

社区发现算法-局部拓展



作者 八刀一闪 (/u/6979be4ade9b) +关注

2016.09.21 22:41 字数 562 阅读 35 评论 0 喜欢 2

(/u/6979be4ade9b)

简介

局部拓展的方法是社区发现中的一大类方法，并且现在也比较活跃。这些方法的一个基本的假设就是社区是围绕着一一些中心结点形成的，它们一般都是向当前社区中添加或删除节点来优化一个函数。

LFM算法

什么是社区

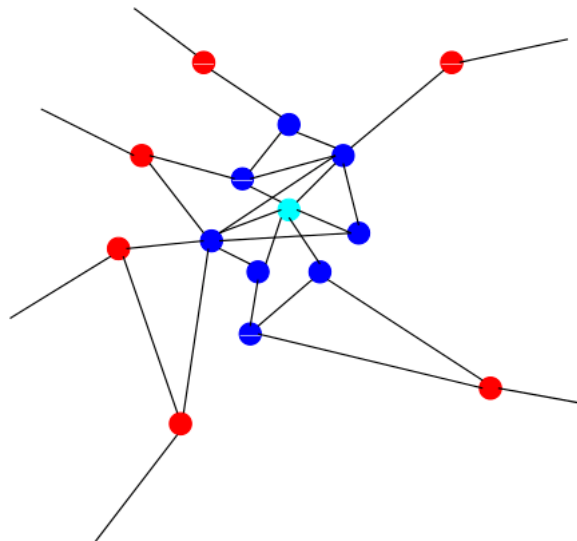
LFM算法首先定义出可以衡量一组结点连接紧密程度的函数，fitness。它的定义如下：

$$f_G = \frac{k_{in}^G}{(k_{in}^G + k_{out}^G)^\alpha},$$

其中k_in表示这些结点内部的度数，也就是内部边的2倍；k_out表示与外部结点相连的度数。那么，一个社区由一组能够使fitness函数最大的结点组成，也就是再向这个社区中添加任何邻居结点都会使fitness减小。而一个结点对于这个社区的fitness定义为包含这个结点的社区的fitness-不包含这个结点的社区的fitness差值：

$$f_G^A = f_{G+\{A\}} - f_{G-\{A\}}.$$

下图中，蓝色和绿色的结点构成社区，而红色的结点对于这个社区的fitness都为负。



算法过程

LFM算法由两个步骤构成，选取种子和拓展种子。它随机地选择一个还没有被分配社区的结点作为种子，通过优化fitness函数的方法拓展它以形成一个社区。迭代这两步直到所有结点都属于至少一个社区为止。由于在拓展社区的时候，即使已经被分配社区的



结点也可能被添加进来，所以LFM算法是可以发现重叠社区的。

算法实现

(<https://github.com/zzz24512653/CommunityDetection/blob/master/algorithm/LFM.py>)

一点心得

LFM算法过程很易于理解，但是由于随机选择种子，导致它其实很不稳定。

GCE算法

GCE与LFM基本只是选取种子不同，GCE选取最大团来最为种子结点，最后需要融合一下相似的社区，因为一些团结构很相似。

参考文献

- 1. LFM: Detecting the overlapping and hierarchical community structure in complex networks
- 2. GCE: Detecting Highly Overlapping Community Structure by Greedy Clique Expansion

 数据挖掘算法 (nb/6412472) 举报文章 © 著作权归作者所有



八刀一闪 (/u/6979be4ade9b)

写了 11111 字，被 8 人关注，获得了 9 个喜欢

(/u/6979be4ade9b)

+ 关注

大数据处理，数据挖掘方向

如果觉得我的文章对您有用，请随意打赏。您的支持将鼓励我继续创作！

赞赏支持

♡ 喜欢 (/sign_in)

|

2



更多分享

(<http://cwb.assets.jianshu.io/notes/images/5922279>)



登录 (/sign_in)

后发表评论

评论

智慧如你，不想发表一点想法 (/sign_in)嘛~

