On the Impact of Antenna Design in the Context of Graph Inference Based On Wi-Fi Metadata

Mandel Oats, Travis Taghavi, Jean-Francois Chamberland, and Gregory H. Huff
Department of Electrical and Computer Engineering
Texas A&M University, College Station, TX

This research project aims to study the interplay between antenna design and graph inference in the context of Wi-Fi monitoring. The premise for this work is that most people carry smartphones with enabled Wi-Fi interfaces. Due to the structure of Wi-Fi protocols, it is possible to monitor and record Wi-Fi packets as they travel between a wireless device and an access point in an inconspicuous fashion. In particular, this can be achieved without having network access or interrupting a user's connection. This third-party monitoring of packets is often referred to as passive data collection. Metadata leaked through this process includes unique and persistent media access control (MAC) address. During the acquisition process, it is straightforward to augment the captured data with a received signal strength indicator, a geographical location, and a timestamp. The gathered information, once placed in a database, can be employed in a number of statistical tasks.

In this research initiative, we are specifically interested in reconstructing graphs of social interactions taking place in the physical world. Our project operates under the assumption that spatiotemporal proximity is a strong indicator of socio-physical relationships. That is, devices that are frequently found in close proximity are likely to belong to users who are interacting in the physical world, as opposed to coincidental proximity attributable to strangers randomly crossing paths. Yet, the limited information afforded by Wi-Fi metadata and the constraints of signal monitoring in the industrial, scientific, and medical radio bands poses many challenges from an engineering viewpoint. The statistical task at hand demands both the careful sampling of areas of interest, as well as tailored inference algorithms.

The technical goal of this project is twofold. First, we seek to better understand how antenna design influences the sampling of active Wi-Fi devices. There is a natural tension between the frequency at which devices are detected and the significance of observations. In particular, antennas with broad patterns are likely to pick up signals from nearby devices repeatedly. Still, two devices are more likely to be nearby if they are sampled concurrently by a monitoring device with a small footprint. Second, we wish to study how antenna footprints affect the performance of machine learning algorithms. It is possible to enhance the performance of graph-inference algorithms by incorporating information about antenna footprints into the structure of the data. For instance, empirical averages can be weighted in a manner that is inversely proportional to antenna footprint. Then, more discriminating events are given more importance, thereby providing a pathway to better performance. This second objective then revolves around studying how machine learning algorithms that integrate information about radiation patterns are affected by the selection of antenna profiles. Overall, this work offers a framework to determine to what degree the social relationships between devices can be accuately predicted by analyzing Wi-Fi metadata through passive monitoring with application-specific antennas.