3-Criptografía teórica: cifrado perfecto y distancia de unicidad

Seguridad Perfecta de los criptosistemas

- En 1949 Claude Shannon publicó (conectar desde el vpn UAM):
  - C. E. Shannon, "Communication theory of secrecy systems," in The Bell System Technical Journal, vol. 28, no. 4, pp. 656-715, Oct. 1949, doi: 10.1002/j.1538-7305.1949.tb00928.x.
- Que tiene una gran importancia en el estudio teórico de la criptografía, en su seguridad y la forma de medirla.
- La herramienta fundamental es probabilidad y teoría de información.
- Es este capítulo de la asignatura discutiremos alguna ideas de este importante trabajo.

Seguridad Perfecta de los criptosistemas

- Hay diferentes enfoques para discutir la seguridad de los criptosistemas:
  - **Seguridad computacional** (se refiere al esfuerzo computacional): Se entiende como que el tiempo necesario para romper un criptosistema es superior a su tiempo de uso.
  - **Seguridad probada**: Cuando se dice que un criptosistema es seguro en función de la seguridad de un problema determinado:
    - Por ejemplo la factorización de dos primos para RSA.
  - **Seguridad incondicional**: Si se supone que tenemos una máquina de recursos ilimitados (infinitos), esta no es capaz de romper el criptosistema.

Seguridad Perfecta de los criptosistemas

- Siempre que se discute la seguridad de un criptosistema hay que especificar el tipo de ataque que se considera.
- Por ejemplo, sustitución y Vigenère (pero Vigenère solo para claves cercanas al tamaño de texto) pueden ser computacionalmente seguros con el conocimiento de A (Ataque con sólo texto cifrado).
- Este capítulo se desarrolla una teoría para generar seguridad incondicional en A (Ataque con sólo texto cifrado).
- En este contexto surge un concepto nuevo que seguridad perfecta, al que todos los criptoistemas deberían aproximarse.
- No podemos utilizar teoría de complejidad ya que el tiempo en seguridad incondicional va a infinito, por tanto necesitamos otra herramienta que es la teoría de probabilidad.
- Por eso vamos a hacer un pequeño repaso de teoría de probabilidad, y vamos a poner el criptosistema en formalismo probabilístico.
- En este contexto de seguridad perfecta, por el momento vamos a asumir que una clave es utilizada solamente para una única encriptación y cada vez que encriptamos de nuevo cambiamos la clave.

Seguridad Perfecta: Repaso de teoría de Probabilidad

- Sea $x \in X=\{x_1, \ldots, x_n\}$ e $y \in Y=\{y_1, \ldots, y_n\}$.
- La probabilidad a priori de $x = x_i$, o $y = y_i$, se escribe como:
  - $P(x = x_i)=P(x)$.
  - $P(y = y_i)=P(y)$.
- La probabilidad conjunta de tener $x= x_i$, e $y= y_i$, se escribe como:
  - $P(x = x_i , y = y_i)=P(x,y)$.
- La probabilidad conjunta de tener $x= x_i$, si conozco $y= y_i$, se escribe como (probabilidad condicional):
  - $P(x|y) = P(x,y)/P(y) = (P(y|x)P(x)) / P(y) = (P(y|x)P(x)) / \sum_x P(y|x)P(x)$. (Formula de Bayes, y teorema de Bayes).

Seguridad Perfecta: Repaso de teoría de Probabilidad

- La probabilidad conjunta se puede poner en función de la condicional, a través de la fórmula de Bayes:
  - $P(x,y) = P(x|y)\,P(y) = P(y|x)\,P(x)$.
- La independencia estadística se puede definir por:
  - $P(x,y) = P(x)\,P(y)$.
- Vamos a ver como hacemos ahora la translación al criptosistema con el formalismo de probabilidad.
- Para ello consideramos los textos planos con una cierta probabilidad.
- Las claves con una cierta probabilidad.
- Y estas probabilidades infieren una probabilidad en los textos cifrados.

Seguridad Perfecta: Probabilidad en el criptosistema

- A continuación hacemos la translación al criptosistema en formalismo probabilístico.
- Dado que tenemos un lenguaje original, podemos expresar la probabilidad a priori del texto original como: $P_p(x)$, $x \in P$.
- Por ejemplo la frecuencia de los caracteres en inglés.
- De igual forma podemos calcular la probabilidad a priori de las claves como: $P_k(k)$, $k \in K$.
- $P_k(k)$ puede ser cualquiera, pero es conocida.
- En general, por seguridad, se asume o se admite que hay independencia estadística entre el espacio P y K, es decir que $P(x,k) = P_p(x)\ P_k(k)$.
- Esta condición es necesaria para que el criptosistema sea resistente a ataques.

Seguridad Perfecta: Probabilidad en el criptosistema

- Como consecuencia de tener una $P_p(x)$, $x \in P$, y una $P_k(k)$, $k \in K$, están inducen una $P_c(y)$, $y \in C$ en los textos cifrados,
- Vamos a calcular como es esta probabilidad:

$$P_c(y) = \sum_{xk} P(x,k,y) = \sum_{xk} P(y \mid x,k)P(x,k) = \sum_{xk} P(y \mid x,k)\, Pp(x)\, Pk(k).$$

$P(x,k) = P_p(x)\, P_k(k)$

- Ahora sabemos que:
- $\begin{cases} P(y \mid x,k) = 1, \ si\ e_k(x) = y \\ P(y \mid x,k) = 0, \ si\ e_k(x) \neq y \end{cases}$, por tanto

$$P_c(y) = \sum_{\forall x,\, k \mid y = e_k(x)} P_p(x)\, Pk(k).$$

Seguridad Perfecta: Probabilidad en el criptosistema

- De la misma forma podemos sacar también la probabilidad de $P_c(y \mid x)$, $y \in C$.
- $P_c(y \mid x) = P(x,y)/P_p(x) = \sum_k P(x,k,y) / P_p(x) = \sum_k P(y \mid x,k)P(x,k) / P_p(x) = \sum_k P(y \mid x,k)Pp(x)\ Pk(k) / P_p(x) = \sum_k P(y \mid x,k)\ Pk(k)$.
- Ahora sabemos de nuevo que:
- $\begin{cases} P(y \mid x,k) = 1, \ si\ e_k(x) = y \\ P(y \mid x,k) = 0, \ si\ e_k(x) \neq y \end{cases}$, por tanto

$P(x,k) = P_p(x)\ P_k(k)$

$$P_c(y \mid x) = \sum_{\forall k \mid y = e_k(x)} P_k(k).$$

Seguridad Perfecta

- Una vez que ya tenemos el criptosistema en el marco probabilístico, podemos definir la seguridad perfecta de un criptosistema, como:
  - $P_p(x \mid y) = P_p(x)$, $\forall\ x \in P$, y $\forall\ y \in C$.

- Recordemos que asumimos que la clave es usada en un único cifrado, y que cada vez que se cifra se cambia a clave de acuerdo a una distribución de probabilidad de claves.
- Esto básicamente quiere decir que no se puede obtener información del texto plano observando el texto cifrado: es decir se puede obtener la misma información del texto plano teniendo el texto cifrado (el texto cifrado nos da nueva información sobre el texto plano).
- Recordemos que esta es una teoría para sistemas de seguridad incondicional, cuando se produce un ataque del tipo A que ya estudiamos.

Seguridad Perfecta

- Ej. Seguridad perfecta. Comprobar si el criptosistema siguiente tiene seguridad perfecta.
  - P={a,b}, $P_p(a)=1/4$, y $P_p(b)=3/4$.
  - K={k_1,k_2,k_3}, $P_k(k_1)=1/2$, y $P_k(k_2) = P_k(k_3) = 1/4$.
  - C={1,2,3,4}.
  - $e_{k1}(a) = 1$, $e_{k1(}b) = 2$, $e_{k2}(a) = 2$, $ek_2(b) = 3$, $e_{k3}(a) = 3$, $ek_3(b) = 4$.
- Sabemos que $P_c(y) = \sum_{\forall x, k \mid y=e_k(x)} P_p(x)\ Pk(k)$, por tanto

  $P_c(1) = P_p(a)\ P_k(k_1) = 1/8$,
  $P_c(2) = P_p(a)\ P_k(k_2) + P_p(b)\ P_k(k_1) = 7/16$,
  $P_c(3) = P_p(a)\ P_k(k_3) + P_p(b)\ P_k(k_2) = 1/4$,
  $P_c(4) = P_p(b)\ P_k(k_3) = 3/16$.

Seguridad Perfecta

- Ahora calculamos $P_p(x \mid y)$ a través del teorema de Bayes:
  - $P_p(x \mid y) = P_p(x)\, P_c(y \mid x) / P_c(y) =$
    $P_p(x) \sum_{\forall k \mid y=e_k(x)} P_k(k) / \sum_{\forall x,\, k \mid y=e_k(x)} P_p(x) P_k(k).$
  - $P_p(a \mid 1) = P_p(a)\, P_k(k_1) / P_c(1) = 1,$
    $P_p(a \mid 2) = P_p(a)\, P_k(k_2) / P_c(2) = 1/7,$
    $P_p(a \mid 3) = P_p(a)\, P_k(k_3) / P_c(3) = 1/4,$
    $P_p(a \mid 4) = 0 / P_c(4) = 0.$
  - $P_p(b \mid 1) = 0 / P_c(1) = 0,$
    $P_p(b \mid 2) = P_p(b)\, P_k(k_1) / P_c(2) = 6/7,$
    $P_p(b \mid 3) = P_p(b)\, P_k(k_2) / P_c(3) = 3/4,$
    $P_p(b \mid 4) = P_p(b)\, P_k(k_3) / P_c(4) = 1.$

Claramente no tiene seguridad perfecta, ya que no se cumple $P_p(x \mid y) = P_p(x)$, $\forall\ x \in P$, y $\forall\ y \in C$.

Por ejemplo:

$P_p(a \mid 2) = 1/7 \neq P_p(a) = 1/4.$

Este hecho da clara información a un criptoanalista.

Seguridad Perfecta del cifrado por desplazamiento

- Vamos a enunciar el siguiente teorema sobre la seguridad del cifrado por desplazamiento.
- Supongamos que m claves en un **cifrado por desplazamiento** se utilizan con la probabilidad de $P_k(k) = 1/m$, entonces podemos afirmar que para cualquier $P_p(x)$ el cifrado por desplazamiento tiene seguridad perfecta.
- Observar que la clave que se escoge para cifrar se elige de una distribución uniforme $P_k(k) = 1/m$ en cada cifrado carácter a carácter.
- Vamos a demostrar este teorema que es muy importante, ya que está relacionado con la seguridad de los cifrados de flujo.

Seguridad Perfecta del cifrado por desplazamiento

- Como habíamos demostrado antes para cualquier criptosistema

  $P_c(y) = \sum_{\forall x,k \mid y=x+k \bmod m} P_p(x)\, P_k(k) =$

  $= \sum_{\forall x,k \mid y=x+k \bmod m} P_p(x)\, (1/m) =$

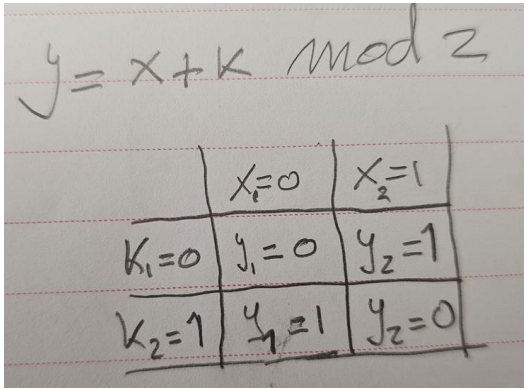
- Para un determinado **y** el sumatorio en **k** tiene m posibles valores, y cada x está relacionado con un y para cada k, por tanto:

  $= (1/m) \sum_{\forall x,k \mid y=x+k \bmod m} P_p(x)$

  $= (1/m) \sum_x P_p(x) = (1/m).$

- Ahora calculamos la condicional

  $P_c(y \mid x) = \sum_{\forall k \mid y=x+k \bmod m} P_k(k)$



Seguridad Perfecta del cifrado por desplazamiento

- La pregunta ahora para ese sumatorio es cuantas claves me relacionan un determinado x con un y, la respuesta es una sola clave por tanto

$P_c(y \mid x) = \sum_{\forall k \mid y = x + k \bmod m} P_k(k) = 1/m.$

- Así para la seguridad perfecta:

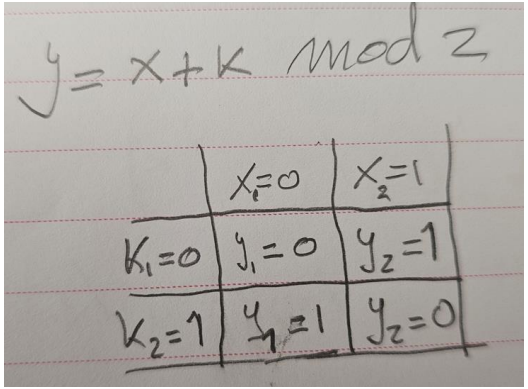
$P_p(x \mid y) = P_p(x)\, P_c(y \mid x) \,/\, P_c(y) =$

$P_p(x) \sum_{\forall k \mid y = e_k(x)} P_k(k) / \sum_{\forall x, k \mid y = e_k(x)} P_p(x) P_k(k) =$

$(P_p(x)\,(1/m)) \,/\, (1/m) = P_p(x).$

Seguridad Perfecta del cifrado por desplazamiento

- Un ejemplo para entender este teorema para m=2 en el cifrado por desplazamiento:
  - P={0,1}, con una cierta $P_p(x)$,
  - K={0,1}, $P_k(0) = P_k(1) = ½$,
  - C={0,1}.
  - La regla de cifrado para todas las posibilidades es $e_0(0) = 0$, $e_1(0) = 1$, $e_0(1) = 1$, $e_1(1) = 0$.



- Este es un cifrado por desplazamiento para m=2, si hacéis el cálculo de las probabilidades para la seguridad perfecta, el resultado es que cumple la seguridad perfecta (hacerlo como ejercicio para casa).
- Observar que si $P_k(k) \neq 1/m$, ya no cumple seguridad perfecta (hacerlo como ejercicio para casa).

Seguridad Perfecta: Cifrado de Vernam.

- Existe un cifrado que es el cifrado de Vernam, que fue utilizado por primera vez en 1917, y que se sabía que era seguro pero no porque.
- Pero hasta 30 años más tarde que Shannon publicó este tratado C. E. Shannon, "Communication theory of secrecy systems," in The Bell System Technical Journal, vol. 28, no. 4, pp. 656-715, Oct. 1949, doi: 10.1002/j.1538-7305.1949.tb00928.x., no se demostró su seguridad incondicional.
- Este cifrado también se denomina OTP (One-time pad).

Seguridad Perfecta: Cifrado de Vernam.

- Este cifrado consiste en lo siguiente:
  - $|P|=|C|=|K|=2$.
  - $M_p = x_1\ x_2 \ldots x_n$, con $x_i \in Z_2$
  - $P_k(ki) = 1/2$.
  - Se selecciona una secuencia de claves aleatorias de la misma longitud del mensaje, con la distribución e probabilidad $P_k(ki) = 1/2$:
    - $K = k_1\ k_2 \ldots k_n$, con $k_i \in Z_2$, y $P_k(ki) = 1/2$.
  - El cifrado se realiza como $y_i = E_{ki}(x_i) = x_i + k_i \bmod 2$.
  - $M_c = y_1\ y_2 \ldots y_n$, con $y_i \in Z_2$
  - Obviamente la clave se debe enviar por un canal seguro y es de la misma longitud del mensaje a cifrar.

Seguridad Perfecta: Teorema del cifrado Perfecto.

- Todos estos resultados nos lleva a enunciar el teorema del cifrado perfecto.
- El **teorema del cifrado perfecto** dice, dado un criptosistema [P, C, K, E_k, D_k], donde $|P|=|C|=|K|$, este proporciona seguridad perfecta si solo si
  1. $P_k(k) = 1/|K| \ \forall \ k \in K$.
  2. Para todo $x \in P$ y para todo $y \in C$, hay una sola clave tal que me relaciona x con y a través de la transformación de cifrado $y = e_k(x)$.
- Demostrar este teorema para casa, la demostración es similar a la anterior (Cryptography : theory and practice. Douglas Robert Stinson).

Claves espúreas y distancia de unicidad

- En este apartado lo que queremos ver es que ocurre con el criptosistema cuando más y más textos se cifran con la misma clave y como un criptoanalista podría llevar a cabo un ataque tipo A, con el suficiente tiempo.
- **El enfoque es diferente al de antes** que cuando una clave es utilizada por una sola encriptación, y que cada vez que se cifra se cambia la clave de acuerdo a una distribución de probabilidad de claves.
- La herramienta utilizada ahora para estudiar en este contexto la seguridad de los criptosistemas es la teoría de información.
- En este campo la medida de incertidumbre o información se calcula basándose en la probabilidad.

Claves espúreas y distancia de unicidad: Incertidumbre

- Qué información ganamos cuando recibimos un mensaje, x_i, de un conjunto de posibles mensajes $x \in X=\{x_1, \ldots, x_n\}$, cuando a esos mensajes tenemos asociadas las posibles probabilidades de ocurrencia $P_p(x) \in \{p_1, \ldots, p_n\}$.
- ¿Qué tiene más información?
  - Que os diga que me estoy comiendo una manzana de postre (más probable).
  - Que os diga que se está produciendo en estos momentos un terremoto en Cantoblanco (altamente improbable).
- Así, la información depende de los posibles mensajes y de sus frecuencias (probabilidad $p_i = P_p(x_i)$ del mensaje).

Claves espúreas y distancia de unicidad: Incertidumbre

- Así, la información debería ser proporcional a la inversa de la probabilidad del suceso que se está midiendo.
- Más aún, diferentes informaciones independientes deberían ser aditivas.
- Pero la probabilidad de mensajes independientes es el producto de las probabilidades.
- Por tanto, necesitamos una función que cumpla la propiedad: $f(xy)=f(x)+f(y)$.
- El logaritmo tiene esas características.

Claves espúreas y distancia de unicidad: Incertidumbre

- De esta manera, ya podemos definir nuestra medida de información como: $I(x_i)=\log_2(1/p_i) = - \log_2(p_i)$. Esto es la ganancia de información para un determinado mensaje.
- La entropía es el promedio de la ganancia de información de todos los mensajes: $H(x)= - \sum_{x=1}^{n} P_p(x_i) \log_2(Pp(x_i))$.
- La entropía de un conjunto de N mensajes igual de probables es: $H(x)= - \sum_{x=1}^{n} P_p(x_i) \log_2(Pp(x_i)) = - \sum_{x=1}^{n} 1/N \log_2(1/N) = \log_2(N)$.
- Es decir la entropía aumenta con el número de mensajes N equiprobables.
- La ganancia de información se puede ver como una reducción de la incertidumbre del sistema: por ejemplo en una urna con bolas 9n+1b si sacamos 1b se elimina completamente la incertidumbre para sacar bolas después de la urna.

Claves espúreas y distancia de unicidad: Entropía de un lenguaje

- Si tenemos un lenguaje de 26 caracteres aleatorio, $P_p(x_i)=1/26$, entonces la entropía de este lenguaje me da $H_L=\log_2(26) = 4{,}70$.
- Si ponemos la distribución del inglés para $P_p(x_i)$ (transparencia 45 del tema anterior), entonces el cálculo de entropía viene determinado por $H_L= 4{,}19$.
- Pero esto es una mala estimación de la entropía de un lenguaje, ya que estamos despreciando todas las correlaciones entre los símbolos del lenguaje a más alto orden:
  - Así por ejemplo hay una estructura diferente en las agrupaciones de dos caracteres. Si denotamos $P_p^2(x_i)$ como la distribución de probabilidad para la agrupación de dos símbolos en el leguaje, en ese caso podemos definir la entropía por símbolo del lenguaje para agrupaciones de dos símbolos como: $H(P_p^2(x_i))/2$, que para el caso del inglés es $\cong 3{,}90$.

Claves espúreas y distancia de unicidad: Entropía de un lenguaje

- Por tanto esto lo podemos extender a estructura en agrupaciones de n-gramas de símbolos del lenguaje obteniendo que la entropía de un lenguaje la podemos definir por:
  - $H_L = \lim_{n\to\infty} \left(\frac{H(P_p^n(x))}{n}\right)$.
- Donde $P_p^n(x)$ representa la distribución de probabilidad de las agrupaciones de n-gramas de símbolos del lenguaje en cuestión.
- Resultados empíricos para el lenguaje del inglés muestran que la entropía de este se encuentra en el rango: $1 \leq H_L \leq 1{,}5$.
- Este valor nos permitirá estimar la distancia de unicidad de un lenguaje.

Claves espúreas y distancia de unicidad: Redundancia de un lenguaje

- La entropía de un lenguaje nos permite definir así la redundancia del mismo en función de este:
  - $R_L = 1 - \left(\frac{H_L}{\log_2 |P|}\right)$, $R_L \in [0, 1]$.
- Cuando $R_L = 1$ decimos que es un lenguaje muy redundante, y cuando $R_L = 0$ decimos que es un lenguaje poco redundante.
- Así un lenguaje aleatorio se caracteriza por tener $H_L = \log_2 |P|$, como ya dijimos todos los símbolos son equiprobables y no tiene estructura a más alto nivel, por tanto $R_L = 0$.
- Por lo tanto un lenguaje aleatorio es que tiene la mínima redundancia.

Claves espúreas y distancia de unicidad: Redundancia de un lenguaje

- Si tomamos el valor medio del intervalo para el idioma inglés $1 \leq H_L \leq 1{,}5$, para la entropía $H_L = 1{,}25$, entonces la redundancia es:
  - $R_L = 1 - \left(\frac{H_L}{\log_2|P|}\right) = 1 - \left(\frac{1{,}25}{\log_2|26|}\right) = 1 - \left(\frac{1{,}25}{\log_2|26|}\right) = 1 - \left(\frac{1{,}25}{4{,}70}\right) \cong 0{,}75.$

- Estos valores los utilizaremos para el cálculo de claves espúreas y distancia de unicidad.

Claves espúreas y distancia de unicidad

- Uno de los objetivos de un criptoanalista es averiguar la clave de cifrado.
- Debido a la estructura de un lenguaje y al tamaño de bloque del cifrado, es posible existan varias claves que al descifrar den un texto con sentido.
- Esto es lo que se denominan claves espúreas, definiéndose como:
  - $e_{k_1}(\overline{x_1}) = \bar{y}$
    $e_{k_2}(\overline{x_2}) = \bar{y}$
    ……
    $e_{k_s}(\overline{x_s}) = \bar{y}$.
- Donde los textos planos $\overline{x_1}$, $\overline{x_2}$, …, $\overline{x_s}$, tienen sentido en lenguaje, siendo x ∈ $P_p^n(x)$, y y ∈ $C_p^n(x)$.
- Al igual que existe una $P_p^n(x)$ que representa la distribución de probabilidad de las agrupaciones de n-gramas de símbolos de los textos planos, esta probabilidad y la probabilidad de las claves inducen en los textos cifrados $C_p^n(x)$ que representa la distribución de probabilidad de las agrupaciones de n-gramas de símbolos de los textos cifrados.

Claves espúreas y distancia de unicidad

- De todas la claves espúreas, solo una es la utilizada para el cifrado y es la verdadera.
- Ej. Por ejemplo supongamos que tenemos el siguiente texto cifrado por desplazamiento: WNAJW.
  - Así en este caso, al menos podemos encontrar dos claves espúreas: la clave F(=5), y la clave W(=22), que dan con textos planos "*river*" y "*arena*", ambos con sentido en el lenguaje original.
- Así a partir de ahora vamos a buscar cual es tamaño e bloque de cifrado, a partir del cual la probabilidad de obtener claves espúreas es mínima.
- Es claro que cuando más grandes sean los bloques de cifrado menos probabilidad de obtener claves espúreas.

Claves espúreas y distancia de unicidad

- Se puede demostrar que dado un texto cifrado de longitud n, el número medio de claves espúreas, $\overline{S_n}$ es siempre mayor o igual que esta cantidad:
  - $\overline{S_n} \geq \frac{|K|}{|\mathrm{P}|^{nR_L}} - 1$, siempre que $|P| = |C|$ y $P_k(k) = \frac{1}{|K|}$.
  - $\overline{S_n} \geq |K| \left(\frac{|\mathrm{P}|^{1-R_L}}{|C|}\right)^n - 1$, siempre que $|P| \neq |C|$ y $P_k(k) = \frac{1}{|K|}$.

POSIBLE Tarea de Evaluación Continúa: Deducir estas cotas para el número medio de claves espúreas. Es decir Estudio analítico teórico de la distancia de unicidad y las claves *espúreas* de un criptosistema basándose en el libro Douglas R. Stinson.

- Según n tiende a infinito la cantidad $\frac{|K|}{|\mathrm{P}|^{nR_L}} - 1$ tiende a cero:
  - Es decir para valores del tamaño de texto de cifrado muy grandes el número de claves espúreas es cero.
- Hay que tener en cuenta que para valores muy pequeños de n, la cantidad $\frac{H(P_p^n(x))}{n}$ podría estar mal estimada (H_L podría ser falsa), y por tanto la cota del número de claves espúreas podría ser mala.

Claves espúreas y distancia de unicidad

- Se define la distancia de unicidad para un criptosistema como el tamaño de texto cifrado mínimo, n_0, para el cual el número medio de claves espúreas se hace cero:
  - $0 \cong \overline{s_n} \geq \frac{|K|}{|\mathrm{P}|^{n_0 R_L}} - 1 \leftrightarrow |\mathrm{P}|^{n_0 R_L} = |K| \leftrightarrow n_0 R_L \log_2|\mathrm{P}| = \log_2|\mathrm{K}| \leftrightarrow \boxed{n_0 = \frac{\log_2|\mathrm{K}|}{R_L \log_2|\mathrm{P}|}}$, siempre que $|P| = |C|$ y $P_k(k) = \frac{1}{|K|}$.
- Calcular para casa la distancia de unicidad para el caso específico de $|P| \neq |C|$ y $P_k(k) = \frac{1}{|K|}$.
- Para el cálculo de la distancia de unicidad se suele utilizar $R_L = 0{,}75$ que corresponde a la redundancia del inglés.

Claves espúreas y distancia de unicidad

- Ej. Vamos a calcular cual es la distancia de unicidad para el cifrado de sustitución:
  - $P = Z_m$, $\mathrm{C} = Z_m$, y K =conjunto de permutaciones.
  - $|P| = |C| = m$, y $|K| = m!$, $R_L = 0{,}75$.
- $n_0 = \frac{\log_2|\mathrm{K}|}{R_L \log_2|\mathrm{P}|} = \frac{\log_2|26!|}{0{,}75 * 4{,}70} = \frac{88{,}3820}{0{,}75 * 4{,}70} \cong 25.$
- Este resultado nos sugiere que dado un texto de cifrado de longitud mayor de 25 caracteres, normalmente una única desencriptación es posible.
- La distancia de unicidad es la longitud de texto mínima que se necesita para romper el cifrado por la fuerza bruta sin ambigüedades debido a las claves espúreas.

Algunas referencias interesantes de este capítulo:

- C. E. Shannon, "Communication theory of secrecy systems," in The Bell System Technical Journal, vol. 28, no. 4, pp. 656-715, Oct. 1949, doi: 10.1002/j.1538-7305.1949.tb00928.x.
- Prediction and entropy of printed English C.E. Shannon January 1951.
- Libro Interesante: Teoría de la información y codificación, Abramson, Norman, Paraninfo (INF/Depósito/2092).
- Entropy and Redundancy in English. (Entropy and Redundancy in English).
- Cryptography : theory and practice. Douglas Robert Stinson. Boca Raton:CRC , © 1995.
- Cryptography : theory and practice. Douglas Robert Stinson. 2nd ed. , Boca Raton:CRC , © 2002.
- Cryptography : theory and practice. Douglas Robert Stinson. 3rd ed. , Boca Raton:CRC , © 2006.
- Cryptography : theory and practice. Douglas Robert. Stinson.-- Maura B Paterson.4th ed. , Boca Raton:CRC , © 2019.