

基于 Owen 值算法的社会网络关键节点问题研究

王学光

(华东政法大学信息科学与技术系 上海 201620) (伦敦大学学院计算机系 伦敦 WC1E 6BT)

摘要 社会网络关键节点发现问题有着许多重要的应用,如何找到网络中具有最大影响力的 K 个节点是一个 NP 问题。考虑到社会网络中普遍存在着社区结构,提出一种新的社会网络的关键节点发现算法,其在两个信息融合模型的基础上利用 Owen 值和 Monte-Carlo 方法得到每个节点的边际贡献,其中边际贡献最大的 K 个节点即为该问题的解。实验结果表明,该算法更适用于网络中存在社区结构的情形,在时间效率上相对于 Greedy 算法有几十倍的提高。

关键词 社区结构,关键节点问题,Owen 值,影响最大化

中图分类号 TP393 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2016.5.008

Research on Critical Nodes in Social Networks Based on Owen Values

WANG Xue-guang

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

(Department of Computer Science, University College London (UCL), London WC1E 6BT, UK)

Abstract Discovering critical nodes in social networks has many important applications and how to find out K critical nodes with the most influence in a social network is a NP problem. Considering the widespread community structure in social networks, this paper presented an algorithm of discovering critical nodes based on two information diffusion models and obtained each node's marginal contribution by using Owen value and Monte-Carlo method. And the solution of the critical nodes problem is the K nodes with the highest marginal contributions. The results show that the algorithm is more suitable for the networks with community structure and the time efficiency of the algorithm is dozens of times than the Greedy algorithm.

Keywords Community structure, Critical node problem, Owen values, Influence maximization

1 引言

传染病在人类或动物网络中的传播是一种融合过程,如禽流感等;类似地,计算机病毒也会在 Internet 上大规模地蔓延而导致严重的损失;一个人使用了一种新的产品后可能会告诉他的朋友们也去使用这种产品,等等。这些现象表明,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。通过这一现象,当计划推出新产品、传播新思想时,就可以利用称之为“word-of-mouth”或“viral marketing”的网络传播特征,即可以先找到一些具有影响力的个体,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等等^[1]。

本文给出一种新的社会网络的关键节点发现算法,并在两个信息融合模型的基础上利用 Owen 值和 Monte-Carlo 方法得到每个节点的边际贡献,最终得到的 K 个节点即为该问

题的解;在两个合成数据集和 6 个实际数据集上验证了其结果。本文第 2 节形式化描述 CNP 问题;第 3 节提出了一种新的社会网络中求解 CNP 问题的算法;第 4 节对本文算法进行了实验验证和性能讨论;最后给出本文的结论。

2 CNP 问题

给定网络 $G=(V, E)$ 、影响扩散模型 M 以及正整数 k , 关键节点发现问题(Critical Node Problem, CNP)是找到 k 个节点,使得在当前模型 M 下由这些节点所影响的节点数目最大,即找到

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

其中, $\sigma(A)$ 表示 k 节点集合 A 在影响扩散过程完成后受影响节点的期望数; S 表示 G 中任意 k 节点的集合。

Domingos 和 Richardson 最先将关键节点发现作为算法问题来进行研究^[2,3]; 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(见算法 1)使得在 $(1-1/e)$ 范围内近似最优解^[4,5]。

算法 1 贪婪算法 Greedy

1. $A = \emptyset$

到稿日期:2015-03-02 返修日期:2015-05-29 本文受国家自然科学基金项目(111BFX125),上海市浦江人才计划项目,上海市、华政公安学科建设项目资助。

王学光(1975-),男,博士,教授,主要研究领域为计算机网络、信息安全、电子证据, E-mail: wangxueguang@ecupl.edu.cn.

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

算法 1 的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, 以较高的概率来获得近似的结果。假定对于任意节点 v , 模拟 $\sigma(A \cup \{v\})$ 过程 R 次, 每次计算 $\sigma(A)$ 的时间花费为 $O(m)$, 则 Greedy 算法的时间复杂度为 $O(knRm)$ 。

信息在网络上传播的模型已经被广泛研究^[2,6,7]。本文考虑两类基本影响扩散模型: 独立级联模型 (Independent Cascade Model, ICM)^[8] 和线性阈值模型 (Linear Threshold Model, LTM)^[9]。

3 一种基于 Owen 值的算法

3.1 Shapely 值与 Owen 值

给定一个有限参与者集合 N , 可转移效用合作博弈 (Co-operative 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), 从具有不同性质的解中找到唯一的配置分配方案^[10], 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

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

$\forall i \in N$

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^[11], 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 $N = \{1, 2, \dots, n\}$ 和 $M = \{1, \dots, m\}$, 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 $\bigcup_{1 \leq k \leq m} N_k = N$, 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 v_S 定义为

$$v_S(Q) = \begin{cases} v(\bigcup_{k \in Q} N_k), & k \notin Q \\ v(\bigcup_{k \in Q(k)} N_k \cup S), & k \in Q \end{cases}$$

博弈 \bar{v}_k 定义为 $\bar{v}_k(S) = Sh_k(v_S)$, 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 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)^[12], 因此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, 将信息融合视为联盟的形成, 可以定义合适的合作博弈来映射社会网络中的信息融合过程, 那么就可以通过信息融合过程中节点的边际贡献来识别关键节点。

社会网络可由相互连接的参与者对表示, 其结构决定了参与者社会的生产率水平。网络博弈由参与者以及值函数构成, 该值函数为参与者构成的每一个可能网络结构指定实值。在参与者之间根据规则分配由它们产生的实值, 分配规则考虑每一个参与者的边际贡献。将上述值函数看作特征函数, 网络博弈可视为转移效用合作博弈^[13]。

定义合作博弈 (N, v) , N 为社会网络中节点的集合, 特征函数 $v: 2^N \rightarrow \mathbb{R}, 2^N$ 表示 N 的所有子集的集合。对于任一 $S \subseteq N$, 如果 S 中的所有节点被激活, 那么 $v(S)$ 表示融合过程结束后被激活节点的期望值, 也就是说 $v(S) = \sigma(S)$ 。因此, 获得每个节点的边际贡献可以利用合作博弈论中的 Owen 值。由于 Owen 值可以看作是 Shapely 值的两次应用, 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 上所有可能排列的集合 Ψ , 令 $\phi \in \Psi$, 定义 $S_i(\phi)$ 为 ϕ 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

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

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 $O((n/e)^n)$ ^[14]。因此, 本文给出计算 Shapely 值的近似方法。随机产生 t 个概率均等的排列集合 Ψ_t (见算法 2), 令排列 $\phi \in \Psi_t, \phi(i)$ 表示排列中的第 i 个节点。首先, 考虑排列中只有节点 $\phi(1)$ 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 $\phi(1)$ 的贡献。接下来, 考虑 $\phi(2)$, 如果 $\phi(1)$ 激活后 $\phi(2)$ 才被激活, 则 $\phi(2)$ 的贡献为 0, 否则激活 $\phi(2)$, 再确定由它激活的节点数目。依此一直计算到 $\phi(n)$ 。对于任一 $\phi \in \Psi_t$, 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值。为提高该值的精确度, 加入 Jacobson 算法来完成, 其中 α 通过验证取 $0 \sim 1$ 间适当的值。之后, 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 3、算法 4 所示。

算法 2 CREAT_ $\Psi_t(n)$

```

t=random(n)
m=random(n)
Select_t(n)
for i=0 to m do
    arr[i]=random()%(n-1)
for j=0 to i do
    if arr[j] <= arr[i] then
        arr[i]=arr[j]+1

```

```

end
end
v=arr[i]
for k=i-1 to j do
    arr[k+1]=arr[k]
    k=k-1
end
arr[j]=v
end
for h=2 to t do
    |Ψh-1|=arr
    n=n-m
    m=random(n)
    Select_t(n)
end
|Ψh|=arr
算法 3 ShapelyValue(v,R)

```

```

n=|ψ|
t=|CREAT_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_b=v(Sj(ψi)∪{j})-v(Sj(ψi))
            tmp[j]=αtmp[j]+(1-α)tmp_b
        end
    end
end
for i=1 to n do
    ShV[i]=tmp[i]/t
end

```

```

算法 4 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^[15]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, 即 $Role = \{R1, R2, R3, R4, R5, R6, R7\}$ 。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^[16]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 $v_i^h \in V_i$ 指定角色信息 $rv_i^h \in Role$ 。

令 $Role' = \{R3, R6\}, i = 1, \dots, l, j = 1, \dots, l, V' = \{v_i^h, v_j^f \mid \exists (v_i^h, v_j^f), v_i^h \in V_i, rv_i^h \in Role', i \neq j\}, E' = \{e_i^h, (v_i^h, v_j^f) \mid e_i^h = (v_i^{ha}, v_i^{hb}) \in E_i, v_i^{ha} \in V_i, v_i^{hb} \in V_i, rv_i^h \in Role', i \neq j\}$, 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的 Shapely 值, 将属于同一社区的所有节点的 Shapely 值的和作为该社区的支付。通常 l 相对较小, 在使用算法 3 实际计算 G' 中节点的 Shapely 值时可减少计算节点边际贡献的次数 (即算法 3 中参数 R), 以缩短算法运行时间。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, 即为该节点的 Owen 值。应用算法 4 即可得到 k 个关键节点。

3.3 算法计算复杂度

下面考虑基于 Shapely 值获得 k 个关键节点的计算复杂度。算法 3 计算每一节点的边际贡献的复杂度为 $O(t(n+m)R)$; 算法 4 对节点边际贡献排序和选择 k 个节点花费的时间复杂度为 $O(n \log(n) + kn)$ 。因此整个方法的复杂度为 $O(t(n+m)R + n \log(n) + kn)$ 。考虑到在实际网络结构中 $n < m$ 的合理性以及 Owen 值是对 Shapely 值的两次应用, 因此通过 Owen 值方法得到 k 个关键节点的计算复杂度为 $O(tmR)$, 其中 t 与 n 是多项式关系。

4 实验

将分别在 2 个合成数据集与 6 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 4GB 和操作系统为 Windows 7 的个人计算机上运行的, 采用 Matlab 2009 和 Microsoft Visual Studio 2010 作为开发工具。

对比本文算法 (Ov)、基于 Shapely 值的方法 (Sv)、贪婪算法 (Greedy, 即算法 1) 和度启发式方法 (Degree)。Greedy 算法作为衡量其他算法的基准, 选择网络中具有 k 个最大度的节点作为初始集合。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。Ov 和 Sv 算法中 t 的取值如表 1 所列。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。

表 1 本文使用数据集对应的 t 值

Datasets	t
BA	500
FF	500
DBLP	10000
Facebook	5000
Enron	4000
Youtube	3500
AS	1000
Power Grid	500

4.1 合成网络中的性能比较

考察社区结构对本文算法的影响。本文采用 BA 模

型^[17]与“森林火灾”模型(The Forest Fire Model)^[18]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)和直径缩减(Shrinking Diameters)且具有明显社区特征的网络。

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,由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。对这一过程重复 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。在不具明显社区特征的 BA 数据集上利用 Sv 算法和 Ov 算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 Ov 算法则明显优于 Sv 算法(见图 1(b))。在图 1 中,Ov 算法与 Greedy 算法所得到的影响扩散范围基本一致。

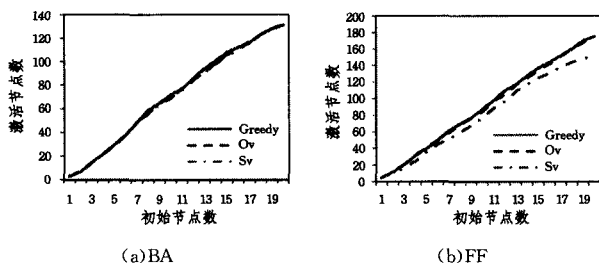


图 1 社区结构对本文算法的影响

4.2 实际网络中的性能比较

本文使用的实际数据集包括 DBLP、Facebook、Enron、Youtube、AS 和 Power Grid 等,其中前 4 个数据集有较为明显的社区结构,而后 2 个数据集的社区结构则不明显。DBLP 数据集^[19]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^[20]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 区域中的朋友关系网络;Enron 数据集^[21]包含 50 万封电子邮件,通过将邮件地址作为节点、两个地址间有邮件通信则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^[22]取 2007 年 1 月 1 日到 2007 年 1 月 15 日的数据来构建网络,由 35468 个节点和 261191 条边组成;AS 数据集^[23]来自于 Internet 上自治系统(Autonomous System)间的拓扑结构,由 11456 个节点和 32759 条边构成;Power Grid 数据集^[24]是由 4941 个节点和 6594 条边构成的北美电力网络。

在 6 个数据集上分别使用 Greedy、Ov、Sv 和 Degree 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,无论是在具有明显社区结构的网络上(见图 2(a)~图 2(d))还是在在不具明显社区结构的网络上(见图 2(e)、图 2(f)),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));而与 Sv 和 Degree 算法相比,本文算法有着较大的优势。

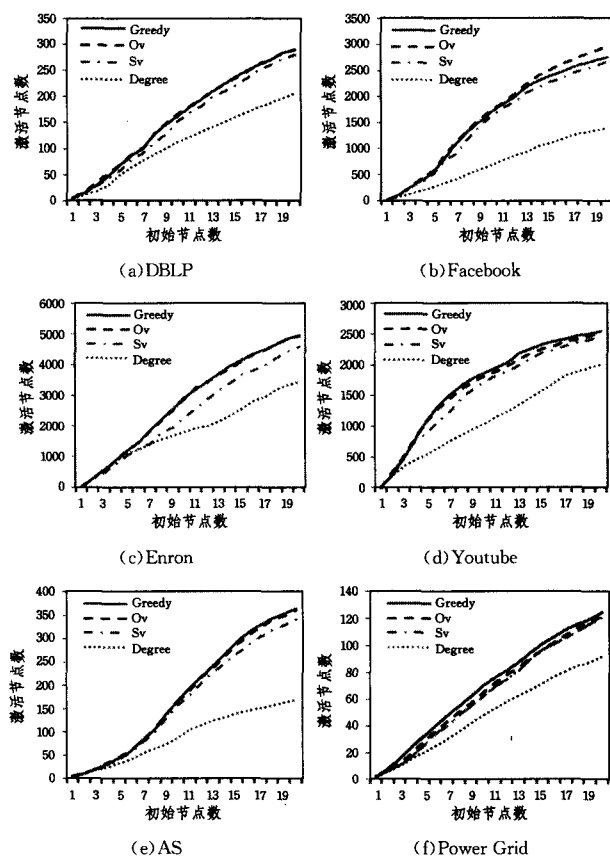


图 2 算法性能分析

4.3 算法时间效率比较

本文使用 FF 模型以 3 组参数(前向概率和后向概率)生成节点都为 1000 的 3 个数据集: $FF_{0.35,0.2}$ 、 $FF_{0.37,0.32}$ 、 $FF_{0.38,0.35}$,它们中的边密度逐渐增加,分别是稀疏图、稠化图和稠密图。考察本文算法与 Greedy 算法在这 3 个数据集上找到 20 个关键节点的运行时间,如图 3 所示。可以看到,在时间效率上 Ov 算法比 Greedy 算法有几十倍的提高。

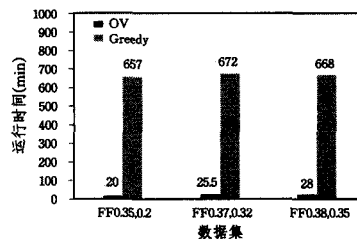


图 3 算法时间效率分析

结束语 考虑社会网络中普遍存在社区结构,基于Owen值,本文提出了一种新的求解 CNP 问题的方法。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 6 个真实数据集上与其他方法的对比结果显示本文方法更为有效;同时,也验证了本文算法较 Greedy 算法在时间效率上亦有几十倍的提高。下一步的工作是研究如何进一步提高获得关键节点的时间效率问题。

参考文献

- [1] He Nan, Li De-yi, Gan Wen-yan, et al. Mining Vital Nodes in Complex Networks[J]. Computer Science, 2007, 34(12): 1-5(in Chinese)

进距离可以自适应地选择较优或次优的邻居节点作为中继节点进行转发。与 OBDR 进行性能比较分析表明, ARPBTD 的平均跳数较小, 平均每跳前进距离较大, 能够使数据分组快速地到达目的节点。

参 考 文 献

- [1] Golmohammadi P, Mokhtarian P, Safaei F, et al. An analytical model of network connectivity in vehicular ad hoc networks using spatial point processes[C]// Proceedings of 2014 IEEE 15th International Symposium on A World of Wireless, Mobile and Multimedia Networks(WoWMoM). Sydney, NSW, 2014; 1-6
- [2] Muhammed Ajeer V K, Neelakantan P C, Babu A V. Network connectivity of one-dimensional networks vehicular ad hoc network[C]// Proceedings of IEEE 2011 International Conference on Communications and Signal Processing(ICCSPP). Calicut, India, 2011; 241-245
- [3] Keykhaie S, Mahmoudifar A. Study of connectivity in a vehicular ad hoc network with random node speed distribution[C]// 6th International Conference on New Technologies, Mobility and Security(NTMS). Dubai, 2014; 1-4
- [4] Rawat D B, Popescu D C, Gongjun Yan, et al. Enhancing VA-

NET performance by joint adaptation of transmission power and contention window size[J]. IEEE Transaction on Parallel and Distributed Systems, 2011, 22(9): 1528-1535

- [5] Soua A, Ben-Ameur W, Afifi H. Broadcast-based Directional routing in vehicular ad-hoc networks[C]// 2012 5th Joint IFIP Wireless and Mobile Networking Conference(WMNC). Bratislava, 2012; 48-53
- [6] Akhtar N, Ergen S C, Ozkasap O. Vehicle Mobility and Communication Channel Models for Realistic and Efficient Highway VANET Simulation[J]. IEEE Transaction on Vehicular Technology, 2015, 64(1): 248-262
- [7] Eiza M H, Owens T, Ni Q. Secure and Robust Multi-Constrained QoS Aware Routing Algorithm for VANETs [J]. IEEE Transaction on Dependable and Secure Computing, 2015(99): 1-13
- [8] Feng Hui-fang, Zhao Liang, Wang Meng-ru. A Routing Algorithm based on Reliability in VANETs[J]. Microelectronic & Computer, 2014, 31(10): 64-68(in Chinese)
冯慧芳, 赵亮, 王梦茹. 一种基于可靠性的车载自组织网络路由算法[J]. 微电子学与计算机, 2014, 31(10): 64-68
- [9] Saleet H, Basir O, Langar R, et al. Region-based location-service-management protocol for VANETs [J]. IEEE Transaction on Vehicular Technology, 2010, 59(2): 917-931

(上接第 50 页)

- 赫南, 李德毅, 淦文燕, 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] 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
- [3] 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
- [4] 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 D C, 2003; 137-146
- [5] Nemhauser G, Wolsey L, Fisher M. An analysis of the approximations for maximizing submodular set functions[J]. Mathematical Programming, 1978, 14(1): 265-294
- [6] Even-Dar E, Shapira A. A note on maximizing the spread of influence in social networks[M]// Internet and Network Economics; Third International Workshop, WINE 2007, San Diego, CA, USA, Dec 12-14, 2007. 2007; 281-286
- [7] 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
- [8] 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
- [9] Granovetter M. Threshold models of collective behavior [J]. American Journal of Sociology, 1978, 83(6): 1420-1443
- [10] 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
- [11] Owen G. Values of games with a priori unions[M]// Henn R, Moeschlin O, eds. Mathematical Economics and Game Theory;

Essays in Honor of Oskar Morgenstern. Springer-Verlag, Berlin, 1977; 76-88

- [12] Scott J. Social Network Analysis: A Handbook (2nd Ed) [M]. Sage: London, 2000
- [13] Caulier J F. Network Games as TU Cooperative Games: The Core, the Shapley Value and Simple Network Games[OL]. http://centres.fusl.ac.be/CEREC/document/seminars/caulier_cerec_feb2009.pdf
- [14] Cormen T H, Leiserson C E, Rivest R L, et al. Introduction to Algorithms(2nd ed)[M]. Cambridge, MA; MIT Press, 2001
- [15] Guimerà R, Amaral L A N. Functional cartography of complex metabolic networks[J]. Nature, 2005, 433: 895-900
- [16] Clauset A, Newman M E J, Moore C. Finding community structure in very large networks[J]. Physical Review E, 2004, 70(6): 066111(6)
- [17] Barabási A L, Albert R. Emergence of scaling in random networks[J]. Science, 1999, 286(5439): 509-512
- [18] 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
- [19] <http://dblp.uni-trier.de>
- [20] 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
- [21] <http://www-2.cs.cmu.edu/~enron>
- [22] 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
- [23] Chen Q, Chang H, Govindan R, et al. The origin of power laws in Internet topologies revisited[C]// The 21st Annual Joint Conference of the IEEE Computer and Communications Societies. Los Alamitos, CA, USA, 2002; 608-617
- [24] Watts D J, Strogatz S H. Collective dynamics of "small-world" networks[J]. Nature, 1998, 393: 440-442