James Bret Michael
Associate Editor in Chief

Security and Privacy for Edge Artificial Intelligence

Edge artificial intelligence (AI) takes decentralization of data and computing to a new level, providing for optimization of resource allocation, and development and functioning of AI, on edge devices. It also introduces opportunities and challenges in the realm of security and privacy.

The convergence of AI, edge computing, and cloud computing impacts our daily lives. For instance, the smart thermostat I recently installed in my home has been learning what my family considers to be a comfortable temperature for different times of the day and night. The thermostat is an edge device that takes as input the sensed room temperature and temperature settings manually input by my family members, refines the machine learning (ML) algorithm on this data, and then makes inferences using the algorithm to select the output control signals to send to my home's furnace. The automation of the temperature-adjusting function with in-situ learning has already resulted in energy savings and the house being neither too warm nor too cool.

Devices like the smart thermostat are part of what is known as *edge AI* or *edge intelligence*. In this "From the Editors" column, I use the former term. But let's back up a moment. What is edge computing? It is a form of distributed computing in which an edge device, or a set of neighboring edge devices, performs computing tasks that would otherwise be done on remote cloud servers. Why is edge computing appealing? With the rapid growth of the Internet of Things (IoT), there is now a vast amount of data being sensed and produced at the edge so enormous in size that it is not technically feasible using the bandwidth of today's Internet to transfer the entirety of the data from the edge devices to cloud servers for storage and processing; even if the bandwidth was available, there would need to be enough data center resources available to handle all of the data. In addition to bandwidth issues, the communication latency incurred by treating edge devices as clients of cloud servers can make it impractical to meet user requirements for fast reaction or response times, such as supporting the real-time decision making performed by collision-avoidance systems embedded in passenger vehicles, buses, and trucks. In other words, there is a need for some degree of federated intelligence. The edge devices, the cars in this example, need local processing for much (but not all of) their activity. Some tasks performed by edge devices may require a combination of local and remote processing.

Edge AI Compared to Edge Computing

What differentiates edge AI from edge computing? Edge AI incorporates AI capabilities on the edge devices. Deng et al. partition edge AI into two categories: AI for edge (also known as *intelligence-enabled edge computing*) and AI on edge.[1] The former is concerned with optimizing the allocation of resources used at the edge, whereas the latter includes "carry[ing] out the entire process of building [and running] AI models on the edge." The smart thermostat serves as an example of AI on edge: this IoT device updates the ML algorithm, makes inferences, and decides what actions to take to shape the behavior of the heating, ventilation, and air-conditioning system. One of the benefits of edge AI technology is that the data needed for refining the algorithm is decentralized. Another advantage is that analysis and decision making can be performed close to the source of the data. From a security and privacy perspective, edge AI can remove attack vectors by minimizing or eliminating the transfer of data between the edge devices and their data centers.

Digital Object Identifier 10.1109/MSEC.2021.3078304
Date of current version: 1 July 2021

In the case of my home thermostat, however, there is likely a continuing connection between my thermostat and some vendor-operated server. Thus, it is likely that my smartphone app is communicating with a central server, which is collecting data and using a session that was previously set up by the thermostat when it connects to the server. I did not set up port forwarding on my router (this has its own security issues).

There are security and privacy issues that arise with the use of edge AI. My thermostat transmits information about its reliability, its performance with optimizing the use of the furnace under multiple constraints, and the parameters of its ML model to cloud servers for use by the company that manufactured the device, ostensibly for improving the company's product line of smart thermostats. How much pattern-of-life information about my family can be deduced from these data? Who has access to them and why? I do not know whether any raw data (for example, unprocessed sensor readings), personally identifiable information about my family and me, or details about my home network security settings (e.g., router password and firewall settings) are shared with the company. I also have no knowledge of how the company transfers those data. Does it use secure connections? Does the company protect this information from side-channel attacks? Finally, does the company update the ML model's parameters from afar (that is, by pushing those parameters from on-cloud servers) without notifying me?

I can remotely communicate with the thermostat from my smartphone via my home's wireless network and the Internet, such as to check on the temperature of my home and change temperature settings, but how much trust should, or can, I place in the authentication and other protocols used with the remote-access functions of the smartphone app for the thermostat? Another issue is that I do not know how much trust to place in my home's wireless infrastructure. I recently received a message from the manufacturer of the routers that I use stating that I need to perform firmware updates. What actually happened when I tried to apply the updates was that the company installed an unwanted network security-scanning application. I lump this into the category of a supply-chain risk. But let's get back to the topic at hand.

In addition, the smart thermostat can operate autonomously. As I mentioned, it learns to self-regulate the temperature inside the home, with no need for attention or input by the user. After about the third week of the thermostat's operation, I stopped fiddling with and checking on the device. I had no situational awareness of what the thermostat was doing, other than that the temperature in the home was comfortable. I stopped logging into the app. I also did not bother to change my password after the initial installation. However, I did set up two-factor authentication. But what if someone hacks my smart thermostat app on my phone and disables the thermostat or hacks the device itself? If someone hacks into my thermostat, then my network is vulnerable, and possibly my family's computing devices become targets—likely of more interest than the thermostat. Could the attacker

Digital Object Identifier 10.1109/MSEC.2020.3044406

start a fire or other significant disruption in the house? (Turning the furnace on and off repeatedly over a short period of time might cause it to fail in unexpected ways.)

Before I purchased the thermostat, I was aware that it and other IoT devices are not immune to attack: they are tempting targets for cybermischief, and hackers have demonstrated the exploitation of vulnerabilities in edge devices designed for the home.[2] I accepted the security and privacy risks for the convenience of automating mundane home-management tasks. I also installed an IoT-based digital door lock. The lock has no AI functionality, so it is an example of edge computing but not edge AI. That is the extent, so far, of my foray into home automation.

Furthermore, with my thermostat falling into the category of AI on edge, there is the possibility that someone may launch an ML adversarial attack, poisoning the data used by the ML algorithm. However, someone could also perform a physical adversarial attack by just leaving a window open near the thermostat. The result of such an attack would likely be the thermostat behaving in an unexpected way, such as causing wide swings in temperature or short-cycling the furnace (that is, causing it to rapidly turn on and off), ultimately resulting in damage to the furnace and an inefficient use of energy. I am not overly concerned about the risks to security and privacy posed by my home-based IoT devices, but I am concerned about those types of issues for industrial and national security uses of edge AI, for which the stakes are higher in the event of compromised security or privacy.

Let's consider a use case for the Industrial IoT. Chemical manufacturing plants, such as those used in the refining of oil and gas, are instrumented with lots of sensors and microcontrollers, with the aim of collecting trusted data to be analyzed as a part of optimizing the processing of the chemicals, maintaining product quality, and monitoring the safety and security of the plant operations. Decentralizing the data provided by the sensors and microcontrollers, in addition to applying AI on edge, is already happening in the chemical manufacturing industry and is viewed as being a vital means for companies to gain an edge—pardon the pun—over their competitors.[3]

IEEE Security & Privacy welcomes submission of articles on this fascinating, rapidly advancing, and game-changing technology.

For this use case, consider the following scenario. A manufacturing plant produces hydrogen sulfide (H_2S), a colorless chalcogen hydride gas also known as *hydrosulfuric acid*. H_2S is poisonous, corrosive, and flammable. Further, let's suppose that the plant employs a federated approach to ML at the edge, with AI-on-edge devices located near sensors that monitor the storage-and-feed, reactor, and clean-up sections of the production unit. In federated ML, each device uses its own "local data to cooperatively train an ML model required by a federated learning (FL) server. They then send the model updates, i.e., the model's weights to the FL server for aggregation. The steps are repeated in multiple rounds until a desirable accuracy is achieved."[4]

What can possibly go wrong? Well, for one thing, it is conceivable that the server acting as the aggregator could leak the trained model or information about the local data sets. In addition, as pointed out by Lim et al., malicious participants (that is, one or more compromised edge devices) could poison the data and model by, for example, "send[ing] incorrect parameters or corrupted models to falsify the learning process during global aggregation."[4] The hacker's intent might be to cause the inference models to improperly manage the cooling water exchangers needed to keep the reactor or clean-up units from overheating, potentially resulting in a mishap if the plant fails to contain the H_2S product. There is the age-old problem of deciding which of the devices can be trusted.

The foregoing example could have been made more complex, allowing for the distributed system to be composed of heterogeneous edge devices (this includes sensors) and cloud servers, along with data sources of varying and possibly unknown levels of quality and trust. What do such systems portend for the specification and implementation of policy and mechanisms for security and privacy?

Edge AI is a burgeoning area of research and development, in part because the enabling technologies are becoming available, such as 5G networks, high-performance AI chips,[5] lightweight AI models, AI-specific service architectures for the edge,[6] co-design methodologies tailored for edge computing and edge AI,[7] and lightweight and leakage-resilient authenticated key exchange protocols for edge AI.[8] Lim et al. are investigating ways to apply edge AI to large-scale

mobile edge networks of heterogeneous devices while preserving the privacy of the data of each of the participants (edge nodes) that are taking part in FL in the presence of one or more malicious participants or aggregators (servers).[4] They have explored several solutions to the problem of not being able to assume that all participants and aggregators can be trusted, such as schemes for secure aggregation and differential privacy. As another example, Libri et al. have demonstrated the application of edge AI by using large-scale sensor networks to detect malware in data centers.[9]

On a personal note, in the mid-1990s, I was a research engineer with the University of California, Berkeley's California Partners for Advanced Transit and Highways program. My colleagues and I were at the forefront of mobile edge computing, demonstrating in 1997 the technical feasibility of safely operating dual-mode automobiles under fully automated control in platoon formations under high-performance driving conditions (e.g., maintaining a velocity of 30 m/s with as little as one car length between vehicles) on dedicated highway lanes.[10] We used classical control system technology to implement the system, with a four-layer hierarchical control architecture: network, link, planning, and regulation.[11] The individual vehicles processed their own sensor data and the data shared among the vehicles. The vehicles also communicated with the instrumented roadway infrastructure. AI on edge would have been helpful, used in concert with the design of the controllers, for developing and continuously improving the models and algorithms used to achieve optimizations, such as for lane-change maneuvers, emergency braking and other safety-related actions under various environmental conditions, minimizing sulfur and nitrogen oxides emissions, and reaching levels of throughput of vehicles that come close to the theoretical capacity of dedicated lanes on the automated highway system.

As edge AI advances and significant progress is made in the evolution from weak to strong AI (that is, from custom AI systems that are tailored to a specific application or limited number of tasks to AI system that have general intelligence abilities), it will be interesting to see how existing security and privacy risks are handled and what new risks and opportunities arise. *IEEE Security & Privacy* welcomes submission of articles on this fascinating, rapidly advancing, and game-changing technology. Please also keep your eyes open for an upcoming call for papers for a theme issue of the magazine on this subject. ■

Disclaimer

The views and conclusions contained herein are those of the author's and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the U.S. government.

References

1. S. Deng, H. Zhao, W. Fang, J. Yin, S. Dustdar, and A. Y. Zomaya, "Edge intelligence: The confluence of edge computing and artificial intelligence," *IEEE Internet Things J.*, vol. 7, no. 8, pp. 7457–7469, 2020. doi: 10.1109/JIOT.2020.2984887.
2. R. Albergotti. "How Nest, designed to keep intruders out of people's homes, effectively allowed hackers to get in." *Washington Post*, Apr. 23, 2019. https://www.washingtonpost.com/technology/2019/04/23/how-nest-designed-keep-intruders-out-peoples-homes-effectively-allowed-hackers-get/ (accessed May 3, 2021).
3. S. Ottewell. "IIoT: Chemical makers approach the edge." Chemical Processing, Mar. 10, 2020. https://www.chemicalprocessing.com/articles/2020/iiot-chemical-makers-approach-the-edge/ (accessed May 3, 2021).
4. W. Y. B. Lim et al., "Federated learning in mobile edge networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 2031–2063, 2020. doi: 10.1109/COMST.2020.2986024.
5. A. James, "The why, what and how of artificial general intelligence chip development," *IEEE Trans. Cogn. Devel. Syst.*, early access, 2021. doi: 10.1109/TCDS.2021.3069871.
6. S. Kum, Y. Kim, D. Siracusa, and J. Moon, "Artificial intelligence service architecture for edge device," in *Proc. 10th Int. Conf. Consumer Electron.*, 2020, pp. 1–3. doi: 10.1109/ICCE-Berlin50680.2020.9352184.
7. C. Hao, J. Dotzel, J. Xiong, L. Benini, Z. Zhang, and D. Chen, "Enabling design methodologies and future trends for edge AI: Specialization and co-design," *IEEE Des. Test.*, early access, 2021. doi: 10.1109/MDAT.2021.3069952.
8. J. Zhang, F. Zhang, X. Huang, and X. Liu, "Leakage-resilient authenticated key exchange for edge artificial intelligence," *IEEE Trans. Dependable Secure Comput.*, early access, 2020. doi: 10.1109/TDSC.2020.2967703.
9. A. Libri, A. Bartolini, and L. Enini, "pAElla: Edge AI-based real-time malware detection in data centers," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 9589–9599, 2020. doi: 10.1109/JIOT.2020.2986702.
10. "National automated highway system consortium technical feasibility demonstration summary report," National Automated Highway System Consortium, Troy, MI, Feb. 1998. Accessed: May 4, 2021. [Online]. Available: https://path.berkeley.edu/sites/default/files/part_1_ahs-demo-97.pdf
11. P. Varaiya, "Smart cars on smart roads: Problems of control," *IEEE Trans. Autom. Control*, vol. 38, no. 2, pp. 195–207, 1993. doi: 10.1109/9.250509.