Computing On Encrypted Data with C++

Max Leibovich

Part 0: Introduction to Cryptography

A Brief History of Cryptography: Symmetric Encryption





- 100 1 A.D.
- 1553
- 1920s
- 1976

- Caesar cipher (Shift cipher)
 - Vigenère cipher (Poly-alphabetic Substitution Cipher)
 - Enigma machine
 - DES (Data Encryption Standard) Symmetric-Key Algorithm

A Brief History of Cryptography: Asymmetric Encryption







- 1976 Diffie, Hellman and Merkle (DH key Exchange, Merkle's Puzzles)
- 1977 Rivest, Shamir and Adleman (RSA Public-Key Cryptosystem)

What else can we do with Encrypted Data?

1978 Rivest, Adleman and Dertouzos: "On data banks and privacy homomorphisms"



What else can we do with Encrypted Data?

1978 Rivest, Adleman and Dertouzos: "On data banks and privacy homomorphisms"



Part 1: Partially Homomorphic Encryption

Where this idea came from?

Lets look at "Textbook" RSA

Choose random k-bit primes p, qKeygen(*k*): Compute $n \coloneqq p \cdot q$ and $\phi(N) = (p-1) \cdot (q-1)$ Choose integer $1 < e < \phi(n)$ which is co-prime to $\phi(n)$, i.e. $gcd(e, \phi(n)) = 1$ $d \coloneqq e^{-1} \mod \phi(n)$, i.e. $d \cdot e \equiv 1 \mod \phi(n)$ Compute Public Key: $pk = \langle n, e \rangle$ Secret Key: $sk = \langle n, d \rangle$ $m \coloneqq c^d \mod n$ $\operatorname{Enc}_{pk=\langle n,e\rangle}(m)$: $c \coloneqq m^e \mod n$ $\operatorname{Dec}_{sk=\langle n,d\rangle}(c)$:

Where this idea came from?

Lets look at "Textbook" RSA

 $\operatorname{Enc}_{pk=\langle n,e\rangle}(m)$: $c \coloneqq m^e \mod n$

 $\operatorname{Dec}_{sk=\langle n,d \rangle}(c)$: $m \coloneqq c^d \mod n$

RSA has the following property:

 $c_1 = \operatorname{Enc}_{pk}(m_1) = m_1^e \mod n \qquad \qquad c_2 = \operatorname{Enc}_{pk}(m_2) = m_2^e \mod n$

 $c_1 \cdot c_2 = [m_1^e \mod n] \cdot [m_2^e \mod n] = [(m_1^e \cdot m_2^e) \mod n] = [(m_1 \cdot m_2)^e \mod n]$

 $= \operatorname{Enc}_{pk}(m_1 \cdot m_2) \qquad \Longrightarrow \qquad \operatorname{Dec}_{sk}(c_1 \cdot c_2) = m_1 \cdot m_2$

Partially Homomorphic Encryptions

• Given groups (G,*) and (H,•) a function $f: G \to H$ is a <u>Homomorphism</u> if it

preserves the operation, i.e. for all $x, y \in G$:

$$f(x * y) = f(x) \circ f(y)$$

• Examples to Homomorphism:

 $|x \cdot y| = |x| \cdot |y|$ $(x \cdot y)^c = x^c \cdot y^c$ $e^{x+y} = e^x \cdot e^y$ $\ln(x \cdot y) = \ln(x) + \ln(y)$

- Multiplicatively Homomorphic Encryptions $Enc(x \cdot y) = Enc(x) \cdot Enc(y)$
 - RSA (1977)
 - ElGamal (1985)

Partially Homomorphic Encryptions

Additively Homomorphic Encryptions

 $\operatorname{Enc}(x+y) = \operatorname{Enc}(x) \cdot \operatorname{Enc}(y)$

- Benaloh(1994)
- Paillier (1999)
- Homomorphic Encryption with respect to XOR $Enc(x \oplus y) = Enc(x) \cdot Enc(y)$
 - Goldwasser-Micali (1982)
- What can we do when we are restricted to a single operation? Not Much!

Part 2: Somewhat Homomorphic Encryption

Models Of Computation

- High-Level Programming Language (e.g. C++)
- Low-Level Programming Language (e.g. Assembly)
- Random-Access Machine
- Turing Machine
- Boolean / Arithmetic Circuits



All are Turing-complete, but there are Time and Space complexity tradeoffs

Why Boolean Circuits?

Because { XOR, AND } is Turing-complete, ANY function can be

computed with a Boolean circuit consisting of only { XOR, AND } gates.





• Over Boolean values we have:

 $AND(a, b) = (a \cdot b) \mod 2$

 $\mathbf{XOR}(a,b) = (a+b) \mod 2$

• Not necessarily the most efficient way to evaluate a function!

Why Boolean Circuits?

- If you can compute products and sums on encrypted bits, you can compute ANY function on encrypted inputs!
- Example: Private Information Retrieval (PIR)

 x_1 x_2 x_3 x_4 x_n

 x_i

<u>Server Input:</u> array of *n* bits $x_1, ..., x_n$ <u>Client Input:</u> index $1 \le j \le n$

<u>Server Output:</u> nothing

<u>Client Output:</u> bit x_j

 $eq(a, b) = \prod_{1 \le k \le \log_2 n} [a_k + b_k + 1]$

Compare Indices $a, b \in \{0,1\}^{\log_2 n}$: PIR function:

$$f(x_1, \dots, x_n, j) = \sum_{1 \le i \le n} [\operatorname{eq}(i, j) \cdot x_i]$$

What objects can we add and multiply?

• Polynomials? $(x^4 + 6x^3 + 2x) + (4x^2 - 3x) = (x^4 + 6x^3 + 4x^2 - x)$ $(5x^2 + 9x + 8) \cdot (7x + 1) = (35x^3 + 68x^2 + 65x + 8)$

• Matrices? $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 & 3 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 6 \\ -1 & 4 \end{pmatrix}$

• Why not Integers?!? 2 + 2 = 4 $7 \cdot 6 = 42$

Example Symmetric Encryption Scheme Over Integers

Keygen(k): Pick a random large k^2 -bit odd integer p as the Secret Key

Enc_p($m \in \{0,1\}$): Pick a random k^5 -bit integer q and compute $q \cdot p$ a large multiple of p

Pick a random small k-bit integer 2r + m,

that is **even** when m = 0, and **odd** when m = 1

Ciphertext will be $c \coloneqq q \cdot p + 2r + m$



Example Symmetric Encryption Scheme Over Integers

 $\operatorname{Dec}_{p}(c \in \mathbb{Z})$:

Compute $c' \coloneqq c \mod^* p$ to recover the "noise".

Where $c \mod^* p$ is the "**Centered Modulo**" operation that returns the integer $c' \in (-p/2, p/2)$ such that p divides c - c'. In other words: 1. $c' \coloneqq c \mod p$;

2. if(c' > p/2) then c' := c' - p;

Return $m \coloneqq c' \mod 2$



Security

- How secure is this?
 - If noise = 0 and we get two encryptions of 0: $\text{Enc}_p(0) = q_1 p$, $\text{Enc}_p(0) = q_2 p$ Recovering the Secret Key p is easy: simply calculate $p = \text{GCD}(q_1 p, q_2 p)$
 - But if there is **noise** the GCD attack doesn't work.

And we believe that neither does any other attack.

This is called the "Approximate GCD Assumption".



Homomorphic Operations

• How do we **XOR** two encrypted bits $c_1 = \text{Enc}_p(m_1)$ and $c_2 = \text{Enc}_p(m_2)$?

 $c_1 = q_1 p + 2r_1 + m_1$ $c_2 = q_2 p + 2r_2 + m_2$

 $c_1 + c_2 = \mathbf{p} \cdot (q_1 + q_2) + \mathbf{2} \cdot (r_1 + r_2) + (m_1 + m_2)$

On decryption, after mod^{*} p: $2 \cdot (r_1 + r_2) + (m_1 + m_2)$



Homomorphic Operations

• How do we **AND** two encrypted bits $c_1 = \text{Enc}_p(m_1)$ and $c_2 = \text{Enc}_p(m_2)$?

 $c_1 = q_1 p + 2r_1 + m_1$ $c_2 = q_2 p + 2r_2 + m_2$

 $c_1 \cdot c_2 = \mathbf{p} \cdot (q_1 c_2 + q_2 c_1 - q_1 q_2) + \mathbf{2} \cdot (r_1 r_2 + r_1 m_2 + r_2 m_1) + (m_1 \cdot m_2)$

On decryption, after mod^{*} p: $2 \cdot (r_1r_2 + r_1m_2 + r_2m_1) + (m_1 \cdot m_2)$



Noise

6

0

• What about the noise? The noise grows after each operation!

•
$$c_1 + c_2 = p \cdot (q_1 + q_2) + 2 \cdot (r_1 + r_2) + (m_1 + m_2)$$

noise $\approx 2 \cdot ($ Initial Noise)

•
$$c_1 \cdot c_2 = p \cdot (q_1c_2 + q_2c_1 - q_1q_2) + 2 \cdot (r_1r_2 + r_1m_2 + r_2m_1) + (m_1 \cdot m_2)$$

noise $\approx (\text{Initial Noise})^2$
The noise $= 2r + m$
 p $2p$ $3p$ \cdots $(q-1)p$ qp $(q+1)p$



So, what did we accomplish?

- We can do lots of additions...
- And some multiplications, until we are no longer able to correctly decrypt...
- This is called: **Somewhat** Homomorphic Encryption.
- It is already enough for some useful applications:
 - PIR over small databases
 - Algorithms implemented as polynomials with logarithmic degree
 - •
- But, we can do much better!

Many C++ Implementations

- This scheme is called <u>DGHV</u>
- There are any many C++ implementations of it in GitHub, e.g.: https://github.com/rinon/Simple-Homomorphic-Encryption https://github.com/bogdan-kulynych/libshe https://github.com/deevashwer/Fully-Homomorphic-Encryption https://github.com/deevashwer/Fully-Homomorphic-DGHV-and-Variants https://github.com/raduMMR/OMP-DGHV https://github.com/raduMMR/OMP-DGHV

Part 3: Leveled Homomorphic Encryption

Noise and Compactness

- Recall that in the scheme over integers we just saw:
 - Noise grows exponentially with the multiplicative depth.

• Scheme is not **compact** as ciphertext size grows with the size of the circuit.

• Let's see how to tackle these problems...



Noise Management Techniques

<u>BGV</u>: A scheme over **polynomial rings** $R_q = \mathbb{Z}_q[X]/(X^d + 1)$ based on the hardness of "Ring Learning With Errors" (RLWE) problem.

• How do objects in polynomial rings $R_q = \mathbb{Z}_q[X]/(X^d + 1)$ look like?

Think of them as a vectors of size d where each element is an integer in \mathbb{Z}_q .

123
$$d-1$$
 d \mathbb{Z}_q \mathbb{Z}_q \mathbb{Z}_q \mathbb{Z}_q \mathbb{Z}_q

• In BGV the noise increases **linearly** with the multiplicative depth! Lets see how...

Modulus-Switching

- Ciphertext $c \in R_Q$ for large modulus Q is an encryption of m.
- Scale *c* by (q/Q) and **round** appropriately with a <u>smaller</u> modulus $q \ll Q$. The resulted ciphertext $c' \in R_q$ is also a valid encryption of *m*.
- This allows to reduce the ciphertext **noise** by a factor $\approx (q/Q)$ without knowing the secret-key!

 $\boldsymbol{c}' = [(q/Q) \cdot \boldsymbol{c}]_q \qquad q$

0

Re-Linearization (Key-Switching)

$$\begin{array}{c} s \otimes s \\ \hline c_1 \\ c_2 \end{array} \begin{array}{c} c^{\star} \\ \end{array}$$

- Given BGV ciphertexts c_1 , c_2 decryptable to messages m_1 , m_2 under key s.
- After multiplication ciphertext $c^* = c_1 \cdot c_2$ has roughly d^2 elements which is decryptable with a **longer** key $s \otimes s$.

Re-Linearization (Key-Switching)





- To reduce the size of c^* back to d we use **Re-Linearization** technique:
- 1. Encrypt s[i] and $s[i] \cdot s[j]$ for $1 \le i, j \le d$ under a new secret key t.
- 2. Place encryptions of $\text{Enc}_t(s[i])$, $\text{Enc}_t(s[i] \cdot s[j])$ for $1 \le i, j \le d$ in **public key**.
- 3. Convert **long** ciphertext c^* to a **short** one c' (decryptable by new key t).
- Tradeoff between long ciphertexts (many) and long secret key (single).

Leveled Homomorphic Encryption

- Apply the **modulus-switching** technique after every multiplication, using a ladder of gradually decreasing moduli $q_L > q_{L-1} > \cdots > q_1 > q_0$.
- Freshly encrypted ciphertexts are over R_{q_L} and on ciphertexts over R_{q_0} we cannot compute anymore.
- After each multiplication preform a **Re-Linearization** on the resulted ciphertext.
- Performing these two operations together is sometimes called:
 "Ciphertext Refresh".



https://github.com/shaih/HElib

• C++ library that implements BGV HE scheme, along with many optimizations.



Initialization



Basic Arithmetic

Ctxt ctxt1(public_key), ctxt2(public_key);

```
ZZX poly1, poly2;
poly1.SetMaxLength(phi m);
poly2.SetMaxLength(phi m);
for (long i = 0; i < phi m; i++) {</pre>
    SetCoeff(poly1, i, RandomBnd(p));
    SetCoeff(poly2, i, RandomBnd(p));
public key.Encrypt(ctxt1, poly1);
public key.Encrypt(ctxt2, poly2);
ctxt1.addCtxt(ctxt2);
Ctxt ctxt3(public key), ctxt4(public key);
public key.Encrypt(ctxt3, to ZZX(14));
public key.Encrypt(ctxt4, to ZZX(80));
                                                              Multiplication
                                                              followed by
ctxt3.multiplyBy(ctxt4);
                                                              ciphertext refresh
cout << "ctxt3 level: " << ctxt3.findBaseLevel() << endl;</pre>
ZZX dec poly1, dec poly2;
secret key.Decrypt(dec poly1, ctxt1);
secret key.Decrypt(dec poly2, ctxt3);
```

Many Operations on Ciphertexts

Ciphertext arithmetic			
negate ()			
operator+= (const Ctxt &other)			
operator-= (const Ctxt &other)			
addCtxt (const Ctxt &other, bool negative=false)			
operator*= (const Ctxt &other)			
automorph (long k)			
operator>>= (long k)			
smartAutomorph (long k)			
automorphism with re-lienarization			
frobeniusAutomorph (long j) applies the automorphsim p^j using smartAutomorphism			
addConstant (const DoubleCRT &dcrt, double size=-1.0)			
addConstant (const ZZX &poly, double size=-1.0)			
addConstant (const ZZ &c)			

void	multByConstant (const DoubleCRT &dcrt, double size=-1.0)
void	multByConstant (const ZZX &poly, double size=-1.0)
void	multByConstant (const zzX &poly, double size=-1.0)
void	multByConstant (const ZZ &c)
void	xorConstant (const DoubleCRT &poly, double size=-1.0)
void	xorConstant (const ZZX &poly, double size=-1.0)
void	nxorConstant (const DoubleCRT &poly, double size=-1.0)
void	nxorConstant (const ZZX &poly, double size=-1.0)
void	divideByP ()
void	multByP (long e=1)
void	divideBy2 ()
void	extractBits (vector< Ctxt > &bits, long nBits2extract=0)
void	multiplyBy (const Ctxt &other)
void	multiplyBy2 (const Ctxt &other1, const Ctxt &other2)
void	square ()
void	cube ()
void	power (long e) raise ciphertext to some power

SIMD (Ciphertext Packing)

- Encrypt and pack **multiple** plaintext values into a **single** ciphertext.
- Main idea: Chinese Remainder Theorem over Polynomial Rings.



• Choose p such that R_p splits into s smaller rings R_{p_1}, \ldots, R_{p_s}

SIMD (Ciphertext Packing)

• This way we can process arrays of values at almost no extra cost.



• In practice: hundreds – thousands of slots in each ciphertext

EncryptedArray: Operations on Arrays of Slots

EncryptedArray ea(context);
long slots = ea.size(); slots: 1680

```
NewPlaintextArray pl(ea), p2(ea);
std::vector<long> v2(slots, 2);
```

```
random(ea, pl); //random values in [0,...,(p^r)-1]
encode(ea, p2, v2);
```

```
Ctxt cl(public_key), c2(public_key);
ea.encrypt(cl, public_key, pl);
ea.encrypt(c2, public key, p2);
```

cl.multiplyBy(c2);

```
Ctxt c3(ZeroCtxtLike, c1);
c3.addCtxt(c1);
```

ea.rotate(c1, 13);

```
std::vector<long> dec1, dec2, dec3;
ea.decrypt(c1, secret_key, dec1);
ea.decrypt(c2, secret_key, dec2);
ea.decrypt(c3, secret_key, dec3);
```

Computing On Integers

• Includes routines for addition/multiplication and comparisons of integers in binary representation (binaryArith.h, binaryCompare.h):

<u>CtPtrs</u>: Unified interface for vector of pointers to ciphertexts

Part 4: Fully Homomorphic Encryption

What have we achieved until now?

- We saw **Somewhat** and **Leveled** Homomorphic Encryption schemes.
- Still unable to compute arbitrary circuits / functions on encrypted data!
- Since suggested in 1978 by Rivest, Adleman and Dertouzos not feasible.
- This was the general situation until October 2008...

when Craig Gentry came up with the first suggested scheme for a Fully Homomorphic Encryption!



The "Bootstrapping method"



The "Bootstrapping method"

|noise| = p/2

N δM δM





The "Bootstrapping method"

|noise| = p/2

Bottomlin@egredleseronidise nevied indreesingsubles.man(b) a limit, use bootstrapping.noisesletvieltindnfix.edtleveEnan(b) riefiseed.until done!





From "Somewhat" to "Fully"

Dec(c)

 C_1

SWH – Can evaluate some circuits

Bootstrappable – Can also evaluate decryption circuits augmented by AND,XOR gates

FHE – Can evaluate all circuits

Augmented Decryption Circuit

 C_2

Enc(sk)

 $\operatorname{Enc}(c_1 \cdot c_2)$

AND

 $\operatorname{Enc}(c_1 \oplus c_2)$

XOR

Augmented Decryption Circuit

Dec(c)

 C_2

Enc(*sk*)

Dec(c)

 C_1

Enc(sk)

Dec(c)

Enc(sk)



https://github.com/lducas/FHEW

• Problem: Bootstrapping is an expensive process (5-30 min. per ctxt in HElib)

- Solution: Bootstrapping Homomorphic Encryption in less than a second!
 - Bootstrapping ciphertexts after every single operation.
 - Use of "cheap" and **functionally complete NAND** gate.
 - Works only over binary plaintext.

FHEW

#include <FHEW/LWE.h>
#include <FHEW/FHEW.h>

int main(int argc, char *argv[]) {

FHEW::Setup();

```
LWE::SecretKey secret_key;
LWE::KeyGen(secret_key);
```

```
FHEW::EvalKey eval_key;
FHEW::KeyGen(&eval_key, secret_key);
```

```
LWE::CipherText c1, c2, c_nand, c_or, c_nor, c_and, c_not;
```

```
LWE::Encrypt(&c1, secret_key, 0);
LWE::Encrypt(&c2, secret_key, 1);
```

```
FHEW::HomNAND(&c_nand, eval_key, c1, c2);
```

```
FHEW::HomNOT(&c_not, c1);
```

```
FHEW::HomGate(&c_or, BinGate::0R, eval_key, c1, c2);
FHEW::HomGate(&c_nor, BinGate::NOR, eval_key, c1, c2);
FHEW::HomGate(&c_and, BinGate::AND, eval_key, c1, c2);
```

```
int res_nand = LWE::Decrypt(secret_key, c_nand); // res_nand = (0 nand 1) = 1
int res_not = LWE::Decrypt(secret_key, c_not); // res_not = not(0) = 1
int res_nor = LWE::Decrypt(secret_key, c_nor); // res_nor = (0 nor 1) = 1
int res_nor = LWE::Decrypt(secret_key, c_nor); // res_nor = (0 nor 1) = 0
int res_not = LWE::Decrypt(secret_key, c_not); // res_not = (0 nor 1) = 0
```



https://github.com/tfhe/tfhe

- Very fast gate-by-gate bootstrapping (≈ 13 milliseconds).
- Supports the homomorphic evaluation of the 10 binary gates (NAND, OR, AND, XOR, XNOR, NOR, etc.), as well as the negation (NOT) and the MUX(a,b,c) = a ? b : c gate.
- Both FHEW an TFHE are based on the <u>GSW</u> cryptosystem.

Additional Implementations & Links

- HEAAN Supports fixed point arithmetics (also with **Bootstrapping**)
- <u>SEAL</u> Well-documented C++ library by Microsoft
- PALISADE General purpose C++ library for lattice cryptography
- <u>cuFHE</u> CUDA (NVIDIA GPU) accelerated FHE library
- Daniele Micciancio FHE Page
- <u>Vinod Vaikuntanathan FHE Page</u>
- <u>FHE Standardization Webpage</u>

Summary

- We have seen:
 - computing over encrypted data is possible via FHE.
 - it is still quite challenging, not trivial, and relatively slow.
 - several C++ implementations of FHE exist.
- There are other methods to compute over encrypted data, e.g.:
 - Secure Multi-Party Computation (MPC)

Thank You!