# EdgeRank

## 介绍

EdgeRank 用于当某个用户查看他的新鲜事时, 决定这些新鲜事先后顺序的一个排序算法. 算法核心是每个事件对这个用户而言的权重 E, 其计算公式是 E = u\*w\*d, 其中

1. u, 事件生产者和观察者之间的亲密度
   * 亲密度主要指的是互动情况, 互动越频繁亲密度越高
   * 互动可以是单向的, 比如 A 经常关注 B, 而 B 不怎么关注 A, 那 u(a,b) 和 u(b, a) 的权重不一样
2. w, 边权重 (主要是事件的类型)
   * 边权重主要受事件类型影响, edgerankchecker 给的关于点击跟其他事件的关系可以大概说明
     + i. Avg Clicks Per Like: 3.103
     + ii. Avg Clicks Per Comment: 14.678
     + iii. Avg Clicks Per Impression: 0.005
3. d, 时间衰减因子
   * 比较少见到讨论时间衰减因子怎么做的, 猜测是类 log 变化
   * 如果将 EdgeRank 公式只保留 d, 则 EdgeRank 退化到经典 timeline 模式

## EdgeRank思想在知乎上的应用

以知乎为例，知乎用户刘看山，他关注了100个人，30个专栏，10个话题。在他打开知乎的一瞬间，将有100个最新动态等着他，你要怎么给这100个新鲜事排序？

从下面几个方面思考：

1. 这些新鲜事对用户而言质量各有高低，喜欢某人的回答而不喜欢某人的回答
2. 确定一个变量作为优化目标，正面目标（点开、赞同、感谢、长时间浏览）和负面目标（没有帮助，短暂停留）

2. 预测用户在哪个新鲜事上更可能触发我们的目标变量，算法(LR,SVM)和feature