《现代密码学》第三讲

密码学基础

上讲内容回顾

- 代换密码
- 置换密码
- Hill密码
- 转轮密码
- 代换密码的惟密文攻击

本章主要内容

- Shannon的通信保密系统
- 熵和无条件保密

- 复杂度理论基础概念
- 计算安全性

本章主要内容

- Shannon的通信保密系统
- 熵和无条件保密
- 复杂度理论基础概念
- 计算安全性

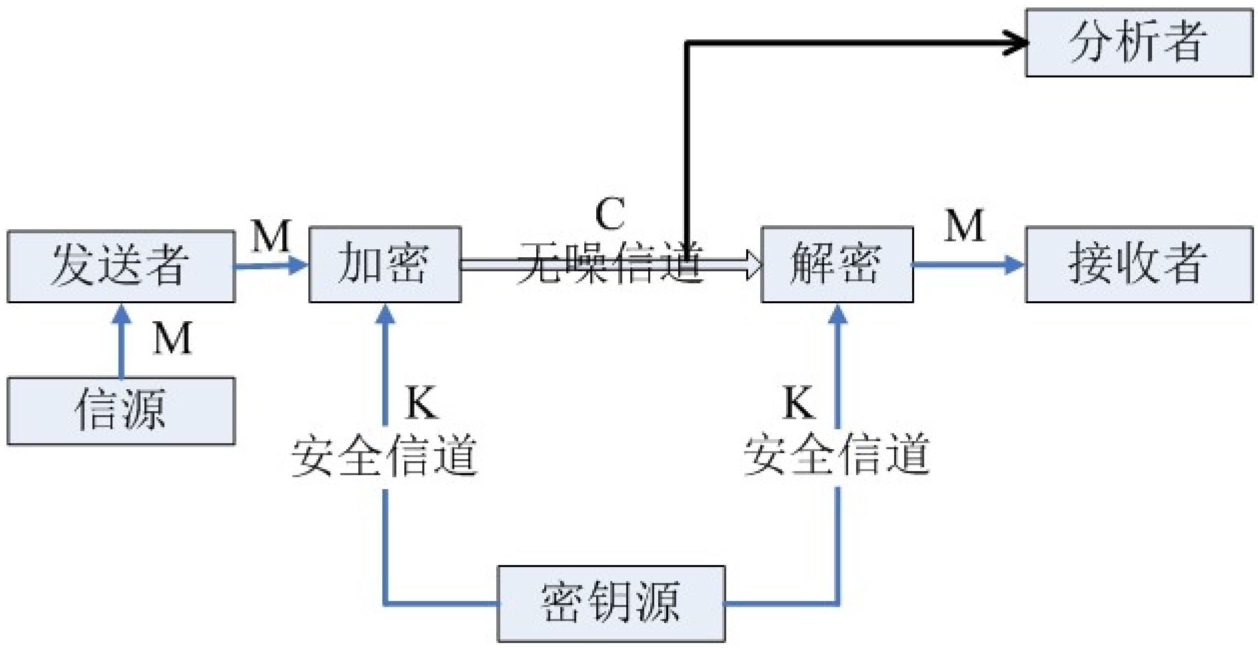
Shannon通信保密系统

C. E. Shannon（香农）----信息论之父

- 1948, A mathematical theory of communication, 奠定了现代信息论的基础.
- 1949, Communication theory of secrecy systems, 定义了保密系统的数学模型, 将密码学由艺术转化为一门科学.

Shannon通信保密系统

Shannon的保密通信系统模型：



Shannon通信保密系统

一个**密码体制**是一个五元组：

$$(\mathcal{P}, \mathcal{C}, \mathcal{K}, \mathrm{E}, \mathrm{D})$$

其中，

$\mathcal{P}$ --明文空间

$\mathcal{C}$ --密文空间

$\mathcal{K}$ --密钥空间

E --加密变换

D --解密变换

Shannon通信保密系统

- 一个加密变换是一个下列形式的映射：

$$E: \mathcal{P} \times \mathcal{K} \rightarrow \mathcal{C}$$

一般对于给定的$k \in \mathcal{K}$, 把E(*, k)记为E_k;

- 一个解密变换是一个与加密E变换相对应的映射：

$$D: \mathcal{C} \times \mathcal{K} \rightarrow \mathcal{P}$$

对于给定的$k \in \mathcal{K}$, 也把D(*, k)记为D_k.

Shannon通信保密系统

重要原则：

一个定义在空间（$\mathcal{P}, \mathcal{C}, \mathcal{K}$)上的密码算法，E和D是一对有效的算法

$$E: \mathcal{P} \times \mathcal{K} \rightarrow \mathcal{C}, \quad D: \mathcal{C} \times \mathcal{K} \rightarrow \mathcal{P}$$

$\forall k \in \mathcal{K}, \ \forall m \in \mathcal{P}$，下式成立

$$\mathcal{D}_k(\mathcal{E}_k(m)) = m.$$

本章主要内容

- Shannon的通信保密系统
- 熵和无条件保密
- 复杂度理论基础概念
- 计算安全性

熵和无条件保密

北京邮电大学 BEIJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS

定义： 设随机变量X={x_i | i=1, 2, ..., n}， x_i出现的概率为Pr（x_i） ≧0， 且$\sum_{i=1}^{n} \Pr(x_i) = 1$ ， 则X的不确定性或熵定义为

$$H(X) = -\sum_{i} p(x_i)\log_2 p(x_i) \quad \geq 0$$

$$H(X) \geqq 0;$$

例1. 若集X为均匀分布时， 即$p(x_i)$=1/n, n≧i ≧1, 则$H(X)$=?

例2. 当X为确定性的事件时, 即X概率分布为 Pr(X=a)=1, Pr(X=a)=1， 则$H(X)$ =?

A： 0 　　B： $\log_2 n$

信息安全中心 BUPT ISC

熵和无条件保密

定义：

设 $X=\{x_i \mid i=1, 2, \cdots, n\}$， x_i出现的概率为 $p(x_i) \geqslant 0$ 且 $\sum_{i=1,\cdots,n} p(x_i)=1$；$Y=\{y_i \mid i=1, 2, \cdots, m\}$， y_i出现的概率为 $p(y_i) \geqslant 0$，且 $\sum_{i=1,\cdots,m} p(y_i)=1$；则集X相对于集Y的条件熵定义为

$$H(X|Y)=\sum_{j=1}^{m} p(y_i)\, H(X|y_j)=-\sum_{j=1}^{m}\sum_{i=1}^{n} p(y_i)p(x_i|y_j)\log_2 p(x_i|y_j)$$

通常将条件熵 $H(X|Y)$ 称作含糊度.

熵和无条件保密

若将X视为一个系统的输入空间，Y视为系统的输出空间，在通信中，X和Y之间的平均互信息定义为：

$$I(X,Y)=H(X)-H(X|Y)$$

它表示X熵减少量。

熵和无条件保密

定义：密码系统（$\mathcal{P}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D}$）是完善保密（无条件保密）的充分必要条件是

$$H(\mathrm{P}\,|\,\mathrm{C}) = H(\mathrm{P}) \text{ 或 } I(P,C) = 0.$$

假设攻击者有无限计算资源，仍然不能从密文得到*明文*任何信息.

熵和无条件保密

一次一密系统： 设n是大于等于1的正整数，

P=C=K=$\{0, 1\}^n$, 对于密钥$K \in$ K, ，$K=\{k_1, k_2, \cdots, k_n\}$.

设明文P=$\{p1, p2, \cdots, pn\}$, 密文C=$\{c1, c2, \cdots, cn\}$.

加密： $\mathcal{E}_K(P)=(p_1 \oplus k_1, p_2 \oplus k_2, \ldots, p_n \oplus k_n)$,

解密： $\mathcal{D}_K(C)=(c_1 \oplus k_1, c_2 \oplus k_2, \ldots, c_n \oplus k_n)$.

一次一密算法由Gilbert Vernam于1917年用于报文消息的自动加密和解密，30年后由Shannon证明它不可攻破.

本章主要内容

- Shannon的通信保密系统
- 熵和无条件保密
- 复杂度理论基础概念
- 计算安全性

复杂度理论基础概念

问题的定义及分类

1 设A=$(a_1, a_2, \cdots, a_n)$是由n个不同的正整数构成的n元组，S是另一已知的正整数．A称为背包向量，S称为背包容积．求A中元素集合A使 $\sum_{a_i \in A'} a_i = S$ ．

2 设背包向量A=(1, 2, 5, 10, 20, 50, 100)，背包容积为177，求向量 $X \in \{0,1\}^7$ ，使得 $\sum_{i=1}^{7} x_i a_i = 177$

复杂度理论基础概念

3 已知整数N，问N是否是一个素数？

4 试问77是否是素数？

5 试问79是否是素数？

6 已知整数N，求N的素分解式.

7 已知整数177，求其素分解式.

复杂度理论基础概念

北京郵電大學
BEIJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS

问题的定义及分类

- 问题：描述参量，陈述解答应当满足的性质（称为询问）.

  参量为具体数值时，称为问题的一个实例.
- 判定问题：回答只有Yes或No.
- 计算问题：从其可行解集合中搜索出最优解.

复杂度理论基础概念

算法复杂度的定义

- 算法（Algorithm）是指解题方案的准确而完整的描述，当其运行时能从一个初始状态和（可能为空的）初始输入开始，经过一系列有限而清晰定义的状态，最终产生输出并停止于一个终态。

A: input -- output

- 确定算法：A$(u)=v$
- 概率算法：A$(u,r)=v$，r是集合$\{0,1\}^n$上的均匀随机变量

复杂度理论基础概念

北京邮電大學 BEIJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS

算法复杂度的定义

例 问x(假设x是小于100的整数)是否是素数?

解答一:

取**2**~$\lfloor \sqrt{x} \rfloor$的所有整数，依次试除***x***，若存在某个整数可以整除***x***，则程序停止，输出***x***为合数，否则输出***x***为素数**.**

最坏试除次数:$\lfloor \sqrt{x} \rfloor$ 存储空间：**0**

解答二：预计算小于**100**的素数存储在寄存器中；然后将***X***与存储器中的元素比较，若存在某个素数等于***x***，则程序停止，输出***x***为素数，否则输出**x**为合数**.**

最坏比较次数**: 100/ln100**, 存储空间**: *100*/ln*100***

复杂度理论基础概念

北京郵電大學 BEIJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS

算法复杂度的定义

- **时间（计算）复杂性：** 考虑算法的主要操作步骤，计算执行中所需的总操作次数.
- **空间复杂性：** 执行过程中所需存储器的单元数目.
- **数据复杂性：** 信息资源.

计算模型——确定性图灵机（有限带符号集合，有限状态集，转换函数）（读写头，读写带）.

复杂度理论基础概念

算法复杂度的定义

$$运行时间 = \frac{基本操作数量}{机器速度}$$

复杂度理论基础概念

算法复杂度的定义

定义 假设一个算法的计算复杂度为$O(n^t)$，其中t为常数，n为输入问题的长度，则称这算法的复杂度是多项式的。具有多项式时间复杂度的算法为多项式时间算法.

函数$g(n)=O(n^t)$表示存在常数$c>0$和$n_0>=0$，对一切$n> n_0$均有$|g(n)|<=c|n^t|$ 成立，也就是说，当n足够大时，$g(n)$存在上界.

定义 非多项式时间算法：算法的计算复杂性写不成$O(P(n))$形式，其中$P(n)$表示n的多项式函数.

复杂度理论基础概念

算法复杂度的定义

例 设x是小于n的某个整数，问x是否是素数？

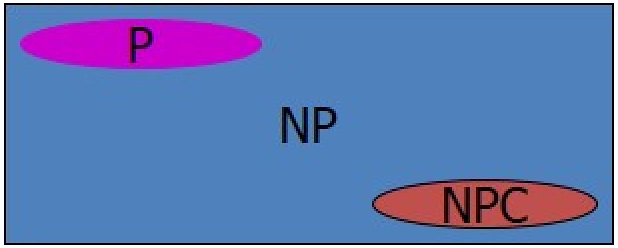
解法1是否是多项式时间算法？

解法2是否是多项式时间算法？

复杂度理论基础概念

P问题和NP问题

- 定义(P问题)如果一个判定问题存在解它的多项式时间的算法，则称该问题属于P类.
- 定义(NP问题) 如果一个判定问题不存在解它的多项式时间的算法，且对于一个解答可以在多项式时间验证其是否正确，则称该问题属于NP类.
- 公开问题：P≠ NP?

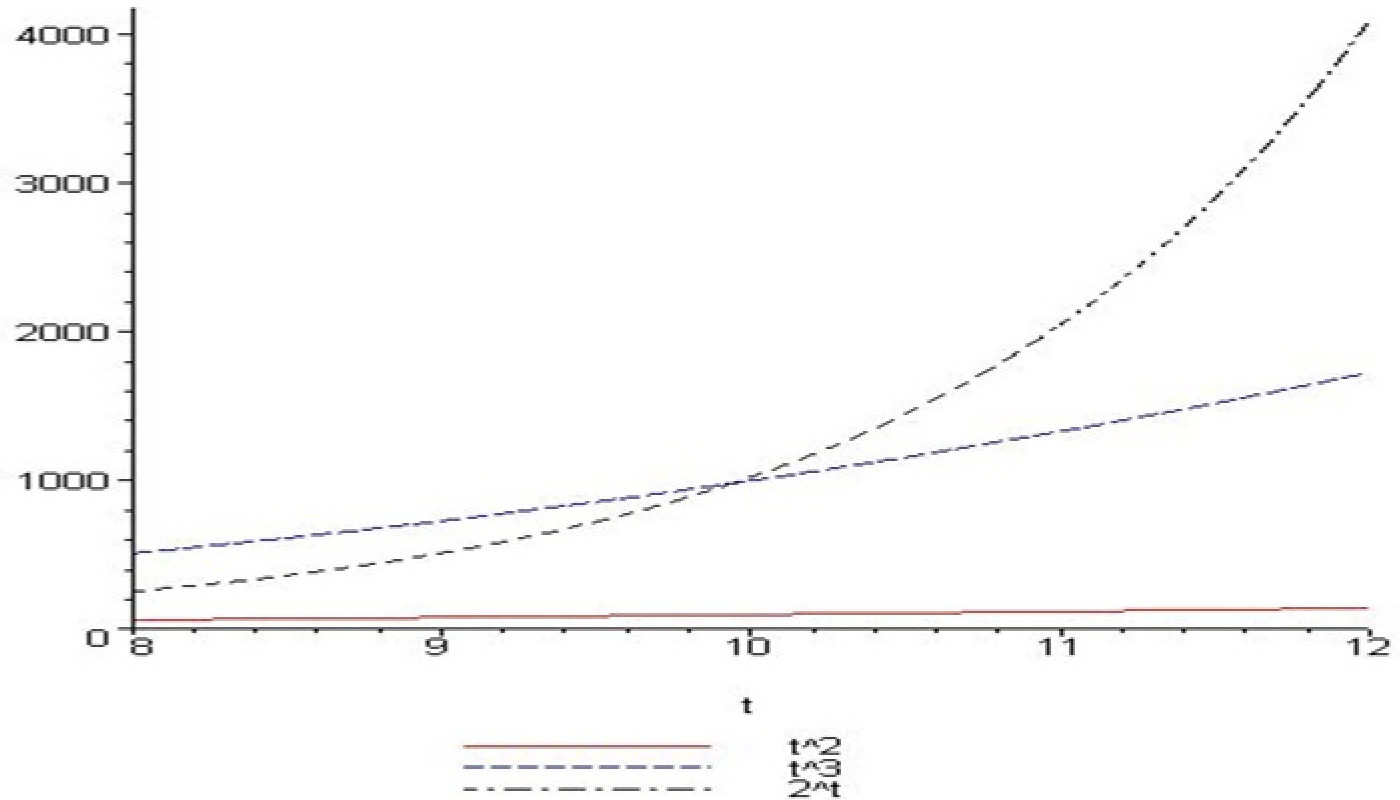


本章主要内容

- Shannon的通信保密系统
- 熵和无条件保密
- 复杂度理论基础概念
- 计算安全性

计算安全性

二次函数、三次函数、2^x函数的示意图



信息安全中心

计算安全性

例：设问题输入长度为n，在一个每秒钟运行百万次的计算机上的运行时间如下：

	10	30	50	60
$T(n)=n^2$	0.0001s	0.0009s	0.0025s	0.0036s
$T(n)=2^n$	0.001s	17.9月	35.7年	366世纪

计算安全性

当问题输入长度足够大，分析密码体制的算法的复杂度较大，可能的计算能力下，在保密的期间内可以保证算法不被攻破，这就是密码体制的计算安全性思想。

计算安全性

- 实际安全是指密码系统满足以下准则之一：
  - 破解该密码系统的成本超过被加密信息本身的价值；
  - 破译该密码系统的时间超过被加密信息的有效生命周期。

本章内容小结

- Shannon的通信保密系统
- 熵和无条件保密
- 复杂度理论基础概念
- 计算安全性

THE END！