Teoria dell'Informazione

Simone Alessandro Casciaro

05 Dicembre 2024

Lezione 13: Parità

Rumore di un Canale

In questo corso ci focalizziamo sul canale con Rumore Bianco, cioè un canale dove gli errori in posizioni differenti sono indipendenti.
La probabilità che accada un errore è fissata a p per ciascun bit e rimane costante nel tempo.
Siano n il numero di bit trasmessi e p la probabilità d'errore di un singolo bit, la probabilità di mandare un messaggio senza errori è $(1-p)^n$. In generale, la probabilità di commettere i errori è definita come:

$$\mathbb{P}(X = i) = \binom{n}{i} p^i (i - p)^{n-i}$$

ovvero una variabile aleatoria binomiale di parametri n, p.

Bit di Parità

Definiamo il bit di parità come

$$x_n = \sum_{i=1}^{n-1} x_i \mod 2$$

Il bit di parità è utile perché, ricevuto un messaggio y, posso controllare il bit di parità:

$$\sum_{i=1}^{n} y_i \mod 2 = \begin{cases} 0 & \text{Errore non rilevato} \\ 1 & \text{Errore rilevato} \end{cases}$$

È importante sottolineare che avere somma uguale a 0 non significa che non è presente un errore, ma solo che non viene rilevato. Infatti, il bit di parità non rileva errori se si commette un numero pari di errore e questo accade con probabilità $\sum_{l=1}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{2l} p^{2l}(1-p)^{n-2l}$

Codice ASCII

Il codice ASCII fa uso del bit di parità. Si converte un carattere in 7 bit e si aggiunge il bit di parità in testa.
Ad esempio, la lettera K dell'alfabeto è la lettera numero 75 in decimale, che convertita in binario equivale a 1001011. Dato che i bit a 1 sono pari, il bit di parità sarà 0.

Come correggiamo gli errori?

Supponiamo di avere il messaggio "Hello NCTU". Aggiungiamo un carattere di parità per tutto il messaggio, che si calcola come il bit di parità ma in verticale.

H	0	1	0	0	1	0	0	0
e	0	1	1	0	0	1	0	1
l	0	1	1	0	1	1	0	0
l	0	1	1	0	1	1	0	0
0	0	1	1	0	1	1	1	1
	0	0	1	0	0	0	0	0
N	0	1	0	0	1	1	1	0
C	0	1	0	0	0	0	1	1
T	0	1	0	1	0	1	0	0
U	0	1	0	1	0	1	0	1
□	0	1	1	0	1	1	1	0

Convertendo 01101110 otteniamo la lettera n, dunque questo è il nostro carattere di parità. In caso di errore, il carattere di parità risulta errato e si può, in alcuni casi, correggere l'errore.

Somma Pesata

Supponiamo di avere un messaggio x di lunghezza n caratteri e chiamiamo x_i tutti i caratteri appartenenti al messaggio $\forall i = 1, \ldots, n$
Definiamo la somma pesata come

$$\sum_{i=1}^{n}(n-i+1)x_i \mod |S|$$

Definiamo lo spazio dei caratteri S come una enumerazione dei caratteri che possono appartenere a un messaggio. Ad esempio, nel caso di un messaggio alfanumerico, i caratteri possibili sono: $0 \ldots 9A \ldots Z\_$, cioè 37 caratteri.

> Esempio
> $x = 3B\_8$ più il carattere di controllo $\square$
> Il carattere più a sinistra, in questo caso il 3. Al carattere B attribuiamo il valore 11 e allo spazio il valore 36.
> Applicando la formula, otteniamo $4 \cdot 3 + 3 \cdot 11 + 2 \cdot 36 + 1 \cdot 8 = 183$
> Vogliamo che $183 + \square = 0 \mod 37$, dunque $\square = 2$
>
> Supponiamo un errore, ad esempio la perdita del carattere spazio.
> Riapplicando la formula si ottiene $3 \cdot 3 + 2 \cdot 11 + 1 \cdot 8 = 61$
> Dato che $61 + \square = 0 \mod 37$, dunque $\square = 13 \neq 2$. L'errore viene rilevato.

Codice ISBN

Il codice ISBN esiste in due versioni diverse, la prima da 10 cifre e la seconda da 13.
Il codice ISBN usa la somma pesata, usando solo caratteri numerici e il modulo 11 (si preferisce avere un numero primo per lavorare su un campo).

Codice UPC (codici a barre)

Anche il codice UPC segue la somma pesata, anche se in maniera diversa.
L'ultima cifra funge da check digit e viene calcolato tramite questo calcolo:

$$3 \cdot (x_1 + x_3 + x_5 + x_7 + x_9 + x_{11}) + (x_2 + x_4 + x_6 + x_8 + x_{10}) + \square = 0 \mod 10$$