

18. Оценка сложности быстрой сортировки при несбалансированном разбиении на постоянные доли

В случае быстрой сортировки при разбиении массива на постоянные доли, например, 70% и 30%, оценка сложности алгоритма будет несколько отличаться от стандартной, где разбиение происходит пополам.

1. Рекуррентное соотношение

При разбиении массива на доли p и $1 - p$ (где $0 < p < 1$), на каждом уровне рекурсии один из подмассивов будет содержать примерно pn элементов, а другой — $(1 - p)n$ элементов. Рекуррентное соотношение для такой ситуации будет выглядеть следующим образом:

$$T(n) = T(pn) + T((1 - p)n) + O(n)$$

где: - $T(pn)$ — время, необходимое для сортировки подмассива размером pn , - $T((1 - p)n)$ — время для второго подмассива размером $(1 - p)n$, - $O(n)$ — время, необходимое для разбиения массива и размещения элементов относительно опорного.

2. Оценка сложности

Для случая, когда массив делится на доли p и $1 - p$ с постоянными значениями p , сложность будет аналогичной стандартному случаю быстрого сортирования, но с дополнительными коэффициентами, определяемыми долями разбиения. Ожидаемая оценка сложности в среднем случае при таком разбиении будет:

$$T(n) = O(n \log n)$$

3. Сложность при неравномерном разбиении

Рассмотрим, например, разбиение на доли 70% и 30%. В этом случае рекуррентное соотношение принимает вид:

$$T(n) = T(0.7n) + T(0.3n) + O(n)$$

Для анализа такого рекуррентного соотношения можно ожидать, что сложность будет стремиться к $O(n \log n)$, даже если доли разбиения отличаются от равных.

4. Влияние долей разбиения

Хотя доли разбиения не равны, сложность по-прежнему остаётся на уровне $O(n \log n)$ при достаточно сбалансированных долях. Однако если разбиение становится сильно несбалансированным, например, 99% и 1%, сложность может ухудшиться до $O(n^2)$, поскольку один из подмассивов будет оставаться практически неизменным, а другой будет значительно уменьшаться. Но при разумных значениях долей, например 70% и 30%, сложность будет оставаться $O(n \log n)$.

5. Заключение

Таким образом, при разбиении на постоянные доли, например, 70% и 30%, сложность быстрой сортировки будет в среднем $O(n \log n)$, как и в случае с равными долями. В худшем случае, при крайне несбалансированном разбиении, сложность может ухудшиться до $O(n^2)$, но в практике алгоритм сохраняет свою эффективность при сбалансированных долях.