DH问题与加密

哈尔滨工业大学

张宇

2024春

概览

1. 循环群与离散对数（Cyclic Groups/Discrete Logrithms）
2. DH假设和应用
3. Elgamal加密方案
4. 椭圆曲线（ECC）加密方案

循环群

循环群（Cyclic Groups）与生成元（Generators）

- $\mathbb{G}$ 是一个群并且一个元素 $g \in \mathbb{G}$通过运算生成一个子群 $\langle g \rangle \stackrel{\text{def}}{=} \{g^0, g^1, \ldots,\} = \{g^0, g^1, \ldots, g^{i-1}\}$。
- g 的阶是最小的正整数 i 令 $g^i = 1$。
- $\mathbb{G}$ 是一个循环群（cyclic group）如果 $\exists\, g$ 有阶 $m = |\mathbb{G}|$. $\langle g \rangle = \mathbb{G}$, g 是 $\mathbb{G}$ 的生成元。注：循环群中存在一个元素通过指数运算可生成整个群中每个元素。
- 例题： 乘法下的$\mathbb{Z}_6^*$, $\mathbb{Z}_7^*$,或 $\mathbb{Z}_8^*$ 是循环群吗? 找到生成元。

离散对数

- 如果 $\mathbb{G}$ 是阶为 q 的循环群，那么 $\exists$ 生成元 $g \in \mathbb{G}$ 使得 $\{g^0, g^1, \ldots, g^{q-1}\} = \mathbb{G}$。
- $\forall h \in \mathbb{G}$, $\exists$ 唯一的 $x \in \mathbb{Z}_q$ 使得 $g^x = h$。
- $x = \log_g h$ 是以g为底h的离散对数（discrete logarithm）。
- 如果 $g^{x'} = h$, 那么 $\log_g h = [x' \bmod q]$。
- $\log_g 1 = 0$ 并且 $\log_g(h_1 \cdot h_2) = [(\log_g h_1 + \log_g h_2) \bmod q]$。

Show an instance of DL problem in $\mathbb{Z}_7^*$

离散对数算法概览

- 给定一个生成元 $g \in \mathbb{G}$ 并且 $y \in \langle g \rangle$，求 x 使得 $g^x = y$.
- 蛮力： $\mathcal{O}(q)$, $q = \mathsf{ord}(g)$ 是 $\langle g \rangle$ 的阶。
- Baby-step/giant-step [Shanks]: $\mathcal{O}(\sqrt{q} \cdot \mathsf{polylog}(q))$.
- Pohlig-Hellman算法：当 q 有较小因子。
- Index calculus 法: $\mathcal{O}(\exp\left(\sqrt{n \cdot \log n}\right))$.
- 已知最好的算法是通用数域筛法：$\mathcal{O}(\exp(n^{1/3} \cdot (\log n)^{2/3}))$.
- 椭圆曲线群 vs. $\mathbb{Z}_p^*$: 在保证安全性相同的同时，更高效。(1024-bit $\mathbb{Z}_p^*$ 和 132-bit 椭圆曲线都需要 2^{66} 步来破解。)

离散对数假设

- 离散对数（discrete logarithm）实验 DLog$_{\mathcal{A},\mathcal{G}}(n)$:
  - 运行一个群生成算法 $\mathcal{G}(1^n)$ 来产生 $(\mathbb{G}, q, g)$，其中 $\mathbb{G}$ 是阶为 q ($\|q\| = n$) 的循环群，并且 g 是 $\mathbb{G}$ 的生成元。
  - 挑选一个 $h \leftarrow \mathbb{G}$. ($x' \leftarrow \mathbb{Z}_q$ and $h := g^{x'}$)
  - 敌手 $\mathcal{A}$ 给定 $\mathbb{G}, q, g, h$，并且输出 $x \in \mathbb{Z}_q$.
  - 实验成功 DLog$_{\mathcal{A},\mathcal{G}}(n) = 1$，如果 $g^x = h$, 否则 0 。
- 定义：离散对数问题相对于群$\mathcal{G}$是难的，如果 $\forall$ ppt 算法 $\mathcal{A}$, $\exists$ negl 使得 $\Pr[\text{DLog}_{\mathcal{A},\mathcal{G}}(n) = 1] \leq \text{negl}(n)$.

DH假设

- 计算性DH（Computational Diffie-Hellman, CDH）问题：
  $\mathsf{DH}_g(h_1, h_2) \stackrel{\text{def}}{=} g^{\log_g h_1 \cdot \log_g h_2}$
- 判断性DH（Decisional Diffie-Hellman, DDH)）问题：区分 $\mathsf{DH}_g(h_1, h_2)$ 与一个随机的群元素 h'.
- 定义：DDH问题与$\mathcal{G}$相关的是难的，如果 $\forall$ ppt $\mathcal{A}$, $\exists$ negl 使得
  $|\Pr[\mathcal{A}(\mathbb{G}, q, g, g^x, g^y, g^z) = 1] - \Pr[\mathcal{A}(\mathbb{G}, q, g, g^x, g^y, g^{xy}) = 1]| \leq \mathsf{negl}(n)$.
- DL, CDH 和 DDH 的难解性：DDH 比 CDH 和 DL 容易。

密钥交换实验

- 密钥交换实验（key-exchange experiment） $\mathsf{KE}^{\mathsf{eav}}_{\mathcal{A},\Pi}(n)$:
  1. 双方持有安全参数 1^n 执行协议 Π。Π 执行的结果为对话记录 (transcript) **trans** 包含双方发送的所有消息，以及各方都输出的密钥 k 。
  2. 选择一个随机比特 $b \leftarrow \{0,1\}$ 。如果 $b=0$ 那么选择 $\hat{k} \leftarrow \{0,1\}^n$ u.a.r；如果 $b=1$ 那么令 $\hat{k} := k$。
  3. 敌手 $\mathcal{A}$ 给定 **trans** 和 $\hat{k}$, 并且输出一个比特 b'。
  4. $\mathsf{KE}^{\mathsf{eav}}_{\mathcal{A},\Pi}(n)=1$ 如果 $b'=b$，否则 0 。
- 定义：一个密钥交换协议 Π 在出现窃听者攻击下是安全的，如果 $\forall$ ppt $\mathcal{A}$, $\exists$ **negl** 使得

  $\Pr[\mathsf{KE}^{\mathsf{eav}}_{\mathcal{A},\Pi}(n)=1] < \frac{1}{2} + \mathsf{negl}(n).$

DH密钥交换

$(\mathbb{G}, q, g) \leftarrow \mathcal{G}$

$x \leftarrow \mathbb{Z}_q$

$h_1 := g^x$ $\xrightarrow{\mathbb{G}, q, g, h_1}$

$y \leftarrow \mathbb{Z}_q$

$\xleftarrow{h_2}$ $h_2 := g^y$

$k_A := h_2^x$ $\qquad$ $k_B := h_1^y$

Q: $k_A = k_B = k = ?$

$\widehat{\mathsf{KE}}^{\mathsf{eav}}_{\mathcal{A},\Pi}$ denote an experiment where if $b = 0$ the adversary is given $\hat{k} \leftarrow \mathbb{G}$.

Theorem 5

If DDH problem is hard relative to $\mathcal{G}$, then DH key-exchange protocol Π is secure in the presence of an eavesdropper (with respect to the modified experiment $\widehat{\mathsf{KE}}^{\mathsf{eav}}_{\mathcal{A},\Pi}$).

Security

Insecurity against active adversaries (Man-In-The-Middle).

使用指数阶的群

- 离散对数问题在质数阶群上是最难的。
- 在质数阶群上找一个生成元很简单。
- 任何非零指数在以质数阶为模下都可逆。
- DDH问题是难题的必要条件是DDH的解与群中随机元素之间是不可区分的。在质数阶群上这基本成立。
- 循环群生成算法：产生一个强质数p，阶为q=(p-1)/2，随机选择一个x \in \mathbb{Z}^*_p，得到生成元g=x^2，输出p, q, g。

  - $y \in \mathbb{Z}_p^*$ 是模p下的二次剩余（quadratic residue modulo），如果 $\exists x \in \mathbb{Z}_p^*$ 使得 $x^2 \equiv y \pmod{p}$
  - 例题：$\mathbb{Z}_7^*$ 下的二次剩余?
  - QR集合是一个子群（满足群条件），阶为 $(p-1)/2$，因为 $x^2 \equiv (p-x)^2 \pmod{p}$。
  - p 是一个强质数（strong prime），如果 $p = 2q + 1$ 且 q 是质数。
  - 强质数下的二次剩余子群是一个循环群，因为群的阶是质数。

课堂练习

$\mathbb{G} = \mathbb{Z}_{11}^*$

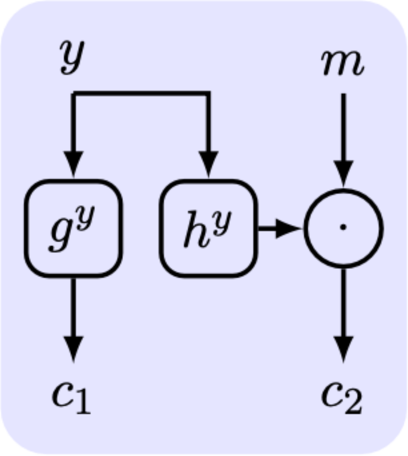
The order $q =?$
The set of quadratic residues ?
Is $g = 3$ a generator?
If $x = 3$ and $y = 4$, what's the message from Bob to Alice?
How does Alice compute the key?
How does Bob compute the key?

完美保密引理

- 引理：$\mathbb{G}$ 是有限群并且 $m \in \mathbb{G}$ 是任意元素。那么选择随机 $k \leftarrow \mathbb{G}$ 并令 $c := m \cdot k$，将得到与随机选择的 $c \leftarrow \mathbb{G}$ 相同的分布，即 $\forall g \in \mathbb{G}$: $\Pr[m \cdot k = g] = 1/|\mathbb{G}|$。
- 证明：$g \in \mathbb{G}$ 是任意的，那么 $\Pr[m \cdot k = g] = \Pr[k = m^{-1} \cdot g]$。由于 k 均匀随机选择，选择 k 的概率与一个固定元素 $m^{-1} \cdot g$ 相同，都是 $1/|\mathbb{G}|$。
- 注：这是一种完美保密的私钥加密方案，将一个元素（明文）与另一个元素（密钥）的运算得到第三个元素（密文），与之前一个字母的移位密码是完美保密是类似的。

ElGamal加密方案

An algorithm $\mathcal{G}$, on input 1^n, outputs a description of a cyclic group $\mathbb{G}$, its order q (with $\|q\| = n$), and a generator g.



Construction 7

- Gen: *run* $\mathcal{G}(1^n)$ *to obtain* $(\mathbb{G}, q, g)$. *A random* $x \leftarrow \mathbb{Z}_q$ *and* $h := g^x$. $pk = \langle \mathbb{G}, q, g, h \rangle$ *and* $sk = \langle \mathbb{G}, q, g, x \rangle$
- Enc: *a random* $y \leftarrow \mathbb{Z}_q$ *and output* $\langle c_1, c_2 \rangle = \langle g^y, h^y \cdot m \rangle$
- Dec: $m := c_2 / c_1^x$

ElGamal加密例子

Encoding binary strings:

- the subgroup of quadratic residues modulo a strong prime $p = (2q + 1)$.
- a string $\hat{m} \in \{0,1\}^{n-1}$, $n = \|q\|$.
- map $\hat{m}$ to the plaintext $m = [(\hat{m}+1)^2 \bmod p]$.
- The mapping is one-to-one and efficiently invertible.

$q = 83$, $p = 2q + 1 = 167$, $g = 2^2 = 4 \pmod{167}$, $\hat{m} = 011101$

The receiver chooses secrete key $37 \in \mathbb{Z}_{83}$.
The public key is $pk = \langle 167, 83, 4, [4^{37} \bmod 167] = 76 \rangle$.
$\hat{m} = 011101 = 29$, $m = [(29+1)^2 \bmod 167] = 65$.
Choose $y = 71$, the ciphertext is
$\langle [4^{71} \bmod 167], [76^{71} \cdot 65 \bmod 167] \rangle = \langle 132, 44 \rangle$.

Decryption: $m = [44 \cdot (132^{37})^{-1}] \equiv [44 \cdot 66] \equiv 65 \pmod{167}$.
65 has the two square roots 30 and 137, and $30 < q$, so $\hat{m} = 29$.

对ElGamal的CCA攻击

Constructing the ciphertext of the message $m \cdot m'$.

Given $pk = \langle g, h \rangle$, $c = \langle c_1, c_2 \rangle$, $c_1 = g^y$, $c_2 = h^y \cdot m$,
Method I: compute $c_2' := c_2 \cdot m'$, and $c' = \langle c_1, c_2' \rangle$.

$$\frac{c_2'}{c_1^x} =?$$

Method II: compute $c_1'' := c_1 \cdot g^{y''}$, and $c_2'' := c_2 \cdot h^{y''} \cdot m'$.

$$c_1'' = g^y \cdot g^{y''} = g^{y+y''} \text{ and } c_2'' =?$$

so $c'' = \langle c_1'', c_2'' \rangle$ is an encryption of $m \cdot m'$.

ElGamal实现问题

- 共享公开参数：$\mathcal{G}$ 产生参数 $\mathbb{G},q,g$。
- 这些参数可以只产生一次并且为所有人所使用（"once-and-for-all"）。
- 可以被多个接收者使用。
- 每个接收者必须选择各自的保密数值 x 并且发布他们自己的公钥包含 $h=g^x$。
- 参数共享：在 Elgamal 的情况下，公开参数可以被共享。在 RSA 情况下，参数可以被共享吗？

椭圆曲线密码学

- 在椭圆曲线群上构造的离散对数问题
- 其他密码学上的应用在1985年被提出
- 类比离散对数，DH密钥交换，ElGamal加密和DSA，在椭圆曲线上有，ECDL，ECDHKE，ElGamal ECC，ECDSA
- 比自然数域上更有效，密钥长度是所需蛮力搜索指数长度的二倍。
- 二倍的原因是，离散对数问题的蛮力搜索所需指数长度是群阶指数长度的一半

椭圆曲线群

- **Elliptic curve group**: points with "addition" operation on a plane algebraic curve in a finite field:

$$y^2 \equiv x^3 + Ax + B \pmod{p}$$

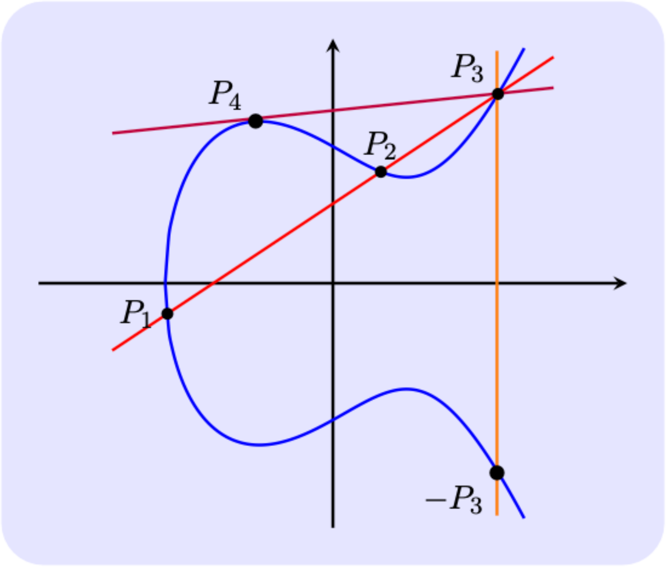
  where $A, B \in \mathbb{Z}_p$ are constants with $4A^3 + 27B^2 \not\equiv 0 \pmod{p}$.
- $\hat{E}(\mathbb{Z}_p)$ is the set of pairs $(x, y) \in \mathbb{Z}_p \times \mathbb{Z}_p$:

$$\hat{E}(\mathbb{Z}_p) \stackrel{\text{def}}{=} \{(x, y) \mid x, y \in \mathbb{Z}_p \wedge y^2 \equiv x^3 + Ax + B \pmod{p}\}$$

- $E(\mathbb{Z}_p) \stackrel{\text{def}}{=} \hat{E}(\mathbb{Z}_p) \cup \{\mathcal{O}\}$, $\mathcal{O}$ is identity, "**point at infinity**".

椭圆曲线上“加法”运算

- 每条直线和曲线有三个交点
  - 一条直线与曲线的切点算2次；垂直线上，无穷远点计做一个点
- 点上的加法：三点成一线，三点之和为无穷远点



Every line intersects the curve in 3 points:

- count twice if tangent.
- count $\mathcal{O}$ at the vertical infinity of y-axis.

"**Addition**" on points:

- $P + \mathcal{O} = \mathcal{O} + P = P$.
- If P_1, P_2, P_3 are co-linear, then $P_1 + P_2 + P_3 = \mathcal{O}$.

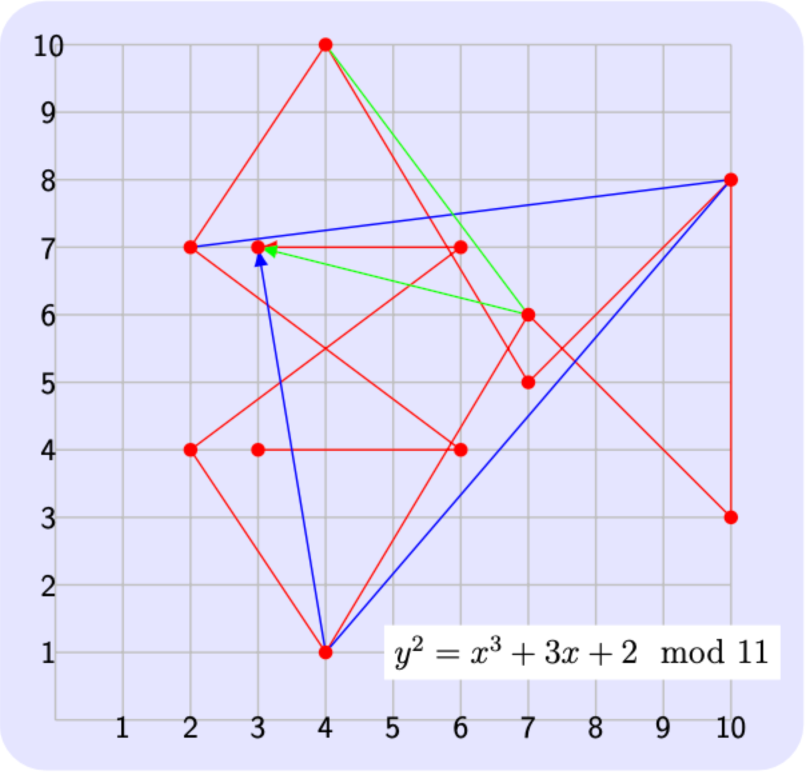
Some equations:

$-P = (x, -y)$, $P_1 + P_2 = -P_3$, $2P_4 = -P_3$, $dP = P + (d-1)P$

Key generation:$sk = (P, d); pk = (P, Q = dP)$

ECC加密例子

- 计算ECDHKE的密钥，这里枚举了生成元为（3，4）的所有指数结果
- Alice的密钥为$a=4$，收到（2，7）
- Alice密钥计算是从（2，7）开始，向后数3个点（乘4=加3次）
- Bob密钥计算是从（4，10）开始（因为$a=4$），向后数2个点（因为$b=3$）



ECC实践

TLS 1.3 (RFC8446) standardizes mandatory-to-implement ECC.

P256 or secp256r1 for DSA and DHKE

- $p := 2^{256} - 2^{224} + 2^{192} + 2^{96} - 1$
- $y^2 = x^3 - 3x + b$, $b :=$ `5ac635d8 aa3a93e7 b3ebbd55 769886bc 651d06b0 cc53b0f6 3bce3c3e 27d2604b`
- It is not clear how b is designed. NOT **twist secure** as the DLP in its twist is not hard. NSA implemented a backdoor into the P256 curve based Dual_EC_DRBG algorithm.

Curve25519 for DHKE

- $p := 2^{255} - 19$
- $y^2 = x^3 + 486662 \cdot x^2 + x$ (Montgomery curve)
- The curve is generated by a point $P = (9, y)$
- It is twist secure and more understandable than P256. And 486662 is a *nothing-up-my-sleeve number*

本节小结

- DHKE，ElGamal加密来自于CDH，DDH问题，后者来自于在指数阶群上的离散对数问题
- 椭圆曲线密码（ECC）更有效并且被广泛使用