IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 17, Issue 1, Ver. III (Jan – Feb. 2015), PP 43-47
www.iosrjournals.org

Modeling and Threshold Sensitivity Analysis of Computer Virus Epidemic

Abidur Rahaman[1], Amran Bhuiyan[2], Md. Ahsan Habib[3] and Z. H. Mozumder[4]

[1](Dept.of ICT,Noakhali Science and Technology University(NSTU),Noakhali-3814,Bangladesh)
[2](Dept.of CSTE,Noakhali Science and Technology University(NSTU),Noakhali-3814,Bangladesh)
[3](Dept. of ETE,Begum Rokeya University,Rangpur,Bangladesh)
[4](Dept. of EEE,University of Dhaka,Dhaka-1000,Bangladesh)

Abstract: *This paper analyzed the methods and techniques used in mathematical modeling of biological epidemics to the domain of information technology. A new epidemic model has been proposed by incorporating a range of parameters rather than using constant parameters for a fixed population network. All individual nodes or computers are classified into four classes- susceptible, exposed, infected and immunized. Changes of population of these four classes with time and higher epidemic threshold are studied. Finally threshold sensitivity analysis has been analyzed to take into account the appropriate measures to control the virus epidemic.*

Keywords:*Computer virus epidemic, epidemic threshold, sensitivity analysis*

I. Introduction

Computer virusespossess a major threat both for standalone and networked computers as they can replicate themselves and spread among computers in the form of malicious programs. Destruction of data by the viruses cause serious problem for individual user and may cause disastrous situation for institutional user, even sometimes destroy the whole computer system [1]. Despite the significant development of anti-virus as a major means of defending against viruses, the computer viruses are still very much a cause of concern in computer network. As a promising alternative of anti-virus technique, the epidemic dynamics of computer viruses aims to understand the way how the computer viruses can spread across network and to work out global policies of inhibiting their prevalence. Analogous behavior of computer viruses and their biological counterparts inspired many researchers to study this new filed computer viruses study. Cohen [2] and Murray [3] evidently suggested exploiting the compartment modeling techniques developed in the epidemic dynamics of biologically infectious disease to study the spread of computer viruses. Later, Kephart and White [4] model the virus spreading in the Internet based on the biological epidemic model. All the state-of-the-arts works related this model is explained below.

In [5-7], it was found that the computer network (i.e. Internet) follows diverse power-law degree distributions. This finding initiates the interest in virus spreading in complex networks, leading to the surprising finding that the epidemic threshold vanishes for scale-free networks with infinite size [8]. To fully understand the spreading mechanism of computer viruses, multifarious virus spreading models, ranging from conventional models to unconventional models such as delayed models, impulsive models and stochastic models have been proposed in [16-17]. All the above mentioned models have been implemented based on fixed parameters.

This paper is intended to develop an epidemic model of computer viruses that can be checked for a range of parameters rather than for a constant parameter. It also shows how epidemic threshold affects the total process and thus decisive measures can be taken to control the viral epidemic.

The rest of the paper has been organized as follows. Section 2 describes of the Mathematical Modeling of Computer Virus Epidemic in detail. Sec. 3 estimates the parameters, Sec.4 reports the result and discussion and finally concluding remarks are drawn in Sec. 5.

II. Mathematical Modeling Of Computer Virus Epidemic

In normal biological epidemics, if each disease carrier infects less than one person, then the infection will eventually die out. This case can mathematically be expressed by the basic reproduction number, $R_0 < 1$. The first full analysis of a computer virus invading a network using similar techniques to those used in analyzing biological epidemics by Kephart and White [4]. It applies a Susceptible-Infected-Susceptible (SIS) model to a random graph of N nodes which represent computers in a network with a general virus propagating throughout. More specific models have been developed since to deal with the different types of computer virus

[15]. We shall now introduce a model developed to analyze the spread of acomputer virus through a network of computers which we shall call the CVE model. It is a modification of a model proposed by Yuan [9].This network is assumed to be a one in which each computers connectivity is identical, thus subject to similar governing principles as biological systems. It is assumed that each individual member of the total population is a single computer and that the computer virus is able to infect but not destroy a host. One further assumption is that once infected, a computer enters an exposed state, described by the class E, before entering the infected and contagiousclass I after a certain length of time. Those in E have the virus but its code has yet to be activated and thus cannot infect susceptible. At the same time, computers are being updated with antivirus software which enables each computer to become immune to the virus modeled by the class R.The model is described by the following system of differential equations:

$$\frac{dS}{dt} = \mu N - \frac{rSI}{N} - (p_{SR} + \mu)S \ldots\ldots (1)$$

$$\frac{dE}{dt} = \frac{rSI}{N} - (\alpha + p_{ER} + \mu)E \ldots\ldots (2)$$

$$\frac{dI}{dt} = \alpha E - (\gamma + \mu)I \ldots\ldots\ldots\ldots\ldots (3)$$

$$\frac{dR}{dt} = p_{SR}S + p_{ER}E + \gamma I \ldots\ldots\ldots\ldots (4)$$

For steady states

$$\frac{dS}{dt} + \frac{dE}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0 \ldots\ldots\ldots\ldots (5)$$

Let us define also epidemic threshold,

$$R_0 = \frac{\alpha\mu r}{(\alpha + pER + \mu)(\gamma + \mu)(pSR + \mu)} \ldots (6)$$

Here S, E, I and R, are all functions of t only and allconstants are assumed to be strictly positive. This model assumes a fixed population as seen through addition of the rates of change of the separate classes. The rate of this replacement can be varied by the coefficient μ in the model and the introduction of these new computers is modeled by the term μN inthe first equation. It isassumed for simplification of the model that newcomputers when introduced are not immune to the virus. Removal of computers is modeled by each of the −μ terms in each equation and since N = S + I + E + R, this removal and replacement cancels out but still has an effect on how the virus spreads.Infection of a computer withthe virus from an infective is assumed to be proportional to the product SI given in (1), which represents connectionsbetween susceptible and infective computers. A fraction r of these will result in a successful passing on of the virus. These initial infections pass into the infected but not contagious class E via the interaction rSI in equation (2).

Individuals in E whose virus code becomes active move to the infected and contagious class I via the term −αE in equation (2). It is assumed that the rate of change of this activation is proportional to the size of E and can be varied by the constant α which is a reasonable assumption to make. Immunization of individual computers from the susceptible, exposed and infected classes to the immunized class R is carried out via the −pSRS, −pER E and −γI terms in the model. pSR S models the effect of real-time immunization, pER E models the effect of immunizing individual computers in the non-infective class E and γI is the recovery rate from I to R. These constants can be changed to model the differing rates of immunization for each class. These individuals pass to the vaccinated class R and remain there unless replaced through the term −μR. Clearly if there is no replacement included in this model (μ = 0) then eventually the entire population of computers will exist in R and will therefore be immune to the virus.

III. Estimation Of Parameters

We have six different constants which all have an impact on the evolution of the model over time. We considered a population of N =100,000 computers and μ = 1/4800 which represents a replacement rate of approximatelymore than half a year (200 days or seven months, time is measured in hours). The constant r represents the average number of newcomputers being infected per new infection. Finding an appropriate figure for r is a difficult task as many viruses or worms infect at differing rates. We can get a reasonable idea by looking back at the analysis of the fastest spreading worms of the recent past. The Code-RedI worm infected over 359,000 unique IP addresses in the 24 hour period after 19th July 2001 with a peak infection rate of 2000 news hosts per minute [10]. One of the fastest worms in history, the 2003 Slammer/Sapphire worm, doubled in size approximately every 8.5 seconds, and after targeting 200,000 servers worldwide, had infected over 75,000 in under 10 minutes or 7,500 per minute [11]. By analyzing all of these including [12] [13], we choose the value of r = 27 as a starting point. Estimation of α, the rate that the virus code is activated is more difficult to calculate, as it can depend from case to case. Some viruses can lay dormant for many days if they require a user to activate

its code, such as in the case of many email viruses. Others will activate almost instantly, spreading their malicious code to other susceptible machines in the locality. For the purposes ofthis model we shall assume that the virusremains inactivefor an average time oft_E = 6 hours.Therefore we must choose (α + pER + μ) such that $\frac{1}{t_E} = (\alpha + p_{ER} + \mu)$.We will choose a value of α = 0.1 for the rate of activation and a rate of immunization of exposed members aspER = 0.066. The only constants left to estimate are the rate of immunization for the infective and susceptible classes, γ and pSR respectively. We shall assume for this model that immunization is carried out through updates which can be acquired via the internet, and that uptake of these updates will be higher in theinfected class,I than in the non-infected class, S. We shall assume that the average time required for an infective to be immunized against the threat is t_I = 1 and thus we shall set γ = 1 .

Finally for pSR , we will assume that the immunization rate is lowest in this class as those with a non-infected computer will not be as aware of the threat and therefore uptake of updates will be lower, and therefore we shall set pSR = 0.02. The initial class sizes are important to have the model set up in a sensible manner. As we are modeling the introduction of a new virus into a group of infective, the exposed and recovered classes will initially be of zero size. Therefore we will set the initial population of infective to be N (0) =99,990 and the infective carriers to have size I(0) = 10. A preliminary list of chosen values for the different constants is given below.

Table (1): Parameter Estimation Table for CVE Model.

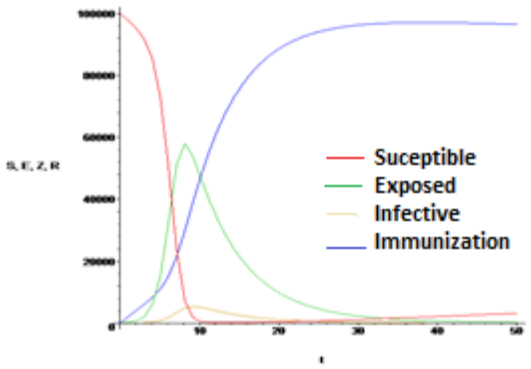
Parameter	Description	Value
N	Total Population Size	100,000
S(0)	Initial size of Susceptible Class	99,990
E(0)	Initial size of Exposed Class	0
I(0)	Initial size of Infective Class	10
R(0)	Initial size of Recovered Class	0
μ	Replacement Coefficient	0.00201
r	Rate of Infection	27
α	Rate of Virus Activation	0.1
pSR	RealTime Immunization-Susceptible	0.02
pER	Rate of Immunization– Exposed	0.066
γ	Rate of Immunization–Infective	1

This initial setup for the CVE model is a reasonable starting point and leads to a reproduction number of R_0 = 1.46 which is the condition for a endemic steady state.

IV. Results And Discussion

After solving the system, we obtain the Population-Time behavior of four classes (Susceptible, Exposed, Infected and Removed) (Fig(1))**.**The combined plot shows the proposedCVE(Computer Virus Epidemic) models evolution for the time t= 50hours (Red-Susceptible, Green-Exposed, Yellow-Infected,Blue-Immunized) (Fig (1)). In the first few hours the susceptible class initially decreased due to the increase in immunized computers. After 2 hours, the rate of infection(r) increased dramatically from approximately 20/hr to well over 100/hr and by the fifth hour the rate of infection peaks at 670/hr. By this time instant the susceptible class shrunk to only one third of its original size and was falling at a higher rate than immunization .Two hours later there were virtually no susceptible left to infect and thus the rate ofinfection suddenly decreased. At the same time the rate of increase in the infective class reached its maximum before becoming zero at approximately 13 hours after the initial introduction of the virus. All the while, the immunized class increased until it reached asymptotically its maximum value at the endemic steady state, $\overline{R} \approx 98{,}900$.The other classes tend to their fixed points$\overline{S}\approx 500$, $\overline{I}\approx 200$ and $\overline{E}\approx 400$(Fig(1)). To study the influence of a high epidemic threshold value (R_0) on the scale and timeframe of an epidemic, the value of $R_0$was gradually increased and the effect of it on different types of population was observed simultaneously (Fig (2) to Fig (5)). As R_0 gradually increases rate of infected population's changes dramatically. The peaks of infected and exposed individuals sharply increased with increasing R_0 and they moved to the left of the graph. As a result susceptible populations decrease sharply. In every moment summation of all four populations remains constant and it is N=100000. It was calculated earlier for the case of steady states. Epidemic threshold depends on six parameters which governs in equation(6) .The sensitivity of R_0 with different parameters was studied .It was seen that R_0 is linearly proportional to the rate of infection(r). So a higher (r) might turn to a higher R_0(Fig(7)). It became clear that rate of immunized-exposed class (pER),rate of immunized-susceptible (pSR)and the rate of immunized- infective class (γ) all influences R_0 significantly. With the increase ofpER,pSR and γ it was seen that R_0 decreased exponentially (Fig (8), (9), and (10)) .On the other hand, increasing virus activation rate (α) evidently increasedR_0. It increased R_0exponentially (Fig(6)). So to control epidemic we have to decrease it.Replacement coefficient (μ) shows its

effect onR_0. First R_0 increased exponentially with increase of (μ) than after a threshold value R_0decreased exponentially with increasing μ. (Fig(11)) Proper selection of μ is an important aspect to control epidemics.



Figure(1): Change in different types of population with time (R_0= 1.46, time 50 hours).

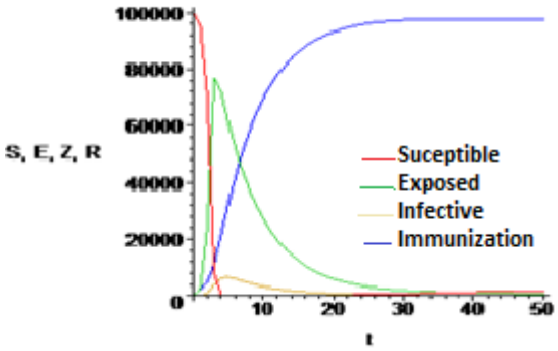


Figure (2): For R_0=2.7, time 50 hours.

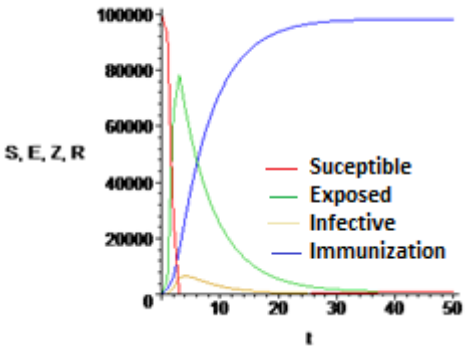


Figure (3): For R_0=8.1, time 50 hours

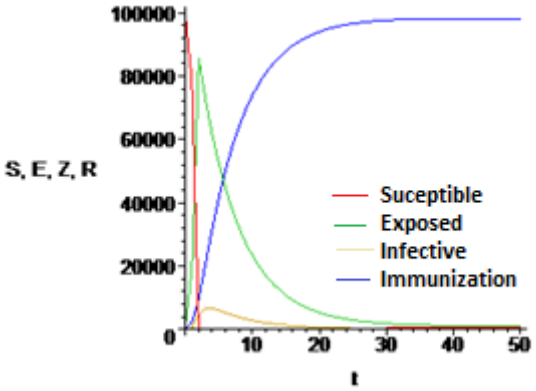


Figure (4): For R_0=13.5, time 50 hours

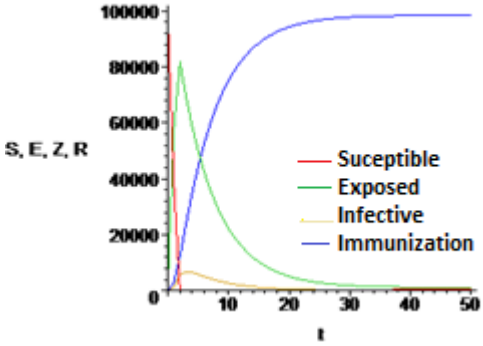


Figure (5): For R_0=21.6, time 50 hours.

a. Sensitivity of R_0 with its parameters

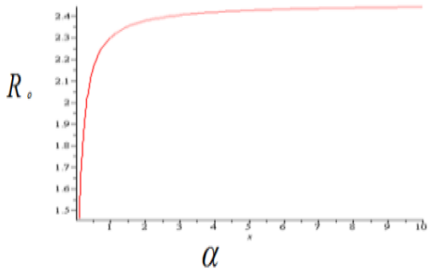


Figure (6): Variations of $R_0$for $\alpha = (0.1 - 10)$

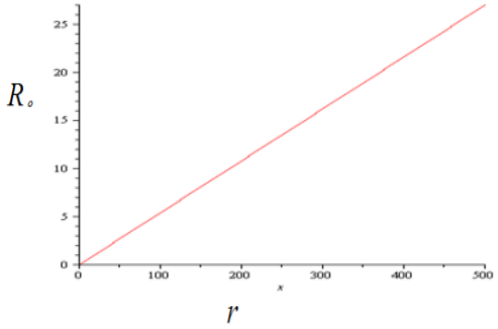


Figure (7): Variations of $R_0$for $r = (0 - 500)$

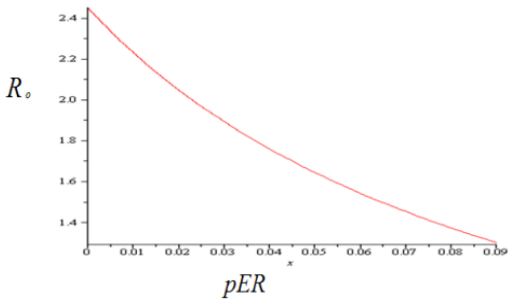


Figure (8): Variations of R_0 for $p_{ER} = (0 - 0.09)$

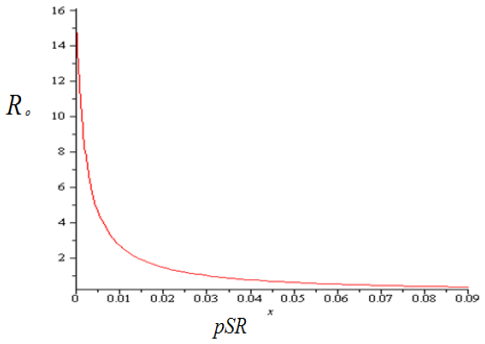


Figure (9): Variation of $R_0$for $p_{SR} = (0 - 0.09)$

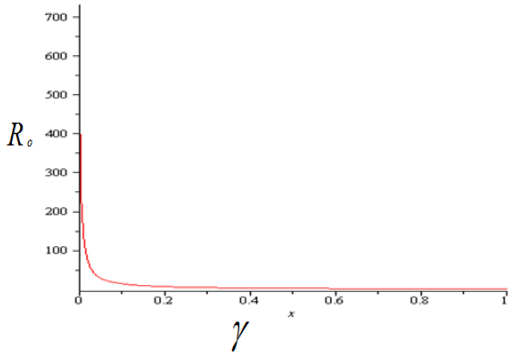


Figure (10): Variation of R_0 for $\gamma = (0 - 1)$

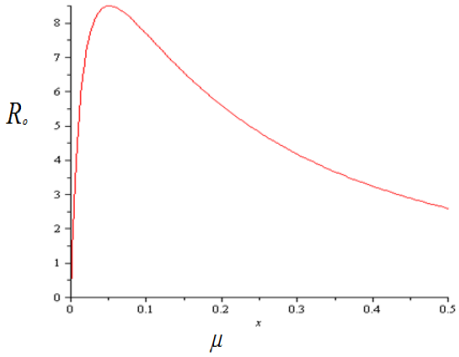


Figure (11): Variation of R_0 for $\mu = (0 - 1)$

V. Conclusion

This works gives a graphical understanding of how a computer virus epidemic evolves and behaves over time. Realization and choosing appropriate parameters for epidemic threshold shows the way to curb an epidemic in cyber world. It could be a good insight for antivirus program developer. In the light of this, further work can be done on immunization strategy or quarantine strategy and application of controltheory may give good result.

References

[1]. Szor P. The Art Of Computer Virus Research And Defense. 1st Ed. Addison-Wesley Education Publishers Inc.: 2005.
[2]. Cohen F. Computer Viruses: Theory And Experiments. Computsecur 1987;6(1):22-35.
[3]. Murray WH. The Application Of Epidemiology To Computer Viruses.Computsecur 1988:7(2) 130-50.
[4]. Kephart, J., White, S., Directed-Graph Epidemiological Models Of Computer Viruses, IEEE Symposium On Security And Privacy (1991).
[5]. Faloutsos M, Faloustsos P, Faloustsos C. On Power-Lae Relationships Of The Internet Topology. ACM SIGCOMM ComputcommunRev 1999,29(4):251-62.
[6]. Albert R, Barabasi A-L. Statistical Mechanics Of Complex Networks. Rev Mod Phys 2002.74(1):47-97.
[7]. Revasz E, Barbasi A-L. Hierarchical Organization In Complex Networks.Phys Rev E 2003,27(2)(Article ID 026112).
[8]. Pastor-Satorras R, Vespignani A. Epidemic Spreading In A Scale-Free Networks.
[9]. Yuan, H. And Chen, G., Network Virus-Epidemic Model With The Point- To-Group Information Propagation, Applied Mathematics And Computation 206 (2008) 357-367
[10]. Moore, D, Shannon, C, Brown, J, Code-Red: A Case Study On The Spread And Victims Of An Internet Worm, Http://Www.Caida.Org/ Publications/Papers/2002/Codered/Codered.Pdf
[11]. Moore, D, Paxson, V, Savage, S, Shannon, C, Staniford,S, Weaver, N, Inside The Slammer Worm, Http://Www.Cs.Ucsd.Edu/~Savage/ Papers/IEEESP03.Pdf
[12]. ConfickerWorm Spikes Infects 1.1million Pcs In 24 Hours, Http://Arstechnica.Com/Security/News Conficker-Worm-Spikes-Infects-1-1-Million-Pcs-In-24-Hours.Ars
[13]. Internet Usage Statistics,Http://Www.Internetworldstats.Com/Stats.Htm
[14]. VBS Love Letter And Variants, Http://Www.Symentec.Com/Security_Response/Writeup.Jsp
[15]. Zou, C. Et. Al, Email Virus Propagation Modeling And Analysis, Technical Report TR-CSE-03-04, Department Of Electrical &Computer Engineering, Univ. Massachusetts.
[16]. Misra BK, Jha N.SEIQRS Model For The Transmission Of Malicious Objects In Computer Network. Appl Math Model 2010,34(3):710-5.
[17]. Zhang C,Yang X, Ren J. Modeling And Analysis Of The Spread Of Computer Virus, Commun Nonlinear Scinumersimul 2012;17(12):5117-24.
[18]. Robinson RC. An Introduction To Dynamic Systems: Continuous And Discrete. 1st Ed. Prentice Hall;2004.