移动P2P网络中的病毒传播建模

冯朝胜[1,2]，秦志光[2]，袁 丁[1]

(1. 四川师范大学计算机科学学院 成都 610101; 2. 电子科技大学计算机科学与工程学院 成都 611731)

【摘要】对移动P2P网络的体系结构和关键协议进行了深入研究分析；在此基础上，根据病毒传播特点并利用流行病学理论，提出了移动P2P网络中病毒的传播数学模型。基于提出的病毒传播模型，使用专门的数字分析软件进行了大量仿真实验，实验主要考查了各P2P参数对病毒传播的影响。实验分析表明，通过控制固定节点的下载率和恢复率这两个影响病毒传播最关键的参数能有效遏制移动P2P病毒传播。

关 键 词 体系结构; 下载协议; 移动P2P网络; 建模; P2P病毒

中图分类号 TP309 **文献标识码** A doi:10.3969/j.issn.1001-0548.2012.01.019

Modeling of Virus Propagation in the Mobile Peer-to-Peer Networks

FENG Chao-sheng[1,2], QIN Zhi-guang[2], and YUAN Ding[1]

(1. School of Computer Science, Sichuan Normal University Chengdu 610066;

2. School of Computer Science and Engineering, University of Electronic Science and Technology of China Chengdu 611731)

Abstract The mathematical model of virus propagation in the mobile peer-to-peer(MP2P) networks is proposed by applying Epidemiology. Based on the proposed model and using the special digital-analysis software Matlab, large-scale experiments are done so as to examine the effect of the MP2P-related parameters on virus propagation. The results of these experiments illustrate that different parameters have different effects on virus propagation, among which the average downloading rate and the average recovery rate of fixed peers have the strongest effects on virus propagation.

Key words architecture; downloading protocols; mobile P2P network; modeling; P2P viruses

据业界专家预测，手机内容共享、手机游戏、移动广告、手机电视、移动社交、移动支付将是未来移动互联网的几大前景看好业务。手机图片、音频、视频等文件共享被认为是未来3G手机业务的重要应用。随着终端、内容和网络3个方面制约问题的解决，手机共享服务将快速发展，用户利用共享服务，可以上传自己的图片、视频至博客空间，还可以用其备份文件，与好友共享，或者公开发布。这样的背景下，集成了P2P网络文件共享诸多优点和移动网络自由灵活优点的移动P2P网络应运而生。

然而，与固定网络中的P2P应用一样，移动P2P网络也面临病毒的严重威胁。目前，在P2P网络和移动互联网络中出现的病毒已有数百种之多[1]。研究病毒的传播与控制，首要任务就是提出能较准确反映蠕虫传播趋势的传播数学模型。本文在分析移动P2P网络的体系结构、关键协议和移动P2P病毒传播特点的基础上，提出了移动P2P病毒的传播模型，并基于该模型分析了各相关参数对病毒传播的影响。

1 相关研究

1.1 P2P病毒传播机制与模型

下载是P2P文件共享网络提供给用户的最重要功能，而P2P病毒正是利用用户正常的下载活动进行传播的。与互联网上的病毒一样，P2P病毒总是依附在文件上，一旦感染了病毒的文件被下载并被执行，病毒就会在共享文件夹中创建多个病毒文件拷贝。为了加快病毒的传播，新生成的病毒文件拷贝通常会用流行文件名来命名。

固定P2P网络上的病毒已经引起了研究人员的关注。文献[2]对P2P文件共享网络上的病毒传播和感染文件传播分别进行了建模。文献[3]对非扫描型P2P蠕虫进行了仿真分析。文献[4]利用数字模拟方

收稿日期：2010 – 04 – 29; 修回日期：2011 – 02 – 04

基金项目：国家自然科学基金(60873075)；国家863高技术研究发展计划基金(2009AA01Z422)；西南交通大学信息编码与传输四川省重点实验室开放研究基金(2010-05)，四川省科技厅应用基础项目(2010JY0125)；四川省教育厅重点课题(10ZA007)

作者简介：冯朝胜(1971 –)，男，博士，主要从事网络与信息安全等方面的研究.

法分析了P2P系统参数对被动式P2P蠕虫传播的影响。

病毒在移动网络上一出现，就引起了人们的重视。通过对跟踪收集的数据进行分析和实验，文献[5]发现蓝牙病毒近几年有可能大规模爆发。文献[6]对手机病毒进行分析并利用流行病学建立了手机病毒传播的数学模型。文献[7]对利用蓝牙技术进行传播的病毒进行深入分析，提出了该类病毒的传播模型。但是，迄今为止，尚无人提出移动P2P病毒传播的数学模型。

1.2　移动P2P文件共享网体系结构和下载协议

要实现移动P2P文件共享系统，首先要解决系统体系结构的设计问题。设计移动P2P网络体系结构需要解决终端能力和网络能力的局限性问题：1) 在网络能力方面，有无线网络传输环境、设计技术等限制；2) 在终端能力方面，有终端大小、处理能力、电池容量等限制。

一种比较实用的移动P2P文件共享系统的体系结构设计[8]如图1所示。在该结构中，超级节点(super peer)构成核心骨干网络，移动节点通过具有无线通信能力的超级节点进行文件的请求和下载。超级节点实际上是移动节点的代理，考虑到代理对带宽、电源和处理能力都有较高要求，代理由固定节点充当。已有的移动P2P网络基本都使用该结构。

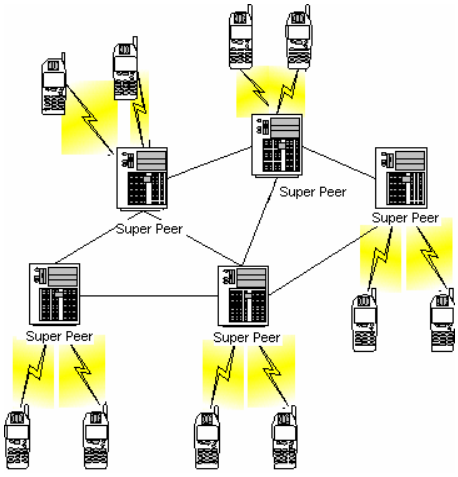

图1　移动P2P网络的体系结构

移动P2P节点下载文件的过程如图2所示。当移动节点要下载文件时，首先向代理节点发出文件搜寻请求，代理节点根据移动节点文件搜寻请求内容，使用固定P2P网络协议在P2P网络中搜寻请求文件。搜寻到相关文件后，代理节点将搜寻到的文件节点列表，通过无线网络返回给移动节点。移动节点从列表中选择下载文件，并向代理节点发送下载文件请求。代理节点在固定P2P网络中执行下载协议，将请求文件下载到本机，之后将请求文件上传给请求文件的移动终端。

移动P2P网络包含固定节点和移动节点两种。在固定节点，每个用户都有一个共享文件夹，用户将所有可共享的文件都放到共享文件夹以便其他用户共享。网络中的任何用户都可以从其他任意一个用户的共享文件夹中下载文件，当用户想要下载某个文件时，会发出搜索文件请求。无论采用哪种搜索算法，请求文件用户最终都会收到与请求相匹配的文件节点列表。获取列表后，用户可以从中选择一个或多个节点下载该文件。下载的文件被放在共享文件夹中，可被网络中其他节点下载。由于网络及移动终端的局限性，文件被移动终端下载后，不能供其他固定或移动节点下载。

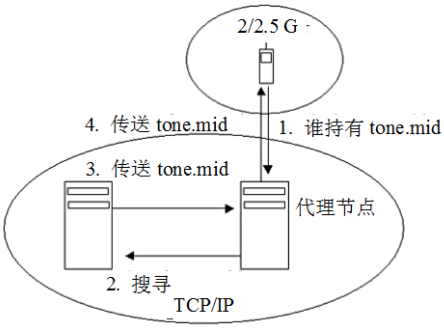

图2　移动节点文件的下载过程

2　P2P病毒传播建模

2.1　建模变量和参数

用m和f分别表示移动节点和固定节点。任何时刻节点都只能处于下面的一种状态之中：

1) 易感的，即节点的P2P共享文件夹中无病毒文件；

2) 暴露的，即节点的P2P共享文件夹中有病毒文件，但病毒尚未被激活；

3) 感染的，即节点的P2P共享文件夹中有病毒文件，且病毒已被激活。

建模时用到的变量和参数如表1所示。

表1　模型中的变量和参数

变量	参　　数
$S_f(t)$	t 时刻易感固定节点数，$S_f(0)$=9 900
$E_f(t)$	t 时刻暴露固定节点数，$E_f(0)$=0
$I_f(t)$	t 时刻感染固定节点数，$I_f(0)$=100
$S_m(t)$	t 时刻易感移动节点数，$S_m(0)$=100 00
$E_m(t)$	t 时刻暴露移动节点数，$E_m(0)$=0
$I_m(t)$	t 时刻感染移动节点数，$I_m(0)$=0
$K_f(t)$	t 时刻固定节点上感染文件数，$K_f(0)$=100 00
$M_f(t)$	t 时刻固定节点上干净文件数，$M_f(0)$=100 000

（续表）

变量	参 数
$h(t)$	t 时刻下载病毒文件的概率，$h(t)=\dfrac{K_f(t)}{M_f(t)+K_f(t)}$
d_f	单位时间内每个固定节点下载文件的平均个数，d_f=0.003
d_m	单位时间内每个移动节点下载文件的平均个数，d_m=0.003
e_f	单位时间内暴露固定节点打开病毒文件的概率，e_f=0.8
e_m	单位时间内暴露移动节点打开病毒文件的概率，e_m=0.8
r_f	单位时间内恢复为易感状态的感染固定节点比例，r_f=0.0001
r_m	单位时间内恢复为易感状态的感染移动节点比例，λ_r=0.0001
f_{e_i}	单位时间内暴露固定节点成功打开病毒文件的概率，f_{e_i}=0.8
m_{e_i}	单位时间内暴露移动节点成功打开病毒文件的概率，m_{e_i}=0.5
c	打开病毒文件时生成的病毒文件拷贝数，c=100
I_0	$I(t)$在0时刻的值，$I_0=I(0)$

2.2 病毒传播模型

下面分别分析易感节点数、暴露节点数、感染节点数等参数的变化率。

2.2.1 易感节点的变化率

t时刻易感移动节点的数量为$S_m(t)$，每个易感移动节点因下载病毒文件而成为暴露节点的概率为 $d_m h(t)$，因而单位时间内有 $d_m S_m(t)h(t)$ 个易感移动节点成为暴露节点。在单位时间内，有 $r_m I_m(t)$ 个感染节点因清除了病毒文件而恢复为易感状态。与此同时，在单位时间内，要打开感染文件的暴露节点数为 $e_m E_m(t)$，而暴露节点不能成功打开病毒文件的概率为 $(1-m_{e_i})$，故单位时间内恢复为易感状态的暴露节点数为 $e_m E_m(t)(1-m_{e_i})$，易感移动节点的变化率为：

$$\frac{\mathrm{d}S_m(t)}{\mathrm{d}t}=-d_m S_m(t)h(t)+r_m I_m(t)+e_m E_m(t)(1-m_{e_i})$$

基于同样的分析，易感固定节点的变化率为：

$$\frac{\mathrm{d}S_f(t)}{\mathrm{d}t}=-d_f S_f(t)h(t)+r_f I_f(t)+e_f E_f(t)(1-f_{e_i})$$

2.2.2 暴露节点的变化率

单位时间内有 $d_m S_m(t)h(t)$ 个易感移动节点转变为暴露节点，而单位时间内有 $e_m E_m(t)$ 个暴露节点打开染毒文件，无论成功与否，它们的状态都要发生转变，所以暴露移动节点的变化率为：

$$\frac{\mathrm{d}E_m(t)}{\mathrm{d}t}=d_m S_m(t)h(t)-e_m E_m(t)$$

暴露固定节点的变化率为：

$$\frac{\mathrm{d}E_f(t)}{\mathrm{d}t}=d_f S_f(t)h(t)-e_f E_f(t)$$

2.2.3 感染节点的变化率

单位时间内有 $e_m E_m(t)$ 个暴露移动节点要打开病毒文件，成功打开病毒文件的概率为 m_{e_i}，所以有 $e_m E_m(t)m_{e_i}$ 个移动节点转变为感染节点；与此同时，有 $r_m I_m(t)$ 个感染移动节点恢复成易感状态，故感染移动节点的变化率为：

$$\frac{\mathrm{d}I_m(t)}{\mathrm{d}t}=e_m E_m(t)m_{e_i}-r_m I_m(t)$$

感染固定节点的变化率为：

$$\frac{\mathrm{d}I_f(t)}{\mathrm{d}t}=e_f E_f(t)f_{e_i}-r_f I_f(t)$$

2.2.4 染毒共享文件的变化率

显然，P2P网络中被共享的病毒文件越多，下载到病毒文件的概率就越大。根据前面的分析，移动节点只能下载文件而不能上传文件，因此，下载病毒文件的概率只与固定节点上的病毒文件数和干净文件数相关，于是有 $h(t)=\dfrac{K_f(t)}{M_f(t)+K_f(t)}$。在时刻$t$，一个固定节点下载病毒文件的个数为 $d_f h(t)$，$S_f(t)$ 个易感固定节点下载的病毒文件数为 $d_f S_f(t)h(t)$，成功打开病毒文件的暴露主机使得增加的病毒文件数为 $e_f E_f(t)f_{e_i}(c-1)$，而打开时被反病毒软件清除的病毒文件数为 $e_f E_f(t)(1-f_{e_i})$；在单位时间内，用户删除或用反病毒软件清除的病毒文件数为 $r_f I_f(t)c$，所以染毒文件的变化率为：

$$\frac{\mathrm{d}K_f(t)}{\mathrm{d}t}=d_f S_f(t)h(t)+e_f E_f(t)f_{e_i}(c-1)-e_f E_f(t)(1-f_{e_i})-r_f I_f(t)c$$

2.2.5 干净文件的变化率

易感固定节点下载病毒文件的概率为 $h(t)$，下载干净文件的概率为 $1-h(t)$。每个固定节点下载的干净文件数为 $d_f(1-h(t))$，$N_f(t)$ 个固定节点共下载 $d_f N_f(t)(1-h(t))$ 个干净文件。于是，固定P2P网络中干净文件的变化率为：

$$\frac{\mathrm{d}M_f(t)}{\mathrm{d}t}=d_f N_f(t)(1-h(t))$$

综合以上对建模变量和参数，以及与建立模型相关的参数变化率的分析，可得移动P2P网络的病毒传播模型为：

$$\frac{\mathrm{d}S_f(t)}{\mathrm{d}t}=-d_f S_f(t)h(t)+r_f I_f(t)+e_f E_f(t)(1-f_{e_i}) \quad (1)$$

$$\frac{\mathrm{d}E_f(t)}{\mathrm{d}t}=d_f S_f(t)h(t)-e_f E_f(t) \quad (2)$$

$$\frac{\mathrm{d}I_f(t)}{\mathrm{d}t}=e_f E_f(t)f_{e_i}-r_f I_f(t) \quad (3)$$

$$\frac{\mathrm{d}S_m(t)}{\mathrm{d}t}=-d_m S_m(t)h(t)+r_m I_m(t)+e_m E_m(t)(1-m_{e_i}) \quad (4)$$

$$\frac{\mathrm{d}E_m(t)}{\mathrm{d}t}=d_m S_m(t)h(t)-e_m E_m(t) \quad (5)$$

$$\frac{\mathrm{d}I_m(t)}{\mathrm{d}t}=e_mE_m(t)m_{e_i}-r_mI_m(t) \tag{6}$$

$$\frac{\mathrm{d}K_f(t)}{\mathrm{d}t}=d_fS_f(t)h(t)+e_fE_f(t)f_{ei}(c-1)-e_fE_f(t)(1-f_{ei})-r_fI_f(t)c \tag{7}$$

$$\frac{\mathrm{d}M_f(t)}{\mathrm{d}t}=d_fN_f(t)(1-h(t)) \tag{8}$$

$$N_f(t)=S_f(t)+E_f(t)+I_f(t) \tag{9}$$

$$h(t)=\frac{K_f(t)}{M_f(t)+K_f(t)} \tag{10}$$

3　实验与分析

3.1　实验说明

考虑到之前所提出的移动P2P网络中的病毒传播数学模型是很难直接求解的非线性微分方程组，而数值分析工具Matlab提供的组件Simulink可以仿真求解非线性微分方程组，因此，可基于本文的病毒传播数学模型，使用Matlab分析各个P2P参数对病毒传播的影响。为了方便考查各参数对病毒传播的影响，将同一参数不同取值对应的传播曲线在同一个图中进行比较。在不作特别说明的情况下，实验中变量的初值和参数的值就是表1所给出的值，实验中的单位时间为1×10^2 min。

3.2　实验结果分析

无论是固定节点的下载率F还是移动节点的下载率M，对病毒传播都有较大影响。固定节点下载率和移动节点下载率对病毒传播的影响分别如图3和图4所示。由图可以看出，固定节点下载率的变化对固定节点和移动节点的感染情况都有较大影响；移动节点的下载率只影响病毒在移动节点中的传播，对病毒在固定节点的传播没有影响。这是因为固定节点下载率的变化会引起下载病毒文件概率$h(t)$的变化，下载病毒文件概率的变化对病毒在固定节点和移动节点上的传播都会产生影响；移动节点并不提供文件上载功能，移动节点下载率的变化不会引起$h(t)$的变化，因而对病毒在固定节点中的传播不会造成影响。但是，移动节点下载率的变化对移动节点的感染情况会造成影响，而且移动节点下载率越大，意味着易感移动节点在单位时间内下载的次数就越多，在单位时间内下载到病毒文件的概率也越大，病毒传播也就越快。

暴露固定节点和暴露移动节点病毒文件打开率对病毒传播的影响分别如图5和图6所示。从直觉上判断，病毒文件打开率对病毒传播有较大影响，而且病毒文件打开率越大，病毒传播越快。但实验结果表明，无论是固定节点病毒文件打开率的变化，还是移动节点病毒文件打开率的变化，对病毒的传播几乎没有影响。

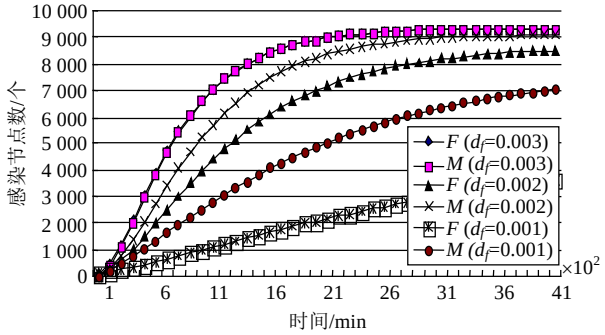

图3　固定节点下载率对病毒传播的影响

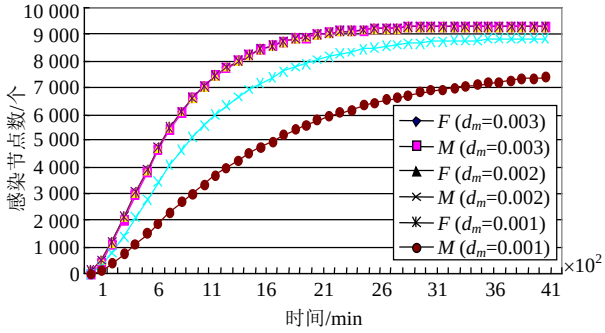

图4　移动节点下载率对病毒传播的影响

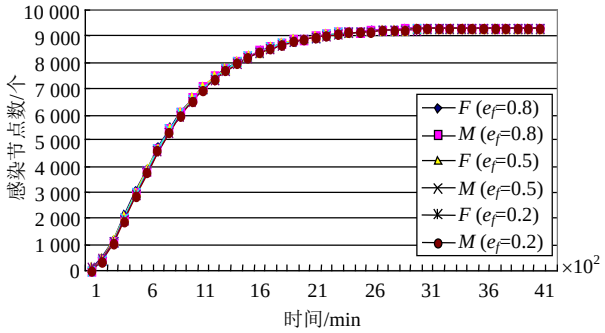

图5　暴露固定节点病毒文件打开率对病毒传播的影响

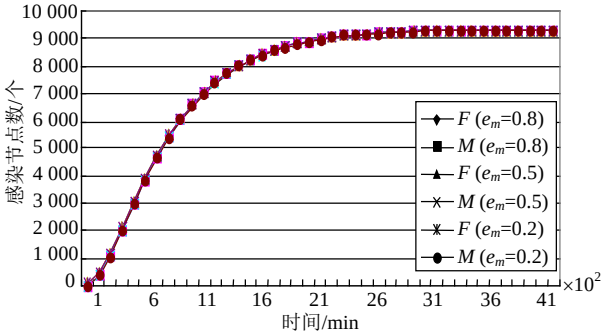

图6　暴露移动节点病毒文件打开率对病毒传播的影响

固定节点和移动节点病毒文件打开成功率对病毒传播的影响分别如图7和图8所示。图5表明，固定节点病毒文件打开成功率对病毒在移动节点和固定节点中的传播的影响都较大。打开成功率越高，病

毒传播就越快，感染高峰值也就越大。图6表明，根据移动P2P网络的特点，移动节点病毒文件打开成功率的变化，对下载病毒文件概率不会造成影响，因而对病毒在固定P2P节点中的传播也没有影响；但成功打开的病毒文件越多，被感染的暴露移动节点也就越多。

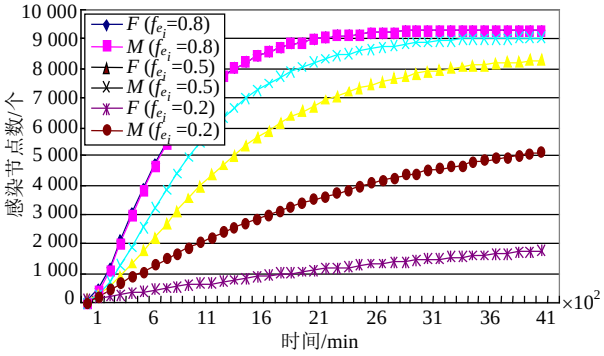

图7 固定节点病毒文件打开成功率对病毒传播的的影响

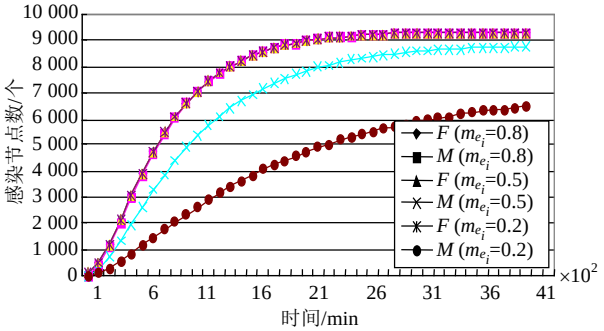

图8 移动节点病毒文件打开成功率对病毒传播的的影响

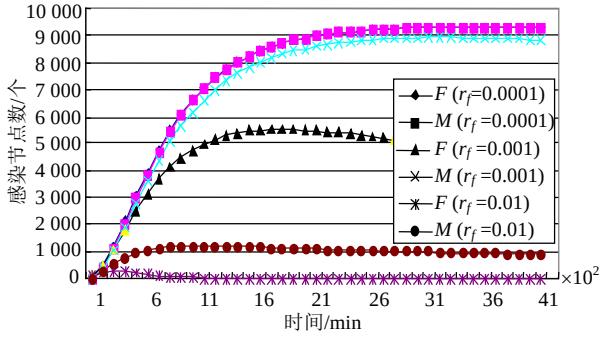

图9 固定节点恢复率对病毒传播的影响

固定节点恢复率和移动节点恢复率对病毒传播的影响分别如图9和图10所示。显然，固定节点恢复率越大，病毒传播就越慢，感染高峰值就越小。从图9可以看出，固定节点恢复率的变化既会影响感染的固定节点数，又会影响感染的移动节点数，固定节点恢复率的影响是通过改变下载病毒文件概率实现的。从图10可以看出，移动节点感染数量会随着移动节点恢复率的变化作相反的变化，但固定节点感染数量则不随移动节点恢复率的变化发生变化，原因还是移动节点恢复率变化不会影响病毒文件的下载概率。

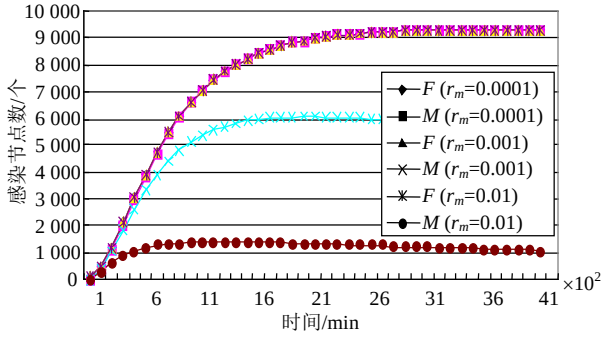

图10 移动节点恢复率对病毒传播的影响

病毒文件打开时所生成的病毒文件拷贝数对病毒传播的影响如图11所示。病毒生成的病毒文件拷贝越多，病毒传播就越快。为了加快病毒传播，病毒编写者可能会增加生成的病毒文件拷贝数量。但随着病毒文件拷贝数量的增加，病毒被用户发现的可能性变大，病毒文件被用户删除的可能性也随之增加。

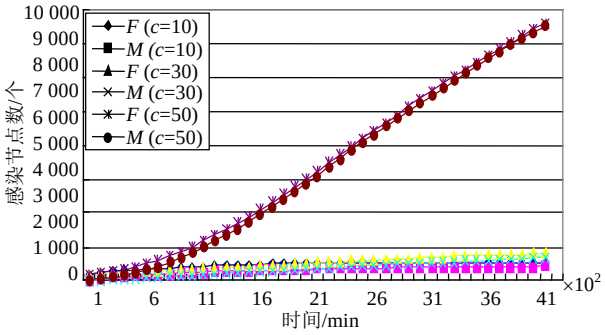

图11 病毒文件打开时所生成病毒文件个数对病毒传播的影响

初始固定感染节点数对病毒传播的影响如图12所示。从图12容易看出，初始感染节点数的变化几乎不会影响病毒(无论是在固定节点中还是在移动节点中)的传播速度和传播峰值。

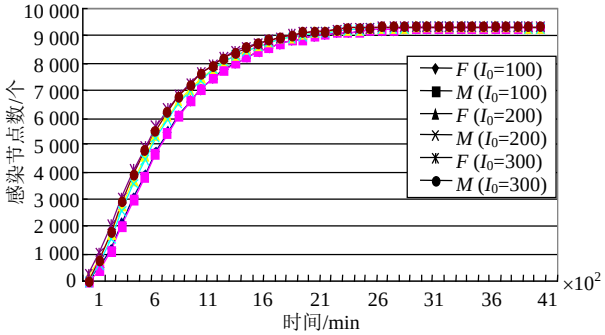

图12 初始固定感染节点数对病毒传播的影响

3.3 病毒传播预测与控制

本文以已有的P2P病毒Darby为例，考察病毒在移动P2P网络中的传播趋势。由于用户每分钟下载的文件数约为0.003[2]，而包含Darby的文件被打开后会

在共享文件夹中生成100多个该病毒文件的拷贝。图12所示就是下载率为0.003和c为100时的情况，大致反映出了该病毒的传播情况。从图中还可看出，无论初始感染节点数(>0)为多大，不到两天时间，该病毒就会使整个网络中9成以上的节点受感染。但是，在实际的移动P2P网络中，随着网络中病毒的增加，病毒被用户发现和清除的概率也会增加(即恢复率会增加)，所以病毒传播的速度要慢些。

实验分析表明，对病毒传播最敏感的参数是固定节点下载率、固定节点恢复率和生成病毒文件数，而前两个参数是用户可以控制的，因此，在P2P网络中出现病毒时，可以通过升级反病毒软件降低下载率，提高恢复率，有效控制和遏制P2P病毒的传播。

4 总 结

移动P2P网络由于兼具P2P网络的优点和移动网络优点，因而应用前景十分广阔。然而，与固定P2P网络一样，移动P2P网络也面临P2P病毒的威胁。病毒利用在共享文件夹中生成病毒拷贝和通过用户的正常下载行为进行传播。对移动P2P网络的体系结构和下载协议进行深入研究分析，根据病毒传播特点并利用流行病学理论，本文提出了移动P2P网络中病毒传播的数学模型，并基于该模型使用专门的数值分析软件进行了大量仿真实验，主要考查了各P2P参数对病毒传播的影响。实验表明，固定节点的下载率、恢复率和生成病毒文件拷贝数是影响病毒传播的关键参数。基于传播模型对已有的P2P病毒Darby在移动P2P网络中的传播进行预测，表明该病毒可在不到两天的时间内感染网络中九成以上的节点。因此，在发现病毒时，要通过尽快降低下载率，提高恢复率遏制病毒传播。

参 考 文 献

[1] HYPPONEN M. Malware goes mobile[J]. Scientific America, 2006, 11: 70-77.

[2] THOMMES R, COATES M. Epidemiological modeling of Peer-to-Peer viruses and pollution[C]//Proceedings of the 25th Annual IEEE Conference on Computer Communications. Barcelona, Spain: IEEE Press, 2006: 15-26.

[3] CHEN G L, GRAY R S. Simulating non-scanning worms on Peer-to-Peer networks[C]//Proceedings of the 1st International Conference on Scalable Information Systems. Hong Kong, China: ACM, 2006.

[4] MA J, CHEN X M, XIANG G L. Modeling passive worm propagation in Peer-to-Peer system[C]//Proceedings of the 2006 International Conference on Computational Intelligence and Security. Guangzhou, China: IEEE, 2006: 1129-1132.

[5] SU J, CHAN K K W, MIKLAS A G, et al. A preliminary investigation of worm infections in a bluetooth environment[C]//Proceedings of the 4th ACM Workshop on Recurring Malcode (WORM'06) Co-located with the 13th ACM Conference on Computer and Communications Security (CCS'06). Alexandria,VA, USA: ACM, 2006: 9-16.

[6] ZHENG HUI, LI DONG, GAO ZHUO. An epidemic model of mobile phone virus[C]//Proceedings of the 1st International Symposium on Pervasive Computing and Applications Proceedings (SPCA'06). Urumchi, China: IEEE Computer Society, 2006: 534-538.

[7] YAN GUAN-HUA, FLORES H D, CUELLAR L, et al. Bluetooth worm propagation: mobility pattern matters[C]// Proceedings of the 2nd ACM Symposium on Information, Computer and Communications Security. Singapore: ACM, 2007: 245-256.

[8] LIU Shu-ping, JIANG Wei-rong, LI Jin-pei. Architecture and performance evaluation for P2P application in 3G mobile cellular systems[C]//Proceedings of International Conference on Wireless Communications, Networking and Mobile Computing. Shanghai, China: IEEE, 2007: 914-917.

编 辑 蒋 晓