

Statické varianty R-stromů

Nechť množina obsahuje N obdélníků a každý uzel může obsahovat až c obdélníků (c může být kapacita uzlu, nebo menší číslo – to když chceme nechat ve stromě volné sloty na později vkládaná data). Chceme sestavit R-strom. Můžeme postupně všechny vkládat do původně prázdného R-stromu. Výhodnější je ale použít statické algoritmy stavby R-stromu (naproti tomu, algoritmy probírané v předchozích lekcích se nazývají dynamické).

Statické algoritmy

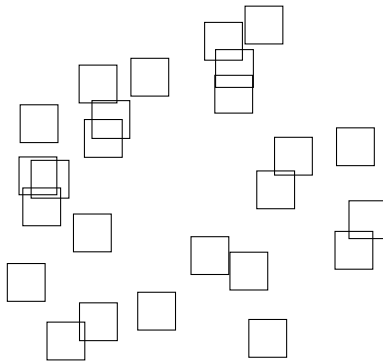
Naivní statický R-strom

Naivní přístup můžeme popsat následovně:

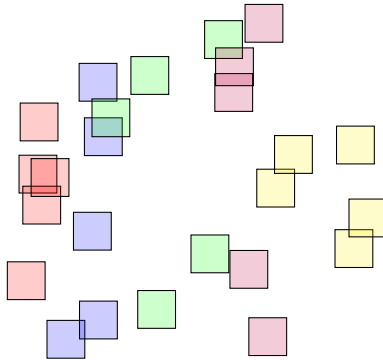
1. Seříd' obdélníky podle x -souřadnice jejich středu (stejně tak může být použita např. x -souřadnice levého dolního rohu).
2. Seskup obdélníky do $\lceil N/c \rceil$ po sobě jdoucích skupin o c obdélnících (až na poslední, ta může obsahovat $1-c$ obdélníků).
3. Najdi MBR každé skupiny vytvořené v předchozím kroku a spoj každou s ukazatelem na uzel, ve kterém bude uložena odpovídajících množin.
4. Rekurzivně pokračuj s MBR z předchozího kroku. Rekurze pokračuje, dokud máme maximálně c obdélníků.

Příklad:

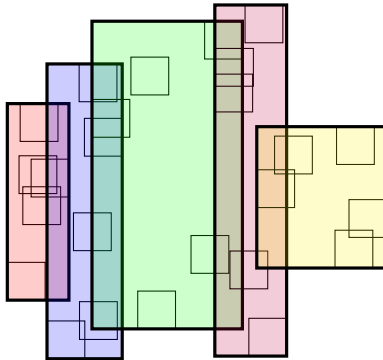
– Zadání: $N = 25, c = 5$



– Kroky 1 a 2: dostáváme $\lceil N/c \rceil = 5$ skupin o 5 obdélnících (na obrázku rozlišeno barvami – červená, modrá, zelená, fialová, žlutá):



– Krok 3: MBR jednotlivých skupin:



– Máme 5 obdélníků, můžeme ukončit konstrukci R-stromu. Kdybychom dostali více obdélníků, pokračujeme rekurzivně pro tyto nové obdélníky.

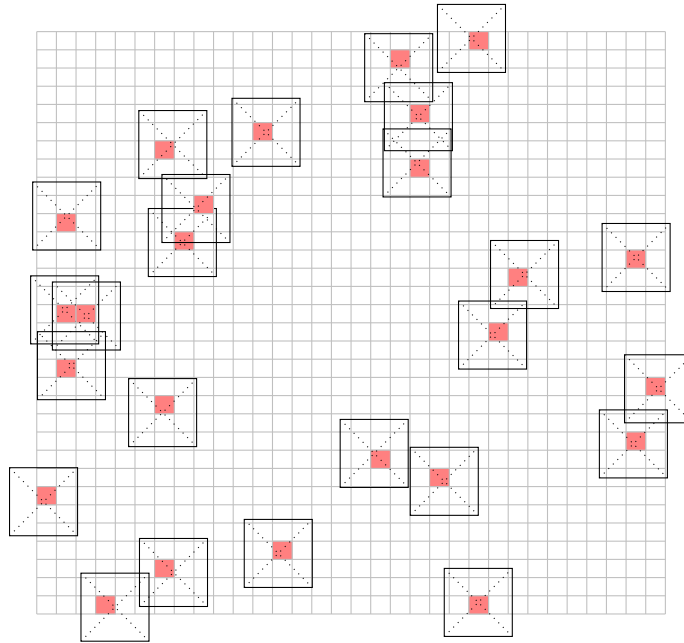
Tento algoritmus nedává dobré výsledky. Data jsou vlastně „rozproužkována“ podle jedné ze souřadnic (x -souřadnice těžiště, některého rohu, ...). Máme ale možnost použít lepší hodnotu než jednu souřadnici: H-hodnotu.

Hilbertův statický R-strom

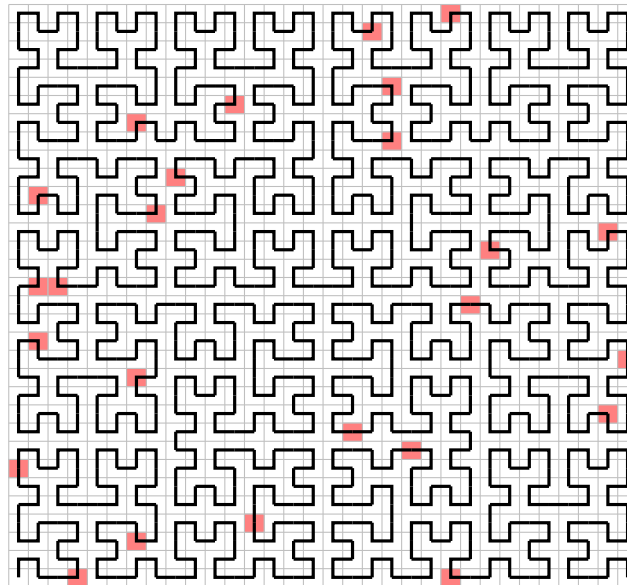
Algoritmus funguje stejně jako naivní algoritmus, ale místo jedné souřadnice používá H-hodnotu (tj. pozici na Hilbertově křivce).

Příklad:

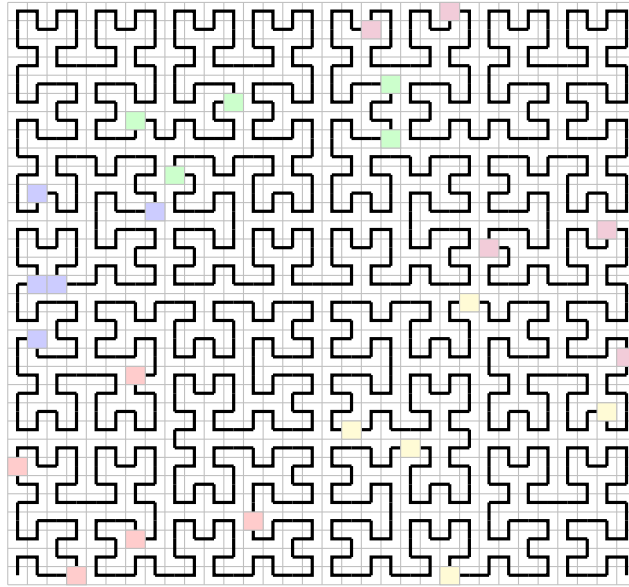
- Stejně zadání jako v prvním příkladě.
- Krok 1: Seřazení podle H-hodnot. 1) Oblast proložíme dostatečně jemnou $2^n \times 2^n$ mřížkou; zde $n = 5$. Namapujeme těžiště obdélníků do čtverců mřížky.



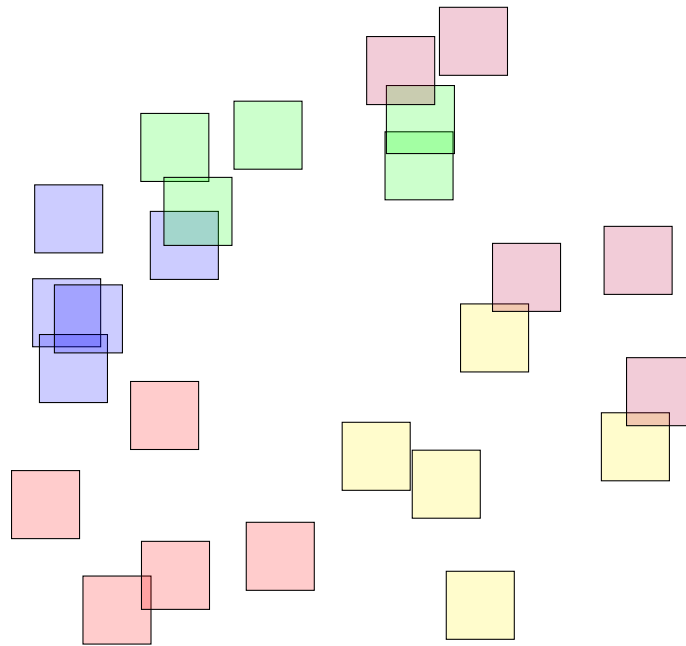
- Mřížku proložíme Hilbertovou křivkou řádu n , tj. 5. (původní obdélníky skryty kvůli přehlednosti, zobrazeny jsou jen čtverce jejich těžišť).



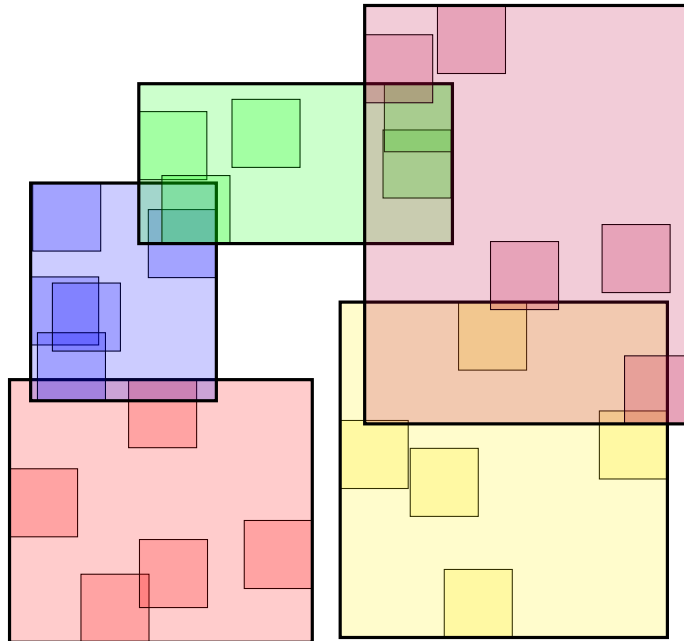
- Krok 2: Obdélníky rozdělíme podle jejich H-hodnoty, tj. pozici na Hilbertově křivce (původní obdélníky opět skryty kvůli přehlednosti, zobrazeny jsou jen čtverce jejich těžišť).



– (naopak skryty těžiště, mřížka a Hilbertova křivka; zobrazeny obdélníky)



– Nalezení MBR jednotlivých skupin.



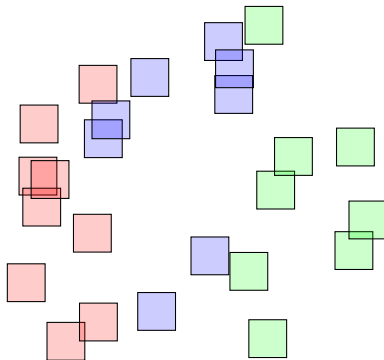
- Máme 5 obdélníků, můžeme ukončit konstrukci R-stromu. Kdybychom dostali více obdélníků, pokračujeme rekurzivně pro tyto nové obdélníky.

STR R-strom

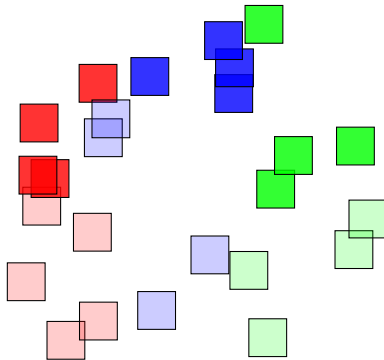
STR (Sort-Tile-Recursive) je algoritmus konstrukce statických R-stromů (popisán pouze pro 2D). Základní myšlenka této metody je nejdříve rozdělit prostor pomocí VS pruhů tak, aby každý obsahoval dostatek obdélníků na vytvoření zhruba $\sqrt{N/c}$ uzlů. Na počátku se stanoví počet listových uzlů, který je $n_l = \lceil N/c \rceil$. Nechť $VS = \sqrt{n_l}$. Obdélníky jsou seřazeny podle x -souřadnice jejich těžiště, a je vytvořeno VS pruhů. V každém pruhu jsou objekty seřazeny dle y -souřadnice těžiště a rozděleny do uzlů (c do každého uzlu).

Příklad:

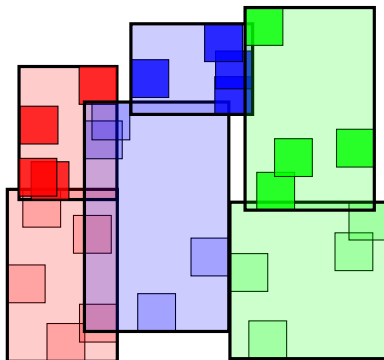
- Zadání stejné jako v předchozích příkladech (tento příklad nebude vycházet úplně hezky kvůli nízkým hodnotám parametrů).
- Máme $n_l = 5$, $VS = \sqrt{2} \approx 2.236$, rozdělme tedy obdélníky rovnoměrně (9,8,8) do tří pruhů dle x -souřadnice těžiště.



– Následně každý z těchto tří pruhů rozdělíme rovnoměrně na dvě množiny:



– Najdeme MBR vzniknuvších množin.

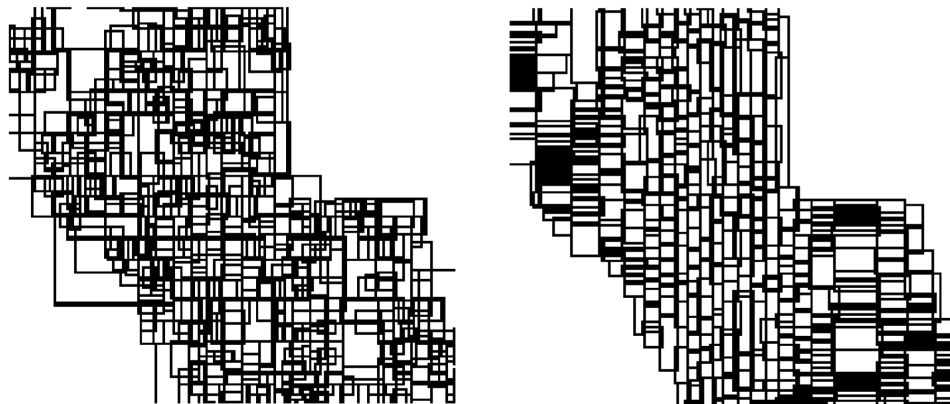


– Dostali jsme 6 množin (více než c), měli bychom rekurzivně pokračovat pro tyto nalezené obdélníky.

Všechny tři algoritmy popsané výše se dají zobecnit na vyšší počet dimenzí:

- U naivního algoritmu je to triviální.
- U Hilbertova statického R-stromu potřebujeme rozšíření Hilbertovy křivky na více dimenzí – není zrovna jednoduché.
- U STR je to snadné: pro $d > 2$ dimenzí, nejdříve rozdělí (nad-)obdélníky na $VS = n_l^{\frac{1}{d}}$ skupin podle x -souřadnice, kde $n_l = \lceil N/c \rceil$. Každá skupina je následně zpracována rekurzivně pro zbývajících $d - 1$ dimenzí.

Grafická ukázka výsledku algoritmu Hilbertova statického R-stromu (vlevo) a STR (vpravo); pouze úroveň listů:



Vidíme, že v STR je méně překryvů MBR.

Podobný závěr můžeme udělat v následujícím obrázku o R-stromu (vlevo), R*-stromu (vpravo) a STR (dole):

