Beyond Virtual Vaccinations

Developing a digital immune system in bits and bytes

By DAMARIS CHRISTENSEN

The fear of new, dangerous viruses sweeping through an unprotected population is not limited to public health officials. Computer researchers have long worried because typical virus-scanning computer programs—which essentially vaccinate machines against known viruses—become outdated as newly created viruses spread over the Internet.

Just as researchers turned to biology in applying the name *virus* to the pesky programs that could make computers sick, several groups have turned to biology for a new model of how to protect computers against unknown viruses. They are focusing on the human immune system.

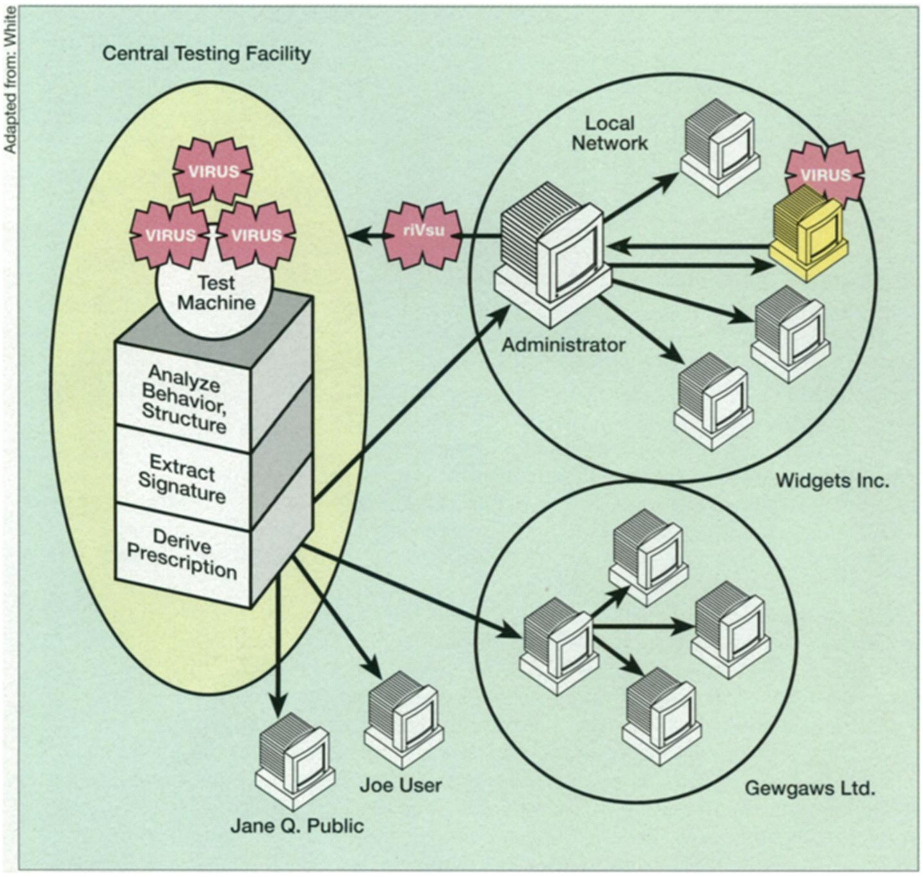
These computer scientists hope to develop a digital system that, like the immune system, can quickly recognize and fight off known infections, identify new intruders and learn how to deter them, and remember all previously encountered pathogens. Such a system also needs to be safe, reliable, and secure.

A computer virus released in March aptly demonstrated the need for more-effective ways of fighting off computer viruses. Although warnings about the Melissa virus went out soon after it was identified, it spread as quickly as the alarms (SN: 5/8/99, p. 303). Within just a few days, the virus had circled the globe, sending countless unwanted E-mail messages across the Internet and clogging E-mail service at hundreds of organizations, forcing them to shut off their Internet connections.

Although Melissa—the first virus to mail itself around the world—merely clogged E-mail systems, virus makers have already launched spin-offs of the virus designed to destroy data.

The risk of computer infections rises as more information is exchanged through E-mail or over the Internet. Likewise, the potential damage that viruses can create multiplies as people send sensitive personal and corporate data over the Internet. Computer security experts also warn that the avenues for viruses to spread multiply dramatically as computers use software that's integrated so that one program can launch another.

"There used to be plenty of time to analyze a virus before it spread, but Internet-borne viruses can spread around the world in hours or days," says Steve R. White of IBM's Thomas J. Watson Research Center in Yorktown Heights, N.Y. "In a world where things can travel this quickly and do this much damage, we have to have automated ways of dealing with them. It is silly to think that we can protect against these viruses manually."



Adapted from: White

A digital immune system: A virus triggers the infected computer (yellow) to forward a sample of the viral code to an administrative machine, which in turn sends an encrypted sample (riVsu) to a central testing facility. There, a test machine lures the virus into replicating. Once the virus is confirmed, other components of the immune system produce prescriptions for identifying and removing the bug. This information then travels back to the administrative machine, which forwards the prescription to the infected computer and then to other computers on the local network. Eventually, the immune system designers envision sending automatic antivirus updates to computers worldwide.

Computer viruses got their name from what White calls "an obvious but deep biological analogy." Like biological viruses, the computer versions replicate by attaching themselves to a host (a computer program rather than a human cell) and then co-opting the host's resources to make copies of themselves. Infection can lead to death: The computer crashes and all program information is irretrievably lost. Infection can also lead to sickness when the virus does not destroy any data but spreads and slows programs and communications. Even seemingly innocuous viruses may taint files and make the computer more likely to crash—like a long-lasting, low-grade infection.

Companies spend several hundred million dollars annually on antivirus products and services, and they lose even more in downtime when they need to take their systems off-line to prevent viral infections from spreading.

Because antivirus programs can only identify the viruses they already know, they aren't effective against the 10 to 15 new viruses created every day. Worst of all, says White, "many users of antivirus software blissfully continue to use antivirus software that is more than a year out of date."

Aside from frequent updates, there are few ways of strengthening this system. Some antivirus programs can monitor a computer system for viruslike behavior, such as making a file bigger without adding new data, but such systems are prone to false alarms and virus makers can take steps to evade such detection systems.

In the early 1990s, White and his colleagues at IBM dreamed of a digital immune system for computers (SN: 7/23/94, p. 63). For a model, they looked to the human immune system, which is constantly bombarded by infectious agents it has never before encountered and yet to which it generally responds quickly.

Computer virus makers often reuse key parts of existing viruses in their new creations, White explains. An immune system should be able to identify previously unrecognized viruses by these short so-called genes, which often are critical to the viruses' function. Although conventional software might contain some of these genelike sequences, the presence of many is typically a sign of viral infection, White says.

When a computer participating in a pilot test of this digital immune system finds virus genes or any other signs of infection, it strips out confidential data and encrypts the rest. The altered file then goes to a central computer facility at IBM to be analyzed. A computer there routes the virus to a test machine that lures the virus into replicating by running a variety of programs. If any of these decoy programs become infected, the test computer attempts to pull out a signature that can identify the virus in other computers.

The signature and a prescription to strip the virus out of infected files is then sent back to the central computer. It adds the new virus to its database and sends the information on detection and treatment back to the infected computer. IBM's automated process typically takes less than 5 minutes to identify a virus signature and derive a prescription, the developers claim.

Uninfected computers will also be "vaccinated," as the IBM team puts it, against infections with this new virus as soon as they check the updated database. Ultimately, White envisions, uninfected computers will be vaccinated automatically.

Later this summer, IBM, in conjunction with a leading antivirus-program developer, Symantec Corp. in Cupertino, Calif., plans to release an antivirus plan that includes such a digital immune system. "This is the first step toward a comprehensive system that can spread a global cure for a virus faster than the virus itself can spread," White says.

The IBM researchers are still trying to develop ways to mimic another trait of the immune system. An infected cell produces chemicals signaling distress, warning neighbor cells to put up barriers to slow the spread of the virus. Thus, when the immune system develops ways of attacking the intruder, it can quickly outpace the spread of the virus.

The biological analogies of computer security may stretch even further than IBM's vision, says Stephanie Forrest of the University of New Mexico in Albuquerque. The human immune system identifies foreign invaders because they don't carry the body's typical flags of "self," not because they resemble other infectious agents. Forrest and her colleagues have found a way for a computer to identify self.

By looking at short sequences of signals between a program and the computer's operating system, she and her colleagues have defined patterns unique to each machine. Abnormal patterns may be a sign of infection. For example, a program making unusual demands on system resources has very likely been coopted by a virus or is being attacked by a hacker, says Forrest.

"We've shown pretty convincingly that looking at these short sequences of self gives good discrimination between what is self and what isn't," she says. Such a system can be very efficient, Forrest points out. The protected computer uses its resources to check only programs and files that it is using.

She and her colleagues have also shown that information packets flowing into and out of a network of computers hooked to the Internet show patterns recognizable as self or nonself.

Like white blood cells in the human body, a digital immune system can create antibodies that recognize foreign material, Forrest says. To minimize the chances that the antivirus program will attack the computer itself, it would always destroy antibodies that flag patterns that are intrinsic to the computer. Using the remaining digital antibodies, the system will periodically check for abnormal patterns that may signify virus infections or intrusions from hackers.

Forrest and her colleagues are working on a system that will allow a computer to continually learn to redefine itself, so the computer can accept new programs without flagging them as viruses. The researchers have not yet explored how to attack viruses once identified.

Forrest says that a self-recognizing system will be practical even for individual computers connected to the Internet and used primarily for E-mail, writing, designing graphic presentations, and perhaps a little programming.

Though still theoretical, Forrest's approach may offer many advantages. A different immune system would run on every computer. Since every computer would create different antibodies, a virus that evaded one computer might not escape detection by another, limiting the spread of the virus. Likewise, a person

Computer viruses: Then and now

The first computer virus, called Brain, appeared in 1987. The people who created the first viruses hitched them to operating systems (such as DOS) or to applications (such as games or editing programs). Some of these viruses are still circulating. With these viruses, when a user turns on an infected computer or runs an infected program, the viral code copies itself into the computer's memory—and from there into any subsequent applications the user runs. These viruses spread only when a computer user shares tainted files and programs with other people.

On the other hand, viruses like Melissa latch onto macros, small programs hidden in word processing software. For example, when an unsuspecting recipient of the Melissa virus opened an infected document written in Microsoft Word, the virus activated and hijacked another program known as Microsoft Outlook. This program E-mailed copies of the infected document to the first 50 people listed in the program's address directory. The virus spread so quickly because so many people use both Word and Outlook.

Until macros became commonplace, viruses couldn't infect data files, including word processing documents and spreadsheets. Macro viruses proliferate rapidly because many people share data files freely, and they do so primarily through E-mail. Once one data file is infected, a virus can infect all other data files of that application as soon as they are opened.

By the end of 1998, programmers and users had identified more than 30,000 viruses. Viruses of all sorts now affect millions of computers every year. *—D.C.*

who broke into one computer network and managed to avoid detection by that system might not be so successful on another network, she says.

"They've taken a much more exact analogy with biology by developing digital antibodies," says White. "But the analogy breaks down. All of my cells come from me, so my immune system can define self. But I put files on my computer every day.... This system may be very good for intrusion detection, but it may not be a good approach for viruses, because it will make too many mistakes. Our approach is more specific for viruses."

Both research groups caution that in nature, no defense system remains perfect forever. Just as white blood cells and viruses engage in a delicate dance, each evolving to outwit the other, so will computer viruses and antivirus technology, White says.

Viruses are getting more dangerous all the time, he says. Several programs for automating the development of macro viruses are circulating, meaning that the virus-writing community can create viruses faster than ever.

There are even some indications that viruses may be evolving on their own, White says. For example, some versions of Microsoft Word may make minor errors when copying viruses. These changes may disable the virus, or they may make the virus harder to spot. Also, if two or more viruses successfully infect a computer, one may accidentally copy itself into the other virus, creating a new kind of bug, he says. While uncommon so far, these scenarios are certainly threatening, White notes.

Whatever the form of the threat, the goal of protecting computer systems remains. "What we would ideally like is for a computer to behave the way the human body does," says Sushil Jajodia of George Mason University in Fairfax, Va. "When we are attacked by a virus, we get sick, but the immune system detects the virus, defeats it, and heals the damage. Computer systems are not like the human body, though, in that we need to provide the technology."

Because programs and operating systems are not usually designed with security in mind, antiviral programs will always be behind the curve, says Jajodia. "It still isn't clear how well this idea [of digital immune systems] will work, but we have no better alternative for detecting virus infections," he says.

Computer users have demanded ease of use but not security, says Forrest. "While people are becoming aware of the issues . . . they don't feel personally threatened yet." She notes that "when the Internet took off in the early '90s, it became evident that the computer-security problem was going to become everybody's problem."

Jajodia, editor-in-chief of the JOURNAL OF COMPUTER SECURITY, says that programmers should address the problem of viruses long before people begin using newly developed software.

Designing computer systems and programs with security in mind is an important first step, he says. More programs should check digital signatures to confirm that transferred files and computer code come from a trusted source. Better encryption systems, which help ensure that information has not been altered in transit from one computer to another, would make it harder for people to design viruses and for viruses to spread, he says.

Computer-security experts warn that no single set of changes will be enough to completely protect increasingly interconnected computer systems. They hope, however, that new security measures, such as digital immune systems, will fend off future epidemics. □

Biology

The early fetus gets the womb

The human egg, once fertilized, apparently has only a short window of time in which to make it from a fallopian tube to the uterine wall. If the fertilized egg doesn't implant there within a week or so of ovulation, scientists find, the chances of a successful pregnancy begin to plummet.

Allen J. Wilcox of the National Institute of Environmental Health Sciences in Research Triangle Park, N.C., and his colleagues recruited 221 women who were about to stop using birth control because they wanted to become pregnant. From the concentrations of certain hormones in urine, the researchers could determine the day a woman ovulated. "We collected about 20,000 urine specimens. That's a lot of women collecting urine every morning and putting it into freezers," laughs Wilcox.

By also detecting the hormone chorionic gonadotropin in urine—the same method that home pregnancy tests use—the scientists could discern when an egg implanted. Cells that will become the placenta make this hormone to halt the menstrual cycle so that the woman doesn't shed the uterine lining and the implanted egg.

As the researchers describe in the June 10 NEW ENGLAND JOURNAL OF MEDICINE, they followed 189 women after conception. In almost all the women, the egg implanted 6 to 12 days after ovulation. The later the time of implantation, however, the more likely it became that the fetus would not survive its first 6 weeks. Indeed, no egg implanting after 12 days endured that initial period, let alone produced a live birth.

Animal studies indicate that the uterine wall becomes less receptive to implantation later in the menstrual cycle, which may explain the findings. Another possibility is that fertilized eggs that journey sluggishly to the uterus may have defects that make them less likely to survive. "It could well be both: There's a limited window of receptivity, and slower conceptuses are more likely to fail," says Wilcox. —*J.T.*

Gene proves to be a pain in the back

Totaling up to 10 pounds per person, the collagens are the most abundant proteins in the human body. "These proteins support our organs and our bones," says Leena Ala-Kokko of the MCP Hahnemann University in Philadelphia. "They hold us together."

Ala-Kokko and her colleagues have now linked a defect in a collagen gene to herniated disks, an excruciatingly painful back problem that afflicts many people. The protein under scrutiny is collagen IX, which represents a relatively minor component of the spongy disks that separate the vertebrae in the back.

Several years ago, a Japanese research group created mice that have a mutation in one of the three genes needed to form collagen IX. As the rodents aged, their disks degenerated.

Curious whether collagen IX mutations trigger back problems in people, Ala-Kokko's team and colleagues in Finland surveyed people with sciatica, a pain that radiates from the lower back to below the knee. A bulging or shattered disk that presses upon a nearby nerve commonly causes sciatica, and the researchers confirmed herniated disks in 157 of the people.

The investigators then examined one of the three genes that together encode collagen IX. In six of the people, the gene had an altered DNA sequence that results in an amino acid switch, tryptophan for glutamine, within the protein. Of 174 people with no sciatica or known disk problems, none had the same alteration, the researchers report in the July 16 SCIENCE.

Finally, they looked at the families of four of the people with the suspicious DNA variation: 26 family members overall had the tryptophan-encoding sequence, and each one had disk problems. "That's when we were convinced," says Ala-Kokko.

The researchers have begun to look for alterations in the two other collagen IX genes. People with collagen IX mutations may wish to avoid other factors, such as obesity, that increase the risk of disk problems, says Ala-Kokko. —*J.T.*