Chapter 6

Conclusions and Future Work

The field of NILM or energy breakdown is more than three decades old. During these three decades, loads of new algorithmic approaches have been proposed. Many start-up companies have leveraged energy breakdown techniques in some of their offerings. However, there were three factors impeding practicality of energy breakdown—lack of comparability, action-ability and scalability. We now conclude our thoughts across these dimensions and also suggest future work.

6.1 Ensuring comparison across approaches

6.1.1 Conclusions

When we began our NILM work, we wanted to use the “best” NILM algorithm and develop applications on top. We realised that finding the “best” NILM algorithm was no trivial task. Different researchers had used different data sets, different benchmark algorithms and different metrics. This made it virtually impossible to compare NILM papers and ascertain the best NILM algorithm. At this point we felt our efforts would be best spent towards making NILM research more standardised. This was also the general consensus of the community as discussed in the NILM workshop. One of our goals was to lower the entry barrier for NILM researchers. We teamed up with researchers from the UK and the US to develop the NILM toolkit. In our experience,
all the engineering effort spent in NILMTK, paid us back many times in terms of research output. We are very satisfied that beyond the core developers, NILMTK has been used by the community. Researchers have contributed their algorithms and data sets to NILMTK.

6.1.2 Future work

1. Despite the positive traction gained by NILMTK, still a vast amount of literature remains hard to compare against. While NILMTK is an important first step towards making NILM algorithms more comparable, significant efforts are needed towards the goal. The image processing community serves as a good example of comparable scientific research. The ImageNet challenge\(^1\) can be attributed to a lot of recent comparable state-of-the-art work in the field. We believe that the energy breakdown community would similarly benefit from such a competition. In fact, one of the NILMTK’s lead developers, Jack Kelly\(^2\), is currently pursuing this thread. There are several other ways in which the community can help, such as mandating code release for any submission. Many conferences encourage code submission for paper submissions. Integrating a Kaggle-like\(^3\) service for standardised tasks (similar to the competition) can greatly help in making the field more standardised. The community will also benefit by integrating their open tools with tools such as NILMTK. An example is a recent household energy simulator called SmartSim [24].

2. While we compared the NILM problem to the image processing problem, which has the ImageNet challenge, there are few important differences. Different NILM researchers focus on different frequency of data collection. The frequency range is huge- ranging from a sample every 15 minutes, to millions of samples every second. NILMTK in its current form is tuned to low frequency data collection. In fact, till date it remains nearly impossible to compare the efficacy

\(^1\)http://image-net.org/
\(^2\)http://jack-kelly.com/
\(^3\)https://www.kaggle.com/
of low-frequency approaches against the high-frequency approaches. This is due to the fact that very few current data sets measure both low-frequency and high-frequency power data, and tools like NILMTK have not been developed for high-frequency data. Future datasets collection should account for such high-frequency and low-frequency parallel data collection so as to support diverse comparison.

6.2 From disaggregation to specific actions

6.2.1 Conclusions

After our NILMTK work, we were faced with two choices - build more accurate NILM algorithms, or, work towards our initial aim, to save energy. The “usefulness” of NILM had also been questioned many times. Thus, we undertook research to understand if energy breakdown can provide specific actionable energy saving insights, over and above the pie-chart energy breakdown. There were two important questions that we needed to answer the applicability of NILM research. First, can we leverage appliance power traces to provide actionable insights? Second, do current NILM approaches provide disaggregated appliance traces with sufficient fidelity to facilitate actionable energy saving insights?

To answer the first question on the utility of appliance level power traces towards actionable energy savings, we need to construct appliance energy models. These appliance energy models should be able to distinguish regular and anomalous operation of the appliance. Based on models and insights developed by domain experts, we created simple models for fridge and HVAC. Our key idea was to use these models to provide insights such as - “your HVAC is set to a wrong temperature, this recommended schedule can save you 10% on your bills”. Our findings indicate that energy saving insights can save up to a quarter of the appliance energy consumption. However, when we investigated the appliance level traces provided by NILM algorithms, we found that the appliance traces produced by current NILM algorithms
show poor feedback accuracy. The same NILM algorithms show good accuracy on conventional NILM metrics such as F1 scores and RMS error. This can be explained by the fact that NILM algorithms do well in general, giving good performance on conventional metrics. But, the cases we care about for appliance feedback are often poorly predicted. Our work suggests that the community take an alternative view of the problem where actionability is a key concern. This would entail development of algorithms with the new set of metrics (focusing on applications).

6.2.2 Future work

We illustrated actionable feedback for two appliances - fridge and HVAC. A large number of appliance categories still need to be covered. In fact, our current approach of manually creating a white-box model for each appliance category may not scale particularly well. One approach could be to develop energy models for classes of appliances, such as - thermostatically controlled, purely resistive, switched-mode based power supply among others. Another possible direction is the development of smart appliances that incorporate actuation capabilities and local intelligence for optimal appliance operation. With the advent of NEST and similar smart appliances, the control and intelligence are increasingly being pushed to the end device. This is where our work could fit well into products. These smart appliances can run algorithms similar to ours and inform the appliance owners about inefficient usage.

6.3 Scaling up energy breakdown

6.3.1 Conclusions

We realised that a great deal of energy breakdown literature could not be scaled today to all homes. This is due to the fact that current energy breakdown solutions require hardware to be installed in each home. Even though smart meters have been rolled out across the US, these smart meters often sample at low rates, which makes most of the NILM literature impertinent. Against this background, we chose to develop
scalable energy breakdown solutions that do not require any hardware to be installed in a test home. We started with the goal of creating an energy breakdown solution that works with whatever data is easily accessible, is able to scale across a large number of homes and requires minimal capital expenditure involved. In order to achieve these objectives, we completely flipped the way we look at the problem. Rather than the existing bottom-up approach of using modelling to identify electrical signatures, we used the top-down approach of using modelling to identify home level characteristics that correlate well with appliance level energy consumption. We showed that such home level characteristics can be easily calculated with static household information and monthly electricity data both of which are readily available. Not only is our approach more scalable, it is also more accurate than state-of-the-art NILM approaches.

6.3.2 Future work

1. Our approach currently faces the challenge of the availability of static information (metadata) along with the power data. Very few public data sets survey such information. Future data set owners should try and obtain as much static household properties as possible. Other NILM approaches have also shown the benefit of such metadata. Our current work on making energy breakdown more scalable works only for homes in the same geographical regions. If we can learn the properties of different regions that cause differences in energy consumption, we can make energy breakdown more scalable. We are currently looking into transfer learning methods for scaling energy breakdown across multiple geographies.

2. The first step towards realising some of the associated benefits from scalable energy breakdown would be to carry out pilot deployments where people are given the energy breakdown estimated by our system. Such large-term studies are needed to truly understand the impact of our technology at scale.