Degree tables

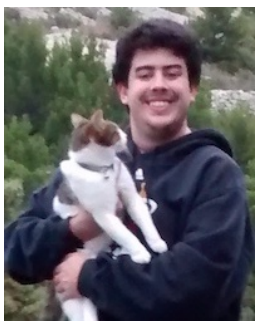
Secure Distributed Matrix Multiplication

Daniel Heinlein
daniel.heinlein@aalto.fi
Aalto University

Aalto University
School of Electrical
Engineering

6. August 2019
Norcom 2019
Copenhagen

Joint work with:



R.G.L. D'Oliveira

S. El Rouayheb

D. Karpuk

Scenario

Client has matrices A, B over $\mathbb{F}_q$ and wants to compute $A \cdot B$ without revealing information about A, B using N servers such that at most T may collude.

Servers are "honest but curious".

Submatrices:

$$A = \begin{pmatrix} A_1 \\ \vdots \\ A_K \end{pmatrix}, \quad B = \begin{pmatrix} B_1 & \cdots & B_L \end{pmatrix} \quad \Rightarrow \quad AB = \begin{pmatrix} A_1B_1 & \cdots & A_1B_L \\ \vdots & \ddots & \vdots \\ A_KB_1 & \cdots & A_KB_L \end{pmatrix}$$

Goal: Minimize N for fixed K, L, T.

Related work

W.-T. Chang, R. Tandon, "On the Capacity of Secure Distributed Matrix Multiplication," *2018 IEEE Global Communications Conference (GLOBECOM)*, 2018.
$\rightarrow$ first presentation of the setting, $K = L$, $N = (K + T)^2$

J. Kakar, S. Ebadifar, and A. Sezgin, "Rate-Efficiency and Straggler-Robustness through Partition in Distributed Two-Sided Secure Matrix Computation," *arXiv:1810.13006*, 2018.
$\rightarrow$ improvement to $N = (K + T)(L + 1) - 1$

H. Yang and J. Lee, "Secure Distributed Computing With Straggling Servers Using Polynomial Codes," in *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 1, pp. 141-150, Jan. 2019.
$\rightarrow$ independently $T = 1$ and $N = (K + 1)(L + 1) - 1$

Scenario

Converted to polynomials (α, β to be chosen, R, S random matrices):

$$A = \begin{pmatrix} A_1 \\ \vdots \\ A_K \end{pmatrix} \quad \Rightarrow \quad f(x) = \sum_{k=1}^{K} A_k x^{\alpha_k} + \sum_{t=1}^{T} R_t x^{\alpha_{K+t}},$$

$$B = \begin{pmatrix} B_1 & \cdots & B_L \end{pmatrix} \quad \Rightarrow \quad g(x) = \sum_{\ell=1}^{L} B_\ell x^{\beta_\ell} + \sum_{t=1}^{T} S_t x^{\beta_{L+t}}$$

Server i gets $f(a_i)$ and $g(a_i)$ and returns $f(a_i) \cdot g(a_i)$ to the client ($i = 1..N$).

Client interpolates $f \cdot g$ and reads $A_k B_\ell$ from its coefficients.

$\Rightarrow$ N is the number of different exponents of x in $f \cdot g$

Degree table

$$f(x) = \sum_{k=1}^{K} A_k x^{\alpha_k} + \sum_{t=1}^{T} R_t x^{\alpha_{K+t}}, \quad g(x) = \sum_{\ell=1}^{L} B_\ell x^{\beta_\ell} + \sum_{t=1}^{T} S_t x^{\beta_{L+t}}$$

	β_1	$\cdots$	β_L	β_{L+1}	$\cdots$	β_{L+T}
α_1	$\alpha_1+\beta_1$	$\cdots$	$\alpha_1+\beta_L$	$\alpha_1+\beta_{L+1}$	$\cdots$	$\alpha_1+\beta_{L+T}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
α_K	$\alpha_K+\beta_1$	$\cdots$	$\alpha_K+\beta_L$	$\alpha_K+\beta_{L+1}$	$\cdots$	$\alpha_K+\beta_{L+T}$
α_{K+1}	$\alpha_{K+1}+\beta_1$	$\cdots$	$\alpha_{K+1}+\beta_L$	$\alpha_{K+1}+\beta_{L+1}$	$\cdots$	$\alpha_{K+1}+\beta_{L+T}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
α_{K+T}	$\alpha_{K+T}+\beta_1$	$\cdots$	$\alpha_{K+T}+\beta_L$	$\alpha_{K+T}+\beta_{L+1}$	$\cdots$	$\alpha_{K+T}+\beta_{L+T}$

$\Rightarrow$ N is the number of different numbers in the table

Degree table

	β_1	$\cdots$	β_L	β_{L+1}	$\cdots$	β_{L+T}
α_1	$\alpha_1+\beta_1$	$\cdots$	$\alpha_1+\beta_L$	$\alpha_1+\beta_{L+1}$	$\cdots$	$\alpha_1+\beta_{L+T}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
α_K	$\alpha_K+\beta_1$	$\cdots$	$\alpha_K+\beta_L$	$\alpha_K+\beta_{L+1}$	$\cdots$	$\alpha_K+\beta_{L+T}$
α_{K+1}	$\alpha_{K+1}+\beta_1$	$\cdots$	$\alpha_{K+1}+\beta_L$	$\alpha_{K+1}+\beta_{L+1}$	$\cdots$	$\alpha_{K+1}+\beta_{L+T}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
α_{K+T}	$\alpha_{K+T}+\beta_1$	$\cdots$	$\alpha_{K+T}+\beta_L$	$\alpha_{K+T}+\beta_{L+1}$	$\cdots$	$\alpha_{K+T}+\beta_{L+T}$

Theorem (D'Oliveira, El Rouayheb, Karpuk (2018))

AB can be retrieved and privacy is guaranteed if at most T servers collude if the numbers in the

1. *blue area are distinct,*
2. *green area are distinct,*
3. *red area are distinct from all numbers in the degree table.*

Summary

	β_1	$\cdots$	β_L	β_{L+1}	$\cdots$	β_{L+T}
α_1	$\alpha_1+\beta_1$	$\cdots$	$\alpha_1+\beta_L$	$\alpha_1+\beta_{L+1}$	$\cdots$	$\alpha_1+\beta_{L+T}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
α_K	$\alpha_K+\beta_1$	$\cdots$	$\alpha_K+\beta_L$	$\alpha_K+\beta_{L+1}$	$\cdots$	$\alpha_K+\beta_{L+T}$
α_{K+1}	$\alpha_{K+1}+\beta_1$	$\cdots$	$\alpha_{K+1}+\beta_L$	$\alpha_{K+1}+\beta_{L+1}$	$\cdots$	$\alpha_{K+1}+\beta_{L+T}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
α_{K+T}	$\alpha_{K+T}+\beta_1$	$\cdots$	$\alpha_{K+T}+\beta_L$	$\alpha_{K+T}+\beta_{L+1}$	$\cdots$	$\alpha_{K+T}+\beta_{L+T}$

Question

For fixed K, L, T, how can α_k, β_ℓ be chosen to minimize N, i.e., the number of distinct entries such that all α_k are distinct, all β_ℓ are distinct, and each number in the red area appears exactly once in the degree table?

Connection to sumsets

By omitting the restriction "... each number in the red area appears exactly once in the degree table ..." we arrive at a classical sumset problem: *inverse problem*

	β_1	$\cdots$	β_L	β_{L+1}	$\cdots$	β_{L+T}
α_1	$\alpha_1+\beta_1$	$\cdots$	$\alpha_1+\beta_L$	$\alpha_1+\beta_{L+1}$	$\cdots$	$\alpha_1+\beta_{L+T}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
α_K	$\alpha_K+\beta_1$	$\cdots$	$\alpha_K+\beta_L$	$\alpha_K+\beta_{L+1}$	$\cdots$	$\alpha_K+\beta_{L+T}$
α_{K+1}	$\alpha_{K+1}+\beta_1$	$\cdots$	$\alpha_{K+1}+\beta_L$	$\alpha_{K+1}+\beta_{L+1}$	$\cdots$	$\alpha_{K+1}+\beta_{L+T}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
α_{K+T}	$\alpha_{K+T}+\beta_1$	$\cdots$	$\alpha_{K+T}+\beta_L$	$\alpha_{K+T}+\beta_{L+1}$	$\cdots$	$\alpha_{K+T}+\beta_{L+T}$

How do $\mathcal{A}, \mathcal{B} \subseteq \mathbb{Z}$ look like if $N = |\mathcal{A} + \mathcal{B}|$ is minimal?[1]

Lemma (well known)

For $\mathcal{A}, \mathcal{B} \subseteq \mathbb{Z}$ nonempty: $N = |\mathcal{A} + \mathcal{B}| \geq |\mathcal{A}| + |\mathcal{B}| - 1$ and "=" iff $\mathcal{A}, \mathcal{B}$ are arithmetic progressions with the same common difference.

Corollary to [Stanchescu (1996)]: If $T \gg K, L$, then there is a $z \in \mathbb{Z}$ such that $\max\{\alpha_i, \beta_j\} \leq z$ and $\min\{\alpha_i\} = \min\{\beta_j\} = 0$.

At least then, the problem is finite ...

[1] $\mathcal{A} + \mathcal{B} = \{a + b \mid a \in \mathcal{A}, b \in \mathcal{B}\}$

Lower bounds

Theorem (D'Oliveira, El Rouayheb, H., Karpuk (2019))

$$N \geq KL + \max\{K, L\} + 2T - 1$$

Call the sets of integers:

A	B
C	D

	β_1	$\cdots$	β_L	β_{L+1}	$\cdots$	β_{L+T}
α_1	$\alpha_1+\beta_1$	$\cdots$	$\alpha_1+\beta_L$	$\alpha_1+\beta_{L+1}$	$\cdots$	$\alpha_1+\beta_{L+T}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
α_K	$\alpha_K+\beta_1$	$\cdots$	$\alpha_K+\beta_L$	$\alpha_K+\beta_{L+1}$	$\cdots$	$\alpha_K+\beta_{L+T}$
α_{K+1}	$\alpha_{K+1}+\beta_1$	$\cdots$	$\alpha_{K+1}+\beta_L$	$\alpha_{K+1}+\beta_{L+1}$	$\cdots$	$\alpha_{K+1}+\beta_{L+T}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
α_{K+T}	$\alpha_{K+T}+\beta_1$	$\cdots$	$\alpha_{K+T}+\beta_L$	$\alpha_{K+T}+\beta_{L+1}$	$\cdots$	$\alpha_{K+T}+\beta_{L+T}$

1. $|A| = KL$ and $A \cap (B \cup C \cup D) = \emptyset$.
2. omit A and C
3. $B \cup D$ is the sumset $\alpha + \{\beta_{L+1}, \ldots, \beta_{L+T}\}$ and hence:

   $$|B \cup D| \geq |\alpha| + |\{\beta_{L+1}, \ldots, \beta_{L+T}\}| - 1 = (K + T) + (T) - 1$$

4. $\Rightarrow N \geq KL + K + 2T - 1$
5. (exchange B and C:) $N \geq KL + \max\{K, L\} + 2T - 1$

Lower bounds

Theorem (D'Oliveira, El Rouayheb, H., Karpuk (2019))

If $3\max\{K, L\} + 3T - 2 < KL$ *or* $2 \leq K = L$, *then*

$$N \geq KL + \max\{K, L\} + 2T$$

Consider the case when the last bound is sharp.

1. omitting C does not remove all occurrences of these integers
   $\Rightarrow C \subseteq B \cup D$
2. $|B \cup D| \geq |\alpha| + |\{\beta_{L+1}, \ldots, \beta_{L+T}\}| - 1 = (K + T) + (T) - 1$ is sharp
   $\Rightarrow \alpha$ and $\{\beta_{L+1}, \ldots, \beta_{L+T}\}$ are arithmetic progression with the same common difference.

$\Rightarrow KL \leq 3\max\{K, L\} + 3T - 2$ and if $K = L$, then $K = L = 1$.

Lower bounds

Theorem (D'Oliveira, El Rouayheb, H., Karpuk (2019))

$$N \geq KL + K + L + 2T - 1 - T\min\{K, L, T\}$$

The proof is more involved but uses:

Lemma (D'Oliveira, El Rouayheb, H., Karpuk (2019))

Let $\alpha = \alpha_{\mathsf{p}}|\alpha_{\mathsf{s}}$, $\beta = \beta_{\mathsf{p}}|\beta_{\mathsf{s}}$ *(of sizes* K, T, L, T*). Then*

$$|\quad (\alpha_i + \beta_{\mathsf{p}}) \quad \cap \quad (\alpha_{\mathsf{p}} + \beta_j) \quad| \leq 1$$

for all $1 \leq i \leq K + T$ *and all* $1 \leq j \leq L + T$.

	β_1	$\cdots$	β_L	β_{L+1}	$\cdots$	β_{L+T}
α_1	$\alpha_1+\beta_1$	$\cdots$	$\alpha_1+\beta_L$	$\alpha_1+\beta_{L+1}$	$\cdots$	$\alpha_1+\beta_{L+T}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
α_K	$\alpha_K+\beta_1$	$\cdots$	$\alpha_K+\beta_L$	$\alpha_K+\beta_{L+1}$	$\cdots$	$\alpha_K+\beta_{L+T}$
α_{K+1}	$\alpha_{K+1}+\beta_1$	$\cdots$	$\alpha_{K+1}+\beta_L$	$\alpha_{K+1}+\beta_{L+1}$	$\cdots$	$\alpha_{K+1}+\beta_{L+T}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
α_{K+T}	$\alpha_{K+T}+\beta_1$	$\cdots$	$\alpha_{K+T}+\beta_L$	$\alpha_{K+T}+\beta_{L+1}$	$\cdots$	$\alpha_{K+T}+\beta_{L+T}$

Construction: GASP_r

(Gap Additive Secure Polynomial codes)

Theorem (D'Oliveira, El Rouayheb, H., Karpuk (2019))

Let wlog. $L \leq K$ and T be given. Fix an integer r with $1 \leq r \leq \min\{K, T\}$. Choose

$$\begin{aligned}\alpha = (&0, 1, \ldots, K-1,\\ &KL, KL+1, \ldots, KL+r-1,\\ &KL+K, KL+K+1, \ldots, KL+K+r-1, \ldots)\end{aligned}$$

of length $K+T$,

$$\begin{aligned}\beta = (&0, K, \ldots, K(L-1),\\ &KL, KL+1, \ldots, KL+T-1).\end{aligned}$$

Then this constructs a degree table.

Note: The suffix of α consists of the first T elements of the generalized arithmetic progression $KL + \{0, \ldots, r-1\} + \{0, K, \ldots\}$.

Construction: GASP_r

Example: $K = L = T = 4$ ($N \geq 28$)

	0	4	8	12	16	17	18	19
0	0	4	8	12	16	17	18	19
1	1	5	9	13	17	18	19	20
2	2	6	10	14	18	19	20	21
3	3	7	11	15	19	20	21	22
16	16	20	24	28	32	33	34	35
20	20	24	28	32	36	37	38	39
24	24	28	32	36	40	41	42	43
28	28	32	36	40	44	45	46	47

$r = 1,\ N = 41$

	0	4	8	12	16	17	18	19
0	0	4	8	12	16	17	18	19
1	1	5	9	13	17	18	19	20
2	2	6	10	14	18	19	20	21
3	3	7	11	15	19	20	21	22
16	16	20	24	28	32	33	34	35
17	17	21	25	29	33	34	35	36
20	20	24	28	32	36	37	38	39
21	21	25	29	33	37	38	39	40

$r = 2,\ N = 36$

	0	4	8	12	16	17	18	19
0	0	4	8	12	16	17	18	19
1	1	5	9	13	17	18	19	20
2	2	6	10	14	18	19	20	21
3	3	7	11	15	19	20	21	22
16	16	20	24	28	32	33	34	35
17	17	21	25	29	33	34	35	36
18	18	22	26	30	34	35	36	37
20	20	24	28	32	36	37	38	39

$r = 3,\ N = 37$

	0	4	8	12	16	17	18	19
0	0	4	8	12	16	17	18	19
1	1	5	9	13	17	18	19	20
2	2	6	10	14	18	19	20	21
3	3	7	11	15	19	20	21	22
16	16	20	24	28	32	33	34	35
17	17	21	25	29	33	34	35	36
18	18	22	26	30	34	35	36	37
19	19	23	27	31	35	36	37	38

$r = 4,\ N = 39$

Construction: GASP_r

What is the number of servers $N_{K,L,T}(r)$ for used in GASP_r for fixed K, L, T?

Consider:

	0	K	$\cdots$	$K(L-1)$	KL	$KL+1$	$\cdots$	$KL+T-1$
0	0	K	$\cdots$	$KL-K$	KL	$KL+1$	$\cdots$	$KL+T-1$
1	1	$K+1$	$\cdots$	$KL-K+1$	$KL+1$	$KL+2$	$\cdots$	$KL+T$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$K-1$	$K-1$	$2K-1$	$\cdots$	$KL-1$	$KL+K-1$	$KL+K$	$\cdots$	$KL+K+T-2$
KL								
$KL+1$								
$\vdots$								
$KL+r-1$								
$KL+K$								
$KL+K+1$								
$\vdots$								
$KL+K+r-1$								
$\vdots$								

The first K rows already contain $KL+K+T-1$ different numbers.

Construction: GASP_r

The last T rows contain $T(L+T)$ numbers $\Rightarrow$ instead of counting different numbers, read these rows row-wise and count the numbers which already appeared.

Define L_i (R_i) to be the number of known entries in the first L (last T) positions of row $K+i-1$.

Lemma (D'Oliveira, El Rouayheb, H., Karpuk (2019))

$$L_i = \begin{cases} \min\{L, 2 + \lfloor (T-1-i)/K \rfloor\} & \text{if } 1 \le i \le r \\ L & \text{if } r+1 \le i \le T \end{cases}$$

and

$$R_i = \begin{cases} \max\{0, K+T-KL-1\} & \text{if } i = 1 \\ \max\{0, T-K+r-1\} & \text{if } 2 \le i \text{ and } i \equiv 1 \pmod{r} \\ T-1 & \text{if } i \not\equiv 1 \pmod{r} \end{cases}$$

Construction: GASP_r

Theorem (D'Oliveira, El Rouayheb, H., Karpuk (2019))

$$\begin{aligned} N_{K,L,T}(r) =& KL + 2K + 3T - 2 - \max\{K, \varphi\} \\ &+ \left\lfloor \frac{T-1}{r} \right\rfloor \min\{T-1, K-r\} \\ &- \gamma(r) + (L-2)\max\{0, \min\{r, r-\varphi\}\} \end{aligned}$$

with

1. $\varphi = T - 1 - KL + 2K$
2. $\gamma(r) = \begin{cases} 0 & \text{if } r \leq \varphi \\ K(x-a)(x+a-1)/2 - ab + xy + x & \text{else} \end{cases}$
   - 2.1 $T - 1 - r = aK + b$, $0 \leq b \leq K-1$,
   - 2.2 $T - 2 - \max\{0, \varphi\} = xK + y$, $0 \leq y \leq K-1$.

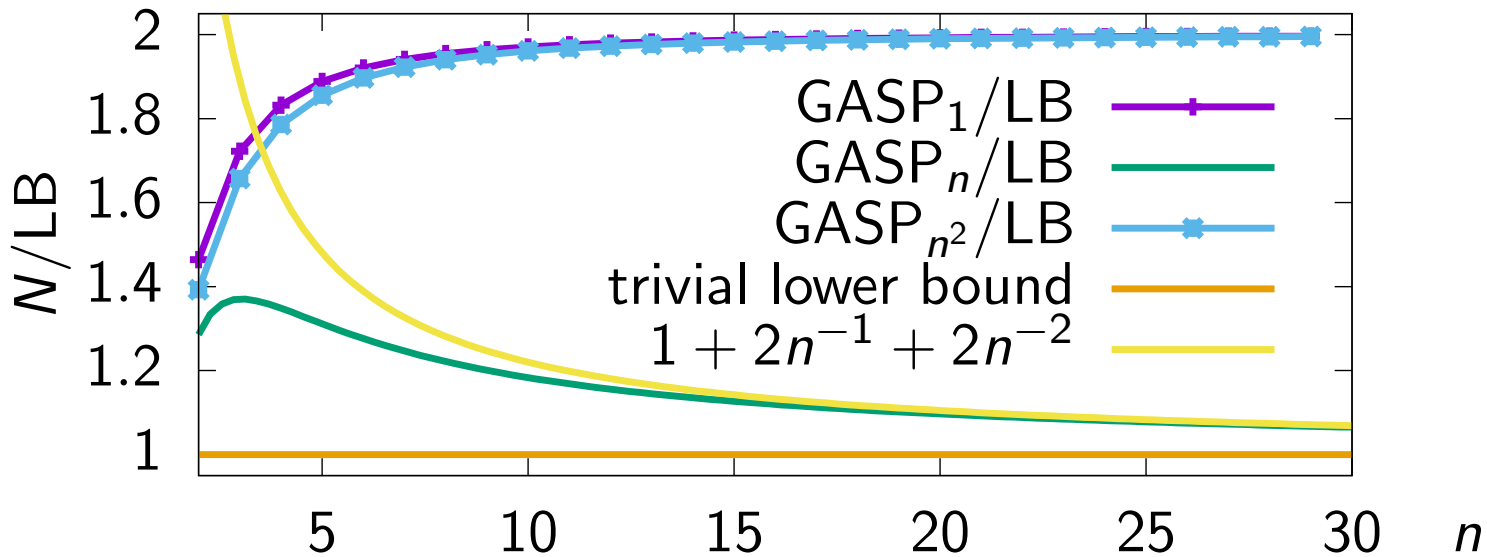
$K = L = T = n^2 \geq 4$

Theorem (D'Oliveira, El Rouayheb, H., Karpuk (2019))

If $K = L = T = n^2 \geq 4$, then the minimizing r is n and

$$n^4 + 3n^2 \leq N_{n^2,n^2,n^2}(n) = n^4 + 2n^3 + 2n^2 - n - 2$$

Moreover, $\frac{N_{n^2,n^2,n^2}(n)}{n^4+3n^2} \leq 1 + 2n^{-1} + 2n^{-2}$ is asymptotically optimal and within 38% of the optimum.



Future work

What r^* minimizes $N_{K,L,T}(r)$?

How large is $N_{K,L,T}(r^*)$/lower bound?

Improve the lower bounds.

Relax the condition that certain coefficients of $f \cdot g$ encode $A_k B_\ell$ such that they encode a linear transformation of them. (Then, the client determines the coefficients and solves a uniquely solvable system of linear equations.)

Thank you.