

一类新型计算机病毒最优控制模型*

赵夫群[†]

(陕西咸阳师范学院教育科学学院, 咸阳 712000)

摘要 计算机病毒的存在,使得很多计算机无法正常运行,并造成巨大的损失.针对这种情况,提出一类新型计算机病毒的最优控制.控制首先对原有的计算机病毒模型进行改进,加入控制项,提出最优控制问题,并证明最优控制的存在性,然后利用庞德里亚金的极小值原理进行理论分析,最后进行数值模拟.数值模拟结果表明,在没有控制的条件下,原模型得出基本再生数大于1,说明存在地方病平衡点,并且最终会导致病毒爆发,而数值模拟揭示,运用有效的控制策略能够更好地抑制计算机病毒的传播.

关键词 计算机病毒模型, 极小值原理, 最优控制, 数值仿真

DOI: 10.6052/1672-6553-2015-036

引言

计算机病毒是指编制者在计算机程序中插入的破坏计算机功能或者破坏数据,影响计算机使用并且能够自我复制的一组计算机指令或者程序代码.计算机病毒和医学上的“病毒”不同,计算机病毒不是天然存在的,是某些人利用计算机软件 and 硬件所固有的脆弱性编制的一组指令集或程序代码.它能通过某种途径潜伏在计算机的存储介质或者程序里,当达到某种条件时即被激活,通过修改其他程序的方法将自己的精确拷贝或者可能演化的形式放入其他程序中.从而感染其他程序,对计算机资源进行破坏,所谓的病毒就是人为造成的,对其他用户的危害性很大.计算机病毒的破坏性及传染性对社会造成极大的伤害,仅2008年,计算机病毒在全球造成的经济损失就高达85亿美元.

计算机病毒的传播类似于生物学中病毒的传播方式,通过各种渠道从已被感染的计算机扩散到为被感染的计算机,导致硬盘数据被清空.计算机病毒具有自我复制并传播给其他计算机的功能,同时又有极大的破坏性,一旦计算机被感染上病毒,将可能面临着计算机瘫痪,更严重的甚至会造成社会巨大的经济损失,因此有效预防和控制计算机病毒的传播就显得尤为重要.生物病毒的最优控制

在理论研究方面已经取得很大的成就.比如,艾滋病病毒的最优控制,乙肝,丙肝病毒的最优控制^[1-6],由于计算机病毒和生物病毒很相似,很多学者将生物学病毒和杀毒软件结合起来研究计算机病毒模型^[7-14].虽然控制理论被广泛应用到流行病传播模型中,但是很少学者有研究将控制理论应用到计算机病毒模型中,因此有必要研究一下计算机病毒的最优控制模型.

本文在文献[14]的基础上进行了改进,将之前的常数参数变量变成时变的控制变量,并为新模型建立了最优控制系统.首先给出控制系统的目标泛函;然后用 Pontryagin 极小值原理^[15-17]推导出最优控制;最后数值仿真验证了所得结论.

1 计算机病毒模型

1.1 计算机病毒模型的提出

考虑计算机网络由若干个节点组成的,并且所有的未感染的计算机都可能受到恶意软件的攻击.当恶意软件攻击计算机网络时,一部分节点会被感染.这些感染的节点又会继续感染健康的节点,因此被感染的节点的数目会大量增加.杀毒软件的使用,能够有效的防治计算机病毒的传播,所以文献[14]提出了如下的计算机模型:

2014-07-29 收到第1稿,2015-05-19 收到修改稿.

* 咸阳师范学院专项科研基金项目(07XSYK224)

[†] 通讯作者 E-mail:13240766843@163.com

$$\begin{cases} \frac{dS}{dt} = A - \beta SM + \gamma IP - dS, \\ \frac{dI}{dt} = \beta SM - \gamma IP - dI - \alpha I, \\ \frac{dM}{dt} = \delta \mu I - \mu_0 MP, \\ \frac{dP}{dt} = \eta I - \eta_0 (P - P_0). \end{cases} \quad (1)$$

1.2 改进的计算机病毒模型

由于现实生活中计算机之间也可以进行病毒传播,所以本文在文献[14]的基础上增加一项 $\beta_1 SI$,得到的改进模型如下:

$$\begin{cases} \frac{dS}{dt} = A - \beta_1 SI - \beta_2 SM + \gamma IP - dS, \\ \frac{dI}{dt} = \beta_1 SI + \beta_2 SM - \gamma IP - dI - \alpha I, \\ \frac{dM}{dt} = \delta \mu I - \mu_0 MP, \\ \frac{dP}{dt} = \eta I - \eta_0 (P - P_0). \end{cases} \quad (2)$$

系统(1)由四部分组成: S 表示没有感染病毒但很容易感染病毒的节点; I 表示已经感染病毒的节点; M 表示网络中恶意对象; P 表示杀毒软件. 系统中参数含义: A 表示计算机的输入率, β_1 表示计算机之间病毒感染率, β_2 表示恶意对象入侵计算机的感染率, γ 表示杀毒率, d 表示断网率, α 表示大量病毒入侵计算机瘫痪率, δ 表示病毒复制能力, μ 表示病毒入侵率, μ_0 表示杀恶意对象的效率, η 表示杀毒软件的增长率, η_0 表示杀毒失败率, P_0 表示计算机现有的杀毒能力.

通过计算,系统(1)能够得到在没有控制的条件下系统(2)的基本再生数 $R_0 = \frac{A(\mu_0 P_0 \beta_1 + \mu \delta \beta_2)}{d \mu_0 P_0 (\alpha + d + \gamma P_0)}$, 当 $R_0 > 1$ 时,系统(1)有地方病平衡点,说明如果没有得到有效控制的话,最终计算机机会爆发地方病,因此,如何有效控制计算机病毒就显得尤为重要.

2 最优控制问题

系统(1)中 γ 和 μ_0 是个常数,本文对此再次改进,将 γ 和 μ_0 改成控制变量函数 $\theta u(t)$ 和 $u(t)$. 因为 γ 和 μ_0 是常数的话,会使得控制的成本大大增加,变成时变的控制变量,可以保证在控制计算机病毒减少到很低甚至到零的成本远远低于常数变量. 带有控制变量的控制模型如下:

$$\begin{cases} \frac{dS}{dt} = A - \beta_1 SI - \beta_2 SM + \theta u(t) IP - dS, \\ \frac{dI}{dt} = \beta_1 SI + \beta_2 SM - \theta u(t) IP - dI - \alpha I, \\ \frac{dM}{dt} = \delta \mu I - u(t) MP, \\ \frac{dP}{dt} = \eta I - \eta_0 (P - P_0). \end{cases} \quad (3)$$

其中 $0 \leq u(t) \leq 1$, θ 表示杀毒率调节系数, $u(t)$ 表示杀毒率,这里假设杀毒率和购买和更新的杀毒软件的数量成正比.

首先,为了建立最优控制系统,对于给定的整数 $T > 0$,给定下列集合作为控制集:

$$U = \{u(t) \in L^2(0, T) : 0 \leq u(t) \leq 1, 0 \leq t \leq T\} \quad (4)$$

其次,目标泛函定义如下:

$$J(u) = \min_{u(t)} \int_0^T (b_1 I^2 + b_2 M^2 + b_3 u^2) dt, \quad (5)$$

这样的目标函数是为了通过控制使得感染的计算机和恶意对象的数量达到最小,并且系统的消耗最少,从而达到最优控制的目的,这也是符合我们最优控制的初衷.

初始条件为:

$$\begin{aligned} S(0) &= S_0, I(0) = I_0, M(0) = M_0, \\ P(0) &= P_0. \end{aligned} \quad (6)$$

下面给出最优控制问题的 Lagrangian 函数,

$$L(I, u) = b_1 I^2 + b_2 M^2 + b_3 u^2, \quad (7)$$

其中 I, M 表示 t 时刻 I 和 M 的数目, u 表示为了减少 I 和 M 所付出的系统消耗. 例如我们用于购买和更新杀毒软件所需要的金钱.

为了寻找目标泛函式(5)的最小值,先定义 Hamiltonian 函数如下:

$$\begin{aligned} H &= b_1 I^2 + b_2 M^2 + b_3 u^2 + \lambda_1 (A - \beta_1 SI - \beta_2 SM + \\ &\quad \theta u(t) IP - dS) + \lambda_2 (\beta_1 SI + \beta_2 SM - \\ &\quad \theta u(t) IP - dI - \alpha I) + \lambda_3 (\delta \mu I - u(t) MP) + \\ &\quad \lambda_4 (\eta I - \eta_0 (P - P_0)). \end{aligned} \quad (8)$$

3 最优控制存在性

最优控制的存在性证明:

由文献[15],且方程(3),(4),(6)满足如下条件:

- 1) 控制集和相对应的状态变量集为非空集合.
- 2) 可测集 U 是闭的凸集.
- 3) 系统(3)方程的右边关于 u 的线性函数,并

且是连续有界的.

4) 目标泛函的被积函数 $L(I, u)$ 在 U 上式凹函数.

5) 存在一个常数 $\rho > 1$ 和两个正数 η_1, η_2 , 满足 $L(I, u) \geq \eta_1 + \eta_2 (|u|)^{\rho/2}$.

综述, 系统(2)存在最优控制.

利用 Pontryagin 的极小值原理^[16], 极小值原理的必要条件如下:

$$6) \begin{cases} \frac{dS}{dt} = A - \beta_1 SI - \beta_2 SM + \theta u(t) IP - dS, \\ \frac{dI}{dt} = \beta_1 SI + \beta_2 SM - \theta u(t) IP - dI - \alpha I, \\ \frac{dM}{dt} = \delta \mu I - u(t) MP, \\ \frac{dP}{dt} = \eta I - \eta_0 (P - P_0). \end{cases}$$

7) $S(0) = S_0, I(0) = I_0, M(0) = M_0, P(0) = P_0.$

$$8) [\lambda'_1 \quad \lambda'_2 \quad \lambda'_3 \quad \lambda'_4]' = - \left[\begin{array}{c} \frac{\partial H}{\partial S} \quad \frac{\partial H}{\partial I} \quad \frac{\partial H}{\partial M} \quad \frac{\partial H}{\partial P} \end{array} \right]' = \left[\begin{array}{c} -\lambda_1 \beta_1 I - \lambda_1 \beta_2 M - \lambda_1 d + \lambda_2 \beta_1 I + \lambda_2 \beta_2 M \\ 2b_1 I - \lambda_1 \beta_1 S + \lambda_1 \theta u(t) P + \lambda_2 \beta_1 S - \lambda_2 \theta u(t) P \\ -\lambda_2 d - \lambda_2 \alpha + \lambda_3 \delta \mu + \lambda_4 \eta \\ 2b_2 M - \lambda_1 \beta_2 S + \lambda_2 \beta_2 S - \lambda_3 u(t) P \\ \lambda_1 \theta u(t) I - \lambda_2 \theta u(t) I - \lambda_3 u(t) M - \lambda_4 \eta_0 \end{array} \right].$$

9) $\lambda_i(T) = 0, i = 1, 2, 3, 4.$

10) $\frac{\partial H}{\partial u} = 2b_3 u(t) + \lambda_1 \theta IP - \lambda_2 \theta IP - \lambda_3 MP = 0.$

得出使目标泛函式(4)到达最小值的一个最优控制为

$$u^*(t) = \max \left\{ \min \left\{ \frac{1}{2b_3} (\lambda_2 \theta IP + \lambda_3 MP - \lambda_1 \theta IP), 1 \right\}, 0 \right\}. \tag{9}$$

4 数值模拟和分析

根据前面的理论分析, 本文通过实验来验证使用适当的控制策略之后, 计算机病毒传播将得到有效控制. 并将杀毒率分别是常数和变量函数进行比较, 从图中看出控制模型能够更加有效地抑制病毒的传播.

给定初始条件如下:

$S(0) = 0.2, I(0) = 0.3, M(0) = 0.3, P(0) =$

$0.2,$ 经过对模型的初始条件及相关参数进行多次试验后, 选取了如下参数:

表1 数值模拟参数

Table 1 Simulation parameters

parameter	value	parameter	value	parameter	value
A	0.8	β_1	0.6	μ_0	0.1
β_2	0.6	θ	0.1	γ	0.1
α	0.8	δ	0.2	d	0.8
μ	0.3	η	0.1	η_0	0.15
P_0	0.1				

同时我们选取最优控制的时间为 100 天, 参数 $b_1 = 5000, b_2 = 5000, b_3 = 5000.$ 下面的图形中, 实线表示加控制时各个变量的变化趋势, 虚线表示不加控制时候各个变量的变化趋势. 采用表格 1 的参数, 在不加控制的时候, 基本再生数 $R_0 = \frac{A(\mu_0 P_0 \beta_1 + \mu \delta \beta_2)}{d \mu_0 P_0 (\alpha + d + \gamma P_0)} =$

5.2447, 显然是大于 1 的, 必然导致地方病的爆发, 也就是会使得计算机病毒的蔓延和传播. 从而理论的结果可以从图 1 到图 4 的虚线的数值模拟得到验证. 图 1 给出了没有控制和有加控制的 S 随时间变化的情况, 在没有控制的情况下, S 的数量随着病毒的增加而逐渐递减; 在有控制的情况下, 病毒得到了更有效的抑制, 因此 S 数量增多. 图 2 给出了没有控制和有控制的 I 随时间变化的情况, 在没有控制的情况下, I 的数量是越来越多; 而有控制的情况下, I 的数量在起始的前几天数量急剧下降, 从而得到了有效的抑制, 从我们的控制曲线很明显的可以看出, 在前几天, 控制的效率很大, 即杀毒软件的利用率很高, 必然会使被感染的计算机病毒的数量大幅度减小, 并最终渐近趋于零, 这样才是真正有效地杀毒. 图 3 给出了没有控制和有控制的 M 随时间变化的情况, 在没有控制的情况下, M 的数量是越来越多; 而有控制的情况下, M 的数量得到了抑制, 并最终渐近趋于零, 计算机中的恶意对象

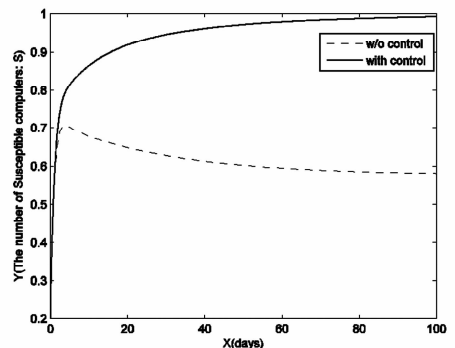


图1 控制前后易感电脑 S 的趋势变化

Fig. 1 The number of susceptible computers with control and without control

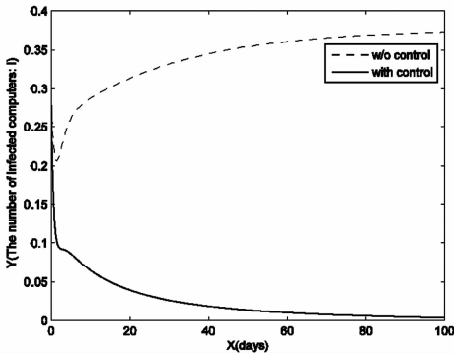
图2 控制前后被感染的电脑 I 的趋势变化

Fig.2 The number of infected computers with control and without control

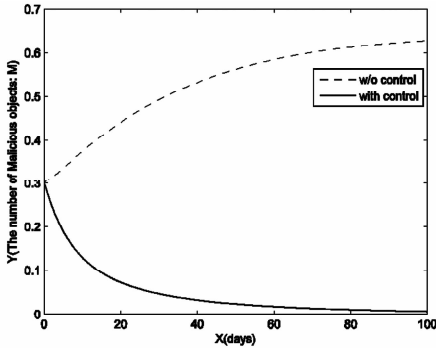
图3 控制前后恶意病毒 M 的趋势变化

Fig.3 The number of malicious objectives with control and without control

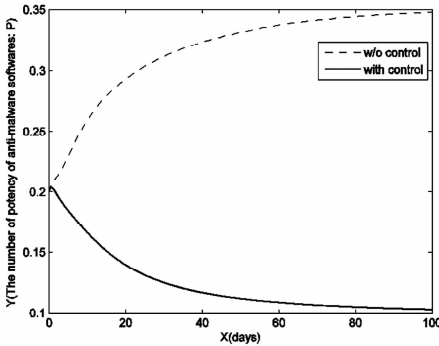
图4 控制前后杀毒软件 P 的趋势变化

Fig.4 The number of anti-malicious software with control and without control

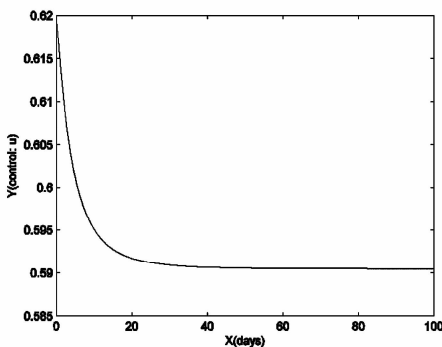


图5 控制变量曲线

Fig.5 The trend of control variable

随时间变化的情况,在没有控制的情况下, P 的数量随着病毒的增加而增加;在有控制的情况下,病毒得到了更有效的抑制,因此 P 数量减少. 将控制策略引入到本文模型中,大大降低了病毒的传播能力,同时,更多的 I 转变为 S ,更多的 M 被有效抑制. 因此,需要最优控制策略 $u(t)$ 来控制 I 和 M 的数目,从而控制病毒的传播,当病毒传播得到一定的控制后,采取控制的数量将会渐渐减少,控制策略 $u(t)$ 曲线图见图 5.

5 结语

通过上面的数值模拟可以很明显的看到,在有最优控制条件下,易感染的计算机的数量和被感染的计算机的数量趋势是符合目标函数的预期,而在没有控制的条件下,计算机病毒的传播无法被控制. 从控制量的走势也可以明显看到,并不需要一直持续的最大效力就可以将计算机病毒的传播控制住,这也是符合最优控制的目标:在费用最小的前提下,被感染的计算机的数量最少. 研究计算机病毒模型是为了更好地控制计算机病毒的传播,在此目的上,本文建立一个新的计算机病毒最优控制模型,通过有效的控制策略,成功地将感染病毒的节点控制在一个较低的水平,利用数值模拟进一步揭示了使用适当的控制方法之后,能够有效控制计算机病毒的传播,并且在有控制的条件下得到的效果比没有控制的条件下的模型的效果好,这也说明研究带有控制的计算机病毒模型更具有实际意义.

参 考 文 献

- 1 Pachpute G, Chakrabarty S P. Dynamics of hepatitis C under optimal therapy and sampling based analysis. *Communications in Nonlinear Science and Numerical Simulation*, 2013,18(8):2202~2212
- 2 Moufo P T, Tewab J J, Mewoli B, Bowong S. Optimal control of a delayed system subject to mixed control-state constraints with application to a within-host model of hepatitis virus B. *Annual Reviews in Control*, 2013,37(2):246~259
- 3 Costanza V, Rivadeneira P S, Biafore F L. Optimizing thymic recovery in HIV patients through multidrug therapies. *Biomedical Signal Processing and Control*, 2013,8(8):90~97
- 4 Lashari A A, Zaman G. Optimal control of a vector borne disease with horizontal transmission. *Nonlinear Analysis: Real World Applications*, 2012,13(1):203~212

- 5 Blayneh K W, Cao Y, Kwon H D. Optimal control of vector-borne diseases: treatment and prevention. *Discrete and Continuous Dynamical Systems-Series B*, 2009,11(3):587 ~ 611
- 6 Jung I H, Kang Y H, Zaman G. Optimal treatment of an SIR epidemic model with time delay. *Biosystems*, 2009,98(1):43 ~ 50
- 7 Mishra B K, Jha N. Fixed period of temporary immunity after run of anti-malicious software on computer nodes. *Applied Mathematics and Computation*,2007,190(2):1207 ~ 1212
- 8 Yang L X, Yang X F. The spread of computer viruses under the influence of removable storage devices. *Applied Mathematics and Computation*, 2012,219(8):3914 ~ 3922
- 9 Feng L P, Han Q, Wang H B, et al. Effective immune measures on P2P botnets. *Journal of Computer Application*, 2012,32(9):2617 ~ 2619,2623
- 10 Billings L, Spears W M, Schwartz I B. A unified prediction of computer virus spread in connected networks. *Physics Letters A*, 2002,297(6):261 ~ 266
- 11 Mishra B K, Saini D. Mathematical models on computer viruses. *International Journal of Applied Mathematics and Computer Science*, 2007,187(2):929 ~ 936
- 12 Yang L X, Yang X F, Zhu Q Y, et al. A computer virus model with graded cure rates. *Nonlinear Analysis: Real World Applications*, 2013,14(1):414 ~ 422
- 13 陈方方,洪灵.一类具有时滞和非线性发生率的SIRS传染病模型稳定性与Hopf分岔分析. *动力学与控制学报*, 2014,12(1):79 ~ 85 (Chen F F, Hong L. Stability and hopf bifurcation analysis of a delayed SIRS epidemic model with nonlinear saturation incidence. *Journal of Dynamics and Control*, 2014,12(1):79 ~ 85 (in Chinese))
- 14 Misra A K, Verma M, Sharma A. Capturing the interplay between malware and anti-malware in a computer network. *Applied Mathematics and Computation*, 2014,229:340 ~ 349
- 15 Kamien M I, Schwartz N L. Dynamics optimization: the calculus of variations and optimal control in economics and management. Netherland: Elsevier Science, 2000
- 16 Qi X H, Huang J Q, Dong H R, Yang Z J. Modern control theory and applications. Beijing: National Defence Industry Press, 2007:186 ~ 191
- 17 张旭龙,杨小帆.计算机病毒的最优控制模型. *计算机应用研究*,2011,28(8):3040 ~ 3043 (Zhang X L, Yang X F. Optimal control for computer viruses. *Applied Research of Computers*, 2011,28(8):3040 ~ 3043 (in Chinese))

ANOPTIMAL CONTROL MODEL FOR A NEW TYPE OF COMPUTER VIRUS*

Zhao Fuqun[†]

(Xianyang Normal University, College of Education Science, Xianyang 712000, China)

Abstract Due to the existence of computer virus, large numbers of computers cannot run normally. It usually results in huge loss. To this end, an control model for a new type of computer virus based on a general computer virus model is proposed in this paper. Firstly, a new type of computer virus model is improved, and the control is taken into account. The optimal control problem is then proposed. Meanwhile, the existence of optimal control is proved. Pontryagin's minimal principal is employed to analyze the optimal control. Finally, the theory analysis is also supplemented as a comparison work. The simulating results shows that under the condition without control, the basic reproductive number is obtained from the original model, and it is bigger than 1. Moreover, above conclusion indicates that there exists the endemic equilibrium, and it leads to the outbreak of the computer virus finally. The numerical modeling shows that through the optimal control strategy, the spread of computer virus can be prevented effectively.

Key words computer virus, minimum principal, optimal control, numerical simulation

Received 29 July 2014, revised 19 May 2015.

* Xianyang Normal University Special Scientific Research Fund Project (07XSYK224)

[†] Corresponding author E-mail:13240766843@163.com