Cybersecurity and Artificial Intelligence

From Fixing the Plumbing to Smart Water

Carl E. Landwehr
Editor in Chief

Ray Kurzweil, inventor and futurist, predicts that by 2040 or 2050, machine intelligence will exceed human intelligence—an event he and others have dubbed the "singularity." Will such intelligent machines be better able to defend themselves than today's relatively unsophisticated ones? Will their intelligence be used for attacks as well?

In their early days, computer security and artificial intelligence didn't seem to have much to say to each other. AI researchers were interested in making computers do things that only humans had been able to do, while security researchers aimed to fix the leaks in the plumbing of the computing infrastructure or design infrastructures they deemed leakproof. Further, AI researchers were often most interested in building systems with behaviors that could change over time through learning or adaptation, and hence were to some degree unpredictable. From the security standpoint, unpredictable system behavior seemed undesirable. But the two fields have grown closer over the years, particularly where attacks have aimed to simulate legitimate behaviors, not only at the level of human users but also at lower system layers.

CAPTCHAs (Completely Automated Public Turing test to tell Computers and Humans Apart) are an amusing example of an intersection of AI and security today. Alan Turing posed as the ultimate AI challenge the ability to create a program that could respond to questions in a way indistinguishable (to a human) from responses generated by other humans— the "Turing test." Today, companies use CAPTCHAs to distinguish humans from bots in a wide variety of commercial contexts. The tests used typically depend on humans' superior ability to recognize distorted character sequences—a much narrower scope than a true Turing test would demand. Improvements in automated character recognition software, which might reasonably be considered an advance in AI technology, could drive the field toward more sophisticated pattern recognition tasks. So, in the process of trying to secure assets, such as online ticket reservations, the commercial security market is in a way stimulating advances in AI.

Although it wasn't identified as AI at the time, Dan Farmer and Wietse Venema's SATAN program, released in 1995, automated a process for finding vulnerabilities in system configurations that had previously required much more human effort. It also highlighted the fact that finding vulnerabilities is of interest to both defenders and attackers.

A branch of AI that has been connected with computer security from relatively early days is automated reasoning, particularly as applied to programs and systems. Among the offshoots of early efforts to teach machines to do mathematics were techniques for theorem proving. These techniques were used to prove that program specifications had particular properties (such as computing a particular function). Researchers also wanted to be able to prove that a program correctly implemented its specifications. When the difficulty of proving functional correctness became apparent, researchers hoped to have better success at proving security properties, which seemed less demanding. Although it turned out that this problem was far from simple, advances in decision procedures, model checking, and, most recently, Boolean satisfiability solvers have already been useful to cybersecurity researchers and might soon help assure, for example, that complex system configurations conform to specified policies.

Machine-learning techniques have been brought to bear on a variety of security-related problems. The earliest intrusion detection systems were motivated by the notion that the behavior of an intruder using a stolen account might differ from that of the authorized user. Accurately identifying abusive behavior has

Letter to the editor

Editors,

The article "Learning by Failing (and Fixing)" (July/August, p. 54) raised a brilliant point: "The initial exposure to insecure systems and mistakes motivates students to learn about building secure systems." Internet security is an important issue in today's computer-centric world, and finding employees that are able to secure code is imperative. In the past, universities might have come under fire when they introduced courses that would task students with creating computer viruses. Some people have concerns that teaching computer-science students how to hack could lure them to the dark side of security and ultimately lead them to cybercrime.

Nowadays, it's necessary to make students look at software with a different perspective. Teaching defense security isn't enough—understanding offensive security is equally important. One cannot really have a defense plan if one doesn't know what the offense is. In other words, if they know how to hack certain systems, they can use that knowledge to tighten up security. This sort of learning is absolutely crucial for writing well-secured applications. It's not so much that universities are teaching hacking, but comprehensive security. Not only are students working toward degrees, but are also looking for useful job experience.

Hackers and security professionals may have the same set of tools at their disposal. It's the knowledge and their moral and ethics that set them apart. I believe it would also be constructive to teach computer ethics course together with security courses. And teaching hacking is a way to understand the risks to corporate networks and personal computers. These courses should be extended to managers and executives as well so they can make better decisions regarding their companies' defenses, especially for software acquisition.

Best regards,
Hong-Lok Li, The University of British Columbia

been difficult, but the problem has attracted substantial efforts to apply AI techniques. Researchers have used Markov models of various sorts, genetic algorithms, neural networks, and other machine learning techniques to detect anomalies at low system levels, for example in packet streams, protocol use patterns, and bit patterns in images. At a slightly higher level, researchers have studied system call sequences to discover abnormal behavior that might indicate compromised software, and, finally, user behaviors as captured in audit logs and databases.

Looking ahead, how might cybersecurity and AI come together for mutual benefit? Several areas show promise.

System security architects have for years aimed to provide mechanisms that could efficiently support fine grained security, so that they could apply security controls minimally and precisely. But when they have tried out such mechanisms, getting users or administrators to specify or even understand the policies needed has been extremely difficult. Perhaps AI techniques can help explain complex policies to users and detect policy settings that are out-of-sync with users' expectations. The areas of potential application include not only file access but also usage-based privacy policies and even network and system configuration policies.

As techniques for finding and eliminating lower-level system flaws improve, we could expect attackers to focus more on the higher levels of systems. Indeed, cross-site scripting attacks seem to have replaced buffer overflows as the most common approach for penetrating systems. Even if we succeed in buttoning up systems completely, attackers might exploit social engineering approaches (as we already see in spear-phishing attacks) to convince users to hand over sensitive data or system access. AI-based techniques that involve natural language understanding, for example, could turn out to be an essential tool in fighting spoofing-based attacks.

More speculatively, we might imagine systems that would have a degree of self-awareness about the data that they process. The notion of reflective systems—systems that can reference and modify their own behavior—has its origins in the AI community. It might be extended to create systems that could reference and modify (or limit) their data flows. Imagine a plumbing system that contained a sort of smart water that could notify the plumber if it found itself dripping out of a hole in a pipe, or perhaps a system of smart pipes that could detect incipient leaks. A cyberinfrastructure that incorporated the analog of smart pipes or smart water (or both) would be of great interest.

The notion of an intelligent attacker, rather than a random natural process, as the underlying generator of risk for a system distinguishes security engineering from other fields. Whether we reach Kurzweil's singularity or not, AI and security concerns will surely be driven together as systems gain intelligence. □

Interested in writing a letter to the editor? Please email lead editor Kathy Clark-Fisher (kclark-fisher@computer.org) with your comments or suggestions. We look forward to hearing from you.