Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, v. 11, n. 1, 2025.

Trabalho apresentado no XLIII CNMAC, Centro de Convenções do Armação Resort - Porto de Galinhas - PE, 2024

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Introdução aos Códigos de Hamming

José A. P. Nogueira[1]
URCA, Juazeiro do Norte, CE
Clarice D. Albuquerque[2]
UFCA, Juazeiro do Norte, CE

Os códigos de Hamming são uma classe específica de códigos corretores de erros, que fazem parte dos códigos lineares e são estudados na Teoria da Informação. Eles foram propostos pelo matemático Richard Hamming na década de 1950. Esses códigos são projetados para detecção e correção de erros em transmissões de dados.

A ideia básica dos códigos de Hamming é adicionar bits de paridade aos dados originais, de modo que a presença de erros possa ser detectada e, em alguns casos, corrigida.

Definição. Os Códigos de Hamming de ordem m são construídos sobre o corpo finito $\mathbb{F}_2 = \{0, 1\}$ e determinados pela matriz teste de paridade H_m, onde as colunas desta matriz são todos os vetores não nulos de $\mathbb{F}_2^m$.

Seja C um Código de Hamming de ordem m, então o comprimento de C é $n = 2^m - 1$ e a dimensão é $k = n - m = 2^m - m - 1$. Assim, dizemos que um código de Hamming é um código linear binário $C(2^m - 1, 2^m - m - 1)$.

Um Código de Hamming $C(2^m - 1, 2^m - m - 1)$ é obtido por uma matriz teste de paridade H_m, na qual associamos a matriz geradora G_m. Dessa forma, para realizar a codificação de uma palavra $c \in \mathbb{F}_2^{2^m-m-1}$ que será transmitida por um canal, basta efetuar o seu produto com a matriz geradora G_m, resultando em $c \cdot G_m \in \mathbb{F}_2^{2^m-1}$. Observamos que depois da codificação serão adicionadas m letras em c, chamadas de redundância, que têm a função de detectar e corrigir os possíveis erros apresentados durante o envio da palavra.

Para realizar a decodificação de uma palavra recebida r, procedemos como descrito abaixo.

Consideremos a palavra r como uma matriz linha de ordem $1 \times (2^m - 1)$. Dada uma matriz teste de paridade H_m, sua ordem é $m \times (2^m - 1)$. Dessa forma, ao calcularmos o produto $H_m \cdot r^t$, o resultado obtido é uma matriz coluna de ordem $m \times 1$.

Quando não houver erro durante o processo de envio da palavra, esta matriz resultante é a matriz coluna nula. No entanto, quando acontece um erro na transmissão, a matriz resultante será uma matriz h_i, $i = 1, ..., 2^m - 1$, que é uma das colunas da matriz teste de paridade H_m. Assim, saberemos que na palavra transmitida o erro aconteceu na coordenada r_i, $i = 1, ..., 2^m - 1$, e dessa forma podemos corrigí-lo e encontrar a codificação exata da palavra c transmitida.

Após determinar r' podemos identificar qual foi a palavra transmitida, uma vez que os primeiros $2^m - m - 1$ dígitos da codificação representam a palavra enviada.

Existem diferentes versões dos códigos de Hamming, mas um dos mais conhecidos é o Código de Hamming (7,4), que é projetado para corrigir um único erro e detectar até dois erros em um bloco de 7 bits de dados. Para entender esse processo vejamos o exemplo a seguir, o qual foi baseado em [2].

[1] augusto.nogueira@urca.br
[2] clarice.albuquerque@ufca.edu.br

Exemplo. Seja $C(2^m - 1, 2^m - m - 1)$, o código de Hamming com $m = 3$, isto é, $C(7, 4)$. Consideremos uma matriz teste de paridade na forma padrão dada por

$$H_3 = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

Suponhamos que recebemos a palavra $r = 1100101$. Assim

$$H_3 \cdot r^t = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1+1+0+0+1+0+0 \\ 1+0+0+0+0+0+0 \\ 0+1+0+0+0+0+1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

Note que a matriz coluna resultante é igual a coluna h_1 de H_3, isto é, a palavra r sofreu um erro na coordenada r_1, portanto a palavra correta é $r' = 0100101$.

Para determinar a palavra enviada, lembramos que em $C(7, 4)$, os quatro primeiros dígitos representam a palavra enviada, e os três últimos são redundâncias, que servem justamente para detecção e correção de erros. Sendo assim, a palavra transmitida foi $c = 0100$.

Os Códigos de Hamming são aplicados em uma variedade de contextos onde a detecção e correção de erros são cruciais para a integridade dos dados. como por exemplo em memórias RAM, comunicações digitais, armazenamento de dados, sistemas de comunicação por satélite, sistemas de transmissão de vídeo e áudio, entre outros.

Essencialmente, o Código de Hamming é aplicado em qualquer cenário em que a precisão dos dados seja crítica e a ocorrência de erros seja uma preocupação, proporcionando uma forma eficaz de detecção e correção de erros.

Vale ressaltar que o Código de Hamming, em sua forma mais simples (como o (7,4)), é projetado para detectar e corrigir um único erro. Isso significa que, se houver apenas um bit incorreto no bloco de dados, o código pode identificar a posição desse bit e corrigi-lo. No entanto, o Código de Hamming não foi originalmente concebido para corrigir mais de um erro em um único bloco de dados. Quando ocorrem mais de um erro, a capacidade de correção do código pode ser excedida, e a detecção torna-se possível, mas a correção não é garantida.

Para lidar com múltiplos erros, podem ser utilizados códigos mais avançados, como os códigos BCH (Bose-Chaudhuri-Hocquenghem) ou códigos Reed-Solomon. Esses códigos são capazes de detectar e corrigir um número maior de erros em comparação com o Código de Hamming.

Para maiores informações sobre códigos lineares, no qual o Código de Hamming faz parte, indicamos as referências [2], [3] e [1]. Estas duas últimas trazem conceitos mais aprofundados sobre os códigos citados no parágrafo anterior, os quais serão foco para estudos subsequentes.

Referências

[1] F. J. MacWilliam e N. J. A. Sloane. **The theory of error correcting codes**. Elsevier, 1977.

[2] J. A. P. Nogueira. "Aplicações Matemáticas em Códigos Corretores de Erros". Dissertação de mestrado. UFCA, 2019.

[3] M. L. T. Villela e A. Hefez. **Códigos Corretores de Erros**. 2a. ed. Rio de Janeiro: IMPA, 2008.