

文章编号: 2096-1472(2016)-01-21-02

一种新的改进的加权k-核分解方法

宋起超

(西南大学计算机与信息科学学院, 重庆 400715)

摘要: k-核分解算法是一种优秀的评估复杂网络节点重要性的方法, 然而该方法对于复杂网络节点的排序还存在一些问题。本文提出了一种改进的加权k-核分解算法, 通过改进节点加权度的计算对已提出的方法进行改进。然后在四个真实网络上利用SIR传染病模型进行了实验仿真。实验结果表明, 改进后的算法比原有方法在评估节点重要性方面更具有优越性。

关键词: 复杂网络; 节点重要度; k-核分解; SIR

中图分类号: TP393.0 **文献标识码:** A

A New Improved Weighted K-shell Decomposition Method

SONG Qichao

(School of Computer and Information Science, Southwest University, Chongqing 400715, China)

Abstract: K-shell decomposition is an excellent method in evaluating the nodes influence of complex network. While this method is not perfect in sorting the nodes importance of complex network. In this paper, an improved weighted k-shell decomposition is proposed. This method improved the method of calculating nodes weight. To evaluate the improved method, we did some experiment using SIR disease spreading model in four real networks. The experiment results show that the improved method is prior to the existing method in evaluating nodes influence.

Keywords: complex network; nodes influence; k-shell decomposition; SIR

1 引言(Introduction)

复杂网络是一门多学科交叉的领域, 研究网络节点重要度在很多领域都具有重要的理论意义和实际意义。经典的评估节点重要度的指标有度中心性、介数中心性、近邻中心性等^[1-3]。最近几年有一些研究学者陆续提出了一些新的评估方法, 比如PageRank、LeaderRank、半局部中心性、k-核分解等^[4]。k-核分解由Kitsak在2010年提出, Kitsak认为节点的重要度不是由节点度来决定的, 也不是由介数决定的, 而是由k-核(k-shell)决定的^[5]。文献[5]的实验表明, k-核分解比度中心性和介数中心性更能有效地评估节点重要度。k-核分解的优越性逐渐被人们认识, 然而该方法也存在一定的缺陷, 研究学者们相继对其进行改进。文献[6]中提出了一种新的加权k-核分解算法, 该算法改进了节点度的计算, 很大程度上有效地解决了k-核分解单调性的问题。本文基于文献[6]进行改进, 改进后的算法一定程度上优越于魏等人的方法。

2 理论方法(Theory and method)

2.1 k-核分解算法

k-核分解是一个层层推进的过程, 好像剥洋葱。第一步, 去掉度为1的节点, 剩下一个子图, 如果该子图中依然有度为1的点则继续删除这些点, 直到最后剩下一个不含度为1的节点的子图 G_1 。那些被删除的节点则属于 $ks=1$ 的核。第二步, 跟第一步类似, 删除子图 G_1 中度为2的节点, 最后得到一

个子图 G_2 , G_2 中所有点的度均大于2。以此类推, 直到所有的点都被分解到某个核中^[5]。图1为k-核分解的示意图。

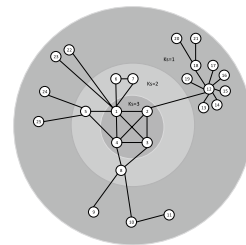


图1 k-核分解示例

Fig.1 An example of k-shell decomposition

2.2 加权k-核分解

魏等人于2015年提出了一种新的加权k-核分解^[6]。在文献[6]中, 魏等人改进了k-核分解中加权度的计算, 计算方法如下:

$$DC_i = \sum_j^n e_{ij} \quad (1)$$

$$\omega_{ij} = DC_i + DC_j \quad (2)$$

$$DC_i^\omega = \alpha DC_i + (1 - \alpha) \sum_{j \in \Gamma_i} \omega_{ij} \quad (3)$$

其中, $e_{ij} = 1$ 表示节点 i 与节点 j 是连接的; 反之, 则表示节点 i 与节点 j 是断开的。 DC_i 表示节点 i 的度。 ω_{ij} 表示节点 i 与节点 j 之间链接边的权重。 DC_i^ω 表示节点 i 的加权度。魏的方法在计算每个节点的加权度之后, 将加权度进行向上取整, 然后按照经典k-核分解对网络进行分解。文献[7]中的实验表明该方

法能有效解决经典k-核分解中单调性等问题。

2.3 改进的加权k-核分解

本文提出了一种改进的加权k-核分解算法，该算法主要改进了文献[6]中对加权度的计算。改进的加权度计算方法如下：

$$DC_i^\omega = DC_i^\alpha (\sum_{j \in \Gamma_i} \omega_{ij})^{1-\alpha} \tag{4}$$

其中， α 为调节因子，其取值范围为[0,1]。当 $\alpha=1$ 时，则变为经典k-核分解；当 $\alpha=0$ 时，则节点完全依赖邻居节点的重要性。计算加权度后，若加权度不为整数则向上取整。然后按照经典k-核分解对网络进行分解。

3 实验结果与分析(Experiment result and analysis)

为探究改进算法的有效性和可行性，我们在四个真实网络上进行了SIR传染病模拟实验^[7]。这四个真实网络分别是Blogs、Email、Netscience和USAir^[8-10]。我们分别将改进的方法所排序的前30个节点和魏的方法排序得出的前30个节点进行SIR传播实验。实验结果如图2和表1所示。图2中红色虚线表示改进方法中感染(Infected)节点的数目，红色实线表示改进方法中恢复(Recovered)节点的数目，蓝色虚线表示魏的方法中感染(Infected)节点的数目，蓝色实线表示魏的方法中恢复(Recovered)节点的数目。表1中 $T_{改进}$ 表示改进方法，SIR模型达到平衡后所需的时间。 $T_{魏}$ 表示魏的方法，SIR模型达到平衡后所需要的时间。 $N_{改进}$ 表示改进的方法，SIR模型达到平衡后，从感染状态恢复到健康状态的所有节点数目。 $N_{魏}$ 表示魏的方法，SIR模型达到平衡后，从感染状态恢复到健康状态的所有节点数目。

表1 基于SIR传染病模型实验数据统计表

Tab.1 The data of SIR spreading

网络	Blogs	Email	Netscience	USAir
$T_{改进}$	9	13	5	10
$T_{魏}$	9	13	5	9
$N_{改进}$	954	749	2.92	247
$N_{魏}$	942	733	2.82	208

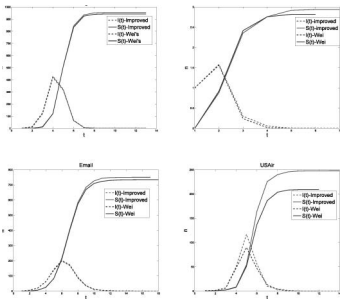


图2 基于SIR模型的传染病模拟实验

Fig.2 The experiment of SIR spreading

由图2和表1可以看出，改进的方法比魏的方法所得前30个节点在SIR模型进行传播实验后得到更多的恢复节点，也就是说改进的方法SIR实验中曾经被感染的节点更多。因此可以看出改进的方法所得到的前30个重要节点有更强的传播能力。由此可以认为改进的节点所排序得出的节点重要性更强。

4 结论(Conclusion)

本文在魏等人的基础上提出了一种改进的加权k-核分解方法。本文的方法主要通过改进计算加权度的方法对魏等人的k-核分解方法进行改进。通过实验进一步验证了该方法的可行性和优越性。当然，任何方法都不是完美的，该方法计算量较大，比较耗时。对于节点重要度的排序算法仍需不断完善和改进。

参考文献(References)

[1] L.C.Freeman,A set of measures of centrality based on betweenness, Sociometry,1977:35-41.

[2] M.E.Newman.The mathematics of networks.The new palgrave encyclopedia of economics 2,2008:1-12.

[3] 穆宝良,李晋.基于自适应仿射传播聚类的社团发现求解[J].软件工程师,2013(6):32-34.

[4] 任晓龙,吕琳媛.网络重要节点排序方法综述[J].科学通报,2014,13:004.

[5] M.Kitsak,L.K.Gallos,S.Havlin,F.Liljeros,L.Muchnik,H.E.Stanley,H.A.Makse.Identification of influential spreaders in complex networks,Nature Physics 6,2010:888-893.

[6] B.Wei,J.Liu,D.Wei,C.Gao,Y.Deng.Weighted k-shell decomposition for complex networks based on potential edge weights,Physica A:Statistical Mechanics and its Applications 420,2015:277-283.

[7] F.Fu,D.I.Rosenbloom,L.Wang,M.A.Nowak.Imitation dynamics of vaccination behaviour on social networks,Proceedings of the Royal Society of London B:Biological Sciences 278,2011:42-49.

[8] Du Y,et al.A new closeness centrality measure via effective distance in complex networks[J].Chaos:An Interdisciplinary Journal of Nonlinear Science,2015,25(3):033-112.

[9] Chen D,et al.Identifying influential nodes in complex networks[J].Physica a:Statistical mechanics and its applications,2012,391(4):1777-1787.

[10] Newman M E J,Forrest S,Baltherop J.Email networks and the spread of computer viruses[J].Physical Review E,2002,66(3):035-101.

作者简介：

宋起超(1989-),女,硕士生.研究领域:复杂网络.