Distancia de Hamming

En teoría de la información se denomina **distancia de Hamming** a la efectividad de los códigos de bloque y depende de la diferencia entre una palabra de código válida y otra. Cuanto mayor sea esta diferencia, menor es la posibilidad de que un código válido se transforme en otro código válido por una serie de errores. A esta diferencia se le llama distancia de Hamming, y se define como el número de bits que tienen que cambiarse para transformar una palabra de código válida en otra palabra de código válida.
Si dos palabras de código difieren en una distancia **d**, se necesitan **d** errores para convertir una en la otra.

Por ejemplo:

- La distancia Hamming entre **1011101** y **1001001** es 2.
- La distancia Hamming entre **2143896** y **2233796** es 3.
- La distancia Hamming entre "**tener**" y "**reses**" es 3.

Detección y corrección de errores

La distancia de Hamming es utilizada para definir algunas nociones esenciales en teoría de códigos, tales como **códigos detectores de errores** y **códigos correctores de errores**. En particular, se dice que un código C detecta k-errores si dos palabras cualesquiera $c_1, c_2 \in C$ que tienen una distancia de Hamming menor que k coinciden. Dicho de otro modo, un código detecta k-errores si y solo si la distancia de Hamming mínima entre dos palabras cualesquiera en él es a lo menos $k+1$.

Se dice que un código C corrige k-errores si para cada palabra w en el subyacente espacio de Hamming H existe al menos una palabra $c \in C$ tal que la distancia de Hamming entre w y c es menos que k. En otras palabras, un código corrige k-errores si y solo si la mínima distancia de Hamming entre dos cualesquiera de sus palabras es por lo menos $2k+1$. Esto es más fácil de comprender geométricamente como que dos bolas cerradas cualesquiera de radio k centradas en distintas palabras son disjuntas. En este contexto se conoce a estas bolas como **esferas de Hamming**.

De esta manera, un código que tiene distancia de Hamming mínima d entre sus palabras puede detectar a lo más $d-1$ errores y puede corregir $\lfloor (d-1)/2 \rfloor$ errores. Este último número es también conocido como el **radio de empaquetado** o la **capacidad de corrección** del código.

Historia y aplicaciones

La distancia de Hamming se denomina así gracias a su inventor Richard Hamming, profesor de la Universidad de Nebraska, que fue el que introdujo el término para establecer una métrica capaz de establecer un *código para la detección y auto-corrección de códigos*. Se emplea en la transmisión de información digitalizada para contar el número de desvíos en cadenas de igual longitud y estimar el error, por esto se denomina a veces como *distancia de señal*.

La distancia de Hamming tiene las siguientes propiedades.

- $d(a,b) = d(b,a)$
- $d(a,b) = 0$ si y solo si $a = b$
- $d(a,b) + d(b,c) \geq d(a,c)$

d es el n.º de bits *p* en que son diferentes el mensaje emitido del recibido.

Si $d \geq p+1$ entonces se puede **detectar** un error de peso *p*

Si $d \geq 2p+1$ entonces se puede **corregir** *p* dígitos.

Ejemplo: Si queremos detectar 3 errores entonces la distancia mínima de Hamming debe ser de $(3)+1=4$. Si queremos corregir 3 errores entonces la distancia mínima de Hamming debe ser de $2*(3)+1=7$.

Véase también

- Código Hamming
- Distancia de Levenshtein

Datos: Q272172