

# 基于进化谱分方法的动态社团检测

付立东 聂靖靖

(西安科技大学计算机学院 西安 710054)

**摘要** 为了有效地分析动态网络中的社团结构功能和特性,在进化时间平滑框架下基于进化聚类方法对模块密度函数和否定平均关联函数进行了优化,论证了理论可行性;在此基础上提出了检测动态网络社团结构的进化谱分算法,并对两类算法进行了详细的谱分分析。分别在计算机合成的动态网络以及真实网络中检验了所提算法的准确性和有效性,并将其与其他算法进行对比。实验结果表明,所提算法对动态网络中的社团检测仍有很高的准确性和有效性。

**关键词** 动态网络,社团结构,模块密度,否定平均关联,进化谱分

中图分类号 TP399 文献标识码 A DOI 10.11896/j.issn.1002-137X.2018.02.030

## Dynamic Community Detection Based on Evolutionary Spectral Method

FU Li-dong NIE Jing-jing

(School of Computer, Xi'an University of Science and Technology, Xi'an 710054, China)

**Abstract** In order to effectively analyze the function and characteristics of the community structure in the dynamic network, the module density function and the negative average correlation function were optimized based on the evolutionary clustering algorithm under the evolutionary time smoothing framework, and the theoretical feasibility was demonstrated. The evolution spectrum algorithm was proposed based on community structure of the dynamic network. The accuracy and effectiveness of the proposed algorithm was verified and compared with other algorithms in the computer synthesis and real dynamic network respectively. The experimental results show that the proposed algorithm is still very accurate and effective in the community detection of dynamic network.

**Keywords** Dynamic network, Community structure, Module density, Negative average correlation, Evolution spectrum

现实世界中实体之间存在的复杂关系可用网络来描述,网络已成为研究复杂系统强有力的工具。研究表明,许多网络中都存在社团结构,如社会网络、Web 网络以及生物网络<sup>[1]</sup>。对复杂网络中的社团结构<sup>[2]</sup>进行检测已成为进一步探索复杂系统结构与功能对应关系的关键。例如,在生物网络中,蛋白质复合物对应着社团结构,往往在生物的生命活动中具备某种特定功能<sup>[3]</sup>。目前已有大量用于测静态网络中的社团结构的聚类算法被提出<sup>[4]</sup>。然而,根据真实世界构建的网络总是随时间不断变化的。在社交、生物等动态网络中检测社团结构,更能帮助人们揭示网络的进化进程。例如,在科学家的合作网络中,科学家的分组随着科学家研究方向的转变而改变<sup>[5]</sup>。在生物网络中,基因表达会随着疾病的发展而进化。因此,在静态网络中进行社团检测难以捕获自然规律及重要的动态特性。

然而,设计一种高效检测动态网络中的社团结构的算法是非常困难的。主要原因有两个:1)依然不清楚如何定义动态社团,难以描述变化着的社团结构;2)动态网络有着多层次和复杂的特性。为解决第一个问题, Li 等人提出了模块密度函数  $D$ <sup>[6]</sup> 来评估静态网络中的社团质量。进化聚类方法常常被用于捕获动态网络的进化特性,这种进化聚类方法通常是

在时间平滑框架下<sup>[7]</sup>假设网络社团结构在一小段时间内没有明显变化的情况下进行的。

目前已有许多用于检测动态社团结构的方法被提出,这些方法大体上可分为两类:基于非进化的方法<sup>[8]</sup>和基于进化的方法<sup>[9-10]</sup>。在基于非进化的方法中,在每个时间步上,不考虑历史社团结构,仅考虑当前的社团结构。这种策略忽视了历史信息对当前社团结构的影响。而基于进化的方法的社团检测算法,不但考虑了当前的社团结构,而且考虑了社团结构的历史信息,使得检测出的社团结构更加合理和准确。传统的动态社团检测方法忽略了后续时间之间的连接步骤,导致了不良的社团划分;使用的优化方式不合理,会导致成本过高,测出的算法的准确性不高。因此,本文通过重构时间平滑度框架,证明了进化聚类方法之间的理论联系,解决了动态社团检测多目标优化问题。

本文在进化框架下将模块密度函数  $D$  以及图划分<sup>[11]</sup>的否定平均关联函数<sup>[12]</sup>进一步推广到动态网络中,以评估社团结构,对它们进行谱分优化,并在此基础上提出了检测动态网络的谱分算法;在计算机合成的动态网络中验证了算法的准确性,在真实网络中测试了算法的有效性,并将其与其他算法进行了比较。

## 1 相关工作

### 1.1 进化方法

进化方法引入了时间平滑性框架来检测动态社团,平衡了两个连续时间点获得的社团,包括快照成本(CS)和时间成本(CT),其中CS量化了当前网络时间步长社团结构的好坏,CT测量了当前时间步长与上一时间步长社团结构相比的变化程度。进化方法的最终目标是获得两个子成本之间的最优交易。使用得最广泛的策略是组合CS和CT的加权线性函数。

$$Cost = \rho CS + (1 - \rho) CT \quad (1)$$

其中, $\rho$ 为控制CS和CT的相关性参数。当 $\rho=1$ 时,算法捕获当前网络的聚类;当 $\rho=0$ 时,返回前一时间步长的网络聚类。通常, $\rho \in [0, 1]$ ,其控制每个子成本的偏好程度。

### 1.2 符号定义

设置 $\{1, 2, \dots, T\}$ 为一组有限的时间步长。对于给定的变量,下标 $t$ 代表 $t$ 时刻变量的值。动态网络图 $G = (V, E, A)$ 为一系列子网络的集合,即 $G = \{G_1, G_2, \dots, G_T\}$ ,其中 $G_t$ 是 $t$ 时刻的网络, $V_t$ 为 $t$ 时刻网络中节点的集合, $E_t$ 代表 $t$ 时刻网络中边的集合, $A_t$ 代表 $t$ 时刻的网络邻接矩阵。动态网络被定义为三维的邻接矩阵,设置 $A = (a_{ijt})_{n \times n \times T}$ ,其中 $n$ 为节点的数目,当网络中的节点 $i$ 和节点 $j$ 之间有连接边时, $a_{ijt}$ 为1,否则为0。 $A = [A_1, A_2, \dots, A_T]$ ,其中 $A_t = (a_{ijt})_{n \times n \times T}$ 为带权网络 $G_t$ 的邻接矩阵。当 $G_t$ 为无向图时, $A_t$ 为对称矩阵。本文假设网络为无向网络。在 $t$ 时刻的网络 $G_t$ 中,第 $i$ 个节点的度被定义为连接到它的所有节点的数目,即 $d_{it} = \sum_{j=1}^n a_{ijt}$ 。动态网络 $G$ 的度被定义为 $D = [D_1, D_2, \dots, D_T]$ , $D_t$ 为 $t$ 时刻网络度序列的对角矩阵,即 $D_t = \text{diag}(d_{1t}, d_{2t}, \dots, d_{nt})$ 。

给定网络 $G_t = (V_t, E_t, A_t)$ ,社团结构对应硬分区 $\{P_{1t}, P_{2t}, \dots, P_{\omega t}\}$ ,当 $i \neq j$ 时, $P_{it} \cap P_{jt} = \emptyset, V_t = \sum_i P_{it}$ 。 $P_{it}$ 是 $t$ 时刻的第 $i$ 个局部社团(聚类)。 $\{P_{it}\}_{i=1}^{\omega}$ 可以用 $n \times \omega$ 的指示矩阵 $Z_t = (z_{ijt})_{n \times \omega}$ 表示,在 $t$ 时刻,当第 $i$ 个节点属于第 $j$ 个局部聚类时, $z_{ijt} = 1$ ,否则 $z_{ijt} = 0$ 。为了方便,通过局部聚类的大小将矩阵 $Z_t$ 的列规范化,即 $z_{.jt} / \sqrt{|P_{jt}|}$ ,其中 $z_{.jt}$ 为 $Z_t$ 的第 $j$ 列。定义归一化指示矩阵为 $\tilde{Z}_t = [\tilde{z}_{.1t}, \dots, \tilde{z}_{.\omega t}]$ ,其中 $\tilde{z}_{.jt} = z_{.jt} / \sqrt{|P_{jt}|}$ 。容易验证 $\tilde{Z}_t$ 为列正交,即 $\tilde{Z}_t^T \tilde{Z}_t = I_{\omega}$ ,其中 $I_{\omega}$ 为单位矩阵, $\tilde{Z}_t^T$ 为 $\tilde{Z}_t$ 的转置矩阵。

### 1.3 NA函数和D模块密度函数

在图划分问题中,通过优化基本图表的测量方式来解决图分割问题。其中最常用的两种方式是最大化平均关联和最小化归一化切割。对于节点集 $V$ 的两个子集,它们之间的关联函数为 $L(P_u, P_v)$ ,那么 $\omega$ 路平均关联为:

$$AA(\{P_{it}\}_{i=1}^{\omega}) = \sum_{i=1}^{\omega} \frac{L(P_{it}, P_{it})}{|P_{it}|} \quad (2)$$

$\omega$ 路归一化切割定义为:

$$NC(\{P_{it}\}_{i=1}^{\omega}) = \sum_{i=1}^{\omega} \frac{L(P_{it}, \bar{P}_{it})}{L(P_{it}, V)} \quad (3)$$

平均关联值越大代表社团内部的关联越紧密;规范化切割值越小代表聚类的内部关系越紧密,聚类间的关系越稀疏。为了统一,将否定平均关联函数NA作为衡量指标,具体公式如下:

$$NA(\{P_{it}\}_{i=1}^{\omega}) = \text{tr}(A_t) - \sum_{i=1}^{\omega} \frac{L(P_{it}, P_{it})}{|P_{it}|} \quad (4)$$

式(4)表明,否定平均关联函数NA越小,则图的划分越好。

网络的模块密度函数D被定义为:

$$D(\{P_{it}\}_{i=1}^{\omega}) = \sum_{i=1}^{\omega} \frac{L(P_{it}, P_{it}) - L(P_{it}, \bar{P}_{it})}{|P_{it}|} \quad (5)$$

其中, $P_{it}$ 是 $t$ 时刻的第 $i$ 个局部社团(聚类), $\bar{P}_{it}$ 为 $P_{it}$ 的补集, $L(P_{it}, P_{jt}) = \sum_{i \in P_{it}, j \in P_{jt}} a_{ijt}$ ,且 $\bar{P}_{it} = V \setminus P_{it}$ , $L(P_{it}, P_{jt})$ 度量了 $t$ 时刻社团内部边的权重和,顶点的度为顶点到其他所有顶点的边数,即 $\text{degree}(P_{it}) = L(P_{it}, V)$ ,有 $L(P_{it}, \bar{P}_{it}) = \text{degree}(P_{it}) - L(P_{it}, P_{it})$ ;  $L(P_{it}, \bar{P}_{it})$ 度量了不同社团之间边的权重; $|P_{it}|$ 是 $t$ 时刻第 $i$ 个社团内顶点的数目,度量了社团的大小。

## 2 谱分优化

### 2.1 D函数

根据式(1),在进化框架下,通过优化模块密度函数D来优化进化聚类的总体成本。函数可以被定义为:

$$Cost_D = \rho CS_D + (1 - \rho) CT_D \\ = \rho(D)_t |Z_t| + (1 - \rho)(D)_{t-1} |Z_t| \quad (6)$$

其中:

$$(D)_t |Z_t| = \sum_{i=1}^{\omega} \left[ \frac{Z_i^T A_t Z_i - Z_i^T (D_t - A_t) Z_i}{Z_i^T Z_i} \right] \quad (7)$$

其中, $Z_i^T Z_i$ 为社团大小, $Z_i^T A_t Z_i$ 与社团内部边的权重和相等, $Z_i^T (D_t - A_t) Z_i$ 的值和社团与其他社团边的权重和相等。式(7)可进一步表示为:

$$(D)_t |Z_t| = \sum_{i=1}^{\omega} z_{.it}^T (2A_t - D_t) z_{.it} \quad (8)$$

将式(8)转化为迹的形式:

$$(D)_t |Z_t| = \text{tr}[\tilde{Z}_t^T (2A_t - D_t) \tilde{Z}_t] \quad (9)$$

同理:

$$(D)_{t-1} |Z_t| = \text{tr}[\tilde{Z}_t^T (2A_{t-1} - D_{t-1}) \tilde{Z}_t] \quad (10)$$

将式(9)和式(10)代入到式(6)中,得:

$$Cost_D = \rho \{ \text{tr}[\tilde{Z}_t^T (2A_t - D_t) \tilde{Z}_t] \} + (1 - \rho) \{ \text{tr}[\tilde{Z}_t^T (2A_{t-1} - D_{t-1}) \tilde{Z}_t] \} \\ = \text{tr}[\tilde{Z}_t^T [ \rho(2A_t - D_t) + (1 - \rho)(2A_{t-1} - D_{t-1}) ] \tilde{Z}_t] \quad (11)$$

因此存在以下关系:

$$\max Cost_D = \max \{ \text{tr}[\tilde{Z}_t^T ( \rho(2A_t - D_t) + (1 - \rho)(2A_{t-1} - D_{t-1}) )^T \tilde{Z}_t] \} \quad (12)$$

### 2.2 NA函数

根据式(1),在PCQ进化框架下,通过NA函数来优化进化聚类的总体成本。函数可以被定义为:

$$Cost_{NA} = \rho CS_{NA} + (1 - \rho) CT_{NA} \\ = \rho(NA)_t |Z_t| + (1 - \rho)(NA)_{t-1} |Z_t| \quad (13)$$

其中:

$$(NA)_t |Z_t| = \text{tr}(A_t) - \sum_{i=1}^{\omega} \left( \frac{Z_i^T A_t Z_i}{Z_i^T Z_i} \right) \\ = \text{tr}(A_t) - \text{tr}(\tilde{Z}_t^T A_t \tilde{Z}_t) \quad (14)$$

同理:

$$(NA)_{t-1} | Z_t = \text{tr}(A_{t-1}) - \text{tr}(\tilde{Z}_t^T A_{t-1} \tilde{Z}_t) \quad (15)$$

将式(14)和式(15)代入式(13)中,有:

$$\begin{aligned} Cost_{NA} = & \rho[\text{tr}(A_t) - \text{tr}(\tilde{Z}_t^T A_t \tilde{Z}_t)] + (1-\rho)[\text{tr}(A_{t-1}) - \\ & \text{tr}(\tilde{Z}_t^T A_{t-1} \tilde{Z}_t)] \\ = & \text{tr}[\rho A_t + (1-\rho)A_{t-1}] - \text{tr}\{\tilde{Z}_t^T [\rho A_t + \\ & (1-\rho)A_{t-1}] \tilde{Z}_t\} \end{aligned} \quad (16)$$

由于  $\text{tr}[\rho A_t + (1-\rho)A_{t-1}]$  为一常数,因此有:

$$\min Cost_{NA} \propto \max_{\tilde{Z}_t} \text{tr}\{\tilde{Z}_t^T [\rho A_t + (1-\rho)A_{t-1}] \tilde{Z}_t\} \quad (17)$$

### 3 算法

最大化  $Cost_D$  和最小化  $Cost_{NA}$  都是 NP 难问题,因此寻找特征值向量较为困难。通过谱分松散计算列矩阵  $\tilde{Z}_t$  的值,对生成的特征向量进行后继处理,得到复杂网络顶点的离散划分,从而找到  $\{P_u\}_{i=1}^{\omega_t}$  社团的最后社团结构。算法的详细描述如下:

输入:动态网络  $G$ ;每时刻  $t$  上的聚类数目  $\{\omega_t\}_{i=1}^T$

输出:动态网络的社团结构  $\{P_t\}_{i=1}^T$

Step1 计算矩阵  $Q_t = \rho(2A_t - D_t) + (1-\rho)(2A_{t-1} - D_{t-1})$  或者  $Q_t = \rho A_t + (1-\rho)A_{t-1}$ 。

Step2 利用兰兹稀疏特征向量分解方法来计算矩阵  $Q_t$  的首个  $\omega$  特征向量  $u_{1t}, u_{2t}, \dots, u_{\omega t}$ 。

Step3 构建元素为  $[u_{1t}, u_{2t}, \dots, u_{\omega t}]^T$  的列矩阵  $\tilde{Z}_t$ 。

Step4 当  $i=1$  时,社团为  $t$  时刻的动态网络本身。当  $2 \leq i \leq \omega_t$  时,重复以下步骤:

Step4.1 生成来自矩阵  $\tilde{Z}_t$  的首个  $i$  列矩阵  $W_{it}$ ;

Step4.2 使用 k-means 聚类算法聚类  $W_{it}$  的行向量。

Step5 执行 Step4,当  $Cost_D$  值不再增大,  $Cost_{NA}$  值不再减小,或者达到最大迭代次数时,  $i$  的值就是  $t$  时刻网络所对应的最优社团划分数目。

首先讨论算法的空间复杂度:给定一个动态网络  $G$ ,三维邻接矩阵  $A_{n \times n \times T}$  需要的空间复杂度为  $O(n^2 T)$ ;给定一个邻接矩阵  $A_t$ ,它的指示矩阵的空间复杂度为  $O(n\omega_t)$ ;矩阵  $A$  的整个分解过程的空间复杂度为  $O(n(\omega_1 + \omega_2 + \dots + \omega_n)) = O(nT \max\{\omega_1, \dots, \omega_n\})$ ,由于  $\omega_i \leq n(1 \leq i \leq T)$ ,因此矩阵分解过程的空间复杂度为  $O(n^2 T)$ 。

接着讨论算法的时间复杂度:针对每个时刻  $t$ ,算法由矩阵的计算分解以及 k-means 聚类两个主要部分组成,矩阵计算分解的时间复杂度为  $O(n^3)$ ,k-means 聚类在  $t$  时刻的时间复杂度为  $O(rn^2 \omega_t)$ ,其中  $r$  为迭代次数。因此,每个时间步长的时间复杂度为  $O(n^3 + rn^2 \omega_t + \omega_{t-1} \omega_t) = O(n^3 + rn^2 \omega_t)$ ,其中  $\omega_t < n$ 。算法的整体时间复杂度为  $O(T(n^3 + rn^2 \max\{\omega_t\}_{i=1}^T))$ 。

## 4 实验

### 4.1 计算机合成网络

为检验算法的准确性,使用归一化互信息(NMI)作为评判算法准确性的标准。给出一个标准网络划分  $M^*$ ,以及由算法获得的网络划分  $M$ ,构建一个混淆矩阵  $N$ ,  $N$  中行对应

$M^*$  中的社团结构,列对应  $D$  中的社团结构。 $N_{ij}$  表示第  $i$  个社团与第  $j$  个社团的重叠节点数目。NMI 可表达为:

$$NMI(M^*, M) = \frac{-2 \sum_{i=1}^{|M|} \sum_{j=1}^{|M^*|} N_{ij} \log \frac{N_{ij} N}{N_i N_j}}{\sum_{i=1}^{|M|} N_i \log \frac{N_i}{N} + \sum_{j=1}^{|M^*|} N_j \log \frac{N_j}{N}} \quad (18)$$

其中,  $|M|$  和  $|M^*|$  分别表示  $M$  和  $M^*$  内社团的数目,  $N_i$  和  $N_j$  分别表示  $N$  中第  $i$  行和第  $j$  列元素的和 NMI 处于 0 和 1 之间,其值越大表示算法的准确性越高。

为了使合成的数据集更接近于真实世界网络的数据, Greene 等人合成了一个具有 4 种基本事件的数据集。每个网络包含 1000 个节点,每个节点的平均度为 15,其中最大度为 50。在首个时间步长中,社团的平均数量为 33,社团之间的边所占的百分比为 0.2。将 4 种基本类型的事件添加到事件发生器中,具体步骤如下。

1) 出生与死亡:在每个时间步长中,移除现有社团的节点,10%的社团被建立,10%的其他现有社团被随机删除。

2) 合并与分拆:在每个时间步长中,10%的社团进行了分裂,其他 10%的社团被选中,对选中的社团随机地两两合并,从而得到新的社团。

3) 扩张与收缩:在每个时间步长中,10%的社团被随机选择,在原来的规模上扩大或缩小 25%。

4) 间歇性社团:在每个时间步长中,10%的社团在此时间步长中从原始网络中消失,并在下一个时间步中重新出现。

每种类型的事件都包含 5 个网络。基于模块密度函数  $D$  和否定平均关联函数  $NA$  的进化谱分算法在 4 个基本事件的不同网络中检测动态网络社团的准确性(见图 1),并与文献[13]中的动态多目标遗传算法(DYNOMGA)做了比较,结果显示:两种谱分算法在 4 种基本事件的动态网络中的表现比较稳定,且都比 DYNOMGA 算法准确。

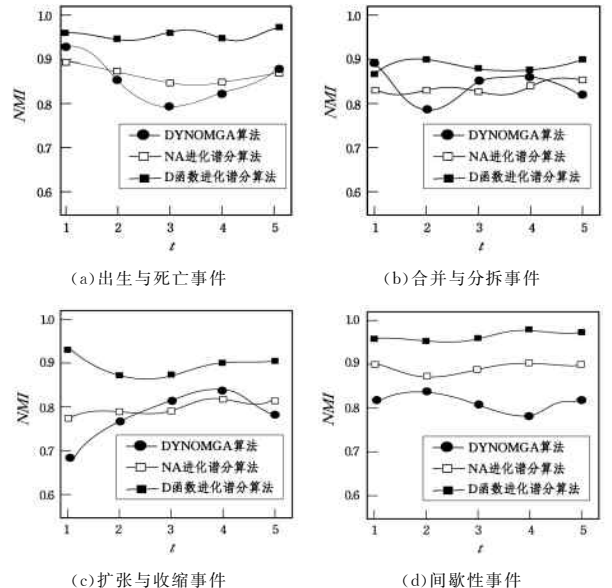


图 1 3 种算法在 4 个基本事件的不同网络中检测动态网络社团的准确性对比

Fig. 1 Comparison of accuracy of three algorithm detecting dynamic network community in different networks including four basic events

## 4.2 基因共表达动态网络

本节将提出的进化谱分算法应用于大规模生物数据集上,以检验其在动态复杂网络上的有效性。采用的生物数据集来自 GEO 数据库(GSE74817),该数据库存储了 2010 年 11 月份流感季节成年人接种 TIV 疫苗后的基因表达的具体数据<sup>[14]</sup>。在一周内的 4 个时间点(疫苗接种后的第 0,1,3 和 7 天)进行了数据收集。针对收集的数据,通过局部等级算法构建 4 个基因共表达网络,每个网络包含 7559 个链接节点。

通过算法获得的 4 个网络中的基因概率来评估聚类性能。算法通过式(19)捕获基因的概率。

$$p(X \geq x) = 1 - \sum_{k=0}^{x-1} \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}} \quad (19)$$

其中, $N$ 是基因的总数, $M$ 是已知功能基因的数目, $n$ 是确定的一个簇(聚类)的候选基因的数量, $k$ 是它们之间共用的基因数量。这里评估基因数至少为 5 的簇。

所提算法与广泛使用的 FaceNet 算法检测社团时的有效性如图 2 所示,其中显示了几种算法预测出的重要进化簇的百分比,也显示了数据库中生物进程类别中的基因本体术语参考值有着明显重叠,同时表明了两种进化谱分算法可以有效地检测重要的功能集群,并且在疫苗接种后的第 3 天得到最好的有效百分比。FaceNet 算法产生了大量的小簇,这些小簇会导致功能基因簇的比例明显降低。

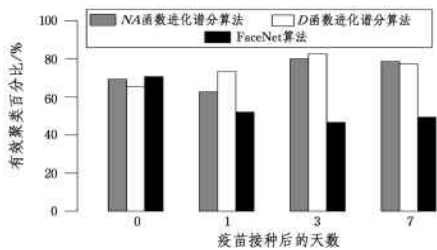


图 2 两种算法在疫苗接种后的基因表达数据集上检测到的有效聚类百分比

Fig. 2 Effective clustering percentage detected by two algorithms on gene expression datasets after vaccination

**结束语** 本文在进化时间平滑框架下对模块密度函数和否定平均关联函数进行了优化,论证了这两类函数作为进化谱分聚类方法检测动态网络中社团结构的理论可行性,提出了检测动态网络社团结构的进化谱分算法。在计算机合成的动态网络以及真实网络中检验了所提算法的准确性和有效性,并与其他可行算法进行了比较。实验结果显示,所提算法对动态网络中的社团检测具有很高的准确性和有效性。

未来准备将本算法应用到更多的大型真实网络中检测其准确性,并且将探索更多的社团评价函数或图划分函数,将其纳入进化时间平滑框架中,进一步研究它们在动态网络社团检测中的特性。

## 参考文献

[1] NOWACKA-WOSZUK J, PRUSZYNSKA-OSZMALEK E, SZYDŁOWSKI M. Nutrition modulates Fto and Irf3 gene tran-

script levels, but does not alter their DNA methylation profiles in rat white adipose tissues[J]. *Gene*, 2017, 610:44-48.

- [2] REN L K, LI H J, JIA C L, et al. Near Linear Time Community Detection Algorithm Based on Dynamical Evolution[J]. *Computer Science*, 2016, 43(6A):395-399, 412. (in Chinese)  
任涿崑, 李慧嘉, 贾传亮, 等. 近似线性时间的社团结构动态演化挖掘算法[J]. *计算机科学*, 2016, 43(6A):395-399, 412.
- [3] CRAENE B D, BERX G. Regulatory networks defining EMT during cancer initiation and progression[J]. *Nature Review Cancer*, 2013, 13(2):97-110.
- [4] LI H J. Fast Algorithm for Detecting Multi-scale Overlapping Community Structure Based on Information Spreading[J]. *Computer Science*, 2014, 41(9):125-131. (in Chinese)  
李慧嘉. 基于信息扩散的多尺度重叠社团快速探测算法[J]. *计算机科学*, 2014, 41(9):125-131.
- [5] YU S Y, WANG H M. Scientific collaboration: a social network analysis based on literature of animal-derived regenerative implantable medical devices[J]. *Regenerative Biomaterials*, 2016, 3(3):197-203.
- [6] LI Z P, ZHANG S H, WANG R S, et al. Quantitative function for community detection[OL]. <https://www.ncbi.nlm.nih.gov/pubmed/18517463>.
- [7] GUPTA M, GAO J, AGGARWAL C, et al. Outlier detection for temporal data: a survey[J]. *IEEE Transaction on Knowledge and Data Engineering*, 2014, 26(9):2250-2267.
- [8] EGIZI A, FEFFERMAN N H, FONSECA D M. Evidence that implicit assumptions of 'no evolution' of disease vectors in changing environments can be violated on a rapid timescale[OL]. <http://resb.royalsocietypublishing.org/content/roytpb/370/1665/20140136.full.pdf>.
- [9] MA X K, DONG D. Evolutionary nonnegative matrix factorization algorithms for community detection in dynamic networks[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2017, 29(5):1045-1058.
- [10] FOLINO F, PIZZUTI C. An evolutionary multi-objective discovery in dynamic networks[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2014, 26(8):1838-1852.
- [11] QIN X, LIANG W, YUAN C A, et al. Image Segmentation Algorithm of Spectral Clustering Optimized by Genetic[J]. *Computer Science*, 2017, 44(1):100-102. (in Chinese)  
覃晓, 梁伟, 元昌安, 等. 基于遗传优化谱聚类的图形分割方法[J]. *计算机科学*, 2017, 44(1):100-102.
- [12] CHI Y, SONG X D, ZHOU D Y, et al. On evolutionary spectral clustering[J]. *ACM Transactions on Knowledge Discovery from Data*, 2009, 3(4):1-30.
- [13] FOLINO F, PIZZUTI C. An evolutionary multi-objective approach for community discovery in dynamic networks[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2014, 26(8):1838-1852.
- [14] WANG P Z, GAO L, MA X K. Dynamic community detection based on network structural perturbation and Topological similarity[J]. *Journal of Statistical Mechanics: Theory and Experiment*, 2017, 1(1):013401.