In Proceedings of the 5th Virus Bulletin International Conference, Boston,
September 20-22, 1995, Virus Bulletin Ltd, Abingdon, England, pp. 165-181.

Computer Viruses: A Global Perspective

Steve R. White, Jeffrey O. Kephart and David M. Chess
High Integrity Computing Laboratory
IBM Thomas J. Watson Research Center
P.O. Box 704
Yorktown Heights, NY 10598

1 Introduction

Technical accounts of computer viruses usually focus on the microscopic details of individual viruses: their structure, their function, the type of host programs they infect, etc. The media tends to focus on the social implications of isolated scares. Such views of the virus problem are useful, but limited in scope.

One of the missions of IBM's High Integrity Computing Laboratory is to understand the virus problem from a global perspective, and to apply that knowledge to the development of anti-virus technology and measures. We have employed two complementary approaches: observational and theoretical virus epidemiology [1, 2, 3, 4, 5, 6]. Observation of a large sample population for six years has given us a good understanding of many aspects of virus prevalence and virus trends, while our theoretical work has bolstered this understanding by suggesting some of the mechanisms that govern the behavior that we have observed.

In this paper, we review some of the main findings of our previous work. In brief, we show that, while thousands of DOS viruses exist today, less than 10% of these have actually been seen in real virus incidents. Viruses do not tend to spread wildly. Rather, it takes months or years for a virus to become widespread, and even the most common affect only a small percentage of all computers. Theoretical models, based on biological epidemiology, can explain these major features of computer virus spread.

Then, we demonstrate some interesting trends that have become apparent recently. We examine several curious features of viral prevalence over the past few years, including remarkable peaks in virus reports, the rise of boot-sector-infecting viruses to account for almost all incidents today, and the near extinction of file-infecting viruses. We show that anti-virus software can be remarkably effective within a given organization, but that it is not responsible for the major changes in viral prevalence worldwide. Instead, our study suggests that changes in the computing environment, including changes in machine types and operating systems, are the most important effects influencing what kinds of viruses become prevalent and how their prevalence changes.

Finally, we look at current trends in operating systems and networking, and attempt to predict their effect on the nature and extent of the virus problem in the coming years.

2 The Status of the Virus Problem Today

Over the past decade, computer viruses have gone from an academic curiosity to a persistent, worldwide problem. Viruses can be written for, and spread on, virtually any computing platform. While there have been a few large-scale network-based incidents to date [7, 8, 9, 10] the more significant problem has been on microcomputers. Viruses are an ongoing, persistent, worldwide problem on every popular microcomputing platform.

In this section, we shall first review briefly our methods for monitoring several aspects of computer virus prevalence in the world. Then, we shall present a number of the most interesting observations. We will attempt to explain these observations in later sections of the paper.

2.1 Measuring Computer Virus Prevalence

We have learned much about the extent of the PC-DOS virus problem by collecting virus incident statistics from a fixed, well-monitored sample population of several hundred thousand PCs for six years. The sample population is international, but biased towards the United States. It is believed to be typical of Fortune 500 companies, except for the fact that central incident management is used to monitor and control virus incidents.

Briefly, the location and date of each virus incident is recorded, along with the number of infected PCs and diskettes and the identity of the virus. From these statistics, we obtain more than just an understanding of the virus problem within our sample population: we also can infer several aspects of the virus problem worldwide. Figure 1 illustrates how this is possible. [1]

From the perspective of one of the organizations that comprises our sample population, the world is full of computer viruses that are continually trying to penetrate the semi-permeable boundary that segregates that organization from the external world. At a rate depending on the number of computer virus infections in the world, the number of machines in the organization, and the permeability of the boundary, a computer virus will sooner or later make its way into the organization. This marks the beginning of a *virus incident*. Assuming that the permeability of the boundary remains constant, the number of virus incidents per unit time per machine within the set of organizations that makes up our sample population should be proportional to the number of computer virus infections in the world during that time period. (In fact, our measure will lag the actual figure somewhat, since incidents are not always discovered immediately.)

2.2 Observations of Computer Virus Prevalence

As shown in Figure 2, there are thousands of DOS viruses today. During the past several years, the rate at which they have appeared worldwide has crept upwards to its present value of 3–4 new viruses a day on average (see Fig. 3).

Note that the number of new viruses is not "increasing exponentially", as is often claimed [11, 3]. The rate of appearance of new viruses in the collections of anti-virus workers has been increasing gradually for several years, at roughly a linear rate. Thus the number of known viruses is growing quadratically at worst. In fact, almost nothing at all about viruses is "increasing exponentially". The problem is significant, and it is growing somewhat worse, but prophets of doom in this field have poor track records.

While there are thousands of DOS viruses, less than 10% of them have been seen in actual virus incidents within the population that we monitor. These are the viruses that actually constitute a problem for the general population of PC users. It is very important that anti-virus software detect viruses that have been observed "in the wild". The remainder are rarely seen outside of the collections of anti-virus groups like ours. Although many of them might never spread significantly, viruses that are not prevalent remain of interest to the anti-virus community. We must always be prepared for the possibility that a low-profile virus will start to become prevalent. This requires us to be familiar with all viruses, prevalent or not, and to incorporate a knowledge of as many of them

[1] Further details about our methods for collecting and interpreting statistics can be found in several references [2, 4, 5, 6].

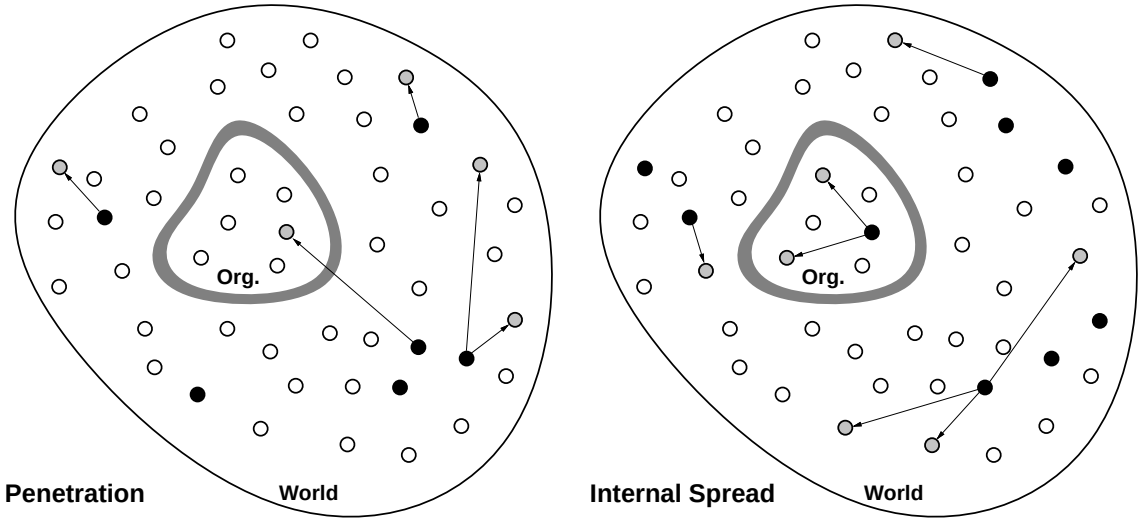

Figure 1: Computer virus spread from an organization's perspective. White circles represent uninfected machines, black circles represent infected machines, and gray circles represent machines in the process of being infected. Throughout the world, computer viruses spread among PCs, many of them being detected and eradicated eventually. **Left:** Occasionally, a virus penetrates the boundary separating the organization from the rest of the world, initiating a virus incident. **Right:** The infection has spread to other PCs within the organization. The number of PCs that will be infected by the time the incident is discovered and cleaned up is referred to as the *size* of the incident.

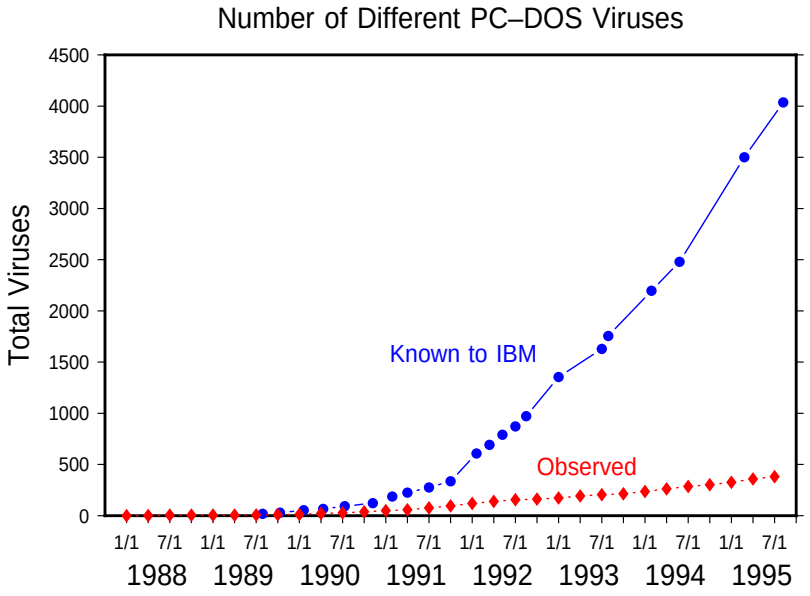

Figure 2: Cumulative number of viruses for which signatures have been obtained by IBM's High Integrity Computing Laboratory vs. time. There are thousands of viruses, but only a few have been seen in real incidents.

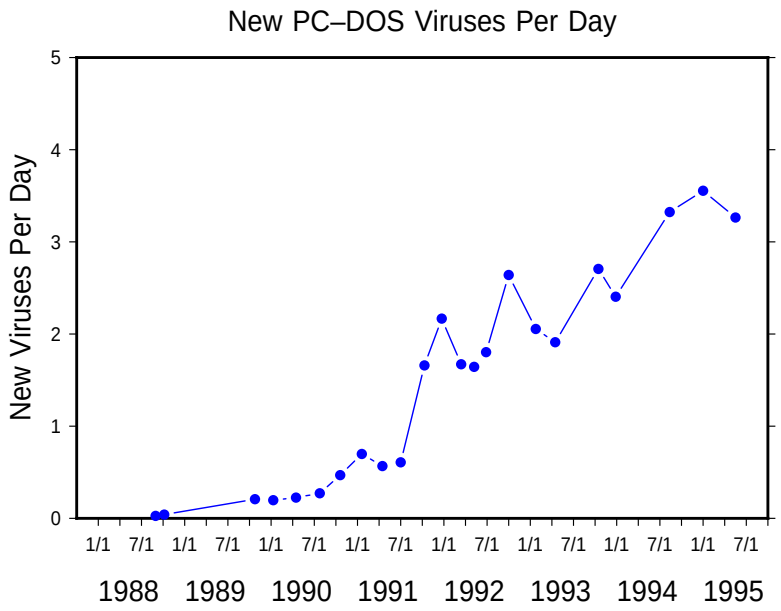

Figure 3: The number of new viruses appearing worldwide per day has been increasing steadily.

as possible into anti-virus software. We continue to monitor the prevalence of *all* viruses, regardless of how prevalent they are at present.

Out of the several hundred viruses that have ever been observed in actual incidents, a mere handful account for most of the problem. Figure 4 shows the relative fraction of incidents caused by the ten most prevalent viruses in the world in the past year. These ten account for over two thirds of all incidents. The one hundred other viruses that have been seen in incidents in the past year account for less than a third of the incidents. Most of these were seen in just a single incident.

Curiously, the ten most prevalent viruses are all boot viruses. Boot viruses infect boot sectors of diskettes and hard disks. When a system is booted from an infected diskette, its hard disk becomes infected. Typically, any non-write-protected diskette that is used in the system thereafter also becomes infected, spreading the virus. The dominance of boot viruses is especially striking when one takes into account the fact that, of the thousands of known DOS viruses, only about 10% are boot sector infectors.

Boot viruses have not always been dominant. Three years ago, the second and third most prevalent viruses were file infectors, as were 4 of the top 10. The total incident rates for boot infectors and file infectors were roughly equal. Figure 5 provides another view of what has happened to the relative prevalence of these two types of viruses over time. Beginning in 1992, the incident rate for boot sector infectors continued to rise, while the incident rate for file infectors began to fall dramatically. We will attempt to explain this phenomenon in a subsequent section.

It is interesting to break up our incident statistics even further into trends for individual viruses. Figure 6 shows the incident rate for selected viruses. Note that some viruses have increased in prevalence, while others have declined.

Figures 2–6 raise several important questions:

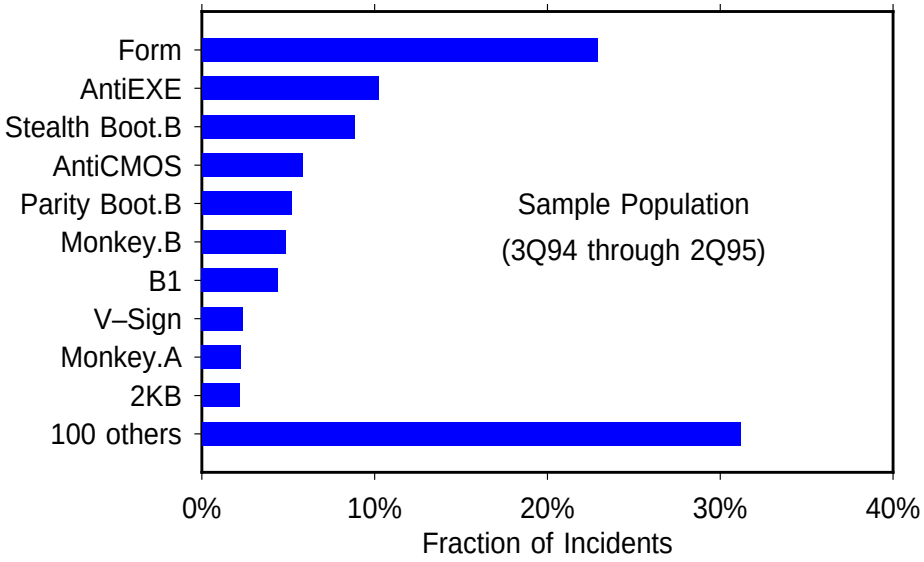

Figure 4: The top ten viruses account for two thirds of all incidents. All of them are boot-sector infectors.

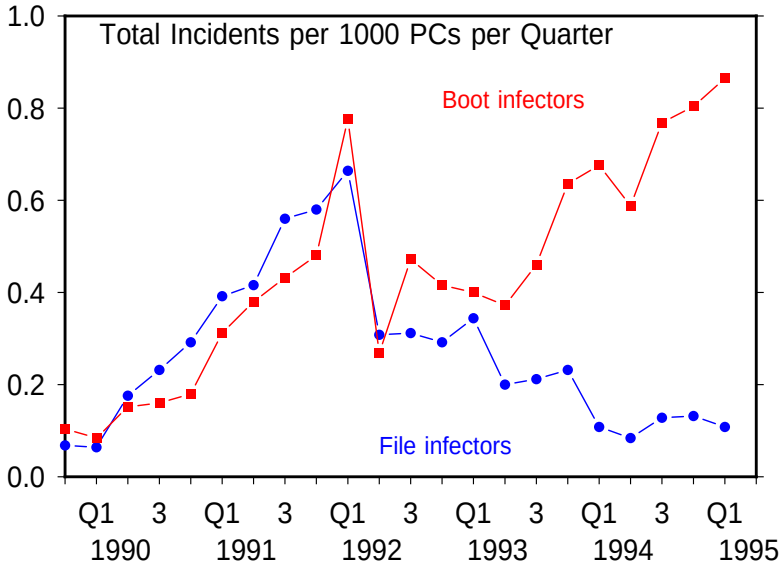

Figure 5: Boot viruses have continued to rise in prevalence, while file viruses have declined.

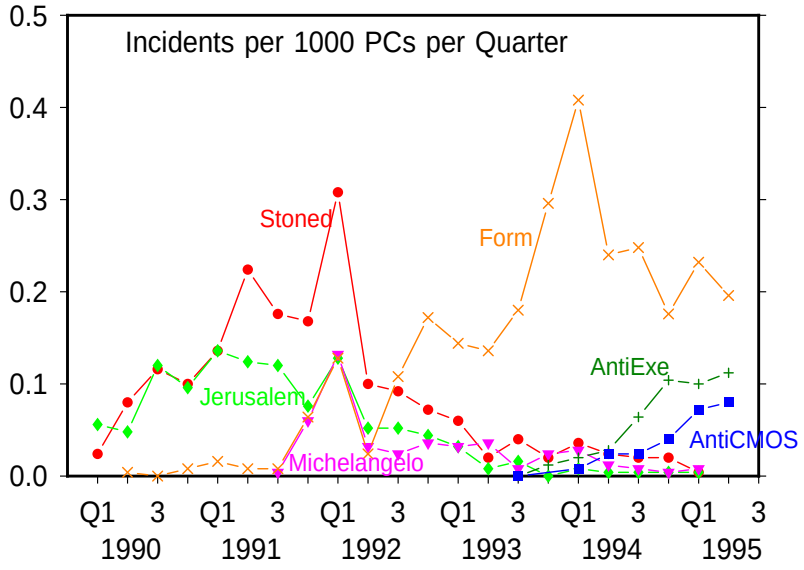

Figure 6: Some viruses have increased in prevalence, while others have declined.

1. Why are some viruses more prevalent than others?
2. Why do some viruses continue to increase in prevalence, while others plateau or decline?
3. Why are boot viruses so prevalent relative to file infectors, and why has their dominance increased over time?
4. Finally, can we predict what viruses are likely to become more prevalent in the future?

To begin to address these questions, we now review some of our previous theoretical work on virus epidemiology.

3 How Viruses Spread

Over the past several years, we have constucted theoretical models of how computer viruses spread in a population, and compared them against the results of an ongoing study of actual virus incidents [1, 2, 3, 4, 5, 6].

Our models are purposefully simple, in an attempt to understand the most important aspects of global virus spread. In these models, a system is either infected or not. If it is infected, there is some probability each day that it will have an infectious contact with some other system in the world, typically via exchange of floppy diskettes or software exchange over a network. This is called the *birth rate* of the virus. Similarly, there is some probability each day that an infected system will be discovered to be infected. When that happens, it is cleaned up, and it returns to the pool of uninfected systems. This is called the *death rate* of the virus.

The birth and death rates are influenced by a number of factors. A virus' birth rate is governed by its intrinsic properties, such as the particular way in which it infects and spreads. Just as for biological diseases, its birth rate is also highly dependent upon social factors, such as the rate of software or diskette exchange among systems. The death rate is determined by how quickly the virus is found and eliminated, which in turn depends on the extent to which people notice the virus, due to its behavior or through the use of anti-virus software. As we shall see, the birth and death rates also depend critically on the nature of the world's computing environment.

All of our models show the same basic characteristics of virus spread. One fundamental insight is that there is an *epidemic threshold* above which a virus may spread, and below which it cannot. If the birth rate of a virus is greater than its death rate, the virus has a chance to spread successfully, although it may die out before it spreads much. If the virus does manage to get a foothold, it will start to rise slowly in prevalence. The rate at which it does so is governed by a number of factors, such as intrinsic characteristics of the virus and the overall rate at which software is exchanged. A second fundamental insight that has emerged from our research is that the growth rate can be much slower than the exponential rate that was predicted by one theory [11]. Our theory shows that, when software sharing is localized, the global rate of spread can be very slow, even roughly linear [1, 2]. At some point, the virus levels off in prevalence, reaching an equilibrium between spreading and being eliminated. Figure 7 illustrates the typical behavior of a system above the epidemic threshold.

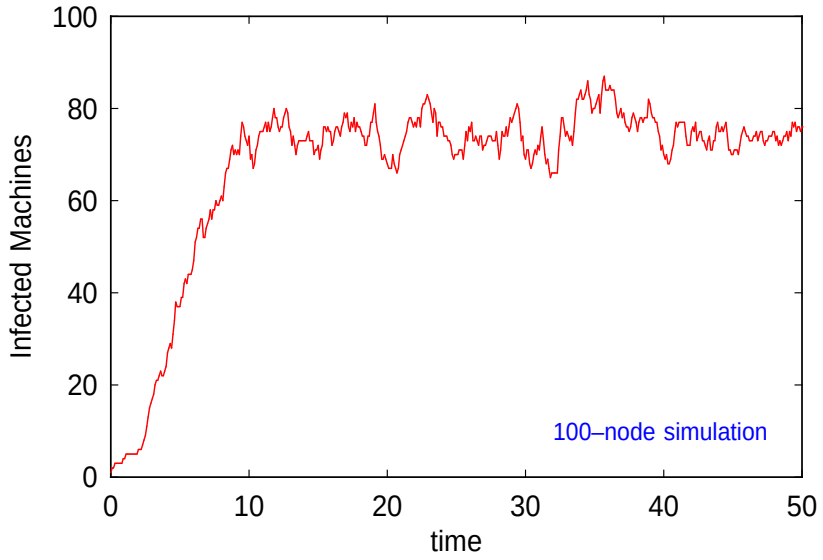

Figure 7: Above the epidemic threshold, a virus rises in prevalence at a rate that depends on a variety of factors, then plateaus at an equilibrium. In this simulation, the birth rate exceeded the death rate by a factor of 5.

If the birth rate is less than the death rate — if the virus is found and eliminated more quickly than it spreads — then the virus cannot spread widely. It may spread to a few machines for a little while, but it will eventually be found and eliminated from the population, becoming "extinct". Figure 8 illustrates this behavior.

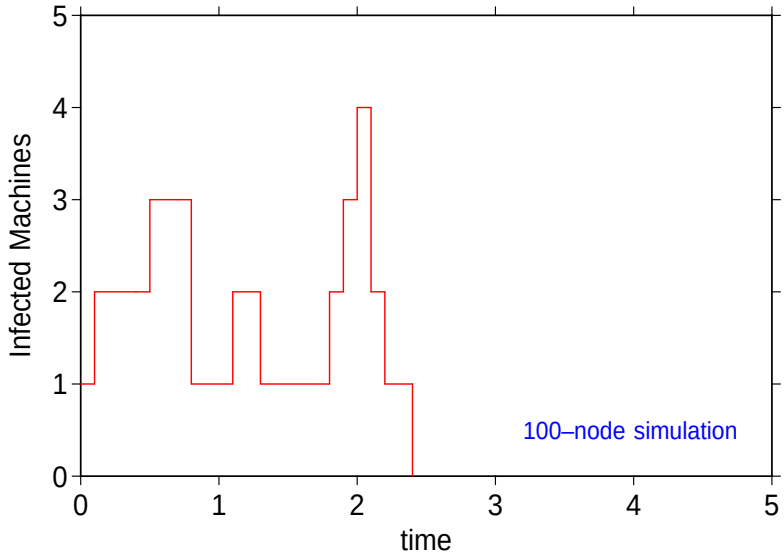

Figure 8: Below the epidemic threshold, very small outbreaks can occur, but extinction of the infection is inevitable. In this simulation, the birth rate was 10% less than the death rate. Note that the vertical and horizontal scales are much different than those of Fig. 7.

4 Virus Case Studies

In this section, we illustrate the interaction between viruses and their environment by narrowing our focus to the behavior of selected, individual viruses. We relate changes and shifts in virus prevalence to theoretical findings and to our knowledge of relevant shifts in the computing environment.

4.1 Michelangelo Madness

The Michelangelo virus was first found in early 1991 in New Zealand. It is a typical infector of diskette boot records and the master boot record of hard disks, with one exception. If an infected system is booted on March 6 of any year, the Michelangelo virus will overwrite parts of the hard disk with random data. This renders the hard disk of the system, and all of its information, inaccessible.

The virus is named Michelangelo not because of any messages in the virus itself, but because one of the first people to analyze it noticed that March 6 is the birthday of the famous artist. The name stuck.

Finding a new virus is not unusual in itself; several dozen new viruses are found each week. Michelangelo was unusual in that it was found in an actual incident, rather than as one of the thousands of viruses gathered by anti-virus workers but as yet unseen in an incident. It was also unusual because it could cause such substantial damage to the information on peoples' PCs, and because that damage would all happen on a single day.

In the weeks that preceded March 6, 1992, something even more unusual happened. In a fascinating interplay between the media and some parts of the anti-virus industry, the Michelangelo virus became a major news event. News stories warning about Michelangelo's destructive potential were broadcast on major television networks. Articles about it appeared prominently in major newspapers.

As March 6 drew nearer, the stories grew ever more hysterical. The predictions of the number of systems that would be wiped out grew to hundreds of thousands, then millions [12, 13].

When the fateful date came, the predictions of doom turned out to have been a bit inflated. The Michelangelo virus was found on some systems, and probably did destroy data on a few of them. But the worldwide disaster did not occur. Indeed, it was difficult to find any verified incident of destruction of data by Michelangelo in most places [14].

This should not have come as a surprise. Our own research at the time showed that the Michelangelo virus was not very prevalent, and certainly not one of the most common viruses. We estimated that about the same number of systems would have their hard disks crash due to random hardware failures on March 6 as would have their data destroyed by the Michelangelo virus. It is important to keep the risks in perspective.

Michelangelo Madness, as we came to call it, did have a dramatic effect, though not the anticipated one. Concerned about the predictions of widespread damage, people bought and installed anti-virus software in droves. In some locations, lines of people waiting to buy anti-virus software stretched around the block. In other places, stores sold out of their entire supply of anti-virus software during the week leading up to March 6. Around the world, a very large number of people checked their systems for viruses in those few days.

Figure 9 illustrates the effect of this activity. In the two weeks before March 6, 1992, reports of virus incidents shot up to unprecedented levels. Naturally, this was not because viruses were spreading out of control during those two weeks. Rather, infections that had been latent for days or weeks were found, simply because people were looking for them. In environments like that of our sample population, where anti-virus software is widely installed and used, it is likely that these same infections would have been caught anyway in subsequent weeks. But, since so many people checked their systems prior to March 6, the infections were discovered then rather than later.

People did find the Michelangelo virus, but they found far more viruses of other kinds. The Stoned virus, for instance, the most prevalent virus at the time, was found about three times more frequently than was the Michelangelo virus.

In the first few months after Michelangelo Madness, fewer virus incidents were reported than in the few month before it. This is easy to understand. First, virus incidents were caught earlier than they might have been because everyone was looking. Viruses found in the beginning of March might have been found in the beginning in April instead. So one would expect fewer virus incidents to be reported shortly after March 6 that year. Second, viruses were probably found and eliminated even in systems that might not have found them for a very long time. In just a few days, the worldwide population of viruses was decreased. We would expect that the virus population, and hence virus incident reports, would increase again in subsequent months.

Virus incidents did increase after that, but in a way that is rather complicated. We will examine this in more detail in a subsequent section.

Despite the beneficial effects of eliminating some viruses temporarily, the hysteria caused by this event was clearly out of proportion to the risk. Individuals and businesses spent vast sums of money and time warding off a threat that was much smaller than they were led to believe. We hope that those involved learned from the experience — that our friends in the anti-virus industry will be more careful in saying that they understand viral prevalence when they do not, and that the media

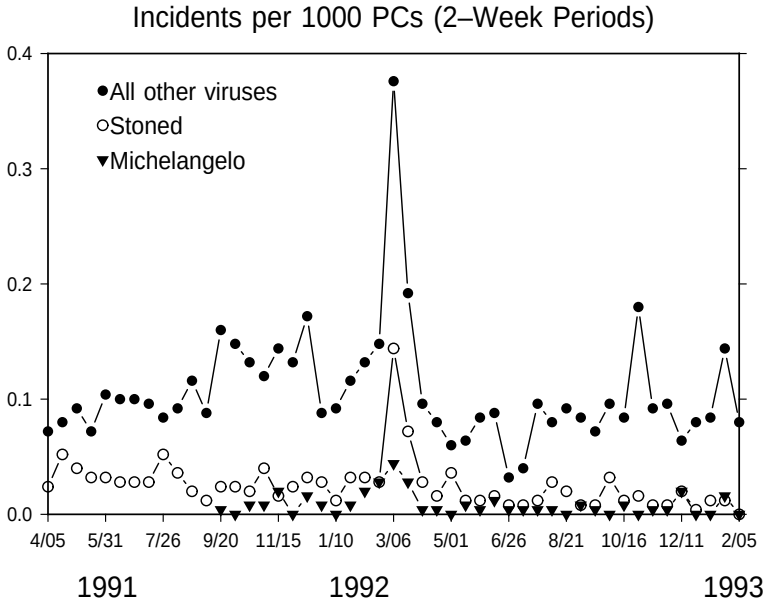

Figure 9: Michelangelo Madness resulted in many people finding viruses of all kinds.

will examine predictions of impending doom with a somewhat more critical eye.

4.2 The Missing Brain

The Brain virus was first observed in October, 1987, making it one of the first DOS viruses seen in the world [15]. It infects diskette boot sectors, and becomes active in a system when that system is booted from an infected diskette. Unlike most boot viruses today, Brain does not infect boot sectors of hard disks.

In the early days of PCs, most PCs were booted from diskettes and did not have hard disks. This provided a perfect medium for Brain to spread. Diskettes used in an infected system became infected themselves, and could carry that infection to other systems. Brain spread around the world in just this way.

Beginning with the introduction of the IBM PC-XT in 1982, the PC industry made a transition to systems that have hard disks. Unlike their predecessors, these systems were not booted from diskettes as frequently. When they were booted from diskettes, it was typically for some special activity, such as system maintenance. Once that activity was concluded, the system was rebooted from the hard disk. It became very uncommon for a system to be booted from a diskette and then used for an extended period of time, with more diskettes being inserted into the system. This denied the Brain virus the opportunity to spread in most cases. The world became a much more difficult place for the Brain virus to spread, and its prevalence declined.

This decline in prevalence occurred before we started gathering accurate statistics about virus incidents, so we cannot illustrate it quantitatively. Anecdotal evidence and our own informal statistics from the late 1980's, however, suggest that the Brain virus was substantially more common than

it is today. While the Brain virus is still seen on rare occasions, it does not spread well today. We sighted the Brain virus several times from mid-1988 until mid-1990, but since 1990 it has only appeared in our sample population once, in early 1992.

4.3 Not Stoned Again

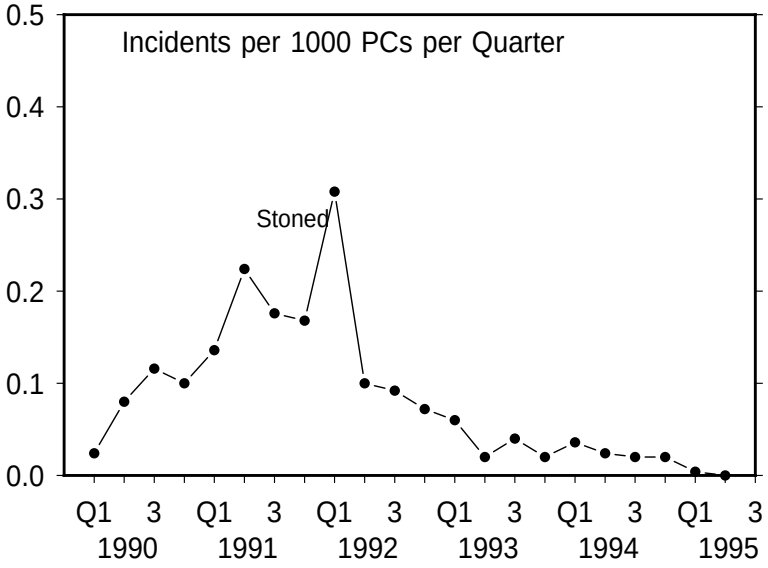

Figure 10: The Stoned virus, a boot infector, rose in prevalence and then declined.

The Stoned virus was first observed in an incident in 1989. It is a typical boot virus, infecting diskette boot records and master boot records of hard disks. One time out of eight that a system is booted from an infected diskette, the message "Your PC is now Stoned!" will appear on the display. The virus has no other effects.

The Stoned virus followed the expected pattern of rising in prevalence through 1991, at which time it had reached a rough equilibrium. After a large peak during Michelangelo Madness, it slowly declined in prevalence over the next several years. Once the most prevalent virus in the world, the Stoned virus is seen much less frequently today.

Its rise in prevalence and subsequent equilibration is what we expect of a virus. Its decline is a bit puzzling at first, until we notice that a system infected with the Stoned virus only spreads that infection to the diskette in the A: drive, not to any other diskette drive. The system became infected in the first place by booting from an infected diskette in the A: drive. The Stoned virus started its life on 5.25-inch diskettes. In spreading from diskette to system to diskette, it could only spread to other 5.25-inch diskettes.

Early in Stoned's life, most systems used 5.25-inch drives, so there was a fertile medium around the world which Stoned could use to spread. In the late 1980s, however, a trend began towards systems that used 3.5-inch drives as their A: drive. The fraction of systems that had 5.25-inch A: drives declined, and has been declining steadily ever since. With fewer and fewer systems that Stoned could infect and spread between, the virus too declined in prevalence.

4.4 Jerusalem's Rise and Fall

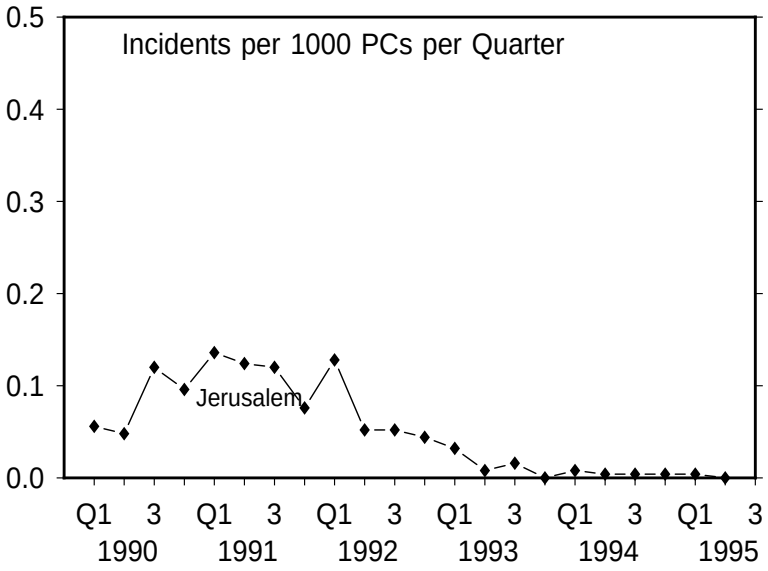

Figure 11: The Jerusalem virus, once quite prevalent, is seen much less often today.

The Jerusalem virus was first observed in December, 1987, in the city of Jerusalem, Israel [15]. In many ways, it is an archetypical file virus. When an infected program is run, the Jerusalem virus installs a resident extension in DOS. Subsequently, when any other program is executed, the virus' resident extension will infect the program file.

Prior to 1992, the Jerusalem virus followed the expected pattern of a virus that is spreading around the world. It rose gradually in prevalence through 1990. At the end of 1990, it had reached an equilibrium level in most of the world. Through 1991, it maintained this same level of prevalence, neither increasing or decreasing.

After 1991, however, an odd thing happened. Fewer and fewer incidents of the Jerusalem virus occurred. What was one of the most prevalent viruses in 1990 declined to one of the least prevalent viruses in 1995. Indeed, we saw only five incidents of the Jerusalem virus in our sample population in 1994, and just a single incident so far in 1995.

What caused this decrease? It was not a change in diskette drive type or the move from floppy diskettes to hard disks. File viruses like the Jerusalem virus spread to files on any kind of diskette, and persist in systems that boot from hard disks. We will return to the cause of this mysterious decrease in a subsequent section of this paper.

4.5 Form Follows Function

The Form virus was first observed in an incident in 2Q90. It infects diskette boot sectors and system boot sectors of hard disks. When the system is booted from an infected diskette or hard disk, the virus becomes active in memory and infects essentially any diskette used in the system thereafter.

Unlike the Brain virus, the Form virus remains on the hard disk and can spread if the system is booted from the hard disk subsequently. Unlike the Stoned virus, the Form virus is capable of

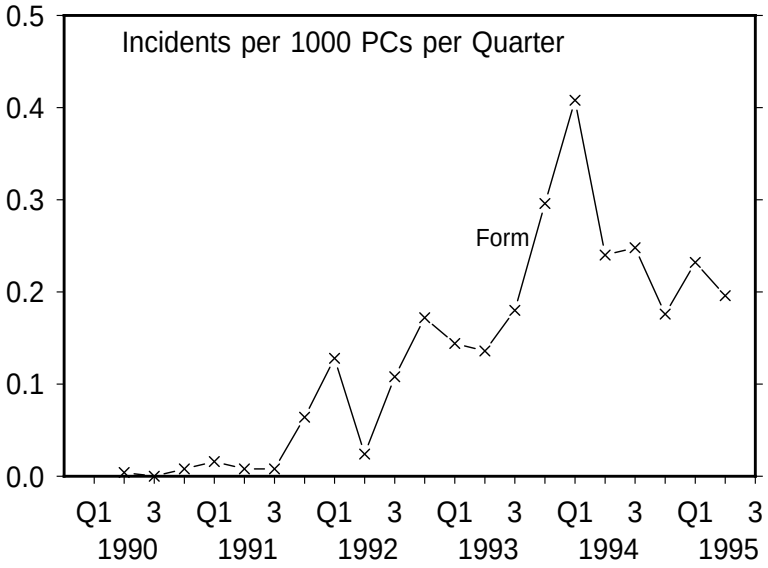

Figure 12: The Form virus, another boot infector, rose steadily in prevalence before reaching equilibrium.

infecting diskettes of any kind in any diskette drive, so it did not remain limited to one kind of diskette. On the 18th of any month, the Form virus will cause a slight clicking when keys are depressed on an infected system. This is often subtle enough to go unnoticed.

The Form virus does not possess the limiting features that caused the Brain and Stoned viruses to have difficulty spreading in the early and middle 1990s. It has exhibited what we expect to be typical behavior for a virus that has found its way into the world. It took over a year before it started rising significantly in prevalence. It rose steadily during 1992 and 1993, becoming the most prevalent virus worldwide. By the end of 1994, it had reached a rough equilibrium at about the same level as other mature viruses such as Jerusalem or Stoned. In the absence of environmental change, we might expect the Form virus to remain about as prevalent as it is today.

5 Why Are Boot Viruses So Common?

Boot viruses are by far the most common viruses today, accounting for nearly 90% of all incidents in 2Q95. File viruses, on the other hand, have decreased in prevalence. This is a remarkable change. Several years ago, file viruses accounted for around 50% of all incidents. What could be responsible for this dramatic change?

Was it Michelangelo Madness? No. That caused only a temporary depletion of viruses of all kinds. Michelangelo Madness explains the large peak in reported incidents, and the subsequent temporary decrease in incidents. It does not account for the difference in prevalence between boot infectors and file infectors.

Is it due to the increased use of anti-virus software? As anti-virus researchers and producers of anti-virus software, we would certainly like to think so. It is tempting to conclude that anti-virus software has made a difference in the world, given our experience with the sample population, in which we have found that widespread usage of anti-virus software and central incident management

substantially reduces the size of incidents within an organization [4, 5, 2, 6] Unfortunately, a closer look at our own data show that, while anti-virus software and policies can make a real difference within organizations, anti-virus software does *not* seem to have made as much of a difference to the world in general. All of the common viruses have been known for quite some time. All of them are detected, even by older anti-virus programs. If anti-virus software was responsible, we would have expected to see a decline in all viruses. The use of anti-virus software does not account for the difference in prevalence between boot infectors and file infectors.

To find the solution to this mystery, we look once again at changes in the computing environment, rather than events associated with the anti-virus industry. The biggest change in the PC computing environment over the past several years has been the change from the use of native DOS to the use of Windows 3.0 and 3.1. Windows 3.0 was released in 1990, and started to become a popular enhancement to the DOS operating system. Windows 3.1, released in 1992, accelerated this trend. Today, a large number of PCs run Windows 3.1.

How does Windows affect the spread of viruses? Experiments carried out at IBM's High Integrity Computing Laboratory demonstrated that Windows is a fragile environment in the presence of typical file viruses. In many cases, if a file virus is resident in the memory of a DOS system, Windows cannot even start. On the other hand, Windows behaves very differently on a system that is infected with a typical boot virus. For many boot viruses, an infected DOS system can not only start Windows, but can spread the virus to diskettes from within Windows.

If Windows users get a file virus, Windows will typically be inoperable. This will cause the users to eliminate the virus one way or another, whether or not they realize that the system is infected. They might use anti-virus software. They might send their system out for repair. They might re-install everything from backups. Whatever they do, they will eliminate the virus because they cannot get back to work until they do.

If Windows users get a boot virus, however, they might not notice it at all. Windows will usually start and function as expected. Unfortunately, the virus will typically spread to non-write-protected diskettes that are accessed from within Windows. In this sense, most boot viruses are not affected by Windows, and spread in just the same way whether the user is running DOS or Windows. Unless users have good anti-virus software, they will not usually have any reason to suspect a problem, and hence no reason to get rid of the virus.

This environmental analysis led us to predict, in 1994, that boot viruses would continue to increase in prevalence, oblivious to the use of Windows. Similarly, we predicted that file infectors would continue to decrease in prevalence. Furthermore, we predicted that boot viruses that were not then very prevalent would become more prevalent, while few file viruses would [16].

This is exactly what has happened. Figure 5 illustrates the dramatic rise of boot virus incidents over the past several years, and the corresponding dramatic decrease in file virus incidents.

Several boot viruses that do spread from within Windows, including AntiEXE and AntiCMOS, were low in prevalence in 1994 but are now substantially more prevalent. As shown in Figure 6, they are approaching the prevalence of more common boot viruses like Form. Once they increase to this level of prevalence, we would expect them to reach equilibrium and not increase further in prevalence.

6 Predicting the Future

We have come to the surprising conclusion that the world's computing environment has been the primary factor in determining the change in prevalence of computer viruses. It is reasonable to assume that this will continue to be the case for some time.

If this is so, we can get some insight into future problems by examining current trends and the expected changes in the computing environment over the next several years. Some of these changes will tend to decrease viral prevalence, while others will tend to increase it.

If there were no changes in the world's computing environment, we might expect to see current trends continue. File viruses would continue to remain very low in prevalence. Boot viruses that have already reached equilibrium, such as the Form virus, would remain at about the same level of prevalence that they have today. Other boot viruses would be expected to start becoming more prevalent, perhaps rising in prevalence until they too reached equilibrium. Since there are several hundred boot viruses, having all of them rise in prevalence to the level that Form has reached would result in a huge rise in virus incidents worldwide.

There are, however, some environmental changes that we might expect over the next few years: 32-bit operating systems and networking. These changes could have a significant effect on the virus problem.

6.1 32-Bit Operating Systems

One of the significant environmental changes will be the transition from DOS to 32-bit operating systems for PCs, such as OS/2 and Windows 95. In the next few years, we expect that more and more systems will run 32-bit operating systems in order to better use the increasing power of newer PCs.

IBM's OS/2 is a 32-bit operating system that lets users run DOS, Windows and OS/2 programs simultaneously. The effects of computer viruses on OS/2 systems is described elsewhere [17]. Boot viruses do not generally spread from within OS/2 itself, though they can spread from systems that have DOS as well as OS/2 installed in separate partitions.

File viruses can often spread to other files when infected programs are run in Virtual DOS Machines (VDM) within OS/2. However, they remain active in the system only as long as the infected VDM is active, which is often only as long as the infected program is running. Some file viruses are likely to not spread in VDMs, simply because of differences between VDMs and DOS. This decreases the rate at which file viruses spread in collections of OS/2 systems [17]. In environments in which OS/2 predominates over DOS, we would expect this to lead to a decline in prevalence of all current DOS viruses.

Microsoft's Windows 95 is a 32-bit operating systems that supports DOS, Windows 3.1 and Windows 95 programs. Recent experiments with a pre-release version of Windows 95 suggest that DOS boot viruses will not in general spread well from Windows 95 systems [18]. File viruses were not tested in these experiments.

Preliminary experiments carried out at the High Integrity Computing Laboratory with a pre-release version of Windows 95 suggest that some DOS file viruses will spread as usual, some might not,

and some might cause system problems. In environments in which Windows 95 predominates over DOS, we would also expect this to lead to a decline in prevalence of all current DOS viruses.

Not all of the news is good, however. Viruses can be written for 32-bit operating systems, and the first few such crude viruses have already appeared [17]. These operating systems offer new facilities that viruses can use to both hide and spread. The transition to these newer operating systems will change the virus problem, perhaps significantly, but it will not eliminate it.

6.2 Networking

As more and more systems are connected to local and wide area networks, networks may become a more common medium for viral spread.

Of particular interest is the inclusion of networking capabilities in newer 32-bit operating systems. If people typically configure their systems to take advantage of these capabilities, and if that leads to more program sharing on local area networks, it could also increase viral spread in these environments. Currently, these capabilities are used primarily for workgroup computing rather than wide area networking, so the increased spread will result primarily in larger incidents, affecting an entire workgroup instead of a single PC, rather than a large increase in worldwide prevalence.

The final trend that bears watching is the rise of the Internet and global computing. This has the ability to increase the virus problem substantially over time.

There have been incidents of DOS viruses being transmitted on the Internet. Sometimes, they are posted to Internet newsgroups, which function much like bulletin board systems for anyone on the Internet. When the infected programs are downloaded and run, they can infect your PC just like any other infected program. So far, vigilance and rapid action have spread the word about infected programs in newsgroups quickly, and eliminated the problems as they have occurred.

The Internet can be used to support wide-area file servers. These are much like file servers on a LAN, but they can be accessed globally. A virus can spread to files on a LAN-based file server, and from there to the other client systems attached to the server. Similar, systems that run programs from wide-area file servers can become infected if the programs on the server are susceptible to infection.

While boot viruses could be transmitted on the Internet as diskette images, which would be downloaded and installed onto diskettes, this seems unlikely to become a common means of transporting information. As more information is exchanged over the Internet instead of on diskettes, and the use of diskettes decreases, we would expect a decrease in the prevalence of DOS boot viruses. We would expect that the increased use of the Internet to interchange and access programs would promote an increase in the prevalence of DOS file viruses.

There have been a few incidents of viruses and worms that are specifically designed to use worldwide networks to spread [7, 8, 9, 10]. These provide dramatic examples of how quickly and how widely viruses can spread on such networks. Fortunately, while these incidents have been rapid and large, they did not usually recur. After a matter of hours or days, when the virus was eliminated from the network and increased defenses put into place, the virus did not continue to spread. Unlike DOS viruses, which have continued to spread around the world for years, Internet viruses have (so far!) been episodic — they come, and then they go. But this need not always be the case.

7 Conclusion

The problem of DOS viruses continues to get slowly worse around the world. There are many more viruses than there were a few years ago, and they are appearing at a slightly higher rate. Virus incidents have also increased slightly, but we have to analyze the changes in prevalence of each individual virus in order to understand this trend.

Fortunately, we have made significant progress in this regard. We have achieved a good basic understanding of the spread of computer viruses. We know that a virus can either spread widely or almost not at all, depending upon how fast the virus spreads and how quickly and infection can be found and eliminated. If a virus does spread worldwide, it will rise slowly in prevalence, until it reaches an equilibrium level in the population.

For DOS viruses, this rise is very slow, often taking months or years. The equilibrium level is also quite low. Well-prepared organizations experience about one virus incident per quarter for every one thousand PCs they have, and this incident rate has not changed substantially for a number of years.

Our ongoing study of actual virus incidents had also demonstrated the remarkable effectiveness of good anti-virus software coupled with central incident management in controlling the virus problem within an organization.

This paper has focussed on the causes of the major changes in viral prevalence worldwide. We conclude, perhaps surprisingly, that the use of anti-virus software does not play a major role in these changes. Rather, they are determined by the way in which specific viruses, and classes of viruses, interact with the world's computing environment.

We examine the history of several specific viruses to understand this interaction between a virus and its changing environment. The Michelangelo virus was never very prevalent, but media attention to it resulted in increased reports of viruses of all kinds, followed by a temporary decrease in reports. The Brain virus, which spread primarily among systems without hard disks, effectively died out as systems with hard disks became the norm. Virtually all file viruses, including the once-prevalent Jerusalem virus, have decreased dramatically in prevalence because of the increased usage of Windows, and because Windows is fragile in the presence of file viruses. The Form virus, along with other boot viruses, have increased substantially in prevalence, to the point where boot viruses account for around 90% of all virus incidents today. Their spread is not unusual. It is the expected behavior of viruses in a population. They have not died off as have file viruses because their spread is not limited by Windows.

If the computing environment did not change, we would expect that file viruses would remain very low in prevalence, while other boot viruses would increase substantially. If dozens of boot viruses became as prevalent as the Form virus is today, the total number of virus incidents would increase substantially.

By examining trends in the computing environment, however, we can analyze how these might affect computer virus prevalence in the next few years.

Increased use of 32-bit operating systems, such as OS/2 and Windows, is likely to cause a decrease in the prevalence of all current DOS viruses. This is not because they were designed to resist viruses. Quite the contrary, viruses can be written for and spread by these operating systems. Rather, the

predicted decrease in DOS virus prevalence is simply because features that current DOS viruses use to spread changed in these newer operating systems.

Increased networking, and global networking in particular, will tend to increase the spread of file viruses and decrease the spread of boot viruses. Viruses written to take advantage of features of 32-bit operating systems, especially local and global networking, could become increasing problems. This is a worrisome prospect, as viruses can spread with remarkable speed on world-wide networks.

The technology required to deal with a world of rapidly spreading viruses will be much more challenging than current anti-virus technology. It will be required to respond very quickly, and globally, to new viruses — probably more quickly than humans can respond. While elements of this technology are working in the lab today [19, 20] the task of creating an immune system for cyberspace will occupy us for some time to come [21].

Acknowledgments

The authors thank Alan Fedeli, Yann Stanczewski and many others for diligently gathering accurate information on worldwide virus incidents for many years. We also thank Joe Wells for his suggestion, later verified experimentally, that most boot viruses can spread from within Windows, while most file viruses cannot.

References

[1] J.O. Kephart and S.R. White, "Directed-Graph Epidemiological Models of Computer Viruses," *Proceedings of the 1991 IEEE Computer Society Symposium on Research in Security and Privacy,* Oakland, California, May 20–22, 1991, pp. 343–359.

[2] Jeffrey O. Kephart and Steve R. White. "Measuring and modeling computer virus prevalence," *Proceedings of the 1993 IEEE Computer Society Symposium on Research in Security and Privacy,* Oakland, California, May 24–26, 1993, 2–15.

[3] J.O. Kephart and S.R. White, "Commentary on Tippett's 'Kinetics of Computer Virus Replication'," *Safe Computing: Proceedings of the Fourth Annual Computer Virus and Security Conference,* New York, New York, March 14–15, 1991, pp. 88–93.

[4] J.O. Kephart and S.R. White, "How Prevalent Are Computer Viruses?," *Proceedings of the Fifth International Computer Virus and Security Conference,* March 12–13, 1992, New York, pp. 267–284.

[5] J.O. Kephart and S.R. White, "Measuring Computer Virus Prevalence," *Proceedings of the Second International Virus Bulletin Conference,* Edinburgh, Scotland, September 2–3, 1992, pp. 9-28.

[6] Jeffrey O. Kephart, Steve R. White, and David M. Chess. *Computers and epidemiology.* IEEE Spectrum, May 1993, 20–26.

[7] Spafford, E. H. 1989. "The Internet worm program: an analysis." *Computer Comm. Review* 19, 1.

[8] Cliff Stoll, "An epidemiology of viruses and network worms," *12th National Computer Security Conference,* 1989, pp. 369–377.

[9] M.W. Eichin and J.A. Rochlis, "With microscope and tweezers: An analysis of the Internet virus of November 1988," *Proc. 1989 IEEE Symp. on Security and Privacy,* Oakland, California, May 1–3, 1989, pp. 326–343.

[10] D. Seeley, "A tour of the worm," *Proc. Usenix Winter 1989 Conference,* San Diego, California, 1989, p. 287.

[11] P.S. Tippett, "The Kinetics of Computer Virus Replication: A Theory and Preliminary Survey," *Safe Computing: Proceedings of the Fourth Annual Computer Virus and Security Conference,* New York, New York, March 14–15, 1991, pp. 66–87.

[12] J. McAfee, quoting expert sources on The MacNeil/Lehrer News Hour, March 5, 1992.

[13] Joshua Quittner, "Michelangelo Virus: No Brush With Disaster," *New York Newsday,* April 5, 1992, pp. 68.

[14] Michael W. Miller, "'Michelangelo' Scare Ends In an Anticlimax," *The Wall Street Journal,* March 9, 1992, pp. B5.

[15] Harold J. Highland, *Computer Virus Handbook,* Elsevier Advanced Technology, Oxford, England, 1990, pp. 32.

[16] Steve R. White, Jeffrey O. Kephart, David M. Chess, "An Introduction to Computer Viruses," *Fourth International Virus Bulletin Conference,* St. Helier, Jersey, UK, September 8–9, 1994.

[17] John F. Morar and David M. Chess. "The Effect of Computer Viruses on OS/2 and Warp," *Proceedings of the Fifth International Virus Bulletin Conference,* Boston, Massachusetts, Sept. 20–22, 1995.

[18] "Viruses on Windows 95," Virus Bulletin, June 1995, pp. 15–17.

[19] Jeffrey O. Kephart, "A biologically inspired immune system for computers," in R. Brooks and P. Maes, editors, *Artificial Life IV: Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems*, pages 130–139. MIT Press, 1994.

[20] Jeffrey O. Kephart, Gregory B. Sorkin, William C. Arnold, David M. Chess, Gerald J. Tesauro, and Steve R. White, "Biologically inspired defenses against computer viruses," to appear in Proceedings of IJCAI '95, Montreal, August 19–25, 1995.

[21] IBM's Massively Distributed Systems home page on the World Wide Web, `http://www.research.ibm.com/massdist`