第37卷 第14期 计 算 机 工 程 2011年7月
Vol.37 No.14 Computer Engineering July 2011

·专栏· 文章编号：1000—3428(2011)14—0005—04 文献标识码：A 中图分类号：TP391

基于结构特征量化的本体模块化划分研究

梁 帅[1,2]，罗强一[2]，毕 诚[3]，黄镇鸿[1]

(1. 解放军理工大学指挥自动化学院，南京 210007；2. 中国电子设备系统工程公司研究所，北京 100141；
3. 北京市遥感信息研究所，北京 100192)

摘 要：针对大规模本体模块化划分问题，从本体结构与复杂网络的相似性分析入手，根据本体语义和网络结构特征将其转换为对应的概念网络。提出衡量本体中节点局部中心度和边的语义连接强度的量化标准，用于选取核心节点和度量边权重，得到显式表示本体语义和结构特征的无向加权层次网络，将其建模为具有不同位势的电路网络，利用最大位势差法划分本体模块。实验结果表明，该方法能够在线性时间内得到质量较高、适合匹配应用的本体划分。

关键词：本体模块化；无向加权层次网络；局部中心度；语义距离；位势差

Research of Ontology Modular Partition Based on Structure Characteristics Quantitation

LIANG Shuai[1,2], LUO Qiang-yi[2], BI Cheng[3], HUANG Zhen-hong[1]

(1. Institute of Command Automation, PLA University of Science and Technology, Nanjing 210007, China;
2. Institute of China Electronic Equipment System Engineering Corporation, Beijing 100141, China;
3. Beijing Institute of Remote Sensing Information, Beijing 100192, China)

【Abstract】As for large-scale ontology modular partition problem, this paper analyzes the similarity between complex networks and ontology structure, converts ontology to the corresponding concept networks according to its semantic and structural characteristics. It proposes a set of quantitative criteria for measuring the local central of nodes and semantic connection strength of edges, uses these quantitative criteria to identify core concept nodes and assign weight to edges, obtains undirected weighted hierarchical networks, which explicit express ontology semantic and structural characteristics, transforms the networks to circuit voltage networks, and partitions it by the maximum potential difference method. Experimental results show that this method can acquire high-quality ontology modules in linear time and fit for the further use of ontology matching application.

【Key words】ontology modularization; undirected weighted hierarchical networks; local centralities; semantic distance; voltage gap

DOI: 10.3969/j.issn.1000-3428.2011.14.002

1 概述

随着本体在信息领域应用的不断深入，本体的规模和数量都在不断扩大，这阻碍了对其进一步开发和质量控制，导致其应用的复杂性明显增加，影响了本体在协作式开发、分布式管理、大规模推理和部分重用等应用中的使用和维护。

模块化划分是解决大规模本体应用问题的有效途径。现有的本体模块化方法主要包括本体划分法和模块提取法2类[1]：(1)本体划分法将本体分割为一组模块，每个模块都是一个本体，所有模块的集合语义等同于原本体。划分结果中模块之间通常不互相覆盖。有代表性的划分法包括SWOOP、PATO和向量空间划分[2]。(2)模块提取法从本体中特定的词汇启动，递归遍历本体中的节点，以获得与特定词汇相关元素组成的模块。PROMPT和KMi采用典型的本体模块提取法。这些方法大多通过语义联系构建模块，忽略了本体结构中隐含的信息，缺少对本体中节点和边组成结构的系统分析。

本文提出基于本体结构特征量化的模块化划分方法，根据本体语义信息对其网络结构特征进行量化，采用势能分析方法对网络结构进行建模，通过迭代得出模块化划分结果。

2 本体网络特征分析

本体可以视为描述现实世界中概念间相互关系的概念网络，表示特定领域中的概念组织结构和概念之间的关系，通过将本体中的概念映射为网络中的节点、概念之间的相互关系映射为网络中的边，可以建立本体的概念网络模型。本体在转换为概念网络时需要对其进行转化：(1)添加功能性关系节点，在OWL图中以边的形式表示概念节点间的函数关系，将功能性关系由边转换为节点表示；(2)OWL中的关系通常是有向的，由于本体中的边大多可以用一对方向相反的互逆的关系类型表示，因此能用无向边代替有向边，用无向边的权重表示关系的依赖强度，从而将本体转换为具有多种类型节点和边的无向层次概念网络。其中，节点分为概念节点和功能节点2类，边分为层次性关系和非层次性关系2类。

根据对大规模本体结构特征的研究[3]，本体对应的概念网络特征主要包括：平均度数较低，大部分节点为叶子节点，存在少部分度数较大的集散节点；平均最短距离较小，较短最短路径所占的比例较大，具有类似的小世界特征；网络的度近似服从幂律分布，具有无标度特性。从而可以得出，本体中节点的重要性差距较大，少数核心概念节点能够系统地

作者简介：梁 帅(1981－)，男，博士研究生，主研方向：数据工程，知识工程；罗强一，研究员；毕 诚，工程师；黄镇鸿，硕士研究生

收稿日期：2011-02-11 **E-mail：**liangshuai12@163.com

反映本体的结构特征。

3 核心节点选取

复杂网络聚类方法通常将节点作为同级别对象处理，而本体网络中的节点之间存在多种语义关系，具有层次关系的不同节点在概念网络中的地位是不同的。本文定义局部中心度衡量节点在本体中的重要程度，充分考虑本体的语义描述能力和全局拓扑结构，选取能够较好反映本体网络特征的节点。具有较高局部中心度的节点既包含较多的语义描述信息，又能够最大限度地覆盖本体的主要组成部分，将这种节点定义为核心节点。节点局部中心度的度量与其所在的层次、密度和覆盖度有关，由于尚缺少形式化的证明，因此本文将这些标准作为假设提出。

假设 1 本体中处于中间层次的节点能更好地反映本体的细节，具有较强的语义描述能力。

节点层次能够反映本体中概念的表示方法，心理语言学家认为分类系统中处于中间级别的概念比更高或更低级别的概念更详细和重要[4]，即一个概念越接近于等级的中间层次，越能更好地分析和表示，从而定义节点层次趋近系数 $mid(c)$，以度量节点所在层次相对于中间层次的趋近程度：

$$mid(c)=1-\left|\frac{D(c)-\frac{H(c)}{2}}{\frac{H(c)}{2}}\right| \tag{1}$$

其中，$H(c)$是从根节点到包含 c 的所有分支的叶子节点中的最长距离；$D(c)$是从根节点到概念 c 的层次数，$mid(c)\in[0,1]$，距离中间层次越近，其值越大。

假设 2 本体不同位置节点分布的稀疏程度不同，节点的全局密度和局部密度越高，其在本体中的地位越重要。

节点密度表示概念节点描述的详细程度，概念细节表示确定度，描述包括概念进一步特化的程度(子类的数量)、超类的数量、直接关系数量、同伴的数量等。$density(C)\in[0,1]$ 描述包括节点的直接子类、属性和实例等。本体中不同区域概念的密集程度一般不同，一些区域可能包含许多密集概念，而另一些区域中的概念相对疏松。根据密度相对区域范围的不同，节点密度可以分为全局密度和局部密度 2 类：全局密度计算节点相对于整个本体的密度，局部密度计算节点相对于邻近概念的密度。

节点全局密度 $Gden(c)$的计算公式为：

$$Gden(c,O)=\frac{\sum_{i=1}^{3}w_i\left|S(c)_i\right|}{\max(\{\forall N\in O\rightarrow\sum_{i=1}^{3}w_i\left|S(N)_i\right|\}} \tag{2}$$

其中，$S(c)=\{S(c)_1,\ S(c)_2,\ S(c)_3\}=\{direct\text{-}subclasses[c], direct\text{-}superclasses[c], Relations[c]\}$，分别表示节点 c 具有的直接超类、直接子类和关系数量；w_i 衡量密度计算中不同类型的关系的重要程度，实验中分别取 1、0.25、0.5。

节点局部密度 $Lden(c)$的计算公式为：

$$Lden(c)=\frac{Gden(c,O)}{\max(\{\forall N\in nearest_k(c)\rightarrow wGD(c,N)\})} \tag{3}$$

其中，$wGD(c,N)=(1-(ratio_D\times dis(c,N)))\times Gden(N,O)$；$nearest_k(c)$返回概念 c 在最短路径 k 范围内包含的所有节点；$ratio_D$ 定义距离影响比例；$dis(c,N)$为 c 到 N 的最短路径长度。下文实验中设定 k=2，$ratio_D$=0.1。

节点密度可以表示为节点全局密度和局部密度的加权和，即：

$$den(c)=w_G\times Gden(c)+w_L\times Lden(c) \tag{4}$$

实验中取 w_G=0.2，w_L=0.8。由节点深度和密度得出节点的局部中心度 $LC(c)$：

$$LC(c)=\frac{w_h\times mid(c)}{\max_{1\leqslant j\leqslant|O|}\{mid(j)\}}+\frac{w_d\times den(c)}{\max_{1\leqslant j\leqslant|O|}\{den(j)\}} \tag{5}$$

其中，$w_h,w_d\in(0,1)$，$w_h+w_d=1$。在实验中取 w_h=0.3，w_d=0.7。

假设 3 本体中指定数量的节点集合对本体概念的覆盖程度越高，对本体的全局描述越准确。

前者考虑本体节点自身重要程度的衡量标准，对于多个核心节点组成的集合，不仅要求返回重要的概念节点，而且这些节点还应具有适当的分布，能够概括本体的主要部分，其衡量标准采用该集合对本体的覆盖度。节点集合的覆盖程度表示集合中有限个节点对本体概念的覆盖程度，核心节点集合的选取应确保达到较高的覆盖度。节点集合的覆盖度可以由节点关联的所有超类、子类和关系节点的数量之和与本体中节点数量的比值得出：

$$coverage(\{c_1,c_2,\text{L}\ ,c_n\})=\frac{\left|\{cov(c_1)\cup cov(c_2)\text{L}\ \cup cov(c_n)\}\right|}{|O|} \tag{6}$$

其中，$cov(c)=c\cup allSubClasses[c]\cup allSuperClasses[c]\cup Relations[c]$，表示单个节点的覆盖范围包括其自身及所有超节点、子节点和关系节点；绝对值符号表示取集合的节点数量；$|O|$表示本体包含的节点数量。在有限节点组成的集合中，单个节点对集合覆盖度的贡献程度可以表示为该节点的覆盖范围与集合中其他节点覆盖范围之差的绝对值：

$$con(c_i,\{c_1,c_1,\text{L}\ ,c_i,\text{L}\ ,c_n\})=\frac{\left|cov(c_i)-(cov(c_i)\cap(\cup_{1\leqslant j\leqslant n\wedge j\neq i}cov(c_j)))\right|}{\left|\cup_{1\leqslant k\leqslant n}cov(c_k)\right|} \tag{7}$$

综合考虑节点中心度和覆盖能力的结构特征，定义节点特征值：

$$sig(c_i,\{c_1,c_2,\text{L}\ ,c_n\})=w_c\times\frac{con(c_i,\{c_1,c_2,\text{L}\ ,c_n\})}{\max con(\{c_1,c_2,\text{L}\ ,c_n\})}+w_{nc}\times nc(c_i) \tag{8}$$

根据以上对节点结构特征的衡量标准，定义核心节点集合的选取算法，输入本体对应的概念网络和核心节点集合大小 n，输出由本体中 n 个核心节点组成的有限集合，算法实现方法如下：

(1)计算 O 中每个节点的节点层次趋近系数 $mid(c)$、节点局部密度和全局密度，得出节点局部中心度 $LC(c)$。

(2)根据节点的局部中心度进行排序，令 S 为 O 中节点中心度最高的 $k(k\leqslant n)$个概念节点，T 为$\{O-S\}$集合中 $n-k$ 个概念节点中具有最高分值的点，当 T 为空时，返回 S 并停止。

(3)对于每个 $c\in\{S\cup T\}$，令 c 为调用函数 $con(c,\{S\cup T\})$的平均值，a 为调用函数 $sig(c,\{S\cup T\})$所得的平均值。

(4)令 w 为 T 中具有最小 $sig(w,\{S\cup T\})$值的节点，R 为集合$\{\{S\cup T\}-\{w\}\}$，如果存在节点 $b\in\{O-\{S\cup T\}\}$满足：1)对每个 $c\in\{R\cup\{b\}\}$调用 $sig(c,\{R\cup\{b\}\})$生成的平均值 a'大于 a；2)对每个 $c\in\{R\cup\{b\}\}$调用 $con(c,\{R\cup\{b\}\})$生成的平均值 c'大于 c，则在$\{S\cup T\}$中交换 w 和 b，返回步骤(3)，否则，返回到$\{S\cup T\}$并结束。

4 边的语义距离量化

本体中不同类型、不同拓扑分布的边所连接的 2 个节点之间的相互作用强度是不同的，需要定义衡量标准对边的连接强度进行量化，使边的权重可以提供更多的语义信息，进一步揭示本体拓扑结构中隐含的语义特征。

本体网络中边所表示的连接强度可以用边的 2 个端点之

间的语义距离衡量，边的2个端点间的语义越相关，语义距离越小，反之，语义距离越大，边的语义相关度与关系类型、边的深度(较深节点的深度)与密度有关见表1。

表1 本体关系类型

关系类型	关系名	关系描述
继承	part-of Is-a kind-of	概念之间部分与整体的关系 概念之间的继承关系，类似于面向对象中的父类与子类之间的关系
关系	object property	概念的实例与概念之间的关系，类似于面向对象中的对象和类之间的关系
属性	attribute-of	某个概念是另一个概念的属性
实例	instance-of	概念与实例的关系

本体中边表示的语义联系主要有2种。

4.1 层次性关系度量

层次性关系由关联的强度系数、所处深度、端点密度等因素决定。层次性关系包括表示继承的Is-a、kind-of关系和表示概念之间部分与整体的part-of关系等，2种连接的相关程度有所不同，建立边的关系类型系数 α_t 进行区别，其中，t=0表示继承关系；t=1表示整体局部关系。由经验可知，继承比部分整体的语义关联度大，由此可定义继承关系的系数为 α_0=0.9，部分整体关系的系数为 α_1=0.8。边的权重与边的拓扑层次有关，同类的关系随着层次深度的加大，边的2个端点之间势差越小，概念之间关系越紧密，两端点之间的语义相关程度越高，由此可以定义边的层次系数 $\beta(c_i,c_j)$ 为：

$$\beta(c_i,c_j)=\frac{2dep(c_{ij})}{dep(c_i)+dep(c_j)} \tag{9}$$

其中，$dep(c_i)$=$dep(c_{ij})$=$dep(c_j)$−1。由式(9)可知，深度越深，层次系数越大。

边中端点的密度越大，边的重要程度越低，针对本体中的边的量化，文献[5]提出基于连接概率的计算方法及结合3种因素的层次性边权重度量方法，根据端点同种关系类型的局部密度和边的深度(较深的节点深度)计算该边2个端点之间的语义距离，其公式为：

$$dis(c_i,c_j)=(1-\beta(c_i,c_j))\times\frac{w(c_i\to_t c_j)+w(c_i\to_t c_j)}{2\max(dep(c_i),dep(c_j))} \tag{10}$$

$$w(X\to_t Y)=1-\frac{1-\alpha_t}{n_t(X)} \tag{11}$$

其中，$n_t(X)$表示 X 参与的类型为 t 的边数量。语义距离与语义相关度相对，两端点的语义距离越短，其语义相关度越高。

4.2 功能性关系度量

功能性关系两节点之间的关系强度具有社会网络的强度正比网络特征[6]，可以利用节点对关系的依赖性衡量功能性关系的重要程度。本文设定所有的功能关系都是二元关系(多元关系可以分解为多个二元关系)，功能性关系的强度主要由功能节点所连接的2个概念节点的连接强度决定。给定功能节点 r_a 分别与概念节点 c_i、c_j 相连，则可以定义概念节点 c_i 与功能节点 r_a 之间边的连接依赖系数，表示为 p_{ia}=1/n，n 为概念节点 c_i 参与的功能性关系数量，可知节点连接的功能性关系越多，对特定关系的依赖程度越低，则以概念节点 c_i 与功能节点 r_a 为端点的功能性边的关系强度定义为：

$$w(c_i,r_a)=\frac{p_{ia}+p_{ja}}{\sum_{k=1}^{n}p_{ik}+p_{c(k)k}} \tag{12}$$

其中，$c(k)$表示与概念节点 c_i 相对的节点；功能节点 r_k 所连接的另一条功能性边端点对应的概念节点 $P_{c(k)k}$ 表示概念节点 $c(k)$的连接依赖系数。

与一般的社会网络不同，本体中的概念具有层次性，在层次性关系中，子类能继承超类的功能性关系，所以，在计算以某一概念节点为端点的功能性边的权重时，要考虑超类的功能性关系，将其作为该概念节点功能性关系进行计算。度量概念节点与功能节点间功能性关系强度的具体方法为：

(1)求出本体中所有含有功能性关系的概念节点的连接依赖系数，步骤如下：

1)确定概念节点 c_i 对应的具有功能关系的最高层超节点 c_{top}；

2)找出路径 $P(c_{\text{top}}, c_i)$中所有节点的功能性关系对应的 m 个功能节点$\{r_1, r_2, \cdots, r_m\}$；

3)将 c_i 视为分别与功能节点$\{r_1, r_2,\cdots, r_m\}$连接，即 c_i 参与 m 个功能性关系，求出 c_i 对应其功能性边的连接依赖系数。

(2)将所求功能性边的相关的概念节点连接依赖系数带入式(12)，即可得到功能性边的关系强度。

5 本体划分算法

由于网络中2个节点间的语义距离可以由其最短路径所包含边的权重和确定，因此考虑将本体对应的层次概念网络建模为复杂电路系统，将网络连接视为线路，将边权重作为线路电阻，基于基尔霍夫方程求解节点位势，根据节点势能分布情况进行划分。具体的启发式规则是：在不同的簇中分别选取2个节点作为正负极后，由于簇间的电阻远远大于簇内电阻，因此同簇节点位势应近似相同，而异簇节点位势应具有明显的差异。

根据位势进行本体划分的思路是在待划分网络 G 的节点集合中任意选取2个核心节点作为极点：势能 V_1=1，V_2=0。基于基尔霍夫方程，其余节点势能为：

$$V_i=\frac{1}{k_i}\sum_{(i,j)\in E}V_j=\frac{1}{k_i}\sum_{j\in G}V_j\boldsymbol{a}_{ij}=\frac{1}{k_i}\sum_{j=3}^{n}V_j\boldsymbol{a}_{ij}+\frac{1}{k_i}\boldsymbol{a}_{i1} \tag{13}$$

其中，i={3, 4, ⋯, n}；j={1, 2, ⋯, n}；k_i 为节点 i 的度数；$\boldsymbol{a}_{ij}$ 为邻接矩阵，定义：

$$\boldsymbol{V}=\begin{pmatrix}V_3\\V_4\\\mathrm{M}\\V_n\end{pmatrix},\ \boldsymbol{B}=\begin{pmatrix}\frac{\boldsymbol{a}_{33}}{k_3}\mathrm{K}\ \frac{\boldsymbol{a}_{3n}}{k_3}\\\frac{\boldsymbol{a}_{43}}{k_3}\mathrm{K}\ \frac{\boldsymbol{a}_{4n}}{k_3}\\\mathrm{M}\quad\mathrm{M}\\\frac{a_{n3}}{k_n}\mathrm{K}\ \frac{a_{nn}}{k_n}\end{pmatrix},\ \boldsymbol{C}=\begin{pmatrix}\frac{\boldsymbol{a}_{31}}{k_3}\\\frac{\boldsymbol{a}_{41}}{k_4}\\\mathrm{M}\\\frac{a_{n1}}{k_n}\end{pmatrix} \tag{14}$$

得到矩阵方程：

$$\boldsymbol{V}=\boldsymbol{B}\boldsymbol{V}+\boldsymbol{C} \tag{15}$$

求解上式，得出其余节点的位势后，因为在2个簇之间的边相对稀疏，簇内阻率相对簇间阻率较大，所以连接处的位势下降较快，在处于中间位置的最大位势间隔可以作为2个簇的分隔，其计算公式为：

$$V_p=\max_{v_i<v_j\wedge v_i,v_j\in(0.5\pm\alpha)\wedge\neg\exists v_k,v_i<v_k<v_j}V_i+1/2(V_j-V_i) \tag{16}$$

其中，α 为区域选择系数，实验中取 α=0.2，表示在位势为0.3~0.7 的范围内选取两相邻节点最大的位势差作为节点位势模块划分边界，将原节点集合一分为二。选用核心节点集合确定簇的个数、规模和位置，核心节点之间的距离确定聚集规模；核心节点集合能够全面反映本体的结构特征，接近于不同簇的中心节点，利用核心节点代替随机节点进行极点选择，能够保证极点处于不同的簇，得出更准确的结果，并能够降低迭代的次数，减少运算时间。

本体网络模块化算法的输入为本体 O 对应的无向加权层次网络 $G(V,E,w)$，输出是本体多个模块化集合 $M(T_1,T_2,\cdots)$，具体步骤为：

(1)由核心节点集合算法求出本体核心节点集合 S；

(2)在待划分的节点集合 T 中任意选取其中的 2 个核心节点 c_i、c_j，赋予势能 V_i=1，V_j=0；

(3)根据基尔霍夫定律，计算其余节点位势 V_k(k 不等于 i、j)；

(4)根据节点位势差间隔确定 T 中节点划分，得到对应于 c_i、c_j 的 2 个节点集 T_1、T_2；

(5)判断 T_1、T_2 包含的核心节点个数：

1)如果 T_1(或 T_2)中有且仅包含一个核心节点，从 T 中清除该集合节点，将该集合加入模块集合 M；

2)如果 T_1(或 T_2)包含多于一个的核心节点，将该集合转到第(2)步；

(6)当 T 中节点为空时，结束循环，输出划分集合 M。

6 实验分析

本体模块化划分用于支持大规模本体的特定应用，下面结合本体匹配应用评价以上算法的本体模块划分质量，本体匹配中模块化划分的目标是通过对 2 个大规模本体的划分，分别得到一系列规模相近、组成相似的模块，实验在开源系统 PATO 基础上进行了扩展，运行环境为 Intel Core2 T6600 2.20 GHz 处理器，2 GB 内存，操作系统为 Windows XP。实验中选取 2 个用于匹配的中等规模的测试本体：旅游业本体 tourismA 和 tourismB，本体结构组成见表 2。

表 2 本体结构信息

本体	节点数	边数			层数
		层次性边	功能性边	合计	
tourismA	340	559	336	895	8
tourismB	474	703	552	1 255	9

理想的本体划分评价方式是预先构建本体的一个参考最优划分，通过计算自动划分与参考最优划分的查全率、查准率和 EdgeSim[7]评价自动划分的质量，区分内部边和外部边，内部边指在同一簇中 2 个节点之间的边，外部边指在不同簇中 2 个节点之间的边，查准率定义为自动划分的内部边对同时在最优划分内部边中的比例，查全率定义为在最优划分中内部边同时是自动划分的内部边的比例，EdgeSim 衡量在 2 个划分中同时是内部边和中间边占所有边的比例，分别将本文方法得到的 2 种划分(边相同与不同)与参考本体划分进行比较，结果见图 1。

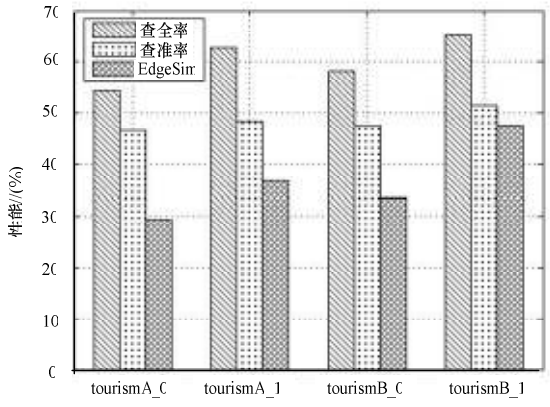


图 1 本文方法与参考最优划分的性能比较

图 1 中标号 0 表示不区分边的强度，标号 1 表示边的强度通过语义距离量化。由结果可知，通过核心节点构建势能分布可以得到较准确的本体划分，而将边的语义距离作为参数引入划分运算能在前者的基础上得到更准确的本体划分。

在大多数情况下，通过领域专家预先建立大规模本体的划分标准是不现实的，并且划分问题通常针对满足应用的特定需要，因此，需要针对应用评价划分质量的好坏。从本体匹配应用角度来讲，2 个同领域本体划分得出的模块在数量、规模、组成方面应该尽量接近。通过对比 2 个不同本体的划分结果，并可以通过 Pajak 进行可视化表示，toursimA 的划分结果核心区域如图 2 所示。

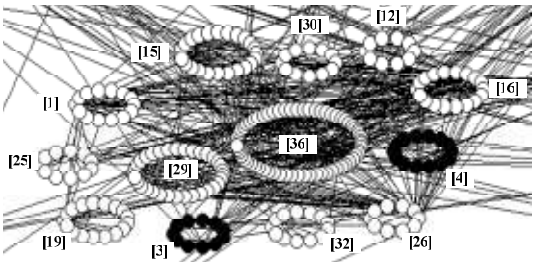


图 2 本体划分结果

对同一领域待匹配的 2 个本体的划分结果进行比较可以得出，由于在核心节点的选择中遵循同样的规则，本划分方法得到的本体模块数量接近，对应核心节点所处模块内部节点和边的组成相近，适用于以核心节点作为锚点、通过相似度扩散建立映射的本体匹配运算，在本体匹配应用中能降低计算复杂度，提高匹配准确率和效率，适用于本体匹配应用。

7 结束语

本体划分是解决大规模本体使用和维护的关键，本文从提取本体结构中隐含的语义信息出发，根据大规模本体具有的复杂网络特性，定义了核心节点的评价标准和边权重的度量方法，将本体结构量化结果建模为势能谱，通过最大电势差划分获得本体模块。仿真结果证明通过结构特征量化进行本体模块化划分的质量较高，其结果适用于本体匹配等应用。

参考文献

[1] Mathieu A, Schlicht A, Stuckenschmidt H, et al. Criteria and Evaluation for Ontology Modularization Techniques[M]. [S. l.]: Springer, 2009.

[2] 徐德智, 陶 克. 大本体的分块与映射方法研究[J]. 计算机工程, 2009, 35(4): 75-77.

[3] 张大陆, 王志晓, 刘 雯, 等. 基于复杂网络的本体结构分析[J]. 同济大学学报, 2009, 37(2): 258-261.

[4] Burt R S. Structural Holes: The Social Structure of Competition[M]. Cambridge, USA: Harvard University Press, 1992.

[5] Sussna M. Word Sense Disambiguation for Free-text Indexing Using a Massive Semantic Network[C]//Proc. of the 2nd International Conference on Information and Knowledge Management. Washington D. C., USA: IEEE Press, 1993.

[6] Stuckenschmidt H, Klein M. Structure-based Partitioning of Large Concept Hierarchies[C]//Proc. of the 3rd International Semantic Web Conference. Hiroshima, Japan: [s. n.], 2004.

[7] Mitchell B S, Mancoridis S. Comparing the Decompositions Produced by Software Clustering Algorithms Using Similarity Measurements[C]//Proc. of the 17th IEEE International Conf. on Software Maintenance. Florence, Italy: IEEE Press, 2001: 744-753.

编辑 张正兴