# Residential Electric Load Disaggregation for Low-Frequency Utility Applications


Guanchen Zhang
Gary Wang
School of Mechatronic Systems Engineering
Simon Fraser University
Surrey, Canada

Hassan Farhangi
Ali Palizban
Group of Advanced Information Technology
British Columbia Institute of Technology
Burnaby, Canada



***Abstract*—Recent load disaggregation approaches take advantage of artificial intelligent techniques and require low sampling frequency. From utility perspective, intrusive data for training are not available due to privacy and the sampling frequency may be too low to recognize meaningful signatures. This paper proposes a 1-hour frequency disaggregation algorithm for real and reactive energy without knowing what appliances are in a specific house/apartment. The proposed algorithm is particularly developed concerning utility's constraints. A database of typical types of appliances in BC, Canada is built. Based on BC Hydro's smart meter data over a year, the most probable appliances in a specific house/apartment are firstly inferred through likelihood maximization and energy consumption matching. The disaggregation is then implemented by an integer multi-objective Genetic Algorithm tuned by appliance dependence rules. The results show that despite of high uncertainty, more than 50% of energy consumption could be disaggregated for random houses/apartments.**



***Index Terms*—Load Disaggregation, Smart Meters Data Mining, Load Forecast.**


## I. INTRODUCTION

The coordinated Volt/VAR optimization (VVO) engine developed at the BC Institute of Technology (BCIT) is capable of predicting the Volt/VAR demand at customer ends, thus adjusting Volt/VAR injection for optimal feeder performance [1]. This paper proposes an electric load disaggregation engine that is able to provide the VVO engine the breakdown of residential real and reactive aggregate energy consumption. Such information have the potential of improving the VVO in term of better forecast of consumption and distributed generation (DG) dispatching, hence improving the energy management of microgrids. Fig. 1 illustrates the framework of the disaggregation system.

For customers, disaggregation could reveal appliance-level information of power usage, advising customers for more efficient power habit and lower energy cost. However, customer-focused disaggregation algorithms may require intrusive data of a period of time for training purpose. These data are usually sampled at less than fifteen minute to detect useful appliance signatures.

From utility's perspective, sub-metered data may violate the privacy of customers and not be widely available. Although artificial intelligence (AI) based unsupervised disaggregation may not require sub-metered data [2]-[3], appliance signatures need to be extracted from relatively higher frequency than one hour to implement unsupervised training. The available data from BC Hydro is measured at 1-hour frequency, which means most appliance signatures may be obscured and state-of-art disaggregation approaches may not apply to such low frequency with high uncertainty.

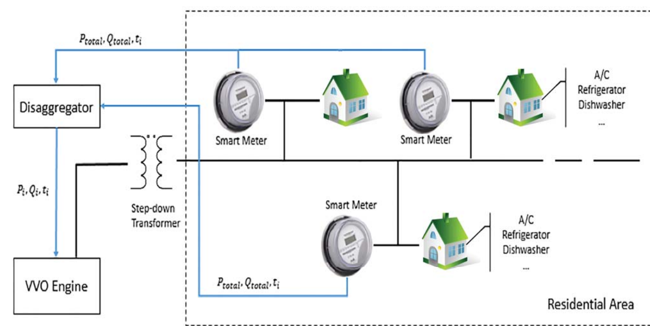


Figure 1. VVO-Disaggregator Framework

Considering the constraints from available data and utility's perspective, this paper proposes an approach that can disaggregate both active and reactive energy at 1-hour frequency. No prior knowledge of appliances in specific residence is needed. However, due to a variety of uncertainties, this approach only generate approximate results. The approach is composed of two steps. The first step infers what appliances are in a specific household based on likelihood maximization and energy consumption matching. The second step uses the results from step one as inputs, and implements an integer multi-objective Genetic Algorithm (GA) to disaggregate. The optimization results are tuned by appliance dependence assumptions and reactive energy matching. By taking advantage of the big data from the Advanced Metering Infrastructure (AMI), this algorithm is intended to apply to houses/apartments in different regions in BC, Canada, accounting for seasonal variance, and different dwelling types, heating types and family sizes.


This project is a part of Project 3.4 – Integrated Data Management and Portals, of NSERC Smart Grid Network (NSMG-Net)

## II. General Methodology

Non-intrusive disaggregation techniques could be categorized as supervised and unsupervised. Supervised disaggregation requires sub-metered or interactive data to tune appliance models, whereas unsupervised disaggregation may only need general appliance models at deployment [2]. Both techniques require a sufficiently high frequency to recognize useful signatures. Classical non-intrusive load monitoring (NILM) methods rely on event-based signature detection, power clustering and statistical inference [4]-[5]. Recent techniques in non-intrusive disaggregation include Hidden Markov Model (HMM) and its variants [2]-[3], [6]-[7]. However, implementing HMM requires sub-metered data to construct appliance models, or the frequency needs to be higher enough for recognizing useful signatures.

Utility constraints in this paper, however, specify that sub-metered data are not available and the sampling frequency is one hour. In addition, appliances in households are not known from utility's point of view. Recent non-intrusive techniques, therefore, may not apply to the scenario in this paper.

To overcome the constraints, this paper sequentially applies appliance inference and disaggregation as two steps (Fig. 2). The appliance inference step determines the most probable appliances based on smart meter data and appliance ownership data. The disaggregation step uses the results from step one to build appliance models. Errors in real and reactive energy matching are concurrently minimized in an integer multi-objective GA, in parallel to the maximization of appliance turn-on probabilities. Appliance dependence rules are integrated into the optimization algorithm to adjust optimization values.

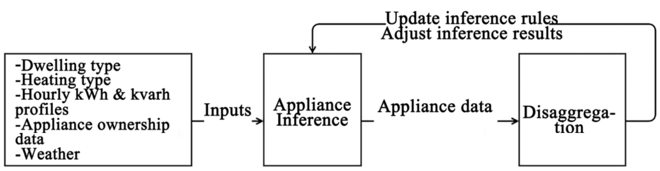


Figure 2. Top-Level Architecture of the Disaggregator

## III. Appliance Inference

This step infers the most probable typical appliances in the target building. Probabilities of having a certain type of appliance and its corresponding usage frequency in weekly basis are derived from surveys in [8]-[9]. The data of each appliance are determined through an integer multi-objective GA. Because the inference based on one single week cannot guarantee the correctness due to behavioral and seasonal variance and other uncertainties, the optimization is performed iteratively on a number of weeks of data. The type of each appliance that has the most appearance count is selected. The top-level flowchart is shown in Fig. 3.

### A. *Inputs*

The AMI data from BC Hydro contain the dwelling type and heating type in addition to real and reactive energy measured at 1-hour frequency over a year. Weekly power consumptions are calculated as another input. Seasonal variance is accounted as variance in on/off probabilities from [10]. The appliance ownership data from [8]-[9] include the probabilities of having a certain type of appliance, quantity and usage frequency, categorized by dwelling type, family size and geographic region.

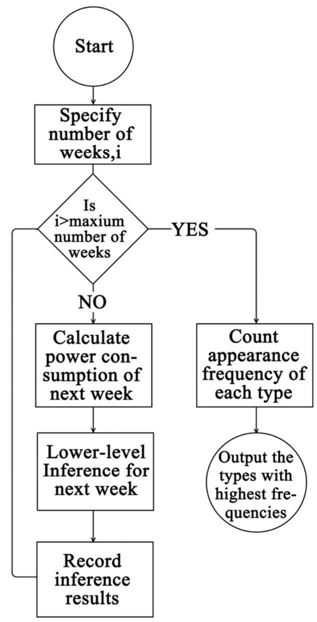


Figure 3. Top-Level Flow of Appliance Inference

### B. *Appliance Models*

Appliances are modeled based on the appliance ownership data in [8]-[9]. The unit energy consumptions (UECs) of typical types of appliances in Canada are available in [8]. In [9], the usage frequency of each appliance, percentage of each type of an appliance, and quantity of each appliance are categorized by housing unit type, i.e. detached and attached houses and apartments with different units. This information, together with the UECs from [8], forms the appliance models.

### C. *Lower-Level Inference Algorithm*

#### *1) Optimizer*

As shown in Fig. 4, each appliance model, based on dwelling type and etc., generates a vector of corresponding UEC and matrices of various probabilities. The most probable type of each appliance is determined by an optimizer that has two objectives:

- Matching the stacked energy consumption with hourly measurement (1)
- Maximizing the combinational likelihood of having a certain type of each appliance (2)

$$f(1) = (P_{week} - \textstyle\sum_{1}^{N} x_i / 52)^2 \qquad (1)$$

where $P_{week}$ is the measured weekly power consumption, $N$ is the number of typical appliances, $x_i$ is the annual UEC of the $i$'th appliance, and 52 is the number of weeks in a year.

$$f(2) = -\textstyle\sum_{1}^{N} p_{type_i} * p_{region_i} * p_{quantity_i} \cdots \qquad (2)$$

where $p$ represents the probability.

The integer multi-objective GA generates integer variables that specify the indices of type and usage frequency of each appliance. The objective functions are evaluated corresponding to the values indexed by the integers.

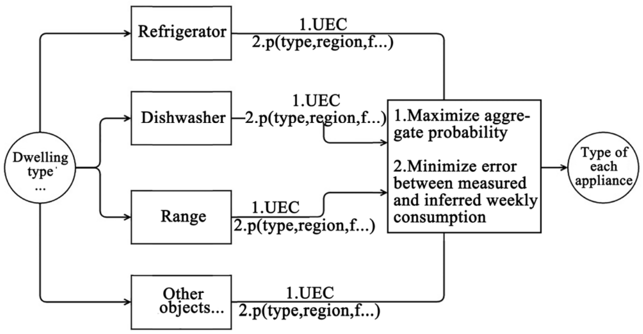


Figure 4. Lower-Level Flow of Appliance Inference

*2) Post-Screening of Optimization Results*

Because of the multi-objective optimization, the Pareto front provides multiple optimal candidates that have different error values and aggregate probabilities. The optimal solution(s) is chosen as the one that satisfy the thresholds for both objective values. The thresholds are chosen as ones that can yield, not necessarily the minimum, but relatively smaller objective values for both objectives.

## D. Energy Consumption of Typical Loads

The power consumption of each week, $P_{week}$ in (1), is calculated by filtering out the portion of non-typical consumption. The breakdown of average residential energy consumption by appliance is given in [11]. It is noted that, in a short term, the power consumption of a household may not have the average end-use portions. However, in a long run, the end-use contributions will most likely converge to the average. Therefore, multiple weeks are used to search the most possible appliances in a household.

The annual and weekly UECs are assumed to account for average usage frequencies. If the usage frequency from the optimization is different from the average usage frequency, the annual and weekly UECs are scaled up/down.

## E. Seasonal Variance

Seasonal variance in power consumption is mainly contributed by variance in space heating and cooling. When calculating the energy consumption of typical loads, the end-use portions of these two loads are adjusted based on the date of measurement. For example, based on [12] in January, space heating may contribute to 60% of total monthly energy consumption, and in March it has 28%.

# IV. Load Disaggregation

The general idea of the disaggregation algorithm is similar to the inference algorithm in term of hourly energy matching and probability maximization. Because that the single-objective optimization of energy matching at each hour does not explicitly relate itself to preceding or previous hours, false on/off detection is highly possible. The multi-objective optimization, on the other hand, implicitly takes into account on durations through on-probability maximization, so optimization at each hour is related to optimizations at other instances. Reactive energy matching is additionally considered in parallel for reactive energy disaggregation. The top-level flow of disaggregation is shown in Fig. 5.

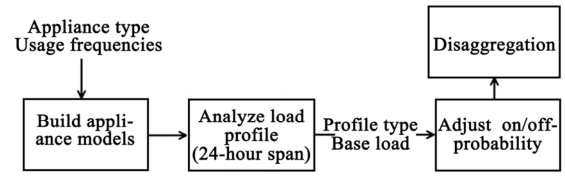


Figure 5. Top-Level Flow of Disaggregation

## A. Appliance Models

*1) Appliance Real Power*

Since the UECs from step one have weekly or yearly basis, converting UECs into kWs needs the usage frequency and on duration of each appliance. These kW values are regarded as the base real powers upon which further adjustment is performed.

The usage frequencies are acquired from step one. The on duration of each appliance is obtained from the Smart Residential Load Simulator from [13]. In the scope of this paper, only the clothes washer and dryer have on duration longer than one hour.

Space heating and cooling, water heating, lighting, miscellaneous electric loads (MELs), and kitchen appliances are additionally modeled in the disaggregation algorithm. Each of these models is assumed to have a fixed base consumption and time-dependent hourly consumptions. The fixed base consumption is assumed to be the power consumption of typical types. The time-dependent consumptions at each hour are derived from hourly profiles of each appliance (e.g. Fig. 6) and have a daily average consumption equal to the base consumption. Time-dependent consumptions account for loads that have various active states (e.g. more lights at night and less during the day).

*2) Appliance Reactive Power*

In this paper, the reactive power of each appliance is simplified as that all types of an appliance have the same constant power factor, and the reactive power is derived from the real power and power factor relation.

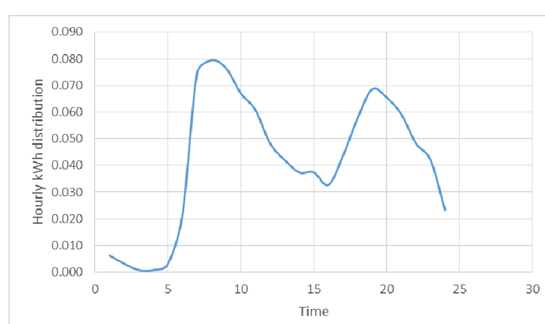


Figure 6. Hourly Profile Example – Water Heater in Winter [10]

*3) On Probabilities*

The on probabilities at each hour, or time-of-use probabilities, are obtained from [10]. The probability distributions are similar to the pattern in Fig. 6. Note that the probability profiles from [10] are derived based on a typical load profile in which two peaks occur in the morning and evening. For random load profiles, however, the probability distributions need to be adjusted to best describe random scenarios. This paper identifies load profiles as typical, peak at night and flat. Another reason for adjusting probability distributions is to blend appliance usage dependence into time-of-use probabilities. This is implemented by adjusting the probability at following time instances if some events occur at the current hour. For example, if clothes washer is turned on at 10 am, the on probability of clothes in next two hours will likely increase. Usage frequencies from step one also affect on probabilities. If clothes dryer is turned on at current hour and its usage frequency is once a week, its on probabilities in following hours/days are lowered. The dependence rules in this paper are summarized in Table I.

TABLE I. APPLIANCE DEPENDENCE RULES

| Type of rules | Applicable appliances |
|---|---|
| On-duration continuity | Dish washer, clothes dryer |
| Usage sequence | Clothes washer and dryer |
| Time dependence | Lighting |
| Appliance dependence | Lighting, clothes dryer |
| Base load detection | All appliances |
| Usage frequency constraint | All not-always-on appliances |

### B. *Multi-Objective Optimization*

#### 1) *Objectives*

The objectives, similar to (1)-(2), are provided in (3)-(5).

$$f(1) = \left(P_{hour,i} - \sum_{1}^{N} P_j\right)^2 \tag{3}$$

where $P_{hour,i}$ is the aggregate power consumption at hour $i$, and $P_j$ is real power of the $j$'th appliance.

$$f(2) = \left(Q_{hour,i} - \sum_{1}^{N} Q_j\right)^2 \tag{4}$$

where $Q_{hour,i}$ is the aggregate reactive power consumption at hour $i$, and $Q_j$ is the reactive power of the $j$'th appliance.

$$f(3) = -\sum_{1}^{N} p_{i,j} \tag{5}$$

where $p_{i,j}$ is the on probability of the $j$'th appliance at hour $i$.

The variable, $\boldsymbol{x}$, of each appliance is an integer representing either on or off.

#### 2) *Optimization Procedures*

The optimization is composed of three steps: a) in-optimization decision tuning b) optimal solution screening c) results validity check and adjust (Fig. 7).

In-optimization tuning is implemented by overwriting objective values if the optimization solution violates the appliance dependence rules. Hence the optimization continues to seek other solutions. Because the multi-objective optimization generates multiple candidates, the screening process chooses the optimal solution in an order of highest probability, best real energy matching and best reactive energy matching. All objective values, however, need to be within a reasonable range at the same time. In case that not all the objective values can fit into the range, the optimal solution is still chosen based on objective ranking, but a post adjustment is carried out according to the error in reactive energy matching. If the disaggregation leads to more reactive consumption than the measurement, the detected appliance from the optimization with high reactive power is removed.

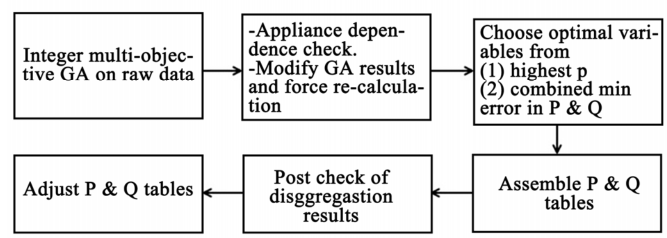


Figure 7. Flowchart of Lower-Level Disaggregation

## V. RESULTS AND DISCUSSION

### A. *Appliance Inference*

For a house in Kootenay, BC with electric heating and annual consumption of 17 MWh, an iteration of 30 weeks is performed. The result is shown in Table II.

TABLE II. Appliance Inference based on 30 Weeks

| Appliance | Type | Usage frequency | Adjusted UEC [kWh/yr] |
|---|---|---|---|
| Refrigerator | 2-door with top freezer, 417 kWh/yr | Always on | 417 |
| Freezer | No freezer | NA | NA |
| Dishwasher | 325 kWh/yr | Once a week | 162.5 |
| Electric range | Self-cleaning with 499.3 kWh/yr | A few times a week | 166.4 |
| Clothes washer | Front loading with 148.3 kWh/yr | 5-9 loads each week | 222.5 |
| Clothes dryer | 925 kWh/yr | Used for some but not all loads of clothes washer | 185 |

Out of 30 iterations, i.e. 30 weeks, 68 potential solutions that pass the screening criteria are obtained. The minimal objective function value of power consumption error could be as low as 0.0044. The average objective values are [8.19 65.27%], which means the inferred appliance types could sum to a weekly total energy consumption that has a difference of 2.86 kWh to the actual measurement. That is, roughly 2% of the actual erngy consumption is not explained.

A set of iterations of 50 weeks on the same house is also simulated to compare with 30 weeks. The results show that the differences situate in appliances having types with similar UECs and having high uncertainty in usage pattern.

### B. *Disaggregation*

Three types of load profile are tested: 1) peaks in the morning and evening, i.e. typical profile 2) peak at night 3) relatively flat profile. As discussed, the shape of load profiles affect the adjustment of on/off probabilities.

Fig. 8 shows disaggregated real energy at each hour. Undisaggregated portions, from 16:00 to 19:00 are due to the post adjustment in Fig. 7. Because of the lack of data on power factors and inaccuracy in reactive power consumption of each appliance, Fig. 9 shows that disaggregation on reactive yields large errors. The post adjustment in Fig. 7 removes the appliance with the highest kVAR if the detected aggregate kVAR is larger than the measurement. For example, clothes dryer is identified as on at 17:00 before post adjustment. However due to its large kVAR and that the aggregate kVARh exceeds the measurement, it is removed and leaves a large portion of kWhs unexplained.

It is believed that if more accurate reactive powers are obtained, the multi-objective optimization will yield better results for both kWh and kVARh disaggregation. Recall that appliance inference in step one provides the usage frequency as well in addition to UECs. The disaggregation algorithm could also set the usage frequency as constraints to further narrow down the searching space. It is also desired to feedback disaggregation results to the appliance inference step to adjust inference accuracy.

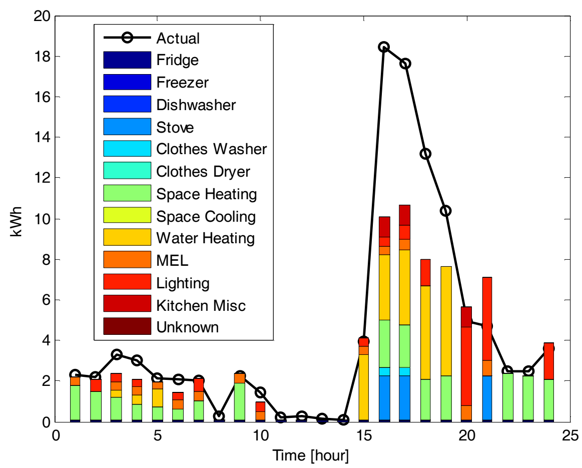


Figure 8. kW Results for Type-1 Profile in Winter

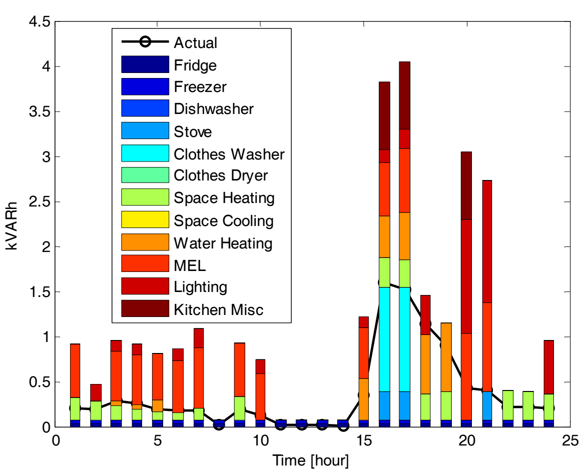


Figure 9. kVAR Results for Type-1 Profile in Winter

## VI. CONCLUSION

A non-intrusive disaggregation algorithm for parallel real and reactive energy at 1-hour frequency is proposed in this paper. Appliance information for different dwelling types are derived based on appliance ownership data through multi-objective optimization of energy matching and likelihood maximization. The disaggregation algorithm implements parallel optimization of real and reactive energy matching and likelihood maximization. The optimal solutions are tuned by appliance dependence rules and usage frequencies. With 1-hour frequency AMI data, this disaggregator is able to disaggregate more than 50% of the aggregate data. Much better performance will be gained if more data on reactive power of each appliance are acquired. The disaggregation results will be used as the prediction of next-day load composition, and provided to the VVO engine for Volt/VAR injection planning.

## REFERENCES


[1] M. Manbachi, H. Farhangi, A. Palizban and S. Arzanpour, "Impact of V2G on Real-Time Adpative Volt/Var Optimization of Distribution Networks," in *IEEE Electrical & Energy Conference*, Halifax, NS, 2013.

[2] O. Parson, "Unsupervised Training Methods for Non-intrusive Appliance Load Monitoring from Smart Meter Data," University of Southampton, Southampton, 2014.

[3] S. Makonin, "Real-Time Embedded Low-Frequency Load Disaggregation," Simon Fraser University, Burnaby, 2014.

[4] John Powers, Bedig Margossian, and Bruce Smith, "Using a Rule-Based Algorithm to Disaggregate End-Use Load Profiles from Premise-Level Data," *IEEE Computer Applications in Power,* vol. 4, no. 2, pp. 42-47, 1991.

[5] G. Hart, "Nonintrusive Appliance Load Monitoring," *Proceedings of IEEE,* vol. 80, pp. 1870-1891, 1992.

[6] J.Z. Kolter and T. Jaakkola, "Approximate Inference in Additive Factorial HMMs with Application to Energy Disaggregation," in *15th International Conference on Artificial Intelligence and Statistics*, La Palma, 2012.

[7] H. Kim, M. Marwah, M.F. Arlitt, G. Lyon, and J. Han, "Unsupervised Disaggregation of Low Frequency Power Measurements," in *11th SIAM International Conference on Data Mining*, Mesa, AZ, 2011.

[8] "Energy Consumption of Major Household Appliance Shipped in Canada," Natural Resource Canada, 2010.

[9] "RESIDENTIAL ENERGY CONSUMPTION SURVEY," U.S. Energy Information Administration, 2009. [Online]. Available: http://www.eia.gov/consumption/residential/data/2009/.

[10] E. Wilson, C. Engebrecht Metzger, S. Horowitz, and R. Hendron, "2014 Building America House Simulation Protocols," National Renewable Energy Laboratory, 2014.

[11] "BUILDINGS ENERGY DATA BOOK," U.S. Department of Energy, [Online]. Available: http://buildingsdatabook.eren.doe.gov/ChapterIntro2.aspx. [Accessed 2014].

[12] Philip McKeen, and Alan S. Fung, "The Effect of Building Aspect Ratio on Energy Efficiency: A Case Study for Multi-Unit Residential Buildings in Canada," *Building,* vol. 4, pp. 336-354, 2014.

[13] "Smart Residential Load Simulator (SRLS)," University of Waterloo, [Online]. Available: https://ece.uwaterloo.ca/~ccanizar/software/srls.htm.