The internet is a boon for business, communication and access to information — but also to cyber criminals and hackers. Researchers in the Computer Security Laboratory of the Concordia Institute for Information Systems Engineering are working diligently, and successfully, at beefing up security.
CONCORDIA'S EXPERTISE

Munir Debabi, who became the CISE's first hired faculty member in 2002 and then its director from 2007 to 2011, recalls receiving a recruitment email while sitting in Panamonic's Atlanta, Ga., research department, where he worked. "It is very rare to have an opportunity to create a department, rather than join one," he explains. "I was also charmed by the fact that it would be research intensive and interdisciplinary around information and systems engineering. I thought, 'This is an opportunity not to be missed.'"

Students agreed. With few other institutions offering graduate programs in information and systems engineering, enrollment in CISE's PhD, two masters and two graduate–certificate programs was up to full capacity at 100 entrants per year a few years after its founding.

Eleven years later, the 19 full-time faculty—computer scientists and electrical and mechanical engineers—work in three primary research areas. One of them, the Computer Security Laboratory (CSL), is home to six professors and more than 60 graduate students. This makes the CSL the country’s largest concentration of information systems security researchers. Rachida Dousouli, a professor and CISE’s founding and current director, says it also has "the highest impact in terms of reputation, as well as faculty such as Benjamin C.M. Fung, a former assistant professor who specializes in data mining, hoping to track down bad guys. (See the sidebar on page 41, “Forensics: Tools to bring criminals to justice.”)

Part of the magnetic attraction for money and students is the CSL’s integral role in the National Cyber-Forensics and Training Alliance Canada (NCFTA). The non-profit organization brings together academic institutions, government and law enforcement, and private companies to share resources, intelligence and expertise to stop emerging cybercrime threats and mitigate existing ones. Since it started in 2008, NCFTA has been headquartered at the CSL, ensuring Concordia scholars work on the latest, most relevant topics—a fast-moving target in the cyber world.

Debabi’s many research interests include network security, cyber- forensics and malicious code detection. "To design less vulnerable systems, you need to detect problems as they occur, prevent them, and also perform more in-depth detection research after the fact—forensics," he explains.

DETECTION: ROOTING OUT ATTACKERS

As president of NCFTA, Debabi has access to information feeds that monitor a wide variety of malware—malicious Internet Protocols (IPs) and domains, reconnaissance and intrusion attempts, and dedicated denial-of-service (DDoS) attacks, a major threat with an interesting local connection.

Back in 2000, a high-school student from Montreal’s West Island made history. Like most hackers of the era, Michael Calce—internet alias: Mafiaboy—was a young man who wanted to show off to other hackers. He had figured out how to control computers remotely by installing viruses via the internet, to link these compromised machines together into powerful networks called "botnets," and to instruct the linked computers to send packets of information to a receiving server simultaneously, thereby overloading that server and crashing it. Testing his method in February 2000, in the space of a week he took down the websites of Yahoo, eBay, E*Trade, Amazon and Dell, plus CNN and its 1,200 auxiliary sites. The attack caused an estimated $4.7 billion in damage and sent the stock market for a ride by demonstrating that in the dot-com boom, the new emperor, online commerce, was not just exposing a little midriff, completely unprotected from unsavoury elements, it was stark naked. Hacker groups worldwide took note.

Calce had perfected the DDoS attack, a powerful cyber weapon that renders one’s opponent inoperable. Anything online—military, banks, utilities—could be compromised by hacktivist groups like today’s Anonymous or even by governments. It is widely believed that Russian hackers protected Estonia’s decision to move a Soviet war memorial by unleashing just such an attack in 2007. The Estonian government, media and financial institutions’ sites all went down, virtually incapacitating a country which, like Canada, is one of the planet’s most wired.

As the CSL’s Mohammad Mannan says, early hackers groups running "botnets" were flashy and showing off—"look, I infected millions of machines in 15 minutes!" Yet as the technology matured, professional gangs monetized it. They shrunk botnets to avoid detection, and now rent them out by the hour to groups attempting DDoS attacks and phishing schemes, who spam to sell real or counterfeit products or spread propaganda, adds Debabi.

Most users automatically delete any spam that slips through the email service’s spam filter—an ad for cheap Viagra, or an ungrammatical help request from a "Nigerian prince" trying to move his money overseas. "But for me, it’s huge," says Debabi. "I can get significant levels of intelligence from spam."

Spam is actually the beginning of a broad–creminal trail to an individual hacker or crime syndicate. If a user responds to the forlorn Nigerian prince, the reply is received by the fake IP that is part of the botnet that sent it. Debabi’s live feed can detect the response and, through correlation, identify where the spammers are and what they are doing in real time.

Through NCFTA, the CSL has worked with Canada’s Department of National Defense and Ericsson Canada to research attack-detection and create attack prediction models. "We need to identify the servers that are phishing and take them down, by deriving the information from spam," says Debabi.

With so many working credit card numbers available that hacker groups sell them to fraudsters for as low as $4 each, and a full ID—date of birth, social insurance number, driver’s licence and photo—for only $8, this is a societal problem. Yet Canada lags behind the U.S. in information sharing for cybercrime mitigation purposes, says Debabi. That’s why this summer the CSL increased its capacity to become a U.S.-style data hub for information that carries little privacy value—spam and viruses—but can help protect us all.
FORENSICS: _TOOLS TO BRING CRIMINALS TO JUSTICE_

While citizens may be wary of their government shopping online into their affairs, there are many cases where government officials, or at least law enforcement, should and must intervene. For example, forensic specialists used to need to spend months sifting over emails, chat logs and text messages to assess evidence against child pornographers. In an age when keyboards have replaced pens, police hoped to identify suspects not by their written script, but by their writing style. That's why Benjamin CM. Fung, formerly a Core Mail Analyst at Information Systems Engineering assistant professor who specialized in data mining, and his former PhD student, Farshid Tohid, had been using email to identify the American company went bankrupt in 2001 after its corporate fraud scandal, Fung and Tohid identified its features and patterns found in other messages, known to be written by the suspect. Repeated grammatical mistakes, punctuation patterns, commonly used words, and spacing between paragraphs are all examples – thousands of unconscious writing habits. Then, they filter out any of these features also found in the emails of other suspects. The combination of remaining features is unique to the author of the message being analyzed. For example, while all suspects may type "he" instead of "her" only one type "the" of "the" or "a" to indicate "the" instead of "of" or "to" thinking "thinks" takes a plural verb. Fung calls this combination the suspect's "signature."

Police can use an IP address to determine the physical address from which electronic communications were sent. But, what if several people live at the same address? To determine who has authored a particular, telltale message from a pool of known authors, Fung and Tohid's software first identifies the features and patterns found in other messages, known to be written by the suspect. Repeated grammatical mistakes, punctuation patterns, commonly used words, and spacing between paragraphs are all examples – thousands of unconscious writing habits. Then, they filter out any of these features also found in the emails of other suspects. The combination of remaining features is unique to the author of the message being analyzed. For example, while all suspects may type "he" instead of "her" only one type "the" of "the" or "a" to indicate "the" instead of "of" or "to" thinking "thinks" takes a plural verb. Fung calls this combination the suspect's "signature."

They tested their software on Enron's email dataset – 200,000 emails from 158 users – which were made public by the American company went bankrupt in 2001 after its corporate fraud scandal. Fung and Tohid achieved an authorship accuracy of 80 to 90 percent, a rate highly valuable to law enforcement, which is already using the tool. And analysis that used to take months now takes mere hours. The next stage of research will be to apply the data-mining method to the even shorter levels of instant messaging, chat rooms and social media. This summer, Fung co-published with his PhD student, Gaby Daphne, a time-saving variant. "Out of all the types of available data in cybersecurity investigation, text data is the most common medium used by scammers, identity thieves and child exploitation criminals," says Fung. "But this type of data is also the most challenging to analyze. So, when a suspect's computer is seized, months may pass before sufficient information to press charges can be extracted from encrypted and messages in the hard drive. As Daphne explains, "In a normal search engine, a user enters some keywords and results can vary widely." Fung and Daphne's search engine, in contrast, captures the suspect's vocabulary, eliminating their slang — and then uses it to improve the speed and accuracy of the search results. The researchers new methods automatically identify the criminal topics discussed in the text conversation, show which participants are most active with the identified criminal topics, and then provide a visualization of the social networks among the participants. What took months now takes mere hours. The search engine can even help uncover hidden messages. Daphne recalls a recent case where the hard drive of a suspect arrested for sexual harassment was found to contain thousands of stolen credit card numbers.

Daphne, a self-described "theoretical guy" who normally works on data mining, encryption and the security of storing information through cloud computing, admits, "Getting real data from real cases is a challenge for the law enforcement agencies and helping them catch these people is very satisfying."