Multivariate Exploration of Non-Intrusive Load Monitoring via Spatiotemporal Pattern Network

Chao Liu[a], Adedotun Akintayo[a], Zhanhong Jiang[a], Gregor P. Henze[b,c], Soumik Sarkar[a,*]

[a] *Department of Mechanical Engineering, Iowa State University, Ames, IA 50011, USA*
[b] *Department of Civil, Environmental and Architectural Engineering, University of Colorado, Boulder, CO 80309, USA*
[c] *National Renewable Energy Laboratory, Golden, CO 80401, USA*

Abstract

Non-intrusive load monitoring (NILM) of electrical demand for the purpose of identifying load components has thus far mostly been studied using univariate data, e.g., using only whole building electricity consumption time series to identify a certain type of end-use such as lighting load. However, using additional variables in the form of multivariate time series data may provide more information in terms of extracting distinguishable features in the context of energy disaggregation. In this work, a novel probabilistic graphical modeling approach, namely the spatiotemporal pattern network (STPN) is proposed for energy disaggregation using multivariate time-series data. The STPN framework is shown to be capable of handling diverse types of multivariate time-series to improve the energy disaggregation performance. The technique outperforms the state of the art factorial hidden Markov models (FHMM) and combinatorial optimization (CO) techniques in multiple real-life test cases. Furthermore, based on two homes' aggregate electric consumption data, a similarity metric is defined for the energy disaggregation of one home using a trained model based on the other home (i.e., out-of-sample case). The proposed similarity metric allows us to enhance scalability via learning supervised models for a few homes and deploying such models to many other similar but unmodeled homes with significantly high disaggregation accuracy.

*Corresponding author.
Email address: soumiks@iastate.edu (Soumik Sarkar)

Keywords: Non-intrusive load monitoring (NILM); spatiotemporal pattern network (STPN); multivariate time-series.

1. Introduction

Non-intrusive load monitoring is a well-established problem that involves disaggregating the total electrical energy consumption of a building into its constituent electric load components without the necessity for extensive metering installations on individual end-uses. Such problems are relevant and challenging from the perspective of software-based detection for control of end-use patterns as well as the future internet of things (IoT). Recently, such problems have been gaining widespread attention due to the potential benefits of energy disaggregation for the purposes of energy efficiency and the development of smart grid systems [1–3], especially in the presence of increasing penetration of renewable energy resource [4]. An overview of load disaggregation concepts was provided in [5], beginning from the pioneering disaggregating technique in [6] that detected sharp changes in signals to optimization of error terms for pattern detection using genetic algorithms [7]. Fourier transforms of end-use signals also provided a useful categorization of their patterns [8]. Authors in [8] evaluated the performance of factorial hidden Markov models (FHMM)[9] for the energy disaggregation problem.

Recent approaches have either attempted determining the hidden states of the FHMM using modified tractable sparse Viterbi algorithm [10], but the methods cannot scale properly for multivariate analysis without conditioning. With conditional Factorial Semi-Hidden Markov Model (CFSHMM) [11], the internal states are conditioned on the additional appliances features, in addition to a suitably chosen prior shape distribution representing the appliances relationships. CFHSMM's maximization step however assumes independence and cannot handle the relational dependencies that we consider here.

In (Dinesh et al. 2016), spectral and energy level information of appliances are determined from the subspace components derived from the active power, after which mean shift is used to classify each from short windows. However, using aided Linear Integer Programming ALIP [12] mainly included extra constraints, some appliance based knowledge and median filtering to improve the computational efficiency of the ALIP on several appliances. The usage of such ALIP is still limited to univariate disaggregation. Indeed, many past work on non-intrusive load monitoring for energy disaggregation

has been primarily based on univariate real power measurements at different sampling intervals [3, 13–20].

With the importance of NILM, there was the need for standardized datasets for benchmarking the large variety of published algorithms. In that light, authors of [21] were motivated to build the Reference Energy Disaggregation Data set (REDD), facilitated by hardware and software systems designed to collect real and reactive power of appliances from multiple homes in Boston, MA in 2011. An open source toolkit, NILMTK was subsequently developed [22] as a common platform to enhance the reproducibility of the algorithms' results on available datasets such as REDD [21], BLUED [23], Smart* [24], and several others. For the purpose of benchmarking the results of the algorithms, several performance evaluation metrics had been proposed [25–27].

The motivation of this present work is the vision of a smart grid [28–31], in which grid operators know what electric load components dominate the demand at each point of service and any point in time. Such capability enables the operators to call on flexible demand side resources for demand response and demand shaping that is perhaps due to large fluctuations of renewable energy sources. This vision of the smart grid involves estimation of the electric load breakdown, not only at the level of a single home but at different levels of aggregations such as residential transformers serving a few homes or distribution feeders serving many hundreds of homes [32]. In this context, a univariate measurement is possibly insufficient for the task of identifying the dominant load contributions such as air-conditioning related demand. Instead, at any level of aggregation, we hypothesize that a multivariate measurement of several concurrent electrical properties may improve the inference on what load components are contributing to the transformer or feeder demand. This may also work for other scenarios such as thermal load monitoring [33], energy intensity of domestic activities [34], and specific saving of appliances [35].

In electric power systems, magnitude and phase angle of the various wave forms can be presented by a complex number referred to as a phase vector, or phasor, and it can be measured by a newly developed phasor measurement unit (PMU) [36]. The wealth of measured multivariate time series data allows for the previously unavailable assessment of the incremental value of measuring not only the real power but several other electric system variables for the purpose of identifying usage patterns and load components. In addition, there are other resources (e.g., weather data, building property, occupancy and usage logs) that can provide multivariate data for the NILM. With this

kind of multi-source information perspective, the performance of the energy disaggregation might be improved.The authors in [11] considered the disaggregation problem by deriving multi-appliance relationship. The authors in [28] proposed a knowledge-based classification of appliances signals based on measured variables into triangular or rectangular sections to perform disaggregation. Here however, data-driven NILM using multivariate relationship to individual end use using information theoretic measure is introduced. A related question that we aim to answer is: If multivariate time series is more useful than univariate time series for energy disaggregation and if so, what are the most valuable measurements to use?

In this context, we apply the recently developed framework of spatiotemporal pattern networks (STPN) that is built on the concept of symbolic dynamics [37]. STPN is proposed to model multivariate time series, via learning atomic patterns (APs, Markov models for individual variables) and relational patterns (RPs, Markov models to model causal interactions between variables) [38, 39]. Patterns of multivariate time series data are formed based on these features (APs and RPs) and then used to study the characteristics of electricity usage.

The state-of-the-art NILM approaches are mostly supervised learning schemes, where measuring the usage of the end-uses is needed [16]. The goal of NILM is thus influenced significantly by the cost of sensor and metering installations. An unsupervised learning method which only employs unlabeled data is presented in [40] via building a set of probabilistic end-use models, yet how to determine the number of labeled end use models is still an on-going problem. However, even in a supervised context, if the number of the sub-metered homes can be reduced while preserving the disaggregation performance on the total number of buildings analyzed, the cost of NILM application will be greatly reduced. Therefore, this work proposes a similarity metric using the STPN concept for out-of-sample disaggregation (a model trained in one home with sub-metered end-uses is applied for disaggregation of another home without end-uses sub-metered). With the similarity metric, one can evaluate the number of homes that need to be sub-metered, and hence the installment cost and disaggregation performance can be estimated.

Contributions: The primary contributions of this paper include: (i) proposing an STPN framework for energy disaggregation using multivariate time-series data, (ii) applying a mutual information based metric to explore the energy consumption patterns and select the most valuable variables for disaggregation, (iii) validation of the STPN scheme in diverse test cases show-

ing that it outperforms the state-of-the-art techniques in NILM, and (iv) proposing a similarity metric for homes/buildings based on STPN for out-of-sample disaggregation, which will enable us to predict end uses power for many homes without necessarily retraining them or expensive submetering in all of them.

Outline: The remaining paper is organized as follows. Section 2 provides a brief background and preliminaries including the definition of STPN and metrics in causality interpretation with STPN. Section 3 presents the approach to explore patterns of multivariate time series to distinguish types of end-uses, the framework of energy disaggregation using multivariate time-series data via STPN, and the similarity metric for out-of-sample disaggregation. Section 4 describes results for the discovered patterns in different kinds of appliances (solar yield, A/C usage, using data sets on Distribution phasor Measurement Units–DMUs), and disaggregation using ECO and RBSAM data sets. Finally, the paper is summarized and concluded with directions for future work in Section 5.

2. Background on Spatiotemporal Pattern Network (STPN)

In the view of multivariate time series data in NILM, features among different signals can be distinguished and used for disaggregation. As the STPN is suitable for extracting spatial and temporal features from time-series data, it has been successfully applied in classification and detection areas [41]. Different from the current NILM approaches that solely use the total power consumption for disaggregation, the proposed setup of STPN that can leverage multiple information sources which can improve the disaggregation performance.

Before introducing STPN, the concept of probabilistic finite state automaton (PFSA) is first defined here as a basis. PFSA is defined in the symbolic space that is generated via time series abstraction (preprocessing and discretization/partitioning [38]). As shown in Fig. 1, the time-series data is discretized into symbol sequences and then state sequences, PFSA is formed using D-Markov machine. More formally, based on a deterministic finite state automaton (DFSA) $\mathcal{D} = (S, Q, \chi)$, a PFSA, an extension to probabilistic setting from DFSA is defined as a pair $\mathcal{P} = (\mathcal{D}, \psi)$, i.e., the PFSA is a 4-tuple $\mathcal{P} = (S, Q, \chi, \psi)$, where:

1. S is a set of finite size for the symbol alphabet and $S \neq \emptyset$;

2. Q is a set of finite size for states and $Q \neq \emptyset$;
3. $\chi : Q \times S \to Q$ is the mapping for state transition;
4. $\psi : Q \times S \to [0, 1]$ is defined as a symbol generation function, i.e., probability morph function which is such that $\sum_{\sigma \in S} \psi(q, \sigma) = 1 \quad \forall q \in Q$, where p_{ij} indicates the probability of the symbol $\sigma_j \in S$ occurring with the state $q_i \in Q$.

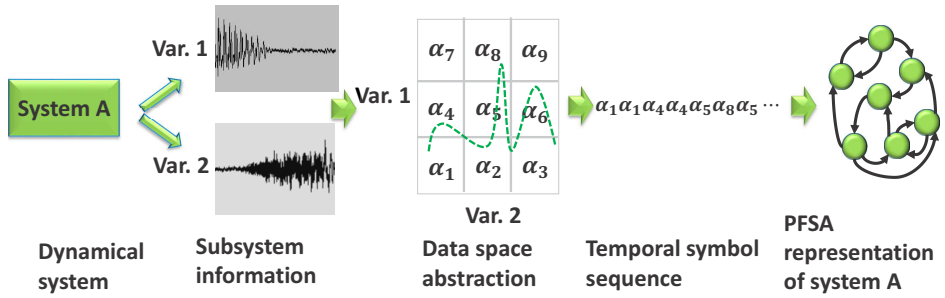


Figure 1: Steps to form a PFSA model with time-series data:

Using the xD-Markov machine defined in [37], a PFSA can be used to explore the relationship of two time-series data as shown in Fig. 2. Let the symbolic system a represent a measurement at the aggregate side (metering data) and symbolic system b represent an end-use i (sub-metering data), the transition matrix Ω^{ab} (relational pattern, RP) is used to characterize features in the total energy consumption due to the end-use i. Atomic pattern (AP) is also shown here that is used to capture the predictability of end-use i based on its past measurements. The metric Λ^{ab} for the patterns (APs and RPs) is used to evaluate the importance of the patterns.

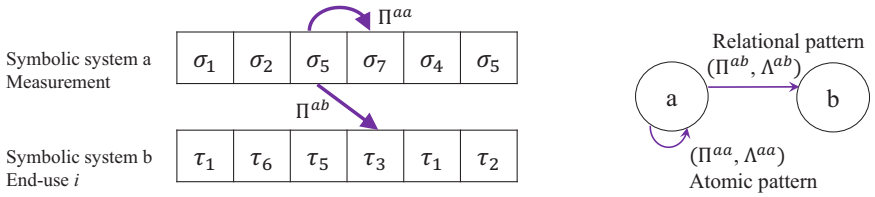


Figure 2: *Extraction of atomic patterns and relational patterns (with D-Markov machine and xD-Markov machine respectively and $D = 1$ for simplicity, i.e., states and symbols are equivalent) to model individual variables and interaction among different variables respectively.*

With the description of PFSA, D-Markov machine, and xD-Markov machine [38], STPN is defined in [39, 42].

Definition. A PFSA based STPN is a 4-tuple $W_D \equiv (Q^a, \Sigma^b, \Omega^{ab}, \Lambda^{ab})$: (a, b denote nodes of the STPN)

(a) $Q^a = \{q_1, q_2, \cdots, q_{|Q^a|}\}$ is the state set corresponding to symbol sequences S^a;
(b) $\Sigma^b = \{\sigma_0, \cdots, \sigma_{|\Sigma^b|-1}\}$ is the alphabet set of symbol sequence S^b;
(c) Ω^{ab} is the symbol generation matrix of size $|Q^a| \times |\Sigma^b|$, the ij^{th} element of Π^{ab} denotes the probability of finding the symbol σ_j in the symbol string s^b while making a transition from the state q_i in the symbol sequence S^a; while self-symbol generation matrices are called atomic patterns (APs) i.e., when $a = b$, cross-symbol generation matrices are called relational patterns (RPs) i.e., when $a \neq b$.
(d) Λ^{ab} denotes a metric that can represent the importance of the learnt pattern for $a \to b$ which is a function of Π^{ab}.

3. Multivariate exploration of NILM via STPN

A pattern discovery framework is first presented in this section using STPN to explore the modeling of multivariate time-series. Then the STPN based framework is proposed for energy disaggregation from multivariate time-series data. Finally, a similarity metric is presented for out-of-sample disaggregation.

3.1. Pattern learning from multivariate time series data via STPN

The primary hypothesis in this paper is that the deployment of multivariate time-series for NILM will improve the energy disaggregation accuracy. The pattern learning scheme presented in this section is intended to validate the hypothesis that some of the patterns among multivariate time-series are more significant in identifying different types of energy generation/consumption than traditional univariate time-series. The DMU data set (details of which are provided in Section 4.1.1) is used to validate the hypothesis. However, due to the lack of ground truth in the DMU data set, we assume that for the same time of day (weekday), the energy generation (by solar panel) and consumption are similar in the same area, while the energy generation and consumption are reasonably different for two different time periods. By extracting the features in two time periods (noted as time

period 1 and 2 respectively), patterns are learnt to represent the type of electric loads in each time period, and the comparisons are carried out for the patterns extracted from different types of electric loads. The goal is to show that multivariate patterns are more effective in identifying the similarity and differences in energy generation/consumption than the univariate patterns.

Formally, for the multivariate time-series, $\mathbf{X} = \{X^{\mathbb{A}}(t),\ t \in \mathbb{N},\ \mathbb{A} = 1, 2, \cdots, f\}$ at the aggregate side (f is the number of time series), the STPN is learned following the definition in Section 2, and the mutual information based metrics are computed. The formulation of STPN provides us the insights of the connections between variables (as shown in Fig. 2), which shows the causal effect from a variable to another. In this context, patterns can be observed in different end-uses in energy disaggregation problem. For time periods 1 and 2 with two kinds of electric loads (noted as $\mathbf{X}_{T1}$ and $\mathbf{X}_{T2}$ respectively), two STPNs can be formed after data abstraction. To evaluate the importance of each pattern $a \rightarrow b$ ($a, b \in \mathbb{A}$) , a mutual information based metric I is defined based on the state sequences q^a and q^b) is as follows.

$$\Lambda^{ab} \triangleq I^{ab} = I(q^b_{k+1}; q^a_{k+1}) = H(q^b_{k+1}) - H(q^b_{k+1}|q^a_k) \tag{1}$$

where,

$$\begin{aligned}
H(q^b_{k+1}) &= -\sum_{i=1}^{Q^b} P(q^b_{k+1} = i) \log_2 P(q^b_k = i) \\
H(q^b_{k+1}|q^a_k) &= \sum_{i=1}^{Q^a} P(q^a_k = i) H(q^b_{k+1}|q^a_k = i) \\
H(q^b_{k+1}|q^a_k = i) &= -\sum_{j=1}^{Q^b} P(q^b_{k+1} = j|q^a_k = i) \cdot \log_2 P(q^b_{k+1} = j|q^a_k = i)
\end{aligned} \tag{2}$$

A detailed description of mutual information based metric in the context of APs and RPs can be found in [38]. An application of the information based metric for STPN is shown in [43].

Using the information based metric in Eq. 1, the importance metric Λ^{ab} of the pattern $a \rightarrow b$ in each time period can be obtained, and the two sets of the metrics are expressed as,

$$\Lambda_{T1} = \{\Lambda^{ab}_{T1}\}\ \forall a, b,\ a, b \in \mathbb{A},$$

$$\Lambda_{T2} = \{\Lambda_{T2}^{ab}\}\ \forall a, b,\ \ a, b \in \mathbb{A},$$

and the variation of the information based metric is obtained (shown in right-panel of Fig. 3),

$$\delta(\Lambda) = |\Lambda_{T1} - \Lambda_{T2}|.$$

The larger the difference is, the more predictive and informative the link (AP or RP) is, which means that this pattern can be used to distinguish the two kinds of end uses.

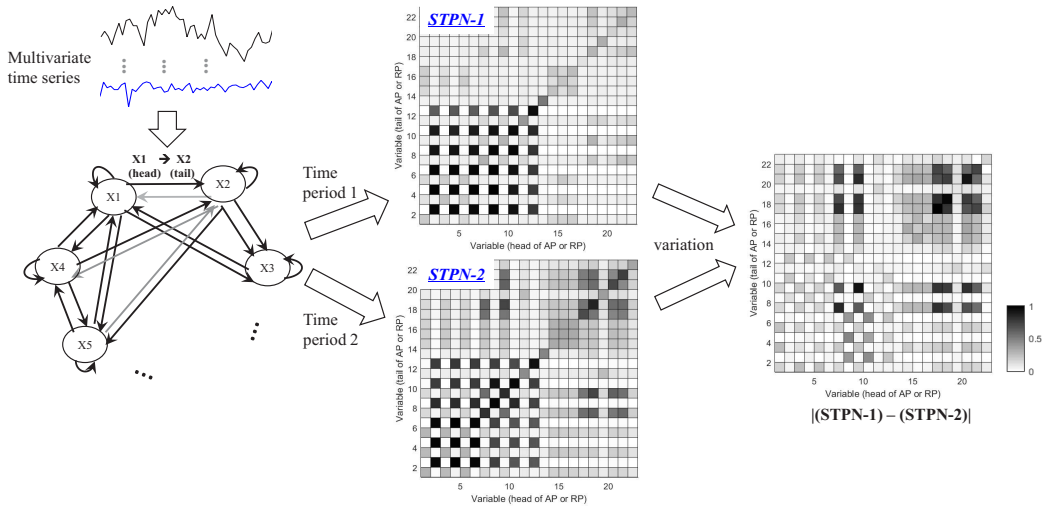


Figure 3: Pattern representation via STPNs. The gray-scale map (mid-panel) represents the links formulated in the left panel, where the variable in x-axis is the head of the AP or RP, and the variable in y-axis is the tail. The value in the links is indicated by the gray scale, the larger value (darker color) shows a stronger connection of the AR or RP, which means strongly predictive and informative. $STPN-1$ and $STPN-2$ (mid-panel) represent two types of end-uses in two time periods, and they are applied to compare the difference between the appliance usage (right-panel), where the variations in the links show the significance of the features (represented by the links) in distinguishing these two kinds of end-uses.

As discussed in Section 1, for the purpose of NILM at different levels (e.g., house-level, transformer-level, feeder-level), the more the load components included in the collected data, the more difficult it is to perform disaggregation, as the problem becomes underdetermined when real power consumption is treated as the sole measured and known variable. The measurement techniques availed us data with more variables, and the STPN formulated here

could answer the question on what the most valuable measurements to collect are, by comparing the significant links in the STPNs. Detail results in this context are presented in Section 4.

3.2. STPN framework for NILM with multivariate measurements

While the previous section applies STPN in exploring patterns in different end-uses, here we introduce the STPN framework for energy disaggregation using multivariate time-series data. The multivariate time-series data can be in two categories: (i) phasor measurements (e.g., PMU) including voltages, currents, power, etc; (ii) power measurement at the aggregate side and other available information (e.g., indoor/outdoor temperature, time of day, etc). Let us denote $\mathbf{X} = \{X^{\mathbb{A}}(t),\ t \in \mathbb{N},\ \mathbb{A} = 1, 2, \cdots, f\}$ for both kinds of multivariate time-series data in a home, and $\mathbf{Y} = \{Y^{\mathbb{B}}(t),\ t \in \mathbb{N},\ \mathbb{B} = 1, 2, \cdots, g\}$ for the power consumption for the end-uses. f is the number of time series at the aggregate side, g is the number of end-uses. Note that the time-series of the aggregate side may have different sampling rates (e.g., the temperature is usually measured hourly, while the power measurement can be a couple of seconds or minutes), although up-sampling techniques can be applied.

For the time-series $\mathbf{X}$ and $\mathbf{Y}$, symbol sequences $\boldsymbol{\sigma} = \{\sigma^{\mathbb{A}}(t)\}$ and $\boldsymbol{\tau} = \{\tau^{\mathbb{B}}(t)\}$ can be obtained after data processing and partitioning. To form an STPN for disaggregation, joint symbol sequences

$$\sigma^J = \sigma^1(t) \oplus \cdots \oplus \sigma^f(t)$$

and

$$\tau^J = \tau^1(t) \oplus \cdots \oplus \tau^g(t)$$

are formed based on the symbol sequences $\boldsymbol{\sigma}$ and $\boldsymbol{\tau}$ respectively. Note, the joint symbol space is generated via the direct sum of the individual symbol spaces. For example, the direct sum $\sigma^a \oplus \sigma^b$ defines the product space of σ^a and σ^b. Here, we use the depth $D = 1$, which means the symbol and the state are equivalent. Then we can have the state sequences Φ and Ψ generated from the joint symbols sequences respectively.

With the setup, the learning stage of STPN is to compute the transition matrix $\boldsymbol{\Omega}(\Phi, \boldsymbol{\tau^J})$ from the states in Φ to the symbols in $\boldsymbol{\tau^J}$ using a frequentist's approach (e.g. counting the number of occurrences). For example, the probability of the state Φ_m to the symbol τ_n^J can be computed by

$Pr(\phi_m, \tau_n^J) = N_{mn}/N_m$, where N_{mn} is the number of times that the symbol $\tau_n^J \in \boldsymbol{\tau^J}$ is emanated after the state $\phi_m \in \Phi$, i.e.,

$$\begin{aligned} N_{mn} \triangleq |\{(\phi(k), \tau^J(k+1)) : \tau^J(k+1) = \tau_n^J \mid \phi(k) = \phi_m\}|, \\ m = 1, 2, \cdots, \prod_{\mathbb{A}=1}^{f} |\sigma^{\mathbb{A}}|, \quad n = 1, 2, \cdots, \prod_{\mathbb{B}=1}^{g} |\tau^{\mathbb{B}}|, \end{aligned} \tag{3}$$

where $\phi(k)$ is the k^{th} state in the state sequence Φ and $\tau^J(k+1)$ is the $(k+1)^{th}$ symbol in the symbol sequence $\boldsymbol{\tau^J}$. The time lag equals to 1 in this case, and

$$N_m \triangleq \sum_{n=1}^{\prod_{\mathbb{B}=1}^{g} |\tau^{\mathbb{B}}|} N_{mn} \tag{4}$$

The real values represented by the symbols $\boldsymbol{\tau^J}$ can be also computed from the training data, and noted as $E\big(\mathbf{Y}|\tau^J = n\big), n = 1, 2, \cdots, \prod_{\mathbb{B}=1}^{g} |\tau^{\mathbb{B}}|$. In the disaggregation stage, only the time-series $\tilde{\mathbf{X}}$ is known, the time-series of end-uses $\tilde{\mathbf{Y}}$ is assumed unknown and used for testing. Similarly, the symbol sequences $\tilde{\boldsymbol{\sigma}}$ and joint symbol sequence σ^J are obtained after partitioning, and the corresponding state sequences are generated using D-Markov and xD-Markov machines. Then the power consumptions of the end-uses $\mathbf{Y}_s$ are obtained,

$$\mathbf{Y}_s = \sum_{n=1}^{\prod_{\mathbb{B}=1}^{g} |\tau^{\mathbb{B}}|} Pr(\tau^J|\tilde{\phi}(t)) \cdot E\big(\mathbf{Y}|\tau^J = n\big) \tag{5}$$

With multivariate time-series data, the disaggregation formed in this work is noted as the STPN framework, as shown in Fig. 4. When there is only one time-series at the aggregate side (usually the total energy consumption), it is a one-to-one relationship between the total energy consumption, we note it as the PFSA approach, which follows the definition in Section 2.

It should be noted that, increasing the number of symbols during partitioning will preserve the more information for disaggregation in terms of keep finer resolution for the expectation in Eq. 5. However, the larger number of symbols will significantly increase the dimension of the transition matrix, and hence cause difficulty in stability of model learning. Also, due to the use of joint states, the number of states will surge when several variables

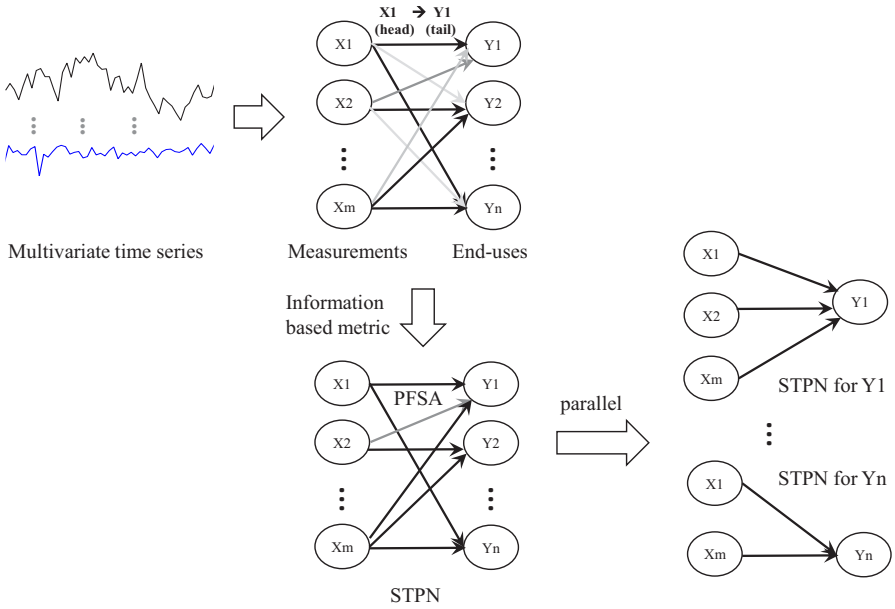


Figure 4: Formulation of STPN for NILM with multivariate measurements. The multivariate time-series (measurement–V_i and end-use–A_j) are represented as nodes in a graphical model (shown in the top panel), the ability of predicting the end-use from each measurement is estimated via information based metric (e.g., mutual information) based on the formulation of PFSA in Section 2 and Fig. 1. The selected time-series are used in the STPN model (shown in the bottom panel) to learn the disaggregation model. The STPN can be implemented in a parallel manner to improve computational efficiency (shown in the bottom-right panel) where an STPN is formed for the disaggregation of one end-use.

are considered. A state merging approach is introduced to avoid the above issues. To implement the state merging, states that are less informative are first identified. The metric $\gamma(r)$ to evaluate the importance of the state is defined as,

$$\gamma(r) = \left|\left|Pr(\phi_r, \tau_n^J) - \overline{Pr}(\phi_r, \tau^J)\right|\right|_1, \quad r = 1, 2, \cdots, \prod_{\mathbb{A}=1}^{f} |\sigma^{\mathbb{A}}| \tag{6}$$

where

$$\overline{Pr}(\phi_r, \tau^J) = \sum_{n=1}^{\prod_{\mathbb{A}=1}^{g} |\tau^{\mathbb{A}}|} Pr(\phi_r, \tau_n^J) / \prod_{\mathbb{B}=1}^{g} |\tau^{\mathbb{B}}|$$

If $\gamma(r) < \eta$ where η is a specified threshold, the state ϕ_r is identified to be merged to other states. Then the relevance $\Gamma(r, s)$ of the two states is defined as,

$$\Gamma(r, s) = \sum_{n=1}^{\prod_{\mathbb{B}=1}^{g} |\tau^{\mathbb{B}}|} \left|\left|Pr(\phi_r, \tau_n^J) - Pr(\phi_s, \tau_n^J)\right|\right|_1, \quad r, s = 1, 2, \cdots, \prod_{\mathbb{A}=1}^{f} |\sigma^{\mathbb{A}}| \tag{7}$$

$\Gamma(r, s)$ can be applied to find out the closest state ϕ_s to be merged, where $\gamma(s) \geq \eta$. Also, we can have the states that $\Gamma(r, s) < \mathcal{D}$ to be merged, where the ϕ_r and ϕ_s are the states with very similar transition probabilities, here $\mathcal{D}$ is a specified threshold.

For the end-uses, it can include all of them in one STPN model or several of them which have strong dependency. Also, it can be used for one-by-one disaggregation, where each STPN model is learnt to predict one end-use. If all end-uses are included in the model, the learning process automatically preserves the property such that the sum of the end-use power is equal to the total power consumption. However, when there is a large number of end-uses, the learning process may become computationally expensive.

For the number of multivariate time-series ($\mathbf{X}$) at the aggregate side, tens of variables may be available in different scenarios, (e.g., DMU measures 26 variables at the transformer). The more variables used, the more information is included in the model, while the computational cost is also increased. Here, a mutual information based metric is applied to select the most valuable variable. Based on Eq. 2, $\mathbf{X}^u$ is the variables that are currently included in STPN, and $\mathbf{X} \setminus \mathbf{X}^u$ are the candidates can be added into STPN model, the

most valuable variable X^e is determined by the added time-series that has the maximal mutual information among all the candidates.

$$X^e = \arg\max\{\Lambda^{(X^u \cup X^e) \to Y}\}, X^e \in \mathbf{X} \setminus \mathbf{X}^u \tag{8}$$

where $\Lambda^{(X^u \cup X^e) \to Y}$ is the mutual information metric (defined in Eq. 1) of the pattern $(X^u \cup X^e) \to Y$–from the aggregate side (with the newly added variableX^e) to the end-use(s).

Remark 3.1. *Our recent work has also reported the relevant methodologies of non-intrusive load monitoring using spatiotemporal pattern network [41]. However, the previous results focused on univariate performance via STPN as Whole Building Electric (WBE) consumption was the only known variable applied to disaggregate other variables. In this paper, multivariate exploration is carried out to further explore relationships among other variables besides WBE such that the disaggregation for each variable is improved compared to the univariate case. Therefore, a new STPN based inference framework is proposed in this work to use the multivariate time-series for prediction.*

3.3. STPN framework for out-of-sample disaggregation

The scheme of out-of-sample disaggregation is shown in Fig. 5, as well as the proposed STPN framework. A similarity metric $\mathbb{C}$ based on STPN is presented to estimate the closeness of the homes based on the aggregate data (the left panel of Fig . 5).

The assumption here is that, the multivariate time-series data $\mathbf{X}_\mathcal{P}$ at the aggregate side are available for all of the homes (with the number of homes P), while only some of the homes have sub-metered data $\mathbf{Y}_\mathcal{P}$ for the end-uses.

Using the multivariate time-series data at the aggregate side, let $\mathbf{X}_M = \{X_M^{\mathbb{A}}\}$ and $\mathbf{X}_T = \{X_T^{\mathbb{A}}\}$ represent the data from home M (for modeling) and home T (for disaggregation). The time series is then symbolized into $S_M = \{S_M^{\mathbb{A}}\}$ and $S_T = \{S_T^{\mathbb{A}}\}$, then state sequences are generated with the STPN formulation, noted by Φ_M and Φ_T. The similarity metric $\mathbb{C}$ is defined that $\mathbb{C}_{M|T}$ is related to the probability of a sequence S_M contains the similar information from the other sequence S_T in terms of the features extracted from the STPN,

$$\mathbb{C}_{M|T} \propto \hat{Pr}(S_M|S_T) \tag{9}$$

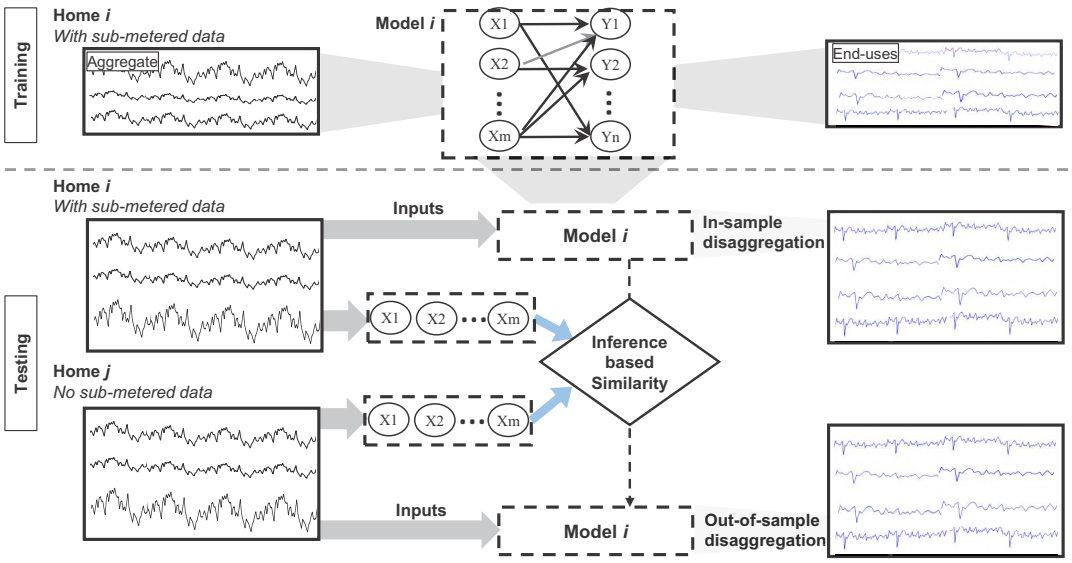


Figure 5: STPN framework for out-of-sample disaggregation. The model i is learned with the training data where the aggregation and end-uses are all measured. In the testing stage, the total energy consumption and other available data are used to compute the similarity metric between two STPN models marked in blue for both homes. The in-sample disaggregation is the prediction of end-uses of home i with model i, while the out-of-sample disaggregation is the prediction of end-uses of home$j$ using model i. The similarity metric presented here is used to evaluate the confidence that whether the model i can be used in the disaggregation of home j in terms of out-of-sample disaggregation performance.

where $\hat{Pr}(S_M|S_T)$ is the conditional probability of the symbol subsequence S_M given the pattern S_T,

$$\hat{Pr}(S_M|S_T) = \prod_{\mathbb{A}=1}^{f} \prod_{m=1}^{|Q_T^{\mathbb{A}}|} \frac{(\mathcal{N}_m^{\mathbb{A}})!(\mathrm{N}_m^{\mathbb{A}} + |\Sigma_M^{\mathbb{A}}| - 1)!}{(\mathcal{N}_m^{\mathbb{A}} + \mathrm{N}_m^{\mathbb{A}} + |\Sigma_M^{\mathbb{A}}| - 1)!} \prod_{n=1}^{|\Sigma_M^{\mathbb{A}}|} \frac{(\mathcal{N}_{mn}^{\mathbb{A}} + \mathrm{N}_{mn}^{\mathbb{A}})!}{(\mathcal{N}_{mn}^{\mathbb{A}})!(\mathrm{N}_{mn}^{\mathbb{A}})!} \tag{10}$$

Also,

$$\mathrm{N}_{mn}^{\mathbb{A}} \triangleq |\{(Q_T^{\mathbb{A}}(k), S_T^{\mathbb{A}}(k+1)) : S_T^{\mathbb{A}}(k+1) = \sigma_{T,n}^{\mathbb{A}} \mid Q_T^{\mathbb{A}}(k) = q_{T,m}^{\mathbb{A}}\}|$$

$$\mathrm{N}_m^{\mathbb{A}} = \sum_{n=1}^{|\Sigma_T^{\mathbb{A}}|} (\mathrm{N}_{mn}^{\mathbb{A}})$$

$$\mathcal{N}_{mn}^{\mathbb{A}} \triangleq |\{(Q_M^{\mathbb{A}}(k), S_M^{\mathbb{A}}(k+1)) : S_M^{\mathbb{A}}(k+1) = \sigma_{M,n}^{\mathbb{A}} \mid Q_M^{\mathbb{A}}(k) = q_{M,m}^{\mathbb{A}}\}|$$

$$\mathcal{N}_m^{\mathbb{A}} = \sum_{n=1}^{|\Sigma_M^{\mathbb{A}}|} (\mathcal{N}_{mn}^{\mathbb{A}})$$

$|Q_T^{\mathbb{A}}|$ is number of states emanated from time series $X_T^{\mathbb{A}}$, $|\Sigma_T^{\mathbb{A}}|$ is number of symbols in symbol sequence $S_T^{\mathbb{A}}$, $\mathbb{A} = 1, 2, \cdots, f$, and $|\Sigma_M^{\mathbb{A}}|$ is number of symbols in symbol sequence $S_M^{\mathbb{A}}$.

For more detailed derivation, please refer to [39, 44]. Note that, joint state sequences Φ_M and Φ_T can also be applied to compute the conditional probability.

With the similarity metric $\mathbb{C}_{M|T} = K(\hat{Pr}(S_M|S_T))$, where K is a proportional constant, the out-of-sample disaggregation can be evaluated. The two homes with higher similarity by $\mathbb{C}$ are expected to contain similar features in energy consumption and may imply the comparable end-uses. Thus, the model trained in one home can be transferred to another home.

Remark 3.2. *The out-of-sample disaggregation is aimed to form a semi-supervised scheme when applying NILM into a new residential area. The proposed similarity metric can serve as a metric for clustering the homes into groups based on the measurements at the aggregate side (no requirement of submetering). After clustering, disaggregation can be implemented in groups where one model is trained per group using a representative home from the group. In this paper, we demonstrate using a real data set that a high degree of disaggregation accuracy can be preserved in this manner while achieving significant scalability.*

4. Results and Discussion

Three data sets are applied in this work: (i) DMU data set–phasor measurements at the transformer level, (ii) ECO data set–multivariate measurements with ground truth of end-uses, and (iii) RBSAM data set–real data in 61 homes with four types of end-uses. Further details of each data set are illustrated in the corresponding sections.

4.1. DMU data set: energy-pattern exploration

Two cases are analyzed in this work for validating the application of STPNs in extracting features from multivariate time series data. The first case intends to find out the unique features when the photovoltaic (PV) systems are working hard and exporting net power to the grid, while the second case tries to determine the features of A/C usage.

4.1.1. DMU data set and problem setup

Phasor measurements that occur at the same time are called synchronized phasors or synchrophasors and the measurement device deployed for this purpose is a phasor measurement unit (PMU). Utilizing the common time source of a GPS radio clock allows for the synchronization of geographically dispersed PMUs. Synchrophasor technology helps electric system operators and planners measure the state of the electrical system and manage power quality. In the past, the majority of PMUs are installed at the transmission level to monitor the dynamic response of the bulk grid.

When the synchrophasor measurements are specifically taken at the low-voltage level of a distribution system, DMUs. The National Renewable Energy Laboratory (NREL) has developed GPS-synchronized DMUs [36], which measure multiple voltages, currents and power quantities and have been installed on the secondary side of distribution transformers in distribution circuits with high solar penetrations in several utility service areas. Real-time data has been collected from those circuits for several years. Through a project with a partner utility company in central California, NREL has installed fifteen DMUs on the secondary side of 50-kVA and 75-kVA pad-mounted split-phase residential transformers [45]. All of these transformers are in the same community, which contains a high concentration of residential rooftop PV generation.

The DMU data are collected in 26 variables (22 of them are listed in Table 1. The remaining four are also analyzed and the results show that

they are not sufficiently informative in terms of the presented patterns). This data set provides the opportunity to explore the disaggregation problem in a multivariate time-series context.

Table 1: Variables measured in DMU.

1-6	$V1$-PM[1], $V1$-PA[2], $V2$-PM, $V2$-PA, $V12$-PM, $V12$-PA
7-13	$I1$-PM, $I1$-PA, $I2$-PM, $I2$-PA, In-PM, In-PA, Frequency
14-19	$V1$-RMS[3], $V2$-RMS, $V12$-RMS, $I1$-RMS, $I2$-RMS, In-RMS
20-22	Apparent Power Magnitude, Real Power, Reactive Power

[1] Phasor Magnitude, [2] Phase Angle, [3] Root Mean Square.

Using STPN, the APs and RPs of the 22 variables are computed with information based metric (as shown in Section 2), the framework to validate the efficacy of STPN in disaggregation is shown in Fig. 3, where two time periods are applied to form two STPNs. The reason for using two time periods is that each data set is collected at the transformer level, including five to ten homes and this makes the ground truth difficult to establish when there are no sub-metering devices in every home. Therefore, the *assumption* here is: the two time periods are chosen based on the fact that the electric loads operating in each time period are reasonably different. For example, there will be very little A/C usage in the spring (except fans), while there will be heavy A/C loads in the summer. Another example is that the characteristics of appliance usage in the morning should be different from that in the night (e.g., 3-6AM).

Based on the assumption above, the patterns extracted from each time period represent specific features of the appliances. The variation of the patterns will reflect the characteristics of different appliances. That is to say, larger variation of a feature (a feature means a block in the right panel of Fig. 3 and is the AP or RP between the variables) indicates the higher possibility of distinguishing between the two types of load patterns using this feature. This will provide evidence that the measurement (used to build the feature) is valuable and beneficial for disaggregation.

4.1.2. Patterns in distinguishing solar yield

For the first case, collected data on July 11, 2013, is used, where two time periods are selected as 3-6AM and 9AM-12PM. The time period 1 (3-6AM) is assumed to have no solar yield, considering the sunrise time in central

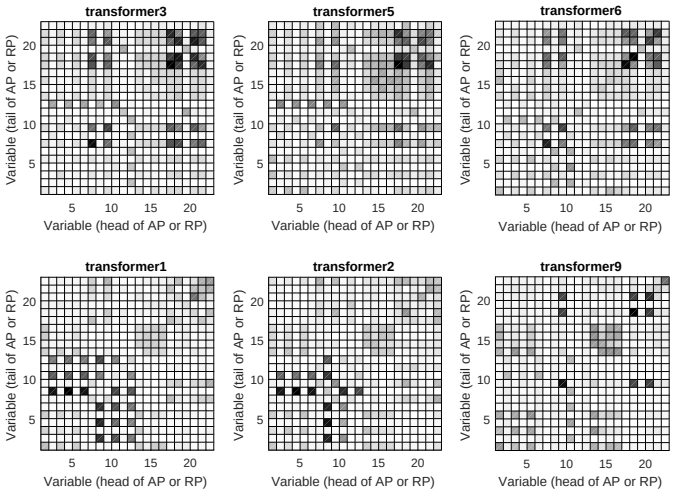


Figure 6: Variation of patterns (APs and RPs) between 3-6AM and 9AM-12PM on July 11, 2013. The patterns are used to compare unique features when PV system outputs net power to the grid.

California (the location where DMUs are installed). On the other hand, there is increasing solar yield during time period 2 (9AM-12PM), if the PV works in good condition.

Using the approach presented in Section 3.1, APs and RPs are extracted from data set of 10 transformers, where six of them are shown in Fig. 6. By comparing the patterns, we find that the patterns in transformers 3, 5, 6 are similar in terms of the top-right corner of APs and RPs (features, $\Lambda^{i,j}, i, j = 17, 18, 20, 21$). The patterns in transformers 1, 2, 9 are similar but very different from the previous three transformers (the features in the top-right corner are different from the previous ones).

According to domain knowledge, the solar yield is a possible reason for the observed differences. To find out the solar yield in this area, the real power in the transformers mentioned above is plotted in Fig. 7. In the top panel of Fig. 7, the solar yield of transformers 3, 5, and 6 increases as the time goes on (where the negative net real power means the PV outputs power to the grid), while transformers 1, 2, and 9 do not present that obvious increasing solar yield. This may be caused by the PV system orientation in those homes that are connected to these transformers.

Comparing between the two sets of the transformers, the features ($\Lambda^{i,j}, i, j = 17, 18, 20, 21$) discussed above are sensitive to the solar yield, and they can be

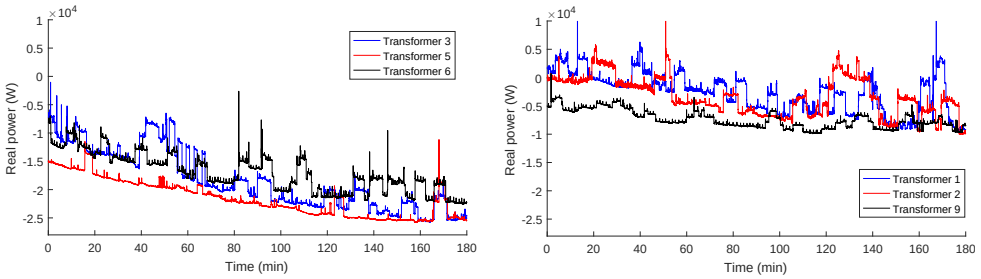


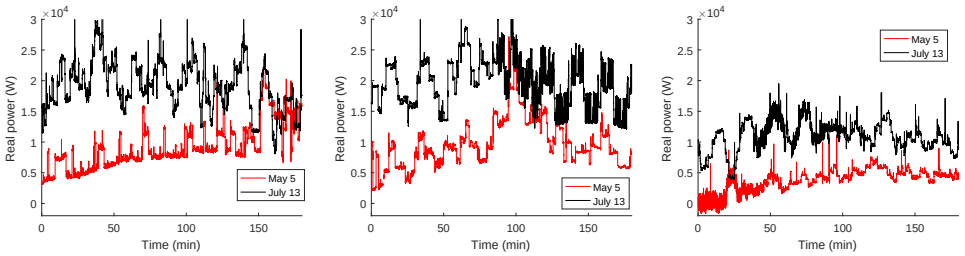
Figure 7: Real power measurements (9AM-12PM, time period 2) of the transformers corresponding to Fig. 6. As there is no solar yield from 3-6AM (time period 1), real power is always positive (consuming power) and not shown here.

applied in distinguishing the generation of solar energy, which corresponds to the four parameters, $I1$ RMS, $I2$ RMS, apparent power magnitude and real power. If we only use the above four variables, a 4×4 matrix is formed in each transformer with APs and RPs representing the information based metric between the variables. The matrices can be used to compute the similarity between transformers. Treating the matrix as n vectors in an image (similar to the definition of structural similarity, SSIM, in image quality assessment and other applications [46, 47]), SSIM is applied. Using transformer 3 as a reference, the similarities of transformers 5, 6, 1, 2, and 9 to transformer 3 are 0.97, 0.81, 0.41, 0.07, 0.29, respectively. This shows that transformers 3, 5, and 6 have similar usage patterns, while transformers 1, 2, and 9 have different usage patterns.

Note that the real power presents the ability to distinguish the solar yield in transformers 3 and 6 (as the AP of real power $\Lambda^{21,21}$ is the most significant one among all of the features), but RPs ($\Lambda^{17,21}$, $\Lambda^{21,17}$) are more informative than $\Lambda^{21,21}$ in transformer 5. This is an evidence that measurements of multiple variables are important in disaggregating the solar yield.

4.1.3. Patterns to distinguish A/C usage

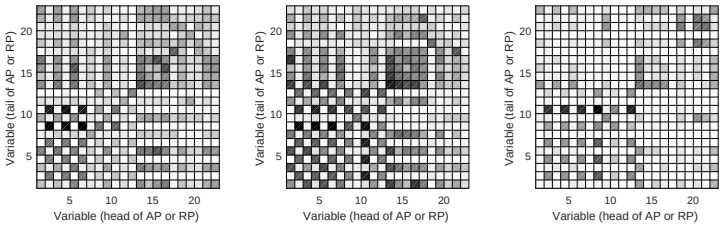
Air-conditioning is a dominant load component during summer, and its disaggregation is important for optimizing grid operation and other applications. With the multivariate time series data collected by DMUs, two time periods are chosen in two dates, one is 7-10PM on May 5, 2013, the other one is 7-10PM July 13, 2013. The sunset time on the two days are 8:02 PM and 8:31 PM, respectively [48]. This means that the solar yield is similar in the two time periods of each transformer, with decreasing yield in the beginning



(a) Transformer 4 (b) Transformer 5 (c) Transformer 9

Figure 8: Real power measurements during 7-10PM May 5 and July 13, 2013 respectively. Although the levels of power consumption are different among transformers, the same observation is that the power consumption significantly increases on July 13.

and no solar energy output later after dark. The temperature on May 5 is 55-75 deg F, and on July 13 is 57-94 deg F [49], where the latter one is typical summer weather. Both days are weekends, and the usage of appliances is assumed to be close, except A/C, because of higher temperatures on the second day. The *assumption* for this situation is that A/C heavily operates during time period 2 (7-10PM, July 13), while there is very little to no A/C operation during time period 1 (7-10PM, May 5). The real power consumption on the two days is shown in Fig. 8. The power consumption is significantly higher on July 13, and this is reasonable based on the assumption of A/C usage.



(a) Transformer 4 (b) Transformer 5 (c) Transformer 9

Figure 9: Variation of patterns (APs and RPs) during 7-10PM between May 5 and July 13, 2013. The patterns are used to compare unique features of A/C usage on typical days of spring and summer.

Based on this assumption, we are looking at the patterns of different features formed by the 22 measured variables, and the results are shown in Fig. 9.

From Fig. 9, RPs are shown to be more informative, e.g., $\Lambda^{2,8}$, $\Lambda^{4,8}$, $\Lambda^{6,8}$, $\Lambda^{2,10}$, $\Lambda^{4,10}$, $\Lambda^{6,10}$ in transformer 4; $\Lambda^{2,8}$, $\Lambda^{4,8}$, $\Lambda^{6,8}$, $\Lambda^{10,2}$, $\Lambda^{10,2}$, $\Lambda^{10,4}$, $\Lambda^{10,6}$, $\Lambda^{10,8}$ in transformer 5; $\Lambda^{10,2}$, $\Lambda^{10,4}$, $\Lambda^{10,6}$, $\Lambda^{10,8}$, $\Lambda^{10,12}$ in transformer 9. These patterns are stronger than APs, which means that the RPs are more beneficial in this case in terms of distinguishing the A/C loads. Note that the RPs are not symmetric (e.g., $\Lambda^{10,2}$ has large value in transformer 9, while $\Lambda^{2,10}$ is much smaller). Also, real power (variable 21) is not shown as an important feature here, although the power consumption on the two days are significantly different.

In this case, the phase angles of $V1$, $V2$, $V12$, $I1$, $I2$, and In present important contributions on distinguishing heavy A/C usage, and this validates that measuring these variables is beneficial for A/C load monitoring.

Based on the DMU data, applications of STPNs in exploring features of solar yield and A/C usage are carried out in this work, with multivariate time series data collected via distribution phasor measurement units. The results show that: (i) STPNs are capable of extracting features in multivariate time series data for energy disaggregation, (ii) in some cases, relational patterns (between variables) are more significant in distinguishing usage patterns than atomic patterns, and this validates that measuring more electrical variables at the transformer level is beneficial for disaggregation, (iii) the most informative variables can be determined by the metric formed in STPN.

It should be noted that, due to the lack of ground truth (metered data of load components is not available, as each transformer has a few homes connected, and each home has several appliances), assumptions were made based on the usual appliance usage with times of a day, days of a week, and weather (temperature, sunlight) conditions. Based on discussions with domain experts, these assumptions are reasonable in the authors' view. Further analysis is being implemented with more validation on the assumptions and collection of ground truth data.

4.2. Case study with multivariate measurements at the aggregate side using ECO data set

4.2.1. Data set and problem setup

Electricity Consumption and Occupancy (ECO) data set was collected in 6 Swiss households during a period of 8 months. For each home, the ECO data set provides: (i) 1 Hz aggregate consumption data. Each measurement contains data on current, voltage, and phase shift for each of the three phases in the household; (ii) 1 Hz plug-level data measured from selected appliances;

(iii) Occupancy information measured through a tablet computer (manual labeling) and a passive infrared sensor (in some of the households) [50, 51]. For the main meter, the ECO data set has 16 metered variables including total power consumption, powers in three phases, neutral current, currents in three phases, voltages in three phases, and phase angles of $V12$, $V13$, $I1$-$V1$, $I2$-$V2$ and $I3$-$V3$. For the sub-metered data, it consists of several kinds of plug usage, such as tablet, dishwasher, fridge, freezer, kettle, lamp, stove, TV, stereo, etc. Using the ECO data set, we have multivariate time-series data for the aggregated power, and the ground truth of the sub-metered appliances, which we can test our algorithm on the energy disaggregation in the circumstance of multivariate measurements at the aggregate level.

In this work, the data set of house 2 is applied during Jan. 2013. There are 12 appliances metered in the house, including tablet, dishwasher, air exhaust, fridge, entertainment, freezer, kettle, lamp, laptops, stove, TV, stereo. 11 are used in this work, as appliance 10 (stove) is not continuously measured and has some missed data during the time period we use. In addition to the metering data, moving average of the total energy consumption and the increase/decrease of the total energy consumption (between the current time step and the previous step) are used.

As shown in the previous section, information based metric is used to extract the significant feature (i.e. the link in STPN) for disaggregation. The total energy consumption (the first measured variable in the ECO data set) is always used, while the other two variables are chosen by the information based metric. The time periods for training and test are 21 days (Jan. 1–Jan. 21, 2013) and 7 days (Jan. 22–Jan.28, 2013), respectively.

Compared approaches

Factorial Hidden Markov Model: Factorial Hidden Markov Model (FHMM) [9] is an extension of Hidden Markov Models that parallelizes multiple Markov models in a distributed manner, and performs some task–related inference to arrive at predicted observation. The application of such models is done by representing each end–use as a hidden state that is modeled by multinomial distribution using $\mathbb{K}$ discrete values, and then sum each appliance meter's individual independent contribution to the expected observation (i.e., the total expected main meter value). AFAMAP [25] variant of FHMM which includes the trends in the hidden states of FHMM has also been reported to be effective in the disaggregation task. In our application of FHMM, the number of hidden states is the number of testing appliances, while $\mathbb{K} = 3$ in order to keep the computational requirements low.

Combinatorial Optimization: Combinatorial optimization (CO) [52] algorithm is a heuristic scheme that attempts to minimize the ℓ_1–norm of the total power at the mains and the sum of the power of the end–uses, given either single or multi–state formulation of the sum. The drawbacks of CO for disaggregation tasks are its sensitivity to transients and degradation with increasing number of devices or similarity in device characteristics.

Note, among many advanced NILM approaches, we choose to compare with the above methods as their software implementations are readily available in the non–intrusive load monitoring toolkit [22]. Also, for FHMM we use the exact inference [9] scheme.

Metrics for evaluating performance: Here, metrics to evaluate the disaggregation performance in different aspects are applied including root mean square error (RMSE), aggregation error (AE), normalized disaggregation error (NDE).

RMSE for ith end-use is defined as, $\sum_{t=1}^{T}\sqrt{(\hat{y}_t^i - y_t^i)^2}$, where $\hat{y}_t^i$ is the prediction for the ith end-use at tth time step, y_t^i is the ground truth for the ith end-use at tth time step.

AE for the ith end-use is defined as (SAE in [13]),

$$\text{AE} = \frac{|\sum_{t=1}^{T}\hat{y}_t^i - \sum_{t=1}^{T} y_t^i|}{\sum_{t=1}^{T} y_t^i}, \tag{11}$$

AE reflects the error of the algorithm in predict the total energy consumption of each end-use in a period of time.

NDE for ith end-use is defined as [13],

$$\text{NDE} = \frac{\sum_{t=1}^{T}(\hat{y}_t^i - y_t^i)^2}{\sum_{t=1}^{T}(y_t^i)^2}, \tag{12}$$

NDE evaluates the performance of the algorithm for predicting the energy consumption at each time step.

Accuracy is defined as [21],

$$\text{Acc} = 1 - \frac{\sum_{t=1}^{T}\sum_{i=1}^{g}|\hat{y}_t^i - y_t^i|}{2\sum_{t=1}^{T}\sum_{i=1}^{g} y_t^i}. \tag{13}$$

The accuracy metric estimates the performance of the algorithm in all of the end-uses at all time steps.

Note that, precision, recall, and F-value metrics [25] are not applied in this work, as the ground truth will be more than on/off states when dealing with continuous types of end-uses (e.g., the end-uses in the RBSAM data set in Section 4.3.

4.2.2. Disaggregation results

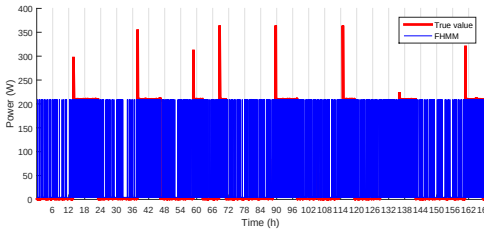
The time-series disaggregated by FHMM, CO, PFSA, and STPN are shown in Fig. 10, where appliance 5–entertainment, and appliance 11–TV are shown. FHMM and CO correctly capture the two states of the appliances, while the switching between on and off occurs all the time and causes numerous errors. PFSA can distinguish some of the states but still get a lot of switching states. STPN performs well in terms of predicting true values and avoiding erroneous switching.

The performance metrics of RMSE, AE, NDE, and accuracy of the four methods are listed in Table 2. The results show that the PFSA and STPN outperform the FHMM and CO.

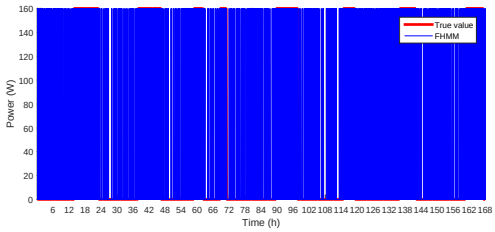
Table 2: Performance of PFSA and STPN with comparison to FHMM and CO.

ID	FHMM	CO	PFSA	STPN
Metrics	RMSE/AE/NDE	RMSE/AE/NDE	RMSE/AE/NDE	RMSE/AE/NDE
1	1.6/0.3/1.0	2.2/0.5/1.9	1.1/0.026/0.504	1.1/0.022/0.522
2	203.0/0.1/0.9	197.4/0.1/0.9	139.3/0.070/0.443	68.6/0.022/0.108
3	30.8/56.1/76.1	42.1/92.4/142.5	4.3/1.314/1.497	6.3/1.708/3.218
4	45.4/0.1/0.9	72.2/0.1/2.3	41.5/0.018/0.772	38.7/0.003/0.673
5	87.1/0.5/0.6	94.2/0.6/0.7	36.1/0.034/0.100	8.4/0.016/0.005
6	37.4/0.1/0.7	71.4/0.3/2.7	35.9/0.006/0.676	35.4/0.032/0.655
7	170.0/3.9/4.1	138.0/2.6/2.7	64.9/0.159/0.601	57.4/0.596/0.470
8	59.6/0.3/0.8	60.3/0.1/0.8	35.3/0.046/0.265	24.2/0.001/0.124
9	19.2/0.3/1.2	30.3/1.5/3.0	15.0/0.142/0.732	15.5/0.087/0.781
11	68.7/0.5/0.6	64.5/0.4/0.6	30.4/0.028/0.124	5.8/0.006/0.004
12	27.1/0.3/0.8	29.4/0.3/1.0	12.4/0.052/0.171	6.1/0.056/0.041
Acc	0.56	0.44	0.70	0.84

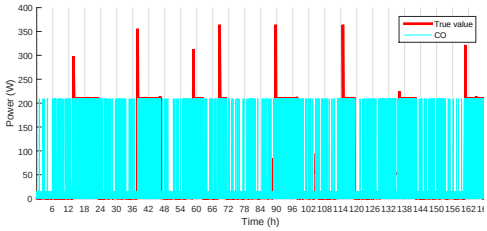
*The values of the three metrics (RMSE/AE/NDE) are listed respectively.
*For appliance 10, performance is not calculated as all values equal to zero.



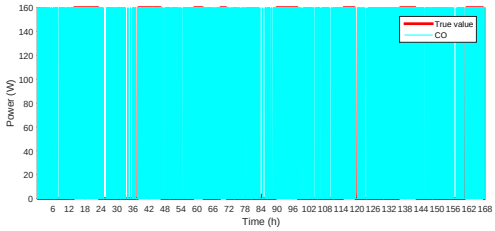
(a) FHMM on appliance 5



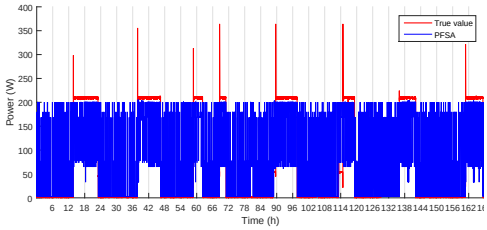
(b) FHMM on appliance 11



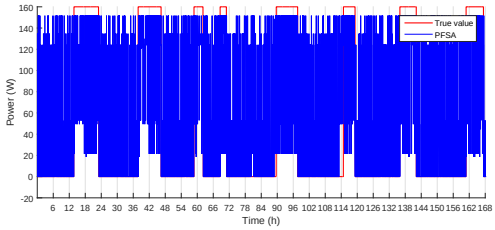
(c) CO on appliance 5



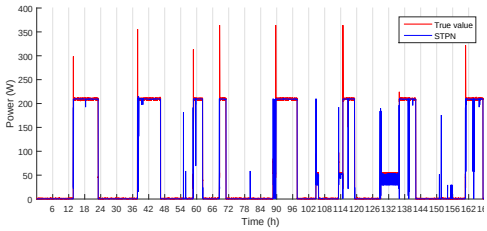
(d) CO on appliance 11



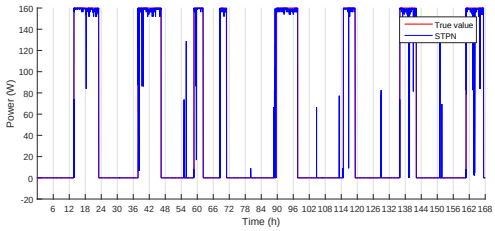
(e) PFSA on appliance 5



(f) PFSA on appliance 11



(g) STPN on appliance 5



(h) STPN on appliance 11

Figure 10: Disaggregation results using FHMM, CO, PFSA, and STPN.

4.3. Case study using mixed types of data in RBSAM data set

4.3.1. Data set and problem setup

The Residential Building Stock Assessment data set is collected based on a field study by Pacific Northwest National Laboratory (PNNL) and Northwest Energy Efficiency Alliance (NEEA) in 2013 [53, 54]. The study was conducted by leveraging existing sub-metering infrastructure in the Residential Building Stock Assessment (RBSA) owner-occupied test bed that is located in the Pacific Northwest and operated by NEEA. RBSA electric consumption data is based on field data from a representative random sample of existing homes, which encompasses 28-months of 15-minute observations within single-family homes in the Pacific Northwest United States [55].

In addition to whole building electricity use, there are typically 25 sub-metered loads per home including various types of heating, ventilating and air conditioning (HVAC) systems, appliances, lighting, entertainment, home office and plug loads. The Northwest had no precedence for a residential field study of this size and nature of the RBSA, and it was thus a new standard for residential characterization studies in the Northwest. The 2009 International Energy Conservation Code classifies RBSA metered homes in IECC Climate Zones 4, 5, and 6.

The RBSA sub-metered data reflects diversity across homes, appliances, occupancy patterns, hour of day, day of week, seasons, holidays, shopping cycles, home chore cycles, vacations, etc. Vigilance around data hygiene is critical during extraction, transformation and loading of data. Some data are out of range (positive and negative) and others are missing; data issues bring into question completeness of acquisition, accuracy of processing, and the possibility that an appliance was set-back or turned off for hours, days, weeks or months.

To form the multivariate time-series information for the disaggregation, the available data are added including outdoor temperature (ODT), indoor temperature (IDT), and the nearest weather station temperature (WST). In addition to the measured variables, moving average of the whole building electric (WBE) and time of day (ToD) are also used. The multivariate time-series are listed in Table 3. With the measured variables, four types of end-uses are formed among 25 kinds of appliances: (i) Appliances (APPL), including domestic hot water heaters, (ii)HVAC, all types of heating, ventilating and air Conditioning systems, (iii) LIGHTS, various but not all interior and exterior lighting, and (iv) MELS, miscellaneous electrical loads (were not

measured directly), they have been calculated as MELS = WBE - APPL - HVAC - LIGHTS.

Table 3: Multivariate time-series information for RBSAM dataset.

ID	Variable	Abbr.
1	Whole building electricity use	WBE
2	Outdoor temperature	ODT
3	Indoor temperature	IDT
4	Weather station temperature (the nearest)	WST
5	Moving average of WBE	MVG
6	Time of Day	ToD

During the training, the sub-metered data for each end-use is applied to learn the transition probabilities between the multivariate time-series data. The information based metric is used to select the candidates for disaggregation. The WBE is always used, while the others are chosen by the information based metric. The time periods for training and test are 21 days (e.g. Jun. 1–Jun. 21, 2012, Mar. 1–Mar. 21, 2013) and 7 days (e.g. Jun. 22–Jun. 28, 2012, Mar. 22–Mar. 28, 2013), respectively.
The approaches compared in this section are the same as those in Section 4.2. Also, the same set of metrics (RMSE, AE, NDE and accuracy) are used here.

4.3.2. Selection of informative variables for disaggregation

The information based metric (mutual information) is presented to select the most useful variables for disaggregation. A case study is shown here to analyze the relationship between mutual information metric and the disaggregation performance (RMSE for an end-use is used here). The metered and sub-metered data in 12 months of a home is used. For each month, two variables (WBE plus another one –ODT, IDT, WST, MVG, or ToD) are applied to disaggregate the end-use (in this case, APPL is shown as an example) with the same training and testing scheme. The RMSE and mutual information are plotted in Fig. 11, where the Pearson correlation coefficient between them is -0.88 and p-value is 1.8×10^{-20}. It is therefore concluded that the mutual information metric is negative relative to the disaggregation error (RMSE), which means that applying the variable with higher mutual information of the pattern (from the measured variable to the end-use) will achieve better performance in disaggregation. Therefore, mutual information

metric can be used to select the next variable for disaggregation to minimize the error and maximize the performance. With this setup, the question can be answered that which are the most valuable measurements to take.

It should be noted that the selection of next variable for disaggregation here is based on a specific end-use, which means different sets of variables can be used for disaggregating different end-uses. And it shows another advantage of using the STPN scheme in Fig. 4 (the right panel), that independent variable set can be used for disaggregation each end-use.

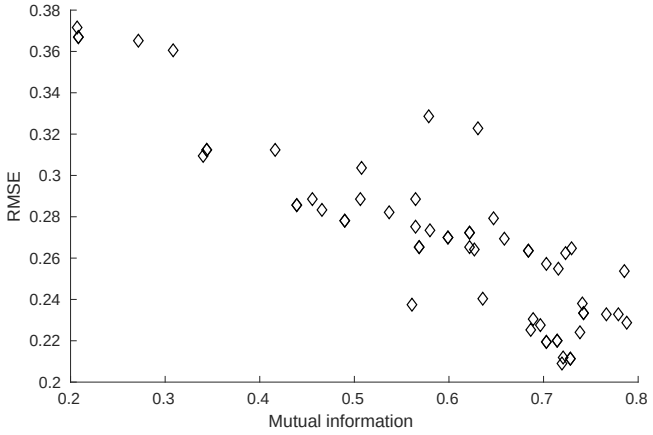


Figure 11: Relationship between mutual information of the pattern (WBE+?→end-use) and the disaggregation performance. Each point represents a testing case using WBE plus another variable (ODT, IDT, WST, MVG, or ToD)) to predict the end-use (APPL is shown here). The same training and testing scheme is applied in all 12 months' data of the same home.

4.3.3. Disaggregation performance with the number of used variables in STPN scheme

With mutual information metric to select the next measurement for disaggregation, comparisons are carried out between the performance and the number of the variables. The results are shown in Fig. 12, where the HVAC is disaggregated during Mar. 2013. The used variables and the disaggregation performance are listed in Table 4.

It can be seen that the disaggregation error decreases with the increasing number variables used. Note that, MVG and ToD are the variables generated without the requirement of additional sensors. In this context, the proposed STPN framework provides us the privilege of improving disaggregation without the requirement of new measurement.

Note that the available data in the applications are different from case to case. The proposed method is flexible and presents a scheme of utilizing multiple information sources, and the information based metric can be applied to choose the subset of the available information sources to optimize the disaggregation performance.

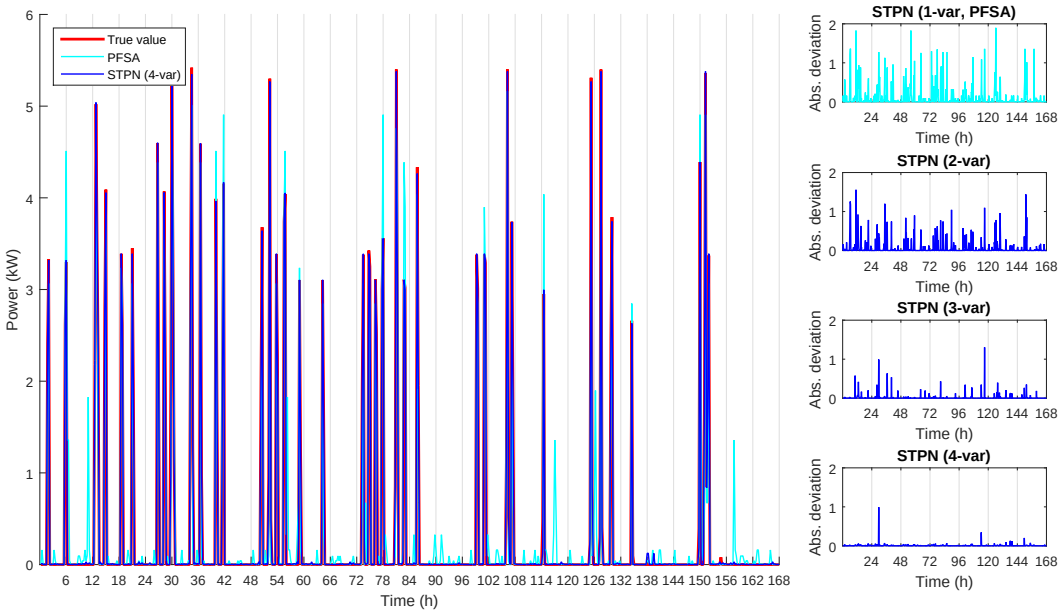


Figure 12: The disaggregation performance increases with the number of used variables. The HVAC disaggregation of Mar. 22-28, 2013 is shown on the left panel with two methods (PFSA–only WBE is applied, and STPN(4-var)–four variables are used). The absolute errors of the four approaches are shown in the right panel where the used variables increase from one to four.

Table 4: Performance of STPN with increasing number of variables used.

Method	Variables	RMSE	AE	NDE
PFSA	WBE	0.2971	0.0370	0.0691
STPN (2-var)	WBE, MVG	0.2139	0.0015	0.0358
STPN (3-var)	WBE, MVG, IDT	0.0896	0.0018	0.0063
STPN (4-var)	WBE, MVG, IDT, ToD	0.0441	0.0017	0.0015

4.3.4. Comparisons to FHMM and CO

The comparisons of the STPN to FHMM and CO are illustrated using the data in Jun. 2013. The disaggregation results and ground truth are shown in Figs. 13, 14. The STPN results are obtained using 3 variables.

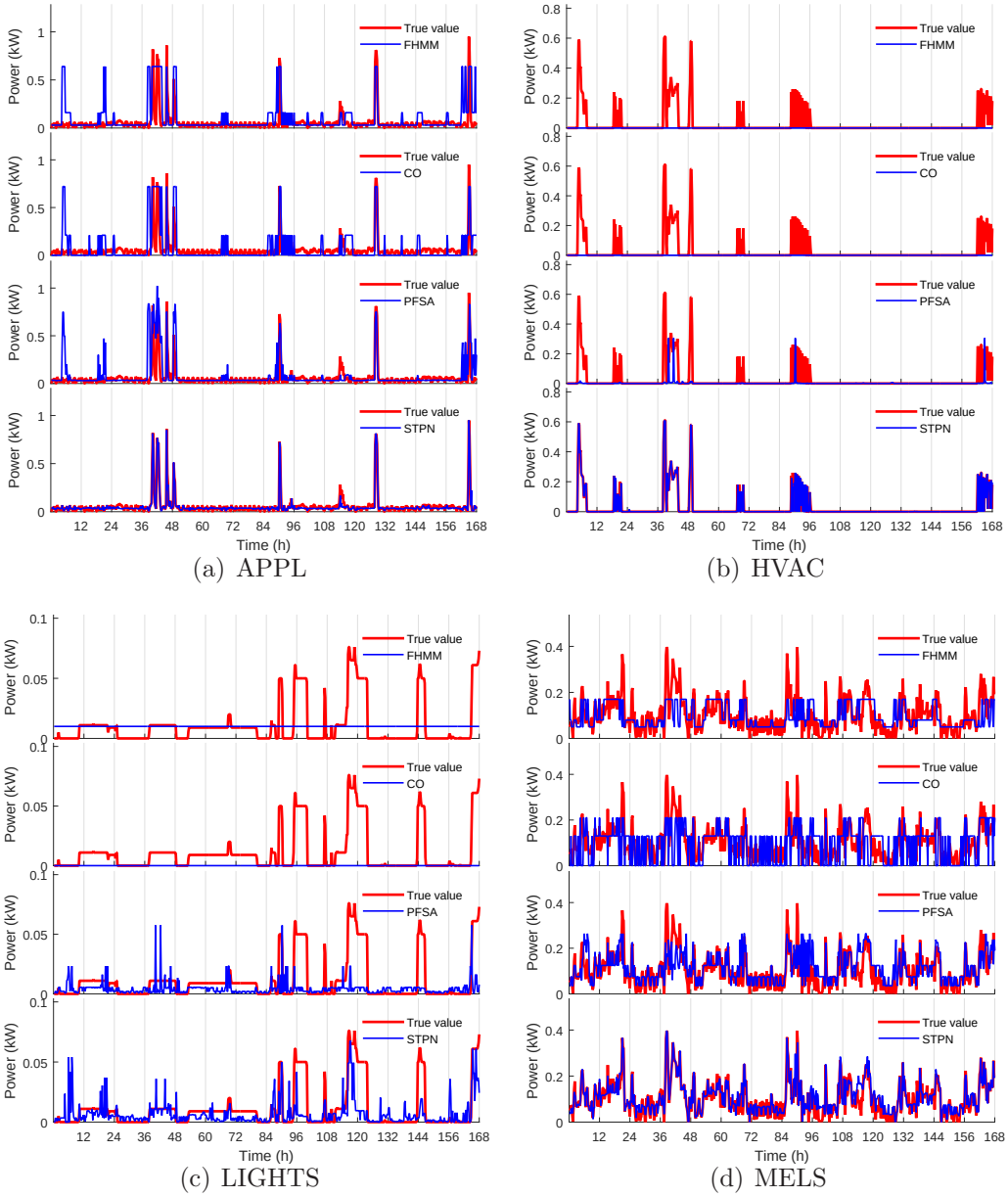
It should be noted that, as NILMTK only applies three states for most of the test cases here, the actual disaggregation results don't reflect the true values of the end-uses well. However, the four types of end-uses are combinations of several appliances, the power usage is more than on/off status. This implies that STPN should get better performance in terms of using more states.

FHMM and CO capture the profiles of energy consumption in some appliances (e.g., Fig. 14 (a)–APPL and (b)–HVAC). However, there are considerable errors in the disaggregation, even only looking at the on/off status. From the plots of PFSA (e.g., the third panels of (a) and (d) in Fig. 13), PFSA can capture the on/off status, and the actual values are finer than FHMM and CO, although there are mis-disaggregation in quite a few cases. When applying more variables by STPN, the on/off status is well captured as well as the actual values. For the performance, the results of RMSE, AE, NDE, and accuracy are listed in Table 5. For home 1, PFSA is better than FHMM and CO, while FHMM gets higher accuracy in home 2 than PFSA and CO. In both homes, STPN outperforms FHMM and CO.

Table 5: Performance of PFSA and STPN with comparison to FHMM and CO.

ID	FHMM	CO	PFSA	STPN
Metrics	RMSE/AE/NDE	RMSE/AE/NDE	RMSE/AE/NDE	RMSE/AE/NDE
1/APPL	0.14/0.58/1.17	0.13/0.15/1.04	0.13/0.41/0.96	0.02/0.05/0.03
1/HVAC	0.10/1.00/1.00	0.10/1.00/1.00	0.10/0.92/0.95	0.01/0.01/0.01
1/LIGHTS	0.02/0.06/0.73	0.02/1.00/1.00	0.02/0.60/0.83	0.02/0.37/0.54
1/MELS	0.06/0.07/0.26	0.08/0.09/0.37	0.06/0.11/0.20	0.03/0.06/0.05
1/Acc	0.63	0.58	0.68	0.90
2/APPL	0.16/0.04/0.46	0.24/0.04/0.99	0.35/1.19/2.11	0.04/0.03/0.03
2/HVAC	0.18/0.13/0.19	0.24/0.08/0.36	0.29/0.63/0.50	0.03/0.01/0.01
2/LIGHTS	0.06/0.55/1.31	0.10/2.17/3.79	0.06/0.75/1.14	0.03/0.06/0.39
2/MELS	0.13/0.12/0.13	0.23/0.58/0.42	0.11/0.13/0.09	0.04/0.004/0.01
2/Acc	0.81	0.60	0.68	0.95

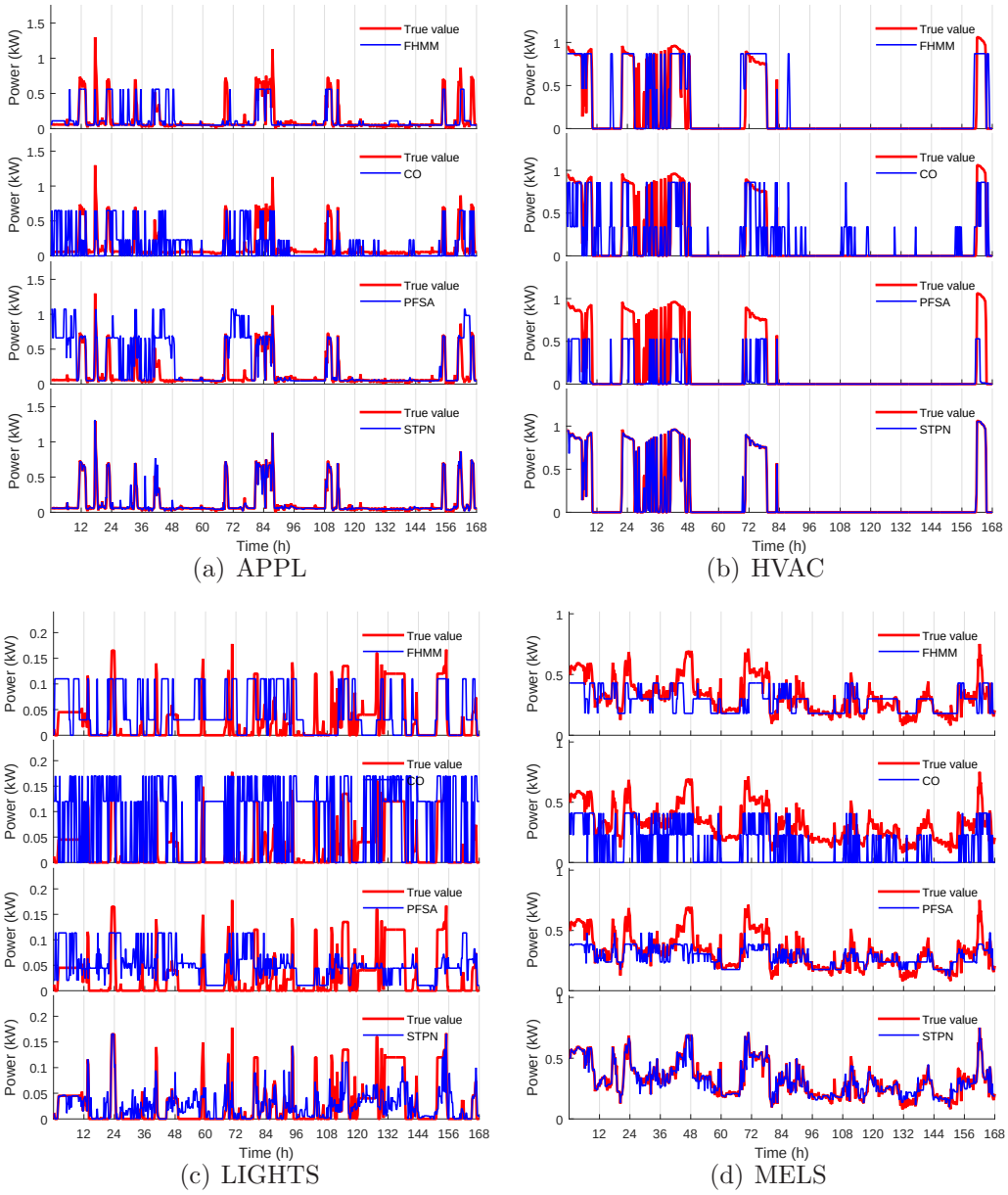
*The values of the three metrics (RMSE/AE/NDE) are listed respectively.



(a) APPL (b) HVAC

(c) LIGHTS (d) MELS

Figure 13: Disaggregation results using FHMM, CO, PFSA, and STPN using data in Jun. 2013 at home 1.



(a) APPL (b) HVAC

(c) LIGHTS (d) MELS

Figure 14: Disaggregation results using FHMM, CO, PFSA, and STPN using data in Jun. 2013 at home 2.

4.4. Out-of-sample case: transferring the trained model to another home

The out-of-sample case study is explained here to show the model transferability of the STPN framework. The relationship between the similarity metric and the disaggregation performance is first evaluated. For an STPN model M_i trained using a home i, applying the model to all of the other homes (60 homes in RBSAM data set here), the similarity between home j and home i is computed based on WBE, ODT, IDT, WST, and MVG, and the accuracy is also obtained and shown in Fig. 15. Note that, we build the similarity metric only based on the available data from all of the homes, no sub-metered data is considered here.

Fig. 15 shows that high accuracy is obtained when applying the trained model to a home which is similar to the model. The Pearson correlation coefficient between the similarity and the accuracy is 0.74 and p-value is 8.4×10^{-12}. It is therefore concluded that the similarity is positive relative to the disaggregation accuracy.

A further test case is formed via estimating the loss of accuracy when applying a model trained for home j to disaggregate energy in home i. The results are shown in Fig. 16, where the red line is the accuracy using a model trained for the same home (i.e., using the sub-metered data of home i), and the blue dashed line are the results with a model trained for a home j (the home j is selected based on the similarity metric). It can be seen that while the accuracy decreases when using a model trained for another home, the loss in accuracy remains within an acceptable range in most cases. Then, the loss of accuracy is shown in Fig. 17, where the x-axis is the percentage population of homes, and the y-axis is the loss of accuracy between the red line and blue line in Fig. 16. It shows that around 54% homes incur less than 5% loss of accuracy while 85% of all homes incur less than 10% loss of accuracy. Therefore, the similarity metric has the potential to cluster the homes into several groups such that each group has comparable end-uses and can be disaggregated by one model with reasonable accuracy.

It should be noted that only WBE, ODT, IDT, WST, and MVG are considered in the similarity metric. While other information (e.g., occupants, appliances, and structure details) might also be quite useful to define the similarity of homes in terms of end-uses. Although such information was not available for this study, the formulation of STPN provides the flexibility to accommodate different types of data [39, 43], and further analysis can be carried out by considering such additional information.

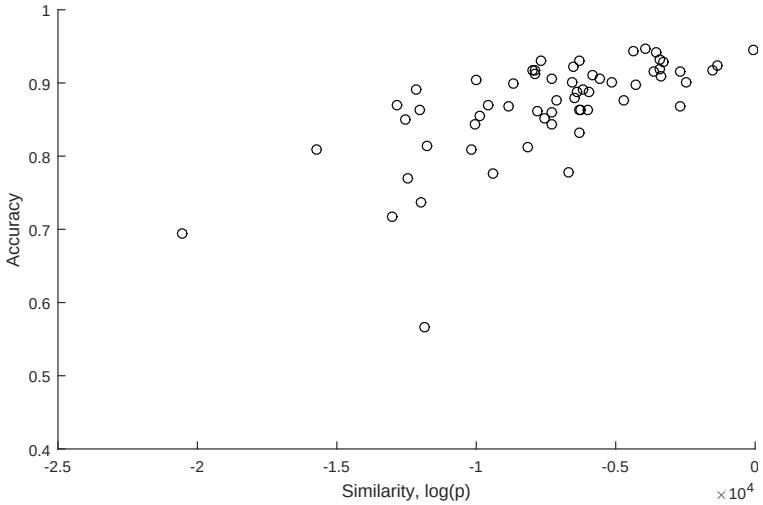


Figure 15: Relationship between similarity of metered data (at the aggregate side) and the disaggregation performance. The similarity is computed between the home (with submetering ground truth) and the other 60 homes using the data at the aggregate side. The accuracy is based on Eq. 13.

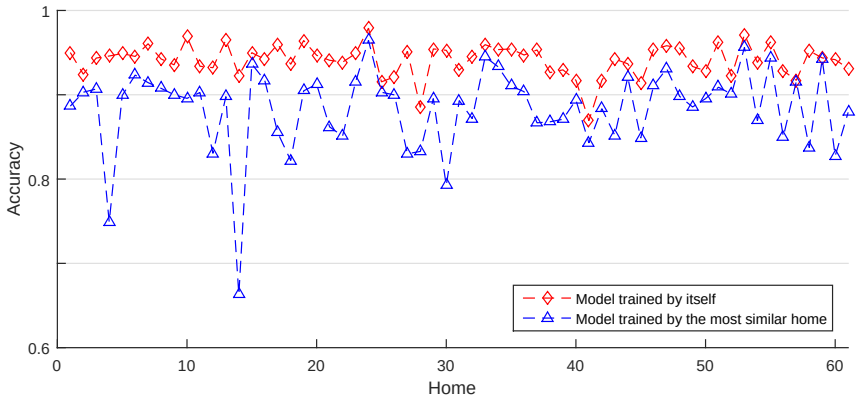


Figure 16: Performance of out-of-sample cases in all 61 homes. The results in red are the disaggregation accuracy using the model trained by itself (with the sub-metered data of this home), and the dash line show the accuracy using a selected model with highest similarity (no need of sub-metered data of this home).

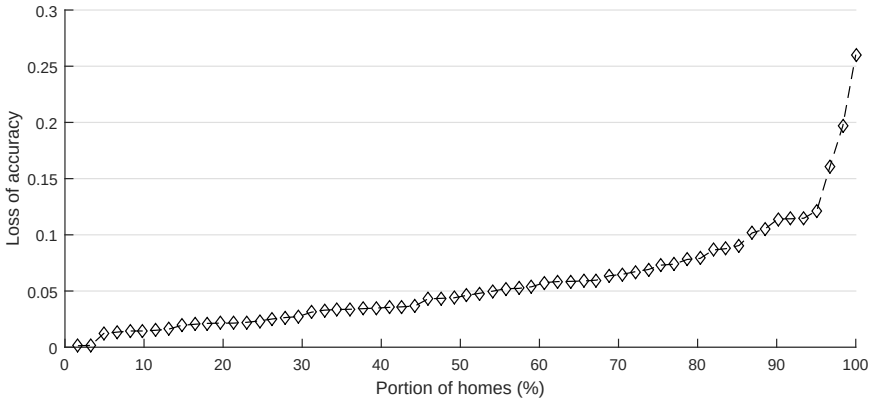


Figure 17: Loss of accuracy in out-of-sample cases. The loss of accuracy is defined as the accuracy loss between the disaggregation model trained by the sub-metered data of this home and the model selected via similarity metric. The x-axis shows the portion of homes (in 61 homes) that having the loss of accuracy less than the corresponding value in y-axis.

4.5. Discussion

With a multivariate view of the NILM problem, this work proposed a novel STPN framework to implement energy disaggregation using multivariate time-series data. The presented scheme is validated using three data sets (DMU, ECO, and RBSAM). The results show that the STPN framework is capable of harnessing rich multivariate time-series data in terms of (i) discovering specific patterns of energy usage/generation in phasor measurements, (ii) disaggregating electricity end-uses with high accuracy using multivariate measurement at the aggregate level (in the ECO data set), and (iii) energy disaggregation with whole building electric (WBE) consumption and other available information (e.g., indoor/outdoor temperature, time of day, and move average of WBE).

The case study of RBSAM data set shows that, although multivariate electrical measurements at the transformer may not always be available, there are other ways to obtain multiple data sources for disaggregation, such as weather station data (important and useful for renewable energy generation (solar, wind, etc.), temperature relative to A/C usage), time of day, house area, occupancy, etc. The framework proposed in this work is successfully applied to this kind of data, and the results show that STPN outperforms incumbent FHMM and CO methods.

The out-of-sample case shows that the proposed similarity metric is effective in comparing the end-use behavior through aggregate (whole building

electric) measurements, allowing us to learn supervised models for energy disaggregation in a few homes and to deploy these models to other previously unseen homes, while retaining high levels of disaggregation performance. Therefore, when applying the NILM technique in a new residential area, significant scalability can be achieved as the proposed framework will not require supervised training and expensive submetering for every home.

5. Conclusions

This work presented a spatiotemporal pattern network (STPN) framework to utilize multivariate time-series data for non-intrusive load monitoring (NILM). The proposed STPN framework is capable of (i) using diverse types of data, (ii) discovering specific patterns of energy usage/generation in phasor measurements, (iii) disaggregating electricital end-uses with high accuracy using multivariate measurement at the aggregate level (in the ECO data set), and (iv) energy disaggregation using whole building electric (WBE) consumption and other available concurrent information (e.g., indoor/outdoor temperature, time of day, and move average of WBE). The out-of-sample disaggregation shows that the proposed similarity metric is effective in comparing the end-use behavior through the aggregate measurements, allowing us to learn supervised models for energy disaggregation in a few homes and deploy these models to other previously unseen homes, while retaining high degree of disaggregation accuracy. Although STPN is able to achieve increasing accuracy with addition of multivariate data, the formulation of the atomic and relational pattern has assumed a depth, $D = 1$ (the memory parameter) for feature extraction. It is conjectured that better accuracy would be achieved if a learning algorithm is able to adaptively determine the signal memory required for feature extraction. Also, this work has not included household behavioral pattern as one of the variables in the use patterns.

Future work will pursue: (i) gathering more building and appliance information (RBSAM data set) to enhance the similarity confidence between homes, (ii) collection of ground truth data at the home/transformer level (DMU data set) for scalability analysis of the proposed approach, and (iii) cluster analysis in a residential area with the purpose of minimizing the installment cost and maximizing the disaggregation performance.

Acknowledgement

This work was supported by the National Science Foundation under Grant No. CNS-1464279. The authors would like to express their sincere gratitude to Dr. Soumalya Sarkar for the support on the mutual information based metric.

References

[1] A. Cominola, M. Giuliani, D. Piga, A. Castelletti, A. Rizzoli, A hybrid signature-based iterative disaggregation algorithm for non-intrusive load monitoring, Applied Energy 185 (2017) 331–344.

[2] R. K. Jain, K. M. Smith, P. J. Culligan, J. E. Taylor, Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance accuracy, Applied Energy 123 (2014) 168–178.

[3] N. Bassamzadeh, R. Ghanem, Multiscale stochastic prediction of electricity demand in smart grids using bayesian networks, Applied Energy 193 (2017) 369–380.

[4] S. Garshasbi, J. Kurnitski, Y. Mohammadi, A hybrid genetic algorithm and monte carlo simulation approach to predict hourly energy consumption and generation by a cluster of net zero energy buildings, Applied Energy 179 (2016) 626–637.

[5] M. Zeifman, K. Roth, Non-intrusive appliance load monitoring (NIALM): Review and outlook, 2011, pp. 1–27.

[6] G. Hart, Nonintrusive appliance load monitoring, Proceedings of IEEE 80 (12).

[7] M. Baranski, J. Voss, Genetic algorithm for pattern detection in nialm systems, in: Systems, man and cybernetics, 2004 ieee international conference on, Vol. 4, IEEE, 2004, pp. 3462–3468.

[8] J. Liang, S. K. Ng, G. Kendall, J. W. M. Cheng, Load signature study - part i: Basic concept, structure, and methodology, IEEE Transactions on Power Delivery 25 (2) (2010) 551–553.

[9] Z. Ghahramani, M. I. Jordan, Factorial hidden markov models, Kluwer Academic Publishers, Boston, MA (1997).

[10] S. Makonin, F. Popowich, I. V. Bajić, B. Gill, L. Bartram, Exploiting HMM sparsity to perform online real-time nonintrusive load monitoring, IEEE Transactions on Smart Grid 7 (6) (2016) 2575–2585.

[11] H. Kim, M. Marwah, M. Arlitt, G. Lyon, J. Han, Unsupervised disaggregation of low frequency power measurements, in: Proceedings of the 2011 SIAM International Conference on Data Mining, SIAM, 2011, pp. 747–758.

[12] M. Z. A. Bhotto, S. Makonin, I. Bajic, Load disaggregation based on aided linear integer programming, IEEE Transactions on Circuits and Systems II: Express Briefs.

[13] M. Zhong, N. Goddard, C. Sutton, Latent bayesian melding for integrating individual and population models, in: C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, R. Garnett (Eds.), Advances in Neural Information Processing Systems 28, Curran Associates, Inc., 2015, pp. 3618–3626.

[14] G. C. Koutitas, L. Tassiulas, Low cost disaggregation of smart meter sensor data, IEEE Sensors Jounal 16 (6) (2015) 1665–1675.

[15] J. Kelly, W. Knottenbelt, Neural nilm: Deep neural networks applied to energy disaggregation, Proceedings of 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environment - Buildsys'15 (2015) 55–64.

[16] H. Shao, M. Marwah, N. Ramakrishnan, A temporal motif mining approcah to unsupervised energy disaggregation applications to residential and commercial buildings, 1st International Workshop on Non-Intrusive Load Monitoring (2012) 1 – 7.

[17] S. Giri, M. Berges, An energy estimation framework for event-based methods in non-intrusive load monitoring, Elsevier Journal of Energy Conversion and Management 90 (2015) 488 – 498.

[18] S. M. Tabatabaei, S. Dick, W. Xu, Toward non-intrusive load monitoring via multi-label classification, IEEE Transactions on Smart Grid 8 (1) (2017) 26–40.

[19] J. M. Gillis, S. M. Alshareef, W. G. Morsi, Nonintrusive load monitoring using wavelet design and machine learning, IEEE Transactions on Smart Grid 7 (1) (2016) 320–328.

[20] C. Dinesh, R. I. Godaliyadda, M. P. B. Ekanayake, J. Ekanayake, P. Perera, Non-intrusive load monitoring based on low frequency active power measurements, AIMS Energy 4 (3) (2016) 414–443.

[21] J. Z. Kolter, M. J. Johnson, Redd: A public data set for energy disaggregation research, in: Workshop on Data Mining Applications in Sustainability (SIGKDD), San Diego, CA, Vol. 25, 2011, pp. 59–62.

[22] N. Batra, J. Kelly, O. Parson, H. Dutta, W. Knottenbelt, A. Rogers, A. Singh, M. Srivastava, Nilmtk: An open source toolkit for non-intrusive load monitoring, 5th International Conference on Future Energy Systems(ACM e-Energy), Cambridge UK (2014) 1 –14.

[23] K. Anderson, A. Ocneanu, D. Benitez, D. Carlson, A. Rowe, M. Berges, Blued: A fully labeled public dataset for event-based non-intrusive load monitoring research, In Proceedings of 2nd KDD Workshop on Data Mining Applications in sustainability, Beijing, China (2012) 12 –16.

[24] S. Barker, A. Mishra, D. Irwin, E. Cecchet, P.Shenoy, J. Albrecht, Smart*: An open data set and tools for enabling research in sustainable homes, In Proceedings of 2nd KDD Workshop on Data Mining Applications in sustainability, Beijing, China.

[25] J. Z. Kolter, T. Jaakola, Approximate inference in additive factorial hidden markov models with application in energy disaggregation, In Proceedings of the International Conference on Artificial Intelligence and Statistics, La Palma, Canary Islands (2012) 1472 – 1482.

[26] M. J. Johnson, A. S. Willsky, Bayesian nonparametric hidden semi-markov models, Journal of Machine Learning Research 14 (Feb) (2013) 673–701.

[27] O. Parson, S. Ghosh, M. Weal, A. Rogers, Non-intrusive load monitoring using prior models of general appliance types, In Proceedings of the 26th AAAI Conference on Artificial Intelligence , Toronto, ON, Canada (2012) 356 – 362.

[28] Z. Wang, G. Zheng, Residential appliances identification and monitoring by a nonintrusive method, IEEE transactions on Smart Grid 3 (1) (2012) 80–92.

[29] Y.-H. Lin, M.-S. Tsai, An advanced home energy management system facilitated by nonintrusive load monitoring with automated multiobjective power scheduling, IEEE Transactions on Smart Grid 6 (4) (2015) 1839–1851.

[30] T. Hassan, F. Javed, N. Arshad, An empirical investigation of vi trajectory based load signatures for non-intrusive load monitoring, IEEE Transactions on Smart Grid 5 (2) (2014) 870–878.

[31] T. M. Lawrence, M.-C. Boudreau, L. Helsen, G. Henze, J. Mohammadpour, D. Noonan, D. Patteeuw, S. Pless, R. T. Watson, Ten questions concerning integrating smart buildings into the smart grid, Building and Environment 108 (2016) 273–283.

[32] H. Wang, P. Xu, X. Lu, D. Yuan, Methodology of comprehensive building energy performance diagnosis for large commercial buildings at multiple levels, Applied Energy 169 (2016) 14–27.

[33] R. Enríquez, M. Jiménez, M. Heras, Towards non-intrusive thermal load monitoring of buildings: Bes calibration, Applied Energy 191 (2017) 44–54.

[34] L. Stankovic, V. Stankovic, J. Liao, C. Wilson, Measuring the energy intensity of domestic activities from smart meter data, Applied Energy 183 (2016) 1565–1580.

[35] D. Murray, J. Liao, L. Stankovic, V. Stankovic, Understanding usage patterns of electric kettle and energy saving potential, Applied Energy 171 (2016) 231–242.

[36] J. Bank, B. Kroposki, Development of a real-time, high-speed distribution level data acquisition system, in: 2012 IEEE PES Innovative Smart Grid Technologies (ISGT), IEEE, 2012, pp. 1–6.

[37] C. Rao, A. Ray, S. Sarkar, M. Yasar, Review and comparative evaluation of symbolic dynamic filtering for detection of anomaly patterns, Signal, Image and Video Processing 3 (2) (2009) 101–114.

[38] S. Sarkar, S. Sarkar, N. Virani, A. Ray, M. Yasar, Sensor fusion for fault detection and classification in distributed physical processes, Frontiers in Robotics and AI 1 (2014) 16.

[39] C. Liu, S. Ghosal, Z. Jiang, S. Sarkar, An unsupervised spatiotemporal graphical modeling approach to anomaly detection in distributed cps, in: Proceedings of the International Conference of Cyber-Physical Systems, (Vienna, Austria), 2016.

[40] O. Parson, S. Ghosh, M. Weal, A. Rogers, An unsupervised training method for non-intrusive appliance load monitoring, Elsevier Journal of Artificial Intelligence 217 (2014) 1–19.

[41] Z. Jiang, C. Liu, A. Akintayo, G. Henze, S. Sarkar, Energy prediction using spatiotemporal pattern networks, Applied Energy 206 (2017) 1022–1039.

[42] C. Liu, S. Ghosal, Z. Jiang, S. Sarkar, An unsupervised anomaly detection approach using energy-based spatiotemporal graphical modeling, Cyber-Physical Systems.

[43] C. Liu, Y. Gong, S. Laflamme, B. Phares, S. Sarkar, Bridge damage detection using spatiotemporal patterns extracted from dense sensor network, Measurement Science and Technology 28 (1) (2017) 014011.

[44] A. Akintayo, S. Sarkar, Hierarchical symbolic dynamic filtering of streaming non-stationary time series data, arXiv preprint arXiv:1702.01811.

[45] P. F. McNutt, J. Hambrick, M. Keesee, D. Brown, Impact of SolarSmart Subdivisions on SMUD's Distribution System, National Renewable Energy Laboratory, 2009.

[46] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, Image quality assessment: from error visibility to structural similarity, Image Processing, IEEE Transactions on 13 (4) (2004) 600–612.

[47] C. Liu, D. Jiang, W. Yang, Global geometric similarity scheme for feature selection in fault diagnosis, Expert Systems with Applications 41 (8) (2014) 3585–3595.

[48] Time and Date AS, https://www.timeanddate.com/sun/usa/sacramento (2016).

[49] US climate data, Climate sacramento - california, http://www.usclimatedata.com/climate/sacramento/california/united-states/usca0967 (2016).

[50] C. Beckel, W. Kleiminger, R. Cicchetti, T. Staake, S. Santini, The ECO data set and the performance of non-intrusive load monitoring algorithms, in: Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings, ACM, 2014, pp. 80–89.

[51] W. Kleiminger, C. Beckel, S. Santini, Household occupancy monitoring using electricity meters, in: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, ACM, 2015, pp. 975–986.

[52] J. C. William, H. C. William, R. P. William, A. SCHRIJVER, Combinatorial Optimization, 5th Edition, Vol. 21, Springer, 2011.

[53] D. Baylon, P. Storm, K. Geraghty, B. Davis, Residential building stock assessment: single-family characteristics and energy use, Northwest Energy Efficiency Alliance.

[54] E. Mayhorn, G. Sullivan, R. Butner, H. Hao, M. Baechler, Characteristics and performance of existing load disaggregation technologies, PNNL-24230.

[55] N. E. E. Alliance, Residential building stock assessment: Metering study, Tech. rep., Report (2014).