

文章编号:1007-130X(2011)12-0066-06

# 基于领域本体的社会网络信息分析与可视化<sup>\*</sup>

## Information Analysis and Visualization of Social Network Based on Domain Ontology

吴 鹏,李思昆

WU Peng, LI Si-kun

(国防科学技术大学计算机学院,湖南 长沙 410073)

(School of Computer Science, National University of Defence Technology, Changsha 410073, China)

**摘 要:** 社会网络分析与可视化是当前的热门研究领域,但是针对社会网络信息的高效理解与组织的研究成果却十分缺乏。本文提出一种针对社会网络信息的领域本体模型,它对社会网络信息领域的客观存在及其关系进行描述。该模型适合于描述各种社会网络分析与可视化方法,并能针对不同社会网络信息可视化应用进行扩展,克服了传统力导引布局算法在社会网络结构分析与可视化上的不足。其可视化结果能够清晰显示子群分布,表现行动者间的密切程度,显示行动者关键属性分布以及子群内部的角色分布等信息。最后通过恐怖活动信息实例,验证了领域本体模型在社会网络信息分析与显示方面的优越性。

**Abstract:** Social network analysis and visualization is an active area of study which lacks good comprehensions and organizations of social network information. This paper proposes a domain ontologies model focusing on social network information, which describes the impersonal existences and the relationships among them in the social network information domain. The model is suitable for describing a wide variety of the social network analysis and visualization methods and can extend easily for different applications of information visualization of the social network. It improves the force directed layout algorithm for overcoming the disadvantage of traditional force directed layout algorithm in analyzing and visualizing the structure of social network. The graphic results intuitively display the layout of subgroup and the close degree between actors, display key attributes and link characteristics etc. The results with the case of terrorist information show that our model can be excellent in analysis and visualization of social network information.

**关键词:** 社会网络;领域本体;信息可视化;力导引布局算法;恐怖活动信息

**Key words:** social network; domain ontology; information visualization; force directed layout algorithm; terrorist information

doi:10.3969/j.issn.1007-130X.2011.12.012

中图分类号:TP391

文献标识码:A

\* 收稿日期:2009-12-14;修订日期:2010-03-20

基金项目:国家 863 计划资助项目(2006AA01Z451)

通讯地址:410073 湖南省长沙市国防科学技术大学计算机学院 606 室

Address: Section 606, School of Computer Science, National University of Defense Technology, Changsha, Hunan 410073, P. R. China

## 1 引言

社会网络信息可视化是当前的热门研究领域,它在反恐、国家安全与军事上的应用受到越来越广泛的关注。目前有多种社会网络信息可视化方法,例如力导引布局算法、邻接矩阵布局算法、树形布局算法等等<sup>[1]</sup>。这些算法能够提供快捷的可视化实现和满意的可视化结果,如最近的多比例多体的力导引算法<sup>[2]</sup>能够在限定时间内对上万节点的网络进行节点连接分布。然而,传统的信息可视化方法缺乏对社会网络信息的理解,如最通用的力导引布局算法把节点和边看成作用力系统,却不涉及社会网络分析理论,这样固然实现简单,能够运用于各类信息的可视化应用,却缺乏对社会网络信息更深层次的理解。因此,它对于社会网络分析无法做出更大贡献,产生的图像中节点经常聚集在屏幕中央,很难分辨诸如子群分组、行动者角色等社会网络分析特征。

社会网络分析学作为一门科学已经存在了近八十年,其产生的理论、方法丰富而有效。如果在可视化方法中简单地将社会网络看成点线连接的物理系统,难免使得对社会网络信息的理解流于简单化和表面化,也无法继承与使用社会网络分析学的成果,这是传统力导引布局算法缺乏对社会网络信息的良好组织与理解的根本原因。

一种比较好的解决办法是采用本体论来描述与组织社会网络信息,在本体论模型的基础上定义分析与布局算法。这样做的好处是:首先,本体论是共享概念的显式规范说明<sup>[3]</sup>,在领域本体模型的帮助下,能够对社会网络分析与可视化领域进行很好的信息组织与规范定义;其次,能够在此框架下引入社会网络分析的多种方法,帮助布局算法的前期分析;最后,在此框架下能够根据新的社会网络分析需求,对布局算法和各种应用分析方法进行方便的扩展与改进。

目前,在社会信息可视化领域也存在一些应用采用本体论来进行信息的组织与可视化,但这些应用对本体论的运用还比较简单,没有更好地发挥本体论的优势。本文提出了针对社会网络分析的领域本体模型,该模型描述了社会网络分析领域的客观存在,并分析了本体间的关系。根据此模型进行应用与扩展,可以更好地对社会网络信息进行组织与理解,并克服传统力导引布局算法在社会网络结构分析与可视化方面的缺陷。

## 2 研究现状

力导引布局算法是 Eades<sup>[4]</sup>于 1984 年提出的。该算法把顶点与边转换为一个作用力系统,通过计算能量最小状态来对节点进行分布。这一算法易于理解和实现,但是计算开销较大,因此最近这些年来有了一系列改进。ODL(OutDegree Layout,简称 ODL)算法<sup>[5]</sup>根据社会网络中每个节点的出度来划分层次,从而简化布局问题。Quigley 提出的算法<sup>[6]</sup>是一种绘制大型无向图的快速算法,它通过基于图的结构进行空间解析,并运用多层及集群思想来加速计算。由于这些算法都缺乏对社会网络数据的透彻理解,因此很难利用已有的一些社会网络分析方法,造成社会网络结构分析与显示上的困难等问题。

目前已经存在采用本体论模型来进行社会网络信息的组织或可视化的应用。例如,RC Joint Research Center 的报告<sup>[7]</sup>针对恐怖主义活动,用本体论建模工具 Protégé 2000 对其进行了简单的本体论建模,并用本体描述语言对本体论模型进行了描述。原型系统从大量网络新闻报道中摘取恐怖主义活动相关的信息,建立了一个基于本体论的知识库,提供基于本体的查询功能。Jung 提出了一种三层的模型,包括社会网络、本体网络以及这些本体所使用的概念之间的网络<sup>[8]</sup>。这一模型采用基于网络特征的分析技术,主要适用于语义网与社会网络分析。Gilson 在他的原形系统 SemViz<sup>[9]</sup>中,使用本体映射技术和概率推理技术,将网络获取的特定领域数据自动转化为可视化图形,主要转化过程是首先将网页数据映射为领域本体,然后使用一种语法桥接本体将领域本体转换为可视化表现本体。Ma 博士的 OntoVis 系统<sup>[10]</sup>是一个用于大型异种网络的可视化工具,该工具基于本体信息进行结构抽象和过滤,采用双图模式:语义图和与之相关的本体图对社会网络进行简化,OntoVis 系统使用的实例包括从 MIPT 反恐知识库得到的恐怖活动网络。在这些例子中,本体论常用于社会网络信息的组织,却没有涉及更深层次的社会网络分析。

因此,为充分利用本体论的优势,我们提出了一种针对社会网络信息分析与可视化的领域本体。在此本体模型下,各种社会网络分析与可视化方法都能得到很好的描述,也能方便地添加新的方法,从而很好地支持了多种社会网络分析应用需求。



恐怖活动信息主要由恐怖组织、恐怖分子以及恐怖袭击组成,这三个应用本体都可以看作是行动者这一领域本体的子本体,因而可以很方便地继承领域本体内定义的属性与分析方法,此外,还为恐怖袭击定义了地点、策略、目标、武器等属性。恐怖分子可以是恐怖组织的主要成员,属性恐怖组织可以是恐怖袭击的负责人。恐怖组织之间还存在隶属、同盟等关系,这些关系是关系本体的子本体。

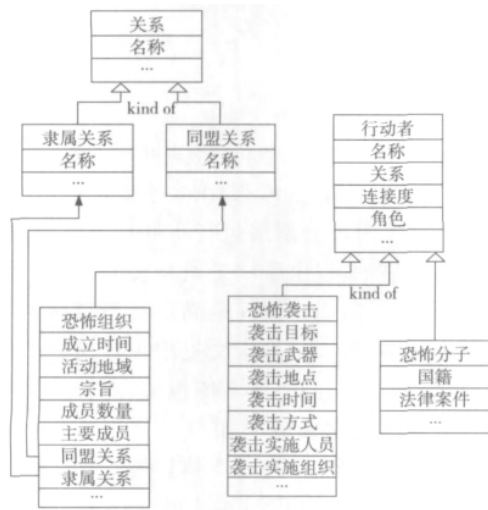


图2 恐怖活动信息本体描述

RC Joint Research Center 的报告<sup>[7]</sup>对恐怖主义活动相关的网络新闻报道的本体描述信息如图3所示,它包含组织、人员、地点和事件四个主要的本体。该本体能够描述恐怖活动信息,帮助进行信息的组织,方便查询。但是,与本文的本体模型相比,其缺乏领域基础,没有扩展余地,而且不涉及信息的分析与可视化,仅适用于信息组织与查询。

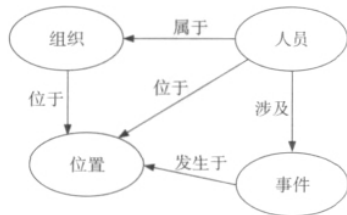


图3 恐怖活动网络新闻本体图

图4中,OntoVis系统<sup>[10]</sup>罗列出现恐怖活动信息中出现的所有关键词作为本体,但是缺乏深入分析和层次化描述。这种本体描述的本质不同也造成了OntoVis系统仅适合于示意性显示,而本文的方法不仅适合于显示,而且适合于进行更加细致的分析。例如,OntoVis中恐怖组织的活动地域和恐怖袭击的袭击地点都描述为国家地域,而在本文的本体描述中这两者分属不同的属性,甚至可以由此定义出不同的关系网络以进行详细分析。因

此,本文的领域本体模型继承性更强,可扩展性更好,更适用于对领域应用进行描述和分析。

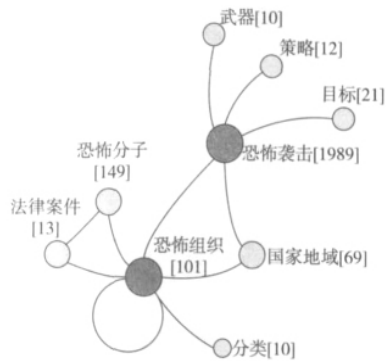


图4 OntoVis本体图

#### 4.2 恐怖组织信息可视化结果比较

可视化过程选取了3类共87个最主要的恐怖组织信息进行,这也是OntoVis系统<sup>[10]</sup>采用传统力引导布局算法实现的例子,因此能够进行横向的比较。在恐怖组织信息中,将恐怖组织看作领域模型中的行动者,它们之间存在的同盟、共享等关系构成关系网络。恐怖组织关系信息中,将盟友、怀疑盟友、怀疑别名、共享成员等关系进行合并,并看作无向关系。

运用本文的领域本体模型及其方法,能够在对社会网络信息有着更深刻的理解前提下,充分利用社会网络分析领域的方法与知识。对于传统力引导布局算法所作的改进有:

- (1)通过对恐怖组织进行角色分析,帮助了解恐怖组织的重要性;
- (2)通过模型方法对恐怖组织进行子群的分析,掌握恐怖组织网络的结构特性;
- (3)分析恐怖组织关键属性对恐怖组织子群分布的影响,挖掘恐怖组织深层关系。

采用本文算法得到的恐怖组织信息可视化结果如图5所示,图中央位置是基地组织(al-Qaeda)为首的恐怖组织子群,其右方是哈马斯组织(Hamas)为首的恐怖组织子群,一些小的恐怖组织如CPN-M,IRA和LVF等则分布在四周。

OntoVis系统<sup>[10]</sup>同样使用恐怖组织信息,结果图像如图6所示,与之相比:

- (1)实例数据基本相同,故可视化结果图形有相似之处,都能够明显看出以基地组织为核心和以哈马斯组织为核心的两个主要子群。但是,图5中子群分布更加规则,显示更加清晰。
- (2)图5中,子群间距离与子群间关键属性相似度相关,因此相似度高的恐怖组织子群间距离较

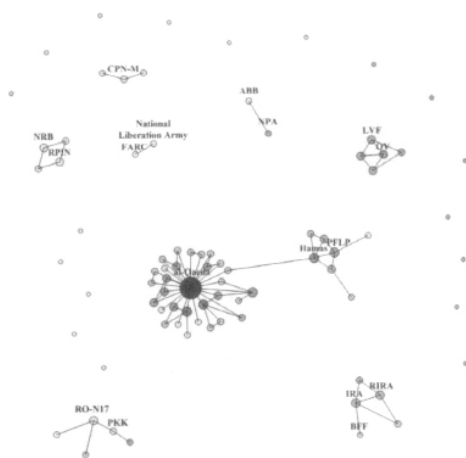


图5 本文恐怖组织信息布局结果

接近,相似度低的恐怖组织子群间距离较疏远。特别是对于孤立恐怖组织来说,传统力导引算法将其与其余恐怖组织适宜距离判定为无穷,因而它们将分布在图像的最外围。而通过本文算法的布局,孤立恐怖组织在图中的位置上会更靠近关键属性类似的恐怖组织子群。

(3)图5中恐怖组织相互间距离反映了行动者间关键属性的相似度。因此,根据可视化结果中的关键属性分布以及行动者间距离能够对子群的内部凝聚力进行直观分析。例如,在图5中以基地组织为核心的子群,其成员绝大多数都是宗教主义恐怖组织,成员间的距离较接近,因而内部凝聚力较强;而以IRA为核心的子群,其成员则是民族主义与国家主义恐怖组织参半,成员间的距离较疏远,因而内部凝聚力较差。

(4)图5中行动者灰度的深浅和半径大小代表了不同规则对等组及其连接度的大小,因此从图中可以看出以基地组织为核心的子群内部有4类角色,颜色越深则表明角色越重要。而在图6中,以基地组织为核心的恐怖组织子群,只有基地组织半径较大易于识别,其余恐怖组织的图形灰度表现完全一样。这就意味着本文算法更适合于显示社会角色分布。

从以上的对比可以看出,在本文提出的社会网络分析领域本体模型的基础上,综合运用角色分析、子群分析、关键属性分析的布局算法优于传统力导引布局算法,利于直观显示社会网络信息结构特性,帮助社会网络分析过程。

## 5 结束语

目前,社会网络信息可视化通常采用的是力导

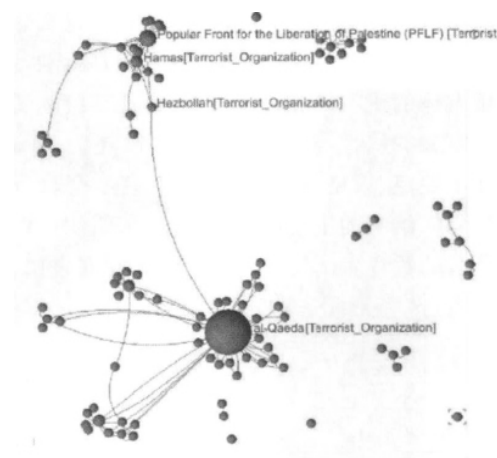


图6 恐怖组织信息传统力导引布局结果

引布局算法,该类算法常采用对行动者进行分层、集群的方法来对社会网络进行布局计算,适合于网络整体分布显示,但由于缺乏对社会网络信息的良好组织与理解,因而难以完全满足社会网络结构分析与可视化需求。为此,本文提出了针对社会网络信息分析和可视化的领域本体模型,对社会网络分析领域以及可视化领域的知识进行描述。并在此领域本体模型基础上,通过本体扩展运用于各种社会网络信息可视化应用。在领域本体模型的方法定义帮助下进行角色分析、子群分析与关键属性分析,然后根据分析结果对传统力导引布局算法进行改进,克服了其在社会网络结构分析与可视化上的不足,使得可视化结果能够清晰显示子群分布,展现行动者关键属性分布,并直观显示行动者在子群内的角色分布与影响力。最后通过恐怖组织信息实例,验证了本文模型在社会网络信息分析与显示方面的优越性。

通过本文中的领域本体模型,能够扩展更多的社会网络信息可视化分析与应用,我们计划基于本体模型来分析多类行动者之间的关系和变化,例如通过恐怖组织之间的关系来分析恐怖袭击之间的联系。

## 参考文献:

- [1] Herman I, Melancon G, Marshall M S. Graph Visualization and Navigation in Information Visualization: A Survey[J]. IEEE Transactions on Visualization and Computer Graphics, 2000, 6(1):24-43.
- [2] Hachul S, Junger M. An Experimental Comparison of Fast Algorithms for Drawing Large General Graphs[C]//Proc of Graph Drawing, 2005:235-250.
- [3] Gruber T R. A Translation Approach to Portable Ontology Specifications[J]. Knowledge Acquisition, 1993, 5(2):199-220.

- [4] Eades P. A Heuristic for Graph Drawing[J]. Congressus Numerantiunt, 1984, 42(5):149-160.
- [5] Chan D S-M, Chua K S, Leckie C, et al. Visualisation of Power-Law Network Topologies[C]//Proc of the Eleventh IEEE International Conference on Networks, 2003:69-74.
- [6] Quigley A, Eades P. FADE: Graph Drawing, Clustering, and Visual Abstraction[C]//Proc of the 8th International Symposium on Graph Drawing, 2001, 1984:77-80.
- [7] Wennerberg P O. Ontology Based Knowledge Discovery in Social Networks[R]. Ispra: Joint Research Center Report, 2005.
- [8] Jung J J, Euzenat J. Towards Semantic Social Networks[C]//Proc of the Semantic Web: Research and Applications, 2007:267-280.
- [9] Gilson O, Silva N, Grant P W, et al. From Web Data to Visualization via Ontology Mapping[J]. Journal in Modern Foreign Literatures, 2008, 27(3):959-966.
- [10] Shen Z Q, Ma K L, Eliassi-Rad T. Visual Analysis of Large Heterogeneous Social Networks by Semantic and Structural Abstraction[J]. IEEE Transactions on Visualization and Computer Graphics, 2006, 12(6):1427-1439.

- [11] 刘军. 社会网络分析导论[M]. 北京: 社会科学文献出版社, 2004.
- [12] Scott J. 社会网络分析法[M]. 刘军, 译. 重庆: 重庆大学出版社, 2007.
- [13] 冯志勇, 李文杰, 李晓红. 本体论工程及其应用[M]. 北京: 清华大学出版社, 2007.



吴鹏(1976-),男,江西南昌人,博士,助理研究员,研究方向为计算机仿真和虚拟现实。E-mail:wupenny@vip.sina.com

WU Peng, born in 1976, PhD, assistant researcher, his research interests include computer simulation, and virtual reality.



李思昆(1941-),男,山东青岛人,教授,博士生导师,研究方向为CAD和虚拟现实。E-mail:lisikun@263.net.cn

LI Si-kun, born in 1941, professor, PhD supervisor, his research interests include computer aided design, and virtual reality.