Linear Codes for Secure Computation

Clément Ducros
Geoffroy Couteau

IRIF

5 avril 2022
Secure Computation: what for?

Classical cryptography goal: protecting communications. But data can be used in computations.
What is Secure Multiparty Computation (MPC)?

Goal: Consider $n$ players, each ones owning a secret value $x_i$. Each player wants to compute the result of a function $f_i$ on the entries $(x_j)_{1 \leq j \leq n}$.

Correctness → Each player should get the correct result.

Security → Any group of players who ally themselves must not learn more than is already implied by their secret entries and their function.
Example 1 : Oblivious Transfer

OT requires **public-key cryptography**
Useful correlation for efficiently computing Boolean circuit.
Example 2: OLE and Vector OLE

**OLE: Oblivious Linear Evaluation**

\[ x, c \]
\[ c = a \cdot x + b \]
\[ x \]
\[ a, b \]

Useful correlation for efficiently computing arithmetic circuit.

**VOLE: Vector Oblivious Linear Evaluation**

\[ x, w \]
\[ w = u \cdot x + v \]
\[ x \]
\[ u, v \]
\[ u, v \]
How to achieve MPC? [GMW87]

Secret sharing of each inputs!

\[ \langle x_0 \rangle_A \oplus \langle x_0 \rangle_B = x_0 \]

\[ \langle x_1 \rangle_A \]
\[ \langle x_0 \rangle_A \]
\[ x_0 \]
\[ x_1 \]
\[ \langle x_0 \oplus x_1 \rangle_A \]
\[ \text{XOR} \]
\[ \langle x_0 \rangle_A \]
\[ \langle x_1 \rangle_A \]
\[ \text{AND} \]
\[ \langle x_0 \land x_1 \rangle_A \]
\[ \text{AND} \]

\[ \langle x_0 \oplus x_1 \rangle_A = \langle x_0 \rangle_A \oplus \langle x_1 \rangle_A \]

\[ \langle x_0 \land x_1 \rangle_A = ? \text{ Requires 2 OT} \]
How to achieve MPC?

Research leads to split the protocol in two phases [Bea95, IKNP03]

First Phase: Preprocessing

$K_0$ \[ \xrightarrow{\text{Correlated randomness}} \] $K_1$

Second Phase

$x_0$ \[ \xrightarrow{\text{Correlated randomness}} \] $x_1$

The first phase is input-independent, and can be done ahead of time
From random VOLE to Pseudorandom Vector OLE

Question: how to generate many random OTs? (or others correlations).

\[ r_x, r_w = r_u \cdot r_x + r_v \]

\[ m_x = x - r_x \]
\[ m_u = u - r_u \]
\[ m_v = m_x \cdot r_u - r_u + v - r_v \]

\[ w = m_u x + m_v + r_w = u \cdot x + v \]
Correlated Randomness generation \[\text{[BCGI18, BCG}^{+19}\text{]}\]

First Phase: Preprocessing

- Silent expansion
- « Small » setup protocol
- Correlated short seeds
- Long, pseudorandom correlated strings

Second Phase

- Online phase
  
  \[f(x_0, x_1)\]

- Very fast online phase
- Few communication to compute
- Downside: we have to do again all the computation when it is done.
Pseudorandom Correlated Functions (PCF) [BCG+20]

PRF: functions that cannot be distinguished from truly random functions
Equivalent for correlation?

Correctness: \((R_0, R_1) \approx \) fresh sample of correlation
Security: against insiders
How to construct a PCF?

- A Weak Pseudo-Random Function = PRF but the adversary can’t chose where to evaluate the functions.
The LPN assumption

Let $A \in \mathbb{F}_q^{m \times k}$, $s \in \mathbb{F}_q^k$, $e \in \mathbb{F}_q^m$, $r \in \mathbb{F}_q^m$, with $\mathcal{HW}(e)$ small.

We mostly focus on the dual version of this assumption, which is equivalent.
The LPN and the VDLPN assumption

The matrix $H$ and the noise $e$ have some structure!

$$H = \begin{bmatrix} \begin{array}{ccc} 2w & 4w & 8w \\ H_1 & H_2 & H_3 \\ 16w & \end{array} \end{bmatrix}$$

Exponentially decreasing density. The noise have the shape of one line of $H$.

$$H_i = H_{i,1} \ldots H_{i,w}$$

Unit vector of $\mathbb{F}_2^i$.
## Linear attacks examples

<table>
<thead>
<tr>
<th>Attacks</th>
<th>Types of attacks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian elimination</td>
<td>Linear</td>
</tr>
<tr>
<td>Statistical decoding</td>
<td>Linear</td>
</tr>
<tr>
<td>Information set decoding</td>
<td>Linear</td>
</tr>
<tr>
<td>BKW</td>
<td>Linear</td>
</tr>
<tr>
<td>Algebraic</td>
<td>Linear</td>
</tr>
<tr>
<td>Statistical Query Algorithm</td>
<td>Non-linear</td>
</tr>
<tr>
<td></td>
<td>Non-linear</td>
</tr>
</tbody>
</table>
Linear attacks for our variants

**Definition (Bias of a distribution)**
Given a distribution $\mathcal{D}$ over $\mathbb{F}_2^n$, a vector $u \in \mathbb{F}_2^n$:

$$\text{bias}_u(\mathcal{D}) = \left| \frac{1}{2} - \Pr_{v \leftarrow \mathcal{D}} [u^\top \cdot v = 1] \right|$$

**Definition (Resistance against linear attacks)**
We obtain the resistance against linear attacks when

$$\Pr_{x^{(1)}, \ldots, x^{(N(\lambda))} \leftarrow \mathbb{F}_2^{n(\lambda)}} [\text{bias}(\mathcal{D}(x) > \epsilon(\lambda))] < \delta(\lambda)$$

where $\epsilon$ and $\delta$ are small depending on the security parameter $\lambda$. 
Analysis of security

\[ \Pr[\text{Attack vector } HW(v) = l] = 1 \approx \frac{1}{2} \]

\[ \Pr[\text{Unit vector of } \mathbb{F}_2^{2i}] \rightarrow \quad \]

\[ \Pr[\text{v}] \quad H_i \]

\[ e_i \]

\[ H_{i,k} \]

\[ e_{i,k} = 1 \approx \frac{1}{2} \]

\[ X_{j,k} \]
Analysis of security

We define $R_{i,l,k} = (v^\top \cdot H_{i,k}) = \left( \bigoplus_{j=1}^{l} X_{j,k} \right)$ and $Z_{i,l,k}$ as $Z_{i,l,k} = |2^{i-1} - R_{l,k}|$.

**Definition ($\delta$-Bad Matrices)**

Let $M \in \mathbb{F}_2^{N \times 2^i}$. We say that $M \in \text{Bad}_{\delta,v}$ with respect to a vector $v \in \mathbb{F}_2^{2N}$ if

$$(v^\top \cdot M) = Z_{l,k} \in \left[ (1/2 - \delta) \cdot 2^i, 2^{i-1} \right].$$

Given vector $v$, we denote $B_{\delta,v} = \#\text{Bad}_{\delta,v}$.

**Lemma**

*For any $v \in S_{i,N}$, there is a constant $C$ such that*

$$\Pr \left[ B_{\delta,v} > \alpha \cdot w \right] \leq 2^{-C \cdot 2^i \cdot w}$$
Analysis of security

We introduce a function $\Phi$

$$\Phi(X_{1,1}, \cdots, X_{l,w}) = 2^{i-1} \cdot w - \sum_{k=1}^{w} Z_{l,k}.$$ 

$$\Pr \left[ B_{\delta,v} \geq \alpha \cdot w \right] \leq \Pr \left[ \Phi(X_{1,1}, \cdots, X_{l,w}) < \gamma \cdot w \cdot 2^i \right],$$

$\Phi$ is 2-Lipschitz : we use the Bounded Difference Inequality.

**Proposition (Bounded Difference Inequality)**

Let $\Phi : [n]^m \rightarrow \mathbb{R}$ be a function satisfying the Lipschitz property with constant $d$, and let $(X_1, \cdots, X_m)$ e independant random variables over $[n]$)

$$\Pr[\Phi(X_1, \cdots, X_m) < \mathbb{E}[\Phi(X_1, \cdots, X_m)] - t] \leq \exp(-\frac{2t^2}{m \cdot d^2})$$
Remains to find an upper bound of $\mathbb{E}[\Phi] \to$ find an upper bound of $\mathbb{E}[Z_{l,k}]$. There was an error in the proof [BCG+20]!

**Correction**

$$\mathbb{E}[Z_{l,k}] = \sum_{j=0}^{2^{i-1}-1} \Pr(R_{l,k} \geq j + 1 + 2^{i-1}) + \sum_{j=0}^{2^{i-1}-1} \Pr(R_{l,k} \leq 2^{i-1} - j - 1)$$

1. We bound the shares that we can with the Generalized Chernoff Inequality.
2. For that we had to prove that the distribution of the $R_{l,k}$ shows some kind of independence.
3. We bound the remaining shares with a trivial bound.

We have to remember the union bound!

This corrects the proof but is highly unpractical, with $w \approx 10^6$. 
Proposition

Let \( n \in \mathbb{N} \) an integer, and let \((Y_1, \cdots, Y_n)\) be independent boolean random variables such that, for some \( \eta \in [0, 1] \) it holds that for every subset \( S \in [n] \),
\[
\Pr \left[ \bigwedge_{q \in S} Y_q \right] \leq \eta^{|S|}.
\]
Then for any \( \kappa \in [\eta, 1] \),
\[
\Pr \left[ \sum_{q=1}^{n} Y_q \geq \kappa n \right] \leq \exp \left( -n \cdot D_{KL} (\kappa \| \eta) \right),
\]
where \( D_{KL} (\kappa \| \eta) \) denotes the relative entropy function, defined as
\[
D_{KL} (\kappa \| \eta) = \kappa \log_\frac{\kappa}{\eta} + (1 - \kappa) \log_\frac{1 - \kappa}{1 - \eta}.
\]
A new proof:

- Simulation to prove that $\mathbb{E}[Z] < \beta^i$ with better $\beta$
- Erasing the corner cases.
- A new idea to bound the bias:

$$\Pr[\text{bias}_v(O_{\text{par}}^i) > B] = \Pr \left[ \prod_{k=1}^{w} Z_{i,l,k} > 2^{(i-1)w} \times (2B) \right].$$

- The sum $\sum_k Z_{i,l,k}$ is minimized when all the terms in the product are equal.

$$\Pr[\text{bias}_v(O_{\text{par}}^i) > B] \leq \Pr \left[ \sum_{k=1}^{w} Z_{i,l,k} > w \cdot 2^{(i-1)} \cdot c \right],$$

- With this sum we can apply again our results with the function $\Phi$. 
We obtain with this new proof $w \approx 350$.

Estimation of the **concrete cost** of the PCFs.

- Seed size: 2.55MB
- PCF evaluation time: $\approx 500$ PCF evaluations per second on a single 3GHz processor.

Another work of [BCG+20] also suggested an improved all prefix variant. No proof of the security for this variant yet, but very promising values.

- Seed size: 0.34MB.
- PCF evaluation time: around 3500 evaluations per second on a single 3GHz processor.
Conclusion and open questions

- Pseudo-random Function achieves very promising parameters.

Open Problems and ongoing works:
- All prefix Variant
- Variable density matrix shapes.
- Ring LPN and variants
Questions ?
References

Elette Boyle, Geoffroy Couteau, Niv Gilboa, Yuval Ishai, Lisa Kohl, and Peter Scholl.
Efficient pseudorandom correlation generators: Silent OT extension and more.

Elette Boyle, Geoffroy Couteau, Niv Gilboa, Yuval Ishai, Lisa Kohl, and Peter Scholl.
Correlated pseudorandom functions from variable-density LPN.

Elette Boyle, Geoffroy Couteau, Niv Gilboa, and Yuval Ishai.
Compressing vector OLE.

Donald Beaver.
Precomputing oblivious transfer.

Oded Goldreich, Silvio Micali, and Avi Wigderson.
How to play any mental game or A completeness theorem for protocols with honest majority.

Yuval Ishai, Joe Kilian, Kobbi Nissim, and Erez Petrank.
Extending oblivious transfers efficiently.