Medians of Median Analysis
Chen Yuhao
NetId Yuhaoc2
Graduate Student

Q1.

Assuming that the length is repeatedly divisible by all the cuts m, we shall start with m=7

Similarly, we will have the first part in the master theorem to be $T(\frac{n}{7})$ for finding the true median of the $\frac{n}{7}$ medians.

Of the $\frac{n}{7}$ groups half of them have their median smaller/bigger than the median of the median, $\frac{n}{14}$. In each group there are 4 elements that are smaller than the pivot. The extreme case would be that the in these groups containing one group with less than seven elements and the one with the true medians in it.

Thus if we look at the elements number smaller/bigger than the true median, the formula will be $4(\frac{1}{14}n)$ giving $\frac{2n}{7}$ elements smaller/bigger than the median of medians.

Therefore, the worst-case will be a $\frac{5n}{7}$&$\frac{2n}{7}$ cut.

Therefore, we have $T(n) \leq T\left(\frac{n}{7}\right) + T\left(\frac{5n}{7}\right) + cn \leq kn$, suppose $T(n) = \mathrm{k}n$ meets the condition, $\mathrm{k} \geq \frac{cn}{\frac{n}{7}}$ will satisfy the condition. that there exists constant k such that $T(n) \leq T\left(\frac{n}{7}\right) + T\left(\frac{5n}{7}\right) + cn \leq kn$, giving that the algorithm running time $T(n) \in O(n)$. For an active constraint,

Similarly, when b=9, the minor side will be $5\left(\frac{n}{9*2}\right)$, giving $\frac{13}{18}n$ in the worst case, therefore, $T(n) \leq T\left(\frac{n}{9}\right) + T\left(\frac{13}{18}n\right) + cn \leq kn$ will require a k meeting the condition that $k \geq \frac{cn}{\frac{3}{18}n}$ will satisfy the condition, for example, $n = 180, k \geq 9c$.

When b=11, $k \geq \frac{cn}{\frac{2}{11}n}$, b=13, $k \geq \frac{cn}{\frac{5}{26}n}$. For a general odd cut m, we will have $T(n) \leq T\left(\frac{n}{m}\right) + T\left[\left(1 - \frac{\frac{m+1}{2}}{2m}\right)n\right] + \mathrm{c}_m n = T\left(\frac{n}{m}\right) + T\left[\left(\frac{3m-1}{4m}\right)n\right] + \mathrm{c}_m n \leq kn$, similarly we have a active condition on k: $\mathrm{k} = \frac{\mathrm{c}_m n}{(1-\frac{1}{m}-\frac{3m-1}{4m})}$

Therefore, the recursive formula will be for $\mathrm{b} > 3, T_m(n) = \frac{\mathrm{c}_m n}{\frac{\frac{m-3}{2}}{2m}} = \frac{4m}{m-3}\mathrm{c}_m n.$

Q2.1

Following the formula provided in the assignment, we have

$$c_m = \tilde{c} * \log(m)\,, m\epsilon\{7,9,11,13\}$$

When m=7, ignoring the constants in Q1 when n is big, we will have $\frac{5n}{7}\&\frac{2n}{7}$ cut, $T_7(n) = 7c_7 n = 7\tilde{c} * \log(7) * n$

When m=9, we will have $\frac{13n}{18}\&\frac{5n}{18}$ cut and $T_9(n) = 6c_9 n = 6\tilde{c} * \log(9) * n$.

When m=11, we will have $T_{11}(n) = 5.5c_{11}n = 5.5\tilde{c} * \log(11) * n$.

When m=13, we will have $T_{13}(n) = 5.2c_{13}n = 5.2\tilde{c} * \log(13) * n$.

As we put the term c_m into the formula of $T_m(n)$, we will have:

$$T_m(n) = \frac{4m}{m-3}c_m n = \frac{4m}{m-3}\tilde{c} * \log(m) * n$$

Q2.2

Using the derived equations in Q2.1, the minimum $T_m(n)$ is dependent on the constant for n. If we compare the constants for $m\epsilon\{5,7,9,11,13\}$, we can see that when m=9, the constant has the smallest value. Therefore, there must be an optimal m for $T_m(n)$ to be smallest.

We can show this conclusion holds in general. That under given n, we can see that the slope function $\frac{4m}{m-3} * \log(m)$ is a tick-shaped curve, and have its minimum at m=9 or m=11, at these two points the function value will be extremely close to each other. The function plot is shown below.