Merkle Trees

Taylor Blair

Reed College

March 20, 2023

A Brief History

- 1979 — The patent a 'Method of providing digital signatures' is filed by Ralph C. Merkle [4].
- 1999 — The original patent expires.
- 2009 — Bitcoin uses Merkle Trees for 'block header commitment.' [3]
- 2023 — Twenty students taking a cryptography class .

Definition

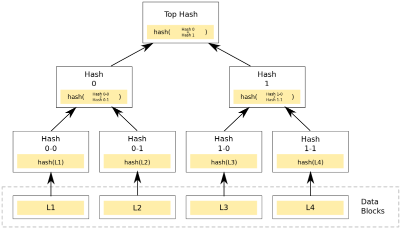

Figure: Basic Merkle Tree [8]

- Merkle trees provide proof of membership that can be publicly verified using a minimal number of hashes.
- The value of the parent for each node is the hash of the left and right child.

Applications

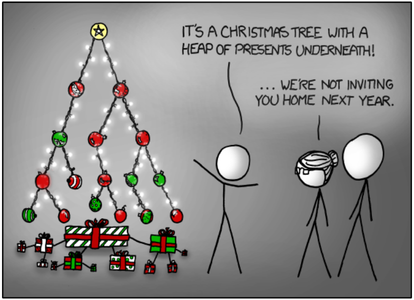

Figure: XKCD: " *Tree*" [6]

Merkle trees are secured data structures whose operations can be used to prove/verify membership of a node in $\mathcal{O}(\log(n))$ hashes.

Proving Membership (*singular*)

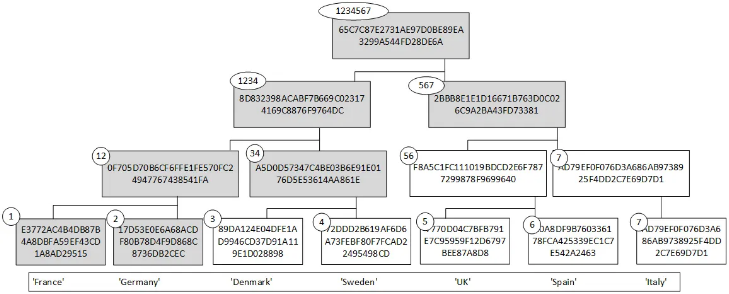

Figure: Show Germany exist in the tree [2]

Proving Membership (*multiple*)

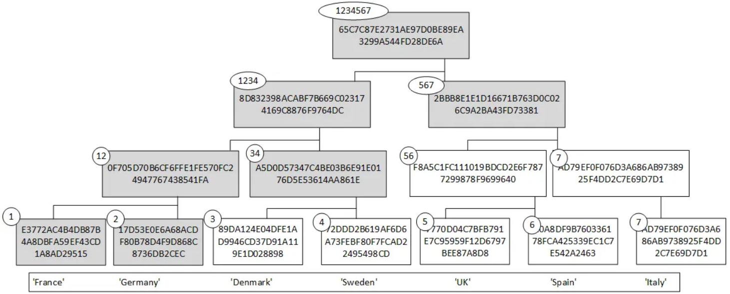

Figure: Show Germany **and** France exist in the tree [2]

Proving non-membership

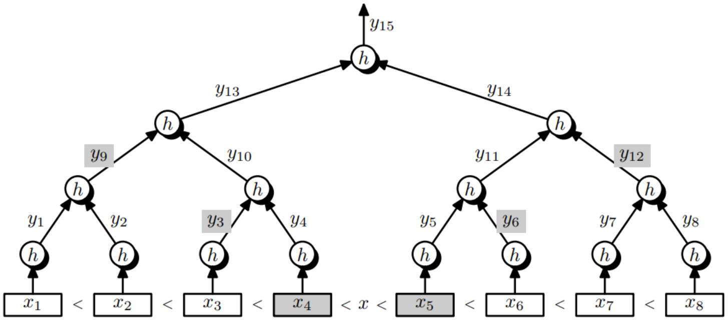

Figure: Show x is not in the tree [1]

Joining trees

See blackboard

Figure: Create a new root node and connect trees A and B [1]

Equality

See blackboard

Figure: Show trees A and B are equal.

Building a tree

input: $x_1, \ldots, x_n \in \mathcal{X}$, where n is a power of 2
output: $y \in \mathcal{Y}$

for $i = 1$ to n: $y_i \leftarrow h(x_i)$ // initialize $y_1, \ldots, y_n$
for $i = 1$ to $n - 1$: $y_{i+n} \leftarrow h(y_{2i-1}, y_{2i})$ // compute $y_{n+1}, \ldots, y_{2n-1}$

output $y_{2n-1} \in \mathcal{Y}$ [1]

Is it a secure authenticated data structure

> *We next define security. We say that an adversary defeats the scheme if it can output a hash value $y \in Y$ and then fool the verifier into accepting two different elements x and x' in X at some position i. [1]*

We assume the underlying hash function h is collision resistant.

Authenticated data structure scheme syntax

An authenticated data structure scheme $\mathcal{D} = (H, P, V)$ defined over $(\mathcal{X}^n, \mathcal{Y})$ is a tuple of three efficient deterministic algorithms:

- H is an algorithm that is invoked as $y \leftarrow H(T)$, where $T := (x_1, \ldots, x_n) \in \mathcal{X}^n$ and $y \in \mathcal{Y}$.
- P is an algorithm that is invoked as $\pi \leftarrow P(i, x, T)$, where $x \in \mathcal{X}$ and $1 \leq i \leq n$. The algorithm outputs a proof π that $x = x_i$, where $T := (x_1, \ldots, x_n)$.
- V is an algorithm that is invoked as $V(i, x, y, \pi)$ and outputs accept or reject.
- We require that for all $T := (x_1, \ldots, x_n) \in \mathcal{X}^n$, and all $1 \leq i \leq n$, we have that

$$V(i, x_i, H(T), P(i, x_i, T)) = \text{accept}$$

[1]

Attack Game

For Merkle tree $D = (H, P, V)$ defined over $(\mathcal{X}^n, \mathcal{Y})$, and a given adversary $\mathcal{A}$:

> *The adversary A outputs a $y \in \mathcal{Y}$, a position $i \in \{1, \ldots, n\}$, and two pairs (x, π) and (x', π') where $x, x' \in \mathcal{X}$.*

$\mathcal{A}$ wins the game if $x \neq x'$ and $V(i, x, y, \pi) = V(i, x', y, \pi') =$accept. Define $\mathcal{A}$'s advantage with respect to $\mathcal{D}$, denoted $\mathrm{ADSadv}[\mathcal{A}, \mathcal{D}]$, as the probability that $\mathcal{A}$ wins the game. [1]

Merkle hash tree scheme is a Secure Authenticated Data Structure Scheme

The Merkle hash tree scheme is a secure authenticated data structure scheme, assuming the underlying hash function h is collision resistant. [1]

"The proof is essentially the same as the proof of a parallel Merkle-Damgård"

8.9 (A parallel Merkle-Damgård). The Merkle-Damgård construction in Section 8.4 gives a *sequential* method for extending the domain of a secure CRHF. The tree construction in Fig. 8.16 is a parallelizable approach: all the hash functions h within a single level can be computed in parallel. Prove that the resulting hash function defined over $(\mathcal{X}^{\leq L}, \mathcal{X})$ is collision resistant, assuming h is collision resistant. Here h is a compression function $h : \mathcal{X}^2 \to \mathcal{X}$, and we assume the message length can be encoded as an element of $\mathcal{X}$. More precisely, the hash function is defined as follows:

input: $m_1 \dots m_s \in \mathcal{X}^s$ for some $1 \leq s \leq L$
output: $y \in \mathcal{X}$

let $t \in \mathbb{Z}$ be the smallest power of two such that $t \geq s$ (i.e., $t := 2^{\lceil \log_2 s \rceil}$)
for $i = s + 1$ to t: $m_i \leftarrow \perp$
for $i = t + 1$ to $2t - 1$:
  $\ell \leftarrow 2(i - t) - 1$, $r \leftarrow \ell + 1$ // *indices of left and right children*
  if $m_\ell = \perp$ and $m_r = \perp$: $m_i \leftarrow \perp$ // *if node has no children, set node to null*
  else if $m_r = \perp$: $m_i \leftarrow m_\ell$ // *if one child, propagate child as is*
  else $m_i \leftarrow h(m_\ell, m_r)$ // *if two children, hash with h*
output $y \leftarrow h\big(m_{2t-1},\ s\big)$ // *hash final output and message length*

Figure: The Merkle-Damgård proof [1].

Lessons from Bitcoin

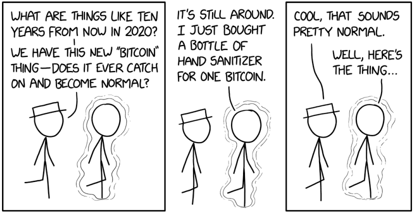

Figure: XKCD: "*2010 and 2020*" [7]

- ~~All cryptocurrencies are Ponzi schemes~~
- The *chain* is actually collection of root nodes.
- Bitcoin incorrectly implemented their merkle trees and it resulted in DOS attacks due to over hashing and duplicate nodes (CVE-2012-2459).

BitTorrent Data Integrity

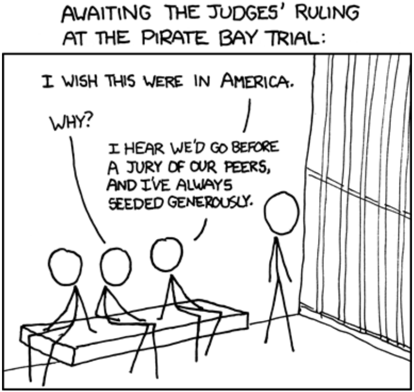

Figure: XKCD: "*Pirate Bay*" [5]

- Finding errors in $\mathcal{O}(\log(n))$!
- Only needing to compare nodes below incorrect nodes.

References I

Boneh, D., & Shoup, V. (2020). A graduate course in applied cryptography. *Draft 0.5.*

Buchannen, B. (2022). Bloom filters, merkle trees and... accumulators. https://medium.com/asecuritysite-when-bob-met-alice/bloom-filters-merkle-trees-and-accumulators-27bc2f7baf5a

Friedenbach, M., & Alm, K. (2017). Fast merkle trees proposal. https://github.com/bitcoin/bips/blob/master/bip-0098.mediawiki

Merkle, R. C. (1979). Method of providing digital signatures. *Patent US4309569A.*

Monroe, R. (2009). Xkcd: Pirate bay.

Monroe, R. (2010). Xkcd: Tree.

Monroe, R. (2020). Xkcd: 2010 and 2020.

References II

Wikipedia contributors. (2022). Merkle tree — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Merkle_tree&oldid=1123544588