Planificación de Redes y Modelado

3º Ingeniería de Telecomunicaciones — UPV/EHU

"Under-promise and over-deliver."

Javier de Martín – 2016/17

Definiciones

R: Cadencia $\vec{p}$: Vector conjunto de probabilidades
X: Variable aleatoria discreta
x_i: Valor de la v.a.
$p(x)$: Probabilidad marginal
$p(x,y)$: Probabilidad conjunta
$p(x/y)$: Probabilidad condicional
$H(X)$: Entropía
$H(X,Y)$: Entropía conjunta
$H(X/Y)$: Entropía condicionada

Tema 1.- Teoría de la Información

Introducción a la Teoría de la Información - Técnicas de Codificación

Cadencia: Por cada k símbolos que entran al codificador llegan n al sistema.

$$R = \frac{k}{n}$$

Probabilidad de error de bit:

$$P_e = \frac{1}{k}\sum_{i=1}^{k} P_e^{(i)}$$

Función Entropía: Representa la cantidad de información proporcionada al conocer X, la medida de incertidumbre o ignorancia de X o la aleatoridad o dispersión de X.

$$H_n(X) = \sum_{i\geq 1} p_i log_n \frac{1}{p_i}$$

Dependiendo del valor de n, la entropía se mide en unidades diferentes:

- $n = 2 \rightarrow$ bits
- $n = e \rightarrow$ nats
- $n = 10 \rightarrow$ Hartleys

Se puede cambiar de base, que significa un cambio de escala:

$$\log_b x = \log_b a \cdot \log_a x$$

La **interpretación de la entropía**:

- El hecho más improbable es el que proporciona mayor información.
- $I(X)$ se puede interpretar como la cantidad de información proporcionada por el suceso $\{X = x\}$.
- $H(X)$ será el valor medio de dicha cantidad de información $I(X)$.

Codificación de Fuentes de Longitud Variable

Códigos de Longitud Variable Únicamente Decodificables

- **Ristra σ de longitud k**: Secuencia ordenada de k elementos de S.
- **Operación concatenación**: Dadas dos ristras, $\sigma = s_1, s_2, ..., s_k$ y $\tau = t_1, t_2, ..., t_l$, sobre S, $\sigma * \tau = s_1, s_2, ..., s_k, t_1, t_2, ..., t_l$.
- Si $\sigma = \sigma_1 * \sigma_2 * \sigma_3$: $\rightarrow$ σ_1 es prefijo, σ_2 es subristra y σ_3 es sufijo.
- **Longitud de una ristra ($|\sigma|$)**: $|\sigma * \tau| = |\sigma| + |\tau|$.

Un **código** es **unívocamente decodificable (UD)** si cualquier ristra c^k tiene una única descomposición posible como concatenación de palabras código.
Un **código** c se denomina **prefijo** si ninguna palabra código es prefijo de otra.

Si un código es prefijo $\rightarrow$ es unívocamente decodificable

Teorema de McMillan: Si un código $c = \{\sigma_0, \sigma_1, ..., \sigma_{r-1}\}$ es un código UD sobre $S = \{0, 1, ..., s-1\}$ entonces:

$$\sum_{i=0}^{r-1} \frac{1}{s^{n_i}} \leq 1/n_i = |\sigma_i|$$

Teorema de Kraft: Dadas r longitudes n_i para $i = 0, 1, ..., r-1$ que cumplan $\sum_{i=0}^{r-1} \frac{1}{s^{n_i}} \leq 1$, existirá algún código prefijo y UD con esas longitudes.

Método para generar ese código prefijo y UD

Reordenar los $n_i \rightarrow n_0 \leq n_1 \leq ... \leq n_{r-1}$ y definir los r enteros w_j:

$$w_0 = 0 \text{ y } w_j = \sum_{i=0}^{j-1} s^{n_j - n_i} \rightarrow w_j = s^{n_j}\sum_{i=0}^{j-1} s^{-n_i} \leq s^{n_j} - 1 \;\; \forall j$$

Cada w_j se codifica con n_j símbolos s-arios. Cada σ_j es la representación s-aria del entero w_j (se rellena con ceros hasta n_j).

- Sea una fuente discreta sin memoria de alfabeto fuente $A = \{0, 1, ..., s-1\}$ definida por $\vec{p} = \{p_0, p_1, ..., p_{r-1}\}$ y definamos un código s-ario de la fuente $\{\sigma_0, \sigma_1, ..., \sigma_{r-1}\}$ sobre $S = \{0, 1, ..., s-1\}$ de tal forma que si queremos transmitir el símbolo fuente i transmitimos σ_i. Si queremos transmitir k símbolos fuente $i_1, i_2, ..., i_k$ seguidos, enviaremos la ristra $\sigma = \sigma_{i1} * \sigma_{i2} * \sigma_{ik}$ y para asegurarnos que el receptor pueda recuperar la cadena de símbolos fuente $i_1, i_2, ...i_k$ usaremos un código UD.
- La **longitud media de un código** $\{\sigma_0, \sigma_1, ..., \sigma_{r-1}\}$ respecto a la fuente $\{p_0, p_1, ..., p_{r-1}\}$ se define como:

$$n = \sum_{i=0}^{r-1} p_i \cdot |\sigma_i|$$

- **Interpretación**: n es una medida de la **carga media del canal** dada por el uso del código,
- Llamaremos $n_s(\vec{p})$ a la **longitud media** de un código s-ario para $\vec{p}$.
- **Límite inferior de la longitud media** n: Para un código c que sea UD, su longitud media **no** es inferior a la entropía s-aria de la fuente.

$$n \geq H_s(\vec{p})$$

Algoritmo de Huffmann

Permite generar un código prefijo, que también será UD.

1. Ordenar las probabilidades de los símbolos de mayor a menor.
2. Reemplazar el vector de probabilidades $\vec{p}$ por $\vec{p'}$ donde se combinan las tres probabilidades más pequeñas.
3. Reemplazar el vector de probabilidades de forma recursiva agrupando s símbolos de menor probabilidad en uno sólo hasta conseguir una fuente reducida de s símbolos.
4. Se codifica cada fuente reducida de forma óptima empezando por la más pequeña hasta llegar a la fuente original.

Con esto se crea el **código óptimo** para un conjunto de símbolos y probabilidades. Hay que tener la **determinación del tamaño** s' de la primera reducción. El valor de s' está definido por dos condiciones:

$$s' \in \{2, 3, ..., s\} \text{ y } s' \equiv r \mod(s-1)$$

Adaptación de Códigos a las Fuentes: Extensión de Fuentes

Sea una fuente definida por $\vec{p}$ sobre el alfabeto A, son necesarios $H_s(\vec{p})$ para representar fielmente la fuente. En cambio, si se utiliza un código UD s-ario se necesitarán, en general, más símbolos.

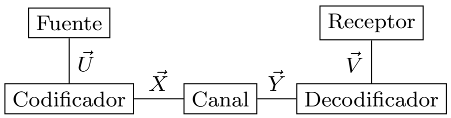
$$n_s(\vec{p}) \geq H_s(\vec{p})$$

Para solucionar esto se extiende la fuente considerando la fuente como secuencia de bloques consecutivos de m símbolos sobre un alfabeto A^m. Tomando un valor de m suficientemente grande, se puede aproximar todo lo que se quiera el número medio de símbolos necesarios para representar la fuente. Sea $\vec{p}^m$ la distribución de probabilidades de la extensión y $n_s(\vec{p}^m)$ la longitud media de un código de dicha extensión:

$$\lim_{m\to\infty} \frac{1}{m} n_s(\vec{p}^m) = H_s(\vec{p})$$

Interpretación: Una fuente $\vec{p}$ puede ser representada fielmente utilizando $H_s(\vec{p})$ símbolos s-arios por símbolo fuente.

Modelado del Proceso de Información



- $\vec{U}$: k salidas consecutivas de la fuente.
- $\vec{X}$: n-tupla que utiliza el codificador para codificar en el canal la información transmitida.
- $\vec{Y}$: Versión ruidosa de n salidas recibidas de $\vec{X}$.
- $\vec{V}$: k-tupla que reproduce de forma estimada la fuente.

En los sistemas de comunicaciones realizables las salidas de cada bloque no dependen de las entradas anteriores, sino que son funciones de la entrada en cada momento. Viene definido por las probabilidades condicionales:

$$p(\vec{y}/\vec{x},\vec{u}) = p(\vec{y},\vec{x}) \quad , \quad p(\vec{v}/\vec{y},\vec{x}) = p(\vec{v}/\vec{y})$$

Teorema del Proceso de Información

"*El procesado de información sólo puede destruir información.*"

$$\left.\begin{array}{l} I(\vec{U},\vec{V}) \leq I(\vec{X},\vec{V}) \\ I(\vec{X},\vec{V}) \leq I(\vec{X},\vec{Y}) \end{array}\right\} \rightarrow I(\vec{U},\vec{V}) \leq I(\vec{X},\vec{Y})$$

Canal Discreto Sin Memoria (*DMC*)

- Discreto: Número de E/S finito.
- Sin Memoria: La salida sólo depende de la entrada en ese dispositivo y no de entradas de dispositivos anteriores.
- Aleatorio: La salida no es una función directa de la entrada sino que está regulado por una matriz de probabilidades.

Se define una matriz de probabilidades $p(y/x)$ de dimensión (r,s) que representa la probabilidad de que se produzca una salida y habiéndose producido una entrada x.

$$p(y_j/x_i) \geq 0 \forall (i,j)$$
$$\sum_j p(y_j/x_i) = 1 \forall i$$

$$Q_{(r,s)} = \begin{pmatrix} p(y_0/x_0) & p(y_1/x_0) & ... & p(y_s/x_0) \\ p(y_0/x_1) & p(y_1/x_1) & ... & p(y_s/x_0) \\ ... & ... & ... & ... \\ p(y_0/x_r) & p(y_1/x_r) & ... & p(y_s/x_r) \end{pmatrix}$$

Modelización de un Canal DMC:

- Se define una variable aleatoria X de rango $R = \{0,1,...,r-1\}$ que modeliza la entrada de un canal DMC según una distribución de probabilidades $Pr\{X = x\} = p(x)$
- Se define también una variable aleatoria Y de rango $S = \{0,1,...,s-1\}$ que representa la salida del canal DMC.

Distribución conjunta: $p(x,y) = p(x) \cdot p(y/x)$ Distribución marginal: $p(y) = \sum_x$

$H(X)$ es la incertidumbre que se tiene de X antes de conocer a Y. Una vez conocida y, la entropía remanente de X es:

$$H(X/y) = \sum_x p(x/y) \cdot \log \frac{1}{p(x/y)}$$

La **entropía condicionada de** X/Y será la esperanza de esta variable aleatoria:

$$H(X/Y) = \sum_y p(y) \cdot H(X/Y = y) = E_y\{H(X/Y = y)\}$$

$H(X/Y)$ representa la cantidad de incertidumbre remanente de X tras conocer Y.

Entropía de Sucesos con Dos Variables Aleatorias: Sea un suceso definido por dos variables aleatorias X w Y con rangos $R_X = \{x_1, x_2, ..., x_r\}$ y $R_Y = \{y_1, y_2, ..., y_s\}$ y con una distribución de probabilidades: **FALTA EQ**

Se define la **entropía conjunta** de X e Y como:

$$H(X,Y) = \sum_i \sum_j p(x_i,y_j) \cdot \log \frac{1}{p(x_i,y_j)} = \sum_x \sum_y p(x,y) \cdot \log \frac{1}{p(x,y)}$$

Si las variables X e Y son independientes:

$p(x,y) = p(x) \cdot p(y) \quad \forall x,y \rightarrow H(X,Y) = H(X) + H(Y)$

La entropía conjunta de dos variables aleatorias independientes es la suma de entropías.

Si las variables son dependientes la suma total será menor que la suma de entropías.

$H(X_1, X_2, ...) \leq H(X_1) + H(X_2) + ...$

Probabilidad Conjunta:

$$p(x,y) = Pr\{X = x, Y = y\}$$

Probabilidades Marginales:

$$p(x) = Pr\{X = x\} = \sum_y p(x,y)$$
$$p(y) = Pr\{Y = y\} = \sum_x p(x,y)$$

Probabilidades Condicionales:

$$p(x/y) = Pr\{X = x/Y = y\} = \frac{p(x,y)}{p(y)}$$
$$p(y/x) = Pr\{Y = y/X = x\} = \frac{p(x,y)}{p(x)}$$

Entropía condicional de X dada por Y: **FALTA EQ**

La entropía condicionada cumple las siguientes condiciones:

$\rightarrow H(X,Y) = H(X) + H(Y/X) = H(Y) + H(X/Y)$
$\rightarrow H(Y/X) \leq H(Y)$
$\rightarrow$ Inecuación de Fano **TERMINAR**:
$H(X/Y) \leq H(P_e) + P_e \cdot log(r-1)$

Siendo X e Y dos variables aleatorias de igual rango de valor r y $P_e = Pr\{X \neq Y\}$

Información Mutua: Si $H(X)$ es la cantidad de incertidumbre a priori de X y $H(X/Y)$ es la incertidumbre de X una vez conocido Y, se puede entender $I(X;Y)$ como la información que Y suministra sobre X.

$$I(X;Y) = I(Y;X) = H(X) + H(Y) - H(X/Y)$$
$$I(X;Y) \geq 0$$

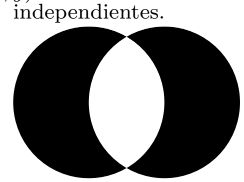
Se define la variable aleatoria $I(x;y)$ como la información mutua de entre los elementos x e y.

$$I(X = x; Y = y) = I(x;y) = \log \frac{p(x/y)}{p(y)} = \log \frac{p(y/x)}{p(y)}$$

Si esta v.a. proporciona valores positivos $\rightarrow$ proporciona información

Si esta v.a. proporciona valores negativos $\rightarrow$ confunde

$I(X;Y)$ es una medida de la dependencia estadística entre las dos variables, será cero cuando las variables aleatorias sean independientes.



TERMINAR

Codificación de Canal

Teorema de la Codificación del Canal

$$r = \frac{k}{n} \leq \frac{C(\beta)}{1 - H_2(\epsilon)}$$

La **cadencia del sistema** ($\frac{k}{n}$) representa el número de bits por uso del canal transmitidos por el sistema de comunicaciones. Tiene un **límite superior** que es función creciente de ϵ.

"*Cuanto más fiablemente se quiera comunicar más lentamente hay que realizarlo*"