The Benefits of a Notification Process in Addressing the Worsening Computer Virus Problem: Results of a Survey and a Simulation Model

Joan L. Aron[1], Michael O'Leary[2], Ronald A. Gove[3], Shiva Azadegan[4] and M. Cristina Schneider[5]

[1]Joan L. Aron, Ph.D., Science Communication Studies, 5457 Marsh Hawk Way, Columbia, MD 21045, USA

[2]Michael O'Leary, Ph.D., Dept. of Mathematics, Towson University, Towson, MD 21252, USA

[3]Ronald A. Gove, Ph.D., Mitretek Systems, Mail Stop F220, 3150 Fairview Park Drive South, Falls Church, VA 22042, USA

Present Address: Mitretek Systems, 7525 Colshire Drive, McLean, VA 22102, USA

[4]Shiva Azadegan, Ph.D., Dept. of Computer Science, Towson University, Towson, MD 21252, USA

[5]M. Cristina Schneider, D.V.M., Sc.D., Consultant, 4 Northrup Court, Rockville, MD 20850, USA

ELSEVIER ADVANCED TECHNOLOGY

Computers & Security
Vol. 21, No. 2, 2002, pp.142-163

Copyright ©2002 Elsevier Science Ltd
Printed in Great Britain
All rights reserved

0167-4048/02US$22.00

Abstract

Computer viruses present an increasing risk to the integrity of information systems and the functions of a modern business enterprise. Systematic study of this problem can yield better indicators of the impact of computer viruses as well as a better understanding of strategies to reduce that impact.

We conducted a Computer Virus Epidemiology Survey (CVES) on the World Wide Web to examine indicators of the impact of computer viruses. A major finding from the CVES is that multiple indicators of the impact of computer viruses reveal a problem growing more severe that affects large, as well as small, organizations. Another important finding is that viruses not detected despite regular updating of antiviral software caused only about 15% to 21% of virus problems reported in workgroups using antiviral software. The possible reasons for failure to detect include improper configuration of software and the inability of all known anti-virus detectors to detect. A related implication is that a substantial amount of damage due to viruses could probably have been prevented by regular updating of antiviral software.

We also used the CVES in the development of a simulation model for the spread of computer viruses in workgroups in order to analyze the effect of a notification process on control. Our major finding is that the process of notification, whether by human behaviour or by technology, substantially reduces the impact of computer viruses in workgroups. For example, if a workgroup has a period of vulnerability when only 80% of its workstations are effectively using antiviral software, then even a 50% probability of notification of a detected virus substantially reduces the burden. An added benefit of maintaining an environment with high effective antiviral software usage and high levels of notification is that greater rates of communication events that can potentially transmit computer viruses within the workgroup actually reduce the impact of computer viruses in the workgroup. Anecdotal observations also indicate that the process of notification is significant in controlling the spread of 'new' viruses not yet detectable by software, although the process of notification from law enforcement authorities to workgroups was not in the simulation model.

More formally, the reduced impact of computer viruses in a workgroup due to a greater rate of communication events that can potentially transmit computer viruses corresponds to a situation when a computer virus introduced into the workgroup produces, on average, less than one copy in the workgroup. This threshold corresponds to the basic reproduction ratio in epidemiology that describes the spread of infectious disease.

Keywords: computer virus; antiviral software; simulation model; survey; notification process

Introduction

Computer viruses are receiving more attention in the mass media [1]. However, systematic surveys of computer viruses and their relationship to the organizational environment are not common. Moreover, computer viruses often receive scant attention in overviews of the management of information security [2] and many organizations fail to maintain an up-to-date security management strategy. The proper assessment of computer viruses in the management of information security and integrity depends on estimates of the risk and impact of computer virus incidents and an

analysis of how they are influenced by various factors in the computing environment. Mathematical or computer simulation models of the transmission and control of computer viruses can be useful in synthesizing available information and providing a theoretical basis for control strategies.

Our major focus in the analysis of the simulation is the effect of the process of notification, which depends on interaction between computer technology and the behaviour of the end user. This focus stands in contrast to much of the marketing literature for the antiviral software industry, which is now replete with biological analogies to immunization. Research in this area has been aimed at the development of software 'immunization' tools to combat computer viruses [3] with little or no involvement by the end user. The stated aim is to have software provide protection without any involvement by the end user beyond receiving that 'flu shot'. We think that an exclusive focus on immunization tools, to the exclusion of notification and the end user, is short sighted and ignores a richer set of analogies with the efforts of the public health community to combat emerging infectious diseases [4]. Vaccines are not the only means for preventing and controlling disease, and the public health community recognizes that behaviour, environment, and host factors all play important roles in the spread of infectious disease. By analogy with this broader view of controlling the spread of infectious agents, advances in antiviral software should be coordinated with user behaviour, organizational behaviour, and the computing environment.

Total reliance on technology to combat computer viruses opens up the vulnerability of systems to errors by human or machine when the technology fails. The reasons for failure include not only inadequate technological products but also poor management practices. An example is in the improper installation, configuration and maintenance of up-to-date antiviral software. In other words, systems may fail due to problems in deployment and utilization of potentially effective antiviral technology. Even systems that are normally running properly may have windows of vulnerability during periods of change (e.g., systems are re-installed after an office move, systems are expanded and new managers are hired as company grows rapidly, etc.). Of even greater concern is system failure when the latest antiviral software, properly installed and used, is ineffective. The development of a digital immune response has also raised concerns about the dangers of "authorizing a centralized system to pull files — including those that may be misidentified as viruses — from a client's computer and then deposit unfamiliar code from a remote server onto the client's system" [5]. For instance, a recent update of McAfee antiviral software using its Superdat facility was found to damage the master boot record of Windows NT 4.0, requiring re-installation of the entire operating system [6].

Another vulnerability that is seldom recognized even in organizations with high usage of antiviral software is the virus signature update process. New viruses are constantly appearing and the vendors of antiviral software respond by providing free updates of the virus signature files. These updates can be as frequent as weekly in some circumstances. There are two problems that arise. First, if the signature update process requires action by the end user, the updates are often not promptly installed or not installed at all because of other work priorities. On the other hand, even if the updates are installed immediately, users often fail to run an immediate virus scan with the new signature. This leaves a window where a virus, undetectable by the old signature file, may remain dormant and later forwarded to another user. This is because the antiviral software will not check the file until it is opened by the user.

While the end user may, and often does, fail to update his or her software, modern software change and configuration management tools

Erratum

Elsevier Science would like apologise to the authors of the paper entitled, "The Benefits of a Notification Process in Addressing the Worsening Computer Virus Problem: Results of a Survey and a Simulation Model" which appeared in *Computers & Security*, Vol. 20, No. 8, pp. 693-714. A number of errors were introduced during the publication process for which the publisher takes full responsibility. This is the corrected version of the paper.

allow for remote updating without user intervention. That is, rather than relying on the end user to update his or her antivirus software, a centralized software management organization can 'push' updates to the end user with no intervention on the user's part. In addition, the newer antiviral software packages may be configured so that in addition to notifying the user that it has detected a virus, it can automatically notify the system administrator (via email) that the user has a virus. Our work shows that 'notification' behaviour, whether initiated automatically or by the end user, is protective in that it has a significant impact on reducing the number of virus incidents without requiring notification to proceed correctly in every instance. Of course, for the reasons stated above, we would also advise system managers not to rely exclusively on an automated system.

In developing our simulation model, we found that existing survey data on computer viruses did not adequately describe the relationships between computer viruses and the computing environment. The most publicized survey on computer viruses is the International Computer Security Association's (ICSA's) annual survey of computer viruses [7, 8]. The ICSA has demonstrated an increase in the number of computer virus incidents every year. However, the specific results of the ICSA surveys did not provide the kind of environmental and demographic data that we needed for our simulation model. We conducted a Computer Virus Epidemiology Survey (CVES) to examine indicators of the impact of computer viruses and to provide reasonable ranges for parameters in the simulation model.

Methods

Surveying the community

Characteristics of the Computer Virus Epidemiology Survey (CVES)

The CVES was accessed over the World Wide Web, where the respondents were anonymous and were not pre-selected for type or size of organization. The CVES placed great emphasis on characteristics of the workgroup and organization with sections on demographics, views, system environment, practices for sharing files, and practices for system protection. The CVES used general questions about computer virus experience in the 12 months prior to the responses. The CVES did not attempt to construct a detailed timeline of computer virus incidents, an activity fraught with problems of recall, or ask about specific brands of antiviral products. [9]

The CVES was kept online from June 1998 to September 1999. The CVES was advertised by links in major US search engines, by links on US information security websites and by email to US information security specialists in academia and business. The information security specialists were asked to notify others about the survey. Responses were numerically encoded and placed into a text data file. Most of the questions were configured to allow a respondent to select at most one response. In some questions the respondent was instructed to check all choices that applied. The data file was periodically transformed into a database compatible with STATA, a statistical software package [10].

A total of 106 respondents submitted survey forms. The respondents were obviously biased in favour of those who used the World Wide Web. That was by design because the World Wide Web is a major concern for the transmission of computer viruses. Within that, the respondents were diverse. Different kinds of organizations ran their computer networks.

Q 1.1. What is the type of organization running your computer network? If you use more than one network, pick one and limit your survey responses accordingly.

ANSWERS: Business (27);
Government (25);
Higher Education (15);

K-12 Education (7);
Other (15);
Skipped (17).

The networks and workgroups reported by the respondents varied in size.

Q 1.2. How large is your entire organization's network? Size is defined as number of enduser workstations, such as portable computers, desktop computers, etc.

ANSWERS: 1-10 (13);
11-100 (8);
101-1000 (18);
More than 1000 (45);
Not sure (6);
Skipped (16).

Q 1.3. How large is your workgroup's network? This includes the systems that you routinely interact with, such as those of your team or division; these need not be at a single geographic location.

ANSWERS: 1-10 (16);
11-100 (31);
101-1000 (22);
More than 1000 (15);
Not sure (0);
Skipped (18).

Many of the respondents were computer administrators, but some were not. Different organizational roles provide different experiences.

Q 1.4. Are you responsible for managing or administering computers other than your own computer?

ANSWERS: Yes (58);
No (30);
Skipped (18).

We retained only 80 responses for this analysis because the remaining 26 neglected to answer large portions of CVES. Dropping these 26 responses eliminated frivolous answers. No information was obtained on those who may have looked at or started CVES but did not actually 'submit' a survey form (done by clicking on a submit button at the end of CVES.) The 80 respondents who remained in the analysis did not necessarily respond to every question.

Analysis of the survey

We analyzed the data using numerous statistical tools [10, 11] and focused on three key areas — the severity of the virus incident as measured by a 'Severity Index', various organizational characteristics, and an estimate of the prevalence of viruses not detected despite regular updating of antiviral software. This group includes possibly 'undetectable' viruses that at a point in time cannot be detected by any anti-virus detector. (The initial appearance of the Melissa virus is the canonical example of this type of an undetectable virus.)

Severity Index

We analyzed the severity of computer virus impacts based on the respondents' answers to six questions in the CVES about computer virus experience in the 12 months prior to the survey. We constructed a 'severity index', scaled from 0 to 6, by incrementing a counter for any of the answers shown in italics below. Each characteristic contributes 0 or 1 to the counter.

C1)Q 3.10. Over the past 12 months, computer virus incidents resulting in at least one infected computer in your workgroup occurred:

ANSWERS: Not at all;
Infrequently (1 to 3 times during the year);
Routinely (every month);
Continuously (every week);
Don't know;

C2) Q 3.12. Over the past 12 months, the WORST computer virus infections in your workgroup were:

ANSWERS: Not a problem (no effect);
A nuisance (effect easily undone);
Minimally disruptive (management attention for a

week);
Moderately disruptive (management attention for a month);
Seriously disruptive (management attention for longer than a month);
Don't know;
Not applicable;

C3)Q 3.13. Over the past 12 months, did any virus infect 10 or more computer workstations in your workgroup?

ANSWERS: *Yes*;
No;
Don't know;

C4) Q 3.14. Over the past 12 months, did any virus problem persist for a month or more?

ANSWERS: *Yes*;
No;
Don't know;

C5) Q 3.15. Over the past 12 months, did any virus infection recur at least one month after an initial cleanup? Recurrence means that the same virus was infecting computers again. Cleanup means that viruses were removed from computers and other physical media, infected files were deleted or infected media were discarded.

ANSWERS: *Yes*;
No;
Don't know;

C6) Q 3.16. Over the past 12 months, what were the effects of computer virus incidents in your workgroup? Check all that apply to the best of your knowledge.

ANSWERS: *Total workgroup disruption*;
Complete loss of one or more workstations' hard drive volumes (requiring complete reinstallation);
Unrecoverable file or data damage;
Recoverable file or data damage.
Interference with work activities, but no loss of data or files;
No damage;
Don't know;
Not applicable;

Organizational characteristics

Organizational characteristics, such as workgroup size, were assigned two categories, each of which was evaluated for presence or absence of a particular computer virus experience among respondents in that category. The counts of the number of respondents in each combination formed a 2x2 table. Fisher's exact test for a 2x2 table [10, 11] was used to test the null hypothesis that there is no association, i.e., the computer virus experience does not vary according to an organizational characteristic. This test produces a number called the p-value, which is the probability that the observed differences could appear by chance under the null hypothesis. A smaller p-value indicates greater significance, i.e., that an apparent association is less likely to be due to chance alone. Associations were selected on the basis of the p-value of Fisher's exact test being less than 5%. We also examined the magnitude of the odds ratio, a concept in epidemiology that is similar to a risk ratio [12]. For example, the use of frequent email with attachments might double, triple or quadruple the risk that an organization experiences any computer virus incidents in a year. Associations were also selected on the basis of an odds ratio that is greater than or equal to 4.

Viruses not detected despite regular updating of antiviral software

We used two ways to approach the problem of estimating the prevalence of viruses not detected despite regular updating of antiviral software. The possible reasons for failure to detect include improper configuration of software and the inability of all known anti-virus software to detect. The first approach considered the

organizations whose respondents reported that every workstation in the workgroup was running antiviral software. This corresponds to the answer shown in italics below.

Q 6.8. How many computer workstations in your workgroup are running anti-virus software?

ANSWERS: None (0%);
Few (around 20%);
Some (around 50%);
Many (around 80%);
All (95% or 100%);
Don't know.

Respondents were selected for further analysis if they answered a question that asked about the factors responsible if the workgroup experienced problems during the past 12 months. This corresponds to the answer shown in italics below.

Q 3.17. If your workgroup experienced virus problems over the past 12 months, what factors do you consider responsible? Check all that apply to the best of your knowledge.

ANSWERS: Contact (files, diskettes, CD-ROMs) external to the workgroup;
Installation of infected computers;
Antiviral detection software ineffective DESPITE regular updates;
Antiviral detection software ineffective BECAUSE of lack of regular updates;
Potentially effective antiviral detection software installed but not used properly;
Antiviral detection software not installed;
Poor recovery procedures;
Poor awareness and monitoring procedures;
Don't know;
Not applicable.

Only respondents reporting the factor that antiviral software was ineffective despite regular updates were considered to have had problems due to viruses not detected despite regular updates of antiviral software. The second approach considered organizations whose respondents reported that antiviral software was updated in the workgroup once a month or once a week. This corresponds to the answers shown in italics below.

Q 6.16. How often is anti-virus software updated in your workgroup?

ANSWERS: Never;
Several times a year;
Once a month;
Once a week;
Don't know.

Respondents were selected for further analysis on the basis of Q 3.17 and analyzed as described above.

Simulation

Description of the simulation

The results of the CVES were used to establish ranges for the simulation parameters of a stochastic simulation model constructed in MODSIM, an object-oriented simulation development environment. The model represented a typical computer based workgroup consisting of 200 computers. (The number of computers in the workgroup is an input parameter for the model. We selected 200 as a representative value that would keep the simulation run time manageable.) The main simulation parameters are shown in Table 1. The simulation is described in more detail elsewhere [13]. The major steps in the simulation are shown in the Appendix.

Selection of parameters

First, we selected base values for all of the parameters based on the survey and other sources. The frequencies of types of computer viruses "in the wild" are taken from the 'WildList' in August 1998 [14, 15]. We then selected the 11 parameters (or parameter clusters) that had the

most significant effect on the transmission of computer viruses in the simulation:

- AV Use;
- Probability of Email Use;
- Probability of Network Use;
- Probability of Floppy Use;
- Probability of Sharing Use;
- Notification Probabilities;
- Cleanup Probabilities;
- Detection Probabilities;
- Exposure Probabilities;
- Re-infection Probabilities (Lingering);
- Scrub Threshold.

See Appendix for additional background information. We selected low, base and high values for each of these parameters. We ran a sequence of simulations, two for each parameter, which had that parameter either at a high value or at a low value, while keeping all of the other parameters at their base values. Based on these results, we chose to study in detail:

- AV Use;
- Communication Rate;
- Exposure Rate;
- Notification Rate.

Key Parameters Overview

We focus especially on three groups of parameters:

(1) Parameters for a changing computing environment (Exposure, Comm);
(2) Parameters for managerial efforts (AV, Notify);
(3) Parameters for unchanging characteristics in the background (Recognize, Cleanup).

The main concern about a changing computing environment is that increasing rates for exposure and communication will increase levels of risk due to the spread of computer viruses. The central focus is the effects of the notification process and the use of antiviral software. In order to aid in understanding these effects, each simulation run holds the parameters fixed. These parameters are described in detail below.

Table 1: Main Parameters of the Stochastic Simulation

Variable	Settings
Run Length	365 Days
Number of virus types	20
Number of computers	200
Probability that computer has active antiviral software (AV)	0.80, 0.95
Average # of communication events per network, per day (Comm)	100, 200, 400, 700, 1000
Probability a communication is email	0.75
Probability a communication is a network connection	0.20
Probability a communication is via a floppy disk	0.05
Probability a transferred file is a Word, Excel, Access, or Executable	0.70, 0.10, 0.01, 0.05 resp.
Probability that a virus recipient notifies sender and administrator (Notify)	0.10, 0.25, 0.50, 0.75, 0.90
Probability that a user who is told about a virus will remove it (Cleanup)	0.85
Probability per workstation per day that a user without AV software will recognize a virus infection (Recognize)	0.05
Probability of outside exposure per workstation per day (Exposure)	001, .005, .02
In the wild frequency Word Macro	0.76
In the wild frequency Excel Macro	0.05
In the wild frequency Generic Boot Sector	0.02
In the wild frequency Generic Executable	0.17

Computing environment

Exposure

Exposure is the probability per workstation per day of a virus exposure from outside of the workgroup. In the simulation, computer viruses arrive from outside the workgroup at one of three possible values for Exposure. Computer viruses are initially assumed to be detectable by properly installed antiviral software. Once the model determines (stochastically) that a computer or workstation is exposed from the outside, it selects the type of virus based on the frequencies reported 'in the wild' [14, 15]. The number of exposures is much higher than the number of infections because effective antiviral software prevents exposures from developing into infections. Once a computer is infected from an exposure originating outside of the workgroup, the virus may spread throughout the workgroup by means of the various communications that take place.

Comm

Comm is the mean daily number of communication events within the workgroup that have the potential to transmit computer viruses. The simulations are run sequentially with one of five values for Comm. Each communication event that can transmit computer viruses takes one of three forms:

- Email
- Network connection (e.g., Network Neighbourhood in Windows, or File Transfer Protocol)
- Floppy disk (a.k.a. 'sneaker net')

The reason for the distinctions is that different types of computer viruses are transmitted by different methods of communication. For example, a boot sector virus is carried by a floppy disk but not sent in an email attachment. If a target workstation, which is selected at random, receives a computer virus through the workgroup, then infection will occur only if the workstation is without effective antiviral software. Antiviral software might be ineffective if it is installed but not used properly.

Managerial efforts

Within a changing computing environment, managerial efforts to combat detectable computer viruses aim for better technology and better use of technology. In the simulation, the parameters for managerial efforts are:

AV

AV is the probability that a workstation in the workgroup has effective antiviral software. In the simulation, the level of antiviral coverage, AV, has one of two values. The higher value (95%) corresponds to an organization that has very good security; the lower value (80%) is considered to be a temporary window of vulnerability for the organization. We selected these values since it is well accepted that organizations should use effective antiviral software. At lower levels of coverage, infection is rampant and masks any effect of the other parameters. Improvement in the use of technology corresponds to enhanced user awareness directed at reporting problems.

Notify

Notify is the probability that a message recipient, who has detected the receipt of a virus from within the workgroup, will notify the sender of the virus and notify the system administrator of this event. In the simulation runs, Notify takes one of five values. For detectable computer viruses, an increase in notification is the central focus of strategies to involve users directly in enhancing the response to system failure. The notification represented by Notify may be initiated by the user or initiated automatically through software change and configuration management tools, scripting, or automatic email generation by the antiviral software itself.

Fixed background

We also mention other characteristics of user behaviour that affect the transmission of

computer viruses, but these characteristics, while used in the simulation, are not the target of managerial efforts. User characteristics that can be set by the user in the simulation but do not vary in the simulations are:

Recognize

Recognize is the probability per workstation per day that a user without effective antiviral software notices the presence of a virus. This parameter is set at a very small value. Although speeding up recognition of viruses would help to reduce their spread, it is difficult to imagine a training program that would be effective for the typical user. The nature of recognition is highly dependent upon the particular behaviour of a computer virus and the user's experience.

Cleanup

Cleanup is the probability that the sender of a virus will successfully clean up his or her infected workstation upon receiving notification from a recipient of the virus within the workgroup. In the simulation runs this parameter is fixed at a moderately high level.

Analytical Approximation

Motivation

A simplified analytical mathematical model aids in understanding the dynamics of the more complex simulation model. The focus is on the parameter Comm, which is the mean daily number of communication events within the workgroup that have the potential to transmit computer viruses. In the simulation, increasing rates of communication (Comm) within a workgroup sometimes increases and sometimes decreases the impact of computer viruses within the workgroup. On the one hand, greater communication can be a risk factor by enhancing the spread of computer viruses. On the other hand, greater communication can be a protective factor by increasing the number of notifications that can halt transmission. It is difficult to tell whether the protective or risk aspect will dominate in any given situation. An analytical model helps to understand this problem by giving an approximation of the simulation dynamics in the region where a greater communication rate switches from being a risk factor to being a protective factor.

Consider the repeated exposure of a workgroup to detectable computer viruses for which only a proportion of the workstations have effective antiviral software. Only the workstations without effective antiviral software will ever become infected. The workstations with effective antiviral software will never become infected and, moreover, they will contribute to the detection and notification of computer viruses sent from workstations without effective antiviral software. Since transmission remains relatively contained, we make the simplifying assumption that infected workstations without effective antiviral software clear the infection before re-infection. That is, workstations are either infected or not.

Description of the analytical model

We formulate the dynamics in terms of y, the prevalence of infection among the workstations without effective antiviral software. The prevalence of infection measures the fraction of such workstations that are infected with a computer virus. The prevalence of infection increases when uninfected workstations become infected. This process occurs due to exposure to computer viruses from outside the workgroup and transmission of computer viruses through communication within the workgroup. The prevalence of infection decreases when workstations with effective antiviral software detect the infected computers, notify the sender and the administrator, and the infections are cleaned up. Less likely, the user of the workstation without effective antiviral software might detect the computer virus based on symptoms alone. The rate of change in the prevalence may then be stated as a simple differential equation that expresses the forces of increase and decrease in y that act simultaneously.

A separate analysis for prevalence of infection applies to each of four different types of computer virus: Word Macro, Excel Macro, Generic Executable, and Generic Boot Sector. These four types were circulating during the period when the Computer Virus Epidemiology Survey was in operation (June 1998 - September 1999). More recent worm-like viruses are functionally similar to Word macro viruses in an environment of an increased Comm rate. Central to the analysis is the formulation of the per-workstation contact parameter, c_V, which measures the average number of effective contacts per day for a workstation infected with a particular virus type, v. The contact parameter, c_V, is the product of the average number of communication events per workstation per day, Comm/200, and the average number of effective contacts per communication event, φ_V, which varies by type of computer virus (see Table 2). In the computer simulation model, 75% of the communication events are email with an average of 3 recipients while the other communication events (network, floppy disk) have a single recipient. Files can be transferred through any mode of communication, but the simulation assumes that 70% are Word files, 10% are Excel files, and 5% are Executable files. Boot sector viruses can only be transmitted through transfer of a floppy disk. Therefore, the average number of communication events that have the potential to transmit a computer virus, φ_V , is (where C represents a communication event):

$$\varphi_V = \sum_C \text{Prob [C] *Average \# Recipients * Prob [C transmits v]} \quad (1)$$

The corresponding values of φ_V are shown in Table 2. The fraction of exposures due to a particular file type, G_V, takes values corresponding to the frequencies 'in the wild' shown in Table 1 (Word Macro, 76%; Excel Macro 5%; Generic Executable, 17%; Generic Boot Sector, 2%).

The dynamics of the prevalence of infection, which is calculated separately for each type of computer virus, are given by the differential equation (2).

$$dy/dt = (c_V (1 - \alpha))xy - (\gamma_V)y + (\lambda G_V)x \quad (2)$$

where:

$x + y = 1$

$c_V = (\text{Comm}/200)\ \varphi_V$

$\alpha = AV$

$\gamma_V = \text{Recognize} + c_V\ \alpha\ (\text{Notify})\ (\text{Cleanup})$

$\lambda = \text{Exposure}$

Table 2: Average Number of Effective Contacts per Communication Event for Four Virus Types

Comm Type (C)	Prob [C]	# Recipients	Prob [C transmits V]			
			Word Macro	Excel Macro	Generic.exe	Boot
Email	0.75	3	0.70	0.10	0.05	0
Network	0.20	1	0.70	0.10	0.05	0
Floppy	0.05	1	0.70	0.10	0.05	1
Average Effective Contacts (φ_V)		1.75	0.25	0.125	0.05	

G_V = fraction of exposures to computer viruses due to a particular virus type

There are three terms on the right hand side of the differential equation (2). The first term on the right hand side shows the rate at which uninfected workstations become infected due to contact with infected workstations within the workgroup. The fraction infected is denoted by y and the fraction not infected is denoted by x. The use of x is a notational convenience to substitute for $(1 - y)$. The daily contact rate c_V is for a particular type of computer virus. The parameter α is the fraction of workstations with effective antiviral software; only the fraction $(1 - \alpha)$ can be infected and become a source of infection. The second term on the right hand side shows the rate at which infected workstations are cleared of infection and revert to the uninfected state. The parameter γ_V sums two components: 1) the rate Recognize at which users of workstations recognize computer viruses on their own without antiviral software or notification; 2) the rate at which users of infected workstations clean up computer viruses in response to notification by those with effective antiviral software (fraction α) who detect computer viruses and send out notification messages. The third term on the right hand side shows the rate at which uninfected workstations become infected due to exposure to computer viruses from outside the workgroup. The rate is the product of the overall exposure λ and the fraction of computer viruses due to a particular type G_V.

Equation (2) has an equilibrium for y, which is determined by setting $dy/dt = 0$. The solution for the equilibrium y is a quadratic equation:

$$R_V y^2 - [(R_V - W_V) - 1] y - W_V = 0 \quad (3)$$

where:

$$R_V = c_V (1 - \alpha) / \gamma_V$$

$$W_V = \lambda G_V / \gamma_V$$

For each type of computer virus, R_V represents the basic reproduction ratio for internal propagation within the network and W_V represents the propagation due to exposure from external sources. The basic reproduction ratio is commonly used in epidemic theory to describe the number of new infections generated directly from an infection introduced into a population. [16]

Results

Survey

Severity

Table 3 provides individual detail about the patterns of responses to CVES questions that meet the criteria for severity described in the methods section above. Forty-two responses meet at least one criterion for severity. Criteria that permit a degree of response are marked extra severe (*XX*) to show the more severe outcome. Extra severity is designated for *continuous* in C1, *seriously disruptive* in C2, and *total workgroup disruption* in C6. The multiple dimensions of the impact of computer viruses are important. Consider, for example, that criterion C2 for disruptive impacts and criterion C3 for a virus infecting 10 or more workstations are not typically reported together. Only 13 responses contain both C2 and C3. The number of responses containing C2 without C3 is 8. Most of those responses come from workgroups with more than 10 workstations (size 11 — 100 had three responses and size 101 — 1,000 had three responses), so that infection of 10 or more workstations is not limited by the small size of the workgroup. Conversely, the number of responses containing C3 without C2 is also 8, indicating that 10 or more infected workstations does not necessarily cause much disruption.

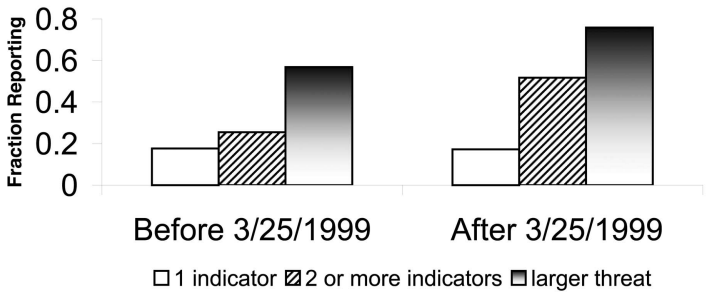
A striking finding is a reported rise in the severity of impacts of computer viruses. Figure 1 shows the proportion of respondents who reported one

Table 3: Type of Organization, Type of Virus, and Responses to Questions Related to Severity

ID	Type of organization	Type of virus	C1	C2	C3	C4	C5	C6
3	Government	Macro + Non-macro	X	X				X
5	Higher Education	Macro + Non-macro	X	X	X		X	X
7	Government	Macro	X	X				X
8	Government	Macro + Non-macro	X		X		X	
11	Business	Macro + Non-macro			X		X	
13	Other	Macro + Non-macro		X				
17	Government	Non-macro						X
19	Government	Macro					X	
21	Business	Non-macro				X		
24	Business	Macro + Non-macro					X	X
25	Business	Macro + Non-macro	X	X	X	X	X	X
27	Government	Macro + Non-macro	X	X	X			
28	Higher Education	Non-macro		X				
29	Higher Education	Non-macro	X				X	X
32	K-12 Education	Macro + Non-macro	X		X		X	
33	Government	Macro + Non-macro					X	
36	Business	Macro					X	
37	Other	Non-macro		XX		X		X
41	Government	Macro + Non-macro		X				
42	Business	Macro					X	
49	Higher Education	Macro + Non-macro	X		X			X
51	Business	Macro + Non-macro		XX	X	X	X	XX
53	K-12 Education		XX	X	X	X		
54	Government	Non-macro		X	X	X	X	
56		Non-macro	X					
57	K-12 Education	Macro + Non-macro		XX		X	X	XX
58			XX	X	X	X	X	X
59			XX	X	X	X	X	X
60	Higher Education	Non-macro		X	X	X	X	
61	K-12 Education	Macro	X	X	X	X		X
63	Business	Macro + Non-macro			X		X	XX
65	Higher Education	Macro + Non-macro				X	X	
69	Government	Macro + Non-macro	X		X	X	X	
70	Higher Education	Macro + Non-macro		XX	X	X	X	XX
71	Other						X	
72	Business	Macro		X				
73	Other	Macro + Non-macro	X	XX	X		X	XX
74	Business	Macro			X			
75	Government	Macro + Non-macro	X	XX	X		X	XX
76	Business	Macro + Non-macro			X		X	X
79	Business	Macro					X	X
80	Government							X

of the criteria of severity and the proportion of respondents who reported two or more of the criteria of severity. The cut-point of 25 March 1999, on which there were no responses, was chosen to divide the study period into two time periods in order to separate the later period that included a worldwide outbreak of the Chernobyl virus and the Melissa virus. Unfortunately, the CVES cannot link responses to specific computer viruses. The later period reveals a higher prevalence of those reporting at least two criteria of severity, a rise from 25% to 52%; Fisher's exact test of the difference in proportions is statistically significant at the 5% level. Figure 1 also shows the proportion of respondents who reported that the threat of computer viruses was larger than

Figure 1. Prevalence of one severity criterion, prevalence of two or more severity criteria, and prevalence of perception that the computer virus threat is getting worse as reported by respondents. The study period is split into the period before March 25, 1999 (51 respondents) and the period after March 25, 1999 (29 respondents). No responses occurred on March 25, 1999.



increased from 57% in the earlier study period to 76% in the later study period; the difference in proportion is statistically significant at the 10% level, but not the 5% level.

Size of organization

Another important observation is that respondents from organizations with over 1000 workstations reported disproportionately high frequencies of experience with macro virus incidents, outcomes with a single severity criterion, and outcomes with at least two severity criteria. The result is counter-intuitive since larger organizations typically put more resources into security. Not surprisingly, a high frequency of sending and receiving attached files was associated with computer virus experience.

Viruses not detected despite regular updating of antiviral software

We used two ways to approach the problem of estimating the prevalence of viruses not detected despite regular updating of antiviral software (see Methods). The possible reasons for failure to detect include improper configuration of software and the inability of all known anti-virus detectors to detect. The first method examined the 38 respondents who reported that everyone in the workgroup used antiviral software and answered a question about factors responsible for problems.

Of these 38 respondents, eight (21%) reported that antiviral software was ineffective despite regular updates. Most of these eight respondents had a positive severity index (see ID# 11, 41, 42, 51, 63, 73, 80 in Table 3). The second method examined the 20 respondents who reported that antiviral software was updated every month or every week and answered a question about factors responsible for problems. Of these 20 respondents, three (15%) reported that antiviral software was ineffective despite regular updates. Two of these three respondents had a severity index of 3 or 5 (see ID# 63, 73 in Table 3).

The number of respondents classified with severe outcomes in Table 3 indicates that many reports of severity are not linked to viruses not detected despite regular updating of antiviral software. A related implication is that a substantial amount of damage due to viruses could probably have been prevented by regular updating of antiviral software.

The simulation

Format of output

The stochastic simulation models the dynamics of detectable computer viruses in a workgroup with 200 workstations. The output of the stochastic simulation is expressed in terms of the number of infected workstations averaged over each day of the year. The stochastic variability of the results is demonstrated by running the simulation with five different sets of seeds for the random number generators. For every combination of parameters, the simulation therefore generates five values for the average daily number of infected workstations generated according to different seeds for the stochastic components. For any set of parameters, the five values are represented graphically as a symbol for the mean value with 'error bars' to the maximum and minimum value. Counts of discrete computer virus incidents are not used because the simulated workgroup experiences periods when a small number of infected computers are in the system every day.

The effect of increasing Comm

Figure 2 shows the effect of increasing communication (Comm) for the lower level of effective usage of antiviral software (AV = 80%). The general pattern is that the level of infection is relatively stable or increasing for low levels of notification (Notify = 10% or 25%), but the level of infection actually decreases with communication rates when notification is more frequent (Notify = 50%, 75% or 90%). Figure 3 shows the effect of increasing communication for the higher level of effective usage of antiviral software (AV = 95%). Under this assumption, the level of infection decreases as the communication rate increases for all levels of notification. This result argues for increasing levels of notification either through user awareness or automated tools. In either case, the enterprise should experience a decrease in the level of computer virus infection. In contrast to the non-monotonic effects of increasing the communication rate, increasing the exposure rate from .001 to .005 to .020 always increases the level of infection. The noisy effects of stochasticity appear to be greatest for the lowest exposure rate of .001, which is used in Figure 2a and Figure 3a. That is, the general patterns appear to be smoother for the higher rates of exposure.

The effect of increasing AV *and* Notify

The level of effective usage of antiviral software (AV) and the level of notification (Notify) have very strong effects. Increasing AV from 80% to 95% as shown in Figures 2 and 3 reduces the average daily number of infected computers by roughly an order of magnitude. Similarly, increasing levels of notification can reduce the average daily number of infected computers by roughly an order of magnitude. In other words, increased user awareness and participation is about as effective as significant improvements in the use of effective anti-virus software. For lower levels of effective usage of antiviral software, higher notification changes the way the system responds to greater levels of

Figure 2. Average daily number of infections for five different seeds versus communication level: AV = 80%; Comm = 100, 200, 400, 700, 1000; Notify = 10%, 25%, 50%, 75%, 90%. (a) Exposure = .001; (b) Exposure = .005; (c) Exposure = .020.

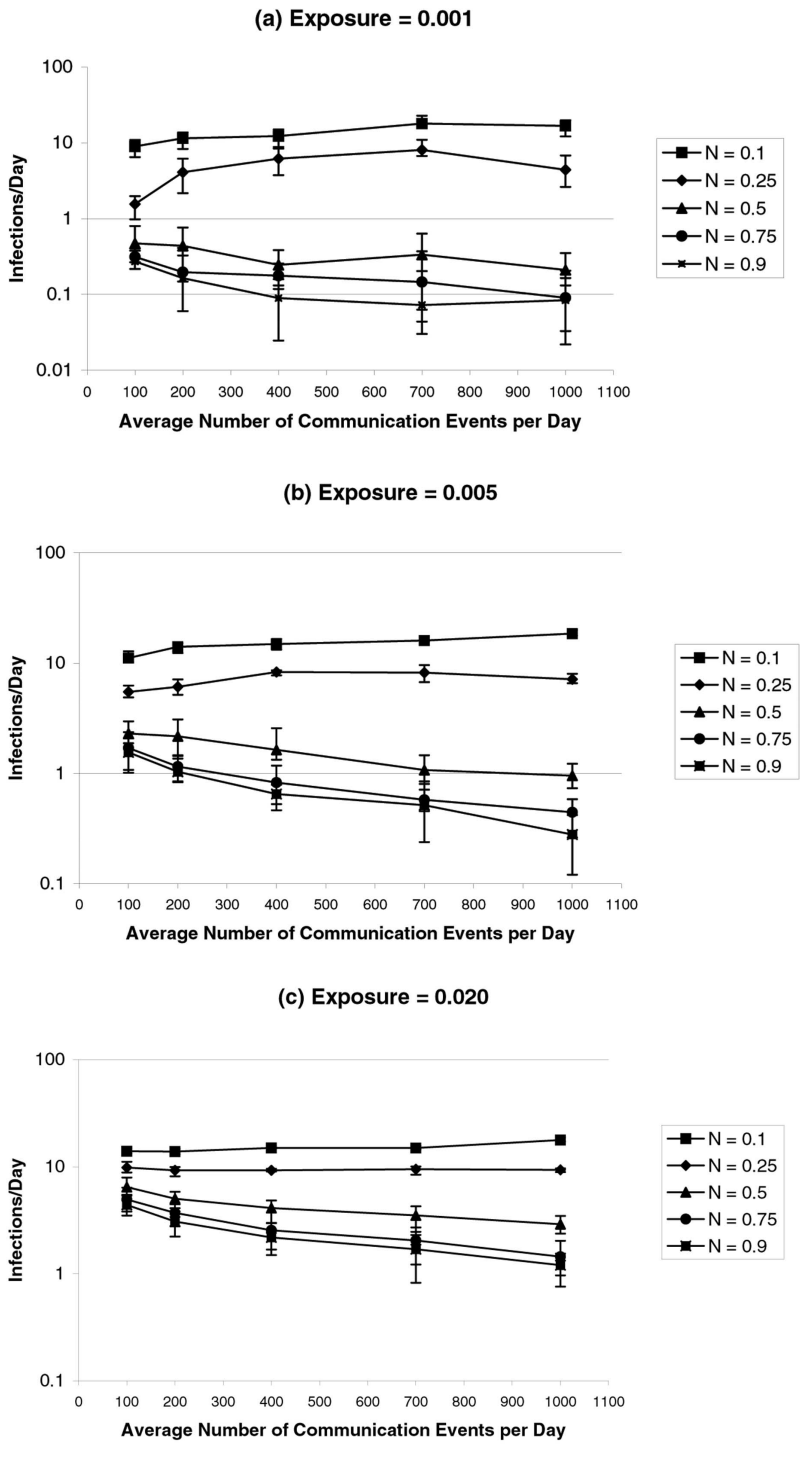
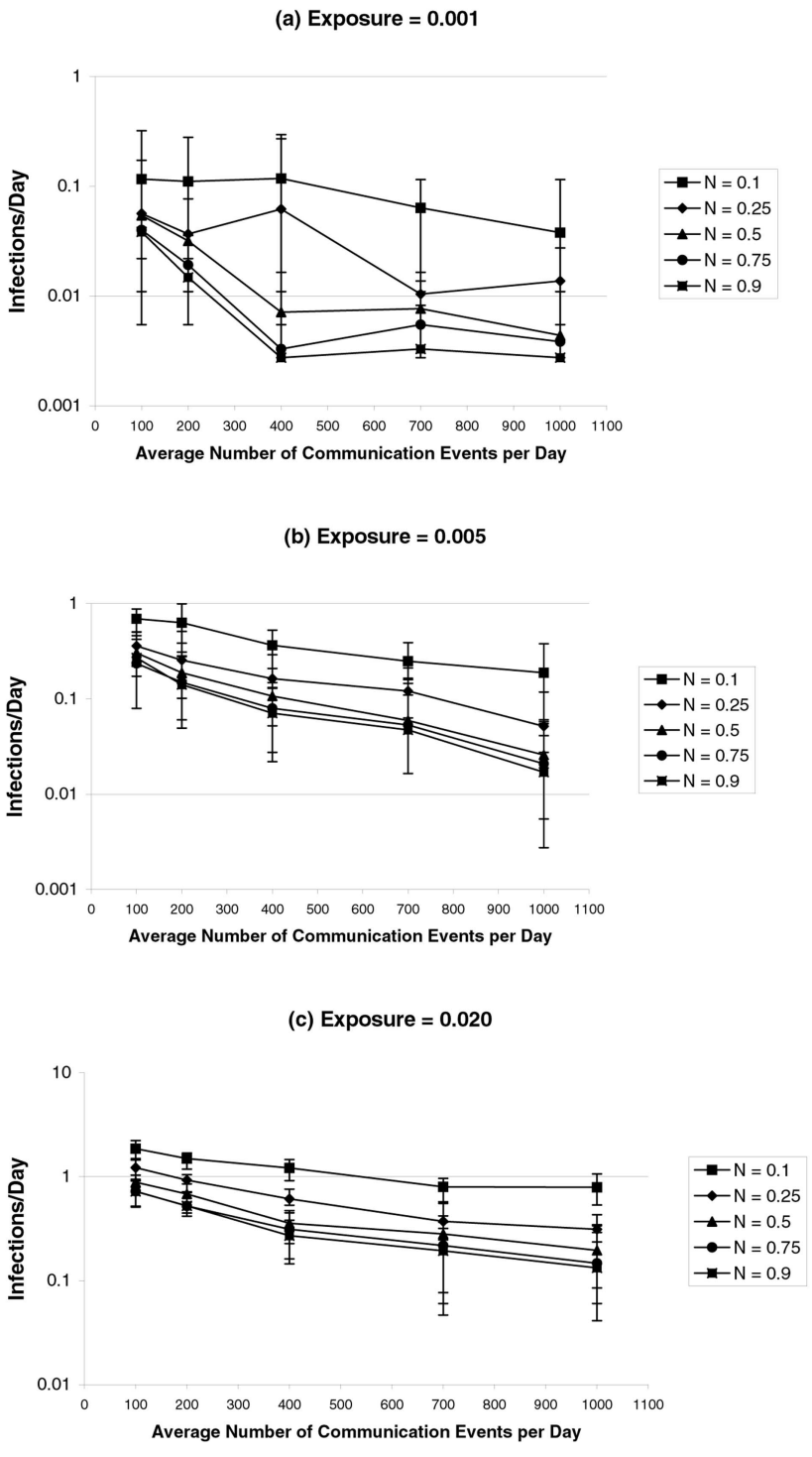


Figure 3. Average daily number of infections for five different seeds versus communication level: AV = 95%; Comm = 100, 200, 400, 700, 1000; Notify = 10%, 25%, 50%, 75%, 90%. (a) Exposure = .001; (b) Exposure = .005; (c) Exposure = .020.



communication. The effect of greater communication switches from an increase in the prevalence of infection to a decrease in the prevalence of infection.

The analytical approximation

Compare to simulation model

The analytical approximation to the simulation aids in understanding when greater communication is a risk factor increasing infection and when greater communication is a protective factor decreasing infection. Communication here refers specifically to communication events within the workgroup that have the potential to transmit computer viruses. Figures 4 and 5 compare the average daily number of infections averaged over runs from five different seeds in the simulation model (solid lines) with the equilibrium number of infections derived in the analytical approximation (dashed line). The calculation of the equilibrium number of infections uses the combined equilibrium prevalence of infection, which adds together the equilibrium prevalence for each of the four types of computer viruses. The combined equilibrium prevalence is multiplied by the number of workstations without effective antiviral software to produce the number of infected workstations for purposes of comparison with the simulation model. Although the approximation in the analytical model does not match exactly the simulation, the analytical approximation does display the general trend of how the prevalence changes with an increasing communication rate. The pattern is that, for the lower levels of antiviral software usage (AV = 80%), the value of the notification rate influences the dependence on the communication rate (Figure 4). For lower levels of notification (Notify = 10% or 25%), the prevalence of infection is fairly flat or increasing as the communication rate increases. For higher levels of notification (Notify ≥ 50%), the prevalence of infection declines as the communication rate increases. For higher levels of antiviral software usage (AV

= 95%), the prevalence of infection always decreases as the communication rate increases (Figure 5).

Basic reproduction ratio

The explanation of this behaviour lies with an understanding of the basic reproduction ratio, $\underline{R}_V$. The following calculations focus on the basic reproduction ratio for Word Macro viruses because they are the most common viruses. Figure 6 displays the calculated values of the basic reproduction ratio for Word Macro viruses for five different values of Comm and five different values of Notify. For AV = 80%, the value of $\underline{R}_V$ is near or above 1 when Notify is 10% or 25%; otherwise, the value of $\underline{R}_V$ is below 1 (Figure 6a). For AV = 95%, the value of $\underline{R}_V$ is below 1 (Figure 6b). Thus, it appears that a basic reproduction ratio below 1, which signifies that computer viruses do not propagate internally, defines a regime in which increases in communication rates are protective. Conversely, when a basic reproduction ratio starts to exceed 1, computer viruses do propagate internally and the increases in communication rates increase the risk of transmission.

Management recommendations

Our recommended strategy to control the impact of computer viruses has a technology component and a management component. The technology component focuses on the AV parameter in the simulation model. The aim is to increase the use of effective antiviral software, which is a conventional recommendation universally acknowledged in information security. The management component focuses on the Notify parameter in the simulation model, which leads to a recommendation not emphasized strongly in the research literature or in practice. Although some managers recognize that the notification process is a part of overall information security, many efforts to improve defences against computer viruses focus exclusively on the

Figure 4. Average daily number of infections averaged over runs from five different seeds (solid line) and the equilibrium number of infections derived in the analytical approximation (dashed line) versus communication level: AV = 80%; Comm = 100, 200, 400, 700, 1000; Notify = 10%, 25%, 50%, 75%, 90%. (a) Exposure = .001; (b) Exposure = .005; (c) Exposure = .020.

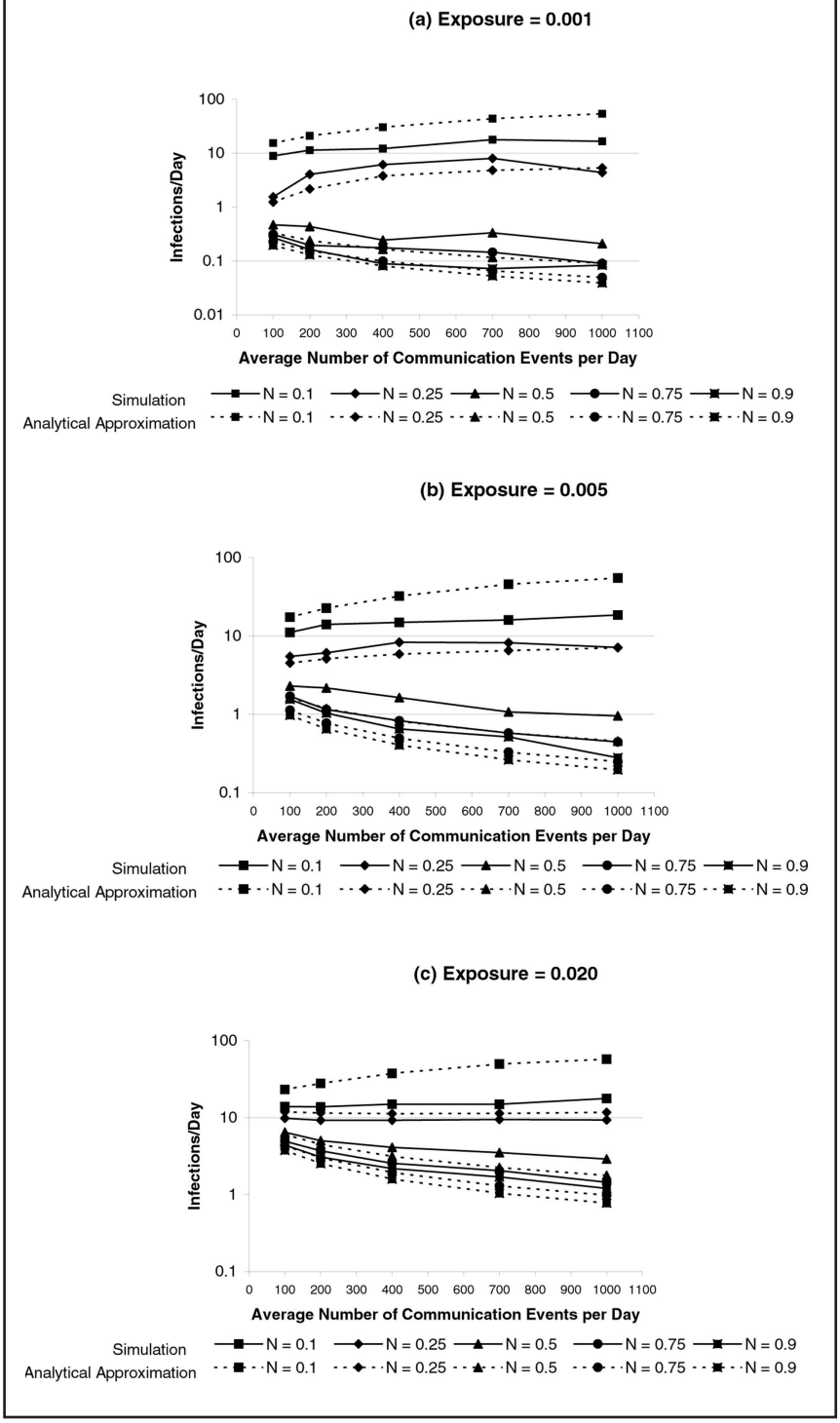
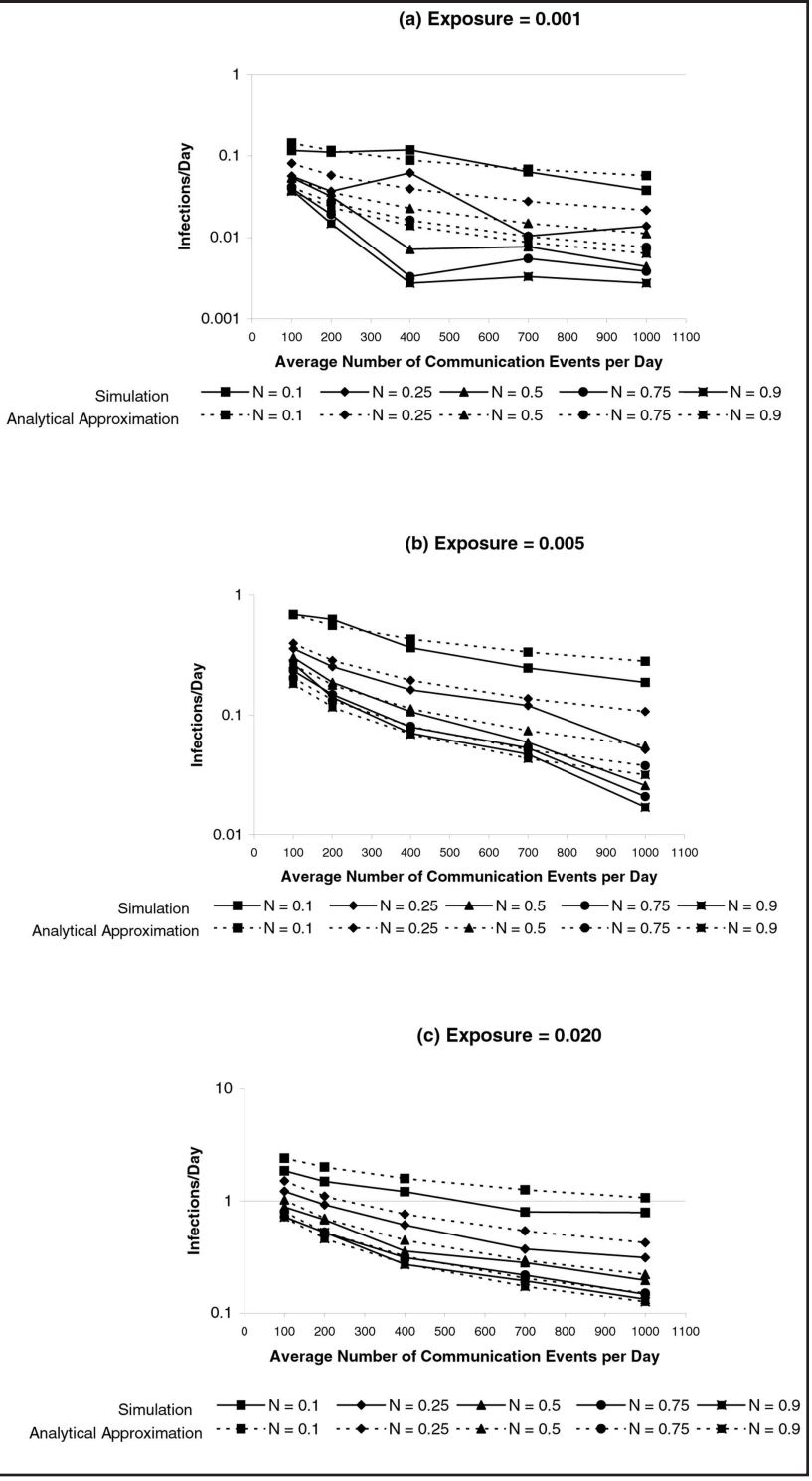


Figure 5. Average daily number of infections averaged over runs from five different seeds (solid line) and the equilibrium number prevalence of infections derived in the analytical approximation (dashed broken line) versus communication level: AV = 95%; Comm = 100, 200, 400, 700, 1000; Notify = 10%, 25%, 50%, 75%, 90%. (a) Exposure = .001; (b) Exposure = .005; (c) Exposure = .020.



technological tools to detect and clean up viruses.

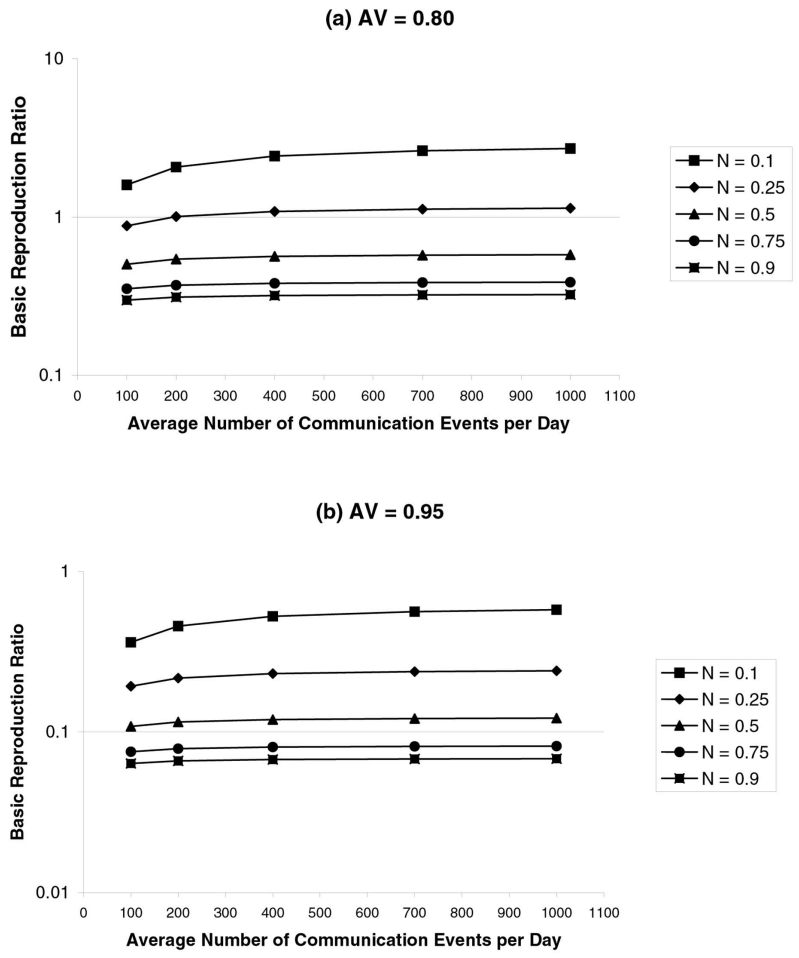
The implementation details of the AV technology recommendation involve increasing the use of effective antiviral software at the desktop, servers and gateways in order to increase the likelihood of stopping the virus at possible entry points. The effect of increasing antiviral software usage from 80% to 95% is a substantial reduction in the prevalence of infected workstations. Software change and configuration management tools, such as Microsoft® Systems Management Server, can greatly enhance the ability of an organization to maintain high levels of effective antiviral software usage.

The implementation details of the Notify management recommendation involve training users to report viruses to the system administrator when they occur. Tools to automate the notification process should be incorporated as much as possible. In simulations using 80% effective usage of antiviral software, increasing the probability of notification of virus detection from 25% to 50% profoundly alters the vulnerability of a workgroup to higher rates of communication that could transmit computer viruses. Notification of the detection of computer viruses within the workgroup can enhance security even with high levels of effective antiviral coverage, but notification has an even more important impact at lower levels of effective antiviral coverage. A strong notification process in effect provides additional protection when there are windows of vulnerability in the technology component to detect and clean up computer viruses. In addition to user training, increasing the management component Notify can be achieved through the use of software change and configuration management technology and other tools. A customer of one of the authors who has implemented software change and configuration management tools as a way to enhance notification and the effective usage of

antiviral software has seen significant cost savings, equivalent to two system administrator-years.

Beyond the simulation model itself, we draw an analogy between a suspicious user who does not permit the spread of a computer virus and the effect of antiviral software. Accordingly, the management component should also train users to be suspicious of attachments and other code whose source they do not know and trust. In order for training to be effective, it will have to become ever more sophisticated to keep pace with virus writers. A virus writer linked with the creation of the Melissa virus posted an essay that shows how to use demand, deception, and infection to spread computer viruses [17, 18]. For example, disguising viruses as patches to problems, packaging viruses in zip files, and causing instant infection upon downloading are suggestions for enhancing spread. As with conventional detection methods, the notification process should be strengthened. Training should establish emergency procedures and alert users to respond promptly to instructions from authorities about new viruses. Michael Vatis, former head of the US Federal Bureau of Investigation's National Infrastructure Protection Center, testified on 15 April 1999 at a hearing of the US House of Representatives science subcommittee that damage from the Melissa virus was "significantly contained" thanks to warnings from law enforcers and the media that spread faster than the virus itself. [19]

Figure 6. Basic reproduction ratio of Word Macro viruses in the analytical approximation versus communication level: Comm = 100, 200, 400, 700, 1000; Notify = 10%, 25%, 50%, 75%, 90%. (a) AV = 80%; (b) AV = 95%.



Discussion

Our main thesis is that training for user awareness will become ever more important for mitigating the impact of computer viruses in a world of increasing interconnectivity. For computer viruses that are detectable by antiviral software, we have shown the importance of the process of users notifying other users and system administrators when they detect a virus transmitted within their network. Notification becomes especially important during windows of vulnerability when antiviral technology may not be fully and properly deployed. For computer viruses that cannot be detected by even the most up-to-date software, economic impacts can be especially large. User awareness directed at greater suspicion about external contacts can be beneficial. We stress that benefits of user awareness do not require an impossible assumption that virtually everybody in a workgroup be extremely attuned to risk. Striking results may be obtained under realistic

objectives of enhancing awareness of half of the users.

We have shown methodologically the utility of analyzing problems of information security with the collection of survey data. However, one should be aware of limitations of self-reporting by self-selected individuals. The problems of obtaining well-defined samples in this area are intrinsic. For example, we had an absolute policy of maintaining the anonymity of any respondent. The ICSA surveys are limited, by design, to companies of a minimum size, eliminating large numbers of computer users. Within a company, the ICSA interviewers seek an information systems manager, a process that may introduce unknown biases since job titles in information security are notoriously difficult to standardize with respect to expertise. Moreover, the perspectives of multiple roles are important in different areas (computer virus software, user behaviour, impact of computer viruses on the organization). Larger studies are needed to obtain sufficiently large numbers of individuals in different roles in order to characterize the different roles involved in computer virus transmission and control.

Nevertheless, we find credible the general result about a worsening computer virus problem. The ICSA surveys, which have different strengths and weaknesses, point in the same direction. In addition, an economic analysis for US businesses in 1998 found the cost of a data loss incident for a personal computer to be $2557, which takes into account costs due to technical support, lost productivity and permanent loss of data; a similar study for European businesses found a value of $2615 [20, 21]. Moreover, many businesses do not recover from severe data losses [21].

We have also shown methodologically the value of simulations in analyzing control options. Further, relatively simple analytical approximations to a complex simulation can provide insights into the key factors that drive the behaviour in the simulation. We are building upon an earlier generation of epidemiological models of computer viruses developed in the early 1990s [22] when transmission was to a large degree limited by transfer of floppies via sneaker net [23, 24]. We also note that analogies to the spread of biological agents apply to various non-biological phenomena, such as the dissemination of information [25].

More work needs to be done to identify specific types of notification processes for viruses that can and cannot be readily detected as well as for more recent worm-like viruses such as "ILoveYou". Notification involves both training of users and technology development. Benefits of notification need to be identified and quantified in order to aid managers in making decisions about where to invest information security resources. The benefits need to address the diversity of impacts of computer viruses. Some computer viruses may be destructive primarily through the cost of lost data while other computer viruses may be disruptive primarily through their effect on using the bandwidth of communication networks. Some strategies to control computer viruses may have their own heavy costs, as in a decision to close down email systems to prevent infection. The field of information security has long recognized the difficulties in creating and maintaining user awareness. Research to integrate security technology and user awareness should re-invigorate this old issue in today's dynamic technological context.

Acknowledgments

We appreciate the Towson University Applied Mathematics Laboratory directed by Dr Martha Siegel that brought together mathematicians and computer scientists at Towson to work on this project. The student team who worked on this project were Shadi Alagheband, Michael R. Connelly, Sarah Faris

and Michael Thomas. CVES website was established due to the work of Jon McKnight in cooperation with the Center for Secure Information Systems at George Mason University and the Secure Web group at Science Applications International Corporation. We also thank Myron Cramer, Cedric Armstrong and Jim Frazer who worked on earlier phases of this project. This work was supported by the US Department of Defense.

Appendix

Main Steps in the Simulation Model

(1) Initialize all of the random number generators from a seed file;
(2) Enter input parameters of the simulation through a graphical user interface or Web interface;
(3) Use the input parameters and built-in probability distributions randomly to configure and assign virus types;
(4) Construct the network as an array of computer objects.

For each computer in the network:

(5) Determine if the antiviral software is on or off for each computer based on the input variable AV;
(6) Start with k infected computers (k is an input value and may be set to 0). Sampling a uniform distribution infects the k computers. For those infected at the start of the simulation, the AV is turned off;
(7) Configure the antiviral software for the computers and for the email server. By default, all virus types are detectable and cleanable by both kinds of software;
(8) Begin simulation process. Simulation time is managed by SIM time, which has been set to hours; so as to more accurately monitor changes in the various objects based on specific events. Each day uses 8 SIM time units to simulate the business day.

The following occurs each day:

(9) Dump the network status to a log file, stating which computers are infected with which virus and the status of the antiviral software on that computer;
(10) Introduce n new outside infections by sampling a binomial distribution. (The parameters are set as inputs);
(11) Check the VirusLinger vector. This variable determines if there is a possibility of reinfection from a previously scrubbed virus. If the computer is to be re-infected, the number of new exposures is selected from a binomial distribution. If that computer has active anti-virus software that can detect that virus, the computer will prevent the infection. All results are written to the log file;
(12) Run Communications. For each communication, do the following:
  (a) Select the method (email, network (e.g. MS Network Neighbourhood, ftp), or floppy disk) and the file type to be transferred based on uniform distributions.
  (b) Choose a sending computer using the uniform distributions. Email communications may have more than one recipient, based on a random selection. The mean value for the distribution is an input parameter.
  (c) For each recipient, see if communication is checked using a uniform random variable. If the communication variable is checked, go through process listed below. If it is not checked, then generate a wait time using exponential distribution with the average being calculated using the input parameters. The following two steps would then be executed at that later SIM time (the wait time). If that SIM time would fall during the following day, these would be executed at the beginning of the next day.

(d) Examine the infection array for the sender computer. In general, if the sender has a virus, the simulation determines if the virus can be transmitted by the given file type across the given connection. If it can, then the receiver is checked to see if it has anti-virus software. If it does, and if the software can detect that virus type, it prevents the infection, and then a uniform random variable is polled to see if the receiver notifies the sender and administration of the virus. If the receiver does not have active anti-virus software, or if the anti-virus software cannot detect the virus, the computer becomes infected. This is all recorded in the log file.

(e) If the sender is notified that it sent a virus, either by an email generated by the anti-virus software, by the receiver of the transferred file, or by a software change and configuration management tool, a uniform random variable is polled to see if it will clean its machine, based on input data. If it does, it will use the anti-virus software to remove all viruses that the anti-virus software can clean from its machine. It does not change its anti-virus active status however, so that if it were not using anti-virus software effectively as preventative before it were notified of the infection, it will not be using it as a preventative measure after the clean up. This is all recorded in the log file.

(13) Record the number of users that notice the presence of a virus on their machine. Noticing is selected from a uniform random variable and the appropriate input parameter. If the virus is noticed, the virus is removed from the computer, and the computer notifies the administration. This is all noted in the log file;

(14) Administrator counts notifications and when a threshold (input parameter) is reached the administration 'scrubs' the network. This is implemented by having each computer use its anti-virus software to remove all the viruses it can. The VirusLinger array is then set, which is used for possibility of re-infection (see above). If a 'scrub' takes place, this is noted in the log file;

(15) Write current status to the log file;

(16) Continue simulation until 365 SIM-days have passed.

References

[1] J. Schwartz and D.A. Vise. 'Love' virus is traced to Philippines: authorities move to seize computers used in attack. Washington Post, May 6, 2000, pp. A1, A11.

[2] General Accounting Office, Executive Guide. Information Security Management: Learning from Leading Organizations (GAO/AIMD—98—68). General Accounting Office, Washington, DC, 1998.

[3] J.O. Kephart, G.B. Sorkin, D.M. Chess and S.R. White, Fighting computer viruses, Scientific American 277 (1997), 88—93.

[4] Centers for Disease Control and Prevention, Preventing Emerging Infectious Diseases: A Strategy for the 21st Century, Centers for Disease Control and Prevention, Atlanta, Georgia, 1998.

[5] K. Zetter. Viruses: The Next Generation, PC World 18 (2000) (December), p. 203.

[6] E. Woodbury. McAfee Anti Virus Update Damages NT 4.0 Master Boot Record. December 18, 2000. ZDNET News http://www.zdnet.com/zdhelp/stories/main/0,5594, 2665878-1,00.html

[7] M. Kabay, P. Tippett and L.M. Bridwell, Fifth Annual ICSA Computer Virus Prevalence Survey 1999, ICSA.net, Reston,VA, 1999.

[8] L.M. Bridwell and P. Tippett, ICSA Labs 6th Annual ICSA Computer Virus Prevalence Survey 2000, ICSA.net, Reston,VA, 2000.

[9] J.L. Aron and R.A. Gove, Application of models from epidemiology to metrics for computer virus risk — a brief update, in: Integrity and Internal Control in Information Systems: Strategic Views on the Need for Control, IFIP TC11 Working Group 11.5 Third Working Conference on Integrity and Internal Control in Information Systems, November 18-19, 1999, Amsterdam, The Netherlands. (M.E. van Biene-Hershey and L.A.M. Strous, eds.), Kluwer Academic Publishers, Boston/Dordrecht/London, 2000, pp. 179-184.

[10] L.C. Hamilton. Statistics with STATA 3, Duxbury Press, Belmont, MA, 1993.

[11] VassarStats http://vassun.vassar.edu/~lowry/VassarStats.html

[12] R. Beaglehole, R. Bonita, T. Kjellstrom. Basic Epidemiology, WHO, Geneva, Switzerland, 1993.

[13] S. Alagheband, M.R. Connelly, S. Faris and M. Thomas, The Spreading of Computer Viruses and their Risk to Local Area Networks, Final Report, May 26, 1999, The Applied Mathematics Laboratory, Towson University, Towson, Maryland, 1999.

[14] Data Fellows http://www.europe.datafellows.com/vir-info/

[15] Virus Bulletin http://www.virusbtn.com/Prevalence/199904.html

[16] J.L. Aron, Mathematical modeling: the dynamics of infection, in: Infectious Disease Epidemiology: Theory and Practice, K.E. Nelson, C.M. Williams, N.M.H. Graham, eds., Aspen Publishers, Gaithersburg, Maryland, 2000, Chapter 6.

[17] J. Deane, VicodinES manifesto: 'Infect the world' , March 31, 1999, ZDNet News http://www.zdnet.com/zdnn/stories/news/0,4586,2235046,00.html

[18] VicodinES, Theory of Better File Virus Distribution, http://www.zdnet.com/zdnn/special/essay.html

[19] Nando Times, April 15, 1999

[20] D.M. Smith, The Cost of Lost Data, Storage Management Solutions 4 (1999), 60-62.

[21] D.M. Smith and C. Bowker. The Cost of Lost Data in Europe, Technical Report, Pepperdine University Graduate School of Business and Management, January 2000.

[22] Nationwide Computing Corporation, Proceedings of the Fourth Annual Computer Virus & Security Conference, New York City, March 14-15, 1991, Nationwide Computing Corporation, 1991.

[23] J.O. Kephart and S.R. White, Directed-graph epidemiological models of computer viruses, in: Proceedings of the IEEE Computer Society Symposium on Research in Security and Privacy, Oakland, CA, May 20-22, 1991, 1991, pp. 343-359.

[24] J.O. Kephart and S.R. White, Measuring and modeling computer virus prevalence, in: Proceedings of the 1993 IEEE Computer Society Symposium on Research in Security and Privacy, Oakland, CA, May 24-26, 1993, 1993, pp. 2-15.

[25] R.W. Pew and A.S. Mavor, eds., Modeling Human and Organizational Behaviour: Application to Military Simulations, National Academy Press, Washington, DC, 1998, Chapter 11.