TRN's
Making the Future report
The State of an Emerging Technology and a Look at What Lies Ahead

Report Number 6 Summer, 2003

Security: Secrets, Intruders and ID Tags

Executive Summary

Security researchers are finding ways to protect computer data, hardware, and software from intruders. They are finding methods of tracking data, objects, and people. And they are exploring ways to protect privacy on the Internet and in the real world.

The technologies involved include cryptography, data hiding methods, data finding methods, information privacy, hardware security, system recovery, Internet security, property protection, intelligent tracking, and biometrics.

Security methods run the gamut from finding twists on encryption schemes, to gleaning secret information from patterns, to constructing cheap plastic tags that can be used to identify objects, to making microscopic glowing beads that can be used to hide information in a document.

Researchers are working to make encryption more secure and more convenient, and finding new ways to split keys and hide data.

At the same time, researchers are working on methods to crack codes. Practical DNA and quantum computers are 5 to 20 years away, but have the potential to solve the problems underlying most of today's security codes. Probabilistic security schemes, including quantum cryptography, however, would be immune to these fantastically fast computers.

Researchers are working to identify stealth computer intrusions and unknown viruses by tracking anomalous patterns. They are working to find ways to coordinate recovery from attacks, especially in large systems that run continuously, like transaction databases. And they are working on finding ways to keep the Internet secure.

Meanwhile, advances in the broad worlds of computing, engineering and materials science are beginning to blur the distinctions between the physical world and the digital. Automatic recognition systems can identify faces, sounds, and handwriting.

As access to the Internet becomes ubiquitous, wireless networks proliferate, and smart devices shrink to the microscopic level, technology research is opening up new forms of security, including biometrics and global tracking. At the same time, researchers are working to balance these new technologies by protecting privacy on the Internet and in the real world.

What to Look For

Cryptography:

Biometric encryption key generation
Public keys email addresses
DNA code cracking computer
Large-scale quantum computer

Quantum cryptography:

Quantum repeaters
High-bandwidth quantum cryptography
Satellite-based quantum cryptography

Data hiding:

Practical print-based data hiding
Practical text-based data hiding
Efficient hidden data detection and decryption

Systems and software:

Self-healing operating systems
Self-healing database systems
Practical anomaly-based intruder detection

Internet:

Practical Internet immunization scheme
Internet immune system
Internet hub protection

Privacy protection:

Of the good for the from the method Widely deployed
Anonymous Web access software
Censor-proof Web publishing

Physical security:

Copy-proof unique physical tokens
Wireless automatic personal authentication
Attention-based automatic camera and sensor targeting
Reliable automatic forgery detection

Knowing Who's Who

Security in the information age involves far more than simply locking the electronic equivalent of a front door. In addition to protecting the devices that make up your home base, you need to protect the words, graphics and sound files that make up your communications, business dealings, and even your identity as information speeds through cyberspace, diced into packets and scattered around the electronic ether.

A simple email message, for instance, encounters several points of vulnerability as it passes from your computer through the Internet to your recipient. The systems involved — your computer, its software, the Internet, and your recipient's computer and software — all must carry out the task you request without allowing an intruder to access your message.

In order to pull this off, the systems involved must successfully determine who to take instructions from — you, or an intruder who attempts to access your message.

Data, systems and objects

Security researchers are working toward three basic aims:

- Protecting data like email, medical records and credit card transactions where the data resides and also as it travels
- Keeping systems like the Internet, databases, and computer operating systems from being tampered with
- Making objects like computers and dollar bills tamper-resistant and copy-proof

Security research includes runs the gamut from finding twists on encryption schemes, to exploiting patterns to glean secret information, to constructing cheap plastic tags that can be used to identify objects, to making microscopic glowing beads that can be used to hide information.

This report takes a look at the following areas of security research:

- Cryptography
- Data hiding methods
- Data finding methods
- Information privacy
- Hardware security
- System page to recovery
- Internet security
- Property protection
- Intelligent tracking
- Biometrics

There are four basic aspects of security: hiding, finding, protecting and tracking. And a large part of security has to do with secrets.

How It Works

Two sides of cryptography

Cryptography is the cornerstone of electronic security. There are two basic cryptography methods: computational and probabilistic.

Computationally secure methods use cryptographic keys that are answers to difficult-to-solve problems. In order to find the answer to the problem, and thus the key, an eavesdropper would have to sift through a very large number of possibilities. If a problem is difficult enough that it would take a state-of-the-art computer millions of years to sift through all the possibilities, chances are the key is secure. Computationally secure cryptography methods are also referred to as conditional.

The drawback to this type of cryptography is that the difficulty of a problem can lessen as technology advances. A problem that would have taken millions of years to solve a decade ago might no longer provide viable security today.

Probabilistically secure methods use cryptographic keys chosen at random from a fast source of random signals. If a source is fast enough that an eavesdropper cannot record all of it, the probability of the eavesdropper gaining access to a key can be very small even given unlimited time for computations. This type of security is also referred to as information-theoretic, or unconditional, security.

The drawback to probabilistically secure cryptography is that it tends not to be as practical as computationally secure cryptography.

Computing security

Encryption schemes commonly in use today use public-private key encryption schemes, a computationally secure method based on one-way problems.

As the term suggests, such problems are easy to solve in one direction, but extremely difficult in the other direction. The RSA security algorithm that underlies many of today's schemes uses a mathematical problem similar to factoring. A number's factors are the numbers that can be multiplied together to get that number. The factors of 15, for instance, are 3 and 5. Given a pair of factors, it is a simple matter to multiply them to get a number. It is much more difficult to find factors for a given number, especially for very large numbers, simply because there are so many possibilities.

Public-private key encryption schemes use the difficult side of a problem — in the case of RSA, a very large number — as a public key, and the easy side of the problem — in the case of RSA, the number's pair of factors — as the private key.

Keeping secrets

There are two ways to keep a secret — make it cryptic enough so it can't be understood by anyone but its intended recipient, or hide it so it can't be found by anyone but its intended recipient. Do both, and your secret is even more secure. One challenge with either method, however, is finding a way to give only the intended recipient the information needed to reveal the secret.

Cryptography, the mathematical scrambling of words and numbers, makes up the foundation of electronic security. Keeping secrets by encrypting them gives people the means to control access to computers, messages, transactions, databases, and even the security systems employed to protect computers and information.

Researchers looking to protect access to data are finding ways of making encryption more secure and more convenient, and finding new ways to hide data. Researchers are also working on methods to crack codes.

Scrambling well

The classic cryptographic scheme involves scrambling data using a mathematical key. If only the recipient has the key and the key is used only once, the message is secure. The challenges are to make sure only the intended recipient gets the key, and to make the whole process easy enough that people are inclined to use it.

One way around the key distribution problem is using one type of key — a public key — to scramble a message, and another key — a private key — to unscramble it. Improving this type of security means making it more difficult to forge a public key; strong public keys make it more difficult for an eavesdropper to fool a sender into encrypting a message with the eavesdroppers key.

Researchers from Lucent Technologies' Bell Labs are tapping the individuality of the human voice to solve both challenges. This scheme uses numbers that model the vocal tract to construct a key. The user speaks a password, and the system looks for both the right password and the right voice speaking it. In the researchers' tests, they couldn't fool the system using recorded or synthesized speech. (See "Voiceprints Make Crypto Keys", page 12)

Researchers from the University of California and Stanford University have devised an encryption scheme that automatically generates a public key using information from the email address of the recipient. The scheme requires a central administrator to authenticate users and assign private keys, however. (See "Address Key Locks Email", page 14)

Researchers from the Naval Postgraduate School have tapped audio for security in a fairly unconventional way. This scheme, based on time-reversal acoustics signal processing, allows users to project an audio signal that can be clearly understood only at a given location. (See "Reverb Keeps Secrets Safe and Sound", page 13.)

When Alice wants to send Bob a message, she looks up his public key and uses it to scramble the message. The message can only be unscrambled using the private key, which is very difficult to figure out even though the public key is freely available.

These schemes can be enhanced by having the two parties frequently exchange new public/private keys, and by having the two parties use a shared secret to encrypt their public keys.

Security in probabilities

Probabilistic security relies on public sources of random signals and assumes that all parties have limited storage capacity so that they cannot record all of a source's random signals. Sources of random signals include radio broadcasts of noise and natural radio signals from space.

The scheme works by having two parties tune into a random signal and independently record portions of the signal at randomly selected intervals. The two compare notes to determine which portions they recorded in common and use these to form an encryption key. As long as the signal from the time of the first recording to the time of the last generates more data than an eavesdropper can store, and as long as the two parties share more than a few recordings, the odds against an eavesdropper finding the key are astronomical.

Quantum cryptography

Quantum cryptography works by a similar principle, but with a twist. Instead of storage capacity limiting an eavesdropper's ability to capture the random signal, quantum cryptography relies on the physical impossibility of accurately copying the quantum states of individual photons.

Photons can be polarized in one of two pairs of orientations — horizontal-vertical and the two diagonals — and each pair can be used to represent a 1or a 0. To measure the polarization of an electron, you must choose which pair of orientations to measure, and, because measuring a photon destroys it, if you guess wrong about which pair holds the information you don't get a chance to measure which of the other pair of orientations the photon was polarized in.

To use a quantum key, a sender transmits a random string of polarized photons and records the orientations. The recipient measures the photons, choosing orientations at random. The sender and recipient compare notes over a public channel, and the sender tells the recipient which orientations he guessed correctly about, and so correctly measured. The correctly measured photons form a string of bits that can be used as an encryption key.

Because measuring a photon destroys the information it contains, an eavesdropper would have

These schemes could be used in practical applications within a year.

Cracking codes

Encryption is more or less strong depending on how many possibilities you have to go through to try them all — whether it's cracking a cryptographic code or coming up with a 10-letter password. Passwords must contain a minimum number of characters and not be a dictionary word in order to increase the possibilities. Today's encryption schemes are only as secure as they are beyond the ability of today's fastest computers to simply go through all the possible combinations.

There are two emerging computer technologies that have the potential to render all of today's encryption software obsolete:

- Computers made from DNA
- Computers made from individual atoms and subatomic particles

The common denominator of these two emerging computing technologies is a kind of massive parallelism that is far greater in scale than the thousand-node multiprocessing supercomputers of today. In the case of DNA, it is theoretically possible to have trillions of individual DNA molecules processing a problem at the same time. In quantum computing, the quirks of quantum physics allow a single string of quantum bits to represent every possible number and for the computer to check every answer to a problem with one set of operations.

DNA computers

Researchers have proved that is possible to compute using DNA, the molecule all biological organisms use to store instructions for coding proteins. Once a DNA computer is set up, DNA molecules carry out computations quickly and automatically.

Researchers from Ruhr University in Germany and Accenture Technology Labs in France, have shown that it is possible to do a type of crude programming that would cut down on the number of possibilities DNA would have to go through in order to solve very large problems like cracking encryption codes. In theory, the design could be used to break a strong public key encryption code in a couple of days rather than the 3,000 years it would take an electronic computer. (See "DNA Could Crack Code", page 15)

Practical DNA computers are about five years away.

Quantum computers

Many researchers are working to find ways to construct quantum computers, which use attributes of particles like atoms, electrons and photons to compute.

Like DNA, computers made from particles could check all possible answers to a problem at once rather than one after to replace the photons she measured, and because she would only be able to correctly measure about half of the photons, she would have to randomly replace the other half. This would mean that an average of 25 percent of the forged photons measured by the recipient would fail to match up with the sender's record. If the sender and recipient notice an unusually high error rate when they compare notes, they can assume the key has been compromised and they can discard it and try again.

Boosting privacy

Probabilistic cryptography schemes are theoretically perfect, but in practice, imperfect equipment, noisy communications lines and human error create vulnerabilities, and users have to assume that an eavesdropper could pick up at least some of their communications.

One way to reduce the risk is to use privacy amplification, which distills partially secure shared information into a smaller amount of more highly secured information. Mathematical formulas, called one-way hash functions, turn two or more of the original bits into a single new bit. Even if an eavesdropper knows some of the original bits, she doesn't have enough information to calculate the new bits. The drawback is that this requires sending more bits to generate a key.

Who to Watch

Cryptography

Mihir Bellare, University of California, San Diego
La Jolla, California
charlotte.ucsd.edu/users/mihir

Dan Boneh, Stanford University
Stanford, California
theory.stanford.edu/~dabo

Peter Gutmann, University of Auckland
Auckland, New Zealand
www.cs.auckland.ac.nz/~pgut001/

Ueli Maurer, ETH Zürich
Zürich, Switzerland
www.crypto.ethz.ch/~maurer

Christof Paar, Ruhr-University Bochum
Bochum, Germany
www.crypto.ruhr-uni-bochum.de/MitarbeiterInnen/paar_eng.html

Adi Shamir, The Weizmann Institute of Science
Rehovot, Israel
www.wisdom.weizmann.ac.il/~oded/wis-cry.html

another. This is made possible by weird quantum traits that make the world of the very small act differently than we are used to at the macro scale. The disadvantage of working with particles is that they are very difficult to handle and control.

Most researchers agree that it will take a couple of decades to build a quantum computer powerful enough to break today's encryption schemes.

Quantum Cryptography

Another line of quantum research, however, is poised to make encryption keys perfectly secure.

Quantum cryptography uses attributes of photons to transmit an encryption key. Photons can be polarized in one of four orientations and can represent digital information by making two of the orientations represent 0 and the other two represent 1. (See TRN's Making the Future report Quantum Cryptography: Potentially Perfect Security)

If each bit of a key is represented by more than one photon, an eavesdropper may be able to siphon off some of the photons from each bit to gain a copy of the key without being detected. If each bit is represented by only one photon, however, it is not possible to look at the information without changing it.

If Alice, for instance, sends Bob a key whose bits are each represented by single photons, they can tell for sure that the key arrived without being looked at. If it has been compromised, they can simply throw away the key and exchange a new one.

Quantum cryptography is a form of probabilistic, or information theoretic, cryptography, in contrast to traditional cryptographic methods that are based on the difficulty of solving certain math problems. (See How It Works, page 2.)

Probabilistic cryptography is theoretically perfect, but its usefulness comes down to how practical it is to implement systems that are close enough to perfect. For quantum cryptography, the goals are the development of room temperature electronic sources of single photons and efficient photon detectors.

Research that promises to improve the systems includes devices that will send and receive single photons, and schemes to transmit keys more quickly and over longer distances. Current prototype systems are limited to a few kilobits per second and tens of kilometers.

Researchers from the French National Scientific Research Center (CNRS) and Ecole Polytechnic in France have made a single photon source from a deliberately contaminated microscopic diamond and tested the prototype in a quantum cryptography set up. (See "Diamonds improve Quantum Crypto", page 19)

Researchers from the University of Geneva in Switzerland and the Swiss company id Quantique have demonstrated a system capable of sending messages over existing fiber-optic phone lines from Geneva to Lausanne, which are 67 kilometers apart. (See "Quantum Secrets Ride Phone Lines", page 20)

Data Hiding

Charles Boncelet, University of Delaware
Newark, Delaware
www.eecis.udel.edu/~boncelet

Jana Dittmann, Otto-von-Guericke-University
Magdeburg, Germany
www.darmstadt.gmd.de/~dittmann

Jessica Fridrich, Binghamton University
Binghamton, New York
www.ws.binghamton.edu/fridrich/fridrich.html

Neil F. Johnson, George Mason University
Fairfax, Virginia
www.jjtc.com/neil/research.html

Stefan Katzenbeisser, Technical University of Munich
Munich, Germany
wwwbrauer.in.tum.de/~katzenbe

Bede Liu, Princeton University
Princeton, New Jersey
www.ee.princeton.edu/bios/liubio.html

System and Software Security

Matt Bishop, University of California, Davis
Davis, California
nob.cs.ucdavis.edu/~bishop

Edward W. Felten, Princeton University
Princeton, New Jersey
www.cs.princeton.edu/~felten

Sushil Jajodia, George Mason University
Fairfax, Virginia
www.ise.gmu.edu/~csis/faculty/jajodia.html

Karl N. Levitt, University of California, Davis
Davis, California
seclab.cs.ucdavis.edu

R. Sekar, SUNY Stony Brook
Stony Brook, New York
seclab.cs.sunysb.edu/sekar

David Wagner, University of California, Berkeley
Berkeley, California
www.cs.berkeley.edu/~daw

Dan Wallach, Rice University
Houston, Texas
www.cs.rice.edu/~dwallach

Privacy, Policies and General Security

Ross Anderson, University of Cambridge
Cambridge, England
www.cl.cam.ac.uk/~rja14

And researchers from the University of Munich in Germany have made a laser diode-based source of entangled photons. This opens away to building quantum cryptographic devices on circuit boards. (See "Quantum Crypto Gear Shrinks", page 21)

Researchers from the Institute of Optics in France and the Free University of Brussels in Belgium have proposed a quantum cryptography method that sidesteps the need to send just single photons to keep a message perfectly secure. It uses the quantum nature of light's amplitude and phase rather than the usual polarization information of a single photon. The multi-photon scheme also allows information to be transmitted about an order of magnitude faster than current single-photon methods. (See "Faster Quantum Crypto Demoed", page 16.)

Researchers have also shown that quantum information can be tapped as a scrambling mechanism.

Northwestern University scientists are using the quantum noise generated by lasers as a means to encrypt data. (See "Fast Quantum Crypto Demoed", page 18)

The Swiss quantum cryptography method is ready for use now. The other developments could be ready for commercial products in two to five years. (For more information about these types of systems, see the TRN Making the Future report entitled "Quantum Cryptography: Potentially Perfect Security.)

Rudimentary quantum cryptography systems are commercially available today, but the systems still use a few photons per bit rather than just one, and are only available over relatively short distances. It is likely to take five to ten years before the technology is practical enough for widespread use.

Annie I. Antón, North Carolina State University
Raleigh, North Carolina
www.csc.ncsu.edu/faculty/anton

Jan Camenisch, IBM Research
Zürich, Switzerland
www.zurich.ibm.com/~jca

Dorothy Denning, Naval Postgraduate School
Monterey, California
www.nps.navy.mil/ctiw/staff/denning.html

Joan Feigenbaum, Yale
New Haven, Connecticut
www.cs.yale.edu/homes/jf/home.html

Anita Jones, University of Virginia
Charlottesville, Virginia
www.cs.virginia.edu/~jones

Peter G. Neumann, SRI International
Menlo Park, California
www.csl.sri.com/users/neumann

Andrew Odlyzko, University of Minnesota
Minneapolis/St. Paul, Minnesota
www.dtc.umn.edu/~odlyzko

Ronald L. Rivest, Massachusetts Institute of Technology
Cambridge, Massachusetts
theory.lcs.mit.edu/~rivest

Aviel D. Rubin, Johns Hopkins University
Baltimore, Maryland
www.cs.jhu.edu/~rubin

Eugene H. Spafford, Purdue Unversity
West Lafayette, Indiana
www.cerias.purdue.edu/homes/spaf

Doug Tygar, University of California, Berkeley
Berkeley, California
www.cs.berkeley.edu/~tygar

Splitting up the secret

One way to make it impossible for any one person — no matter how trusted — to compromise security, is to split a key into pieces. Secret-sharing schemes require that all or some percentage of the pieces be brought together to form the key.

Scientists from IBM Research have shown that adding a quantum component to secret-sharing cryptographic protocols would make it harder for the people holding the pieces to cheat or be coerced into revealing the secret. (See "Quantum Code Splits Secrets", page 23)

Hiding data in plain sight

Another way to make data secure is to hide it.

One common scheme, dubbed steganography, is to hide information within a digitized picture. Many images contain more than one million bits, and colors can be altered so slightly that the human eye cannot detect the changes.

Researchers from Ben Gurian University in Israel have taken the idea a step further with a way to hide an image within a printed picture. The hidden image can represent encoded information; this allows both digital and hard copies of an image to retain hidden information. (See "Printed Pictures Hide Images", page 22.)

Researchers from Perdue University have devised a way of hiding information within text using word substitution and syntactical changes. (See "Watermarks Hide in Plain Text", page 24.)

Finding hidden data

The flip side of hiding is finding, and researchers are working on this as well. A researcher from Dartmouth College has devised a method that finds messages hidden in digital images by comparing the statistical profile of a compressed file with similar compressed profiles from a library of images that have not been tampered with. (See "Fast Quantum Crypto Demoed", page 18.)

These data-hiding and data-finding methods could be used in commercial products within one or two years.

Securing systems

As important as keeping data secure is securing the systems that store, manipulate and send data.

One major challenge to securing systems is guarding against the ill effects of unauthorized access by humans and by computer viruses — small bits of computer code that can propagate wherever there is contact between computers.

A large part of making a system secure is being able to tell who should be granted access to a system and who should not. Software exists today that blocks intrusions, and that senses and eradicates known viruses.

Researchers are working on software that identifies stealth intrusions and unknown viruses by tracking anomalous patterns. Anomaly detection systems sense when normal patterns of communication change as a way to identify viruses, worms and unauthorized users, but have historically had impractically high false positive rates.

Researchers from the University of California at Davis have built a system that achieves lower error rates by employing a pattern-recognizing technique that is usually used to classify text. (See "Text Software Spots Intruders", page 28.)

Researchers from the University of South Carolina are working on a similar system that taps methods used to tease data from nuclear experiments. (See "Physics Methods May Spot Intruders", page 31.)

Recovering from an attack

Another challenge is coordinating recovery from attacks, especially in large systems that run continuously, like transaction databases.

A team from Pennsylvania State University has written software that keeps a database running during the recovery process. The software monitors the database in real-time, maintains especially detailed log files, detects intrusions, contains any damage, and repairs each corrupted data object by restoring its most recent undamaged backup copy. (See "Software System Heals Itself", page 26.)

Another aspect of database security is making sure users are getting exactly what they have requested. Researchers from the University of California at Davis and Stubblebine Consulting have come up with a scheme that allows users to ensure that documents retrieved from an Extensible Markup Language (XML) database are authentic. XML is widely used on the Internet. (See "Data Protected on Unlocked Web Sites", page 29.)

These systems could be ready for commercial use within three years.

Securing the 'Net

Keeping systems secure also means securing the networks that connect them, including the Internet. For several years now, researchers have studied the Internet with growing interest; the more that is known about the way networks grow, the better the growth process can be managed.

The Internet makes for an especially interesting study because it is so big and because its growth is not centralized. This makes it similar in some ways to biological networks like the web of interactions among chemicals used in the body.

The Internet is really two networks — the physical servers and communications lines that connect one computer to another, and the network of links that connect Web pages to each other. Researchers have shown that both parts of the network are scale-free, meaning they contain a few highly connected nodes and many that have few connections.

Researchers from Ohio State University have pointed out that although the physical structure of the Internet was originally distributed because it was designed to withstand failure and provide service under adverse conditions, the structure has lately grown to resemble the hub-and-spoke topology used by the airlines. The Internet's big hub cities include Los Angeles, New York City, Atlanta, Dallas and Chicago. This type of structure is less costly to build, but makes the network more vulnerable to attack. (See "Hubs Increase Net Risk", page 32.)

Researchers from Clarkson University and Bar-Ilan University in Israel have showed mathematically that when five percent of large hubs are methodically targeted, even large, scale-free networks like the Internet can be broken up into separate islands. (See "Five Percent of Nodes Keep Net Together", page 35.)

The Internet's hubs also make it more vulnerable to attack from viruses and worms. Viruses attach themselves to or replace existing software. Worms, which are less common, are self-contained programs.

A pair of physicists from the University of Notre Dame and the Polytechnic University of Catalonia in Spain have applied condensed-matter physics, which examines the collective behavior of matter, to map the ways viruses traverse Internet's labyrinth of connections.

The researchers showed that the Internet has become more vulnerable to software viruses in much the same way that human populations that are large and crowded are more likely to fall prey to biological viruses. The Internet is even more vulnerable, however, because connections among computers tend to be more numerous than the human connections that allow biological viruses to spread. The researchers showed that the random inoculation strategy employed for human epidemics does not work on the Internet, but a strategy that targets large hubs does work. (See "Net Inherently Virus Prone", page 36; "Hubs Key to Net Viruses", page 33.)

Exploiting the Internet

Researchers are also looking into ways the Internet can be co-opted in order to keep people from misusing it.

Researchers from the University of Notre Dame have shown that it is possible to exploit a basic operation carried out by all servers connected to the Internet to steal computer power from the computers. This type of attack, dubbed parasitic computing, is similar to distributed computing schemes like the SETI at Home project because it uses processing power from computers connected to the Internet, though unlike SETI it took processing power without the computer owners' knowledge or consent. (See "Scheme Harnesses Internet Handshakes", page 34.)

Privacy

One aspect of protecting individual security has to do with being free to go where you want without anyone watching. The digital age has spawned a world town square in the form of the Internet, which in turn has raised the issue of being followed in cyberspace — whether by a stalker, an oppressive government, or a commercial data mining operation.

Researchers from the Massachusetts Institute of Technology have found a way to guarantee that users can access data on the Internet in such a way that their actions cannot be monitored. (See "Scheme Hides Web Access", page 38.)

Meanwhile, researchers from the Virginia Polytechnic Institute and State University, Purdue University and the University of Minnesota have found that the information people often unconsciously reveal about their associations when participating in recommendation systems can be used to trace individuals. (See "Rating Systems Put Privacy at Risk", page 40.)

Another aspect of individual security is free speech. Researchers from AT&T Labs and New York University have devised a scheme that uses a shared encryption key in order to allow Web publishing that is both anonymous and difficult to remove. (See "Fault-tolerant Free Speech", page 42.)

These methods can be used now.

Protecting property

Security in the context of technology usually means protecting electronic data and communications, but technology also plays a role in physical security, ranging from anti-theft measures to authentication to surveillance.

Radio frequency identification tags have garnered a lot of attention lately because manufacturers and retailers are poised to use the devices to track inventory from the factory to the consumer. The small devices resonate at radio frequencies and transmit unique identification numbers when tag readers hit them with radio waves. Advances in plastic circuits could lead to RFID tags that cost pennies or less.

Advances in materials science and engineering have opened a wide range of other possibilities for tracking objects, as well. There has also been a burst of activity in using the optical properties of materials structured at the micro- or even nanoscale to make counterfeit-proof identification tags.

Researchers from Corning, Inc. have formed bar coded beads that are barely visible — small enough to be embedded in ink in order to tag currency and other documents to protect against counterfeiting. (See "Glowing Beads Make Tiny Bar Codes", page 42.)

Massachusetts Institute of Technology researchers have devised a way to make inexpensive, unique ID tags that can't be copied. The tags are made from small pieces of plastic that contain tiny glass spheres that produce unique patterns of light when lasers shine through them. (See "Plastic Tag Makes Foolproof ID", page 43.)

In an age where important data often resides on portable machines is important to make sure data cannot be stolen simply by taking the machine that contains it.

Researchers from the University of Michigan have combined a pair of well-known security techniques to better secure data on a laptop. The scheme calls for a user to enter a password into the laptop at the start of the day. The user also wears a token, which can be embedded in a watch or piece of jewelry. The token and computer communicate via encrypted radio signals. Files remain unlocked until the computer is turned off, but only as long as the token remains within a few feet of the computer. (See "Radio ID Locks Lost Laptops", page 45.)

The glass beads could be ready for commercial use in three to six years, the ID tags in one year, and the token scheme within four years.

High visibility

Technological advances are also allowing people to be more easily tracked and monitored. Security cameras, motion detectors and other sensors are becoming more intelligent, which makes them more efficient and also opens the possibility of implementing privacy protocols that would limit their reach.

University of Illinois researchers have made a camera system that copies the way a barn owl hunts. The system homes in on the location most likely to contain whatever is making the most interesting noise at any given time. (See "Sounds Attract Camera", page 46.)

Researchers at Duke University have devised a simple tracking method that dramatically reduces the computing resources needed for computer vision tracking systems. The method maps the angles of light radiating from a source through a set of pipes and onto a set of light detectors. As an object moves across the field of view, light reflecting from the object triggers some detectors but not others. (See "Light pipes track motion", page 47.)

Researchers from the University of Colorado at Boulder have found a way to use networks of tiny sensors to gain useful traffic statistics but at the same time preserve privacy to cloaking local information for any given individual. (See "Sensors guard privacy", page 37.)

Biometrics

Researchers are employing systems that automatically recognize patterns to identify things like faces, sounds, and handwriting. These systems include cameras that collect data and use a considerable amount of processing power to sift through and identify salient data.

Researchers from the Mayo Clinic and Honeywell Laboratories have devised a way to track heat changes in a person's face in order to tell whether the person is lying. According to tests the researchers carried out at the U.S. Department of Defense Polygraph Institute, the method is as accurate as a polygraph. (See "Hot Spots Give away Lying Eyes", page 46.)

Researchers from the University of Buffalo are working on a pattern-recognizing system designed to identify people using samples of their handwriting. (See "Software Spots Forged Signatures", page 48.)

Perpetually pursuing perfect security

Security research is the stage where the age-old struggle between those trying to protect information and those trying to thwart these protections is played out.

The cryptography arms race, in particular, is poised for a significant shift in the balance of power because theoretically perfect security schemes based on the laws of nature like quantum physics and chaos are beginning to become practical. Code makers have generally been able to stay ahead of code breakers. The arrival of practical quantum cryptography in the next five to ten years could cement that lead for decades to come.

Meanwhile, advances in the broad worlds of computing, engineering and materials science are beginning to blur the distinctions between the physical world and the digital. Computers are increasingly able to see, hear and feel the world

around them. And wireless communications, radio frequency ID tags and microscopic bar codes are making people and objects move visible to computers.

This visibility, combined with the ever-increasing power of computers to carry out large-scale data sorting, pattern recognition and statistical analysis, is dramatically shrinking privacy.

In general, the high-tech realm is advancing both sides of the timeless cat-and-mouse games of intruder detection and secret messages. As access to the Internet becomes ubiquitous, wireless networks proliferate, and smart devices shrink to the microscopic level, technology research is opening up new forms of security, including biometrics and global tracking. This technology is spawning a new cat-and-mouse game between those who want to track information and individuals, and those who seek to maintain privacy.

Recent Key Developments

Advances in cryptography:

- A system that combines voice recognition and passwords to generate cryptographic keys (Voiceprints Make Crypto Keys, page 12)
- A method of targeting sound wave communications to a particular location (Reverb Keeps Secrets Safe and Sound, page 13)
- A scheme to use email addresses as public keys for public key encryption (Address Key Locks Email, page 14)
- A scheme for programming DNA computers to crack an encryption code (DNA Could Crack Code, page 15)
- A random number generator that is based on shooting a laser through liquid crystal, Kent State University, September 2002
- A scheme for combining multiple random sources to reduce the risks from compromised sources, IBM Research, January 2002

Advances in quantum cryptography:

- A scheme for entangling photons to allow multiple photons per bit, University of Oxford and University of California at Santa Barbara, May 2002
- A scheme to carry out entangling laser beams to provide continuous-variable cryptography, University of Erlangen-Nuremberg, April 2002
- A scheme for continuous-variable cryptography using weak laser beams (Faster Quantum Crypto Demoed, page 16)
- A method of encrypting data using the quantum nature of laser beams (Fast Quantum Crypto Demoed, page 18)
- A quantum cryptographic system that uses a solid-state, room-temperature single-photon source (Diamonds Improve Quantum Crypto, page 19)
- A quantum cryptographic system demonstration over ordinary phone lines between Swiss cities (Quantum Secrets Ride Phone Lines, page 20)
- A small laser diode that generates entangled photons (Quantum Crypto Gear Shrinks, page 21)

Advances in data hiding:

- A method for embedding 3D holographic watermarks in 3D holograms, University of Connecticut, April 2003
- A method for extracting watermarks from digital images to recover the original images, University of Rochester and Xerox Corporation, September 2002
- A method for synchronizing digital video to preserve watermarks, Purdue University, January 2002
- A method for hiding data in printed pictures (Printed Pictures Hide Images, page 22)
- A scheme for splitting encryption keys using quantum communications (Quantum Code Splits Secrets, page 23)
- A method for embedding secret data in text (Watermarks Hide in Plain Text, page 24)
- A method for detecting hidden data using statistical analysis (Statistics Sniff out Secrets, page 25)

Advances in systems and software defenses:

- A method for combining multiple timeline records to assure a continuous history of system operations, Stanford University, Aug. 2002
- A method for containing and repairing damage to database systems without taking the systems off-line (Software System Heals Itself, page 26)
- A method of detecting intruders that is based on text classification software (Text Software Spots Intruders, page 28)
- A method of detecting intruders that is based on based on nuclear physics analysis (Physics Methods May Spots Intruders, page 31)
- A method of securing XML documents on unsecure servers (Data Protected on Unlocked Web Sites, page 29)

Advances in Internet vulnerabilities:

- An analysis that shows that the evolving hub structure of the Internet puts it at risk (Hubs Increase Net Risk, page 32)
- An Internet virus inoculation scheme that targets large nodes (Hubs Key to Net Viruses, page 33)
- An analysis that shows that the largest five percent of Internet nodes keep the net together (Five Percent of Nodes Keep Net Together, page 35)
- An analysis that shows that the Internet's structure makes it prone to viruses (Net Inherently Virus Prone, page 36)
- A scheme showing it is possible to steal computer processor time using standard Internet protocols (Scheme Harnesses Internet Handshakes, page 34)

Advances in privacy protection:

- A scheme for programming traffic-monitoring sensor networks that cloaks locations of specific individuals (Sensors Guard Privacy, page 37)
- A method for providing anonymous Web access (Scheme Hides Web Access, page 39)
- A vulnerability in online rating systems that could be used to identify users (Rating Systems Put Privacy at Risk, page 40)
- A method for protecting online publications by encrypting them and breaking them into pieces spread across multiple servers (Fault-Tolerant Free Speech, page 42)

Advances in physical security:

- Tiny glass fluorescent barcode beads that can be mixed into paints and inks (Glowing Beads Make Tiny Bar Codes, page 42)
- Tamper-proof plastic tags whose microscopic structures scatter light in uniquely identifiable ways (Plastic Tag Makes Foolproof ID, page 43)
- A system that locks access to laptop files when a user's wireless token moves out of range of the computer (Radio ID Locks Lost Laptops, page 45)
- A lie detection technique based on heat sensors measuring blood flow in the face (Hot Spots Give Away Lying Eyes, page 46)
- An inexpensive object tracking sensor that uses photodetectors and light angle calculations (Light Pipes Track Motion, page 47)
- A camera that aims at a sound's source (Sounds Attract Camera, page 46)
- Automatic signature forgery detection software (Software Spots Forged Signatures, page 48)

Cryptography

Voiceprints Make Crypto Keys

By Kimberly Patch, Technology Research News
October 16/23, 2002

As we rely on computers for tasks like handling money and keeping secrets safe, it has become increasingly important to give our desktops, laptops and PDAs the means to know for sure who they are dealing with. The classic solution is to lock up the data, and give the user a cryptographic key.

The main challenge to improving this type of security is to make it more difficult to steal or reconstruct the keys, but easier for legitimate users to access computing resources.

Researchers from Lucent Technologies' Bell Labs have tapped the individuality of the human voice to generate unique cryptographic keys for computer users. Under the researchers' scheme, a user speaks a password, and the system listens for both the correct word and the correct voice.

The method uses the random variability of people's voices to add a layer of security to even a simple password, said Fabian Monrose, a member of technical staff at Bell Labs. "The randomness of [a] key is drawn from both the pass-phrase that is spoken and the speech patterns of the user... speaking it," he said. The more randomness contained in the information the key is constructed from, the harder the key is to figure out.

The scheme uses cepstral coefficients, which are numbers that model the vocal tract, to help construct the key. These coefficients are also commonly used in speech and speaker recognition software. They are robust, meaning they contain a lot of information, and reliable, meaning they are fairly consistent for a single speaker, but vary across a population.

The researchers' prototype software, which runs on a Compaq Ipaq PDA, uses 60 different features from a given voice sample to form a mathematical descriptor, then uses the descriptor to construct cryptographic keys and verify whether keys generated by users are correct. "We've been... generating 60-bit keys from a few seconds of speech," said Monrose. "Our studies suggest that the techniques... enable significant randomness from pass-phrase utterances."

Because there is variability even in the way a single user says a password, the method allows for some legitimate errors in the 60 parameters used. These errors are due to background noise or changes in vocalization. "Since the biometric readings are hardly exact across successive measurements, we typically need to correct... five errors on average for the legitimate user," Monrose said. "An adversary speaking the password, however, will cause a far greater number of errors," he said.

The scheme also includes software to protect reverse-engineering of the key in the event that the device being protected is captured, said Monrose. Information about the scheme stored on the device is protected using a secret-sharing scheme, which divides a secret into two or more pieces. The secret is revealed only when the pieces are combined. "The key is regenerated from scratch in each reconstruction attempt, and no speaker-specific information is stored in the clear," said Monrose.

To make the prototype work, the researchers needed to make sure that the system did a good job of processing the user's speech in order to minimize error correction for the legitimate user, and they had to devise secret-sharing schemes and reconstruction algorithms that allowed the system to recognize a legitimate user in a reasonable amount of time, according to Monrose. "The challenge... is to find the right balance of eliminating environmental effects early via signal processing versus relying on the error correction in the key generation step to compensate for the effects of noise and silence that may occur in the user's utterance," he said.

The researchers' attempts to fool the system using recorded and synthesized speech did not work, said Monrose. "Cut-and-paste attacks of a user's speech, and text-to-speech attacks... did not significantly outperform random guessing," he said.

This could change as speech synthesis and audio sorting tools get better, however, Monrose said. As advances are made in speech synthesis and in tools for automatically finding phonemes in an utterance, these types of attacks will become more successful, he said. "We're actively exploring effective countermeasures against such attacks," he added.

The work is an efficient way to use a natural user interaction to provide personal information security, said Philip Robinson, a researcher at the University of Karlsruhe in Germany.

Although biometric techniques like speech or fingerprints are readily available and therefore easy to use, there is a potential downside — you can't change speech and fingerprints if the security is compromised. The researchers' method addresses this problem, said Robinson. The randomness associated with a spoken password is increased by basing the key regeneration process on the variation in a user's speech pattern, he said.

Finding novel ways of facilitating usability while maintaining strong security is a major underlying theme in ubiquitous computing security research, Robinson added.

The researchers' prototype has proved the plan plausible, but does not achieve especially strong security. Their next step is to strengthen the method, said Monrose. "Our immediate goals are more extensive user trials, which will involve analyzing ways to increase the strength of the derived keys," he said. The researchers are aiming to achieve key lengths of 80 bits or longer, he said. The strength of the cryptography programs used by today's business community generally range from 128 to 8,192 bits.

It will take a couple of years for the researchers to determine if the scheme is capable of generating strong cryptographic keys that can be used in commercial applications, said Monrose.

Monrose's research colleagues were Michael K. Reiter of Carnegie Mellon University, and Qi Li, Daniel P. Lopresti and Chilin Shih of Bell Labs. They published the research in the Proceedings of the 11th Usenix Security Symposium, which was held August 5-9, 2002 in San Francisco. The research was funded by Bell Labs.

Timeline: Unknown
Funding: Corporate
TRN Categories: Cryptography and Security; Computer Science; Human-Computer Interaction
Story Type: News
Related Elements: Technical paper, "Toward Speech-Generated Cryptographic Keys on Resource Constrained Devices," Proceedings of the 11th Usenix Security Symposium, August 5-9, 2002 in San Francisco.

TRN

Reverb Keeps Secrets Safe and Sound

By Kimberly Patch, Technology Research News
May 29/June 5, 2002

Encryption usually means disguising data using a numerical formula. Researchers from the Naval Postgraduate School have come up with a scheme for encrypting sound that protects the information by taking advantage of the way sound waves propagate.

The scheme focuses a clear audio signal at only one point in space, making it impossible to listen in from any other point. "It is possible to create a signal that will focus in a unique location," said Kevin Smith, an associate professor of physics at the Naval Postgraduate School. At locations other than the focus, the various sound waves that make up an audio signal arrive at different times, producing interference instead of a coordinated signal.

The technique can also be used to improve sound quality in any room, including home theaters, according to Andres Larraza, a physics professor at the Naval Postgraduate School.

The scheme is based on time-reversal acoustics signal processing, which can be used to transmit audio clearly in environments — like rooms with cement walls or underwater — where reflections of sound off different surfaces at different times cause reverberation or echo.

Audio signals get garbled when sound waves diverge and overlap, causing the listener to hear the same sound at several different times. "A simple pulse distorts and spreads in time due to the variety of paths between source and receiver," said Smith.

A good demonstration is to clap inside the type of tubular slide commonly found in playgrounds, said Larraza. "The apparent long duration of the clap is due to all the different propagation paths [from] multiple reflections of sound inside the slide," he said.

The overlap makes it difficult to distinguish different sounds, said Smith. "When a sequence of symbols is transmitted... this multipath propagation may cause the various symbols to overlap, degrading the ability of the receiver to distinguish the information," he said.

Time reversal acoustics fixes the problem by transmitting sound to a point and noting exactly when the parts of the signal — or different paths — arrive, then transmitting the same signal with the arrival times reversed, said Larraza. "This allows the slowest path a head-start and the fastest path brings up the rear." This way the different sound paths arrive back at the destination simultaneously, bringing the sound back into focus.

Another way of looking at the scheme is as an encryption method for all the points except for those where the sound is in focus. The inherent multipath environment of water "provides for a natural encryption whereby only a single location can receive the unscrambled message," said Smith.

The researchers realized that the reverberant environment of an enclosure could be used in a similar way as a natural encryption method. The encryption code is inherent in the structure of the enclosure, and each enclosure provides a unique type of scramble. Time-reversal acoustics encryption is "always unique to the environment and independent of [a] signaling scheme," said Smith.

To demonstrate the method, the researchers used a speaker as a sound source and two microphones to record at different positions in a concrete chamber 2.59 meters long, 2.41 meters wide and 2.83 meters high.

This type of chamber causes a lot of reverberation, said Larraza. If a few notes of Beethoven's Fifth Symphony were piped in, for instance "the first note... would still be playing into the last quarter of the fourth note, overlapping all along with the other two notes in between, resulting in cacophony," he said.

Applying time-reversal acoustics to the first note in Beethoven's Fifth would allow it to play its intended duration at the focal location, while anywhere else in the room it would reverberate longer, said Larraza.

The researchers tested their time-reversal based encryption algorithm by sending sets of signals representing binary bits — the ones and zeros of digital communications. They transmitted each signal with enough time between transmissions for all the multipath signals to arrive at each receiver, then used the time records for each receiver to build symbols out of the time-reversed signals, he said. "In a room with complex geometry, the time record of each reception measured by each receiver is unique."

In the researchers' experiment, the source transmitted simultaneously to each of the microphones its own unique message. "This would be equivalent to applying from the same source Beethoven's Fifth and the Beatles "All You Need

Is Love" and being able to listen... to each one in their unique.... location; at the Beethoven spot, "All You Need Is Love" [would] not be playing. Anywhere else in the room there [would be] noise," said Larraza.

The method could be combined with traditional encryption for added security, said Smith. It may also prove useful for enhancing sound quality in home theaters and concert halls, said Larraza.

The research is interesting, novel and potentially useful, according to Manuel Torres, a researcher from the Superior Council of Scientific Research in Spain, and Jose-Luis Aragon, a researcher at the National University of Mexico. "The application of this technique to encrypt some messages is a clever and original idea," said Torres.

The technique also looks promising as a method to enhance sound quality in buildings, Torres said. "The ability to focus a full message in space and time... and simultaneously send multiple messages from one source to different locations in [an] enclosure makes the technique potentially applicable in architectural acoustics."

In addition, because these messages are destroyed if they are intercepted before they reach their destination points, the method could find application in military underwater communications, said Torres.

In general, waves of any kind are a potentially powerful alternative to numerical encryption techniques, Torres added. Waves are the result of periodic disturbances in any medium or in space. Sound waves, for instance, result from vibrations in elastic media, including air, water or the earth, that can be sensed by the human ear.

The technique could be applied immediately, according to Larraza. "Using time reversal acoustics as a diagnostic tool for enhancing sound quality is technologically plausible at this time," he said.

Larraza's and Smith's research colleague was Michael G. Heinemann. They published the research in the January 28, 2002 issue of Applied Physics Letters. The research was funded by the Office of Naval Research (ONR).

Timeline: Now
Funding: Government
TRN Categories: Applied Technology; Physics
Story Type: News
Related Elements: Technical paper, "Acoustic Communications in an Enclosure Using Single-Channel Time-Reversal Acoustics," Applied Physics Letters, January 28, 2002.

TRN

Address Key Locks Email

By Chhavi Sachdev, Technology Research News
October 31, 2001

When you pay for a book online or check stock quotes from your mobile phone, your password and credit card number are kept secure by an encryption scheme; one of the most widely used ways to spy-proof transactions is to use encryption keys.

In this type of encryption, each party has two keys: one to lock, or encrypt, messages and the other to unlock, or decrypt, them. If I wanted to send you a confidential message, I would look up your public key, use it to encrypt the message and then send my message to you. The only way to decipher the coded message would be your private decryption key.

Looking up a public key takes time and requires the receiver to first set one up. A pair of researchers has made the process easier with a scheme that automatically generates public keys using something most people have already made publicly available: an email address.

Using a person's unique email address as a public key makes it possible to send encrypted messages without having to look anything up.

Common encryption schemes like RSA can also use names to generate public keys, but not everyone can get a key based on a name because only one John Smith can use the John Smith name key; also, like getting a phone number that spells your name on a phone key pad, there is a certain amount of overlap. Using unique email addresses solves this problem.

"What we have tried to do is to create a new public key encryption scheme... designed so that... every user will get a valid key," said Matthew Franklin, an acting associate professor of computer science at the University of California at Davis.

All public key algorithms are based on difficult mathematical problems, said Franklin. The security of RSA, for instance, depends on a mathematical problem that is closely related to factoring large numbers. Two factors multiply together to produce a number. For example, 3 and 5 are factors of 15. Finding the particular factors of a really large number is very difficult because there are so many possibilities. RSA uses the large number as the public key and the two factors make up the private key.

The researchers' algorithm uses mathematics based on the Weil Pairing, a mathematical function that takes as input two points on an elliptical curve. Although the mathematics is different, "the speed of encryption and decryption and [the] size of keys and ciphertexts for our scheme is comparable to... popular public key encryption schemes such as RSA and ElGamal," said Franklin.

To send an encrypted email message, the sender would use an email program that incorporated the encryption scheme

and could automatically generate the public key using the email address of the recipient, said Franklin.

The system's drawback is that it requires a central administrator who authenticates users and assigns private keys, he said. "When the recipient gets the encrypted email, she won't be able to decrypt it until she registers with the proper authorities to get her private decryption key," said Franklin.

Registering is a one-time burden for the recipient. "Once she has her private decryption key installed in her mail program, she can read any encrypted email that comes to her from any sender," he said.

The catch to having a central administrator is that someone would be privy to all encrypted email. The master key, however, could be split among several parties. "The functionality of the master key can be split among many parties — geographically distant, mutually suspicious — which greatly decreases the chances that its power will be abused," Franklin said.

The work is novel and potentially useful, said Andrew Odlyzko, director of the Digital Technology Center at the University of Minnesota. The researchers have provided "a clean solution to a famous problem... that has been open for a long time," he said.

"Key management is a very complex problem with conventional cryptosystems, and public key cryptography was invented largely to solve its difficulties. However, it turns out that public key systems also have their own... difficulties. The authors' system is a nice solution," Odlyzko said.

The reliance on a central authority means identity-based crypto systems are not an easy sell, however, and any new scheme is not likely to be accepted quickly, he said. "Known public key systems tend to be preferred, and new ones are slow to be accepted."

Although identity-based cryptography has been proposed before, this research is excellent, said Ronald Rivest, one of the creators of the RSA encryption scheme and a professor of computer science at the Massachusetts Institute of Technology.

While there are no technical barriers to implementing the proposal immediately, "it would be prudent to give the cryptographic community more time to assess the strengths and weaknesses of our proposal," Franklin said.

Franklin's research colleague was Dan Boneh of Stanford University. They presented the research at the 21st Annual International Cryptology Conference held at the University of California at Santa Barbara from August 19 to 23, 2001. Boneh was funded by the Defense Advanced Research Projects Agency (DARPA) and the Packard Foundation; Franklin was funded by the National Science Foundation (NSF).

Timeline: Now
Funding: Government

TRN Categories: Cryptography and Security; Internet
Story Type: News
Related Elements: Technical paper, "Identity-Based Encryption from the Weil Pairing," presented at the 21st Annual International Cryptology Conference, University of California at Santa Barbara, August, 2001; Demo: crypto.stanford.edu/ibe/

TRN

DNA Could Crack Code

By Kimberly Patch, Technology Research News
October 24, 2001

Knowledge that electrical engineers have gained from laying out components on circuit boards could make it easier to coax DNA molecules to do computations. The result may make it possible to crack a code that requires 3,000 years to solve on today's computers.

DNA computers use the same type of molecules that make up the genetic code for all life on earth and are potentially very powerful because they can perform computations on many molecules at once.

A strand of DNA is a long string of phosphates, each attached to one of four bases: adenine, cytosine, guanine and thymine. Various types of enzymes can cut the long molecules in places where the bases appear in a certain order, causing the strands to reassemble. This setup can be co-opted to perform the logic of computing.

Researchers have already plotted the minimum requirements needed to build a DNA computer. This involves a set of standard DNA molecules, or tiles, that each handle two inputs and two outputs. The tiles compute by interacting with each other, and the answer is extracted from the resulting structural changes.

Researchers from Ruhr University in Germany and Accenture Technology Labs in France are proposing to redesign the DNA tiles that make up a DNA computer to look more like the layout of an electronic circuit.

"Because the computation is done through the spatial arrangement of DNA tiles, you have to be really careful in the way you design your tiles," said Andre Weimerskirch, a graduate student at Ruhr University. "It turns out that the schoolbook layout of an electronic circuit designed to perform multiplication can be easily translated into a design for DNA tiles," he said.

These more complicated tile designs, which the researchers equate to DNA programming, would make DNA computers easier to use, Weimerskirch said. "To make it simple, imagine a jigsaw puzzle. There's only one way [the pieces] all fit together, because there are rules governing their matching," he said. The DNA tiles are similar to the pieces of a jigsaw puzzle. "Your knowledge of the problem is encoded in the pieces, and the results of the computation is the whole jigsaw

puzzle. The beauty is that you don't need to assemble the jigsaw puzzle yourself, it self-assembles," he said.

The researchers designed a multiplication tile, then went on to design even more complicated tiles. "We realized that we could do even more complicated operations... with few modifications," Weimerskirch said.

In theory, the design could be used to break a strong public key encryption system in a couple of days rather than the 3,000 years it would take an electronic computer, said Weimerskirch. The key needed to decrypt the NTRU Cryptosystems, Inc. encryption scheme is one in a very large number of possibilities: around 1,460 billion billion billion billion billion, which can also be represented as 2 to the 160th power.

The DNA computing scheme is fast; in addition, it can find an answer without having to try every number until one fits, according to Weimerskirch. Using the different types of tiles, the researchers can logically cut down on the possibilities in order to reduce the number that must be weeded out using brute force. This is essentially a type of programming. "The design of our tiles gives us the flexibility to... program the attack. The type of programming we can do is still rather crude, but we think this is... an important step in the right direction," said Weimerskirch.

There are several hurdles to carrying out the scheme on real DNA, according to the researchers.The first step is to make DNA tiles that conform to the designs. "This should not be too difficult and is within the reach of current technology," said Weimerskirch. The second challenge has to do with the error rate, he said. "We're currently performing simulations that will hopefully help us understand the problem better and hopefully allow us to give advice to the experimentalists" who may want to carry out the scheme, he said.

The researchers are also looking to find new tile designs that will solve even more complex problems, Weimerskirch said.

Doing computations using self-assembly is a very powerful method, said Nadrian Seeman, a chemistry professor at New York University. "The work... takes advantage of the notion of computation by self-assembly. However, the work remains a theoretical suggestion and its ultimate value will depend on its experimental implementation," he said.

It looks feasible to build tiles like the researchers have suggested, however, said Seeman. "We're not at this time building tiles exactly like those suggested by the authors, but we may well be able to do so in the foreseeable future," he said.

Because DNA computing is inherently more powerful than electronic computing, it could eventually be applied to many difficult problems like scheduling and cryptanalysis, said Weimerskirch. It is likely to take at least five years to overcome the experimental hurdles, he said.

Weimerskirch's research colleague was Oliver Pelletier of Accenture Technology Labs. The research was funded by Accenture Technology Labs.

Timeline: 5 years
Funding: Corporate
TRN Categories: Biological, Chemical, DNA and Molecular Computing;Cryptography and Security
Story Type: News
Related Elements: Technical paper, "Algorithmic Self-assembly of DNA Tiles and Its Application to Cryptanalysis," posted on the arXiv physics archive at http://xxx.lanl.gov/abs/cs.CR/0110009

TRN

Quantum Cryptography

Faster Quantum Crypto Demoed

By Eric Smalley, Technology Research News
January 29/February 5, 2003

There has been tremendous progress in quantum cryptography in recent years, and one system is already available commercially. But there's a long way to go before the technology matches its promise, and one of the biggest issues is coming up with devices that reliably generate and detect single photons at high speeds.

Using single photons isn't the only path to perfectly secure communications, however. Working out how to use only standard telecommunications gear to transmit cryptographic keys could dramatically improve quantum cryptography's paltry performance.

Cryptographic keys are numbers used to mathematically scramble and unscramble secret messages. The trick to using secret keys to encrypt messages is making sure they get to the intended recipient and no one else.

A group of researchers from the Institute of Optics in France and the Free University of Brussels (ULB) in Belgium has demonstrated a quantum cryptographic system that uses ordinary, weak laser pulses of several hundred photons each rather than single-photon pulses that entail special equipment. "Producing and detecting single photon pulses is not easy," said Phillippe Grangier, research director at the Institute of Optics. "Our scheme avoids these difficulties by using a completely different technique [that is] closer to standard optical telecommunications," he said.

The multi-photon method taps the nature of pulses instead of individual photons to guard against eavesdropping.

Initial results from the researchers' table-top prototype show that the multi-photon scheme should allow for faster information transmission than single-photon schemes over distances shorter than ten kilometers, said Grangier. In simulations of transmissions over short distances, the researchers were able to transmit secret keys as fast as 1.7

megabits, or millions bits, per second, according to Grangier. Today's single-photon quantum cryptographic systems transmit secret keys at speeds of a few hundred to a few thousand bits per second.

At the equivalent of just over 15 kilometers, the system transmitted data at about 75 kilobits per second, said Grangier. "Our system should allow much higher secret bit rates over short distances, but might be more sensitive to losses [and thus] may not work well over large distances," he said.

Single-photon systems use the polarizations of individual photons to represent each bit of a random string of 1s and 0s. Photons can be polarized in one of four paired orientations: horizontal and vertical and the two diagonals. The horizontal orientations can represent 0 and the vertical orientation can represent 1, for example.

The Heisenberg uncertainty principle explains why the system provides perfect security: a photon cannot be measured for both pairs of orientations. In order to eavesdrop on single photons, an eavesdropper — Eve — would have to replace the photons she intercepts, but because she would only be able to correctly measure half of them she would have to guess at the other half. The sender and receiver could compare a few of their photons, and if a quarter of them failed to match they would be tipped off to Eve's presence. They would discard those photons and try again until they got an unobserved string of photons. Over an ordinary communications line, Alice tells Bob which photons he measured correctly and they use those corresponding bits as a key.

The researchers' method uses weak laser pulses that contain a few hundred photons each, but taps the quantum nature of the pulses' amplitude and phase to secure the information they carry, said Gangier. Quantum physics describes the rules for individual atoms, photons and other particles. Lasers are quantum devices because their photons flow in lockstep, and the state of a laser pulse can only be measured according to quantum rules, including the Heisenberg uncertainty principle, said Grangier.

Instead of using the polarization of single photons to represent the bits of a key, a sender — Alice — would randomly choose an amplitude and phase for each weak laser pulse she sends to Bob, Grangier said. The method works because amplitude and phase are non-commuting quantum variables, which means they, like the pairs of polarizations, cannot be measured simultaneously. "This can be used to hide the secret key," he said.

The researchers' scheme also contains a twist. Its process of comparing bits reverses Alice's and Bob's usual roles. "The trick of reverse reconciliation is that the correct key is not what was sent by Alice, but rather what was received by Bob," said Grangier.

Reversing the process incorporates transmission errors into the key, said Grangier. This is more secure because Eve can only intercept messages before they get to Bob. "Eve always knows less about what was received by Bob than about what was sent by Alice," he said. This gives Alice and Bob a larger advantage over Eve compared to standard quantum key distribution systems, he added.

The difficult part of putting together the scheme was evaluating how much information is shared by the sender, receiver and eavesdropper, said Gangier. "While this is simple and well-known for photon counting, it was not clearly done for [multi-photon systems] before our work," he said.

The researchers have proved the system is secure in some situations, but have work to do to prove that it is perfectly secure. The initial results look good, however, said Grangier. "Our feeling is that the protocol is also robust against more general attacks, but this remains to be proven."

The researchers are not the first to propose a multi-photon quantum key distribution system, but the other methods generally rely on unusual, hard-to-produce light beams.

The work is "undoubtedly... the most advanced [multi-photon] quantum crypto demonstration," said Nicolas Gisin, a professor of applied physics at the University of Geneva in Switzerland.

The distance the scheme can be used over is limited, however, said Gisin. "It is not clear how realistic the proposal is for significantly longer distances." More work also needs to be done before the security of the system is clear, he added.

The researchers are working on testing the protocol at the wavelengths used by today's telecommunications systems, said Grangier. "The first challenge [is] using telecom integrated modulators and optical fibers at [a wavelength of] 1,550 nanometers instead of our present bulk optics set-up at 780 nanometers," he said.

They are also looking into quantum repeaters, he said. "We're also interested in... quantum repeaters that would do for quantum key distribution what optical repeaters can do for usual telecommunications, i.e., reaching arbitrarily large distances [using] optical fibers," he said.

How close the scheme is to practical depends on exactly how it will be used, said Grangier. "Since quantum key distribution setups are now on the market, the problem is no [longer] only technological," said Gangier. "One should decide what is ultimately needed: high bit [rates] over short distances [or] smaller bits rates over larger distances; fiber systems or free space systems?"

Gangier's research colleagues were Frédéric Grosshans, Jérôme Wenger and Rosa Brouri from the Institute of Optics, and Gilles Van Assche and Nicolas J. Cerf from the Free University of Brussels in Belgium. They published the research in the January 16, 2003 issue of the journal *Nature*. The research was funded by the European Union Information Society Technologies program (IST).

Timeline: Unknown
Funding: Government

TRN Categories: Quantum Computing and Communications; Cryptography and Security
Story Type: News Related Elements: Technical paper, "Quantum Key Distribution Using Gaussian-Modulated Coherent States," *Nature*, January 16, 2003.

TRN

Fast Quantum Crypto Demoed

By Eric Smalley, Technology Research News
November 27/December 4, 2002

Researchers at Northwestern University are tapping the laws that govern subatomic matter and energy in order to securely encrypt data and transmit it at 250 megabits per second over fiber-optic lines.

Sending a secret message involves two steps: encrypting, or scrambling, the message so that only the right person can read it, and making sure that the key to unscramble the message gets to the intended recipient and no one else.

Most quantum cryptography research has focused on the second step: finding ways to securely exchange encryption keys, which are random strings of numbers used to encrypt and decrypt messages.

The approach uses individual photons to represent each bit of a string of random bits. Because individual photons cannot be observed without altering their quantum states, an eavesdropper cannot look at a message sent using single photons without tipping off the sender and receiver to his presence. If the sender and receiver are sure that no one spied on their bit string, they can use it as an encryption key.

The Northwestern researchers have addressed the first step with a way to use quantum physics in the encryption process itself, with the premise that the parties communicating have already exchanged an encryption key. "No one has been addressing actual data encryption," said Prem Kumar, a professor of electrical and computer engineering and physics at Northwestern University.

The researchers' technique uses the quantum noise present in ordinary lasers. There is always a certain amount of randomness at the level of atoms and subatomic particles because the quantum world is ruled by probabilities. This randomness is known as quantum noise.

Quantum noise in lasers is caused by random fluctuations in the number of photons generated over a given amount of time, said Kumar. For example, if a laser produces 10,000 photons a nanosecond, or billionth of a second, it would have a fluctuation proportional to the square root of 10,000, which is 100, he said. "There will be a spread around the mean value of 10,000 and that spread would have a standard deviation of 100 photons," he said. "So sometimes I will have 10,100 photons, sometimes 10,090 photons."

The researchers' scheme mixes the message and the encryption key with this random quantum noise to produce a pattern that appears to be random to anyone without the encryption key.

The researchers demonstrated the technique by sending an encrypted message at a speed of 250 megabits per second over 4 kilometers of optical fiber in the laboratory and across a 2-kilometer fiber-optic line between two buildings.

Several research teams have used the single-photon quantum key distribution method to demonstrate communications systems that are impervious to eavesdroppers. But this method has major drawbacks: it is exceedingly slow and is limited to short distances, said Kumar.

Because each bit of the key is used to encrypt one bit of the message, and each message requires a new key, messages can be transmitted only as fast as keys can be sent.

The light sources that generate single photons can only produce data at about one kilobit, or thousand bits, per second, said Kumar. This translates to about 40 words a second, or about five single-spaced pages of text per minute. It would take about 17 minutes to transmit a print-quality image that takes up about one megabit, or million bits, and about 70 hours to transmit 250 megabits. "In communications networks we talk about tens of gigabits per second," Kumar said. A gigabit is one billion bits.

And because photons can't be copied without changing their quantum states and thus destroying the information they represent, quantum cryptographic transmissions can't be sent through the repeaters used to boost signals over longer distances in ordinary communications lines.

By sidestepping the issue of secure key exchange, the researchers were able to send secure data using the faster, standard lasers used in optical communications rather than devices that emit just one photon at a time, said Kumar.

The signals can also pass through amplifiers unharmed, and thus traverse long distances, Kumar said. "The encrypted communications can theoretically span thousands of kilometers, he said.

There are many efforts under way to study implementations of quantum cryptography using common, practical equipment, rather than single-photon generators, said Nicolas Gisin, a professor of applied physics at the University of Geneva in Switzerland. Several research teams have been developing quantum key distribution systems based on ordinary laser beams, he said. "We are also working on this."

It is too early, however, to judge the impact of the work, Gisin said. "It is not even clear whether these new schemes are really secure," he added. Researchers have not been able to analyze how the new schemes would handle the most powerful eavesdropping attacks, he said.

The Northwestern University researchers' claim for security against eavesdropping has to be tested, said Anton Zeilinger, a professor of physics at the University of Vienna in Austria. "If it is valid then this is a very significant [development]," he said. "We are analyzing that claim in my group," he added.

The researchers next plan to boost the data rate of their system to 2.5 gigabits per second, and increase the distance it covers, said Kumar. They also plan in the next year or two to tackle the problem of using ordinary laser beams to distribute perfectly secure encryption keys, he said.

The researchers expect to demonstrate the feasibility of their data encryption scheme in the next couple of years, and practical applications could be possible in five years, said Kumar.

Kumar's research colleagues were Horace Yuen, Geraldo Barbosa, Eric Corndorf and Chuang Liang. The research was funded by the Defense Advanced Research Projects Agency (DARPA).

Timeline: 5 years
Funding: Government
TRN Categories: Quantum Computing and Communications
Story Type: News
Related Elements: Technical paper, "Secure communications using coherent states," scheduled to appear in the Proceedings of QCMC'02: Quantum Communication, Measurement, and Computing; "Continuous Variable Quantum Cryptography Using Coherent States," Grosshans and Grangier, Physical Review Letters, February 4, 2002

TRN

Diamonds Improve Quantum Crypto

By Eric Smalley, Technology Research News
September 18/25, 2002

Scientists have thoroughly demonstrated that the quirks of quantum physics can secure secret messages, and one company is already selling a commercial quantum cryptography system.

But there's still plenty of room for improvement. Prototype quantum key distribution systems are slow and only work over relatively short distances. The main challenge is coming up with a light source that reliably fires off one and only one photon per pulse.

Today's quantum cryptography prototypes use lasers that are so heavily filtered that most of the pulses contain no photons, a few contain a single photon and fewer still contain two photons. Sending a cryptographic key means waiting for enough single photon pulses to be generated, and compensating for pulses that contain too many photons or none at all.

Researchers from the French National Scientific Research Center (CNRS) and Ecole Polytechnic in France have bettered the usual weak laser pulse method with a deliberately dirtied microscopic diamond: a 40-nanometer diamond nanocrystal with a nitrogen atom embedded next to an atom-size gap in the center. A nanometer is one millionth of a millimeter, and an atom measures about one tenth of a nanometer.

The nanocrystal emits light by fluorescence. When hit by a laser, the nanocrystal absorbs energy, then gives it off in the form of a single photon.

"We have developed an efficient, stable, all solid-state, room temperature single-photon source [and] we have used this single-photon source in a quantum cryptography setup," said Alexios Beveratos, a researcher at CNRS.

When the researchers used the setup as a light source to transmit quantum encryption keys through the open air, they were able to transmit 9,000 secure bits per second over a distance of 50 meters. "The limiting factor for the distance is that we didn't have a longer corridor. We should be able to span larger distances," said Beveratos.

An encryption key is a string of numbers used to lock and unlock encrypted messages sent over unsecure communications lines.

The researchers' goal is to communicate perfectly secure keys between the Earth and satellites. Researchers at Los Alamos National Laboratory aiming for the same goal have demonstrated a quantum cryptographic system that spans 10 kilometers using weak lasers, which is roughly equivalent to sending photons up through the thinner upper atmosphere to reach satellites hundreds of kilometers above the Earth. The next challenge is being able to aim single photons precisely enough to hit satellites.

The efficiency of single-photon detectors limits the distance that quantum cryptographic systems can operate over. Detector efficiency is affected by thermal noise and so detectors are usually cooled. Noise produces false positives, or signals when no photon is present.

The researchers' diamond-based device emits a single photon about two percent of the time it is stimulated, and can deliver as many as 116,000 single-photon pulses per second. Weak lasers can also generate 116,000 single-photon pulses per second but the diamond only generates 90 two-photon pulses during that time compared to 1,300 for the weak laser, said Beveratos. Two-photon pulses compromise security, and minimizing the risk they pose lowers the efficiency of the device.

Because two-photon pulses are inevitable, quantum cryptographic schemes use privacy amplification, which reduces a string of bits that includes some that have been exposed to an eavesdropper to a smaller string of secret bits. Privacy amplification converts two or more of the original bits into a single, new bit. Even if an eavesdropper knows some of the original bits, she is highly unlikely to be able to figure out the new bit. The more two-photon pulses a light source emits, the more original bits have to be used to make one secret bit.

Quantum cryptography involves a trade-off between data rate and distance, meaning the further apart hypothetical correspondents Alice and Bob are, the fewer secure bits they can send to each other. Because it generates fewer two-photon pulses than weak lasers, the researchers' light source requires

fewer pulses to make secret bits and so can span longer distances, said Beveratos.

Quantum cryptography allows users to tell for sure whether the encryption key they are using to encrypt and decrypt a message has been compromised.

Quantum cryptography schemes send encryption keys by representing each bit with only one photon. If there were two or more photons per bit, an eavesdropper could siphon off extra photons in order to copy the key without being detected. Using only one photon per bit means that an eavesdropper would have to replace the photons she intercepted, but the laws of physics make it impossible to replicate all of the photons correctly.

In the race to develop reliable single-photon light sources, several possibilities have surfaced. Certain molecules work well for a time, but "after having emitted a certain amount of photons, they photobleach, which means that they are not optically active anymore and do not emit any more photons," said Beveratos.

Quantum dots, which are microscopic specks of semiconductor that trap one or a few electrons, don't bleach, but they only emit single photons at very low temperatures, which requires cumbersome and expensive cryogenic equipment, said Beveratos.

The researchers' device, with its reduced multiple-photon rate, is the first to show a better secret bit rate than weak laser pulses, said Richard Hughes, a physicist at Los Alamos National Laboratory who has built a quantum key distribution prototype spanning 10 kilometers. "This type of light source and other similar ones will lead to improvement in the efficiency with which quantum key distribution [systems] can generate secret sharing keys," he said.

In order to reach satellites, the researchers will need to improve the device's efficiency from two percent to ten percent, said Beveratos. The researchers plan to test their system over longer distances and outdoors, he said.

Developing the quantum cryptographic systems for practical satellite communications will take at least five years, said Beveratos. The device would need to be miniaturized to fit on a satellite, he said. It will take less time to ready the device for use between two points on Earth, he said.

Beveratos' research colleagues were Rosa Brouri, André Villing, Jean-Philippe Poizat and Philippe Grangier of CNRS, and Thierry Gacoin of Ecole Polytechnique. The research was accepted for publication in the journal Physical Review Letters. The research was funded by the European Union.

Timeline: 5-6 years
Funding: Government
TRN Categories: Cryptography and Security; Quantum Computing and Communications
Story Type: News
Related Elements: Technical paper, "Single photon quantum cryptography," European Quantum Information Processing and Communications workshop in Dublin, September, 2002

Quantum Secrets Ride Phone Lines

By Eric Smalley, Technology Research News
August 7/14, 2002

The ability to safeguard secret messages using the quirks of quantum physics has been thoroughly demonstrated in the laboratory. Now field tests of quantum cryptography are showing that the technology can withstand the rigors of real-world communications.

Researchers in Switzerland have used this type of cryptography, which represents bits of information using single photons, to send theoretically perfectly secure messages between the cities of Geneva and Lausanne, which are 67 kilometers apart.

Quantum cryptography provides perfect security because it allows users to tell for sure whether the key they are using to encrypt and decrypt a message has been compromised.

Researchers at Los Alamos National Laboratory previously proved that a quantum signal could travel 50 kilometers. But that was over a spooled fiber-optic line contained in a laboratory, said Nicolas Gisin, a physics professor at the University of Geneva. "In our case the two end points were really spatially separated," he said.

More importantly, the Swiss experiment used existing fiber-optic phone lines. The fibers were "nothing special," said Gisin. They were not in commercial use during the experiment, but were part of a cable containing many fibers that were, he said.

Key encryption schemes use a unique mathematical key to mask each message. The sender and intended recipient use the key to encrypt a message, send it over unsecured channels, then decrypt it. The trick to keeping the message secret is making sure no one but the sender and receiver have access to the key.

The quantum cryptography scheme sends encryption keys over fiber-optic lines in a perfectly secure way by representing each bit with only one photon. Using two or more photons per bit makes it possible for an eavesdropper to siphon off some extra photons in order to peek at the key without being detected. Using only one photon per bit means that an eavesdropper would have to replace the photons she intercepted, but it is impossible to replicate all of the photons correctly.

This is because any given photon, or particle of light, can have one or more attributes, including polarization, which has to do with how the photon vibrates, and wave phase.

The researchers' quantum cryptography scheme generates photons in one of four states based on their wave phases. The system splits each photon, sends the halves down short pieces of fiber of slightly different lengths, and then joins the two halves. Because the halves travel different distances, their waves are out of phase, meaning the crests and troughs are out of sync by a particular amount.

The photons' four phase states come in two types: those whose waves match or are exactly opposite, and those whose waves are half way out of phase with one wave ahead of the other. Each type can be used to represent the 1s and 0s of digital information.

It is a quirk of quantum physics — the Heisenberg uncertainty principle — that makes the scheme perfectly secure: you can't look for both of the pairs of states at the same time, and you only get one look before the photon disappears. If you measure a photon to see if it is a 1 or 0 based on one pair of states, but it was generated in one of the other two states, you're out of luck. Your measuring device has absorbed the photon during your first look so you will never know whether it represented a 1 or 0.

This means an eavesdropper would only be able to correctly measure half of the photons he intercepts and would have to guess at the other half to produce substitutes. And he would only get about half the missing half right by chance, meaning one quarter of the substitute bits would be wrong.

The sender and receiver can check the error rate and so detect the eavesdropper by comparing a few bits. If the key has been compromised, they can throw it out and send another until they get an uncompromised key to encrypt their data. To form a key, the receiver measures the photons by randomly picking one of the two sets of states. Then they compare notes and the sender tells the receiver which photons he measured correctly. They then use those bits as the key.

The researchers' quantum key distribution system can only be used across relatively short distances because its performance drops off as the distance increases. At 10 kilometers the system can transmit quantum keys at 4,000 bits per second. At 20 kilometers the bit rate drops to 1,500 per second, and at 50 kilometers it drops to 100 bits per second. An ordinary modem transmits 56,000 bits per second. Once the users have an uncompromised key, however, the encrypted data can be sent over fast communications lines that include repeaters.

Today's fiber-optic communication systems compensate for diminishing signal strength — and thus span great distances — by using repeaters, which copy and retransmit fading light pulses. Repeaters can't be used to send quantum keys because they would intercept photons in the same manner as an eavesdropper.

The company id Quantique in Geneva, a spinoff from Gisin's laboratory, is marketing the quantum key distribution system. It consists of a pair of 18-inch-wide boxes that connect to personal computers via USB ports, and to each other over a fiber-optic line.

Gisin's research colleagues were Damien Stucki and Hugo Zbinden of the University of Geneva, and the Olivier Guinnard and Grégoire Ribordy of id Quantique SA. They published the research in the July 12, 2002 issue of the journal New Journal of Physics. The research was funded by the European Union.

Timeline: Now
Funding: Government
TRN Categories: Quantum Computing and Communications; Cryptography and Security
Story Type: News
Related Elements: Technical paper, "Quantum Key distribution over 67 km with a plug & play system," New Journal of Physics, July 12, 2002

TRN

Quantum Crypto Gear Shrinks

By Eric Smalley, Technology Research News
October 3, 2001

Researchers around the world are closing in on realizing the centuries-old dream of being able to send secret messages that are perfectly secure against any possible code breaking attempt. Before the average computer user can protect messages using quantum cryptography, however, the bulky laboratory equipment involved must be redesigned to fit on a few computer chips.

The key to this miniaturization is figuring out how to produce pairs of entangled photons using small, low-power lasers.

Researchers from Ludwig Maximilians University in Germany have taken a step in this direction by producing entangled photons using a small laser diode. The advancement opens the way to building quantum cryptographic devices on circuit boards.

Quantum particles like photons can be entangled, or linked so that they have properties like polarization in common. Particles can remain entangled regardless of the distance between them.

Entangled photons are the main ingredient of quantum cryptography. Two people who want to secretly communicate can, in theory, split a series of entangled photon pairs. They can then measure their photons and use the results as a key to encrypt a message that can be read by the person holding the other half of the entangled photons.

Because an eavesdropper cannot look at the photons without disturbing them, any intrusion can be detected immediately and the compromised key discarded.

Photons entangled using polarization work like this: the electric field of light vibrates in a plane perpendicular to the direction the light is traveling. When light is polarized, its electric field vibrates in one of four directions on that plane: horizontal, vertical or one of the two diagonals.

Entangled photons occur in some mix of the four possible polarization orientations, but when one of the particles is measured both particles snap into one polarization, regardless of the physical distance between them.

The preferred method for producing polarization-entangled photons is shining a laser through a special crystal that can turn a single ultraviolet or blue photon into a pair of entangled infrared photons. But this process is very inefficient, said Jürgen Volz, now a graduate student at the University of Munich.

"Usually this problem is overcome by the use of intense laser beams," he said. "Only large-frame ion lasers can be used... because only these lasers are able to create an adequate power output. These lasers are quite large and need enormous amounts of electrical energy... and possibly also water cooling. This makes standard entangled-photon-pair sources very expensive," said Volz.

The researchers got around the power problem by taking advantage of a basic principle of lasers. Lasers work by stimulating the atoms of a gas, which causes the atoms to emit photons. A pair of mirrors facing each other at opposite ends of the laser's gas chamber keeps these photons bouncing back and forth through the gas. As the photons bump into the gas atoms, they stimulate the emission of more photons.

"We placed an optical resonator around the nonlinear crystal," said Volz. In this case, the photons that bounce back and forth pass through the crystal rather than hitting the atoms of a gas. "With each pass, entangled photon pairs are created. So we can use much lower laser powers," he said.

Lower laser power means the lasers can be much smaller. "We use a simple laser diode, which could be operated from a simple battery," Volz said. "That makes our source much cheaper and quite compact in contrast to those based on ion lasers."

The researchers' entangled photon source generated about 10,000 pairs of entangled photons per second. Although this is only a tiny fraction of the astronomically large number of photons generated by even low-power lasers, it is sufficient for many quantum cryptography schemes, according to Volz.

The solid-state entangled photon source could be used for quantum cryptography in a few years, said Volz. "Two to five years seems possible," he said.

Volz's research colleagues were Christian Kurtsiefer and Harald Weinfurter of Ludwig Maximilians University. They published the research in the August 6, 2001 issue of the journal Applied Physics Letters. The research was funded by the German Research Foundation and the European Union.

Timeline: 2-5 years
Funding: Government
TRN Categories: Quantum Computing; Optical Computing, Optoelectronics and Photonics
Story Type: News
Related Elements: Technical paper, "Compact All-Solid-State Source of Polarization-Entangled Photon Pairs," Applied Physics Letters, August 6, 2001

TRN

Data Hiding
Printed Pictures Hide Images

By Chhavi Sachdev, Technology Research News
January 2, 2002

It's not hard to hide information in a digital image, but when the image is printed out, the hidden information is usually left behind on the computer.

A pair of researchers from Ben Gurion University in Israel have come up with a way to hide an image within a printed picture, allowing hard copies of an image to retain hidden information.

The scheme hides one half-tone picture in another so that scanners like those in supermarkets can unlock and view the information. The technique could be used to hide barcodes in product labels and fingerprints in ID pictures.

The researchers' system puts together two data files for two images, "one that we want to print as an observable picture and the other that we want to conceal within the observable picture," said Joseph Rosen, associate professor of electrical and computer engineering at Ben Gurion University.

The second, hidden image is coded into the first picture. The researchers scramble the mathematical representation of the hidden image with a mathematical key. Once an image is encoded, "only an authorized person who has the key... can reveal the hidden image," Rosen said.

The composite image can be printed on any printer; the print can then be read by a conventional optical scanner and processed by a PC to access the hidden image, said Rosen.

The method works because almost every printed picture is a halftone image, meaning it is actually a series of discrete dots. But because these dots are too small for the naked eye to see, a black and white picture appears many shades of gray. "A gray-tone printed picture is a collection of many white-only dots on a black background. The size of the dots, or their density, defines the level of the gray-tone that we see," said Rosen.

The hidden image is encoded in the position of the visible image's dots. Each dot can be shifted without significantly changing the visible image, and the altered positions can be mapped as a mathematical code.

Although the researchers have only hidden gray-tone pictures, the same method could be used for color, he said. A color picture is "still a composite of dots but this time there are 3 color dots — red, green and blue," he said.

While the visible image looks slightly fuzzy, there is no way for the eye to discern the hidden image. The deciphering depends on the key. "The correlation is an operation of scrambling the key... with one image in order to get the other image," he said.

This system is optical and different from digital watermarking, which hides information within the numbers

that make up the digital image, said Rosen. "Instead of having digits to play with, we have binary dots that we manipulate," Rosen said.

Printed copies of the researchers' pictures still contain the hidden file, but printouts of digital files that contain watermarks do not retain those watermarks. "When you print out a [digital] file the numbers become gray levels and the hidden information gets lost," he said.

Like a hologram, the hidden image is concealed in a global manner, said Rosen. "Every part of the [visible image] contains information on the entire hidden image, such that if you cover or destroy part of the concealogram you can still recover the entire hidden image from the rest," he said. The hidden image can be elicited even when 55 percent of the halftone picture is damaged or missing, he said.

The method can be used to conceal any image, including barcodes and fingerprints. College students, for instance, could have their meal-plan barcodes embedded in their ID photos; the labels of prescription medicine bottles could include information for the pharmacist about counter-indications, compliance, and pricing in one easily scanned label, said Rosen.

The work looks good, said Chris Honsinger, a senior research scientist at Kodak. While it seems similar to digital watermarking, this work is the "first real optical implementation I have seen. Most watermarking techniques require analog to digital conversion so that they can perform operations only possible using digital computation," he said.

The technique could be in practical use in a year, Rosen said. The researchers next plan to improve resolution of the images. "With a better PC, printer, scanner and a software engineer devoted to the project, we can reach a much better quality," said Rosen.

Rosen's research colleague was Bahram Javidi of the University of Connecticut. They published the research in the July 10, 2001 issue of the journal *Applied Optics*. The research was funded by the Ben Gurion University.

Timeline: < 1 year
Funding: University
TRN Categories: Cryptography and Security; Computer Vision and Image Processing
Story Type: News
Related Elements: Technical paper, "Hidden Images in Halftone Pictures," *Applied Optics*, July 10, 2001

TRN

Quantum Code Splits Secrets

By Eric Smalley, Technology Research News
October 10, 2001

IBM researchers have shown that tapping the weird quantum properties of particles like atoms and photons would improve on a classic technique that allows a group of people to hold pieces of a secret that can only be revealed by combining the pieces.

When a secret is too important for any one person to know, secret-sharing cryptographic protocols provide a way to break up the secret into parts held by several or even many people. The protocols keep the secret until all or most of the parts are assembled.

Adding a quantum component to this scheme would make it harder for the people holding the pieces to cheat or be coerced into revealing the secret.

The IBM scheme is a step in that direction. "We haven't done anything so sophisticated in the quantum version" as splitting a secret into many parts, said David P. DiVincenzo, a physicist at IBM Research. "We've just been investigating the simple case of splitting a secret into two."

The quantum secret-sharing scheme is similar to quantum cryptography and quantum computing because it relies on the quantum mechanical condition of entanglement.

Particles like atoms are usually either spin up or spin down, meaning that the axes they spin around point either up or down relative to the magnetic field around the atoms. But when atoms or other particles are isolated from the environment and cannot be observed, they enter the quantum mechanical state of superposition, which means they are in some mixture of both spin up and spin down.

Two or more particles in superposition can be entangled so that even if they are separated, when one of them is measured and becomes either spin up or spin down the other particle immediately leaves superposition and assumes the same spin regardless of the distance between them.

There are four possible combinations of spins for a pair of entangled particles. One combination, called a singlet, stands out from the other three, which are called triplets.

The quantum secret-sharing scheme represents a bit of information by creating a string of entangled pairs of particles. An odd number of singlets in the string represents a one, and an even number of singlets represents a zero.

Because the two particles have to be together in order to tell whether they form a singlet or a triplet, two people sharing a secret this way couldn't simply measure their halves of the string and compare notes to tell whether the bit is a one or a zero. This makes quantum versions of secret-sharing protocols more secure than classical versions.

"If the parts of the secret are actually pieces of a quantum state, then even communication — at least communication of the ordinary, classical sort — can be insufficient for them to reconstruct the secret," said DiVincenzo. "They need to do something stronger. They need some kind of additional quantum technology in order to unlock the secret," he said.

The needed quantum technology could be a quantum communications channel. If the polarization of photons were used rather than the spin of atoms, the photons could be transmitted while preserving their quantum states.

In order to carry out the scheme, however, there must be a way to store the quantum states of particles for long periods of time.

"This scheme is not something that can be realized in the immediate future, except as a demonstration," said Daniel Gottesman, a fellow at the Clay Mathematics Institute and a visiting scholar at the University of California at Berkeley. "You need to store the quantum states until it comes time to open the secret, and it will be a while until we can do that reliably."

Quantum secret sharing "would require a good quantum memory and the ability to measure qubits. Some of the rudiments of what are needed in this scheme are available today," said DiVincenzo.

Practical quantum secret sharing will also require the development of quantum repeaters in order to send quantum information over distances greater than the roughly 10 kilometers possible today. Repeaters boost signals traveling along communications lines.

Quantum repeaters could be developed in about six years but quantum memory will probably take longer, said DiVincenzo. "That gets into the cloudy future," he said.

It also remains to be seen whether the added property of requiring quantum communications makes for a more useful form of secret sharing, Gottesman said.

It should be possible to make a practical form of the quantum secret-sharing scheme before large-scale quantum computers can be built, said DiVincenzo. Large-scale quantum computers are probably more than 20 years away, according to many researchers.

DiVincenzo's research colleagues were Barbara M. Terhal and Debbie W. Leung of IBM Research. They published the research in the June 18, 2001 issue of the journal Physical Review Letters. The research was funded by the National Security Agency (NSA), the Army Research Office and IBM.

Timeline: Unknown
Funding: Government; Corporate
TRN Categories: Cryptography and Security; Quantum Computing
Story Type: News
Related Elements: Technical paper, "Hiding Bits in Bell States," Physical Review Letters, June 18, 2001

TRN

Watermarks Hide in Plain Text

By Ted Smalley Bowen, Technology Research News
June 6, 2001

Writers are known as much for how they word things as for what they have to say. Putting this maxim to more prosaic use, a group of researchers at Purdue have devised a way of using word substitution and syntactical changes to watermark text documents.

Watermarking a document usually involves embedding an image or symbol that is invisible to most viewers in order to establish authorship and/or ownership. It is common to watermark bitmapped multimedia files, which are literally maps of individual pixels, by changing some of those pixels. Sound files are commonly watermarked by making frequency changes.

Text formats, which store strings of characters, offer fewer opportunities to embed visual watermarks. The alternatives include making bitmaps of text files or changing spacing between letters, words and lines.

The Purdue scheme differs from other watermarking techniques in its use of the actual words in a document to create patterns that, taken together, function as a distinguishing mark. In order to read the watermark, the user needs the unique encryption key used to create it. While not watermarking in the sense of embedding discernible images in a document, the Purdue natural language scheme is functionally similar.

An advantage of the scheme is the word patterns can withstand edits and revisions, according to the researchers. "Even if a watermark-carrying sentence is modified, the watermark [bits] it stores will survive some changes, such as replacing words by their synonyms, and a watermark bit has a 50-percent chance of surviving a drastic change to the sentence," said Mikhail Atallah, professor of computer science at Purdue.

Future implementations of the scheme could survive translation to other human languages, according to Atallah.

The scheme does have some drawbacks, however. Because it changes some of the language of the document, it is inherently incompatible with text, like creative writing, whose meaning requires precise and unique syntax. The changes include inserting an extra phrase in a sentence, splitting a sentence, adding transitional words or converting sentences to that bane of English teachers and editors — the passive voice.

"It's designed for situations where style is not so important, such as in government documents [and] user manuals, [although] the current prototype is fine for precise technical writing," said Atallah.

It also requires that the document be at least several dozen sentences long in order to provide enough text in which to sequester the watermarks. "The watermark is hidden in a [relatively] small number of sentences. So the technique is not suitable for watermarking very short text," Atallah said.

To test the scheme, the researchers wrote a program that uses computer-based natural language processing to determine which words, phrases or arrangements of words can be altered without changing the document's meaning, and then makes the changes to watermark the text.

The changed passages are represented as syntactic tree diagrams, which are in turn converted mathematically to streams of bits. In their experiments, the researchers were able to hide a 26-bit watermark in text 50 sentences long.

In the next phase of the work, the researchers plan to use more complicated text meaning representation (TMR) trees, which would permit them to associate the watermarks with the meanings of words rather than their structures. Because semantics is not directly determined by syntax, the scheme would be more resistant to changes in the arrangement of words, according to the researchers.

As a key-based system, the prototype is vulnerable, according to Dan Wallach, an assistant professor of computer science at Rice University. "In general, they're taking what I would describe as a reasonable approach to their problem, [but] their security analysis... ignores the effect of an attacker deciding to insert a new watermark using [the] same system. I'd imagine that would make the system much easier to defeat," he said.

Although an attacker who doesn't have the key won't necessarily be able to determine which sentences are original and which were modified, the attacker could make more widespread and potentially unpleasant changes to the text to wipe out all the modified sentences, said Wallach.

Another hurdle is modifying the scheme for different languages, he said. Because the system hinges on being able to find relatively unimportant sentences in the text that it can safely mutate, it depends on a fairly complex language model. "This technique would need a separate model for every language on which it was to work, and heaven only knows what it would do with slang or with mixed-language writing," said Wallach.

Atallah's colleagues at Purdue were Victor Raskin, Michael Crogan, Christian Hemplemann, Florian Kerschbaum, Dina Mohamed, and Sanket Naik. The researchers presented their work at the International Information Hiding Workshop in Pittsburgh, April 25-27, 2001. The research was funded by Purdue's Center for Education and Research in Information Assurance Security (CERIAS).

Timeline: Now
Funding: University
TRN Categories: Cryptography and Security
Story Type: News
Related Elements: Technical paper, "Natural Language Watermarking: Design, Analysis, and Proof-of-Concept Implementation" published in the Proceedings of the 4th International Information Hiding Workshop, Pittsburgh, Pennsylvania, April 25-27, 2001

TRN

Statistics Sniff out Secrets

By Kimberly Patch, Technology Research News
September 26, 2001

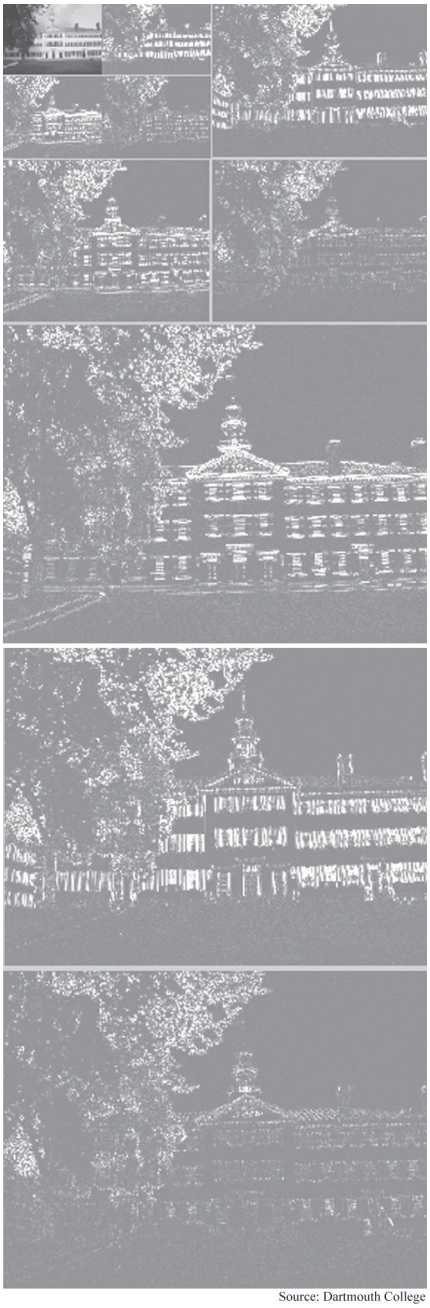
As digitized pictures, audio and text proliferate, people are exploring ways to exploit these media by hiding messages within the information, which leads others to try to detect these hidden messages.

Although steganography — the practice of hiding a secret message in written or audio information — is hardly new, computers and the Web have added a new twist simply because the volume of information that makes up digitized media is so large. This provides for historically large haystacks that easily obscure needles.

A Dartmouth College researcher has found a method that makes it easier to detect hidden messages in digital images, which can contain a megabyte — a string of one million ones and zeros — of information, or more.

Digital images are made up of pixels, or dots of color. Especially with high-resolution digital images that have one million or more different shades of color, it's easy to hide a message by slightly altering these colors in ways that are imperceptible to the human eye.

In an image that has not been tampered with, however, the information that



Source: Dartmouth College

These pictures show different aspects of the wavelet mathematics used to compress images. The fainter images show diagonal information only; they're more ghostly because this particular image has many strong horizontal and vertical lines.

makes up the image is not simply random. The key to the Dartmouth detection method is creating a statistical profile of the compressed data files that make up natural, or undisturbed images, then checking a given image against the profile, according to Hany Farid, an assistant professor of computer science at Dartmouth. "In order to detect hidden messages in an image we need to start by characterizing the statistics of natural images. The hope, then, is that when a message is hidden in an image, these statistics are disturbed," he said.

When images are compressed so they can be stored as smaller files, the digital information that indicates the color of each pixel is changed into wavelet information. Wavelet mathematics includes functions like spatial position, orientation and scale. Wavelets allow for compression because all the information that makes up a wavelet can be reconstructed from only a portion of that information. An image is compressed by storing only the portion that is needed to reconstruct the whole.

Farid collected two types of wavelet statistics: variations like mean, variance, skewness and kurtosis in the coefficients, or numbers that make up the wavelets, and information about the rate of errors that occur when reconstructing full wavelets from compressed information.

Variance shows how spread out the data is from the mean, or average; skewness shows how evenly distributed the data is on either side of the mean, and kurtosis shows how peaked the distribution of data is around the mean, said Farid.

He then combined the variation and error rate statistics into a vector — a mathematical construction that is like a virtual sculpture with 70 to 100 dimensions rather than the usual three.

By comparing the statistical vector information with the same information in an individual image, Farid was able to tell if the image had been disturbed with a hidden message, he said.

The practice of information hiding, or steganography is related to, but different from cryptography. In cryptography a message is encrypted and then transmitted. If you saw the transmission you wouldn't be able to decipher the message, but you would know the sender and receiver might be trading secrets. The goal of information hiding is to go a step further by camouflaging the transmission entirely, said Farid.

The statistical vector method only detects hidden messages, and cannot read or remove them, but may eventually be adapted to do so, said Farid. "This work cannot obviously be adapted to remove or decipher the hidden message. I do believe, however, that it is possible to do so," he said.

The technique is an extension of previous steganography detection schemes, said Neil F. Johnson, associate director of the Center for Secure Information Systems. "It is potentially useful if the techniques for detection are repeatable," he said.

In addition to determining if there's information embedded in a message, it is also useful for a detection method to identify the steganography technique used to hide the information, said Johnson. Another goal is to be able to extract the embedded information, he said. Steganography tools exist that can do this in at least some cases, he added.

Steganography has many applications, both good and bad, said Farid. "It can be used to protect copyrights in digital media, for unobtrusive military and intelligence communication, covert criminal communication, trafficking of illegal pornography, and for the protection of civilian speech against repressive governments."

An unfortunate side effect of research that reveals hidden messages, is "repressive governments could use this research to limit civilian speech," Farid said. Because of the possible unpleasant applications, "some will be very critical of this research, possibly with good reason," said Farid. "Nevertheless, I believe that the development of these techniques are inevitable and... will lead to better techniques for hiding information, which in turn will lead to better detection schemes and so on. My larger research vision is in authenticating digital media so that [neither] the 'good-guys' [nor] the 'bad-guys' will... be able to manipulate digital sound, image or video to suit their needs," he said.

The method can also eventually be applied to analyzing works of art to detect forgeries or to determine if more than one artist painted a single painting, Farid said.

The method could be used practically in less than two years, said Farid.

Farid's research was funded by the National Science Foundation (NSF) and the Department of Justice (DOJ).

Timeline: < 2 years
Funding: Government
TRN Categories: Cryptography and Security; Pattern Recognition
Story Type: News
Related Elements: Technical paper, "Detecting Steganographic Messages in Digital Images," posted at www.cs.dartmouth.edu/farid/publications/tr01.html

TRN

Systems and Software Defense

Software System Heals Itself

By Kimberly Patch, Technology Research News
November 27/December 4, 2002

The human body is largely self-healing — you don't have to consciously orchestrate the process of platelets forming a scab and skin healing over after you cut your finger. Your body senses what needs to be done and does it automatically,

freeing you up to complain all you want about the pain and inconvenience.

A computer system, however, usually has to stop what it's doing in order to recover from a virus or hacker attack. This can range from annoying to expensive: some database systems process thousands of financial transactions per second.

Researchers from Pennsylvania State University are working on a database system that automates the process of recovering from attacks, and keeps the database running during the recovery process.

Although it is difficult to prevent intrusions from unauthorized users, many intrusions can be detected soon after they occur. The key is to contain the damage so it does not spread, according to Peng Liu, an assistant professor of information sciences and technology at Penn State University. With this in mind, the researchers built a system that "monitors its environment and its health status in real-time," said Liu.

One requirement for database health is that the percentage of corrupted data objects should be small, said Liu. Whenever the status reaches a certain threshold, the system adapts its behavior to make sure its health does not worsen, he said.

The researcher' self-healing system detects intrusions, contains the part of the database that has been damaged, locates the corrupted data, and repairs each corrupted data object by restoring its most recent undamaged backup copy, said Liu.

A key part of the system is that its algorithms only replace corrupted objects, and it allows the rest of the database to keep processing transactions while this takes place.

The Penn State process is less interruptive than existing schemes, which take an entire database off-line to restore it, then must re-create transactions that occurred between the backup version and when the damage occurred, according to Liu. Traditional recovery mechanisms address the problem using complete rollbacks, which undo the work of benign transactions as well as malicious ones, he said.

The researchers have implemented a prototype system using an Oracle database running on the Windows NT operating system. The software "can already support many real-world Oracle database applications," said Liu.

The researchers' software includes a separate log file that keeps track of transactions written to the database in a more detailed way than the standard Oracle database log. The software includes an algorithm that mediates every user transaction in order to collect log information and make the system aware of the status of transactions.

An intrusion detection algorithm receives an alert when a new event is recorded in the log, and uses log information to identify bad transactions. If a bad transaction is active when it is identified, the software will abort the transaction. If the transaction is already committed, the system puts it on a list of bad transactions and sends an alert to a repair manager algorithm so that the damage can be assessed and quickly repaired.

The repair algorithm is based on traditional recovery mechanisms, according to Liu. The challenge was keeping the database working during the healing process, he said.

Each repair algorithm has static and dynamic versions. The dynamic version allows the database to keep running as the repairs are made. The repair manager keeps tabs on the growing log of on-the-fly histories and marks any bad or suspect transactions. The repair manager builds an undo transaction for every bad or suspect transaction and submits it to a scheduler, which schedules the operations to generate a correct on-the-fly history.

There is a drawback to using the dynamic repair manager algorithm, however, according to Liu. The researchers' tests show that it backs out, or reverts, more good transactions than the static version. The advantage of the static version is that less work must be done reprocessing transactions, but the database is inaccessible while repairs are taking place.

The work is novel, said Karl Levitt, a professor of computer science at the University of California at Davis. "Rather than using checkpointing, it relies on identifying the bad transactions and generating anti-transactions," he said.

And it uses "a clever method of syntactic dependency to identify, from the many transactions that occur after a single, attack-causing transaction, those that are linked to the bad transaction," said Levitt.

One potential drawback to the system is that it relies on the identification of a single attack transaction, Levitt said. "Some intrusion detection systems detect the occurrence of a bad state but do not identify the single transaction, if, indeed, there is just one, that caused the bad state," he said.

Such a system also has to deal with false positives that get carried over from the intrusion detector, said Madhavi Gandhi, a researcher at the University of California at Davis. "Repairing falsely accused bad transactions may have greater impact than simply falsely detecting them," she said.

The performance hit a database will have to take to run the self-healing software could also be an issue, said Gandhi. Performance generally degrades when audit logs are collected; "using triggers in databases also typically drains performance and is recommended for limited use. The design of the attack recovery system here seems to use them on all tables," she said.

The self-healing software slows the database 10 to 30 percent, according to Liu. "Using a lot of triggers does have negative impact," he said. This may be remedied by integrating the algorithms into the database program, however, he said. "If our algorithms [were to be] built into the database kernel, the performance impact should be near zero percent," he said.

This type of work is needed, but rare, Levitt added. "This is one of just a few papers that suggest work beyond just detecting attacks." The industry needs ways of automatically

responding to attacks "using state restoration, as this paper suggests; stopping the attack if it is a fast-moving worm; fighting back, [which is] very controversial; fixing the bug or misconfiguration that permitted the attack in the first place; [and] deception, to slow down the attack," he said.

The researchers are looking to give the system the ability to adapt during the healing process so that it will be less vulnerable to the same damage a second time, said Liu. "We will extend the work from self-healing to self-regenerative, where the database system can be generated... into an even stronger system after self-healing from some attacks," he said.

The researchers are aiming for software that will adapt to its circumstances in ways similar to living beings, he said. "We ultimately aim for a system as autonomic and resilient as human bodies."

The system should be ready for practical use in two to four years, according to Liu.

Liu's research colleagues were Paul Ammann and Sushil Jajodia. They published the research in the September, 2002 issue of IEEE Transactions on Knowledge and Data Engineering. The research was funded by The Defense Advanced Research Projects Agency (DARPA), the U.S. Air Force and the National Science Foundation (NSF).

Timeline: 2-4 years
Funding: Government
TRN Categories: Databases and Information Retrieval; Cryptography and Security
Story Type: News
Related Elements: Technical paper, "Recovery from Malicious Transactions," IEEE Transactions on Knowledge and Data Engineering, September, 2002

TRN

Text Software Spots Intruders

By Kimberly Patch, Technology Research News
October 30/November 6, 2002

The computer anti-virus programs in common use today use signature detection schemes that can only protect a machine from viruses that have been previously identified and entered into the programs' virus databases.

Anomaly detection systems, however, sense when normal patterns of communications change in order to stop new viruses — or any other system intruders like worms or unauthorized users — in their tracks.

The trouble is, existing anomaly detection schemes all generate high error rates — they cry wolf so often that they are impractical. In order to identify the real intrusions, system managers must spend time checking out every possibility.

Researchers from the University of California at Davis have taken an unusual tack in anomaly detection by adapting text classification techniques to intrusion detection. Their initial results suggest that the technique could produce an anomaly detection system with a reasonable error rate.

The idea to apply text classification to intrusion detection began with a conversation about categorizing Web pages into clusters that share a given property, said V. Rao Vemuri, a professor of applied science and computer science at the University of California at Davis, and a scientist at Lawrence Livermore National Laboratories.

Instead of categorizing Web pages, however, the researchers used the classification system to categorize computer users into just two groups — authorized users and intruders. "The problem is to decide what 'text'" to use for the problem, Vemuri said. We wanted some objective way of characterizing a user that the user... cannot consciously influence" in order to prevent an intruder from fooling the system, he said.

They turned to system calls to characterize a user. System calls are the internal requests various pieces of software make to each other in the course of carrying out a user's instructions. "The system calls are generated by the computer, and the user cannot really influence the sequence in which they are generated," said Vemuri. The scheme treats each system call as a word and each sequence of system calls as a document, and classifies each document as one generated during normal activity or intrusive activity, he said.

The nearest-neighbor text categorization technique the researchers used categorizes Web pages based on how they are linked. The nearest neighbors in terms of links also tend to be closer in terms of content.

The researchers' detection scheme characterizes an authorized user by building a profile of activities. "For example, in the course of my normal life, I use email, browse Web pages, use Word, PowerPoint [and] printers," said Vemuri. "Let's suppose that I rarely, if ever, use Java or C++. I rarely use root privileges. If someone logging onto my machine uses these, that departure from normal usage should signal... abnormal, [possibly] intrusive activity," he said.

The problem turned out to be easier than categorizing Web pages, said Vemuri. "Usually we use many categories. In our example, we have only two categories — authorized or intruder, and in the worst-case three" if the system has to resort to classifying activity as unknown.

In addition, Web pages can be very long and the size of the English vocabulary is around 50,000 words, which makes categorizing Web pages a computer-intensive task. "In our case, the vocabulary — distinct system calls — rarely exceeds 100, and the size of the 'pages', [or groups of calls], is also very small," he said.

Short sequences of system calls have been used before to characterize a person's normal behavior, but this requires building a database of normal sequences of system calls for each program a person uses. The text categorization technique,

however, calculates the similarities between program activities, which involves fewer calculations.

This allows the system to detect an intruder as the intruder is affecting the system, said Vemuri. "The computational burden in our case is much smaller, to the extent we started to dream about the possibility of detecting an intruder in real-time," like the way contestants called out titles as songs played on the TV show "Name That Tune", he said.

The researchers' current implementation is almost real-time, said Vemuri. "We have to wait until [a] process, terminates, or halts" before completing the classification, he said. Intrusive attacks, however, are usually conducted within one or more sessions, and every session contains several processes, said Vemuri. Because the classifier method monitors the execution of each process, it's likely that an attack can be detected while it is happening, he said. The researchers are also working on allowing the system to make a classification before a process terminates, he added.

The researchers tested their scheme with 24 attacks within a two-week period. The method detected 22 of 24 attacks, and had a relatively low false-positive rate of 31 false alarms out of 5,285 events, or 0.59 percent, according to Vemuri.

The method shows promise, said Bennet Yee, an assistant professor of computer science and engineering at the University of California at San Diego. "The novelty is noticing that text classification techniques can be adapted to intrusion detection, and doing the experiments that validate it," he said.

If it proves practical and is widely deployed, the technique could help prevent malicious software like the Internet worms Code Red and Klutz, Yee said. "It should be able to recognize new attacks as anomalous behavior and raise alarms earlier [than] signature detection schemes where a database of bad behavior must be compiled first," he said.

There is still work to do to determine if the method can be improved to a low enough false positive rate, however, said Yee. A practical anomaly detection system must have a very low false positive rate in order to be commercially useful because if system administrators spend too much time chasing down false alarms, they "will not want to use the system and will turn the intrusion detector off," he said.

Even a false positive rate of 0.44 percent could mean 23 false alarms per day if there are 5,285 events per day, Yee said. "Most people will not want to handle a false alarm per hour per machine," he said. The researchers' method is an improvement over earlier anomaly detector designs, but "further improvements are still necessary for broader use," he added.

It is theoretically possible to use the method today, said Vemuri. The researchers are working on proving that the method can be used without raising too many false alarms, he said.

To cut down on false alarms, the researchers are looking to make a redundant system "where we use different methods on different data sets, combine the results of both those methods, or use a best of three voting system," he said. One method could use system call data, for instance, while another could analyze instructions used, he said.

The researchers hope to have their anomaly detection system worked out and supported with performance data within a few years, said Vemuri.

Vemuri's research colleague was Yihua Liao. They published the research in the Proceedings of the 11th Usenix Security Symposium, which was held in San Francisco August 5 through 9, 2002. The research was funded by the Air Force Office of Scientific Research (AFOSR).

Timeline: 2-3 years
Funding: Government
TRN Categories: Cryptography and Security; Computer Science; Internet; Networking
Story Type: News
Related Elements: Technical paper, "Using Text Categorization Techniques for Intrusion Detection," Proceedings of the 11th Usenix Security Symposium, San Francisco August 5-9, 2002

TRN

Data Protected on Unlocked Web Sites

By Chhavi Sachdev, Technology Research News
December 19/26, 2001

Looking up information on the Internet is easy, but is it sometimes too good to be true? How do you know that a posted investment history, for instance, is correct and complete?

Existing technology allows an author to use a digital signature to authenticate a document. The author signs the document using a private key program, which performs a mathematical calculation on the document. To view the signature, the reader downloads the authors public key, which can be posted in a publicly available place.

But existing signature schemes only work with specific sets of data. To request the last two years of that investment history, for example, you might have to download the entire record to get an authenticated copy.

A team of researchers has come up with a signature scheme that allows portions of signed documents that are stored in Extensible Markup Language (XML) databases to be retrieved and authenticated. "The existing XML signature standard won't let you do that. You can only authenticate an entire document, not parts of it," said Premkumar Devanbu, an associate professor of computer science at the University of California at Davis.

Using the researchers' TruthSayer scheme, an author can also sign an XML document and give it to someone else to store and post, said Devanbu. In other words, the author

would not have to be the publisher in order to authenticate the material. This means that anyone, from a government agency to the Mafia, could have a Web site that published authenticated data from multiple sources, and the receiver would be able to verify the origin of the documents, Devanbu said.

When the originator of the data uses the scheme to sign a document, the system processes the data involved, including its indexes, which are pieces of software that handle queries from clients and speed up searches, said Devanbu. "Typically, only a tiny fraction... of these indexes need to be looked at to answer the client's query. It is actually this index that is digested in a special way, to compute the database signature in our scheme," he said.

The secure data is then sent to an untrusted publisher;

"When the publisher gets a signed [answer] from the owner, he checks to see if that's right using the owner's public key," said Devanbu. When anyone queries the data, the publisher provides the response and a verification code to prove that the accompanying answer is accurate and complete, he said.

When an untrusted online site gets a client query, it searches through the indexes, keeping track of which parts of the index were searched, and returns those parts along with the answer, Devanbu said. "The client now runs a [verification] program over the answer [and] the returned parts of the index."

The verification program compares the publicly available author's key with the publisher's certificate. "The critical thing about the verification [code] is that it doesn't depend on any keys at all. It uses a... digesting operation to prove that the answer that was sent by the publisher was the same as the answer the owner would have given," said Devanbu.

If the comparison proves a match, the client knows the data has not been compromised. If there is a discrepancy, she knows the data has been changed by someone other than the author.

"If a bad guy replaces a publisher's copy of the owner's public key with a forged public key, then the bad guy can make the publisher trust an invalid root hash value, and deceive the publisher into publishing bad data," said Devanbu. "But as long as the clients have the correct copy of the owner's public key, they won't believe this deceived publisher."

To digest documents, the signature system uses the Merkle hash tree mathematical function. The function starts with a set of data and computes until there is only one root value left, which is the key the author uses when he signs a document, said Devanbu.

The scheme could be used to retrieve authenticated portions of published data, from traffic citations and court proceedings to Freedom of Information Act requests, "all of which are either already or soon will be in XML," said Devanbu. In short, "any situation where correctness of data and efficiency of access is important."

"Suppose the government signs a large XML document containing all discussions within the Department of Labor on some topic, and gives it to another agency to handle responses to FOIA queries," said Devanbu. "Someone in the Department of Labor who wanted to hide something might try to coerce the person at the agency handling FOIA queries to hide some details in responses to queries. With [Truthsayer,] a false or incomplete answer to queries on the XML document would be detected immediately," he said.

Another advantage of this encryption scheme is that the owner of the data does not have to be online. "If the owner is physically disconnected, he cannot be hacked, and no one can steal his private key. So his signature is not forgeable," said Devanbu. This type of system is called an 'air gap' and is used by many Defense Department systems, he said.

This work is elegant and efficient and could spur further developments in this area, said Andrew Odlyzko, a professor of mathematics and the director of the Digital Technology Center at the University of Minnesota.

The most important feature of this scheme is that it could "provide authenticated information access through untrusted intermediaries," Odlyzko said. People might, however, opt for simpler solutions than this one because the threat the authors scheme guards against is probably not all that serious, he said.

The researchers are getting ready to test the scheme with a realistic, open-source database system, said Devanbu. It could be ready for practical use in 4 to 6 years, he said.

Devanbu's research colleagues were Michael Gertz, April Kwong, Chip Martel, Glen Nuckolls, and Philip Rogaway from the University of California at Davis, and Stuart G. Stubblebine of Stubblebine Consulting, LLC.

They presented the research at the 8th ACM Conference on Computer and Communications Security held in Philadelphia between November 5 and 8, 2001 and is scheduled to be published in the Computer Security Journal, 2001. The research was funded by the National Science Foundation (NSF), and the Defense Advanced Research Project Agency (DARPA).

Timeline: 4-6 years
Funding: Government
TRN Categories: Cryptography and Security; Internet; Databases and Information Retrieval
Story Type: News
Related Elements: Technical paper, "Flexible Authentication of XML Documents," in the 8th ACM Conference on Computer and Communications Security in Philadelphia, November, 2001; Technical paper, "Authentic Re-Publication by Untrusted Servers: A Novel Approach to Database Survivability," presented at the Third Information Survivability Workshop 2000, October 24-26, 2000, in Boston

Physics Methods May Spot Intruders

By Kimberly Patch, Technology Research News
December 5, 2001

The key to detecting uninvited visitors is recognizing them.

This gets difficult in crowded situations, like large networks, because there is a lot of normal traffic, or noise, that can cover an intruder's comparatively quieter signal. What's even more difficult, however, is detecting a new type of intrusion the first time it happens. Essentially what's needed is a way to detect what you don't know you're looking for.

Researchers from the University of South Carolina have tapped the methods of nuclear experiments to map network traffic and extract patterns of typical network behavior. When scientists looking into the makeup of matter cause nuclear particles to collide, hundreds of detectors monitor every facet of the complicated reaction to capture any slight derivation that may point to an unknown phenomenon.

Analyzing network traffic data this way makes it easier to tease out derivations that point to known network intruders, said Vladimir Gudkov, a physics research professor at the University of South Carolina. "If... almost complete monitoring and data collection [of nuclear events] is possible in physics, why not try to find a way to do similar things in network monitoring?" he said.

The research could also eventually be adapted to the really difficult problem of detecting new methods of intrusion as they are happening, said Gudkov. "We have an opportunity to detect even unknown intrusions in the reconnaissance stage of an attack," he said.

When a file is transmitted over a network it is first broken up into many small packets, which traverse the network using whatever route is available and are reassembled when they arrive at their destination.

To closely monitor a network, the researchers track all the properties of these packets, including how they change over time. Routers, the specialized computers the control traffic around the Internet, put time stamps and other marks on the packets. The advantage of using this time-dependent information is it provides a complete description of the process. "This is exactly what we need for reliable numerical analysis," Gudkov said.

The researchers translate this information into mathematical functions in order to use the complex systems theory that physicists use to extract information from large, changing sets of data, said Gudkov.

The method captures raw data from a network node, then on a separate system plots the mathematical functions in two or three-dimensional imaginary space, and uses pattern recognition to find deviant signals. The result is an "ability to optimize signal-to-noise ratio and to analyze signals in real-time," Gudkov said.

This makes the faint tracks of an intruder more apparent. "The basic idea is to define the normal network behavior using the complete network monitoring. The deviation from the normal traffic behavior will give an alert for possible... intrusions," he said.

In plotting the signals the researchers also found something surprising: some of the ways information flows in these imaginary spaces are independent of how a network is laid out and what system software the computers are running. "This looks natural [to] me now, but some months ago we did not even suspect that... characteristics like the dimension of information flow in the parameter space are... not sensitive to network topology [or] operating systems," Gudkov said.

The researchers are working on a test model of a system that will detect known intrusions as they are happening, said Gudkov. If the research goes as expected, a model for detecting unfamiliar types of intrusions could be available within a year, and a practical working system a couple years after that, Gudkov said.

The researchers are also working on finding a way to detect unfamiliar intrusions by analyzing all the data rather than just looking for known intrusion patterns. The challenge is finding a method of pattern recognition that will work in real-time data plotted in imaginary spaces that have more than three dimensions, according to Gudkov. "The next step for this is the study of multidimensional pattern recognition methods based on wavelet analysis," he said. Wavelets are a form of compressed data.

The researchers' idea of modeling network traffic characteristics as functions is an interesting one, but "the question of whether such a view is meaningful, or if it would lead to useful results," cannot be answered without testing the method on real networks, said R. Sekar, an assistant professor of computer science at the State University of New York at Stony Brook.

It is also difficult to predict whether it will be possible to find unfamiliar intrusions this way, according to Anita Jones, a professor of engineering and applied science at the University of Virginia. "Any mathematical approach depends upon detecting some properties that distinguish the intrusive traffic from normal traffic. Just as in real life, what is harmful can often be masked to appear benign. Such traffic can sometimes be very hard to distinguish from normal traffic," she said.

Gudkov's research colleague is Joseph E. Johnson of the University of South Carolina. The research was funded by the Defense advanced research projects agency (DARPA) and the Air Force Research Laboratory.

Timeline: 3 years
Funding: Government
TRN Categories: Networking; Internet
Story Type: News
Related Elements: Technical paper, "New Approach for Network Monitoring and Intrusion Detection," posted on the arXiv physics archive at xxx.lanl.gov/abs/cs.CR/0110019

Internet Vulnerabilities

Hubs Increase Net Risk

By Kimberly Patch, Technology Research News
January 1/8, 2003

The Internet has much in common with air travel, according to researchers from Ohio State University. This does not bode well, considering how disruptive storms can be to the airlines.

The commercial Internet has shifted from its original distributed structure toward a hub-and-spoke topology similar to those the airlines use to plot routes, and that shift has made the network more vulnerable, said Tony Grubesic, an Ohio State researcher who is now an assistant professor of geography at the University of Cincinnati.

A handful of cities, including Los Angeles, New York City, Atlanta, Dallas and Chicago, have become central to the Internet and have many more backbone connections than other locations, said Grubesic. These cities are "in effect, acting as hub cities," he said. Chicago, for example, had 23 direct connections to other cities on the AT&T network in the year 2000, versus three for Salt Lake City.

Although the hub-and-spoke topology is cheaper to build, hubs make the network more vulnerable to attack in the same way bad weather in a major hub city can affect flights all over the country, said Grubesic. "Where Internet survivability is concerned, this type of network topology is not a particularly effective one because it forces large volumes of traffic through a handful of cities," he said. "If one of the major points of presence in a city should fail, [for example] the metropolitan area exchange in Dallas, traffic would be disrupted nationwide."

The original topology of the Internet was more distributed, and was designed to withstand failure and provide service under adverse conditions - even a nuclear attack.

As the Internet has grown, however, the competitive nature of the Internet backbone provider industry has caused many providers to shift to the more vulnerable hub-and-spoke system in search of the most economically efficient network topology, according to Grubesic.

The researchers' analysis showed the overall vulnerability of the hub-and-spoke system for 41 network backbone providers. The most susceptible networks have the greatest reliance on hub-and-spoke configurations. The networks most susceptible to disconnection are AT&T, GTE, and Multacom, which would suffer significant performance hits and leave many smaller spoke cities without service with the loss of any one of eight, seven or six of the 14 largest hubs, respectively, according to the analysis.

In contrast, there are 11 network providers that use network topologies that resemble a mesh rather than a hub and spokes; these providers are robust enough to survive the loss of any of the largest hubs. These mesh-like topologies are more expensive to construct, but clearly have advantages where survivability is concerned, according to Grubesic.

To carry out the study, the researchers integrated information about a large set of Internet backbone networks into a geographic information system. "This allowed us to simulate a wide range of Internet disruptions and failures [and] examine... the topological and spatial impacts simultaneously," said Grubesic.

The researchers simulated what would happen if there were a catastrophic failure of the Internet at a hub city or an equally important backbone link. "We simulated the failure of four things: complete loss of a node, or city; loss of a backbone [provider]; loss of a single network node; loss of selected backbone links," said Grubesic.

If an entire hub were knocked out, service to the city in question would be impossible for any backbone. This is a fairly improbable scenario, especially because providers tend to maintain multiple connections in large cities, according to Grubesic. It is a vulnerability, however.

In one portion of the results, the researchers simulated the availability of the network of one provider — Multacom — after a complete node failure.

The city of Washington is the most accessible node on the Multacom network. The researchers showed that if all connections into Tampa failed, Washington would lose access to Tampa plus one other city — Miami. However, if all connections to New York failed, the ramifications for Washington would be much greater; Washington would lose access to New York, Chicago, Denver, San Jose, Portland and Seattle, according to Grubesic.

Worse, the simulation showed that if Atlanta, the most important node on the Multacom backbone, lost all its connections, Multacom communications would cease between Dallas, Los Angeles, Miami and Tampa and 10 other cities each, and between Chicago, Denver, New York, Portland, San Jose, Seattle and Washington and five other cities each.

This scenario is particularly problematic for spoke cities, which rely on the nearest hub.

The second scenario, the loss of a backbone provider, would leave cities serviced by a single provider completely without Internet service. Spoke cities would again be hard hit, according to Grubesic.

The third possibility, failure of a single network node within a city, is a smaller problem. Although this eliminates service to that node from a single provider, other backbones will remain, allowing traffic to continue, according to Grubesic.

But even a single network node failure would be problematic for spoke cities, because it effectively eliminates the delivery of all traffic destined for the node in question, said Grubesic. The large hubs, and cities served by several providers would do much better because they can reroute traffic.

In the fourth scenario, where select links in a network are severed, isolated nodes would lose service, but nodes that connect to more than one backbone would remain functional.

The methodology can also be applied to other types of networks, including critical infrastructure networks like electric, gas and oil, said Grubesic.

Two of the challenges in carrying out the study were creating code that simultaneously simulated node and link failure for a geographic information system, and developing intuitive ways to interpret the results, Grubesic said.

The researchers' analysis methods can be applied now to the Internet and other types of networks, said Grubesic. "One of our primary goals... was to provide a clear and understandable methodology for estimating the spatial impacts of node link failure for the Internet. This methodology can be revisited, duplicated and perhaps improved by other research teams interested in questions of Internet survivability," he said.

Grubesic's research colleagues were Morton E. O'Kelly and Alan T. Murray. The results are slated to be published in the February, 2003 issue of *Telematics and Informatics*. The research was funded by the National Science Foundation.

Timeline: Now
Funding: Government
TRN Categories: Internet; Computers and Society
Story Type: News
Related Elements: Technical paper, "A Geographic Perspective on Commercial Internet Survivability," *Telematics and Informatics*, February, 2003.

TRN

Hubs Key to Net Viruses

By Kimberly Patch, Technology Research News
November 7, 2001

When the Internet linked distant computers 30 years ago, its founders were probably not thinking about protecting the machines from infecting each other. Today's exponentially larger Internet, however, is vulnerable to software viruses in much the same way that large, crowded human populations are more likely to fall prey to biological viruses.

The Internet has a scale-free structure, meaning it has a few pages, or nodes with many connections to other pages and many with just a few connections. Researchers looking into how bits of disruptive code spread on the Internet have found that this structure isn't conducive to the conventional practices of inoculating large populations.

The researchers did, however, find an inoculation strategy that promises to protect computers more effectively.

When they applied an immunization strategy that's commonly used for biological populations to a simulated scale-free network, it simply didn't work, said Alessandro Vespignani, a research scientist at the Abdus Salaam International Center for Theoretical Physics in Italy.

The researchers inoculated progressively larger numbers of nodes, expecting the epidemic to eventually die out, he said. It did not even when they inoculated more than 90 percent of the nodes, he said. "Surprisingly, in scale-free networks we observed that infection survived... in the presence of massive vaccination campaigns involving the majority of the population. We realized that random... schemes were practically useless in scale-free networks."

The Internet is generally more vulnerable than human populations because the connections among computers are both more numerous and structured differently than many of the human connections that allow viruses to spread. Scale-free networks have some nodes — large portals, for instance — that contain more connections to other pages than even the most widely-traveled people could possibly have with other people.

The researchers eventually caused the epidemic to die out by targeting nodes that had a high number of connections rather than inoculating individuals randomly.

Using this scheme, the researchers sharply lowered the network's vulnerability to epidemic attacks, Vespignani said. "We have tested this recipe on a real map of the Internet [with] a targeted immunization involving all the most-connected individuals. In this case, by immunizing [less] than one percent of the total population, the cyber infection cannot propagate," he said.

The research explains why, though antivirus software is very successful in protecting individual computers, it does not prevent computer infection from becoming endemic. "The 'I love you' virus is still in the top list of most frequent viruses more than a year after its introduction... because the global implementation of antivirus [software] is practically equivalent to a random... vaccination," Vespignani said.

Ironically, this scheme could also be useful in the biological world where some of the paths viruses take to propagate in a human population have some similarities to the Internet. A map of human sexual relations, for instance, has scale-free properties, said Vespignani. The research implies that epidemics spread this way could be prevented more effectively by targeted vaccination of the few promiscuous individuals, he said.

This type of targeted vaccination would also prove to be much cheaper than the random kind, Vespignani said. "Instead of massive vaccination campaigns, we can think of identifying the network connectivity hierarchy." Controlling the hubs that spread the infection more quickly is both more effective and requires relatively few inoculations, he said. "The strategy is... particularly convenient in terms of economical and practical resources."

The problem in both the Internet and biological networks that harbor a scale-free nature is identifying the large hubs, said Vespignani. "The difficulty... is... detailed knowledge of

the network connectivity. This is not always possible for privacy and economical reasons. It is very difficult to obtain a complete map of the Internet because many providers do not want to share publicly their information. As well in the case of sexual diseases we have to rely on people's concerns about their own sexual habits," he said.

This strategy "looks reasonable. It is consistent with my experience," said Gene Spafford, a computer science professor at Purdue University. "I'm surprised no one else has noted this property in research... either in networks or in epidemiology," he said.

One complication that the model leaves out is the notion of workgroups, or local area networks where each machine is connected to all the other machines in that group, and an infection of one infects all the others, Spafford added.

It is hard to estimate when the research could be used to actually inoculate networks, said Vespignani. "The use of these results is strictly related to social factors — individuals' privacy — and the existence of control agencies." These make estimating the time frame difficult, he said.

Vespignani's research colleague was Romualdo Pastor-Satorras of the Technical University of Catalonia in Spain. The research was funded by the European Community, the Spanish Ministry of Education and Culture, the Abdus Salaam International Center for Theoretical Physics (ICTP) and the Technical University of Catalonia (UPC).

Timeline: Unknown
Funding: Government, Private
TRN Categories: Internet
Story Type: News
Related Elements: Technical paper, "Optimal Immunization of Complex Networks," posted in the Los Alamos physics archive at arXiv.org/abs/cond-mat/0107066

TRN

Scheme Harnesses Internet Handshakes

By Eric Smalley, Technology Research News
September 12, 2001

Whenever you click on a link on a Web site, your computer sends a message to the site's Web server and the server responds. Billions of such network handshakes take place on the Internet every day. Although individually these handshakes are insignificant, large numbers of them can add up to an impressive amount of computer processing power.

A team of researchers at the University of Notre Dame has figured out a way to use Web server handshakes to compute small pieces of a mathematical problem by disguising the pieces as ordinary Web browser messages.

The researchers' parasitic computing scheme uses the processing power of unwitting Web servers by exploiting one of the most basic operations carried out by all computers connected to the Internet, the Transmission Control Protocol (TCP) checksum, said Vincent W. Freeh, an assistant professor of computer science and engineering at the University of Notre Dame.

TCP breaks messages from one computer to another into small pieces, or packets, sends them over the network and reassembles them on the receiving end. The TCP checksum adds up the number of bits in the message and attaches the result to the message. On the receiving end, the computer adds up the number of bits received and compares it to the checksum number to make sure the message arrived intact.

The researchers performed their experiment with a type of math problem that can only be solved by examining each possible solution until the right solution is found. They encoded each candidate solution as a Web browser request for a web page so that the TCP checksum was actually checking to see if the message contained the correct solution.

The Web servers that received requests treated the messages that contained failed solutions to the math problem as corrupted messages and discarded them. The Web servers treated messages that contained the correct solution as a request for a Web page that did not exist and sent the standard 'page not found' error message to the researchers' computer.

Although the parasitic computing scheme demonstrates a principle, it is not a useful tool because the amount of computer resources used to implement the scheme far exceeds the amount that would be needed to solve the problem on the researchers' computer by itself, said Freeh.

"For a general communication protocol, I think the probability [of developing an efficient version of the scheme] is very remote," he said. "By design, the receiver doesn't have to do that much. However, I think people are [already] exploiting specific Web sites."

Web servers that run interactive applications and process forms are good candidates for this kind of scheme, said Freeh. "This is where lots of host cycles can be gotten," he said.

The parasitic computing scheme raises the possibility that computers on the Internet can be used in ways their owners are unaware of, which raises ethical and legal issues about the use of publicly available computer resources.

Though the Web server resources used in the Notre Dame implementation were barely measurable, if the scheme were used aggressively it could have a similar effect to denial-of-service attacks in which one or more Web servers are flooded with messages and effectively shut down, said Ian Foster, a computer science professor at the University of Chicago and a senior scientist at Argonne National Laboratory.

Each tiny message in the parasitic computing scheme is by itself indistinguishable from any other Web page request, said Freeh. "The way to tell is by seeing many such messages and deducing what is happening," he said. The researchers

have configured an intrusion detection system to detect their parasitic computing scheme and they are working on configuring the system to detect variations of the scheme, he said.

"The instance of parasitic computing that [the researchers] demonstrate... is totally inefficient, returning a minuscule amount of computation for great effort," said Foster.

The question is whether there are more efficient versions of such a scheme, he said. "Within the Internet infrastructure [it] seems very unlikely to me, given its fundamental simplicity." It's more likely, though still doubtful, that someone could develop an efficient scheme to exploit peer-to-peer networks like Gnutella, he said. "I wouldn't discount it totally, especially as these infrastructures evolve."

Even if computationally efficient versions of the scheme can be developed, it remains to be seen if it can perform useful work, said Miron Livny, a computer science professor at the University of Wisconsin.

"It's a creative idea [but] it's not clear to me how it will work if you really care about the result," said Livny. The problem is the scheme counts on messages that do not generate a reply to indicate that the message did not contain the correct solution, but failing the TCP checksum is not the only reason a message might not be returned. "The biggest challenge in distributed systems is to understand why somebody is not responding, because there [are] many, many reasons why you didn't hear back," Livny said.

The researchers tested the reliability of their scheme by repeatedly sending out the correct solution. They got the correct answer back at rates that varied from about 99 out of 100 to about 16,999 out of 17,000 times, according to Freeh.

Freeh's research colleagues were Albert-Laszlo Barabsi, Hawoong Jeong and Jay B. Brockman of Notre Dame. They published the research in the August 30, 2001 issue of the journal Nature. The research was funded by the National Science Foundation (NSF).

Timeline: Now
Funding: Government
TRN Categories: Internet; Distributed Computing
Story Type: News
Related Elements: Technical paper, "Parasitic Computing," Nature, August 30, 2001

TRN

Five Percent of Nodes Keep Net Together

By Kimberly Patch, Technology Research News
May 23, 2001

Because the Internet is a distributed network with no central server directing information flow, there are many potential paths from any given point on the network to any other point. This makes it a robust network that is difficult to shut down.

The Internet is also a scale-free, or power-law network, meaning it harbors a small number of very large hubs with many connections to other nodes, and a large number of nodes with only a few connections. This concentration of connections, a trait the Internet shares with large social and biological networks, makes it more vulnerable to intentional attack, however, than a network with more evenly distributed node sizes.

Researchers from Bar-Ilan University in Israel and Clarkson University are examining just how vulnerable the Internet's scale-free nature makes it. Knowing more about scale-free networks' vulnerabilities may point the way to both protecting the Internet from attacks and providing better strategies for attacking biological networks in order to fight disease.

While the Internet is made up of computers that are connected via communications lines to other computers, a typical biological scale-free network is made up of the molecules a cell uses. In this case, the network connections are interactions among molecules. The large hubs in a cell's chemical communications network include water and the cellular fuel ATP, which are used in many more reactions then most of the molecules it uses.

The researchers work shows that large scale-free networks are fairly impervious to random node breakdowns, but if large hubs are targeted methodically, even large scale-free networks can be broken up into separate islands. "We've studied the problem mathematically. According to our findings, while networks like the Internet are resilient to random breakdown of nodes, they're very sensitive to intentional attack on the highest connectivity nodes," said Shlomo Havlin, a physics professor at Bar-Ilan University.

This is because a scale-free network's stability depends on the state of its large hubs, he said.

In scale-free networks as large as the Internet, "there are just enough high connectivity nodes to keep the network connected under any number of randomly broken nodes," he said. "A random breakdown of nodes will leave some... highly connected sites intact, and they will keep a large portion of the network connected," he said.

An attack that targets about five percent of these highly connected sites, however, has the capacity to totally collapse the Internet, "very rapidly [breaking] down the entire network to small, unconnected islands," containing no more than 100 computers each, Havlin said.

The researchers cannot pinpoint the breakdown threshold any more precisely than near five percent, Havlin noted, because the exact distribution of nodes on the Internet can only be roughly estimated.

To find the threshold, the researchers used a branch of mathematics known as percolation theory, which was originally developed to predict how much oil can be pumped

from a reservoir. "Since oil can only flow through holes in the ground, this is similar to data flowing through... computers on the Internet," said Havlin.

Another way to picture percolation theory is to draw a square lattice of dots on a piece of paper. If you remove a small number of the dots, you can still connect the rest of the dots around the ones you have removed. "However, after removing the critical fraction [of dots] there's no continuous paths from side to side," said Havlin.

In terms of the Internet, "as long as we're above the threshold, there will be a large connected structure with size proportional to that of the entire Internet. Below the threshold, there will only be small unconnected islands of sizes in the dozens [of nodes] each," he said.

The researchers' work offers the theoretical basis for calculating the threshold for the breakdown of any complicated network, said Albert-László Barabási, a physics professor at the University of Notre Dame. "By offering a method to calculate... the number of nodes required to be removed in order to destroy the network by breaking it into isolated clusters, it will be of great use [in] fields ranging from Internet research to drug delivery, where the goal is, [for example,] to destroy some microbes by gene removal. I expect this result will have a lasting impact on our understanding of the resilience of complex networks in general," he said.

The researchers' aim is to find ways to design networks that are more resilient to both random error and intentional breakdown, said Havlin. The work may also lead to better understanding of network traffic and virus propagation on the Internet, he said.

Havlin's research colleagues were Reuven Cohen and Keren Erez of Bar-Ilan University in Israel, and Daniel ben-Avraham of Clarkson University. They published the research in the April 16, 2001 issue of Physical Review Letters. The work was funded by the Bar-Ilan University and the Minerva Center.

Timeline: Now
Funding: Institutional, University
TRN Categories: Networking
Story Type: News
Related Elements: Technical paper, "Breakdown of the Internet under Intentional Attack," Physical Review Letters, April 16, 2001

TRN

Net Inherently Virus Prone

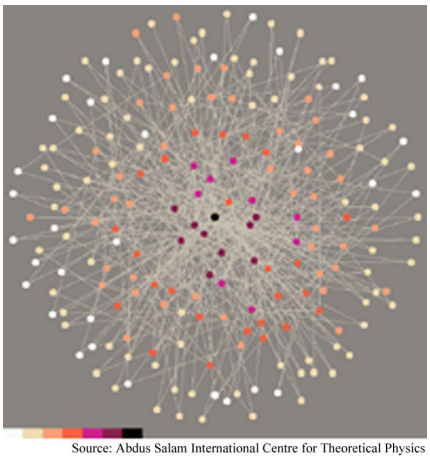
By Kimberly Patch, Technology Research News
March 21, 2001

The Internet's sheer size and large central hubs make it an efficient communications network, but those same traits make it vulnerable to the uninvited bits of code that are the computer's equivalent of biological viruses.

Two physicists have applied their understanding of condensed matter physics, which examines the complex, collective behavior of matter, to mapping how viruses traverse the Internet's complicated labyrinth of connections.

What they have found is that the Internet's efficient communications structure may make it vulnerable to even the weakest of viruses.

Standard epidemiological models look at how virulent biological viruses are. The more virulent, or easily spread a virus is, the larger the risk that it can spark an epidemic. If the virulence falls below a certain threshold, however, the infection will die out exponentially fast and therefore cannot spread fast enough to become a threat.



Source: Abdus Salam International Centre for Theoretical Physics

This color density plot shows the average number of times each node on a 200-node, scale-free network was hit with a virus. The hit rate is directly proportional to the number of connections a given node has, with very large hubs having the highest probability of being infected.

In order to study virus spread within the Internet, the physicists took into account the network's scale-free structure. The Internet harbors a few extremely large hubs, or nodes with huge numbers of connections and, many nodes with only a few connections.

In contrast, the hubs in social connections are more limited in size. "Real viruses can be transmitted only by close physical contact, and so diffuse in the community in a series of short hops between infected and uninfected individuals," said Alessandro Vespignani, a research scientist at the Abdus Salam International Centre for Theoretical Physics in Italy. "The crucial difference is that computer viruses spread on the Internet, which has a very special branching structure so on the Internet viruses can always pervade the system," he said.

The researchers found that this type of structure allows the epidemic threshold to fall below zero, meaning that no matter how low a virus' virulence, it won't necessarily die out. "Strikingly, we found that the Internet lacks any epidemic threshold. The Internet is prone to the spreading and the persistence of infections [no matter how low their] virulence," said Vespignani.

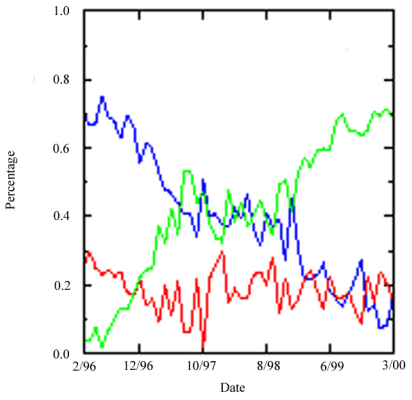
The Internet's structure also explains why computer viruses affect only portions of the Internet. Despite spreading easily and being able to survive for a long time, any single

virus has a low probability of infecting the bulk of the Internet, according to the researchers.

Ironically, "the ideal world for data sharing and fast communications... is also an ideal environment for viruses, which easily find... ways to rapidly infect new hosts through the intricate digital highways," he said. "The connections between computers on the Internet have enormous fluctuations and intricate structure that has to be included in the theoretical and experimental study of digital epidemics."

The lack of a threshold is surprising and potentially important for other scale-free networks as well, said Albert-László Barabási, a physics professor at Notre Dame University.

Network models have until now shown that viruses invariably die out if they're not too contagious, said Barabási. "However, [these models were] based on... outdated ideas on the topology of real networks," he said. In recent years it has become increasingly clear that many networks, including the Internet, are scale-free and have inhomogeneous topologies dominated by a few highly connected hubs, he said.



Source: Abdus Salam International Centre for Theoretical Physics

This graph shows the prevalence of the three major types of computer viruses. The red line shows viruses that reside in applications, the blue line viruses that hide on the boot sector of a hard drive, and the green line macro viruses that infect data files word processor documents.

The lack of a threshold for virus spreading in scale-free networks "is highly unexpected and it will have a significant effect on a number of fields," said Barabási.

The model could be used to understand epidemic dynamics in scale-free networks like "food webs, power grids and social networks," said Vespignani. It could also be applied to problems like the spread of polluting agents, he said.

There's another issue worth looking at that may make understanding how computer viruses spread more complicated, however, said Jon Kleinberg, an assistant professor of computer science at Cornell University. "It's a subtle issue. What is really the network on which these viruses spread?" he said.

Instead of the full, scale-free network of the Internet, many computer viruses actually spread on subsets that look more like social networks that have limitations on how large a hub can be, he said. It is very common, for instance, for a computer virus to spread by sending itself to the e-mail addresses listed in an infected machine's address book.

While the large, central hubs on computer networks can have tens of thousands or even millions of connections, the size of even the largest e-mail address books are limited. "The real network on which viruses spread is an invisible network of who talks to whom sitting on top of the Internet, and that's a network that we have less ability to measure at the moment," said Klineberg.

The researchers are fine-tuning their models to see how effective things like immunization could be on the Internet. "We're introducing more details and realism in the model. For instance, we're considering the presence of immunization, latency times, [and] detailed Internet maps," Vespignani said.

The researchers are headed toward a general theory of epidemiology and complex networks, he added. Such a model could help in devising algorithms to protect the Internet from a virus epidemic, he added.

Ultimately, the Internet needs a global immunization organization in order to establish the optimal policies of immunization and antivirus implementation, said Vespignani. "We claim that the Internet needs a digital immune system... that automatically detects and submits viruses to some central control laboratory," he said.

Vespignani's research colleague was Romualdo Pastor-Satorras of the Polytechnic University of Catalonia in Spain. The research was funded by The International Centre for Theoretical Physics, the Polytechnic University of Catalonia, the European Community Network and the Spanish Ministry of Education and Culture.

Timeline: Not Available
Funding: Government, Private
TRN Categories: Internet; Networking; Cryptography and Security
Story Type: News
Related Elements: Technical paper, "Epidemic Spreading in Scale-Free Networks," scheduled to appear in the April 2, 2001 issue of Physical Review Letters and is posted on the Computing Research Repository (CoRR) at arXiv.org/abs/cond-mat/?0010317

TRN

Privacy Protection
Sensors Guard Privacy

By Kimberly Patch, Technology Research News
July 16/23, 2003

In a world where sensor networking and location tracking technology is becoming increasingly sophisticated and prevalent, preserving privacy is an increasingly difficult challenge.

Researchers from the University of Colorado at Boulder have addressed the problem with a way to set up networks of tiny sensors that allows users to gain useful traffic statistics

but preserves privacy by cloaking location information for any given individual.

"We realized that privacy policies tend to get incredibly complex because they need to define in detail under which circumstances who should get access to what information," said Marco Gruteser, a researcher at the University of Colorado.

At the same time, privacy is poised to get even more complicated, said Gruteser. Researchers are developing wireless networks of tiny, cheap, powerful sensors. "Sensor network technology promises to enable vast arrays of sensors monitoring many aspects of our daily lives," he said.

Such high-precision data can give away the identity of the people being monitored, said Gruteser. For example, an array of infrared sensors, which key off the heat human bodies emit, may track movements throughout an office building. "If such movements [were] correlated with knowledge of individuals' office locations, most monitored subjects could be identified simply by checking... which place they spend most of their time," Gruteser said.

The researchers' software uses the computational abilities built into today's sensors to automatically adjust the precision of location data and removes obvious identifiers like names in order to make such correlations more difficult.

The researchers' algorithm defines the strength of privacy as the number of people any given individual is indistinguishable from. The software can be set to a minimum level of privacy, for instance, to make it impossible to distinguish a given person from five of her colleagues. "The algorithm monitors the overall number of people and adaptively changes the precision of reported locations—say, from a room level to a building floor level—to maintain the predefined minimum level of privacy," he said.

The researchers' key idea was realizing that networks of sensors offer enough computational power to allow for such privacy-enhancing algorithms, said Gruteser. The main challenge to turning sensor nodes into a trustworthy, privacy-protecting network was designing an algorithm that distributed the needed computation among a large number of sensors, he said.

Distributing the information among nodes yields another advantage, said Gruteser. "If an adversary succeeds in compromising one sensor node, he does not gain access to all information collected by the network," he said.

Putting privacy protection in the sensor network means there are fewer places where information can be hacked. Such protection is usually set up further downstream, in the applications or databases that use the information collected by sensors. "Our work applies... anonymity or depersonalization on-the-fly to a stream of location data before it can be stored in a database that might be exposed to inside attacks or... inadvertent data disclosures," said Gruteser.

The drawback to the approach is that it is only suitable for applications that do not require user identification and can deal with less precise location data, said Gruteser.

Appropriate applications include monitoring the use of facilities, tracking the availability of meeting rooms and offices, collecting retail store traffic statistics to improve store design and product shelving, and tracking vehicles in order to better manage traffic, he said.

Privacy-enabled sensor networks would allow service providers to avoid privacy software and policies at the database and application level, he said.

To ensure privacy, an organization that wishes to collect data would have the sensor networks certified by a third-party agency, said Gruteser. The process could be similar to the federal information processing computer security standard for cryptographic modules, he said. "This should increase user trust in the deployment and decrease organizations' exposure to privacy liabilities," he said.

The scheme is a collection of techniques that have been applied separately in other contexts, and a nice piece of design that shows how to avoid aggregation, said Gene Spafford, a professor of computer science at Purdue University. "This is the first time I've seen that all combined into the context of sensor systems," he said.

The scheme addresses a general issue of privacy concerns, said Spafford. "It is difficult to preserve privacy and also meet necessary accuracy constraints," he said.

But preserving privacy at the sensor level may turn out to be too vague a scheme for many applications, said Spafford. "For instance, if I'm installing a building alarm system for after-hours use, I want to know exactly how many people are in each room," he said.

The scheme also requires that all sensors can be trusted and that they can't be tempered with, said Spafford. "That may or may not be realistic in actual application," he said.

The researchers are ultimately looking to provide a way for users to remain private if they wish to, said Gruteser. "Research in sensor and wireless networking will, as a side-effect, dramatically increase the potential for data collection and surveillance," he said. "Our research seeks to provide a toolkit of techniques that enable users to protect their location privacy when desired," he said.

The privacy-aware location sensor networks could be used practically in 3 to 6 years, according to Gruteser.

Gruteser's research colleagues were Graham Schelle, Ashish Jain, Rick Han and Dirk Grunwald. They presented the work at Usenix HotOS IX: 9TH Workshop on Hot Topics in Operating Systems, in Lihue, Hawaii, May 18-21, 2003. The research was funded by the National Science Foundation (NSF).

Timeline: 3-6 years
Funding: Government

TRN Categories: Cryptography and Security; and Data Acquisition; Computers and Society
Story Type: News
Related Elements: Technical paper, "Privacy-Aware Location Sensor Networks," presented at at Usenix HotOS IX: 9TH Workshop on Hot Topics in Operating Systems, in Lihue, Hawaii, May 18-21, 2003, and posted at www.usenix.org/events/hotos03/tech/full_papers/gruteser/gruteser.pdf

TRN

Scheme Hides Web Access

By Ted Smalley Bowen, Technology Research News
October 2/9, 2002

The ringing declaration that information wants to be free often bounces off a hard reality — the free flow of information can attract interference. The reality online is that censorship and surveillance are widespread and growing.

The everyday flow of ordinary Internet traffic, however, could provide cover for political dissidents, whistleblowers, or anyone else who wants to access censored information online without the activity being recorded or blocked by others.

Researchers from the Massachusetts Institute of Technology have come up with a scheme that could guarantee users access to data in such a way that their actions could not be monitored.

The development follows an age-old pattern. Strictures on communication traditionally provoke workarounds, from prisoners tapping on cell bars to con men gaming early telegraph systems to get the jump on stock market or horse race results.

Latter-day examples have played out on the Internet for years. Proxy software allows users to surf anonymously, covering virtual tracks by masking Internet Protocol addresses and other personal information; and the Web's hypertext transfer protocol — HTTP — allows users to encrypt requests for information. But these solutions have not proved watertight.

Proxy software, which serves as an intermediary to let people access Web pages anonymously, can draw attention and be blocked by censorship software. Common security protocol software can also fail to protect users' identities, and it can be stymied by firewall software.

The MIT researchers' scheme, dubbed Infranet, allows Internet users to navigate using standard hypertext transfer protocol without being noticed.

The key to the scheme's ability to allow users to avoid monitoring is that it handles covert communications without adding a conspicuous amount of traffic. To be useful, a covert Internet communications system needs to cloak transmissions well enough to foil most would-be detectors, but must also be efficient enough to permit reasonably speedy browsing.

Infranet consists of software for Web servers and browsers. The scheme's responder software runs on public Web servers that store or are able to access data that is blocked or banned for some parts of the Web. Its requester software runs on systems seeking secure access to that data.

The software employs a transmission cloaking method, tried-and-true public-private key and shared session key encryption mechanisms, and existing data-hiding schemes.

Public-private key encryption allows anyone to use a receiver's freely-available public key to encrypt a message so that only the receiver's private key can decrypt the message and access its contents.

A shared session key is a single key that can be used to decrypt the messages it was used to encrypt.

To gain access to blocked data using Infranet, the requester begins a session by sending a shared session key using a responder's public key. "As long as either the requester or responder know how to communicate with the other initially, they can come to agreement on the session key," said Nick Feamster, a researcher at MIT's Laboratory for Computer Science.

The responder then uses the session key to send code to the requester that translates hypertext transfer protocol traffic into a kind of alphabet that will allow the requester to hide ensuing transmissions to the responder within ordinary requests for non-censored Web pages.

This coded alphabet is made up hypertext transfer protocol requests for pages on the responder's Web site, and the code is different for each requester. A request for a covert Web page consists only of a series of requests for permissible Web pages on the server.

The order and timing of the requests for openly available pages determines the covert request. "If the requester and responder agree on how visible HTTP traffic maps to hidden messages, then everything works," said Feamster.

The responder uses the shared session key to encrypt the requested information, uses separate data-hiding techniques to embed the encrypted information in non-censored material, and sends that material to the requester as ordinary hypertext transfer protocol traffic.

The scheme currently calls for hiding the data served to the requester in JPEG's, one of several types of image files that can be transferred using the hypertext transfer protocol. In theory, responders can hide data in many types of files served up by Web computers, including MPEG video streams, said Feamster. "Our basic philosophy is to leverage existing steganography and data hiding techniques for the downstream communication," he said. In downstream communication served to the requester, "we're dealing with a pretty traditional data hiding problem," he said.

Although the researchers chose to conceal the requested information in JPEGs, and embed requests in the order and timing of hypertext transfer protocol requests, the method could work with any number of bi-directional

communications, said Feamster. "Many possibilities exist: instant messaging, news feeds, stock tickers, satellite radio, online games, just to name a few," said Feamster.

The main qualification of a suitably innocuous scheme is that the communications be largely unidirectional, with more downstream than upstream traffic. The cloaked requests need only contain small amounts of information, while the responses pack the censored data into larger, more ordinary files that are openly sent to the requester. This fits well with the uneven nature of most Web communications: requests for data typically require much less bandwidth than serving up that data.

The researchers tested Infranet by subjecting it to passive attacks by monitors that logged all transactions and packets passing through a given segment of the Internet, and to active attacks by detection schemes that mimicked Infranet systems.

The process of covertly requesting and then serving up data hidden within other files turns out to be reasonably efficient. Half of the researchers' tested requests fit in six or fewer served files, and 90 percent of the requests required ten or fewer files. The requested files could be concealed in typical Web images by adding about 1 kilobyte of hidden data to each ordinary transmission, which typically range between 5 and 50 kilobytes.

One potential drawback of with this type of scheme is that users might suspect that the scheme itself is a surveillance tool. This can probably be addressed by including existing mechanisms that ensure that users can trust downloaded software, Feamster said.

Another issue is how to conceal the initial download of the Infranet software, a problem the researchers are currently addressing, said Feamster. Physically distributing the software via disks is one way to minimize the risk of disclosure.

For a scheme like Infranet to succeed, the responder software would have to be installed on a considerable number of public Web servers. "We're thinking of starting with something on the order of 50 to 100," Feamster said. If the responder software were bundled with a Web server like Apache, active participants would be much harder to detect, according to Feamster. The researchers' requester prototype is an Apache module.

"The trick is that you need to allow clients to discover the responders," Feamster said. "But if it's too easy to discover all of them, the censor can simply block them. Thus, we have to have enough to make it difficult for the censor to keep up with where all of the responders are."

In the cat-and-mouse contest that pits censorship and surveillance against the free flow of information, time works against such schemes, according to Avi Rubin, a secure systems researcher at AT&T Labs. "[It] illustrates an arms race. Once the adversary, in this case, a censoring government, knows about Infranet and how it works, they can attempt to detect and block it," he said.

Infranet is an impressive, novel scheme, said Rubin. "This is a big step forward towards evading that kind of censorship," he said. "It's actually going to be a bit of work for the censoring bodies to counter this, so it forces them to put in some additional effort, thus raising the cost of censoring."

Infranet could probably be optimized to allow more information to be exchanged without detection, Rubin said. "They could eventually develop high-bandwidth covert channels," he added.

Feamster's MIT colleagues were Magdalena Balazinska, Greg Harfst, Hari Balakrishnan, and David Karger. The researchers presented the work at the 11th USENIX Security Symposium in San Francisco, August 5 through 9, 2002.

Timeline: < 6 months
Funding:
TRN Categories: Computers and Society; Computer Science; Cryptography and Security; Internet
Story Type: News
Related Elements: Technical paper, "Infranet: Circumventing Web Censorship and Surveillance," Proceedings of the 11th USENIX Security Symposium, San Francisco, California, August 5-9, 2002, and posted at www.usenix.org/publications/library/proceedings/sec02/feamster.html

TRN

Rating Systems Put Privacy at Risk

By Ted Smalley Bowen, Technology Research News
July 25, 2001

The Internet has given us new ways of carrying out activities as diverse as shopping and political agitation, and many of these new modes share a strong dependency on the medium's shaky guarantees of privacy and anonymity. This uncertainty has led to a variation on the trap of guilt by association: the threat of exposure by indirect association.

The chance you take when you use a Web recommender is typical of this new jeopardy, which researchers at three U.S. universities have quantified into basic equations of risk and benefit.

A Web recommender, or recommendation system, is a consumer rating system popular with online buyers of books, movies, and other items whose merits are a matter of taste.

A Web recommender may, for example, suggest to a person who has rated only books about baseball that he might also like a book about ballet. The recommender would have this information if another person had rated books on both topics. The recommendation system could unearth this connection using a nearest neighbor algorithm, which searches for the query point, or data point nearest the reference.

In this example, the recommendation system, while supplying a form of advice, has also showed the baseball fan a weak tie, which in social network theory is a connection

between groups that don't ordinarily interact. A malicious user could exploit this seemingly incidental piece of information, according to the researchers.

On the Web, weak ties can be combined with other information to trace individual users' identities. Such tracing robs users of the option to act anonymously, and can be used to mine personal, financial, political and other information and affiliations.

Even though the risks are intuitively apparent, it's difficult to quantify the odds of weak tie exposure.

Toward that end, a group of computer scientists from the Virginia Polytechnic Institute and State University, Purdue University and the University of Minnesota has analyzed the risks of exposure by mapping the types of connections users make — often unconsciously — when participating in recommendation systems.

"Our main goal was to quantifiably assess the benefits and risks," said Naren Ramakrishnan, a professor of computer science at Virginia Tech. Everybody talks of risks in terms of 'don't disclose credit card', 'don't disclose age and address'. But we hope to identify more subtle forms of risk involving seemingly harmless information," he said.

The researchers did this using graph-theoretic models, which show relationships and connections among entities in a way similar to family trees, highway maps and organization charts. By mapping exposure risk, the researchers quantified the risks and benefits of recommendation systems in general.

"In our case, we use a graph-theoretic model to represent the connections between people and the artifacts they rate," Ramakrishnan said. Recommendation systems make connections between people based on their common recommendations. Such connections, or jumps, move beyond the common items to the people who rated them.

These jumps can be represented as social network graphs, which depict people and how they are related. Recommender graphs go a step further and include the artifacts, or items that people have rated in common. With this information, it's possible to find the connection between a user making a query and one who has rated the item of interest, according to Ramakrishnan.

Although it's laborious, a user could game the system and sift for connections that can be traced back to individuals, said Ramakrishnan. "By varying the ratings, you might notice that the recommendations change," he said. "In addition, you might notice that a particular recommendation of book X happens only for some specific values for ratings. If you know something about the algorithm behind the recommender system, then you could reverse-engineer the rating by inspecting the behavior of the algorithm."

To calculate the risk and benefit inherent in a given recommendation system, the researchers drafted a rough formula: benefit = w/l2, where w is a connection or connections between people who have rated the same item or items and l is a sequence of such connections.

"The... higher the w, the higher the benefit. The lower the l, the higher the benefit. The "squared" is there to make the second statement a little stronger than the first," Ramakrishnan explained.

This formula applies to any recommender system that works by making connections, which is how most of today's e-commerce recommender systems work, said Ramakrishnan. "Its limitations are that it might have to be adjusted for individual domains. The formula as it stands is a good qualitative measure, nevertheless," he said.

The key is presenting risk in terms of how a person relates to the larger social context of a recommender system, he said. "Thus, the same person with the same ratings may not be at risk in a recommender system where he is just like everybody else; it is his uniqueness [within a given system] that is posing the risk."

The risk equation can be likened to the way an individual can be singled out in a crowd, said Ramakrishnan. "If you look like everybody else, nobody can single you out. If you wear crazy clothes, you will be immediately spotted. Similarly, if you rate like everybody else, sure you get along and there is no danger," he said. "If you rate crazily, on the one hand you provide a lot of benefit to the recommender, but then you are at risk."

The researchers are aiming to demonstrate the risks inherent in such rating systems and broaden the context in which they are considered, said Ramakrishnan. "We're still studying this area," he said. They are looking into the causes of weak links, looking for other ways of quantifying benefit and risk and are looking to derive new ways to manage recommendation systems, he said.

The use of social network theory to study Web dynamics is compelling, although the seriousness of these risks is debatable, said David Madigan, a professor of statistics at Rutgers University.

"Making the connection with the social network literature is fascinating. [But] is the privacy threat real? I don't think so," Madigan said. The researchers' example of identifying someone through their ratings seems "far fetched in the context of large-scale e-commerce," he said.

A more likely threat comes from old-fashioned violations of privacy agreements, according to Madigan. "While I might trust, say, amazon.com, a less trustworthy e-tailer might try my name and password on lots of other sites and get a complete picture of all the stuff I buy," he said.

Ramakrishnan's colleagues were Benjamin J. Keller and Batul J. Mirza of Virginia Tech, Ananth Y. Grama of Purdue University, and George Karypis of the University of Minnesota.

Timeline: Now
Funding: University
TRN Categories: Internet
Story Type: News

Related Elements: Technical paper, "When being Weak is Brave: Privacy Issues in Recommender Systems," posted on the Computing Research Repository at http://xxx.lanl.gov/abs/cs.CG/0105028

TRN

Fault-Tolerant Free Speech

By Kimberly Patch, Technology Research News
July 12, 2000

Free speech on the World Wide Web sometimes lasts only as long as it takes a secret police agent, judge, or corporate lawyer to swing into action.

In an effort to strengthen free cyber speech, two scientists from AT&T labs and a New York University grad student have put together software that allows for Web publishing that is both anonymous and difficult to remove.

The software, called Publius after the pen name of the Federalist Papers authors, encrypts a file and publishes it on many sites. Because it is encrypted, however, the sites carrying it cannot read it. Publius then breaks up the encryption key and sends single pieces, or shares, around to the sites. But the publishing sites still can't read the encrypted document.

"The Web servers have no idea what's being stored at their site because they only get one share and the encrypted file — they don't get information on where the other shares are," Avi Rubin, a research scientist at AT&T Labs.

To read a Publius document, a person must download a copy of it, then download a certain number of shares to reconstruct the key, then use the key to decrypt the document. The publisher can set the number of key shares needed to reconstruct the key. The default is three of twenty available shares.

This way, shutting down one, or even most of the servers carrying the document will not preclude people from continuing to access it. "If somebody shuts down 15 of the servers as long as there are still five or even three [available], then the key can be reconstructed," said Rubin.

Rubin said the researchers' goal is to make the Web more censor resistant. "There are numerous examples where there might be pressure on someone to take down a Web site that somebody else doesn't like," Rubin said. "Imagine a very powerful chemical company that's dumping chemicals a river," said Rubin. "If you want to make people aware of it... but you don't want any retribution and you don't want [the company] to be able to take it down... then you might want to publish it on Publius," he said.

The researchers are making the software available for a two-month Internet trial starting July 28. It can be downloaded at cs.nyu.edu/waldman/publius.

After the trial, the researchers will make adjustments to the software and either continue the trial or do a new deployment, said Reuben. "Our goal is to have this thing existing and widespread on the Web," he added.

Avi Reuben's colleagues in the research are AT&T Labs Research Scientist Lorrie Cranor and New York University Ph.D. student Mark Waldman. They are presenting a paper on the subject at the Usenix Security Symposium in August.

The project was funded by a Usenix student grant awarded to Waldman.

Timeline: Now
Funding: Association; Corporate
TRN Categories: Internet; Computers and Society
Story Type: News
Related Elements: Anonymous Publishing Website

TRN

Physical Security

Glowing Beads Make Tiny Bar Codes

By Kimberly Patch, Technology Research News
April 9/16, 2003

Researchers from Corning, Inc. have found a way to form tiny barcoded beads that are small enough to be embedded in ink and attached to DNA molecules.

The beads measure 100 by 20 by 20 microns, which is just at the edge of invisible. A micron is one thousandth of a millimeter.

The researchers made the coded beads by fusing together glass doped, or mixed, with lanthanide metal oxide ions. These metal oxides glow at certain wavelengths under ultraviolet light. Stripes of oxide that glow different colors can be used to make codes.

The researchers have proved that 100 billion unique barcodes are possible using the method, said Joydeep Lahiri, manager of biochemical sciences at Corning. "This could be pushed further," he added

The microbeads could be be embedded in inks as a way to tag currency and other documents to protect against counterfeiting, said Lahiri. They could also be used for security purposes in everything from automobile paint to explosives, he said.

The beads can also be used to keep track of different types of DNA or other molecules in drug discovery experiments, according to Lahiri.

The researchers made the beads by fusing together glass doped with lanthanide, drawing the mixture into a fiber, etching the fiber with a laser, then breaking the beads along the cuts by putting them in an ultrasonic water bath, said Lahiri.

There were three keys to developing the beads, said Lahiri.

The first was developing brightly-fluorescent glasses with good surface chemistry that did not interfere with organic labels, he said. DNA is often tagged with dye and identified by shining light on the dye and measuring the wavelength of the resulting glow.

It was a challenge to figure out which doped glasses "have distinguishable fluorescence to enable their decoding, but also do not interfere with the fluorescence emitted from biological materials," said Lahiri.

The second key was finding a way to fuse and consistently draw miles of banded ribbon fiber, he said. "Not only are they rectangular ribbons, but [at 20 microns] these are probably the thinnest structured glass fibers ever drawn," Lahiri said.

The third was being able to scribe the thin fibers. The researchers used a laser that put out light pulses that lasted only a few million billionths of a second.

Making the beads required the researchers to combine their knowledge of specialty glassy materials, optical fiber, surface chemistry and biochemistry, said Lahiri.

The researchers tested the microbeads in a gene expression assay, which determines which genes are expressed by a cell, said Lahiri.

The researchers' next step is to synthesize DNA and peptides on the beads. Biological assays, or experiments, like studies of gene expression or drug-protein interactions, can then be performed on the attached organic molecules, Lahiri said. "If we do the synthesis of the DNA or peptides on the coded microbeads [scientists can] order DNA attached to the encoded beads," he said.

The researchers have done some neat work that expands the still-limited repertoire of encoded bead technologies, said Shuming Nie, an associate professor of biomedical engineering at the Georgia Institute of Technology and Emory University.

The researchers have found "a novel method for fabricating microbarcodes," said Nie. "The most striking feature is perhaps the fiber bundling and pulling process, a new procedure that would not be anticipated from previous barcoding studies," he said.

The microbarcodes will be useful for applications like security tagging, but it is not yet clear if there are biological applications for the relatively large microbarcodes, Nie added.

The material has some drawbacks that may limit its practical use, Nie said. It emits light at multiple wavelengths, is relatively inefficient at absorbing light, and its long excited-state lifetimes will limit how quickly the codes can be read out, he said.

The technology could be ready for commercial use in three to six years, according to Lahiri.

Lahiri's research colleagues were Matthew J. Dejneka, Alexander Streltsov, Santona Pal, Anthony G. Frutos, Christie L. Powell, Kevin Yost, Po Ki Yuen and Uwe Muller. The research appeared in the January 6, 2003 issue of the *Proceedings of the National Academy of Sciences*. The research was funded by Corning.

Timeline: 3-6 years
Funding: Corporate
TRN Categories: Biotechnology; Materials Science and Engineering
Story Type: News
Related Elements: Technical paper, "Tiny Glowing Barcode Beads," *Proceedings of the National Academy of Sciences*, January 6, 2003

TRN

Plastic Tag Makes Foolproof ID

By Eric Smalley, Technology Research News
October 2/9, 2002

Shine a flashlight through a shattered window and you'll project a unique pattern onto any surface beyond the window. Move the flashlight to a new angle and you'll get another unique pattern, but one that looks more like the first than one produced by shining the light through a different shattered window.

A scheme that leverages this principle could make counterfeiting and forgery much harder to pull off.

Researchers at the Massachusetts Institute of Technology have made inexpensive identification tags, or tokens, that cannot be copied or altered by any known means. The tokens are small pieces of plastic containing tiny glass spheres that produce unique patterns of light when lasers shine through the tokens.

The tokens are "low-cost... unique, tamper-resistant and unforgeable identifiers," said Ravikanth Pappu, one of the MIT researchers who is now a founding partner at ThingMagic. "Everyday objects — envelopes, bank notes, passports, credit cards, et cetera — could have... tokens attached to them and thereby obtain a unique identity," he said.

At 10 by 10 by 2.5 millimeters, the tokens are about the size of an extra-thick thumb tack. They contain several hundred glass spheres that are less than a millimeter in diameter and spaced a tenth of a millimeter apart. The cost of the materials for the token is about one cent, according to Pappu.

The spheres scatter laser light, yielding speckle patterns that can be captured with a digital camera and mathematically converted into binary numbers. Each pattern is intricate enough to yield a 2,400-digit binary number.

The researchers' light-scattering scheme is a physical version of the one-way mathematics functions used to encrypt sensitive information like passwords and credit card numbers. One-way functions are easy to calculate in one direction, but extremely difficult to reverse.

The multiplying two numbers, for instance, is easy. Reversing the process to find the original two numbers from the answer, however, is much harder. The larger the answer, the more two-number combinations there are that could have been the originals.

The token presents a similar barrier. It is impossible to determine the exact arrangement of the spheres in the token by looking at the speckle patterns, but without knowing the exact structure of the token it is impossible to come up with the right patterns.

The token is not simply a bar code containing a single 2,400-digit binary number, however. Each time a laser beam passes through the plastic it produces a different number, even when it passes through at nearly the same angle. What makes each token unique is that the numbers produced by shining laser beams at very nearly the same angle are more similar to each other than to numbers produced by shining laser beams at the same angle through different tokens. Two numbers generated by different tokens differ by 50 percent, but two numbers generated by the same token differ by only 25 percent, said Pappu.

This means that comparing two numbers will show whether they were produced by the same token. A number from a token can be stored in a database that registers the identity of a token attached to an object. Verifying the identity of the object would entail shining a laser through the token at the same angle as the laser used to derive the number in the database in order to get another number, and comparing that number to the number in the database. Using two or more laser angles provides additional points of comparison.

Once a token has been verified, the number it supplied and the comparison number from the database are thrown out. Each token is capable of generating 1011, or 100 billion, different 2,400-digit binary numbers, enough to provide 1,000 numbers a day for 280,000 years, said Pappu.

The theoretical limit to the number of numbers a single token can generate is 1070, which is a much larger number, but increasing the number of possible numbers would also increase the cost of the system, said Pappu. 1070 can also be written as a 1 followed by 70 zeros. That number is 50 orders of magnitude larger than the estimated 1020 stars in the universe.

The linchpin of the scheme is the security of the token. Copying a token would be extremely difficult because matching the exact positions of the spheres in the token is far beyond the capabilities of today's technology, said Pappu. Getting the spacing of the particles wrong by less than a thousandth of a millimeter would change the entire speckle pattern, he said.

Even reproducing the patterns using other lighting techniques is impractical, and simulating them on a computer is currently impossible, said Pappu. Simulating light scattering off of even a single particle would require a supercomputer.

In addition, tampering with a token renders it unusable, according to Pappu. The researchers drilled a half-millimeter diameter hole one millimeter into a token, and found that the numbers produced afterward differed by 46 percent from the numbers produced before.

The researchers' proposal is a clever idea ideally suited for specific uses like arming nuclear weapons or storing code keys in home satellite receivers, said Eugene Spafford, a professor of computer sciences at Purdue University. "It won't supplant [software] methods, but it is a useful addition to the security tool box," he said.

There are several drawbacks to using a physical token, including the possibility that it will be lost or stolen and used by others, said Spafford. Shock, vibration, heat, cold and radiation could also the degrade the physical key to the point where it no longer works, he said. "The material chosen is important, as is the packaging," he said.

The physical one-way token is a promising idea, but it is probably only useful for authenticating physical items and transactions carried out in person, not for electronic transactions, said David Wagner, an assistant professor of computer science at the University of California at Berkeley. "It's not good for authenticating the identity of someone across a network, but it could be a valuable defense against counterfeiting," he said.

It will take time to validate how secure the researchers' proposal is, said Wagner. "Security is a conservative discipline. It takes years of analysis to build confidence in a defensive measure," he said.

The researchers are working on making the system practical and applying it to authentication problems, and are working out the theoretical connections between physical one-way functions and mathematical one-way functions, Pappu said.

The tokens could be used in practical applications within 12 to 18 months, said Pappu. "The system is quite simple," he said. "Most of the technical challenges are centered around packaging the token in the context of the application, and building readers to read those tokens," he said.

Pappu's research colleagues were Ben Recht, Jason Taylor and Neil Gershenfeld. They published the research in the September 20, 2002 issue of the journal Science. The research was funded by the MIT Media Lab Things That Think Consortium, the National Science Foundation (NSF), the MIT Media Lab, and IBM.

Timeline: 1-1 1/2 years
Funding: Corporate, Government, University
TRN Categories: Cryptography and Security; Optical Computing,
Optoelectronics and Photonics
Story Type: News
Related Elements: Technical paper, "Physical One-Way Functions," Science, September 20, 2002; TRN Letters page

commentary by Ross Anderson of Cambridge University
www.trnmag.com/Letters/Letters.html

TRN

Radio ID Locks Lost Laptops

By Eric Smalley, Technology Research News
September 4/11, 2002

The best security is the kind you don't have to think about. Researchers at the University of Michigan have taken that adage as their guide in developing an encryption system that could reduce the security risk from lost or stolen laptops.

The researchers' Zero-Interaction Authentication system combines two well-known security techniques: a hardware token that authorizes the person holding it to use a particular computer, and encryption software that locks and unlocks files on a computer. The user wears the token in the form of a watch or piece of jewelry.

Although most people would agree that securing data on a laptop is a good idea, if the system requires them to periodically re-enter their passwords or otherwise interrupt their work, "they'll figure out ways to work around it, or turn it off," said Brian Noble, an assistant professor of electrical engineering and computer science at the University of Michigan. "One of our philosophical touchstones is to make sure that there's no reason for the user to know [the security system] is there," he said.

Although ID cards with magnetic stripes are a good way to control access to buildings and rooms, when the technique is used for computers, many people simply leave the card in their computer's card reader, said Noble.

Under the researchers' scheme, the user enters a password into his laptop or handheld computer at the start of the day to link his token to the computer. Until the computer is turned off and as long as the token remains within a few feet of the computer, the files remain unlocked.

The computer and token communicate via radio signals, which are encrypted to prevent anyone from eavesdropping on and duplicating them. The token transmits encryption keys, which are binary numbers, that unlock a second set of encryption keys on the laptop. Those keys lock and unlock the files on the computer.

The computer continuously checks for the presence of the token, and if it fails to receive the token's signal, it locks all the files. The files lock within five seconds of the user walking away, and unlock in just over six seconds once he comes back into range. These times are short enough to keep the security system from entering the user's awareness, according to Noble.

The two-part key process is central to keeping the locking times short. Because the communications link between the token and the computer is slow, it would take too much time for the token's keys to lock and unlock the files directly. It takes much less time to lock and unlock an encryption key than an entire data file.

The linchpin of the system is, of course, the token, so if the user loses it he's locked out of his own data. "If you lose the token and you haven't escrowed the keys, then the [data on the] laptop is junk," said Noble.

You can leave a copy of the token's keys in escrow, say with your system administrator, and the escrow authority can generate a new token for you, he said. "In the meantime, the laptop is not usable," he added.

Similar technologies exist, according to Dan Wallach, an assistant professor of computer science at Rice University. "The main advantage here is the focus on usability, making the security happen where the user doesn't even notice it," he said.

The researchers' technology cannot work alone; it requires techniques for encrypting software, Wallach pointed out.

Practical applications for the technology will take between one and five years to develop, said Noble. The biggest challenge is probably going to be building a small enough token with a long enough battery life, he said.

The researchers' also plan to expand the idea to applications and other services beyond the file system, said Noble. This brings up a number of questions, he said. For example, in a ubiquitous computing environment where everything from your car to your whiteboard is computerized and networked together, how do the rules of the game change if you have a token that authenticates you in a 10-meter bubble, he said. "Just what are the implications of having authentication be a very short-term and transient property?"

Noble's research colleague is Mark Corner. They are scheduled to present the research at the International Conference on Mobile Computing and Networking (Mobicom '02) during the week of September 23rd in Atlanta. The research was funded by Intel Corporation, Novell, Inc., the National Science Foundation (NSF) and the Defense Advanced Research Projects Agency (DARPA).

Timeline: 1-5 years
Funding: Corporate, Government
TRN Categories: Cryptography and Security; Wireless Communication
Story Type: News
Related Elements: Technical paper, "Zero-Interaction Authentication," International Conference on Mobile Computing and Networking, Atlanta, September 23-28, 2002

TRN

Hot Spots Give Away Lying Eyes

By Kimberly Patch, Technology Research News
January 23, 2002

You can't always see a person blush, but a computer that tracks heat changes can sense even subtle shifts in the amount of blood in the capillaries that lie just under the skin.

Researchers from the Mayo Clinic and Honeywell Laboratories have come up with a way to measure these heat changes in a person's face in order to tell whether the person is lying or telling the truth.

The system consists of a high-definition thermal imaging camera and a computer. The camera takes pictures of heat emanating from a subject's face, and the computer provides a quick analysis of any changes.

Monitoring blood changes in the face is similar to the traditional polygraph exam, according to James Levine, a consultant at the Mayo Clinic. The polygraph lie detector test measures changes in a subject's breathing, pulse rate and blood pressure. It also measures sweating by sensing changes in skin conductance via electronics attached to the skin.

The thermal method measures infrared lightwaves, or heat, around the person's face. The infrared light shows up on the computer screen as red areas. The theory behind using thermal changes in a person's face to detect lying is similar to the principal behind the polygraph. When someone is not telling the truth there is likely to be instantaneous warming around the eyes, which is probably a natural response produced by the sympathetic nervous system, said Levine.

The thermal method is as accurate as traditional polygraph tests, according to Levine. It is also faster than polygraph tests and doesn't require the subject to be connected to a device, he said.

The researchers tested their theory and the system at the U.S. Department of Defense Polygraph Institute. Twenty volunteers were randomly assembled into two groups. One group was instructed to stab a mannequin, take $20 from it, then lie about what took place.

The thermal imaging system correctly identified six of eight of the subjects who were lying, and 11 of 12 who were innocent. The subjects were also put through traditional polygraph tests, which correctly identified the same number of guilty subjects, but correctly identified only eight of the 12 innocent subjects.

Thermal imaging operators would not need the type of training to carry out the tests that traditional polygraph tests require, according to Levine.

"The technique sounds interesting and promising," said Christoph Koch, a professor of cognitive and behavioral biology at the California Institute of Technology. "For mass security and screening applications, you need a technology that can rapidly, at low-cost and with a low false alarm rate, screen people."

The accuracy of polygraph methods is controversial, said Koch. However, if the thermal imaging technique is faster than polygraphs and if it is less prone to label truthful statements lies it is worth investigating further, he said.

The researchers are continuing to test the method and are aiming to turn it into a practical security application, said Levine.

Levine's research colleagues were Ioannis Pavlidis From Honeywell Laboratories and Norman L. Everhardt from the Mayo Clinic. They published the research in the January 3, 2002 issue of Nature.

Timeline: 2-4 years
Funding:
TRN Categories: Computer Vision and Image Processing; Applied Computing
Story Type: News
Related Elements: Technical paper, "Seeing through the Face of Deception," Nature, January 3, 2002

TRN

Sounds Attract Camera

By Chhavi Sachdev, Technology Research News
July 25, 2001

When Steven Stills penned the lyrics, "Stop! Hey, what's that sound? Everybody look what's going down," he was describing a natural phenomenon that we seldom think about consciously — sounds make us look. When people clap, shout, or whistle to get our attention, our heads instinctively swivel towards them. Imagine the potential of a robot that reacts the same way.

University of Illinois researchers are taking steps towards that goal with a self-aiming camera that, like the biological brain, fuses visual and auditory information.

In time, machines that use vision systems like this one could be used to tell the difference between a flock of birds and a fleet of aircraft, or to zoom in on a student waving her arm to ask a question in a crowded lecture hall.

The self-aiming camera consists of a video camera, two microphones, a desktop computer that simulates a neural network, and a second camera that records chosen information. The microphones are mounted about a foot apart to mimic the dynamics of an animal's ears.

The heart of the researchers' system is a software program inspired by the superior colliculus, a small region in vertebrates' brains that is key in deciding which direction to turn the head in response to visual and auditory cues. The colliculus also controls eye saccades — the rapid jumps the eyes make to scan a field of vision.

In determining where to turn the camera, the system uses the same formula that the brain of a lower level vertebrate like a barn owl uses to select a head-turning response, said Sylvian Ray, a professor of computer science at the University of the Illinois at Urbana-Champagne.

The computer picks out potentially interesting input and calculates the coordinates where sound and visual motion coincide, Ray said. "The output [is] delivered to a turntable [under] a second camera. The turntable rotates to point the second camera at the direction calculated by the neural network."

Source: University of Illinois

The self-aiming camera uses the microphones and camera in the foreground to find sounds and motion, which it then records with the camera in the top right corner.

In this way, the second, self-aiming camera captures the most interesting moving object or source of noise on film for further analysis, saving a human operator the chore of sifting through all the data.

The camera re-aims every second toward the location most likely to contain whatever is making the most interesting noise. Because it always chooses an estimate of the best location for all available input, the system works even if several motions or noises happen at once, according to the researchers.

Source: University of Illinois

The camera aims in the direction of the sound of a researcher clapping his hands.

The self-aiming camera could be used to pare down meaningless data captured by surveillance systems that use several cameras to take pictures of their surroundings, according to Ray. It could be used as an intelligent surveillance device in hostile environments, and for ordinary security, he said.

To have the system differentiate among different types of inputs, the researchers plan to add specializations that will give certain inputs more weight. In nature, different vertebrates respond to particular targets; a cat likes different sounds and motions than an owl, for instance, said Ray. "One specialization of the self-aiming camera would be to train it to like to look for human activity," he said.

"This is a nice example of exploiting ideas from biology to better engineer systems since this pairing of stimuli increases the reliability and robustness of the self-aiming camera," said John G. Harris, an associate professor of electrical and computer engineering at the University of Florida. A better understanding of the underlying neural mechanism is still needed, he said.

Eventually the system will have to deal with multiple objects as well as noise and room reflections, Harris said. "Such a system needs an attention mechanism in order to attend to objects of interest while ignoring others. This is a higher level behavior that requires different sets of neurons and is beyond the scope of the current demonstrated system," he said.

The system could be in practical use in two to three years, according to the researchers. Ray's research colleagues were Thomas Anastasio, Paul Patton, Samarth Swarup, and Alejandro Sarmiento at the University of Illinois. The research was funded by the Office of Naval Research (ONR).

Timeline: 2 -3 years
Funding: Government
TRN Categories: Neural Networks; Computer Vision and Image Processing
Story Type: News
Related Elements: Technical paper, "Using Bayes' Rule to Model Multisensory Enhancement in the Superior Colliculus," *Neural Computation,* 12: 1141-1164

TRN

Light Pipes Track Motion

By Eric Smalley, Technology Research News
July 2/9, 2003

Researchers at Duke University have devised a simple tracking method that promises to dramatically reduce the computing resources needed for computer vision systems that allow computers and robots to sense their surroundings.

The technique bridges the gap between full-blown computer vision systems, which precisely track moving objects but are computer-intensive, and simple, inexpensive motion detectors, which are much less precise.

Traditional computer vision systems use relatively sophisticated software and camera equipment; they are also limited to fairly simple models of physical space-camera relationships, said David Brady, a professor of electrical and computer engineering at Duke University.

The researchers' method dispenses with the complicated software and lenses and instead maps the angles of light radiating from a source by channeling the light through set of pipes onto a set of light detectors. As an object moves across the field of view, light reflecting from the object triggers some detectors but not others.

The method relies on a rapid prototyping system, which uses computer-controlled lasers to harden liquid plastic or fuse powdered metal, to make a custom set of pipes. The researchers calculate the necessary pipe angles for a certain task and use the rapid prototyping system to produce the structure.

The researchers made a prototype that monitored a moving light source at a distance of three meters. The 25.2-millimeter prototype has eight viewing angles, eight detectors and 36 pipes. Each pipe channels light from a given angle to a detector. Seven of the eight detectors monitor four angles and the remaining one monitors all eight. Each of the eight viewing angles spans five degrees, giving the device a 40-degree field of view.

When an object is in one position within the field of view, for example, it triggers detectors five, six, seven and eight, and when it moves to the next position it triggers detectors three, four, six and eight. A computer controlling the device simply has to know which combination of triggered detectors corresponds to which position.

In contrast, computer vision systems analyze every pixel in each digital video frame—usually 15 to 30 frames per second—to determine the borders of objects in the scene. The software tracks motion by comparing from one frame to the next the position of an object's pixels relative to background pixels.

At the other end of the scale, motion detectors like those that turn on backyard lights simply detect motion and don't track the positions of objects. They detect rapid changes in the intensity of infrared light hitting first one and then the other of a pair of side-by-side light detectors. The infrared light is typically produced by the heat of a human body, and the sequential triggering of the detectors is typically caused by a person moving across the motion detector's field of view.

The separate angles of the field of view through the researchers' structure allow for basic digital representations of moving objects, and the relatively low-tech detector array cuts down the amount of information a computer must sift through, according to Brady. "These sensors may be capable of reducing the data load in tracking... systems by several orders of magnitude," he said.

The lightened computational load could make object tracking much cheaper. The researchers are working on using the method to track vehicles and people in real-time, and have produced a prototype that tracks cars at a distance of 15 meters. "The sensors may also be useful in developing spatially-aware robots," said Brady.

The researchers are working with commercial partners to develop simple motion tracking systems using the technology, according to Brady. The system should be ready for practical use in the next year, he said.

Brady's research colleagues were Prasant Potuluri, Unnikrishnan Gopinathan and James R. Adelman. The work appeared in the April 21, 2003 issue of *Optics Express*. The research was funded by the Defense Advanced Research Projects Agency (DARPA).

Timeline: 1 year
Funding: Government
TRN Categories: Computer Vision and Image Processing
Story Type: News
Related Elements: Technical paper, "Lensless Sensor System Using a Reference Structure," April 21, 2003, Optics Express.

TRN

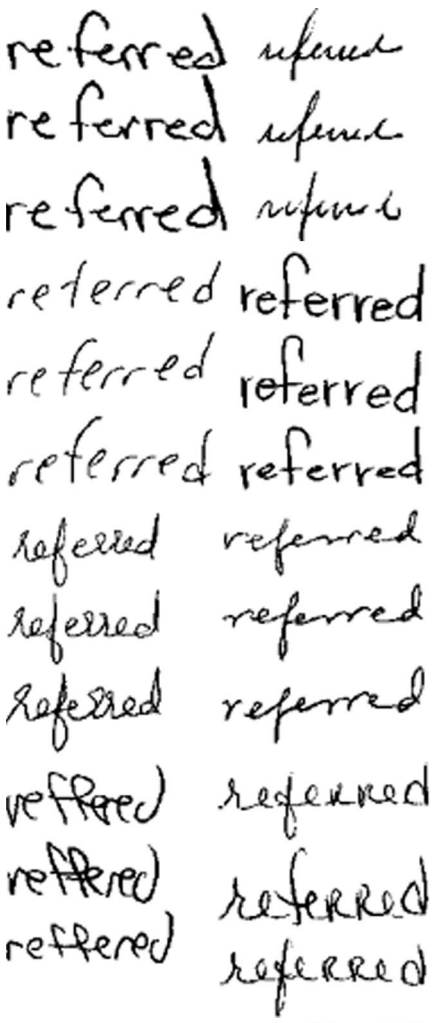
Software Spots Forged Signatures

By Kimberly Patch, Technology Research News
March 21, 2001

Researchers from the University of Buffalo are working on a handwriting analysis system aimed at determining who penned a ransom note or forged a check.

The move was motivated by several high court rulings that required expert testimony to be substantiated by scientific evidence, according to Sargur Srihari, a professor of computer science and engineering at the University of Buffalo and director of the school's Center for Document Analysis and Recognition. "Since individual handwriting had not been subjected to such study, we undertook this work," he said.

The researchers' handwriting recognition program extracts handwriting features like character shape, line and word separation, and stroke slant and thickness. To compare handwriting from different



Source: University of Buffalo

This shows a variety of handwriting styles. Handwriting analysis software compares features like loops and arches in handwriting samples to determine if the same writer wrote both samples.

documents, the program computes the differences in these features between the two samples measured, then determines if the differences fall within the limits of an individual writer's variability, according to Srihari.

In tests to determine whether or not two documents were written by the same person, the results were 95 percent accurate, said Srihari. In a second type of handwriting task, where the program determines which member of a known group of writers wrote a certain document, the performance varies from 98 percent for two writers to 89 percent for 1,000 writers, he said.

The program is based on a pair of well-known pattern recognition algorithms: Artificial Neural Network and Nearest Neighbor. It differs from conventional handwriting recognition in a key way, however. "Our work is about determining the variability between writers and within writers, [while] in handwriting recognition, the goal is to identify the message by averaging out the differences between writers," said Srihari.

The researchers' next steps are to capture finer features of handwriting in order to increase the accuracy. "For example, we currently are not measuring very many word-level features such as ascenders or descenders, or the presence of garlands or arcades," Srihari said. Ascenders and descenders are the parts of lowercase letters that extend above or below most other lowercase letters. Garlands resemble circles and arcades, arches.

Capturing features like these should make it possible to boost accuracy past 99 percent, he said.

The research is scientifically thorough, and is certainly useful, but is aiming for a difficult goal, said Nasser Sherkat, an associate professor of real-time machine vision at Nottingham Trent University in England.

"Variability within a writer' s [handwriting] is very high, especially when time elements and conditions of writing are taken into account. The ultimate question is whether we can tell if the writer is the same when the constraints [used in the study] have been removed," Sherkat said.

The program can eventually be used as the basis for human experts testimony in court, and eventually as a purely objective handwriting evaluation tool that doesn't require human intervention, said Srihari. This would be useful, said Sherkat. "Checking for forgeries manually is expensive and we could do with automating at least part of the process," he said.

The software could be available in less than a year, said Srihari.

Srihari's research colleague is Sung-Hyuk Cha of the University of Buffalo. The research was funded by the National Institute of Justice.

Timeline: < 1 year
Funding: Government
TRN Categories: Applied Computing
Story Type: News
Related Elements: Technical paper, "Handwriting Identification: Research to Study Validity of Individuality of Handwriting and Develop Computer-assisted Procedures for Comparing Handwriting," downloadable from www.cedar.buffalo.edu/NIJ/publications

TRN

Index

TRN's Making The Future Report is published 10 times a year by Technology Research News, LLC. Each 20- to 40-page package assesses the state of research in a field like biochips, data storage or human-computer interaction.

Single reports are $300 to $500. A one-year subscription is $1,600. To buy a report or yearly subscription, go to www.trnmag.com/email.html.

We welcome comments of any type at feedback@trnmag.com. For questions about subscriptions, email mtfsubs@trnmag.com or call (617) 325-4940.

Technology Research News is an independent publisher and news service dedicated to covering technology research developments in university, government and corporate laboratories.

Kimberly Patch
Editor
kpatch@trnmag.com

Eric Smalley
Editor
esmalley@trnmag.com

Ted Smalley Bowen
Contributing Editor
tbowen@trnmag.com

Chhavi Sachdev
Contributing Writer
csachdev@trnmag.com