# 多智能体系构架下的属性图分布式聚类算法

边宅安 李慧嘉 陈俊华 马雨晗 赵 丹
(中央财经大学管理科学与工程学院 北京 100081)

**摘 要** 近年来属性图聚类受到了广泛关注,其目的是将属性图中的节点划分到若干簇中,使得每一个集群都有紧密的簇内结构和均匀的属性值。现有的理论主要是假设属性图中的节点或对象是为了协助优化某个给定的方程,而忽略了它们在现实生活中本身的属性。同时,一些开放性问题尚未得到有效解决,如异构信息集成、计算成本高等。为此,把属性图聚类问题理解为自身节点代理的集群形成博弈。为了有效地整合拓扑结构和属性信息,提出了基于紧密性和均匀性约束的节点代理策略选择。进一步证明了博弈过程将会收敛到弱帕累托纳什均衡。在实证方面,设计了一个分布式和异构的多智能体系统,给出了一个快速的分布式学习算法。该算法的主要特点是结果分区的重叠率可以由一个事先给定的阈值控制。最后,在现实社交网络上进行了模拟实验,并与目前先进方法进行比较,结果证实了所提算法的有效性。

**关键词** 属性图聚类,集群形成博弈,紧密性和均匀性约束,分布式学习算法,多智能体系统
**中图法分类号** TP393 **文献标识码** A

## Distributed and Heterogeneous Multi-agent System for Attributed Graph Clustering

BIAN Zhai-an LI Hui-jia CHEN Jun-hua MA Yu-han ZHAO Dan
(School of Management Science and Engineering, Central University of Finance and Economics, Beijing 100081, China)

**Abstract** Recent years have witnessed a renewed attention towards attributed graph clustering, which aims to divide the nodes in the attribute graph into several clusters, so that each cluster has a densely connected intra-cluster structure and homogeneous attribute values. Existing methods ignore nodes/objects selfish nature in real-life contexts. Meanwhile, some open problems, such as heterogeneous information integration, high computational cost, etc., have not been effectively resolved yet. To this end, we considered the attribute graph clustering problem as the cluster formation game of selfish node-agents. To effectively integrate both topological and attributive information, we proposed both tightness and homogeneity constraints on node-agents' strategy selection. To be specific, the game process will converge to weakly Pareto-Nash equilibrium almost surely. In the aspect of implement, we carefully designed a distributed and heterogeneous multiagent system, based on which, a fast distributed learning algorithm is also given. The main feature of the proposed algorithm is that the overlap rate of the resulted partition can be well controlled by a pre-specified threshold. Finally, we conducted a set of simulation experiments on real-life social networks and comparisons are listed.

**Keywords** Attributed graph clustering, Cluster formation game, Tightness and homogeneity constraints, Distributed learning algorithm, Multiagent system

## 1 引言

许多现实中的信息系统是由大量高度关联的参与者或对象组成的,如在线社会网络、无线传感器网络和众包平台。这些系统可以被属性图很好地模拟出来,其中节点代表组件对象,属性向量用以描述对象的特征,边缘表示它们的联系。最有趣并具有挑战性的研究课题之一,就是把分割成若干个属性图的同一集群的内部结构和同类的属性值紧密联系起来,这项工作被称为属性图聚类(AGC)[1]。AGC 目前已经具有广泛的应用,如检测在线社交网络[2]社团、降低无线传感器网络能耗[3]、优化众包平台的任务分配[4]等。

在过去的几十年中,人们已经对 AGC 做过大量的研究,但是仍然存在一些问题,这需要我们投入更多的精力。1)已有的大多数 AGC 研究,都假设现实生活中信息系统的合作优化对象是一个给定的目标函数,却忽略了它们"自私"的性质。比如,在众包平台上,人们总是理性的,他们加入团队的唯一动机是实现自身利益的最大化。2)拓扑和属性信息是两个看似独立的异构数据,很难有效集成。文献[7,9,19]中虽然已经提出了一些线性组合方法,但是很难调整权重参数以及解释综合距离函数。此外,考虑所有节点的属性可能导致维数灾难[17]问题,这也是多维关联聚类领域的一个难点。3)大多数AGC 方法划分的节点是没有交集的,然而现实中一

本文受国家自然科学基金项目(71401194,71401188),中央财经大学"青年英才"培育支持项目(QYP1603)资助。
边宅安(1995-),男,硕士生,主要研究方向为复杂网络;李慧嘉(1985-),男,博士,副教授,主要研究方向为社会网络、数据挖掘、运筹学,E-mail:Hjli@amss.ac.cn;陈俊华(1975-),男,副教授,主要研究方向为复杂网络;马雨晗(1994-),女,硕士生,主要研究方向为复杂网络;赵 丹(1995-),女,硕士生,主要研究方向为复杂网络。

些应用程序中的对象往往可能同时属于多个集群。一个好的AGC框架应自然结合重叠的簇[20]。4)现有的AGC算法的计算是非常复杂的,特别是当考虑聚类过程中所有节点的属性时,如何协调聚类的有效性和时间复杂度仍是一个悬而未决的问题,我们应做进一步的研究。

## 2 问题的描述和预备知识

设 $G=(V,E,X)$ 是一个属性图,$V:=\{v_1,\cdots,v_n\}$ 是节点集,$E:=\{e_{ij}\}$ 是边集,$X$ 是属性矩阵的行的集合,即 $X_i$ 可以被看作是节点的属性向量。本文将专注于二进制属性。若列属性存在于行的节点向量,那么 $x_{ij}=1$,否则 $x_{ij}=0$。每个节点通常与一组相邻连接节点连接,节点的相邻节点集被称为 $N_i$。显然,$|N_i|+K_i$ 在图中没有自我循环(其中 $K_i$ 是节点的度),AGC问题被归结为从属性图 $G$ 中寻找良好集群($P=\{C_1,C_2,\cdots,C_n\}$)的问题。属性图的基本框架应该产生具有紧簇内结构和齐次节点属性的集群。为了更好地解释这两个特性,应仔细研究具体的 $C_k$ 聚类并引入以下概念。

定义1(紧密性) 集群 $C_k$ 的紧密性记为 $T_k$,表示如下:

$$T_k=\frac{2L_k^{in}}{n_k^2}-\frac{L_k^{out}}{n_k(n-n_k)} \tag{1}$$

其中,$n_k$ 是 $C_k$ 聚类中包含的节点个数,$L_k^{in}$ 是 $C_k$ 簇内边的数量,$L_k^{out}$ 是 $C_k$ 簇间边的数量。由于 $n_k$ 可以被忽略不计,因此做如下定义。

定义2(调整紧密性) $C_k$ 聚类的调整紧密性定义为:

$$\overset{\wedge}{T}_k=n_k(n-n_k)T_k=\frac{2(n-n_k)}{n_k}L_k^{in}-L_k^{out} \tag{2}$$

其中,因子 $n_k(n-n_k)$ 对于很大或者很小的集群会有平衡解。

定义3(同质化) 基于Havrda-Charvat广义熵离散概率集群 $C_k$ 分布的均匀性 $H_k$ 定义为:

$$H_k=-\frac{1}{p}\sum_{j=1}^{p}C_{kj}(1-C_{kj}) \tag{3}$$

其中,$C_k=(C_{k1},\cdots,C_{kp})$ 表示 $C_k$ 聚类中的属性向量,元素 $C_{kj}$ 表示第 $j$ 个元素的属性值为1的概率。$H_k$ 越大,表示集群 $C_k$ 中的节点具有的属性值越类似。

## 3 集群形成博弈

本节将属性图聚类理解为节点代理的集群形成博弈(CFG)。每一个节点只能加入唯一的集群,并且它的收益取决于所有节点所采取的策略集合。CFG可以重复多个阶段。每个节点都允许同步更新它们的策略,它们可以通过之前的阶段来收集有效信息。此外,假设每个节点代理既可以单独提高其在一定约束机制下的效用,也可以以一定的概率维持现状,即使它有更好的选择。

### 3.1 博弈的形成

给定一个属性图 $G=(V,E,X)$,它包含 $n$ 个节点和 $m$ 条边。$V$ 中的每个节点都被称为节点代理,由一个 $p$ 维的二进制属性向量 $X_i$ 表示。要定义一个CFG,必须指定节点代理的策略集合 $S_i$ 和效用函数 $u_i(\cdot):S\to R$。据此,$v_i$ 的策略 $s_i$ 就可以被一个列向量 $(\rho_{1i},\rho_{2i},\cdots,\rho_{Ki})^{\mathrm{T}}$ 表示,其中 $\rho_{Ki}$ 表示节点加入集群 $C_k$ 的概率。

假设1(唯一的集群成员身份) 每个节点代理在CFG中只能加入唯一的集群。

通过分析,$s_i$ 中只有一个为1,其余 $K-1$ 个都为0。即 $\rho_{Ki}=1$,且对于任意的 $\tilde{k}\neq k$,$\rho_{\tilde{k}i}=0$。所以,$v_i$ 的策略集合将为 $S_i=\{e_1,e_2,\cdots,e_K\}$,其中 $e_K$ 是 $K$ 维的单位向量,它的第 $K$ 个分量是1,其余分量为0。

定义效用函数的直觉来自于从众心理,从众是集体现象[5]。许多复杂因子(如组大小、一致性、凝聚力、地位、事先承诺和公众舆论等)会影响个体显示出来的从众程度。为了简化这个问题,对节点代理的从众行为做出如下假设。

假设2(局部从众) 个人的行为主要受邻居和朋友的影响。不难发现,在现实中,一个人从朋友那里听到了不同的意见,如果其中一个意见被他的朋友们广泛接受,并且接受的比率远远超过预期,那么他也会很容易接受这个观点。因此,对于一个策略组合 $s\in\prod_{i=1}^{n}S_i$,令 $\delta_k(s)$ 表示集群 $C_k$ 占领程度的百分比,即 $\frac{\sum_{s_j=e_k}k_j}{2m}$,那么加入集群 $C_k$ 的 $v_i$ 中的相邻节点的实际个数 $\bar{\tau}_{ki}$ 可以被确定为 $k_i\delta_k(s)$,加入集群 $C_k$ 的 $v_i$ 中的相邻节点的实际个数为 $\tau_{ki}(s)=|\{j\in N_i\mid s_j=e_k\}|$。在集群形成博弈中,节点代理 $v_i$ 的策略 $s_i$ 的效用函数只取决于两者差值的加权平均。于是,定义 $v_i$ 的效用为:

$$u_i(s)=\sum_{k=1}^{K}\rho_{ki}(\tau_{ki}(s)-k_i\delta_k(s))=\sum_j(A_{ij}-\frac{k_ik_j}{2m})s_i^{\mathrm{T}} \tag{4}$$

其中,$[A_{ij}]$ 是 $G$ 的邻接矩阵,如果节点 $v_i$ 和 $v_j$ 之间有边,$A_{ij}=1$,否则为0。基于假设2,可进一步得到节点代理的可行策略空间的定义。

定义4(可行策略空间) 给定一个除 $v_i$ 外的包含所有节点代理的策略集 $s_{-i}$,节点代理 $v_i$ 的可行策略集为:

$$F_i(s_{-i}):=\{e_k\mid\forall j\in N_i,s_j=e_k\} \tag{5}$$

引理1 给定一个除 $v_i$ 外的包含所有节点代理的策略集 $s_{-i}$,最优回复策略属于它的可行策略空间,即 $S_i^*(s_{-i})\in F_i(s_{-i})\in S_i$。

为使效用最大,每个节点代理只能从它们状态依赖的策略空间 $F_i(s_{-i})$ 中选择策略,并且 $F_i(s_{-i})$ 只由 $v_i$ 相邻点的策略决定,即 $v_i$ 只能加入至少有一个与其相邻节点的集群。

引理2 属性图聚类是一个潜在博弈,并且存在一个势函数:

$$\emptyset(s)=\frac{1}{2}\sum_i u_i(s) \tag{6}$$

可以观察到,式(6)中定义的势函数与评价指标模块化[8]存在线性关系。给定一个策略集 $s\in\prod_{i=1}^{n}S_i$,与该策略集相关的模块化定义为:

$$Q_{os}=\frac{1}{2m}\sum_i\sum_{j\neq i}(A_{ij}-\frac{k_ik_j}{2m})s_i^{\mathrm{T}}s_j=\sum_{k=1}^{K}\frac{L_k^{in}}{m}-(\frac{D_k}{2m})^2 \tag{7}$$

其中,$L_k^{in}$ 表示集群 $C_k$ 中的节点个数,$D_k$ 是集群 $C_k$ 中节点的度。将势函数乘以常数 $1/m$,得到:

$$\frac{1}{2m}\sum_i u_i(s)=\frac{1}{2m}\sum_i\sum_{j\neq i}(A_{ij}-\frac{k_ik_j}{2m})s_i^{\mathrm{T}}s_j-\frac{\sum_i k_i^2}{4m^2}=Q_s-\frac{\sum_i k_i^2}{4m^2} \tag{8}$$

### 3.2 集群形成博弈中的紧密性和均匀性约束

模型优化已经被证明存在“分辨率限制”问题[6],我们不能辨别出小型集群。文献[6]通过合并两个直观集群获得收益 $\Delta Q$,如果 $\Delta Q>0$,则这两个直观集群的模块将不能被优化,因为合并它们会增加模块化。也就是说,通过合并 $C_i$ 和

$C_j$ 得到收益 $\Delta Q=\frac{L_{ij}^{inter}}{m}-2\frac{D_i}{2m}\frac{D_j}{2m}$，其中 $L_{ij}^{inter}$ 为两个集群公共边的个数。对于任意的 $L_{ij}^{inter}>\frac{D_iD_j}{2m}$，$\Delta Q$ 都是正数。不等式右侧可以被理解为 $C_i$ 和 $C_j$ 之间的期望边数。现实中 $m$ 是无穷大的，因此合并两个集群是有问题的，即使两个集群的直观公共边只有一条。

拟合的势函数与模块呈线性关系（见式(8)），若采用最优回复策略从而达到一个 PNE，那么 CFG 也会受到“分辨率限制”，缓解这个问题的一种方法是增加额外信息（如节点属性或结构相似之处）的原始图的权重。然而这不是一个可靠的解决方案，原因有两个：1）在现实中还没有好的方法来解决小的簇间权重；2）计算边缘权重将会额外增加计算量，这个问题在大规模分布图中尤为突出。属性图中一个好的集群划分，对于任意 $C_k\in P$，$\hat{T}_k$ 和 $H_k$ 应该尽可能大。为实现这两个目标，对于节点代理的策略选择，我们尝试把紧密性和均匀性约束加入到 CFG。这两种约束可以很好地聚焦于“分辨率限制”。在这里，考虑一般场景，如图 1(a)所示。给定一套可行的策略集 $s$，策略集合 $F_i(s_{-i})$ 对 $v_i$ 是有效的，它选择的策略或集群应满足以下 3 个标准：

标准 1　$e_k$ 应该是节点代理 $v_i$ 的更优回复策略之一，$e_k\in\widetilde{S}_i(s)$；

标准 2　一旦 $v_i$ 与 $C_k$ 结合，新 $C_k$ 的调整紧密性不应该减少，即 $\hat{T}_{k\cup i}\geqslant\hat{T}_k$；

标准 3　如果最终 $v_i$ 与 $C_k$ 结合，新集群均匀性下降，$H_{k\cup i}$ 应该不重要，即 $H_k-H_{k\cup i}\leqslant\varepsilon$，$\varepsilon$ 是一个充分小的值。


(a)　　(b)

图 1　节点代理周围的候选集群

为了满足标准 1，每个节点代理 $v_i$ 可以交替采用更优回复策略来提高效用，这样 CFG 可以最终达到 PNE。这里，我们把注意力集中在如何满足标准 2 和标准 3 上。

根据图 1(b)，给定一个策略集 $s$，假设一个集群 $C_k(s)$，在 $v_i$ 附近。$\beta_k(s)$ 和 $C_k(s)$ 的边界区域紧密联系，$\beta_k(s):=\{v_j\mid s_j\neq e_k\wedge s_j=e_k\wedge A_{ij}=1\}$，接下来可以定义 3 种边界类型：在集群内部（表示为 $L_k^{in}(s)$）；在 $C_k(s)$ 和 $v_i$ 之间（表示为 $\tau_{ki}(s)$）；在 $C_k(s)$ 和其他 $\beta_k(s)$ 中的节点之间（表示为 $L_{k\tilde{i}}^{out}(s)$）。记 $L_k^{out}(s)=\tau_{ki}(s)+L_{k\tilde{i}}^{out}(s)$。为了简化计算过程，依据 $L_k^{in}(s)$ 和 $k_i$（$v_i$ 的程度）来表示外部链接的数量：$\tau_{ki}(s)=aL_k^{in}(s)=bk_i$；$L_{k\tilde{i}}^{out}(s)=cL_k^{in}(s)$，满足 $a>\frac{1}{L_k^{in}(s)}$，$b\geqslant\frac{1}{k_i}$（因为 $\beta_k(s)$ 中的任意 $v_i$ 在 $C_k(s)$ 中至少有一个相邻点），$c\geqslant0$。根据上述内容，集群的调整紧密性可以写为：

$$\hat{T}_k(s)=\frac{n-n_k(s)}{n_k(s)}2L_k^{in}(s)-(a+c)\ L_k^{in}(s)\tag{9}$$

其中，$n_k(s)$ 为 $C_k(s)$ 的大小。一旦 $v_i$ 加入 $C_k(s)$，调整紧密性便变为：

$$\hat{T}_{k\cup i}(s)=\frac{n-n_k(s)-1}{n_k(s)+1}2L_k^{in}(s)(1+a)-(cL_k^{in}(s)+k_i-\tau_{ki}(s))\tag{10}$$

定义 5（调整紧密性增量，ATI）　给定一个 CFG 的策略集 $s$，选择策略 $s_i(s_i=e_k)$ 对节点代理 $v_i$ 的集群 $C_k\mid s_i=e_k$ 的调整紧密性增量定义为：

$$T_i^{s_i}=\hat{T}_{k\cup i}(s)-\hat{T}_k(s)=2n\frac{\tau_{ki}(s)n_k-L_k^{in}(s)}{n_k(s)(n_k(s)+1)}-k_i$$
$$\text{s. t. } s_i=e_k\tag{11}$$

因此，为了满足标准 2，当且仅当 $T_i^{e_k}\geqslant0$ 时 $v_i$ 可以加入 $C_k\mid s_i=e_k$。接下来，用类似的办法来分析标准 3。一旦节点代理 $v_i$ 加入集群 $C_k(s)$，均匀性的增量可被定义为：

$$\Delta H_{k\cup i}(s)=H_{k\cup i}(s)-H_k(s)=$$
$$\frac{1}{p}\sum_{j=1}^{p}\frac{n_k(s)(2c_{kj}(s)-1)(x_{ij}-c_{kj}(s))+c_{kj}(s)(1-c_{kj}(s))}{(n_k(s)+1)^2}\tag{12}$$

如果节点代理 $v_i$ 通过判断 $\Delta H_{k\cup i}(s)$ 是否大于 0 来决定它是否可加入 $C_k(s)$，就存在冷启动问题。考虑一种情况：每个节点代理都存在于单独的集群，这种情况下联合策略都被固定在 $s^0=\{e_1,\cdots,e_n\}$，任意集群的矢量中心 $c_k(s^0)(k\neq i)$ 正好是 $v_k$ 的属性向量，即 $c_k(s^0)=X_k$。则 $\Delta H_{k\cup i}(s^0)$ 可被表示为：

$$\Delta H_{k\cup i}(s^0)=\frac{1}{4p}\sum_{j=1}^{p}(2x_{kj}-1)(x_{ij}-x_{kj})\tag{13}$$

根据式(13)，可以发现如果 $v_i$ 加入任意一个集群 $c_k(s^0)$，$k\neq i$，那么 $\Delta H_{k\cup i}(s^0)\leqslant0$。当且仅当 $v_i$ 正好和 $v_k$ 或 $c_k(s^0)$ 有相同的属性向量时，两者相等。

我们认为 $\Delta H_{k\cup i}(s^0)\geqslant0$ 太严格，避免冷启动的一个可行方法是放宽这个条件。式(12)可以被进一步写成：

$$\Delta H_{k\cup i}(s)=\frac{1}{p}\sum_{j=1}^{p}\frac{n_k(s)x_{ij}(2c_{kj}(s)-1)}{(n_k(s)+1)^2}=$$
$$\frac{1}{p}\sum_{j=1}^{p}\frac{n_k(s)c_{kj}(s)(1-2c_{kj}(s))+c_{kj}(s)(1-c_{kj}(s))}{(n_k(s)+1)^2}\tag{14}$$

定义 6（调整均匀性增量，AHI）　根据式(14)中定义的均匀性增益，将最大化 $\Delta H_{k\cup i}(s)$ 简化为最大化假设 1。给定一个策略集 $s$，由于我们只关心假设 1 的极性，集群 $C_k\mid s_i=e_k$ 对于选择策略 $i(s_i=e_k)$ 的节点代理 $v_i$ 的调整均匀性增量被定义为：

$$H_i^{s_i}=\sum_{j=1}^{p}x_{ij}(2c_{kj}(s)-1),\text{ s. t. } s_i=e_k\tag{15}$$

定义 7（弱帕累托纳什均衡，WPNE）　在 CFG 中，策略集 $s^*=\{s_1^*,\cdots,s_n^*\}\in S$ 若构成 WPNE，则需满足对于任意一个 $v_i\in V$，$s_i^*\in s^*$，并且

$$u_i(s_i^*,s_{-i}^*)>u_i(s_i',s_{-i}^*)$$
$$\text{s. t. }\forall\, s_i'\neq s_i^*,T_i^{s_i'}\geqslant0,H_i^{s_i'}\geqslant0\tag{16}$$

定义 8（更优约束回复策略，CBRD）　给定一个策略集 $s$，节点代理 $v_i$ 的更优约束回复策略被定义为：

$$\Psi_i(s)=\{s_i'\mid s_i'\in\widetilde{S}_i(s),T_i^{s_i'}\geqslant0,H_i^{s_i'}\geqslant0\}\tag{17}$$

显然，任何 $\Psi_i(s)$ 中现有的策略都满足以上 3 个标准。因此，更优约束回复机制在 CFG 中被定义为策略更新规则，当且仅当 $\Psi_i(s)\neq\emptyset$ 时，每个节点代理 $v_i$ 在每个离散点被随机选中，它可以通过更优约束回复机制及时改变其当前的策略。

引理 3　在 CFG 中，CBRD 总是在有限时间内收敛到 WPNE。

证明：因为 CBRD 是特殊的更优回复策略，同时后者的

收敛总是满足潜在博弈，所以这里略去证明。

### 3.3 约束下的 JSFP 与惯性

为了在潜在博弈模型中达到帕累托纳什均衡(PNE)，文献中已经提及大量的集中性规则(如最优回复策略和自适应播放算法)。虽然这些规则能保证收敛性，但它们要求每个玩家参照其他玩家的策略，同时所有玩家都被严格限制实时更新其策略。在一个大型博弈中，这两个要求可能难以实现。

本文考虑 CFG 的重复或分段，也就是说，在每一阶段 $t=\{1,2,\cdots\}$，每个节点 $v_i\in V$ 视其他所有节点为自己的对手，并根据对手随机联合策略(但根据一定的联合经验频率)的假设选择自己的策略。从联合策略虚拟博弈(JSFP)入手，令 $\check{p}_{-i}^{\check{s}_{-i}}(t)$ 表示在这一阶段中 $t-1$ 时刻节点 $v_i$ 的对手们选择联合策略 $\check{s}_{-i}\in S_{-i}$ 的百分比：

$$\check{p}_{-i}^{\check{s}_{-i}}(t)=\frac{1}{t-1}\sum_{\tau=1}^{t-1}I(s_{-i}(\tau)=\check{s}_{-i}) \tag{18}$$

其中，$s_{-i}(\tau)\in S_{-i}$ 是时间 $\tau$ 的联合策略选择情况，$I(\cdot)$ 是指示函数。在第 $t$ 阶段，节点 $v_i$ 选择的策略 $s_i\in S_i$ 的期望效用 $\bar{u}_i(s_i,p_{-i}(t))$ 可以表示为：

$$\bar{u}_i(s_i,p_{-i}(t))=\sum_{s_{-i}\in S_{-i}}u_i(s_i,s_{-i})p_{-i}^{s_{-i}}(t) \tag{19}$$

把式(18)代入(19)，可以得到：

$$\bar{u}_i(s_i,p_{-i}(t))=\frac{1}{t-1}\sum_{\tau=1}^{t-1}u_i(s_i,s_{-i}(\tau)) \tag{20}$$

把 $\bar{u}_i(s_i,p_{-i}(t))$ 缩写为 $\bar{u}_i^{s_i}(t)$，以下递归总是成立：

$$\bar{u}_i^{s_i}(t+1)=\frac{1}{t-1}u_i(s_i,s_{-i}(\tau))+\frac{t}{t+1}\bar{u}_i^{s_i}(t) \tag{21}$$

CFG 的每个节点都可以用以上递归公式在更新一个可行策略后预测自己的效用。注意，在 JSFP 和式(20)里定义的预测效用是节点 $v_i$ 可接受的平均效用，然而在现实中，每个节点可能更注重观察最近的信息。则节点 $v_i$ 选择策略 $s_i\in S_i$ 的记忆衰退期望效用定义为：

$$\begin{aligned}&\bar{u}_i^{s_i}(1)=u_i(s_i,s_{-i}(1))\\&\bar{u}_i^{s_i}(t+1)=\alpha u_i(s_i,s_{-i}(t))+(1-\alpha)\bar{u}_i^{s_i}(t),t\geqslant 1\end{aligned} \tag{22}$$

其中，$0<\alpha\leqslant 1$ 是用来测度当前策略方案 $s(t)$ 重要性的加权参数。当 $\alpha=1$ 时，只有近期最重要的信息影响决策；当 $\alpha=\frac{1}{t+1}$ 时，所有过去的信息在 JSFP 中同等重要。对节点更新策略的意愿做进一步假设。

假设 3(惯性反应) 在 CFG 中，节点不愿选择一个更优的策略。令 $\beta_i(t)$ 为节点 $v_i$ 在 $t$ 时刻愿意更新其目前策略的可能性，对于任意 $v_i\in V$，有常量 $Q$：

$$0<\underline{\varepsilon}<\beta_i(t)<\bar{\varepsilon}<1 \tag{23}$$

假设 3 意味着在 CFG 中的每个节点 $v_i$ 会以 $1-\beta_i(t)$ 的可能性保持先前的策略，既使它有更好的选择来提高自己的效用。这个假设与现实情况一致，即个人决策总是有滞后性。回顾假设 2，每个节点的策略选择主要被相邻节点影响，假设在 $t$ 时刻大多数节点改变策略，那 $v_i$ 在 $t+1$ 时刻的期望效用会有很大波动，因此它会更愿意思考是否要改变现有策略。另一方面，如果在 $t$ 时刻只有很少的相邻节点改变策略，那么它会更愿意选择维持之前的策略。基于以上分析，本文对 $\beta_i(t)$ 进行如下定义：

$$\begin{aligned}&\beta_i(1)=1\\&\beta_i(t)=\frac{|\{j\mid j\in N_i,s_j(t)\neq s_j(t-1)\}|}{k_i},t\geqslant 2\end{aligned} \tag{24}$$

换句话说，$t\geqslant 2$ 时，$\beta_i(t)$ 定义了 $v_i$ 的相邻节点在 $t$ 时刻改变策略的比例。

定义 9(具有惯性约束的 JSFP，CJSFPR) 根据假设 3，当且仅当 $\bar{\Psi}_i(t+1)=\emptyset$，在概率 $\beta_i(t)$ 下，CFG 中的每个节点愿意改变现有策略，在 $t(t\geqslant 2)$ 时刻从策略集里选择一个策略 $s_i(t)$：

$$\begin{aligned}&\bar{\Psi}_i(t+1)=\{s_i'\mid s_i'\in F_i(t),\bar{u}_i^{s_i'}(t+1)\geqslant\bar{u}_i^{s_i}(t+1)\},\\&T_i^{s_i'}(t)\geqslant 0,H_i^{s_i'}(t)\geqslant 0\end{aligned} \tag{25}$$

注意，$F_i(t)$ 是 $v_i$ 在 $t$ 时刻的可行策略空间，$T_i^{s_i'}(t)$ 和 $H_i^{s_i'}(t)$ 分别是集群 $C_k\mid s_i'=e_k$ 调整紧密性和均匀性的增量。这个过程叫做惯性约束下的 JSFP。

引理 4 在重复的 CFG 中，如果在 $t\geqslant 1$ 的任何时刻 CJSFPR 产生的策略 $s(t)$ 都满足 WPNE，那么对于所有 $\tau>0$，有 $s(t)=s(t+\tau)$。

证明：由于 $s(t)$ 是 WPNE，对于每一个 $s_i'\in F_i(t)/s_i(t)$，如果 $T_i^{s_i'}(t)\geqslant 0$ 并且 $H_i^{s_i'}(t)\geqslant 0$，则有：

$$u_i(s_i(t),s_{-i}(t))\geqslant u_i(s_i',s_{-i}(t))$$

$$\bar{u}_i(s_i(t),p_{-i}(t))\geqslant\bar{u}_i(s_i',p_{-i}(t))$$

我们根据 $p_{-i}(t)$ 和 $s_{-i}(t)$ 写出 $p_{-i}(t+1)$，得到：

$$\bar{u}_i(s_i,p_{-i}(t+1))=\alpha u_i(s_i',s_{-i}(t))+(1-\alpha)\bar{u}_i(s_i',p_{-i}(t))$$

$$\bar{u}_i(s_i(t),p_{-i}(t+1))=\alpha u_i(s(t))+(1-\alpha)\bar{u}_i(s_i(t),p_{-i}(t))$$

因此，有 $\bar{u}_i(s_i(t),p_{-i}(t+1))>\bar{u}_i(s_i,p_{-i}(t+1))$。另一方面，如果 $s_i''\in F_i(t)/s_i(t)$，并且 $T_i^{s_i'}(t)\geqslant 0$ 或者 $H_i^{s_i'}(t)\geqslant 0$，那么在 $t+1$ 时刻它将不会被 $v_i$ 选择，即使 $\bar{u}_i(s_i(t),p_{-i}(t))<\bar{u}_i(s_i'',p_{-i}(t))$。因此，$s(t)$ 是对 $p_{-i}(t+1)$ 最好的回复并且是稳定的。

引理 5 在重复的 CFG 中，由 CJSFPR 生成的策略集 $s(t)$ 几乎全部收敛于 WPNE。

### 3.4 重叠分区细化

在现实中，根据各自的专业、收入、经历等，人们往往属于不同的社交圈子。虽然 CFG 假定每个节点只可以加入唯一的集群(见假设 1)，但我们强调它可以自然地吸收重叠集群。

定义 10(重叠概率) 给定一个策略组合 $s(t)$，节点 $v_i$ 的重叠概率定义为：

$$\gamma_i(t)=-\sum_{e_k\in F_i(t)}\frac{\tau_{ki}(t)}{k_i}\log_{|F_i(t)|}\frac{\tau_{ki}(t)}{k_i} \tag{26}$$

其中，$|F_i(t)|$ 指 $v_i$ 在 $t$ 时刻的可选策略空间的大小。注意，$\gamma_i(t)$ 的值在[0,1]之间，如果 $v_i$ 是重叠节点，可能会有更大的取值。因此，将重叠集群定义如下：

定义 11(重叠集群标签集) 如果 $v_i$ 的重叠概率大于某个阈值 $\gamma$，则在 $t$ 时刻与它关联的集群标签集可定义为：

$$\begin{aligned}&\Lambda_i(t)=\{k\mid e_k\in F_i(t),u_i(e_k,s_{-i}(t))\geqslant 0\},\\&T_i^{e_k}(t)\geqslant 0,H_i^{e_k}(t)\geqslant 0\end{aligned} \tag{27}$$

定义 11 表明，如果：1)$C_k$ 位于 $v_i$ 的邻域，2)在 $t$ 时刻加入 $C_k$ 的效用不为负，3)在 $t$ 时刻 $v_i$ 加入 $C_k$ 的 ATI 不为负，4)在 $t$ 时刻合并 $v_i$ 时 AHI 不为负，那么候选簇 $C_k$ 对 $v_i$ 有潜在诱惑力。在实际应用中，可以指定参数 $\gamma$ 来控制重叠率导致的集群分区，即，当 $\gamma$ 被设置为 1 时，重叠集群的每个节点代理的标签的集合为空，因此最终分区非重叠；相反，如果 $\gamma$

被指定为 0，所有节点代理都将与非空集合的重叠集群标签相关联，使得最终聚类分区高度重叠。

## 4　分布式和异构的多智能体系统

大多数现有的 AGC 方法旨在发展集中式的算法来确定属性图中的集群。这些算法通常需要一个权威来保证对图有一个完全的了解，并尝试全局优化一个特定的目标函数。然而，在许多现代的应用中，这两个要求可能不适用或不可取。因此，本文开发一个分布式和异构多智能体系统（简称 DH-MAS）来解决 AGC 问题。

### 4.1　DH-MAS 的环境

拟议的 DH-MAS 被定义为：$\Xi=\{V,P,CA,t\}$，其中 $V:=\{v_1,\cdots,v_n\}$ 是一组节点集，$P:=\{C_1,\cdots,C_K\}$ 是集群代理集，CA 是用来管理整个系统的中央代理，$t$ 是系统时间。$\Xi$ 中的每个节点 $v_i$ 可以被一个十二元组 $\langle X_i,F_i(t),s_i(t),\beta_i(t),\overline{\Psi}_i(t),\gamma_i(t),\Lambda_i(t),\vec{\tau}_i(t),\vec{u}_i(t),\vec{u}_i(t),\vec{T}_i(t),\vec{H}_i(t)\rangle$ 表示。同样地，每个集群代理 $C_k$ 被一个五元组 $\langle C_k(t),n_k(t),\delta_k(t),L_k^{in}(t),c_k(t)\rangle$ 表示。以上元素代表的含义均已在前文提到。不同于现有的每个节点可以与其他节点进行交互的图聚类或群体检测方法，在我们的方法中，每个节点代理和集群代理视域都限于局部的范围，我们认为这样更接近现实情况。例如，节点 $v_i$ 在 $t$ 时刻只能加入与它邻近的局部范围集群 $\{C_k\mid e_k\in F_i(t)\}$，而不能加入自己已经存在的集群 $\{v_i\mid v_i\in C_k(t)\}$。DH-MAS 分配给中央代理的任务很简单，它只需要更新系统时间和监控所有节点代理，从而进一步更新 $C_k(t)$ 的每个集群代理的策略。

### 4.2　分布式学习算法

为了实际运用图聚类方法，根据 DH-MAS 工作框架提出了分布式的学习算法。给定属性图 $G:=(V,E,X)$、重要性权重参数 $\alpha$ 和概率阈值 $\gamma$，详细程序在分布式学习算法中进一步描述。

**算法 1**　DH-MAS 初始化

Input：$\mathcal{G}:=(V,E,X)$

1. $t\leftarrow 1$；
2. $K\leftarrow n$；
3. Paralleled Execution $\forall C_k\in\mathbb{P}$；
4. $C_k(t)=v_k$；$n_k(t)=1$；$\xi_k(t)=\frac{k_k}{2m}$；$L_k^{in}(t)=0$；$c_k(t)=X_k$；
5. Paralleled Execution $\forall v_i\in\mathbb{V}$；
6. $X_i=\mathcal{G}.X_i$；$F_i(t)=\{e_k\mid v_k\in N_i\}$；$s_i(t)=e_i$；$\beta_i(t)=1$；$\overline{\Psi}_i(t)=\varnothing$；$\gamma_i(t)=0$；$\Lambda_i(t)=\varnothing$；
7. For $\forall e_k\in F_i(t)$
8. $T_{ki}(t)=1$；$u_i^{e_k}(t)=-k_i\xi_k(t)$；$\bar{u}_i^{e_k}(t)=-k_i\xi_k(t)$；$T_i^{e_k}(t)=n-k_i$；$H_i^{e_k}(t)=\sum_{j=1}^{p}x_{ij}(2x_{kj}-1)$；

在分布式学习算法的步骤 1 中，首先应该初始化 DH-MAS。CFG-CJSFPR 使用位于它自己的集群中的每个节点代理的初始配置。初始化步骤的细节在算法 1 中给出。集群和节点同步初始化。CFG-CJSFPR 与这些初始配置可以被解释为无监督学习方法，其中所有代理活动都基于它们各自的决策过程。此外，如果某些部分的信息是可用的，例如一些节点确实共享一个集群，则可以将这部分的信息集成到初始配置，并且不允许 CJSFPR 改变集群中节点代理的标签，因此 CFG-CJSFPR 是一种半监督的学习方法。在分布式学习算法步骤 2－步骤 5 中，节点代理同时更新其期望效用和一组约束下的更优策略（式（22）、式（25）），并以一定概率转换到更优的策略。在步骤 6－步骤 7 中，中央代理 CA 基于当前的策略 $s(t)$ 更新系统时间 $t$ 和所有集群配置，即对于任意 $C_k\in p$，$C_k(t)=\{v_i\mid s_i(t)=e_k\}$。在步骤 8－步骤 9 中，除 $C_k(t)$ 外的每个集群代理的组成同步更新。同样，每个节点代理的组成在步骤 10－步骤 11 同步更新。以上步骤会一直重复，直到没有节点再改变策略。然后，在步骤 3－步骤 16 中，每个节点代理同步更新 $\gamma_i(t)$ 和 $\Lambda_i(t)$（式（26）、式（27））。最后，步骤 17－步骤 19 细化中央代理，在步骤 20 中输出重叠分区的结果。

### 4.3　计算复杂度

为了降低预计的计算复杂度，施加一个单一的线程环境，对每个节点和集群循环执行。给定一个有 $n$ 个节点和 $m$ 个边缘的属性图 $G$，令 $\langle p\rangle$ 为节点的属性向量中非零元素的平均数量，$\langle k\rangle=\frac{2m}{n}$ 为属性图的平均度。在算法 1 的步骤 1 中，初始化 $n$ 个集群代理和 $n$ 个节点代理将分别需要 $n\langle p\rangle$ 和 $n\langle p\rangle\langle k\rangle$ 的时间。因此，第一阶段的费用是 $O(m\langle p\rangle)$。初始化完成后，算法 1 中每个迭代的运行时间（即从步骤 2－步骤 11）由 3 个部分组成，分别来自节点代理、集群代理和中央代理的更新操作。假设在 $t$ 时刻，每个节点代理所消耗的时间主要来自更新 $\vec{H}_i(t)$，其费用为 $O(|F_i(t)|\langle p\rangle)$。对于每个集群代理 $C_k$，时间主要用来更新 $L_k^{in}(t)$ 和 $c_k(t)$，分别为 $n_k(t)\langle k\rangle^2$ 和 $n_k(t)\langle p\rangle$。由中央代理耗掉的时间主要花费在步骤 7，为 $O(m)$。将所有的时间相加，第 $t$ 次迭代的成本是 $O(\sum_i|F_i(t)|\langle p\rangle+n(\langle k\rangle^2+\langle p\rangle))$。令 $\delta$ 为均衡时的迭代次数，本文模拟中在 10～12 之间取值。因此，所有迭代的成本是 $O(\sum_{\tau=1}^{\delta}\sum_i|F_i(t)|\langle p\rangle+\delta n(\langle k\rangle^2+\langle p\rangle))$。注意，每个节点代理 $v_i$ 在时刻 $t$ 的可行策略空间的大小通常远小于 $k_i$，因此在最坏的情况下，上述时间成本也可以简化为 $O(\delta m\langle k\rangle+m\langle p\rangle)$，简化后，假定 $\delta$ 为常数，变为 $O(m\langle k\rangle+m\langle p\rangle)$。一旦 DH-MAS 达到均衡，每个节点代理将更新 $\gamma_i(t)$ 和 $\Lambda_i(t)$，其中最坏情况下的成本为 $O(m)$。之后，中央代理进行分区细化也需要 $O(m)$。因此，算法 1 单一的线程环境的最终时间成本之和为：$O(m\langle p\rangle+(m\langle k\rangle+m\langle p\rangle)+m)=O(m\langle k\rangle+m\langle p\rangle)$。

## 5　实验分析

图聚类算法的有效性可以通过 6 个指标来评估，其中 4 个是无监督的，其余 2 个是有监督的。4 个无监督指标分别是指覆盖率（Cover）、重叠率（OvLap）、归一化的紧密性（NorTi）和归一化的均匀性（NorHo）：前 3 个指标从拓扑结构的角度划分质量，最后一个指标从熵理论的角度划分质量。其他 2 个监督指标是 F1 分数的平均值（AvgF1）和 Ja 的平均值（AvgJa），并且这两个指标使用的是基准数据集的集群知识。给定一个集群分区 $P:=\{C_1,\cdots,C_K\}$ 和真实的聚类划分 $P^*:=\{C_1^*,\cdots,C_K^*\}$，这些指标定义为：

$$Cover=|\{v_i\mid\exists C_k\in P,v_i\in C_k\}|/n \tag{28}$$

$$OvLap=(\sum_k n_k)/|\{v_i\mid\exists C_k\in P,v_i\in C_k\}| \tag{29}$$

$$NorTi=\frac{1}{\sum_k n_k}\sum_k n_k\left(\frac{2L_k^{in}}{n_k^2}-\frac{L_k^{out}}{n_k(n-n_k)}\right) \tag{30}$$

$$NorHo=\frac{1}{p\sum_k n_k}\sum_k n_k\left(-\sum_{j=1}^{p} c_{kj}(1-c_{kj})\right) \tag{31}$$

$$AvgF1=\frac{1}{2K}\sum_{C_i\in P}\max_{C_j^*\in P^*}F1(C_i,C_j^*)+\frac{1}{2K^*}\sum_{C_i\in P}\max_{C_j^*\in P^*}F1(C_i,C_j^*) \tag{32}$$

$$AvgJa=\frac{1}{2K}\sum_{C_i\in P}\max_{C_j^*\in P^*}Ja(C_i,C_j^*)+\frac{1}{2K^*}\sum_{C_i\in P}\max_{C_j^*\in P^*}Ja(C_i,C_j^*) \tag{33}$$

其中，$F1(A,B)$和 $Ja(A,B)$分别表示 F1 分数和 Ja 在两组之间的相似性。期望更高的度量值意味着更好的图形分区。数据集 Facebook，Twitter，Google＋是 3 个属性图，每个节点被一个 0/1 向量描述。例如，节点属性在 Facebook 和 Google＋的特点是有用户的个人简介，如性别、年龄、工作、事业等。此外，选择了没有节点属性的四大节点图：Amazon，Delp，YouTube 和 LiveJournal。这些网络由斯坦福网络分析平台(SNAP)提供，并且局部集群是由 ego-networks 的持有者手动标记的。表 1 的＃COMS 列总结了数据集的一些统计特征，＊表示局部真实簇的数量，其他符号在之前已经提到过。

表 1 实验数据集

| Network | $n$ | $m$ | $\langle k\rangle$ | $p$ | $\langle p\rangle$ | ＃*Coms.* * | *Cover** | *OvLap** | *NorTi** | *NorHo** |
|---|---|---|---|---|---|---|---|---|---|---|
| Facebook | 4039 | 84243 | 41.7 | 175 | 4.2 | 146 | 0.708 | 1.459 | 0.308 | −0.0130 |
| Twitter | 76245 | 1242397 | 32.6 | 33208 | 2.6 | 3170 | 0.289 | 2.223 | 0.439 | −0.0004 |
| Google＋ | 102100 | 122113501 | 237.3 | 805 | 1.6 | 438 | 0.228 | 2.691 | 0.203 | −0.0068 |
| Amazon | 334863 | 925872 | 5.5 | — | — | 75149 | 0.571 | 7.128 | 0.00001 | — |
| Dblp | 317080 | 1049866 | 6.6 | — | — | 13447 | 0.626 | 3.205 | 0.00135 | — |
| YouTube | 1134890 | 2987624 | 5.3 | — | — | 8385 | 0.042 | 2.401 | 0.158 | — |
| LiveJournal | 3997962 | 34681189 | 17.3 | — | — | 287512 | 0.273 | 5.847 | 0.006 | — |

在一系列的实验中，通过从 3 个属性图(Facebook，Twitter，Google＋)检测重叠的簇来探讨 CFG-CJSFER 的性能。首先专注于参数分析，然后用不同的方法比较 CFG-CJSFER。

在 CFG-CJSFPR 中，有两个预先指定的参数：式(22)中的 $\alpha$ 表示最近的策略集的重要性权重；定义 11 中的概率阈值 $\gamma$ 控制最后分区的重叠率。

对于参数 $\alpha$，指定 $\gamma$ 为 0.5，在 CFG-CJSFPR 中说明参数 $\alpha$ 的影响，从 0.1 到 1 线性增加 $\alpha$。运行 CFG-CJSFPR 100 次，生成 100 个分区。图 2 显示了 $\alpha$ 影响的 4 个指标：NorTi，NorHo，AvgF1 和 AvgJa，得到：1)在所有观察到的属性图中，NorTi 和 NorHo 对 $\alpha$ 都不敏感；2)AvgF1 和 AvgJa 指标是呈正相关的，Twitter 和 Google＋中的增长可以增加 AvgF1 和 AvgJa，而 Facebook 却有着相反的趋势。对于要追求良好的分区的大属性图，接下来的实验中设置 $\alpha=1$，这表明只有最近的策略集才会影响 DH-MAS 的节点代理。

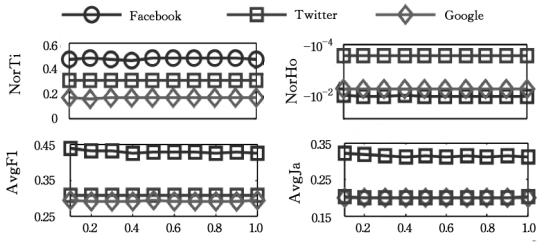


图 2 最近策略集的重要性权重 $\alpha$

对于参数 $\gamma$，指定 $\alpha$ 为 1，采用相同的方法分析参数 $\gamma$ 的影响。从 0 到 1 线性增加 $\gamma$(见图 3)，得到：1)在大多数情况下，最终的 Cover 和 OvLap 分区要高于局部真实的分区；而且，NorHo 价值和局部真实值是同一量级。在 Facebook 中，NorTi 要比局部真实值大得多，但是它们略低于 Twitter 和 Google＋的地面真实值。2)当 $\gamma$ 在[0，0.9]的区间时，前半部分设置的所有指标保持一致。进一步提高 $\gamma$，其他 5 个指标开始下降时 NorHo 持续性增长，当 $\gamma=0.5$ 时，这一阶段的转换发生在 Facebook 中；对于 Twitter and Google＋来说，各自的拐点分别在 0.8 和 0.7。3)当 $\gamma$ 取 1 时，分区结果是碎片化的，即对于所有数据集，OvLap=1。在这种情况下，NorHo 和 NorTi 达到最大值，但是 Cover，AvgF1 和 AvgJa 的值却是最低的。

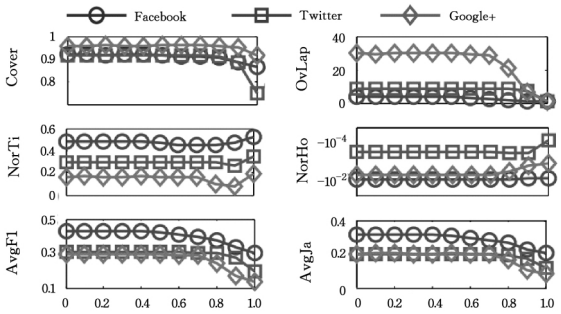


图 3 概率阈值 $\gamma$

因此，为了更好地匹配真实的分区并平衡集群间的紧密性和均匀性的效果，在以下实验中默认 CFG-CJSFPR 的参数设置为 $\alpha=1$，$\gamma=0.5$。

接下来分析 CFG-CJSFPR 的收敛性，并用不同的算法比较分区的质量。

(1)利用最优回复策略的集群形成博弈(CFG-BRD)：在每个迭代中，每个节点代理都是按顺序选择最好的应对策略。CFG-BRD 重复这个过程，直到没有节点可以增加它们的效用值。

(2)利用更优约束回复策略的集群形成博弈(CFG-CBRD)：CFG-CBRD 重复每个节点的代理过程，在每个迭代中，按顺序选择一个被约束的更优的应对策略，直到没有节点可以增加它们的效用值。

在这次实验中，CFG-BRD 和 CFG-CBRD 的参数都设置为 0.5。在每次迭代后，监控模块化的值、检测簇的数量和平均簇的大小。图 4 示出了 3 个属性图测试算法的收敛性质。如图 4 的底部显示，在几次迭代之后，3 种算法的收敛速度都加快，CFG-BRD 具有最高的模块化的值。

根据＃Coms，3 种算法的值随迭代过程减小(见图 4 的中间行)；然而 CFG-BRD 只能识别少部分的簇，比如，在 Twitter 中通过 CFG-BRD 检测簇的数量只有 1576，远远少于真实值 3170。但在 3 个数据集中，与其他两个算法相比，通过

CFG-BRD 的检测簇的平均数量是最大的，如图 4 的顶行所示。

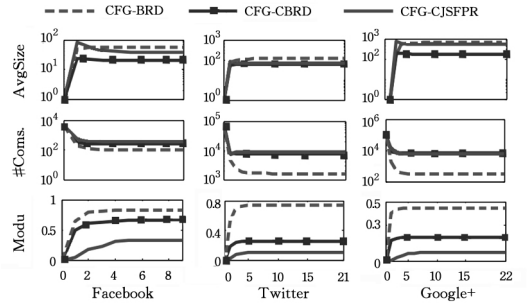


图 4　测试算法的收敛性

在 CFG-BRD 中，没有紧密性和均匀性的约束，每个节点所考虑的唯一问题就是如何最大化它们的效用。在此背景下，CFG-BRD 将进化成为模块化优化算法。因此，CFG-BRD 能在 3 个数据集中获得最高模块化的值。最后，比较表 2 中的 3 种算法的分区质量（表现好的用粗体表示），从中可以发现，CFG-CJSFPR 的整体效果是最好的，而 CFG-CBRD 是次优选择。

表 2　3 个属性图的内部比较

| Algorithms | NorTi | NorHo | AvgF1 | AvgJa |
|---|---|---|---|---|
| | | Facebook | | |
| CFG-BRD | 0.207 | −0.0127 | 0.386 | 0.286 |
| CFG-CBRD | 0.408 | −0.0120 | 0.380 | 0.273 |
| CFG-CJSFPR | 0.475 | −0.0127 | 0.423 | 0.309 |
| | | Twitter | | |
| CFG-BRD | 0.064 | −0.0001 | 0.158 | 0.097 |
| CFG-CBRD | 0.274 | −0.0003 | 0.300 | 0.192 |
| CFG-CJSFPR | 0.298 | −0.0003 | 0.307 | 0.198 |
| | | Google+ | | |
| CFG-BRD | 0.024 | −0.0018 | 0.078 | 0.047 |
| CFG-CBRD | 0.258 | −0.0090 | 0.275 | 0.181 |
| CFG-CJSFPR | 0.161 | −0.0066 | 0.293 | 0.201 |

结束语　本文精心设计了一个分布式异构的多智能体系统，在此基础上，针对属性图聚类提出了一种快速的分布式学习算法。具体而言，每个节点可以被理解为一个代理，它们都会在预定义的约束机制下“自私”地提高自己的效用，但是它也可以以一定的概率维持现状，即使它有更好的选择。此外，还证明了所有节点代理可以同步重复上述过程，几乎可以确定博弈会收敛到一个弱帕累托纳什均衡。在现实生活中的社交网络上的大量实验证明了所提出的方法的高效性。我们未来的工作方向之一是扩展本文方法来处理一般类型的节点属性（如分类的或者数值的）。此外，还应该深入探讨如何融合一部分知识，从而能够处理动态的属性图。

## 参 考 文 献


[1] BOTHOREL C, CRUZ J D, MAGNANI M, et al. Clustering attributed graphs: Models, measures and methods[J]. Network Science, 2015, 3(3): 408-444.

[2] LI Q, ZHI Q, WANG M. Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks [J]. Journal of Software, 2006, 29(12): 2230-2237.

[3] Fortunato S. Community detection in graphs[J]. Physics Reports, 2010, 486(3-5): 75-174.

[4] CLAUSET A, NEWMAN M E, MOORE C. Finding communitystructure in very large networks[J]. Physical Review E, 2005, 70(6Pt 2): 264-277.

[5] BLONDEL V D, GUILLAUME J L, LAMBIOTTE R, et al. Fastunfolding of communities in large networks[J]. Journal of Statistical Mechanics Theory & Experiment, 2008, 2008(10): 155-168.

[6] BARTHELEMY M, FORTUNATO S. Resolution limit in community detection[J]. Proceedings of the National Academy of Sciences of the United States of America, 2007, 104(1): 36-41.

[7] ALDECOA R, MARŁN I. Surprise maximization reveals the community structure of complex networks[J]. Scientific Reports, 2013, 3(1): 173-185.

[8] CLAUSET A. Finding local community structure in networks [J]. Physical Review E Statistical Nonlinear&-Soft Matter Physics, 2005, 72(2): 254-271.

[9] LUO F, WANG J Z, PROMISLOW E. Exploring local community structures in large networks[J]. Web Intelligence & Agent Systems, 2006, 6(4): 387-400.

[10] HUANG J, SUN H, LIU Y, et al. Towards online multiresolution community detection in large-scale networks[J]. Plos One, 2011, 6(8): 492-492.

[11] LI K, PANG Y. A vertex similarity probability model for finding network community structure[C]// PAKDD. 2012: 456-467.

[12] CHEN H H, GOU L, ZHANG X, et al. Discovering missing links in networks using vertex similarity measures[C]// ACM Symposium on Applied Computing. 2012: 138-143.

[13] COMBE D, LARGERON C, EGYED-ZSIGMOND E, et al. Combining relations and text in scientific network clustering[C]// International Conference on Advances in Social Networks Analysis and Mining. 2012: 1248-1253.

[14] ZHOU Y, CHENG H, YU J X. Graph Clustering Based on Structural/Attribute Similarities[J]. Proceedings of the Vldb Endowment, 2009, 2(1): 718-729.

[15] CHENG H, ZHOU Y, YU J X. Clustering large attributed graphs: A balance between structural and attribute similarities [J]. ACM Transactions on Knowledge Discovery from Data, 2011, 5(2): 190-205.

[16] ZHOU Y, CHENG H, YU J X. Clustering large attributed graphs: An efficient incremental approach[C]// 2013 IEEE 13th International Conference on Data Mining. 2010: 689-698.

[17] III J J P, MORENO S, FOND T L, et al. Attributed graph models: Modeling network structure with correlated attributes[C]// Proceedings of the 23rd International Conference on World Wide Web. ACM, 2014: 831-842.

[18] ZANGHI H, VOLANT S, AMBROISE C. Clustering based on random graph model embedding vertex features[J]. Pattern RecognitionLetters, 2010, 31(9): 830-836.

[19] XU Z, KE Y, WANG Y, et al. A model-based approach to attributed graph clustering[C]// Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data. ACM, 2012: 505-516.

[20] GUNNEMANN S, FARBER I, BODEN B, et al. Subspace clustering meets dense subgraph mining: A synthesis of two paradigms[C]// IEEE International Conference on Data Mining. 2010: 845-850.