Identification and Signature Based on MQ Problem

27 February, 2023

Finite Field

- A field consists of a set of elements where we can perform addition, subtraction, multiplication and division.
  - Example: rational numbers, real numbers, ...
- A field having a finite number of elements.
  - Example: $\mathbb{Z}_p = \{0, 1, \ldots, p-1\}$ for some prime p.
- We'll denote a finite field by $\mathbb{F}_q$ and consider $\mathbb{Z}_7$ as an example.

Polynomials

$$f^{(1)}(x_1,\ldots,x_n)=\sum_{i=1}^{n}\sum_{j=i}^{n}\alpha_{ij}^{(1)}\cdot x_i x_j+\sum_{i=1}^{n}\beta_i^{(1)}\cdot x_i+\gamma^{(1)}$$

$$f^{(2)}(x_1,\ldots,x_n)=\sum_{i=1}^{n}\sum_{j=i}^{n}\alpha_{ij}^{(2)}\cdot x_i x_j+\sum_{i=1}^{n}\beta_i^{(2)}\cdot x_i+\gamma^{(2)}$$

$$\cdots$$

$$f^{(m)}(x_1,\ldots,x_n)=\sum_{i=1}^{n}\sum_{j=i}^{n}\alpha_{ij}^{(m)}\cdot x_i x_j+\sum_{i=1}^{n}\beta_i^{(m)}\cdot x_i+\gamma^{(m)}$$

$m =$ no. of equations; $n =$ no. of variables
For all i, j and k, $\alpha_{ij}^{(k)}, \beta_i^{(k)}, \gamma^{(k)} \in \mathbb{F}_q$
Degree of the polynomials in the system is $d = 2$.

Why Quadratic Polynomials

- The set of polynomials form the public key of Multivariate Public Key Cryptosystem (MPKC).
- No. of terms with degree $d = \binom{n+d-1}{d}$
- No. of terms with degree $\leq d = \binom{n+d}{d}$
- For $d \geq 2$ the public key size becomes huge.
- For efficiency, MPKC usually restrict to quadratic polynomials.
  - Why not $d = 1$?

The Hard Problem

- **MQ Problem:** You are given $\mathbf{y} = (y_1, y_2, \ldots, y_m) \in \mathbb{F}_q^m$. Your task is to find $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{F}_q^n$ such that $f^{(k)}(x_1, x_2, \ldots, x_n) = y_k$, for $1 \leq k \leq m$; if such an $\mathbf{x}$ exists.
- The decision version of the MQ Problem is known to be NP Hard.
- MQ Problem is believed to be intractable for both classical and quantum conputers.
  - Forms the basis of MPKC.

Quadratic Polynomials

$$f^{(1)}(x_1,\ldots,x_n) = \sum_{i=1}^{n}\sum_{j=i}^{n}\alpha_{ij}^{(1)}\cdot x_ix_j + \sum_{i=1}^{n}\beta_i^{(1)}\cdot x_i + \gamma^{(1)}$$

$$f^{(2)}(x_1,\ldots,x_n) = \sum_{i=1}^{n}\sum_{j=i}^{n}\alpha_{ij}^{(2)}\cdot x_ix_j + \sum_{i=1}^{n}\beta_i^{(2)}\cdot x_i + \gamma^{(2)}$$

$$\cdots$$

$$f^{(m)}(x_1,\ldots,x_n) = \sum_{i=1}^{n}\sum_{j=i}^{n}\alpha_{ij}^{(m)}\cdot x_ix_j + \sum_{i=1}^{n}\beta_i^{(m)}\cdot x_i + \gamma^{(m)}$$

$m =$ no. of equations; $n =$ no. of variables

For all i, j and k, $\alpha_{ij}^{(k)}, \beta_i^{(k)} \in \mathbb{F}_q$

We take $\gamma^{(k)} = 0$ (without any loss of security).

Polar Form

- $\mathcal{P} : \mathbb{F}_q^n \to \mathbb{F}_q^m$ be some MQ system.
- Define its polar form $\mathsf{G} : \mathbb{F}_q^n \times \mathbb{F}_q^n \to \mathbb{F}_q^m$ as:

$$\mathsf{G}(\mathbf{x}, \mathbf{y}) = \mathcal{P}(\mathbf{x} + \mathbf{y}) - \mathcal{P}(\mathbf{x}) - \mathcal{P}(\mathbf{y})$$

  where $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$.
- G is a bilinear map. [Exercise]
  - $\mathsf{G}(\mathbf{x} + \mathbf{z}, \mathbf{y}) = \mathsf{G}(\mathbf{x}, \mathbf{y}) + \mathsf{G}(\mathbf{z}, \mathbf{y})$
  - $\mathsf{G}(\mathbf{x}, \mathbf{y} + \mathbf{z}) = \mathsf{G}(\mathbf{x}, \mathbf{y}) + \mathsf{G}(\mathbf{x}, \mathbf{z})$

Three Pass IDS

Recall the definition: Key-Gen, P and V together with a challenge space ChS.

- Key-Gen: It takes as input a security parameter κ and outputs a public and private key pair (pk, sk).
- Identification: The interactive execution of $\mathsf{P}(pk, sk)$ and $\mathsf{V}(pk)$ is as follows.
  1. Prover P first sends a commitment ct to V.
  2. Verifier V then picks $ch \in_R \mathsf{ChS}$ and sends it to P.
  3. In the final pass, P sends a response rsp to V.

  Finally, V decides to accept or reject based on pk and $trans$, where $trans = (ct, ch, rsp)$ is called transcript of the protocol.

IDS Security: Zero-Knowledge

$\mathrm{IDS} = (\text{Key-Gen}, \mathsf{P}, \mathsf{V})$ is said to be honest verifier zero-knowledge (HVZK), if there exists a simulator Simu such that the following two distributions are indistinguishable.

1. $\{(pk, sk, \pi) : (pk, sk) \leftarrow \text{Key-Gen}(); \pi \leftarrow \text{Trans}(\langle \mathsf{P}(pk, sk), \mathsf{V}(pk) \rangle)\}$
2. $\{(pk, sk, \pi) : (pk, sk) \leftarrow \text{Key-Gen}(); \pi \leftarrow \text{Simu}(pk)\}$

IDS Security: Proof of Knowledge

Soundness: $\mathsf{IDS} = (\mathsf{Key\text{-}Gen}, \mathsf{P}, \mathsf{V})$ is said to be sound with knowledge error k if for all PPT adversary $\mathcal{A}$ we have

$$\Pr[b = 1 : (pk, sk) \leftarrow \mathsf{Key\text{-}Gen}; b \leftarrow \langle \mathcal{A}(pk, \perp); V(pk) \rangle] \leq k + \epsilon$$

where ϵ is negligibly close to 0.

- We want the knowledge error k to be negligibly close to 0.
- If k is non-negligible, then by running IDS r-times in parallel (IDS^r), the knowledge error becomes k^r.
    - negligible for sufficiently large r.

3-special Soundness

Consider a 3-pass identification scheme (Key-Gen, P, V) with $|\mathrm{ChS}| \geq 3$ and public-key pk.
Suppose we have three accepted transcripts of the following form:

1. $\pi = (\mathrm{ct}, \mathrm{ch}, \mathrm{rs})$
2. $\pi' = (\mathrm{ct}, \mathrm{ch}', \mathrm{rs}')$
3. $\pi'' = (\mathrm{ct}, \mathrm{ch}'', \mathrm{rs}'')$

Note: $\mathrm{ch} \neq \mathrm{ch}' \neq \mathrm{ch}'' \neq \mathrm{ch}$.

The IDS satisfies (3-special) soundness property if given these three transcripts we can have an extractor Extr that can efficiently compute a witness sk.

Sakumoto-Shirai-Hiwatari IDS

$\underline{\mathsf{P}((\mathcal{P}, \mathbf{v}), \mathbf{s})}$:		$\underline{\mathsf{V}(\mathcal{P}, \mathbf{v})}$:
pick $\mathbf{a}_0, \mathbf{b}_0 \xleftarrow{\$} \mathbb{F}^n$ and $\mathbf{c}_0 \xleftarrow{\$} \mathbb{F}^m$		
set $\mathbf{a}_1 = \mathbf{s} - \mathbf{a}_0$ and $\mathbf{b}_1 = \mathbf{a}_0 - \mathbf{b}_0$		
$\mathbf{c}_1 = \mathcal{P}(\mathbf{a}_0) - \mathbf{c}_0$		
$\mathsf{ct}_0 = H(\mathbf{a_1} \| G(\mathbf{b}_0, \mathbf{a}_1) + \mathbf{c}_0)$		
$\mathsf{ct}_1 = H(\mathbf{b}_0 \| \mathbf{c}_0)$		
$\mathsf{ct}_2 = H(\mathbf{b}_1 \| \mathbf{c}_1)$		
set $\mathsf{ct} = (\mathsf{ct}_0, \mathsf{ct}_1, \mathsf{ct}_2)$		
	$\xrightarrow{\quad \mathsf{ct} \quad}$	
		$\mathsf{ch} \xleftarrow{\$} \{0, 1, 2\}$
	$\xleftarrow{\quad \mathsf{ch} \quad}$	

SSH IDS (Contd.)

$\underline{\mathsf{P}((\mathcal{P}, \mathbf{v}), \mathbf{s})}$:

if ch $= 0$, set rs $= (\mathbf{a}_0, \mathbf{b}_1, \mathbf{c}_1)$
if ch $= 1$, set rs $= (\mathbf{a}_1, \mathbf{b}_1, \mathbf{c}_1)$
if ch $= 2$, set rs $= (\mathbf{a}_1, \mathbf{b}_0, \mathbf{c}_0)$

$$\xrightarrow{\text{rs}}$$

$\underline{\mathsf{V}(\mathcal{P}, \mathbf{v})}$:

if ch $= 0$, parse rs as $(\mathbf{a}_0, \mathbf{b}_1, \mathbf{c}_1)$ and check,
if $\mathrm{ct}_1 \stackrel{?}{=} H(\mathbf{a_0} - \mathbf{b_1} || \mathcal{P}(\mathbf{a_0}) - \mathbf{c_1})$
and $\mathrm{ct}_2 \stackrel{?}{=} H(\mathbf{b}_1 || \mathbf{c}_1)$

if ch $= 1$, parse rs as $(\mathbf{a}_1, \mathbf{b}_1, \mathbf{c}_1)$ and check,
if $\mathrm{ct}_0 \stackrel{?}{=} H(\mathbf{a}_1 || \mathbf{v} - \mathcal{P}(\mathbf{a}_1) - G(\mathbf{b}_1, \mathbf{a}_1) - \mathbf{c}_1)$
and $\mathrm{ct}_2 \stackrel{?}{=} H(\mathbf{b}_1 || \mathbf{c}_1)$

if ch $= 2$, parse rs as $(\mathbf{a}_1, \mathbf{b}_0, \mathbf{c}_0)$ and check,
if $\mathrm{ct}_0 \stackrel{?}{=} H(\mathbf{a}_1 || G(\mathbf{b}_0, \mathbf{a}_1) + \mathbf{c}_0)$
and $\mathrm{ct}_1 \stackrel{?}{=} H(\mathbf{b}_0 || \mathbf{c}_0)$

Security: HVZK

Note: For any challenge $\mathsf{ch} \in \{0, 1, 2\}$, the components of the response rs are uniformly distributed.

A simulator Simu generates proofs as follows.

- Simu chooses $\mathsf{ch} \xleftarrow{\$} \{0, 1, 2\}$ and a fake secret $\mathbf{s}' \in \mathbb{F}_q^n$.
- If $\mathsf{ch} \in \{0, 2\}$ then $\mathsf{ct} = (\mathsf{ct}_0, \mathsf{ct}_1, \mathsf{ct}_2)$ and rs are generated exactly like the protocol (but using the fake secret $\mathbf{s}'$).
- If $\mathsf{ch} = 1$, then rs and ct_2 are generated as in protocol but ct_0 is computed as is done during verification:

$$\mathsf{ct}_0 = H(\mathbf{a}_1 || \mathbf{v} - \mathcal{P}(\mathbf{a}_1) - \mathsf{G}(\mathbf{b}_1, \mathbf{a}_1) - \mathbf{c}_1)$$

Assuming H behaves like a random function, ct does not leak any information of $\mathbf{s}'$ and the corresponding response rs.

Hence, the original proof and the simulated proof are indistinguishable for the adversary.

3-Special Soundness

Assume that H is a CRHF.

- Suppose we have three accepted transcripts:

$$(\mathsf{ct}, 0, \mathsf{rs}_0); (\mathsf{ct}, 1, \mathsf{rs}_1); \quad \text{and} \quad (\mathsf{ct}, 2, \mathsf{rs}_2)$$

  - $\mathsf{ct} = (\mathsf{ct}_0, \mathsf{ct}_1, \mathsf{ct}_2)$
  - $\mathsf{rs}_0 = (a_0^{(0)}, b_1^{(0)}, c_1^{(0)})$
  - $\mathsf{rs}_1 = (a_1^{(1)}, b_1^{(1)}, c_1^{(1)})$
  - $\mathsf{rs}_2 = (a_1^{(2)}, b_0^{(2)}, c_0^{(2)})$
- Being valid transcripts, they must satisfy the corresponding verification equations.

3-Special Soundness (Contd.)

- For $\mathrm{ch} = 1$: $\mathrm{ct}_0 = H(a_1^{(1)} \| v - \mathcal{P}(a_1^{(1)}) - G(b_1^{(1)}, a_1^{(1)}) - c_1^{(1)})$.
- For $\mathrm{ch} = 2$: $\mathrm{ct}_0 = H(a_1^{(2)} \| G(b_0^{(2)}, a_1^{(2)}) + c_0^{(2)})$.

From this equality:

$$a_1^{(1)} = a_1^{(2)}$$

$$\begin{aligned} v &= \mathcal{P}(a_1^{(1)}) + G(b_1^{(1)}, a_1^{(1)}) + c_1^{(1)} + G(b_0^{(2)}, a_1^{(2)}) + c_0^{(2)} \\ &= G(b_0^{(2)} + b_1^{(1)}, a_1^{(1)}) + \mathcal{P}(a_1^{(1)}) + c_1^{(1)} + c_0^{(2)} \end{aligned}$$

3-Special Soundness (Contd.)

1. For ch $= 0$: $\mathrm{ct}_1 = H(a_0^{(0)} - b_1^{(0)} || \mathcal{P}(a_0^{(0)}) - c_1^{(0)})$
2. For ch $= 2$: $\mathrm{ct}_1 = H(b_0^{(2)} || c_0^{(2)})$

From this equality:

$$a_0^{(0)} = b_0^{(2)} + b_1^{(0)}$$

$$\mathcal{P}(a_0^{(0)}) = c_0^{(2)} + c_1^{(0)}$$

3-Special Soundness (Contd.)

1. For $\mathsf{ch} = 0$: $\mathsf{ct}_2 = H(b_1^{(0)} || c_1^{(0)})$.
2. For $\mathsf{ch} = 1$: $\mathsf{ct}_2 = H(b_1^{(1)} || c_1^{(1)})$.

From this equality:

$$b_1^{(0)} = b_1^{(1)} \text{ and } c_1^{(0)} = c_1^{(1)}$$

Plugging in the expression for v:

$$v = G(a_0^{(0)}, a_1^{(1)}) + \mathcal{P}(a_0^{(0)}) + \mathcal{P}(a_1^{(1)}) = \mathcal{P}(a_0^{(0)} + a_1^{(1)})$$

So, $a_0^{(0)} + a_1^{(1)}$ is a witness.

Knowledge Error

[Claim] Any cheating prover P can impersonate with probability at most $2/3$.

- P will choose some $s' \neq s$ as witness and simply execute the protocol.
- P can successfully impersonate if $\mathsf{ch} \neq 1$ as $v = \mathcal{P}(s)$ is not involved in the verification.
- P fails to convince V when $\mathsf{ch} = 1$.

Proof of Knowledge

Claim: Assuming that the MQ-problem is intractable, the 3-pass IDS is a proof-of-knowledge (with knowledge error $2/3$).

- Suppose there is an adversary $\mathcal{A}$ who can impersonate with probabilty $2/3 + \epsilon$ for some non-negligible ϵ.
- We will rewind $\mathcal{A}$ on the same commitment ct but three different challenges $\mathsf{ch} = 0, \mathsf{ch} = 1$ and $\mathsf{ch} = 2$.
- Finally we get three accepted transcripts of the form:

$$(\mathsf{ct}, 0, \mathsf{rs}_0); (\mathsf{ct}, 1, \mathsf{rs}_1); \text{ and } (\mathsf{ct}, 2, \mathsf{rs}_2)$$

- We have seen that given three such accepted transcripts one can extract the witness s.
- Hence we can solve the MQ problem.

Fiat-Shamir Transformation

Recall from previous lecture:

- The signer runs the IDS Key-Gen algorithm to generate $\langle pk, sk \rangle$. Makes pk public in addition with a cryptographic hash function.
- To sign some message M:
  1. Generate the ct of IDS.
  2. Compute $\mathrm{ch} = H(M, \mathrm{ct}, pk)$
  3. Generate rs as per the IDS protocol.
  4. The signature on M is $\sigma = \langle \mathrm{ct}, \mathrm{rs} \rangle$.
- For signature verification, first compute $\mathrm{ch} = H(M, \mathrm{ct}, pk)$ and check whether $(\mathrm{ct}, \mathrm{ch}, \mathrm{rs})$ forms a valid transcript or not.

MQDSS

- MQDSS is the signature scheme that was constructed from the 5-pass IDS of [SSH'11].
- Here we consider the signature scheme obtained from 3-pass IDS.
- To make the IDS knowledge error negligible we consider r rounds of IDS run in parallel so that the knowledge error becomes negligible: $(2/3)^r$.
- Signer secret key is $sk = \mathbf{x}$ and the corresponding $pk = (\mathcal{P}, v)$.
- Public information also includes another hash function $\mathcal{H}_2 : \{0,1\}^* \to \{0,1,2\}^r$.

Signing

1. G: polar form of the system $\mathcal{P} : \mathbb{F}_q^n \to \mathbb{F}_q^m$.
2. pick $\mathbf{a}_{0,i}, \mathbf{b}_{0,i} \xleftarrow{\$} \mathbb{F}^n$ and $\mathbf{c}_{0,i} \xleftarrow{\$} \mathbb{F}^m$ for $i \in [r]$
3. set $\mathbf{a}_{1,i} = \mathbf{x} - \mathbf{a}_{0,i}$, $\mathbf{b}_{1,i} = \mathbf{a}_{0,i} - \mathbf{b}_{0,i}$ and $\mathbf{c}_{1,i} = \mathcal{P}(\mathbf{a}_{0,i}) - \mathbf{c}_{0,i}$ for $i \in [r]$
4. for each $i \in [r]$, compute:
   4.1 $\mathrm{ct}_{0,i} \longleftarrow H(\mathbf{a}_{1,i}, \mathrm{G}(\mathbf{b}_{0,i}, \mathbf{a}_{1,i}) + \mathbf{c}_{0,i})$
   4.2 $\mathrm{ct}_{1,i} \longleftarrow H(\mathbf{b}_{0,i} || \mathbf{c}_{0,i})$
   4.3 $\mathrm{ct}_{2,i} \longleftarrow H(\mathbf{b}_{1,i} || \mathbf{c}_{1,i})$
5. set $\mathbf{ct} = (\mathrm{ct}_{0,1}, \mathrm{ct}_{1,1}, \mathrm{ct}_{2,1}, \ldots, \mathrm{ct}_{0,r}, \mathrm{ct}_{1,r}, \mathrm{ct}_{2,r})$ and compute $\mathbf{ch} = \mathcal{H}_2(M, v, \mathbf{ct})$
6. parse $\mathbf{ch}$ as $(\mathrm{ch}_1, \ldots, \mathrm{ch}_r)$ and for each $i \in [r]$:
   6.1 if $\mathrm{ch}_i = 0$, set $\mathrm{rs}_i = (\mathbf{a}_{0,i}, \mathbf{b}_{1,i}, \mathbf{c}_{1,i})$
   6.2 if $\mathrm{ch}_i = 1$, set $\mathrm{rs}_i = (\mathbf{a}_{1,i}, \mathbf{b}_{1,i}, \mathbf{c}_{1,i})$
   6.3 if $\mathrm{ch}_i = 2$, set $\mathrm{rs}_i = (\mathbf{a}_{1,i}, \mathbf{b}_{0,i}, \mathbf{c}_{0,i})$
7. set $\mathbf{rs} = (\mathrm{rs}_1, \ldots, \mathrm{rs}_r)$ and return $\sigma = (\mathbf{ct}, \mathbf{rs})$

Security

If the IDS is “secure” then the signature scheme is unforgeable.
If one can produce a forgery in the signature scheme then that amounts to a valid transcript of the underlying IDS.

Exercise

Study the 5-pass IDS of Sakumoto-Shirai-Hiwatari and figure out the knowledge error. Is there any advantage/disadvantage of using the 5-pass IDS in the signature construction?