Supplement Explanation

107030028 劉騏鋒

環境:python3.7

```
import random
from math import log10
```

補充說明:

做完頻率分析後做可讀性的計算

可讀性分析利用資訊理論的 information entropy 來計算，以四個字母為一組，將四個字母組成的 quadgrams 的出現機率取 log 然後將資訊加起來。加完的結果是負的，負越多代表可讀性越低，負越少則可讀性越高。

exp. ATTACK:

$$\log\left(p(\mathrm{ATTACK})\right) = \log\left(p(\mathrm{ATTA})\right) + \log\left(p(\mathrm{TTAC})\right) + \log\left(p(\mathrm{TACK})\right)$$

演算法會隨機對調 key 的兩個字母，如果對調後可讀性提高則繼續延續下去對調。循環 1500 次後可讀性會來到峰值，此時將進行句首分析，如辨別不出五個英文字則結果無效，重新迭代，若找的到五個英文字母則為可行解。

為避免只找到局部最佳解，演算法會進行多次迭代已找到全域最佳解，給的密文越短需要迭代的次數越多。