第2章 密码学基础

罗文坚

主要内容

公开密钥密码

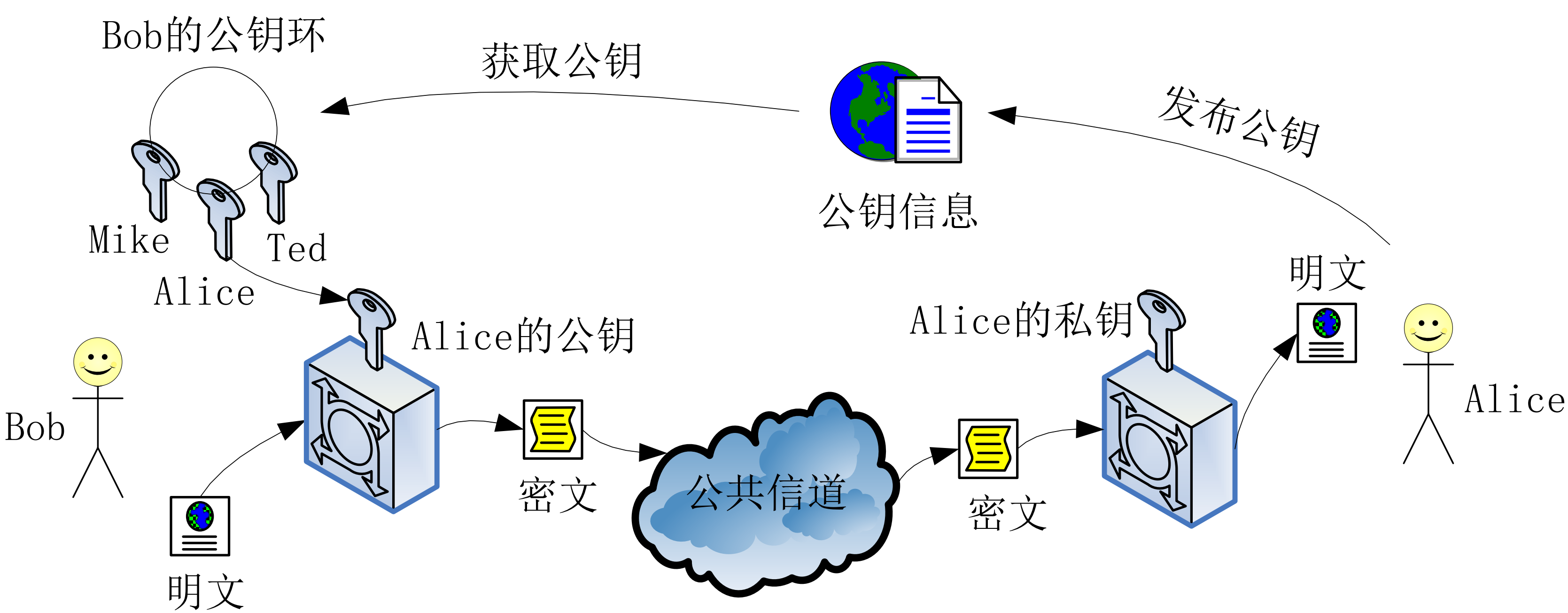
- 公开密钥密码体制是现代密码学最重要的发明，也可以说是密码学发展史上最伟大的革命。

- 两大特点：
  1. 与之前的所有密码不同，其算法不是基于代替和置换，二是基于数学函数。
  2. 与使用一个密钥的传统的对称密钥密码不同，公开密钥密码是非对称的，使用两个独立的密钥。

- 一般认为，密码学就是保护信息传递的机密性，其实这仅仅是现代密码学主题的一个方面。
  - 对信息发送人与接收人的真实身份的验证、事后对所发出或接收信息的不可抵赖性，以及保障数据的完整性是现代密码学主题的另一方面。

公开密钥密码

- 公开密钥密码，又称非对称密钥密码或双密钥密码。
  - 加密密钥和解密密钥为两个独立密钥。
  - 公开密钥密码的通信安全性取决于私钥的保密性。



公开密钥密码的模型

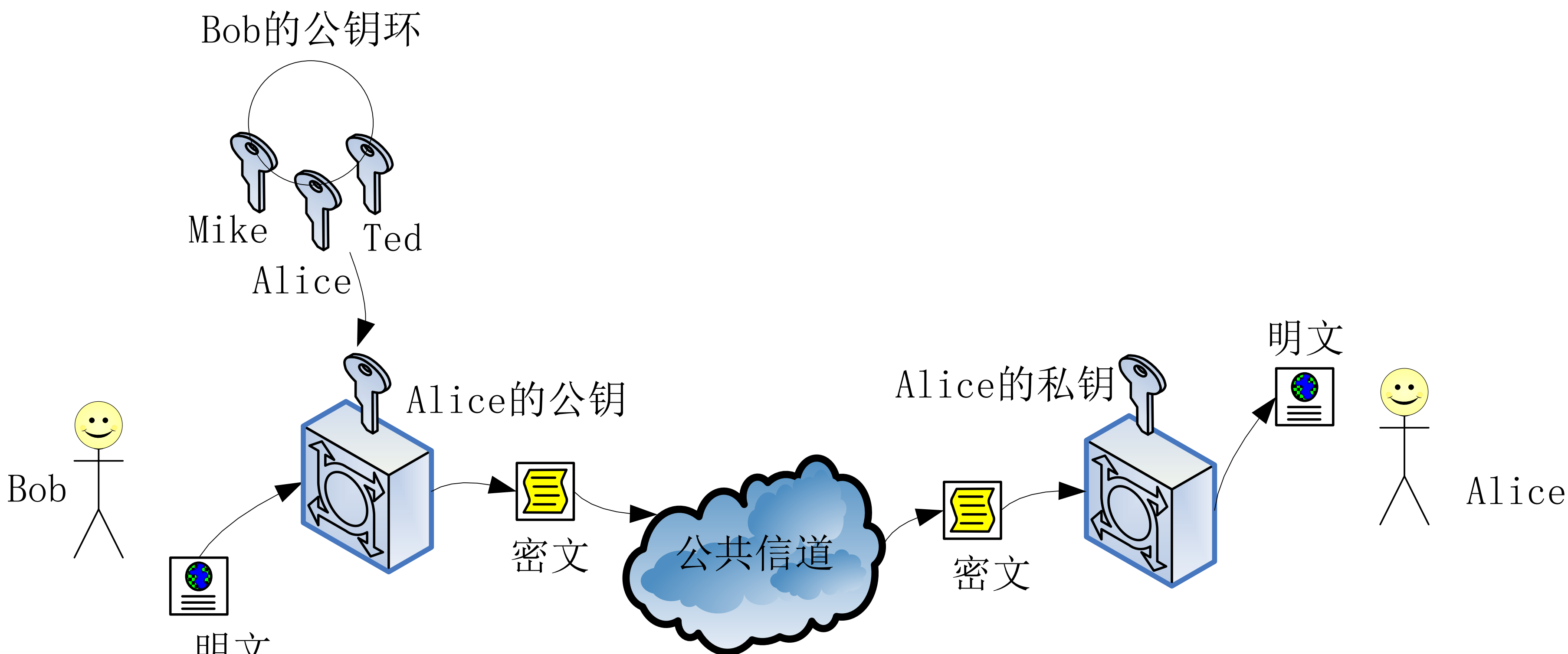
公开密钥理论基础

- 公开密钥密码是1976年由Whitfield Diffie和Martin Hellman在其“密码学新方向”一文中提出的。

- 单向陷门函数f(x)，必须满足以下三个条件。
  ① 给定x，计算y=f(x)是容易的；
  ② 给定y，计算x使y=f(x)是困难的（所谓计算$x=f^{-1}(y)$困难，是指计算上相当复杂已无实际意义）；
  ③ 存在δ，已知δ时，对给定的任何y，若相应的x存在，则计算x使y=f(x)是容易的。

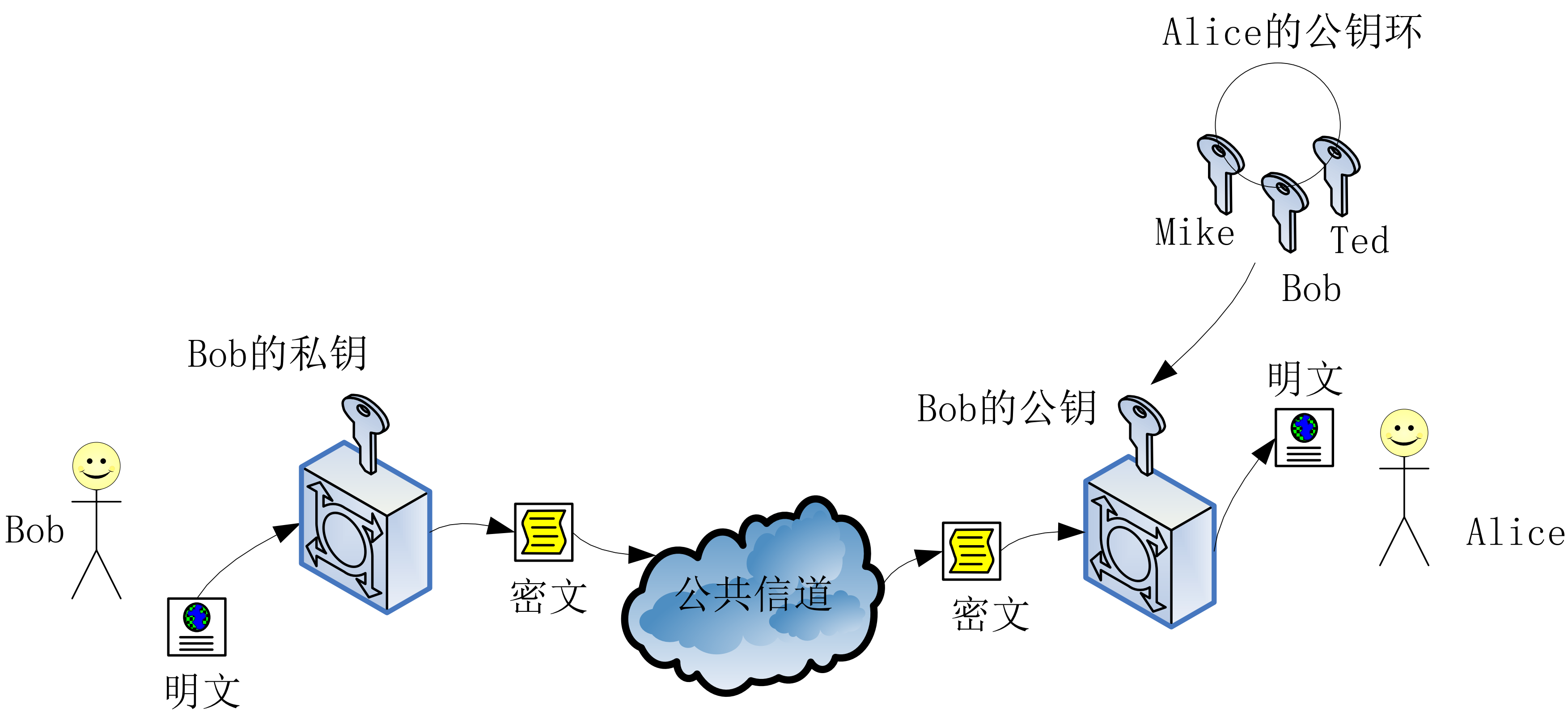
  – 仅满足前2条的称为单向函数；第3条称为陷门性，δ称为陷门信息。

公开密钥的应用：加密模型



公开密钥密码的加密模型

公开密钥的应用：认证模型



公开密钥密码的认证模型

Diffie-Hellman密钥交换算法

- 数学知识：原根
  - 素数p的原根（primitive root）的定义：如果a是素数p的原根，则数a mod p, a^2 mod p, …, a^{p-1} mod p是不同的并且包含从1到p-1之间的所有整数的某种排列。对任意的整数b（mod p ≠ 0），可以找到唯一的幂i，满足b≡a^i mod p，且1 ≤ i ≤ p-1。
  - 注："b≡a mod p"等价于"b mod p = a mod p"，称为"b与a模p同余"。

Diffie-Hellman密钥交换算法

- 数学知识：离散对数
  - 若a是素数p的一个原根，则相对于任意整数b（mod p ≠ 0），必然存在唯一的整数i（1 ≤ i ≤ p-1），使得$b \equiv a^i \bmod p$，i称为b的以a为基数且模p的幂指数，即离散对数。
  - 对于函数$y \equiv g^x \bmod p$，其中，g为素数p的原根，y与x均为正整数，已知g、x、p，计算y是容易的；而已知y、g、p，计算x是困难的，即求解y的离散对数x是困难的。
  - 注：离散对数的求解为数学界公认的困难问题。

Diffie-Hellman密钥交换算法

- Alice和Bob协商好一个大素数p，和一个大的整数g，$1<g<p$，g是p的原根。p和g无须保密，可为网络上的所有用户共享。

- 当Alice和Bob要进行保密通信时，他们可以按如下步骤来做：
  ① Alice选取大的随机数$x<p$，并计算$Y=g^{x}(\bmod P)$；
  ② Bob选取大的随机数$x'<p$，并计算$Y'=g^{x'}(\bmod P)$；
  ③ Alice将Y传送给Bob，Bob将Y′传送给Alice；
  ④ Alice计算$K=(Y')^{X}(\bmod P)$，Bob计算$K'=(Y)^{X'}(\bmod P)$

- 显而易见，$K=K'=g^{xx'}(\bmod P)$，即Alice和Bob已获得了相同的秘密值K。

RSA公开密钥算法

- 欧拉函数：对于一个正整数n，由小于n且和n互素的正整数构成的集合为Z_n，这个集合被称为n的完全余数集合。Z_n包含的元素个数记做φ(n)，称为欧拉函数。
  - φ(1)被定义为1，但是并没有任何实质的意义。
  - 如果两个素数p和q，且n = p×q，则φ(n) =(p-1)(q-1)。
  - 欧拉函数是欧拉定理的核心概念。
- 欧拉定理的具体表述：
  - 正整数a与n互素，则$a^{\varphi(n)} \equiv 1 \bmod n$。

- 推论：给定两个素数p和q，以及两个整数m、n，使得n = p×q，且0<m<n，对于任意整数k下列关系成立：

$$m^{k\,\varphi(n)+1} = m^{k(p-1)(q-1)+1} = m^* (m^{(p-1)})^{k(q-1)} \equiv m \bmod n。$$

大整数因子分解

- 大整数因子分解问题：
  - 已知p、q为两个大素数，则求N=p×q是容易的，只需要一次乘法运算；但已知N是两个大素数的乘积，要求将N分解，则在计算上是困难的，其运行时间复杂程度接近于不可行。
- 算法时间复杂性：
  - 输入规模为n时，若算法运行时间复杂度为O(n)，称此算法为线性的；运行时间复杂度为$O(n^k)$，其中k为常量，称此算法为多项式的；若有某常量t和多项式h(n)，使算法的运行时间复杂度为$O(t^{h(n)})$，则称此算法为指数的。
- 一般说来：
  - 在线性时间和多项式时间内被认为是可解决的，比多项式时间更坏的，尤其是指数时间被认为是不可解决的。
  - 注：如果输入规模太小，即使很复杂的算法也会变得可行的。

RSA密码算法

- **RSA密码体制：**
  - 明文和密文均是0到n之间的整数，n通常为1024位二进制数或309位十进制数。
  - 明文空间M=密文空间C={x∈Z|0<x<n, Z为整数集合}。

- **RSA密码的密钥生成具体步骤如下：**

  ① 选择互异的素数p和q，计算n=pq，$\varphi(n) = (p - 1)(q - 1)$；

  ② 选择整数e，使$\gcd(\varphi(n), e) = 1$，且$1 < e < \varphi(n)$；

  ③ 计算d，$d \equiv e^{-1} \bmod \varphi(n)$，即d为模φ(n)下e的乘法逆元；

- 公钥Pk = { e, n }，私钥Sk = { d, n, p, q }
- 加密：$c = m^e \bmod n$；解密：$m = c^d \bmod n$。

RSA举例

- 选定p=101，q＝113，则n=11413，φ(n)=100×112＝11200。
- 选定e = 3533，可求得$d \equiv e^{-1} \bmod 11200 \equiv 6597 \bmod 11200$，d = 6597。
- 公开n=11413和e=3533。
- 若明文为9726：
  - 计算$9726^{3533} \bmod 11413 = 5761$，发送密文5761。
- 收到密文5761时：
  - 用d＝6597进行解密，计算5761^{6597}（mod 11413)=9726。

RSA的安全性

- RSA的安全性是基于单向函数$e_k(x)=x^e(\mathrm{mod}\ n)$，求逆计算不可行。
- 解密的关键是了解陷门信息，即能够分解n=pq，知道$\varphi(n)=(p-1)(q-1)$，从而解出解密私钥d。
- 如果要求RSA是安全的，p与q必为足够大的素数，使分析者没有办法在多项式时间内将n分解出来。
- RSA开发人员建议，p和q的选择应该大约是100位的十进制素数，模n的长度要求至少是512bit。
  - 国际数字签名标准ISO/IEC 9796中规定n的长度为512bit。

RSA的安全性

- 为了抵抗现有的整数分解算法，对RSA模n的素因子p和q还有如下要求：
  ① $|p-q|$很大，通常p和q的长度相同。
  ② $p-1$和$q-1$分别含有大素因子p_1和q_1。
  ③ p_1-1和q_1-1分别含有大素因子p_2和q_2。
  ④ $p+1$和$q+1$分别含有大素因子p_3和q_3。

RSA的安全性

- 为了提高加密速度，通常取e为特定的小整数。

- 例如，EDI（Electronic Data Interchange）国际标准中规定$e=2^{16}+1$，ISO/IEC 9796甚至允许取$e=3$。这时，加密速度一般比解密速度快10倍以上。

- 模n的求幂运算的效率问题：
  - 著名的“平方-和-乘法”方法将计算$x^c \bmod n$的模乘法的次数缩小到至多为2*l*，*l*是指数c二进制表示的位数。
  - 若n以二进制形式表示有*k*位，$l \le k$，则$x^c \bmod n$能在$O(k^3)$时间内完成。

其他公开密钥密码简介

- 基于大整数因子分解问题:
  - RSA密码、Rabin密码
- 基于有限域上的离散对数问题:
  - Diffie-Hellman公钥交换体制、ElGamal密码
- 基于椭圆曲线上的离散对数问题:
  - Diffie-Hellman公钥交换体制、ElGamal密码。

其他公开密钥密码简介

- Rabin密码算法是M. Rabin设计的，是RSA密码算法的一种改进。
  - RSA是基于大整数因子分解问题，Rabin则是基于求合数的模平方根的难题。

- Elgamal算法是Taher Elgamal发明的，既能用于数据加密，也能用于数字签名，其安全性依赖于计算有限域上离散对数这一难题，其不足之处是它的密文成倍扩张。

- 大整数分解算法的发展，计算机速度的提高和网络的发展，RSA的密钥长度需要不断增加。
  - 但是，密钥长度的增加，导致了加密、解密的速度大为降低。
  - 需要新的算法来代替RSA！

其他公开密钥密码简介

- **1985年，Koblitz和Miller**分别独立提出将椭圆曲线用于密码算法，其根据是椭圆曲线上的离散对数问题（**ECDLP，Elliptic Curve Discrete Logarithm Problem**）。

- 椭圆曲线密码体制（**ECC，Elliptic Curve Cryptosystems**）相比于**RSA**的优势：
  - 相同的密钥长度，**ECC**抗攻击性比**RSA**强很多倍。
  - 计算量小，处理速度快。
  - 存储空间小。**ECC**的密钥尺寸和系统参数比**RSA**要小得多。
  - 带宽要求低。对于短消息加密，**ECC**带宽要求比**RSA**低得多。带宽要求低使得**ECC**在无线网络领域具有广泛的应用前景。

- **ECC**是新一代安全电子交易**(SET)**协议中缺省的公钥密码算法。

补充内容

不经意传输协议

百万富翁问题

不经意传输协议

- 例如，Alice是机密的出售者，Alice列举了很多问题，意欲出售各个问题的答案；Bob想买其中一个问题的答案，但又不想让Alice知道自己买的是哪个问题的答案。

- 2选一不经意传输协议
  - 1985年，Shimon Even, Oded Goldreich和Abraham Lempel 提出。
  - 模型：Alice有两条信息（m1、m2），Bob提供一个输入，并根据输入获得其中一个信息。在协议结束后，Bob得到了自己想要的那条信息（m1或者m2），而Alice并不知道Bob最终得到的是哪条。

2选一不经意传输协议

- 假定：
  - Alice有两对公开密钥密码，PK1和SK1，以及PK2和SK2。
  - Bob想得到秘密m1。

1. Bob产生一个随机数x，用PK1加密，得到密文c, 发给Alice。
2. Alice用SK1和SK2分别解密c，得到x1和x2。
   - 其中，x和x1相等，x2是无意义的随机数。
3. Alice将$x1 \oplus m1$和$x2 \oplus m2$ 发给Bob。
4. Bob通过$x \oplus x1 \oplus m1$得到m1，另一个$x \oplus x2 \oplus m2$是随机数。

百万富翁问题

- 假设现有两个百万富翁Alice和Bob，他们各有钱a和b。他们想知道谁的钱多，但是又不想把各自的钱的具体数量告诉对方，怎么样才能比大小呢？

- 考虑简化的问题：假如a和b是1到10之间的数。

百万富翁问题

1. 首先Bob挑选一个非常大的整数x，然后用Alice的公钥加密，得到 k=Enc（x），然后把k-b+1发给Alice。

2. Alice计算以下这些数：

   Dec{k-b+1, k-b+2, … , k-b+b, …, k-b+10}

3. 然后，除以一个素数p取余得到：

   z1 = Dec{k-b+1}(mod p), z2, z3, ... z10

4. 因为Alice的钱是a，将下列数字发给Bob：

   z1, z2, z3, ... Za, Z(a+1) = Z(a+1)+1, ...

5. Bob观察第b个数字。如果第b个数字等于x（mod p）就说明a>=b，否则说明a<b；然后，bob把结果返给Alice。

作业

- 课后阅读：欧拉定理相关材料，理解欧拉定理的证明过程。

1. 习题3（2）。Alice和Bob使用Diffie_Hellman协议协商共享密钥，得知使用的素数q=13，原根a=2。如果Alice传递给Bob YA=12，则Alice的随机数XA是多少？如果Bob传递Alice YB=6，则共享的密钥K是多少？
2. 习题3（3）。如果攻击者截获了Alice发给Bob的消息C为10，并得知加密密码是RSA（公钥：e=5，n=35），那么明文M是什么？