An Epidemiological View of Worms and Viruses

Thomas M. Chen
Dept. of Electrical Engineering
Southern Methodist University
PO Box 750338
Dallas, TX 75275-0338 USA
Tel: 214-768-8541
Fax: 214-768-3573
Email: tchen@engr.smu.edu
Web: www.engr.smu.edu/~tchen

1. Introduction

The communal nature of the Internet exposes organizations and home computer users to a multitude of worms, viruses, and other malicious software (malware) threats such as spyware and Trojan horses. Viruses are program fragments attached to normal programs or files that hijack the execution control of the host program to reproduce copies of the virus. Worms are automated self-replicating programs that seek out and copy themselves to vulnerable new targets over the Internet. In the same way that germs are quickly shared among people, worms can spread rapidly among networked computers. In the second half of 2004, Symantec reported 7,360 new Windows worms and viruses, an increase of 63 percent over the number of new worms and viruses in the first half of 2004 [1]. The most prevalent worms were variants of Netsky, MyDoom, Beagle, and Sober. In the 2005 CSI/FBI Computer Crime and Security Survey, 75 percent of the surveyed organizations reported being hit by worm and virus attacks [2]. Worms and viruses were the most frequent and costly type of attack, despite the use of antivirus software and firewalls by 96 percent of the surveyed organizations.

Biologists tackle infectious diseases at both microscopic and macroscopic levels. However, very little effort is spent to treat worms and viruses at the macroscopic or epidemiological level. Today the security industry focuses on the treatment of worms and viruses exclusively at the "microscopic" level, analogous to the microbiological approach to infectious diseases. Antivirus companies collect samples of worms and viruses through donations and honeypots. The malicious code is disassembled into a more human readable format to study its programmatic structure and develop a new antivirus signature. The new signatures are downloaded to update antivirus software programs.

Epidemiology is more interested in the dynamics of diseases spreading through populations than their biochemical mechanism. In the long history of medicine, epidemiology has been a relatively recent development. The foundations of epidemiology are often traced to Dr. John Snow who studied an outbreak of cholera in London in 1848 [3]. In treating patients, he became convinced that the disease was spread by ingesting germs from polluted water. At the time, many physicians did not believe in germs as the cause of infectious diseases. To avoid controversy, Snow described the cause of cholera as a "poison" that had the ability to "multiply itself" within cholera victims, before being spread to new victims through polluted water. He came across a district supplied by two private water companies. Snow collected a vast amount of statistical evidence that linked a high mortality rate to people supplied by one of the water companies, and a much lower mortality rate to the other water company. Unfortunately, Snow

was the first person to make use of a survey of the statistical incidence and distribution of an epidemic in an effort to determine its cause, and his evidence was not believed by other doctors.

In 1853, another outbreak of cholera occurred in a neighborhood close to Snow's home in the London district of Soho. He traced the water supplied to cholera victims to a water pump on Broad Street. Snow was able to convince the Board of Guardians to turn off the pump, and the local cholera outbreak quickly ended. When Snow died in 1858, his theory about the spread of cholera still had not been accepted. The germ theory of disease did not gain acceptance until the 1860s after it was demonstrated by the chemist Louis Pasteur. In historical perspective, Snow's important contribution was his persistent efforts to determine how cholera was spread by means of statistical and mapping methods which have become standard methods in epidemiology.

2. Successes of Epidemiology

The practical usefulness of epidemiology was demonstrated by the successful eradication of smallpox. Smallpox is an acute contagious disease caused by the variola virus. It is believed to have originated over 3,000 years ago in India or Egypt. For centuries, devastating epidemics have swept across continents, decimating populations. In the absence of vaccination, humans are universally susceptible to infection. No effective treatment has ever been developed, and the mortality rate is about 30 percent. Survivors are often left with scars or blindness.

The mathematician Daniel Bernoulli made a major contribution to epidemiology by mathematically proving that variolation (inoculation with a live virus obtained from a victim with a mild case of smallpox) was beneficial. Variolation usually resulted in immunity from smallpox. Bernoulli was able to formulate differential equations to show that variolation could reduce the death rate [4].

In 1798, Edward Jenner demonstrated inoculation with cowpox. The smallpox vaccine contains live vaccinia virus, which is closely related to the variola virus. Vaccine administered up to 4 days after exposure to the virus, and before the rash appears, provides protective immunity that can prevent infection for many years or at least reduce the severity of an infection.

In the 1950s, there were an estimated 50 million cases of smallpox in the world each year. Smallpox vaccination became part of the mission of the Center for Disease Control and Prevention (CDC), originally established in 1946 as the Communicable Disease Center led by Dr. Joseph Mountin within the U.S. Department of Health and Human Services [5]. Its broad mission is to monitor the prevalence of infectious diseases, develop public health policies, enact strategies for disease prevention, and investigate problems of public health. Dr. Mountin envisioned the CDC as a center for epidemiology responsible for all infectious diseases. Dr. Alexander Langmuir joined when the Korean War in 1950 posed the threat of biological warfare, to establish the CDC's Epidemic Intelligence Service (EIS). Medical epidemiologists were rare at the time, and the EIS was instrumental in training epidemiologists. In the 1950s, the CDC was instrumental in overseeing the polio inoculation program and developing a national vaccination program for a major influenza epidemic in 1957.

The CDC established a smallpox surveillance unit in 1962. It worked to refine a smallpox vaccination and introduce the vaccine to millions of people in Central and West Africa. The CDC established the application of scientific principles of surveillance to the problem. In 1967, the World Health Organization followed the success of the CDC and resolved to intensify their plan to eradicate smallpox. The WHO had passed an earlier resolution for global eradication of smallpox in 1959 but had not dedicated much resources. The intensified program consisted of a combination of mass smallpox vaccination campaigns and surveillance and containment of

outbreaks. Through the success of the global eradication campaign, smallpox was finally pushed back to the horn of Africa and then to a single last natural case in Somalia in 1977. The global eradication of smallpox was certified by the WHO in 1980.

3. Role of an Epidemiology Center for Worm Control

Today, no counterpart of the CDC exists for worm/virus control or prevention. Although analogies can be drawn between worm outbreaks in the Internet and disease outbreaks in the human population, there is no national-level organization responsible for coordinating and responding to worm outbreaks. Given the success of the CDC for human diseases, an argument could be made by analogy for the need for a national "center for worm control." The establishment of a national center for worm control could have several benefits to network security.

First, the prominence of a national center would elevate the worm problem to a national priority. Although the importance of infectious diseases effecting public health is obviously a national priority, the health of the Internet is not currently seen as a problem concerning the federal government. It might be argued that the Internet has evolved to the point of becoming a critical infrastructure essential for national productivity, and even national security. However, the Internet is generally viewed as a commercial enterprise, although its genesis began as a DARPA-funded research project. It is somewhat loosely administered by the ISOC (Internet Society), a professional membership society with over 100 organization and 20,000 individual members in 180 countries [6]. It includes the Internet Architecture Board (IAB) and Internet Engineering Task Force (IETF) responsible for Internet infrastructure standards. The ISOC is really a facilitator to coordinate the efforts of various stakeholders in the Internet. The Internet is really administered by the many companies and organizations that own parts of the Internet. Worms, and network security problems in general, are viewed as problems of the separately administered networks.

Second, a national center for worm control could be instrumental in developing an Internet-wide "health policy" to maintain the security and integrity of the Internet, in the same way that the CDC devises public health policies. Health policies could include standard practices for software patching, antivirus software updates, sharing worm information among companies and organization, and coordination of local responses to new worm outbreaks. Today worms are not treated as a single Internet-wide problem. Instead, individual networks are responsible for their own protection and defense. By design, the Internet is highly distributed and decentralized. Consequently, worm protection and defense is carried out in a piecemeal manner. However, worm infections of one network obviously have effects on other networks. An infected network not only increases the chances of infecting another network, but could also substantially increase the level of congestion with worm traffic. Therefore, it is not difficult to see the advantage of a national network security health policy that enforces consistency among security practices for the benefit of all networks.

Third, a national center for worm control could facilitate the collection and sharing of worm samples and information. Today, antivirus companies collect their own worm samples through donations and honeypots, and informally share samples with each in a limited way. They publish their own libraries of worm information. In addition, there are informal vendor-neutral groups such as AVIEN (Anti-Virus Information Exchange Network) for exchanging worm/virus information among security specialists [7]. However, there is no centralized repository which makes it difficult for anyone else to obtain worm samples or detailed information, without

subscribing to a proprietary service. Obviously, security researchers depend on access to real worm code, and the lack of data availability is a hindrance to further research. In addition to making worm samples available for research, a central repository could have additional benefits: (1) consistency in worm/virus names and terminology (2) pooling of information about specific worms (3) consistent and safe practices for worm code sharing.

Fourth, the idea of information sharing could be taken further to propose that a national center could provide an early warning for new worm outbreaks. Current approaches to early warning, like the approaches for information sharing, are either proprietary or grassroots. A well-known example of a proprietary approach is Symantec's DeepSight Threat Management System [8]. It collects log data from 24,000 sensors (firewalls, intrusion detection systems, honeypots, and hosts running Symantec antivirus) distributed throughout 180 countries, in addition to 2 million decoy e-mail accounts. The log data is correlated and analyzed for signs of attacks including worm outbreaks. The wide geographic coverage of the DeepSight System enables it to theoretically detect a new worm outbreak that might originate anywhere in the world. Another example is AT&T's Internet Protect Service which monitors traffic going through AT&T IP backbone routers. These backbone routers handle a considerable fraction of the total Internet traffic. The traffic data is correlated and analyzed for signs of worms, viruses, and denial of service attacks. An example of a grassroots early warning system is AVIEWS (Anti-Virus Information and Early Warning System), an outgrowth of the AVIEN information sharing network.

Fifth, a national center for worm control could coordinate real-time responses to new worm outbreaks. Due to the decentralized nature of the Internet, responses today are piecemeal and ad hoc. System administrators are generally responsible for protecting their own networks. When a new worm outbreak is discovered, they respond in a variety of ways, such as configuring firewalls, patching systems, updating antivirus programs, and taking systems off-line. Unfortunately, there is little coordination among system administrators of different networks.

Lastly, a national center for worm control could promote the scientific principles of epidemiology that have been successful for human diseases and apply them to worms. Little epidemic theory has been developed for worms. The idea of epidemiology for worms was suggested as early as 1993 but has not been pursued far [9].

4. Goals of Worm Epidemiology

How can epidemiology apply to worms and what can be learned? The so-called "simple epidemic model" fits random scanning worms fairly well [4,10]. The vulnerable hosts in the Internet are viewed as a fixed size population, all initially in a "susceptible" (vulnerable but not infected) state. A small number of infected hosts are introduced. After contact with a worm from an infected host, susceptible hosts will change state to "infected" and subsequently remain permanently in the infected state. An infected host makes contacts with susceptible hosts at a certain "infectious contact rate" that depends on the scanning rate of the worm and the likelihood that any scan will reach a susceptible (and not already infected) host.

A more complicated "general epidemic model" adds another "removed" state to factor in the possibility of worm disinfection. That is, system administrators are assumed to be removing the worm from infected hosts by patching software or running antivirus. Infected hosts may change state to "removed" and subsequently remain permanently in the removed state, immune from future re-infection. The transitions from infected to removed state occur at a certain "removal rate."

One of the obvious goals of epidemiology is to predict how far a worm outbreak can spread as a function of time. This is important knowledge because it always takes some time to detect and respond to a new worm outbreak. In the meantime, a new worm might spread without any constraint. Containment of the outbreak to a given infection level would require a response time that can be calculated.

Another goal of epidemiology is to quantify the effectiveness of immunization. Hosts can be protected against infection by keeping software patches and antivirus software up to date. In practice however, it is difficult to keep up patching and antivirus updates on all hosts in a network. Epidemiology can predict how a given level of immunization can slow down a worm outbreak.

Still another goal of epidemiology is modeling of active responses such as quarantining [11]. Quarantine of worms works in the same way as quarantine of human diseases. The idea is to prevent infected hosts from making contacts with susceptible hosts. Epidemic models can be used to evaluate different quarantine strategies by proper selection of infectious contact rates between pairs of hosts.

5. Conclusions

We have made a case arguing for the success of biological epidemiology and the need to further develop a similar body of theory for worms. A national-level center for worm control, analogous to the CDC for human diseases, could be instrumental in fostering and applying this theory.

References

[1] D. Turner, et al., "Symantec Internet security threat report: trends for July 2004 - December 2004," available at http://www.symantec.com.

[2] L. Gordon, et al., "2005 CSI/FBI Computer crime and security survey," available at http://www.goscsi.com.

[3] W. Winterton, "The Soho cholera epidemic of 1854," *History of Medicine*, vol. 8, 1980, pp. 11-20.

[4] D. Daley, J. Gani, *Epidemic Modeling: An Introduction*, Cambridge University Press, 1999.

[5] Centers for Disease Control and Prevention home page, available at http://www.cdc.org.

[6] Internet Society (ISOC) home page, available at http://www.isoc.org.

[7] AVIEN home page, available at http://www.avien.org.

[8] Symantec DeepSight Threat Management System, available at http://tms.symantec.com.

[9] J. Kephart, D. Chess, S. White, "Computers and epidemiology," *IEEE Spectrum*, vol. 30, May 1993, pp. 20-26.

[10] D. Moore, C. Shannon, J. Brown, "Code-Red: a case study on the spread and victims of an Internet worm," *ACM Internet Measurement Workshop*, Nov. 6-8, 2002, pp. 273-284.

[11] D. Moore, et al., "Internet quarantine: requirements for containing self propagating code," *IEEE Infocom 2003*, pp. 1901-1910.