integration with privacy-preserving storage access [3, 23] to avoid leakage stemming from access patterns.

Finally, ClaimChain or its component data structures can have applications to use cases beyond key distribution. The claim map data structure, for example, can be applied in similar settings when a verifiable dictionary with cryptographic access controls for its lookup keys is needed.

ACKNOWLEDGMENTS

This research is funded by NEXITLEAP project¹³ within the European Union's Horizon 2020 Framework Program for Research and Innovation (H2020-ICT-2015, ICT-10-2015) under grant agreement 688722. We thank Holger Krekel, Azul, and Harry Halpin for their feedback and discussions.

¹³<https://nextleap.eu>

REFERENCES

- [1] Lucian Armasu. 2017. Google Abandons 'End-To-End' Email Encryption Project, Invites Community To Take It Over. <https://www.tomshardware.com/news/google-abandons-end-to-end-email-encryption,33745.html>. Last accessed: August 29, 2018.
- [2] J. Callas, L. Donnerhacke, H. Finney, D. Shaw, and R. Thayer. 2007. OpenPGP Message Format. RFC 4880. <https://rfc-editor.org/rfc/rfc4880.txt>
- [3] Benny Chor, Oded Goldreich, Eyal Kushilevitz, and Madhu Sudan. 1995. Private Information Retrieval. In *36th Annual Symposium on Foundations of Computer Science, Milwaukee, Wisconsin, USA, 23-25 October 1995*. IEEE Computer Society, 41–50. <https://doi.org/10.1109/SFCS.1995.492461>
- [4] Scott A. Crosby and Dan S. Wallach. 2011. Authenticated Dictionaries: Real-World Costs and Trade-Offs. *ACM Trans. Inf. Syst. Secur.* 14, 2 (2011), 17:1–17:30. <https://doi.org/10.1145/2019599.2019602>
- [5] V. Dukhovni. 2014. Opportunistic Security: Some Protection Most of the Time. RFC 7435. <https://rfc-editor.org/rfc/rfc7435.txt>
- [6] Amos Fiat and Adi Shamir. 1986. How to Prove Yourself: Practical Solutions to Identification and Signature Problems. In *Advances in Cryptology - CRYPTO '86, Santa Barbara, California, USA, 1986, Proceedings (Lecture Notes in Computer Science)*, Andrew M. Odlyzko (Ed.), Vol. 263. Springer, 186–194. https://doi.org/10.1007/3-540-47721-7_12
- [7] Matthew K. Franklin and Haibin Zhang. 2013. Unique Ring Signatures: A Practical Construction. In *Financial Cryptography and Data Security - 17th International Conference, FC 2013, Okinawa, Japan, April 1-5, 2013, Revised Selected Papers (Lecture Notes in Computer Science)*, Ahmad-Reza Sadeghi (Ed.), Vol. 7859. Springer, 162–170. https://doi.org/10.1007/978-3-642-39884-1_13
- [8] Andy Greenberg. 2017. After 3 Years, Why Gmail's End-to-End Encryption Is Still Vapor. <https://www.wired.com/2017/02/3-years-gmails-end-end-encryption-still-vapor/>. Last accessed: August 29, 2018.
- [9] Paul Grubbs, Jiahui Lu, and Thomas Ristenpart. 2017. Message Franking via Committing Authenticated Encryption. In *Advances in Cryptology - CRYPTO 2017 - 37th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 20-24, 2017, Proceedings, Part III (Lecture Notes in Computer Science)*, Jonathan Katz and Hovav Shacham (Eds.), Vol. 10403. Springer, 66–97. https://doi.org/10.1007/978-3-319-63697-9_3
- [10] Bryan Klimt and Yiming Yang. 2004. Introducing the Enron Corpus. In *CEAS 2004 - First Conference on Email and Anti-Spam, July 30-31, 2004, Mountain View, California, USA*. <http://www.ceas.cc/papers-2004/168.pdf>
- [11] Thomas Kluyver, Benjamin Ragan-Kelley, Fernando Pérez, Brian E. Granger, Matthias Bussonnier, Jonathan Frederic, Kyle Kelley, Jessica B. Hamrick, Jason Grout, Sylvain Corlay, Paul Ivanov, Damián Avila, Safia Abdalla, Carol Willing, and et al. 2016. Jupyter Notebooks - a publishing format for reproducible computational workflows. In *Positioning and Power in Academic Publishing: Players, Agents and Agendas, 20th International Conference on Electronic Publishing, Göttingen, Germany, June 7-9, 2016*, Fernando Loizides and Birgit Schmidt (Eds.). IOS Press, 87–90. <https://doi.org/10.3233/978-1-61499-649-1-87>
- [12] B. Laurie, Langley A., and E. Kasper. 2013. Certificate transparency. RFC 6962. <https://rfc-editor.org/rfc/rfc6960.txt>
- [13] Wendy Lee. 2017. Yahoo, Google still working on end-to-end encryption for email. <https://www.sfchronicle.com/business/article/Yahoo-Google-still-working-on-end-to-end-10872573.php>. Last accessed: August 29, 2018.
- [14] Jure Leskovec, Kevin J. Lang, Anirban Dasgupta, and Michael W. Mahoney. 2009. Community Structure in Large Networks: Natural Cluster Sizes and the Absence of Large Well-Defined Clusters. *Internet Mathematics* 6, 1 (2009), 29–123. <https://doi.org/10.1080/15427951.2009.10129177>
- [15] Marcela S. Melara, Aaron Blankstein, Joseph Bonneau, Edward W. Felten, and Michael J. Freedman. 2015. CONIKS: Bringing Key Transparency to End Users. In *24th USENIX Security Symposium, USENIX Security 15, Washington, D.C., USA, August 12-14, 2015*, Jaeyeon Jung and Thorsten Holz (Eds.). USENIX Association, 383–398. <https://www.usenix.org/conference/usenixsecurity15/technical-sessions/presentation/melara>
- [16] Silvio Micali, Michael O. Rabin, and Salil P. Vadhan. 1999. Verifiable Random Functions. In *40th Annual Symposium on Foundations of Computer Science, FOCS '99, 17-18 October, 1999, New York, NY, USA*. IEEE Computer Society, 120–130. <https://doi.org/10.1109/SFCS.1999.814584>
- [17] Satoshi Nakamoto. 2008. Bitcoin: A peer-to-peer electronic cash system.
- [18] Dimitrios Papadopoulos, Duane Wessels, Shumon Huque, Moni Naor, Jan Vcelák, Leonid Reyzin, and Sharon Goldberg. 2017. Can NSEC5 be practical for DNSSEC deployments? *IACR Cryptology ePrint Archive* 2017 (2017), 99. <http://eprint.iacr.org/2017/099>
- [19] Torben P. Pedersen. 1991. Non-Interactive and Information-Theoretic Secure Verifiable Secret Sharing. In *Advances in Cryptology - CRYPTO '91, 11th Annual International Cryptology Conference, Santa Barbara, California, USA, August 11-15, 1991, Proceedings (Lecture Notes in Computer Science)*, Vol. 576. Springer, 129–140. https://doi.org/10.1007/3-540-46766-1_9
- [20] petlib [n. d.]. a Python library that implements a number of Privacy Enhancing Technologies. <https://github.com/gdanezis/petlib>. Last accessed: August 29, 2018.

- [21] Claus-Peter Schnorr. 1991. Efficient Signature Generation by Smart Cards. *J. Cryptology* 4, 3 (1991), 161–174. <https://doi.org/10.1007/BF00196725>
- [22] O. Tange. 2011. GNU Parallel - The Command-Line Power Tool. *The USENIX Magazine* 36, 1 (Feb 2011), 42–47. <http://www.gnu.org/s/parallel>
- [23] Raphael R. Toledo, George Danezis, and Ian Goldberg. 2016. Lower-Cost ϵ -Private Information Retrieval. *PoPETs 2016*, 4 (2016), 184–201. <https://doi.org/10.1515/popets-2016-0035>
- [24] Philip R. Zimmermann. 1995. *The Official PGP User's Guide*. MIT Press, Cambridge, MA, USA.

A UNIQUE-RESOLUTION KEY-VALUE MERKLE TREE

Our unique-resolution key-value Merkle tree data structure is composed of two types of nodes:

Internal = (pivot, left : $H(\text{Node})$, right : $H(\text{Node})$)

Leaf = (key, value)

We denote the root of a tree as MTR. Each Internal node contains a pivot string and the hashes of its two children. The invariant of the structure is that any nodes in the left sub-tree will have pivots or leaf keys smaller than the parent pivot, and any nodes to the right sub-tree have pivots or leaf keys equal or larger than the parent pivot. As in a normal Merkle tree, the hash of the root node is a succinct authenticator committing to the full sub-tree (subject to the security of the hash function).

A proof of inclusion, or authentication proof, of a key-value pair in the tree involves disclosing the full resolution path of nodes from the root of the tree to the sought leaf. We show that such path is indeed a proof of inclusion, and, moreover, is unique in Section A.2.

A.1 Algorithms

Building the tree. To build a tree from a set of key-value pairs $S = \{\dots, (k_i, v_i), \dots\}$ we run the BUILDTREE procedure (Algorithm 1) The procedure take as input a set of claims S and a content-addressable store. It constructs the tree nodes and saves them to the store. Finally, it returns the hash of the root node of the resulting tree.

Querying the tree. The tree querying procedure QUERYTREE is described in Algorithm 2. It takes as input the tree root MTR and store that contains the tree nodes. The procedure traverses the tree starting from the root. For each intermediate node, the procedures follows a left or right sub-tree depending on the pivot field. It continues until it ends up in a leaf node. If the leaf node has the correct key, QUERYTREE returns the corresponding value, otherwise it returns \perp .

A.2 Unique resolution

For a given key, only one value can be stored in the tree. Any violation of this invariant will be detected when the tree is queried—thus the creator of the tree does not need to be trusted to enforce this invariant. More formally, for a given key k it is only possible to successfully prove the inclusion of one leaf node in the tree with root MTR. We capture this notion in the UniqRes game in Experiment 1. The following theorem states that no adversary can win this game.

Theorem 1. For any probabilistic polynomial time adversary \mathcal{A} it holds that $\Pr_{\mathcal{A}}[b = 1] = \text{negl}(\lambda)$, where the bit $b \in \{0, 1\}$ is the output of the UniqRes game (Experiment 1).

Algorithm 1 Tree construction

```

procedure BUILDTREE( $S$ , store)
  if  $|S| = 1$  then
     $\{(k, v)\} \leftarrow S$ 
    leaf  $\leftarrow$  Leaf( $k, H(v)$ )
    PUT(store, leaf)
    PUT(store,  $v$ ) ▷ Put the value itself into the store
    return  $H(\text{leaf})$ 
  else
     $(k^*, v^*) \leftarrow_{\S} S$  ▷ Pick the pivot arbitrarily
     $(S^-, S^+) \leftarrow$  PARTITION( $k^*, S$ )
    left  $\leftarrow$  BUILDTREE( $S^-$ , store)
    right  $\leftarrow$  BUILDTREE( $S^+$ , store)
    node  $\leftarrow$  Internal( $k^*$ , left, right)
    store.PUT(node)
    return  $H(\text{node})$ 

procedure PARTITION( $k^*, S$ )
   $S^-, S^+ \leftarrow \{\}, \{\}$ 
  for  $(k, v)$  in  $S$  do
    if  $k < k^*$  then ▷ Lexicographic comparison of strings
       $S^- \leftarrow S^- \cup \{(k, v)\}$ 
    else
       $S^+ \leftarrow S^+ \cup \{(k, v)\}$ 
  return  $(S^-, S^+)$ 

```

Algorithm 2 Querying the tree

```

procedure QUERYTREE(MTR,  $k$ , store)
   $\pi \leftarrow$  GETPATH(MTR,  $k$ , store)
   $[\dots, \text{Leaf}(k', v)] \leftarrow \pi$ 
  if  $k' = k$  then
    return  $\perp$ 
  else
    return  $v$ 

procedure GETPATH( $h, k$ , store)
  node  $\leftarrow$  store.GET( $h$ )
  if node is Leaf then
    return [node]
  else if node is Internal(pivot, left, right) then
    if  $k < \text{pivot}$  then
       $\pi \leftarrow$  GETPATH(left,  $k$ , store)
    else
       $\pi \leftarrow$  GETPATH(right,  $k$ , store)
  return [node] +  $\pi$  ▷ Prepend the current node to the list  $\pi$ 

```

PROOF. Assume \mathcal{A} wins the game. Then it is able to construct two stores such that there are two different valid paths:

$$\begin{aligned} \pi &\leftarrow \text{GETPATH}(\text{MTR}, k, \text{store}) \\ \pi' &\leftarrow \text{GETPATH}(\text{MTR}, k, \text{store}') \end{aligned}$$

that start with the same root MTR, but end with different leaves containing (k, v) and (k, v') respectively.

Experiment 1 Unique Resolution

 $\text{UniqRes}^{\mathcal{A}}(\lambda)$ $MTR, k, \text{store}, \text{store}' \leftarrow \mathcal{A}()$ **if** $\text{store} = \text{store}'$ **then****return** 0 $v \leftarrow \text{QUERYTREE}(MTR, k, \text{store})$ $v' \leftarrow \text{QUERYTREE}(MTR, k, \text{store}')$ $b \leftarrow v \neq v'$ **return** b .

First, assume one of the paths, w.l.o.g. π , consists of a single leaf node t with (k, v) . Then the other path π' can contain either another leaf t' with (k', v) , or start with an internal node t' . This implies a hash collision, since $t \neq t'$, but $MTR = H(t) = H(t')$. By the collision resistance property of the cryptographic hash function H , this happens with negligible probability.

Now, assume that the paths have a common beginning. Let t, t' be the first nodes along the paths that differ, and let $t^* = \text{Internal}(p^*, h_l^*, h_r^*)$ be their common parent. Then, there are four possible options:

- Both t and t' are a left child of t^* . In this case, $H(t) = H(t') = h_l^*$. This implies a hash collision, which we assume to happen with negligible probability.
- Both t and t' are a right child of t^* . This is analogous to the previous case.
- The children t and t' are respectively the left child and the right child of t^* . This situation cannot happen, because Get-Path decides which child to follow based on the value of the pivot p^* and the lookup key k . Since the parent is common, the procedure will always choose either the left, or the right child.
- The children t and t' are respectively the right child and the left child of t^* . This is analogous to the previous case.

Thus, the probability that \mathcal{A} wins the game, $\Pr_{\mathcal{A}}[b = 1]$, equals the probability of a hash collision and is therefore negligible. \square

B SECURITY OF THE CLAIMCHAIN DATA STRUCTURE

B.1 Privacy

Here we formally describe the privacy properties of ClaimChains.

Claim privacy. The adversary cannot learn anything about the content claims for which it does not have the corresponding capabilities.

We formalize this in Experiment 2 using an indistinguishability game. The game models that the adversary cannot distinguish between a claim containing one of two equal-length messages of its choice. The experiment is executed by a challenger that plays a game with the adversary \mathcal{A} .

The game starts with creating a user that represents an honest reader, and another user that represents the challenger. We then provide the adversary with an oracle access that allows it to create users and request them to extend their chains with adversary-supplied claims and access control sets (see Algorithm 3). Moreover, the adversary is allowed to modify store.

Algorithm 3 Add user and extend chain oracles

▷ Add a new user

procedure AU(id) $(\text{sk}_{\text{sig}}^{\text{id}}, \text{pk}_{\text{sig}}^{\text{id}}) \leftarrow \text{Sig.KeyGen}(1^\lambda)$ $(\text{sk}_{\text{DH}}^{\text{id}}, \text{pk}_{\text{DH}}^{\text{id}}) \leftarrow \text{DH.KeyGen}(1^\lambda)$ $(\text{sk}_{\text{VRF}}^{\text{id}}, \text{pk}_{\text{VRF}}^{\text{id}}) \leftarrow \text{VRF.KeyGen}(1^\lambda)$ $\text{keys}^{\text{id}} \leftarrow (\text{sk}_*^{\text{id}}, \text{pk}_*^{\text{id}})$ $(\text{sk}'_{\text{sig}}^{\text{id}}, \dots) \leftarrow \text{keys}^{\text{id}}$

▷ Separately record the signing key

▷ Extend the chain of an existing user

procedure EC(id, data, claims, acs, store)**if** user id does not exist **then return** \perp $\text{ptr}^{\text{id}} \leftarrow \text{EXTENDCHAIN}(\text{data, claims, acs, keys}^{\text{id}} \cup \text{sk}'_{\text{sig}}^{\text{id}}, \text{ptr}^{\text{id}}, \text{store})$ **return** ptr^{id}

Eventually, the adversary outputs two claims (l_0, m_0) and (l_1, m_1) . The challenger flips a random coin b , and constructs a challenge block containing claim (l_b, m_b) , readable by the honest reader, but not by the adversary. The adversary then has to guess which of the two challenge claims were included in the challenge block. It may make further oracle queries.

Note that this definition implies that the adversary cannot learn anything about the claim neither from the claim encoding itself, not from any of the capabilities. Additionally, the adversary could have access to the claim in the past, but not in the challenge block.

The proof of knowledge π in the claim encoding c depends on the claim key k and other public values, making it difficult to prove directly that the adversary cannot learn anything about the bit b . Therefore, in one of the steps we replace this proof π with a completely random proof. The following lemma states that we may do so.

LEMMA 2. *To any distinguisher that does not know the value $k_\pi \in \{0, 1\}^{2\lambda}$, the proof π in ENCCCLAIM is indistinguishable from a randomly drawn proof in the random oracle model for H_q .*

PROOF. Without loss of generality, we focus on a simpler proof with only a single conjunct, writing m for $l \parallel \text{nonce}$:

$$\pi \leftarrow \text{SPK}\{(\text{sk}_{\text{VRF}}) : \text{pk}_{\text{VRF}} = g^{\text{sk}_{\text{VRF}}} \wedge h = \text{VRF.Eval}(\text{sk}_{\text{VRF}}, m)\}(k_\pi).$$

Which abbreviates the following proof:

$$\pi \leftarrow \text{SPK}\{(\text{sk}_{\text{VRF}}) : \text{pk}_{\text{VRF}} = g^{\text{sk}_{\text{VRF}}} \wedge h = H_{\mathbb{G}}(m)^{\text{sk}_{\text{VRF}}}\}(k_\pi).$$

To construct this proof, pick a randomizer $r_{\text{sk}} \leftarrow \mathbb{Z}_q$, and compute

$$R_{\text{pk}} = g^{r_{\text{sk}}}$$

$$R_h = H_{\mathbb{G}}(m)^{r_{\text{sk}}}$$

$$c = H_q(g \parallel H_{\mathbb{G}}(m) \parallel \text{pk}_{\text{VRF}} \parallel h \parallel R_{\text{pk}} \parallel R_h \parallel k_\pi)$$

$$s_{\text{sk}} = r_{\text{sk}} + c \cdot \text{sk}_{\text{VRF}}.$$

The proof is then given by (c, s_{sk}) . To verify the proof, compute

$$R'_{\text{pk}} = g^{s_{\text{sk}}} \text{pk}_{\text{VRF}}^{-c}$$

$$R'_h = H_{\mathbb{G}}(m)^{s_{\text{sk}}} h^{-c},$$

Experiment 2 Claim privacy

 $\text{ClaimPriv}^{\mathcal{A}}(\lambda)$

.....
 Setup

AU('reader') ▷ Initialize reader's chain

AU('challenger') ▷ Initialize challenger's chain

.....
 Content to include in the challenge block

$(l_0, m_0), (l_1, m_1), \text{data}, \text{claims}, \text{acs}, \text{store} \leftarrow$
 $\leftarrow \mathcal{A}^{\text{EC}(\cdot), \text{AU}(\cdot)}(\text{pk}_{\text{DH}}^{\text{reader}'})$

if l_0 or l_1 in acs or $|m_0| \neq |m_1|$ **then return 0**

.....
 Challenge block

$b \leftarrow \$_\{0, 1\}$

$\text{claims}' \leftarrow \text{claims} \cup \{(l_b, m_b)\}$

$\text{acs}' \leftarrow \text{acs} \cup \{(\text{pk}_{\text{DH}}^{\text{reader}'}, l_b)\}$ ▷ Give the reader the access to l_b

$\text{ptr}_C \leftarrow \text{EC}(\text{'challenger'}, \text{data}, \text{claims}', \text{acs}', \text{store})$

.....
 Response

$\hat{b} \leftarrow \mathcal{A}^{\text{EC}(\cdot), \text{AU}(\cdot)}(\text{ptr}_C)$

return $\hat{b} = b$

and verify that c equals $H_q(g \parallel H_{\mathbb{G}}(m) \parallel \text{pk}_{\text{VRF}} \parallel h \parallel R'_{\text{pk}} \parallel R'_h \parallel k_\pi)$.

Suppose that the adversary does not know k_π . To randomly generate the proof, draw $(c', s'_{\text{sk}}) \leftarrow \$_{\mathbb{Z}_q^2}$ at random. Since the adversary does not know k_π it can never query the random oracle H_q with the correct value for k_π , therefore it cannot distinguish the fake proof (c', s'_{sk}) from a real proof (c, s_{sk}) . \square

Theorem 3 (Claim privacy). For any probabilistic polynomial time adversary \mathcal{A} it holds that $\Pr[b = 1] \leq \frac{1}{2} + \text{negl}(\lambda)$, where $b \in \{0, 1\}$ is the result of CLAIMPRIV game (Experiment 2) run with \mathcal{A} .

PROOF. We construct a sequence of games and show that \mathcal{A} can distinguish between them with negligible probability, starting with $G_0 = \text{ClaimPriv}^{\mathcal{A}}(\lambda)$.

First, we show that the adversary cannot extract any information about b from the capability entry for l_b because of security of the Diffie-Hellman key exchange and the encryption scheme.

Recall from the ENCCAP (Figure 2) and EXTENDCHAIN procedures (Figure 3) that the corresponding capability lookup key i_{cap} and the encryption key k_{cap} are given by:

$$\begin{aligned} i_{\text{cap}} &= H_3(s \parallel l_b \parallel \text{nonce}) \\ k_{\text{cap}} &= H_4(s \parallel l_b \parallel \text{nonce}), \end{aligned}$$

where s is the shared DH secret.

G_1 In this game we substitute the shared Diffie-Hellman secret s with the random string $\alpha \leftarrow \$_{\{0, 1\}^\lambda}$ in all capabilities for reader $\text{pk}_{\text{DH}}^{\text{reader}'}$ in all blocks on the challenger's chain. In particular, we set:

$$\begin{aligned} i_{\text{cap}} &= H_3(\alpha \parallel l_b \parallel \text{nonce}) \\ k_{\text{cap}} &= H_4(\alpha \parallel l_b \parallel \text{nonce}), \end{aligned}$$

G_2 In this game, we substitute the capability key k_{cap} with a random string $\beta \leftarrow \$_{\{0, 1\}^{2\lambda}}$. The capability becomes:

$$\text{cap} = \text{Enc}(\beta, h \parallel k \parallel k_\pi).$$

G_3 In this game, we substitute the lookup index i_{cap} with a random string $\gamma \leftarrow \$_{\{0, 1\}^{2\lambda}}$ as well.

G_4 In this game, we substitute the plaintext $h \parallel k \parallel k_\pi$ with a random string γ of the same length:

$$\text{cap} = \text{Enc}(\beta, \gamma).$$

The games G_0 and G_1 are indistinguishable by the decisional Diffie-Hellman assumption. Games G_1 and G_2 are indistinguishable by the pseudorandomness of the hash function H_4 . The indistinguishability of G_2 and G_3 follows from the pseudorandomness of H_3 . Since the encryption key β is random, distinguishing between G_3 and G_4 can be trivially reduced to the IND-CPA security for the encryption scheme. Therefore, games G_3 and G_4 are indistinguishable as well.

The adversary is not allowed to give access to labels l_0, l_1 to any user (honest or not). As a result, no other capability entries depend on the challenge bit b .

Since G_4 replaces the real plaintext with a random plaintext, the adversary also does not learn anything about k and k_π .

Now we show that the adversary cannot extract information about b neither from the claim encoding, nor from the claim lookup key. We use the IND-CPA security of the encryption scheme and pseudorandomness of the VRF scheme.

Recall from the ENCCCLAIM (Figure 2) and EXTENDCHAIN procedures (Figure 3) that here the encoded claim c is given by:

$$c = \text{Enc}(k, \pi \parallel m_b) \parallel \text{com}.$$

G_5 In this game, we replace the non-interactive zero-knowledge proof π with a uniformly random proof π' that does not depend on any of the secret values, nor on any of the public values.

G_6 In this game, we replace the commitment com by a random commitment $\text{com}_R \leftarrow \mathbb{G}$.

Games G_4 and G_5 are indistinguishable because of Lemma 2. Since π_{com} no longer depends on the randomness r , the commitment com is perfectly hiding. Therefore, games G_5 and G_6 are indistinguishable as well.

Next, we change the claim encryption key k to a random key. Note that because of the changes made in G_4 , the adversary does not learn anything about k from the capability cap .

G_7 In this game, we generate a random encryption key δ and use it to replace k :

$$c = \text{Enc}(\delta, \pi' \parallel m_b) \parallel \text{com}.$$

G_8 In this game, we replace the plaintext $\pi' \parallel m_b$ with a random message μ of the same length:

$$c = \text{Enc}(\delta, \mu) \parallel \text{com}_R.$$

Games G_6 and G_7 are indistinguishable since the adversary learns nothing about k because of earlier transformations. Games G_7 and G_8 are indistinguishable because of the CPA security of the encryption scheme.

Experiment 3 Capability-reader unlinkability

CapReaderUnlink ^{\mathcal{A}} (λ)

..... Setup

AU('challenger') ▷ Initialize challenger's chain

..... Content to include in the challenge block

$\text{pk}_{\text{DH}}^0, \text{pk}_{\text{DH}}^1, l, m, \text{data}, \text{claims}, \text{acs}, \text{store} \leftarrow \mathcal{A}^{\text{EC}(\cdot), \text{AU}(\cdot)}()$

if pk_{DH}^0 or pk_{DH}^1 not a honest user **then return 0**

..... Challenge block

$b \leftarrow \$ \{0, 1\}$

$\text{claims}' \leftarrow \text{claims} \cup \{(l, m)\}$

$\text{acs}' \leftarrow \text{acs} \cup \{(\text{pk}_{\text{DH}}^b, l)\}$

$\text{ptr}_C \leftarrow \text{EC}(\text{'challenger'}, \text{claims}', \text{acs}', \text{store})$

..... Response

$\hat{b} \leftarrow \mathcal{A}^{\text{EC}(\cdot), \text{AU}(\cdot)}(\text{ptr}_C)$

return $\hat{b} = b$

The final dependency on the bit b is in the claim lookup key $i = H_1(h_b)$, see ENCCLAIM (Figure 2). We remove this final reference.

G_9 In this game, we substitute h_b in i with a random value $q' \leftarrow \$ \mathbb{G}$:

$$i = H_1(q')$$

The changes in games G_4 and G_5 ensure that the adversary does not learn anything about h_b directly. Also, indirectly the adversary cannot learn about h_b . The adversary can learn *other* VRF values by adding claims and giving itself access to them. However, the pseudorandomness property of the VRF ensures that even if the adversary makes many VRF queries, the remaining values remain pseudorandom. Hence, the adversary cannot distinguish G_8 from G_9 .

In game G_9 none of the values depend on the challenge bit b , hence, the adversary cannot have advantage better than random guessing. \square

Capability-reader unlinkability. The adversary should not be able to determine who has been given access to a claim, i.e., for which honest user a capability has been created. We model this using the indistinguishability game in Experiment 3. The adversary can create users (using the AU oracle) and extend their chains (using the EC oracle). It then outputs the public keys pk_{DH}^0 and pk_{DH}^1 of two honest users it created using the AU and a description of a claim with label l on which it wants to be challenged. The challenger picks one of the honest users at random, and adds a capability to l for that user. The adversary must decide which user has been given the capability.

Theorem 4. For any polynomially-bounded \mathcal{A} it holds that $\Pr[b = 1] \leq \frac{1}{2} + \text{negl}(\lambda)$, where $b \in \{0, 1\}$ is the result of CapReaderUnlink game (Experiment 3).

Experiment 4 Intra-block non-equivocation

BlockNonEq ^{\mathcal{A}} (λ)

$\text{sk}_{\text{DH}}, \text{pk}_{\text{DH}} \leftarrow \text{DH.KeyGen}(1^\lambda)$

$\text{sk}'_{\text{DH}}, \text{pk}'_{\text{DH}} \leftarrow \text{DH.KeyGen}(1^\lambda)$

$l, \text{ptr}, \text{store}, \text{store}' \leftarrow \mathcal{A}(\text{pk}_{\text{DH}}, \text{pk}'_{\text{DH}})$

$m \leftarrow \text{GETCLAIM}(\text{sk}_{\text{DH}}, l, \text{ptr}, \text{store})$

$m' \leftarrow \text{GETCLAIM}(\text{sk}'_{\text{DH}}, l, \text{ptr}, \text{store}')$

return $m \neq m' \wedge m \neq \perp \wedge m' \neq \perp$

PROOF. We show that the adversary cannot extract any information about b from the capability entry for l . The adversary may have given other readers access to label l , but the corresponding capabilities are independent of the bit b , so we ignore them. We focus instead on the capability for reader pk_{DH}^b . Recall from the ENCCAP (Figure 2) and EXTENDCHAIN procedures (Figure 3) that the corresponding capability lookup key i_{cap} and the encryption key k_{cap} are given by:

$$i_{\text{cap}} = H_3(s \parallel l_b \parallel \text{nonce})$$

$$k_{\text{cap}} = H_4(s \parallel l_b \parallel \text{nonce}),$$

where s is the DH secret between the chain owner and the reader pk_{DH}^b . We apply the sequence of games G_0, \dots, G_4 in the proof of Theorem 3. The indistinguishability of the games proves that the adversary does not learn anything about the bit b . Therefore, we have capability-reader unlinkability. \square

B.2 Non-equivocation

Intra-block non-equivocation. Within a given block, a Claim-Chain owner cannot include two different claims having the same label to different readers.

We model this in Experiment 4. The adversary's task is to produce a block pointed to by ptr and a label l such that the two readers pk_{DH} and pk'_{DH} derive different claims m and m' .

Theorem 5 (Intra-block non-equivocation). For any polynomially-bounded \mathcal{A} it holds that $\Pr[b = 1] \leq \text{negl}(\lambda)$, where $b \in \{0, 1\}$ is the result of BlockNonEq game (Experiment 4).

PROOF. We first prove that both store and store' must contain the same block B . Suppose not, i.e., store contains block B whereas store' contains a different block B' that both hash to the same head ptr . Then the adversary breaks the collision resistance of H . Since H is a cryptographic hash function, this happens with negligible probability.

The remainder of this proof is also by contradiction. Assume adversary \mathcal{A} wins Experiment 4. We use the uniqueness of the VRF, first for the claim key k , then for the lookup key h , to derive a contradiction, i.e., that $m = m'$. Both readers pk_{DH} and pk'_{DH} first compute the capability lookup key (step 3), see GETCLAIM procedure (Figure 3), retrieve the capability (step 4) and decode it (step 5). Capabilities are per reader, and therefore different. We continue the proof from step 5.

Let i and i' be the claim lookup keys derived in step 5 of the GETCLAIM() call by respectively the first and second user. We first

consider the case where $i = i'$. By the unique resolution property of the tree (see Experiment 1), we know that in step 6 both `GETCLAIM()` calls must then derive the same claim encoding c with overwhelming probability.

Since the adversary wins, the derived messages m and m' are different and not \perp , therefore the calls to `DECCLAIM()` in step 7 returned different messages $m \neq m'$:

$$\begin{aligned} m &\leftarrow \text{DECCLAIM}(\text{pk}_{\text{VRF}}^O, l, h, k, k_\pi, c, \text{nonce}) \\ m' &\leftarrow \text{DECCLAIM}(\text{pk}_{\text{VRF}}^O, l, h', k', k'_\pi, c, \text{nonce}). \end{aligned}$$

Since the encoding c is the same for both m and m' , this situation is not possible by the binding property of the commitment scheme. Indeed, the users verify proofs π , respectively π' , in step 6, which verify the commitment com.

We now consider the case where the readers derive different lookup keys i and i' in step 5. Since $i \neq i'$ and by the collision resistance of H_1 , we have that the corresponding VRF values h and h' must be different as well. However, by uniqueness of the VRF, this cannot happen. More precisely, both users successfully verify the proofs π , respectively π' , in step 6, which prove that $h = \text{VRF.Eval}(\text{sk}_{\text{VRF}}, l \parallel \text{nonce}) = k'$, respectively $h' = \text{VRF.Eval}(\text{sk}_{\text{VRF}}, l \parallel \text{nonce})$, and therefore $h = h'$, contradicting the assumption that $i \neq i'$. \square

Detectable inter-block equivocation. The game in Experiment 5 models that a claim owner cannot make a non-consistent reference, yet produce a proof of consistency that validates using `CHECKCONSISTENCY()` (see Figure 6). More precisely, the adversary outputs valid blocks on two chains: the blocks $\{O_i\}_1^n$ on its own chain, and the blocks $\{C_i\}_1^t$ on the referenced chain. Moreover, the adversary outputs a label l for the referenced chain, and a valid consistency proof π_{consist} .

To win, the adversary also outputs a pointer ptr to one of its own blocks such that the challenger has access to label l . The adversary wins if the cross-referenced block m differs from the legitimate cross referenced blocks $\{C_i\}_1^t$.

Theorem 6. For any polynomially-bounded stateful \mathcal{A} it holds that $\Pr[d = \top] = \text{negl}(\lambda)$, where $d \in \{\top, \perp\}$ is the result of `DET EQ` game (Experiment 5).

PROOF. Suppose the adversary wins the game. Let i be the index such that ptr corresponds to block O_i . Since the adversary wins,

$$m = \text{GETCLAIM}(\text{sk}_{\text{DH}}^{\text{challenger}}, l, \text{ptr}, \text{store}')$$

returned a message $m \notin \{C_i\}_1^t$. Let h be the VRF hash that it computes in step 5, and let $c_i = \bar{c}_i \parallel \text{com}_i$ be the encoded claim that this algorithm retrieves in step 6. In step 7, the algorithm calls `DECCLAIM()`, to verify the proof π . Since the proof is valid, com_i commits to $H_q(m)$ and h_i is the VRF hash of $l \parallel \text{nonce}_i$.

We now show that `CHECKCONSISTENCY()` retrieves the same commitment com_i together with a proof that the committed value $x' \in \{H_q(C_i)\}_1^t$, contradicting the binding property of the commitment scheme.

The proof π_{consist} contains the VRF hash h'_i of $l \parallel \text{nonce}_i$ and the proof of correctness $\pi_h^{(i)}$. Since the proof verified, h'_i is the VRF hash of $l \parallel \text{nonce}_i$, and therefore $h'_i = h_i$. By the unique resolution

Experiment 5 Detectable inter-block equivocation

`InterBlockEqDetection` ^{\mathcal{A}} (λ)

..... Setup

`AU`(‘challenger’) ▷ Initialize the challenger’s chain

..... Adversary-supplied blocks and validation of consistency

$\{O_i\}_1^n, \{C_i\}_1^t, \text{store}, l, \pi_{\text{consist}} \leftarrow \mathcal{A}^{\text{AU}(\cdot), \text{EC}(\cdot)}()$

if `VALIDATEBLOCKS`($\{O_i\}_1^n$) = \perp **then return 0**

if `VALIDATEBLOCKS`($\{C_i\}_1^t$) = \perp **then return 0**

if `CHECKCONSISTENCY`($l, \{O_i\}_1^n, \{C_i\}_1^t, \pi_{\text{consist}}$) = \perp **then return 0**

..... Final read phase

$\text{ptr}, \text{store}' \leftarrow \mathcal{A}()$

if `GET`(store' , ptr) $\notin \{O_i\}_1^n$ **then return 0**

$m \leftarrow \text{GETCLAIM}(\text{sk}_{\text{DH}}^{\text{challenger}}, l, \text{ptr}, \text{store}')$

return $m \notin \{C_i\}_1^t$

property of the tree, `CHECKCONSISTENCY()` therefore derived the same encoded claim $c_i = \bar{c}_i \parallel \text{com}_i$ as the challenger did by calling `GETCLAIM()`. Moreover, the proof $\pi_{\text{ref}}^{(i)}$ proves that com_i commits to x' such that $x' \in \{H_q(C_i)\}_1^t$.

This contradicts the binding property of the commitment scheme or the soundness of the zero-knowledge proofs. \square