Master’s Project Outlines

**Background**

The Quail Ridge Automated Animal Tracking (QRAAT) system provides ecological researchers with tracking data from small wild animals such as birds and rodents. The system consists of transmitters that are attached to each animal, various radio tower sites that are stationary located throughout the Quail Ridge Reserve, and a central processing server that computes locations and stores data. Transmitters transmit signal pulses that are received by the radio towers. Radio towers are equipped with a direction finding antenna and sends the received pulse data to a central processing server. Central processing server then uses pulse records collected from each radio towers to calculate the most likely position for the pulse.

Due to the carrying capacity of small animals, the battery of the transmitters are generally small and causing the researchers to reduce signal strength in exchange for a longer battery life. The signals with reduced strength are hard to discriminate from noises and the resulting position can be significantly off to the habitats of the targeting animals. We need an algorithm to differentiate the real pulse and the noise before we preform the position calculation.

**Problem description**

Each pulse have nine parameters being measured: bandwidth at 3dB (band3), bandwidth at 10dB (band10), frequency, eigenvalue confidence (ec), total nose power (tnp), eigenvalue decomposition signal power (edsp), fourier decomposition signal power (fdsp), eigenvalue decomposition signal to noise ratio (edsnr), and fourier decomposition signal to noise ratio (fdsnr). It is assumed that the pulse has a different pattern from the noise in the nine parameters. The goal of the project is to develop an algorithm that takes in a signal record and determines whether it is a pulse or a noise.

**State of the Art**

Machine Learning is one of the major research areas in computer science. The problem we are trying to solve can be considered as a classification problem where we are trying to determine whether a signal is a real pulse or a noise. One classification algorithm we can try is Naïve Bayes Classifier. The algorithm learn the statistics of the parameters on the training set and then uses Bayes theorem to find the probability of the signal being a real pulse and the probability of it being a noise. With the two quantities in hand, the algorithm marks the signal as the class with the higher likelihood.

**Approach**

To implement Naïve Bayes Classifier, one has to first collect training set with tag of classes for the algorithm to learn. In our QRAAT database we already have few sets of data that we have both GPS and transmitter records. We can use the GPS as the ground truth and calculate the correct bearing angel the receiver should be getting. If the receiver is getting a likelihood of more than a certain threshold at the bearing, we will mark it as a real signal. Else, we will mark it as a noise. With all the data being marked, we will randomly keep a testing set separate from the training set for evaluation purpose. After we have the training set for the algorithm we can start develop the algorithm. In general, the algorithm will go through the training set and find the statistics of a given class using the properties from each parameter. Finally, we will feed the testing set that we have kept to evaluate the algorithm by counting the false positive and false negative rate. The false positive and false negative rate will be used to evaluate candidate

**Data** **Sets**:

* Deployment 60 – stationary beacon with known location. Timestamp from 1383012615 to 1384222215 (10/28/2013 - 11/11/2013). 4,874,724 est records.
* Deployment 51 – moving beacon with GPS tracking. Timestamp from 1376427178 to 1376434127 (8/13/2013). 39,164 est records and 6,864 gps records.
* Deployment 57 – stationary wood rat transmitter with know location. Timestamp from 1382252400 to 1385366400 (10/20/2013 – 11/25/2013). 2,407,324 est records.
* Deployment 61/62 – moving wood rat transmitter with GPS tracking. Two transmitters moved together during the two tracking with gps records. First tracking from 1391276584 to 1391285374 (2/1/2014). Deployment 61 has 11,191 est records, deployment 62 has 22,422 est records, and 8,388 gps records. Second tracking from 1396725597 to 1396732326 (4/5/2014). Deployment 61 has 11,672 est records, deployment 62 has 9,331 est records, and 6,438 gps records.

**Timeline**

2/1 – Try out different algorithms and start evaluations of algorithms.

2/29 – Need to have all implementations done and evaluated.

3/21 – First draft of the write up done and reviewed by committees.

4/4 – Final draft done and submitted.

5/27 - Final deadline for File Thesis with Graduate Studies.