1 关于按键使用均衡性和输入效率的理解

1.1 按键使用均衡性

在第一问中，我统计出了使用全拼输入法输入给定中文文章时的26个字母按键的使用频率，并绘制出热力图（图-1）。

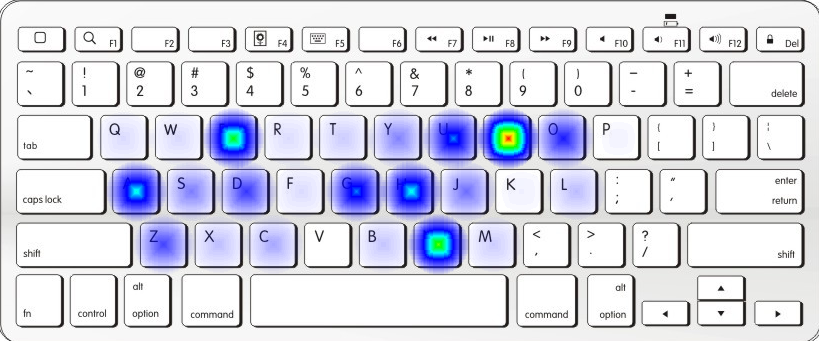


图- 1 26字母键使用热力图

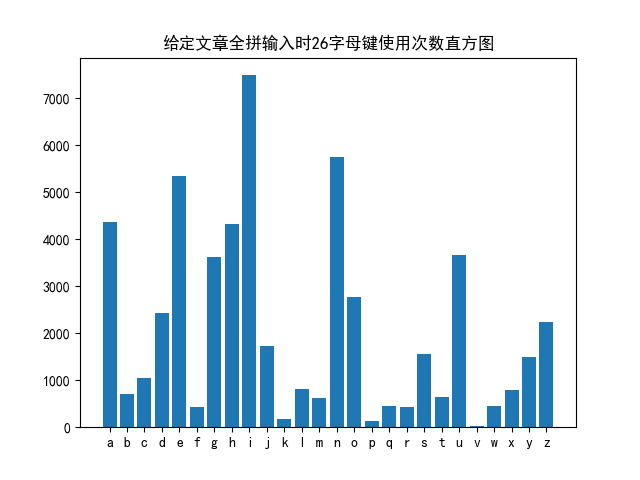


图- 2 26字母键使用次数直方图

通过对统计结果进一步计算处理，可以得到：26字母键的**总使用次数**为53489次，按键的**平均使用次数**为2057次。其中使用次数**低于1000次**的按键有12个，占46.15%；使用次数**高于4000次**的按键有5个，占19.23%。

因此可以发现在进行全拼输入时，不同按键的使用次数差距较大，出现了部分按键的使用次数较多、而部分按键的使用次数较少的情况，不同按键的使用分布情况不均匀。所以在我看来**按键使用均衡性**是用来衡量在输入某一指定文章时不同按键的使用分布的均匀程度；若不同按键的使用频率之间的差距越小，则其按键使用均衡性便越好。

1.2 输入效率

继续对统计数据进行计算处理，可以得到：指定文章的总字数为18118字，**逐字输入的平均按键次数**为2.95次/字。这里的逐字输入的平均按键次数指的是使用全拼输入给定文章时平均每输入一个字所用的按键次数。很明显，如果逐字输入的平均按键次数越小，输入同一篇文章所用的总按键次数也越小，输入效率便越高。

此外，考虑到目前主流的键盘打字方式，其指法大致如图-3所示：



图- 3 打字指法示意图

初始状态下，除去两个大拇指之外的其它手指是放在‘A’、‘S’、‘D’、‘F’、 ‘J’、‘K’、‘L’按键上，如果遇到这7个字母则直接敲击即可，而其他的字母则需要移动手指去进行敲击。因此这里考虑，如果能减少手指的移动次数，那么便可以在一定程度上提高打字输入的效率。

2 量化指标的确定

2.1 按键使用均衡性的量化指标

根据1.1中的分析，这里决定使用26字母按键数据的**标准差**来作为按键使用均衡性的量化指标。

假设为全拼输入给定文章时的26个字母键的使用次数数据集，为该数据集的平均值，那么标准差的计算公式为：

（1）

根据前面的分析可知，如果26字母按键次数的标准差越小，则按键使用均衡性越高。

2.2 输入效率的量化指标

根据1.2中的分析，这里决定使用**逐字输入的平均按键次数**以及**平均手指移动次数**来作为输入效率的量化标准。

1. **逐字输入的平均按键次数**

假设全拼输入给定文章时所用的总按键次数为，给定文章的总字数为，那么根据1.2的说明可得逐字输入的平均按键次数的计算公式为：

（2）

1. **逐字输入的****平均手指移动次数**

假设为全拼输入给定文章时的26个字母键的使用次数数据集，其中表示某一字母键的使用次数；为键盘上26个字母键的手指移动次数数据集，其中表示使用某一字母键时的手指移动次数。

那么可得逐字输入的平均手指移动次数的计算公式为：

（3）

另外根据图-3所示的指法示意图，这里可以假设对于按键‘A’、‘S’、‘D’、‘F’、‘J’、‘K’、‘L’来说，其手指移动次数均为0；而其它的按键则紧挨着这8个按键，所以假设它们的手指移动次数均为1。

根据前面的分析可知，如果逐字输入的平均按键次数和平均手指移动次数越小，则输入效率越高。

3 针对全拼输入方案的评价结果

这里使用2中的量化评价标准对附件文章的全拼输入方案进行统计计算，从而计算出评价指标（26字母按键数据的标准差、逐字输入的平均按键次数和平均手指移动次数）的数据。计算过程使用python语言进行编程实现，运行结果如图-4所示：

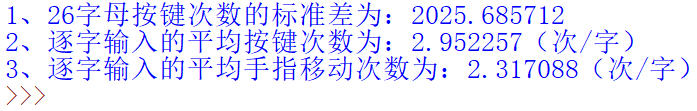


图- 4 运行结果