

B-дерево. B+-дерево.
B*-дерево.

Цель

- Цель: разработать и описать реализацию алгоритмов B-tree, B+-tree и B*-tree. Протестировать данные алгоритмы с помощью автоматических тестов и сравнить результаты.

Основные операции

$\text{insert}(x)$ – вставить элемент x , если его ещё не существует в дереве;

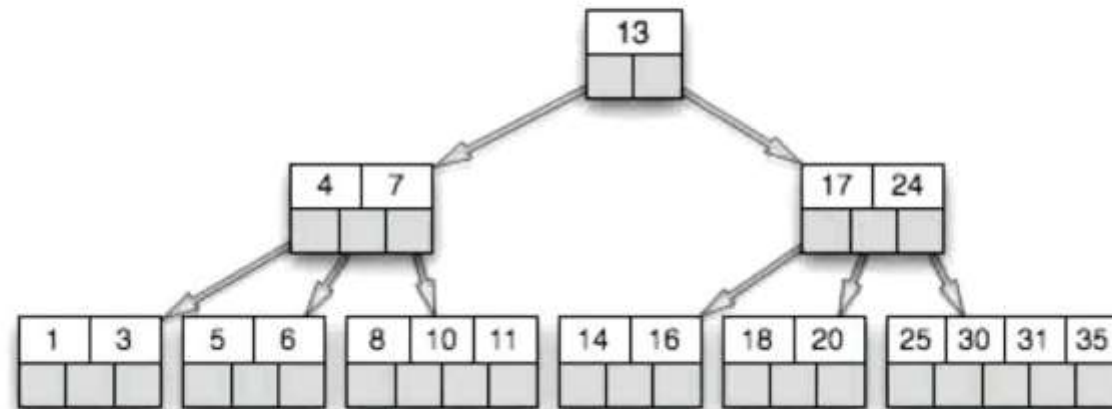
$\text{delete}(x)$ – удалить элемент x , если он есть в дереве;

$\text{search}(x)$ – найти элемент x , если он существует в дереве.

В-дерево

- В-tree — структура данных, сбалансированное, сильно ветвистое дерево поиска, предложенное Р. Бэйером и Э. МакКрейтом в 1970 году для хранения данных во внешней памяти.

B-Tree



Преимущества В-дерева

- Минимизируют количество операций чтения-записи с диском;
- Сохраняет ключи в отсортированном порядке для последовательного обхода;
- Использует частично заполненные блоки для ускорения вставок и удалений.



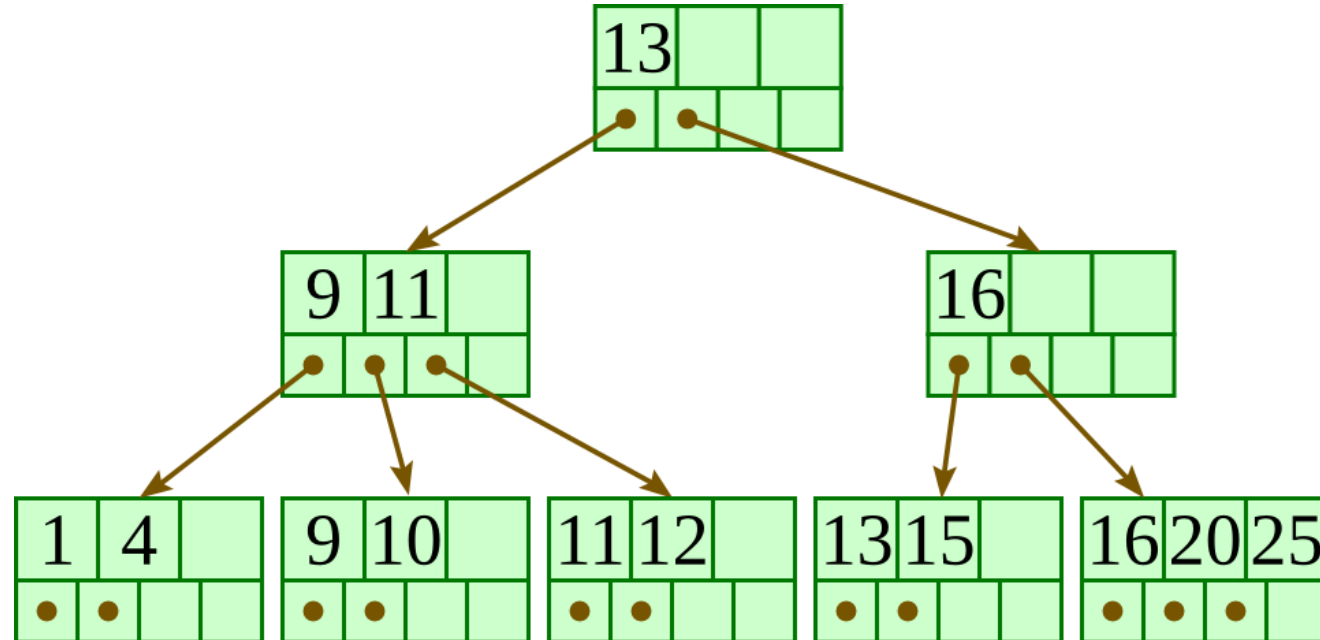
Недостатки В-дерева

- Невозможен последовательный доступ;
- Внутренние узлы долго удаляются;
- Поиск не эффективен полностью.



В+-дерево

- В+-tree — структура данных на основе В-дерева, сбалансированное n -дерево поиска с переменным, но зачастую большим количеством потомков в узле.



Преимущества В+-дерева

- Поиск эффективен, так как записи хранятся в конечных узлах;
- Удаление элементов происходит быстро;
- Возможен последовательный доступ к элементам.



Недостатки В+-дерева

- Имеет большую ширину по сравнению с высотой;
- Может быть избыточность в появлении ключей.



B^* -дерево

- B^* -дерево — разновидность B -дерева, в которой каждый узел дерева заполнен не менее чем на $\frac{2}{3}$ (в отличие от B -дерева, где этот показатель составляет $\frac{1}{2}$).

Преимущества В*-дерева

- Компактнее в сравнении с В-деревом, так как каждый узел используется полнее.



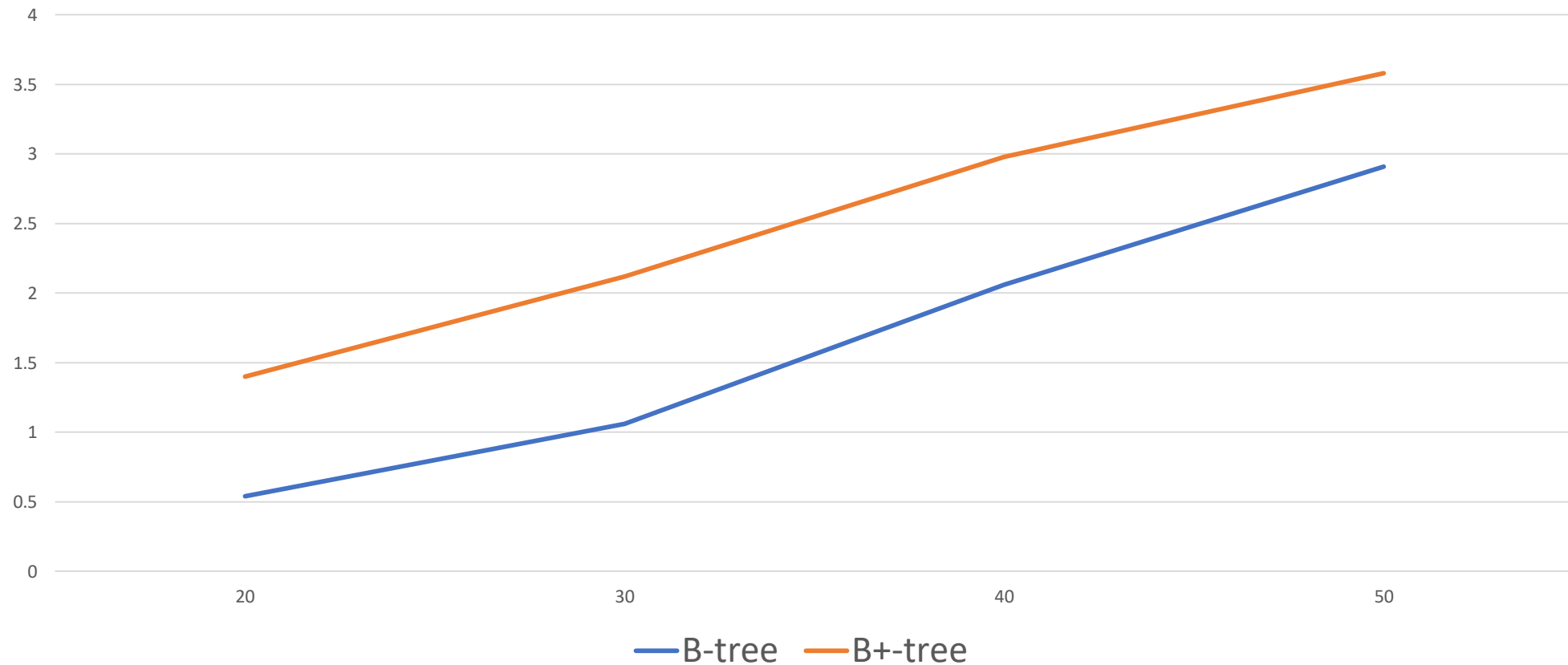
Структура узла

```
struct BTreeNode {  
    int *data;  
    BTreeNode **child_ptr;  
    bool leaf;  
    int n;  
}
```

Гипотеза

Структура данных B-tree быстрее выводит строение дерева на экран, но удобнее в использовании всё же структура данных B+-tree.

Сравнение вывода деревьев на экран в зависимости от количества элементов



Подтверждение гипотезы

Гипотеза подтверждена: в процессе выполнения практической работы было выявлено множество не особо существенных различий, от которых все же зависит производительность алгоритмов.

Реализация и тестирование

Физические характеристики текущей системы:

- Язык программирования алгоритма — C++;
- Объём написанного кода для В-дерева — 10 Кб;
- Объём написанного кода для В-дерева в строках – 365 строк;
- Объём папки с проектом — 91,2 Мб.

Заключение

- Изучено строение трёх алгоритмов, их отличие друг от друга и производительность алгоритмов на примерах. Углублены знания в области сбалансированных деревьев поиска, повышены навыки в области скорости написания кода.

