Asymmetric Encryption
Information Security Principles

tgeng@email.sc.edu

September 14, 2021
Objective

- Diffie-Hellman Key Exchange
- RSA
Asymmetric Encryption

Asymmetric encryption, or public-key encryption, is a cryptographic system that uses pair of keys. Each pair consists of a public key and a private key.

- public key: may be known to others
- private key: may not be known by anyone except the owner

The generation of public key and private key on mathematical problem, and they are one-to-one mapping.

\[ E : M \times K_r \rightarrow C \]
\[ D : C \times K_u \rightarrow M \]
Number Theory Background

Number Theory is a branch of pure mathematics devoted primarily to the study of the integers. It is the study of the set of positive whole numbers

$$1, 2, 3, 4, 5, 6, 7, \ldots$$

which are often called the set of natural numbers.

Since ancient times, people have separated the natural numbers into a variety of different types:

- odd: $1, 3, 5, 7, 9, 11, \ldots$
- even: $2, 4, 6, 8, 10, \ldots$
- square: $1, 4, 9, 16, 25, 36, \ldots$
- prime: $2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, \ldots$
- perfect: $6, 28, 496, \ldots$
- Fibonacci: $1, 1, 2, 3, 5, 8, 13, 21, \ldots$
1 (modulo 4): 1, 5, 9, 13, 17, 21, 25, . . .
A number is said to be congruent to 1 (modulo 4) if it leaves a remainder of 1 when divided by 4.

If two numbers $b$ and $c$ have the property that their difference $b - c$ is integrally divisible by a number $m$ such that $(b - c)/m$ is an integer, then $b$ and $c$ are said to be congruent modulo $m$.

The statement “$b$ is congruent to $c$ (modulo $m$)” is written as

$$b \equiv c \pmod{m}$$
Binary Operations

\(+, -, \times, \div\) are four binary operations in mathematics, but actually only two in reality.

- \(a - b\) is equal to \(a + (-b)\)
  \(\quad\) 
  \(a\) plus the additive inverse of \(b\)

- \(a/b\) is equal to \(a \times \frac{1}{b}\)
  \(\quad\) 
  \(a\) multiply by the multiplicative inverse of \(b\)
Group

Group is a set of elements together with a binary operation that fulfills certain properties.

**Definition**

A *group* is a 2-tuple \((G, \cdot)\) where \(G\) is a set, \(\cdot\) is a binary operation, and they satisfy the following four properties:

1. **Closure**: \(\forall x, y \in G : x \cdot y \in G\).
2. **Identity**: There exists an neutral element in \(G\), often denoted by \(1\) (or \(e\)), such that \(\forall x \in G : 1 \cdot x = x \cdot 1 = x\).
3. **Inverses**: \(\forall x \in G, \exists y \in G : x \cdot y = y \cdot x = 1\).
4. **Associativity**: \(\forall x, y, z \in G : (x \cdot y) \cdot z = x \cdot (y \cdot z)\).
Multiplicative Group

It is called multiplicative to denote that the group operation is multiplication (as opposed to being addition).

Definition

A *multiplicative group* is a 2-tuple \((G, \star)\) where \(G\) is a set, \(\star\) is a binary operation, and they satisfy the following four properties:

1. **Closure**: \(\forall x, y \in G : x \star y \in G\).
2. **Identity**: There exists an neutral element in \(G\), often denoted by 1 (or \(e\)), such that \(\forall x \in G : 1 \star x = x \star 1 = x\).
3. **Inverses**: \(\forall x \in G, \exists y \in G : x \star y = y \star x = 1\).
4. **Associativity**: \(\forall x, y, z \in G : (x \star y) \star z = x \star (y \star z)\).

Sometimes, a group is denoted as \(G\) only when the operator is clear in the context.
Cyclic Groups

- One very common type of group is the cyclic groups. This group is isomorphic to the group of integers \((\text{modulo } n)\), is denoted \(Z_n\), and is defined for every integer \(n > 1\).
- A cyclic group is a group that can be generated by a single element \(g\). It contains an element \(g\) such that every other element of the group may be obtained by repeatedly applying the group operation or its inverse to \(g\).
- Each element can be written as a power of \(g\) in multiplicative notation, or as a multiple of \(g\) in additive notation.
- This element \(g\) is called a \textit{generator} of the group.
- The name “cyclic” is kind of misleading, it’s possible to generate infinitely many elements and not form any literal cycles.
Cyclic Groups

A trivial example is the group $Z_n$, the additive group of integers \textit{modulo} $n$. In $Z_n$, 1 is always a generator:

\begin{align*}
1 &\equiv 1 \text{ mod } n \\
1 + 1 &\equiv 2 \text{ mod } n \\
1 + 1 + 1 &\equiv 3 \text{ mod } n \\
1 + 1 + 1 + \ldots + 1 &\equiv n \equiv 0 \text{ mod } n
\end{align*}
Cyclic Group $\mathbb{Z}_5$

Specifically, $\mathbb{Z}_5$ can be shown as \{0, 1, 2, 3, 4\}:

\[
\begin{align*}
1 & \equiv 1 \mod 5 \\
1 + 1 & \equiv 2 \mod 5 \\
1 + 1 + 1 & \equiv 3 \mod 5 \\
1 + 1 + 1 + 1 & \equiv 4 \mod 5 \\
1 + 1 + 1 + 1 + 1 & \equiv 0 \mod 5 \\
1 + 1 + 1 + 1 + 1 + 1 & \equiv 1 \mod 5 \\
1 + 1 + 1 + 1 + 1 + 1 + 1 & \equiv 2 \mod 5 \\
\vdots
\end{align*}
\]
Cyclic Group $\mathbb{Z}_5^*$

$\mathbb{Z}_n^*$ is the multiplicative group *modulo* $n$, and $\mathbb{Z}_5^*$ is cyclic multiplicative group *modulo* 5. 2 is one generator for this group.

\[
\begin{align*}
2^1 & \equiv 2 \mod 5 \\
2^2 & \equiv 4 \mod 5 \\
2^3 & \equiv 8 \equiv 3 \mod 5 \\
2^4 & \equiv 16 \equiv 1 \mod 5 \\
2^5 & \equiv 32 \equiv 2 \mod 5 \\
2^6 & \equiv 64 \equiv 4 \mod 5 \\
2^7 & \equiv 128 \equiv 3 \mod 5 \\
\ldots
\end{align*}
\]

If $\mathbb{Z}_5^*$ is cyclic and $g$ is a generator of $\mathbb{Z}_n^*$, then $g$ is also called a **primitive root modulo** $n$.  

Diffie-Hellman Key Exchange

Diffie-Hellman key exchange is a method of securely exchanging secret key over a public channel, and it was one of the first public-key protocols. It was conceived by Ralph Merkle and named after Whitfield Diffie and Martin Hellman.

- It is based on exponentiation in a finite field, also called Galois field (modulo a prime or a polynomial). This is easy to calculate.
- The security of Diffie-Hellman relies on the difficulty of computing discrete logarithms (similar to factoring).
DHKE Cryptographic Explanation

The simplest and the original implementation of DH uses the multiplicative group of integers modulo $p$, where $p$ is prime and $g$ is a primitive root modulo $p$. 
DHKE Steps

Alice and Bob need to exchange a secret key securely over an insecure channel. Secret key should never be on the insecure channel.

1. Alice and Bob to agree on a large prime $p$ and a nonzero integer $g$ (where $p$ is large (typically at least 512 bits) and $g$ is a primitive root modulo $p$)
2. Alice and Bob make the values of $p$ and $g$ public knowledge.
3. Alice chooses a large random number $a$ as her private key and Bob similarly chooses a large number $b$
4. Alice then computes $A = g^a \mod p$ and sends $A$ to Bob; Bob computes $B = g^b \mod p$ and sends $B$ to Alice
5. Both Alice and Bob compute their shared key $K = g^{ab} \mod p$, which Alice computes as

$$K = B^a \mod p = (g^b \mod p)^a \mod p$$

and Bob computes as

$$K = A^b \mod p = (g^a \mod p)^b \mod p$$

6. Alice and Bob can now use their shared key $K$ to exchange information.
Discrete Logarithm

A potential eavesdropper (Eve) would first need to obtain $K$ with the knowledge of $p, g, A, B$ only if she wants to do the eavesdropping.

\[
K = B^a \mod p = (g^b \mod p)^a \mod p
\]

\[
K = A^b \mod p = (g^a \mod p)^b \mod p
\]

The only thing Eve doesn’t know is $a$ or $b$. It is a discrete logarithm problem to get $a$ from $A = g^a \mod p$ or get $b$ from $B = g^b \mod p$

Every element $h$ in $G$ can be written as $g^x \mod p$ for some $x$. Brute force on the key is to find the value of $x$. 

Discrete Logarithm

If \( p, g \) is small, so easy to compute the \( x(a, \text{ or } b) \).

However, the length of \( p \) is typically very large, for example:

2519590847565789349402718324004839857142928212620403202777713783604366202070759
5556264018525880784406918290641249515082189298559149176184502808489120072844992
6873928072877767359714183472702618963750149718246911650776133798590957000973304
5974880842840179742910064245869181719511874612151517265463228221686998754918242
2433637259085141865462043576798423387184774447920739934236584823824281198163815
010674810451660377306056201619676256133844143603839044149526344321901146575444
5417842402092461651572335077870774981712577246796292638635637328991215483143816
7899885040445364023527381951378636564391212010397122822120720357
Algorithm for Discrete Logarithm

$n$ stands for length of the key, such as 1024 bits, 2048 bits.

- Pollard’s rho algorithm for logarithms: $O(x^{n/2})$
- Exhaustion: $O(x^n)$
  - If the key used in DHKE is 1024, that would be $2^{1024}$, or around $10^{300}$.
  - The total number of atoms in the whole universe is less than $10^{100}$. 
Example of DHKE

- Shared $p = 353$, and $g = 3$
- Alice chooses $a = 97$ and Bob chooses $b = 233$
- Alice computes $A = 3^{97} \mod 353 = 1.908805632e46 \mod 353 = 40$, and Bob computes $B = 3^{233} \mod 353 = 1.476564251e111 \mod 353 = 248$
- $K = A^b \mod 353 = B^a \mod 353 = 160$
RSA

- Ron Rivest, Adi Shamir, Leonard Adleman
- An RSA user creates and publishes a public key based on two large prime numbers
- Security of RSA relies on the practical difficulty of factoring the product of two large prime numbers, there are no published methods to defeat the system if a large enough key is used
- RSA is a relatively slow algorithm. Because of this, it is not commonly used to directly encrypt user data. More often, RSA is used to transmit shared keys for symmetric key cryptography, which are then used for bulk encryption and decryption.
If two integers have no common factor other than 1, we say these two integers are coprime. For example, 15 and 32 are coprime.

We have the following properties:

1. Any two primes $p$ and $q$ are coprime such as 13 and 61
2. One number $p$ is prime, as long as the other one is not multiple of $p$, they are coprime such as 3 and 10
3. The bigger number in two numbers is prime, then they are coprime
4. 1 is coprime with any natural number
5. If $p$ is an integer larger than 1, then $p$ and $p - 1$ are coprime
6. If $p$ is an odd number larger than 1, then $p$ and $p - 2$ are coprime
Given any positive integer $n$, how many integers less than or equal to $n$ are coprime with $n$? For example, between 1 and 8, how many integers are coprime with 8?

Euler’s totient function can answer this problem, and the result is denoted by $\varphi(n)$. For example, $\varphi(8) = 4$ since 1, 3, 5, 7 are coprime with 8.
Euler’s totient function

1. If $n = 1$, then $\varphi(1) = 1$. Since 1 is coprime with any number, including itself.

2. If $n$ is prime, then $\varphi(n) = n - 1$. Since a prime number is coprime with any numbers smaller than it (property 3).

3. If $n$ is a power of one prime $p^k$, then $\varphi(p^k) = p^k - p^{k-1}$.

   $\varphi(8) = \varphi(2^3) = 2^3 - 2^2 = 8 - 4 = 4$

   The reason of this is that only if a number is not a multiple of $p$, it will be coprime with $n$ ($n$ is some powers of $p$), and the number which is multiple of $p$ are $1 \times p$, $2 \times p$, $3 \times p$, ..., $p^{-1} \times p$.

4. If $n$ can be factorized as the product of two primes $n = p_1 \times p_2$, then

   $\varphi(n) = \varphi(p_1 p_2) = \varphi(p_1) \varphi(p_2)$

Common form of Euler’s totient function:

$$\varphi(n) = n \left(1 - \frac{1}{p_1}\right) \left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_r}\right)$$

where $n = p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r}$.
Euler’s theorem

Definition

If two positive number $a$ and $n$ are coprime, then

$$a^\varphi(n) \equiv 1 \mod n$$

For example, 7 and 10 are coprime, then we have $7^{\varphi(10)} \equiv 1 \mod 10$. We know that $\varphi(10) = 4$, so $7^4 \equiv 1 \mod 10$. 
**Modular multiplicative inverse**

**Definition**

For two positive numbers $a$ and $n$ are coprime, then there must exist

$$ab \equiv 1 \mod n$$

then $b$ is the *modular multiplicative inverse* of $a$.

For example, 3 and 11 are coprime, then the modular multiplicative inverse of 3 is 4, since $3 \times 4 - 1$ is a multiple of 11. Of course, its modular multiplicative inverse has more: 

$$\{..., -18, -7, 4, 15, 26, ...\}$$
Key generation

1. Choose two different prime numbers $p$ and $q$
2. Compute $n = pq$
3. Compute $\varphi(n) = (p-1)(q-1)$
4. Randomly choose $e$ such that $1 < e < \varphi(n)$ and $e$ is coprime with $\varphi(n)$
5. Compute the modular multiplicative inverse of $e$ such that $ed \equiv 1 \mod \varphi(n)$
6. Encapsulate $(n, e)$ as public key, $(n, d)$ as private key.
Encryption

$m$ is the plaintext message, and $c$ is the ciphertext.

- Encryption with public key $(n, e)$

$$m^e \equiv c \mod n$$

- Decryption with private key $(n, d)$

$$c^d \equiv m \mod n$$
Reliability

During the key pair generation, we have the following:

- $p$
- $q$
- $n$
- $\varphi(n)$
- $e$
- $d$

Only $n$ and $e$ are public, the other 4 are not public, the most important one or most secret should be $d$ since $(n, d)$ is the private key, since once $d$ is leaked, the private key is leaked.

Is it possible for the attacker to compute $d$ with the knowledge of $n$ and $e$?
Product factorization

- $ed \equiv 1 \mod \varphi(n)$, only we know $e$ and $\varphi(n)$, we can get $d$
- $\varphi(n) = (p-1)(q-1)$, only we know $p$ and $q$, we can get $\varphi(n)$
- $n = pq$, only we factorize $n$, we can get $p$ and $q$

3233 = 61 * 52 is easy to do, but it is not easy to do the factor of:

1230186684530117755130494958384962720772853569595334792197322452452151726400507263
6575187452021997864693899564749427740638459251925573263034537315482685079170261
221429134616704292143116022212240479274737794080665351419597459856902143413
Largest encryption key cracked

- In 2019, an RSA key that is 795 bits (240 decimal digits, RSA-240) got factored.
  - Clusters of computers in France, Germany, and the US
  - Equivalent of single computer core running for 35 million hours, or almost 4000 years
- Before this one, RSA cracking record is RSA-232 768 bits
Application of Public Key Encryption

- Key establish (Key exchange)
  - In practice, asymmetric encryption is firstly used to establish the key
  - The established key is then used in subsequent communication through inexpensive symmetric cryptography

- Digital Signature
Summary

- The idea of public key encryption is proposed by Ralph Merkle
- RSA, ElGamal, DSA, ECDSA (Elliptic Curve)