

**И. С. Колосова**

магистрант кафедры вычислительных систем и сетей

**А. В. Аксенов**

научный руководитель

## РАЗРАБОТКА И АНАЛИЗ КРИТЕРИЯ ВРАЩЕНИЯ В ДРЕВОВИДНОЙ СТРУКТУРЕ ДАННЫХ С ПРИОРИТЕТАМИ ЭЛЕМЕНТОВ

В результате раннего исследования был разработан и реализован принцип балансировки двоичных деревьев поиска с приоритетами элементов. [1] При этом были описаны алгоритмы малых вращений и критерий вращения, основанный на разнице суммы весов элементов в поддеревьях. [2]

В продолжение исследования был разработан принцип большого вращения. Реализация такого вращения состояла в последовательности двух малых вращений. Для примера, при большом правом вращении относительно узла Н (рис. 1), сначала выполнится малое левое вращение относительно левого дочернего узла - D, далее выполнится малое правое вращение относительно узла Н (рис. 2).

Экспериментально было выявлено, что прирост производительности в таком случае необъяснимо мал, а временами даже приводит к худшим результатам, чем балансировка только малыми вращениями. Критерием производительности в данном случае служил AWD (Average Weight Depth) - средневзвешенный путь до узла. Чем меньше значение AWD, тем меньше средний путь до произвольного узла от корня, следовательно, время доступа к этому узлу сокращается. Результаты экспериментов представлены в таблице.

Таблица

Результаты эксперимента

Количество входных символов	67	43520	54059	86016	213847
Количество узлов дерева	10	251	256	256	256
	AWD				
Балансировка только малыми вращениями	2,43	3,13	7,14	4,02	7,18
Балансировка большими и малыми вращениями	2,57	3,41	7,06	4,15	7,07

В данной работе произведена попытка объяснения полученных результатов. Сравним соотношение весов поддеревьев до и после малого и большого вращения. Для начала представим идеально сбалансированное дерево (см. рис.1).

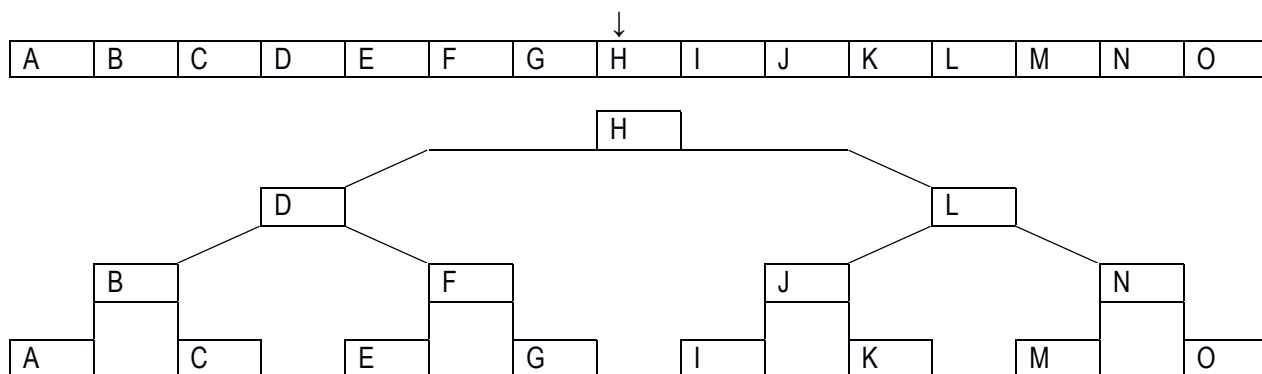


Рис. 1. Сбалансированное дерево

Соотношение весов поддеревьев в нем определяется формулой (1). Примерное равенство означает возможность наличия дочерних узлов на более низких уровнях. При правильном балансе их вес

незначителен и не влияет на приведенные ниже соотношения.

$$\overline{AE} + FG \cong \overline{IO} \quad (1)$$

При сериализации дерева, то есть выстроении его узлов в отсортированных массив, наглядно видно соотношение весов.

После некоторого изменения весов и достижения некоего дисбаланса, выполнится большое вращение, и дерево поменяет свою структуру (см. рис. 2).

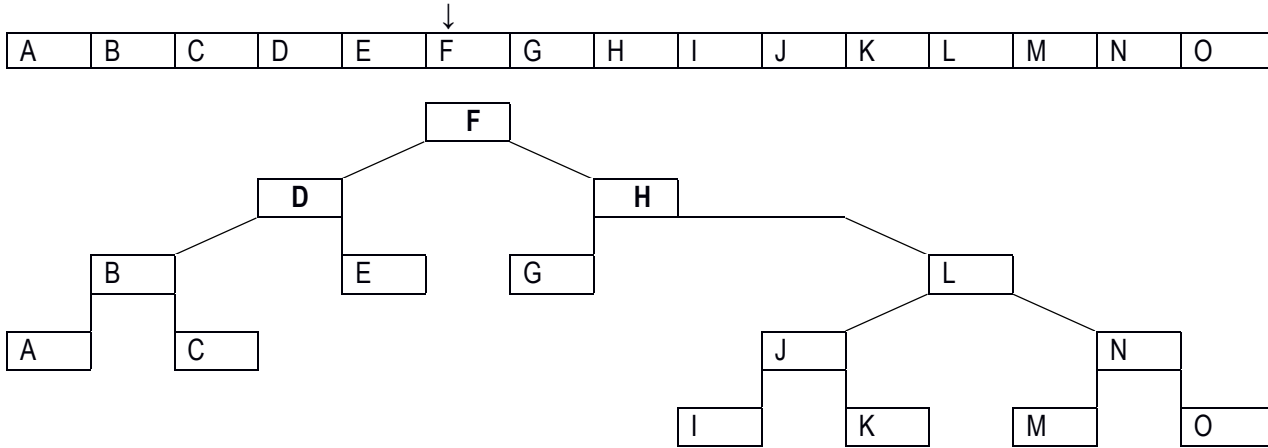


Рис. 2. Структура дерева после большого вращення

Жирным выделены узлы, которые требуется загрузить из памяти при выполнении вращення. Соотношение весов поддеревьев в таком дереве будет определяться формулой (2).

$$\overline{AE} \cong \overline{IO} + GH \quad (2)$$

Значение дисбаланса в дереве для совершения вращення должно достичь величины (3)

$$G + \frac{F+H}{2} + \varepsilon \quad (3)$$

где  $\varepsilon$  – величина, не позволяющая дереву выполнять одинаковые последовательные левые и правые вращення при добавлении элементов в разные поддеревья.

Аналогично рассмотрим ситуацию для малого вращення (рис.3).

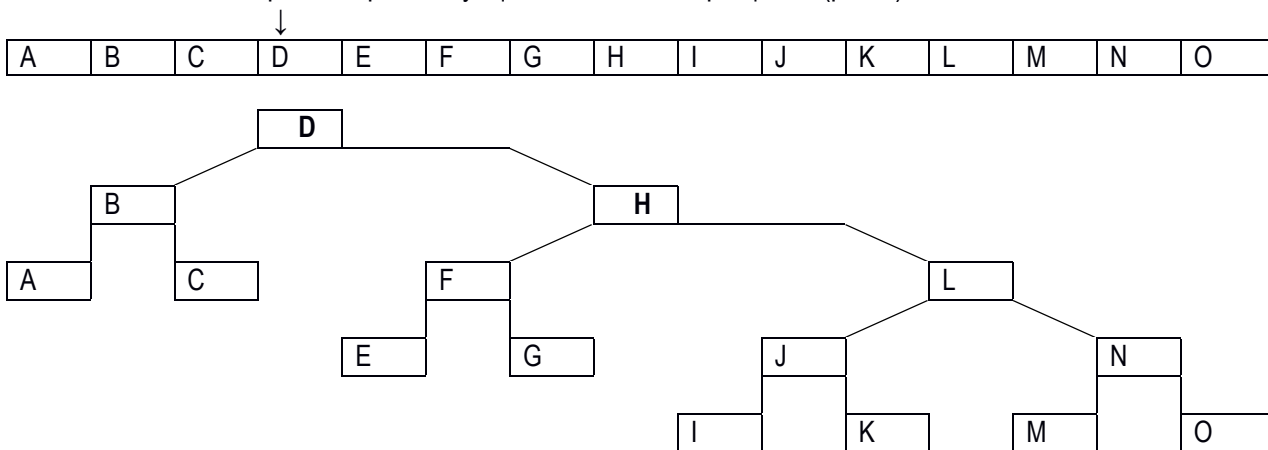


Рис. 3. Структура дерева после малого вращення

В таком случае, соотношение весов поддеревьев в первоначальном дереве будет определяться формулой (4),

$$\overline{AC} + DEFG \cong \overline{IO} \quad (4)$$

а после выполнении вращения – формулой (5).

$$\overline{AC} \cong EFGH + \overline{IO} \quad (5)$$

Дисбаланс в дереве, в этом случае будет равен (6).

$$EFG + \frac{D+H}{2} + \varepsilon \quad (6)$$

Значение (3) очевидно меньше, чем значения (6), следовательно, большее вращение спровоцировать легче, чем малое, а после его выполнения необходимость в малом может отпасть в связи с изменением структуры дерева.

Может показаться, что механизм малых вращений избыточен и не будет использоваться при соблюдении вышеприведенных принципов.

Для увеличения быстродействия балансировки, оговариваются условия работы с памятью. При вращениях используются только те узлы, которые уже были загружены в память при поиске. Это также накладывает ограничение на оптимальность структуры дерева.

Представим ситуацию, что большое количество элементов последовательно добавляются только в ветку AC идеально сбалансированного дерева (см. рис. 1). Таким образом, сначала выполнится большое вращение, а потом малое, что займет больше времени, а также приведет к менее выгодной структуре дерева. Также при выполнении большого вращения будут затронуты узлы, которые не были пройдены при поиске, значит, их придется загружать из памяти, что займет время.

Рассмотрим направление движение по дереву при балансировке. Баланс на нижних уровнях дерева не влияет на баланс верхних уровней, однако влияет на возможность вращения в ту или иную сторону. Однако, баланс на верхних уровнях влияет на расположение узлов и баланс на более низких уровнях. При сравнении рисунков 2 и 3 видно, что после большого вращения, в отличие от малого, баланс правого поддерева сильно нарушен. Это нарушение структуры неоправданно, так как более тяжелые узлы находятся в левом поддереве. К тому же, после такого вращения, при поиске узла потребуются спускаться далее в левое поддерево, следовательно, баланс правого поддерева восстановить не удастся. Поэтому, при балансировке на спуске, после больших вращений баланс нижних уровней может быть нарушен. Кроме того, применение больших вращений на спуске приведет к постоянной смене направления движения по дереву, что связано с их реализацией.

Разумнее выполнять балансировку на спуске, в таком случае, баланс более нижних уровней не будет поврежден, а также понадобится лишь один проход по дереву для добавления элемента и одновременной с этим балансировки, что вдвое сократит время выполнения ее время выполнения.

В результате данного исследования было выявлено, что использование больших вращений нерационально при данных ограничениях на быстродействие по причине загрузки из памяти большого количества узлов, а также невозможно при балансировке на спуске.

### **Библиографический список**

1 Аксенов А.В., Колосова И.С. Анализ эффективности балансируемых структур данных с приоритетами элементов, Сборник докладов Научной сессии ГУАП, посвященной Всемирному дню космонавтики, СПб, ГУАП, 2013 г.

2 Аксенов А.В. Балансировка деревьев в статистической модели энтропийного кодировщика, Сборник докладов Научной сессии ГУАП, посвященной Всемирному дню космонавтики, СПб, ГУАП, 2011 г.