Computers & Security, 16 (1997) 412-415

A History Of Computer Viruses — Introduction

Harold Joseph Highland FICS, FACM

Editor-in-Chief Emeritus

The following series of articles are taken from Harold's Computer Virus Handbook, published by Elsevier Advanced Technology in 1990. Viruses have moved on a long way since then, but the extracts published here provide a useful background in virus development, and contain much information that is still relevant today. It is also interesting to note that Harold introduces the Macro Virus concept a few years before it became more widely identified as a major problem.

In this section we shall present detailed information about a number of computer viruses, specifically when and where the virus was found, how it behaved and a technical report on how it works as well as other relative information. We had hoped to present these data in historical perspective. However, it is too early to prepare a comprehensive history of computer viruses.

This volume is about DOS computer viruses, that is computer viruses that have been found in systems using either IBM-DOS or MS-DOS. No attempt has been made to cover the many other viruses that have surfaced to infect Macintosh microcomputers. Nor are any of the VAX viruses included. Furthermore, although we shall include detailed data about a number of viruses, we are not willing to 'put into print' some of the material and purported research reports currently available. So that the reader is better able to understand our viewpoint, we shall elaborate on some of the problems prior to the detailed reports about specific viruses.

A Matter of Definition

First, there is the question of a definition of a computer virus. There is currently no agreement in the computer community. To the general reader differences may appear slight but to the technician they are major.

There are many who consider computer viruses as the offspring of Dr. Frederick B. Cohen. He created a virus, as part of his doctoral thesis, in an effort to find ways to defend computer systems from self-replicating programs. There are others who claim that computer viruses existed well before 1984 when Dr. Cohen did his research. The debate about the appearance of the first virus will probably continue far into the future. Currently it does not appear likely that computer scientists will agree upon an 'official' definition of the term.

3160167-4048/97$17.00 © 1997 Elsevier Science Ltd

Dr. Cohen first made his research public at the 1984 National Computer Security Conference. He made his findings known to an international audience during his presentation that same year at the International Federation for Information Processing Computer Security Conference in Toronto, Canada, IFIP/Sec '84.

That conference was sponsored by IFIP Technical Committee 11 responsible for information processing security. It was attended by several hundred computer security specialists from all over the world. We often tell our lecture audiences about the reaction to his presentation at that meeting.

Later in the day, after Dr. Cohen presented his paper, we met with several computer security directors from Europe and Asia. Most of them felt that Dr. Cohen's report was interesting but esoteric. One security director from a major multinational corporation remarked that it was most interesting to him that an American university would provide a young man with a laboratory "to play games." He could see no "practical" application of the research and felt that it too would disappear among the many "useless, academic studies."

Dr. Cohen's reports, made in the United States and Canada, received little, if any, coverage in the European press. It was not until a presentation by Rudiger Dierstein of the Deutsche Forschungs und Versuchsanstalt fur Luft - und Raumfahrt [DFVLRl at SECURICOM in Paris the next year that the European press began to report about computer viruses.

Is It Really a Virus?

We have followed a conservative approach to the acceptance of computer virus claims. Unless we have been able to obtain a copy of the virus, disassemble it and see it in action, we have steadfastly refused to accept unsupported claims made by others.

For example, early in 1988 one of the anti-virus product producers reported that he had found a new computer virus that "destroyed" the hard disk. To obtain additional information I spoke a few weeks later with the individual who had reported the 'virus' to him. The 'virus' had appeared several months earlier on her system. What she found was that when backing up a file to a floppy disk using the DOS COPY command or even using her text editor, the backup copy was sometimes incomplete — part of the copy just vanished. Having read about the producer's appeal to report computer viruses, she telephoned him. At his request she sent him a copy of her hard disk.

During our conversation she admitted that she had not reformatted her disk and reloaded it with clean programs. Almost five months after the press ran the producer's report she was still operating as before. She still encountered the difficulty at infrequent intervals. Because the 'virus' was found only at one site and not reported elsewhere, we filed that report for future consideration.

Because we were busy with other viruses, we did not find time to follow up that story for many months. However, on November 14, 1988, Dr. P. M. Adams of the Computer Science Department of Nova University [Florida] issued a research report, "Hardware-Induced Data Virus: Floppy Diskette Controller Design Flaw." In it he explained that there was a basic flaw in INTEL's chip 8272A that had been used on the floppy disk controller board in roughly 25 million microcomputers. According to Dr. Adams, INTEL had sent a release on May 2, 1988 to its customers stating:

> "It has been found that the 8272A cannot detect a DMA underrun on the last byte of a write operation to a sector. If the 8272A is preempted during a DMA [1] transfer, and an overrun occurs on the last byte of a sector, the following occurs: the underrun flag does not get set, the last byte written to the disk is made equal to the previous byte written and CRC [2]

[1] DMA is direct memory access, a technique that allows peripheral devices to gain direct access to the microcomputer's main memory. This causes the processor to stop all activity along a bus, a communications line along which the data are transmitted — HJH

[2] CRC is cyclic redundancy check a method used for detecting errors in the transfer of data — HJH

> is generated on the ALTERED data. The result is that INCORRECT DATA is WRITTEN to the disk and VALIDATED by the 8272A."

Although we do not agree with Dr. Adams's use of the term, "hardware-induced data virus," it appears likely that the earlier reported 'virus' may well have been a hardware defect. In any event the so-called virus had not destroyed her hard disk.

The Numbers Game

An oft-repeated question by the press during an interview with anyone working with computer viruses is "How many computer viruses are there?" An answer that we are not certain but we have 15 in our laboratory, sends the interviewer off to find a 'better' source.

There appears to be a competition among some working in the computer virus field to announce a greater number than anyone else. One researcher who distributes his findings on a bulletin board announced that he had collected and examined 48 computer viruses. Another whom we heard at a conference in the late spring of 1989 told the audience that he had already collected more than 160. It appears that the more computer viruses one can list, the greater an authority he is on the subject.

There are viruses and there are often mutations of these viruses. For example, one attribute of the Brain virus is to write "Brain" as the label on an infected disk. If another virus is found that writes "HA-HA" as the disk label but is identical in every other respect, does one count this as "a new" virus? The code of both are the same but only five ASCII characters have been changed.

We deal with these mutations in a simple way. So long as critical code in the virus has not been altered, we call the other virus a variant or mutant. On the other hand, many researchers have taken an easier way out. If there is any change, no matter how slight, they count the other as a new virus. The policy we have followed thus far leads to some problems.

- For example, a virus that will attack all .COM programs except COMMAND.COM appears in an altered form so that even COMMAND.COM is attacked. Although the modification of the code of that virus is no more difficult than the change of the label name, this altered form is different. The action of the virus has been modified.

- Similarly if a virus that attacks only 5 1/4-inch floppy disks appears so that it is capable of attacking a hard disk drive, do we consider it to be a new virus?

We feel that so long as any two viruses have identical code and do not behave differently, they are variants of the original virus. However, if their actions have been modified they should be classed as a new form of the original virus. We are not interested in amassing numbers. However we feel that a logical, scientific approach to virus taxonomy is needed.

Virus Identification

Each time a virus appears in a new location, the finder often believes he has a new virus. When we received a virus from an associate in the Middle East we accepted the name as the Ping-Pong virus. Our first reports from England late in 1988 about that virus called it the Italian virus. Later some researchers there renamed it as the "1803" virus. Since then we have seen it called the Bouncing Ball virus and the Turin virus.

The virus specialist, who does not have a copy of this virus and/or is unable to confirm that the versions are identical, is too often misled. He is likely to consider counting each as a separate virus. Even if he goes through the many reports from the different centres he might not be fully informed.

Most serious researchers have called for a protocol whereby specialists in different parts of the world can compare the viruses they have without the need to send the actual virus and/or its disassembled code. Most are reluctant to send either for fear of spreading the virus. Making source code or a copy of a disk with a virus available is dangerous. It takes little effort on the part of a skilled programmer to modify the trigger and/or

action portions of a virus once one has a workable copy.

There have been calls by researchers to establish a central clearing house for computer viruses. In most cases the researcher feels that his site should serve as that center. We have long felt that there is need for a method by which researchers can exchange information without sending the actual virus and/or the disassembled code. Charles M. Preston, a computer security specialist and virus researcher in Anchorage [Alaska] and we have discussed the need for creating a computer virus directory. That directory would provide specific information about each virus; among some of the data would be:

- its size in number of bytes,
- the medium which it attacks,
- a hexadecimal or ASCII checksum of its actual code,
- the signature, if any, that the virus uses to avoid reinfection,
- a listing of ASCII strings in the viral code and their location, and
- detailed information about the replication procedure, the trigger mechanism, and action taken.

Source of Virus Data

In line with the conservative policy we have followed since the computer virus explosion in late 1987 we have included virus analysis in the following few sections based on the following sources:

[1] Computer viruses we have in our laboratory. These viruses have been received from sites that have been attacked as well as from associates in different parts of the world. In addition to our own analysis we have supporting information from Bill Kenny, a highly-skilled programmer and analyst with Digital Dispatch Inc. of St. Paul [Minnesota] and Dr. Jon David of Systems Research and Development of Tappan [New York]. We should also acknowledge the assistance from several computer security specialists in different parts of the world, ranging from Australia to the United Kingdom to Finland and Sweden.

[2] Substantiated reports from reliable researchers. Although we have a number of computer viruses and mutations, we do not physically have copies of all the viruses that have been found in the world. Many researcher reports cannot be confirmed and others have analyses of viruses that do not conform with our findings; these were not used.

[3] Finally we should note that the material presented in the section of laboratory viruses has come from sources that cannot be publicly identified. In each case, however, we have thoroughly examined the data and investigated the integrity of the source.