Hi Dr. Yan and everyone here, my name is Anna Nguyen and I’m a 5th year electrical engineering and mathematics student, and my thesis topic is the automatic power system event detection via power quality meters.

So to give some context into my topic, a power system is a whole network of equipment which generates, distributes and transmits power to anywhere, such as our houses or to university. As always, these systems are prone to failure and faults, and these faults can be caused by many things, such as manufacturing error and equipment aging over time, but also things such as lightning strikes, or tree branches brushing up against power lines.

Faults like these that are temporary and have low severity are considered to be low risk, as most power systems have failsafes to protect against these short temporary faults caused by the external environment. However, faults that are more severe occur more often and for longer periods of time and these are often an indicator of potential permanent damage and outages, and these need to be identified and addressed as soon as possible.

As consumers, we always want to have consistent and reliable power being delivered to us, but power quality is also extremely important for distributors. For example, energy distributors and retailers have an obligation and responsibility to provide energy to consumers at determined standards. When these standards are not met, distributors are required to make guaranteed service level (GSL) payments to consumers. In the 2018-19 financial year, it was reported that Ergon Energy and Energex paid $1.2 million and $3.4 million in GSL payments respectively in QLD alone. Over 90% of the respective payments came from reliability interruptions, so there is a high motivation to be able to detect and classify events as early as possible.

The data was gathered from power quality meters on the secondary side of a 11kV-415V transformer located on the UQ St Lucia campus near the P3 & P4 carparks which was commissioned in 1965. The PV solar panels located on top of these carparks also lead into the system that this transformer is located in. Ideally, the data from the PQM should look like &&&&&FIGURE&&&&&.We’re mainly noting the sinusoidal nature of the voltage and current waveforms of each phase of a typical power system.

However, the data collected from our PQM looks like &&&&& FIGURE &&&&&&&. We are able to see the distortions present in the current waveform as well as a transient event towards the right of screen. It was noted that it doesn’t seem like these events or distortions are present in the PV waveforms. We are assuming that the PQMs are calibrated and operational so any events are occurring are due to the load that the transformer is supplying and not due to faulty equipment. I was supplied with approximately 19GB of data from the power quality meters, which consists of waveform data sampled at a rate of around 4000Hz for all phases of voltage and current.

In the past semester, I have researched many different types of algorithms to implement for event detection within power systems. So I first had a look at what types of methods are currently being used for event detection and came across a competition by drivendata and Schneider Electric for the detection of anomalies in power usage. The first place winner implemented a solution using the isolation forest, or iForest method.

The crux of this method is that data points which are considered ‘normal’ will have many similar features, where as anomalous points are very few and different and can be more easily separated from those that are normal. The iForest method is a group or a forest of decision trees, where each branch of a decision tree is analogous to a partition or separation in the data, which can be more easily visualised in 2D space in the figure here. We can see that the more ‘strange’ a data point is, the easier it is to partition it from the ‘normal’ data, where as points that are normal require more branches of the decision tree to separate it.

Although this method was accurate and reliable in detecting anomalies in the power data, the output of this method only returns a binary result as to whether a particular point is anomalous or not. It returns little to no information about any features of the event, for example a time duration or severity, and as a result, becomes difficult to extract further information from any particular event, such as why it was caused.

There were further complications which became more obvious as I spoke to Dr. Yan about this method, such as

* Not being able to detect harmonic distortions in the waveform
* Mismatching data format, since the data provided for the competition was one energy reading every half an hour, compared to the multi phase readings

As a result, I had to look into a method that was a bit more versatile and could extract features for further analysis for classification.

The most successful method I have implemented so far was the use of a recurrent neural network, or RNN. The general structure for a neural network is a system of layers, where the hidden layers consist of functions, weights and biases. It trains by receiving training data, where the data is input into the model, then compared with the output that it should be producing. The difference in this output is called the error function and with every epoch, the weights and biases in the hidden layer are updated so that the error function is minimised.

A recurrent neural network is unique in that the hidden layers feedback on themselves, where as typical neural networks are only feedforward. This makes RNNs more effective in predicting data inputs which are unsegmented and connected such as our power quality meter data. However, one common problem that occurs with feedback neural networks is the vanishing gradient problem, where the change in the error function is so small and the weights and biases in the hidden layer do not get changed, which can result in the model no longer training properly.

To overcome this, I implemented the Long Short Term Memory architecture to overcome the vanishing gradient issue. One of the most notable aspects of this architecture is that a ‘forget’ gate is introduced into the hidden layer, which allows information that is unimportant to be forgotten or ignored.

The following gifs are visualisations of the prediction of the data signal after training for a certain number of epochs.

To detect anomalies and events in the power system from the data given, I trained an RNN with a sample of data and used the network to predict future behaviours. The difference between the predicted value and the PQM data was calculated and the anomaly scores were returned. Plotting the various curves on the same set of axes results in the following figures:

* No anomaly – none detected
* Anomaly – detected

The accuracy and reliability of using an RNN to predict and detect possible events is dependent on the training period – more specifically the type of training data provided and for how long the program trains for. At the moment, I cannot conclude a numerical rate of accuracy, given the limitations of my computer and the amount of time taken to train and test the RNN. However, the model seems to be accurate and has a high anomaly score for the events that I have been able to put through the model.

Now we consider the harmonic distortions that can be seen in the waveform. We have gathered approximately 19GB of data from the power quality meters and this results in well over 10,000,000 data points. As a result, it is just not feasible for me to be able to see if this distortion in the current waveform occurs for all the data given, or whether this is a temporary issue. Since this is considered abnormal behaviour for a power system, this also needs to be flagged as an anomaly to the system if it is not a steady state occurrence.

One potential solution I have researched is the discrete wavelet transform, which transforms a time domain signal into a time-frequency domain. The original signal is decomposed into multiple frequency bands, halving the frequency range with each pass. This is similar to the Fourier transform, however, the wavelet transform retains temporal information about the signal, which is important when classifying the type of event. By decomposing the signal in this way, I would be able to determine whether the distortion is present in the waveform as well as where in time it occurs.

With each pass of the DWT, we remove half the frequencies from the signal, and from Nyquist’s rule, we also remove half the samples. We can see a sample of our data after 5 passes of the discrete wavelet transform to remove as many of the harmonics as possible to see the behaviour of the signal at fundamental frequency.

The resultant signal is undersampled by 24 or 16 and this results in some observations:

* When the amount of data points is reduced, it results in a much faster neural network training and testing time.
* However, when the signal is undersampled, it reduces the time resolution of the signal, which may remove important information of the event such as its duration or magnitude.

In terms of future research and development, both of these observations will be investigated and other ideas will be further developed in the future.

The most immediate action that I plan to take is to set certain thresholds on the anomaly scores for the neural network to determine whether or not a change in power system is deemed to be an event.

The next major step in the future research and development plan is the classification of events. IEEE defines power quality disturbances in Standard 1159 based on its characteristics and indices as shown here. My goal for the classification stage is to be able to extract a detected event and categorise it automatically. Once an event has been successfully detected, I plan to consider all its specific features such as its duration and magnitude or previous events to try and find a potential cause.

There are many current methods used to classify PQ disturbances, but many require the event to be deterministic and modelled. For example, Uyar and others are able to classify power quality events by matching it to one of the below rules. However, given the distortions in the waveform, this may not be possible, so one potential method is the probabilistic neural network, which uses statistical methods to determine the probability of an event. It’s noted that the success of this method will be highly dependent on the quality and amount of training data provided.

It is also possible that once the discrete wavelet transform has been implemented to remove the harmonic distortions and reduce the signal down to only its fundamental frequency, that its waveform may match closely with the rules here. These observations and ideas are all part of future research and development that I plan on undertaking in the next semester.

Finally, I have been and plan on continuing to keep a logbook which contains my ideas, research and notes with timestamps and references to any literature I have used. Any code I write is being version controlled using my personal git account and committed regularly to avoid risk of losing code. This is so I have a strong record of my progress and will be able to look back on when I start writing my thesis. As always I plan on meeting with my supervisor fortnightly to discuss progress and any potential problems.

I really appreciate everyone’s time in listening to my seminar presentation and now I’m more than happy to answer any questions in regards to my thesis topic.