密码学原理与实践

（第三版）

第二章 Shannon理论

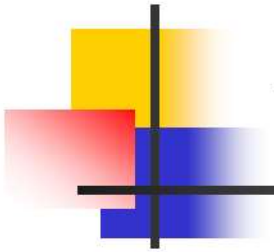
苏 明

[加] Douglas R. Stinson 著

冯登国 等译

概览

- 2．1 引言
- 2．2 概率论基础
- 2．3 完善保密性
- 2．4 熵
  - 2．4．1 Huffman编码
- 2．5 熵的性质
- 2．6 伪密钥和唯一解距离
- 2．7 乘积密码体制



引言

- 计算安全性(computational security)

--计算复杂度，讨论**攻击可行性**

- 可证明安全性(provable security)

--**规约**为已知的困难问题(NP完全)

- 无条件安全性(unconditional security)

--即使提供无限计算资源，无法获取**优势**

概率论基础

定义 2.1 一个离散的随机变量，比方说 $\mathbf{X}$ ，由有限集合 X 和定义在 X 上的概率分布组成。我们用 $\Pr[\mathbf{X}=x]$ 表示随机变量 $\mathbf{X}$ 取 x 时的概率。如果随机变量是固定的，我们有时缩写成 $\Pr[x]$。对任意的 $x\in X$，有 $0\leqslant \Pr[x]\leqslant 1$，并且

$$\sum_{x\in X}\Pr[x]=1$$

定义 2.2 假设 $\mathbf{X}$ 和 $\mathbf{Y}$ 是分别定义在有限集合 X 和 Y 上的随机变量。联合概率 $\Pr[x, y]$ 是 X 取 x 并且 Y 取 y 的概率。条件概率 $\Pr[x\mid y]$ 表示 $\mathbf{Y}$ 取 y 时 $\mathbf{X}$ 取 x 的概率。如果对任意的 $x\in X$ 和 $y\in Y$，都有 $\Pr[x, y]=\Pr[x]\Pr[y]$，则称随机变量 $\mathbf{X}$ 和 $\mathbf{Y}$ 是统计独立的。

概率论基础

$$\Pr[x, y] = \Pr[x \mid y]\Pr[y]$$

$$\Pr[x, y] = \Pr[y \mid x]\Pr[x]$$

定理 2.1(Bayes 定理) 如果 $\Pr[y] > 0$，那么

$$\Pr[x \mid y] = \frac{\Pr[x]\Pr[y \mid x]}{\Pr[y]}$$

完善保密性

- 完善保密性就是攻击者不能从观察密文中获取明文的任何信息

定义 2.3 一个密码体制具有完善保密性，如果对于任意的 $x \in \mathcal{P}$ 和 $y \in \mathcal{C}$，都有 $\Pr[x \mid y] = \Pr[x]$。也就是说，给定密文 y，明文 x 的后验概率等于明文 x 的先验概率。

完善保密性

利用全概率公式、Bayes公式进行推导

- $C(K)$代表密钥是K时所有可能的密文 $C(K)=\{e_K(x):x\in\mathcal{P}\}$

$$\Pr[\mathbf{y}=y]=\sum_{\{K:y\in C(K)\}}\Pr[\mathbf{K}=K]\Pr[\mathbf{x}=d_K(y)]$$

$$\Pr[\mathbf{y}=y\mid\mathbf{x}=x]=\sum_{\{K:x=d_K(y)\}}\Pr[\mathbf{K}=K]$$

$$\Pr[\mathbf{x}=x\mid\mathbf{y}=y]=\frac{\Pr[\mathbf{x}=x]\times\sum_{\{K:x=d_K(y)\}}\Pr[\mathbf{K}=K]}{\sum_{\{K:y\in C(K)\}}\Pr[\mathbf{K}=K]\Pr[\mathbf{x}=d_K(y)]}$$

完善保密性

定理 2.3 假设移位密码的26个密钥都是以相同的概率 1/26 使用的，则对于任意的明文概率分布，移位密码具有完善保密性。

证明：

$$\Pr[\mathbf{y}=y]=\sum_{K\in\mathbf{Z}_{26}}\Pr[\mathbf{K}=K]\Pr[\mathbf{x}=d_K(y)]$$

$$=\sum_{K\in\mathbf{Z}_{26}}\frac{1}{26}\Pr[\mathbf{x}=y-K]$$

$$=\frac{1}{26}\sum_{K\in\mathbf{Z}_{26}}\Pr[\mathbf{x}=y-K] \quad =\frac{1}{26}$$

$$\Pr[y\mid x]=\Pr[\mathbf{K}=(y-x)\bmod 26]$$

$$=\frac{1}{26}$$

$$\Pr[x\mid y]=\frac{\Pr[x]\Pr[y\mid x]}{\Pr[y]}=\Pr[x]$$

完善保密性

定理 2.4 假设密码体制 $(\mathcal{P}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D})$ 满足 $|\mathcal{K}| = |\mathcal{C}| = |\mathcal{P}|$。该密码体制是完善保密的，当且仅当每个密钥被使用的概率都是 $1/|\mathcal{K}|$，并且对于任意的 $x \in \mathcal{P}$ 和 $y \in \mathcal{C}$，存在唯一的密钥 K 使得 $e_K(x) = y$。

证明： **左边到右边**（充分性）：

对于 $x \in \mathcal{P}$ 和 $y \in \mathcal{C}$，刚好存在一个密钥 K 使得 $e_K(x) = y$。

使用 Bayes 定理，我们有

$$\Pr[x_i \mid y] = \frac{\Pr[y \mid x_i]\Pr[x_i]}{\Pr[y]}$$

考虑完善保密的条件 $\Pr[x_i \mid y] = \Pr[x_i]$。从这里，我们有 $\Pr[K_i] = \Pr[y]$，$1 \leqslant i \leqslant n$。

$$= \frac{\Pr[\mathbf{K} = K_i]\Pr[x_i]}{\Pr[y]}$$

对任意的 $K \in \mathcal{K}$，$\Pr[K] = 1/|\mathcal{K}|$。

右边到左边（必要性）：类似前面证明。

完善保密性

- 一次一密（Gilbert Vernam, 1917）

比如流加密系统

密码体制 2.1 一次一密

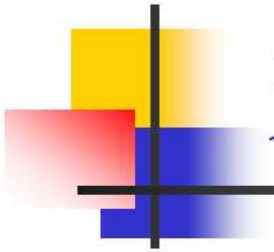
假设 $n \geq 1$ 是正整数，$\mathcal{P} = \mathcal{C} = \mathcal{K} = (\mathbb{Z}_2)^n$。对于 $K \in (\mathbb{Z}_2)^n$，定义 $e_K(x)$ 为 K 和 x 的模2的向量和（或者说是两个相关比特串的异或）。因此，如果 $x = (x_1, x_2, \cdots, x_n)$ 并且 $K = (K_1, K_2, \cdots, K_n)$，则

$$e_K(x) = (x_1 + K_1, \cdots, x_n + K_n) \bmod 2$$

解密与加密是一样的。如果 $y = (y_1, \cdots, y_n)$，则

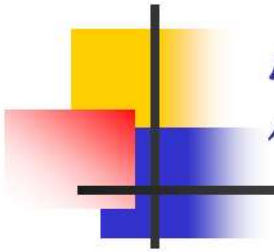
$$d_K(y) = (y_1 + K_1, \cdots, y_n + K_n) \bmod 2$$

大家可以自行验证 $|\mathcal{K}| = |\mathcal{C}| = |\mathcal{P}|$



完善保密性

- 一次一密：每个密钥仅用一次
- 密钥管理的挑战: 密钥（通信）开销巨大
- 折衷的解决办法：
  - ✓ PRSG(伪随机序列发生器)
  - ✓ 重用密钥：如分组加密



熵

- **熵：信息的度量，通过概率分布函数来计算**

- 直观的说起来，不确定性的度量（含有的信息量）

信息度量

- 我们可以知道：

  普遍公认的事实含有的信息量少；

  但可能性小的事件一旦发生，含有的信息量多

信息度量

- 启发我们两点：

若定义事件m的信息度量为I(m), 那么

1. I(m)= f(p(m));
2. 而且

$p(m_1)<p(m_2)$, 那么$I(m_1)>I(m_2)$;

信息度量

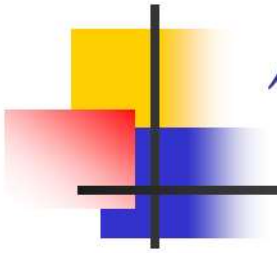
- 另外，从我们引入的度量考虑，应该有

  3. I(m)>=0;

- 此外还应有可加性；组合事件具有的信息量应该是其中不相干(独立)的成分事件信息量之和。

比如事件A,B独立, 组合事件C由A,B构成，那么I(C)=I(A)+I(B), i.e.,

4. f(p(C))=f(P(A)p(B))=f(p(A))+f(p(B)).



信息度量

综上可以推导出：

I(m)=f(p(m))=-log p(m).

条件自信息

- 定义在给定$Y=y_j$的情况下，事件$X=x_i$的的条件自信息(conditional self information)为

$$I(x_i \mid y_j) = \log\left(\frac{1}{P(x_i \mid y_j)}\right)$$

因此，$I(x_i ; y_j) = I(x_i) - I(x_i \mid y_j)$.

其中互信息 $I(x_i; y_j) = \log\left(\frac{P(x_i \mid y_j)}{P(x_i)}\right)$

平均互信息和熵

- **随机变量X和Y的平均互信息(average mutual information)**为

$$I(X;\ Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} P(x_i, y_j) I(x_i; y_j) = \sum_{i=1}^{n}\sum_{j=1}^{m} P(x_i, y_j) \log \frac{P(x_i, y_j)}{P(x_i)P(y_j)}$$

$I(X;\ Y) \geq 0$, with equality if and only if X and Y are statistically independent.

- **随机变量X 的平均自信息(average self information or entropy)**定义为

$$H(X) = \sum_{i=1}^{n} P(x_i) I(x_i) = -\sum_{i=1}^{n} P(x_i) \log P(x_i)\ .$$

熵

定义 2.4 假设随机变量 **X** 在有限集合 X 上取值，则随机变量 **X** 的熵定义为

$$H(\mathbf{X}) = -\sum_{x \in X} \Pr[x] \mathrm{lb} \Pr[x]$$

假设X是一个定义在n个信源符号上的随机变量。何时熵H(X)取到最大值？

熵的性质

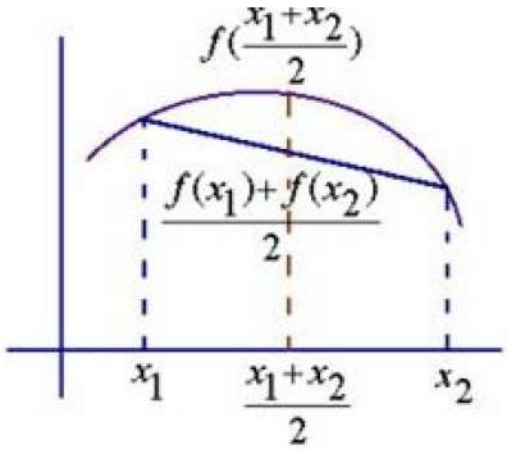
- Preliminaries

定义 2.5 一个区间 I 上的实值函数 f 为凸函数，如果对任意的 $x, y \in I$ 满足

$$f\left(\frac{x+y}{2}\right) \geqslant \frac{f(x)+f(y)}{2}$$

称 f 是区间 I 上的严格凸函数，如果对任意的 $x, y \in I$，$x \neq y$ 满足

$$f\left(\frac{x+y}{2}\right) > \frac{f(x)+f(y)}{2}$$



熵的性质

- 如何判断凸函数？
- 一阶导数严格递减
- $\log_2(x)$ 在区间 $(0,+\infty)$是凸函数吗？

熵的性质

定理 2.5（Jensen 不等式） 假设 f 是区间 I 上的连续的严格凸函数，且

$$\sum_{i=1}^{n} a_i = 1$$

其中 $a_i > 0$，$1 \leqslant i \leqslant n$。那么

$$\sum_{i=1}^{n} a_i f(x_i) \leqslant f\left(\sum_{i=1}^{n} a_i x_i\right)$$

其中 $x_i \in I$，$1 \leqslant i \leqslant n$。当且仅当 $x_1 = \cdots = x_n$ 等式成立。

熵的性质

定理 2.7 $H(\mathbf{X},\mathbf{Y}) \leqslant H(\mathbf{X}) + H(\mathbf{Y})$，当且仅当 $\mathbf{X}$ 和 $\mathbf{Y}$ 统计独立时等号成立。

定理的直观解释： 随机变量(X,Y)含有的信息量
不超过随机变量X的信息量+随机变量Y的信息量

熵的性质

证明：

假设 $\mathbf{X}$ 取值 x_i，$1 \leqslant i \leqslant m$，$\mathbf{Y}$ 取值 y_j，$1 \leqslant j \leqslant n$。记 $p_i = \Pr[\mathbf{X} = x_i]$，$1 \leqslant i \leqslant m$；$q_j = \Pr[\mathbf{Y} = y_j]$，$1 \leqslant j \leqslant n$。记 $r_{ij} = \Pr[\mathbf{X} = x_i, \mathbf{Y} = y_j]$，$1 \leqslant i \leqslant m$，$1 \leqslant j \leqslant n$

于是 $p_i = \sum_{j=1}^{n} r_{ij}$ $q_j = \sum_{i=1}^{m} r_{ij}$

$$H(\mathbf{X}) + H(\mathbf{Y}) = -\left(\sum_{i=1}^{m} p_i \mathrm{lb}\, p_i + \sum_{j=1}^{n} q_j \mathrm{lb}\, q_j \right)$$

$$= -\left(\sum_{i=1}^{m} \sum_{j=1}^{n} r_{ij} \mathrm{lb}\, p_i + \sum_{j=1}^{n} \sum_{i=1}^{m} r_{ij} \mathrm{lb}\, q_j \right)$$

$$H(\mathbf{X},\mathbf{Y}) - H(\mathbf{X}) - H(\mathbf{Y}) = \sum_{i=1}^{m} \sum_{j=1}^{n} r_{ij} \mathrm{lb} \frac{1}{r_{ij}} + \sum_{i=1}^{m} \sum_{j=1}^{n} r_{ij} \mathrm{lb}\, p_i q_j$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} r_{ij} \mathrm{lb} \frac{p_i q_j}{r_{ij}}$$

$$\leqslant \mathrm{lb} \sum_{i=1}^{m} \sum_{j=1}^{n} p_i q_j = 0$$

等式成立当且仅当 $r_{ij} = p_i q_j$，也就是

$$\Pr[\mathbf{X} = x_i, \mathbf{Y} = y_j] = \Pr[\mathbf{X} = x_i] \Pr[\mathbf{Y} = y_j]$$

熵的性质

- 条件熵

定义 2.6 假设 $\mathbf{X}$ 和 $\mathbf{Y}$ 是两个随机变量。对于 $\mathbf{Y}$ 的任何固定值 y，得到一个 $\mathbf{X}$ 上的(条件)概率分布；记相应的随机变量为 $\mathbf{X}\mid y$。显然

$$H(\mathbf{X}\mid y)=-\sum_{x}\Pr[x\mid y]\mathrm{lb}\,\Pr[x\mid y]$$

定义条件熵 $H(\mathbf{X}\mid\mathbf{Y})$ 为熵 $H(\mathbf{X}\mid y)$ 取遍所有的 y 的加权平均值。计算公式为

$$H(\mathbf{X}\mid\mathbf{Y})=-\sum_{y}\sum_{x}\Pr[y]\Pr[x\mid y]\mathrm{lb}\,\Pr[x\mid y]$$

条件熵度量了 $\mathbf{Y}$ 揭示的 $\mathbf{X}$ 的平均信息量。

熵的性质

- 练习

定理 2.8 $H(\mathbf{X},\mathbf{Y}) = H(\mathbf{Y}) + H(\mathbf{X}|\mathbf{Y})$。

$$\mathrm{H}(X,Y) \leq \mathrm{H}(X) + \mathrm{H}(Y).$$

推论 2.9 $H(\mathbf{X}|\mathbf{Y}) \leqslant H(\mathbf{X})$，等式成立当且仅当 $\mathbf{X}$ 和 $\mathbf{Y}$ 统计独立。

直观：不确定性事件熵值(Entrophy)更大

伪密钥和唯一解距离

- 密钥分析的目的是确定密钥，因此研究H(K|C)—密钥含糊度

定理 2.10　设 $(\mathcal{P}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D})$ 是一个密码体制，那么

$$H(\mathbf{K} \mid \mathbf{C}) = H(\mathbf{K}) + H(\mathbf{P}) - H(\mathbf{C})$$

伪密钥和唯一解距离

证明：

$$H(\mathbf{K}\mid\mathbf{C}) = H(\mathbf{K},\mathbf{C}) - H(\mathbf{C})$$
$$= H(\mathbf{K},\mathbf{P},\mathbf{C}) - H(\mathbf{C})$$

注意到 $H(\mathbf{K},\mathbf{P},\mathbf{C}) = H(\mathbf{K},\mathbf{P}) = H(\mathbf{K}) + H(\mathbf{P})$

伪密钥和唯一解距离

- 密码分析的目的是确定密钥

- 假设Oscar获得密文串WNAJW(移位密码)
- 可能对应river(加密密钥F(=5))
- 可能对应arena(加密密钥W(=22))

- 希望推导伪密钥的期望数的下界

伪密钥和唯一解距离

- 考虑自然语言的特性：每个信源符号平均信息的度量
- P^n为所有n字母组的概率分布构成的随机变量

定义 2.7 假设 L 是自然语言，语言 L 的熵定义为

$$H_L = \lim_{n\to\infty} \frac{H(\mathbf{P}^n)}{n}$$

语言 L 的冗余度(redundancy)定义为

$$R_L = 1 - \frac{H_L}{\mathrm{lb}\,|\mathcal{P}|}$$

伪密钥和唯一解距离

- 比如在英语中：
- $H(P) \approx 4.19$; $\frac{H(P^2)}{2} \approx 3.90$;
- 各种实验已经得到一个经验性的结果

$$1.0 \leq H_L \leq 1.5$$

若取H_L=1.25， 冗余度R_L?

75%

伪密钥和唯一解距离

给定 $y \in \mathbf{C}^n$，定义

$$K(y)=\{K \in \mathcal{K}: \exists x \in \mathcal{P}^n \text{ 使得 } \Pr[x]>0, e_K(x)=y\}$$

伪密钥的平均数目(在所有可能的长为 n 的密文串上)记为 $\overline{s}_n$，

计算可以得到

$$\begin{aligned}\overline{s}_n &= \sum_{y \in C^n} \Pr[y](|K(y)|-1) \\ &= \sum_{y \in C^n} \Pr[y]|K(y)| - \sum_{y \in C^n} \Pr[y] \\ &= \sum_{y \in C^n} \Pr[y]|K(y)| - 1\end{aligned}$$

伪密钥和唯一解距离

定理 2.11 假设 $(\mathcal{P}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D})$ 是一个密码体制，$|\mathcal{C}|=|\mathcal{P}|$ 并且密钥是等概率选取的。设 R_L 表示明文的自然语言的冗余度，那么给定一个充分长(长为 n)的密文串，伪密钥的期望数满足

$$\overline{s}_n \geqslant \frac{|\mathcal{K}|}{|\mathcal{P}|^{nR_L}} - 1$$

伪密钥和唯一解距离

定义 2.8 一个密码体制的唯一解距离定义为使得伪密钥的期望数等于零的 n 的值，记为 n_0，即在给定的足够的计算时间下分析者能唯一计算出密钥所需密文的平均量。

令 $\overline{s}_n = 0$，　　$n_0 \approx \dfrac{\mathrm{lb}|\mathcal{K}|}{R_L \mathrm{lb}|\mathcal{P}|}$

考虑代换密码。在这种密码体制中 $|\mathcal{P}| = 26$，$|\mathcal{K}| = 26!$。如果取 $R_L = 0.75$，就得到唯一解距离的估计为

$$n_0 \approx 88.4/(0.75 \times 4.7) \approx 25$$

乘积密码体制

- 利用简单密码体制“组合”形成新的密码体制
- 乘积密码

设 $S_1=(\mathcal{P},\mathcal{P},\mathcal{K}_1,\mathcal{E}_1,\mathcal{D}_1)$， $S_2=(\mathcal{P},\mathcal{P},\mathcal{K}_2,\mathcal{E}_2,\mathcal{D}_2)$

$S_1\times S_2$ 定义为

$$(\mathcal{P},\mathcal{P},\mathcal{K}_1\times\mathcal{K}_2,\mathcal{E},\mathcal{D})$$

$K=(K_1,K_2)$，其中 $K_1\in\mathcal{K}_1$， $K_2\in\mathcal{K}_2$。

e_K 定义为 $e_{(K_1,K_2)}(x)=e_{K_2}(e_{K_1}(x))$

d_K 定义为 $d_{(K_1,K_2)}(y)=d_{K_1}(d_{K_2}(y))$

乘积密码体制

- 乘积运算是可结合的：

$$(S_1 \times S_2) \times S_3 = S_1 \times (S_2 \times S_3)$$

- 一个密码体制是幂等的，如果$S^2 = S$

比如移位密码、代换密码、仿射密码...

乘积密码体制

- 如果$M \times S = S \times M$,那么称密码体制M和S是可交换的。
- 比如：M是乘法密码（密钥a等概率选取）

密码体制 2.2　乘法密码

设$\mathcal{P} = \mathcal{C} = \mathbb{Z}_{26}$，并且

$$\mathcal{K} = \{a \in \mathbb{Z}_{26} : \gcd(a, 26) = 1\}$$

对于$a \in \mathcal{K}$，定义

$$e_a(x) = ax \bmod 26$$

和

$$d_a(y) = a^{-1}y \bmod 26$$

$(x, y \in \mathbb{Z}_{26})$。

- S是移位密码（密钥K等概率选取）

乘积密码体制

乘积密码 $M \times S$ 的密钥具有 (a, K) 的形式，并且

$$e_{(a,K)}(x) = (ax + K) \bmod 26$$

考虑 $S \times M$ 。这个密码的密钥具有形式 (K, a)，并且

$$e_{(K,a)}(x) = a(x + K) \bmod 26 = (ax + aK) \bmod 26$$

- 双射

$M \times S$中的密钥(a,K) $\leftrightarrow$ $S \times M$中的密钥($a^{-1}K, a$)

乘积密码体制

- 考虑合理的乘积组合方式

如果S_1和S_2都是幂等的，并且是可交换的，则$S_1 \times S_2$也是幂等的。

$$
\begin{aligned}
(S_1 \times S_2) \times (S_1 \times S_2) &= S_1 \times (S_2 \times S_1) \times S_2 \\
&= S_1 \times (S_1 \times S_2) \times S_2 \\
&= (S_1 \times S_1) \times (S_2 \times S_2) \\
&= S_1 \times S_2
\end{aligned}
$$

所以$(S_1 \times S_2)^n, n \geq 2$没有意义。