Epidemics of Mobile Worms

Yuriy Bulygin
Intel Corporation
Security Expertise Center of Excellence
JF4-318, 2111 NE 25th Ave, Hillsboro, OR 97124-5861, U.S.A.
yuriy.bulygin@intel.com

Abstract

Mathematical epidemiology has been developing for over a hundred years. Mathematical models developed for spreading of viruses in human populations take into account various effects influencing the epidemic. Recently results of mathematical epidemiology have also been applied to modeling of epidemic viruses in computer networks. The current work applies results of mathematical epidemiology to the research of dynamics of potential epidemics of worms in mobile networks. The work simulates epidemics of two types of mobile worms that we have already run into - MMS and Bluetooth worms. [1]

1 Introduction

This work uses basic models of mathematical epidemiology for simulation of spread of Bluetooth and MMS worms in mobile networks. These models give us a very rough picture of spread dynamics of those worms in mobile networks. Currently there are no Bluetooth or MMS worms known in the wild that do not require user action to infect a vulnerable mobile device. However remotely exloitable vulnerabilities in wireless LAN/WAN and Bluetooth device drivers or firmware may become fruitful targets and can result in the spread of the worm in wireless networks.

The main goal of the current work is to provide estimations of potential scale and timing characteristics of worm propagation over MMS and Bluetooth capable mobile devices. Epidemiological models used in this paper for simulating MMS and Bluetooth worms also apply to the analysis of worms exploiting vulnerabilities in wireless LAN/WAN devices.

[1] The results of this work do not represent the views of Intel Corporation.

2 Bluetooth Worms

We will briefly describe the mathematical model applied to the analysis of epidemic of Bluetooth worms, such as Worm.SymbOS.Cabir and Worm.SymbOS.Lasco.

2.1 The Susceptible-Infectious Model

The *Suceptible→Infectious (SI)* model is the simplest model of the dynamics of viral epidemics [7, 2]. In the given model the individual is either healthy and vulnerable to infection or infected and thus infecting others. The size of the population is constant and equals $N = S(t) + I(t)$ where in the moment of time t there are $S(t)$ vulnerable and $I(t)$ infected individuals. Density of infected individuals $i(t) = \frac{I(t)}{N}$ changes proportionately to the density of vulnerable individuals $s(t) = 1 - i(t)$, the average number of contacts the vulnerable individual has in a unit of time $\langle k \rangle$, and the probability of infection during the entire time. The probability of infection, in turn, is equal to the product of probability of infection during contact with the infected individuals λ and the density of those infected $i(t)$. The change in the number of infected individuals is thus described by the equation known as the *Verhulst equation*:

$$\frac{di(t)}{dt} = \lambda \langle k \rangle [1 - i(t)] i(t) \tag{1}$$

The solution is *logistics function*, which reflects exponential growth of the number of infected individuals in a two-component population, where $\tau = \frac{1}{\lambda \langle k \rangle}$ is the characteristic time of propagation of the epidemic:

$$i_{SI}(t) = \frac{i_0 e^{\frac{t}{\tau}}}{1 - i_0 + i_0 e^{\frac{t}{\tau}}} \tag{2}$$

2.2 Epidemics of Bluetooth Worms

What is the meaning of the quantity λ in epidemics of network or mobile worms? In the case of the disease that is

contracted by respiratory means that causes Respiratory Viral Disease in an individual, λ means that during the contact with the person ill with Respiratory Viral Disease we become ill with the probability of λ. It depends on a multitude of factors, including the virulence of the disease, immunity of the individual, etc., but for the mathematical model of the viral epidemic, the value of this quantity may be estimated based on statistical data from past epidemics.

Bluetooth worms have the following characteristics of spreading:

1. The range of infection is limited by the range of the Bluetooth connection, i.e. 10-20 meters.

2. Due to this the Bluetooth worm cannot infect the next victim selectively, such as from a prepared list or by a randomly generated mobile phone number. That is, infection spreads spontaneously: if the vulnerable individual is discovered in the infection range, then the attempt at infecting him is carried out.

Thus, modeling the dynamics of the Bluetooth worm epidemic may on one hand be considerably simplified, but on the other is a very complex pursuit. It is possible to simplify modeling by using the model and parameters used for the analysis of transmission dynamics of respiratory viral diseases, such as the human influenza virus, its range of infection being limited to several meters.

The order of exponent $\lambda \langle k \rangle$ in the *SI* model is none other than the number of infected individuals in a unit of time. In the literature on mathematical epidemiology for the human influenza virus, this index varies within the limits of $0.3 - 2$ per day or $100 - 700$ of infected individuals per year. We use $\lambda = 1.37$ for modeling Bluetooth worms, which constitutes 500 infected individuals per year. We will take the size of the vulnerable population to be equal to the number of mobile phones in Moscow, Russia, i.e. 10^7, and that the initial infection is carried out from one mobile phone $I_0 = 1$. Thus, the parameters of the *SI* model of the Bluetooth worms epidemic are as follows:

$$\begin{aligned} (\lambda \langle k \rangle)_{Bluetooth} &= 1.37 \\ N &= 10^7 \\ i_0 &= \frac{1}{10^7} \end{aligned}$$

Dynamics of the number of infected mobile devices is shown in Figure 1.

It can be seen from the Figure 1 that in the absence of factors that contain the epidemic, the Bluetooth worm needs a little over 15 days to infect all of the mobile phones located in Moscow, Russia.

However, even though the features of Bluetooth worms have allowed us to apply the classical model of epidemic with parameters of Respiratory Viral Diseases for their modeling, they are the ones that complicate research of the realistic picture of an epidemic spreading. They complicate it because the applied *SI* model only shows the temporal dynamics of the epidemic but does not take into account for the spreading of the worm in space.

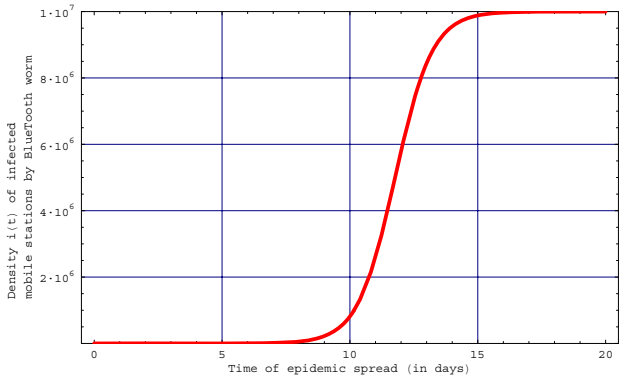

Figure 1. Dynamics of the number of infected mobile devices by the Bluetooth worm

In the case of Bluetooth worms, the spatial dynamics is important due to the limited infection range. Infection follows the travel of the infected phone, and that is why the epidemic will develop in pockets, starting and expanding in various spots of the country and the world as some infected phones move for example by air transport.

3 MMS Worms

Epidemiological *SI* model showed sufficiently good correspondence with real results of spreading of network worms, such as Code Red I v2, SQL Slammer/Sapphire, Witty [6, 5, 8]. Unlike Bluetooth worms that have a limited infection range, MMS worms completely correspond to their computer counterparts due to the following features:

1. The range of infection by an MMS worm is not limited, and infection may reach any mobile device capable of receiving MMS messages.

2. Thus the infection is not carried out by a victim who randomly enters the range, but selectively through a prepared list, by a generated or located mobile phone number.

3.1 Random Scanning MMS Worms

As in the last case, let us use the *SI* model to analyze the dynamics of a possible epidemic of *random scanning* MMS worms, i.e. infecting the mobile phone that has a randomly generated number.

Let the infected mobile phone be scanning the next victim (generate a random number and send it an MMS message) once every 5 seconds. This means that the number of contacts of the infected phone per second is $\langle k \rangle = 0.2$. Let us take the possible size of the vulnerable population to be equal to the product of the number of mobile phones in Moscow, Russia $9 \cdot 10^6$, the portion of these phones that support MMS 0.5, and the possible prevalence of vulnerability, due to which the MMS worm can infect a phone without the owner's knowledge $0.17 - 0.92$. In order to estimate the last factor, we used results from the analysis of prevalence of smartphones operating systems, which was recently conducted by Kaspersky Lab [3]. Since the random scanning version of MMS worms generates the mobile phone number randomly, the total number of possible numbers comprises the product of the number of seven-digit phone numbers and the number of codes of various communication carriers $10 \cdot 10^7 = 10^8$. If, as in the previous case, infection starts from one phone, then parameters of the model will look like this:

$$\begin{aligned} \langle k \rangle_{MMS} &= 0.2 \\ N = 0.17 \cdot 0.5 \cdot 9 \cdot 10^6 &= 7.65 \cdot 10^5 \\ \lambda_{MMS} = 7.65 \cdot 10^5 \cdot 0.2 \;/\; 10^8 & \\ i_0 = \frac{1}{7.65 \cdot 10^5} & \end{aligned}$$

Dynamics of the number of infected mobile devices by the random scanning version of the MMS worm is shown in Figure 2. The Figure shows that the random scanning MMS worms spread significantly faster than Bluetooth worms. Less than 4 hours may be necessary to infect the entire vulnerable population of more than $750,000$ mobile devices.

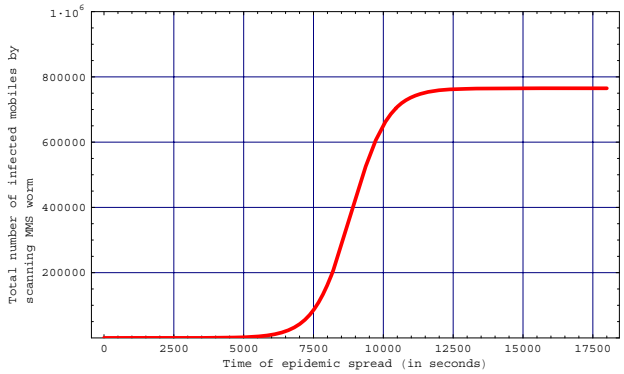

Figure 2. Dynamics of the number of infected mobile devices by the random scanning version of the MMS worm

3.2 Infection During a Limited Time (Model SIR)

The epidemic model of the *SI* epidemic that we have used up until now assumes an unlimited period of infectiousness. However, immunity is produced during infection with the virus, which may be retained after recovery either temporarily, as in the human influenza virus, or for life, as in "children's" viral diseases. One may also die as a result of the disease, which also means the end of infection. Events of recovery with lifelong immunity and death are taken into account in the *Susceptible→Infected→Recovered (SIR)* epidemic model [4, 1].

Mathematically this means that the *SIR* model extends the *SI* model by adding the flow of events of recovery (or death) with constant rate μ, and in the equation of the evolution of the number of infected individuals a decaying member $-\mu i(t)$ is added. The *SIR* model is recorded with the following system of differential equations:

$$\frac{ds(t)}{dt} = -\lambda \langle k \rangle i(t) s(t) \tag{3}$$

$$\frac{di(t)}{dt} = \lambda \langle k \rangle i(t) s(t) - \mu i(t) \tag{4}$$

$$\frac{dr(t)}{dt} = \mu i(t) \tag{5}$$

A new state is added in the normalization equation, which may be used instead of any systems equations. This is the state in the moment of time t with density of $r(t)$ individuals with expired period of infectiousness (they have either acquired immunity or died):

$$s(t) + i(t) + r(t) = 1 \tag{6}$$

The solution describing dynamics of the number of infected individuals relative to the *SI* model is as follows:

$$i_{SIR}(t) = i_{SI}(t)\, e^{-\mu t} \tag{7}$$

Thus, after some time the epidemic begins to decay exponentially because of the decrease in the number of infectious individuals.

3.3 Address-Book Scanning MMS Worms

What makes the present model interesting from the standpoint of network or mobile worm epidemics, or more precisely when does a computer or mobile phone stops infecting others? The loss of infectiousness may occur as a result of the following events:

- Disconnection from the network for some reason, due to a large volume of transmitted data in the network caused by the epidemic, which was observed during the epidemic of the SQL Slammer worm.

- Immunization of the computer or mobile phone: for example, installation of an antivirus containing the signature of the given worm.

- The end of infection by the worm itself. For example, a number of worms attempt to infect only a limited number of computers. Another example is *e-mail worms* that send copies only to a limited number of victims whose addresses were found by the worm on the infected computer.

Existing MMS worms, such as Worm.SymbOS.Comwar, also do not infect for an unlimited period of time, but send their copies only to contacts found in the address book. The above estimates are only valid for random scanning MMS worms. Let us use the *SIR* model to estimate the time of the epidemic of MMS worms that use address-book scanning for infection.

Before determining parameters of the model, let us consider the empirical estimate of the speed of infection. It may seem that the speed of spreading remains the same, and the picture will differ only by infection of a portion of the vulnerable population, and not the entire population, as well as further decay of the epidemic. However, this is not the case due to the following reason. Random scanning MMS worms randomly generate mobile phone numbers, but of all possible 10^8 telephone numbers only $9 \cdot 10^6$ really exist. As a result, the majority of attempts at infecting are made in vain. During infection by address-book scanning, however, it can be assumed that all contacts exist, and the attempt at infecting may be wasted only due to other causes - the phone is not vulnerable, etc. Taking this fact into account, $\lambda_{MMS} = 0.17 \cdot 0.5 \cdot 0.2$. The parameter of the epidemic decay μ is calculated from the period of infectiousness of the mobile phone, and the phone is infectious until the worm is done sending its copy to all of the address-book contacts. If the average number of address-book contacts is assumed to be 100 contacts in mobile phone, then the average period of infectiousness will comprise 500 seconds, and thus $\mu_{MMS} = 0.002$. As a result, the model of the *SIR* epidemic of non-random-scanning version of the MMS worm has the following parameters:

$$\begin{aligned} \langle k \rangle_{MMS} &= 0.2 \\ N = 0.17 \cdot 0.5 \cdot 9 \cdot 10^6 &= 7.65 \cdot 10^5 \\ \lambda_{MMS} &= 0.17 \cdot 0.5 \cdot 0.2 \\ \mu_{MMS} &= 0.002 \\ i_0 &= \frac{1}{7.65 \cdot 10^5} \end{aligned}$$

Dynamics of the epidemic of the MMS worm is presented in Figure 3.

As seen from the Figure 3, the maximum number of infected mobile devices is greater than 100 thousand, which comprises only about 1% of the population of Moscow, Russia, but it is reached during a very short period of time - 15 minutes, after which the epidemic begins to decay and will end in less than an hour.

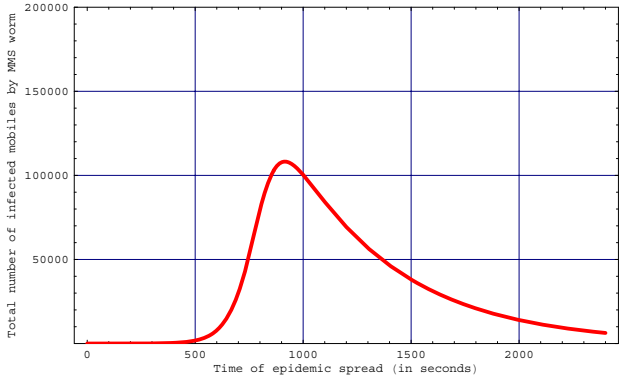

Figure 3. Dynamics of the number of infected mobile devices by the MMS worm

4 Acknowledgement

The author would like to thank Yury Mashevsky from Kaspersky Lab for interesting discussions about this work.

References

[1] R. M. Anderson and R. M. May. Population biology of infectious diseases: Part 1. *Nature*, (208):361–367, 1979.

[2] R. M. Anderson and R. M. May. *Infectious diseases of humans. Dynamics and control*. Oxford University Press, Tran. from Eng. by Romanyukha A. A., Ed. by Marchuk G. I., 1991.

[3] A. Gostev. Wardriving in China. *Kaspersky Lab analytics*.

[4] W. O. Kermack and A. J. McKendrick. A contribution to the mathematical theory of epidemics. *Proc. Roy. Soc. Lond.*, A(115):700–721, May 1927.

[5] D. Moore, V. Paxson, S. Savage, C. Shannon, S. Staniford, and N. Weaver. Inside the Slammer Worm. *IEEE Security & Privacy*, 1(4), July/August 2003.

[6] D. Moore, C. Shannon, and J. Brown. Code-Red: a case study on the spread and victims of an Internet worm. *2nd ACM Internet Measurement Workshop 2002*, pages 273–284, November 2002.

[7] J. D. Murray. *Mathematical Biology: I. An Introduction*. Springer-Verlag, 2003.

[8] C. Shannon and D. Moore. The Spread of the Witty Worm. *IEEE Security & Privacy*, 2(4), July/August 2004.