Introducción a la Teoría de Códigos

M.A. García, L. Martínez, T. Ramírez

Facultad de Ciencia y Tecnología. UPV/EHU

Resumen Teórico
Apartado 3 del Tema 2:
Distancia de Hamming

Mayo de 2017

3 Distancia de Hamming

Hemos introducido la distancia de Hamming que nos mide en cuántas componentes difieren dos palabras de la misma longitud. El utilizar el nombre de "distancia" está justificado desde el punto de vista topológico, ya que es fácil demostrar que:

Proposición 3.1 *Sea $A = \{a_1, \ldots, a_m\}$ un alfabeto, T_n el conjunto de las palabras sobre el alfabeto A de longitud n, esto es, $T_n = \{x_1 \ldots x_n \,|\, x_i \in A, i = 1, \ldots, n\}$ y $d : T_n \times T_n \to \{0, 1, 2, \ldots, n\}$ la aplicación definida por*

$$d(x_1 \ldots x_n, y_1 \ldots y_n) = |\{i \in \{1, \ldots, n\} \,|\, x_i \neq y_i\}|,$$

para todo $x_1 \ldots x_n, y_1 \ldots y_n \in T_n$. Entonces, (T_n, d) es un espacio métrico.

Al ser (T_n, d) un espacio métrico tiene sentido construir sus bolas cerradas. Recordemos que, dado $r \in \mathbb{N} \cup \{0\}$ y $\mathbf{x} \in T_n$, se llama **bola cerrada de centro x y radio** r al conjunto

$$\overline{B}(\mathbf{x}, r) = \{\mathbf{y} \in A^n \,|\, \mathrm{d}(\mathbf{x}, \mathbf{y}) \leq r\}.$$

Obviamente, si $r \geq n$, $\overline{B}(\mathbf{x}, r) = T_n$. Más aún, es sencillo determinar el número de elementos de una bola cerrada cuando $r = 0, 1, \ldots, n$, tal y como nos lo indica el siguiente resultado:

Lema 3.2 *Sea $A = \{a_1, \ldots, a_m\}$ un alfabeto, $\boldsymbol{x} = x_1 \ldots x_n \in T_n$ y $0 \leq r \leq n$. Entonces,*

$$|\overline{B}(\boldsymbol{x}, r)| = \sum_{i=0}^{r} \binom{n}{i} (m-1)^i.$$

El Lema anterior va a ser útil para demostrar la conocida Cota de Hamming:

Proposición 3.3 (Cota de Hamming) *Sea $C \subseteq T_n$ un código de longitud n sobre el alfabeto $A = \{a_1, \ldots, a_m\}$ con distancia mínima d. Entonces,*

$$|C| \sum_{i=0}^{\left[\frac{d-1}{2}\right]} \binom{n}{i} (m-1)^i \leq m^n.$$

Para probarlo simplemente hay que darse cuenta que

$$\overline{B}(\mathbf{c}, \left[\frac{d-1}{2}\right]) \cap \overline{B}(\mathbf{c}', \left[\frac{d-1}{2}\right]) = \emptyset, \quad \forall \mathbf{c}, \mathbf{c}' \in C \text{ tales que } \mathbf{c} \neq \mathbf{c}'$$

y usar el Lema 3.2.

Relacionamos ahora el número de errores que puede detectar y corregir un código de bloque de longitud n con su distancia mínima. Se puede demostrar que:

CC BY NC SA

Proposición 3.4 *Sea $C \subseteq T_n$ un código de longitud n sobre el alfabeto A con distancia mínima d. Entonces*

1. *C puede detectar hasta $d-1$ errores.*
2. *C puede corregir hasta $\left[\frac{d-1}{2}\right]$ errores.*

Ejemplo Consideramos el código $C \subseteq \mathbb{F}_2^7$ cuyas palabras son

$$\begin{aligned} C \ = \{&0000000, 0110100, 0011010, 0001101, \\ &1000110, 1001011, 1011100, 0010111, \\ &1010001, 1110010, 0111001, 1111111, \\ &0101110, 1101000, 0100011, 1100101\}. \end{aligned}$$

Este código C tiene distancia mínima 3, ya que haciendo las distancias de pares de palabras distintas de C, se observa que el valor mínimo es este valor. Si le aplicamos ahora la Proposición 3.4, deducimos que puede detectar hasta dos errores y corregir uno. Por ejemplo, si recibimos la palabra $\mathbf{x} = 1000111$ vemos que no está en el código, por tanto se ha(n) producido error(es) en la transmisión. Pero nos fijamos que 1000110 sí pertenece al código y difiere solo en un dígito con $\mathbf{x}$, luego podemos deducir que es la que se ha emitido cuando hemos recibido $\mathbf{x}$, ya que este código corrije un error. De hecho, si calculamos la distancia que hay entre $\mathbf{x}$ y las diferentes palabras de C, vemos que la única palabra de C que dista 1 de $\mathbf{x}$ es 1000110. Más aún, C alcanza la cota de Hamming ya que

$$|C| \sum_{i=0}^{\left[\frac{d-1}{2}\right]} \binom{n}{i} (m-1)^i = 2^4 \sum_{i=0}^{1} \binom{7}{i} 1^1 = 2^4 2^3 = 2^7.$$

Estudiamos los códigos que alcanzan la Cota de Hamming en el siguiente apartado y veremos qué significa alcanzar esta cota en términos de poder determinar la palabra emitida cuando recibimos una palabra de $\mathbb{F}_2^7$.

CC BY NC SA