Introdução aos códigos corretores de erros e aplicações

Introduction to error correcting codes and applications

Belatrice da Rocha Bolzani[1], Erika Patricia Dantas de Oliveira Guazzi[2]

RESUMO

Os códigos corretores de erros constituem um campo essencial no âmbito da tecnologia da informação e da comunicação. Este trabalho tem o objetivo de estudar sobre esses códigos em blocos, explorando seus fundamentos, aplicações e a importância na sociedade atual. Os códigos lineares representam uma classe importante de códigos corretores de erros. Eles utilizam conceitos da álgebra linear, como matrizes geradoras e de verificação de paridade, para detectar e corrigir erros, também conhecidos como ruídos, que são inerentes à transmissão de informações em sistemas de comunicação digital. As matrizes geradoras desempenham um papel fundamental na criação dos códigos lineares. Elas convertem uma sequência de bits de dados originais em uma sequência de bits codificados. As matrizes de verificação de paridade, por sua vez, são responsáveis por verificar se a sequência de bits codificados contém erros, assegurando a integridade dos dados. São apresentadas algumas pesquisas que apresentam aplicações, a fim de destacar a importância desses códigos na atualidade pois a sua utilização abrange diversas áreas, tornando-os essenciais para atender as necessidades de comunicação confiável em um mundo cada vez mais interconectado.

PALAVRAS-CHAVE: código corretor de erro; comunicação; sistema de informação.

ABSTRACT

Error correcting codes constitute an essential field within information and communication technology. This work aims to study these block codes, exploring their foundations, applications and importance in today's society. Linear codes represent an important class of error correcting codes. They use concepts from linear algebra, such as generating and parity checking matrices, to detect and correct errors, also known as noise, that are inherent to the transmission of information in digital communication systems. Generating matrices play a fundamental role in creating linear codes. They convert a sequence of original data bits into a sequence of encoded bits. The parity check matrices, in turn, are responsible for checking whether the encoded bit sequence contains errors, ensuring data integrity. Some research is presented that presents applications, in order to highlight the importance of these codes today as their use covers several areas, making them essential to meet the needs of reliable communication in an increasingly interconnected world

KEYWORDS: error corrector code; communication; information system.

INTRODUÇÃO

Neste trabalho, uma breve introdução à Teoria dos Códigos é apresentada, com ênfase especial nos códigos corretores de erros, particularmente os códigos de blocos lineares. A importância desses códigos em várias áreas do conhecimento será destacada, juntamente com a exploração de algumas de suas aplicações. Esse enfoque é relevante em um mundo cada vez mais interconectado, onde a transmissão confiável de dados desempenha um papel fundamental e necessário em diversas esferas da sociedade contemporânea. Nessa direção, ressalta-se que os códigos corretores de erros

[1] Bolsista da Fundação Araucária. Universidade Tecnológica Federal do Paraná, Campo Mourão, Paraná, Brasil. E-mail: bolzanibelatrice24@gmail.com. ID Lattes: https://lattes.cnpq.br/8706149475423457
[2] Docente no Departamento Acadêmico de Matemática – campus Campo Mourão. Universidade Tecnológica Federal do Paraná, Campo Mourão, Paraná, Brasil. E-mail: erikapatricia@utfpr.edu.br. ID Lattes: 0006548791243526.

desempenham um papel essencial na garantia e precisão da transmissão de dados, uma vez que a interferência de erros é inevitável no processo de comunicação. Isso levanta a questão de como lidar com esses erros, dando origem à teoria dos códigos corretores de erros, cujo objetivo principal é abordar esses problemas e assegurar a confiabilidade dos dados transmitidos. Essa teoria teve seu início na década de 1940, quando os computadores eram maquinas de alto custo para serem mantidas e restrito a grandes instituições.

Um marco relevante nessa teoria é o trabalho de Claude Shannon, intitulado *"A Mathematical Theory of Communication"*, publicado em 1948, no qual ele explorou diferente abordagens para codificar informações de forma mais eficiente, (SHANNON, 1948).

De maneira geral, os códigos corretores de erros buscam recuperar a informação caso elas sejam afetadas por ruídos durante o processo de transmissão, garantindo que uma mensagem seja transmitida ou armazenada de maneira confiável.

METODOLOGIA

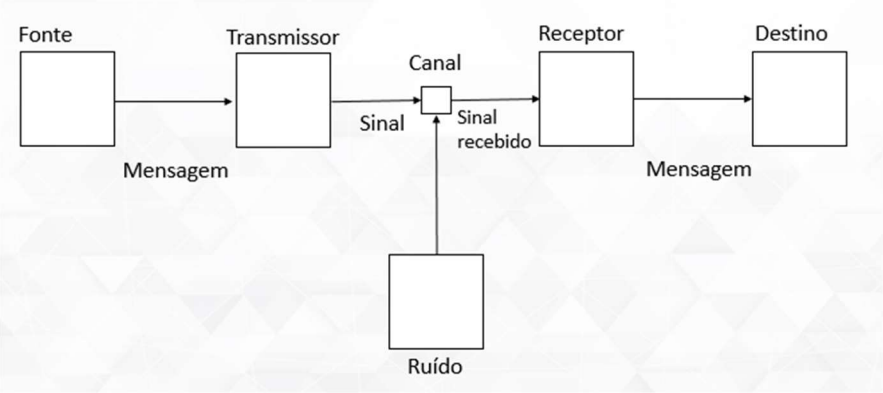
O presente trabalho consiste em estudo teórico sobre os códigos corretores de erros por meio do uso de diversas referências bibliográficas e, a partir disso, pesquisar sobre aplicações e a relevância dos códigos em diversas áreas. Em outras palavras, realizou-se revisão da literatura relacionada a Teoria do Códigos por meio de livros e artigos.

RESULTADOS E DISCUSSÕES

ASPECTOS TEÓRICOS DOS CÓDIGOS

Inicialmente, é essencial compreender alguns conceitos elementares relacionados ao sistema de comunicação. De forma geral, um sistema de comunicação estabelece uma conexão entre uma fonte e um destinatário por meio de um canal, veja a Figura 1.

Figura 1 – Esquema de transmissão de sinais



Fonte: Elaborada pelos autores (2023).

Um código é definido como um conjunto de símbolos empregados na transmissão e recepção de mensagens. Por sua vez, o alfabeto é um conjunto finito de elementos

representado por $F_q = \{0,1,2,\ldots,q\}$. Um código *q-ário* é formado por um conjunto $C = \{c_1, c_2, \ldots, c_m\}$, em que cada elemento c_i é chamado de palavra código, no qual cada elemento c_i é composto por uma sequência de símbolos do alfabeto F_q. Mais especificamente, a palavra-código é uma sequência finita de dígitos, cujo comprimento é determinado pelo número de dígitos que a compõem. Em geral, utiliza-se o alfabeto $F_2 = \{0,1\}$, para se referir aos códigos binários, onde cada elemento do conjunto é conhecido como bit, (SOUZA, 2018).

É importante destacar a existência de dois tipos de códigos corretores de erros: os códigos de bloco, que são sem memória, e os códigos convolucionais, que possuem memória. Em relação aos códigos de blocos, a codificação é um método em que o codificador transforma a informação em blocos de mensagens de comprimentos fixos. Mais especificamente, cada bloco de mensagem, denotado por u, consiste de k dígitos de informações. Assim, o codificador converte cada bloco de mensagens de k dígitos de informações, em uma palavra código de n dígitos, acrescentando bits redundantes a mensagem, os quais são adicionados para auxiliar a detecção e, em alguns casos, a correção de erros. Em geral, denota-se o código de bloco por $C = (n, k)$. Os códigos de bloco ainda são caracterizados pelo fato do codificador não possuir memória, isso devido à palavra código composta por n dígitos depender unicamente da mensagem de entrada de k bits correspondentes, (NICOLETTI, 2015; LIN, 1983).

Por outro lado, os códigos convolucionais, representados como $C = (n, k, m)$, se destacam principalmente por serem códigos com memória, uma vez que dependem de m blocos de mensagens anteriores para sua codificação. O conjunto de sequências codificadas por um codificador de k *bits* da sequência de informação e n dígitos de saída com uma ordem de memória m é chamado de código convolucional, (NICOLETTI, 2015; SABADIN, 2019; LIN, 1983).

Na transmissão de dados reais, ou seja, em sistemas de comunicação real, é inevitável que os erros ocorram. Diante disso, em canais sem memória o ruído impacta cada símbolo transmitido de maneira independente, resultando em erros aleatórios. Por sua vez, em canais com memória, os erros ocorrem em surtos ou sequências de vários erros consecutivos, sendo esses tipos de erros conhecidos como erros em rajadas (*burst errors*), (LIN, 1983).

A seguir, os códigos de blocos são apresentados em maiores detalhes. Esses códigos podem ser caracterizados por meio de uma transformação linear e/ou por uma matriz geradora, que desempenham um papel crucial na execução da codificação e decodificação de mensagens.

Assim, um código de bloco $C = (n, k)$ possui 2^k palavras código de comprimento n e é considerado um código linear se, e somente se, suas 2^k palavras códigos formam um subespaço vetorial de dimensão k do espaço vetorial V constituído de todas a n-uplas cujos elementos pertencem ao conjunto $\{0,1\}$, (LIN, 1983).

A matriz geradora atua na geração e na descrição dos elementos dos códigos. Desse modo, a matriz geradora G é uma matriz, de dimensão $k\ x\ n$, cujas linhas consistem nas palavras códigos e formam uma base para o código. Por exemplo, dado o código $C = (n, k)$ e as palavras códigos $u_1 = (u_{10}, u_{11}, \ldots, u_{1,n-1})$, $u_2 = (u_{20}, u_{21}, \ldots, u_{2,n-1})$, $\ldots$, $u_{k-1} = (u_{k-1,0}, u_{k-1,1}, \ldots, u_{k-1,n-1})$, segue que a matriz geradora desse código pode ser expressa como:

$$G_{k\times n} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_k \end{pmatrix} = \begin{pmatrix} u_{10} & \cdots & u_{1,n-1} \\ \vdots & \ddots & \vdots \\ u_{k-1,0} & \cdots & u_{k-1,n-1} \end{pmatrix} \quad (1)$$

A matriz dada na Eq. (1) está associada à base $B = \{u_1, u_2, \ldots, u_k\}$. As linhas da matriz geradora G são vetores linearmente independentes que geram todas as palavras código do código C. Consequentemente, ao multiplicarmos um vetor mensagem de dimensão k pela matriz geradora, obtemos uma palavra código de comprimento n.

Cada código linear possui sua própria matriz geradora, bem como um conjunto correspondente de palavras códigos. Contudo, a matriz geradora pode ser escrita de diferentes formas. Diante disso, destaca-se a matriz geradora em forma padrão, a qual envolve a subdivisão da matriz em duas partes, onde Id_k é a matriz identidade de ordem $k\ x\ k$ e P é a submatriz de ordem $k\ x\ (n-k)$, (NICOLETTI, 2015). Matematicamente, a matriz geradora em forma padrão pode ser escrita como:

$$G = (Id_k \quad \vdots \quad P) \quad (2)$$

Uma outra matriz que desempenha um importante papel na verificação se uma dada mensagem recebida está correta é a matriz de verificação de paridade H. Em outras palavras, a matriz H é empregada para verificar se uma mensagem recebida é ou não uma palavra código, (LIN, 1983; NICOLETTI, 2015). Diante disso, ao considerar a matriz geradora $G = (Id_k \quad \vdots \quad P)$ do código linear $C(n,k)$, a matriz de verificação de paridade H de ordem $(n-k)\ x\ n$ com $n-k$ linhas linearmente independentes, pode ser representada por:

$$H = (P^t \quad \vdots \quad Id_{n-k}) \quad (3)$$

denominada a forma padrão da matriz H, (LIN, 1983).

Diante do exposto, o código linear pode ser reescrito como:

$$C = \{ y \in (F_q)^n : \ H \cdot y^t = 0\}, \quad (4)$$

e consequentemente, uma palavra código $y \in C$ se, e somente se,

$$H \cdot y^t = 0. \quad (5)$$

APLICAÇÕES PRÁTICAS

A partir da visão geral de um sistema de comunicação e dos principais conceitos sobre os códigos corretores de erros, buscou-se pesquisar sobre as aplicações e usos dos códigos. Diante disso, apresentam-se alguns estudos relevantes que utilizaram os códigos corretores de erros.

Mais especificamente, na pesquisa de Vianna (2016), os códigos corretores de erros convolucionais foram empregados na área das comunicações submarinas, com o propósito de aprimorar a conectividade entre submarinos.

O estudo conduzido por Mobilon (2003) concentrou-se na aplicação dos códigos corretores de erro em sistemas de comunicações ópticas. Em particular, foi realizada uma análise da técnica de correção de erros avançada, conhecida como Forward Error

Correction (FEC), que habilita o decodificador não apenas para a detecção, mas também para a correção de erros originados de interferências no canal de transmissão.

O trabalho desenvolvido por Bassi (2019) explorou a conexão entre os códigos corretores de erros e a genética. O estudo implementou um código capaz de identificar e reproduzir sequências de DNA, a fim de contribuir para diagnósticos de doenças, análises de mutações, produção de fármacos e melhoramento genético de maneira mais eficiente, reduzindo tempo e custos laboratoriais.

Diante desses trabalhos é possível vislumbrar como as aplicações são abrangentes e impactantes em diversas áreas, demonstrando sua versatilidade e relevância na sociedade contemporânea.

CONCLUSÃO

Diante do exposto, vislumbra-se que os códigos de bloco linear desempenham um papel crucial na capacidade de adicionar redundância aos dados, permitindo a detecção e correção de erros. Por meio do uso de ferramentas matemáticas como a matriz geradora e a matriz de verificação de paridade, esses códigos se mostram eficazes na identificação e correção de erros em blocos de dados, desempenhando um papel fundamental na preservação da integridade dos dados em sistemas de comunicação, seja na transmissão e/ou no armazenamento dos dados.

Além disso, diante de alguns trabalhos exibidos, possibilitou-se visualizar a versatilidade e utilidade desses códigos na detecção e correção de erros em diversas áreas.

Assim, esses sistemas são essenciais para garantir uma comunicação confiável e eficiente em diversos domínios da sociedade moderna. A compreensão desses conceitos desempenha um papel central na criação e aprimoramento de soluções tecnológicas que atendem às demandas contemporâneas. Em outras palavras, essa área de pesquisa continua atual e necessária para o desenvolvimento e projeção de sistemas mais robustos e eficientes, capazes de enfrentar os desafios relacionados a erros de transmissão e armazenamento de dados do que os atuais.

Agradecimentos

Agradecemos a UTFPR-CM e a Fundação Araucária, pelo apoio financeiro.

Conflito de interesse

Não há conflito de interesse.

REFERÊNCIAS

BASSI, Mariana Venezian Musto. **Análise de sequências de DNA através de Códigos Corretores de Erros.** São João de Boa Vista: UNESP. 2019.

LIN, Shu; COSTELLO, Daniel Joseph Jr. **Error Control Coding:** Fundamentals and Applications. Englewood Cliffs, New Jersey: Prentice-Hall. 1983.

MOBILON, Eduardo. **Análise Experimental das Aplicações de Códigos Corretores de Erro em Sistemas de Comunicações Ópticas.** Campinas: UNICAMP. 2003.

NICOLETTI, Everton Rodrigo. **Aplicações de álgebra linear aos códigos corretores de erros e ao ensino médio**. Rio Claro: UNESP. 2015.

SABADIN, Graça Aparecida Prestes. **Códigos Corretores de Erros**. Florianópolis: UFSC. 2019.

SHANNON, Claude Elwood. A Mathematical Theory of Communication**. The Bell System Technical Journal**. july, october 1948. v. 27. p. 379–423, 623–656.

SOUZA, Natália Pedroza. **Sobre Códigos Corretores de Erros.** Rio de Janeiro: UFRJ. 2018.

VIANNA, Maria Luiza Costa; MARQUES, João Pedro Kappes; CAMPOS, Marcello L. R. **O Uso de Códigos Corretores de Erros em Comunicações Acústicas Submarinas**. In: XXII Simpósio Brasileiro de Telecomunicações. Santarém. 2016.