

# Lab1 说明文档

经过对 insertion sort 和 merge sort 算法的单独测试，得到 excel 表中的数据。

根据数据生成的图像，insertion sort 符合  $n^2$  的时间复杂度，merge sort 符合  $n \log n$  的时间复杂度，并且计算出  $c_1 = 2.52e-6$   $c_2 = 1.14e-4$

第三问的理论分析：f(n, k)的推导过程 k 的理论值为 319

combine sort

$$T(n) = 2T\left(\frac{n}{2}\right) + C_2 n$$

$C_2 n$   
 $C_2 n$   
 $C_2 n$   
 $C_2 n$   
 $C_2 n$

$\rightarrow \frac{n}{k} T(k)$

$$C_2 n \lg \frac{n}{k} + \frac{n}{k} T(k) = C_2 n \lg \frac{n}{k} + \frac{n}{k} \cdot C_1 k^2$$
$$= C_2 n \lg \frac{n}{k} + C_1 k n = f(n, k)$$
$$\frac{\partial f}{\partial k} = C_2 n \cdot -\frac{n}{k^2} \cdot \frac{1}{\frac{n}{k} \ln 2} + C_1 n$$
$$= n \left( C_1 - \frac{C_2}{k \ln 2} \right) \quad k = \frac{C_2}{C_1 \ln 2}$$

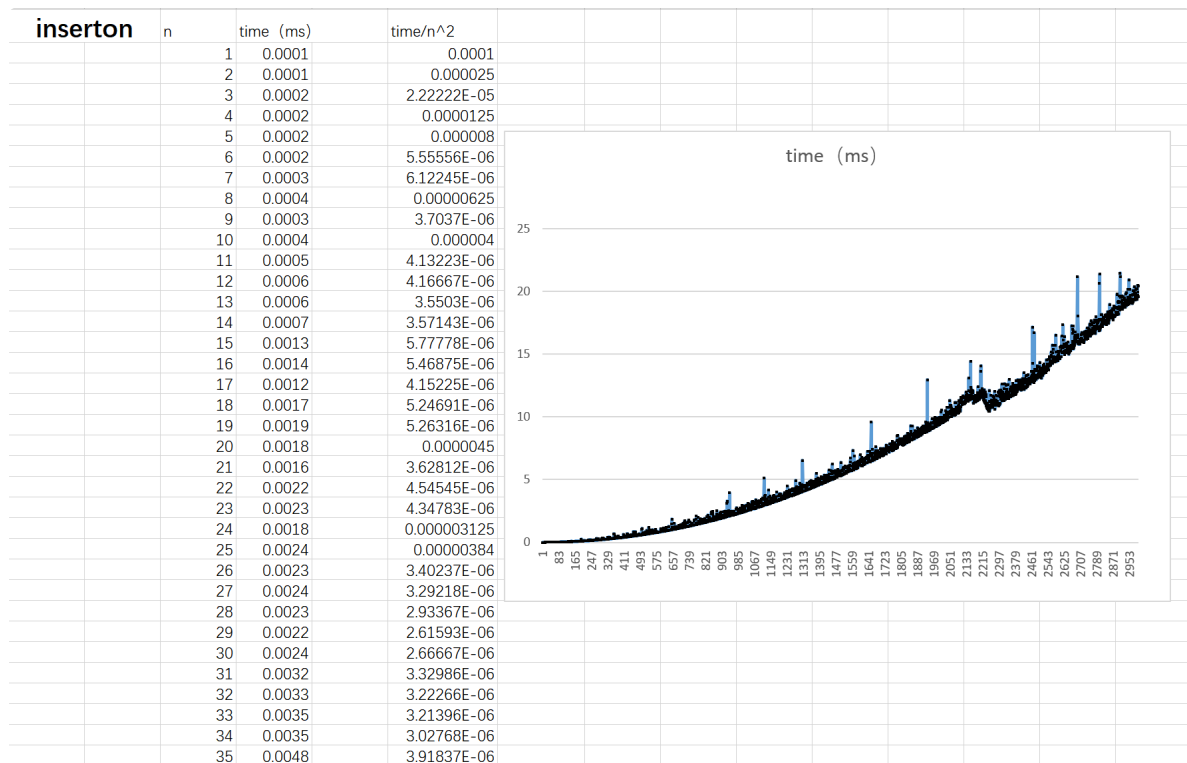
测得理论值  $C_1 = 2.52 \times 10^{-6}$   
 $C_2 = 1.14 \times 10^{-4}$  }  $\Rightarrow$  k 理论值为 319

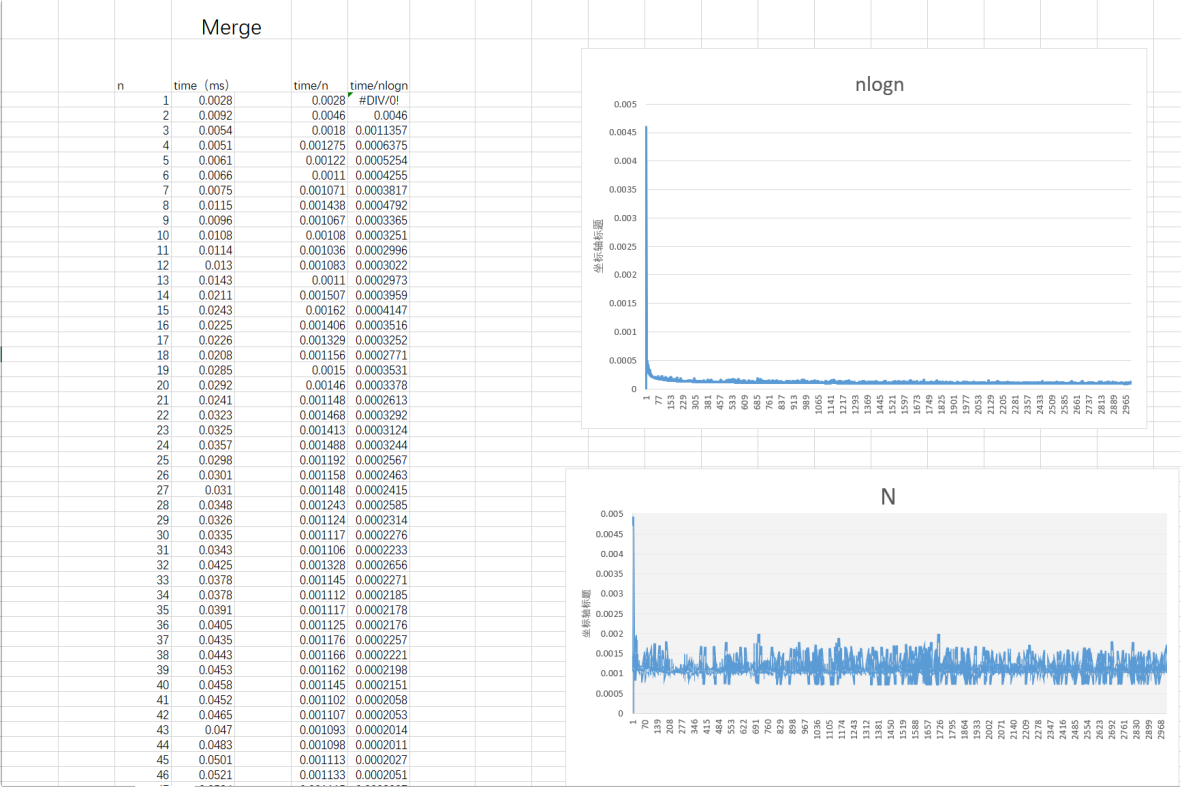
第四题中，经过实验测得 k 的实验值（n = 3000）是 670（数据在 excel 表中）

并且 SelectK()函数采用的方法是枚举法

因为：事实上，经过多次实验发现最佳 k 的波动较大，并且难以寻找其中的规律，于是采取枚举法并且多次实验取平均值，使得找到的 k 更加精确。

附：Excel 表格截图（部分）





		Combine	best k
			446
			777
			571
			856
			723
			792
			704
			439
			708
			769
			840
			749
			669
			881
			577
			743
			730
			657
			853
			428
			731
			415
			717
			741
			780
			740
			877
			721
			692
			214
			360
			508
			708
			594
			465
			665
			531
			887
			872
		最优的k的平均值	670
		n=3000	

