

# 第五届数学中国数学建模网络挑战赛

---

#1465

## 第五届“认证杯”数学中国

### 数学建模网络挑战赛 承 诺 书

我们仔细阅读了第五届“认证杯”数学中国数学建模网络挑战赛的竞赛规则。

我们完全明白，在竞赛开始后参赛队员不能以任何方式（包括电话、电子邮件、网上咨询等）与队外的任何人（包括指导教师）研究、讨论与赛题有关的问题。

我们知道，抄袭别人的成果是违反竞赛规则的，如果引用别人的成果或其他公开的资料（包括网上查到的资料），必须按照规定的参考文献的表述方式在正文引用处和参考文献中明确列出。

我们郑重承诺，严格遵守竞赛规则，以保证竞赛的公正、公平性。如有违反竞赛规则的行为，我们将受到严肃处理。

我们允许数学中国网站([www.madio.net](http://www.madio.net))公布论文，以供网友之间学习交流，数学中国网站以非商业目的的论文交流不需要提前取得我们的同意。

我们的参赛队号为：1465

参赛队员（签名）：

队员 1：李 哲

队员 2：孙立环

队员 3：裴学伟

参赛队教练员（签名）：姜玉山

参赛队伍组别：本科组

# 第五届数学中国数学建模网络挑战赛

---

#1465

第五届“认证杯”数学中国

数学建模网络挑战赛

编号专用页

参赛队伍的参赛队号：（请各个参赛队提前填写好）：

1465

竞赛统一编号（由竞赛组委会送至评委团前编号）：

---

竞赛评阅编号（由竞赛评委团评阅前进行编号）：

# 第五届数学中国数学建模网络挑战赛

#1465

## 2012 年第五届“认证杯”数学中国 数学建模网络挑战赛

题 目 社交网络广告营销的信息传播分析

### 及其费用的优化模型

关 键 词 社交网络 信息传播 小世界模型 无标度网络 多目标函数优化

#### 摘 要:

本文针对社交网络的广告营销问题,根据Twitter 社交网站用户之间的链接关系(follow 关系)数据为基础,分析Twitter 用户的链接分布。依照不同的假设建立了三个模型,分别解决了两个问题。前两个模型是为解决第一个问题而建立的,即探究奥运会开始后,一条含有企业广告的奥运会新闻能够被多少人获知,这为社交网络的广告营销战略奠定了基础。第三个模型解决了在能够达到理想的广告覆盖率的同时用人成本最低的最优化问题,为日后企业在研究社交网络的广告营销的信息传播遍及度以及其成本问题之间的权量提供了模板。

数据处理:首先运用 office2010 (excel) 处理了 Twitter 社交网站用户之间的链接关系数据,得到一个 835541\*2 的电子表格。再运用数据透视表处理得出表示用户数及其粉丝数之间的关系,将其导入 matlab 中构建出了 twitter 上用户与用户之间的关系模型。并计算出了关系的重复率为 42%。

模型一:将广告作为一种信息在网络上传播,且满足简易微分模型。然后通过进一步分析可知,信息传播强度  $k$  或总人数  $m$  增加时,  $t_1$  都会变小与实际情况相吻合。之后我们对该模型中部分用户对信息无法确定归属即可能具有免疫性的情形进一步分析,对以前的模型进行了加权处理,较好地弥补了模型一的不足。

模型二鉴于现实世界的网络有着较小的平均距离和较高的聚集系数,为弥补模型一的不足,我们引入了 NW 小世界模型,然后再根据 NW 小世界网络模型的构造算法,用 matlab 进行仿真实现。网络的连接也用邻接矩阵  $a_{nm}$  表示。而又由于网络中每天都有新的加入者,信息的引用也在不断地增长,故边的加入不是随机的,而是有选择的,并且考虑重复率。所以我们又引入了 BA 模型,根据对其传播概率分析代入 matlab 仿真分析可以得到一个较接近现实的数值。

综上可得出一条含有企业广告的奥运会新闻在奥运会期间可以被约44万人观看到。

模型三中我们为了使企业广告的信息推广率及其成本同时达到最优,考虑到时间问题,故在问题一的仿真模型的基础之上加入了时间变量以及信息新鲜度,对其进行多目标函数最优化分析,使用 lingo 软件从而求出了最优解,为 109 个专业社交网络推广和 15 个兼职宣传者。

参赛队号 1465

所选题目 C 题

参赛密码 \_\_\_\_\_  
(由组委会填写)

# 第五届数学中国数学建模网络挑战赛

---

#1465

## Abstract

This paper concentrates on the problem of advertising and marketing for social networking, to analyze the link distribution of Twitter users, which is based on the data of the link relationship (follow relationship) of the Twitter social networking site users. We establish three models to solve two problems based on different assumptions. The first two models are established to solve the first problem, namely to explore the number of people informed with Olympic News with company's ads in it after the Olympic Games, which laid the foundation for social network advertising and marketing strategy. The third model solves the optimization problem that how to achieve the desired coverage of advertising with lowest cost of employing, which provides a model to deal with how to balance the popularity of the social network's advertising and marketing information and its costs for the enterprises.

**Data processing:** firstly we deal with the link relationships of Twitter social networking site users with the help of office2010 (excel), getting a  $835541 * 2$  electronic form. Then we get the relationship of the users and its fans using the data pivot table, and import it into matlab to construct the relation model between users on Twitter. Then calculated the repeat rate of relationship.

**Model one:** advertising on the network as an information dissemination, and satisfies the simple differential model. And then the further analysis shows that with the increase in information dissemination strength  $k$  or the total number of  $m$ ,  $t_1$  will become smaller, which is consistent with the actual situation. We can see the arrival of the peak number of Olympic advertising  $t_1$  simultaneously, which is useful for development and forecast of the company. Then we make a further analysis of the situation where some users unable to determine the attribution of information, namely it may has immunity. Here we weigh the previous model to compensate for deficiencies in model one.

**Model two:** in view of the real-world network has a smaller average distance and high clustering coefficient, we introduce the NW small-world model to compensate for deficiencies in model one. And then using matlab simulation to achieve it, which is based on the construction algorithm of NW small-world network model. The connection of the network is presented by adjacency matrix  $ann \times n$ . Due to the every-day new entrants on the internet, references to information are constantly growing, fans join in frequently. Secondly, the join of edge is selective, not random, and should take the repetition rate into consideration. Thus there will be a scale-free network over time. We introduce the BA model, According to the analysis of the spread of the probability and computer simulation, we can get a value which is closer to the reality.

**Model three:** In order to make the rates and costs of corporate advertising information achieve optimization at the same time, we add time variable as well as information freshness base on the simulation in model one, taking into account a matter of time. We make multi-objective function optimization analysis, using the lingo software to find the optimal solution.

# 第五届数学中国数学建模网络挑战赛

#1465

## 1 问题重述

信息的人际传播实质上是信息在人际网络上扩散的过程。在信息的口口相传时代,无疑人际传播在信息传播中扮演了重要角色。而在信息科技高度发达的今天,Internet的出现,使得人际关系已经被复制到了虚拟空间中,网络信息传播的优点:(1)极大丰富(2)形态多样(3)传递迅速及时(4)全球传播(5)互动交流。由于网络信息传播的优点,网络广告也越来越受广告主欢迎。在数字化背景下,越来越多的企业通过社交网络在奥运会期间进行企业宣传。同时,受众获得奥运信息的途径不再仅仅通过一个电视频道;而且新渠道因为用户生成内容,因为个性化的社交网络,被分割成了无数个屏幕,使奥运营销大战实现了从一个屏到85亿屏的转变。

随着奥运会的发展和影响,企业看到了借助奥运传播广告潜力。通过建立模型,预测一条企业广告的奥运会新闻在奥运开始后借助社交网络传播能被多少人看到。在离奥运开幕100天之前,一家大型企业通过网络进行广告宣传,可以雇佣专业推广者和兼职宣传者,考虑价格和粉丝增长率,给出合理方案,实现广告覆盖率为潜在用户的40%,并且使成本最低。

## 2 问题初步分析

对于问题一:广告作为一种信息在网络上传播时,人的接受和传播信息受到两种因素的影响,一是由其自身特征所决定内部力量,二是由信息本身的吸引力或确信度及其周围人影响力构成的外部力量。通过构建包含网络和人接受广告信息的影响因素这两个方面要素的信息传播模型,并在计算机上进行模拟。

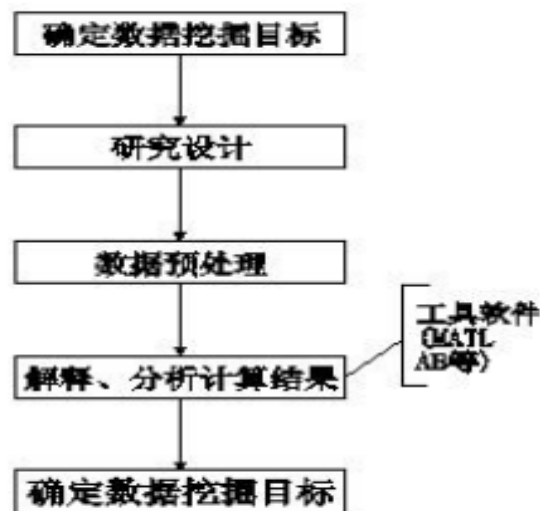
对于问题二:这是一个多目标函数优化问题,通过多目标函数优化算法进行约束,并借助lingo进行求解。

## 3 数据处理

数据是对事实、概念或指令的一种表达形式,数据经过解释并赋予一定的意义之后,便成为信息。数据处理的基本目的是从大量的、可能是杂乱无章的、难以理解的数据中抽取并推导出对于某些特定的人们来说是有价值、有意义的数据。即数据挖掘。

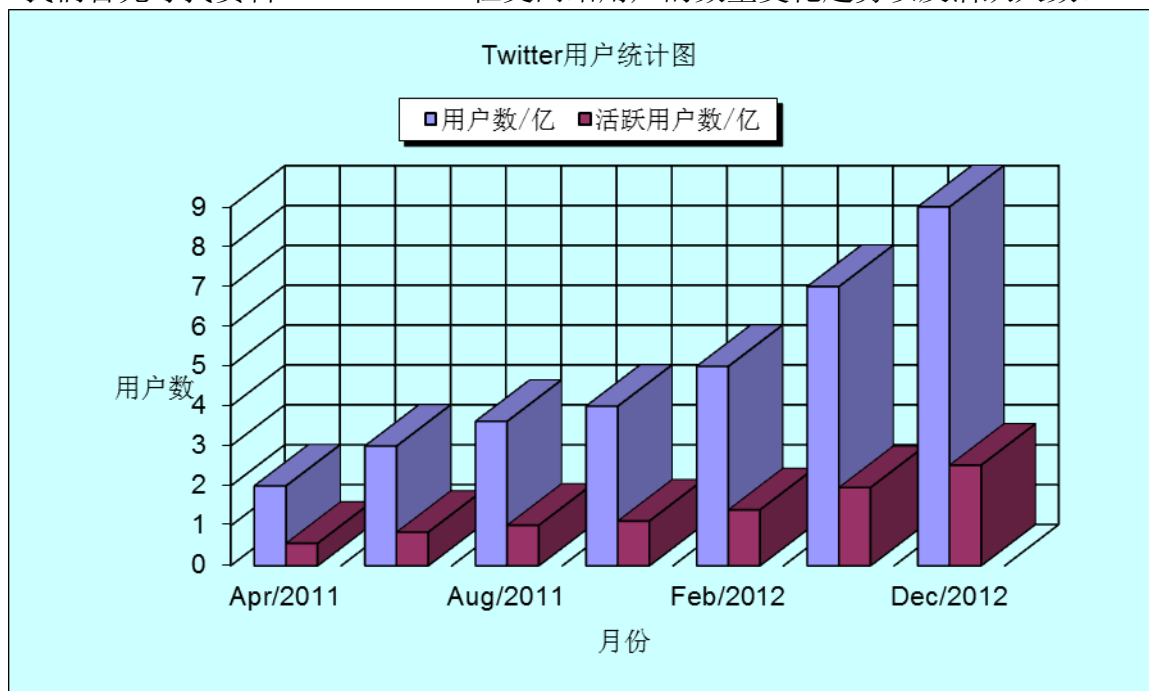
# 第五届数学中国数学建模网络挑战赛

#1465



而面对本题中多达835,541个数组的庞大数据，即附件中Twitter 社交网站用户之间的链接关系（follow 关系）数据，用于发现用户组，以及分析Twitter 用户的链接分布。如何处理好这些数据并挖掘是我们要解决的第一个问题。

我们首先寻找资料——Twitter 社交网站用户的数量变化趋势以及活跃人数。



然后我们运用office2010 (excel) 处理了Twitter 社交网站用户之间的链接关系数据，得到一个835541\*2的电子表格。再运用数据透视表处理得出表示用户数及其粉丝数之间的关系，将其导入matlab中构建出了twitter上用户与用户之间的关系模型（编程见附录）。并计算出了关系的重复率为42%。

# 第五届数学中国数学建模网络挑战赛

#1465

## 4 问题假设

- (1) 每个看到奥运广告的人单位时间内传递信息给其它人数与这时未看到奥运广告的人数成正比, 即  $k_0 = kg(t)$ 。
- (2) 总人数为  $m$ , 即  $h(t) + g(t) = m$ 。
- (3) 粉丝会把推广者发布的和奥运会相关的所有信息都分享给自己的粉丝。
- (4) 粉丝每天的增长数为  $U_i$ ;
- (5) 普通粉丝的下一层粉丝对其所分享的信息的关注度为  $\lambda$ 。

## 5 符号说明

- (1) 用  $h(t)$ ,  $g(t)$  表示  $t$  时刻看到奥运广告的人数和未看到广告的人数, 且  $h(0) = h_0$
- (2)  $p_{ii}^j$  表示节点  $j$  在  $[t, t + \Delta t]$  时段内处于未感染状态的概率,
- (3) 可设  $i$  代表用户身份定位;
- (4) 距2012年4月18号 (距伦敦奥运会100天) 天数为  $j$ ;
- (5) 当天新加入的人数为  $x_{ij}$ ;
- (6) 其初始粉丝数为  $S_i$ ;
- (7)  $p_{is}^j$  表示节点  $j$  从未感染状态转移到传播状态的概率且  $p_{is}^j = 1 - p_{ii}^j$ ;
- (8)  $g = g(t)$  表示在  $t$  时刻节点  $j$  的邻居中传播节点的数量;
- (9) 公司在网络宣传所花费的费用为  $w$

## 6 问题一模型的建立与求解

### 6.1 奥运会企业广告新闻被多少人观看到的模型——简易微分模型

#### 6.1.1 模型建立

由假设 (1) (2) 得微分方程

$$\frac{dh(t)}{dt} = kg(t)h(t) \quad (1)$$

$$h(t) + g(t) = m \quad (2)$$

$$h(0) = h_0 \quad (3)$$

#### 6.1.2 模型求解

通过以上三个微分方程解得

# 第五届数学中国数学建模网络挑战赛

#1465

$$h(t) = \frac{m}{1 + \left( \frac{m}{h_0} - 1 \right) e^{-kmt}} \quad (4)$$

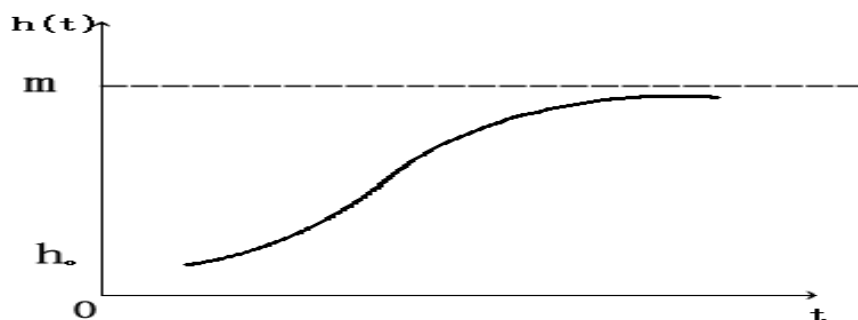


图1 看到奥运广告的人数与时间的关系曲线图

$$\frac{dh}{dt} = \frac{kn^2 \left( \frac{n}{n_0} - 1 \right) e^{-kmt}}{\left[ 1 + \left( \frac{m}{h_0} - 1 \right) e^{-kmt} \right]^2} \quad (5)$$

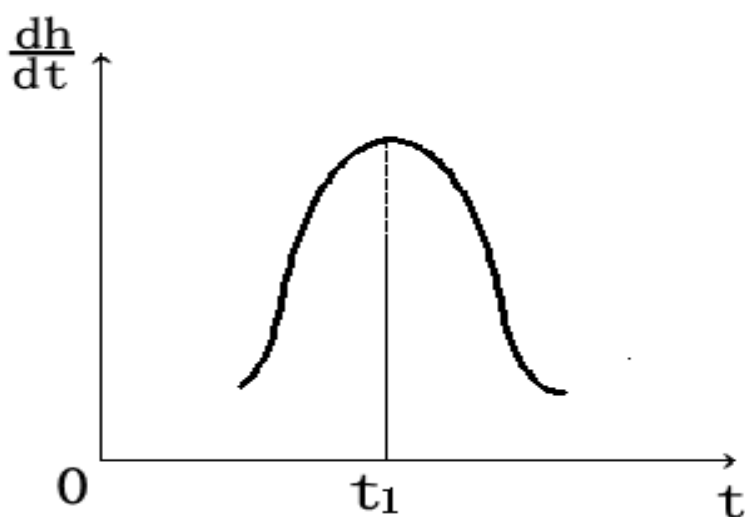


图2 看到奥运广告的人数增加率与时间的曲线图



# 第五届数学中国数学建模网络挑战赛

#1465

$$\frac{d^2 h(t)}{dt^2} = 0 \quad (6)$$

$$t_1 = \frac{\ln\left(\frac{n}{h_0} - 1\right)}{Km} \quad (7)$$

## 6.1.3 模型的评价与改进

由此可知，信息传播强度 $k$ 或总人数 $m$ 增加时， $t_1$ 都会变小与实际情况相吻合。同时，可预报看到奥运广告人数高峰 $t_1$ 的到来，这对公司发展及预测很有益处。为了与实际更加吻合，对数学模型要进一步修改。这要考虑信息的中断，所以，作出新的假设，建立新的模型。

## 6.2 奥运会企业广告新闻被多少人观看到的模型——信息传播模型

### 6.2.1 理论准备

#### 6.2.1.1 加权产业共生局域网络中信息传播仿真

网络在演化的过程中所具有的两个重要特征：一是网络持续增长性；二是节点择优连接机制。在产业共生网络中，节点表示个体企业，边表示的是节点个体之间发生的各种关系，信息通过这些边进行传播，不同权重的边，信息传播速率则不同。在信息传播模型中，各企业节点对信息的处理通常只有两种结果，已接受和易接受，且信息被接受的节点企业将对该信息不具有免疫性，即无法消除信息对其影响。将网络仿真拓扑图中节点之间连接的紧密程度用节点个体之间边的权重大小来衡量。权重不同将直接影响网络规模的紧密程度，从而间接影响信息在节点企业之间的传播速率。假设网络中边的权重越大，则节点企业间连接的紧密程度就越商，信息的传播速度也越大，反之亦然。同时假设两节点 $i$ 和 $j$ 之间的权重为 $w_{ij}$ ，其中 $i, j=1, 2, \dots$ 。模型的初始条件：假设初始网络规模为 $m_0$ ， $e_0$ 条边。 $e_0$ 条边不存在重连和自连现象，将每条边随机赋初始权重值为 $w_0$  ( $w_0 \in (0, 1)$ )。连接概率参数 $P \in (0, 0.05)$ 、 $P_2 \in (0, 0.05, 0.3)$ 、 $P \in (0.3, 0.65)$ 、 $m=3$ 和 $a=5$ ，网络规模稳定状态为

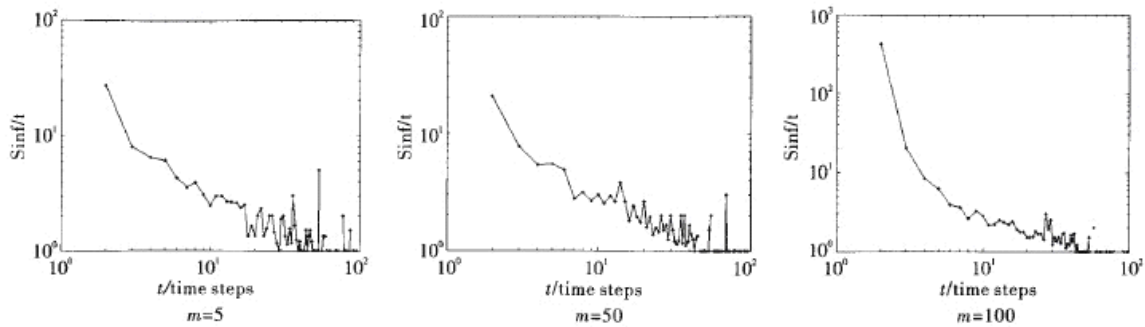
1) 信息传播过程中节点平均强度变化在加权网络中。不同权重的节点在信息传播过程中发挥的作用截然不同。为了验证此观点正确性，定义传播新节点在 $t$ 时刻的“平均强度” $I(t)$ 为

# 第五届数学中国数学建模网络挑战赛

#1465

$$S_{inf}(t) = \frac{\sum_s [I_s(t) - I_s(t-1)]}{I(t) - I(t-1)} \quad (8)$$

其中 $I_s(t)$ 为强度为 $s$ 的被外来信息同化的节点业数量。图2分别表示为在 $M=5$ 、 $M=50$ 、 $M=100$ 三种状态下新的被同化节点平均强度随时间 $t$ 的变化态势。



M 取不同值时加权局域网络中传播节点的平均强度变化

从图可以看出，节点平均强度函数呈现幂率分布，说明节点平均强度越大，被外来信息同化后对该信息的传播影响就越大。初始传播节点的选择对信息传播速度的影响初始网络中的节点一开始均处于易被信息同化的状态( $s$ )

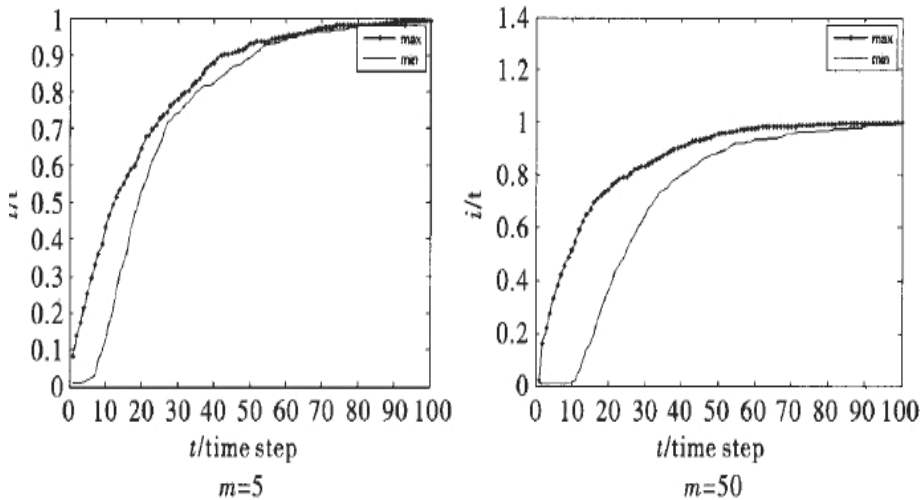


图 4 M 取不同值时网络中强度最大和强度最小节点对信息传播速度的影响

从图可以看出，初始局域网络规模 $M$ 越大，则两条曲线的差异则越大，这说明随着网络规模 $M$ 值的增加，个网络的非均匀性逐渐增强，初始网络中传播度最大的节点企业能加快信息在整个网络上的传播，同时可以看出信息在初始传播度比较小的节点企业中传播速度明显缓于 $M$ 值较小的网络。

局域世界规模 $M$ 的选择对信息传播速度的影响由于初始局域世界网络规模 $M$ 的选择对信

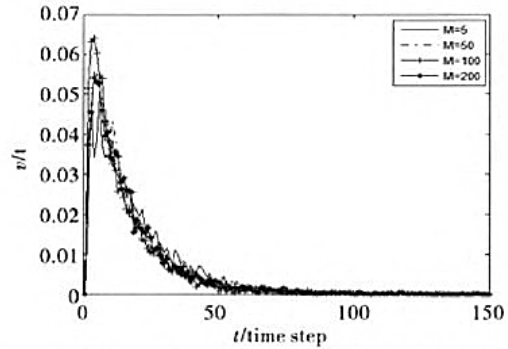
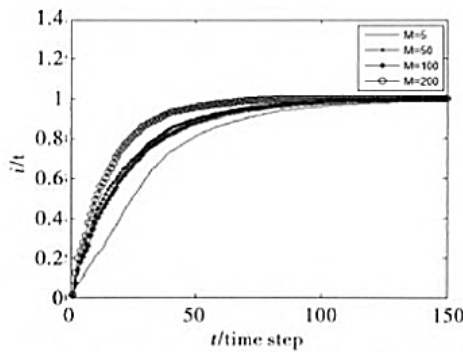
# 第五届数学中国数学建模网络挑战赛

#1465

息在网络中的传播有着重要的影响，定义信息传播速率为网络中被信息同化的节点变化

速率<sup>[2]</sup>:

$$v_{\text{inf}}(t) = \frac{di(t)}{d(t)} \approx \frac{I(t) - I(t-1)}{N} \quad (9)$$



M 取不同值时网络达到稳态时的传播密度比较 M 取不同值时网络达到稳态时信息传播速度比较  
初始局域网络规模M的值越大，则网络中信息的传播速度越快。

结论：初始局域网络规模M的值越大，同时节点强度较大的网络，则信息在网络中的传播速度较快。随着M值的不断增加，网络将会呈现一定的不均匀性，这也将会导致信息在网络中的传播速度较快。

## 5.2.1.2 信息传播模型IDM( $\alpha, \beta$ )<sup>[3]</sup>

$\alpha \in [0, 1]$ 表示消息的敏感度，其值是由消息的来源，内容及表现形式决定的； $\beta \in \{-1, 0, 1\}$ 是模型的另一个参数，用来决定选择转发信息的节点的方式。在这个模型中，网络中的节点只可能处于3种状态，一是未激活状态，即不知道消息I，记为S<sub>0</sub>，二是被消息I激活但选择不作为状态，记为S<sub>1</sub>，三是被消息I激活后对消息进行转发的状态，记为S<sub>2</sub>。时间步序列是一个离散的等间隔（如，5 min）时间序列。模型定义如下：

1) 初始时刻，将网络中所有节点的状态初始化为S<sub>0</sub>。在时间步T=0，随机选择一个节点v作为发布消息I的源节点。节点v的出度是dout<sub>v</sub>。2) 在时间步T=T+1，关注节点v的dout<sub>v</sub>个节点被激活，其中转发信息I的节点数量为M(M ≤ dout<sub>v</sub>)，M计算如下： $M = \text{dout}_v \cdot \alpha$ ，

2) 设P<sup>T</sup>是时间步T内选择转发的节点的集合，即P<sup>T</sup> = {u<sub>ju</sub>}，其中j, u的状态为S<sub>2</sub>。

3) 对关注节点v且不属于集合P<sup>T</sup>-1的n(n ≤ dout<sub>v</sub>)个节点的出度进行排序，按照从小到大的顺序记为Q = {u<sub>(1)</sub>, u<sub>(2)</sub>, ..., u<sub>(dout<sub>v</sub>)</sub>}。若1, 顺序选择序列Q中前M个节点，将它们状态设为S<sub>2</sub>，放入集合P<sup>T+1</sup>；若1, 顺序选择序列Q最后M个节点，将它们状态设为S<sub>2</sub>，放入集合P<sup>T+1</sup>；若b β = 0，从序列Q中随机选出M个节点，将它们状态设为S<sub>2</sub>，放入集合P<sup>T+1</sup>。将节点v从集合P<sup>T</sup>删去。

4) 对每一个节点v ∈ P<sup>T</sup>递归执行2)和3)，直到集合P<sup>T</sup>为空。

由式(1)可知，微博网络中节点v对消息传播贡献是与它的出度及消息的敏感度成正比的，这与现实情况中用户的粉丝越多，其发布（转发）的消息被进一步转发的可能性就越大及消息的敏感度越高，吸引力越大，其转发率越大是吻合的。在该模型中，利用

# 第五届数学中国数学建模网络挑战赛

#1465

参数来选择节点 $v$ 的“粉丝”节点中会进一步转发消息 $I$ 的节点,当 $b\beta=-1$ 时,选择出度小的节点对信息进行转发,当 $\beta=1$ 时,选择出度大的节点对信息进行转发,当 $b\beta=0$ 时,随机选择出转发节点。

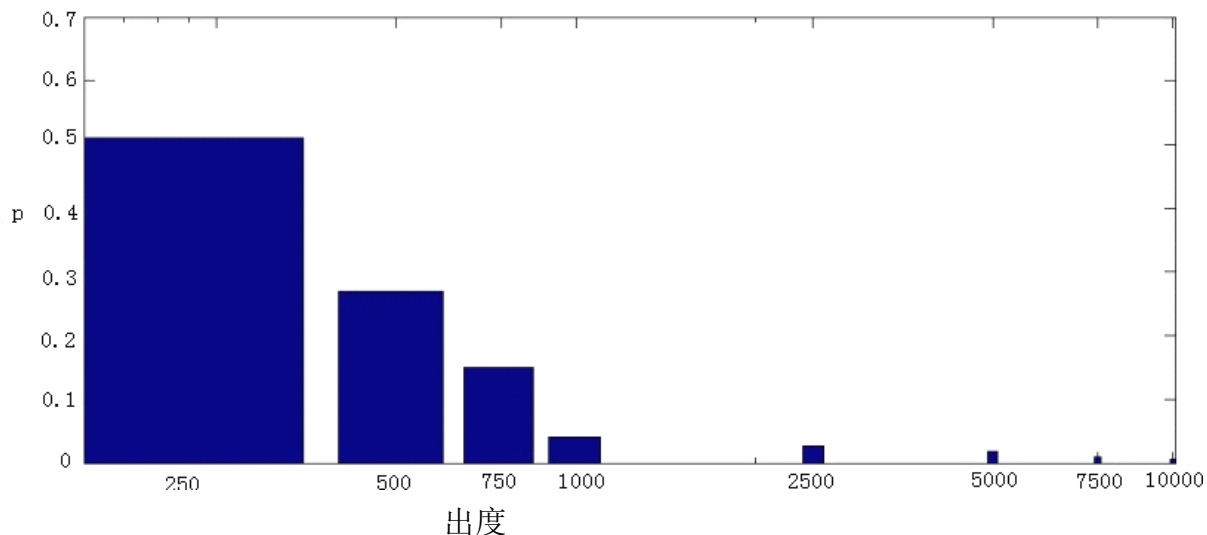
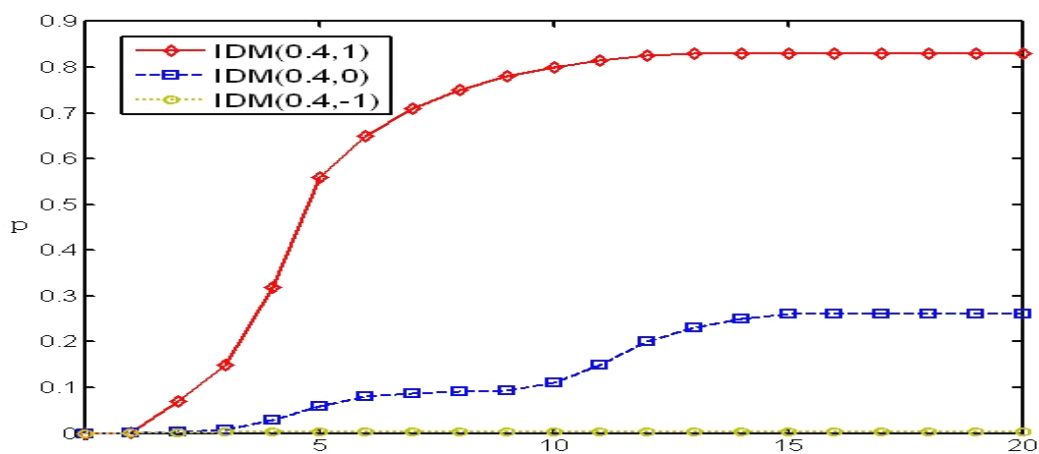


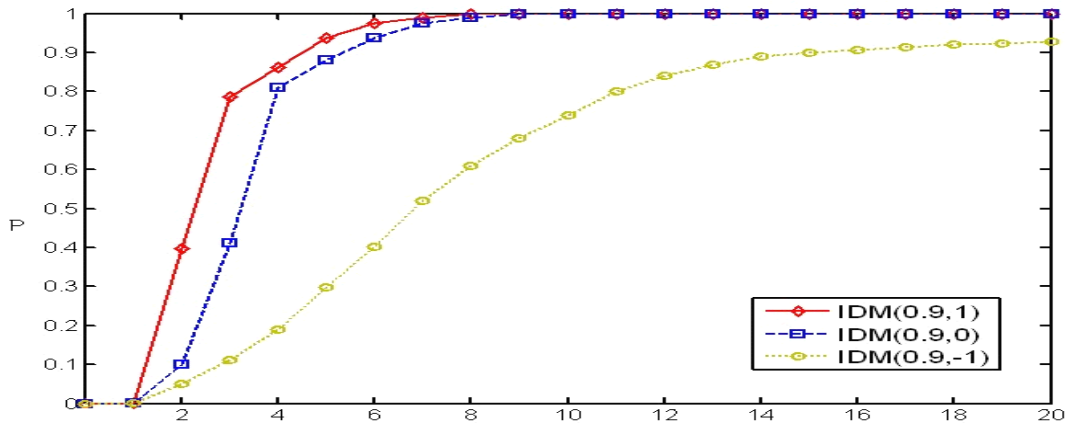
图7 累积出度分布



(a)  $\alpha=0.4$

# 第五届数学中国数学建模网络挑战赛

#1465



(b)  $\alpha=0.9$

图 8 信息扩散密度随时间的演化

图显示的是信息敏感点选择策略，信息在网络上的扩散面积和扩散耗时的比较，图 8 中空点，倒三角，叉分别代表取 1, 0, -1 时信息在网络中的扩散比例，每一条曲线是通过对每对参数 50 次仿真结果的统计平均得到的。从图 8 可以看出当信息敏感度较小时，在信息传播过程中扮演着意见领袖的角色，对舆论的传播起着绝对的推动作用；但当信息敏感度大于某一临界值后，明星用户对提高信息传播速度、扩大信息传播面积的作用逐渐减小；当信息敏感度趋于 1 时，明星用户在信息传播中的优势趋于 0。

## 5.1.2.3 NW 小世界模型

现实世界的网络有着较小的平均距离和较高的聚集系数。小世界网络模型就是为了同时把握这两个性质而建立的模型。首先要有一个具有周期边界条件的最近邻耦合网络，然后在所有与该节点不相邻的节点中随机选择一个节点重新建立连接，这样建立的网络就是 NW 小世界模型<sup>[4]</sup>。因为一个完全规则的最邻近耦合网络是有较高聚集系数的，而当 P 较小时，重新连线后得到的网络与原始的规则网络的局部性质差别不大，故而网络的聚集系数也变化不大。规则的最近耦合网络的平均路径长度很大，但当以不大的概率 P 对边重连时，网络的平均路径长度会显著减小。所以，NW 模型既具有较小的平均路径长度又具有较高的聚集系数。

### 平均最短路径

网络的平均最短路径 (average shortest paths) 可以对网络的连通性进行较好地描述。网络中两个节点 i 和 j 之间的距离  $d_{i,j}$  定义为连接这两个节点的最短路径上的边数。网络的平均最短路径长度 L 定义为任意两点之间的最短路径的平均值，即

$$L = \frac{1}{\frac{1}{2}N(N-1)} \sum_{i>j} d_{i,j} \quad (10)$$

### 聚类系数

聚类系数 (clustering coefficient) 表征的是网络的聚类特性，也就是群落特性。一般假设网络中的节点 i 与  $k_i$  条边关联，即与另外  $k_i$  个节点相连。显然，在这  $k_i$  个节点之间最多可能有  $k_i(k_i - 1) / 2$  条边。而这  $k_i$  个节点之间实际存在的边数

# 第五届数学中国数学建模网络挑战赛

#1465

是 $E_i$ 。那么这 $k_i$ 个节点之间实际存在的边数是 $E_i$ 与总的可能的边数 $k_i(k_i - 1)/2$ 之比就定义为节点 $i$ 的聚类系数 $C_i$ ，即

$$C_i = \frac{2E_i}{k_i(k_i - 1)} \quad (11)$$

而对网络中所有节点的聚类系数取平均值，就是整个网络的聚类系数 $C$ ，即

$$C = \frac{1}{N} \sum_{i=1}^N C_i \quad (12)$$

考虑一个含有 $N$ 个节点的最近邻耦合网络，它们围成一个环，其中每个节点都与它左右相邻的各 $k/2$ 个节点相连， $k$ 是偶数，也就是节点的度；然后进行随机化加边：以概率 $p$ 在随机选取的一对节点之间加上一条边。其中任意两个不同节点之间至多只能有一条边，并且每一个节点都不能有边与自身相连。在NW小世界网络模型中， $p = 0$ 对应于原来的最近邻耦合网络， $p = 1$ 是规则网和随机网的叠加。



图9 随机化加边

根据NW小世界网络模型的构造方法，用 $\text{rand}$ 函数进行仿真实现。网络的连接也用邻接矩阵 $a_{n \times n}$ 表示，其中 $n$ 表示网络的节点数。 $a_{i,j} = 1$ ，当节点 $i$ 和节点 $j$ 有边连接时； $a_{i,j} = 0$ ，当节点 $i$ 和节点 $j$ 无边相连时。利用均匀随机数产生器产生连续的和离散的随机数来代替。并且也使用稀疏矩阵的方法： $\text{sparse}(a)$ 来减少内存空间的使用。图5表示NW小世界网络模型随着随机加边概率 $p$ 的增加，平均最短路径和聚类系数的变化情况。与WS网络相同， $L(p)$ 表示网络的平均最短路径， $C(p)$ 表示网络的聚类系数，并且也使用规则网络的 $L(0)$ 和 $C(0)$ 进行归一化处理。从图中可以看到，当 $p$ 足够小的时候，NW小世界网络等同于WS小世界网络，随着 $p$ 的增加，平均最短路径急剧下降，而聚类系数下降十分缓慢；当 $p$ 比较大的时候，聚类系数开始快速下降；当 $p = 1$ 时，网络成为一个规则网络和随机网络的叠加网络，平均最短路径 $L(p)$ 和聚类系数 $C(p)$ 同时到达最小值。NW小世界网络模型的度分布如图所示，由于NW小世界网络用随机化加边代替了WS小世界网络的随机化重连，因此NW小世界网络每个节点的度至少是原来规则网络的 $k$ 值，所以度分布图为 $k \geq 4$ 部分的泊松分布<sup>[5]</sup>。



# 第五届数学中国数学建模网络挑战赛

#1465

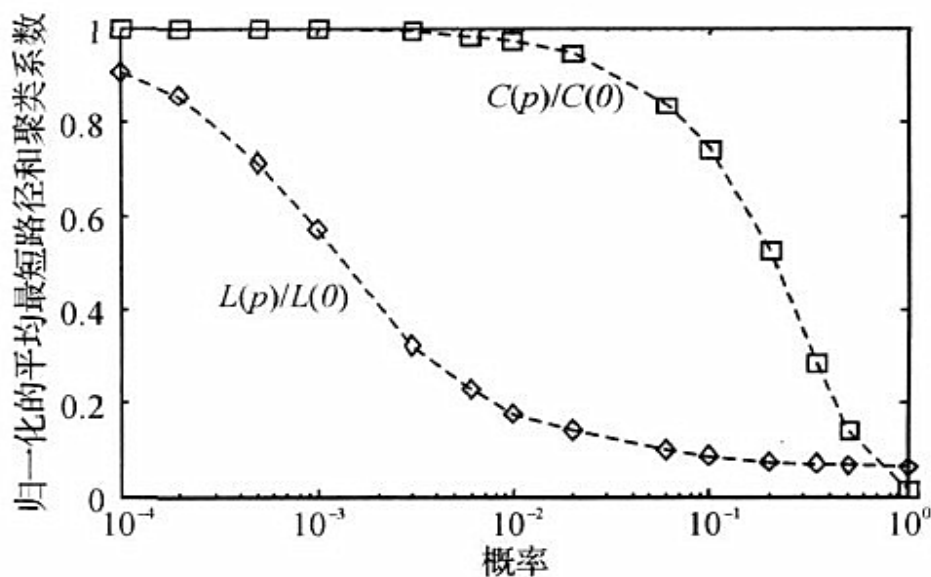


图10  $n=1000, k=10$ 的WS小世界网络 $L(p)$ 与 $C(p)$ 的归一化图

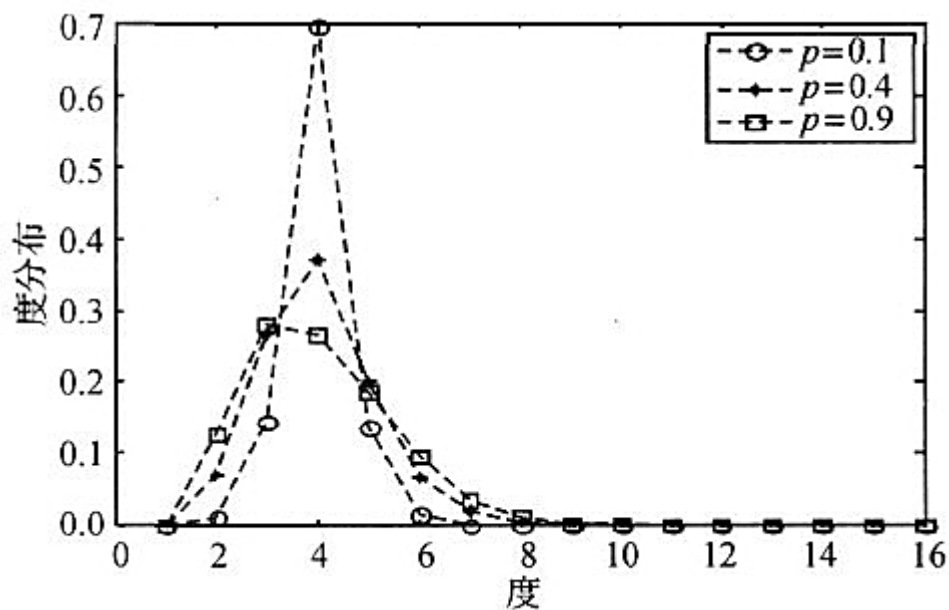


图11  $n=1000, k=4$ 的小世界网络度分布图

## 6.2.1.4 BA模型<sup>[6]</sup>

万维网这样的网络节点度没有明显的特征值，或者说没有一个值能够代表所有节点的特征，节点度的值变化很大。信息的被引频次呈现“长尾现象”，其服从帕累托分布，也就是幂率分布。它的度分布服从幂率分布，可由幂指数为3的幂率函数近似描述。其要点在于“增长”和“优先链接”。首先网络规模不是一成不变的，而是不断增长的，随着时间的推移，点和边会不断地加入。这一点与现实是一样的，网络中每天都会有新的加入者，信息的引用也在不断地增长，因为总有粉丝加入。其次边的加入不是随机的，

# 第五届数学中国数学建模网络挑战赛

#1465

而是有选择的。这样随时间推移，就会产生一个无标度网络。无标度网络的平均路径长度<sup>[7]</sup>是

$$L \sim \frac{\log n}{\log \log n} \quad (13)$$

所以BA模型同样具有小的平均路径长度。

## 5.2.2 模型建立

网络上广告信息的传播兼具以上几种模型的特征，但以上几种模型都有自己的缺点，不能全面的描述网络广告信息的传播。综合信息传播模型IDM( $\alpha, \beta$ )、NW小世界模型及BA模型，根据广告在网络中的传播规律<sup>[8]</sup>，我们把网络的节点分为三类：传播节点、未感染节点、免疫节点。传播节点表示该节点接受了来自其邻居节点的广告信息，并具有传播该广告的能力。未感染节点表示该节点没有接受过来自其邻居节点的信息，并有机会接受信息，即有概率被感染。免疫节点表示该节点已经接受了其邻居节点的信息，但是不具有传播能力，节点在传播状态、未传播状态和免疫状态之间的转移不仅依赖于节点自身的状态，还与它的邻居节点的状态相关，定义一下传播规则：

- 1) 如果一个传播节点与一个未传播节点接触，则未感染节点会以概率  $p_1$  成为传播节点。
- 2) 如果一个传播节点与一个免疫节点接触，则传播节点会以概率  $p_2$  成为免疫节点。
- 3) 传播节点不会无休止地传播下去，会以一定的速度  $v$  停止传播而变为免疫节点，且无需与其他节点接触。

假设节点  $j$  含有  $k$  条边， $g$  是具有如下二项分布的随机变量：

$$\Pi(g, t) = \binom{k}{g} \omega(k, t)^g (1 - \omega(k, t))^{k-g} \quad (14)$$

$\omega(g, t)$  表示在  $t$  时刻从具有  $k$  条边的未感染节点连接到一个传播节点的概率。 $\omega(g, t)$  可写成如下形式：

$$\omega(k, t) = \sum_{k'} p(k'|k) p(s_{k'}|i_k) \quad (15)$$

$$\approx \sum_{k'} p(k'|k) \rho^s(k', t) \quad (16)$$

式中， $p(k'|k)$  为度相关函数，表示度为  $k$  的节点与度为  $k'$  的节点邻接的条件概率： $p(s_{k'}|i_k)$  表示一个拥有  $k'$  条边的节点在它连接到一个度为  $k$  的未感染节点条件下，处于传播状态的概率： $\rho^s(k', t)$  表示在  $t$  时刻度为  $k'$  的传播节点的密度。

所以，得到度为  $k$  的节点在  $[t, t + \Delta t]$  时段内处于未感染状态的平均转移概率  $p^-$



# 第五届数学中国数学建模网络挑战赛

#1465

$i_i(k, t)$  如下:

$$\begin{aligned}\overline{p_{ii}}(k, t) &= \sum_{g=0}^k \binom{k}{g} (1 - p_1 \Delta t)^g \omega(k, t)^g \\ &\quad \times (1 - \omega(k, t))^{k-g} \\ &= (1 - p_1 \Delta t \omega(k, t))^k\end{aligned}\tag{17}$$

将 (3) 式代入上式得到

$$\overline{p_{ii}}(k, t) = \left( 1 - p_1 \Delta t \sum_{k'} p(k'|k) \rho^s(k', t) \right)^k\tag{18}$$

同理, 假设  $p_{ss}^j$  表示节点  $j$  在  $[t, t + \Delta t]$  时段内从传播状态转移到免疫状态的概率,

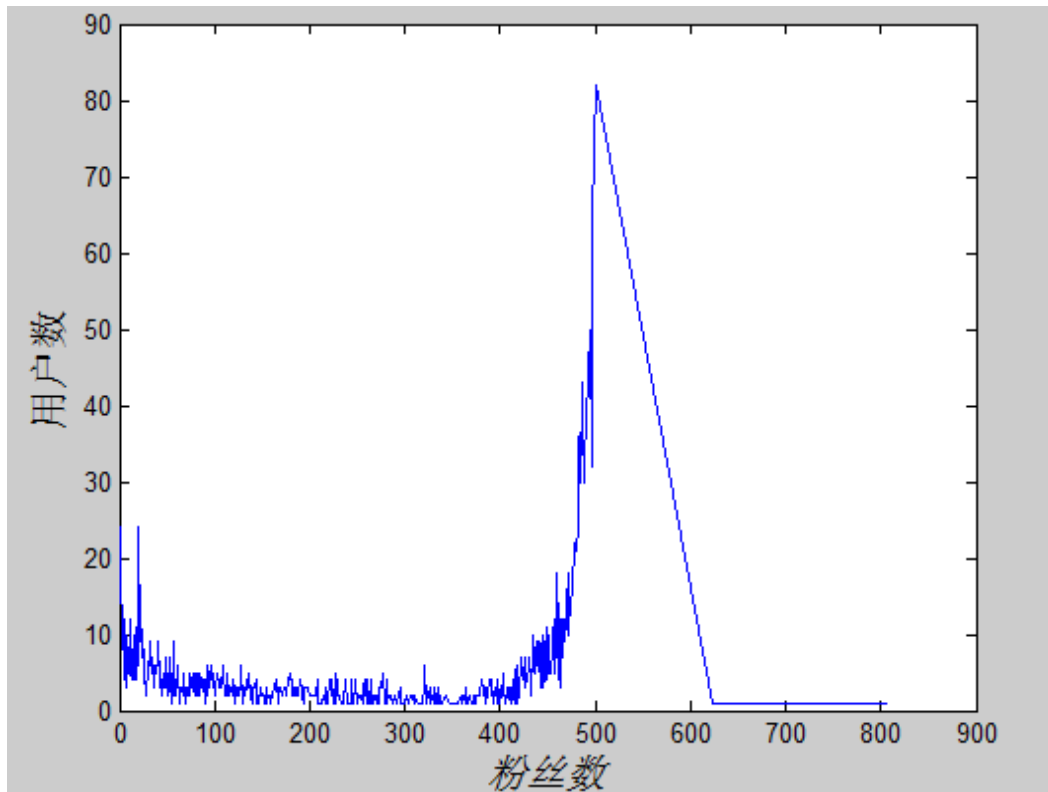
$p_{ss}^j$  表示节点  $j$  保持传播状态的概率且  $p_{ss}^j = (1 - \Delta t p_2)^g (1 - \nu \Delta t)$

## 6.2.3 模型求解

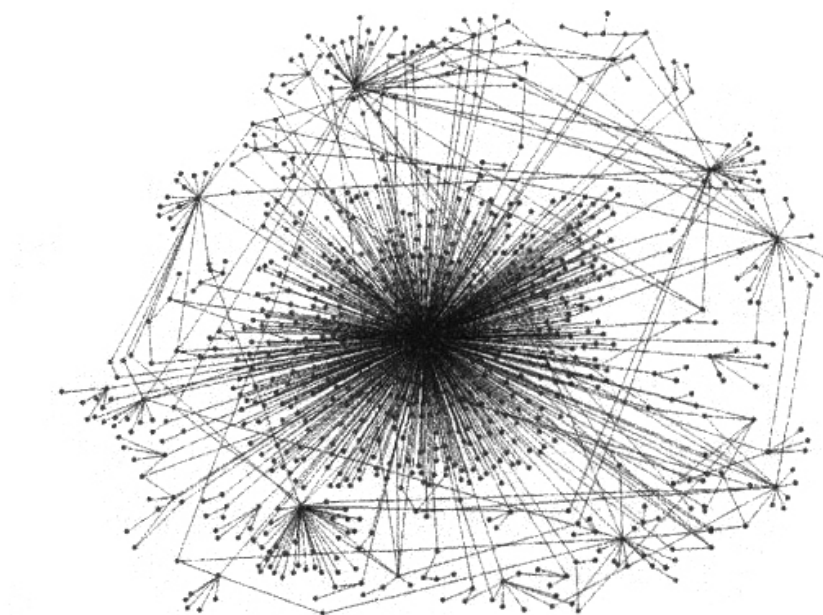
通过MATLAB对数据进行模拟得到用户数与粉丝数的关系模型, 如下图

# 第五届数学中国数学建模网络挑战赛

#1465



以及个人关系图:



得到度为  $k$  的节点在  $[t, t + \Delta t]$  时段内处

# 第五届数学中国数学建模网络挑战赛

#1465

$$\begin{aligned}
 S(k, t + \Delta t) &= S(k, t) + I(k, t) \left( 1 - \bar{p}(k, t) \right) \\
 &\quad - S(k, t) \left( 1 - \bar{p}(k, t) \right) \\
 &= S(k, t) + I(k, t) \\
 &\quad \times \left[ 1 - \left( 1 - p_1 \Delta t \times \sum_{k'} \rho^s(k', t) p(k'|k)^k \right) \right] \\
 &\quad - S(k, t) \left[ 1 - (1 - p_2 \Delta t) \times \sum_{k'} \rho^r(k', t) p(k'|k)^k (1 - \nu \Delta t) \right] \\
 &\quad R(k, t + \Delta t) \\
 &= R(k, t) + S(k, t) \left( 1 - \bar{p}_{ss}(k, t) \right) \\
 &= S(k, t) \left[ 1 - (1 - p_2 \Delta t) \times \sum_{k'} \rho^r(k', t) p(k'|k)^k \times (1 - \nu \Delta t) \right] \\
 &\quad + R(k, t)
 \end{aligned}$$

9)

由 (6), (7) 式可以得到

$$\begin{aligned}
 \frac{I(k, t + \Delta t) - I(k, t)}{N(k, t) \Delta t} &= - \frac{I(k, t)}{N(k, t) \Delta t} \\
 &\quad \times \left[ 1 - \left( 1 - p_1 \Delta t \times \sum_{k'} \rho^s(k', t) p(k'|k)^k \right)^K \right]
 \end{aligned}$$

当  $\Delta t \rightarrow 0$  时, 对 (10) 式右侧进行泰勒展开式得到

# 第五届数学中国数学建模网络挑战赛

#1465

$$\frac{\partial \rho^i(k, t)}{\partial t} = -kp_1 p^i(k, t) \sum_{k'} \rho^s(k'|k) p(k'|k)$$

同理，由（8）式可以得到

$$\begin{aligned} \frac{\partial \rho^s(k, t)}{\partial t} = & kp_1 p^i(k, t) \sum_{k'} \rho^s(k', t) p(k'|k) \\ & - kp_2 \rho^s(k, t) \sum_{k'} \rho^r(k', t) p(k'|k) - \nu \rho^s(k, t) \end{aligned}$$

由（9）式得到

$$\frac{\partial \rho^r(k, t)}{\partial t} = kp_2 \rho^s(k, t) \sum_{k'} \rho^r(k', t) p(k'|k)$$

通过以上推理并借助MATLAB求得一条含有企业广告的奥运会新闻可以被44万人看到。

## 6.2.4模型的评价与改进

本模型考虑了节点度的影响，当初始传播节点的度较大时，广告在网络中的传播速度较快。相反，当初始传播节点的度较小时，广告的传播具有滞后性，即广告从没有传播到开始传播需要一定的时间。但是当时间  $t$  趋向无穷大时，不论初始传播节点的度为多少，最终网络中免疫节点密度都趋向于 1，这是由于在线社交网络具有较高连通性的缘故。一些专业推广者及具有众多粉丝的名人大大加快了广告在网络中的传播速度。虽然这些节点能加快广告的传播，但是一旦他们成为免疫节点，它们同样能够有效地阻碍信息的传播但具有不同度数的节点在网络中所占比例不同，且它们呈现出相同的变化趋势。但实际网络的选型连接性有一些呈现协调混合，有一些呈现非协调混合本模型可进一步研究网络信息传播的基础，也为如何在动态网络中对传播行为进行建模提供了方向。

## 7 问题二模型的建立与求解

### 7.1 理论准备

# 第五届数学中国数学建模网络挑战赛

#1465

## 7.1.1 多目标函数优化算法

无约束的多目标函数优化的数学模型为：

$$\text{Min } F(X) = \{f_1(X), f_2(X), \dots, f_n(X)\} \quad X \in R \quad (1)$$

定义1 ( 点的优劣性, Dominance/ Inferiority) 如果目标向量  $u$  和  $v$  满足关系：

$$u_i \leq v_i \quad \forall i \in (1, 2, \dots, n) \text{ 且 } u_i < v_i \quad \exists i \in (1, 2, \dots, n) \quad (2)$$

则称目标向量  $u = \{u_1, u_2, \dots, u_n\}$  优于  $v = \{v_1, v_2, \dots, v_n\}$ , 记为  $u_p < v$ , 反过来称目标向量  $v$  劣于  $u$ , 记为  $v_p > u$ 。

定义2 ( 非劣最优解, Pareto 最优解) 决策变量  $x_u \in R^n$  ( $n$  为上角标) 称为多目标问题的非劣最优解或 Pareto 最优解, 当且仅当不存在决策变量  $x_v \in R^n$  ( $n$  为上角标), 使得相应的目标向量  $v = f(x_v \in R^n) = \{v_1, v_2, \dots, v_n\}$  优于  $u = f(x_u \in R^n) = \{u_1, u_2, \dots, u_n\}$ , 即  $u_p < v$ 。多目标最优化问题的数学模型为：

$$\begin{aligned} \min_{x \in R^n} F(x) \\ G_i(x) &= 0 \quad i = 1, \dots, m_e \\ G_i(x) &= 0 \quad i = m_e + 1, \dots, m \\ x_l &\leq x \leq x_u \end{aligned}$$

式中  $F(x)$  为目标函数向量。

由于多目标最优化问题中各目标函数之间往往是不可公度的, 因此往往没有惟一解, 此时引进非劣解的概念 (非劣解又称为有效解或帕累托解)。

定义 若  $x^* (x^* \in \Omega)$  的邻域内不存在  $\Delta x$ , 使得  $(x^* + \Delta x) \in \Omega$ , 且

$$\begin{aligned} F_i(x^* + \Delta x) &\leq F_i(x^*) \quad i = 1, \dots, m \\ F_j(x^* + \Delta x) &< F_j(x^*) \quad \text{对于某些 } j \end{aligned}$$

则称  $x^*$  为非劣解

## 7.1.1 加权分析

将多目标向量问题转化为所有目标的加权求和的标量问题, 即

$$\min_{x \in \Omega} f(x) = \sum \omega_i \cdot F_i(x)^2$$

该问题可以用标准的无约束最优化算法进行求解。

## 7.1.2 进行约束约束

为克服了加权分析的某些凸性问题。它对目标函数向量中的主要目标 进行最小化, 将其他目标用不等式约束的形式写出：

# 第五届数学中国数学建模网络挑战赛

#1465

$$\min_{x \in \Omega} F_p(x)$$

$$\text{sub. } F_i(x) \leq \varepsilon_i \quad i=1, \dots, m \quad i \neq p$$

## 7.1.3 目标达到

目标函数系列为  $F(x) = \{F_1(x), F_2(x), \dots, F_m(x)\}$ , 对应地有其目标值系列  $F^* = \{F_1^*, F_2^*, \dots, F_m^*\}$ 。允许目标函数有正负偏差, 偏差的大小由加权系数向量  $W = \{W_1, W_2, \dots, W_m\}$  控制, 于是目标达到问题可以表达为标准的最优化问题:

$$\min_{\gamma \in R, x \in \Omega} \gamma$$

$$\text{sub. } F_i(x) - \omega_i \gamma \leq F_i^* \quad i=1, \dots, m$$

指定目标  $\{F_1^*, F_2^*\}$ , 定义目标点 P。权重向量定义从 P 到可行域空间  $\Lambda(\gamma)$  的搜索方向。在优化过程中,  $\gamma$  的变化改变可行域的大小, 也是边界变为惟一解点  $F_{1s}, F_{2s}$ 。

## 7.2 模型建立

在该问题中我们将模型中的参与的用户分为三类: 专业网络社交推广者、兼职宣传者、普通粉丝。当天新加入的人数为  $i = \begin{cases} 1(\text{专业推广者}) \\ 2(\text{兼职推广者}) \\ 3(\text{普通推广者}) \end{cases}$ ;

则

$$\text{目标函数 1: } \min W = \sum_{i,j} S_i X_{ij} \cdot j$$

可获得信息的总人数为: z 则

$$\text{目标函数 2: } \max Z = \sum_{i=1}^n \sum_{j=1}^n (S_i \cdot j \cdot u_i \cdot N + S_i \cdot j \cdot N + S_i \cdot j) x_{ij}$$

## 7.3 模型求解

使用 lingo 软件进行最优化求解得到: 109 个专业社交网络推广和 15 个兼职宣传者。达到最优。

## 7.4 模型评价与改进

目标达到法的一个好处是可以将多目标最优化问题转化为非线性规划问题, 但是, 在序列二次规划 (SQP) 过程中, 一维搜索的目标函数选择不是一件容易的事情, 因为在很多情况下, 很难决定是使目标函数变大好还是使它变小好。这导致许多目标函数创建过程的提出。可以通过将目标达到问题变为最大最小化问题来获得更合适的目标函数。

# 第五届数学中国数学建模网络挑战赛

---

#1465

## 参考文献

- [1] Yah Gang. Zhou Tan, Wang Ji, Epidemic spread inweighted scale-free networks[J],Chinese Physics letters, 22(2): 510, 2005.
- [2] 李慧嘉, 马英红, 加权局域网络上的病毒传播行为研究[J], 计算机工程与应用, 45(35): 80\_83, 2009.
- [3] 郑蕾, 李生红, 基于微博网络的信息传播模型, 通信技术, 02 (45): 39-40. 2012.
- [4] 毛秉毅, 基于目标树的组卷算法的研究[J], 计算机工程与应用, 23: 245-247, 2002.
- [5] 王小凡, 李翔, 陈关荣, 复杂网络理论及其应用[ M], 北京: 清华大学出版社, 2006.
- [6] Maciej, Lewenstein, Andrzej Nowak, Bibb Latan6. Statistical mechanics of social Impact, Phys. Rev. A, Vol. 45, PP. 763-776, 1992.
- [7] M. F. Laguna, Guillermo Abramsom, Damián H. Zanette, Vector opinion dynamics in a model for social influence[J], Physica A: Statistical Mechanics and its Applications, Vol 329, PP. 459-472, 2003. 11.
- [8] 刘云, 张海峰, 基于在线社交网络的信息传播模型, 物理学报, 03 (7), 2010.

# 第五届数学中国数学建模网络挑战赛

#1465

## 附 录

### 附录一：程序

#### 程序源代码之一

##### 算法代码

```
function [d,r] = floyd(a)
n = size(a,1);
d = a;
for i = 1:n
    for j = 1:n
        r(i,j) = j;
    end
end

for k = 1:n
    for i = 1:n
        for j = 1:n
            if d(i,k)+d(k,j) < d(i,j)
                d(i,j) = d(i,k) + d(k,j);
                r(i,j) = r(i,k);
            end
        end
    end
    k;
    d;
    r;
end
```

#### 程序源代码之二

##### 问题一模型一求解代码

```
clc
clear
close all;

load A.txt %节点的坐标
load oneTo.txt %相邻节点的信息
load FaAnLv.txt %发案率的信息

data = A;

data23 = data(:,2:3);

data23 = data23./10; %单位换算

temp23 = data23'; %求各个路口之间的距离
len = length(data23(:,1));

lenOneTo = length(oneTo);

distanceA = zeros(len,len); %初始化距离矩阵
```



# 第五届数学中国数学建模网络挑战赛

#1465

```
for i = 1:len
    for j = 1:len
        if i == j
            distanceA(i,j) = 0;
        elseif i ~= j
            distanceA(i,j) = inf;
        end
    end
end

for i = 1:lenOneTo
    beginNum = oneTo(i,1);
    endNum = oneTo(i,2);
    if beginNum <= len && endNum <= len
        distanceA(beginNum,endNum) = sqrt((data23(beginNum,1) -
data23(endNum,1))^2 + (data23(beginNum,2) - data23(endNum,2))^2);
        distanceA(endNum,beginNum) = sqrt((data23(beginNum,1) -
data23(endNum,1))^2 + (data23(beginNum,2) - data23(endNum,2))^2);
    end
end

[d,r] = floyd(distanceA);
distance1No2Temp = d(1:20,:);
disXiongYaLi = d(1:20,:);

dis1No1 = d(1:20,:);
dis1No3Final = d;
belongWhere = zeros(1,len);

for i = 1:len
    disTemp = inf;
    pos = 0;
    for j = 1:20
        if dis1No1(j,i) < disTemp
            disTemp = dis1No1(j,i);
            pos = j;
        end
    end
    for j = 1:20
        if j ~= pos
            dis1No1(j,i) = 0;
        elseif j == pos
            dis1No1(j,i) = 1;
        end
    end
    belongWhere(1,i) = pos;
end

for i = 1:len
    hold on;
    plot(data(i,2),data(i,3),'b. ');
    % text(data(i,2)+1,data(i,3),num2str(i),'fontsize',6);
end
```

# 第五届数学中国数学建模网络挑战赛

#1465

```
for i = 1:lenOneTo %开始画相邻的线
    beginNum = oneTo(i,1);
    endNum = oneTo(i,2);
    if beginNum <= len && endNum <= len
        plot([data(beginNum,2) data(endNum,2)], [data(beginNum,3)
data(endNum,3)]);
    end
end

for i = 1:20
    plot(data(i,2),data(i,3),'ro');
end

flag = strvcats('^k','sy','dg','pm','hc','^y','sg','dm','pc','hk','^g',
'sm','dc','sk','pk','^h','dc','pm','sc','dy');

for i = 1:len
    for j = 1:20
        if dis1No1(j,i) == 1
            text(data(i,2),data(i,3),num2str(j));
            if j ~= 16
                plot(data(i,2),data(i,3),strcat(flag(j,:),''));
            end
        end
    end
end
end
```

## 程序源代码之三

### 问题一模型一改进算法代码

```
clear;clc

Zone = 'A';
v = 60e3/60; % m/min
timelimit = 3; % min
distlimit = v * timelimit;

[node,path,station,DIST] = ZoneData(Zone,v);

%VoronoiGraph(node,station,path);

figure
[node, station,path] = PathGraph2(node, station, path, DIST, v);

%PathGraph(node, station, path, DIST);

xlswrite(strcat(Zone,'.xls'), {'num','station','journey',...
    'max n-n time','max s-n time'}, 'sheet1', 'A1:E1')
xlswrite(strcat(Zone,'.xls'), [station.stalab], 'sheet1', 'A2')
xlswrite(strcat(Zone,'.xls'), [station.nodlab], 'sheet1', 'B2')
xlswrite(strcat(Zone,'.xls'), [station.journey], 'sheet1', 'C2')
xlswrite(strcat(Zone,'.xls'), [station.maxtwonodestime], 'sheet1', 'D2')
xlswrite(strcat(Zone,'.xls'), [station.maxStaNodeTime], 'sheet1', 'E2')
```

# 第五届数学中国数学建模网络挑战赛

#1465

## 程序源代码之四

### 代码

```
function [x,y] = zhipai(n,A)
a = zeros(n*2,n*n); % 花费最小指派问题
for i = 1:n
    a(i,(i-1)*n+1:n*i) = 1; % A 是指派的人和任务的花费关系矩阵
    a(n+i,i:n*n) = 1; % x 是最小花费的指派矩阵
end % y 是最小花费的值
b = ones(n*2,1); % A=A(:);
P = zeros(n,n*n);
for i = 1:n-7
    for j = 1:n
        P(i,20*(i-1)+j) = A(j,i);
    end
end
[x,y] = bintprog(max(P),[],[],a,b);
X = reshape(x,[n,n]);
```

## 程序源代码之五

### 代码

```
A = xlsread('A2:E93');
plot(A(:,2),A(:,3),'.b');
text(A(:,2),A(:,3),num2str(A(:,1)));

B = xlsread('A2:B145');
[n,m] = size(B);
d = ones(n,1)*inf;
for i = 1:n
    if B(i,1) < 93 & B(i,2) < 93
        X = [A(B(i,1),2),A(B(i,2),2)];
        Y = [A(B(i,1),3),A(B(i,2),3)];
        hold on;
        plot(X,Y,'b');
        hold off;
        d(i) = sqrt((X(1)-X(2))^2 + (Y(1)-Y(2))^2);
    end
end
l = ones(93,93)*inf;
for k = 1:n
    i = B(k,1);
    j = B(k,2);
    if i < 93 & j < 93
        l(i,j) = d(k);
        l(j,i) = d(k);
    end
end

C = xlsread('A2:B21');
[n,m] = size(C);
for i = 1:n
    X = A(C(i,1),2);
    Y = A(C(i,1),3);
    hold on;
    plot(X,Y,'or');
```

# 第五届数学中国数学建模网络挑战赛

#1465

```
hold off;
end

D = xlsread('A2:C18');
N = 13;
for i = 1:n
    X = A(D(i,3),2);
    Y = A(D(i,3),3);
    hold on;
    plot(X,Y,'*r');
    hold off;
end

l_min = [];
for i = 1:20
    [S,P] = minRoute1(C(i,1),93,1,1);
    l_min = [l_min;P];
end

d_min = [];
for i = 1:13
    d_min = [d_min,l_min(:,D(i,3))];
end

I = zeros(20,7);
d_min = [d_min,I];

[x,y] = zhipai(20,d_min);
x2 = x(:,2);
x(:,2) = x(:,3);
x(:,3) = x2;
P = zeros(1,13);
for i = 1:13
    for j = 1:20
        if x(j,i) == 1
            P(i) = j;
        end
    end
end

t_s = x.*d_min;
t_s = max(t_s);
t_s = max(t_s);
t = t_s/10;
hold on;
plot(A(D(1,3),2),A(D(1,3),3),'^k');
plot(A(P(1),2),A(P(1),3),'^k');
plot([A(D(1,3),2),A(P(1),2)], [A(D(1,3),3),A(P(1),3)], 'm');
hold off;

hold on;
plot(A(D(2,3),2),A(D(2,3),3),'sy');
plot(A(P(2),2),A(P(2),3),'sy');
plot([A(D(2,3),2),A(P(2),2)], [A(D(2,3),3),A(P(2),3)], 'm');
```

# 第五届数学中国数学建模网络挑战赛

#1465

```
hold off;

hold on;
plot(A(D(3,3),2),A(D(3,3),3),'dg');
plot(A(P(3),2),A(P(3),3),'dg');
plot([A(D(3,3),2),A(P(3),2)], [A(D(3,3),3),A(P(3),3)], 'm');
hold off;

hold on;
plot(A(D(4,3),2),A(D(4,3),3),'pm');
plot(A(P(4),2),A(P(4),3),'pm');
plot([A(D(4,3),2),A(P(4),2)], [A(D(4,3),3),A(P(4),3)], 'm');
hold off;

hold on;
plot(A(D(5,3),2),A(D(5,3),3),'hc');
plot(A(P(5),2),A(P(5),3),'hc');
plot([A(D(5,3),2),A(P(5),2)], [A(D(5,3),3),A(P(5),3)], 'm');
hold off;

hold on;
plot(A(D(6,3),2),A(D(6,3),3),'^y');
plot(A(P(6),2),A(P(6),3),'^y');
plot([A(D(6,3),2),A(P(6),2)], [A(D(6,3),3),A(P(6),3)], 'm');
hold off;

hold on;
plot(A(D(7,3),2),A(D(7,3),3),'sg');
plot(A(P(7),2),A(P(7),3),'sg');
plot([A(D(7,3),2),A(P(7),2)], [A(D(7,3),3),A(P(7),3)], 'm');
hold off;

hold on;
plot(A(D(8,3),2),A(D(8,3),3),'dm');
plot(A(P(8),2),A(P(8),3),'dm');
plot([A(D(8,3),2),A(P(8),2)], [A(D(8,3),3),A(P(8),3)], 'm');
hold off;

hold on;
plot(A(D(9,3),2),A(D(9,3),3),'pc');
plot(A(P(9),2),A(P(9),3),'pc');
plot([A(D(9,3),2),A(P(9),2)], [A(D(9,3),3),A(P(9),3)], 'm');
hold off;

hold on;
plot(A(D(10,3),2),A(D(10,3),3),'hk');
plot(A(P(10),2),A(P(10),3),'hk');
plot([A(D(10,3),2),A(P(10),2)], [A(D(10,3),3),A(P(10),3)], 'm');
hold off;

hold on;
plot(A(D(11,3),2),A(D(11,3),3),'^g');
plot(A(P(11),2),A(P(11),3),'^g');
plot([A(D(11,3),2),A(P(11),2)], [A(D(11,3),3),A(P(11),3)], 'm');
```

# 第五届数学中国数学建模网络挑战赛

#1465

```
hold off;

hold on;
plot(A(D(12,3),2),A(D(12,3),3),'sm');
plot(A(P(12),2),A(P(12),3),'sm');
plot([A(D(12,3),2),A(P(12),2)], [A(D(12,3),3),A(P(12),3)], 'm');
hold off;

hold on;
plot(A(D(13,3),2),A(D(13,3),3),'dc');
plot(A(P(13),2),A(P(13),3),'dc');
plot([A(D(13,3),2),A(P(13),2)], [A(D(13,3),3),A(P(13),3)], 'm');
hold off;
```

## 程序源代码之六

### 模型三的代码

```
A = xlsread('A2:E93');
plot(A(:,2),A(:,3),'.b');

B = xlsread('A2:B145');
[n,m] = size(B);
d = ones(n,1)*inf;
for i = 1:n
    if B(i,1) < 93 & B(i,2)<93
        X = [A(B(i,1),2),A(B(i,2),2)];
        Y = [A(B(i,1),3),A(B(i,2),3)];
        hold on;
        plot(X,Y,'b');
        hold off;
        d(i) = sqrt((X(1)-X(2))^2 + (Y(1)-Y(2))^2);
    end
end
l = ones(93,93)*inf;
for k = 1:n
    i = B(k,1);
    j = B(k,2);
    if i < 93 & j < 93
        l(i,j) = d(k);
        l(j,i) = d(k);
    end
end

C = xlsread('A2:B21');
[n,m] = size(C);
for i = 1:n
    X = A(C(i,1),2);
    Y = A(C(i,1),3);
    hold on;
    plot(X,Y,'or');
    hold off;
end

D = xlsread('A2:C18');
N = 13;
```

# 第五届数学中国数学建模网络挑战赛

#1465

```
for i = 1:n
    X = A(D(i,3),2);
    Y = A(D(i,3),3);
    hold on;
    plot(X,Y,'*r');
    hold off;
end

l_min = [];
for i = 1:20
    [S,P] = minRoute1(C(i,1),93,1,1);
    l_min = [l_min;P];
end

for i = 1:20
    for j = 1:93
        if l_min(i,j) > 30
            l_min(i,j) = inf;
        end
    end
end
l_min = l_min(:,21:93);
m_num = min(l_min);

p = [1:20,28,38,61,92];
u_min = [];
for i = 1:24
    [S,P] = minRoute1(p(i),93,1,1);
    u_min = [u_min;P];
end
for i = 1:93
    for j = 1:24
        if (i>=1 & i<=20 | i==28 | i==38 | i==61 | i==92)
            u_min(j,i) = inf;
        end
    end
end

for i = 1:24
    for j = 1:93
        if u_min(i,j) > 30
            u_min(i,j) = inf;
        end
    end
end

for i = 1:24
    for j = 1:93
        if u_min(i,j) ~= inf
            u_min(i,j) = A(j,5);
        end
    end
end
```

# 第五届数学中国数学建模网络挑战赛

#1465

```
u_s = zeros(1,92);
for i = 1:20
    u_s(i) = i;
end

u_s(28) = 28;
u_s(38) = 38;
u_s(61) = 61;
u_s(92) = 92;

u_p = zeros(24,1);
for i = 1:20
    u_p(i) = A(i,5);
end
u_p(21) = A(28,5);
u_p(22) = A(38,5);
u_p(23) = A(61,5);
u_p(24) = A(92,5);

u_b = zeros(24,93);
for i = 1:24
    for j = 1:93
        if u_min(i,j) ~= inf
            u_b(i,j) = 1;
        end
    end
end

u_k = zeros(93,1);
for i = 1:93
    u_k(i) = sum(u_b(:,i));
end

for i = 1:93
    if u_k(i) == 1
        j = find(u_b(:,i));
        u_p(j) = u_p(j) + A(i,5);
    end
end

u_l = find(sum(u_b));
[u_n,u_m] = size(u_l);
for i = 1:u_m
    u_x = u_b(:,u_l(i));
    u_r = f_min(u_p,u_x);
    u_p(u_r) = u_p(u_r) + A(u_l(i),5);
    for j = 1:24
        u_b(j,u_l(i)) = 0;
    end
    u_b(u_r,u_l(i)) = 1;
end

for i = 1:20
    u_b(i,i) = 1;
```



# 第五届数学中国数学建模网络挑战赛

---

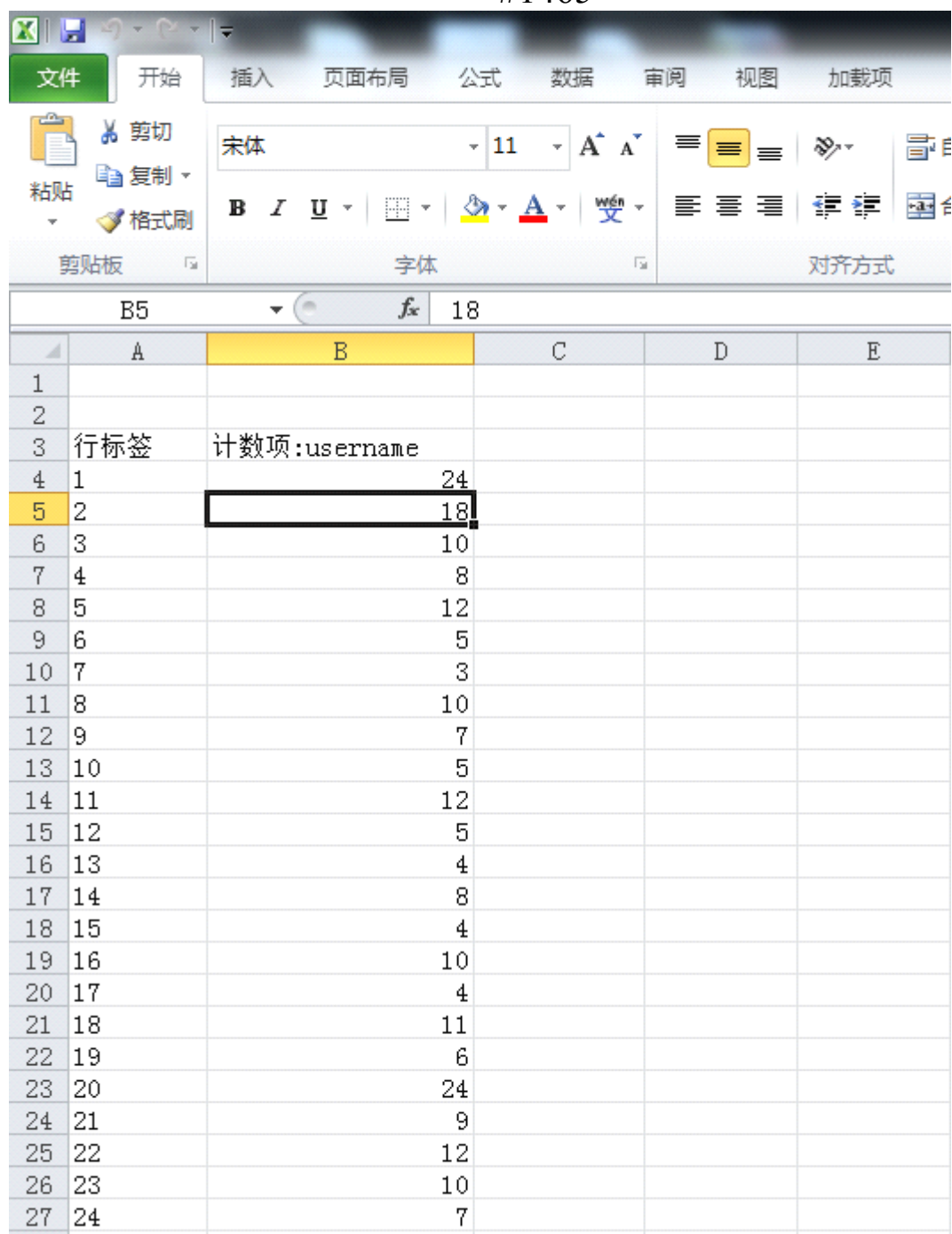
#1465

```
end  
u_b(21,28) = 1;  
u_b(22,38) = 1;  
u_b(23,61) = 1;  
u_b(24,92) = 1;
```

附录二 表格

# 第五届数学中国数学建模网络挑战赛

#1465



	A	B	C	D	E
1					
2					
3	行标签	计数项:username			
4	1	24			
5	2	18			
6	3	10			
7	4	8			
8	5	12			
9	6	5			
10	7	3			
11	8	10			
12	9	7			
13	10	5			
14	11	12			
15	12	5			
16	13	4			
17	14	8			
18	15	4			
19	16	10			
20	17	4			
21	18	11			
22	19	6			
23	20	24			
24	21	9			
25	22	12			
26	23	10			
27	24	7			

# 第五届数学中国数学建模网络挑战赛

#1465

	A	B	C	D	E
408	477	20			
409	478	19			
410	479	22			
411	480	21			
412	481	21			
413	482	24			
414	483	36			
415	484	31			
416	485	30			
417	486	43			
418	487	37			
419	488	30			
420	489	32			
421	490	39			
422	491	40			
423	492	42			
424	493	47			
425	494	43			
426	495	50			
427	496	32			
428	497	68			
429	498	70			
430	499	74			
431	500	82			
432	624	1			
433	807	1			
434	总计	2503			

B13		fx		
	A	B	C	D
1	月份	用户数/亿	活跃用户数/亿	
2	Apr/2011	2	0.56	
3	May/2011	3	0.84	
4	Aug/2011	3.62	1.0136	
5	Oct/2011	4	1.12	
6	Feb/2012	5	1.4	
7	Aug/2012	7	1.96	
8	Dec/2012	9	2.52	
9				
10				
11				
12				