

一种基于关联关系的有向网络关键节点挖掘算法

梁莹莹^{1,2} 黄 岚^{1,2,3} 王 喆^{1,2}

(吉林大学计算机科学与技术学院 长春 130012)¹

(符号计算与知识工程教育部重点实验室(吉林大学) 长春 130012)²

(吉林大学珠海学院计算机系教育部符号计算与知识工程实验室 珠海 519041)³

摘 要 关键节点在网络中的重要程度高于其他大部分节点,关键节点挖掘是网络分析的重要研究内容,对网络结构和网络中的关系等研究而言都具有非常重要的意义。已有的关键节点挖掘算法从不同的侧重点进行节点关键性评价,文中基于网络中节点的局部性信息,结合节点与其一阶邻居节点的关联关系,提出了一种有向网络关键节点挖掘算法。该算法在关注节点所处局部环境的同时考虑关联节点间的关联强度及重要性影响,根据局部重要性和关联重要性共同进行关键节点的评价。在实验网络上的影响力传播实验表明,相比于经典的度中心性等关键节点评价算法,所提算法挖掘得到的关键节点对影响力的传播能力更强,说明了算法的准确性。

关键词 有向网络,中心性,关联关系,关键节点,影响力传播

中图法分类号 TP302 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2017.12.004

Key Nodes Mining Algorithm Based on Association with Directed Network

LIANG Ying-ying^{1,2} HUANG Lan^{1,2,3} WANG Zhe^{1,2}

(College of Computer Science and Technology, Jilin University, Changchun 130012, China)¹

(Key Laboratory of Symbolic Computation and Knowledge Engineering (Jilin University), Ministry of Education, Changchun 130012, China)²

(Zhuhai Laboratory of Key Laboratory of Symbol Computation and Knowledge Engineering of Ministry of Education,

Department of Computer Science and Technology, Zhuhai College of Jilin University, Zhuhai 519041, China)³

Abstract The importance of key nodes in the network is higher than most other nodes, and key nodes mining is an important research area of network analysis. It is of great significance for study of network, such as study of network structure, network relations and so on. Many key nodes mining algorithms evaluate key nodes from different emphases. In this paper, we proposed a key nodes mining algorithm based on association with directed network, combining nodes' partial information and association with neighbors in the network. This algorithm calculates the local centrality of nodes at first. In addition, it also calculates the association centrality between associated nodes based on nodes' local centrality and the strength of association at the same time. In order to measure the accuracy of the key nodes mining algorithm, we designed influence propagation experiment to test the algorithm. Results show that compared with the classical centrality and other key nodes mining algorithms, key nodes mined by our algorithm have higher ability to spread influence in experimental network, which also illustrates the accuracy of our algorithm.

Keywords Directed network, Centrality, Association, Key nodes, Influence propagation

1 引言

近年来,关于网络的结构、功能和网络中关系的研究得到了广泛的关注^[1],电力网络、通信网络、社交网络等网络上的研究表明,对其中关键节点的挖掘对于网络功能的提升具有极大作用,对实际生产具有十分重要的意义^[2]。在通信网络中,若对其中的关键部分采取预防和保护措施,当网络遭到攻击时,可以最大限度地降低损失,提高网络的抗毁性。在社交网络中,挖掘关键用户可以为高效营销、谣言控制等工作提供

帮助,维护网络安全^[3-4]。因此,如何识别网络中的关键节点具有非常大的研究价值和现实意义^[5-6]。

中心性是度量关键节点的重要指标,是指派给网络中每一个节点数值的函数,可以从许多不同的角度来计算节点的中心性^[7-8]。近年来,结构化中心性方法得到了较多关注,如度中心性^[7]、介数中心性^[8]等,这类中心性计算方法简单直观,利用网络拓扑结构信息挖掘关键节点的方法已经被广泛应用^[9-10]。

然而,结构化中心性方法也存在一些不足,如度中心性^[7]方法由于考虑因素过于局限,因此有些情况下不能有效反映

收稿日期:2016-10-11 返修日期:2016-11-12 本文受国家自然科学基金项目(61373051,61472159),吉林省电子商务关键支撑技术工程实验室创新能力项目资助。

梁莹莹(1993—),女,硕士,主要研究方向为社交网络分析;黄 岚(1974—),女,博士,教授,主要研究方向为数据挖掘和商务智能;王 喆(1974—),男,博士,副教授,主要研究方向为社交媒体挖掘, E-mail: wz2000@jlu.edu.cn(通信作者)。

网络中节点的重要程度;介数中心性^[8]的计算复杂度过高,难以用于大规模网络中等。因此,随着研究的深入,研究人员也提出了很多新的结构化中心性方法^[11-17]。这些方法都从某一方面对关键节点挖掘工作做出了改进,取得了较好的效果,但也存在一些不足。文献[11]首次提出节点重要性依赖于其在整个网络中的位置的思想,并提出了 k-shell 指标,该指标的计算复杂度低,可应用于大规模网络,但其仅考虑剩余度的影响,将具有相同剩余度的节点看作同等地位,忽略了现实世界中节点的出度和入度分别代表着不同含义,显然在某些情况下不能合理地反映节点在网络中的重要程度。文献[12]提出了一种基于半局部信息的节点重要性排序方法,该方法仅对节点的四阶邻居进行数量统计,对节点间的关联关系没有做深入研究,如关联方向、关联强度等;此外,该方法将不同层次的邻居节点视作同等地位,不能反映出不同层次的邻居节点对节点中心性的影响。文献[13]提出了基于度与聚集系数的算法,综合考虑了节点的度及其邻居节点间的紧密程度,并且只需考虑节点的局部性信息,其适用于大规模网络;但该算法将研究的网络视作无向的,不能反映真实网络中的情况,影响了关键节点评价的准确性。文献[14-15]均利用了重要性贡献矩阵的思想,考虑了关联节点间的重要性影响,但对节点间重要性贡献比例的计算方法较粗糙。另一方面,在有向网络的关键节点挖掘领域也出现了很多新的算法,著名的 PageRank 算法认为:从任意网页出发,采用输入网址来访问其他网页的概率相同,但这不符合现实中人们在浏览网页时的访问习惯;此外,PageRank 算法中参数 c 的选取往往需要通过实验获得,并且在不同的应用背景下最优参数不具有普适性。文献[16]提出了 LeaderRank 算法,该算法很好地解决了这两个问题^[9],在相同的实验背景下,LeaderRank 算法收敛速度更快,健壮性更强,对网络中关键节点的识别效果更好。文献[17]提出了 ClusterRank 算法,该算法基于节点的局部性信息,与 Page-Rank 和 LeaderRank 相比在运行时间和最终效果上具有明显优势,但是该算法丢失了节点间的信任程度信息,没有对邻居节点间的关联强度进行深入的研究和探讨。

可以看出,尽管结构化中心性方法在计算复杂度等方面表现突出,但其进行关键节点评价时依赖于网络拓扑结构信息,可用信息有限;部分结构化中心性方法忽略了节点间关联的方向信息或关联强度信息,不能合理地反映现实世界中的真实情况,从而影响了关键节点挖掘的效果。针对这一问题,本文提出了一种基于关联关系的有向网络关键节点挖掘算法,综合考虑节点的局部性信息和节点间的关联关系,并计算关联强度,利用局部中心性和关联中心性共同完成关键节点评价,取得了较好的效果。

2 基于关联关系的有向网络关键节点挖掘

文献[18]提出,网络中节点的行为与其邻居节点密切相关,节点间的关联关系将影响节点的行为及其在网络中的中心性。以社交网络为例,节点的入度代表其受欢迎程度,出度代表合群程度^[19],这些都对节点的中心性产生影响,将其视为无向网络时势必会丢失这些信息,影响关键节点挖掘的准确性。本算法基于有向网络,根据节点所处局部环境和节

点间的关联关系进行关键节点挖掘。

本算法首先根据节点所处局部环境计算其局部中心性,然后计算关联节点间的关联强度,从而计算关联中心性,再综合两部分得出节点最终的中心性评价结果。其中,关联关系的强度对节点中心性和节点间信息的传播起着重要的作用,对诸如链接预测等各种网络分析工作具有非常重要的意义^[20-23]。

定义基于关联关系的有向网络关键节点挖掘算法 Relativity Centrality, $RelaCentrality(j)$ 表示节点 v_j 通过本算法评价所得的最终中心性,为了方便,此后用 $RC(j)$ 代表 $RelaCentrality(j)$ 。设 $G=(V, E)$ 是一个有向网络,其中 $V=\{v_1, v_2, \dots, v_n\}$ 是网络中所有节点的集合, $E=\{e_1, e_2, \dots, e_m\}$ 是网络中边的集合, $E \subseteq V \times V$, 其中 $e_{ij}=(v_i, v_j)$, $deg^{out}(i)$ 表示节点 v_i 的出度, $deg^{in}(i)$ 表示 v_i 的入度, $I(i)$ 表示节点 v_i 的局部中心性。

2.1 局部中心性

设节点 v_i 的局部中心性为 $I(i)$ ^[22], 计算方法如式(1)所示:

$$I(i) = \frac{deg^{in}(i)}{deg^{out}(i) + deg^{in}(i)} \quad (1)$$

该指标结合节点所处位置信息,反映了节点在其一阶邻居中的影响力,代表了节点的局部重要性,且只需分别统计节点的出度和入度,计算复杂度低。

2.2 关联中心性

为了度量关联关系对节点中心性的影响,定义节点间的关联中心性。如图 1 所示,假设节点 x 、节点 y 和节点 z 均代表网络 G 中的用户,分别用 v_x 、 v_y 和 v_z 代表。可以看到, v_x 和 v_y 都有指向 v_z 的有向边,此后当出现形如这种关联关系时,我们称 v_x 和 v_y 为 v_z 的粉丝节点, v_z 为被关注节点。

有向网络中的关联中心性代表了关联关系对节点中心性的影响,是指通过节点间的关联关系使被关注节点在网络中的中心性获得提升的情况。关联中心性对节点中心性的影响程度受多方面因素影响,如关联关系的强弱程度、粉丝节点在网络中的中心性等。例如,在学术文献网络中,当某文献 A 被另一影响因子较高的文献 B 引用时,文献 A 可能因此被更多的研究人员关注,使其在学术文献网络中的中心性获得提升,而文献 A 的中心性提升程度与文献 B 的中心性以及文献 A 对文献 B 的重要程度等因素有关。

相似性反映了节点间在基本信息或兴趣爱好等方面的相似程度,能够反映节点间相互信任或相关的程度,因此本算法使用节点间的相似性代表节点间的关联强度。当前已有许多度量节点间相似性的指标^[20],本算法使用经典的共同邻居指标度量节点间相似性,该指标的计算复杂度低,健壮性强。如图 1 所示,共同邻居相似性指标(Common Neighbors Index, CN)认为,若 e_{xz} 存在且 v_x 和 v_z 之间存在共同邻居节点,则 v_x 和 v_z 之间的相似度与 $CN(x, z)$ 的值成正比^[24]。令 $\tau(x)$ 表示节点 v_x 的邻居集合,一般的 CN 指标定义如式(2)所示:

$$CN(x, y) = |\tau(x) \cap \tau(y)| \quad (2)$$

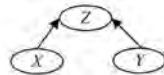


图1 关联中心性示意图

本算法认为节点只能对其一阶邻居(关联)节点产生影响,且这种关联对被关注节点的中心性的影响程度是有限的,影响程度由节点间的关联强度即相似性决定。以 v_x 关注 v_z 为例, $CN(x, z)$ 越大,代表 v_x 与 v_z 的相似程度越高,相比于 v_x 关注的其他节点, v_z 对 v_x 的重要性更高, v_z 基于这一关联关系获得的中心性提升也就越大。

节点间关联中心性影响比例的计算方法如式(3)所示:

$$C_{i,j} = \frac{1 + CN(i, j)}{\deg^{out}(i) + \sum_{k=1, k \neq i}^n CN(i, k)} \quad (3)$$

其中, $C_{i,j}$ 表示粉丝节点 v_i 对其关注的某节点 v_j 的关联中心性在 v_i 自中心性中的占比, $\sum_{k=1, k \neq i}^n CN(i, k)$ 表示 v_i 与其关注的所有节点间的共同邻居总数。可以看出, v_i 对其关注的所有节点都会产生关联中心性影响。以节点 v_i 和节点 v_j 为例, v_i 对 v_j 的关联中心性影响程度与 $CN(i, j)$ 成正比, $CN(i, j)$ 越大, v_i 对 v_j 的关联中心性的影响占 v_i 自中心性的比例越大,代表 v_i 对 v_j 的信任程度越高,则 v_j 对 v_i 的影响力越大。

节点最终中心性由局部中心性和关联中心性两部分构成,计算方法如式(4)所示:

$$RC(j) = I(j) + \sum_{i=1, i \neq j}^n C_{i,j} \times I(i) \quad (4)$$

其中, $\sum_{i=1, i \neq j}^n C_{i,j} \times I(i)$ 表示 v_j 的所有粉丝节点与 v_j 间的关联中心性,此后用 $R_i(j)$ 表示用这一方法计算得到的节点 v_j 的关联中心性。

3 实验及分析

设 top- k 节点表示算法挖掘得到的关键节点集合,本文使用影响力传播模型中的独立级联模型^[25]度量 top- k 节点在网络中的影响力,并将其作为算法表现效果的度量。将各个算法得到的 top- k 节点作为种子节点集合,分别测试种子节点被激活后网络中最终可能被激活的节点总数。在影响力传播中,传播结束后被激活的节点个数的期望值是研究热点,由于各个传播过程是随机的,为了避免结果的不稳定性,本文选用蒙特卡洛方法,取各算法在每个数据集上独立运行 20 次的均值作为最终激活节点数的期望值,将该期望作为算法在相应实验数据集上挖掘出的关键节点的最终影响力。

3.1 案例网络上的实验及分析

文献[17]提出了 ClusterRank 算法,该算法综合考虑了节点的邻居数、邻居节点性质以及邻居节点集合的内聚程度,是近年来提出的有向网络关键节点挖掘的有效算法,用 $S(i)$ 表示节点 v_i 在该算法下的重要度,计算方法如式(5)所示:

$$S(i) = f(c_i) + \sum_{j \in \tau(i)} (k_j^{\text{in}} + 1) \quad (5)$$

其中, $f(c_i)$ 表示节点 v_i 的局部聚集系数的影响,即当某节点的局部聚集系数越大,该节点的影响力越难以向外传播,从而降低了其在网络中的重要程度。 $f(c_i) = 10^{-c_i}$, c_i 表示节点 v_i 的聚集系数,有向网络中节点的聚集系数计算方法如式(6)所示:

$$c_i = \frac{e_i}{k_i \times (k_i - 1)} \quad (6)$$

其中, e_i 表示节点 v_i 与其任意两个邻居节点之间所形成的三

角形的个数, k_i 表示节点 v_i 的度。

文献[17]中给出的有向网络如图 2 所示,该网络中共有 38 个节点,110 条有向边,在该网络上对本算法与 ClusterRank 算法的关键节点挖掘结果进行对比。经过实验得到该网络中重要性排名前 13 位的节点,如表 1 所列。可以看出,本算法与 ClusterRank 算法识别出了相同的前 5 名的关键节点,但对后续节点的中心性排序存在差异。

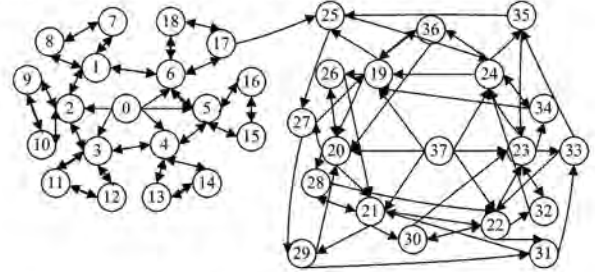


图 2 文献[17]中的有向网络示例

表 1 本文算法与 ClusterRank 算法对示例网络关键节点的挖掘结果

节点排序	算法	
	RelaCentrality	ClusterRank
1	19	19
2	20	20
3	21	21
4	22	22
5	23	23
6	1	24
7	2	6
8	3	37
9	4	1
10	5	2
11	24	3
12	6	4
13	17	5

我们测试 top- k 节点在网络中的影响力传播情况,分别取两种算法挖掘得到的前 10 位和第 6—13 位节点作为种子节点,进行网络中的影响力传播实验,得到网络中影响力传播情况,分别如图 3、图 4 所示。其中,横坐标代表时刻,纵坐标代表网络中被种子节点激活的节点数期望值。在图 3 中,本算法在各个传播时刻挖掘得到的 top-10 关键节点在网络中激活的节点数均多于 ClusterRank 算法。在图 4 中,不考虑两算法共同的 top-5 节点后,取后续节点进行影响力传播实验,最终本算法在网络中激活的节点数的期望值高于 ClusterRank 算法,即本算法挖掘出的关键节点在网络中对影响力传播的能力更强,从而印证了本算法的准确性。

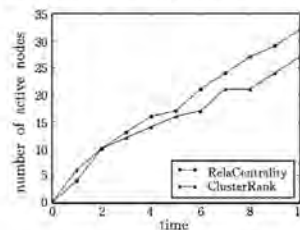


图 3 top-10 节点在示例网络中的影响力传播情况

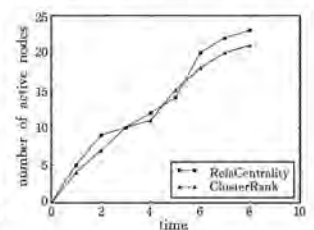


图 4 第 6—13 位关键节点在示例网络中的影响力传播情况

3.2 真实网络数据集实验及分析

本文选择了 3 个真实的有向无权的网络数据集^[26-28]作为

实验网络,并进行关键节点的影响力传播实验,数据集信息如表 2 所列。

表 2 有向网络数据集描述

数据集名称	Taro	HighSchool	Physician
数据集描述	Papuan village 中,家庭间的礼物馈赠网络	1957 年和 1958 年,经调查得到的某高中男生间的朋友网络	1966 年,调查得到的生活在 Illinois 等 5 个地方的 246 个物理学家间的朋友关系网络
节点数	22	70	241
边数	78	366	1098
top-k 节点比例/%	33	33	16.7
top-k 节点个数	8	24	41

除 ClusterRank 外,实验所选对比算法还包括经典的度中心性算法和近年新出现的半局部中心性算法,下面简要介绍这两种算法。

度中心性(Degree Centrality, DC)是刻画网络中节点重要性的最简单的指标^[7]。该指标认为一个节点的度越大,其能直接影响的邻居数目就越多,从而节点越重要。节点 v_i 的度记为 k_i ,是指与 v_i 直接相连的节点数目,由于有向网络中节点的度有出、入度之分,因此节点的 v_i 归一化度中心性指标为:

$$DC(i) = \frac{k_i}{n-1} \quad (7)$$

其中, n 为网络中节点总数, $n-1$ 为网络中节点可能取得的最大度值。度中心性是一个简单有效且健壮性很强的指标,但其对节点所处环境的考虑有限,有些情况下不能准确评价关键节点。

文献^[12]提出了一种基于半局部信息的节点重要性排序方法,简称半局部中心性(Semi-Local Centrality, SLC)。该指标首先定义 $N(\omega)$ 为节点 v_ω 的两层邻居度,其值等于在出发 2 步内可到达的邻居数目,然后定义:

$$Q(j) = \sum_{\omega \in \tau(j)} N(\omega) \quad (8)$$

其中, $\tau(j)$ 表示节点 v_j 的一阶邻居节点的集合,节点 v_i 的半局部中心性为:

$$SLC(i) = \frac{Q(\omega)}{n-1} \quad (9)$$

该方法考虑了节点的 4 阶邻居信息,计算复杂度较低,其实验效果优于度中心性和介数中心性。

在表 2 所列数据集上分别测试各算法挖掘得到的关键节点的影响力传播情况,实验结果如图 5—图 7 所示。可以看出,在算法运行的各个时刻,本算法得到的关键节点在网络中激活的节点数期望值均大于其他 3 个算法,证明本算法挖掘的关键节点在网络中的中心性更高,影响力更强,从而印证了本算法的有效性。

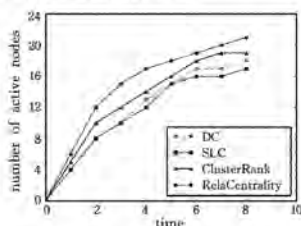


图 5 在 Taro 数据集上 top-8 节点的影响力传播情况

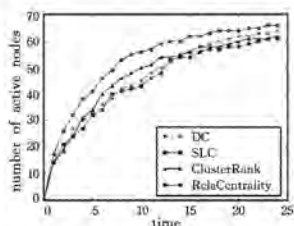


图 6 在 HighSchool 数据集上 top-24 节点的影响力传播情况

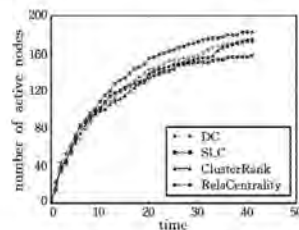


图 7 在 Physician 数据集上 top-41 节点的影响力传播情况

4 算法各参数的影响情况分析

本算法的局部中心性指标有多种选择,而节点间的关联中心性也具有多种计算方法,有向网络中节点的度分为入度和出度,本文选择节点的度中心性作为局部中心性指标,计算方法如式(10)所示:

$$DC(i) = \frac{deg^{in}(i) + deg^{out}(i)}{n-1} \quad (10)$$

进行关联中心性计算时,若不考虑节点间的关联强度,则认为粉丝节点 v_i 将其关注的所有节点视作同等地位,而不关注与这些节点间的亲密程度。因此,不考虑关联强度的节点中心性计算方法如式(11)所示:

$$SRC(j) = I(j) + \sum_{i=1, i \neq j, \exists e_{ij}} \frac{1}{deg^{out}(i)} \times I(i) \quad (11)$$

该方法尽管考虑了节点间的关联信息,但不能表现出关联强度,关联中心性的计算粒度较粗,此后用 $R_2(j)$ 表示按式(11)计算的节点 v_j 的关联中心性。

综上,局部中心性和关联中心性两部分指标可能的计算方法组合如表 3 所列。

表 3 算法两部分参数的组合情况

算法名称	计算方法
RelaCentrality	$I(i) + R_1(i)$
RelaCentrality_1	$DC(i) + R_1(i)$
RelaCentrality_2	$I(i) + R_2(i)$
RelaCentrality_3	$DC(i) + R_2(i)$

使用上述不同组合的算法进行关键节点挖掘,并对比实验结果,在表 2 所列的数据集上进行实验,得到 top-k 节点在实验网络上的影响力传播情况,如图 8—图 10 所示。可以看出,本算法对应的指标组合识别出的关键节点在网络中对影响力的传播能力更强,印证了本算法所选指标的合理性。

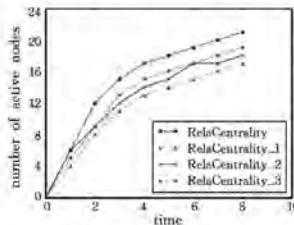


图 8 Taro 数据集上的算法参数

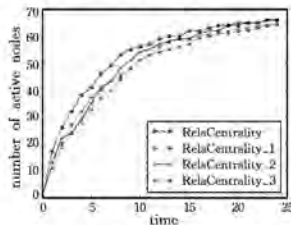


图 9 HighSchool 数据集上的算法参数对比图

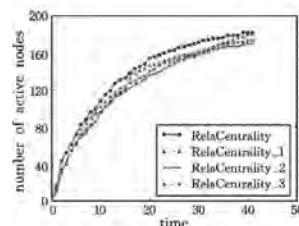


图 10 Physician 数据集上的算法参数对比图

结束语 关键节点挖掘是保护网络中重要节点、提升网络性能的重要依据,也是网络分析中的重要部分,对理论研究和实际应用都有十分重要的意义,能对网络中的推荐、链接预测等工作提供帮助,起到一定的推动作用。因此,本文面向有向网络,综合考虑节点在网络中的所处位置及周围环境,基于节点的局部位置信息和节点间的关联强度共同进行关键节点挖掘,提出了一种基于关联关系的有向网络关键节点挖掘算法。影响力传播实验表明,与度中心性等经典指标相比,本算法挖掘得到的关键节点在网络中对影响力的传播能力更强,即关键节点评价更为准确。本算法还存在一定的不足,如关联中心性的计算方法以及节点间关联中心性比例的计算方法较粗糙、实验数据集不够完善等。未来的工作包括优化节点的局部中心性在内的各部分计算方法、增加对比实验等,以提升指标计算的合理性和精确度,从而提高关键节点挖掘的准确度。

参考文献

- [1] NEWMAN M. Networks: An Introduction[M]. Oxford University Press, Inc. 2010.
- [2] CALLAWAY D S, NEWMAN M E, STROGATZ S H, et al. Network robustness and fragility: percolation on random graphs[J]. Physical Review Letters, 2000, 85(85): 5468-5471.
- [3] CRUCITTI P, LATORA V, MARCHIORI M, et al. Error and attack tolerance of complex networks[J]. Nature, 2000, 406(6794): 542-542.
- [4] WENG J, LIM E P, JIANG J, et al. TwitterRank: finding topic-sensitive influential twitterers[C]//International Conference on Web Search and Web Data Mining (WSDM 2010). New York, Ny, Usa, February, 2010: 261-270.
- [5] BODENDORF F, KAISERISER C. Detecting opinion leaders and trends in online social networks[C]//ACM Workshop on Social Web Search and Mining. ACM, 2009: 65-68.
- [6] GAO C, LAN X, ZHANG X, et al. A bio-inspired methodology of identifying influential nodes in complex networks[J]. Plos One, 2013, 8(6): e66732.
- [7] FREEMAN L C. Centrality in Social Networks: I. Conceptual Clarification[J]. Social Networks, 1979, 1(3): 215-239.
- [8] SABIDUSSI G. The centrality index of a graph[J]. Psychometrika, 1966, 31(4): 581-603.
- [9] REN X L, LV L Y. Review of ranking nodes in complex networks[J]. Chinese Science Bulletin (Chin Ver), 2014, 59: 1175-1197. (in Chinese)
任晓龙,吕琳媛. 网络重要节点排序方法综述[J]. 科学通报, 2014, 59(13): 1175-1197.
- [10] LIU J G, REN Z M, GUO Q, et al. Node importance ranking of complex networks[J]. Acta Physica Sinica, 2013, 62(17): 178901. (in Chinese)
刘建国,任卓明,郭强,等. 复杂网络中节点重要性排序的研究进展[J]. 物理学报, 2013(17): 178901.
- [11] KITSACK M, GALLOS L K, HAVLIN S, et al. Identification of influential spreaders in complex networks[J]. Nature Physics, 2010, 6(11): 888-893.
- [12] CHEN D, LV L, SHANG M S, et al. Identifying influential nodes in complex networks[J]. Physica a: Statistical Mechanics and Its Applications, 2012, 391(4): 1777-1787.
- [13] REN Z M, SHAO F, LIU J G, et al. Node importance measurement based on the degree and clustering coefficient information[J]. Acta Physica Sinica, 2013, 62(12): 128901. (in Chinese)
任卓明,邵凤,刘建国,等. 基于度与集聚系数的网络节点重要性度量方法研究[J]. 物理学报, 2013, 62(12): 128901.
- [14] ZHOU X, ZHANG F M, LI K W, et al. Finding vital node by node importance evaluation matrix in complex networks[J]. Acta Physica Sinica, 2012, 61(5): 50201. (in Chinese)
周濛,张凤鸣,李克武,等. 利用重要度评价矩阵确定复杂网络关键节点[J]. 物理学报, 2012, 61(5): 50201.
- [15] ZHAO Y H, WANG Z L, ZHENG J, et al. Finding most vital node by node importance contribution matrix in communication networks[J]. Journal of Beijing University of Aeronautics and Astronautics, 2009, 35(9): 1076-1079. (in Chinese)
赵毅寰,王祖林,郑晶,等. 利用重要性贡献矩阵确定通信网中最重要节点[J]. 北京航空航天大学学报, 2009, 35(9): 1076-1079.
- [16] LÜ L, ZHANG Y C, YEUNG C H, et al. Leaders in social networks, the delicious case[J]. PloS One, 2011, 6(6): e21202.
- [17] CHEN D B, GAO H, LÜ L, et al. Identifying Influential Nodes in Large-Scale Directed Networks: The Role of Clustering[J]. Plos One, 2013, 8(10): e77455.
- [18] KLEIN A, AHLF H, SHARMA V. Social activity and structural centrality in online social networks[J]. Telematics & Informatics, 2015, 32(2): 321-332.
- [19] SHAO F, GUO Q, ZENG S Q, et al. Research Progress of the Microblog System Structures[J]. Journal of University of Electronic Science and Technology of China, 2014(2): 174-183. (in Chinese)
邵凤,郭强,曾诗奇,等. 微博系统网络结构的研究进展[J]. 电子科技大学学报, 2014(2): 174-183.
- [20] LÜ L, ZHOU T. Link prediction in complex networks: A survey[J]. Physica A Statistical Mechanics & Its Applications, 2011, 390(6): 1150-1170.
- [21] BORGATTI S P, MEHRA A, BRASS D J, et al. Network analysis in the social sciences[J]. Science, 2009, 323(5916): 892-895.
- [22] KOSSINETIS G. Effects of missing data in social networks[J]. Social Networks, 2010, 28(3): 247-268.
- [23] WU X D, LI Y, LI L. Influence Analysis of Online Social Networks[J]. Chinese Journal of Computers, 2014, 37(4): 735-752. (in Chinese)
吴信东,李毅,李磊. 在线社交网络影响力分析[J]. 计算机学报, 2014, 37(4): 735-752.
- [24] LIN D. An Information-Theoretic Definition of Similarity[C]//Fifteenth International Conference on Machine Learning. Morgan Kaufmann Publishers Inc., 1998: 296-304.
- [25] KEMPE D, KLEINBERG J, TARDOS É. Maximizing the spread of influence through a social network[C]//Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2003: 137-146.
- [26] http://konect.uni-koblenz.de/networks/moreno_taro.
- [27] http://konect.uni-koblenz.de/networks/moreno_highschool.
- [28] http://konect.uni-koblenz.de/networks/moreno_innovation.