Masters Thesis

This paper served as a good starting point for me to explore WSN’s. Many aspects of the design and implementation were covered, including real-world constraints, network architecture, communication protocols, sensor algorithms, etc. My biggest takeaways are:

* Transitioning from an ideal environment to a real-world implementation requires a lot of planning:
  + Sensor topology: How the sensors will communicate, like hierarchy, P2P
  + Location topology of where the sensors will be deployed
  + Varying “normal” conditions of various locations
* We can use simulations to test the network performance and transmitted data
* There is a delicate balance between data availability and sensor battery life
  + Communication drains battery the most, so constant data transmission would drain the battery very quickly

The code and implementation were hard to understand at times because some of it was in a language I presume to be specialized for the sensor (Section titled MIC for Apple Scab Module). It was nice to see C!

Fog Computing

The transition from the thesis to this paper is similar to a shift from a centralized computing approach, where sensors simply send data to the cloud or machine which does computations based on that, to shifting some of the computation to the sensors themselves. The objective is to perform operations on that data which allow vital information to be condensed and thus reduce the power spent on communication. The power spent on extra computation must be less than the power saved on communication for this to be a viable strategy.

I was exposed to edge mining tools and related algorithms. These were covered very briefly as this was a short paper so this paper assumes some understanding of these. The graphics which showed the testing results were very useful for me to understand how to gauge performance. I wonder what the challenges might be in implementing the AAL. One could be variations in body movement and position for different users when doing the same activity which makes it difficult to classify accurately for each user.

Genetic Algorithms

I came upon this paper when I was looking for a more in-depth look into genetic algorithms, which were introduced in the Fog Computing paper. I focused on Section 3, which provides an overview of genetic algorithms. It was really cool to see how it simulates the combination of randomness and improvement in evolution. The formation of fitness criteria seems like an interesting challenge because genetic algorithms attempt to optimize something but we don’t necessarily know what an optimal result would look like.

Edge Mining Algorithms

This paper provided more detail on how edge mining algorithms (SIP, ClassAct, BN) work. The most valuable part to me was the pseudo code because I knew precisely what the algorithms would do. The primary method of evaluation was message reduction because that is the reason why edge mining is useful. All methods use a definition of an event: a data point which cannot be predicted using the given methods. The predictions change overtime based on previous data, meaning that the sensors can continuously learn.

Leveraging Fog Analytics

This paper served as a transition from regression-based SIP algorithms into classification problems. The classification problem in this case was the activity of cows. There were many instructive pieces in this paper:

* There are often data points which overlap among classes, leading to some amount of error due to the indistinguishability of measurements.
* Sensors can “work together” by communicating their respective states, further improving the accuracy of the classifications. Nodes of closer proximity are obviously more relevant to the classifications. However, there is a trade-off because having more nodes collaborate requires many more signals to be sent -> greater power usage
* Training the decision tree model:
  + Analysis of data is performed to determine which measurements are most useful for differentiating between states. In this paper, the axes of the sensor accelerometers were compared.
  + The window of previous measurements to include and weights of previous measurements should be compared.
  + Epsilon trade-off: lower epsilon means more sensitivity, but this could lead to false positives of state change. Finding the right epsilon is important.

(This article provides an explanation of the C5.0 classification used in the paper) <https://jmsallan.netlify.app/blog/classifying-with-decision-trees-using-c5-0/>