



Análise de Complexidade - USEI07: KD-Tree Construction and Statistics

Resumo

Este documento apresenta a análise detalhada da complexidade algorítmica da implementação da *User Story* **USEI07 - KD-Tree Construction and Statistics**. A solução utiliza uma **KD-Tree 2D balanceada** para indexação espacial das estações ferroviárias (Latitude/Longitude). A complexidade de **construção é $O(N \log N)$** e as estatísticas (Tamanho, Altura) são reportadas com eficiência, garantindo que a base para buscas espaciais subsequentes seja robusta e otimizada.

1. Introdução

A *User Story* USEI07 é o pilar das funcionalidades de pesquisa espacial (USEI08, 09, 10). O seu foco é na **construção eficiente** de uma estrutura de dados espacial, a KD-Tree, e na validação do seu balanceamento e integridade através da apresentação de estatísticas.

1.1. Arquitetura da Solução

A estrutura central é uma **KD-Tree (K-Dimensional Tree)** com $K=2$ (Latitude e Longitude), construída de forma balanceada.

Características da KD-Tree

Estrutura	Propriedades
KD-Tree	Árvore binária que particiona o espaço 2D.
Balanceamento	A construção é feita recursivamente usando a mediana como nó de divisão, o que limita a altura da árvore a $O(\log N)$.
Dimensões Alternadas	O nó de raiz divide pelo eixo da Latitude. Os seus filhos dividem pelo eixo da Longitude, e assim sucessivamente (dimensões alternadas: D0, D1, D0, D1, ...)

2. Análise de Complexidade Temporal

A complexidade temporal da USEI07 é dominada pelo algoritmo de construção da KD-Tree.

2.1. Construção da KD-Tree Balanceada (`build2DTree`)

2.1. Construção da KD-Tree Balanceada (`build2DTree`)

O método de construção balanceada é crucial para garantir que as buscas futuras (USEI08, 09, 10) atinjam a complexidade sub-linear $O(N^{1/2})$ no caso médio.

Complexidade de Construção

Passo	Complexidade	Justificativa
1. Pré-processamento	$O(N \log N)$	O pré-processamento/ordenação do <i>dataset</i> (por Latitude e Longitude) necessário para o balanceamento domina o custo.
2. Construção Recursiva	$O(N)$	O custo de inserção de todos os N elementos na árvore tem uma complexidade linear.
Complexidade Total (Build)	$O(N \log N)$	

2.2. Cálculo e Apresentação de Estatísticas

As estatísticas são calculadas através de uma travessia da árvore.

Complexidade das Estatísticas

Estatística	Complexidade	Justificativa
Tamanho da Árvore (Size)	$O(N)$	Requer a travessia de todos os N nós para obter a contagem total, a menos que seja um valor <i>cached</i> ($O(1)$).
Altura da Árvore (Height)	$O(N)$	O cálculo da altura requer a travessia completa da árvore para encontrar o maior caminho.
Distribuição de Buckets	$O(N)$	Requer uma travessia dos nós folha (<i>buckets</i>) da árvore para recolher as contagens.

3. Análise de Complexidade Espacial

A complexidade espacial é determinada pela memória necessária para armazenar a estrutura de dados.

Complexidade Espacial

Componente	Complexidade	Justificativa
Armazenamento da KD-Tree	$O(N)$	A KD-Tree armazena um número de nós linearmente proporcional ao número de estações, $5N$.
Pilha de Recursão	$O(\log N)$	A profundidade máxima da pilha de recursão durante a construção é logarítmica, garantida pelo balanceamento da árvore.

4. Conclusão

A implementação da USEI07 alcança a **eficiência e balanceamento** necessários para as buscas espaciais futuras:

- **Eficiência de Construção:** A complexidade $O(N \log N)$ é o padrão para estruturas balanceadas.
- **Validação:** A capacidade de reportar estatísticas de $O(N)$ (Tamanho e Altura) é essencial para validar o balanceamento da árvore, preparando o terreno para buscas eficientes (USEI08/09/10).