

COMP4913 Capstone Project

Blockchain Unleashed: A Secure E-Voting Application

Student Name: Yiu Kam Wing

Student ID: 20028987D

Why use blockchain?

Traditional E-Voting Architecture

Traditional E-Voting Architecture

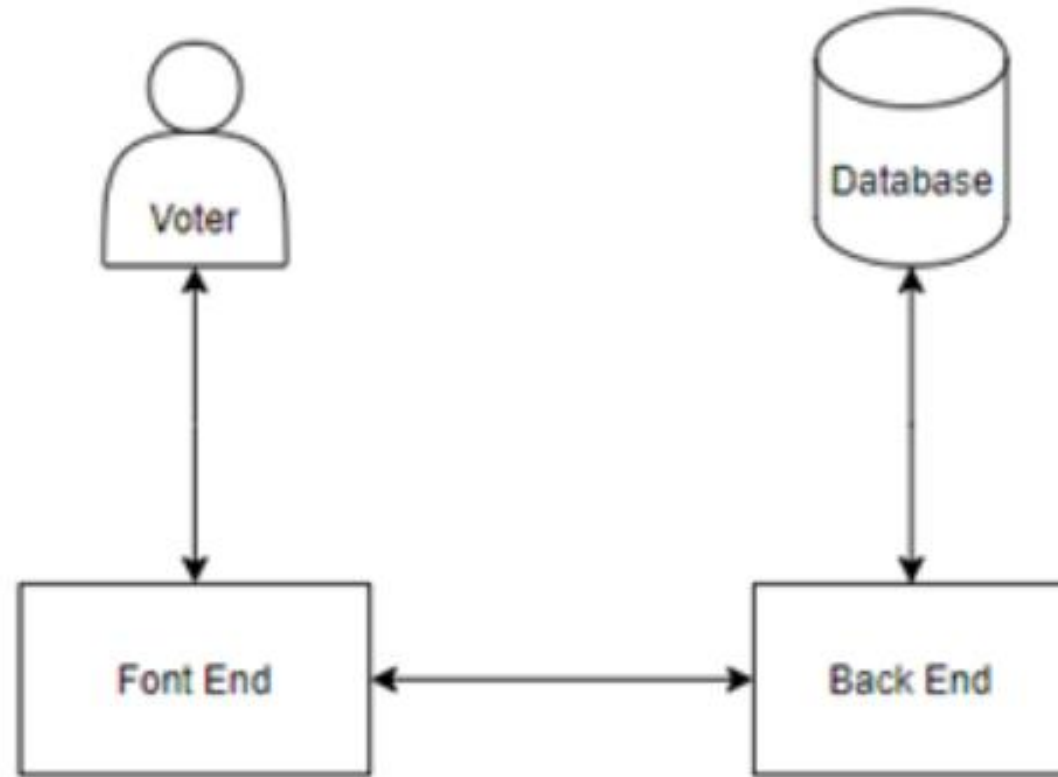
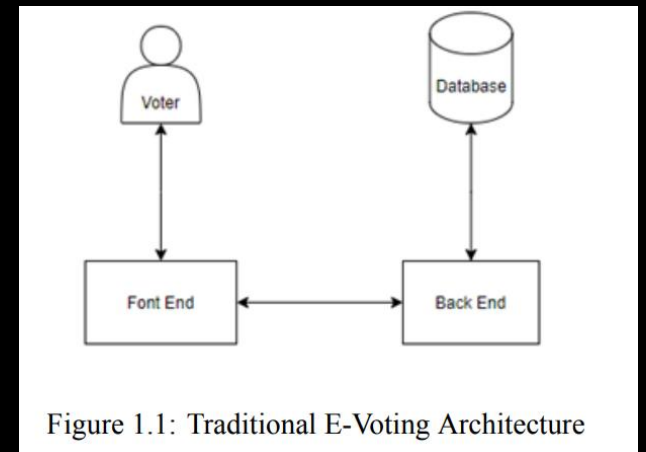


Figure 1.1: Traditional E-Voting Architecture

Problems of Traditional E-Voting Architecture

- Opacity
 - organizations or authorities have complete control over the database and system
- Single point of failure
 - The database and the server can be compromised by the hacker without anyone knowing.

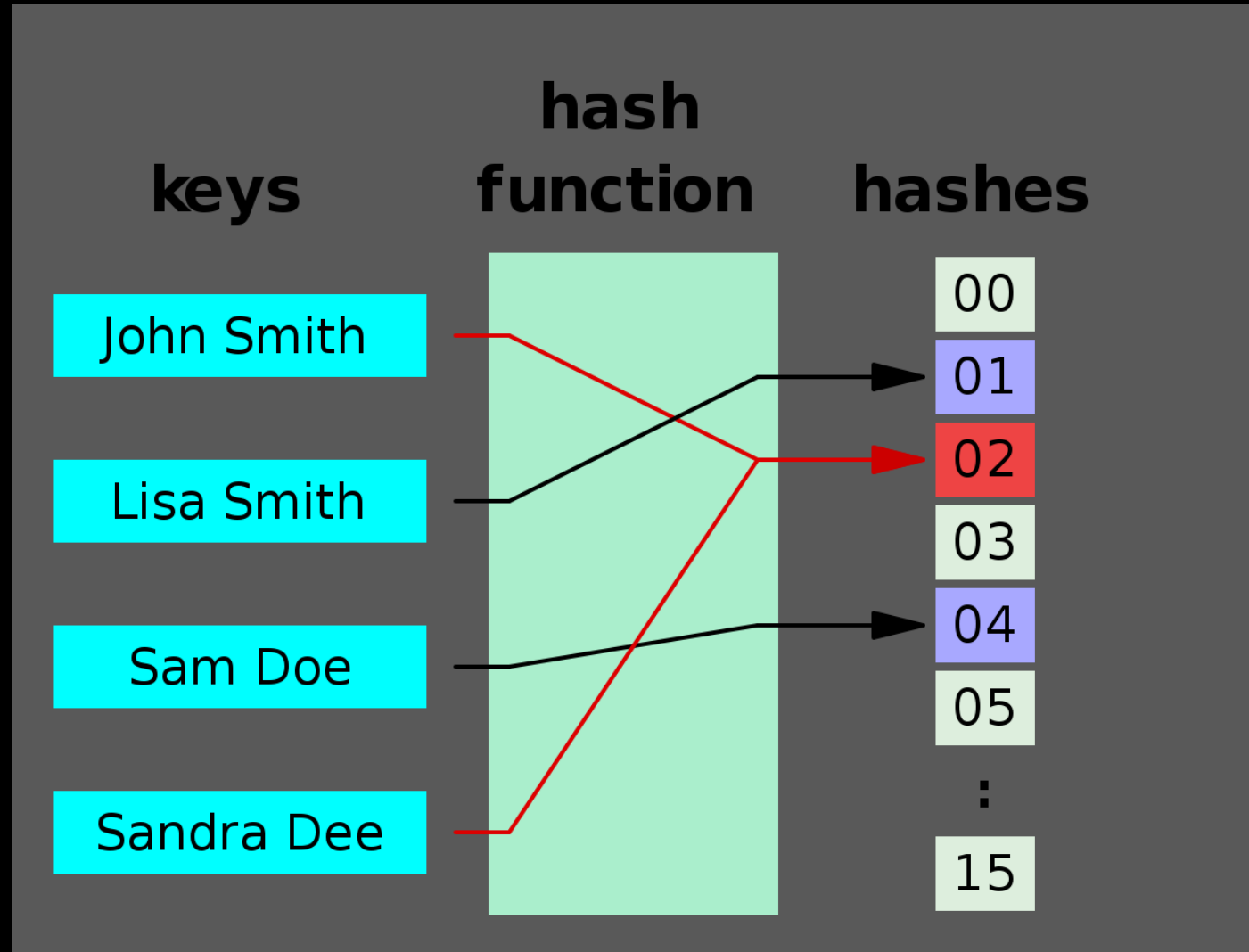
X: the credibility of the voting result



Blockchain

Hashes

- Same arbitrary length input
-> same fixed-length output
- Hard to find input from output
- Hard to find two different inputs
have same output
- efficient to compute the output



Merkle Tree

- Hash of hashes
- Leaf node:
transaction in the blockchain
- Root Hash:
summarize all the transactions
in the blockchain.

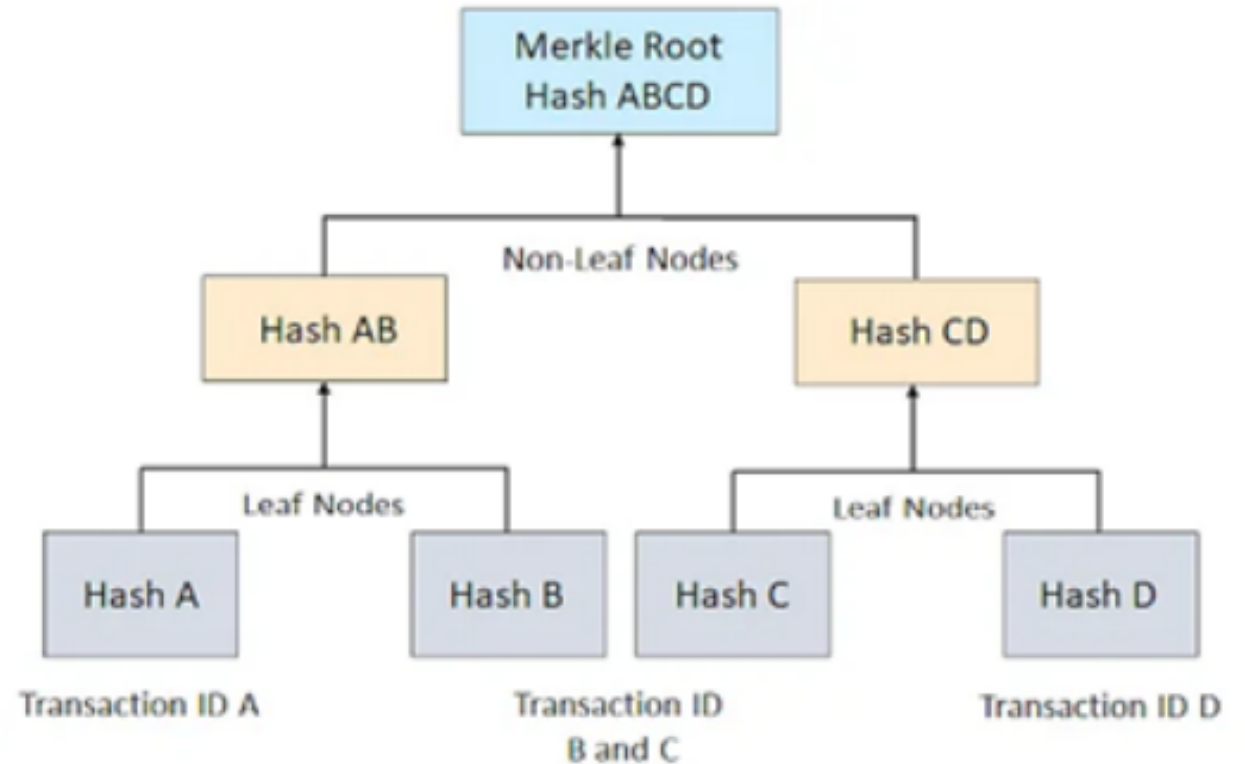
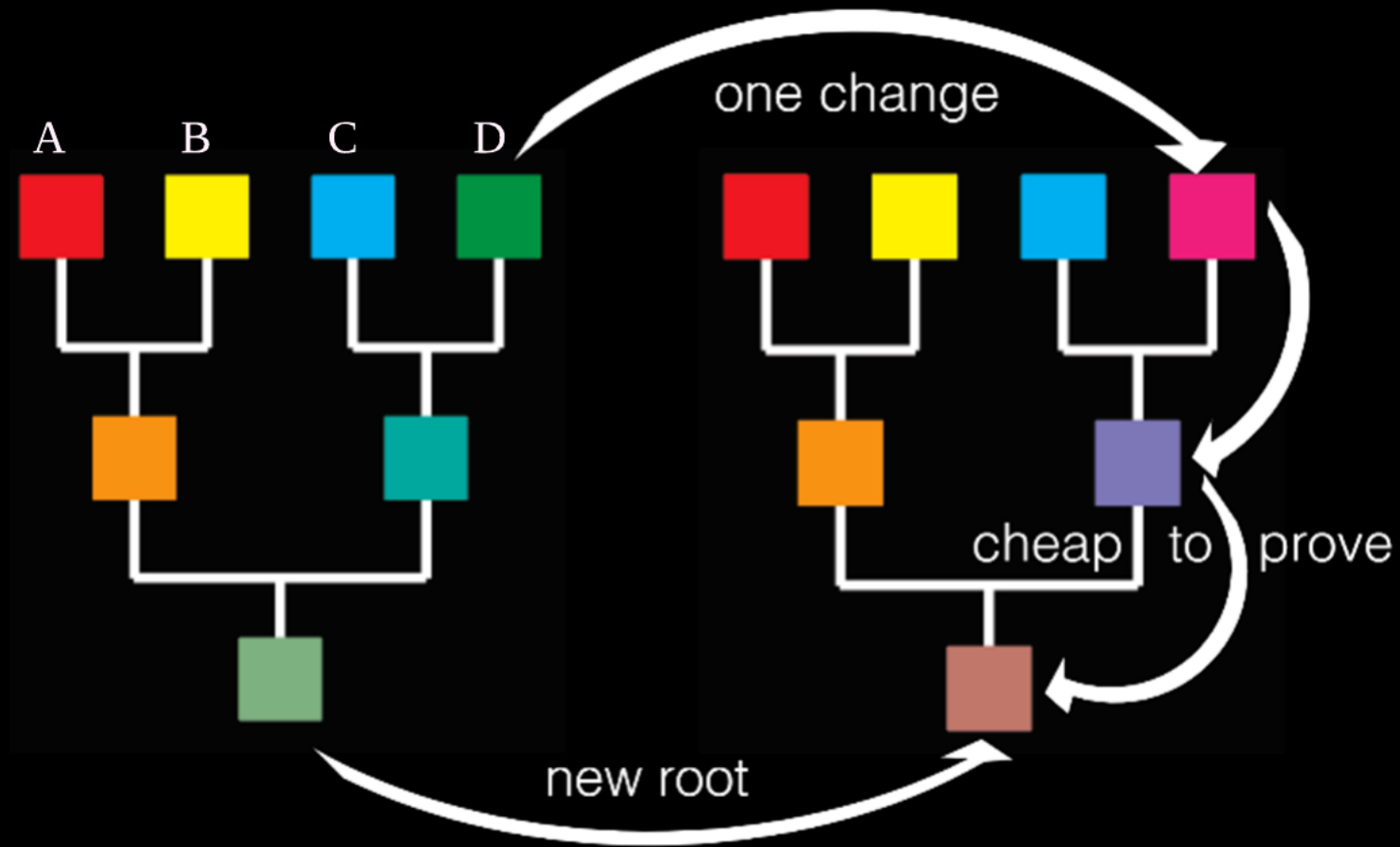


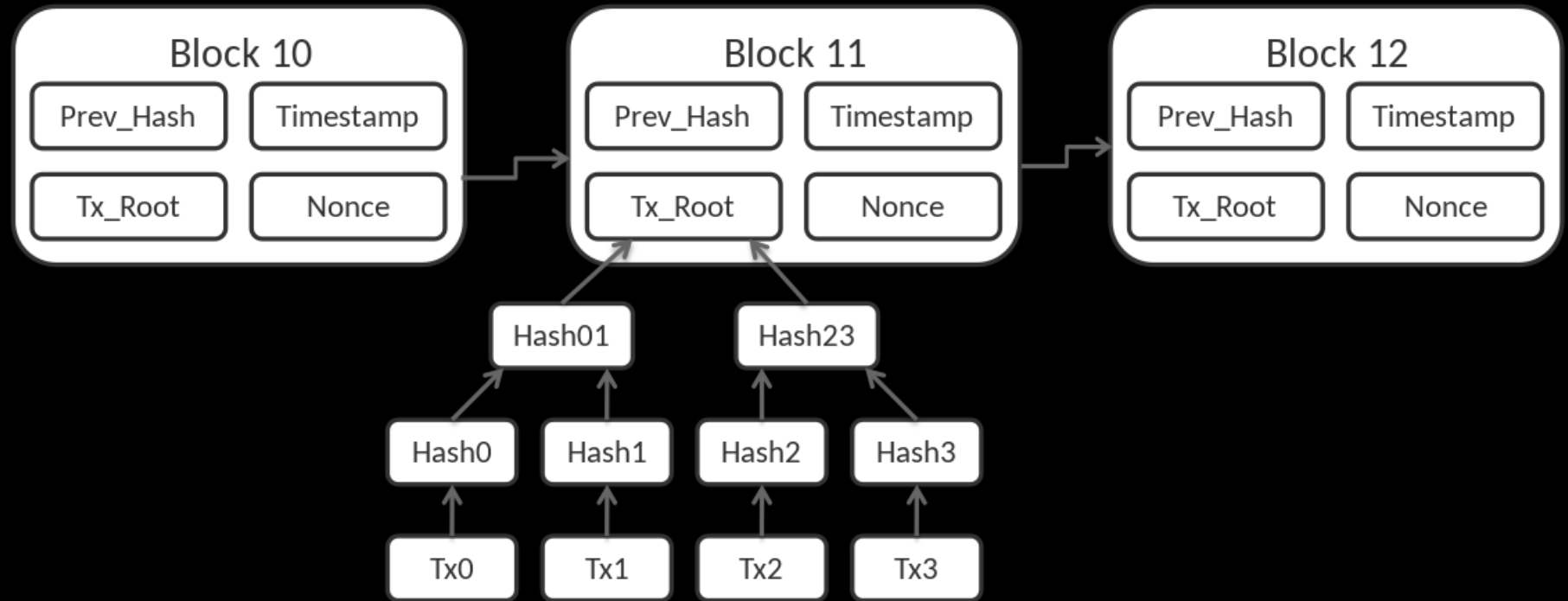
Figure 1.2: Merkle binary tree with 4 transactions

Merkle Tree



Blockchain

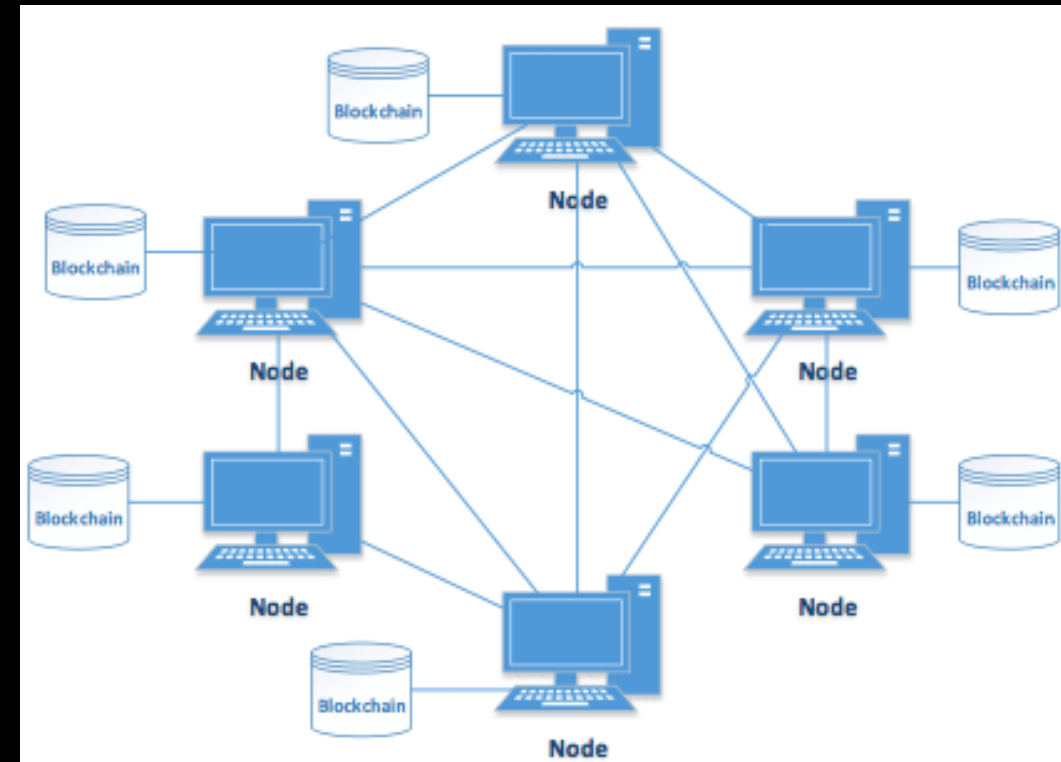
- Block of chain



Blockchain

- P2P network
 - Node: owns a copy of data in the network and can update the data in the network.
 - any update in the blockchain must be agreed upon by the consensus algorithm.

X : Dos Attacks

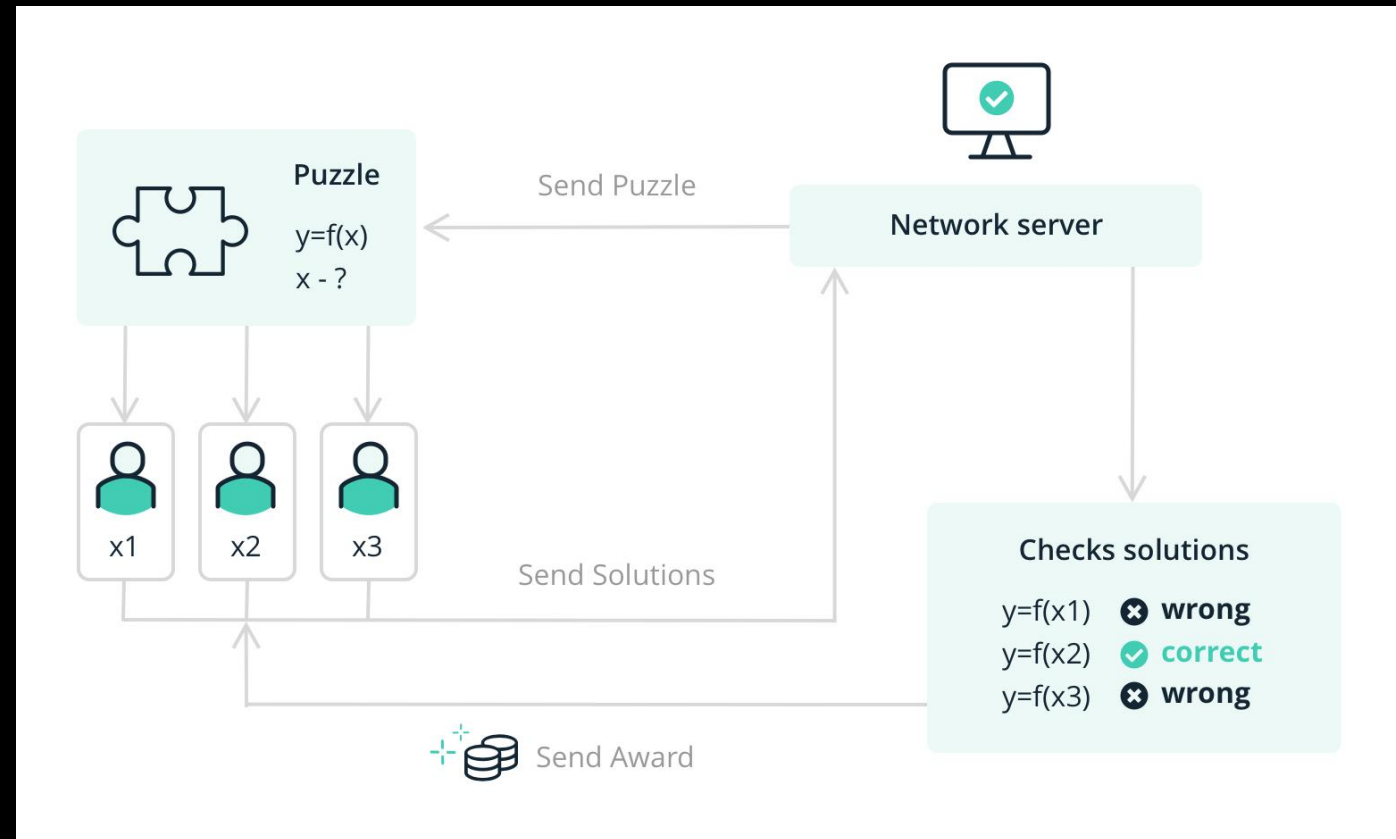


Blockchain

- Consensus Algorithm (POW)

- Puzzle:
 - Hard to solve
 - Easy to verify
- *the majority decision is represented by the longest chain*

✓ : Immutability



<https://www.ledger.com/academy/blockchain/what-is-proof-of-work>

Blockchain

- Ethereum
 - Permissionless
 - Smart Contract
 - program stored on a blockchain
 - Accounts:
 - Externally Owned Account(EOA)
 - Contract Account(CA)

Cryptography Tools

Cryptography Tools

- Public Key Cryptosystem
 - Elliptic Curve Cryptography (ECC)
- Participant Registration
 - Schnorr's Protocol
- Anonymous and Unique Voting
 - Linkable Ring Signature (LRS)
- Validating Message Sender
 - Elliptic Curve Digital Signature Algorithm (ECDSA)

Cryptography Tools

- Encryption / Decryption
 - ElGamal Encryption
 - Verifiable Decryption:
 - Chaum-Pedersen Protocol
- Key Distribution
 - Threshold Cryptosystem

Elliptic Curve Cryptography (ECC)

Elliptic Curve Cryptography (ECC)

- Elliptic Curve over Finite Field F_p
 - a plane algebraic curve that contains points $\{x, y\}$
 - Equation:
 - $y^2 = x^3 + ax + b \pmod{p}$
 - where p is a prime, $4a^3 + 27b^2 \neq 0$, $a, b, x, y \in F_p$ and an extra point O at "infinity".

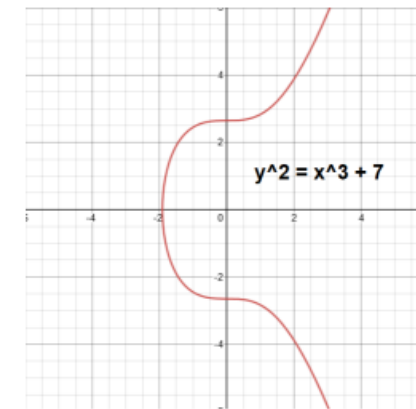


Figure 3.2: elliptic curves $E : y^2 = x^3 + 7$

Elliptic Curve Cryptography (ECC)

- Order n
 - total number of points on $E(F_p)$
- Subgroup h
 - the points on the curve are divided into h number of subgroups
 - where the order of each subgroups is r
- Generator / base point G
 - a point on $E(F_p)$ for generating other points on its subgroup by rG

Elliptic Curve Cryptography (ECC)

- Private Key
 - an integer k
- Public Key
 - point $P = kG$
- Elliptic-Curve Discrete Logarithm Problem (ECDLP)
 - computational infeasible to find k that $P = kG$

Schnorr's Protocol

Schnorr's Protocol

- prover proves the knowledge of a where $A = aG$ is public without revealing a .

Schnorr's Protocol

Prover:

Input: secret a .

1. generate random $r \in \mathbb{Z}_n$ and compute point $R = rG$.
2. compute random $c = H(G, R, A)$ where $H()$ is a cryptographic hash function.
3. compute $m = r + ac \pmod{n}$ and send $\{R, c, m\}$ to verifier.

Verifier:

Input: A and the proof $\{R, c, m\}$.

1. check $R \stackrel{?}{=} mG - cA = (r + ac)G - cA = rG + acG - cA = rG + acG - acG = rG$.

Linkable Ring Signature (LRS)

Linkable Ring Signature (LRS)

- Ring Signature
 - a group signature without a group manager and cooperation between group members that allows a signer to sign a message on behalf of the group without revealing which group member signed this message.
- LRS
 - a modification of a ring signature that detects whether the same signer generates two signatures.

LRS

Public Parameters: a list of public keys of the group members $L = \{pk_1, pk_2, \dots, pk_z\}$ where $pk_i = sk_i G$.

Signature Generation:

Input: the message $m \in \mathbb{Z}_n$, the signer's secret key $sk_i \in \mathbb{Z}_n$, L .

1. compute $H = H_2(L)$ and $\boxed{K} = sk_i H$ where $H_2()$ maps an integer to an elliptic curve point.
tag
2. generate random $c \in \mathbb{Z}_n$ and compute $u_{i+1 \pmod z} = H_1(L, K, m, cG, cH)$ where $H_1()$ is a cryptographic hash function.
3. For $j \in [1, z)$,
 - (a) compute $k = i + j \pmod z$.
 - (b) generate random $v_k \in \mathbb{Z}_p$.
 - (c) compute $u_k = H_1(L, K, m, v_k G + u_k pk_k, v_k H + u_k K)$.
4. compute $v_i = c - sk_i u_i \pmod p$.
5. return signature $\{u_1, v_1, v_2, \dots, v_z, K\}$.

LRS

Signature Verification:

Input: signature $\{u_1, v_1, v_2, \dots, v_z, K\}$, message m , and public keys L .

1. compute $H = H_2(L)$.
2. For $j \in [1, z)$,
 - (a) compute $u_{j+1} = H_1(L, K, m, v_j G + u_j pk_j, v_j H + u_j K)$.
3. check $u_1 \stackrel{?}{=} u_z$.

Elliptic Curve Digital Signature Algorithm (ECDSA)

ECDSA

- offers the functionalities of a handwritten signature and data integrity
 - verifier can determine whether the message was modified
 - and the signer of the signed message

ECDSA

Public parameters: signer(Alice)'s public key $P_a = aG$, where $a \in \mathbb{Z}_n$.

Sign (by Alice):

Input: the message M and the signer's private key a .

1. compute the message hash h by using cryptographic hash function $H()$: $h = H(M)$, where $h \in \mathbb{Z}^+$.
2. generate a random integer $k \in \mathbb{Z}_n$.
3. compute random point $R = kG$ and $r = R_x = x$ coordinate of R .
4. compute the signature proof $s = k^{-1} \times (h + ra) \pmod{n}$.
5. return signature $\{s, r\}$.

ECDSA

Verify:

Input: the message M , signature $\{s, r\}$, and the Alice's public key.

1. compute the message hash h' by using cryptographic hash function $H() : h' = H(M)$, where $h' \in \mathbb{Z}^+$.
2. compute $s_{inv} =$ the modular inverse of $s = s^{-1} \pmod{n}$.
3. recover random point $R' = (h's_{inv})G + (rs_{inv})P_a$.
4. $r' = R'_x = x$ coordinate of R' .

ElGamal Encryption

ElGamal Encryption

- asymmetric encryption scheme that encrypts a message by a one-time-key

ElGamal Encryption

Public parameters: recipient(Bob)'s public key $P_b = bG$, where $b \in \mathbb{Z}_n$.

Encryption:

1. choose a random $k \in \mathbb{Z}_n$.
2. compute $C = kG$ as public key.
3. compute $C' = kP_b$
4. map message M as point P_M on E inversely. $M = x$ coordinate of P_M
5. the ciphertext of $M = (C, D = C' + P_M)$.

Decryption (by Bob):

1. compute $C' = bC$.
2. retrieve $P_M = D - C' = C' - C' + P_M$.
3. obtain the M from P_M that $M = x$ coordinate of P_M .

Chaum-Pedersen Protocol

Chaum-Pedersen Protocol

- proof of knowledge for the equality of discrete logarithms
 - $\log_G(xG) = \log_H(xH)$
 - where G, H are two different generators on curve E

Chaum-Pedersen Protocol

Public Parameters: base points G, H , points $A = xG, B = xH$, and $n \in \mathbb{Z}_n$ where only prover knows x .

Prover:

Input: base points G, H and secret x .

1. generate random $k \in \mathbb{Z}_n$ and compute points $K = kG, L = kH$.
2. compute $c = H(K, L)$.
3. compute $r = k - xc \pmod{n}$
4. send the proof $\{r, c\}$ to verifier.

Verifier:

Input: base points G, H , points A, B , and the proof $\{r, c\}$.

1. compute $K' = rG + cA = rG + cxG = (k - xc)G + cxG = kG$.
2. compute $L' = rH + cH = rH + cxH = (k - xc)H + cxH = kH$
3. check $c \stackrel{?}{=} c' = H(K', L')$.

Threshold Cryptosystem

Polynomial

- A polynomial $f(x)$ is a mathematical expression in the form expression in the form:
 - $a_n x^n + a_{n-1} x^{n-1} + a_{n-2} x^{n-2} + \dots + a_0$
- Coefficients
 - $a_n, a_{n-1}, a_{n-2}, \dots, a_0$
- Degree
 - the highest exponent of x

Polynomial

- Polynomial Interpolation

Theorem 1 (Polynomial interpolation). *Given $d + 1$ points $(x_1, y_1), (x_2, y_2), \dots, (x_{d+1}, y_{d+1})$ where x_1, x_2, \dots, x_{d+1} are distinct numbers, there is only one polynomial $f(x)$ of degree $\leq d$ that $f(x_i) = y_i$ for $i \in [1, d + 1]$.*

- Lagrange interpolation

Theorem 2 (Lagrange interpolation). *Given $d + 1$ points $(x_1, y_1), (x_2, y_2), \dots, (x_{d+1}, y_{d+1})$ where x_1, x_2, \dots, x_{d+1} are distinct numbers, the unique polynomial $f(x)$ of degree $\leq d$ is $f(x) = \sum_{i=1}^{d+1} y_i \lambda_i(x)$ where*

$$\lambda_i(x) = \prod_{j=1, j \neq i}^{d+1} \frac{x - x_j}{x_i - x_j}.$$

Shamir Threshold Scheme

Distribution: A dealer picks a random polynomial

$$f(x) = a_tx^t + a_{t-1}x^{t-1} + a_{t-2}x^{t-2} + \dots + a_0 \pmod{p}$$

, where the coefficients a_t, a_{t-1}, \dots, a_0 and $f(x) \in \mathbb{Z}_p$ and the secret $s = f(0) = a_0$. Then the dealer sends $s_i = f(i)$ to participants P_i for $i \in [1, m]$.

Reconstruction: Any set of $t + 1$ participants can use their shares s_i to reconstruct secret s by Lagrange interpolation:

$$\sum_{i \in Q} s_i \lambda_i(0), \text{ where } \lambda_i(0) = \prod_{j \in Q, i \neq j} \frac{0 - j}{i - j} = \prod_{j \in Q, i \neq j} \frac{j}{j - i} \pmod{p}$$

Shamir Threshold Scheme

- Even t participant pool their shares together, they still cannot know the secret

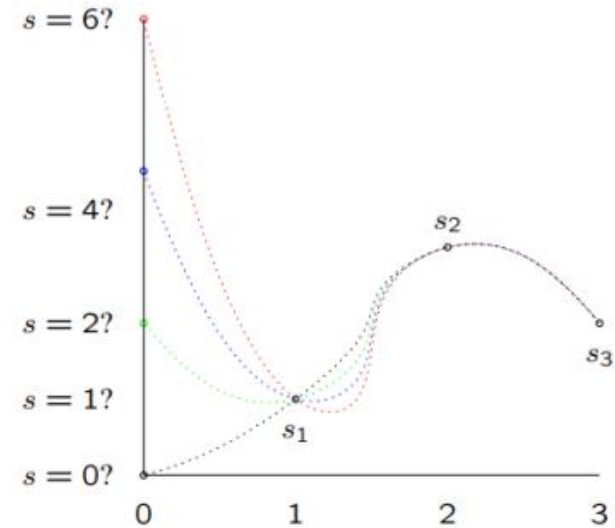


Figure 3.7: Security of Shamir's scheme illustrated: 5 degree-3 polynomials that can interpolate s_1, s_2, s_3 .

Shamir Threshold Scheme

- Problem
 - A dealer knows the secret
 - A dealer sends incorrect shares to some or all participants

Feldman VSS

- an extension of Shamir's secret sharing scheme
- the dealer not only sends the share s_i to participant P_i but also broadcasts a verification value to all participants such that the participants can use them to validate their shares

Feldman VSS

Distribution: A dealer picks a random polynomial

$$f(x) = a_t x^t + a_{t-1} x^{t-1} + a_{t-2} x^{t-2} + \dots + a_0 \pmod{n}$$

, where the coefficients a_t, a_{t-1}, \dots, a_0 and $f(x) \in \mathbb{Z}_p$ and the secret $s = f(0) = a_0$. Then the dealer sends $s_i = f(i)$ to participants P_i for $i \in [1, m]$. Furthermore, the dealer broadcasts commitments $A_j = a_j G$ for $j \in [0, t]$ to all participants. Upon receipt of share s_i , the participant P_i can verify the correctness of the share by checking the following equation:

$$s_i G = a_t i^t G + a_{t-1} i^{t-1} G + a_{t-2} i^{t-2} G + \dots + a_0 G$$

$$= A_t \cdot i^t + A_{t-1} \cdot i^{t-1} + A_{t-2} \cdot i^{t-2} + \dots + A_0$$

$$= \sum_{j=0}^t A_j \cdot i^j$$

Feldman VSS

- Problem
 - A dealer knows the secret

Threshold Cryptosystem

- No dealer
 - Each participant plays the role of the dealer

Threshold Cryptosystem

Distributed Key Generation Protocol

The distributed key generation protocol is defined as follows:

1. Each participant P_i generates a random polynomial $f_i(x) = a_{i,t}x^t + a_{i,t-1}x^{t-1} + a_{i,t-2}x^{t-2} + \dots + a_{i,0} \pmod{n}$ of degree t where all coefficients $a_{i,j} \in \mathbb{Z}_p$ and $f_i(x) \in \mathbb{Z}_n$, and broadcasts a commitment $A_{i,j} = a_{i,j}G$ for $j \in [0, t]$.
2. Each participant P_i computes the public key $H = \sum_{j=1}^m A_{j,0}$.
3. Each participant P_i executes Feldman's VSS scheme once that lets $a_{j,0}$ as the secret value. P_i plays the role of the dealer and P_j plays the role of the participant for $i \neq j$ and $j \in [1, m]$.
4. Each participant P_i receives $f_j(i)$ for $j \in [1, m]$ [Table 3.2]. Then, P_i computes share $f(i) = \sum_{j=1}^m f_j(i)$. Participant P_i verifies $f_j(i)$ by checking $f_j(i)G = \sum_{k=0}^t A_{j,k} \cdot i^k$. The public verification key $h_i = f(i)G$.

Threshold Cryptosystem

Threshold Decryption Protocol

The threshold decryption protocol works as follows:

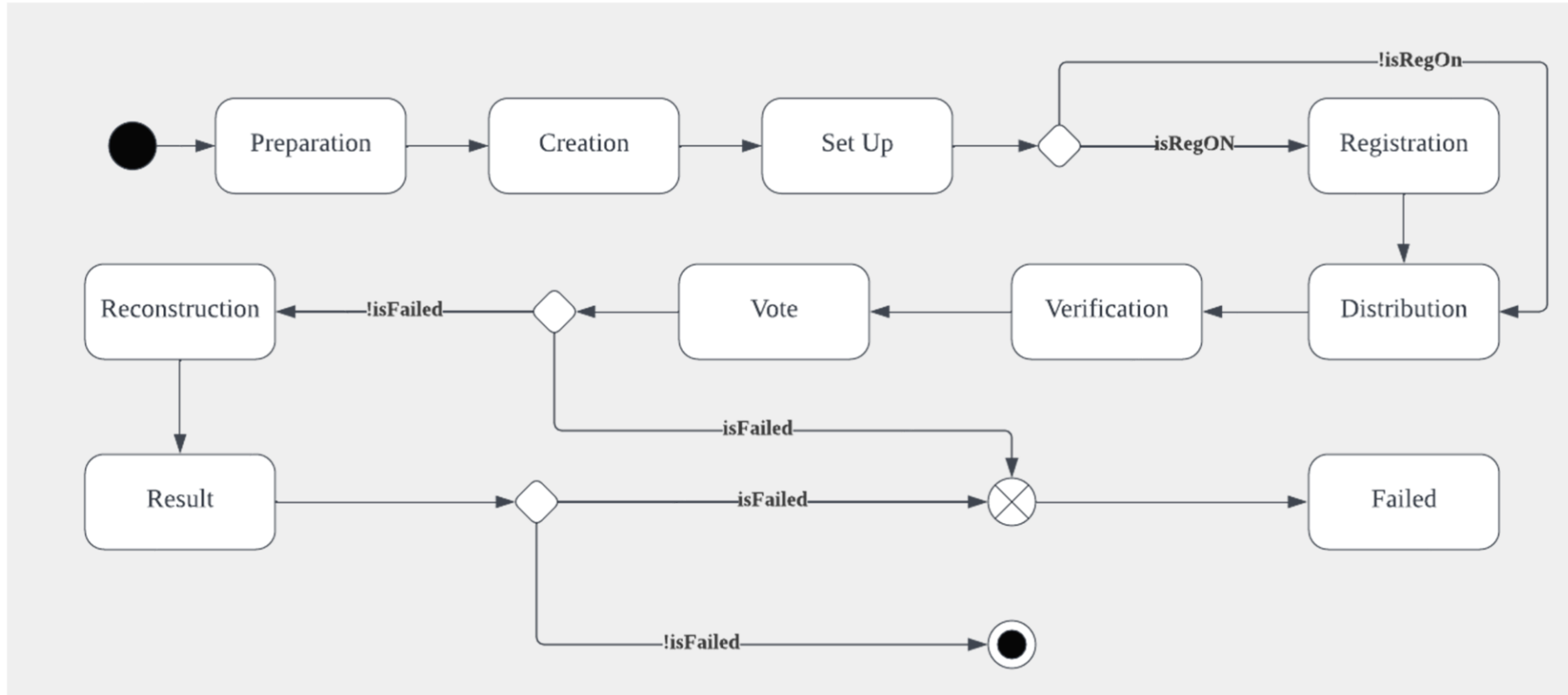
1. Each participant publishes share $f(i)$ with a proof that shows $f(i)G = h_i$.
2. A Q is a set of $t + 1$ participants publishes valid shares $f(i)$. Then the private key S can be recovered by using lagrange interpolation:

$$S = \sum_{i \in Q} f(i) \lambda_i(0), \text{ where } \lambda_i(0) = \prod_{j \in Q, i \neq j} \frac{j}{j - i} \pmod{n}$$

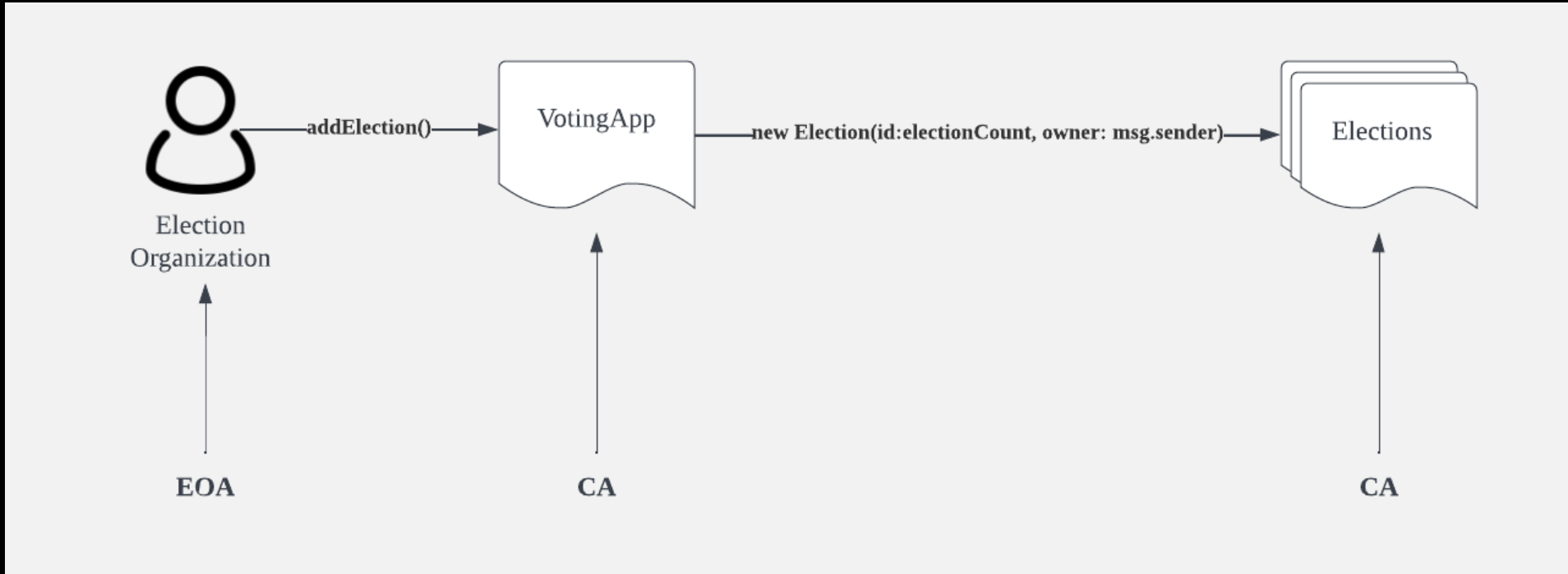
3. The ciphertext can be decrypted by using S .

Blockchain based E-Voting

An Overview of E-Voting Stages



Creation Stage



Setup

- Candidates C_i for $i \in [1, d]$
- Participants' public keys pk_i for $i \in [1, m]$
- minShares t
- isRegOn
- regInfo
 - $R_i = r_i G$ for $i \in [1, h]$ where r_i is register i 's personal data
- Timers
- ...

Registration Stage

- Register
 - Generate a key pair (sk_{m+1}, pk_{m+1})
 - Find his R_i
 - generate a schnorr proof $p = \text{schnorrProve}(r_i)$
 - Send (pk_{m+1}, i, p)
- Election
 - Verify schnorr proof $\text{schnorrVer}(R_i, p)$
 - Add pk_{m+1}
 - Update $m = m + 1$

Distribution Stage

- Each participant
 - generates a random polynomial $f_i(x)$ of degree $t - 1$ where coefficient is $a_{i,j}$ for $j \in [1, t - 1]$ and compute commitment $A_{i,j} = a_{i,j} \cdot G$
 - $\text{sig}A_{i,j}$: ECDSA signature of $A_{i,j}$ for $j \in [1, t - 1]$
 - $F_i(k)$: ElGamal ciphertext of $f_i(k)$
 - $F_i(k)$ can only be decrypted by participant P_k using his secret key sk_k
 - $\text{sig}F_i(k)$: ECDSA signature of $F_i(k)$ for $k \in [1, m]$
 - Send each $A_{i,j}, \text{sig}A_{i,j}, F_i(k), \text{sig}F_i(k)$

Distribution Stage

- Election
 - Verify signature $\text{sig}A_{i,j}$ and store $A_{i,j}$
 - Verify signature $\text{sig}F_i(k)$ and store $F_i(k)$

Verification Stage

- Report Non-contributed Participant
- Report Malicious Participant

Report Non-contributed Participant

- Election
 - check each participant P_i whether he sends his commitment $A_{i,j}$ for $j \in [1, t - 1]$
 - check each participant P_i whether he sends $F_i(k)$ to participant P_k for $k \in [1, m]$
 - Put them in disqualified participants set $\{P_i\} \cup Q$

Report Malicious Participant

- Each Participant P_i
 - Decrypt $F_j(i) = (C, D = C' + f_j(i))$ and verify $f_j(i)$ by checking $f_j(i)G = \sum_{k=0}^{t-1} A_{j,k} \cdot i^k$ for $j \in [1, m]$ and $P_j \notin Q$
 - compute proof of correct decryption key $p = \text{cpProve}(G, F_j(i).C, sk_i)$ and the decryption key is $C' = sk_i \cdot F_j(i).C$ if he receives incorrect $f_j(i)$
 - Send (C', j, i, p)

Report Malicious Participant

- Election
 - $\log_G(sk_i G) = \log_{F_j(i).C}(C') ?$
↓
 - Verify the decryption key $cpVer(G, pk_i, F_j(i).C, C')$
 - Decrypt $F_j(i)$ by using C'
 - Verify $f_j(i)$ by checking $f_j(i)G = \sum_{k=0}^{t-1} A_{j,k} \cdot i^k$ for $j \in [1, m]$ and $P_j \notin Q$
 - Put P_j in disqualified participants set $\{P_j\} \cup Q$ if $f_j(i)$ is incorrect

Vote Stage

- Set Vote Public Key
- Vote

Set Vote Public Key

- Election
 - set honest participants $P' = P - Q$
 - If $|P'| < t$, -> failed stage
 - participants' public key $pk = pk - pk_j$ for $j \in [1, m], P_j \in Q$
 - Compute public key $H = \sum_{pk} pk_i$

Vote

- Participant

- \mathbf{B}_b : $\text{elgamalEnc}(C_j, H)$

- $\text{sig}_{h\mathbf{B}_b}$: $\text{lrsSign}(h\mathbf{B}_b, \text{pk}, \text{sk}_i) = (\mathbf{u}_b \mathbf{1}, V_b, K_b)$

- Send \mathbf{B}_b and $\text{sig}_{h\mathbf{B}_b}$

of
ballots

- Election

- Verify double voting $K_b \in K$?

- Verify signature $\text{lrsVer}(h\mathbf{B}_b, \text{sig}_{h\mathbf{B}_b})$

- Store \mathbf{B}_b and $\text{sig}_{h\mathbf{B}_b}$

- $\{K_b\} \cup K$ and $b = b + 1$

Reconstruction Stage

- Participant P_i
 - $f_j(i)$: $elgamalDec(F_j(i), sk_i)$ for $j \in [1, m]$ and $P_j \notin P'$
 - Send $f_j(i)$ for $j \in [1, m]$ and $P_j \notin P'$
- Election
 - Set $f(i) = 0$
 - for $j \in [1, m]$ and $P_j \notin P$
 - $f(i) += f_j(i)$ if $f_j(i)G = \sum_{k=0}^{t-1} A_{j,k} \cdot i^k$
 - Otherwise, -> error
 - Store $f(i)$

Result Stage

- Recover Private Key
- Tally the Ballots

Recover Private Key

- T : a set of t number of participants who submitted their shares
- $S = \sum_{i \in T} f(i) \lambda_i(0)$, where $\lambda_i(0) = \prod_{j \in T, i \neq j} \frac{j}{j-i} \pmod{n}$
- Send S

Tally the Ballots

- If $|f(i)| < t$, \rightarrow failed stage
- Verify Private Key $H = S \cdot G$?
- For each B_i ,
 - $C_j = \text{elgamalDec}(B_i, S)$
 - $C_j.\text{votecount} += 1$ if $C_j \in [1, d]$

DEMO

DEMO

- Timers off
- isRegOn: True
- Participants: P_0, P_1, P_2, P_3
- # of register: 1
 - Become P_4
- Candidates: a, b
- minShares: 2

DEMO

- Non-contributed Participant: P_0
- Malicious Participant: P_2
- Vote: P_1
- Submit Shares: P_1, P_3

DEMO

- Participant Public Keys

1	04c34c400b969c4d363c8e6d248ec41805ac0169e80c6c01cfd3a657628adaad26c2c339ebdeb2c33f0d4775df425fef0d56262a9bff153692d91426161d2e95e9
2	047a981b79c0990b49fe997fe9085db734f98ba7f7bea4d018f51393dc8dc711524161183f3e2a5c543178ed4759dca01cafe5978f9f5a9c6fd6a9effdf2af1a3a
3	047c0ee3153ae36908ce2c4c0559ca07567cbbd81b0fdc97282a01a49e1f62039e116a5a71ae58ae257d46266dbac66847130bf87a87596b9656d937c6a32fc759
4	045691d781ab41d080ce17098f76f4da37d190186ebb40de293f4ef606cd85233f3b6e4513b75fd0fe2025c85a5d9df01ca5824c70659380f3ff0847709cf317d5

DEMO

- Participant Private Keys

1	Private Keys:
2	(0) 734991b8c8d3e68685bd25d65c4e2e7e842af6670841e6464803b0495fffd978
3	(1) 372d8d8646ab21fd4d22bd206fed6c0ac362a9bd583b3ed1b9f0e62a3c47be2b
4	(2) 963a8c27a8eeb06a8d082b08ee07b4c61bd5640b0b42544072a4cceb5b37e154
5	(3) 995896dd302796abf4b7b35ee49a071ceb8cd8a7aa13d2aae1b41f2959a0ddde
6	(4) 72281352b23573aa3426b5a2e625419e022db98b58afa4fae382a2e5b0e09a33