Defending Computer Networks against Attack

ASSAULTS on stand-alone and networked computers, called cyber attacks, are escalating in frequency and severity as the world relies increasingly on World Wide Web applications for commerce, defense, research, education, and health care. In particular, government operations have come to depend on the Internet and are, therefore, vulnerable to a variety of attacks. As a result, cyber security has become a top national priority requiring the best computer experts in government, academia, and business. Some of these experts are working on a Lawrence Livermore project whose goal is to develop a fundamentally new approach for cyber defense.

"A large-scale computer network is the most complicated thing humans have ever developed," says Livermore engineer Jim Brase, who helps oversee several Laboratory efforts in cyber defense. Brase notes that commercial products such as antivirus software are useful at thwarting cyber attacks for stand-alone computers. However, they are inadequate for defending a network of thousands of computers from attacks orchestrated by groups of expert programmers as well as solitary hackers.

Forms of attack vary but most are attempts to read, alter, or destroy data or to compromise a computer's operating system to take control of the machine. Most computer users are aware of the possible danger from computer viruses, worms, and "phishing," in which an attacker sends an e-mail purporting to come from a valid bank or credit card company and requests personal information. They are also aware that simply surfing the Web can result in "drive-by downloads," in which malicious software (malware) is unknowingly installed on the user's computer.

Some cyber attacks are the work of solitary hackers simply yearning for notoriety. However, far more sophisticated threats exist, in particular from overseas groups, designed to steal important military and business data as well as personal banking information. U.S. computer experts estimate that more than 60,000 machines per day are co-opted into loose networks of computers, called botnets, some of which are operated by foreign professionals.

Cyber Attacks Are Unrelenting

Lawrence Livermore's unclassified computer network, which includes about 40,000 machines, is under continuous siege from cyber attacks. Detecting these attacks is a daunting challenge for Laboratory cyber security experts because the volume (several trillion bytes, or terabytes, of data per day) and diversity of legitimate traffic make it difficult to identify the relatively small amount of malicious activity. "Furthermore, the increasingly

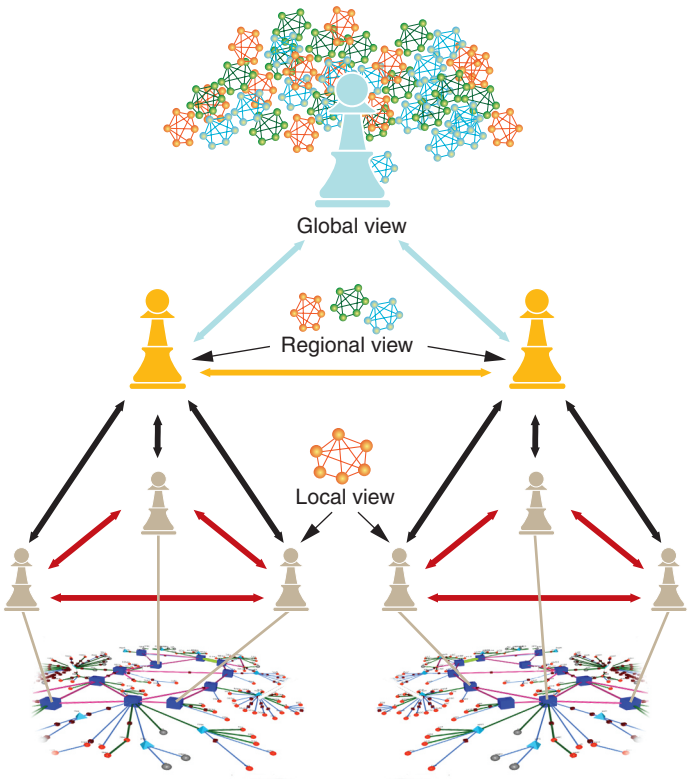
sophisticated attacks are designed to be undetectable," says computer scientist Celeste Matarazzo.

Matarazzo is leading a Laboratory Directed Research and Development–funded project called the Supercomputing Enabled Transformational Analytics Capability (SETAC). The project's goal is to dramatically increase the ability to detect, characterize, and combat malicious attacks on large computer networks. The three-year effort focuses on establishing situational awareness—that is, a state of continual awareness—of network behavior to better detect malicious intrusions in real time (or nearly real time), while being respectful of individuals' privacy. In this way, human analysts will have the opportunity to respond to threats immediately.

Matarazzo says that, to date, effective situational awareness of computer networks has been challenging because of the problem's scale and complexity. For example, at the perimeter of Livermore's unclassified computer network, terabytes of traffic data are collected each day, containing hundreds of millions of connection records. Furthermore, the cyber threat is extremely dynamic as adversaries can continually change the Internet Protocol (IP) addresses from which they conduct their operations, making detection difficult. Another challenge is malware that changes its behavior over time. Finally, the Web is always evolving, and computing environments blur the lines between personal computers and applications that reside on networks.

As part of the situational awareness effort, SETAC researchers are using complex algorithms distributed throughout the network together with novel hardware architectures derived from supercomputers. The researchers plan to deploy the algorithms (also called software sensor agents) to collect, analyze, and share data across the unclassified Livermore network. These distributed software sensor agents will also access data provided by commercial tools such as antivirus software and "learn" to quickly recognize suspicious behavior and take any necessary protective actions.

The sensors are designed for use at specific locations: firewalls (the part of a network that attempts to block unauthorized access), intermediate routers (devices that join smaller networks), and individual computers. Basic sensors report data to manager sensors that have more sophisticated analysis capabilities and a regional view of the network. In turn, manager sensors report to director sensors that have full analytic capabilities and a global view of the network. Currently, for testing purposes, about 50 sensor agents have been deployed on host machines of development team members and interested parties. Thousands of sensor agents will be deployed by the project's completion in September 2011.



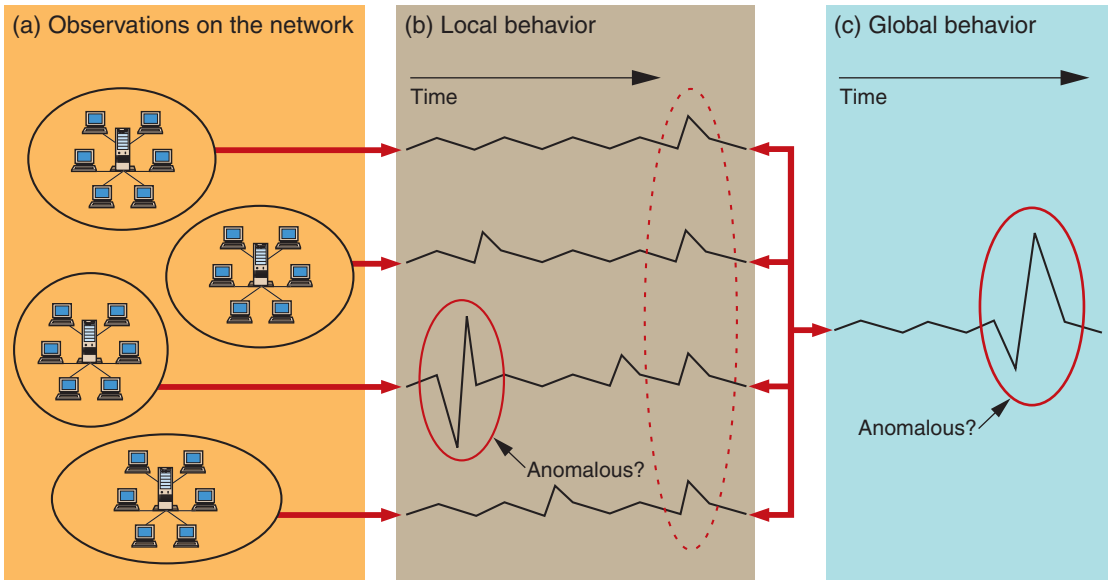
The Supercomputing Enabled Transformational Analytics Capability (SETAC) will deploy thousands of software sensor agents (algorithms) that collect, analyze, and share data across the unclassified Lawrence Livermore network. Basic sensors (gray) report data to manager sensors (gold) that have a regional view of the network. In turn, manager sensors report to director sensors (blue) that have a global view of the network. In addition, information is shared among all levels for decentralized control.

Searching for Anomalies

Sensor agents look for anomalous scenarios such as multiple machines simultaneously performing the same action or an unauthorized action, a large amount of data suddenly being sent outside the Laboratory, a computer accessing a supercomputer that it has never previously accessed, or a computer communicating with an outside server that frequently changes its IP address. Should a sensor detect an anomalous scenario, it shares this information with other sensors to determine if similar behavior is occurring elsewhere on the network; if so, a security analyst would be alerted.

Matarazzo emphasizes that SETAC does not displace human cyber security personnel. Rather, "SETAC enhances an analyst's ability to make timely decisions," she says. By gathering network data and performing analyses in real time, the sensors permit human analysts to respond to threats as they unfold, thereby preventing identity theft, data collection, and installation of malware.

The three primary characteristics of SETAC—distributed decision making, an emphasis on behavior modeling using machine learning approaches, and real-time analysis and detection—are novel features of cyber defense. SETAC's situational awareness emphasis is preferable to today's typical cyber defense, which relies on commercial software at the organization's network perimeter and analysis after an attack



SETAC is designed to detect both local and global patterns of behavior. (a) In this example, sensors continuously observe the behavior of four computers. (b) A spike in activity of one computer indicates a possible anomalous local behavior. (c) A simultaneous spike in activity of all four computers indicates a possible anomalous global behavior. The information collected is then shared with other sensors, and a security analyst is alerted for further investigation.

has occurred and damage has been done. "Current cyber defense has limitations," says Matarazzo. For example, once an internal machine is breached, the entire network is at risk, because of the difficulty in preventing the spread of malware from within the organization. Also, intrusion detection typically means searching for known attack signatures, such as IP addresses previously identified as belonging to hackers. Intrusions are temporal in nature and are often identified only through analyses that cannot be performed in real time. These limitations frequently make contemporary cyber defense a job of cleanup and repair following an attack. Matarazzo adds, "Commercial tools are moving in directions similar to SETAC, although they deploy proprietary algorithms geared toward solving specific problems at limited scale."

SETAC's continual monitoring is analogous to credit card companies monitoring their clients' credit card use, with deviations from typical activity triggering a warning of possible unauthorized use. To detect anomalies of computer usage, however, SETAC developers must better understand what Livermore network activity looks like on a "normal" workday and weekend. "There are many dimensions of 'normal,'" says Matarazzo. For example, different departments and groups, and even occupations, may have different normal patterns of behavior.

Cyber Defense Taps Many Specialties

Livermore's research team includes computer scientists, statisticians, mathematicians, and engineers. Academic partners include the University of California at Riverside and Davis and Carnegie Mellon University. "Some interesting cyber defense research is under way in academia," says Matarazzo. "Partnering with academia allows us to access that knowledge." The SETAC team is also collaborating with experts at Sandia National Laboratories, Pacific Northwest National Laboratory, and Cisco, Inc.

SETAC is developing new approaches that contribute to national efforts such as the Comprehensive National Cyber Security Initiative, one of the largest single national security research and development investments in the U.S. today. "Livermore and the other national labs have much to offer the nation," says Brase. "We're expert at developing large-scale systems and monitoring how information flows."

In the meantime, researchers have instituted a test bed for sensor and system evaluation using controlled experiments with real cyber security data. "We don't have to make up a dangerous environment to test our system," says Matarazzo.

Because Livermore's network is not unusual in size or structure, a broad range of businesses and government agencies could soon benefit from SETAC cyber defense innovations. Within a few years, many large networks in the U.S. may well be armed with defenses pioneered at Livermore and aimed at thwarting a world full of clever "bad guys."

—Arnie Heller

Key Words: Comprehensive National Cyber Security Initiative, cyber security, hacker, Internet, malware, Supercomputing Enabled Transformational Analytics Capability (SETAC).

For further information contact Celeste Matarazzo (925) 423-9838 (matarazzo1@llnl.gov).