互联网的结构

2021年10月31日 11:19



网络计量学: 旨在定量分析网络现象, 链接分析成为主要内容



▲▲ 链接的类型

入链是指指向某页面的链接。一般而言,这个链接应该 是外部指向某特定集合的页面。

出链是指从某页面指出的链接。一般而言,这个链接应 该指向某个特定集合以外的页面。

自链是指从某页面指向该页面自身的链接。

互链通常是指两个不同网站之间的链接、也叫站间链接。

共入链(CO-LINKED): 如果两个网页都含有来自第三个页面的入链,则这两个页面共入链。

共出链(CO-LINKING): 如果两个网页都含有指向第三个网页的出链,则这两个网页共出链。



• 地域分布不均衡

- > 东部地区的网络信息资源大小居全国领先
- > 中部、东北部地区次之
- ▶ 西部地区远远落后于中东部地区
- 时间分布不均匀
 - > 信息更新不够及时
 - > 信息资源的时效性差

网络超链接的宏观结构实际上是现实世界中 权利、商业关系和文化心理复杂作用的结果。

Internet自发产生有序的超链接结构这一事实表明, Internet网络中微观上看似杂乱无章的随机行为, 如网络超链接的设置, 经过网络传播系统, 复杂的人机互动, 产生了聚合效应, 在宏观上可以涌现出大尺度的有序结构, 这证明Internet并非均匀一致的随机网络。

深网

- > 存储在网络数据库里
- ▶ 不能通过超链接访问的资源集合
- ▶ 需要通过动态网页技术访问
- ▶ 网站内容也受登录信息保护或者要求付费



深网和表层网页信息搜索的区别:

- 1. 搜索结果不同,表层网页是web网页,深网是结构化的数据记录,深网数据获取的力度更小,能够更好的满足用户的需求,无需再次检索
- 2. 搜索引擎的查询接口不同: 表层网页较为简单,只需输入一个或多个关键字,深网带有复杂的领域查询接口
- 3. 查询结果排序方式不同: 表层网页根据预先设定的评价标准,即计算网页与查询结果的相似度,根据相似度由大到小排序,深网的查询结果按照多种方式进行统计分析和排序