Тураев Тимур, 504 (SE)

15.10.2013

**6.5\*. Разбиение неотсортированных постов на Хабрахабре.**

Можно поступить следующим образом: ищем за линию k-ю порядковую статистику. Затем бы нам хотелось найти все посты, время которых не меньше, чем найденная статистика, так и сделаем: только делать это будем вместе с алгоритмом поиска k-квантилей, т.е. в момент поиска среднего k-квантиля (см. соответствующую задачу): в том алгоритме все равно будет линейный поиск по всему массиву – ну в этот момент и найдем все элементы не большие найденной статистики – то есть заполним первую страницу.

Затем отсортируем посты. Очевидно, что время на генерацию этой страницы равно **.**

В итоге мы нашли первую страницу, плюс «поделили» массив.

Дальше будем искать конец второй страницы уже в левой половине массива (2k статистика), при этом повторяя второй шаг поиска к-квантилей: делим левую часть массива на две половины, делая линейный поиск, но уже по половине массива и таким образом генерируя вторую страницу. Снова не забудем ее отсортировать. Ясно, что сложность будет та же . И так далее.

Получается, мы объединяем поиск статистик с поиском к-квантилей, причем второй шаг облегчает первый: не нужно делать просмотр всего массива.

Таким образом сокращается суммарное время на генерацию всех страниц: формула будет очень похожа на формулу поиска к-квантилей: O(n)+2\*O(n/2) + 4\*O(n/4)+… = O(N\*logN), так как делается два линейных поиска по всему массиву (сначала статистику, потом “средний” квантиль), потом два поиска по левой половине( то же самое) и по правой – когда до нее дойдет алгоритм, 4 раза по четвертям и так далее. Плюс, не забыть про суммарное время на сортировку, но это слагаемое поглощается: