

基于合作博弈论的社会网络关键节点发现研究

王学光

(华东政法大学信息科学与技术系 上海 201620)

摘要 社会网络关键节点发现问题有着许多重要的应用,同时也越来越受到了研究机构和学者的关注。为了找到社会网络中 top K 个关键节点,提出了一种基于合作博弈理论的方法。该方法考虑到社会网络普遍存在社区结构,利用 Owen 值得到每个节点的边际贡献,由此得到该问题的解。在 2 个合成数据集和 4 个实际数据集上验证了本方法的可行性和较其他方法的有效性。

关键词 社会网络,关键节点问题,合作博弈论,Owen 值

中图分类号 TP393 **文献标识码** A

Discovering Critical Nodes in Social Networks Based on Cooperative Games

WANG Xue-guang

(Department of Computer Science, East China University of Political Science and Law, Shanghai 201620, China)

Abstract Discovering critical nodes in social networks has many important applications and more and more institutions and scholars have been attended. For finding out the top K critical nodes in social networks, this paper presented a method based on cooperative games to obtain each node's marginal contribution by using Owen value and considering the widespread community structure in social networks. And then we can get the solution of the critical nodes problem. The feasibility and effectiveness of our method were verified on two synthetic datasets and four real datasets.

Keywords Social networks, Critical node problem, Cooperative games, Owen values

1 引言

信息的扩散、融合以及级联行为是发生在网络中的基本过程。考虑商业公司计划推出新产品的场景,利用网络的这种传播特征(或称之为“word-of-mouth”、“viral marketing”效应),可以先找到一些具有影响力的关键个体,让他们推荐产品给他们的朋友们,以使得这样的级联在人群中最大范围地传播。如何选择这些具有影响力的个体被称为“关键节点问题(Critical Node Problem, CNP)”。该问题的有效解决具有重要的实用价值。比如,在罪犯关系网络中,可以迅速定位犯罪团伙的头目,集中警力进行布控;在电力网络中,对重要的断路器、发电单元等进行保护,可以有效防止由相继故障引起的大范围停电;在传染病网络中,可以有针对性地先治疗、隔离病源,有效防止病毒的传播和扩散;在谣言传播网络中,可以发掘出始作俑者,避免“蝴蝶效应”等等^[1]。

本文给出一种解决该问题的方法,即利用合作博弈理论的联盟和解概念为社会网络每个社区中的节点赋予边际贡献,据此对网络中的所有节点进行排序并按照一定规则获取相应节点,它们即为关键节点。本文第 2 节介绍两个基本的融合模型以及相关背景知识;第 3 节提出一种基于 Owen 值的社会网络关键节点发现算法;第 4 节对本文算法进行实验验证和性能讨论;最后给出本文的结论。

2 背景知识

2.1 融合模型

信息在网络上传播的模型已经被广泛研究^[2-4],本文考虑两类基本的模型:独立级联模型(Independent Cascade Model, ICM)^[5]和线性阈值模型(Linear Threshold Model, LTM)^[6]。在介绍上述两种模型前,首先给出一些必要的定义和假设。

网络可以由图 $G=(V, E)$ 表示,其中 V, E 分别表示节点和边的集合。在社会网络中,节点表示个体,边表示节点间的关系。一个节点可以有两种状态:激活(active)和非激活(inactive),这意味着一种产品、思想等信息是否被个体接受或采纳。假定节点可以由非激活态转换为激活态,反之则不能;先期处于非激活态的节点在之后的某个时间可以由处于激活态的邻居激活,同样,处于激活态的该节点也可以去激活它的一些处于非激活态的邻居;网络中激活节点的扩散表示了信息的传播。

• 独立级联模型

在该模型中,对于每一条边 $(u, v) \in E$ 都有相对应的传播概率 $p_{u,v}$,它表示当 u 处于激活态后 v 被其激活的概率。假定网络中初始激活节点集为 A_0 ,网络上的影响扩散过程以离散时间步骤 $i \geq 0$ 展开。当一个节点 u 在时间 i 被激活,它仅有一次机会以概率 $p_{u,v}$ 激活它的非激活态邻居节点 v ,激活的顺序可以是任意的。如果 u 成功激活 v ,则节点 v 在时间 $i+1$

变为激活态;否则 u 不能再去尝试激活节点 v ,直到扩散过程结束。影响扩散过程直到没有更多的节点被激活而结束。

• 线性阈值模型

在该模型中,节点 v 以权重 $w_{v,u}$ 受其邻居节点 u 影响且满足条件 $\sum_v w_{v,u} \leq 1$ 。同时每个节点都有一预定义阈值 $\theta_v \in [0,1]$,它意味着节点 v 的激活态邻居去激活它时的困难程度,通常阈值以均匀分布随机选取。假定网络中初始激活节点集为 A_0 ,网络上的影响扩散过程以离散时间步骤 $i \geq 0$ 展开。所有在时间 i 被激活的节点在时间 $i+1$ 仍保持激活态,任一未激活节点 v 由其处于激活态邻居的权重和来决定是否被激活,即需满足条件 $\sum_v w_{v,u} \geq \theta_v$ 。影响扩散过程直到没有更多的节点被激活而结束。

在 ICM 中,对于某一节点的激活尝试都是独立于其他已激活邻居节点的,而在 LTM 中,某一节点是否被激活受其所有已激活邻居节点的权重和的影响。

2.2 问题描述

关键节点发现问题定义如下:

输入:图 $G=(V,E)$ 和整数 k ;

输出: $A = \arg \max_{S \subseteq V, |S| \leq k} \sigma(S)$ 。

对于 k -CNP 问题,希望找到具有 k 个元素的集合 A 使得 $\sigma(A)$ 最大化。这是一个 NP-hard 问题,但是 Nemhauser 等人已经证明存在一个贪婪算法(见算法 1)使其在 $(1-1/e)$ 范围内近似最优解^[7]。

算法 1 贪婪算法 Greedy

1. $A = \emptyset$
2. for $i=1$ to k do
3. $v_i = \arg \max_{u \in V \setminus A} (\sigma(A \cup \{u\}) - \sigma(A))$
4. $A = A \cup \{v_i\}$
5. end for

在算法 1 中一个关键问题是需要计算函数 $\sigma(A)$,然而目前还没有有效的方法得到其精确解。但是,可以采用 Monte-Carlo 方法多次模拟影响扩散过程,从而以较高的概率获得近似的结果。

2.3 相关工作

复杂网络理论提出了大量的统计属性可用于社会网络分析,比如节点的度、聚类系数、介数等^[8]。PageRank^[9] 和 HITS 算法^[10] 应用特征向量中心性(Eigenvector Centrality)对 Web 页面排列;White 和 Smyth 使用马尔科夫中心性(Markov Centrality)计算节点的相对重要性^[11];Shetty 和 Adibi 采用图熵(Graph Entropy)方法发现电子邮件网络中的重要节点^[12];Li 等人将人工神经网络应用到有影响力的 Bloggers 的识别中^[13];Lappas 等人基于有向图,利用树结构和动态线性规划方法找到 Effectors^[14]。许多学者也针对特定社会网络并利用网络相关信息进行关键节点的挖掘^[15-20]。

值得关注的是,Domingos 和 Richardson 最先将关键节点发现作为算法问题来进行研究^[21,22];Kempe 等人形式化该问题为离散优化问题,证明了该问题是 NP-Hard 问题并提出了贪婪近似算法^[2]。然而,他们方法的效率严重制约了它的可扩展性。Leskovec 等人提出了一个优化方法 CELF(Cost-Effective Lazy Forward),其能够加速这些个体的选择^[23]。Chen 等人另辟蹊径,着眼于在网络结构局部范围内计算节点之间的相互影响从而得到具有影响力的节点,而不是去改进

贪婪算法本身。首先,他们针对基于独立级联模型的融合过程进行了优化,提出了 Degree Discount 启发式算法^[24]。该方法影响扩散的精确度接近于贪婪算法,而其运行时间较贪婪算法提高了 6 个数量级。随后,他们针对通用独立级联模型,构建节点的局部树结构来限制影响传播的范围,提出了 MIA (Maximum Influence Arborescence)模型及其算法^[25]。该方法能够处理大规模网络的关键节点发现问题,并具有较好的效率。同时,他们也提出了一个 LDAG(Local Directed Acyclic Graph)方法来处理基于通用线性阈值模型的关键节点发现问题,取得了较好的效果^[26]。他们也研究了消极意见出现和传播时的关键节点发现问题^[27]。

3 算法

3.1 合作博弈与 Owen 值

给定一个有限参与者集合 N ,可转移效用合作博弈(Cooperative Game with Transferable Utility)可以表示为有序数对 (N, v) ,特征函数 $v: 2^N \rightarrow \mathbb{R}$ 且 $v(\emptyset) = 0$ 。对于 $\forall i \in N$,如果支付向量(payoff vector)满足 $x_i \geq v(\{i\})$ 且 $\sum_{i=1}^N x_i = v(N)$,则称该支付向量为 (N, v) 的一个配置。合作博弈的解就是一种配置规则,其分配给每一个参与者一个支付,代表了评价参与者在博弈中谈判能力的一种方法。Shapely 提出了一个解概念(Solution Concept),即从具有不同性质的解中找到唯一的配置分配方案^[28],考虑根据每一个参与者对该博弈的重要程度来分配参与者的支付。参与者 i 在博弈 (N, v) 中的 Shapely 值为

$$Sh_i(v) = \sum_{\{S \subseteq N | i \in S\}} \frac{(n-s)! (s-1)!}{n!} (v(S \cup \{i\}) - v(S)), \quad \forall i \in N$$

其中, $n = |N|, s = |S|$ 。

然而,Shapely 值的计算并没有考虑到联盟结构(Coalition Structure)的影响。因此,Owen 对其进行了扩展^[29],首先各个同盟(union)之间进行博弈,每一个同盟都获得支付,然后这些支付在同盟内部各个成员之间通过内部博弈进行分配。同盟之间进行博弈所获支付和同盟内部各成员之间所获支付都由 Shapely 值来确定。

设 $N = \{1, 2, \dots, n\}, M = \{1, \dots, m\}$,那么 N 上的每一种分隔 $P = \{N_1, N_2, \dots, N_m\}$ 就是一种联盟结构。其中,任一 N_k 为一同盟,且 $\bigcup_{k=1}^m N_k = N$,当 $l \neq k$ 时, $N_l \cap N_k = \emptyset$ 。对于任意 $i \in N, k(i)$ 表示包含参与者的同盟索引,则有 $i \in N_{k(i)}$ 。对于任意同盟 $k \in M$,假定 $S \subseteq N_k$,对于所有 $Q \subseteq M$,博弈 \hat{v}_S 定义为

$$\hat{v}_S(Q) = \begin{cases} v(\bigcup_{h \in Q} N_h), & k \notin Q \\ v(\bigcup_{h \in Q \setminus \{k\}} N_h \cup S), & k \in Q \end{cases}$$

博弈 \bar{v}_k 定义为 $\bar{v}_k(S) = Sh_k(\hat{v}_S)$,那么,博弈 $\bar{v}_{k(i)}$ 中参与者 $i \in N$ 的 Owen 值为

$$Ow_i(v, P) = Sh_i(\bar{v}_{k(i)})$$

3.2 基于 Owen 值的关键节点发现算法

前文中贪婪算法的基本思想是在每一迭代过程中基于社会网络融合模型找到一个具有最大影响的节点,本质上是选择信息扩散过程中具有最大边际贡献(marginal contribution)

的节点。由于社会网络中普遍存在社区结构 (community structure)^[30], 因此需分别考虑每一社区和社区内部节点对该社区在信息扩散方面的影响。将社会网络中的节点视为合作博弈论中的参与者, 信息融合视为联盟的形成, 因此可以将社会网络中的信息融合映射为可转移效用合作博弈, 在文献 [31] 中提供了这种映射的形式化描述。因此, 可以利用合作博弈论中的 Owen 值获得每个节点的边际贡献。由于 Owen 值可以看作是 Shapley 值的两次应用, 因此首先需要计算节点的 Shapely 值。

给定节点 $i \in N$ 、子集 $S \subseteq N$ 且 $i \notin S$, 定义节点 i 的边际贡献为 $v(S \cup \{i\}) - v(S)$ 并满足 $\forall S \subseteq N \setminus \{i\}$ 。考虑 N 上所有可能排列集合 Ψ , 令 $\psi \in \Psi$, 定义 $S_i(\psi)$ 为 ψ 中节点 i 之前出现的所有节点的集合, 那么参与者 i 对给定博弈的平均边际贡献为

$$\frac{1}{n!} \sum_{\psi \in \Psi} [v(S_i(\psi) \cup \{i\}) - v(S_i(\psi))]$$

需要注意的是, 上述方法必须考虑 $n!$ 个排列, 其计算复杂度为 $O((n/e)^n)$ ^[32]。因此, 本文给出计算 Shapely 值的近似方法。我们随机产生 t 个排列的集合 Ψ_t , 令排列 $\psi \in \Psi_t$, $\psi(i)$ 表示排列中的第 i 个节点。首先, 考虑排列中只有节点 $\psi(1)$ 处于激活状态时运行融合模型后有多少个节点被激活, 即为 $\psi(1)$ 的贡献。接下来考虑 $\psi(2)$, 如果 $\psi(1)$ 激活后, $\psi(2)$ 才被激活, 则 $\psi(2)$ 的贡献为 0; 否则激活 $\psi(2)$, 再确定由它激活的节点数目。依此一直计算到 $\psi(n)$ 。对于任一 $\psi \in \Psi_t$, 重复上述过程 R 次。然后, 计算每个节点在影响扩散过程中的贡献均值。通过排序可以得到最具影响力的 k 个节点, 同时保证它们互不相邻, 如算法 2、算法 3 所示。

算法 2 ShapelyValue(v)

```

n = |ψ|
t = |Ψt|
tmp[1...n] = 0
ShV[1...n] = 0
for i = 1 to t do
    for r = 1 to R do
        for j = 1 to n do
            tmp[j] = tmp[j] + v(Si(ψr) ∪ {j}) - v(Si(ψr))
        end
    end
end
for i = 1 to n do
    ShV[i] = tmp[i] / (R * t)
end

```

算法 3 TOP_K(V)

```

TopK[1...k] = 0
AsceSort(V);
TopK[1] = V[1]
i = 1
j = 2;
while i < k do
    if V[j] is not adjacent to TopK[1...i-1] then
        TopK[i] = V[j]
        i = i + 1
    end
    j = j + 1
end

```

由上述算法, 根据 3.1 节描述给出 Owen 值的计算方法。

Roger Guimerà 等人^[33]根据节点社区内度 (within-module degree, z) 和社区间参与系数 (participation coefficient, P), 研究了网络社区结构中节点的角色并将其划分为 7 类 $Role = \{R1, R2, R3, R4, R5, R6, R7\}$ 。根据 Owen 值的思想, 考虑 Non-hub connector node ($z < 2.5, 0.62 < P \leq 0.80$, 角色标识为 $R3$) 和 Connector hub ($z \geq 2.5, 0.30 < P \leq 0.75$, 角色标识为 $R6$) 这两类节点, 与它们连接的节点大多数属于其他社区。首先利用 CNM 算法^[34]将网络 $G = (V, E)$ 划分为 l 个社区 $C = \{C_1, C_2, \dots, C_l \mid C_i = (V_i, E_i), i = 1, \dots, l\}$, 并为社区中的节点 $v_i^j \in V_i$ 指定角色信息 $rv_i^j \in Role$ 。

令 $Role' = \{R3, R6\}, i = 1, \dots, l, j = 1, \dots, l, V' = \{v_i^j, v_j^i \mid \exists (v_i^j, v_j^i), v_i^j \in V_i, rv_i^j \in Role', i \neq j\}, E' = \{e_i^j, (v_i^j, v_j^i) \mid e_i^j = (v_i^a, v_i^b) \in E_i, v_i^a \in V_i, v_i^b \in V_i, rv_i^a \in Role', i \neq j\}$, 定义 $G' = (V', E')$ 为社区博弈网络。由此, 可以计算该网络中每一节点 Shapely 值, 将属于同一社区的所有节点 Shapely 值的和作为该社区的支付。然后, 将每一社区作为一个独立的网络计算其中所有节点的 Shapely 值, 社区中的节点根据归一化后的 Shapely 值对该社区的支付进行分配便得到该节点的 Owen 值。然后, 应用算法 3 即可得到 k 个关键节点。

4 实验

我们将分别在 2 个合成数据集与 4 个真实网络数据集上验证本文方法的性能。本文所有实验都是在主频为 Intel CPU 3.2GHz、内存 4G 和操作系统为 Windows 7 的个人计算机上运行, 采用 Matlab 2009 和 Microsoft Visual Studio 2010 开发工具。

4.1 数据集

本文采用 BA 模型^[35]与“森林火灾”模型 (The Forest Fire Model)^[36]分别产生具有 5000 个节点的合成数据集, 称它们为 BA 和 FF。BA 模型将幂率分布的产生归结为两个重要特性: 增长 (Growth) 和优先连接 (Preferential Attachment)。该模型能够产生幂指数为 3 的无标度 (Scale-free) 网络且不具有明显的社区特征。“森林火灾”模型能够产生具有度幂率分布、满足稠化定律 (Densification Laws)、直径缩减 (Shrinking Diameters) 并具有明显社区特征的网络。

实际数据集包括 DBLP、Facebook、Enron 和 Youtube 等。本文的 DBLP 数据集^[37]以 1997 年到 2006 年发表在计算机各领域重要期刊和会议上的文章作为基本数据, 构建具有 143276 个节点和 359812 条边的共作者网络; Facebook 数据集^[38]使用从 2007 年 1 月 1 日到 2008 年 12 月 31 日之间所涉及的 60567 位用户和 583766 条连接来构建 New Orleans 区域中朋友关系网络; Enron 数据集^[39]包含 50 万封电子邮件, 通过将邮件地址作为节点, 以两个地址间有邮件通信就建立一条无向边的方式构成有 36692 个节点和 367662 条边的网络。Youtube 数据集^[40]取 2007 年 1 月 1 日到 2007 年 1 月 15 日的的数据来构建网络, 由 35468 个节点和 261191 条边组成。

4.2 实验结果

我们对比本文算法、贪婪算法 Greedy (即算法 1)、度启发式方法 Degree 以及随机选择法 Random。Degree 算法选择网络中具有 k 个最大度节点作为初始集合, Random 算法随机选择 k 个节点作为初始集合。为了获得各算法的影响扩散范围, 对于每一初始集合分别运行 ICM 和 LTM 模型 10000 次

并取其均值。在 ICM 模型中采用同一激活概率 $p=0.05$, LTM 模型中节点的边权重设置为其度的倒数。假定初始集合的大小从 1 到 20 依次增加。在实验中,对于 ICM 和 LTM 两个模型得到了一致结论。下面仅讨论使用 ICM 的情形。

首先,考察社区结构对本文算法的影响。我们分别产生 BA 和 FF 网络数据集,根据上节所述对每一数据集计算其节点的 Shapely 值和 Owen 值并得到初始激活节点集合,根据初始节点集合在对应的数据集上运行 ICM 模型得到激活节点数目。重复这一过程 100 次,并得到不同规模大小的初始节点集合对应的激活节点数目均值,如图 1 所示,其中虚线 Sv 表示基于 Shapely 值的关键节点发现方法,实线 Ov 则表示基于 Owen 值的方法。在不具明显社区特征的 BA 数据集上,利用 Shapely 方法和 Owen 方法得到的初始节点集合所导致的激活节点数几乎相同(见图 1(a)),而在具有社区结构的 FF 数据集上,基于 Owen 值的方法则明显优于基于 Shapely 值的方法(见图 1(b))。

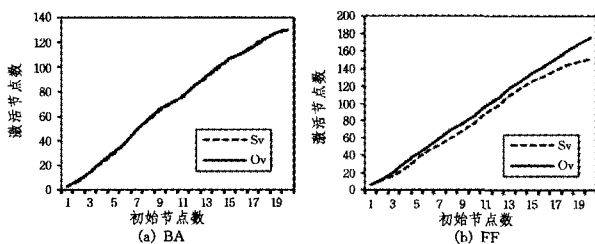


图 1 社区结构对本文算法的影响

其次,分析本文算法的精确度。在 4 个数据集上分别使用 Greedy、Ov、Degree 和 Random 算法找到初始节点集合,考察不同规模的初始节点集合基于 ICM 模型时的激活节点数目,图 2 展示了本文的实验结果。从图 2 中可以看到,本文算法 Ov 在发现关键节点的精确度上与 Greedy 算法相似,有时甚至优于 Greedy 算法(见图 2(b));与 Degree 和 Random 算法相比,本文算法 Ov 有着较大的优势。

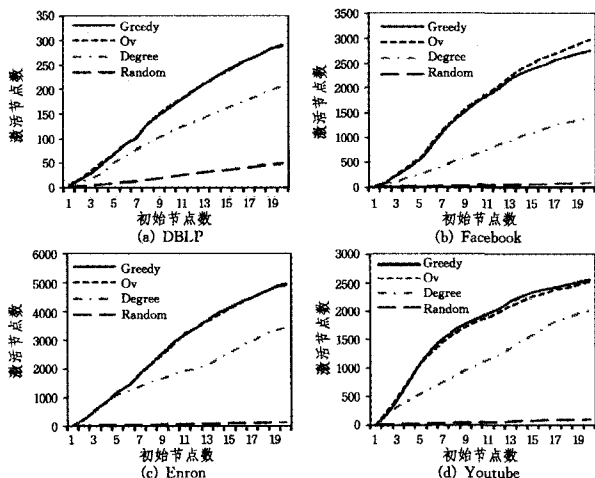


图 2 算法性能分析

结束语 为了解决关键节点发现问题,考虑到社会网络中普遍存在社区结构,基于合作博弈理论提出了一种利用 Owen 值得到该问题解的方法。在 2 个合成数据集上验证了本文方法更适用于网络中存在社区结构的情形;在 4 个真实数据集上通过与其他方法的对比结果显示,本文方法更有效。如何提高获得关键节点的时间效率需要进一步考虑和研究。

[1] 赫南,等. 复杂网络中重要性节点发掘综述[J]. 计算机科学, 2007,34(12):1-5

[2] Kempe D, Kleinberg J M, Tardos é. Maximizing the spread of influence through a social network[C]//The 9th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD). Washington,DC,2003:137-146

[3] Even-Dar E, Shapira A. A note on maximizing the spread of influence in social networks[C]//WINE 2007, LNCS 4858. 2007: 281-286

[4] Lahiri M, Cebrian M. The genetic algorithm as a general diffusion model for social networks[C]//The Twenty-Fourth AAAI Conference on Artificial Intelligence. Atlanta, Georgia, 2010: 494-499

[5] Goldenberg J, Libai B, Muller E. Talk of the network: A complex systems look at the underlying process of word-of-mouth [J]. Marketing Letters,2001,12(3):211-223

[6] Granovetter M. Threshold models of collective behavior[J]. American Journal of Sociology,1978,83(6):1420-1443

[7] Nemhauser G, Wolsey L, Fisher M. An analysis of the approximations for maximizing submodular set functions[J]. Mathematical Programming,1978,14(1):265-294

[8] Costa L F, Rodrigues F A, Travieso G, et al. Characterization of complex networks-A survey of measurements[J]. Advances in Physics,2007,56(1):167-242

[9] Brin S, Page L. The anatomy of a large-scale hypertextual Web search engine [J]. Computer Networks and ISDN Systems, 1998,30(1-7):107-117

[10] Kleinberg J M. Authoritative sources in a hyperlinked environment[J]. Journal of the ACM,1999,46(5):604-632

[11] White S, Smyth P. Algorithms for estimating relative importance in networks[C]//The ninth ACM SIGKDD international conference on Knowledge discovery and data mining (KDD). Washington,DC,2003:266-275

[12] Shetty J, Adibi J. Discovering Important Nodes through Graph Entropy The Case of Enron Email Database[C]//The 3rd international workshop on Link discovery. Chicago, Illinois, 2005:74-81

[13] Li Y M, Lai C Y, Chen C W. Identifying Bloggers with Marketing Influence in the Blogosphere[C]//The 11th International Conference on Electronic Commerce (ICEC). Taipei, Taiwan, 2009:335-340

[14] Li Y M, Lai C Y, Lin C H. Discovering Influential Nodes for Viral Marketing[C]//The 42nd Hawaii International Conference on System Sciences (HICSS). Waikoloa,2009:1-10

[15] Lappas T, Terzi E, Gunopulos D, et al. Finding Effectors in Social Networks[C]//The 16th ACM SIGKDD conference on Knowledge discovery and data mining (KDD). Washington,DC, 2010:1059-1068

[16] Java A, Kolari P, Finin T, et al. Modeling the spread of influence on the blogosphere[C]//The 15th International World Wide Web Conference. Edinburgh, UK, 2006

[17] Hussain D M A. Terrorist Networks Analysis through Argument Driven Hypotheses Model[C]//The Second International Conference on Availability, Reliability and Security. Vienna, Austria,2007:480-492

[18] Zhai Z, Xu H, Jia P. Identifying Opinion Leaders in BBS[C]//

- The IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology. Sydney, Australia, 2008;398-401
- [19] Kargar M, An A. Discovering top-k teams of experts with without a leader in social networks[C]//The 20th ACM International Conference on Information and Knowledge Management (CIKM). Glasgow, Scotland, UK, 2011;895-994
- [20] Zhang Y, Zhou J, Cheng J. Preference-based Top-K Influential Nodes Mining in Social Networks[C]//The IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications. Changsha, China, 2011;1512-1518
- [21] Domingos P, Richardson M. Mining the network value of customers[C]// The 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD). San Francisco, CA, 2001;57-66
- [22] Domingos P, Richardson M. Mining knowledge-sharing sites for viral marketing[C]//The 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD). Edmonton, Canada, 2002;61-70
- [23] Leskovec J, Krause A, Guestrin C, et al. Cost-effective outbreak detection in networks[C]//The 13th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD). San Jose, California, 2007;420-429
- [24] Chen W, Wang Y, Yang S. Efficient influence maximization in social networks[C]// The 15th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD). Paris, France, 2009;199-207
- [25] Chen W, Wang C, Wang Y. Scalable influence maximization for prevalent viral marketing in large scale social networks[C]// The 16th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD). Washington, DC, 2010;1029-1038
- [26] Chen W, Yuan Y, Zhang L. Scalable influence maximization in social networks under the linear threshold model[C]//The 2010 IEEE International Conference on Data Mining (ICDM). Sydney, Australia, 2010;88-97
- [27] Chen W, Collins A, Cummings R, et al. Influence Maximization in Social Networks When Negative Opinions May Emerge and Propagate[C]// 2011 SIAM International Conference on Data Mining. Mesa, Arizona, USA, 2011
- [28] Shapley L S. A value for n-person games[M]// Kuhn H W, Tucker A W, eds. Contributions to the Theory of Games II. Princeton University Press, 1953;307-317
- [29] Owen G. Values of games with a priori unions[M]// Henn R, Moeschlin O, eds. Essays in mathematical economics and game theory. Springer-Verlag, Berlin, 1977;76-88
- [30] Scott J. Social Network Analysis(2nd ed)[M]. A Handbook, Sage, London, 2000
- [31] Caulier J-F. Network Games as TU Cooperative Games[M]. The Core, the Shapley Value and Simple Network Games, 2009
- [32] Cormen T H, Leiserson C E, Rivest R L, et al. Introduction to Algorithms(2nd ed)[M]. Cambridge, MA; MIT Press, 2001
- [33] Guimerà R, Amaral L A N. Functional cartography of complex metabolic networks[J]. Nature, 2005, 433:895-900
- [34] Clauset A, Newman M E J, Moore C. Finding community structure in very large networks[J]. Physical Review E, 2004, 70(6):066111 (6)
- [35] Barabási A L, Albert R. Emergence of scaling in random networks[J]. Science, 1999, 286(5439):509-512
- [36] Leskovec J, Kleinberg J, Faloutsos C. Graphs over time; Densification laws, shrinking diameters and possible explanations[C]// The 11th ACM SIGKDD international conference on Knowledge discovery in data mining (KDD). Chicago, Illinois, USA, 2005;177-187
- [37] <http://dblp.uni-trier.de/>
- [38] Viswanath B, Mislove A, Cha M, et al. On the evolution of user interaction in Facebook[C]//The 2nd ACM SIGCOMM Workshop on Social Networks (WOSN). Barcelona, Spain, 2009;37-42
- [39] <http://www-2.cs.cmu.edu/~enron/>
- [40] Mislove A, Marcon M, Gummadi K P, et al. Measurement and analysis of online social networks[C]// The 7th ACM SIGCOMM conference on Internet measurement conference (IMC). San Diego, California, USA, 2007;29-42

(上接第 126 页)

- [11] Sure Y, Staab S, Studer R. On-To-Knowledge Methodology (OTKM). Handbook on Ontologies[M]. New York, Springer, 2004;117-132
- [12] Pinto H S, Staab S, Tempich C. Diligent; Towards a Fine-grained Methodology for Distributed, Loosely-controlled and Evolving Engineering of Ontologies[C]//ECAI 2004;393-397
- [13] De Nicola A, Missikoff M, Navigli R. A Software Engineering Approach to Ontology Building[J]. Information Systems, 2009, 34:258-275
- [14] Fernández López M. Overview of Methodologies for Building Ontologies[C]//Proceedings of the IJCAI-99 Workshop on Ontologies and Problem-Solving Methods (KRR5). Stockholm, Sweden, August 1999;41-53
- [15] Corcho O, Fernandez-Lopez M, Gomez-Perez A. Methodologies, Tools and Languages for Building Ontologies. Where Is Their Meeting Point? [J]. Data & Knowledge Engineering, Elsevier Science B. V, 2003, 46(1):41-64
- [16] Sure Y, Staab S, Studer R. Ontology Engineering Methodology. Handbook on Ontologies(2nd Edition)[M]. London, Springer, 2009;135-152
- [17] Noy N F, McGuinness D L. Ontology Development 101; A Guide to Creating Your First Ontology[OL]. http://Protege.stanford.edu/publications/Ontology_development, 2012-05
- [18] Guarino N, Welty C A. An Overview of OntoClean. Handbook on Ontologies (2nd Edition)[M]. London, Springer, 2009;201-220
- [19] Gangemi A, Presutti V. Ontology Design Patterns. Handbook on Ontologies(2nd Edition)[M]. Springer, 2009;221-244
- [20] McGuinness D L, van Harmelen F. OWL Web Ontology Language Overview[S]. W3C Recommendation, February 2004
- [21] Mizoguchi R, Kozaki K. Ontology Engineering Environments. Handbook on Ontologies(2nd Edition)[M]. London, Springer, 2009;315-336
- [22] Khondoker M R, Mueller H P. Comparing Ontology Development Tools Based on an Online Survey[C]//Proceedings of the World Congress on Engineering, WCE 2010. London, U. K., 2010;978-988