

# 社会网络影响力最大化的核重构算法及传播模型

刘钰峰 鄧欢欢 周喻鑫

(湖南大学信息科学与工程学院 湖南 长沙 410082)

**摘要** 社会网络影响最大化问题是当前的研究热点之一。针对 SI( Susceptible-Infected) 信息传播模型未考虑节点间亲密关系对信息传播的影响,提出一种 ESI( Extended Susceptible-Infected) 信息传播模型。为避免由于挖掘的初始节点之间的距离选择不当,陷入局部最优影响力,提出一种新的启发式算法——核重构算法 CRA( Core Reconstitutions Algorithm)。该算法引入了 k 阶核心集和重合率的概念,通过重合率合理控制初始节点的影响范围,依次找出影响力最优的节点。基于新浪微博的实验表明,ESI 传播模型优于 SI 传播模型,CRA 算法比现有启发式算法具有更优的全局影响效果。

**关键词** 社交网络 信息传播模型 k 阶核心集 重合率 CRA

中图分类号 TP393 文献标识码 A DOI: 10.3969/j.issn.1000-386x.2018.06.051

## CORE RECONSTRUCTION ALGORITHM AND PROPAGATION MODEL FOR MAXIMIZING SOCIAL NETWORK INFLUENCE

Liu Yufeng Zhi Huanhuan Zhou Yuxin

( College of Information Science and Engineering, Hunan University, Changsha 410082, Hunan, China)

**Abstract** The issue of maximizing social network influence is one of the current research hotspots. In view of SI ( Susceptible-Infected) information propagation model does not consider the impact of intimacy between nodes on information dissemination, an ESI ( Extended Susceptible-Infected) information propagation model was proposed. In order to avoid the improper selection of the distance between the initial nodes of the mining, and to fall into the local optimal influence, a new heuristic algorithm—Core Reconstitutions Algorithm ( CRA) was proposed. The algorithm introduced the concept of k-order core set and coincidence rate. Through the coincidence rate, the initial node's influence range was reasonably controlled, and the most influential nodes were found in turn. Experiments based on Sina Weibo show that the ESI propagation model is superior to the SI propagation model, and the CRA algorithm has better global effect than the existing heuristic algorithm.

**Keywords** Social network Information propagation model K-order core set Coincidence rate CRA

## 0 引言

近年来,大规模社会网络(如 facebook、新浪微博、人人网等)迅速发展,这为“口碑”与“病毒营销”提供了机遇,随之也产生了一个关键的问题,即如何从大规模的社会网络中挖掘一些节点,并使这些节点在网络中传播影响最大化<sup>[1-4]</sup>。

在社会网络中,判断初始节点的影响力需要借助相应的传播模型。独立级联传播模型 IC( Independent

Cascade Model) 和线性阈值传播模型 LT( Linear Threshold Model) 是最基本的传播模型,除此之外,在社会网络传播过程中通用的信息传播模型还有 SI、SIR、SIS<sup>[5-7]</sup>等。Gopalan 等<sup>[8]</sup>提出 SI 信息传播模型在社交网络中运用广泛,且效果比较好。Zhang 等<sup>[9]</sup>经过对社会网络结构分析,根据节点的度将节点进行分类,提出了改进的 SI 信息传播模型。Xu 等<sup>[10]</sup>接着提出了 SIS 的信息传播模型。Hacid 等<sup>[11]</sup>基于 SIR 信息传播模型和 SIS 信息传播模型进行影响力的传播预测。本文针对 SI 信息传播模型未考虑节点间亲密关系对信

收稿日期:2017-09-03。刘钰峰,博士,主研领域:社会网络 and 智能搜索引擎。鄧欢欢,硕士生。周喻鑫,硕士生。

息传播的影响,提出了一种 ESI 信息传播模型。

社会网络节点影响力最大化问题的研究由来已久, Richardson 等<sup>[12]</sup>最先将社会网络节点影响力最大化作为一个算法问题提出,选取特定初始节点使影响力最大化。之后,相关研究者<sup>[13-23]</sup>先后提出了不同的算法来解决社会网络节点影响最大化问题。Kempe 等提出了贪心爬山算法<sup>[16]</sup>,影响结果达到最优解的 63%。但该算法运算量大且时间复杂度较高,不适用于大型网络。Goyal<sup>[17]</sup>等对贪心算法改进后提出了一种影响最大化算法 Celf++。Chen 等<sup>[18]</sup>提出了一种启发式算法 DegreeDiscount, 当一个节点被选为初始节点后,那么计算这个节点的邻居节点的度时就打一定的折扣。Kitsak<sup>[19]</sup>等提出了一种基于覆盖的最大核算法和最大度算法。Zhou 等<sup>[20]</sup>基于节点的位置特征提出了一种效率较高的影响最大化算法。Estevez 等<sup>[21]</sup>提出了 SCG(Slow Crack Growth) 算法,选取度大的初始节点,考虑到度大的初始节点会存在重复邻居节点,所以选取的初始节点需要尽可能的分散。曹玖新等<sup>[22]</sup>提出了基于  $k$  核的影响最大化算法 CCA(Core Covering Algorithm), 用一个距离因子  $d$  控制初始节点的距离。Zhou<sup>[23]</sup>基于 SI 信息传播模型,利用社团划分来选择影响最大化的节点(本文称之为 MSI 算法)。对于影响最大化问题,许多研究者提出了时间复杂度相对较低的启发式算法,但其节点影响力效果却不如贪心算法。针对该问题,本文提出了一种影响效率高且时间复杂度较低的启发式算法 CRA。

## 1 信息传播模型

### 1.1 SI 信息传播模型

SI 信息传播模型是最重要的社会网络信息传播模型之一,该模型认为每个节点只有两种状态:影响状态是  $S(i)$  与未影响状态  $I(i)$ ,且节点只能从未影响状态向影响状态单向转变。SI 传播模型基本思想如下:首先得到节点  $t-1$  时的影响状态,根据邻居节点的影响状态与已影响的邻居节点的影响概率  $\alpha$ ,得到所有邻居节点对该节点的影响概率,最后得到  $t$  时该节点的影响状态。

### 1.2 ESI 信息传播模型

在 SI 信息传播模型中,节点与已影响邻居节点间的影响程度用一个常数  $\alpha$  来进行计算,但在真实社会网络节点影响传播过程中,节点间的关系不同,被影响的程度也不同。通常两个节点之间越亲密,就会更加关注对方的消息,节点间的影响概率就会越大。比如

两个相邻节点表示的用户关系比较亲密,那么其中一个用户发的消息就可能很快影响另一个用户,而对于相邻的是个几乎没有联系的用户,那么就不一定会被影响,所以在信息传播的过程中,要考虑节点间的这种影响因素。基于此,本文提出了 ESI 信息传播模型。

#### 1.2.1 节点影响过程

本文引入节点间亲密度的概念表示节点间互动的亲密程度,用  $A_{ij}$  表示节点  $i$  和节点  $j$  的亲密度,则节点间亲密度计算如下:

$$A_{ij} = \frac{S_i \cap S_j}{S_i \cup S_j} \quad i \in V \quad j \in Z_i \quad (1)$$

式中:  $S_i$  表示该节点  $i$  的消息集合,  $S_j$  表示节点  $i$  的邻居节点  $j$  的消息集合,  $Z_i$  表示节点  $i$  的邻居节点集合,用  $e_{ij}$  表示邻居节点  $j$  对节点  $i$  的影响概率,则其计算如下:

$$e_{ij} = c\alpha + (1-c)A_{ij} \quad i \in V \quad j \in Z_i \quad (2)$$

式中:假设  $\alpha$  是节点起初可能被已影响节点影响的概率,属于  $[0, 1]$ 。  $c$  表示一个调和因子,属于  $[0, 1]$ 。

以下考虑节点被影响的过程,本文使用  $P_{it}$  表示节点  $t$  时刻被影响的概率。 $P_{it}$  主要与三个因素有关:1) 该节点的邻居节点  $t-1$  时刻到  $t$  时刻对其影响的概率;2) 该节点  $t-1$  时刻的影响概率;3) 该节点的邻居节点的影响状态。

本文使用  $\chi_{it}$  来表示节点  $i$  在  $t-1$  时刻到  $t$  时刻未被已影响的邻居节点  $j$  影响的概率:

$$\chi_{it} = ((1 - e_{ij}) p_{j,t-1}) \quad i \in V \quad j \in Z_i \quad (3)$$

邻居节点  $j$  在  $t-1$  时刻未影响的概率为  $(1 - p_{j,t-1})$ 。

$\delta_{it}$  用来表示节点  $i$  在  $t-1$  时刻到  $t$  时刻未被邻居节点  $j$  影响的概率为:

$$\delta_{it} = \chi_{it} + (1 - p_{j,t-1}) = ((1 - e_{ij}) p_{j,t-1}) \quad i \in V \quad j \in Z_i \quad (4)$$

节点  $i$  存在的邻居节点  $j$  有多个,所以用  $\beta_{it}$  来表示所有邻居节点  $j \in Z_i$  从  $t-1$  时刻到  $t$  时刻未能影响节点  $i$  的概率为:

$$\beta_{it} = \prod_{j \in Z_i} \delta_{it} = \prod_{j \in Z_i} ((1 - e_{ij}) p_{j,t-1} + (1 - p_{j,t-1})) \quad i \in V \quad (5)$$

相反,用  $\phi_{it}$  来表示节点  $i$  在  $t-1$  时刻到  $t$  时刻被它的邻居节点影响的概率为:

$$\phi_{it} = 1 - \beta_{it} = 1 - \prod_{j \in Z_i} ((1 - e_{ij}) p_{j,t-1} + (1 - p_{j,t-1})) \quad i \in V \quad (6)$$

节点  $i$  在  $t-1$  时刻未被影响的概率为  $(1 - p_{i,t-1})$ ,节点  $i$  从  $t-1$  时刻到  $t$  时刻未被影响的概率为

$(1 - \varphi_{i,t})$ , 所以节点  $t$  时刻被影响的概率为:

$$p_{i,t} = 1 - (1 - p_{i,t-1})(1 - \phi_{i,t}) \quad i \in V \quad (7)$$

由式(6)代入式(7)可得:

$$p_{i,t} = 1 - (1 - p_{i,t-1})(1 - (1 - \prod_{j \in Z_i} ((1 - e_{i,j}) p_{j,t-1} + (1 - p_{j,t-1}))) = 1 - (1 - p_{i,t-1})(\prod_{j \in Z_i} (1 - e_{i,j} p_{j,t-1})) \quad i \in V \quad (8)$$

将  $\prod_{j \in Z_i} (1 - e_{i,j} p_{j,t-1})$  进行分解得到:

$$\begin{aligned} \prod_{j \in Z_i} (1 - e_{i,j} p_{j,t-1}) &= 1 - \sum_{j \in Z_i} (e_{i,j} p_{j,t-1}) + \sum_{j_1, j_2 \in Z_i} (e_{i,j_1} p_{j_1,t-1} \cdot e_{i,j_2} p_{j_2,t-1}) - \\ &\quad \sum_{j_1, j_2, j_3 \in Z_i} (e_{i,j_1} p_{j_1,t-1} \cdot e_{i,j_2} p_{j_2,t-1} \cdot e_{i,j_3} p_{j_3,t-1}) + \dots \quad i \in V \end{aligned} \quad (9)$$

式(9)后面的值非常小,可忽略不计,得到:

$$\prod_{j \in Z_i} (1 - e_{i,j} p_{j,t-1}) = 1 - \sum_{j \in Z_i} (e_{i,j} p_{j,t-1}) \quad i \in V \quad (10)$$

由式(8)和式(10)得到:

$$p_{i,t} = 1 - (1 - p_{i,t-1})(1 - \sum_{j \in Z_i} (e_{i,j} p_{j,t-1})) = p_{i,t-1} + \sum_{j \in Z_i} (e_{i,j} p_{j,t-1}) - p_{i,t-1} \sum_{j \in Z_i} (e_{i,j} p_{j,t-1}) \quad i \in V \quad (11)$$

同理,  $p_{i,t-1} \sum_{j \in Z_i} (e_{i,j} p_{j,t-1})$  的值非常小,舍去,简化得:

$$p_{i,t} = p_{i,t-1} + \sum_{j \in Z_i} (e_{i,j} p_{j,t-1}) \quad i \in V \quad (12)$$

式中:  $\sum_{j \in Z_i} (e_{i,j} p_{j,t-1})$  表示节点  $i$  的所有邻居节点影响概率。节点  $i$  与其邻居节点  $j$  间的边用  $b_{i,j}$  表示。在社会网络中,如果两节点为邻居节点,则存在边  $b_{i,j}$ ,值为1,否则为0。得到:

$$p_{i,t} = p_{i,t-1} + \sum_{j \in V} (e_{i,j} b_{i,j} p_{j,t-1}) \quad i \in V \quad (13)$$

式(13)表示了一个节点从  $t-1$  到  $t$  影响概率变化的过程,也反映了节点的影响状态的变化。

### 1.2.2 模型框架

用向量  $P_t = [p_{1,t}, p_{2,t}, \dots, p_{n,t}]$  表示所有节点  $i \in V$  在  $t$  时刻的影响概率,  $p_{i,t} = 1$  表示节点  $i$  在  $t$  时刻已被影响,  $p_{i,t} = 0$  表示节点  $i$  在  $t$  时刻未被影响。用矩阵  $E$  表示节点与邻居节点影响的概率矩阵,表示如下:

$$E_{ij} = \begin{cases} e_{i,j} b_{i,j} & i \in V, j \in Z_i \\ 0 & i \in V, j \notin Z_i \end{cases} \quad (14)$$

由传播过程得:

$$P_t = P_{t-1} + (EP_{t-1})^T \quad (15)$$

在  $t$  足够大时,节点的概率可能大于1,所以运用

控制概率函数如下:

$$G([x_1, x_2, \dots, x_n]) = [\min\{x_1, 1\}, \min\{x_2, 1\}, \dots, \min\{x_n, 1\}] \quad (16)$$

将式(14)代入式(15)得到:

$$P_t = G(P_{t-1} + (EP_{t-1})^T) \quad (17)$$

式(17)主要表示了ESI信息传播模型中所有节点随时间被影响的变化过程,在  $t=0$  时,初始节点的影响概率为1,通过任意时间  $t$ ,可得到整个网络节点的影响状态。

社会网络影响最大化问题除了借助相应传播模型外,最重要的是初始节点的挖掘。

## 2 基于 $k$ 阶核心集的影响最大化算法

### 2.1 $k$ 阶核心集

不同于曹玖新等<sup>[22]</sup>提出的  $k$  核,本文提出  $k$  阶核心集的概念。对网络中的每一个节点,其  $k$  阶核心集的定义是与之  $k$  级关联的节点集合,而  $k$  核是度值不少于  $k$  的节点推出的最大连通子图。以下是具体定义:

**定义1**  $k$  阶。在图  $G=(V, S)$  社交网络拓扑结构中,节点  $i(i \in V)$  会存在直接相连的节点,本文将其直接相连的节点称为第一阶节点,将它的间接节点称为第二阶节点,以此类推可以得到  $k$  阶的节点。

**定义2**  $k$  阶核心集。在图  $G=(V, S)$  社交网络拓扑结构中,根据节点级联特征可以得到节点  $i(i \in V)$  的  $k$  阶所有节点,即  $k$  阶核心集节点用  $S^k(i)(i \in V)$  表示。选取初始节点时,节点核心集需满足  $S^1(i) > ((\bigcup_{i \in V} S^k(i))/k)$ ,  $k' \in \{1, 2, \dots, k\}$ , 节点  $i$  为  $k$  阶核心集的核心节点。

如图1所示一个网络节点无向图,节点1的一阶节点是4、8、21,二阶节点是2、5、7、12、13、14、16,三阶节点是9、11、15、18。当  $k=2$  时,可以得到  $S^2(1) = \{4, 8, 21, 2, 5, 7, 12, 13, 14, 16\}$ 。

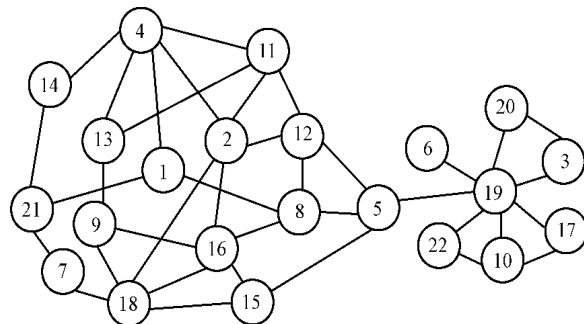


图1 网络节点关系示意图

**定义3** 重合率定义为节点  $i$  与邻居节点  $j$  的  $k$  阶共同节点集合与两节点  $k$  阶核心集占比,使用  $P(i, j)$

表示如下:

$$P(i, j) = \frac{(\bigcup_{i \in M} S^k(i)) \cap S^k(j)}{(\bigcup_{i \in M} S^k(i)) \cup S^k(j)} \quad i \in M \quad j \in V \quad (18)$$

式中:  $M$  表示已选择的初始节点集合,  $\bigcup_{i \in M} S^k(i)$  表示已选择的初始节点的  $k$  阶核心集。

重合率的范围控制初始节点的影响状况  $P$  太高, 易局部影响最大化;  $P$  太低, 初始节点  $k$  阶核心集边缘节点的连通性不好, 不易被影响。所以本文提出了合适的重合率计算方法如下:

$$P(i, j) \in (\alpha, \beta) \quad \alpha, \beta \in (0, 1) \quad (19)$$

$$\alpha = \frac{S^k(j)/2}{(\bigcup_{i \in M} S^k(i)) \cup S^k(j)} \quad i \in M \quad j \in V \quad (20)$$

$$\beta = \sqrt{\frac{S^k(j)}{(\bigcup_{i \in M} S^k(i)) \cup S^k(j)}} \quad i \in M \quad j \in V \quad (21)$$

本文讨论的 CRA 算法主要根据节点  $k$  阶核心集初步选取, 再通过重合率确定。当  $k$  为 1, 通过重合率范围得到的新的初始节点一般处于上一个初始节点一阶节点附近, 这样得到的初始节点集没有足够的扩散, 进而可能造成局部影响最大化, 所以一般  $k$  从 2 开始选取。

## 2.2 CRA 算法

CRA 算法的基本思想简述如下: 1) 根据  $k$  阶核心集, 选择核心集数量大且一阶核心集不小于平均每阶核心集的节点; 2) 将选择满足条件的节点进行标记, 并将其一阶核心集进行标记, 表示不能再被选择; 3) 重合率范围判断, 将新选节点与所有已选初始节点进行重合率计算。重合率的控制给各初始节点的  $k$  阶核心集边缘节点提供了多次被影响的机会。

**算法 1** 基于核重构的社会网络影响最大化算法

输入: 社会网络节点  $G = (V, S)$ , 节点个数  $N$ , 初始节点个数  $m$

输出: 初始节点集合  $M$

1. Initialize  $M = \emptyset$
2.  $Get(MaxS^k(x_1))$  and  $(S^1(x_1) > ((\bigcup_{i \in V} S^k(x_1))/k))$
3.  $x_1 \in M$  And  $mark(x_1, C(x_1))$
4.  $get(MaxS^k(x_2))$  and  $(S^1(x_2) > ((\bigcup_{i \in V} S^k(x_2))/k))$
5. If  $x_2 \in (Cv(x_1 \cup C(x_1)))$
6. Calculate  $P(x_1, x_2)$ , Calculate  $\alpha$ , Calculate  $\beta$
7. If  $P(x_1, x_2) \in (\alpha, \beta)$
8.  $x_2 \in M$  And  $mark(x_2, C(x_2))$
9. Else
10. mark  $x_2$  And Return step 4

11. End If

12. End If

该算法中的  $get(MaxS^k(x_i))$  表示选择出  $k$  阶核心集数量最大的节点,  $mark(x_1, C(x_1))$  表示选择满足条件的节点进行标记, 并将其一阶核心集进行标记。  $Cv(x_1 \cup C(x_1))$  表示节点  $V$  中未标记节点。

## 2.3 时间复杂度分析

CRA 算法节点核心集比较过程是对每一个节点的  $k$  阶核心集进行对比, 故时间复杂度为  $O(n)$ ; 节点标记的过程是只对选择的初始节点的一阶的核心集进行标记, 故时间复杂度小于  $O(n)$ ; 重合率计算与范围比较过程是只对选择的初始节点进行比较, 所以 CRA 算法的时间复杂度是  $O(n)$ 。

## 3 实验设计

### 3.1 数据集

本文通过新浪微博开放的 API 接口<sup>[24]</sup>, 利用网络数据挖掘器挖掘出部分用户和对应的粉丝, 以及用户转发的微博。通过爬虫程序预处理总共得到 11 080 个节点, 1 288 045 条边, 以及节点在 2014 年 8 月 1 日到 12 月 31 日对应转发的微博消息 806 854 条。

### 3.2 实验对比

为了验证 CRA 算法, 与其他算法进行实验对比。对比实验算法如表 1 所示。

表 1 对比算法

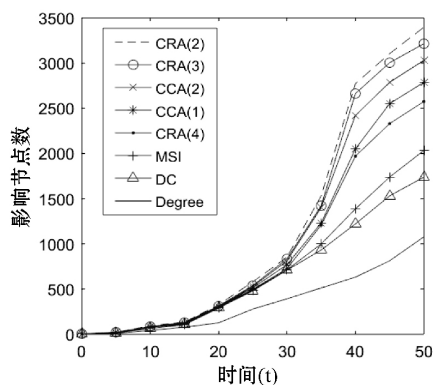
算法	算法描述
Degree	选取 $m$ 个度值最大节点的算法
Degree Cover <sup>[19]</sup>	初始节点的邻居节点进行标记的启发式算法, 简称 DC
MSI <sup>[23]</sup>	基于社团划分选取最终影响数量最多的 $m$ 个初始节点
CCA(d) <sup>[22]</sup>	通过核数层次和距离因子得 $m$ 个初始节点的启发式算法 $d$ 设为 1 和 2
CRA(k)	本文提出的根据 $k$ 级关联节点和重合率 $p$ 得 $m$ 个初始节点的启发式算法 $k$ 设为 2、3、4

### 3.3 结果分析

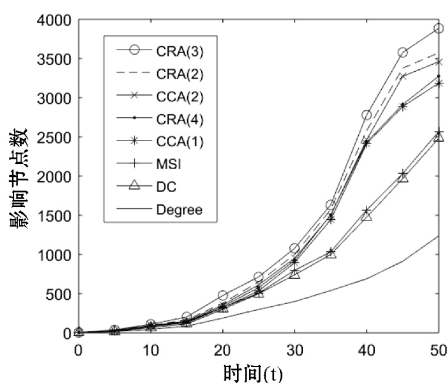
本文的实验结果评价指标可概括为以下两点: (1) 影响范围。通过相同的时间, 节点最终被影响的数量。(2) 算法的运行时间。如何用较短的时间找到初始节点。在传播影响节点图中,  $x$  轴表示信息传播

模型中的时间步  $t$ ,  $y$  轴表示节点的影响数量。

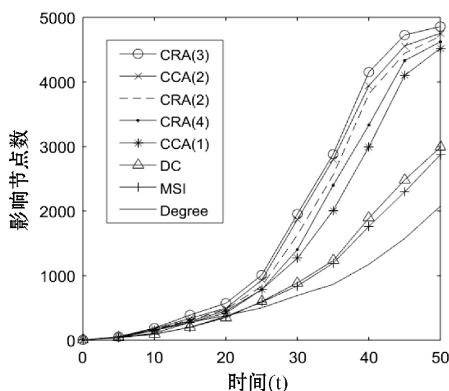
本文将初始节点  $m$  分别为 5 和 10 进行实验。在 SI 信息传播模型和 ESI 信息传播模型上,  $\alpha$  分别为 0.01、0.02 和 0.03。实验结果如图 2、图 3 所示。



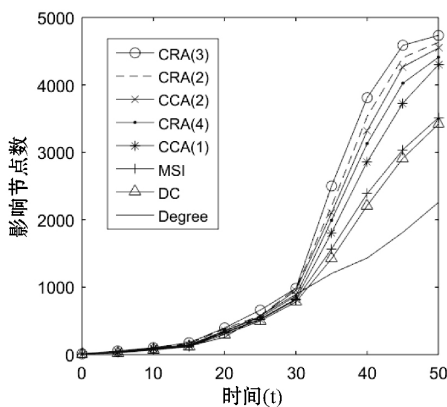
(a)  $m=5$   $\alpha=0.01$



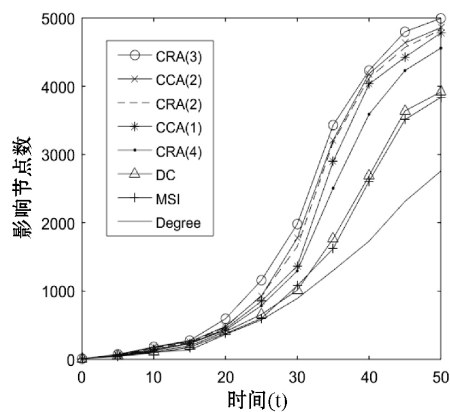
(b)  $m=5$   $\alpha=0.02$



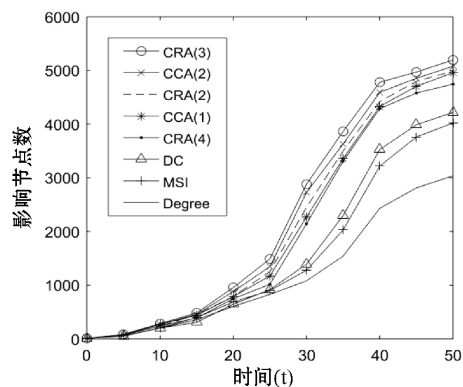
(c)  $m=5$   $\alpha=0.03$



(d)  $m=10$   $\alpha=0.01$

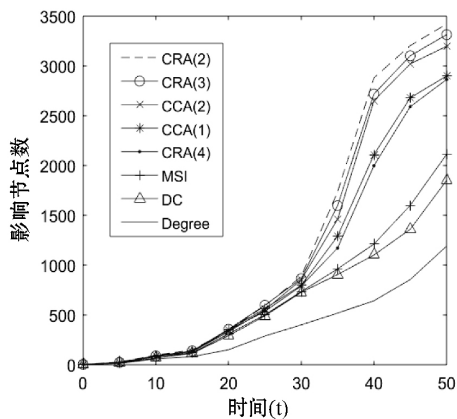


(e)  $m=10$   $\alpha=0.02$

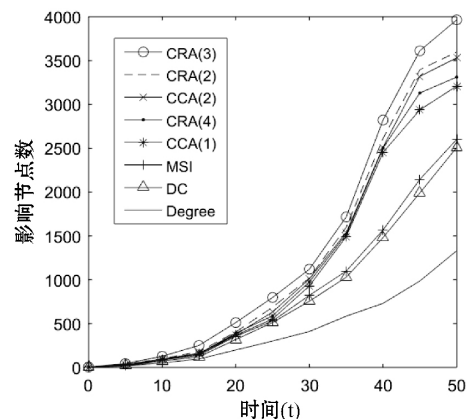


(f)  $m=10$   $\alpha=0.03$

图2 SI 信息传播模型上各算法的影响效果



(a)  $m=5$   $\alpha=0.01$



(b)  $m=5$   $\alpha=0.02$

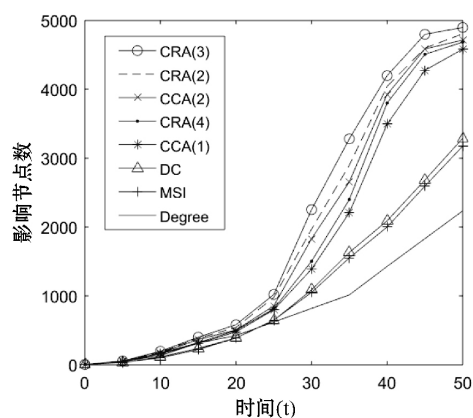
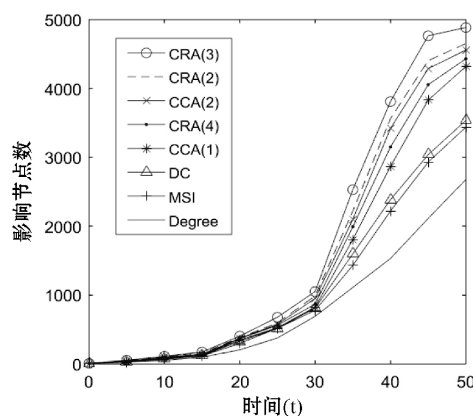
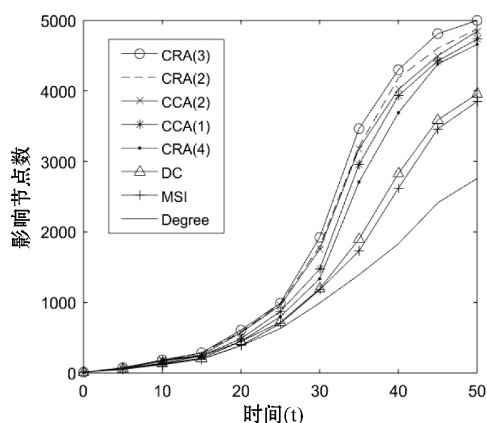
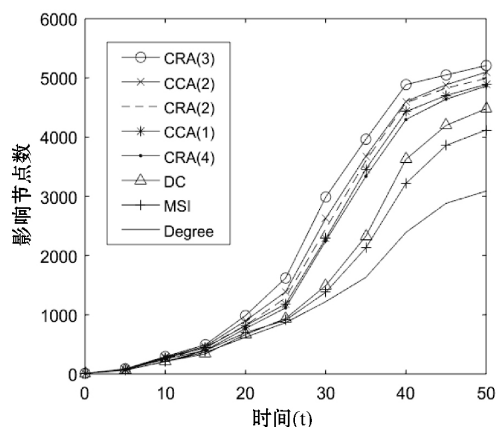
(c)  $m=5$   $\alpha=0.03$ (d)  $m=10$   $\alpha=0.01$ (e)  $m=10$   $\alpha=0.02$ (f)  $m=10$   $\alpha=0.03$ 

图3 ESI 信息传播模型上各算法的影响效果

从图2和图3的对比可以看出,算法在 ESI 信息传播模型上节点影响效果要优于 SI 信息传播模型,其中 CRA、CCA、Degree 算法比较明显,这表明节点间的亲密关系对节点的传播也起很重要的作用。而且本文提出的 CRA 算法比其他四种算法的最终的影响效果明显更好。在影响数量上,相同的初始节点  $m$  和影响概率  $\alpha$  情况下, CRA 的最终影响节点数量都是最多的。在影响速度上, CRA 算法影响节点速度也是最快的。如图2(b-f)和图3(b-f)所示, CRA(3)比 CCA(2)算法的性能更好,这充分证明考虑节点的  $k$  级节点关联特性比节点的层次性更重要。如图2(e-f)和图3(e-f)所示, DC 算法的影响效果优于 MSI 算法,同时 Degree 算法影响效果变的趋于平缓。这表明节点的影响最大化过程一定要考虑初始节点的距离因素。在本文实验中, CRA(2)、CRA(3)比 CRA(4)算法的影响效果更优。这主要是因为 CRA(4)算法考虑的是四阶邻居节点特征,在节点的传播过程中,离初始节点远的节点不易被影响。一旦存在未影响的节点,那么这些未影响节点的邻居节点更不易影响。虽然重合率保证存在一些重复的边缘节点,使它们可以多次被影响,但是并不能保证这些重复节点一定被影响。所以并非阶数  $k$  越大, CRA 算法性能越好。阶数  $k$  需要根据总节点数适宜选取。

## 4 结 语

本文针对社交网络节点影响最大化问题,根据节点的  $k$  阶核心集与重合率,提出了 CRA 算法。针对 SI 信息传播模型未考虑邻居节点间的亲密度的问题,对其进行改进得出 ESI 传播模型。分别在 SI 信息传播模型和 ESI 信息传播模型进行实验,与其他几种算法进行比较,得到如下结论:1) 所有算法在 ESI 信息传播模型下,节点影响效果更好。2) 在不同的传播概率下, CRA 算法的影响效果都是最好的,传播过程中影响节点的速度也是最快的。在接下来的研究中,一方面是通过机器学习的方法来更加准确地预测节点间的影响概率。另一方面是推算数据集节点总数与  $k$  阶核心集特性对  $k$  值的影响,并推导使节点影响最大化的  $k$  值。

## 参 考 文 献

- [1] Chen W, Wang Y, Yang S. Efficient influence maximization in social networks[C]// ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2009: 199-208.

- [2] Zhang H, Dinh T N, Thai M T. Maximizing the Spread of Positive Influence in Online Social Networks [C]// IEEE, International Conference on Distributed Computing Systems. IEEE, 2013: 317-326.
- [3] Jiang B, Hegde N, Massoulie L, et al. How to Optimally allocate your budget of attention in social networks [C]// INFOCOM 2013 Proceedings IEEE. IEEE, 2013: 2373-2381.
- [4] Borgs C, Brautbar M, Chayes J, et al. Maximizing social influence in nearly optimal time [C]// ACM-Siam Symposium on Discrete Algorithms. Society for Industrial and Applied Mathematics, 2014: 946-957.
- [5] Wei Z, Ye Y, Tan H, et al. Information Diffusion Model Based on Social Network [C]// Proceedings of the 2012 International Conference of Modern Computer Science and Applications. Springer Berlin Heidelberg, 2013: 145-150.
- [6] Xu B, Liu L. Information diffusion through online social networks [C]// IEEE International Conference on Emergency Management and Management Sciences. IEEE, 2010: 53-56.
- [7] Lu Z, Wen Y, Cao G. Information diffusion in mobile social networks: The speed perspective [C]// INFOCOM, 2014 Proceedings IEEE. IEEE, 2014: 1932-1940.
- [8] Gopalan A, Banerjee S, Das A K, et al. Random mobility and the spread of infection [C]// INFOCOM, 2011 Proceedings IEEE. IEEE, 2011: 999-1007.
- [9] Zhang H, Dinh T N, Thai M T. Maximizing the Spread of Positive Influence in Online Social Networks [C]// IEEE, International Conference on Distributed Computing Systems. IEEE, 2013: 317-326.
- [10] Xu B, Liu L. Information diffusion through online social networks [C]// IEEE International Conference on Emergency Management and Management Sciences. IEEE, 2010: 53-56.
- [11] Hacid H, Hacid H. A predictive model for the temporal dynamics of information diffusion in online social networks [C]// International Conference on World Wide Web. ACM, 2012: 1145-1152.
- [12] Richardson M, Domingos P. Mining knowledge-sharing sites for viral marketing [C]// Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2002: 61-70.
- [13] Zhang Wei, Ye Yanqing, Tan Hanlin, et al. Information Diffusion Model Based on Social Network [C]// Proceedings of the 2012 International Conference of Modern Computer Science and Applications. Springer Berlin Heidelberg, 2013: 145-150.
- [14] Tang Y, Xiao X, Shi Y. Influence maximization: near-optimal time complexity meets practical efficiency [C]// SIGMOD'14 Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data. 2014: 75-86.
- [15] Kimura M, Saito K. Approximate Solutions for the Influence Maximization Problem in a Social Network [C]// Knowledge-Based Intelligent Information and Engineering Systems, International Conference, Kes 2006, Bournemouth, UK, October 9-11, 2006, Proceedings. DBLP, 2006: 937-944.
- [16] Kempe D, Kleinberg J. Maximizing the spread of influence through a social network [C]// ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2003: 137-146.
- [17] Goyal A, Lu W, Lakshmanan L V S. CELF++: optimizing the greedy algorithm for influence maximization in social networks [C]// International Conference Companion on World Wide Web. ACM, 2011: 47-48.
- [18] Chen W, Lu W, Zhang N. Time-Critical Influence Maximization in Social Networks with Time-Delayed Diffusion Process [J]. Chinese Journal of Engineering Design, 2012, 19(5): 340-344.
- [19] Kitsak M, Gallos L K, Havlin S, et al. Identification of influential spreaders in complex networks [J]. Nature Physics, 2011, 6(11): 888-893.
- [20] Zhou T, Cao J, Liu B, et al. Location Based Influence Maximization in Social Networks [C]// ACM International Conference on Information and Knowledge Management. ACM, 2015: 1211-1220.
- [21] Estevez P A, Vera P, Saito K. Selecting the Most Influential Nodes in Social Networks [C]// International Joint Conference on Neural Networks. IEEE, 2007: 2397-2402.
- [22] 曹玖新, 董丹, 徐顺, 等. 一种基于k-核的社会网络影响最大化算法 [J]. 计算机学报, 2015, 38(2): 238-248.
- [23] Zhao J. Initial spreaders in online social networks [C]// Communication, Control, and Computing. IEEE, 2017: 180-186.
- [24] <http://open.weibo.com/>.
- ~~~~~
- (上接第205页)
- [12] 蔡伯峰, 蔡伟达, 王宜怀. KW01-ZigBee 无线传感网应用开发平台的研制 [J]. 电子技术应用, 2017, 43(3): 55-56, 58.
- [13] 王志超, 刘波, 花於锋. 基于移动4G与ZigBee无线传感网的网关设计 [J]. 计算机测量与控制, 2014, 22(3): 863-865.
- [14] 张建国. 中国移动NB-IoT部署策略研究 [J]. 移动通信, 2017, 41(1): 25-30.
- [15] 岑荣滢, 姜琴, 扈健玮, 等. 面向智能家居应用的ZigBee-WiFi网关 [J]. 计算机系统应用, 2017, 26(1): 232-235.