

An Improved Computer Multi-Virus Propagation Model with User Awareness

Xiaoqin ZHANG ^{a,*}, Shuyu CHEN ^b, Huawei LU ^a, Fenggui ZHANG ^c

^a*School of Computer Science, Chongqing University, Chongqing 400030, China*

^b*School of Software Engineering of Chongqing University, Chongqing 400030, China*

^c*XiChang Satellite Launch Center, XiChang 615000, China*

Abstract

In this paper, an improved computer virus propagation model with user awareness is proposed and how its parameters are related to network characteristics is explained. Unlike other existing computer virus propagation models, we consider an important factor-user awareness. Furthermore, the stability and dynamics of the model are studied, and the simulation results show that the virus-free equilibrium and virus-endemic equilibrium are globally stable, whenever the population of network increases or decreases. Finally, the problem how to control the virus propagation is discussed, and we obtain that user awareness is closely related to the virus propagation.

Keywords: Computer Virus Propagation; U-SLIR Model; User Awareness; Stability Analysis

1 Introduction

The propagation of computer viruses can bring critical threat to the security of computer networks [1]. Computer viruses are programs of codes, once computers are attacked, the viruses can copy themselves quickly and it may be latent, which is the same as biology virus [2-3]. When our computers are attacked, maybe we should buy anti-virus programs to protect the security of systems, however not all users have the crisis awareness to detect viruses. Consequently, trials on better understanding computer viruses spreading dynamics is an important matter, in order to improve the safety and reliability in computer systems and networks [3, 4], and user awareness is very important for the control of computer viruses propagation [6, 7].

Models for the spreading of computer viruses based on epidemic models started to be reported in the late 1980s[2], which shows the long-term behavior of a virus propagation can be derived considering the network connections graph. From the biological point of view, the spread of virus is similar to sexually disease. Kephart used the epidemic models to find out the rule in the computer virus [7] and he focused his attention on the effect of topological structure of the

*Corresponding author.

Email address: happy1984zxq@gmail.com (Xiaoqin ZHANG).

network on the spread of the virus. The macroscopic approach is applied to propagation of e-mail viruses, and the original SIR model has been studied extensively [3], Jose R.C. Piqueira have studied SAIR model and analyzed its stability and dynamic behavior [8]. Z. Constantinescu and T. Gerthory proposed that user awareness is important to prevent the propagation of computer viruses by simulation [5].

Considering computer virus is latent, the SIR model can be modified including a block representing the computers that have the virus but are not infective to SEIR (Susceptible-Exposed-Infected-Recovered) model [9], and here we take user awareness into account the probability of taking anti-virus measures. And then, an improved SEIR model called u-SLIR (Susceptible-Latent-Infected-Recovered) model is proposed and studied with some clues about preventive attitudes being derived.

In section 2, the u-SLIR model is proposed, expressed by a system of first-order non-linear differential equations and the corresponding parameters are explained. In section 3, the virus-free equilibrium and the virus-endemic equilibrium are studied, and the stability of equilibriums is analyzed. In section 4, numerical experiments are conducted indicating that how to control the virus propagation.

2 U-SLIR Model

The model proposed here is a modification of SEIR model, and we assume that if one computer is infected and the virus is exposed, the user will install anti-virus programs with a proportion and it will become recoverable computer. If one computer is not infected or exposed, the user is probable to install anti-virus programs; the probability here is called user awareness.

The total population N is divided into four groups: S of non-infected computers subjected to possible infection; L of infected computers but the virus is latent; I of infected computers; R of computers equipped with anti-virus programs and will not be infected by the same virus.

We assume that there are some new computers joining into the existing network every time interval, and the item is described by rN ; $\lambda(t)$ represents the transition rate from S to L and $\mu(t)$ represents the transition rate from S to R , where $\lambda(t) = \omega I/N$ and $\mu(t) = \beta I/N$. $\lambda(t)$ describes the probability of susceptible computers to establish effective communications with infected ones, with proportion factor represented by ω , and $\mu(t)$ describes user awareness that is related to the infected population, with proportion factor represented by β , and we assume that the more the infected ones, the higher user awareness[6]. β represents the transition rate from L to I and $\mu(t)$ represents the transition rate from L to R . represents proportion coefficient for the mortality rate, not due to the virus.

Considering these facts, the model can be formulated as follows:

$$\begin{aligned}\dot{S} &= rN - \lambda(t)S - \mu(t)S - \mu S \\ \dot{L} &= \lambda(t)S - (\alpha + \mu(t) + \mu)L \\ \dot{I} &= \alpha L - (\beta + \mu)I \\ \dot{R} &= \mu(t)S + \mu(t)L + \beta I - \mu R\end{aligned}\tag{1}$$

Here in our model, if $r = \mu$, the total population N is stable; else the total population N increases or decreases. Fig. 1 describes the model as follows:

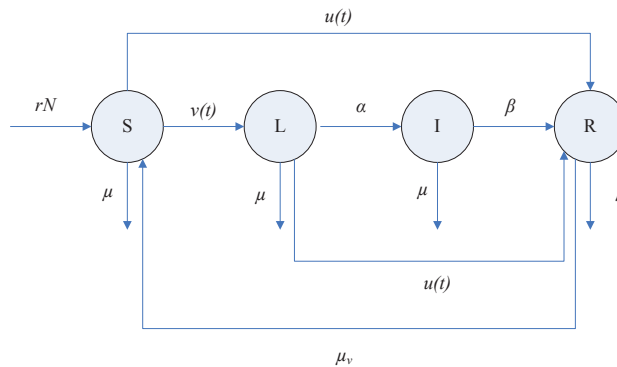


Fig. 1: U-SLIR model

It is important to notice that the first three equations in (1) do not depend on the fourth equation, consequently, the state-space dimension of the state-space is 3, and R can be expressed by $R = N - S - L - I$.

3 Virus-Free and Endemic Equilibrium

3.1 Virus-free equilibrium

Virus-free equilibrium is a state corresponding to the absence of infected population ($I = 0$), and we can calculate the steady point with this condition. Hence, we have the virus-free equilibrium:

$$E_f = \{S_f^*, L_f^*, I_f^*, R_f^*\} = \{rN/\mu, 0, 0, 0\} \tag{2}$$

Considering the equilibrium point, the corresponding Jacobian matrix is given by:

$$J_f = \begin{bmatrix} -\mu & 0 & -(\omega + \beta)r/\mu \\ 0 & -(\alpha + \mu) & \omega r/\mu \\ 0 & \alpha & -(\beta + \mu) \end{bmatrix} \tag{3}$$

The corresponding eigenfunction of J_f is:

$$(\mu + \lambda)(\lambda^2 + (\alpha + \mu + \beta + \mu)\lambda + (\beta + \mu)(\alpha + \mu) + \alpha\omega r/\mu) = 0 \tag{4}$$

By the Routh-Hurwitz criterion, if all eigenvalues are negative, the virus-free equilibrium will be locally asymptotically stable. Therefore, the following conditions should be satisfied:

$$\alpha\omega r < \mu(\beta + \mu)(\alpha + \mu) \tag{5}$$

Let us define the basic reproduction number of the infection as

$$K = \alpha\omega r / (\mu(\beta + \mu)(\alpha + \mu)) \tag{6}$$

Then we have the following lemma regarding the stability of the virus-free equilibrium.

Lemma 1 *The virus-free equilibrium E_f is locally asymptotically stable if $K < 1$ and unstable if $K > 1$.*

Proof By the stability theory [10], the sufficient condition for the three-dimension system to be asymptotically stable is that $\lambda_i < 0$, for $i = 1, 2, 3$. It is easy to see that $\lambda_1 = -\mu < 0$, and by the Routh-Hurwitz criterion, $\lambda_2, \lambda_3 < 0$ is equal to $(\beta + \mu)(\alpha + \mu) - \alpha\omega r/\mu > 0$, which is exactly the sufficient condition in the lemma.

Further, we can have the following theorem.

Theorem 1 *The virus-free equilibrium E_f is globally asymptotically stable if $K \leq 1$*

Proof The following inequality can be obtained from the first equation of system (1):

$$S'(t) \leq rN - \mu S(t) \quad (7)$$

Thus

$$S(t) \leq rN/\mu - (S(0) - rN/\mu) \exp^{-\mu t} \quad (8)$$

when $t \rightarrow \infty$, we obtain

$$S(t) \leq rN/\mu \quad (9)$$

From the second equation of system (1) we obtain

$$L'(t) \leq \lambda(t)S(t) - (\alpha + \mu)L \leq \omega rI/\mu - (\alpha + \mu)L \quad (10)$$

Thus

$$L(t) \leq \omega rI/(\mu(\alpha + \mu)) - (L(0) - \omega rI/(\mu(\alpha + \mu))) \exp^{-(\alpha + \mu)t} \quad (11)$$

when $t \rightarrow \infty$, we obtain

$$L(t) \leq \omega rI/(\mu(\alpha + \mu)) \quad (12)$$

Let us consider the following Lyapunov function defined by

$$V(t) = I(t) \quad (13)$$

The time derivative of $V(t)$ along the solution of the third equation of system (1) is given by

$$V'(t) = I'(t) = \alpha L - (\beta + \mu)I \leq \alpha\omega rI/(\mu(\alpha + \mu)) - (\beta + \mu)I = I(\alpha\omega r/(\mu(\alpha + \mu)) - (\beta + \mu)) \quad (14)$$

We can obtain that $V'(t) \leq 0$ with the condition $K \leq 1$.

Thus, we prove that the virus-free equilibrium E_f is globally stable.

3.2 Endemic equilibrium

Endemic equilibrium points are characterized by the existence of infected nodes in the network ($I \neq 0$), for the case of $I \neq 0$, we can get the virus-endemic equilibrium

$$E_e = \{S_e^*, L_e^*, I_e^*, R_e^*\} = \{(N/\omega)^*[(\beta + \mu)/\alpha]^*(\alpha + \mu + \beta^* I_e^*/N), (\beta + \mu)^* I_e^*/\alpha, I_e^*, R_e^*\} \quad (15)$$

And the corresponding Jacobian matrix at E_e is

$$J_e = \begin{bmatrix} -(\omega + \beta)I_e^*/N - \mu & 0 & -(\omega + \beta)S_e^*/N \\ \omega I_e^*/N & -(\alpha + \mu) - \beta I_e^*/N & \omega S_e^*/N \\ 0 & \alpha & -(\beta + \mu) \end{bmatrix} \quad (16)$$

The eigenfunction of J_e is $f(\lambda) = a_0\lambda^3 + a_1\lambda^2 + a_2\lambda + a_3$, where

$$\begin{cases} a_0 = 1, \\ a_1 = (\beta + \mu) + (\alpha + \mu) + \beta I_e^*/N + (\omega + \beta)I_e^*/N + \mu, \\ a_2 = ((\beta + \mu) + (\alpha + \mu) + \beta I_e^*/N)((\omega + \beta)I_e^*/N + \mu), \\ a_3 = (\omega I_e^*/N)(\alpha(\omega + \beta)S_e^*/N) \end{cases} \quad (17)$$

Lemma 2 *As long as $K > 1$ holds, the endemic equilibrium is E_e locally asymptotically stable.*

Proof By the Routh-Hurwitz criterion, if we can verify a_0, a_1 and $a_1a_2 - a_0a_3$ have the same sign, then the three eigenvalues all have negative real parts. Obviously $a_0 > 0$ and $a_1 > 0$, and $a_1a_2 - a_0a_3 > 0$ holds by computation of a_0, a_1 and a_2 . Therefore, the Routh-Hurwitz stability conditions are satisfied, which implies that the virus-epidemic equilibrium is locally asymptotically stable.

4 Numerical Analysis

In this section, some numerical experiments are conducted by MATLAB with two different cases being considered: $K \leq 1$ and $K > 1$. $N = 10000$ and $\mu = 0.0025$ are given, and other parameters are not fixed to change the value of K , and our simulation is carried out to find out the influence of parameters on computer virus propagation. In all figures, horizontal axis represents time and vertical axis represents the number of computers.

4.1 The case of virus-free equilibrium

In our simulation, $K \leq 1$ should be satisfied. And $S(0), L(0)$, and $I(0)$ represent the initial number of computers of each state, and we assume that the initial total population is $N = 10000$, at the same time, $\alpha = 0.5$ and $\omega = 0.6$ are given. We first consider the case that $r = \mu$ which means the total population is stable. Fig. 2 shows that as long as $K \leq 1$, the virus-free equilibrium is globally stable, and S reaches the maximum N , and the equilibriums of L, I and R are 0.

When $r > \mu$, the total population of the network is larger than the initial population, and Fig. 3 shows the virus-free equilibrium is also globally stable in this case, and we can believe that the total population can be stable in any case.

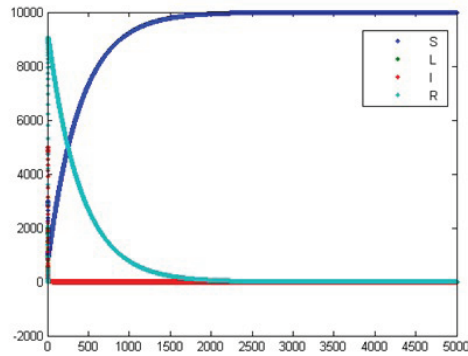


Fig. 2: The virus propagation results with $K < 1$ and $r = \mu$.

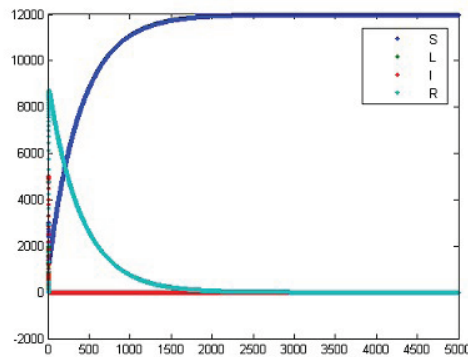


Fig. 3: The virus propagation results with $K = 0.5955$ and $r > \mu$.

4.2 The case of virus-endemic equilibrium

In this case, we should modify the corresponding parameters to satisfy $K > 1$, and Fig. 4 shows that as long as $K > 1$, the virus-endemic equilibrium is globally stable. S and R can reach stable equilibriums, and L and I are relatively low, respectively. Fig. 5 shows that when K reaches a certain value, the recoverable computers will be more than the susceptible ones, and the latent and infected computers are a little more, respectively.

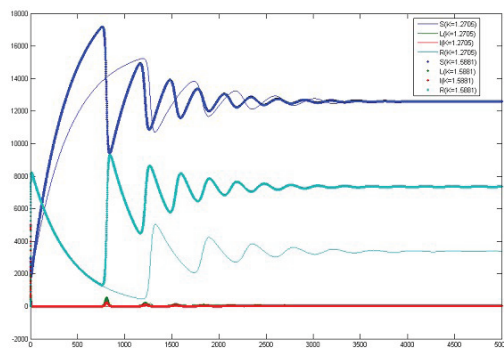


Fig. 4: The virus propagation results with $K > 1$ and $r > \mu$.

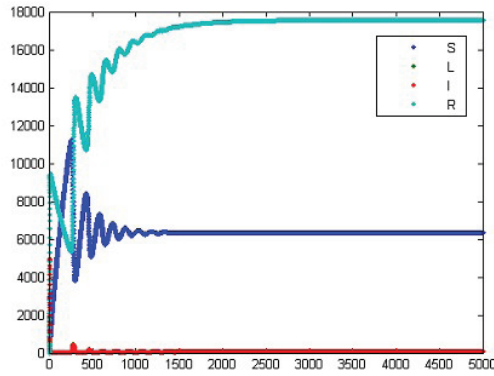


Fig. 5: The virus propagation results with $K = 3.8019$ and $r > \mu$.

4.3 Virus endemic control

When $K > 1$, our experiments show that the virus-endemic equilibrium is globally stable, and we can control the number of latent and infected computers by controlling K . We can learn that the value of K is closely related to user awareness proportion β from the equation (6), therefore, we should improve the awareness of users to make the quantity of latent and infected computers low. And we can also modify other parameters to control the virus-endemic equilibrium. Fig. 6 shows that the bigger K is, the more the latent and infected computers are, therefore, we can control the virus propagation by controlling K , which means the corresponding β can be modified.

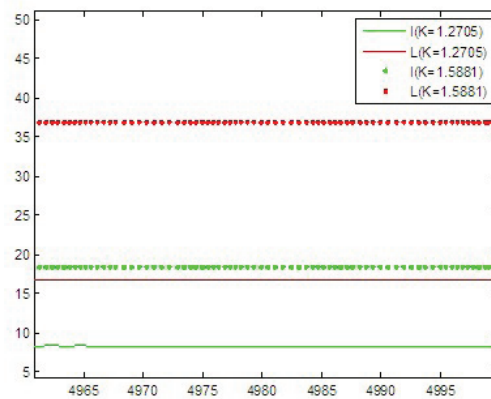


Fig. 6: The virus propagation results with $K > 1$ and $r > \mu$.

5 Conclusion

The stability analysis and simulation results have revealed that the virus was able to pervade if $K > 1$, and the virus endemic would be died out when $K \leq 1$. And when influx rate r is not equal to mortality rate μ , the total population is also globally stable. Furthermore, when K reaches a certain value, the recoverable computers will be more than the susceptible ones, and the results also show that it is difficult to control the long-term virus propagation equilibrium with the recoverable computers are more and the latent and infected computers are less at the same time. Finally, we have discussed the influence of user awareness on controlling the virus-endemic

equilibrium, and we should improve the awareness of users to prevent the propagation of virus. The future work will center on the more realistic model, including network connectivity, more complex network scenarios and manual factors.

References

- [1] P. J. Denning, *Computers Under Attack*, Addison-Wesley, Reading, Mass, 1990.
- [2] J. O. Kephart, S. R. White, D. M. Chess, Computers and epidemiology, *IEEE Spectrum* (1993) 20 – 26.
- [3] J. O. Kephart, G. B. Sorkin, M. Swimmer, An immune system for cyberspace, in: *Proceedings of the IEEE International Conference on Systems, Men, and Cybernetics*, Orlando, CA, October 1997, pp. 879 – 884.
- [4] Chunming Zhang, Xiaofan Yang, Qingyi Zhu, Wanping Liu. Optimal Control in a Novel Computer Virus Spread Model, *Journal of Information and Computational Science*, 2011 (8): 1929 – 1938.
- [5] Z. Constantinescu, T. Gerthory, The dynamics of computer virus infection, *System dynamics*, May, 2002.
- [6] C. Jin, J. Liu, Q. H. Deng, A Novel Email Virus Propagation Model, *Power Electronics and Intelligent Transportation System*, 2008.
- [7] J. Kephart, S. White, Measuring and modeling computer virus prevalence, in: *1993 IEEE Symposium on Security and Privacy*, 1993, pp. 2 – 15.
- [8] J. R. C. Piqueira, V. O. Araujo, A modified epidemiological model for computer viruses, *Applied Mathematics and Computation* 213 (2009) 355 – 360.
- [9] R. M. Anderson, R. M. May, *Infectious diseases of humans: dynamics and control*. Oxford: Oxford University Press, 1991, pp. 304 – 318.
- [10] R.C. Robinson, *An Introduction to Dynamical System: Continuous and Discrete*, Prentice Hall, USA, 2004.