To collect cellular data, a Raspberry Pi 4 was fitted with a Raspberry Pi 3G/4G. The portable hardware setup shown below (with a portable battery and antennas attached) runs a Python script which collects a suite of cellular data.

SIM Card Usage in Drones

- SIM card usage in drones is increasingly popular because it provides rangeier connection than Wi-Fi or Bluetooth.
- The problem with this trend is that Base Station antennas are optimized for terrestrial coverage and drone SIM usage can cause interference as well as expose networks to security issues.
- Technology exists to accommodate T-Mobile needs in a way to identify SIM card usage in drones is increasingly popular because it provides rangeier connection than Wi-Fi or Bluetooth.

Differential: Signal Power and Signal Quality

- The scatterplot below shows data points from field testing where blue points mark data taken on open-air rooftops and red points mark all other collection scenarios.
- The scatterplot shows a clear differentiation in signal behavior in the two scenarios.
- The tower on the left of the foreground of the photo below shows a site where data was collected.

Data Collection Scenarios

- Two types of data need to be taken: data at high-altitude in open-air, i.e., drone data, and data in as many other scenarios as possible.
- Thus far, drone data has been collected from open-air rooftops at altitudes greater than 20 meters above the ground.
- Non-drone data has been collected inside buildings at ground level, outside at ground-level, and inside buildings at altitudes of 20 meters above ground or greater.
- The tower on the left of the foreground of the photo below shows a site where data was collected.

Logistic Regression Algorithm

- Logistic Regression takes both categorical and numerical data as input, and outputs a categorical classification as an output based on a simple mathematical model.
- Our Logistic Regression model was able to correctly classify data as drone or non-drone in over 98% of test data.
- Successful classifications depended on avgRSRP and number of connected towers, while RSSI and avgRSRQ were not very useful in classifications.
- Collected data was split 80/20 between training and testing.

Decision Tree Algorithm

- The Decision Tree Algorithm uses complex algorithms to create simple decision rules to split, and then classify the input data.
- While the Decision Tree Algorithm performed slightly worse than the Logistic Regression algorithm, it still had an accuracy above 96%.
- Similarly to the Logistic Regression model, number of towers and average RSRP among connected towers were important classification factors.
- Collected data was split 80/20 between training and testing.

Future Work, References, and Acknowledgments

- The accuracy of our models both exceeded our goal of 85% classification (based on our limited, gathered data).
- Garnering large datasets with a real drone is our next goal.
- Using positional characteristics may differentiate between a SIM on an open-air high-altitude rooftop and a SIM in a drone.
- More scenarios need to be tested, e.g., rural areas, vehicles, etc. to ensure model accuracy.

Faculties: Jim Ritcey, Payman Arabshahi
Industry Sponsor: Precia Carraway
Students: Brian Arnold, John Keller, Yu-Jen Chen

Citations

1. Google Maps U District Skyline Screenshot