

一种感知关系平衡性的子团获取算法

李力恒^{1,2}, 张健沛¹, 杨 静¹, 王 勇¹

(1. 哈尔滨工程大学 计算机科学与技术学院, 150001 哈尔滨; 2. 黑龙江中医药大学 信息中心, 150040 哈尔滨)

摘 要: 为了解决社会化计算中基于感知关系平衡的子团获取问题,在关系平衡性的基础上,并根据 2-mode 网络环境的特点,提出了平衡云集子团和环境中的个体的平衡云集系数,用于群体行为的分析.使用平衡云集性阈值的方法,转化了计算过程中的迭代寻优问题,并给出 CABS-2mode 算法用于子团的划分, CABC-2mode 算法用于平衡云集性系数的获取.实验表明:算法在划分平衡云集子团的基础上,利用时间情境与社会情境融合的参数,获取到环境中个体的关系平衡性感知数值,并分析出群体中子团的平衡性强度波动过程和个体的行为过程.

关键词: 社会感知计算;平衡云集子团;平衡云集系数;数据挖掘

中图分类号: TP391

文献标志码: A

文章编号: 0367-6234(2013)11-0093-08

Aware relationship-balance of the subgroup acquisition algorithm

LI Liheng^{1,2}, ZHANG Jianpei¹, YANG Jing¹, WANG Yong¹

(1. College of Computer Science and Technology, Harbin Engineering University, 150001 Harbin;

2. Information Center, Heilongjiang University of Chinese Medicine, 150001 Harbin, China)

Abstract: To solve the problem of aware relationship-balance subgroup of context-acquisition of social computing, on the basis of the relationship-balance, and according to the characteristics of the 2-mode network environment, the relationship-balance gather subgroup and context entity of balance gather coefficient is put forward and used for group behavior analysis. The balance gather of threshold method is used to transform iterative optimization problem in the computing process and put forward the CABS-2mode algorithm used in the group clustering. The CABC-2mode algorithm is used in the acquisition of balance gather coefficient. Balance experiments show that by using the algorithm based on relationship-balance subgroup and the time-context and social-context fusion parameters, the aware relationship-balance of the entity in the context value is accessed, and the subgroup of balance to the intensity fluctuation process and the entity of behavior process are analyzed.

Key words: social aware computing; balance gather subgroup; balance gather coefficient; DM

社群的产生是一种涉及多种变量和高度复杂性的社会活动^[1].例如通过社会情境和时间情境的感知计算,获得社群内部平衡性的变化量,并以用户的视角解释量化值的含义.这对理解社群关系的演化过程,预判 k -派系、 k -核和子团的产

生,以及对社会情境推理过程有着重要的意义.因此在多情境环境下对演变社会的平衡关系云集结构分析,是群体行为研究的主要工作之一.

在以往的研究中主要利用社会网络分析技术、数据挖掘等方法分析群体的交互行为,如群体形成、组织结构、群体活动、交流模式及其动态演化等^[2-3].研究者们通过对三方组的研究,构建了平衡理论的系列模型,如平衡模型、可云集模型、等级云集性模型等^[4-7].在此基础上采用改进的贪婪算法从实体关系中生成等级分类结构^[8-9],基本思想是,根据实体之间的相似度,利用预定义的阈值来确定层级关系.进一步提出了上下文感知

收稿日期: 2013-06-10.

基金项目: 国家自然科学基金资助项目(61073043,61073041);高等学校博士学科点基金资助项目(20112304110011, 20112304110012).

作者简介: 李力恒(1977—),男,高级工程师,博士研究生;

张健沛(1956—),男,教授,博士生导师;

杨 静(1962—),女,教授,博士生导师.

通信作者: 张健沛, 26003885@qq.com.

的情境计算,利用位置信息、活动信息、环境参数等进行综合计算,为用户提供感知信息,用于社会行为的分析^[12-15].

但是目前大多数社会感知计算的研究中,忽略了在时间情境和社会情境下,基于平衡关系子团划分的感知计算.针对这一问题,本文提出了基于时序扩展的 2-mode 平衡云集子团,并根据平衡云集子团演变参数的特点进行综合计算,把社会情境和时间情境进行融合,获取到 2-mode 社会网络环境中的平衡云集子团和环境中的个体的平衡云集系数.

1 时序感知型 2-mode 社群网络问题描述

在现实的社会环境中,2-mode 网络的存在是一种普遍现象,如微博、评价网络等.若用户需要感知群体的关系环境变化,就需要对关系强度的变化进行分析.一个静态的社群网络数据集可以被视为是动态演化过程的时序快照,在基于时序的 2-mode 形式下的社群网络,需要对实体和关联强度进行形式化定义.

定义 1(时序 2-mode 社群网络 G^t) 时序社群网络 $G^t = (V^t, E^t, A^t)$ 由实体集 V^t 、有向弧集 E^t 和关系强度 A^t 组成.

其中: $t \in (t_1, \dots, t_k)$ 表示不同的时间, k 为时序序列个数;若 t_1 时刻 i 到 j 有关联,记为 $e_{i,j}^{t_1}$; i 为 $e_{i,j}^{t_1}$ 的始点记为 v_i , j 为 $e_{i,j}^{t_1}$ 的终点记为 v_j ; i 到 j 的关联强度 ω , 记为 ω_e , 它们组成的向量集记为 $A^{t_1} = [\dots, \omega_e, \dots]^{t_1}$; g 表示连通子图, 有 $g \subseteq G$.

1.1 实体关系的平衡与聚集问题

图 1 给出了实体之间三角关系.其中有实体 AB 和 AC 的 ω 均大于等于零,以实线表示. BC 的 ω 小于零, 以虚线表示. 把 1-mode 形式转换成 2-mode 形式的 AB 、 AC 和 BC 之间形成了关系循环.根据社会网络平衡理论,假设图 2 表示实体对事物的观点,其中 AB 和 AC 之间的观点一致, BC 之间的观点不一致.通过 ω 大小,实体 ABC 之间感知到整体环境的不平衡.

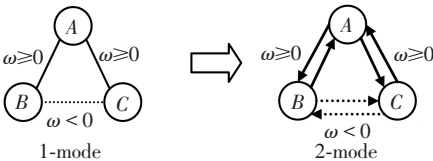


图 1 三角关系

平衡理论将社群分为两个类集合,即平衡关系集合和非平衡关系集合,实体通过 ω 的强弱感知关系的平衡性.但是现实社会中需要将社群内的实体进行多种分类,因此在平衡理论基础上扩展出聚合性问题.实体可以通过聚合子团的描述,

感知整个社群中其他团体的存在和团体之间的平衡关系.

定义 2(聚合子团 g_c) 全部由 $\omega_e \geq 0$ 的 e 组成的连通子图 g_c , 其中 $\omega_e \geq 0$ 的 e 都在 g_c 内, $\omega_e < 0$ 的 e 在 g_c 之间,称 g_c 是 G 的聚合子团.

聚合子团的划分属于寻优问题,需要进行多次分割操作,找出一种在 g_c 之间含有 $\omega_e \geq 0$ 或 $\omega_e < 0$ 最少弧的分割方式.

1.2 基于时序扩展社群平衡云集问题

为了转化社群划分过程中的迭代寻优问题,同时体现实体关联强度的平衡性和云集性变化,可以采用设定阈值的划分方法.在社会网络分析技术中,有多种度量指标可以用于设定阈值的指标,如节点度、中介性、中心势等.但是由于问题领域的不同,需要根据具体问题选择适合的基础指标.本文主要对实体的内部变化和实体之间的关系强度变化进行分析,因此采用 ω_e 为平衡云集性的基础指标.

定义 3(平衡云集阈值 $(P_\omega^+)^t$) 取不同时序 t 关联强度 $\omega_e \geq 0$ 或 $\omega_e < 0$ 的平均值 P_ω^+ , 称 P_ω^+ 为 G 的平衡云集阈值.形式化描述如下:

$$(P_\omega^+)^t = \left\{ \overline{\sum \omega_e \mid \omega_e \in A, \omega_e \geq 0} \right\}^t, \quad (1)$$

$$(P_\omega^-)^t = \left\{ \overline{\sum \omega_e \mid \omega_e \in A, \omega_e < 0} \right\}^t. \quad (2)$$

采用层次聚类的思想,从平衡性强弱的角度,以 $(P_\omega^+)^t$ 为基础对划分子团,并进行分类,平衡性较强为一种类别,平衡性较弱的为另一种类别,表 1 给出了 4 种类型平衡云集子团的形式化描述,分别为正阳型 g_n^+ 、正阴型 g_n^- 、负阳性 g_m^+ 和负阴型 g_m^- .

表 1 平衡云集子团形式化描述

类型	形式化描述
正阳型 g_n^+	$g_n^+ = \{(e_{i,j})_n \mid \omega_e \geq P_\omega^+, e_{i,j} \in E, \omega_e \in A\}$
负阳型 g_n^-	$g_n^- = \{(e_{i,j})_n \mid \omega_e \leq P_\omega^-, e_{i,j} \in E, \omega_e \in A\}$
正阴型 g_m^+	$g_m^+ = \{(e_{i,j})_m \mid 0 > \omega_e > P_\omega^+, e_{i,j} \in E, \omega_e \in A\}$
负阴型 g_m^-	$g_m^- = \{(e_{i,j})_m \mid 0 > \omega_e > P_\omega^-, e_{i,j} \in E, \omega_e \in A\}$

通过平衡云集阈值 P_ω^+ 的划分,同一个实体 v 在不同的时序 t 中有可能同时划分到不同类型的平衡云集子团中,因此需要对实体 v 的平衡云集系数 $C(v)$ 进行求解,为用户提供时间情境下的社会关系强度的波动过程.求解过程主要计算含实

体 v 的 ω_e 在不同类型 $g_{n,m}$ 所占的比重, 并把这些比重值进行叠加得到实体 v 相对 G' 的平衡云集系数 $C(v)$, 其中 $C(v)$ 分为正平衡云集系数 $C(v^+)$ 和负平衡云集系数 $C(v^-)$. $C(v^+)$ 越大或 $C(v^-)$ 越小表示平衡云集性越强, 反之亦然. 在基于时序的社群网络中, 主要观测 $C(v)$ 的变化进行群体行为的演变分析.

定义 4(时序扩展的平衡云集系数 $C(v)^t$)
设时序 T 中平衡云集子团实体 v 的弧集分别为 $e(v_n^+)$ 、 $e(v_n^-)$ 、 $e(v_m^+)$ 、 $e(v_m^-)$, 与之对应的关联强度向量集分别为 $\omega_e(e(v_n^+))$ 、 $\omega_e(e(v_n^-))$ 、 $\omega_e(e(v_m^+))$ 、 $\omega_e(e(v_m^-))$, 则实体 v 的 $C(v^+)$ 和 $C(v^-)$ 如下:

$$C(v^+)^t = C(v_n^+)^t + C(v_m^+)^t, \tag{3}$$

$$C(v_n^+)^t = \left(\frac{\sum \omega_e(e(v_{n_1}^+))}{\sum \omega_e(g_{n_1}^+)} + \frac{\sum \omega_e(e(v_{n_2}^+))}{\sum \omega_e(g_{n_2}^+)} + \dots + \frac{\sum \omega_e(e(v_{n_i}^+))}{\sum \omega_e(g_{n_i}^+)} \right)^t, \tag{4}$$

$$C(v_m^+)^t = \left(\frac{\sum \omega_e(e(v_{m_1}^+))}{\sum \omega_e(g_{m_1}^+)} + \frac{\sum \omega_e(e(v_{m_2}^+))}{\sum \omega_e(g_{m_2}^+)} + \dots + \frac{\sum \omega_e(e(v_{m_i}^+))}{\sum \omega_e(g_{m_i}^+)} \right)^t, \tag{5}$$

$$C(v^-)^t = C(v_n^-)^t + C(v_m^-)^t, \tag{6}$$

$$C(v_n^-)^t = \left(\frac{\sum \omega_e(e(v_{n_1}^-))}{\sum \omega_e(g_{n_1}^-)} + \frac{\sum \omega_e(e(v_{n_2}^-))}{\sum \omega_e(g_{n_2}^-)} + \dots + \frac{\sum \omega_e(e(v_{n_i}^-))}{\sum \omega_e(g_{n_i}^-)} \right)^t, \tag{7}$$

$$C(v_m^-)^t = \left(\frac{\sum \omega_e(e(v_{m_1}^-))}{\sum \omega_e(g_{m_1}^-)} + \frac{\sum \omega_e(e(v_{m_2}^-))}{\sum \omega_e(g_{m_2}^-)} + \dots + \frac{\sum \omega_e(e(v_{m_i}^-))}{\sum \omega_e(g_{m_i}^-)} \right)^t. \tag{8}$$

$C(v)^t$ 系数可以理解为关系平衡性的感知系数, 它与 $(P_\omega^\pm)^t$ 阈值和 $(g_{n,m}^\pm)^t$ 子团之间的关系是相互承接的. 图 2 给出 3 种感知量的关系模式, 其中 $(P_\omega^\pm)^t$ 取自 G^t , $(g_{n,m}^\pm)^t$ 子团以 $(P_\omega^\pm)^t$ 为依据进行划分, 最后 $C(v)^t$ 通过划分的 $(g_{n,m}^\pm)^t$ 子团获得个体的感知数值.

在时序社会情境环境中, $(P_\omega^\pm)^t$ 、 $(g_{n,m}^\pm)^t$ 子团和 $C(v)^t$ 数具有不同的情境含义, 表 2 给出了 3 种感知量的具体含义.

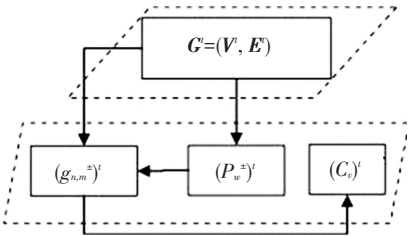


图 2 3 种感知量的与 G' 的关系模式

表 2 3 种感知量的情境含义

平衡云集值	情境含义
$g_n^+、g_n^-、g_m^+、g_m^-$	描述情境环境的主体, 在 t_k 时序情境环境下与 t_{k+1} 之间实体和实体之间关联位置的变化情况的集合.
$(P_\omega^+)^t、(P_\omega^-)^t$	描述情境环境的整体状态, 是关系环境下的情境阈值.
$C(v^+)^t、C(v^-)^t$	描述情境环境中实体 v 在不同 t_k 时序中的平衡性状态所具有相似性或吸引性的数值.

2 感知关系平衡性的云集子团划分算法

算法是把时序 2-mode 网络情境信息进行分类, 组成高层的情境对象. 图 3 给出了本文的算法的计算过程.

Step 1 首先对各个时序 2-mode 社会网络的计算情境求解, 获得 $(P_\omega^\pm)^t$; 然后使用 $(P_\omega^\pm)^t$ 获得平衡云集子团 $g_n^+、g_n^-、g_m^+、g_m^-$.

Step2 求解时序扩展的平衡云集系数 $C(v)^t$, 获得基于时间情境和计算情境下感知关系平衡的云集实体变化值.

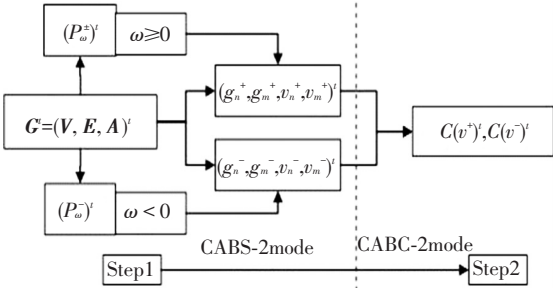


图 3 平衡云集子团和系数获取过程

其中 Step1 主要使用式 (1)、(2) 求解平衡云集阈值 $(P_\omega^\pm)^t$, 然后使用 $(P_\omega^\pm)^t$ 进行平衡云集子团的计算.

图 4 给出了算法 CABS-2mode (Context Acquisition of Balance Subgroup) 求解过程.

根据平衡云集子团的计算方式, CABS-2mode 算法有效转化了凝聚子群算法的迭代寻优问题. 这一过程与层次聚类相似, 而经典的层次聚类 Chameleon 算法进行相似度函数阈值选择时需要

人为给出 K , 在没有先验知识的情况下, 人为确定此类参数难度较大. 而 CABS-2mode 算法采用平衡云集性阈值为指标, 进行子团的获取, 有效转化了寻优问题和先验参数的选取问题.

```

输入:  $G^t$ 
输出:  $(P_{\omega}^{\pm})^t, (g_n^+, g_n^-, g_m^+, g_m^-)^t$ 
1. For  $(t_1, \dots, t_k) \in t // \text{时间序列}$ 
2.  $(P_{\omega}^+)^t = \{ \sum \omega_e \mid \omega_e \in A, \omega_e \geq 0 \}^t; // \text{式(1)}$ 
3.  $(P_{\omega}^-)^t = \{ \sum \omega_e \mid \omega_e \in A, \omega_e < 0 \}^t; // \text{式(2)}$ 
4. Do Case
5. Case  $A(\omega)^t \geq (P_{\omega}^+)^t (g_n^+)^t = \{ (e_{i,j})_n \mid e_{i,j} \in E, \omega_e \in A \}^t;$ 
   // 获取  $g_n^+$ 
6. Case  $0 > A(\omega)^t > (P_{\omega}^+)^t (g_m^+)^t = \{ (e_{i,j})_m \mid e_{i,j} \in E, \omega_e \in A \}^t; // \text{获取 } g_m^+$ 
7. Case  $A(\omega)^t < (P_{\omega}^-)^t (g_n^-)^t = \{ (e_{i,j})_n \mid e_{i,j} \in E, \omega_e \in A \}^t;$ 
   // 获取  $g_n^-$ 
8. Case  $0 > A(\omega)^t > (P_{\omega}^-)^t (g_m^-)^t = \{ (e_{i,j})_m \mid e_{i,j} \in E, \omega_e \in A \}^t; // \text{获取 } g_m^-$ 
9. EndCase;
10. Return  $(P_{\omega}^{\pm})^t, (g_n^+, g_n^-, g_m^+, g_m^-)^t;$ 
11. End For;

```

图 4 算法 CABS-2mode 求解过程

Step2 主要是通过 Step1 得到的平衡云集子团, 使用式(3)、式(6)求解平衡云集系数.

图 5 给出了算法 CABC-2mode (Context Acquisition of Balance Coefficient) 求解过程.

```

输入:  $(g_n^+, g_n^-, g_m^+, g_m^-)^t$ 
输出:  $C(v^+)^t, C(v^-)^t$ 
1. For  $(t_1, \dots, t_k) \in t$ 
2. For  $N_E(e(v_n^+, v_n^-, v_m^+, v_m^-)^t) // N \text{ 个实体 } v \text{ 的弧集}$ 
3. Do Case // 取  $C(v_n^{\pm})^t$ 
4. Case  $(e(v_n^+) \in g_n^+)^t$ 
 $C(v_n^+)^t = C(v_n^+)^t + (\omega_e(v_n^+) / \sum \omega_e(g_n^+))^t;$ 
5. Case  $(e(v_m^+) \in g_m^+)^t$ 
 $C(v_m^+)^t = C(v_m^+)^t + (\omega_e(v_m^+) / \sum \omega_e(g_m^+))^t;$ 
6. Case  $(e(v_n^-) \in g_n^-)^t C(v_n^-)^t = C(v_n^-)^t +$ 
 $(\omega_e(v_n^-) / \sum \omega_e(g_n^-))^t;$ 
7. Case  $(e(v_m^-) \in g_m^-)^t C(v_m^-)^t = C(v_m^-)^t +$ 
 $(\omega_e(v_m^-) / \sum \omega_e(g_m^-))^t;$ 
8. End Case;
9. End for;
10.  $C(v^+)^t = C(v_n^+)^t + C(v_m^+)^t;$ 
11.  $C(v^-)^t = C(v_n^-)^t + C(v_m^-)^t;$ 
12. Return  $C(v^+)^t, C(v^-)^t;$ 
13. End for;

```

图 5 算法 CABC-2mode 求解过程

CABC-2mode 算法能够计算出不同时刻下每个实体在环境中的平衡云集系数, 可以观测平衡云集系数的变化, 感知个体之间平衡关系的演变过程, 以及个体在群体中的行为过程.

3 实验和结果分析

3.1 实验数据及运行参数

实验主要进行了算法执行时间分析、感知关系平衡性的子团分析和个体行为分析. 其中采用由 Pajek 仿真生成的 300、500 和 1 000 个节点的网络数据集, 用于算法的执行时间分析. 然后采用文献[16-17]中的修士网络数据集进行平衡云集子团分析和个体行为分析.

实验的硬件环境为 Intel Core™ 2 CPU 6320 @ 1.86 GHz 2GB 内存, 操作系统 Microsoft Windows XP SP3, 算法通过 Matlab 7.10 和 Pajek2.05 混编环境下实现.

3.2 执行时间分析

CABS-2mode 算法的主要作用是获取平衡云集子团, 与层次聚类 Chameleon 算法相似. 图 6 给出了两种算法在 300、500 和 1 000 个节点数据环境下的执行时间的结果, 两种算法都随节点的数量增大, 执行时间也增大. 其中在 CABS-2mode 算法的计算过程中, 采用阈值 $(P_{\omega}^{\pm})^t$ 为依据获取子团, 这一过程转化了传统方法中需要计算任意两个簇的互连性和紧密性的迭代寻优过程, 因此在执行时间上, CABS-2mode 算法低于 Chameleon 算法的执行时间.

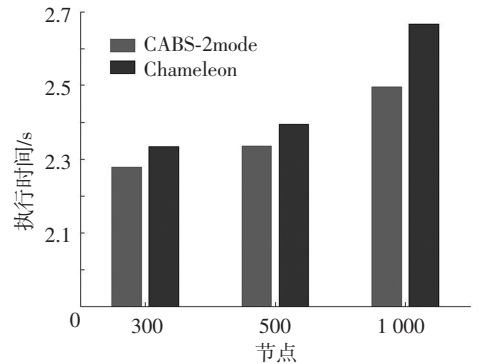


图 6 CABS-2mode 算法和 Chameleon 算法执行时间

CABC-2mode 算法主要作用是在平衡云集子团中计算个体的平衡云集系数, 使用仿真数据集由 CABS-2mode 算法划分的云集子团个数分别为 26、53 和 72 个. 图 7 给出了算法的执行时间, 其中算法随子团个数的增大执行时间也随之加大. 因为在计算每个实体所用的时间消耗, 与所占不同类型的平衡云集子团个数有关. 所以子团个数少执行时间短, 子团个数多执行时间长.

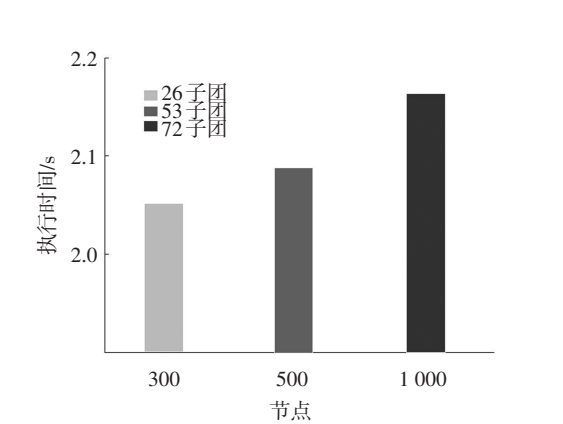


图 7 CABC-2mode 算法执行时间

3.3 感知关系平衡性的云集子团分析

群体的聚集性研究主要关注的是实体之间关系的平衡性演化,当存在大部分实体之间关系达到平衡时,子团也随之产生.

感知关系平衡性云集子团分析的实验数据一共包括了 25 个修道士,在 5 个不同时间点采集的感情关系数据.实体之间喜欢关系的连线被赋予关系值,3 为首选关系,2 为次选关系,1 为第 3 选择.不喜欢关系被赋予负的连线值,-3 为首选,-2 为次选,-1 为第 3 选择.在使用 CABS-2mode 算法进行平衡性阈值计算时发现上述网络属于实体间正负双向选择型 2-mode 关系网络,因此不同时态的 $(P_{\omega}^{\pm})^t$ 变化反映了整个网络在不同时期关系平衡性的波动情况.图 8(a) 给出了正平衡性阈值 $(P_{\omega}^+)^t$ 的波动过程,其中在 $t_2 \rightarrow t_3 \rightarrow t_4 \rightarrow t_5$ 时刻发生了波动现象,并且 t_3 和 t_4 阶段的阈值 $(P_{\omega}^+)^t$ 在平均线以下.这一现象说明整个网络的在 $t_2 \rightarrow t_3 \rightarrow t_4$ 时期的正平衡云集性趋于下降的趋势,原有的正平衡云集性关系已经开始发生变动,在 $t_4 \rightarrow t_5$ 阶段才恢复原有的正平衡云集性强度.而这一波动过程,同时也影响着负平衡关系的变化.

根据负平衡云集性的计算特点,负平衡阈值越小,负平衡关系越强,说明整个网络的反向关系或内部矛盾性比较明显.从图 8(b) 负平衡云集性波动过程中发现,从 t_1 时刻开始负平衡阈值就一直趋于增强的趋势,直至 t_3 时刻以后负平衡云集性强度一直保持最强的趋势.因此从图 8 中可以看出,整个网络的关系中正与负的平衡云集性是相互影响的,并随时间的变化而变化.

上述 $(P_{\omega}^{\pm})^t$ 阈值的波动过程,直接影响着划分子团数量.根据 CABS-2mode 算法特点,至少把目标网络划分为 4 种类型的平衡云集子团.表 3 给出了在不同时期 4 种类型的子团数量.

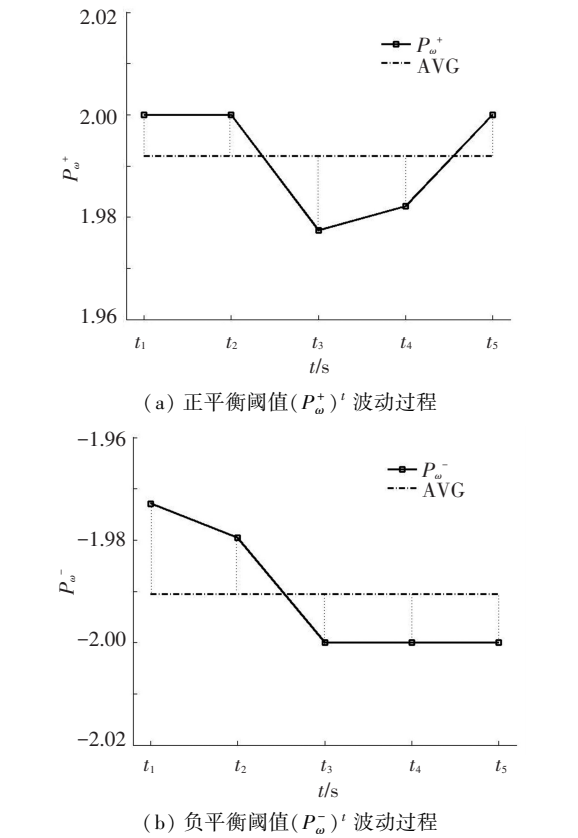


图 8 平衡阈值 $(P_{\omega}^{\pm})^t$ 波动过程

表 3 各类平衡云集子团数量

t/s	g_n^+	g_n^-	g_m^+	g_m^-	SUM
t_1	1	1	2	2	6
t_2	1	1	1	3	6
t_3	1	1	2	4	8
t_4	1	2	3	5	11
t_5	1	1	1	1	4

从云集子团的变化数量上可以看出,正阳型 g_n^+ 子团数量一直保持平稳,说明其成员一般不受关系变化的影响.而负阴型 g_m^- 子团的数量变化较为明显,说明在整个修士网络中 g_m^- 子团的成员最容易受到关系平衡性变化的影响,从而改变整网的关系结构.

图 9 给出了 5 个时态的平衡云集子团拓扑,其中正平衡云集子团 g_n^+ 和 g_m^+ 的连线值为正值,用实线连接.负平衡云集子团 g_n^- 和 g_m^- 的连线值为负值,用虚线连接.结果显示,同一个实体具有在同一时间属于不同子团的现象,这一现象说明实体之间对平衡关系环境的感知选择是多种类型的,各种类型云集子团的规模影响着不同个体在下一个时间区间的行动.如在 t_1 时刻“Bruno”分属在不同的云集子团中,其中在图 9(d) 中“Bruno”选择在一个仅含有 3 个实体相对小规模的关系子团中,随时间的变化“Bruno”最终离开了修道院,而 Peter、Bonaventure 和 Berthold 自始至终在修道院进行修行.

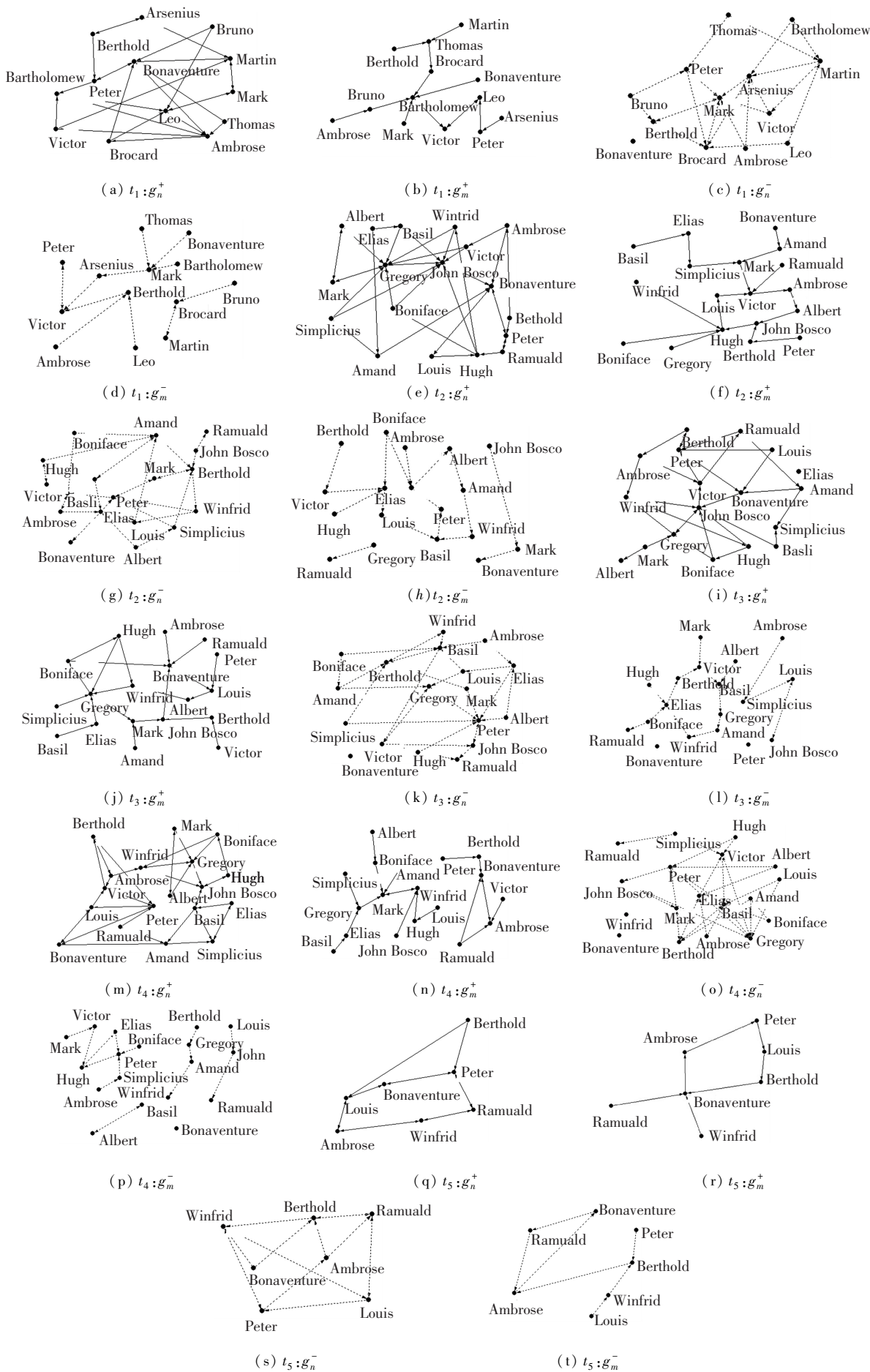
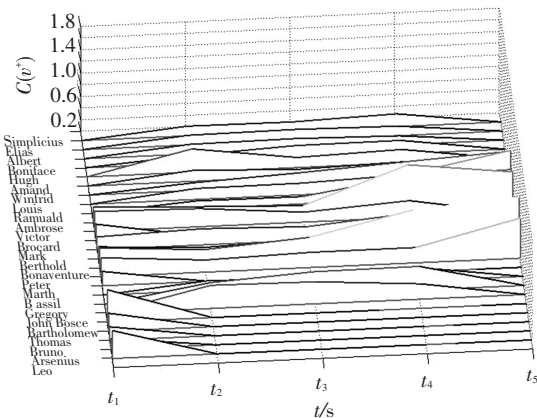


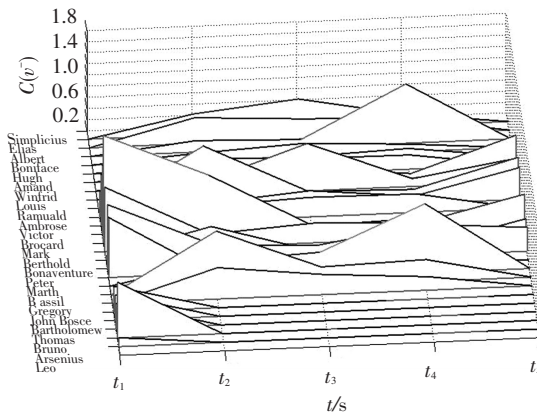
图 9 修士网络平衡云集子团

3.4 感知关系平衡性的个体行为分析

对感知的关系平衡性的个体行为进行分析, 需要把每个实体的平衡性数值化, 通过 CABC-2mode 算法的计算可以得到每个实体的平衡云集系数. 图 10 给出了修士网络中每个实体在不同时期的平衡云集系数 $C(v^*)^t$ 的波动过程.



(a) 实体 $C(v^+)^t$ 波动过程



(b) 实体 $C(v^-)^t$ 波动过程

图 10 平衡云集系数 $C(v^*)^t$ 波动过程

从图 10(b) 中可以看出每个实体的负平衡云集系数 $C(v^-)^t$ 比图 10(a) 中的正平衡云集系数 $C(v^+)^t$ 的波动过程明显, 说明 $C(v^-)^t$ 中包含的个体在整个网络变化过程中个体的不平衡性和矛盾关系比较明显, 因此不同个体有着各自不同的行为过程, 例如从图 7 中可以看出在 Leo 和 Arsenius 在 t_1 时刻就选择了离开修道院.

另外通过具体数值和子团的观测, 可以得到个体行为过程的感知因素. 表 4 给出了修士网络中部分平衡云集系数的计算数值, 其中 Bonaventure 在 t_3 和 t_4 时刻的负平衡云集系数为“0”, 说明在这两个时刻 Bonaventure 并没有对不喜欢关系进行选择. 从 Bonaventure 各个时序的平衡系数可以看出, 除 t_2 时刻负平衡系数大于正平衡系数外, 其余时刻正平衡系数都大于负平衡系

数, 因此可以感知 Bonaventure 的在整个群体中与其他修道士的关系比较融洽, 最终选择一直留在修道院.

表 4 平衡云集子团系数

t/s		Bruno	Peter	Bonaventure	Berthold
t_1	$C(v^+)$	0.230 7	0.323 0	0.338 4	0.200 0
	$C(v^-)$	0.461 5	1.153 8	0.076 9	1.153 8
t_2	$C(v^+)$	Null	0.207 7	0.229 4	0.109 9
	$C(v^-)$	Null	0.355 4	0.525 0	0.233 9
t_3	$C(v^+)$	Null	0.258 7	0.360 3	0.173 0
	$C(v^-)$	Null	0.262 5	0.000 0	0.354 2
t_4	$C(v^+)$	Null	0.261 7	0.753 6	0.221 0
	$C(v^-)$	Null	0.560 9	0.000 0	0.279 5
t_5	$C(v^+)$	Null	0.685 7	0.885 7	0.428 5
	$C(v^-)$	Null	0.428 6	0.428 6	0.714 3

Bruno 在 t_1 时刻离开了修道院, 而与之连接的 Peter 和 Berthold 在 t_1 时刻的负平衡云集系数较大, 并且由图 9(d) 可知 Bruno 在负阴型子团 g_m^- 中表现较为孤立, 而该子团内的成员容易受到平衡性关系的影响, 因此, 这种情况可能导致了 Bruno 最终选择离开了修道院.

通过上述实验与分析, 可以看出采用 CABS-2mode 算法可以在不需要先验执行的情况下, 直接环境阈值, 并使用阈值对群体进行了平衡云集性子团划分, 有效的转化传统方法中的迭代寻优问题, 结合 CABC-2mode 算法, 可以计算出每个实体的平衡云集系数, 最终获得 3 种关系平衡性的感知量, 其中 $C(v^*)^t$ 和 $(P_\omega^*)^t$ 以数值的形式体现, $g_n^+, g_n^-, g_m^+, g_m^-$ 以关系集合形式的形式体现. 它们有效地反映出实体关系的平衡性和波动过程.

4 结 论

1) 在传统研究方法的基础上, 提出平衡云集子团和基于时序扩展的平衡云集系数, 用于群体行为的分析. 并提出 CABS-2mode 算法和 CABC-2mode 算法用于平衡云集性的计算.

2) 算法使用阈值有效转化了寻优过程中的迭代问题, 通过执行时间分析, 优于基于层次聚类的 Chameleon 算法.

3) 实验表明, 采用 CABS-2mode 算法可以有效地对子团的平衡云集性进行划分, 同时采用 CABC-2mode 算法获得到个体在群体环境中的平衡云集系数.

参考文献

- [1] 於志文, 於志勇, 周兴社. 社会感知计算: 概念、问题及其研究进展[J]. 计算机学报, 2012, 35(1): 16–26.
- [2] GONZALEZ M C, HIDALGO C A, BARABASI A L. Understanding individual human mobility patterns[J]. Nature, 2008, 453(5): 779–782.
- [3] EAGLE N, PENTLAND A S. Eigenbehaviors: Identifying structure in routine[J]. Behavioral Ecology and Sociobiology, 2009, 63(7): 1057–1066.
- [4] BU Zhan, XIA Zhengyou, WANG Jiandong. A sock puppet detection algorithm on virtual spaces[J]. Knowledge-based Systems, 2013, 37: 366–377.
- [5] AMBARTSUMYAN A A. Network-centric control based on Petri nets in the structured discrete-event system[J]. Automation and Remote Control, 2012, 73(7): 1227–1241.
- [6] NIKOLAEV A G, JACOBSON S H, TAM-CHO W K, *et al.* Balance optimization subset selection (boss): an alternative approach for causal inference with observational data[J]. Operations Research, 2013, 61(2): 398–412.
- [7] LI Jianfeng, FREIREICH B J, WASSGREN C R, *et al.* Experimental validation of a 2-D population balance model for spray coating processes[J]. Chemical Engineering Science, 2013, 95: 360–365.
- [8] 方红, 杨海蓉. 贪婪算法与压缩感知理论[J]. 自动化学报, 2011, 37(12): 1413–1421.
- [9] 郭永红. 基于贪婪追踪的压缩感知重建算法研究[D]. 成都: 电子科技大学, 2012.
- [10] WANG Chifeng, HU Wen-Hsiang, BAGHERZADEH N. Scalable load balancing congestion-aware Network-on-Chip router architecture[J]. Journal of Computer and System Sciences, 2013, 79(4): 421–439.
- [11] PITKANEN M, KARKKAINEN T, OTT J. SCAMPI: service platform for social aware mobile and pervasive computing[J]. Computer Communication Review, 2012, 42(4): 503–508.
- [12] ZIGNANI M, GAITO S, ROSSI G. Extracting human mobility and social behavior from location-aware traces[J]. Wireless Communication & Mobile Computing, 2013, 13(3): 313–327.
- [13] WANG Yinghe, MA Yue, WANG Yali. Evolving network model with local-area preference for mobile Ad Hoc network[J]. China Communications, 2013, 10(5): 146–155.
- [14] 李蕊, 李仁发. 上下文感知计算及系统框架综述[J]. 计算机研究与发展, 2007, 44(2): 269–276.
- [15] FAN Jialu, CHEN Jiming, DU Yuan. DelQue: a socially aware delegation query scheme in delay-tolerant networks[J]. IEEE Transactions on Vehicular Technology, 2011, 60(5): 2181–2193.
- [16] De NOOY W, MRVAR A, BATAGELJ V. Exploratory Social Network Analysis with Pajek[M]. England: Cambridge University Press, 2011.
- [17] DOREIAN P, MRVAR A. A partitioning approach to structural balance[J]. Social Networks, 1996, 18(2): 149–168.

(编辑 张 红)