



Sessão nº4

Chaves espúrias e Distância de unicidade



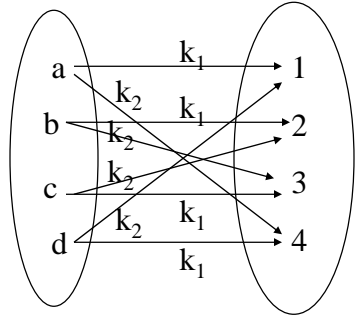
Introdução



- Considere o sistema criptográfico (P, C, K, E, D) onde $P=C=K=\mathbf{Z}_2=\{0,1\}$, as regras E e D correspondem à **cifra de substituição** e a mesma chave é usada na cifra de vários textos.
 - Como varia o equívoco da chave $H(K|C)$ com o número de criptogramas observados?



Análise do sistema Vernam usando a mesma chave para dois textos



Seja
 $a \circ 00 \quad b \circ 01 \quad c \circ 10 \quad d \circ 11$
 $1 \circ 00 \quad 2 \circ 01 \quad 3 \circ 10 \quad 4 \circ 11$

Sabendo que
 $p(0) = 1/4, \quad p(1) = 3/4$
 $P(k_1) = 1/2, \quad p(k_2) = 1/2$

Verifica-se que
 $H(K) = 1$ [bit/chave]

Depois de observarmos dois criptogramas, $H(K|C) = ?$

$$H(K|C) \approx 0,67$$

é ainda menor do que $H(K|C)$ observando um só criptograma!



Análise do sistema Vernam usando a mesma chave para dois textos: expressões genéricas



Considere-se que a sequência de n textos em claro depois de cifrados têm b zeros e $n-b$ uns. Seja ainda $p \equiv p(0)$ e $q \equiv p(1)$, então genericamente

$$p(c) = 1/2 p^b q^{n-b} + 1/2 p^{n-b} q^b$$

$$p(k,c) = \begin{cases} 1/2 p^b q^{n-b} & \text{se } K=k_1 \\ 1/2 p^{n-b} q^b & \text{se } K=k_2 \end{cases}$$

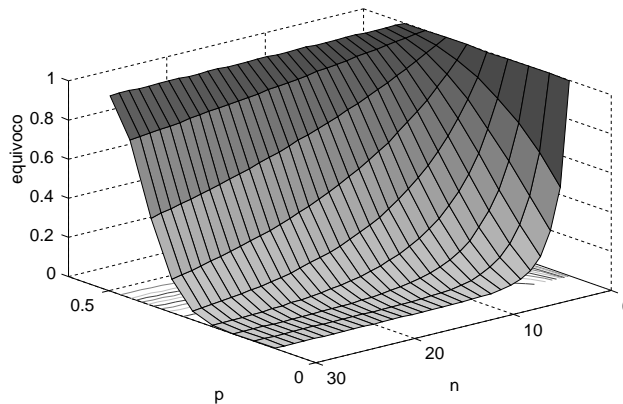
$$p(k|c) = \begin{cases} (p^b q^{n-b}) / (p^b q^{n-b} + p^{n-b} q^b) & \text{se } K=k_1 \\ (p^{n-b} q^b) / (p^b q^{n-b} + p^{n-b} q^b) & \text{se } K=k_2 \end{cases}$$

$$H(K|C) = \sum_{b=0}^n \binom{n}{b} p^b q^{n-b} \log_2 [(p^b q^{n-b} + p^{n-b} q^b) / (p^b q^{n-b})]$$

c.q.d.



Equívoco da chave na criptoanálise



- Equívoco na criptoanálise de uma **cifra de substituição** aplicada a um conjunto de 2 textos em claro (**p** é probabilidade dum dos textos e **n** é o número de criptogramas observados).



Chaves espúrias



- Exemplo: consideremos como linguagem o **Inglês** e ainda o criptograma **WNAJW** obtido pela **cifra de substituição**.
 - Existem apenas dois textos possíveis (com significado): **river** e **arena**. Por isso são apenas duas as **chaves possíveis**. Aquela que não é a verdadeira designa-se por **chave espúria**.



Número de chaves espúrias



- Seja o sistema criptográfico (P, C, K, E, D) onde $|C| = |P|$ e as chaves são equiprováveis. Dado um conjunto de n símbolos do criptograma, com n grande, o **número esperado de chaves espúrias** satisfaz

$$E[s_n] \approx (|K| / |P|^{n R_L}) - 1$$

- R_L representa a redundância da linguagem; toma valores entre 0 e 1
- conforme o n aumenta $E[s_n] \rightarrow 0$
- para n pequeno $E[s_n]$ é uma estimativa pouco precisa porque também o é a da probabilidade logo a de R_L



Distância de unicidade dum sistema criptográfico



- Define-se **Distância de Unicidade** como sendo o valor de n (n_0) para o qual o número esperado de chaves espúrias se torna 0.

$$n_0 \approx \log_2 |K| / (R_L \times \log_2 |P|)$$

- é a quantidade média de criptograma necessária para um adversário ser capaz de calcular a chave (única), dado o tempo de cálculo suficiente.





- Exemplo: seja o sistema de **cifra por substituição**, tal que $|\mathbf{P}|=26$ e $|\mathbf{K}|=26!$. Admitindo que a linguagem dos textos é o Inglês e que $R_L=0,75$ então uma estimativa da **distância de unicidade** é

$$\begin{aligned}n_0 &\approx \log_2 |\mathbf{K}| / (R_L \times \log_2 |\mathbf{P}|) \\ &\approx 88,4 / (0,75 \times 4,7) \approx 25\end{aligned}$$

- em média depois de observados pelo menos 25 letras do criptograma é possível determinar a chave

