



链滴

关于文章排序算法引发的思考

作者: [someone61489](#)

原文链接: <https://ld246.com/article/1615640323859>

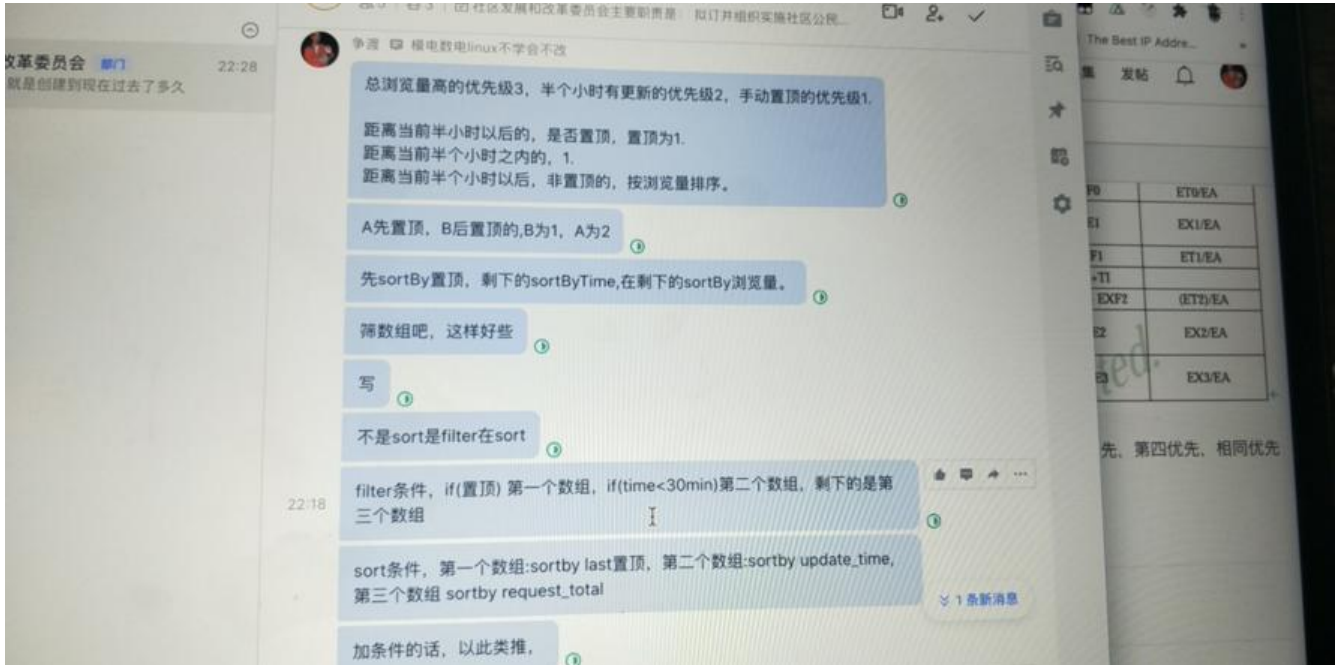
来源网站: 链滴

许可协议: [署名-相同方式共享 4.0 国际 \(CC BY-SA 4.0\)](#)

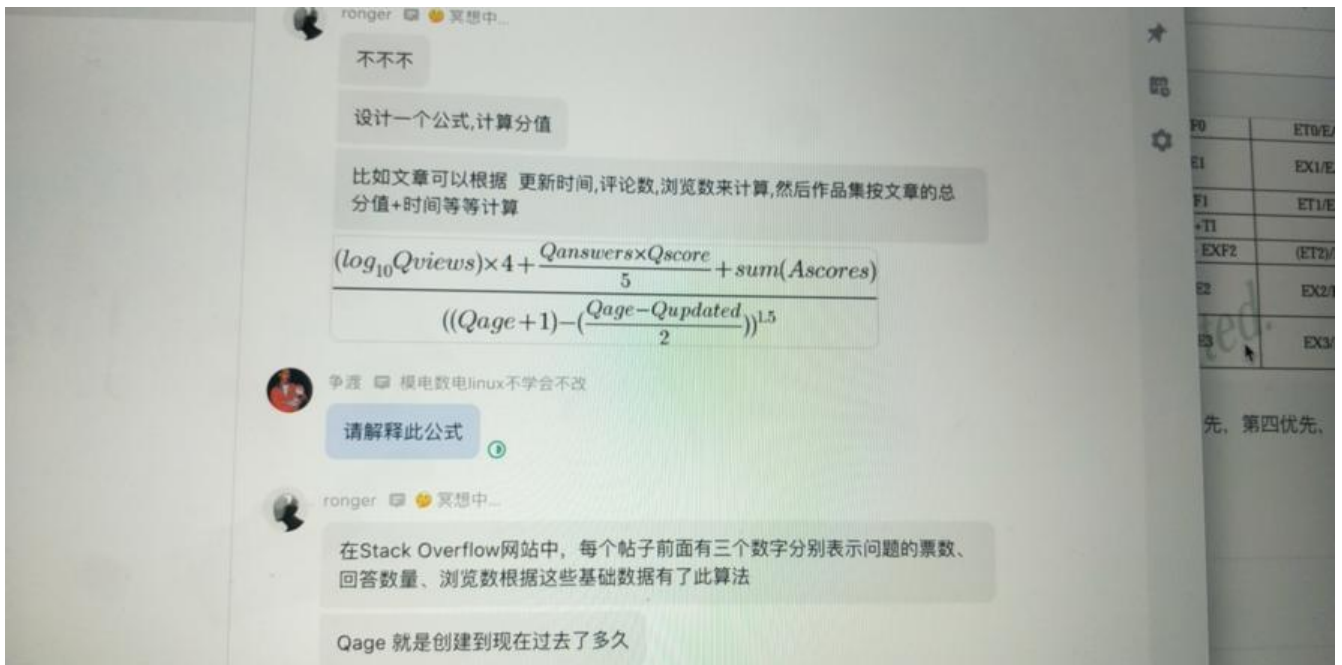
Question

事情的经过大概是这样子

- 前几天R社站长问作品集应该怎么排序?
- 作为写过几天增删改查的我就叭叭了一堆, 先这样这样, 在那样那样, 最后这样这样, 就排好了



- 然后站长贴了一个排序算法, 不同于冒泡选择排序算法, 经过一阵思考和分析后, 感觉妙啊 ~ ~ ~



- 同时也觉得自己太过头发短见识短, 孤陋寡闻了, 沉浸在各种框架里殊不知解决核心问题的还是数

Answer

b3的markdown editor数学公式支持 <https://katex.org/> Katex

$$\frac{(\log_{10}Qviews*4)+\frac{Qanswers*Qscore}{5}+\sum(Ascores)}{((Qage+1)-(\frac{Qage-Qpdated}{2}))^{1.5}}$$

- Qviews = question views, 通常这个数会很大, 几万至几十万
- Qanswers = question answers, 这个数相对于浏览量来说就非常小了, 5-10 等等
- Qscore = question score, 评分, 完全是搜索问题的人觉得这个回答值多少分就点一分, 可为负
- sum(Ascores)不知道, 不考虑
- Qage = question ages, 问题创建了多久了, 年龄, 可能是3年, 10年, 5年等
- Qupdated = question updated, 问题最近一次更新是多久了, 可能是1年前, 2年前等等

好, 我们带入具体数据分析一下

- 假设浏览量为10W, 回答数量为10个, 分数为5分, 10年前, 最后一次更新5年前
- log10 很好的控制了 10W 这么大的数字, 以10为底10W的对数为 4,最后得出16, 10W直接缩减了16这个数, 可见log针对大浏览量是很好用的一种缩小数的算法.
- 在看第二部分, (回答数量*分数)/5, 带入数据最后得出10, 假如5个, 2分呢, 数就算出了2, 回答和分数相互牵制作用非常大, 回答多可能分数是负的, 标志这个问题应该属于灌水问题, 回答多分数说明很多人碰到了这个问题, 所以这个地方大, 分子会大, 整体排序优先级会上去等等...
- 占整体分值比重最高的就是下面的时效数, 时间线拉的越长, 越长时间的问题不经过更新, 代表这问题已经过时了, 也就没有必要排在前面了, 所以时间线拉的越长, 分母会越来越大, 整体排序分值会降的很低

不得不说, 算法排序, 比机械排序优秀太多了, 可以带入许多理念, 许多实际的问题通过一个公式解掉了。就挺妙的。