

基于社会网络分析的社会化标签网络分析与个性化信息服务研究 *

易 明 王学东 邓卫华

摘 要 在对现有的社会化标签序化方法特点与局限性进行分析的基础上,借鉴社会网络分析(SNA)思想,提出四种基于SNA的社会化标签网络分析方法:点、线、密度分析,中心性分析,凝聚子群分析,以及结构等价分析。同时,利用SNA软件实现基于社会化标签网络的个性化信息服务模型的具体应用,进而为个性化信息服务实践的推进提供一种新的思路。未来需要通过进一步实证分析来检验、完善社会化标签网络的理论与实践应用。图7。参考文献12。

关键词 社会化标签 社会网络分析 SNA 兴趣视图 个性化信息服务

分类号 G250.72

ABSTRACT The exiting tag ordering methods have many limitations. The tag network analysis based on SNA can reveal the internal relation among tags from the aspect of network structure, which can organize the tags efficiently. This approach not only overcomes the shortcomings of traditional methods, but can also be used to construct a personalized information service model, which may help to realize the ultimate intent of organizing tags. Furthermore, the personalized information service model based on the tag network can be put into practice by the SNA software, which provides a new idea for the development of practices of personalized information service. 7 figs. 12 refs.

KEY WORDS Tag. Social network analysis. SNA. Interest view. Personalized information service.

CLASS NUMBER G250.72

1 引言

作为Web2.0的主要表现形式之一,社会化标签(Tag)实现了自由分类的思想,鼓励Web用户根据自己的需要和理解对Web资源进行标注。然而,由于社会化标签的大众化特征,Web用户标注的社会化标签可能是杂乱无章的,从而引发了对社会化标签序化的研究。学者们围绕社会化标签的序化问题进行了积极探索,并提出了标签云、标签聚类、标签概念空间等解决方法^[1]。事实上,社会化标签序化的本质就是对社会化标签这种特定信息进行组织。若从信

息组织的角度来判断,这些方法还未达到有效组织社会化标签的目标。首先,标签云主要依据社会化标签被使用的频次来组织社会化标签,但无法揭示社会化标签之间的内在逻辑关联,而这恰恰是信息组织的核心;其次,标签聚类主要按照“物以类聚”的思想对社会化标签进行序化,能够将内在关联较强的一组社会化标签组织在一起,但是这种组织形式比较单一,不能多视角地揭示社会化标签之间的内在关联;最后,标签概念空间旨在构造一个反映社会化标签层级关系的概念空间^[2],通过这种概念空间的语义层级关系序化社会化标签。但是,社会化标签的一个重要特征就是打破“知识之树”

* 本文系国家社会科学基金项目“社会网络嵌入性视角的虚拟团队中知识共享理论与实证研究”(项目编号:08BTQ023)和教育部人文社会科学青年基金项目“面向用户的点击流信息资源开发与利用研究”(项目编号:08JC870005)研究成果之一。

的分类结构,没有层级结构,只存在平行关系。在实践过程中,Web 用户并没有依据某种特定概念空间去选择社会化标签,如果强行将 Web 用户使用的社会化标签映射到某种标准的概念空间,是不是具有充分的理论与现实依据则是一个值得深入思考的问题。而且,很多社会化标签都是极具个性的词语,也无法划分到特定的概念空间之中。

当前,引起学者们广泛关注的社会网络分析(Social Network Analysis, SNA)起源于 20 世纪二三十年代英国人类学的研究^[3]。经过几十年的发展, SNA 已经成为一种具有独特概念体系和测量工具的研究范式。它不仅是一种对关系或结构进行分析的技术,而且还是一种理论方法——结构分析思想^[4]。由此,本文提出基于 SNA 构建社会化标签网络来组织 Web 用户(或 Web 站点)的社会化标签的思想,即针对 Web 用户(或 Web 站点)的社会化标签网络,应用 SNA 进行结构分析以揭示社会化标签之间的内在逻辑关联,进而实现社会化标签的有效组织。虽然 SNA 作为一种跨学科的研究方法已经广泛应用于自然科学和社会科学的诸多研究领域,但是很少有学者将其暗含的结构分析思想应用到社会化标签的序化研究之中。Begelman 等人曾经依据社会化标签同时出现的频率来构建加权无向图(图的顶点表示标签,边线的权重表示同时出现的次数),但是他们仅利用聚类技术对其进行分析,并未应用 SNA 及其结构分析的思想^[1]。由此本文的研究不仅能为社会化标签序化问题的研究提供一种新的视角,而且还实现了社会化标签序化与个性化信息服务的有机融合。从实践应用来看,这种方法不仅能够一定程度上弥补标签云、标签聚类等方法的不足,而且还可用于构建 Web 用户局部兴趣视图,进而为个性化信息服务提供关键支持,达到满足 Web 用户个性化信息需求的目的。

2 Web 用户的社会化标签网络模型

对于单个 Web 用户 u 而言,其社会化标签系统中存在着 Web 资源和社会化标签两类对

象,由此可以建立 Web 用户 u 的社会化标签系统模型(见图 1)。

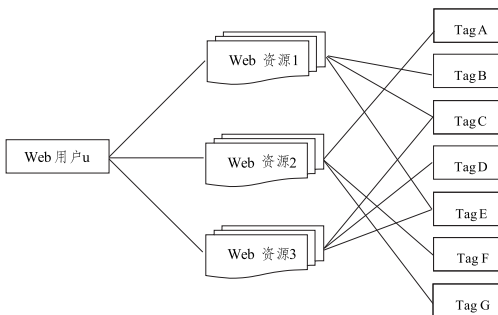


图 1 Web 用户的社会化标签系统示例

定义 1: $G = (R_c, T_c, L_c)$, 为 Web 用户 u 的社会化标签系统网络图^[5]。其中, $R_c = \{r_{G1}, r_{G2}, \dots, r_{GM}\}$, 是 Web 用户 u 的社会化标签系统中的 Web 资源集合; $T_c = \{t_{G1}, t_{G2}, \dots, t_{GN}\}$, 是 Web 用户 u 的社会化标签系统中的社会化标签集合; $L_c = \{l_{ij} = (r_{Gi}, t_{Gj}) \mid i \in [1, M], j \in [1, N]\}$, 是 Web 资源与社会化标签之间的连接关系集合, M 为 Web 用户 u 标注的 Web 资源数目, N 为对应的社会化标签数目。

在 Web 用户的社会化标签系统中,如果两个社会化标签同时出现在同一 Web 资源中,那么表明它们之间存在着重要的内在关联,可以建立连接关系。而且它们同时出现的频率越高,说明两者的关联度越大,彼此的连接关系也就越强。由此,依据图 1 可以导出如图 2 所示的 Web 用户 u 的社会化标签网络集合,它反映了 Web 用户 u 使用的社会化标签之间的内在关联。图 2 中, Tag A、Tag F、Tag G 组成的社会化标签网络与 Tag B、Tag C、Tag D、Tag E 组成的社会化标签网络是不连通的,因为分属不同网络的节点之间没有出现在同一个 Web 资源。事实上这些彼此分割的社会化标签网络就分别对应了 Web 用户不同领域的兴趣。

定义 2: $UG' = \{G'_1, G'_2, \dots, G'_X\}$, 为 Web 用户 u 的社会化标签网络集合。其中, G' 代表每个彼此分割的社会化标签网络, X 为包含的社会化标签网络的数目。

定义 3: $G' = (T'_c, L'_c, W'_c)$, 为 Web 用户 u

的一个社会化标签网络图。其中, $T'_G = \{t'_{G1}, t'_{G2}, \dots, t'_{GN'}\}$, 是对应的节点集合, $N' \leq N, T'_G \subseteq T_G$; $L'_G = \{l'_{ij} | i, j = 1, 2, \dots, N'\}$, 是对应的关系集合; $W'_G = \{w'_{ij} | i, j = 1, 2, \dots, N'\}$, 是对应的关系强度集合。

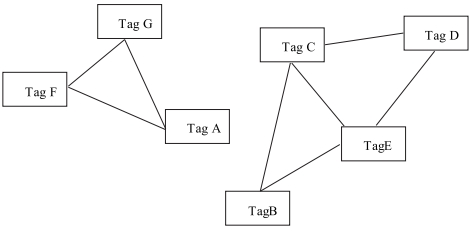


图2 Web 用户的社会化标签网络集合示例

定义4: Web 资源节点集合 R_c 中与社会化标签 $t' \in T'_c$ 相连的 Web 资源节点集合为 Rt' 。若 $Rt'_i \cap Rt'_j \neq \emptyset$, 则社会化标签 t'_i 和 t'_j 之间可以建立连接关系 l'_{ij} , 关系强度 w'_{ij} 等于 $|Rt'_i \cap Rt'_j|^{[6]}$ 。

如果将每个 Web 用户的全部社会化标签网络进行整合, 便可以得到 Web 站点的社会化标签网络集合 $WG = \{WG_1, WG_2, \dots, WG_Y\}$, Y 为包

含的社会化标签网络的数目。

3 基于 SNA 的社会化标签网络分析

基于 SNA 的社会化标签网络分析过程, 就是从网络结构的角度揭示 Web 用户 u (Web 站点) 的社会化标签之间内在逻辑关联以有效组织社会化标签的过程。需要说明的是, 由于 Web 站点的社会化标签网络集合 WG 和 Web 用户的社会化标签网络集合 UG' 在应用 SNA 方面具有类似性, 所以本文将重点研究 Web 用户的社会化标签网络集合 UG' 。依据 SNA 的基本框架, 可以从以下四个层次对集合 UG' 中的社会化标签 G 进行分析。本文以某 Web 用户在豆瓣网标注的 36 本图书共计 73 个社会化标签为研究对象, 利用社会网络分析软件 UCINET 6. 232, 可以得到该 Web 用户的社会化标签网络集合 (见图 3)。这个集合包含了 3 个社会化标签网络 $G1$ 、 $G2$ 和 $G3$, 分别对应了该 Web 用户现有的三个兴趣领域: “社会网络与知识管理”、“组织行为学”和“人力资源管理”。

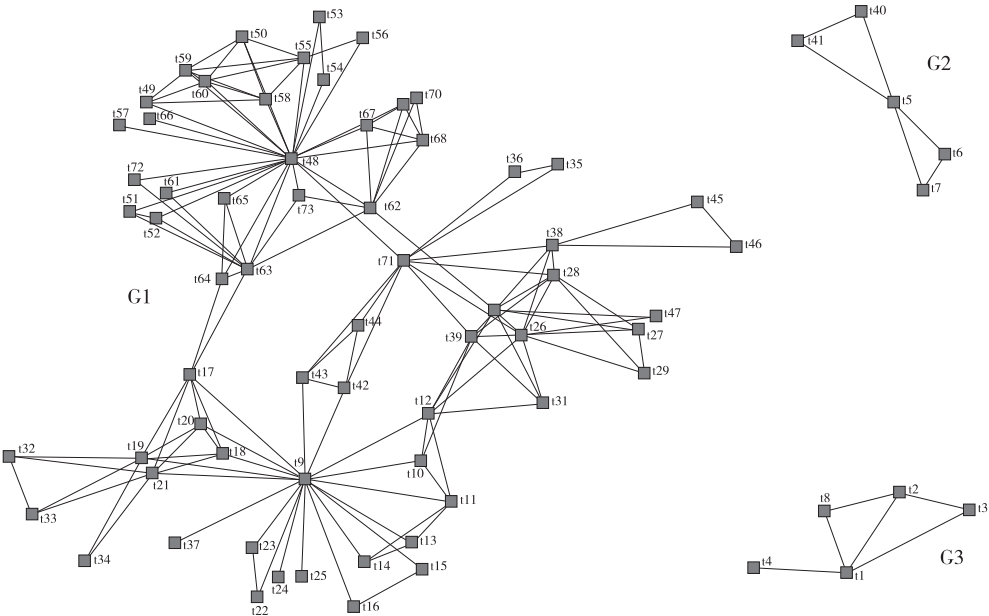


图3 豆瓣网用户的社会化标签网络集合示例

3.1 点、线、密度分析

点、线、密度是 SNA 中最基础的概念。对于社会化标签网络分析而言,比较重要的是线和密度,即社会化标签之间的关系和社会化标签网络的密度。由于 Web 用户 u 的社会化标签网络图是一个无向图,所以社会化标签之间的关系之间没有方向的区分,只有强度的差异。依据 Web 用户 u 的社会化标签网络图 $G' = (T'_c, L'_c, W'_c)$, 社会化标签之间的关系强度等于它们同时出现在同一个 Web 资源的次数 $|Rt'_i \cap Rt'_j|$ 。彼此之间关系强度越强的社会化标签,它们共同出现在同一个 Web 资源的次数就越多,同时也意味着它们之间的内在关联越强。由此,关系强度分析是对社会化标签进行有序组织的一个重要方法。

在 SNA 中,密度是为了汇总各个连线的总分布,以便测量该分布与完备图的差距,即社会网络中实际存在的连线与可能数量的连线的比例。由此,密度越大意味着社会网络中各个节点之间联系程度越紧密;反之则较稀疏。在社会化标签网络中,密度分析还可以用于判断 Web 用户 u 所使用的社会化标签之间的总体关联程度。若密度较大,则表明社会化标签网络中各个节点之间存在的连线就越多,也就意味着同时出现在同一 Web 资源的节点对就越多,节点之间的总体关联度就相对较强;反之,则说明 Web 用户 u 使用的社会化标签之间的总体关联度较弱。比如, G1 中共有 63 个节点,可能存在的连线有 3906 条,实际的连线是 299 条,由此密度只有 0.0765,说明 G1 中社会化标签之间的总体关联度较弱;虽然 G2 和 G3 的密度都是 0.6,相对较大,但是由于网络规模较小,密度分析的实际意义不大。

3.2 中心性分析

中心性分析是 SNA 中的重点之一,也被称为权力的量化分析,主要包括程度中心性分析、中介中心性分析和接近中心性分析。对于社会化标签网络分析而言,比较有意义的是程度中心性分析。程度中心性来自于社会计量学中“明星”这个概念,是用来衡量社会网络中谁是

最主要的中心人物的指标。程度中心性最简单的测量方法就是计算社会网络图中节点的度数,即与该节点直接相连的其他节点的个数。如果节点具有较高的程度中心性,那么处于该节点的个体在社会网络中的影响力和权力就越大^[7],与其他社会化标签有着更紧密的关联。需要强调的是,由于程度中心性的计算是依据社会化标签的度数,所以它能够体现标签云所反映的社会化标签使用频率。比如, G1 中度数为 5 以上的节点如图 4 所示,它们都按照节点度数大小依次排列。其中, t48、t9、t71、t26、t63、t19、t62 等分别代表的“社会网络”、“电子商务”、“高等教育出版社”、“信息管理”、“知识管理”、“数据挖掘”、“知识转移”等社会化标签属于“明星”标签。

FREEMAN'S DEGREE CENTRALITY MEASURES:

Diagonal valid?	NO		
Model:	SYMMETRIC		
Input dataset:	G1 (C:\Documents and Settings\yiming\桌面\G1)		
	1	2	3
	Degree	NormDegree	Share
38 t48	25.000	40.323	0.083
1 t9	19.000	30.645	0.063
61 t71	10.000	16.129	0.033
18 t26	10.000	16.129	0.033
53 t63	10.000	16.129	0.033
11 t19	8.000	12.903	0.027
52 t62	8.000	12.903	0.027
13 t21	8.000	12.903	0.027
22 t30	8.000	12.903	0.027
9 t17	7.000	11.290	0.023
20 t28	7.000	11.290	0.023
4 t12	7.000	11.290	0.023
31 t39	7.000	11.290	0.023
30 t38	7.000	11.290	0.023
45 t55	6.000	9.677	0.020
50 t60	6.000	9.677	0.020
49 t59	6.000	9.677	0.020
48 t58	6.000	9.677	0.020
3 t11	5.000	8.065	0.017
10 t18	5.000	8.065	0.017
59 t69	5.000	8.065	0.017
12 t20	5.000	8.065	0.017
57 t67	5.000	8.065	0.017
58 t68	5.000	8.065	0.017
40 t50	5.000	8.065	0.017
60 t70	5.000	8.065	0.017

图 4 G1 中度数为 5 以上的节点

3.3 凝聚子群分析

在 SNA 中,凝聚子群意味着局部网络中节点之间具有相对较强的、紧密的关系。在 SNA 中,凝聚子群分析可以采用多种方法,如以距离计算,在一定距离内可以达到的节点视为一个子群^[8]。由此发现,凝聚子群分析实质上就是数据挖掘中的聚类分析。对于 Web 用户的社会

化标签网络,凝聚子群分析就是将 Web 用户的社会化标签进行聚类,进而将那些内在关联较强的一组社会化标签组织在一起,类似于学者们提出的标签聚类方法。针对 G1,以 n-clique 为计算方法,同时设定最大距离为 1、子群节点

数最小为 5 的严格条件,结果见图 5。其中,第一个子群“t48、t50、t55、t58、t59、t60”对应的社会化标签分别是“社会网络、刘军、格兰诺维特、历史、博特、林南”。

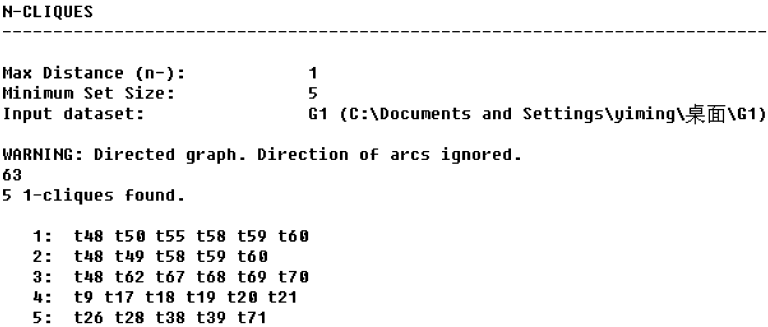


图 5 G1 的凝聚子群分析结果示例

3.4 结构等价分析

结构等价的概念于 20 世纪 70 年代早期由洛兰和怀特引入社会网络^[9] 研究中。如果一个节点与网络中其他节点的连接状态和另一节点与网络中其他节点的连接状态完全相同,则称这两个节点结构等价。然而,在任何社会网络中几乎找不到完全结构等价的节点,所以通常会采用阿基米德距离等方法来衡量结构等价。比如,节点和节点之间的阿基米德距离 d_{ij} 为^[4]:

$$d_{ij} = \sqrt{\sum_{k=1, k \neq i, j}^n [(x_{ik} - x_{jk})^2 + (x_{ki} - x_{kj})^2]}$$

其中, x_{ik} 表示的是节点 i 与节点 k 的连接状况。 d_{ij} 值越小,表明节点 i 和节点 j 之间越相似,结构上越等价。由此,在 Web 用户 u 的社会化标签网络中,结构等价分析可以找到那些结构等价的社会化标签。它们具有类似的结构属性,彼此之间具有重要的内在逻辑关联。针对 G3,各个节点之间的阿基米德距离如图 6 所示。其中,t2(“人才测评”)和 t8(“职业生涯管理”)的阿基米德距离是 1.414,相对较小,但是它们在结构上并非完全等价。

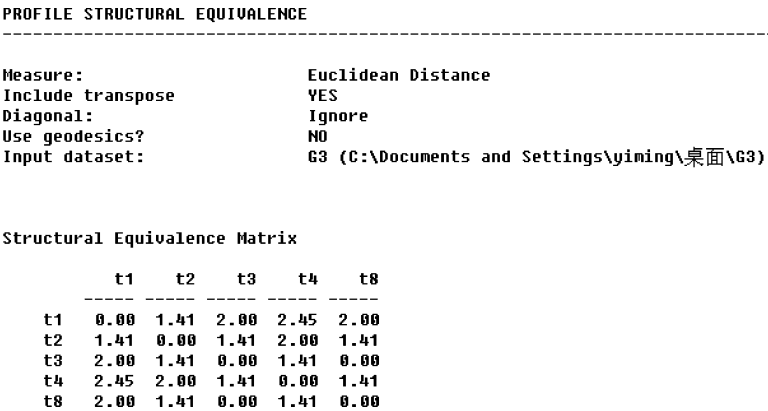


图 6 G3 的结构等价分析结果

4 基于社会化标签网络的个性化信息服务模型

Web 用户 u 使用社会化标签的直接目的是希望能够有效组织 Web 资源,最终目的则是为了更好地开发与利用 Web 资源。利用 SNA 不仅可以有效揭示 Web 用户 u 所使用的社会化标签之

间的内在逻辑关联进而实现社会化标签的序化,而且还能构建 Web 用户 u 的局部兴趣视图,以此为基础可以在 Web 用户 u 开发与利用 Web 资源过程中提供个性化信息服务,从而真正达到有效序化社会化标签的目的。图 7 描述了基于社会化标签网络的个性化信息服务模型,主要包括 Web 用户 u 局部兴趣视图的构建和基于局部兴趣视图的个性化信息服务两个关键环节。

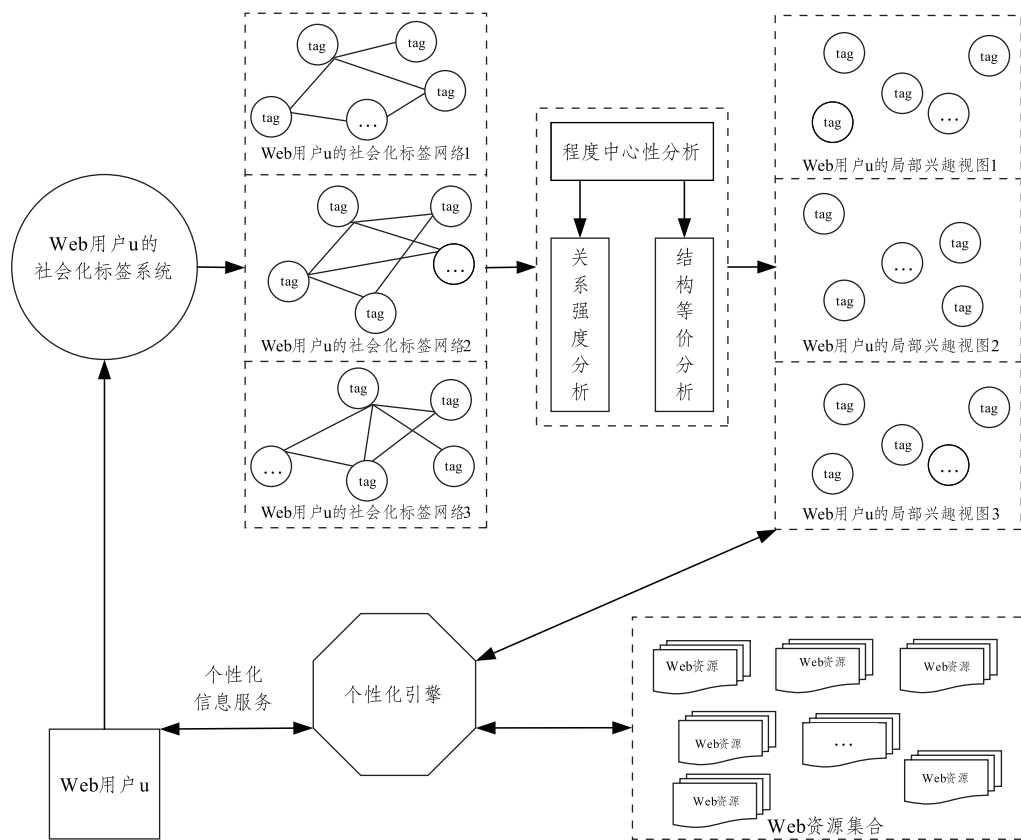


图7 基于社会化标签网络的个性化信息服务模型

4.1 Web 用户 u 局部兴趣视图的构建

Web 用户 u 的社会化标签系统,可以导出不同的社会化标签网络,反映用户不同的兴趣领域。由此,针对不同社会化标签网络进行程度中心性分析、关系强度分析和结构等价分析,可以构建 Web 用户 u 在不同领域的局部兴趣视图。

首先利用程度中心性分析寻找 Web 用户 u 社会化标签网络中的“明星”节点。由于这些“明星”出现的频次较高,而且与网络中的其他社会化标签有着更紧密的联系,所以可以利用它们来形成 Web 用户 u 局部视图的雏形^[10]。以 G1 为例,若将“明星”节点的最小度数设为 8,那么“社会网络”、“电子商务”、“高等教育出

版社”、“信息管理”、“知识管理”、“数据挖掘”、“知识转移”、“数据仓库”、“泰斗”等社会化标签都是“明星”节点,它们便形成了该豆瓣网用户在“社会网络与知识管理”领域的局部兴趣视图雏形。然后,通过关系强度分析,寻找与这些“明星”节点关系强度较高的“邻居”节点,它们与“明星”节点将逐步形成 Web 用户 u 的局部兴趣视图。比如,“社会网络”节点的重要“邻居”就有“刘军”、“罗家德”、“斯科特”、“格兰诺维特”、“组织结构”等社会化标签,关系强度均为 2。最后,通过结构等价分析寻找与“明星”节点具有一定结构等价特征的节点。虽然这些节点与“明星”节点并非完全结构等价,但是它们与“明星”节点之间的关联度也相对较高,能够进一步完善 Web 用户 u 的局部兴趣视图。由于 G1 的密度太小,利用社会网络分析软件 UCINET 6.232 计算得到的节点间阿基米德距离都比较大,没有找到合适的结构等价节点。由此,便可以得到该豆瓣网用户在“社会网络与知识管理”领域的局部兴趣视图。值得说明的是,这一局部兴趣视图基本符合该豆瓣网用户在“社会网络与知识管理”领域的兴趣,从而在一定程度上证明了方法的有效性。

定义 5: $TS = \{ts_1, ts_2, \dots, ts_y\}$, 为 Web 用户 u 的局部兴趣视图集合。其中 $ts_i = \{ts_{i1}, ts_{i2}, \dots, ts_{iz}\}$, 是以社会化标签集合的形式表示的 Web 用户 u 的局部兴趣视图 i , $ts_i \subseteq T_c, z \leq N$ 。

需要指出的是,这种以社会化标签集合形式表示的兴趣视图可以灵活地适应时间的变化^[11]。如果社会化标签网络收集的是 Web 用户最近标注的 Web 资源所对应的社会化标签,那么由此构建的兴趣视图将反映 Web 用户的近期兴趣;如果社会化标签网络收集的是 Web 用户在较长一段时间内先后标注的 Web 资源所对应的社会化标签,那么兴趣视图反映的是 Web 用户的长期兴趣。

4.2 基于局部兴趣视图的个性化信息服务

基于局部兴趣视图的个性化信息服务主要是由个性化引擎利用相关算法来完成。比如,在 Web 用户 u 浏览 Web 站点过程中,个性化引

擎可以将 Web 资源集合中所有 Web 资源与 Web 用户 u 的各个局部兴趣视图进行匹配,那些相似度较高的 Web 资源将被推送给 Web 用户 u ;或者是当 Web 用户 u 检索 Web 资源时,个性化引擎可以依据 Web 用户 u 的每个局部兴趣视图对检索结果进行过滤和排序,以实现个性化检索。由此,个性化引擎需要完成的一个重要任务就是计算 Web 资源集合中每个 Web 资源 r 与 Web 用户 u 的局部兴趣视图 ts_i 之间的相似度 $Sim(r, ts_i)$ 。

如果 Web 资源 r 已经被 Web 用户 u 标注,那么说明 Web 用户 u 已经拥有 Web 资源 r ,在个性化信息服务环节就可以暂时不考虑。所以,在计算 $Sim(r, ts_i)$ 时,只需分析 Web 资源 r 还未被 Web 用户 u 标注的情形。考虑到 Web 资源 r 与 Web 用户 u 局部兴趣视图 ts_i 之间的联系桥梁就是社会化标签,如果其他 Web 用户在标注 Web 资源 r 时用到的社会化标签都是 ts_i 中的社会化标签,那么就可以认为 Web 资源 r 与 Web 用户 u 局部兴趣视图 ts_i 比较相似。由此可以得到计算 $Sim(r, ts_i)$ 的一种方法:

定义 6: $D = (U, R, T, A)$, 为 Web 站点社会化标签系统模型^[12]。其中, $U = \{u_1, u_2, \dots, u_k\}$, 为 Web 站点的 Web 用户集合; $R = \{r_1, r_2, \dots, r_m\}$, 为 Web 站点的 Web 资源集合; $T = \{t_1, t_2, \dots, t_n\}$, 为 Web 站点的社会化标签集合; $A = \{a = (u, r, t) | u \in U, r \in R, t \in T\}$, 为 Web 用户、Web 资源和社会化标签之间的连接关系。

定义 7: $Sim(r, ts_i)$ 等于所有 Web 用户利用 ts_i 中的社会化标签来标注 Web 资源 r 的次数与所有 Web 用户利用 T 中的社会化标签标注 Web 资源 r 的次数的比值,即:

$$Sim(r, ts_i) = \frac{|\{a = (u, r, t) | u \in U, r \in R, t \in ts_i\}|}{|\{a = (u, r, t) | u \in U, r \in R, t \in T\}|}$$

比如,该豆瓣网用户还未标注《构建面向 CRM 的数据挖掘应用》一书。其他豆瓣网用户标注《构建面向 CRM 的数据挖掘应用》一书时用到的社会化标签主要有:数据挖掘(19)、客户关系(8)、客户关系管理(6)、电子商务(3)、应用(2)、用户研究(2)、技术(1)、管理(1)。由此可以得到两者的相似度: $(19 + 6 + 3) / (19 + 8 +$

$6 + 3 + 2 + 2 + 1 + 1) = 0.667$ 。至此,剩下的一个重要任务就是确定相似度最小阈值 α 。只有那些未被 Web 用户 u 标注过,同时 $Sim(r, ts_i)$ 值不低于 α 的 Web 资源,才能成为个性化引擎针对 Web 用户 u 提供个性化信息服务关键资源。对此,需要利用定义 7 的计算方法对 Web 用户 u 已经标注的 Web 资源进行测试,通过比较实验的方法,确定针对 Web 用户 u 的合理相似度最小阈值 α 。Web 用户 u 对这些 Web 资源进行标注,就意味着这些 Web 资源与其兴趣视图的相似度较高,由此利用这些 Web 资源对应的相似度 $Sim(r, ts_i)$ 来确定最小相似度阈值 α 。比如,该豆瓣网用户已经标注了《镶嵌:社会网与经济行动》一书,这就意味着这本书与该豆瓣网用户在“社会网络与知识管理”领域的局部兴趣视图的相似度较高。其他豆瓣网用户在标注这本书时用到的社会化标签主要有:社会学(55)、社会网络(43)、经济学(18)、弱连接(11)、新经济社会学(11)、格兰诺维特(7)、sns(7)、行为经济学(7)。由此可以得到两者的相似度: $(43 + 7) / (55 + 43 + 18 + 11 + 11 + 7 + 7 + 7) = 0.314$ 。利用同样的方法,可以得到该豆瓣网用户标注的其他 Web 资源对应的相似度 $Sim(r, ts_i)$,进而通过加权或者是直接选择最小值等方法确定针对该豆瓣网用户的相似度最小阈值 α 。

5 结语

本文借鉴社会网络分析思想,提出了以社会化标签网络的形式有效组织社会化标签的方法。相对于已有的标签云、标签聚类、标签概念空间等方法,这种方法不仅具有独特的优势,而且还能构建 Web 用户局部兴趣视图,以此为基础建立的个性化信息服务模型也能较好地满足 Web 用户的个性化需求。由于本文研究的重点是 Web 用户的社会化标签网络,所以关于 Web 站点的社会化标签网络分析与应用将是研究的重要内容。此外,通过进一步实证分析来检验、完善社会化标签网络的理论研究与实践应用也是今后研究的重点。

参考文献:

- [1] Begelman G, Keller P, Smadja F. Automated tag clustering: Improving search and exploration in the tag space [C]. In Proceedings of the Collaborative Web Tagging Workshop at WWW'06, 2006.
- [2] Kome S H. Hierarchical subject relationships in folksonomies [D]. North Carolina: the University of North Carolina, 2005.
- [3] 伊凡艾琳达. 社会网络分析在信息科学中的应用和发展 [OL]. [2009-05-30]. <http://blog.zjol.com.cn/1185/viewspace-243513>.
- [4] 林聚任. 社会网络分析:理论、方法与应用[M]. 北京:北京师范大学出版社,2009:44,161.
- [5] 刘臣,张庆普,单伟,等. 学科知识流动网络的构建与分析[J]. 情报学报,2009,28(2):257-265.
- [6] 赵鹏,蔡庆生,王清毅. 交联网络中的可重叠社团结构分析算法[J]. 华南理工大学学报(自然科学版),2008,36(5):19-23.
- [7] Wasserman S, Faust K. Social network analysis: Methods and applications [M]. Cambridge: Cambridge University Press, 1994.
- [8] 罗家德. 社会网分析讲义[M]. 北京:社会科学文献出版社,2005:134-140.
- [9] 李晓佳,张鹏,狄增如,等. 复杂网络中的社团结构[J]. 复杂系统与复杂性科学,2008,5(3):19-42.
- [10] Fan Bei, Liu Lu, Li Ming, et al. Knowledge recommendation based on social network theory [C]. In 2008 IEEE Symposium on Advanced Management of Information for Globalized Enterprises, 2008: 322-324.
- [11] 王翠英. 基于 Folksonomies 的用户偏好挖掘研究[J]. 现代图书情报技术,2009,25(6):37-43.
- [12] Shepitsen A, Gemmell J, Mobasher B, et al. Personalized recommendation in social tagging systems using hierarchical clustering [C]. In Proceedings of the 2008 ACM conference on Recommender systems, 2008: 259-266.

易明 华中师范大学信息管理系副教授,硕士生导师。通讯地址:武汉。邮编 430079。

王学东 华中师范大学信息管理系教授。通讯地址同上。

邓卫华 华中农业大学经济管理学院信息管理系讲师。通讯地址:武汉。邮编 430070。

(收稿日期:2009-08-30;
最后修回日期:2009-10-20)