CSE 594 : Modern Cryptography 03/23/2017

Lecture 14: Hardness Assumptions

Instructor: Omkant Pandey *Scribe: Hyungjoon Koo, Parkavi Sundaresan*

1 Modular Arithmetic

Let $\mathbb{N}$ and $\mathbb{R}$ be set of natural and real numbers respectively. Let $\mathbb{Z}$ be set of integers. $\mathbb{Z}^+$ and $\mathbb{Z}^-$ represent positive and negative integers. For $n \in \mathbb{N}$, $\mathbb{Z}_N$ denotes set of integers *modulo* N as following:

$$\mathbb{Z}_N := \{0, 1, 2, ..., N-1\}$$

In this setting, we can perform "arithmetic in $\mathbb{Z}_N$" - addition, mulplication and division:

$$(a+b) \bmod N = (a \bmod N) + (b \bmod N) \bmod N$$

$$(a \times b) \bmod N = (a \bmod N) \times (b \bmod N) \bmod N$$

$r = a \bmod N$ when integer a is divided by N and r is the *remainder in* $\mathbb{Z}_N$

We say that "a is *congruent* to b modulo N" if a, b have the same remainder and write:

$$a \equiv b \bmod N$$

Thus $a \equiv 0 \bmod N$ if and only if $N|a$, where N divides a.

2 Greatest Common Divisor (GCD)

If a, b are two integers, $\gcd(a, b)$ denotes their GCD, greatest common divisor. If two integers a, b are non-zero and have no common factors (i.e., $\gcd(a, b) = 1$), a, b are *relatively prime*. It is easy to compute gcd for any two integers a, b using *Extended Euclidean*:

$$\exists a, b \in \mathbb{Z} \Rightarrow \exists x, y \in \mathbb{Z} \text{ such that } ax + by = \gcd(a, b)$$

If a, b are relatively prime, we can write it as following:

$$ax + by = 1 \Rightarrow ax \equiv 1 \bmod b$$

Now let $\mathbb{Z}_N^*$ be set of integers mod N that are relatively prime to N:

$$\mathbb{Z}_N^* = \{1 \leq x \leq N-1 : \ \gcd(x, N) = 1\} \Rightarrow \exists a \in \mathbb{Z}_N^* \ \exists x : \ ax = 1 \bmod N$$

Such an x is called the *inverse* of a.

3 Integers modulo a prime

We are of special interest when N is a prime number, say p, which defines:

$$\mathbb{Z}_p = \{0, 1, 2, ..., p-1\}$$
$$\mathbb{Z}_p^* = \{1 \leq x \leq N-1 : \ \gcd(x, p) = 1\} = \{1, 2, ..., p-1\}$$
$$|\mathbb{Z}_p^*| = p - 1$$

3.1 Fermat's Little Theorem

If p is a prime, then for any $a \in \mathbb{Z}_p^*$:

Theorem 1 *Fermat's Little Theorem*

$$a^{p-1} \ mod \ p = 1$$

3.2 Euler's Generalization

Fermat's Little Theorem can be generalized by Euler's theorem.

Theorem 2 *Euler's Theorem*

$$For \ any \ N \in \mathbb{N} \ and \ a \in \mathbb{Z}_p^* :$$
$$a^{\phi(N)} \ mod \ N = 1 \ where \ \phi(N) = |\mathbb{Z}_N^*|$$

$\phi(N)$ is Euler's totient function, which denotes the number of $n \in \mathbb{Z}_N^*$ that is relative to N. Every interger N can be written as the multiplication of its factors as the following theorem.

Theorem 3 *Fundamental Theorem of Arithmetic*

$$N = \prod_{i=1}^{k} p_i^{e_i}$$

$$for \ prime \ factors \ p_1 < p_2 < ... < p_k \ and \ positive \ integers \ e_i > 0$$

$$This \ factorization \ is \ unique \ with \ empty \ product \ taken \ to \ be \ 1$$

$$\phi(N) = N \cdot \prod_{i=1}^{k} (1 - \frac{1}{p_i})$$

If $N = pq$ for distict primes p and q, then $\phi(N) = (p-1) \cdot (q-1)$ because there are q and p multiples for p and q respectively in $N = pq$.

4 Groups

A set G denotes a group with a "group operation": $G \times G \to G$, satisfying the following properties:

- Closure: $\forall a, b \in G, a \bigodot b \in G$
- Identity: $\exists e \in G$ (identity) $s.t.\ \forall a \in G:\ a \bigodot e = e \bigodot a = a$
- Associativity: $\forall a, b, c \in G:\ (a \bigodot b) \bigodot c = a \bigodot (b \bigodot c)$
- Inverse: $\forall a \in G,\ \exists b \in G\ s.t\ a \bigodot b = b \bigodot a = e$ (identity)

In particular, a group with *commitative* property is called "Abelian group" where $\forall a, b \in G : a \bigodot b = b \bigodot a$. For example, $(\mathbb{Z}_N, +), (\mathbb{Z}_N^*, \times)$ are "additive" and "multiplicative" groups for all N.

Theorem 4 *Corollary of Lagrange's Theorem*

$$x^{|G|} = e$$

If the set $\{g, g^2, ...\} = G$, $g \in G$ is a *generator* of G. The set of all generators of G is denoted by Gen_G.

5 Discrete Logarithm Problem

An instance (p, g, y) of the discrete logarithm problem consists of a large prime p and two elements $g, y \in \mathbb{Z}_p^*$. Since p is prime, difference between $\mathbb{Z}_p$ and $\mathbb{Z}_p^*$ is that $\mathbb{Z}_p$ includes 0.

The task is to find an x such that $g^x = y \ mod\ p$. This task is believed to be hard - No known algorithm that can break the problem - except for few special cases. This include:
$g = 1$
p has some special structure like $p = 2^k + 1$
$p - 1$ has many small factors

However, if g is a generator, the problem is believed to be hard. Normally we want to work with a group such that order of the group, $|G|$,the number of elements in G, is prime. However, $\mathbb{Z}_p^*$ has order $(p - 1)$, which is not prime. If $p = 2q + 1$ and q is also a prime, p is called Sophie Germain prime or a safe prime. This subgroup has order q which is prime. The practical method for sampling safe primes is simple: first pick a prime q as usual, and then check whether $2q + 1$ is also prime.

Number of elements in $\mathbb{Z}_p^*$,

$$|\mathbb{Z}_p^*| = p - 1 = 2q$$

Consider a subset $G_q = \{x^2 : x \in \mathbb{Z}_p^*\}$. This set G_q is a set of quadratic residues *mod p*. G_q has an order of $q = (p - 1)/2$. It is easy to prove that G_q is a group of prime order q, by proving that all properties of a group are satisfied by G_q.

For example, we can compute inverse $\forall a$ in G_q by:

$\forall a \in G_q$, if at G_q $\exists x : x^2 = a \bmod p$,

$$\begin{aligned} a^q \ mod \ p &= a^{(p-1)/2} \ mod \ p \\ &= x^{2^{(p-1)/2}} \\ &= x^{(p-1)} \ mod \ p \\ &= 1 \ \ (By \ Fermat's \ Little \ Theorem) \\ a.a^{(q-1)} \ mod \ p &= 1 \\ So, \ Inverse \ of \ a &= a^{(q-1)} \ mod \ p \end{aligned}$$

Since G is a prime order group, cycle through all q elements of G by applying the group operation to the generator over and over again. Other ways to construct prime order groups include group formed by points on an appropriate elliptic curve. It is hard to compute discrete log in prime order groups.

Assumption 1 *Discrete Log Assumption*

If G_q is a group of prime order q, then, for every non-uniform PPT A, there exists a negligible function μ such that:

$$Pr\Big[q \leftarrow \Pi_n; g \leftarrow Gen_{G_q}; x \leftarrow \mathbb{Z}_q : A(1^n, g^x) = x\Big] \leqslant \mu(n).$$

Adversary $A^{'}$s advantage in solving the discrete logarithm problem is negligible. Π_n is the set of n-bit prime numbers. It is important that G is a group of prime-order. So, the normal multiplicative group $\mathbb{Z}_p$ has order $(p-1)$ and therefore does not satisfy the assumption.

6 Diffie-Hellman Problems

6.1 Computational Diffie-Hellman

The adversary gets $X = g^x \ mod \ p$ and $Y = g^y \ mod \ p$ and (p, g). Let G_q be a cyclic group of prime order q. The CDH problem is :

Given (g, q, g^x, g^y), compute $g^{xy} \in G_q$,
where, x and y are random and all computations are in G_q

Assumption 2 *CDH Assumption is that Adversary $A^{'}$s probability of solving the computational Diffie-Hellman problem is negligible.*

$$\begin{gathered} \forall \ non-uniform \ PPT \ A, \ \exists \ negligible \ \mu \ such \ that, \\ \forall \ n: \ A \ solves \ CDH \ problem \ with \ probablility \ atmost \ \mu(n) \end{gathered}$$

Note : When working with a safe $p = 2q + 1$, g can be generator for order q subgroup, and computations can be *modulo p*.

6.2 Decisional Diffie-Hellman

g^{xy} looks indistinguishable from a random group element.

Assumption 3 *Decisional Diffie-Hellman Assumption is the following ensembles are computationally indistinguishable:*

$$\{p \leftarrow \widetilde{\Pi}_n, y \leftarrow Gen_q, a, b \leftarrow \mathbb{Z}_q : g, p, g^a, g^b, g^{ab}\}_n \approx$$
$$\{p \leftarrow \widetilde{\Pi}_n, y \leftarrow Gen_q, a, b, z \leftarrow \mathbb{Z}_q : g, p, g^a, g^b, g^z\}_n$$

where x,y,z are all random and all computations are in G_q. Adverasary is unable to say if its g^{xy} or random g^z.

$\forall$ non-uniform PPT distinguishers, D, $\exists$ negligible μ such that $\forall n : D$ solves the DDH problem with probability at most $1/2 + \mu(n)$

It is crucial for the DDH assumption that the group within which we work is a prime-order group. In a prime order group, all elements except the identity have the same order.

7 RSA

RSA = Rivest, Shamir, Adleman Let p,q be large random primes of roughly the same size and $N = pq$. N is called a RSA modulus.

$\phi(N)$ is the order of $\mathbb{Z}_N^*$, which is $(p-1)(q-1)$, where,

$$\mathbb{Z}_N^* = \{x \in \mathbb{Z}_N : gcd(x, N) = 1\}$$

Let e be an odd number between 1 and $\phi(N)$ such that

$$gcd(e\ , \phi(N)) = 1$$

. Therefore, $e \in \mathbb{Z}^*_{\phi(N)}$.

By using Extended Euclidean Algorithm, we can find a d such that:

$$e.d\ =\ 1\ mod\ \phi(N).$$

If $\phi(N)$ is known, you can compute d.But if $\phi(N)$ is not known, d seems hard to compute. Therefore, $\phi(N)$ must be kept secret.

Let N,e,d be as before so that $e.d\ =\ 1\ mod\ \phi(N)$. d can be used to compute e-th root of numbers modulo N.
Suppose that $y\ =\ x^e\ mod\ N$, then:

$$\begin{aligned} y^d\ mod\ N &= x^{ed}\ mod\ N \\ &= x^{ed\ mod\ \phi(N)}\ mod\ N \\ &= x\ mod\ N \end{aligned}$$

RSA Assumption is that, without d, it seems hard to compute e-th root of numbers modulo N. So even if (N, e) are given it would be hard to find $e - th$ roots; can be used as public key. d and factors of N should be kept secret; d can be used as secret trapdoor.

Assumption 4 *RSA Assumption : For every non-uniform PPT A there exists a negligible function μ such that for all $n \in \mathbb{N}$:*

$$Pr\left[\begin{array}{l} p,q \leftarrow \Pi_n; N \leftarrow pq; \\ e \leftarrow \mathbb{Z}^*_{\phi(N)}; y \leftarrow \mathbb{Z}^*_N; : x^e = y \ mod \ N \\ x \leftarrow A(N,e,y) \end{array}\right] \leqslant \mu(n).$$

For N, e as above, the following is called the RSA function:

$$f_{N,e}(x) = x^e \ mod \ N$$

The RSA Function actually yields a collection of trapdoor one-way permutations.

8 Learning With Errors (LWE)

Let S be a vector of length n, where each element belongs to $\mathbb{Z}_q^n$, some modulus q, greater than 2, and parameter n.

$$s = (s_1, ..., s_n) \in \mathbb{Z}_q^n$$

Suppose you are given many equations for known a values.

$$a_1.s_1 + a_2.s_2 + ... + a_n.s_1 = b_1(mod\ q)$$

$$a_1^{'}.s_1 + a_2^{'}.s_2 + ... + a_n^{'}.s_1 = b_2(mod\ q)$$

etc

This can be solved by using Gaussian elimination. But, it may not work if the equation contain some errors. In particular, solving a system of equations, each with some noise becomes hard when all equations have independent error,according to the Normal Distribution, with standard deviation $\alpha q > \sqrt{n}$.

$$a_1.s_1 + a_2.s_2 + ... + a_n.s_1 \approx b_1(mod\ q)$$

$$a_1^{'}.s_1 + a_2^{'}.s_2 + ... + a_n^{'}.s_1 \approx b_2(mod\ q)$$

etc

Errors should be picked, from Gaussian or Normal distribution, such that error is small enough for problem to be unique but large enough for problem to be hard to solve.