ŽILINSKÁ UNIVERZITA V ŽILINE

Fakulta riadenia a informatiky

Algoritmy a údajové štruktúry 2

**Semestrálna práca č. 2**

**R - tree**

František Kajánek 5ZZS15 2015/2016

# Zadanie:

Problém semestrálnej práce nadväzuje na prvú semestrálnu prácu, ktorej výsledkom bolo vytvorenie centrálneho systému na evidenciu nehnuteľností. Keďže Váš demonštračný systém sa ukázal ako kvalitný a funkčný, bol vybraný ako víťazný a pristúpilo sa k jeho dopracovaniu a nasadeniu do reálnej prevádzky.

Súkromná spoločnosť v danej krajine chce ako prvá prísť na trh so zariadením pre geodetov, ktoré bude schopné pracovať s dátami nového centrálneho systému. Z tohto dôvodu vyhlásila súťaž na vytvorenie systému, ktorý bude použitý na špecializovanom zariadení. Firma sa chce na trhu presadiť s najlacnejším, spoľahlivým a zároveň rýchlym zariadením.

Zariadenia musia umožňovať presné zameranie jednotlivých nehnuteľností podľa signálu GPS a dát získaných z centrálneho systému, ktoré budú na zariadení uchovávané. Keďže musí byť zariadenie schopné pracovať aj bez prístupu k mobilnej sieti (pokrytie signálom mobilného internetu nie je dobré) je nutné, aby si dáta z centrálneho registra dokázalo načítať (resp. aktualizovať) pri pripojení k počítačovej sieti. Vytvorte demonštračnú verziu softvéru pre informačný systém záznamov o nehnuteľnostiach, pričom je požadované, aby sa tento systém (dáta) nachádzal na pamäťovej karte (príp. pevnom disku), teda nie v operačnej pamäti. Systém má uložené len dáta o nehnuteľnostiach.

Pre každú nehnuteľnosť evidujte nasledovné údaje:

• súpisné číslo (jedinečné celé číslo v danom katastrálnom území)

• jedinečný názov katastrálneho územia (reťazec max. dĺžky 20 znakov)

• popis (reťazec max. dĺžky 60 znakov)

• zoznam maximálne 90 pozícií GPS ohraničujúcich nehnuteľnosť, pričom tento obsahuje atribúty:

• šírka (jediný znak - S alebo J)

• pozícia šírka (double hodnota)

• dĺžka (jediný znak - V alebo Z)

• pozícia dĺžky (double hodnota)

V tejto demonštračnej verzii spoločnosť požaduje, aby boli k dispozícii nasledujúce operácie:

• Vyhľadanie nehnuteľnosti podľa informácií – podľa súpisného čísla nehnuteľnosti a mena katastrálneho územia sa vyhľadajú a vypíšu všetky informácie o nehnuteľnosti

• Vyhľadanie nehnuteľností podľa polohy – podľa zadanej pravouhlej oblasti sa nájdu všetky nehnuteľnosti, ktoré do nej patria (patrí do nej aspoň jedna pozícia ohraničujúca nehnuteľnosť) a pre každú sa vypíšu všetky informácie

• Pridanie – na základe vstupných údajov pridá záznam do evidencie (nie je potrebné kontrolovať jedinečnosť súpisného čísla v danom katastrálnom území)

• Vyradenie – na základe súpisného čísla a mena katastrálneho územia vyradí záznam o nehnuteľnosti z evidencie

• Zmena – umožní meniť jednotlivé údaje o nehnuteľnosti (vrátane súpisného čísla, mena katastrálneho územia, GPS pozícií...)

Najfrekventovanejšou operáciou, ktorá sa vykonáva je vyhľadanie, a preto je potrebné, aby táto operácia pracovala s najväčšou možnou rýchlosťou. Po nájdení záznamu je potrebné zobraziť všetky evidované informácie.

Keďže použité zariadenie disponuje malou RAM pamäťou a nie je stále napájané, je potrebné, aby čo najviac údajov bolo neustále uložených na pevnom disku, resp. pamäťovej karte. Zabráňte zbytočnej redundancií dát na disku. Pre účely testovania implementujte aj operáciu zobrazenia celého aktuálneho obsahu databázy (vrátane prepojenia blokov, preplňujúceho súboru, prípadne ďalších súborov) – obsah všetkých súborov aplikácie bude možné v aplikácií sekvenčne vypísať tak, aby bolo vidieť, čo jednotlivé bloky obsahujú (vrátane ich interných atribútov). V semestrálnej práci je potrebné použiť B + strom, dynamické hešovanie alebo R strom. Implementujte efektívny manažment prázdnych blokov v súboroch. V dokumentácii uveďte presný počet prístupov do súboru pri jednotlivých operáciách (v rôznych situáciách). Nezabudnite na všeobecné požiadavky semestrálnych prác (napr. generátor na naplnenie databázy...). Export do textových súborov nie je potrebný. Program musí umožňovať opätovné zapnutie s možnosťou práce so skôr vytvorenými dátami. Dôležitou časťou je dôsledné oddelenie jadra aplikácie od GUI.

Operáciu „Vyhľadanie nehnuteľnosti podľa informácií“ implementujete iba pri použití B + stromu alebo dynamického hešovania. Operáciu „Vyhľadanie nehnuteľností podľa polohy“ implementujete iba pri použití R stromu. Pre zisk 4 bodov za priebežnú kontrolu je potrebné najneskôr v jedenástom týždni predviesť naprogramovanú a otestovanú (pomocou generátora operácii) základnú štruktúru (B+ strom, dynamické hešovanie, R strom), teda musia byť funkčné operácie vlož (1 bod), vymaž (2 body) a vyhľadaj (1 bod) na ľubovoľných dátach (bodový zisk závisí aj na kvalite implementácie).

Pracujte každý samostatne!

Maximálny bodový zisk zo semestrálnej práce bude závisieť od jej náročnosti:

• využitie dynamického hešovacieho súboru – strata 4 bodov pri odovzdaní práce

• využitie B+ stromu – strata 2 bodov pri odovzdaní práce

• nie je kvalitne, prípadne vôbec implementované mazanie (aj s korektným spájaním blokov) – strata 5 bodov pri odovzdaní práce (v tomto prípade nie je potrebné umožniť editáciu kľúčových atribútov)

# Analýza:

Je potrebná údajová štruktúra, ktorá umožňuje efektívne uloženie na pevnom médiu (HDD). Preto som zvolil R-strom.

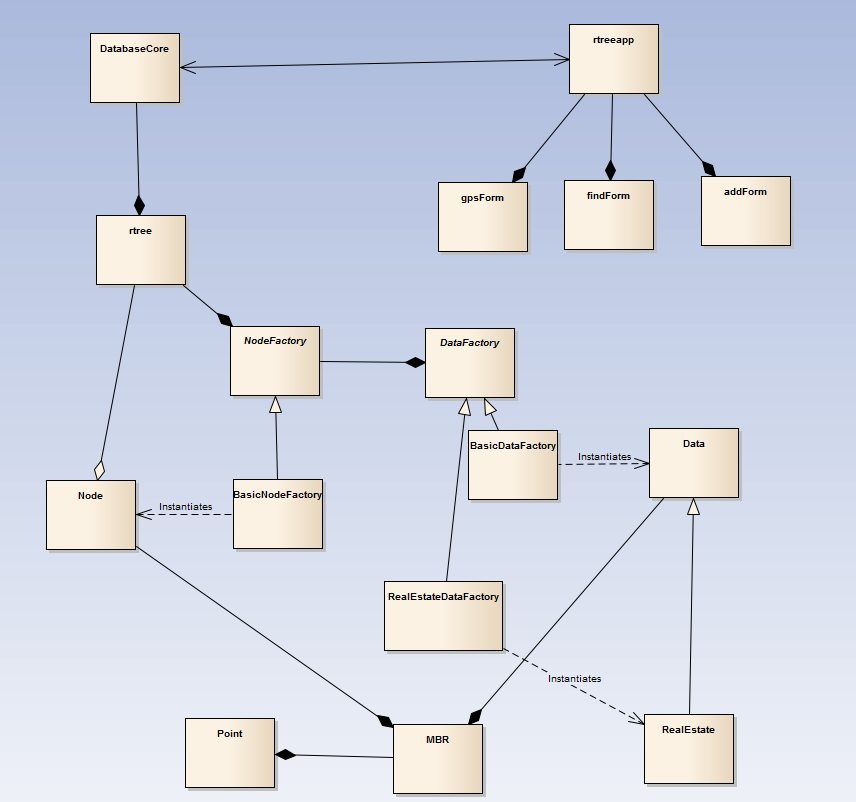
Táto údajová štruktúra musí umožňovať použitie ľubovoľných dát, ľubovoľného počtu rozmerov, a ľubovoľný typ rozmerov.

Taktiež treba implementovať zreťazený zoznam voľných blokov v súbore, samotné zreťazenie bude priamo na disku.

Na záver je potrebný generátor prvkov vložiteľných do stromu a algoritmus na bruteforce test štruktúry – t.j. naplnenie x prvkami, kontrola či x prvkov je v štruktúre, vymazanie x prvkov, kontrola či žiaden z x prvkov nezostal v štruktúre.

# Riešenie:

# Diagram tried:



# Výpočtové zložitosti:

**Vyhľadanie nehnuteľnosti podľa polohy:**

**Pridanie:**

**Vyradenie:**

**Zmena:**