

Indexação Espacial

Contents

- Introdução
- Feições Pontuais
- Feições Lineares
- Feições de Área
- Índice Espacial no PostGIS
- Dados Matriciais

Introdução

- Métodos de acesso espacial, ou índices espaciais, são estruturas de dados auxiliares, mas essenciais para o processamento eficiente de consultas espaciais.
- Busca e Inserção de objetos.
- As consultas não consideram todas as instâncias do banco, mas são otimizadas pela redução do universo da amostra, sobre a qual seriam aplicados algoritmos geométricos computacionalmente complexos e custosos.
- Indexação unidimensional: ordem alfabética ou numérica crescente.
- Complexidade logarítmica X complexidade polinomial;

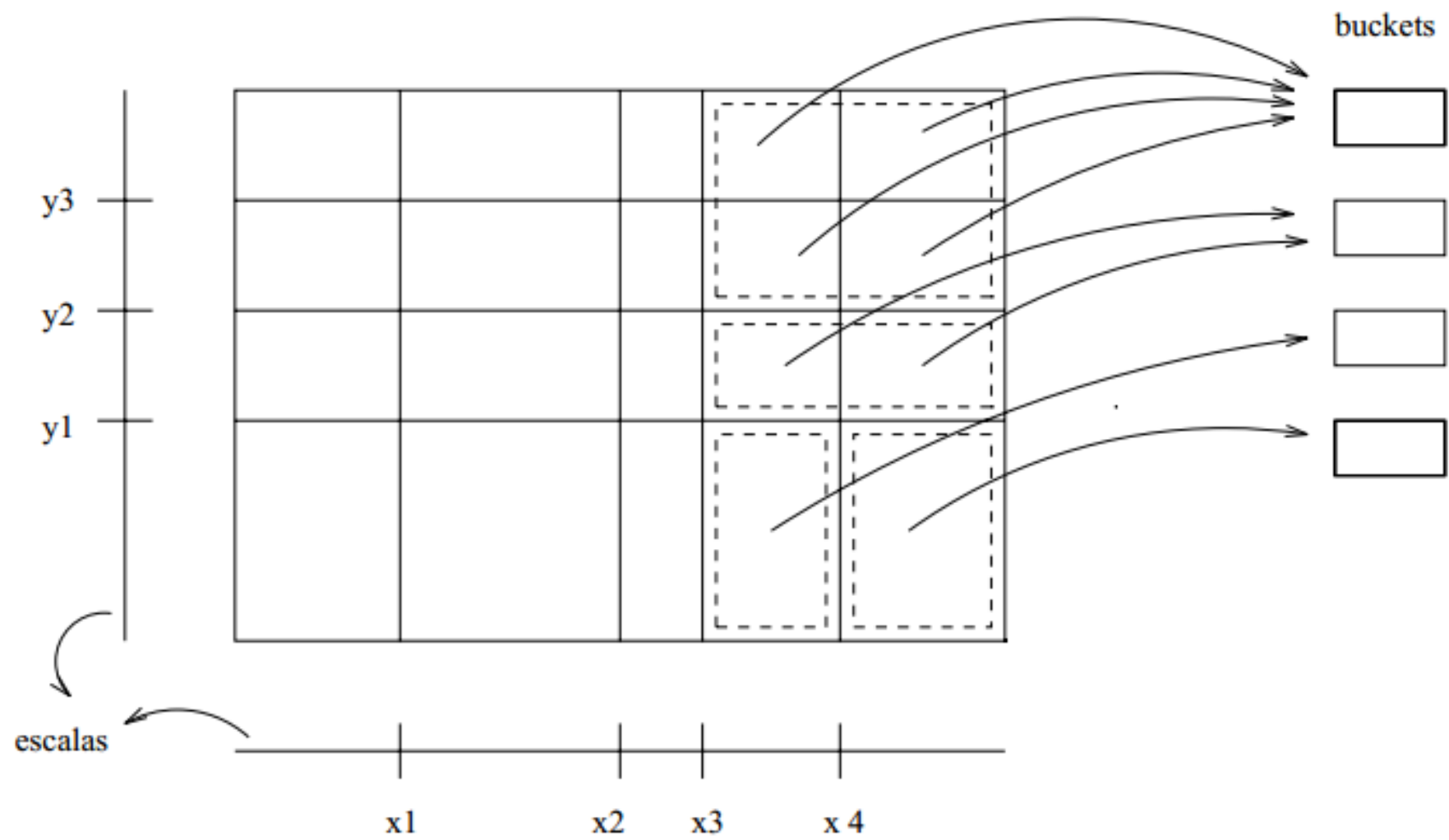
Feições Pontuais

Grade

Particiona a região de interesse em células através de retas paralelas aos eixos e *espaçadas irregularmente*. Um arquivo em grade consiste de *duas escalas e um diretório*.

A escala associada ao eixo X é uma lista contendo as *abcissas em que as retas paralelas ao eixo Y cruzam o eixo X* (e similarmente para a outra escala). Sempre que possível, as escalas são dispostas para criar grades regulares.

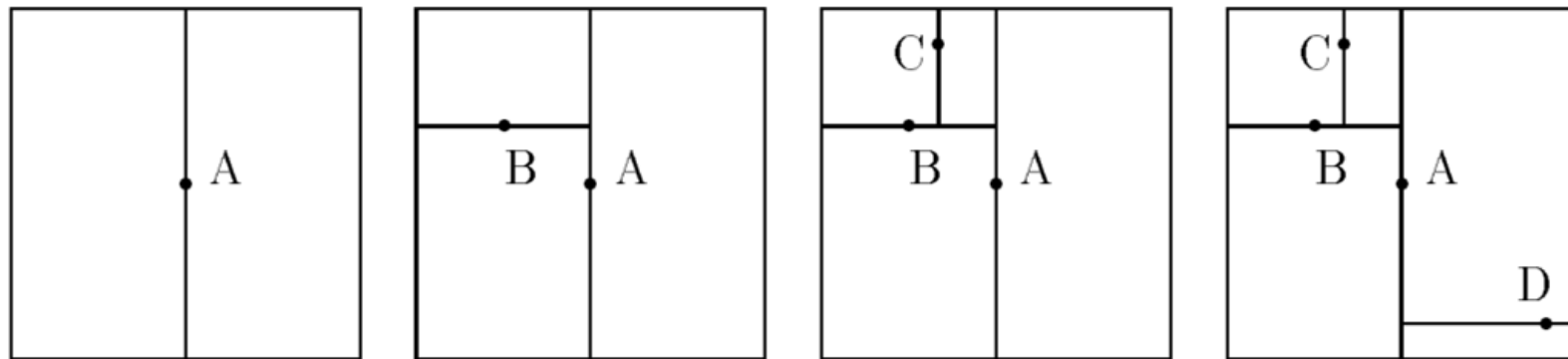
O diretório é uma *matriz bidimensional*, com as mesmas dimensões da grade, tal que cada elemento do diretório corresponde a uma célula e aponta para a página física (ou *bucket*) que contém todas as coordenadas dos pontos localizados na célula.



Feições Pontuais

Árvore *kd*

É uma árvore de busca binária tal que os nós interiores em cada nível contêm valores referentes a um único eixo (X ou Y, alternadamente) e as folhas apontam para páginas físicas. O processo de subdivisão pára quando a célula contém um conjunto de pontos que possa ser armazenado em uma única página física.

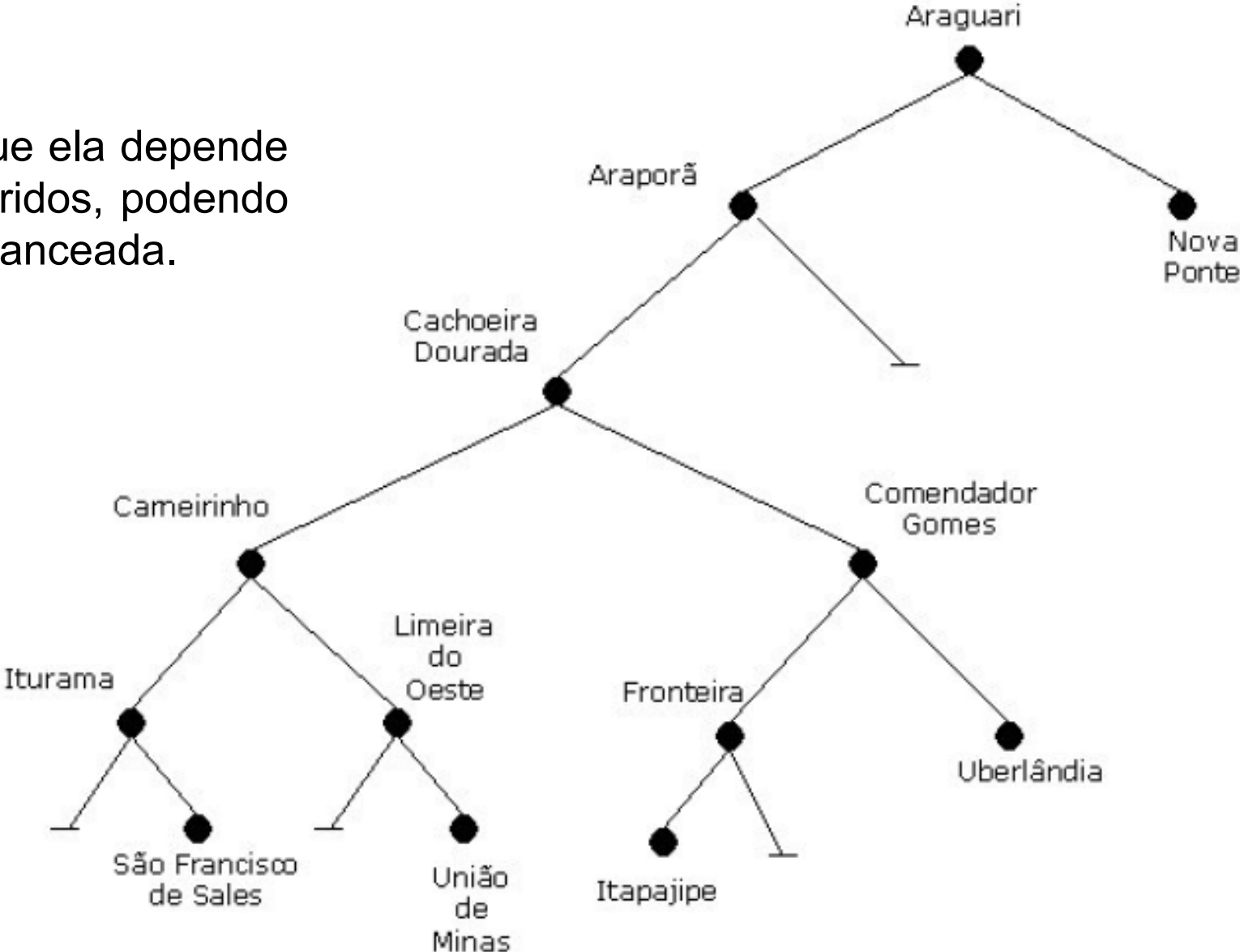


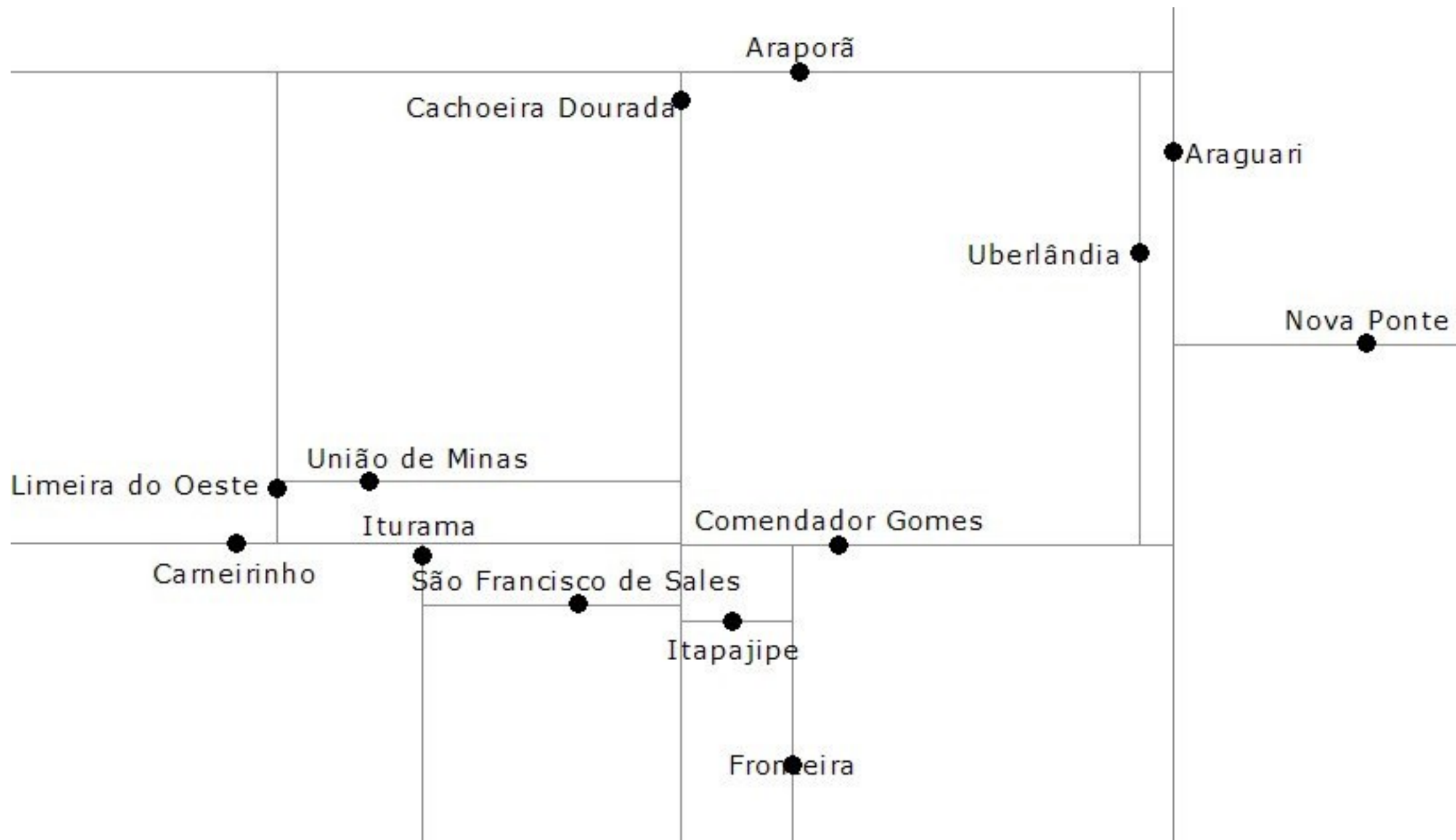


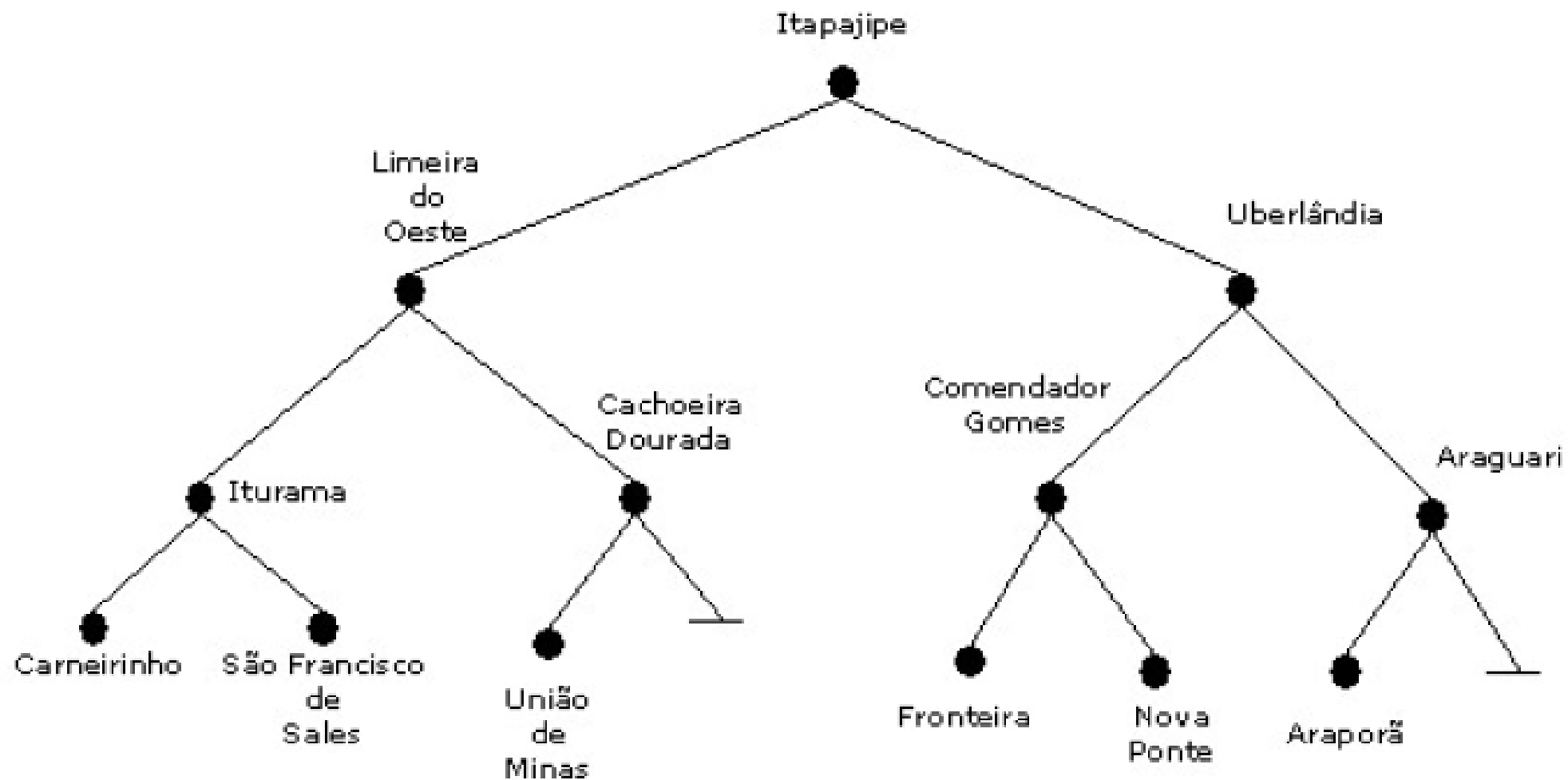
Feições Pontuais

- Considerando Araguari a primeira cidade a ser inserida na árvore, esta ocupará a raiz;
- Sendo Araporã, a próxima, como ela possui um valor de coordenada “x” menor do que a de Araguari, esta é inserida à esquerda. Aqui, como a raiz ocupa um nível par (0), comparamos os valores da componente “x” das coordenadas das sedes municipais.
- Nova Ponte possui um valor “x” maior, e, portanto, é inserida à direita.
- No caso de Cachoeira Dourada, ela possui um valor de “x” menor do que Araguari (comparação em um nível par), e um valor de “y” menor do que Araporã (comparação no nível ímpar – 1), e por isso é inserida à esquerda desta.
- Esse processo é realizado até que todas as cidades estejam armazenadas na árvore.

Um dos problemas da *k-d tree* é que ela depende da ordem em que os nós são inseridos, podendo obter uma árvore totalmente desbalanceada.



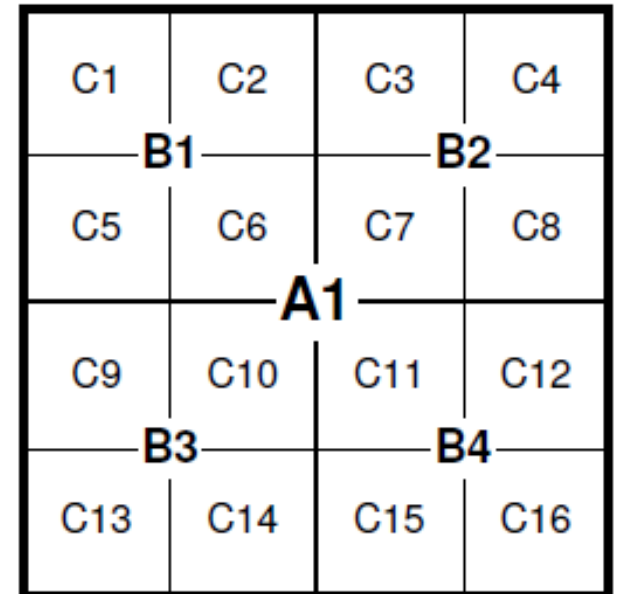




Feições Pontuais

Quad-tree

- Tipo de estrutura de dados organizada em árvore, em que cada “nó” ou “tronco” gera sempre (e exatamente) quatro “folhas”;
- Cada região será subdividida, novamente em quatro partes, gerando mais um nível na árvore, e assim sucessivamente, até que se chegue a ter um ou nenhum objeto geográfico dentro dos quadrados resultantes da subdivisão;
- O resultado final será uma árvore que terá um número variável de níveis, de acordo com a região geográfica: regiões mais densas formarão mais níveis, e regiões menos densas terão menos níveis.

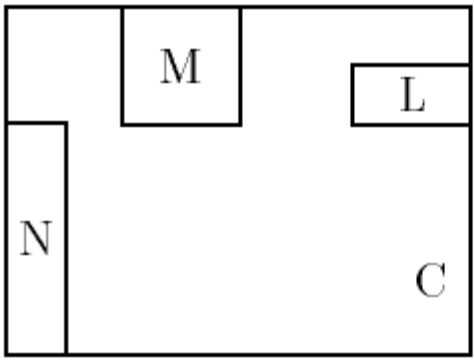
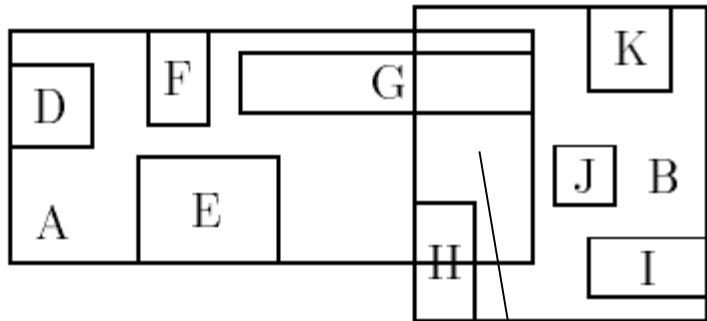


Feições Lineares

Retângulos

Transformação de retângulos em pontos: consiste em transformar um retângulo de duas dimensões em uma quádrupla (um ponto de 4 dimensões) e utilizar um método de armazenamento e indexação espacial de quádruplas, como arquivos em grade e árvores *kd* quadri-dimensionais.

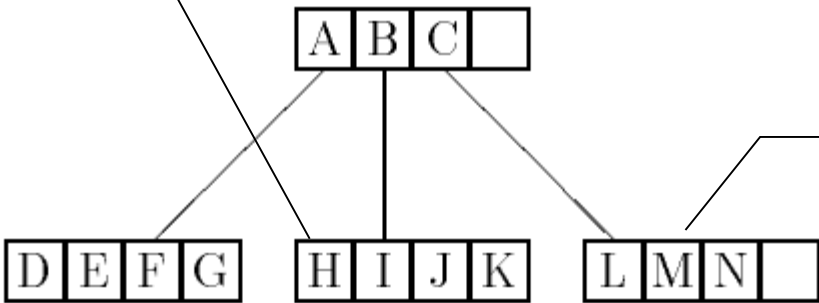
Decomposição do espaço com superposição: abandona a ideia de decompor o espaço em regiões que não se superpõem e utiliza a noção de retângulo envolvente. O *retângulo envolvente mínimo (r.e.m.)* de um conjunto de objetos no \mathcal{R}^2 é o menor retângulo com lados paralelos aos eixos *X* e *Y* que contém todos os objetos no conjunto. Sugere-se o uso de árvores *R* (*r-tree*), onde todos os ramos possuem o mesmo comprimento, ou seja, todas as folhas estão no mesmo nível, como numa árvore de busca balanceada.



Superposição

Espaço não representado

Balanceamento da árvore

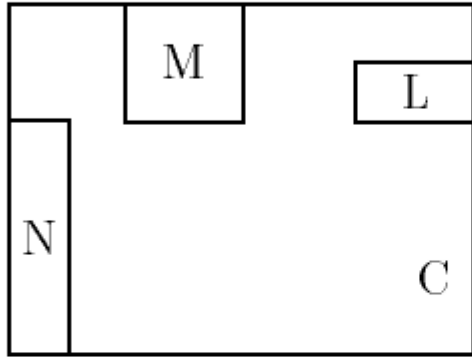
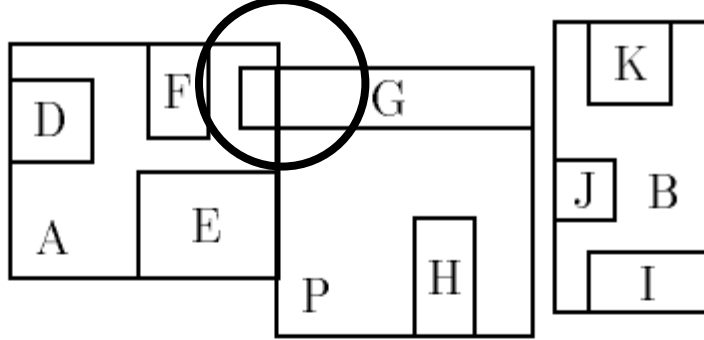


Estrutura dinâmica altera-se a cada inserção ou remoção de retângulo

Feições Lineares

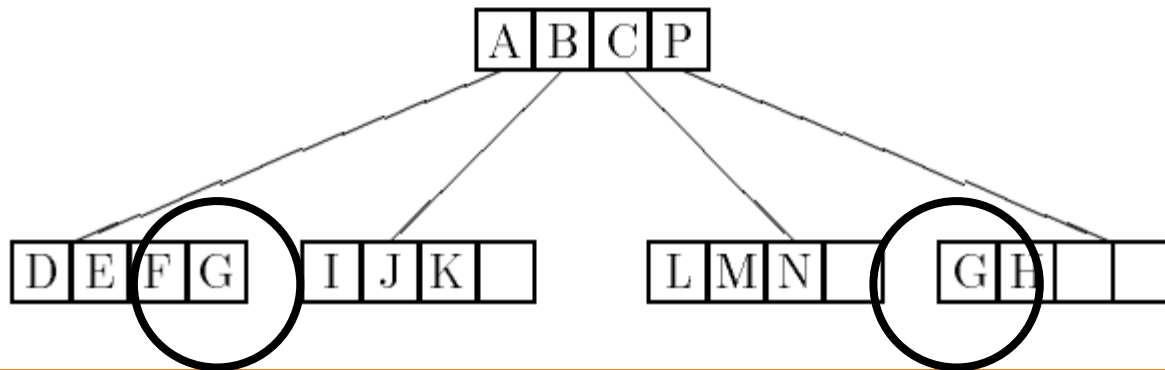
- *Retalhamento de retângulos*: mantém a ideia de decompor o espaço em regiões que não se superpõem. Porém, cada retângulo ao ser inserido, é retalhado em um conjunto mínimo de retângulos tal que:
 - os retângulos não se superpõem;
 - a união dos retângulos reconstrói o retângulo original;
 - cada retângulo intercepta exatamente uma das regiões em que se decompõe o espaço.

Resulta na replicação de objetos na estrutura, pois a inserção de um único retângulo acarreta a inserção de uma entrada para cada região que intercepta.



As árvores R+

Variante de árvores R em que, ao se inserir uma nova entrada, se o seu retângulo intercepta os retângulos de dois ou mais nós, o retângulo é recortado como anteriormente descrito e cada um dos novos retângulos é inserido nos nós apropriados



Feições Lineares

Sejam v_0, v_1, \dots, v_n n pontos no plano.

Sejam $s_0 = v_0v_1, s_1 = v_1v_2, \dots, s_{n-2} = v_{n-2}v_{n-1}$ uma seqüência de $n - 1$ segmentos, conectando estes pontos.

Estes segmentos formam uma poligonal L se, e somente se:

- a interseção de segmentos consecutivos é apenas o ponto extremo compartilhado por eles (i.e., $s_i \cap s_{i+1} = v_{i+1}$)
- segmentos não consecutivos não se interceptam (i.e., $s_i \cap s_j = \emptyset$ para todo i, j tais que $j \neq i + 1$), e
- $v_0 \neq v_{n-1}$, ou seja, a poligonal não é fechada.

Feições Lineares

- decomposição do espaço em regiões que não se superpõem: cada linha é quebrada em um conjunto mínimo de novas linhas tal que:
 - as novas linhas não se superpõem;
 - a união das novas linhas resulta na linha original;
 - cada nova linha intercepta exatamente uma das regiões em que se decompõe o espaço.

Este enfoque permite adaptar para linhas qualquer método desenvolvido para armazenar e acessar pontos.

- Um segundo enfoque consiste em aproximar cada linha pelo seu REM e utilizar os métodos para indexar esses retângulos.

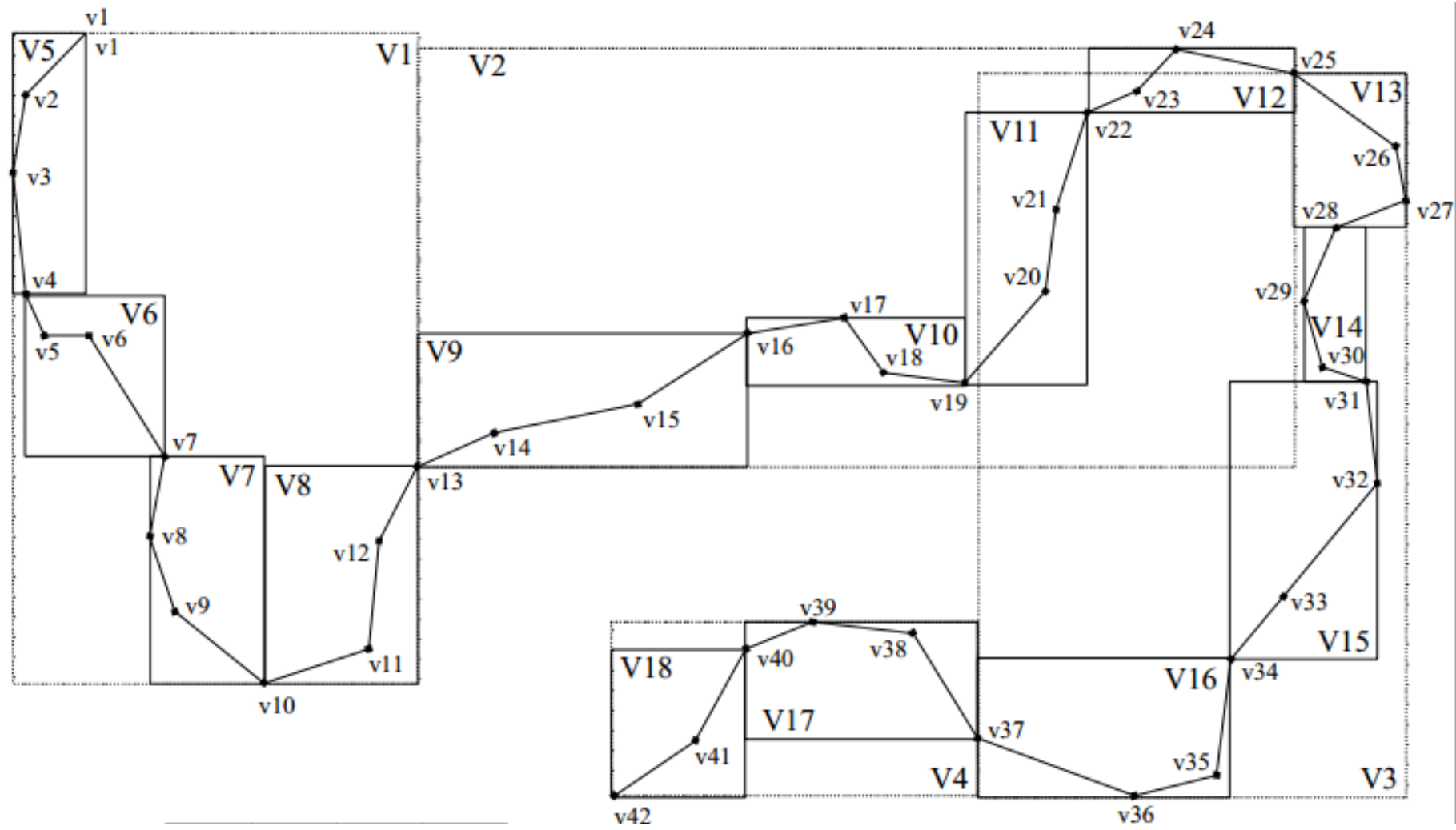
Este enfoque não é adequado para o caso de linhas muito longas ou quando o REM não é uma boa aproximação para as linhas.

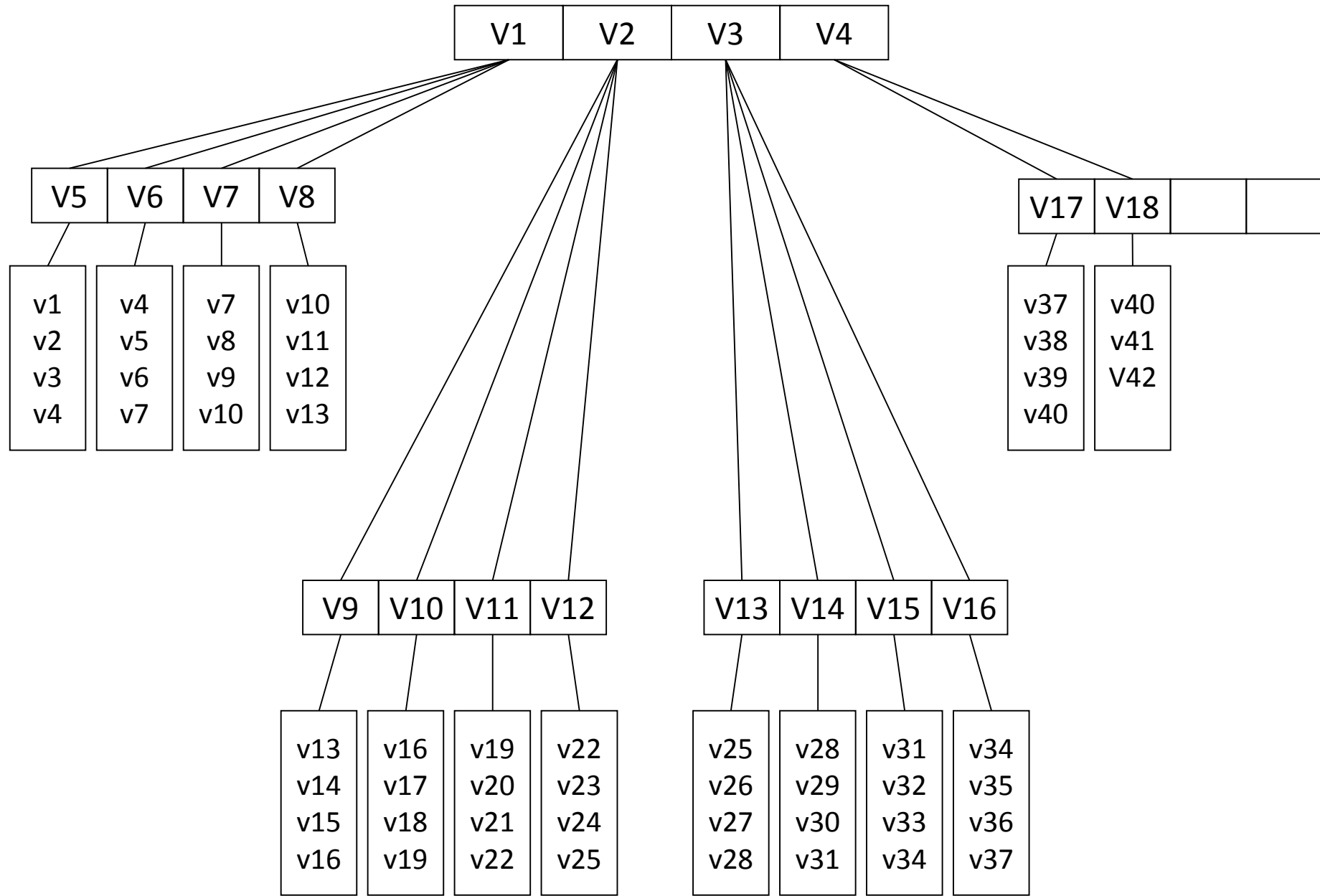
Feições Lineares

Árvores VR

Uma árvore V consiste em uma árvore R construída com base nos REM correspondentes à segmentação de uma linha L .

Cada linha é armazenada em uma árvore V separada e aproximada pelo seu retângulo envolvente mínimo; o conjunto destes retângulos é então armazenado em uma árvore R , cujas folhas apontam para as raízes das árvores V . A estrutura resultante, composta desta árvore R e das várias árvores V , forma então uma árvore VR.





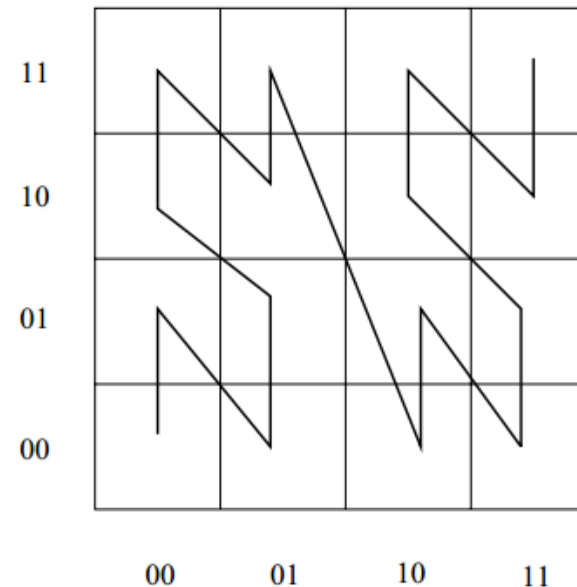
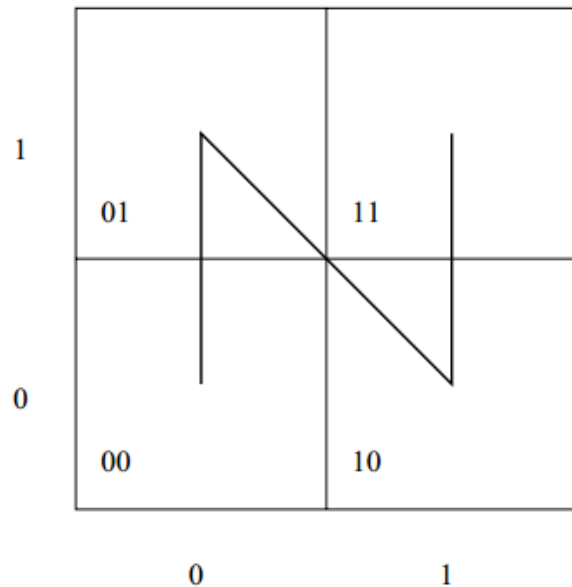
Feições de Área

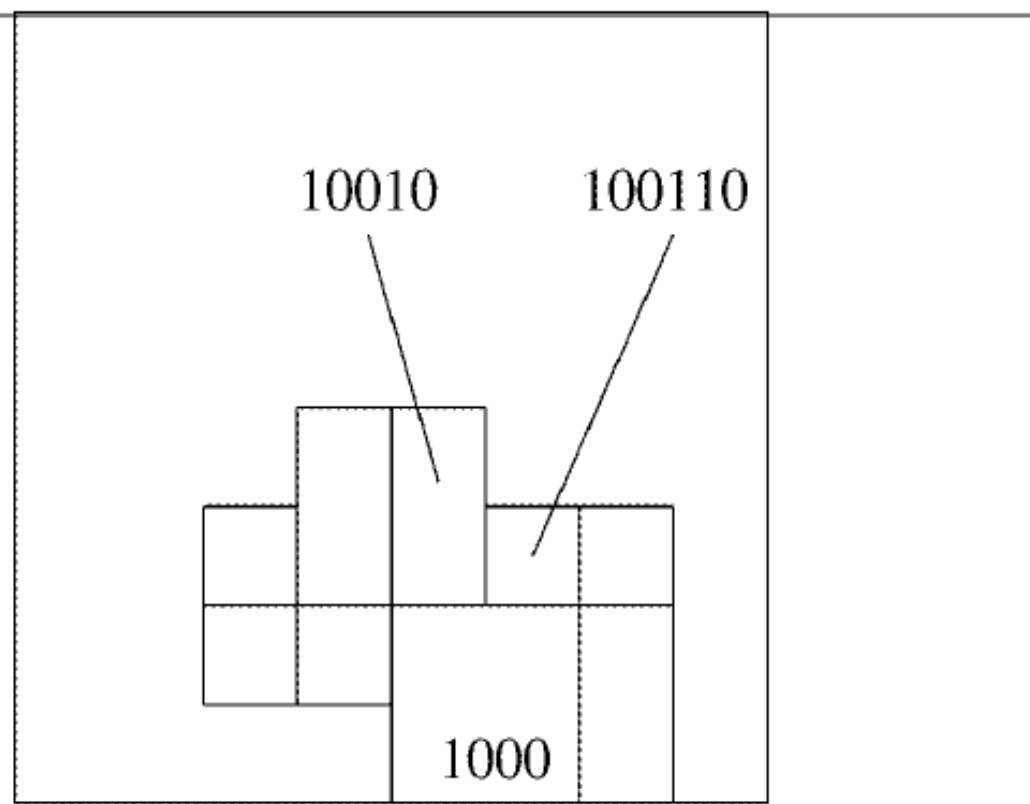
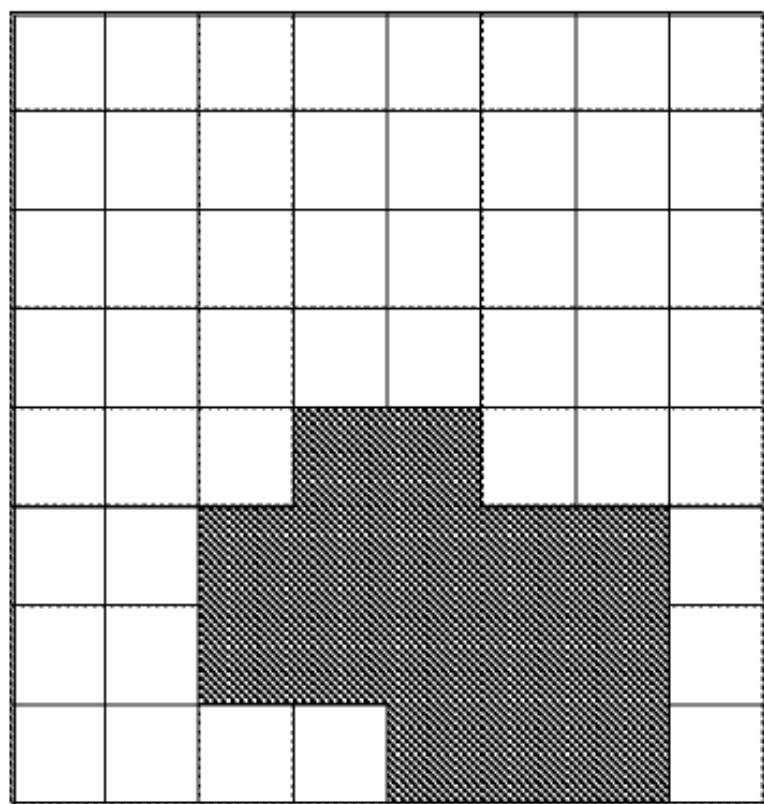
Polígonos: região do plano limitada por uma linha poligonal fechada.

- aproximações para a geometria dos polígonos: O retângulo envolvente mínimo (r.e.m.) de um polígono oferece uma aproximação bastante simples e cômoda.
- decomposição dos polígonos em polígonos simples: significa dividi-lo em várias componentes simples - triângulos, trapézios, e outros - de tal forma que, durante o processamento de uma consulta espacial, apenas um pequeno número destas componentes precise ser examinado.

- método de decomposição por grades: o espaço dividido hierarquicamente em células regulares. As células geradas a cada nível podem ser colocadas em z-ordem através do processo de intercalação de bits, proposto originalmente por Morton.

Em cada quadrante, as células são conectadas em z-ordem e depois os grupos de células de cada um dos quatro quadrantes são novamente conectados em z-ordem. Cada célula em cada nível da hierarquia possui um endereço dado por uma cadeia de bits cujo comprimento corresponde ao nível ao qual a célula pertence.





Índices Espaciais no PostGIS

- GIST: Generic Index Structure;

O PostGIS, por *default*, usa árvores binárias para indexação (*B-tree*), considerando a geometria detalhada de cada objeto. O emprego da estrutura genérica implica a indexação dos REM de cada geometria.

- BRIN: Block Range Index;

Dados Matriciais

- para matrizes mais simples: armazenar, nas páginas físicas em memória secundária, uma linha após a outra (ou uma coluna após a outra).

Este método é adequado para processar operações que exigem visitar os elementos da matriz linha a linha, ou sem nenhuma ordem preferencial como, por exemplo, no cálculo de um histograma.

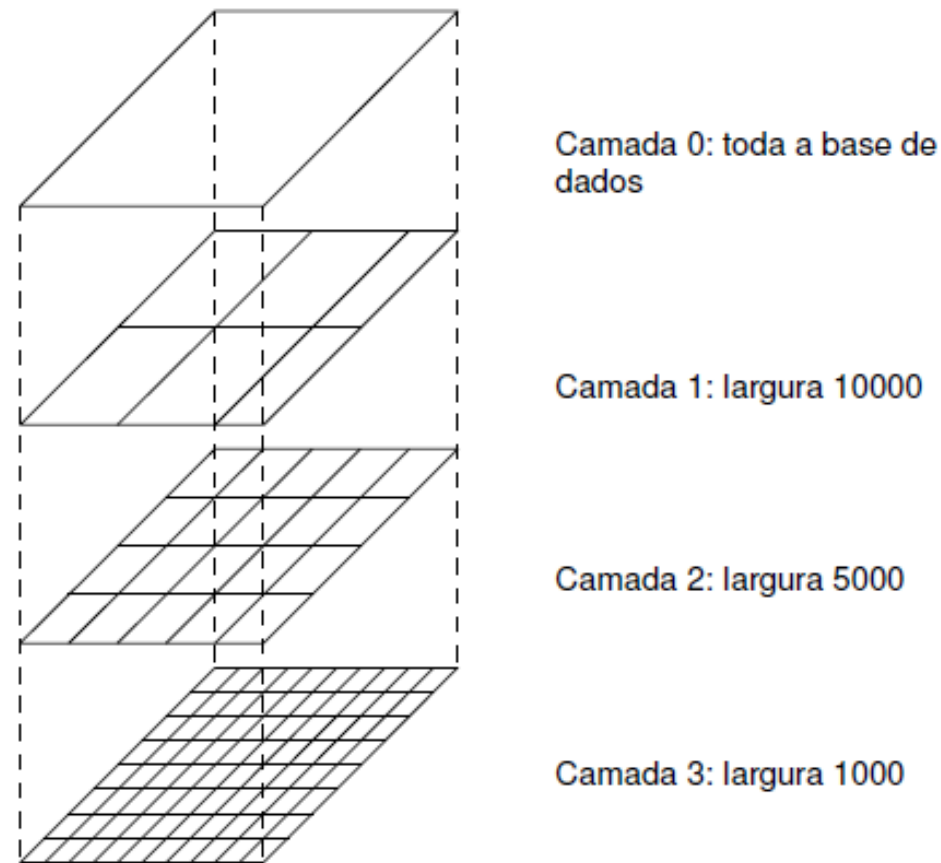
Ele não é apropriado, no entanto, quando se deseja recuperar os elementos da matriz cobertos por uma janela definida por um intervalo de índices em cada dimensão, por exemplo.

Dados Matriciais

- para matrizes maiores: aplicar o processo de *tiling*, ou seja, subdividir a área em grades regulares de dimensões fixas.

No momento do *display*, testa-se os limites da janela de visualização contra a lista de ladrilhos, selecionando aqueles que tenham pelo menos uma extremidade contida no retângulo.

OGC WMTS Simple Profile: Web Map Tile Service



a.tile.openstreetmap.org

X=9798
Y=14664



mt1.google.com

X=9799
Y=14664



otile1.mqcdn.com/tiles/1.0.0/

X=9800
Y=14664



www.ogc.uab.cat/SITiled/GeoEye_Haiti_M

X=9798
Y=14665



ecn.10.tiles.virtualearth.net

X=9799
Y=14665



tile.openstreetmap.org/cycle/

X=9800
Y=14665

