

复杂社会网络节点重要性可拓聚类动态分析方法

严家萌^{1,2} 庞超逸² 许立波^{2*}

¹(浙江大学工程师学院 浙江 杭州 310015)

²(浙江大学宁波理工学院计算机与数据工程学院 浙江 宁波 315100)

摘要 目前大多数复杂网络关键节点影响力识别方法都是建立在静态分析基础之上的,缺乏对节点动态变化的考量。采用可拓聚类方法来量化分析社会网络的动态变化,以多属性决策法测量各节点静态重要性属性值;分别划分等级,取经典域和节域,计算各节点在对应属性上的关联度,并用熵权法取得综合关联度;分析其在同一社会网络不同时间点的变化,给出变动的量和方向。该方法不仅能刻画出节点重要性的变化趋势,而且给出了变化程度。最后,通过实例验证了该方法的可行性和有效性。

关键词 复杂网络 节点重要性 多属性 可拓学 可拓聚类

中图分类号 TP301.6 文献标识码 A DOI:10.3969/j.issn.1000-386x.2019.07.014

EXTENSION CLUSTERING DYNAMIC ANALYSIS METHOD FOR IMPORTANCE OF COMPLEX SOCIAL NETWORK NODES

Yan Jiameng^{1,2} Pang Chaoyi² Xu Libo^{2*}

¹(Polytechnic Institute, Zhejiang University, Hangzhou 310015, Zhejiang, China)

²(School of Computer and Data Engineering, Ningbo Institute of Technology, Zhejiang University, Ningbo 315100, Zhejiang, China)

Abstract At present, the identification method of the influence of key nodes in complex networks is all based on static, which is lacking in consideration for dynamic analysis. In this paper, the expansion clustering method was used to quantitatively analyze the social network under the dynamic change. The value of each node importance was measured by multi-attribute decision. Then, the grade was divided respectively, the classical domain and acceptable domain were taken, and the correlation degree of each node on the corresponding attribute was calculated. Meanwhile, the entropy weight was used to obtain comprehensive correlation degree. We analyzed the node importance of the same social network at different time points, and gave the amount and direction of change. This method can not only describe the trend of changing, but also work out the amount of it. At last, an example is given to verify the feasibility and effectiveness of the proposed method.

Keywords Complex network Node importance Multi-attribute Extenics Extension clustering

0 引言

随着信息化网络化走向纵深领域,对复杂网络^[1]的研究越来越显现其巨大的实用价值。在人类活动的参与下,网络表现为社会化,相比于其他类型网络,社会网络体现出更接近现实的复杂度,且往往要求对关

键和核心节点的认定。比如,交通拥堵路口识别,疾病传播关键节点发现,国际环境中的大国影响力等,都离不开对影响力节点的刻画。发现复杂网络中的关键节点,有针对性的有效利用有着重要意义。在对复杂社会网络节点重要性度量时,应该考虑时间和空间两个维度,即此时此刻和某时某刻。而当前,对复杂网络的研究大都是建立在静态基础之上的,缺乏动态分析和

考量。在复杂网络节点重要性背景下,无论是综合考虑节点连接度和经过该节点最短路径数的节点收缩法^[2],还是借鉴排名算法的有向加权评估方法^[3],亦或是评价网络中每个节点与“核心节点”的关联度的定量评估法^[4],以及通过计算多重属性到理想的接近程度的综合评价指标^[5]等,基本上都是静态分析方法。

在国外,复杂网络节点重要性中心性指标用来对节点的传播影响力进行刻画,主要有度中心性(Degree Centrality)^[6]、介数中心性(Betweenness Centrality)^[7]、紧密度指标(Closeness Centrality)^[8]、K-Shell(KS 指标)^[9]等。文献[10]利用度中心性来度量节点的影响力,节点度描述了节点间的直接通信能力,直观反映节点的重要性,度表示为与节点直接相连的个数,在复杂社会网络中度数较大的节点的传播影响力相对较大。文献[11]利用介数中心性来衡量社会网络节点重要性,刻画了复杂网络中节点对于信息流动的影响力,以其经过某个节点的最短路径的数目来描述节点的重要性,网络节点介数定义为经过节点的路径数目占最短路径总数的比例。介数中心性认为,在复杂网络中如果一个节点是其他节点对之间联通的必经之路,则其在网络中具有重要地位,相应的影响力也就越大。文献[12-13]利用紧密度指标刻画某个节点到其他节点的难易程度,用其到网络中其他所有节点距离之和的倒数来刻画,节点的紧密度值越大,表明该节点越居于网络中心位置,相应地也就越重要。文献[14-15]提出通过 K-shell(KS)分解来确定最具有影响力的单源节点,操作上相当于剥去了最外面一层壳。首先去除网络中度数小于 k 的所有节点及其连边,如果在剩下的节点中仍有度值小于 k 的节点,那么就继续去除这些节点,直至网络中剩下的节点的度值不小 k ,依此类推,直到复杂社会网络中所有的节点都被赋予 KS 值。

总体来看,目前大多数研究对复杂网络节点重要性度量时的方法均是静态的,能很好地阐释复杂网络节点内部相互连接的逻辑,区分重要性程度。然而,对变化了的节点重要度如何度量缺乏研究,更没有给出变化的量。本文意在尝试提出一种新的基于可拓聚类的节点重要性动态分析方法。

1 可拓聚类方法说明

1.1 变量说明

对文中所涉及主要变量作如下定义和说明。

D_c 为度中心性	B_c 为介数中心性
C_c 为紧密度中心性	q_{ij} 节点原始属性值
p_{ij} 为标准化后的属性值	d_r 为第 r 个等级
λ_j 为属性权重	u_i 为节点属性
T_i 为时点	$k(p_{ij})$ 为关联度
$K(i)$ 为综合关联度	$\langle a_r, b_r \rangle$ 为 r 等级的取值区间

1.2 引入可拓学

可拓学^[16]是通过定性和定量相结合的手段去处理矛盾问题的新兴学科,其中可拓集理论是研究事物动态性和可变性的有力工具。杨国为等^[17]利用可拓模式识别建立起神经网络模型,推出模式可拓识别分类器。Xu 等^[18]提出可拓关联结合后悔理论,应用于决策分析领域。周玉等^[19]对可拓神经网络的基本思想、算法思路、应用研究进行了系统分析,并提出了有待进一步研究的方向和问题。而将可拓聚类分析方法应用于复杂网络节点重要性方面尚未见到。本文试图将可拓集理论引入复杂网络节点影响力研究领域,提出一种基于可拓聚类的复杂网络节点重要性动态分析方法。将复杂网络节点重要性加上时间维度,引入可拓关联函数分析关联度变化,从而判断变化方向,给出变化的量。

1.3 节点重要性多属性值

在社会网络中,可用图模式量化,用图中节点表示个体,用边表示各个体之间的连接关系。设复杂社会网络 $G = (V, E)$ 是一个无向无权图,其中 $V = \{v_1, v_2, \dots, v_n\}$ 是网络中所有节点的集合, $|V| = n$; $E = \{e_1, e_2, \dots, e_m\}$ 是节点间边的集合, $|E| = m$, $a_{ij} = 1$, 其中 $a_{ij} = 1$ 表示节点 i 和节点 j 相连, 否则 $a_{ij} = 0$, 如图 1 所示。

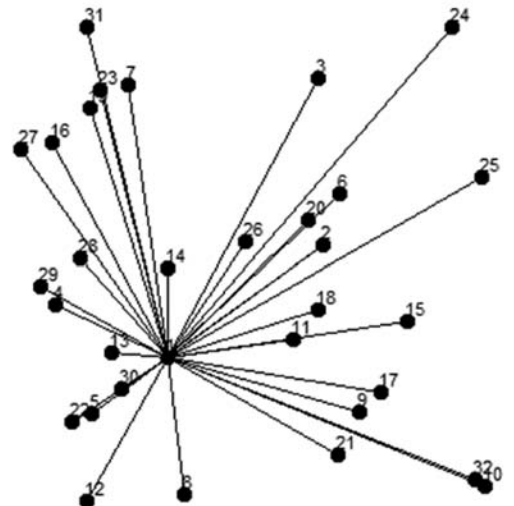


图 1 根据社会网络人员熟识与否模拟复杂网络图 (节点 1 为例)

度中心性用 D_c 表示为:

$$D_c = d(v) \quad (1)$$

式中: d 为节点 v 的度, 即直接相连的边数。

节点的介数 B_c 为:

$$B_c = \sum \frac{S_{jk}(i)}{S_{jk}} \quad (2)$$

式中: S_{jk} 表示源节点 j 到目标节点 k 的最短路径数; $S_{jk}(i)$ 表示节点 j 和 k 之间经过节点 i 的最短路径数。

紧密度中性具体 C_c 表示为:

$$C_c = \frac{N-1}{\sum_{j=1}^n d_{ij}} \quad (3)$$

式中: d_{ij} 表示以节点 i 为起点, 以 j 为终点的最短路径中所含边的数量; N 为节点总数。

在评估复杂网络各节点重要性时, 为了克服单一属性的片面性, 本文采用多属性(即度中心性、介数中心性、紧密度中心性和 KS 指标)综合考量, 在计算完每个节点在对应属性上的关联度后, 取综合关联度为该节点重要性的最终量值。

多属性分析方法要求对属性值进行标准化, 每个节点在对应的属性上的取值规范化为:

$$p_{ij} = \frac{q_{ij}}{\max |q_{ij}|} \quad i=1, 2, \dots, n, j=1, 2, \dots, m \quad (4)$$

式中: q_{ij} 为第 i 个节点第 j 个属性值; p_{ij} 为规范化后第 i 节点第 j 个属性值。

2 可拓聚类模型

2.1 模型经典域及节域设置

以网络节点 O 为对象, c_i 为节点各属性特征, O 关于 c_i 的量值 v_i 构成的有序三元组 $M = (O, c_i, v_i)$ 作为描述节点重要性的基元; 而网络节点 O 的 n 个特征属性 c_1, c_2, \dots, c_n 及 O 关于 $c_i (i=1, 2, \dots, n)$ 对应的量值 $v_i (i=1, 2, \dots, n)$ 所构成的矩阵 M_1 称为 n 维物元^[20]。

$$M_1 = \begin{bmatrix} O & c_1 & v_1 \\ & c_2 & v_2 \\ & \vdots & \vdots \\ & c_n & v_n \end{bmatrix} = (O, c_i, v_i) \quad (5)$$

根据论域 U 中一个对象对于某特征的量值符合要求的程度, 可分为满意区间 $X_0 = \langle a_0, b_0 \rangle$ 和可接受区间 $X = \langle a, b \rangle$, 即经典域和节域, 显然 $X_0 \subset X$ 。

$$(O, c_j, v_j) = \begin{bmatrix} O & c_1 & v_1 \\ & c_2 & v_2 \\ & \vdots & \vdots \\ & c_n & v_n \end{bmatrix} = \begin{bmatrix} O & d_1 & \langle a_1, b_1 \rangle \\ & d_2 & \langle a_2, b_2 \rangle \\ & \vdots & \vdots \\ & d_n & \langle a_n, b_n \rangle \end{bmatrix} \quad (6)$$

式中: O 表示网络节点, $d_r (r=1, 2, \dots, n)$ 表示对应的第 r 个等级; $\langle a_r, b_r \rangle (r=1, 2, \dots, n)$ 表示节点 O 对应等级特征的量值范围。相应地, 有 $\langle c_r, d_r \rangle$ 表示节点 O 对应等级特征的节域。

2.2 关联函数

定义可拓距^[21]为某节点与经典域节域区间的距离。给定区间 $X_0 = \langle a, b \rangle$, 当最优点 x_0 在右侧(最优点在 $x_0 = b$ 处), 则 x 与区间 X_0 关于 x_0 的右侧距为:

$$\rho_r(x, x_0, X_0) = \begin{cases} a-x & x \leq b \\ \frac{a-x_0}{b-x_0}(b-x) & x \in \langle x_0, b \rangle \\ x-b & x > b \end{cases} \quad (7)$$

反映各节点重要性与各经典域的隶属程度的函数为关联函数^[22]。设区间 $X_0 = \langle a, b \rangle$, $X = \langle c, d \rangle$ 且 $X_0 \subset X$, 则 x 关于 x_0 及区间 X_0 和 X 的关联函数为(文中 x 即为属性值 p_{ij}):

$$k(x) = \frac{\rho(x, x_0, X)}{D(x, x_0, X_0, X)} \quad D(x, x_0, X_0, X) \neq 0 \text{ 且 } x \in X \quad (8)$$

$$D(x, x_0, X_0, X) = \begin{cases} \rho(x, X) - \rho(x, X_0) & \rho(x, X) \neq \rho(x, X_0) \text{ 且 } x \notin X_0 \\ \rho(x, X) - \rho(x, X_0) + a - b & \rho(x, X) \neq \rho(x, X_0) \text{ 且 } x \in X_0 \\ a - b & \rho(x, X) = \rho(x, X_0) \end{cases}$$

式中: $\rho(x, x_0, X)$ 为可拓距; $D(x, x_0, X_0, X)$ 称为位值, 用来描述点 x 关于 x_0 与 X_0 和 X 组成的区间套的位置关系。

2.3 熵权及综合关联度

在度量节点重要性综合关联度时, 将根据每个属性的不同重要程度取合理的权系数计算得到。在复杂社会网络中, 节点综合关联度有整体变化的趋势, 因此利用等权值并不合适。根据网络节点不同属性的变化程度来设置不同的权值将更具可靠性。而熵权是根据特征参数值变化程度所反映的信息量大小来确定权值的客观赋权法。计算各节点的熵值 e_j :

$$e_j = -\frac{1}{\ln n} \sum_{i=1}^n (p_{ij} \times \ln(1 + p_{ij})) \quad (9)$$

计算各节点的熵权 λ_j :

$$\lambda_j = \frac{1 - e_j}{\sum_{j=1}^m (1 - e_j)} = \frac{1 - e_j}{m - \sum_{j=1}^m e_j} \quad (10)$$

所有权值向量 λ 满足:

$$\lambda = (\lambda_1, \lambda_2, \dots, \lambda_m), \sum_{j=1}^m \lambda_j = 1 \quad (11)$$

最后根据各属性关联度及熵权系数计算每个节点的综合关联度:

$$K(i) = \sum_{j=1}^m \lambda_j k(p_{ij}) \quad (12)$$

式中: i 为节点序号; j 为属性序号; n 为节点数; m 为属性数量。

2.4 知识发现

可拓集理论可用来描述复杂网络节点重要性动态变化趋势和方向。定义当 $T = (T_K, T_U) = (e, e)$ 时,根据关联函数值把论域分为标准正域、负域和零界,即有表达式: $V_+ = \{u | u \in U, k(u) > 0\}$, $V_- = \{u | u \in U, k(u) < 0\}$, $V_0 = \{u | u \in U, k(u) = 0\}$ 。当 $T = T_K, T_U \neq (e, e)$ 时,把论域 U 划分为关于变换 T 的正质变、负质变、正量变、负量变域和拓界,可分别形式化表示为:

$$V_{+}(T) = \{u | u \in U, y = k(u) \leq 0; T_u(u) \in U, y' = T_k k(T_u u) > 0\} \quad (13)$$

$$V_{-}(T) = \{u | u \in U, y = k(u) \geq 0; T_u(u) \in U, y' = T_k k(T_u u) < 0\} \quad (14)$$

$$V_{+}(T) = \{u | u \in U, y = k(u) > 0; T_u(u) \in U, y' = T_k k(T_u u) > 0\} \quad (15)$$

$$V_{-}(T) = \{u | u \in U, y = k(u) < 0; T_u(u) \in U, y' = T_k k(T_u u) < 0\} \quad (16)$$

$$V_0(T) = \{u | u \in U; T_u(u) \in U, y' = T_k k(T_u u) = 0\} \quad (17)$$

3 实验与分析

3.1 算法流程

对于给定的复杂社会网络,可拓聚类算法首先分析节点的多重属性并计算其重要性。然后按照实际需求划分为不同的等级,统计每个等级的取值范围及其在全体数据中的取值范围分别构成经典域和节域。通过关联函数计算每个节点关于对应等级的关联度。利用熵权值取综合关联度。最后比较不同时间点每个节点综合关联度变化,计算关联差,给出等级变动。利用这种形式化模型刻画复杂社会网络,根据关联函数定量分析各个节点隶属于某区间的范围,并在不同时间点上给出动态变动。

3.2 节点重要性多属性值

实验数据来自 Freeman 观测的由美国国家科学基金会赞助的计算机会议的参加者构成的社会网络,该网络呈现出复杂网络的基本特征。会议的参加者都是社会网络研究领域从事新兴学科专业研究的研究者,选取其中 32 个人构成的网络数据,通过熟识与否来定义网络的边取值,并在两个时间点上进行了测量,两个时间点距离 9 个月。

根据该网络图测量节点的多重重要性属性,分别计算基于度中心性 u_1 、介数中心性 u_2 、紧密度 u_3 以及 K-shell 分解值 u_4 的影响力重要性量值。标准化后得到如表 1、表 2 所示的量值矩阵。

表 1 第一时点多属性节点重要性指标

参加者	u_1	u_2	u_3	u_4
n_1	1.00	0.45	1.00	0.74
n_2	0.87	0.26	0.89	0.68
n_3	0.90	0.26	0.91	0.74
n_4	0.81	0.15	0.84	0.74
n_5	0.84	0.17	0.86	0.74
\vdots	\vdots	\vdots	\vdots	\vdots
n_{28}	0.81	0.21	0.84	0.71
n_{29}	0.87	0.25	0.89	0.74
n_{30}	0.65	0.15	0.74	0.58
n_{31}	0.87	0.33	0.89	0.68
n_{32}	0.52	0.03	0.67	0.52

表 2 第二时点多属性节点重要性指标

参加量	u_1	u_2	u_3	u_4
n_1	1.00	0.27	1.00	0.87
n_2	0.97	0.24	0.97	0.87
n_3	0.90	0.16	0.91	0.87
n_4	0.97	0.20	0.97	0.87
n_5	0.84	0.07	0.86	0.87
\vdots	\vdots	\vdots	\vdots	\vdots
n_{28}	0.94	0.18	0.94	0.87
n_{29}	0.97	0.20	0.97	0.87
n_{30}	0.84	0.12	0.86	0.77
n_{31}	1.00	0.17	1.00	0.87
n_{32}	0.87	0.14	0.89	0.81

3.3 等级划分及经典域节域

节点重要性度量后,在每个属性上取前 25% 组成高影响力等级节点的经典域 X_0 ,取全部属性上的范围为节域 X ,如表 3、表 4 所示。随着节点间的相互熟悉程度的提升,会出现属性值整体提高的可能,因此,在取经典域节域时在两个时点上分别取值,各自考察其高影响力等级节点的变化。此处的属性都是越大越好的效益型属性,最优点在经典域的最大值处取得,即右侧距中 $x_0 = b$ 处。

表3 第一时点各个属性高影响力经典域及节域

范围	u_1	u_2	u_3	u_4
经典域	$\langle 0.90, 1.00 \rangle$	$\langle 0.28, 0.45 \rangle$	$\langle 0.91, 1.00 \rangle$	$\langle 0.74, 0.74 \rangle$
节域	$\langle 0.29, 1.00 \rangle$	$\langle 0.00, 0.45 \rangle$	$\langle 0.58, 1.00 \rangle$	$\langle 0.29, 0.74 \rangle$
最优点	1.00	0.45	1.00	0.74

表4 第二时点各个属性高影响力经典域及节域

范围	u_1	u_2	u_3	u_4
经典域	$\langle 0.97, 1.00 \rangle$	$\langle 0.20, 0.27 \rangle$	$\langle 0.97, 1.00 \rangle$	$\langle 0.87, 0.87 \rangle$
节域	$\langle 0.42, 1.00 \rangle$	$\langle 0.02, 0.27 \rangle$	$\langle 0.63, 1.00 \rangle$	$\langle 0.42, 0.87 \rangle$
最优点	1.00	0.27	1.00	0.87

3.4 关联度及熵权

根据式(7)、式(8)运用关联函数公式计算每个节点的各属性的关联度,例如,有 $p_{21} = 0.87, p_{42} = 0.15$, 所以 $k_{21}(x) = \frac{\rho(x, x_0, X)}{D(x, x_0, X_0, X)} = \frac{0.90 - x}{\rho(x, x_0, X) - \rho(x, x_0, X_0)} = \frac{0.90 - 0.87}{(0.29 - 0.87) - (0.90 - 0.87)} = -0.05, k_{42}(x) = \frac{\rho(x, x_0, X)}{D(x, x_0, X_0, X)} = \frac{0.28 - x}{\rho(x, x_0, X) - \rho(x, x_0, X_0)} = \frac{0.28 - 0.15}{(0.00 - 0.15) - (0.28 - 0.15)} = -0.46$ 。

结果如表5、表6所示。

表5 第一时点各属性节点重要性关联度

参加者	u_1	u_2	u_3	u_4
n_1	0.14	0.38	0.21	0.00
n_2	-0.05	-0.07	-0.06	-0.13
n_3	0.00	-0.07	0.00	0.00
n_4	-0.15	-0.46	-0.21	0.00
n_5	-0.10	-0.39	-0.15	0.00
\vdots	\vdots	\vdots	\vdots	\vdots
n_{28}	-0.15	-0.25	-0.21	-0.07
n_{29}	-0.05	-0.11	-0.06	0.00
n_{30}	-0.41	-0.46	-0.52	-0.36
n_{31}	-0.05	0.11	-0.06	-0.13
n_{32}	-0.62	-0.89	-0.73	-0.49

表6 第二时点各属性节点重要性关联度

参加者	u_1	u_2	u_3	u_4
n_1	0.05	0.28	0.08	0.00
n_2	0.00	0.16	0.00	0.00
n_3	-0.13	-0.22	-0.18	0.00
n_4	0.00	0.00	0.00	0.00

续表6

参加者	u_1	u_2	u_3	u_4
n_5	-0.24	-0.72	-0.32	0.00
\vdots	\vdots	\vdots	\vdots	\vdots
n_{28}	-0.05	-0.11	-0.09	0.00
n_{29}	0.00	0.00	0.00	0.00
n_{30}	-0.24	-0.44	-0.32	-0.22
n_{31}	0.05	-0.17	0.08	0.00
n_{32}	-0.18	-0.33	-0.24	-0.13

根据式(9)和式(10)计算各个属性熵权值(如表7所示)。

表7 各属性不同时点熵权值

熵权	u_1	u_2	u_3	u_4
λ_1	0.32	0.09	0.34	0.25
λ_2	0.32	0.06	0.33	0.29

3.5 综合关联度及关联差

综合关联度的变化反映了节点重要性的变化,根据式(12),计算其前后两个时间节点 T_1 、 T_2 的综合关联度 $K(i)$ 和关联差 Sum_i , 数据结果如表8所示。

表8 节点重要性综合关联度及关联差

参加者	$T_1:K(i)$	$T_2:K(i)$	Sum_i	$V_i(T)$	趋势
n_1	0.15	0.06	-0.09	正量变	-
n_2	-0.08	0.01	0.09	正质变	+
n_3	-0.01	-0.11	-0.10	负量变	-
n_4	-0.16	0.00	0.16	正质变	+
n_5	-0.12	-0.23	-0.11	负量变	-
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
n_{28}	-0.16	-0.05	0.11	负量变	+
n_{29}	-0.05	0.00	0.05	正质变	+
n_{30}	-0.44	-0.27	0.17	负量变	+
n_{31}	-0.06	0.03	0.09	正质变	+
n_{32}	-0.65	-0.19	0.46	负量变	+

在这里,不仅能看到节点重要性变化方向,而且给出了变动的量。在趋势一栏里,节点重要性有所下降则其符号为负,相应地,节点重要性有所提高则符号为正。比较第一时点与第二时点综合关联度,如果符号相同则为量变,符号不同则为质变。更具体地,如果第一时点符号为负第二时点为正则等级提升,该节点变为高影响力节点。相反,如果第一时点符号为正第二时点为负则等级下降,该节点变为非高影响力节点。例如,节点2的综合关联度由负号变为正号,说明其重

要性等级一定有所提升,由非高影响力节点转变为高影响力节点,发生了质变,即“正质变”,且其变化量是0.09。

将各节点重要性可视化表示如图2所示,利用坐标系能直观反映各节点重要性变化方向,变动量则为其差值。可见,零值以上部分对应着前25%高影响力等级节点,在第二时点有部分节点由高影响力节点转变为非高影响力节点,也有部分节点由非高影响力节点转变为高影响力节点,还有部分节点只有同一等级内的量变。

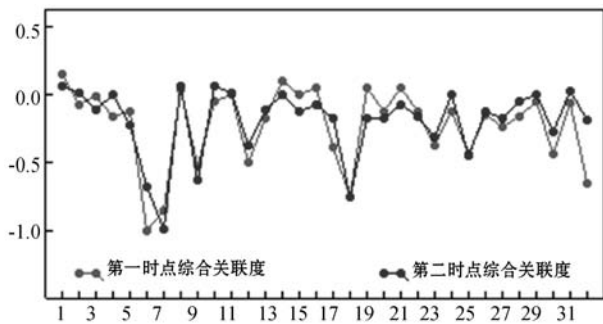


图2 各节点在不同时点的综合关联度变化

4 算法比较

假设共有 N 个节点, r 个类别等级, 各节点有 m 个特征, 则 K 均值算法针对每一个特征首先选取 k 个样本节点, 然后计算每个节点与各个样本节点聚类中心的距离, 时间复杂度为 $O(N \times k)$, 共 m 个特征, 则总复杂度为 $O(N \times k \times m)$ 。而 DBSCAN 密度聚类算法, 针对每一个特征, 首先随机选取一个节点, 然后检查其领域半径内是否达到密度要求, 若是则新建聚类, 把所有点加入候选集, 在候选集中对没处理过的对象继续检查其领域, 如此往复, 时间复杂度为 $O(N \times m)$ 。而可拓聚类算法将每个特征上的节点重要性划分为 r 个类别等级, 时间复杂度均为 $O(m \times r)$ 。由于 $N \gg r$, 所以可拓聚类时间复杂度较低。当然, 可拓聚类中的等级个数主要是由根据实际问题或者经验知识划分得到, 粒度相对较粗, 但可拓聚类的特点在于能够描述网络节点影响力在动态变化下的量变和趋势, 其他的聚类算法则不具备。

5 结语

复杂网络往往具有多对多的复杂关系, 对复杂网络的把握有利于掌握舆论热点和对关键的认定。而对复杂社会网络节点重要性动态观测, 从变化的视角审视网络, 本文给出了可拓聚类分析方法。该方法根据

多属性计算节点重要性, 首先通过聚类分析划分节点重要性等级, 构造经典域和节域。然后根据关联函数计算节点与对应等级关联度, 进而计算综合关联度, 最后分析不同时点的变化量。通过可拓聚类方法实现对节点重要性变化的动态把握, 使网络节点及各指标均能以形式化的模型更直观地展现。该方法不仅能判别节点重要性的变化方向, 而且给出了变化的量。接下来工作包括: (1) 对可拓聚类预测方法在更大数据集的预测效果作进一步探索; (2) 在节点重要性划分等级时尝试利用其他方法, 而不是根据需要主观划分; (3) 在一个有向带权网络图上利用该方法, 验证是否具有相同结果等。

参考文献

- [1] Barabasi A L, Albert R, Jeong H, et al. Power law distribution of the world wide web[J]. Science, 2000, 287(5461): 2115a.
- [2] 谭跃进, 吴俊, 邓宏钟. 复杂网络中节点重要度评估的节点收缩方法[J]. 系统工程理论与实践, 2006, 26(11): 79-83.
- [3] 张琨, 李配配, 朱保平, 等. 基于 PageRank 的有向加权复杂网络节点重要性评估方法[J]. 南京航空航天大学学报, 2013, 45(3): 429-434.
- [4] 张益. 一种定量评估复杂网络节点重要度的算法[J]. 计算工程, 2011, 37(20): 87-88.
- [5] 于会, 刘尊, 李勇军. 基于多属性决策的复杂网络节点重要性综合评价方法[J]. 物理学报, 2013, 62(2): 020204-1-9.
- [6] Albert R, Jeong H, Barabasi A L. Attack and error tolerance of complex networks[J]. Nature, 2000, 406: 378-382.
- [7] Freeman L C. A set of measures of centrality based on betweenness[J]. Sociometry, 1977, 40(1): 35-41.
- [8] Wehmuth K, Ziviani A. DAC CER: Distributed Assessment of the Closeness Centrality Ranking in complex networks[J]. Computer Networks, 2013, 57(13): 2536-2548.
- [9] Kitsak M, Gallos L K, Havlin S, et al. Identification of influential spreaders in complex networks[J]. Nature Physics, 2010, 6(11): 888-893.
- [10] Albert R, Jeong H, Barabasi A L. Attack and error tolerance of complex networks[J]. Nature, 2000, 406: 378-382.
- [11] Freeman L C. Centrality in social networks: Conceptual clarification[J]. Social Networks, 1979, 1: 215-239.
- [12] Opsahl T, Agneessens F, Skvoretz J. Node centrality in weighted networks: Generalizing degree and shortest paths[J]. Social Networks, 2010, 32(3): 245-251.
- [13] Sabidussi G. The centrality index of a graph[J]. Psychometrika, 1966, 31: 581-603.
- [14] Clauset A, Shalizi C R, Newman M E. Power-law distributions in empirical data[J]. SIAM Review, 2009, 51(4):

661 - 703.

- [15] Carmi S, Havlin S, Kirkpatrick S, et al. A model of Internet topology using k-shell decomposition[J]. Proceedings of the National Academy of Sciences of the United States of America, 2007, 104(27): 11150 - 11154.
- [16] 杨春燕, 蔡文. 可拓学[M]. 北京: 科学出版社, 2014.
- [17] 杨国为, 王守觉. 模式可拓识别及其神经网络模型[J]. 哈尔滨工业大学学报, 2006, 38(7): 1130 - 1132.
- [18] Xu L B, Li X S, Pang C Y. Uncertain Multiattribute Decision-Making Based on Interval Number with Extension-Dependent Degree and Regret Aversion[J/OL]. Mathematical Problems in Engineering, 2018, 2018. <https://doi.org/10.1155/2018/6508636>.
- [19] 周玉, 钱旭, 张俊彩, 等. 可拓神经网络研究综述[J]. 计算机应用研究, 2010, 27(1): 1 - 2.
- [20] 蔡文, 杨春燕, 陈文伟, 等. 可拓集与可拓数据挖掘[M]. 北京: 科学出版社, 2008.
- [21] 杨春燕. 可拓创新方法[M]. 北京: 科学出版社, 2017.
- [22] 蔡文, 杨春燕. 可拓学的基础理论与方法体系[J]. 科学通报, 2013, 58(13): 1190 - 1199.

(上接第7页)

5 结 语

本文针对移动云环境下计算任务迁移的响应时间和终端设备能耗综合优化问题进行了研究。针对延迟比较敏感且计算要求相对不高的多个相互依赖的任务, 提出了一种基于遗传算法思想的迁移方法。该方法根据计算任务的不同类型和不同微云的分类处理能力描述来计算迁移方案的效用值并完成选择操作, 基于各个任务之间的依赖关系来完成交叉和变异操作, 从而通过多次迭代最终生成综合效用值较优的任务迁移方案。仿真实验结果表明, 本文方法能有效减少移动终端的响应时间和能耗。

本文的研究仍然存在不足和值得进一步研究的问题, 例如, 如果移动用户处于移动状态, 而且所应用到的无线网络也是会发生变化的。将用户的移动性和网络的时变性加入到迁移模型中, 更加贴合现实生活, 再在此基础上研究迁移方法, 是未来的工作之一。

参 考 文 献

- [1] 张文丽, 郭兵, 沈艳, 等. 智能移动终端计算迁移研究[J]. 计算机学报, 2016, 39(5): 1021 - 1038.
- [2] Liu F, Shu P, Jin H, et al. Gearing resource-poor mobile devices with powerful clouds: Architectures, challenges, and applications[J]. IEEE Wireless Communications, 2013, 20(3): 1334 - 1354.
- [3] Ge C, Sun Z, Wang N. A survey of power-saving techniques on data centers and content delivery networks[J]. IEEE Communications Surveys & Tutorials, 2013, 15(3): 1334 - 1354.
- [4] Kumar K, Lu Y H. Cloud computing for mobile users: Can offloading computation save energy[J]. Computer, 2010, 43(4): 51 - 56.
- [5] Ge Y, Zhang Y, Qiu Q, et al. A game theoretic resource allocation for overall energy minimization in mobile cloud computing system[C]//Proceedings of the IEEE International Symposium on Low Power Electronics and Design. California, USA, 2012: 279 - 284.
- [6] Chen X. Decentralized computation offloading game for mobile cloud computing[J]. IEEE Transactions on Parallel and Distributed Systems, 2015, 26(4): 974 - 983.
- [7] Gao B, He L, Lu X, et al. Developing energy-aware task allocation schemes in cloud-assisted mobile workflows[C]//Proceedings of the IEEE International Conference on Computer and Information Technology. California, USA, 2015: 1 - 6.
- [8] 史雯隼, 武继刚, 罗裕春. 针对移动云计算任务迁移的快速高效调度算法[J]. 计算机科学, 2018, 45(4): 94 - 99.
- [9] Zhang W, Wen Y, Wu D. Collaborative task execution in mobile cloud computing under stochastic wireless channel[J]. IEEE Transactions on Wireless Communications, 2015, 13(1): 1536 - 1276.
- [10] Zhang W, Wen Y, Wu D. Energy-efficient scheduling policy for collaborative execution in mobile cloud computing[C]//Proceedings of the IEEE INFOCOM. Turin, Italy, 2013: 190 - 194.
- [11] Liu X, Yuan C W, Li Y, et al. A lightweight algorithm for collaborative task execution in mobile cloud computing[J]. Wireless Personal Communications, 2015, 86(2): 579 - 599.
- [12] Shiraz M. A Lightweight Active Service Migration Framework for Computational Offloading in Mobile Cloud Computing[J]. The Journal of Supercomputing, 2013, 68(2).
- [13] Magurawalage C M S, Yang K, Hu L, et al. Energy-efficient and network-aware offloading algorithm for mobile cloud computing[J]. Computer Network, 2014, 74(9): 22 - 33.
- [14] Chen M, Hao Y, Li Y, et al. On the computation offloading at ad hoc cloudlet: architecture and service modes[J]. IEEE Communications Magazine, 2015, 53(6): 18 - 24.
- [15] Cardellini V, Vittoria De Nitto Personé, Valerio V D, et al. A game-theoretic approach to computation offloading in mobile cloud computing[J]. Mathematical Programming, 2016, 157(2): 421 - 449.
- [16] 叶苗, 王宇平, 魏静莹. 基于多移动节点和遗传算法的传感器网络覆盖修复策略[J]. 通信学报, 2014, 35(12): 45 - 52.