

# 第五届数学中国数学建模网络挑战赛

地址：内蒙古数学学会  
电话：0471-4969085

邮编：010021

网址：[www.tzmcm.cn](http://www.tzmcm.cn)  
Email: 2012@tzmcm.cn

---

## 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：2553

参赛队员（签名）：

队员 1：程莉

队员 2：王瀚

队员 3：傅梅芬

参赛队教练员（签名）：

参赛队伍组别：大学本科组

# 第五届数学中国数学建模网络挑战赛

地址：内蒙古数学学会  
电话：0471-4969085

邮编：010021

网址：[www.tzmcm.cn](http://www.tzmcm.cn)  
Email: 2012@tzmcm.cn

---

## 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）：

2553

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

地址：内蒙古数学学会  
电话：0471-4969085

邮编：010021

网址：[www.tzmcm.cn](http://www.tzmcm.cn)  
Email: 2012@tzmcm.cn

## 2012 年第五届“认证杯”数学中国 数学建模网络挑战赛

题 目 基于 SNS 的消息推广效应及规划模型

关 键 词 社交网络 消息传播模型 分布密度函数

### 摘 要:

本文以碎片化趋势下的奥运会商业模式为背景，研究了 SNS 社交网络中的消息传播模式，用户对消息的反应行为，分析了消息传播过程与机制，。结合本文结合传播理论、网络拓扑结构将消息传播问题抽绎成社交网络中的信息流动模型，对网络节点及其间的消息传播过程进行了划分，并在此基础上基于微分方程模型、图论、传染病动力学构造了一个社交网络消息传播模型。基于总体功能与架构，将整个模型划分为二次传播模型，消息蔓延模型，规划模型三个子模型。

通过构建二次传播中消息转发比率演化模型，考虑了传播过程中的边际转发率递减规律，结合微分方程、传染病动力学等，研究了消息未知用户，传播用户，仅接收用户分布密度增长加速度。通过节点度的引入，用 SPSS 分析处理出网络结构特征数据，带入模型中结合 C++编程得到中间数据，再导入 MATLAB 中分析，拟合出了传播节点、未知节点、只读节点，三类节点分布密度随时间的变化趋势图。

消息蔓延模型以二次传播模型为基础，从传播过程和时间两个维度对模型进行分解，得出基于不同基数的求解公式。并通过 C++分层模块化实现，拟出基于粉丝基数、粉丝增加数，参数化的实现过程，并进行迭代求总的蔓延效应。通过对不同粉丝基数和粉丝增加数的对比分析，发现粉丝基数对于消息蔓延效应的影响大于粉丝增加数。本部分还通过 SPSS，excel 数据分析，统计计算出其粉丝重复度，带入模型消除重复情况对结果的影响。

基于上述两个子模型，提出了规划模型，带入数据用 MATLAB 求解出最优方案。此外文章还提出了利用幂律分布、BA 模型，均匀分布等对模型进行改进的方案，以及随时间变化的消息蔓延模型的思路。最后文章基于分析结果，对如何利用社交网络平台进行宣传 and 营销提出了建议。

参赛队号 #2553

所选题目 C

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

地址：内蒙古数学学会  
电话：0471-4969085

邮编：010021

网址：[www.tzmcm.cn](http://www.tzmcm.cn)  
Email: 2012@tzmcm.cn

## Promotion effect and programming model of SNS-based news

### Abstract

In the background of fragmentation of the Olympic Games business. This article research SNS social network information dissemination model, the behavior of user reaction to the news, analyze the information dissemination process and mechanism. Combined with this paper, communication theory, network topology information dissemination pumping unravel into a model of the flow of information in social networking, network nodes between the information dissemination process has been divided, and on this basis, based on the differential equation model, graph theory, infectious disease dynamics, the propagation model of a social networking message. Based on the overall function and structure, the entire model for the secondary transmission model, the message spread model, programming model three sub-models.

Evolution model by constructing the message forwarding secondary transmission ratio, considered marginal in the communication process forwarding rate law of diminishing, combined with differential equations, infectious disease dynamics, the message of unknown users, and dissemination of user receives only the user density growth acceleration. Node introduction of processed using SPSS analysis of network structure characteristics data, combined with the C programming to get the intermediate data into the model, and then imported into MATLAB, analysis, fitting out of the spread of nodes, unknown nodes, the read-only node, three types of nodes distribution of density with time trend map.

The news spread model based on the model of secondary transmission from the communication process and the time the two dimensions of the model decomposition, obtained based on different base formula to solve. And C hierarchical modular draw up a fan base, fans increase the number of parameters of the implementation process, and the contagion effects of the iterative total. The news spread effect is greater than a fan to increase the number of different fan base and fans to increase the number of comparative analysis and found that fan base. This section also SPSS, excel data analysis, statistical computing its fans repeat, into the model to eliminate duplication of results.

Based on these two sub-models, programming model into the data using MATLAB to solve the optimal solution. In addition, the article also raised the power-law distribution of BA model, the uniform distribution of the model to improve the program, as well as changes over time, news spread of the model program. The last article on how to use social networking platform based on the results of the analysis, publicity and marketing recommendations.

Key words: social network, news spread model, density function

## 目录

一、 问题重述.....	2
二、 符号说明.....	2
三、 问题假设.....	3
四、 问题分析.....	3
4.1 社交网络中信息传播模式分析.....	3
4.2 社交网络中节点群演化模式.....	4
4.3 模型选择与解题思路.....	5
五、 模型建立.....	6
5.1 二次传播中消息转发比率演化模型（简称二次传播模型）.....	6
5.1.1 模型的建立.....	6
5.1.2 模型的求解.....	9
5.2 消息蔓延效应预测模型.....	11
5.2.1 模型的建立.....	11
5.2.2 编程实现.....	14
5.3 消息覆盖规划模型.....	14
5.3.1 重复率 $L$ 分析.....	14
5.3.2 线性规划模型.....	16
七、 模型的综合评价.....	17
7.1 模型优点.....	17
7.2 模型的不足点.....	18
八、 模型的改进方案.....	18
8.1 改进一：用幂律分布和 BA 模型来改 P1.....	18
8.1.1 不足分析.....	18
8.1.2 P1 的幂律分布和 BA 模型的演化算法.....	18
8.2 改进二：用均匀与幂律混合分布演化算法改进网络中粉丝新增率.....	19
8.2.1 不足分析.....	19
8.2.2 粉丝新增速度的均匀与幂律混合分布演化算法改进.....	19
九、 基于社交网络平台营销提出的几点建议.....	20
十、 参考文献.....	21
十一、 附录.....	21
附录一：二次传播模型中 C++ 部分程序代码.....	21
附录二：消息蔓延效应预测模型的 C++ 部分程序代码.....	24
附录三：求解 LP 模型最优解的 MATLAB 代码.....	26

# 基于 SNS 的消息推广效应及规划模型

## 一、 问题重述

随着网络的不断开放和技术的不断先进，数字化已成为当今潮流，它让传播渠道、受众的注意力、品牌营销方式甚至一切都碎片化了，观众不在只关注电视，他们利用社交网络可以获得更加丰富的比赛信息和网友的评论。这也为更多的企业提供了在奥运期间宣传自己的机会。

一家企业想利用社交网络在奥运会期间为企业做宣传，在距离奥运会开幕的100天，一个专业推广者新增粉丝的平均速度为500 个/天，这些粉丝会把消息转发分享给自己的粉丝们，普通网络用户新增粉丝的平均速度为20 个/天。

第一个需要解决的问题是建立相关的数学模型以及分析Twitter 用户的链接分布来预测奥运会开始后，一条含有企业广告的奥运会消息可以被观看到的人数。

第二个需要给企业制定一份合理的用人方案。假设企业产品在社交网络中潜在用户大约有2 亿人，要求广告宣传覆盖潜在用户的40%，需要雇佣的专业社交网络推广者人数；假设专业推广者工资是500 元/天，网络上可以雇佣兼职宣传者，工资为50 元/天，相应的新增粉丝的平均速度为35 个/天，综合考虑经济效益，制定出最佳用人方案。

## 二、 符号说明

符号	说明
$P_1$	未知节点转发消息，转化为传播节点的概率；
$1-P_1$	未知节点转化为只读节点的概率；
$P_{ww}^j$	表示节点 j 在 $[t, t + \Delta t]$ 时段内处于未知状态的概率；
$P_{wc}^j$	表示节点 j 从未知状态转移到转发状态的概率；
$P_{wg}^j$	表示节点 j 从未知状态转移到只读状态的概率；
$H$	表示在 t 时刻节点 j 的邻接节点中转发节点数量；
$L(m, t)$	表示在 t 时刻从具有 m 条边的未知节点连接到一个转发节点的概率；
$p(C_m   W_m)$	表示一个拥有 $m'$ 条边的节点在它连接到一个度为 m 的未知节点的情况下，处于转发状态的概率；
$p^c(m', t)$	表示在 t 时刻度为 $m'$ 的转发节点的密度；
$N(m, t)$	在 t 时刻社交网络中度为 m 的节点总数；
$w(m, t)$	在 t 时刻网络中度为 m 的未知节点的数量；
$c(m, t)$	在 t 时刻网络中度为 m 的转发节点的数量；
$g(m, t)$	在 t 时刻网络中度为 m 的只读节点的数量；
$m$	普通网络用户和兼职宣传员的原始粉丝基数；
$m'$	专业推广者的原始粉丝基数；
$\Delta S_t$	第 t 天由普通网络用户引起的新增消息转发者；
$\Delta K_t$	第 t 天由普通网络用户引起的新增只读者；

$\Delta T_t$	由第 $t$ 天新增的 500 粉丝在剩余的 $100-t+1$ 天内总的消息接收者;
$P(m')$	度为 $m'$ 的节点的分布函数
$J$	总的普通网络用户接收者的数目;
$X$	招聘专业推广者人数;
$Y$	兼职宣传者人数

### 三、问题假设

1. 由转发引起的时间延迟可忽略, 用户接收到携带广告的奥运新闻后可以及时将此消息转发, 且他的粉丝可以及时收到该条新闻并看到其中的广告;
2. 假设粉丝消息接收方式仅为通过状态更新看到奥运新闻广告, 即不考虑进入收听者主页, 通过查看其历史状态记录等方式。
3. 普通网络用户转发消息当天, 新增粉丝行为发生在转发消息之前, 即当天新增粉丝都可收到其动态 (转发消息)。
4. 假设所有网络普通用户的转发行为都发生在第一次看到该条新闻之后, 此后看到同一条新闻, 用户将不再转发此新闻。因此, 消息转发的时间都为首次看到新闻的时间。专业推广者和兼职宣传者因任务要求, 需要每天都发布或者转发该条新闻。
5. 假设专业推广者的粉丝均为普通网络用户。
6. 假设企业具有选择偏好, 都偏向于雇佣具有较多粉丝的专业推广者。专业推广者分布于粉丝较多的网络用户群中。兼职宣传者粉丝基数和普通网络用户一样; 专业推广者, 兼职宣传者, 普通网络用户的差异表现在原有粉丝基数和粉丝增长速度。

### 四、问题分析

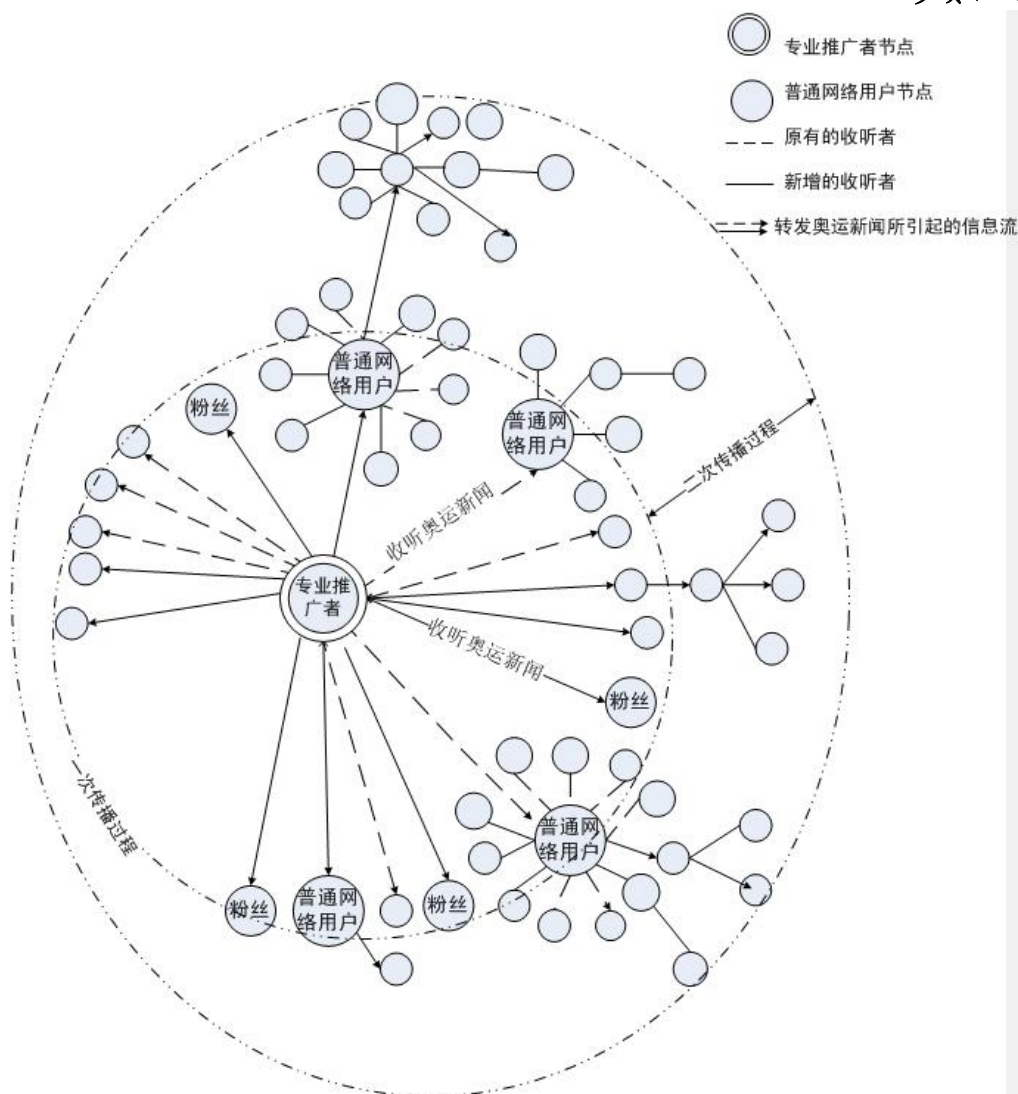
#### 4.1 社交网络中信息传播模式分析

携带广告的奥运新闻在社交网络中的传播模型可以抽象为, 信息在社交网络中节点 (专业推广者, 普通网络用户) 间的流动过程。社交网络中的节点分为专业推广者节点和普通网民节点。消息的传播方式也有两类。

①一次传播。以专业推广者为中心, 由专业推广者发布或者转发奥运新闻。从而将广告信息直接推向其粉丝 (普通网民)。这种奥运新闻及广告由专业推广者直接到普通网民的流动过程, 叫做消息的一次传播过程。专业推广者  $\longrightarrow$  普通网络用户。

②二次传播。普通网民通过转发该条新闻, 从而将广告信息推给其粉丝, 即由专业推广者引起的, 经普通网民转发的普通网民对消息的接收, 叫做二次传播过程。  
专业推广者  $\cdots \rightarrow$  普通网络用户  $\longrightarrow$  普通网络用户。





社交网络传播模式

上图专业推广者一次转发所引起的传播效益图。图中“虚线”代表专业推广者或者普通用户与其原有粉丝的收听关系。“实线”代表专业推广者或普通用户与其新增粉丝的新增收听关系。信息的推动过程表现为新闻的首次发布或者转发引起的动态更新。

专业推广者的首次发布信息过程，其接收者为所有粉丝（粉丝基数+新增粉丝），转发者考虑范围也为所有粉丝，而在之后的重复转发过程中，其接受者仍为所有粉丝，但转发者考虑范围仅为新增粉丝（基于假设4）。

#### 4.2 社交网络中节点群演化模式

在社交网络中，一个人发布的消息会被其粉丝看到，并以一定的概率转发、传播。同时，若其粉丝对其内容不感兴趣则只阅读新闻并不转发消息，此条信息流到此终止。本文把 SNS 网络上的所有相关用户抽象为节点，用户之间的收听关系则可以抽象地用节点之间的边来表示，消息沿着边传播。根据用户节点是否阅读到携带广告的奥运新闻，以及是否转发该条新闻，将社交网络中的所有用户节点分为三类。

- ①未知节点，该节点还没有接收过某条信息，并有可能接收消息。
- ②只读节点，该节点已经接收了来自于其邻接节点转发的消息，但是不准备转发消



息。

③转发节点，该节点接受了来自其收听节点的消息，并通过转发将此消息传播给他的邻接节点；

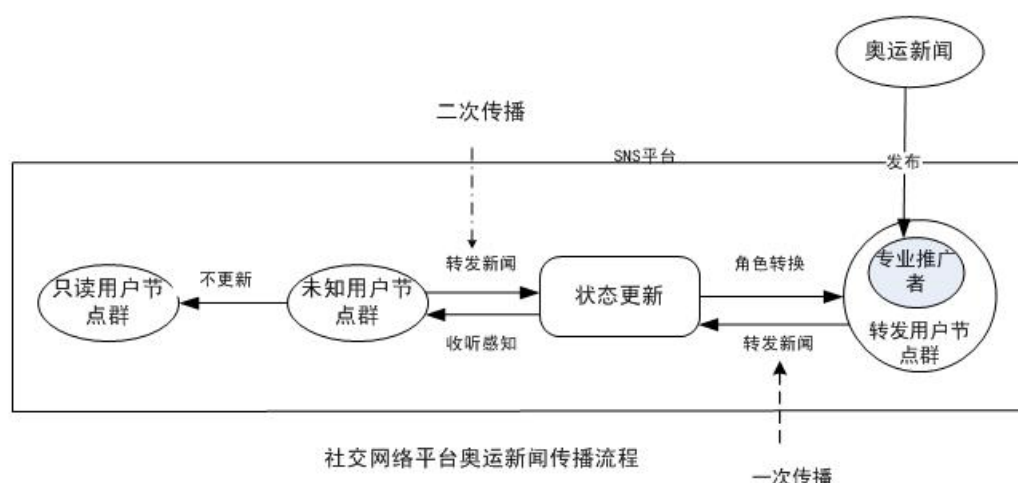
这三类节点分别对应三类用户群。

① 未知用户节点群。该集合中的元素为没有阅读到带广告的奥运新闻的用户。

② 只读用户节点群。该集合中的所有用户均阅读过新闻，但不转发新闻。

③ 转发用户节点群。该群中所有用户均阅读过并且转发了携带广告的新闻。

初始时，企业刚刚在奥运新闻中植入广告，SNS（社交关系网络）中所有用户均属于未知用户群。企业通过专业推广员在线发布新闻，使其进入 SNS 平台。新闻以专业推广员动态更新的方式传送到其粉丝（即一次传播过程）。专业推广员的粉丝在收听感知新闻后，状态激活。若该粉丝不转发此新闻则其由原来的未知状态转为只读状态，进入只读用户群。若该粉丝转发此新闻，将新闻推送给他自己的粉丝（二次传播过程）后，其状态变为转发态，其角色也转化为转发用户进入转发用户群。



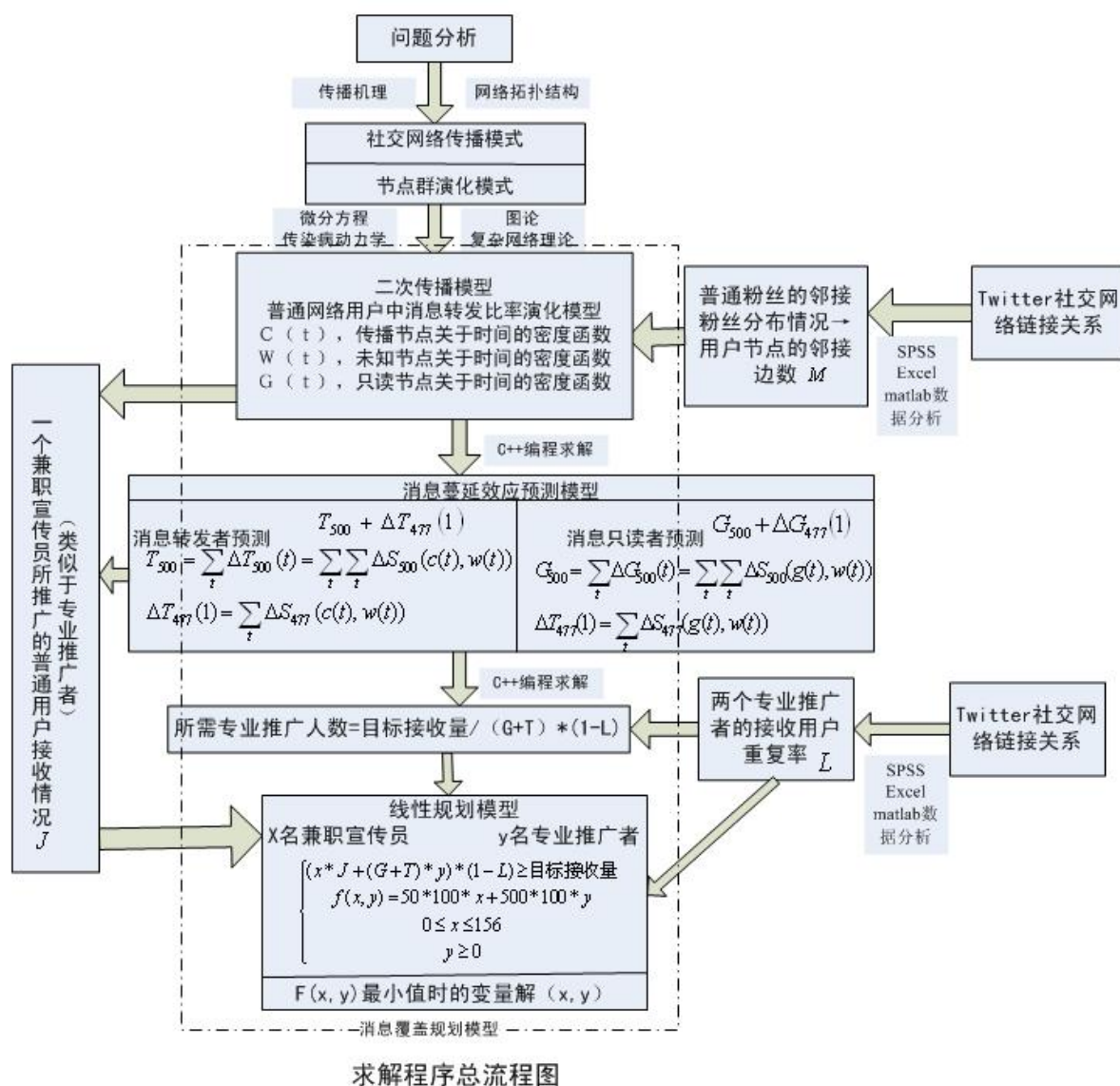
### 4.3 模型选择与解题思路

在前述社交网络信息传播模式和节点群演化模式的基础上，考虑到节点间的相互转化关系，信息的流动过程类似于传染性疾病的传染模型（SIR 模型）。以转发为方式的消息传播类似于以感染为方式的传染病传播模型。未知用户群可以看做未感染人群，具有感染疾病（接收到消息）和传播疾病（转发消息）的可能。只读用户群类似于免疫人群，他们感染过疾病（接受到过消息）但具有免疫力不会在传播疾病（转发消息）。转发用户群类似于传染人群，他们通过转发将消息传播给临近节点。

在 SIR 模型的基础上考虑到节点度（用户链接关系数）和传播机理（边际转发效率递减），结合网络拓扑结构和传染病动力学，构建出了消息传播的动力学方程组。在此方程组的基础上通过 matlab、C++编程等模拟出未知节点、只读节点和转发节点关于时间  $t$  的密度函数  $W(t)$ 、 $G(t)$ 、 $C(t)$ ，即二次传播过程中消息转发比率演化模型，简称二次模型。

基于此二次模型，以一次传播过程的基数为依据，将一次传播分为原有粉丝引发的二次传播和新增粉丝引发的二次传播，并对他们分别代入消息蔓延效应预测模型中求出消息收听者的蔓延情况。由于消息转发密度  $C(t)$  和只读密度  $G(t)$  演化情况不同，消息蔓延效应模型又分为消息转发者预测模型和消息接受者预测模型。到此已求出以推广能力和原有粉丝数为基数的消息蔓延模型。

接着将专业推广者和兼职宣传者的推广能力和粉丝基数代入模型中求解出单个推广人员或者兼职宣传者分别引起的消息蔓延规模。最后再以单个消息蔓延规模为基础，结合粉丝重复率，成本约束条件等构建线性规划模型求解最优组合。此过程称作消息覆盖规划模型。



## 五、模型建立

### 5.1 二次传播中消息转发比率演化模型（简称二次传播模型）

#### 5.1.1 模型的建立

##### 5.1.1.1 概念界定

节点：二次传播过程中的普通网络用户。分为未知节点，传播节点，只读节点。

邻接/边：节点之间的收听关系。分为节点与其收听者的主动收听关系和与粉丝间的被动收听关系。

度：粉丝节点的邻接数量。包括与收听者的消息入度和与粉丝的消息出度。

传播概率：未知节点接受到消息后，转发消息，变成转发节点的概率。

二次传播模型：即二次传播中消息转发比率演化模型，此模型主要研究普通用户间消息传播过程（即二次传播过程）中，传播比率随时间的演化关系。

消息传播规则：一个转发节点将消息转发给一个未知节点后，未知节点会以  $P_1$  的概率转发此条消息，转化为传播节点，以  $1-P_1$  的概率转化为只读节点。

#### 5.1.1.2 建立模型

节点  $j$  在  $t$  时刻处于未知状态， $P_{ww}^j$  表示节点  $j$  在  $[t, t + \Delta t]$  时段内处于未知状态的概率， $P_{wc}^j$  表示节点  $j$  从未知状态转移到转发状态的概率， $P_{wg}^j$  表示节点  $j$  从未知状态转移到只读状态的概率，且  $P_{wc}^j = 1 - P_{wg}^j$ ，其中

$$P_{ww}^j = (1 - \Delta t p_1)^h \quad (1)$$

$h = h(t)$  表示在  $t$  时刻节点  $j$  的邻接节点中转发节点数量。节点  $j$  含有  $m$  条边， $h$  是具有如下二项式分布的随机变量：

$$\Pi(h, t) = \binom{m}{h} L(m, t)^h (1 - L(m, t))^{m-h} \quad (2)$$

$L(m, t)$  表示在  $t$  时刻从具有  $m$  条边的未知节点连接到一个转发节点的概率，

$$L(m, t) = \sum_{m'} p(m' | m) p(c_{m'} | w_m) \approx \sum_{m'} p(m' | m) p^c(m', t) \quad (3)$$

$$P(m' | m) = \frac{m' P(m')}{\bar{m}}$$

$P(m')$  是度为  $m'$  的节点的概率分布函数， $\bar{m}$  为样本网络空间的节点平均度。 $P(m' | m)$  为度相关函数，表示度为  $m$  的节点邻接度为  $m'$  的节点的条件概率。 $p(c_{m'} | w_m)$  表示一个度为  $m'$  的节点邻接到一个度为  $m$  的未知节点的条件概率，转化为转发节点的概率， $p^c(m', t)$  表示在  $t$  时刻度为  $m'$  的转发节点的概率密度。

度为  $m$  的未知节点在  $[t, t + \Delta t]$  时段内处保持未知状态的平均转移概率  $p_{ww}^-(m, t)$ 。

$$p_{ww}^-(m, t) = \sum_{h=0}^m \binom{m}{h} (1 - p_1 \Delta t) L(m, t)^h (1 - L(m, t))^{m-h} = (1 - p_1 \Delta t L(m, t))^m$$

上面两式联合得到：

$$p_{ww}^-(m, t) = \left( 1 - p_1 \Delta t \sum_{m'} p(m' | m) p^c(m', t) \right)^m \quad (4)$$

设  $N(m, t)$  为在  $t$  时刻社交网络中度为  $m$  的节点密度， $w(m, t)$ ， $c(m, t)$ ， $G(m, t)$  分别为在  $t$  时刻网络中度为  $m$  的未知节点、转发节点和只读节点的概率密度。

$\Delta w(m, t), \Delta c(m, t), \Delta g(m, t)$  分别表示相应密度的变化量。

$$\begin{aligned} w(m, t) + c(m, t) + G(m, t) &= N(m, t) \\ \Delta w(m, t) &= \Delta c(m, t) + \Delta G(m, t) \end{aligned} \quad (5)$$

网络中度为  $m$  的未知节点的数量在  $[t, t + \Delta t]$  时段内的变化情况如下

$$\begin{aligned} w(m, t + \Delta t) &= w(m, t) - w(m, t)(1 - p_{ww}^-(m, t)) \\ &= w(m, t) - w(m, t) * \left[ 1 - \left( 1 - p_1 \Delta t * \sum_{m'} p^c(m', t) p(m' | m) \right)^m \right] \end{aligned} \quad (6)$$

以此类推, 可得到度为  $m$  的转发节点和只读节点的数量在  $[t, t + \Delta t]$  时间段内的变化情况, 如下:

$$\begin{aligned} C(m, t + \Delta t) &= C(m, t) + p_1 * W(m, t)(1 - p_{ww}^-(m, t)) \\ &= C(m, t) + p_1 * W(m, t) * \left[ 1 - \left( 1 - p_1 \Delta t * \sum_{m'} p^c(m', t) p(m' | m) \right)^m \right] \end{aligned} \quad (7)$$

$$\begin{aligned} G(m, t + \Delta t) &= G(m, t) + (1 - p_1) * W(m, t)(1 - p_{ww}^-(m, t)) \\ &= G(m, t) + (1 - p_1) * W(m, t) * \left[ 1 - \left( 1 - p_1 \Delta t * \sum_{m'} p^c(m', t) p(m' | m) \right)^m \right] \end{aligned} \quad (8)$$

由 (5), (6) 公式得到,

$$\frac{W(m, t + \Delta t) - w(m, t)}{N(m, t) \Delta t} = - \frac{W(m, t)}{N(m, t) \Delta t} \left[ 1 - \left( 1 - p_1 \Delta t * \sum_{m'} p^c(m', t) p(m' | m) \right)^m \right] \quad (9)$$

当  $\Delta t$  趋近于 0 时, 将 (9) 号公式右侧进行泰勒公式展开得到, 未知节点的密度随时间的变化关系  $w(t)$ :

$$\frac{\partial p^w(m, t)}{\partial t} = -m p_1 p^w(m, t) \sum_{m'} p^c(m', t) p(m' | m), \quad (10)$$

以此类推, 联立 (5), (7) 公式在  $\Delta t$  趋近于 0 可以得到, 转发节点的密度随时间变化的关系  $c(t)$ :

$$\frac{\partial p^c(m, t)}{\partial t} = m p_1^2 p^w(m, t) \sum_{m'} p^c(m', t) p(m' | m), \quad (11)$$

联立(5),(8)公式在  $\Delta t$  趋近于 0 可以得到, 只读节点的密度随时间变化的关系  $g(t)$ :

$$\frac{\partial p^g(m,t)}{\partial t} = mp_1 * (1 - p_1) p^w(m,t) \sum_{m'} p^c(m',t) p(m'|m) \quad (12)$$

公式(10), (11), (12)分别描述了未知节点密度函数  $W(t)$ , 转发节点密度函数  $C(t)$ , 只读节点密度函数  $G(t)$  随时间的变化情况。这三个公式称为节点的传播动力学演化方程组, 用于刻画整个网络中节点的演化力度

## 5.1.2 模型的求解

### 5.1.2.1 数据支撑

根据 twitter 社交网络的连接关系数据, 对第一列中的用户账号和其对应的粉丝账号导入 SPSS 中进行描述统计分析, 得出其频数数据。并以此为基础对频数的分布进行统计处理。spss 和 excel 对参考数据中的 835541 条记录进行分析, 统计出用户 ID 数和粉丝频数分布图等, 部分数据截图如下, 部分数据见附件。

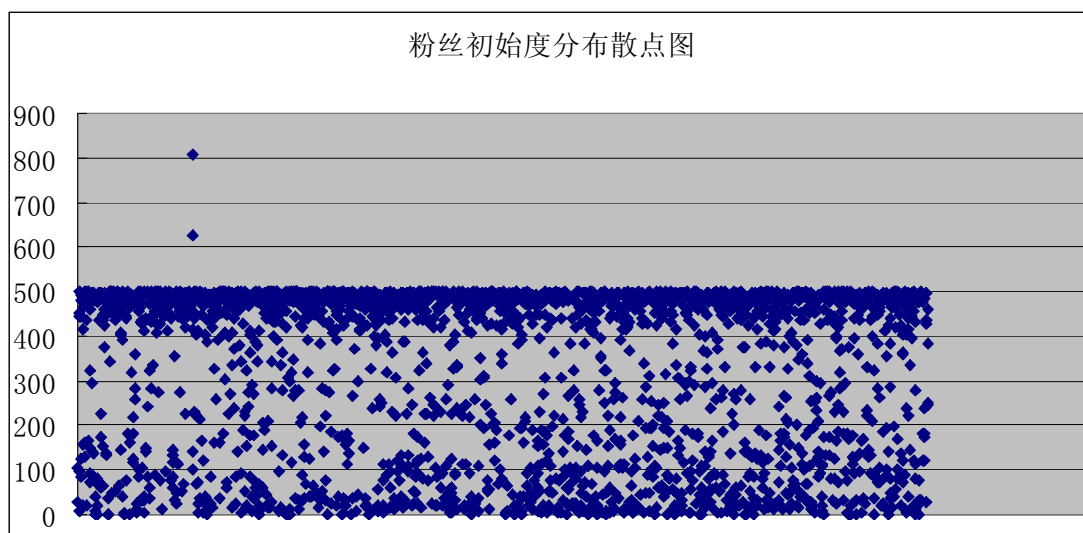
Frequencies					
Statistics					
V1					
N	Valid	835541			
	Missing	0			

粉丝数范围	频数分布
0~100	522
101~200	241
201~300	150
301~400	130
401~500	1458
600以上	2
总计:	2503

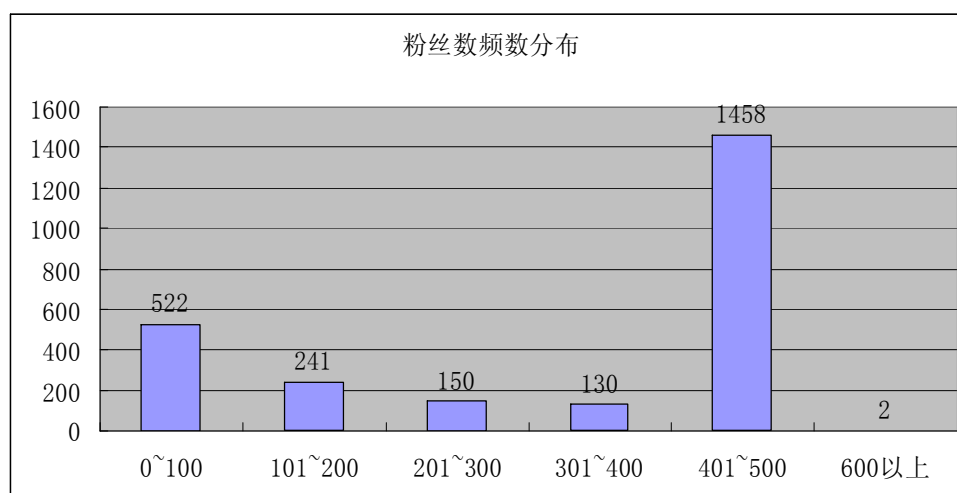
  

	Frequency	Percent	Valid Percent
Valid	104	.0	.0
_Breeanne	27	.0	.0
_wiinter	8	.0	.0
0_0kimjy	96	.0	.0
00__00	500	.1	.1
000webhosting	443	.1	.1
001dabrown	453	.1	.1
005dabrown	444	.1	.1
005matt	481	.1	.1
007brasil	85	.0	.0
007joectms	498	.1	.1
007Lifestyle	127	.0	.0
007LouiseOB	491	.1	.1
009onthebeach	466	.1	.1
00aidenrich	417	.0	.0
00joe	488	.1	.1
00SVT	20	.0	.0
00sw	157	.0	.0
01countrygirl01	131	.0	.0
01Natalie	481	.1	.1
01watch	500	.1	.1
03dentchamp	161	.0	.0
040Milad	485	.1	.1
07BondGirlXX	435	.1	.1
07forcada	22	.0	.0
099c	471	.1	.1



基于以上处理结果，对粉丝账号的整体分布情况求均值，得每个节点的平均粉丝度  $m = 835541/2503 = 333.8158 \approx 334$ ，此为普通网络用户和兼职宣传者的粉丝基数。考虑到企业的选择偏好（假设 7），企业偏向于雇佣粉丝数较多的专业推广者，故专业推广者主要分布于粉丝高频段集中段。从粉丝频数分布图可看出，频数为 401~500 的高频集中段用户数量高达 1458，占整个用户的 58.3%。故可将受雇佣的专业推广者化为归于此部分，求的其平均粉丝度为  $m=477$ 。此部分数据即为  $P(m')$ （度为  $m'$  的节点的分布函数）的函数数据，以二维数组  $(m', P(m'))$  的形式为后面编程求解所用。

每个用户名 ID 进行粉丝数分段点频数分布统计如下表：

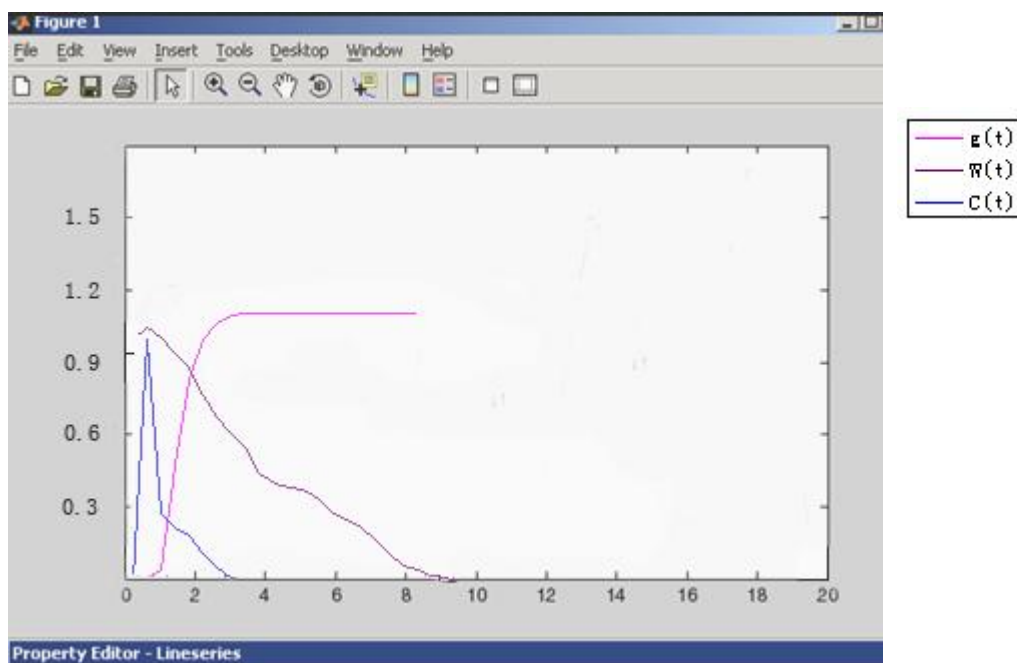


### 5.1.2.2 函数模拟

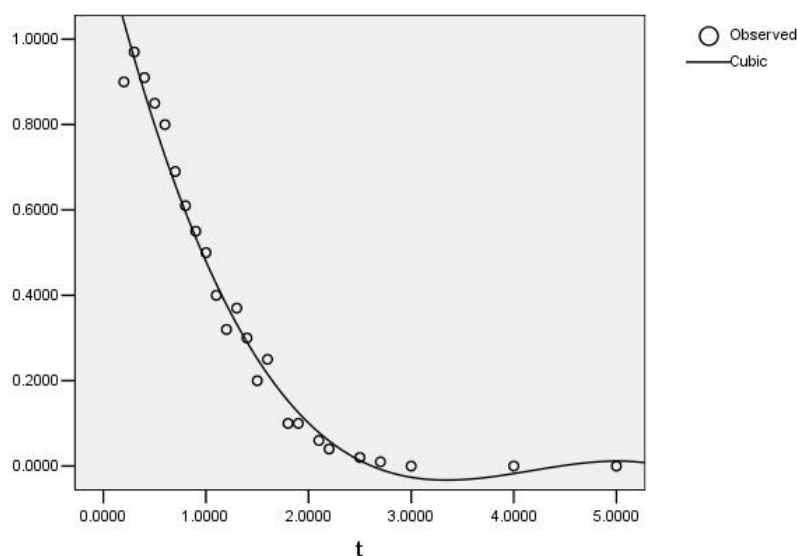
每个节点首次受到消息后转发概率  $P_1$ ，根据问题情景中海怪号（Mar Mostro）帆船发布所引起的 2.5 万互动信息中，60%转发了这张照片，此处  $P_1$  近似取值 0.6。  $P(m')$ （度为  $m'$  的节点的分布函数）的函数数据，为上述二维数组  $(m', P(m'))$ ，  $\bar{m}$  为数组中  $m'$  的平均值。对于  $m$ ，若为专业推广者则取上述高频分布段均数 477，普通用户和兼职宣传员则取全部分布段 334。



基于以上基础数据通过 C++编程, 求的  $L(m,t)$ ,  $p^w(m,t)$ ,  $\frac{\partial p^w(m,t)}{\partial t}$ ,  $\frac{\partial p^e(m,t)}{\partial t}$  等中间函数的对应数据。再将上述中间数据导入 MATLAB 中, 迭代运行模拟出  $C(t), W(t), g(t)$  随时间的变化情况, 即不同节点密度随时间的变化关系: ( $W(t)$  为未知节点的密度,  $C(t)$  为转发节点的密度,  $g(t)$  为只读节点的密度)。(部分代码见附件)



W(t)



## 5.2 消息蔓延效应预测模型

### 5.2.1 模型的建立

此部分模型以前述二次模型为基础, 以一次传播过程的基数为依据, 将一次传播分



为原有粉丝引发的二次传播和新增粉丝引发的二次传播，并对他们分别代入消息蔓延效应预测模型中求出消息收听者的蔓延情况。由于消息转发密度  $C(t)$  和只读密度  $G(t)$  演化情况不同，消息蔓延效应模型又分为消息转发者预测模型和消息接受者预测模型。

根据 twitter 社交网站的用户之间的链接关系数据，分析出普通网络用户的原始粉丝分布平均数为  $M$ ，设  $\Delta S_1$  为第一天新增的 500 名粉丝，所引起的粉丝消息蔓延效应（粉丝间消息不停转发传播）。在第  $t$  天时新增的消息传播者； $\Delta K_t$  第  $t$  天由普通网络用户引起的新增只读者，由一个推广者在第  $t$  天新增的 500 粉丝在未来的  $100-t+1$  天中所带来的新增转发者（转发节点数）为  $\Delta T_t$ ； $C_t = C(t)$ ，即第  $t$  天消息二次传播模型的传播节点密度函数； $W_t = W(t)$ ，即第  $t$  天消息二次传播模型的未知节点密度函数，且  $1 \leq t \leq 100$ 。

专业推广者第 1 天新增的 500 粉丝后，在未来的  $100-t+1$  天里由这些粉丝所引起的新增消息转发者情况：

第 1 天：

$$\Delta S_1 = 500 * C_1 * W_1 * (M + 20)$$

第 2 天：

$$\Delta S_2 = \Delta S_1 * C_2 * W_2 * (M + 20) = 500 * C_1 C_2 W_1 W_2 * (M + 20)^2$$

第 3 天：

$$\Delta S_3 = \Delta S_2 * C_3 * W_3 * (M + 20) = 500 * C_1 C_2 C_3 W_1 W_2 W_3 * (M + 20)^3$$

第 4 天：

$$\Delta S_4 = \Delta S_3 * C_4 * W_4 * (M + 20) = 500 * C_1 C_2 C_3 C_4 W_1 W_2 W_3 W_4 * (M + 20)^4$$

.....

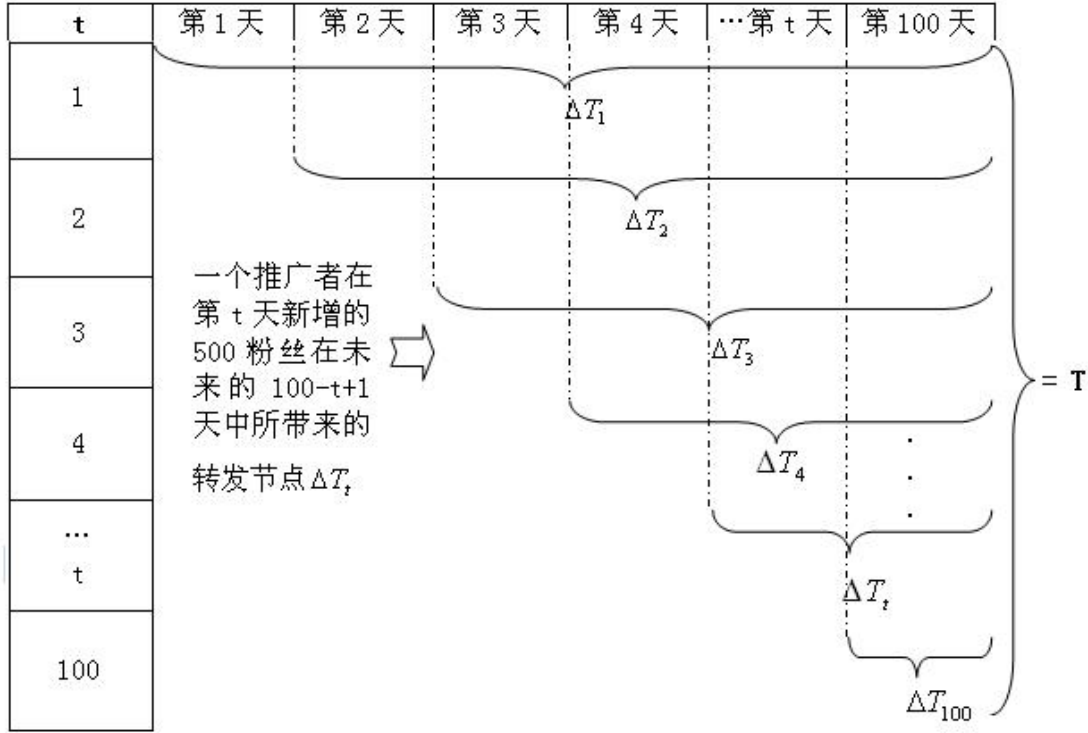
第  $t$  天：

$$\Delta S_t = \Delta S_{t-1} * C_t * W_t * (M + 20) = 500 * C_1 C_2 C_3 \dots C_t W_1 W_2 W_3 \dots W_t * (M + 20)^t$$

第 100 天：

$$\Delta S_{100} = \Delta S_{99} * C_{100} * W_{100} * (M + 20) = 500 * C_1 C_2 C_3 \dots C_t C_{100} W_1 W_2 W_3 \dots W_t W_{100} * (M + 20)^{100}$$

下图为计算转发节点总数  $T$  的过程图：



通过第 1 天新增的 500 粉丝在未来的 100 天内总的消息转发者为：

$$\Delta T_1 = \Delta S_1 + \Delta S_2 + \Delta S_3 + \dots + \Delta S_{100}$$

相应的，通过第 2 天新增的 500 粉丝在剩余的 100 天内总的消息转发者为：

$$\Delta T_2 = \frac{(\Delta S_2 + \Delta S_3 + \dots + \Delta S_{100})}{C_1 W_1 (M + 20)}$$

通过第 3 天新增的 500 粉丝在剩余的 100 天内总的转发者为：

$$\Delta T_3 = \frac{\Delta S_3 + \Delta S_4 + \dots + \Delta S_{100}}{C_1 C_2 W_1 W_2 (M + 20)^2}$$

...

通过第 t 天新增的 500 粉丝在剩余的 100 天内总的转发者为：

$$\Delta T_t = \frac{\Delta S_t + \Delta S_{t+1} + \dots + \Delta S_{100}}{C_1 C_2 \dots C_{t-1} W_1 W_2 \dots W_{t-1} (M + 20)^{t-1}}$$

显然通过第 100 天新增的 500 粉丝在当天内总的转发者为：

$$\begin{aligned} \Delta T_{100} &= \frac{\Delta S_{100}}{C_1 C_2 \dots C_{99} W_1 W_2 \dots W_{99} (M + 20)^{t-1}} \\ &= 500 * C_{100} * (M + 20) \end{aligned}$$

则这 100 天内产生的总的转发者为：

$$T = \Delta T_1 + \Delta T_2 + \Delta T_3 + \dots + \Delta T_i + \dots + \Delta T_{100}$$

到此 T 即为专业推广者在 100 内，每天新增的 500 名粉丝在所引起的消息转发者规模。再以专业推广者  $m=477$  的粉丝基数，替代上述过程中的 500，求的新的  $\Delta T_{477}(1)$ ，即为第 1 天收到消息的原始基础粉丝所引发的消息蔓延效应，持续 100 天后达到的转发者规模。

将上述转发节点密度函数  $C_t$  替换为只读节点密度函数  $G_t$ ，可得消息蔓延引起的只读者扩散情况。再以  $m=477$  的粉丝基数，带入 G 求解过程中，替代 500，得到新的  $\Delta G_{477}(1)$ ，即为第 1 天收到消息的原始基础粉丝所引发的消息蔓延效应，持续 100 天后达到的只读者规模。

经过 100 天的传播后，总的普通网络用户接收者的数目为： $T_{500} + \Delta T_{477}(1) + G_{500} + \Delta G_{477}(1)$ ，即经过一名专业推广者推广后，一条含有企业广告的奥运新闻可以被  $T_{500} + \Delta T_{477}(1) + G_{500} + \Delta G_{477}(1)$  个人看到。

将整个上述过程的（推广能力，粉丝基数）由专业推广者的（500，477）替换为兼职宣传者的（35，334）即可得到兼职宣传者的推广效益蔓延模型  $T_{35} + \Delta T_{334}(1) + G_{35} + \Delta G_{334}(1)$ 。

## 5.2.2 编程实现

根据以上推导过程，用 C++ 编写程序实现。大致思路如下：

根据模型一模拟出的  $C(t)$ ,  $W(t)$ ,  $g(t)$  函数，考虑到 C++ 运行过程对小数的识别精度问题（例如 0.000076 在运算中直接识别为 0）以及数据本身的区分度问题，做如下简化：8 天以内单独取值（第  $t$  天的  $C(t)$ ,  $W(t)$ ,  $g(t)$  值取当天的平均值）。8~30 天取第 8 天的值。30 天以后取趋势值（即极限）。

分别编写函数  $S(\text{num}, k, t)$ ,  $T(\text{num}, k, t)$  等重复迭代求解。代码见附件。

带入数据后求的：

$$T_{500} + \Delta T_{477}(1) + G_{500} + \Delta G_{477} = 518557$$

$$T_{35} + \Delta T_{334}(1) + G_{35} + \Delta G_{334} = 193791$$

初略处理后数据如下表：

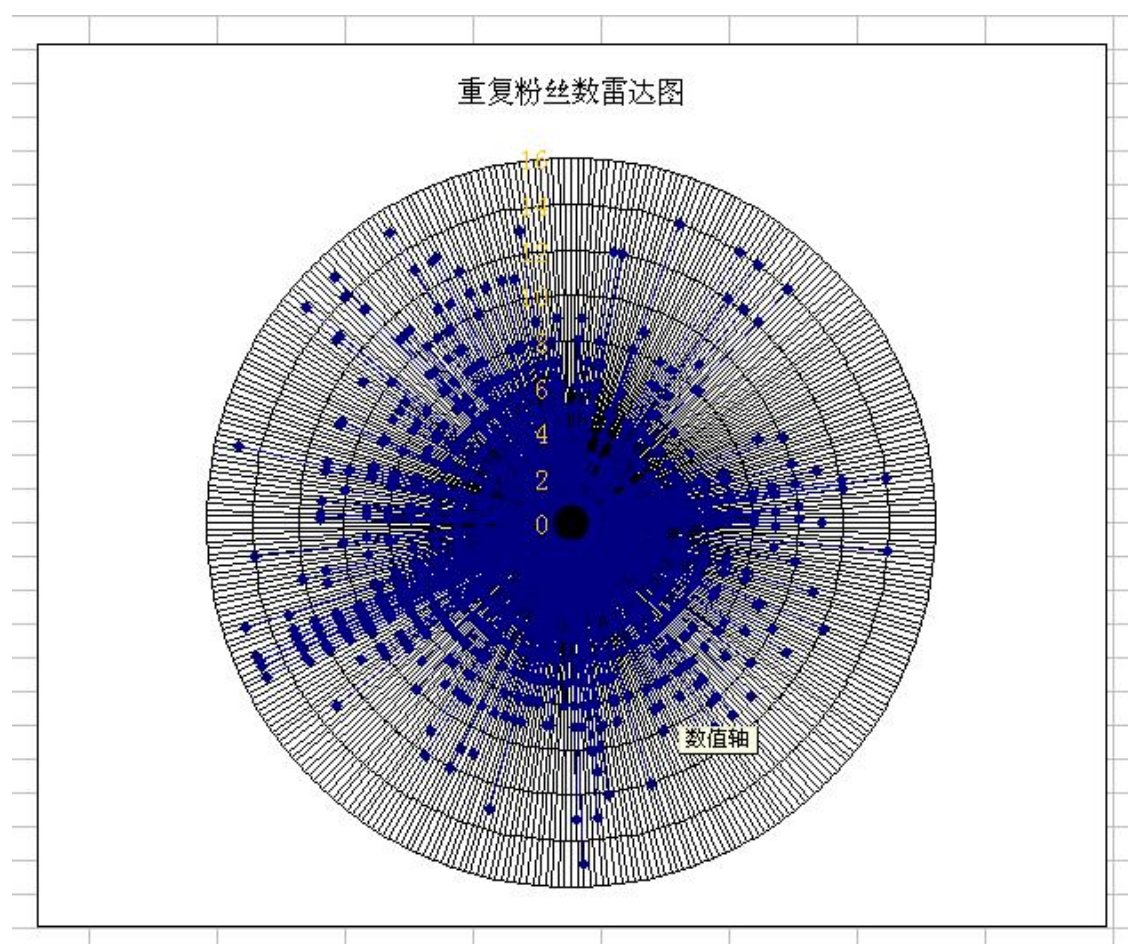
t	1	2	3	4	5	6	7	8~30	30 以后
C(t)均值	0.720 95	0.38 72	0.1704 5	0.069 2	0.026 7	0.0108	0.0044 5	0.001 3	0
W(t) 均值	0.563 19	0.06 71	0.0051 61	0.002 26	0.001 61	0.0003 87	0.0003 23	0.000 26	0
g(t)均值	0.484 5	0.59 63	0.8298 5	0.936 05	0.979 5	0.9938 5	0.9990 5	0.999 2	1

## 5.3 消息覆盖规划模型

### 5.3.1 重复率 L 分析

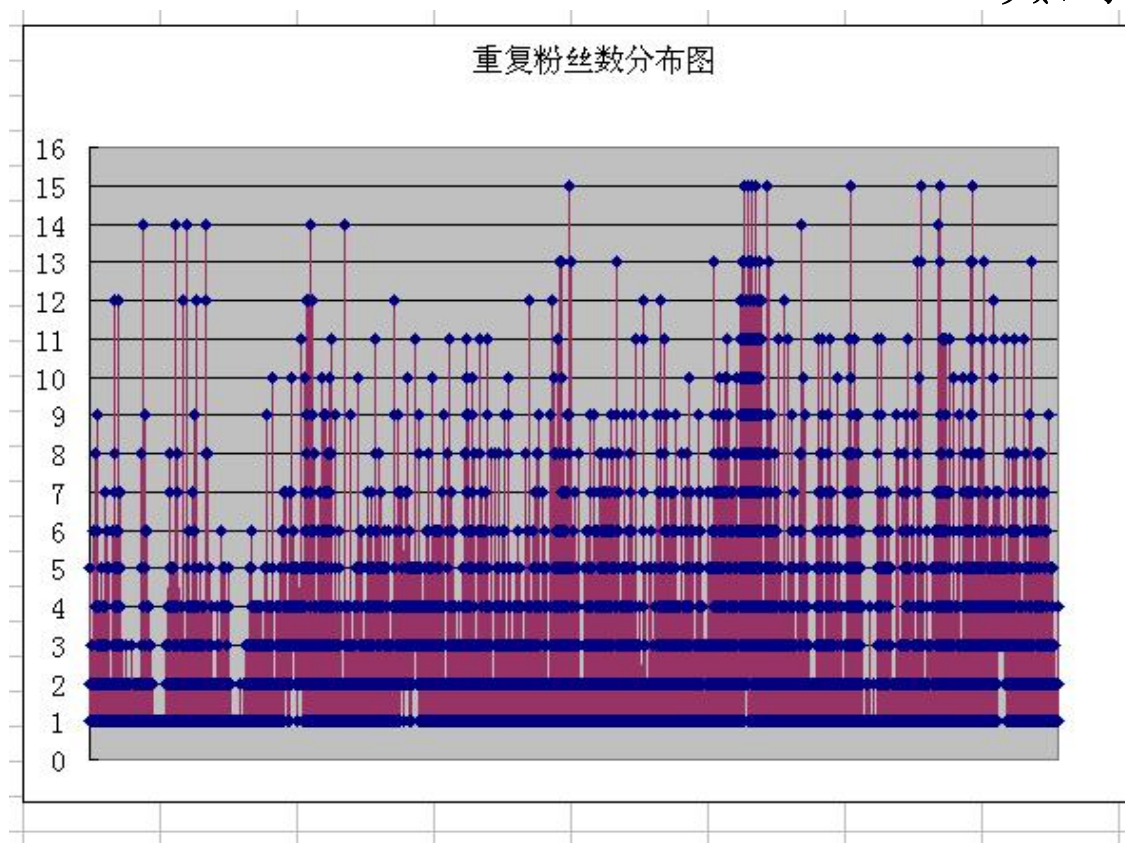
L 为各个专业推广者的所有的消息接受者的重复率，该重复率根据 twitter 社交网站用户之间的链接关系数据分析研究得出。

选取 6 万粉丝数据为统计样本，用 excel 进行重复度分析，得到样本粉丝在重复度上分布情况，表示样本粉丝重复的次数，通过 excel 中的函数 countif ( ) 和演化，得到以下雷达图和折线图：



通过用软件 spss 分析样本数据，重复度在 1~15 之间的频数分析如下：

重复度 ( $n_k$ )	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
频数 ( $f_k$ )	41456	8268	4182	2696	1195	822	462	344	216	130	99	48	52	14	15



则平均重复度为：

$$\bar{n} = \frac{\sum_{k=1}^{15} n_k * f_k}{\text{样本数量}}$$

$$\bar{n} = \frac{1*41456 + 2*8268 + 3*4182 + 4*2696 + 5*1195 + 6*822 + \dots + 13*52 + 14*14 + 15*15}{60000}$$

那么，重复率

$$L = \frac{\bar{n} * \text{样本数量} - \text{样本数量}}{\text{样本数量}} = \bar{n} - 1$$

通过计算得出：L=0.737%

此处计算出的重复率 L 是在样本数据量的范围下推拟出的一个值。

### 5.3.2 线性规划模型

企业产品的潜在用户大约有2 亿人，他们都在使用社交网络，专业推广者每天的工资是500 元，还可以从网络上雇佣兼职宣传者，每天仅需要付50 元的工资，但他们平均每天新增的粉丝数仅为35 人广告宣传需要覆盖企业潜在用户40% 的人群，求出需要雇佣专业社交网络推广者的最小人数；在成本最优的条件下为企业制定一份合理的用人方案

基于前述模型结果，为使宣传覆盖 40%用户企业至少需要的专业推广者的人数为：

$$2 * 10^8 * 40\% / [T_{500} + \Delta T_{477}(1) + G_{500} + \Delta G_{477} * (1 - L)] = 8 * 10^7 / [518557 * (1 - 0.737\%)] = 155.4 \approx 156$$

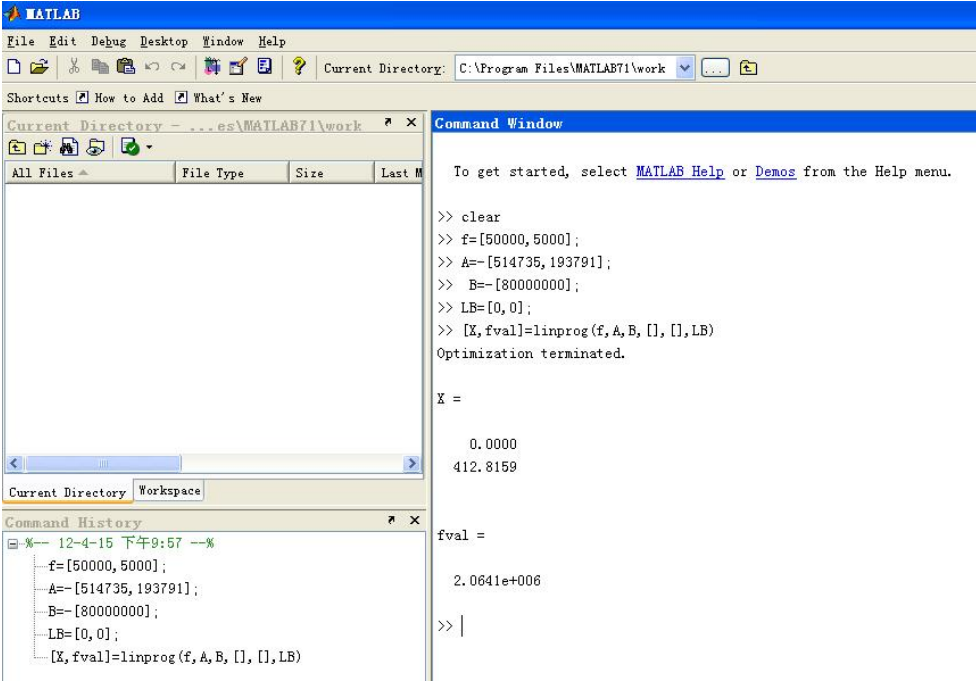
$$(T_{500} + \Delta T_{477}(1) + G_{500} + \Delta G_{477}) * (1 - L) = 514735$$

$$(T_{35} + \Delta T_{334}(1) + G_{35} + \Delta G_{334}) * (1 - L) = 193791$$

根据原问题情景，列出线性规划模型如下：

$$\begin{aligned} \min Z &= 500 * x + 50 * y \\ s.t. &\begin{cases} 514735 * x + 193791 * y \geq 8 * 10^7 \\ 0 \leq x \leq 156 \\ 0 \leq y \leq 413 \end{cases} \end{aligned}$$

用 MATLAB 求线性规划的最优解：MATLAB 命令的标准形是求目标函数的最小值：



The screenshot shows the MATLAB Command Window with the following commands and output:

```
>> clear
>> f=[50000,5000];
>> A=[514735,193791];
>> B=[80000000];
>> LB=[0,0];
>> [X,fval]=linprog(f,A,B,[],[],LB)
Optimization terminated.

X =
    0.0000
   412.8159

fval =
   2.0641e+006

>> |
```

The Command History window shows the same commands entered:

```
f=[50000,5000];
A=[514735,193791];
B=[80000000];
LB=[0,0];
[X,fval]=linprog(f,A,B,[],[],LB)
```

得到用人方案为：雇佣 413 名兼职推广者，所需成本为：2064100

## 七、模型的综合评价

### 7.1 模型优点

本模型是在研究了社交网络信息流动的基础上，以传染病模型（SIR），微分方程模型为参考，结合网络拓扑结构和传播理论的基础上构造出的一个基于 SNS 网络的消息传播模型。总模型又分为二次传播模型，消息蔓延预测模型，线性规划模型三个子模型。该模型具有如下特点：

- ① 该模型采用了自上而下的结构化方法，基于总体功能，将整体化为三个版块。各模块按层次相互依存，嵌套。各模块的参数化，独立化便于根据网络特征参数，人员推广能力，粉丝技术重新调整。这使得模型可适用于不同网络，不同的推广员。
- ② 该模型结合了网络拓扑结构，通过节点度的引入，考虑了网络自身的结构特征，使模型更符合特定 SNS 网络的特征。
- ③ 模型考虑了企业的选择偏好，将专业推广者划入高频粉丝用户段，区别对待于普通网络用户。这种对企业更乐意雇佣粉丝较高的推广员的情况更符合实际。
- ④ 模型在对连接情况重复频数分析的基础上，考虑了不同推广者引发消息推广中的粉丝重复度。
- ⑤ 模型考虑了传播过程中的边际转发率递减规律，结合微分方程、传染病动力学等考虑了分布密度增长长度对于时间的变化情况。

## 7.2 模型的不足点

本模型在抽象处理，数据处理过程中，简化考虑了一些影响因素。主要有：

- ① 模型假设不完全成立，存在特殊情况。没有考虑时间延迟，除状态更新方式外的其他接收情况等。
- ② 没考虑由转发者再非首次看到消息而转发的情况，而引起的时间变化。
- ③ 对于个体的单次转发频率，取得是经验统计数据，缺乏理论支撑。
- ④ 为简化运算对部分数据进行了一定放缩，影响了结果的精确程度。
- ④ 模型中时间变量取值均为 100 天，没考虑基于时间变化的消息扩散模型。在

$T_{500} = \sum_{t=1}^{100} \Delta T_{500}(t) = \sum_{t=1}^{100} \sum_{t=1}^{100} \Delta S_{500}(C(t), W(t))$  系列公式中，若将整个迭代过程中的 100 替换为变量 t, 即可演化完成改进。此处基于时间关系和复杂度不予考虑。

## 八、模型的改进方案

### 8.1 改进一：用幂律分布和 BA 模型来改 P1

#### 8.1.1 不足分析

P<sub>1</sub>: 未知节点转发消息，转化为传播节点的概率,即普通用户接受消息后转发该条消息的概率。前述“消息在普通用户间的二次传播模型”求解中，考虑到问题的复杂度，对于 P<sub>1</sub> 的处理是依据问题情景中海怪号案例“一张海怪号帆船照片引起的 2.5 万条互动信息中，有 60%是转发这张照片”，从而近似取值 0.6。这种近似经验取值法有欠妥当，对于 P<sub>1</sub> 的理论数值可通过引入幂律分布和 BA 模型求解。

#### 8.1.2 P1 的幂律分布和 BA 模型的演化算法

##### 8.1.2.1 情景分析

题中社交网络的专业推广者平均每天可以新增 500 个粉丝，这些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们，普通网络用户平均每天可以新增



20 个粉丝。从题中可看出，专业推广者和普通用户的粉丝增长方式类似指数增长及幂律分布。粉丝对收听者的选择基本符合无标度网络模型（BA模型），在这个社交网络下的信息传播过程中，收听者和传播者都是不断增加的，而收听者总是倾向于选择具有很多收听者的发布者。即该网络中满足，网络的节点是不断的增加的；其二就是优先连接机制，这个特性是指网络当中不断产生的新的节点更倾向于和那些连接度较大的节点相连接。

### 8.1.2.1 BA模型的生成机制

设初始网络中只有 $m$ 个节点；

(1) 生长：每隔一天，在网络中专业推广者和普通网络用户都会增加一定量的粉丝（分别为500,20），这些新增的粉丝以一定概率对奥运会新闻进行转发；

(2) 偏好连接：粉丝或用户节点 $i$ 转发新闻的概率为  $p_i = k_i / \sum_j k_j$ ，其中  $k_i$  为节点 $i$  的度， $\sum_j k_j$  为当前网络中所有节点的度之和。t个时步后，模型产生一个拥有  $N = m_t + t$  个节点和  $m_t$  条边的随机社交网络群体，随着t的增大，该社交网络群体演化进入标度不变状态。其度分布 $p(k) \sim k^{-r}$ ，其中幂律指数 $r$ 与模型中唯一的参数 $m$ 无关。

## 8.2 改进二：用均匀与幂律混合分布演化算法改进网络中粉丝新增率

### 8.2.1 不足分析

原问题情景中，对于普通网络用户粉丝新增率的取值为平均每天新增20个粉丝，即网络中每个节点的邻接节点增长速度为20个/天。此处，每一时刻向网络中添加新节点与新边的数目总是固定的，这与现实世界的某些网络的生成过程并不十分相符。实际上，每一天无论专业推广者还是普通网络用户的粉丝增加量并不是固定的，而与很多因素有关，例如当时网络的网速、客户端的硬件设备和浏览网络人员的喜好等。因此，每一天粉丝的增加数量是一个随机数似乎更确切一些。

### 8.2.2 粉丝新增速度的均匀与幂律混合分布演化算法改进

#### 8.2.2.1 情景分析

社交网络中，每一时刻加入网络的粉丝也不一定都会和当前网络中固定数目的原有普通网络用户之间建立某种关系，因为关系的建立取决于双方的意愿和其他多种客观因素。因此，每一时刻新粉丝与原有普通用户之间建立关系的数目是一个随机数似乎也更确切一些。随机数在一个区间上取值的概率相等，从而形成的概率分布是一种均匀分布。故在SNS网络幂律分布演化过程的基础上引入均匀分布。

#### 8.2.2.2 引入均匀分布的改进BA模型

对BA模型进行了如下改进：在粉丝的增长过程中，设置一个随机数 $m$ ，每天当新粉丝加入网络时，对网络中的新闻进行转发。而偏好连接机制则与BA模型完全相同，本研究中把这种演化算法暂时称为均匀与幂律混合分布演化算法，其具体算法步骤如下：

设初始社交网络为  $m_0$  个节点的完全图；

(1)生长：每隔一天，社交网络中增加一定量的粉丝数，这些粉丝将对新闻进行转发

(2)偏好连接：普通网络用户或粉丝*i*转发新闻的概率为  $\Pi_i = K_i / \sum_j k_j$ ，其中  $K_i$  为节点*i*的度，及对转发新闻的接收用户的数量， $\sum_j k_j$  为当前网络中所有看到新闻的网络用户的数量之和。经过时间*t*后，模型生成一个拥有  $N = m_0 + t$  个节点。

幂律度分布的理论分析在BA模型逐渐演化过程中，绝大多数节点保留着原有的普通网络用户及粉丝数连接或只得到了极少数新的普通网络用户或新增粉丝数的连接，因此，BA模型中绝大多数节点的度值都比较接近于其初始度值*m*。在均匀与幂律混合分布演化算法中，每个节点*i*的度值  $K_i$ （随机变量）可以分为2部分： $K_{i0}$  和  $K_{i1}$ ，其中  $K_{i0}$  为基数，即原有的普通网络用户及粉丝数， $K_{i1}$  为增长度，即新增加的普通网络用户或新增粉丝数目， $K_i = K_{i0} + K_{i1}$ 。

### 8.2.2.3 粉丝新增速度的均匀与幂律混合分布分析

#### ①均匀分布律分析

因为  $K_{i0} = m$ （每个节点的初始普通网络用户及粉丝数），在所提出的均匀与幂律混合分布演化算法中，*m* 是  $[1, m_0]$  中的随机数，因此， $K_{i0}$  也必是  $[1, m_0]$  中的随机数或随机变量，根据随机变量在一个区间上取值概率相等，所形成的概率分布是一种均匀分布的事实，很容易得到  $K_{i0}$  服从均匀分布。再由BA模型的演化规律可知，绝大多数节点必保留着原有的普通网络用户及粉丝数连接或只得到了极少数新的普通网络用户或新增粉丝数的连接，因此，该算法下绝大多数节点的度值也应该是  $[1, m_0]$  中的随机数。

②幂律分布分析在均匀与幂律混合分布演化算法中，每个节点度值的另一部分  $K_{i1}$  的增长机制与BA模型节点度值的增长情况相同，即新增普通网络用户或粉丝转发新闻的概率与原有普通网络用户或粉丝转发新闻的概率成比例，因此，随机变量  $K_{i1}$  的分布应满足幂律分布。

综合上述分析结果，可以得出结论：粉丝新增速度具有均匀和幂律双重分布2种特征。

## 九、基于社交网络平台营销提出的几点建议

根据以上的模型分析，通过查阅相关文献，我们做了一些扩展，SNS 网站的营销有自己的独特优势，它可以充分利用它的社会化关系网络和一对多的便捷沟通方式，它有真人的群组，有朋友之间的互动关系，企业和用户之间也可以建立便捷的沟通方式，对于广告主而言，也可更加容易看到品牌推广到效果。针对这些，我们总结了几点关于如何成功的开展社交网络营销的建议如下：

1、把产品或者品牌植入 SNS 用户的交互网络中。例如：腾讯公司的互赠虚拟礼物，虚拟礼物跟现实生活中的用品相对应，比如饰品、服装、化妆品、书籍等等都可以，如果说推广某款饮料，可以依照现实的模样做一个虚拟礼物，带上品牌商标置于礼物列表中，用户可以把它当作免费礼物送出，通过用户在社交网络中的相互赠送，品牌效应与产品形象就深刻于用户之心了。

2、在 SNS 的有核心影响力的节点处建立产品和品牌的群组。例如：在 msn 上建立可口可乐群，让产品和品牌在无形之中被用户接受。

3、商标饰品营销。此类营销针对博客、微博、QQ、MSN 等社交网络平台。例如：在个人空间的头像上或者是发表的微博中，若用户加入带有品牌商标的特色饰

品图案在头像和微博中，得到相应的奖励。

4、通过社交网站游戏与电子商务企业产品结合通。比如“抢车位”游戏，通过与现实中的汽车品牌结合，从产品外形、定位、价格等方面植入到具体游戏中，这里不用吸引用户进来，因为用户已经进来了，在用户喜欢做的事情上，添加一些营销素材，跟用户利益结合，用户在这样的过程中，更加容易接受产品，也更加容易做口碑，也更加容易导向购买行为。

5、通过分析社交网络上庞大的人的数据。这里的数据不仅包括人的属性，还有人的互动数据，比如说参与的投票、测试，分享的信息等，这些信息被分析后能够告诉营销决策者可能的潜在目标用户群及其特征，便于实施营销计划。

6、利用社交网络平台分享的特点，促成传染病扩散营销策略。利用 SNS 作为启动传染病营销的源点。两个步骤最关键：一是选好传播的人；二是选好传播的内容。传播的人主要是 SNS 内的有影响力用户；传播的内容要有趣并创新，且要有意义。

## 十、参考文献

- [1] 宋志刚，谢蕾蕾，何旭洪，《SPSS16 实用教程》，北京：人民邮电出版社，2008. 10
- [2] 张贤明，*《MATLAB 语言及应用案例》* ISBN，江苏：东南大学出版社，2010. 9
- [3] 董肇君，《系统工程与运筹学》，北京：国防工业出版社，2009. 1
- [4] 裴伟东，夏玮，王全来，赵子平，马希荣，《具有均匀与幂律双重混合分布的复杂网络动态演化模型的研究》，天津师范大学学报。第 30 卷第 4 期，2010. 10
- [5] 张彦超，刘云，张海峰，程辉，熊菲，《基于在线社交网络的信息传播模型》，物理学报，2010. 9
- [6] 董臻圃，《数学建模方法与实践》，北京：国防工业出版社，2006. 8
- [7] 熊启才，《数学模型方法及应用》，重庆：重庆大学出版社，2005. 3
- [8] 郭海霞，《新型社交网络信息传播特点和模型分析》，现代情报，第 32 卷第一期：56-60 页，2011. 10
- [9] 郭涛，《传统企业如何成功开展社交网络营销》，  
[http://promote.yidaba.com/201105/311017001007100100000325773\\_3.shtml](http://promote.yidaba.com/201105/311017001007100100000325773_3.shtml),  
2012. 4
- [10] Vázquez A, Weigt M 2003 Phys. Rev. E 67 027101
- [11] Fu F, Chen X J, Liu L H, Wang L 2007 arXiv: 0701323 [physics]

## 十一、附录

### 附录一：二次传播模型中 C++部分程序代码

```
using namespace std;
const int M=334;//常量 m;
const double P1=0.6;//常量 P1
const int line=1000;//数组行数
double Lmt; //L(m,t)的输出结果
double num;//推广能力
double Pwmt(double t)
{
```

```

        return pow(1-P1*t*Lmt,M);
    }

double Wt(double t)//Δw
{
    return 1-M*P1*Pwmt(t)*Lmt;
}
double Ct(double t)//Δc
{
    return P1*Wt(t);
}
double Gt(double t)//Δg
{
    return (1-P1)*Wt(t);
}
double St(double t)//Δs
{
    double tem=num*pow(M+20,t);
    for(int m=1;m<=t;m++)
        tem*=Ct(m);
    return tem;
}
double Tt(double t)//Δt
{
    double tem1=0;
    double tem2=pow(M+20,t-1);
    for(int m=100;m>=t;m--)
        tem1+=St(m);
    for(int n=1;n<=t-1;n++)
        tem2*=Ct(n);
    return tem1/tem2;
}

double T()//T
{
    double tem;
    for(int m=1;m<=100;m++)
        tem+=Tt(m);
    return tem;
}

double St_g(double t)//Δs_g
{
    double tem=num*pow(M+20,t);
    for(int m=1;m<=t;m++)
        tem*=Gt(m);
    return tem;
}

```

```

double Tt_g(double t)//Δt_g
{
    double tem1=0;
    double tem2=pow(M+20,t-1);
    for(int m=100;m>=t;m--)
        tem1+=St(m);
    for(int n=1;n<=t-1;n++)
        tem2*=Gt(n);
    return tem1/tem2;
}
double G()//T
{
    double tem;
    for(int m=1;m<=100;m++)
        tem+=Tt_g(m);
    return tem;
}

int main()
{
    ifstream in("matrix.txt");//由文件存储数组数据
    double Mp[line][2]={0}; //定义数组
    double temp=0; //临时变量
    // double Lmt=0; //L(m,t)的输出结果
    int i=0;
    int j=0;
    for(i=0;j=0;!in.eof()&&i<line;i++)
    {
        in>>Mp[i][j]>>Mp[i][j+1];
        // cout<<Mp[i][j]<<" "<<Mp[i][j+1]<<endl;
    }
    in.close();
    ifstream inf("Pcmt.txt");
    double Pcmt[line]={0}; //PC(m,t)函数输出为 Pcmt[i]
    for(i=0;!inf.eof()&&i<line;i++)
    {
        inf>>Pcmt[i];
        //cout<<Pcmt[i]<<endl;
    }
    inf.close();
    for(i=0;i<line;i++)
        temp+=Mp[i][0];
    double M_ave=temp/line;
    for(i=0,j=0,temp=0;i<line;i++)
    {
        temp+=Mp[i][j]*Mp[i][j+1]*Pcmt[i]; //
    }
    Lmt=temp/M_ave;
    // cout<<"函数 L(m,t)输出结果为"<<Lmt<<endl;
}

```

```
double t;//时间 t
do{
    cout<<"请输入时间 t(1<=t<=100) ";
    cin>>t;
    //Pwmt(t);
}while(t<1||t>100);

double num;//推广能力 35 or 500
do{
    cout<<"请输入推广能力 num(35 或者 500) ";
    cin>>num;
    //Pwmt(t);
}while(num!=35&&num!=500);
double result=0;
    result=G()+T();
cout<<"接收者人数为"<<result<<endl;
system("PAUSE");
return 0;
}
```

## 附录二：消息蔓延效应预测模型的 C++部分程序代码

```
double c (int t)
{
    int i;
    if(t<1||t>100)    printf("t wrong");
    switch (t){
    case 1: i=0.9098;break;
    case 2: i=0.5321;break;
    case 3: i=0.2419;break;
    case 4: i=0.099;break;
    case 5: i=0.0394;break;
    case 6: i=0.014;break;
    case 7: i=0.0076;break;
    default:break;
    }
    if(t<8) return (i);
    if(t<31) return 0.0013;
    if(t>30) return 0;
}
```

```
double g (int t)
{
    int i;
    if(t<1||t>100)printf("t wrong");
    switch (t){
    case 1: i=0.0497;break;
    case 2: i=0.4348;break;
    case 3: i=0.7578;break;
```

```

        case 4: i=0.9019;break;
        case 5: i=0.9702;break;
        case 6: i=0.9888;break;
        case 7: i=0.9989;break;
        default:break;
    }
    if(t<8) return (i);
    if(t<31)return(0.9992);
    if(t>30)return(1);
}

double z1 (int t){
    double j=1.0;
    int i;
    for(i=1;i<=t;i++) j*=c(t);
    return (j);
}
double z2 (int t){
    double j=1.0;
    int i;
    for(i=1;i<=t;i++) j*=g(t);
    return (j);
}
double s (int num1,char k1,int t){
    double j=1.0;
    if(k1=='c') j=num1*z1(t)*pow(354,t);
    if(k1=='g') j=num1*z2(t)*pow(354,t);
    if(k1!='c' && k1!='g') {
        printf("wrong");
    }
    return (j);
}

double T (int num2,char k2,int t){
    int i;
    double n=1.0;
    double j=1.0;
    double m=1.0;
    for(i=t;i<=100;i++) n+=s(num2,k2,i);
    if(k2=='c') for(i=1;i<=t-1;i++) m*=z1(i);
    if(k2=='g') for(i=1;i<=t-1;i++) m*=z2(i);
    if(k2!='c' && k2!='g') {
        printf("wrong");
    }
    j=n/(m*pow(354,t-1));
    return (j);
}

int main()

```



```
{
    int i,num;
    double j=1.0;
    char k;
    cin>>num;
    cin>>k;
    for(i=1;i<=100;i++)
        j+=T(num,k,i);
    cout<<j<<endl;

    return 0;
}
```

### 附录三：求解 LP 模型最优解的 MATLAB 代码

用 MATLAB 求线性规划 LP 模型最优解的代码如下：

```
>> clear
>> f= [50000,5000];
>> A=-[514735,193791];
>> B=-[800000000];
>> LB= [0,0];
>> [X,fval]= linprog (f,A,B,[],[],LB)
Optimization terminated.
```

X =

```
    0.0000
  412.8159
```

fval =

```
  2.0641e+006
```

```
>>
```