

面向 P2P 特定信息的传播动力学模型研究

丁军平 蔡皖东

(西北工业大学计算机学院 西安 710072)

摘要 鉴于 P2P 特定信息传播与传染病传播的相似性,传播动力学是 P2P 特定信息传播的最新研究方向。针对现有传播动力学模型都不能准确模拟 P2P 特定信息传播过程的问题,对现有的 SEIR 传播动力学模型进行改进,建立了 SEInR 模型。该模型的主要特点包括:将传统的感染者(I)分为 n 个子类,每个子类赋予不同的模型参数;建立潜伏者(E)与移除者(R)之间的转换关系等。应用现代数学中的矩阵理论,得到 SEInR 模型的基本再生数计算公式,并对其进行分析。仿真结果表明,所提出的 SEInR 模型比传统 SEIR 模型能够更准确地模拟 P2P 特定信息传播过程,得到的基本再生数计算公式能够准确反映 P2P 特定信息的传播阈值。

关键词 P2P 特定信息,传播动力学,SEIR 模型,基本再生数,复杂网络

中图分类号 TP393.0 **文献标识码** A

Spreading Dynamics Model Research on P2P Specific Information

DING Jun-ping CAI Wan-dong

(College of Computer Science, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract Spreading of P2P specific information and epidemic are similar to each other, thus spreading dynamics becomes new research direction of P2P specific information spreading. Since the existing spreading dynamics model cannot accurately simulate P2P specific information spreading process, we improved the SEIR spreading dynamics model, and established the new SEInR spreading model. Main features of this model are as follows: dividing the traditional infected group(I) into n sub-groups, each of which has different parameter; and establishing the transformation relationship between Exposed (E) and Removed (R). Matrix theory was applied to get formula of basic reproductive number for SEInR model, and the basic reproductive number was analyzed. The simulation result indicates that the improved SEInR model can much more accurately simulate P2P specific information spreading process than traditional SEIR model; and the basic reproductive number can accurately reflect P2P specific information threshold.

Keywords P2P specific information, Spreading dynamics, SEIR model, Basic reproductive number, Complex network

1 引言

P2P 特定信息传播规律的研究是监控 P2P 网络不良信息的基础,现有的监控技术主要通过被动监测技术对 P2P 网络流量进行整体识别和整体封堵,没有对特定信息进行区分,也没有进行特定信息的传播规律研究。在 P2P 文件共享系统中,用户节点既可作为文件需求者来下载文件,又可作为文件供应者来上传文件,文件的共享与传播具有很好的效率。特定信息的大范围传播过程与传染病传播过程具有极大的相似性,传播动力学是 P2P 特定信息传播的最新研究方向。但是,通过对 P2P 文件共享系统和传播动力学理论深入分析后,发现已有的传播动力学模型都不能很好地完全模拟出 P2P 系统中节点在特定信息传播过程中的状态转换。针对 P2P 特定信息的传播特点,本文在研究现有 SEIR 模型的基础上,对该模型进行改进,提出了 SEInR 模型,对模型的基本再生数进行了推导,并使用数据仿真的方法对模型进行验证,仿真结果表明:SEInR 模型能够准确描述 P2P 特定信息的传

播过程。

2 研究现状

随着复杂网络结构研究的迅猛发展,人们逐步认识到不同事物在真实系统中的传播现象,例如,传染病的流行、计算机病毒在网络上的蔓延以及谣言在社会网中的扩散等都可以看作是服从某种规律的网络传播行为。如何去描述这些事物的传播过程,揭示它们的传播特性,是传播动力学理论的研究内容,也是复杂网络结构研究的最终目的之一。

流行病传播的数学模型能够很好地描述复杂网络的传播特性,是复杂网络传播动力学研究的基础。早在 1760 年,数学家 D. Bernoulli 就用数学模型刻画传染病的传播,并定量地回答了传染病的传播速度、传播规模等问题。Kermack 和 McKendrick 为了研究 1665—1666 年黑死病在伦敦的流行规律以及 1906 年瘟疫在孟买的流行规律,于 1927 年建立了著名的 SIR 仓室模型^[1],于 1932 年提出了 SIS 仓室模型^[2],在分析所建立模型的基础上,提出了区分流行病的“阈值理论”,

为传染病动力学的研究奠定了基础。从那时起,科学家们便致力于建立传染病数学模型去研究各种流行疾病的传播行为。目前研究最为广泛的经典传播模型是 SI 模型、SIS 模型、SIR 模型以及 SEIR 模型^[3-7]。SEIR 模型是由 Aron 等人^[8]最早提出的,Earn 等人^[9]利用 SEIR 模型并引入时间序列在理论上模拟麻疹的传播,Michael 等人^[10]着重从数学上分析了 SEIR 模型变化情况的局部平衡和全局稳定性。

在 SEIR 模型中,所有的个体被抽象为几类,每一类对应于一个典型的状态。其基本状态包括:易染状态(Susceptible, S)、潜伏状态(Exposed, E)、感染状态(Infective, I)和被移除状态(Removed, R)。则 SEIR 模型的转换过程如图 1 所示。

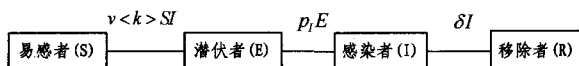


图 1 SEIR 模型传播过程

在该模型中,感染者(I)对易感者(S)进行感染,使其转换为潜伏者(E),潜伏者(E)发展一段时间后转换为感染者(I),感染者(I)被治愈后转换为移除者(R),移除者(R)不对其它状态的个体产生影响。其中: v 表示感染者(I)对易感者(S)的感染概率, $\langle k \rangle$ 表示所有个体的平均度值, p_I 表示潜伏者(E)根据潜伏时间转换为感染者(I)的概率, δ 表示感染者(I)被治愈后转换为移除者(R)的概率。 $S(t)$ 、 $E(t)$ 、 $I(t)$ 和 $R(t)$ 分别表示在 t 时刻系统中易染者、潜伏者、感染者和移除者的数量, $N_B(t)$ 表示所有个体的数量总和,即 $N_B(t) = S(t) + E(t) + I(t) + R(t)$ 。相应地,令 $S_R(t)$ 、 $E_R(t)$ 、 $I_R(t)$ 和 $R_R(t)$ 表示 $S(t)$ 、 $E(t)$ 、 $I(t)$ 和 $R(t)$ 占 $N_B(t)$ 的比率,即 $S_R(t) + E_R(t) + I_R(t) + R_R(t) = 1$ 。利用平均场理论可得该模型的动力学方程组为:

$$\begin{cases} \frac{dS_R}{dt} = -v\langle k \rangle S_R I_R \\ \frac{dE_R}{dt} = v\langle k \rangle S_R I_R - p_I E_R \\ \frac{dI_R}{dt} = p_I E_R - \delta I_R \\ \frac{dR_R}{dt} = \delta I_R \end{cases} \quad (1)$$

3 模型改进描述

3.1 现有 SEIR 模型的不足

将 SEIR 模型应用于 P2P 特定信息传播时,根据 P2P 文件传输的特点,对模型中的参数做了如下规定和定义:

$S_R(t)$ 表示 P2P 网络中还没有开始特定信息下载的点,但有可能对特定信息发生兴趣,并进行下载;

$E_R(t)$ 表示 P2P 网络中正在进行特定信息下载的点,这些点只有特定信息的部分内容,既下载文件内容也上传已有的文件内容;

$I_R(t)$ 表示 P2P 网络中拥有完整的特定信息,并在线共享,让其它节点进行特定信息下载的点;

$R_R(t)$ 表示 P2P 网络中不共享特定信息,也确认不会下载特定信息的点,这些点的来源主要有 2 种:已经完整下载特定信息,但对文件进行删除、更名、移动等操作,不提供特定信息下载的点;只下载了部分内容,由于其它原因(下载

速度过慢或者已下载的部分与预期不符等)不想继续下载的点;

$I_R(0)$ 表示 P2P 网络中初始拥有特定信息并提供下载的点;

在 P2P 文件共享系统中,当节点查询与特定信息相关的节点列表时,返回的节点列表既有拥有部分内容的节点也有拥有全部内容的节点,并且它们返回的概率是相同的,也就是说,潜伏者与感染者对易感者的感染概率是相同的;

由于每个用户重复下载同一个文件的概率极低,本模型不考虑节点对特定信息的重复下载,也就是说,不考虑节点状态由移除状态(R)转换为易感状态(S)。

由于 P2P 特定信息传播的自身特点,如果直接使用现有 SEIR 模型对 P2P 特定信息传播过程进行分析,会存在以下问题:

SEIR 模型中的潜伏者是不能感染易感者的,但是在 P2P 特定信息传播过程中,正在下载文件的潜伏节点也可以进行文件上传,对易感节点进行感染;

SEIR 模型中的潜伏节点是不能转换为移除节点的,但是在 P2P 特定信息传播过程中,部分潜伏节点,由于其它原因(下载速度过慢或者已下载部分与预期不符等)不想继续下载,该节点将转换为移除者(R);

SEIR 模型没有考虑 P2P 网络中新节点的加入和旧节点的永久退出;

SEIR 模型中所有感染者被治愈的概率是相同的,即对感染者不进行区分。但是由于 P2P 节点对特定信息共享意愿以及节点类型的不同,特定信息被共享的时间是不同的,也就是说不同的感染者被治愈的概率是不同的;

现有 SEIR 模型只考虑均匀网络上的特定信息传播,而 P2P 网络是复杂网络,而且 DHT 网络已经被验证属于无标度网络,因此需要考虑不同的度值对传播模型的影响。Pastor-Satorras 等人^[11]在 SIS 模型中将节点按照其度值进行分类,研究了 SIS 模型在无标度网络上的传播特征。但是在 P2P 特定信息的传播过程中,由于节点状态转换时,节点度值也会发生相应的变化,与 Pastor-Satorras 等人所研究的具有固定度值的无标度网络模型存在较大的区别。

3.2 SEInR 模型描述

根据以上对 P2P 特定信息传播特点的分析,需要对 SEIR 模型进行改进,使之能够更精确地反映 P2P 特定信息传播过程中的状态转换。改进后的模型为 SEInR 模型,它的主要改进内容包括:

(1)建立潜伏者对易感者的感染机制,使得易感者被感染时不但要考虑感染者的因素,还需要考虑潜伏者的因素;

(2)单位时间内,潜伏者以概率 p_{ER} 直接转换为移除者,以概率 p_{EI} 转换为感染者;

(3)考虑 P2P 网络中新节点的加入和旧节点的退出,新加入的节点都为易感节点(S),并且所有状态节点的永久退出概率 p_F 相同。为了保持节点总数的一致性,新节点加入的概率为 p_F ;

(4)根据 P2P 节点对特定信息共享时间的不同,将感染者(I)节点划分为 n_I 个子类: I_1, I_2, \dots, I_{n_I} ,并给每个子类设置不同的被治愈概率: $\delta_1, \delta_2, \dots, \delta_{n_I}$,并且在不同的感染者子类中使用不同的度分布函数 $H_{k_i}, 1 \leq i \leq n_I$;

(5) 不仅考虑特定信息在均匀网络上的传播,还需要考虑特定信息在无标度网络上的传播。根据 P2P 特定信息的传播特点:不能将节点按照度值进行分类研究,本模型提出在易感者被感染时,使用度分布函数 H_k 代替平均度值 $\langle k \rangle$ 进行分析。

改进后的 SEInR 模型的转换过程如图 2 所示。

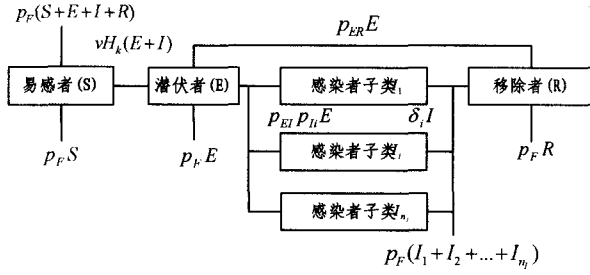


图 2 SEInR 模型传播过程

使用 SEInR 模型描述 P2P 特定信息传播过程的内容为:易感节点(S)发送文件请求,获取节点列表信息,返回的节点列表既包含潜伏节点也包含感染节点;易感节点与得到的节点列表进行连接并获取文件信息,易感节点变为潜伏节点(E),潜伏节点既下载文件信息也上传文件信息;潜伏节点在下载的过程中以概率 p_{ER} 对特定信息失去下载需求,变为移除节点(R);潜伏节点以概率 p_{EI} 下载完成后变为感染节点(I);感染节点(I)对特定信息进行共享,共享完成后,感染节点变为移除节点(R)。每个感染者子类根据该类的平均共享时间计算该类的被治愈概率 δ_i , ($1 \leq i \leq n_i$), n_i 表示感染者子类数量。单位时间内新加入的节点率为常数 p_F , 数量为 $p_F N_B(t)$ 。每个状态中的节点都会以概率 p_F 永久退出 P2P 网络。根据以上对 P2P 特定信息传播过程的描述,SEInR 模型的动力学方程组为:

$$\begin{cases} \frac{dS_R}{dt} = p_F - vS_R(H_k E_R + \sum_{i=1}^{n_i} H_{k_i} I_{R_i}) - p_F S_R \\ \frac{dE_R}{dt} = vS_R(H_k E_R + \sum_{i=1}^{n_i} H_{k_i} I_{R_i}) - (p_{EI} + p_{ER} + p_F) E_R \\ \frac{dI_{R_i}}{dt} = p_{EI} p_{E_i} E_R - (\delta_i + p_F) I_{R_i} \\ \frac{dR_R}{dt} = p_{ER} E_R + \sum_{i=1}^{n_i} \delta_i I_{R_i} - p_F R_R \end{cases} \quad (2)$$

式中, H_k 表示潜伏节点的度分布函数, H_{k_i} 表示感染者子类 I_i 中的节点度分布函数,不同的度分布函数会对传播动力学模型产生不同的性质。 I_{R_i} 表示感染者子类 I_i 中的节点数量占所有节点的比例, p_{E_i} 表示一个新感染节点属于感染者子类 I_i 的概率, $1 \leq i \leq n_i$, 并且 $I_R = \sum_{i=1}^{n_i} I_{R_i}$, $\sum_{i=1}^{n_i} p_{E_i} = 1$, δ_i 表示感染者子类 I_i 中的节点被治愈的概率。

4 模型基本再生数分析

通过对 SEInR 模型的动力学方程组的分析,不难看出,总节点 $N_B(t)$ 满足下面的微分方程:

$$\frac{dN_B(t)}{dt} = p_F N_B(t) - p_F (S(t) + E(t) + I(t) + R(t)) = 0 \quad (3)$$

由于在易感节点(S)、潜伏节点(E)、感染节点(I)的变化过程中没有移除节点(R)的变化因素,略去 SEInR 模型的动力学方程组中有关移除节点(R)的变化方程,并不影响对整

个系统的研究。简化后的动力学方程组为:

$$\begin{cases} \frac{dS_R}{dt} = p_F - vS_R(H_k E_R + \sum_{i=1}^{n_i} H_{k_i} I_{R_i}) - p_F S_R \\ \frac{dE_R}{dt} = vS_R(H_k E_R + \sum_{i=1}^{n_i} H_{k_i} I_{R_i}) - (p_{EI} + p_{ER} + p_F) E_R \\ \frac{dI_{R_i}}{dt} = p_{EI} p_{E_i} E_R - (\delta_i + p_F) I_{R_i} \end{cases} \quad (4)$$

记集合 $D_C = \{(S_R, E_R, I_R) | S_R \in [0, 1], E_R \in [0, 1], I_R \in [0, 1], 1 \leq i \leq n_i\}$, 显然 D_C 是简化动力学方程组的一个正向不变集,以下仅在集合 D_C 内对方程组进行分析。

模型中的基本再生数 R_0 是分析传播动力学模型性能的重要参数。文献[12]介绍了有关基本再生数计算的方法:考虑一个传染病的仓室模型,在该模型中,所有的个体被分类到 N_r 个仓室。令 $z = (z_1, z_2, \dots, z_i, \dots, z_{N_r})$, z_i 表示第 i 个仓室的个体数量并且 $z_i \geq 0$, 为了分析方便,有感染者的仓室排在前面,没有感染者的仓室排在后面。令 $\tilde{F}_i^+(z)$ 表示仓室 i 中出现新的感染者的速率(只包括新增加的感染个体,不包括仓室之间个体的移动), $\tilde{V}_i^+(z)$ 表示其它方式进入仓室 i 的速率, $\tilde{V}_i^-(z)$ 表示移出仓室 i 的速率,且 $\tilde{V}_i(z) = \tilde{V}_i^-(z) - \tilde{V}_i^+(z)$ 。假设每个变量都是连续可微的,那么疾病的传播模型可以由下面方程系统所组成:

$$\frac{dz_i}{dt} = f_n(z) = \tilde{V}_i^-(z) - \tilde{V}_i^+(z) \quad (5)$$

式中, $f_n(z)$ 表示第 i 个仓室中的个体数量变化方程。在传播模型中,定义的函数满足以下定理。

定理 1 如果 $z \geq 0$, 那么,对于 $1 \leq i \leq N_r$, 有: $\tilde{F}_i^+(z)$, $\tilde{V}_i^+(z)$, $\tilde{V}_i^-(z) \geq 0$ 。

$\tilde{F}_i^+(z)$, $\tilde{V}_i^+(z)$, $\tilde{V}_i^-(z)$ 表示仓室的个体移动函数,该定理表示函数值是非负的。

定理 2 如果 $z_i = 0$, 那么 $\tilde{V}_i^-(z) = 0$ 。

该定理表示如果一个仓室已经没有个体存在,那么该仓室就没有个体以任何的形式移出。

定理 3 如果可能包含感染者的仓室数量为 N_m , 当 $i > N_m$ 时, $\tilde{F}_i^+(z) = 0$ 。

该定理表示在没有被疾病感染的仓室,不会有感染的个体。

定理 4 当 $j \leq N_m$, 且 $z_j = 0$ 时,对于所有的仓室,其 $\tilde{F}_i^+(z) = 0, 1 \leq i \leq N_r$ 。

该定理表示当个体所在的仓室都属于无病状态的仓室时,不会有新的感染者进入。

定理 5 令 $Df_r(z_0)$ 表示 $f_n(z)$ 在点 z_0 处的 Jacobian 矩阵。如果 $\tilde{F}_i^+(z) = 0, 1 \leq i \leq N_r$ 时, $Df_r(z_0)$ 有负的实部。

该定理表示当模型中没有新的感染个体出现时,那么人群会保持稳定状态。若人群在无病平衡点附近保持稳定,根据微分方程的性质和定理可知 $Df_n(z_0)$ 有负的实部。

定理 6 若 E_0 是模型的无病平衡点,且 $f_n(z)$ 满足定理 1 到定理 5, 那么 $D\tilde{F}(E_0)$ 和 $D\tilde{V}(E_0)$ 可被分别划分为:

$$D\tilde{F}(E_0) = \begin{pmatrix} F & 0 \\ 0 & 0 \end{pmatrix} \text{ 和 } D\tilde{V}(E_0) = \begin{pmatrix} V & 0 \\ J_3 & J_4 \end{pmatrix}$$

FV^{-1} 为再生矩阵,基本再生数 R_0 是 FV^{-1} 的谱半径长度,即 $R_0 = \rho(FV^{-1})$ 。

其中, $\rho(FV^{-1})$ 为求 FV^{-1} 的谱半径长度。

按照上述计算基本再生数 R_0 的方法,首先将仓室按照

节点的感染顺序进行排列,排列后的仓室顺序为:

$$E, I_i, S$$

按照动力学方程组求出 \tilde{F} 和 \tilde{V} :

$$\begin{aligned} \tilde{F} &= [\tilde{F}_1(z), \tilde{F}_2(z), \dots, \tilde{F}_{n_i+2}(z)]^T = (vS_R(H_k E_R + \sum_{i=1}^{n_i} H_{ki} I_{Ri}), 0, \dots, 0, 0)^T \\ \tilde{V} &= [\tilde{V}_1(z), \tilde{V}_2(z), \dots, \tilde{V}_{n_i+2}(z)]^T \\ &= ((p_{EI} + p_{ER} + p_F) E_R, (\delta_i + p_F) I_{Ri} - p_{EI} p_{Ii} E_R, \dots, \\ &\quad (\delta_{n_i} + p_F) I_{Ri} - p_{EI} p_{Ii} E_R, vS_R(H_k E_R + \sum_{i=1}^{n_i} H_{ki} I_{Ri}) + \\ &\quad p_F S_R - p_F)^T \end{aligned}$$

由于系统处于平衡点时,所有仓室中的节点数量变化率为 0,也就是说:

$$\begin{cases} p_F - vS_R(H_k E_R + \sum_{i=1}^{n_i} H_{ki} I_{Ri}) - p_F S_R = 0 \\ vS_R(H_k E_R + \sum_{i=1}^{n_i} H_{ki} I_{Ri}) - (p_{EI} + p_{ER} + p_F) E_R = 0 \\ p_{EI} p_{Ii} E_R - (\delta_i + p_F) I_{Ri} = 0 \end{cases} \quad (6)$$

容易发现上式存在无病平衡点 $E_0 = (S_R=1, E_R=0, I_R=0)$, 根据 Jacobian 矩阵的计算公式, 求出 $D\tilde{F}(E_0)$ 和 $D\tilde{V}(E_0)$ 的值。

$$D\tilde{F}(E_0) = \begin{pmatrix} vH_k & vH_{k1} & \dots & vH_{kn_i} & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

$$D\tilde{V}(E_0) = \begin{pmatrix} p_{EI} + p_{ER} + p_F & 0 & \dots & 0 & 0 \\ -p_{EI} p_{I1} & \delta_1 + p_F & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ -p_{EI} p_{In_i} & 0 & \dots & \delta_{n_i} + p_F & 0 \\ vH_k & vH_{k1} & \dots & vH_{kn_i} & p_F \end{pmatrix}$$

令 $\omega_E = p_{EI} + p_{ER} + p_F$ 表示潜伏节点的减少概率, $p_{EIi} = p_{EI} p_{Ii}$ 表示潜伏节点进入感染者子类 I_i 的概率, $\omega_{Ii} = \delta_i + p_F$ 表示感染者子类 I_i 中的节点减少概率。从 $D\tilde{F}(E_0)$ 和 $D\tilde{V}(E_0)$ 中提取出 F 和 V 分别为:

$$F = \begin{pmatrix} vH_k & vH_{k1} & \dots & vH_{kn_i} \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

$$V = \begin{pmatrix} \omega_E & 0 & \dots & 0 \\ -p_{EI1} & \omega_{I1} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ -p_{EIIn_i} & 0 & \dots & \omega_{In_i} \end{pmatrix}$$

计算 FV^{-1} 的谱半径长度即可求得基本再生数 R_0 为:

$$\begin{aligned} R_0 &= \frac{v}{\omega_E} (H_k + \sum_{i=1}^{n_i} \frac{p_{EIi} H_{ki}}{\omega_{Ii}}) \\ &= \frac{v}{(p_{EI} + p_{ER} + p_F)} (H_k + \sum_{i=1}^{n_i} \frac{p_{EI} p_{Ii} H_{ki}}{(\delta_i + p_F)}) \end{aligned} \quad (7)$$

5 仿真分析

P2P 文件共享网络通常由上百万个节点构成, 本文通过数据仿真的方式对提出的 SEInR 模型进行分析。将感染节

点(D)划分为 3 个子类: 长时间在线共享子类(I_1)、临时共享子类(I_2)和搭便车子类(I_3)。设定特定信息的文件大小为 600M, 节点的平均下载速度为 100kbyte/s, 则完成特定信息的下载需要 100min, 单位监测时间设定为 10min, 则潜伏节点的平均寿命为 10 个单位时间, 相应地, 单位时间内从潜伏节点转换为下载节点的概率 p_{EI} 为 0.1。其它参数的缺省值设置如下: $v=0.015$, $H_k=10$, $p_{ER}=0.001$, $p_F=0.001$, $S(0)=99999$, $I_1(0)=1$, $I_2(0)=0$, $I_3(0)=0$, $H_{k1}=20$, $H_{k2}=10$, $H_{k3}=2$, $p_{I1}=0.01$, $p_{I2}=0.39$, $p_{I3}=0.6$, $\delta_1=0.0001$, $\delta_2=0.3$, $\delta_3=0.95$ 。各个状态的节点比率变化如图 3 所示, 感染者子类中的节点比率变化如图 4 所示。从图 3 可以看出: 易感节点(S)的比率开始下降较慢, 随后有一个快速下降的过程, 最后的下降过程逐渐平缓, 并最终趋于平衡点中的 S_R^* ; 移除节点(R)的比率开始上升较慢, 随后快速上升, 最后的上升过程逐渐平缓, 并最终趋于平衡点中的 R_R^* ; 潜伏节点(E)与感染节点(I)的比率变化过程都是先上升后下降的过程, 并最终趋于平衡点中的 E_R^* 和 I_R^* 。从图 4 可以看出: 由于 I_1 中的节点长时间在线, 被治愈概率和进入概率都很低, 因此, 节点比率上升后下降缓慢, 基本维持在高位水平 I_{R1}^* ; I_2 中的节点被治愈概率和进入概率都居中, 相应的节点比率快速上升和下降, 最终稳定在平衡点中的 I_{R2}^* ; 虽然 I_3 中的节点进入概率很高, 但是被治愈概率也很高, 属于下载完成后不进行共享的“搭便车”节点, 因此, I_3 中的节点比率高峰值较 I_2 中的要小, 并且最终在平衡点中的值 I_{R3}^* 要小于 I_{R2}^* 。

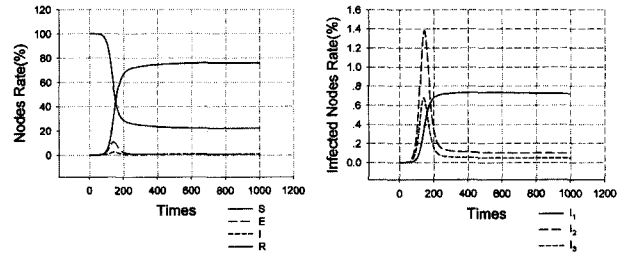


图 3 各个状态节点比率变化图 图 4 感染者子类节点比率变化图

$R_0=1$ 是判断特定信息传播规律的重要参数, 图 5 展示了 R_0 不同取值时, S_R 与 I_R 的变化情况。从图中可以看出: 当 $R_0 \leq 1$ 时, $I_R \rightarrow 0$, $S_R \rightarrow 1$, 也就是趋向于模型的无病平衡点, 并且是全局稳定的; 当 $R_0 > 1$ 时, I_R 趋向于大于 0 的数值, 并且是全局稳定的; 当 R_0 越大时, I_R 越大, 传播速度越快。

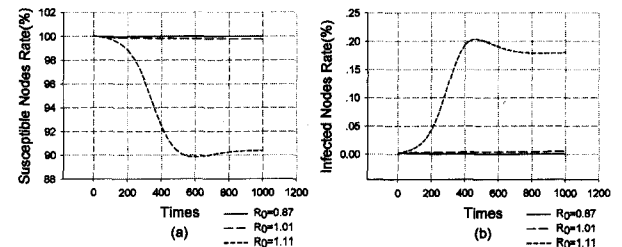


图 5 R_0 对模型的影响

为了验证 SEInR 传播模型是否能够准确描述 P2P 特定信息的传播过程, 本文使用 NS2 仿真平台对 P2P 特定信息的传播过程进行模拟, 并与传统 SEIR 模型和 SEInR 模型的传播过程进行比较, 结果如图 6 所示。从图 6 中可以看出: 传统

(下转第 113 页)

当且仅当对于任意 $w' \in W$, 有
 $w' \models \mathbf{B}(b) \wedge \mathbf{P}(d) \wedge \mathbf{r}(b, d) \wedge \mathbf{C}(c)$

因此, 对于任意 $w \in W$,

$M, v, w \models \neg i(c, b) [a] i(c, b)$

当且仅当 $M, v, w \models \neg i(c, b)$ 蕴含对于任意的 $w', (w, w') \in R_a, M, v, w' \models i(c, b)$ 。

结束语 本文给出了一个描述 Web 服务的一阶动态逻辑。它与传统的一阶动态逻辑不同, 后者只是用来形式化程序的动态行为。 PDL_{ws} 的逻辑语言、语法和语义也不同于传统的一阶动态逻辑, 因为 PDL_{ws} 中的动作包括程序和其他的非程序类型的服务, 例如购书服务; 并且, PDL_{ws} 包括服务常量和变量, 服务变量是用 WSMO 中的本体来描述的。

下一步的工作包含对 PDL_{ws} 的逻辑分析: 它的公理表达式、完备性和可靠性; 描述 Web 服务的动态描述逻辑, 通过逻辑地分析这些形式化, 设计一个描述 Web 服务的、通用的一阶动态逻辑, 并且使其与现有 Web 服务的形式化兼容。在描述 Web 服务的动态描述逻辑中, 服务变量可以用服务概念表示, 这个与对象概念相同。

参考文献

[1] de Bruijn J, Lausen H, Polleres A, et al. The Web Service Modeling Language (WSML): an overview[C]//Sure Y, Domingue J,

eds. ESWC 2006, LNCS 4011. 2006; 590-604

[2] Eiter T, Lukasiewicz T, Schindlauer R, et al. Combining answer set programming with description logics for the semantic Web [C]//Proc. of the International Conference of Knowledge Representation and Reasoning (KR04). 2004

[3] Harel D, Kozen D, Tiuryn J. Dynamic Logic [M]. The MIT Press, 2000

[4] Horrocks I, Patel-Schneider P F, Boley H, et al. SWRL: a Semantic Web Rule Language Combining OWL and RuleML [EB/OL]. W3C Member Submission, <http://www.w3.org/Submission/2004/SUBM-SWRL-20040521/>, 21 May 2004

[5] Klusch M, Fries B, Sycara K. OWLS-MX: a hybrid semantic Web service matchmaker for OWL-S services, Web Semantics: Science, Services and Agents on the WorldWideWeb 7[S]. 2009; 121-133

[6] Martin D, Burstein M, Hobbs J, et al. OWL-S: Semantic markup for Web services [EB/OL]. Member Submission 22, W3C, <http://www.w3.org/Submission/2004/SUBM-OWL-S-20041122/>, November 2004

[7] OWL-S. Semantic markup for Web services [EB/OL]. W3C member submission, <http://www.w3.org/Submission/2004/SUBM-OWL-S-20041122/>, 2004-11-22

[8] Roman D, Keller U, Lausen H, et al. Web service modeling ontology[J]. Applied Ontology, 2005, 1(1): 77-106

(上接第 99 页)

SEIR 模型模拟的传播过程与实际情况差距较大, 而 SEInR 传播模型能够较为准确地模拟 P2P 特定信息的传播过程, 能够应用于 P2P 特定信息的传播规律分析。

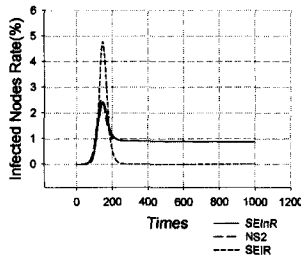


图 6 SEInR 模型效果验证

结束语 P2P 特定信息的大范围传播过程与传染病传播过程具有极大的相似性, 传播动力学是 P2P 特定信息传播的最新研究方向。本文针对 P2P 特定信息的传播特点, 对现有 SEIR 模型进行改进, 提出了 SEInR 模型。该模型的改进内容包括: 将传统的感染者 (I) 分为 n 个子类, 每个子类赋予不同的模型参数; 建立潜伏者 (E) 与移除者 (R) 之间的转换关系; 考虑复杂网络的不同度值对传播模型的影响等。应用现代数学中的矩阵理论, 对模型性能的关键参数——基本再生数 R_0 进行了推导。最后, 使用数据仿真的方法对模型的传播过程进行了分析, 仿真结果表明: 所提出的 SEInR 模型比传统 SEIR 模型能更准确地模拟 P2P 特定信息传播过程, 得到的基本再生数计算公式能准确地反映 P2P 特定信息的传播阈值。

参考文献

[1] Kermack W, McKendrick A. Contributions to the mathematical theory of epidemic[J]. Proc. Roy. Soc., 1927, A115: 700-721

[2] Kermack W, McKendrick A. Contributions to the mathematical theory of epidemic[J]. Proc. Roy. Soc., 1932, A138: 55-83

[3] 马知恩, 周义仓, 王稳地, 等. 传染病动力学的数学建模与研究 [M]. 北京: 科学出版社, 2004

[4] 张达敏, 蔡绍洪, 周海平, 等. 复杂网络上的一种综合型 SIRS 模型[J]. 兰州大学学报, 2009, 45(3): 122-126

[5] 刘浩. P2P 网络的若干关键问题研究[D]. 广州: 华南理工大学, 2010

[6] 谭永基, 蔡志杰, 俞文. 数学模型 [M]. 上海: 复旦大学出版社, 2005

[7] 王方伟, 张运凯, 王长广, 等. Witty 蠕虫的传播模型与分析[J]. 华中科技大学学报, 2008, 36(2): 50-54

[8] Aron J, Schwartz I. Seasonality and period-doubling bifurcations in an epidemic model[J]. Theory Bio., 1984, 110: 665-679

[9] Earn D, Rohani P, Bolker B, et al. A simple model for complex dynamical transitions in epidemics[J]. Science, 2000, 287: 667-670

[10] Michael L, Wang L. Global stability in some SEIR epidemic models[J]. IMA Volumes in Mathematics and its Applications, 2001, 126: 295-311

[11] Pastor-Satorras R, Vespignani A. Epidemic Spreading in Scale-Free Networks[J]. Physical Review Letters, 2001, 86: 3200-3203

[12] 俞嘉地. BitTorrent 对等网文件共享系统关键技术研究[D]. 上海: 上海交通大学, 2007