In this research project, the student will investigate the privacy leakage of smart home speakers. A smart home speaker, such as Amazon Echo, is a voice-controlled device, which is capable of playing music, controlling other IoT devices and providing real-time information, such as weather and traffic. Over 16.8 million smart home speakers were sold in the second quarter of 2018. The market value of smart home speakers is expected to reach $30 billion in 2024. While the technology of smart home speakers is transforming millions of smart homes in the U.S., its security and privacy capabilities have not been rigorously investigated, which leaves critical concerns to users’ privacy and national security.

In this project, the student will work with a group of graduate students at the mentor’s lab and investigate a new attack, named voice command fingerprinting attack, on smart home speakers. Specifically, the student will collaborate with graduate students and demonstrate that a passive attacker, who is only able to eavesdrop encrypted traffic on a public channel between a smart home speaker (e.g., Amazon Echo) and a service server (e.g., Amazon server), can infer the content of a voice command and compromise users’ privacy. The intuition of this passive attack is that every voice command and its response, although being encrypted, have a unique traffic pattern in terms of packet size and packet order. For instance, the traffic patterns of the two following commands, “Alexa, what is the weather” and “Alexa, tell me a Thanksgiving joke”, are very different from each other and can be easily distinguished.

The student will have weekly meetings with the mentor and graduate students. The main tasks for the student will include: collecting traffic traces of
voice commands from different smart home speakers (e.g., Amazon Echo, Google Home, Apple HomePod); applying several machine learning algorithms (in Python) to classify traffic traces; contributing to the writing of a research paper. The student is expected to have some basic background and skills in programming.