



UNIVERSITY OF LEEDS

This is a repository copy of *Temporal Self-Regulation of Energy Demand*.

White Rose Research Online URL for this paper:

<http://eprints.whiterose.ac.uk/157104/>

Version: Accepted Version

Article:

Nambi, SNAU, Pournaras, E and Venkatesha Prasad, R (2016) Temporal Self-Regulation of Energy Demand. *IEEE Transactions on Industrial Informatics*, 12 (3). pp. 1196-1205. ISSN 1551-3203

<https://doi.org/10.1109/tii.2016.2554519>

© 2016, IEEE. This is an author produced version of a paper published in *IEEE Transactions on Industrial Informatics*. Uploaded in accordance with the publisher's self-archiving policy.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Temporal Self-regulation of Energy Demand

Akshay Uttama Nambi S. N., Evangelos Pournaras and Venkatesha Prasad R

Abstract—The increase in deployment of smart meters has enabled collection of fine-grained energy consumption data at consumer premises. Analysis of this real-time energy consumption data bestows new opportunities for better demand-response (DR) programs. This work offers a new perspective to study energy demand and helps in designing novel mechanisms for decentralized demand-side management. Specifically, a new concept of finding the demand states using energy consumption of consumers over time and, feasible transitions therein, are introduced. It is shown that the orchestration of temporal transitions between the demand states can meet broad range of Smart Grid objectives. An online demand regulation model is developed that captures the temporal dynamics of energy demand to identify target consumers for different DR programs. This methodology is empirically evaluated and validated using data from more than 4000 households, which were part of a real-world Smart Grid project. This work is the first one to comprehensively analyze the temporal dynamics of demands.

Index Terms—Temporal analysis, demand regulation, data-driven model, Smart Grid

I. INTRODUCTION

THE introduction of smart meters in large-scale offers new opportunities for fine-grained real-time data collection. This enables new demand-response (DR) programs wherein consumers can self-regulate their energy demands with minimum interventions from utility companies [1]. Data collected can be used to understand consumption behavior and adjust the demands to decrease energy cost, facilitate the use of renewable energy resources, or prevent black-outs [2]–[4].

This work offers a new perspective to study energy demand enabling the design of novel mechanisms for decentralized demand-side energy management. Rather than only optimizing the demand levels of each household so that it meets available supply, the concept of computing the demand states of each household and feasible transitions between these states are introduced. The demand states measure one of the following demand features: (i) demand level, (ii) demand variation and (iii) demand peaks. In contrast to the related work [2]–[4], it is shown that the orchestration of temporal transitions between the demand states can meet a broad range of Smart Grid objectives set by the utility companies. A *generalized data-driven* methodology based on clustering of historic consumption data (time-series) from each household is designed for a local computation of the demand states at different aggregation

Manuscript received Month Date, Year; revised Month Date, 2016.

Copyright 2016 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org

Akshay Uttama Nambi S. N and Venkatesha Prasad are with the Embedded Software Group, EEMCS, Delft University of Technology, The Netherlands (email: Akshay.Narashiman@tudelft.nl, R.R.VenkateshaPrasad@tudelft.nl).

Evangelos Pournaras is with Professorship of Computational Social Science, ETH Zurich, Zurich, Switzerland (email: epournaras@ethz.ch).

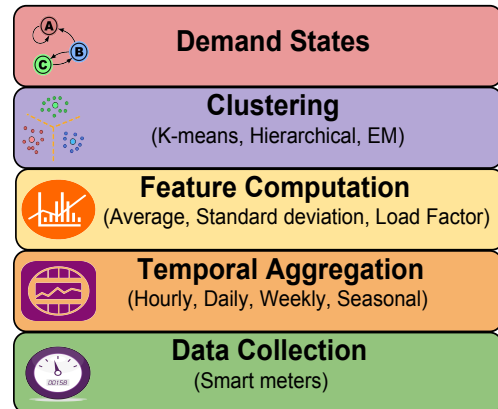


Fig. 1: Computation of demand states.

granularity, e.g., daily, weekly, etc. This methodology can capture the temporal dynamics of demand and can be used to identify target consumers for DR programs. The proposed methodology is decentralized, highly scalable and privacy-preserving. This can be used to build effective real-time recommendations for the self-regulation of demand. The data-driven methodology is generic, domain-independent and can be applied to a broad range of time series data. This can be further applied to other time series data, especially resource consumption data such as water and gas. This methodology is evaluated and validated using data from a real-world Smart Grid project consisting of more than 4000 households.

The main contributions of this work are the following: (i) a generalized, domain-independent data-driven model and methodology for the computation of demand states; (ii) four metrics for measuring and evaluating demand adjustments; (iii) evaluation of the methodology using demand data from a real-world Smart Grid project and quantitative comparison with related work; and (iv) an online self-regulation model for the adjustment of demands by targeted consumers and its validation using survey responses of consumers.

This paper is organized as follows: Temporal dynamics of demand is modeled in Section II. Four metrics for measuring and evaluating demand adjustments is introduced in Section III. An online demand regulation model is illustrated in Section IV. Experimental evaluation of the proposed models is given in Section V. Comparison of the proposed model with related work is in Section VI and finally, conclusion and vistas for future explorations are outlined in Section VII.

II. MODELING TEMPORAL DYNAMICS OF DEMAND

A generalized data-driven model and methodology (illustrated in Fig. 1) for computing local demand adjustments for each household is introduced in this section. An outline

TABLE I: List of mathematical symbols used.

Symbols	Description
l	Number of historic demand time series
\mathbf{d}_j	A demand time series at time j
T	Number of measurements of a demand series
j	Time index, e.g., j th aggregation time period
$\hat{\mathbf{d}}_{l+1}$	Forecasted demand series
$\tilde{\mathbf{d}}_{l+1}$	Regulated demand series
\mathbf{q}_j	Quality of a demand series \mathbf{d}_j
u	Quality feature e.g., AVG, RSD, LF
m	Number of quality features
p_j^u	Demand property of feature u at time j
\mathbf{c}_u^o	Cluster o for feature u
k	Number of clusters computed
s_j^u	Demand state of a feature u at time j
\hat{s}_{l+1}^u	Forecasted demand state for feature u
\tilde{s}_{l+1}^u	Regulated demand state for feature u
\mathbf{a}_j^u	Demand adjustment for feature u at time j
z	Number of state transitions
$\mathbf{T}_{a \rightarrow b}^u$	Transition probability from demand state a to b for feature u
θ	Transition probability threshold
\mathbf{M}_o^u	Temporal membership of a demand state o for feature u
$\mathbf{A}(\mathbf{a}_z^u)$	Temporal adaptability on a demand adjustment \mathbf{a}_z^u
$S_{x,y}^u$	Temporal similarity between two consumers x and y for a demand feature u
β, γ, σ	Binary variables to compute temporal metrics
J	Jaccard similarity coefficient

of mathematical symbols used in this article is given in Table I. Each household is assumed to be equipped with an information system that collects and stores real-time demand measurements using smart meters [8]. The collected data are aggregated at different granularity levels – daily, weekly or seasonal. The information system manages l samples of historic demand series $\mathbf{d}_1, \dots, \mathbf{d}_l$, with \mathbf{d}_1 being the most recent historic time series and \mathbf{d}_l is the earliest. Each demand series consists of $T = |\mathbf{d}_j|$ demand measurements, therefore, T is the number of measurements aggregated for a certain granularity.

The information system serves the DR program of the utility companies by turning the *forecasted* demand series $\hat{\mathbf{d}}_{l+1}$ to the *regulated* demand series $\tilde{\mathbf{d}}_{l+1}$. Such an adjustment is achieved by mining the historic demand series to infer and reason about possible demand changes observed in each household. Utility companies may introduce one or more features for characterizing and assessing the quality of the forecasted and regulated demand. The quality of the demand represents the characteristics extracted from the demand time series. For example, a household demand with low load factor shows occasional high demand peaks resulting in low quality of demand.

The quality \mathbf{q}_j of a demand series \mathbf{d}_j is defined by a set of m measurable features $\mathbf{q}_j = \{p_j^1, \dots, p_j^m\}$, where p_j^u is the property of a demand series \mathbf{d}_j according to the feature u at time j . A property p_j^u is defined as $p_j^u = f_u(\mathbf{d}_j)$, where $f_u(\mathbf{d}_j)$ is a function performed over the demand time series.

This paper focuses on $m = 3$ quality features of demand: (i) *average* (AVG), (ii) *relative standard deviation* (RSD) and (iii) *load factor* (LF).

The average (AVG) feature is defined as,

$$p_j^1 = f_1(\mathbf{d}_j) = \frac{1}{T} \sum_{t=1}^T d_j^t, \quad (1)$$

where, $d_j^t \in \mathbf{d}_j$ is the demand measured at time t within the demand time series \mathbf{d}_j . This feature indicates the aggregate demand over the time period T and does not provide information about how demand is distributed over T . In contrast, relative standard deviation (RSD) feature computes the homogeneity of demand over the time period T and is defined as,

$$p_j^2 = f_2(\mathbf{d}_j) = \frac{1}{p_j^1} \sqrt{\frac{1}{T} \sum_{t=1}^T (d_j^t - p_j^1)^2}, \quad (2)$$

where, $d_j^t \in \mathbf{d}_j$ is the demand measured at time t within the demand time series \mathbf{d}_j . Note that the average demand over the time period T in demand time series \mathbf{d}_j is indicated in (1) by the property p_j^1 . Finally, the load factor (LF) [5] determines the scale of demand peaks and is computed by the ratio of average demand and maximum demand measured over a time period T and is defined as,

$$p_j^3 = f_3(\mathbf{d}_j) = \frac{p_j^1}{\max \mathbf{d}_j}, \quad (3)$$

where, the property p_j^1 denotes the average demand over the time period T . The $\max \mathbf{d}_j = \max(d_j^t), \forall t \in \{1, \dots, T\}$ denotes the maximal element that corresponds to the maximum demand peak during the time period T for the demand time series \mathbf{d}_j .

Demand, and its quality features, can be forecasted by analyzing the historic demand time series. For example, the average demand p_{l+1}^1 at time period $l + 1$ can be predicted by using the average demand p_1^1, \dots, p_l^1 during the past l time periods. Although a broad range of data mining and machine learning algorithms can be used for predicting future demands, the main focus here is on clustering because of the following reasons: (i) clustering is an unsupervised method that does not require labeling of the demand data; (ii) future demand predictions can be determined by analyzing the centroids of the clusters and their corresponding sizes [6]; (iii) the possible states, in which a feature of demand may be, can be extracted via clustering. For example, by clustering the past average demand p_1^1, \dots, p_l^1 into *three* clusters, the centers of the clusters ranked from low to high indicate the low, medium and high demand states of a household; and (iv) clustering provides information about the temporal transitions between different demand states that represent the center of the clusters. In this way, the temporal dynamics of demand are modeled, since clustering reasons about whether or when certain demand transitions are feasible by each household.

Given l demand properties p_1^u, \dots, p_l^u of a feature u , clustering to k clusters is defined as,

$$\bigcup_{o=1}^k \mathbf{c}_u^o = p_1^u, \dots, p_l^u, \quad (4)$$

where, \mathbf{c}_u^o is the cluster o containing demand properties for the feature u . For each cluster \mathbf{c}_u^o , the center c_u^o is computed by the centroid or medoid [9]. *Expectation Maximization* (EM) clustering [10] is employed here to determine the number of clusters based on the demand properties.

When a demand property changes its membership from one cluster to another, this is defined as a *transition*. A *demand*

state $s_j^u = o \in \{1, \dots, k\}$ is defined by the cluster index to which the demand property p_j^u belongs. States s_{l+1}^u and \hat{s}_{l+1}^u represent forecasted and regulated demand states, respectively for a feature u . A sequence of z transitions defines a *demand adjustment* observed or triggered at time j and is given by,

$$\mathbf{a}_j^u = \{s_j^u, \dots, s_{j+z}^u\}, \quad (5)$$

where, \mathbf{a}_j^u is a sequence of transitions starting from state s_j^u of feature u at time point j to state s_{j+z}^u with $z = |\mathbf{a}_j^u|$.

III. MEASURING DEMAND ADJUSTMENT

This section defines the following four metrics to measure and evaluate demand adjustments, *viz.*, (i) *transition probability*, (ii) *temporal membership*, (iii) *temporal adaptability* and (iv) *temporal similarity*.

A. Transition probability

It measures the probability of moving from a certain demand state to another demand state. Given a quality feature u , the average transition probability $\mathbb{T}_{a \rightarrow b}^u$ from demand state a to b is defined as,

$$\mathbb{T}_{a \rightarrow b}^u = \frac{1}{l-1} \sum_{j=1}^{l-1} \beta_j; \quad \beta_j = \begin{cases} 1 & \text{if } s_{j+1}^u = b \mid s_j^u = a \\ 0 & \text{if } s_{j+1}^u \neq b \mid s_j^u = a \end{cases} \quad (6)$$

where, β_j is a binary variable that equals '1' if a transition from demand state a to b occurs at time j or '0' otherwise. It holds that $\mathbb{T}_{a \rightarrow b}^u \in [0, 1]$.

B. Temporal membership

This metric evaluates the probability of a certain demand state occurring over time. The temporal membership M_o^u of a demand state o for feature u is defined as,

$$M_o^u = \frac{1}{l} \sum_{j=1}^l \gamma_j; \quad \gamma_j = \begin{cases} 1, & s_j^u = o \\ 0, & s_j^u \neq o \end{cases} \quad (7)$$

where, γ_j is a binary variable that equals '1' if the demand state s_j^u occurs at time j or '0' otherwise.

C. Temporal adaptability

This metric measures the probability of a demand adjustment occurring over time. Temporal adaptability $A(\mathbf{a}_{l+1}^u)$ of a demand adjustment \mathbf{a}_{l+1}^u for feature u and with size $|\mathbf{a}_{l+1}^u| = z \leq l$ is defined as,

$$A(\mathbf{a}_{l+1}^u) = \frac{1}{l-z+1} \sum_{j=1}^{l-z+1} \sigma_j; \quad \sigma_j = \begin{cases} 1, & \mathbf{a}_j^u = \mathbf{a}_{l+1}^u \\ 0, & \mathbf{a}_j^u \neq \mathbf{a}_{l+1}^u \end{cases} \quad (8)$$

where, σ_j is a binary variable that equals '1' if the demand adjustment \mathbf{a}_{l+1}^u defines the same sequence of transitions as the sequence of the demand states s_j^u, \dots, s_{j+z}^u . Otherwise it holds $\sigma_j = 0$.

D. Temporal Similarity

This metric evaluates the similarity between the demand states of two consumers. Temporal similarity $S_{x,y}^u$ between the demand states of consumer x and consumer y for a feature u is defined by the Euclidean distance as,

$$S_{x,y}^u = \sqrt{\sum_{j=1}^l (s_{j,x}^u - s_{j,y}^u)^2} \quad (9)$$

where, $s_{j,x}^u, s_{j,y}^u$ represent the demand states of two households x and y , respectively.

IV. ONLINE SELF-REGULATION OF DEMAND

A model that improves the quality of demand by a transition from the forecasted state s_{l+1}^u to the regulated state \hat{s}_{l+1}^u is introduced in this section. Demand quality is improved by adjusting one of the demand properties (see Section II), e.g., performing a transition to a demand state with reduced demand, lower variation in demand or lower demand peaks. A heuristic is presented to select consumers who can perform such a transition. The heuristic employs the temporal adaptability metric to quantify the probability of each consumer to perform such a transition. The criterion for selection of target consumers is governed by the *threshold* θ . For example, if a consumer has $\theta = 0.2$ and $A(\mathbf{a}_{l+1}^u) = 0.25 > \theta$, the model reasons that this consumer can self-regulate its demand, i.e., it can perform the change to regulated state using the forecasted state. Otherwise, if $A(\mathbf{a}_{l+1}^u) < \theta$ the consumer remains in the forecasted state. This threshold can be selected by the utility companies, each consumer or it can even be the result of a negotiation between the two parties. For example, utility companies can provide monetary incentives to consumers for lower values of θ so that they increase the likelihood of participation in DR programs in case of a high overload in the power grid.

Algorithm 1 A heuristic for online self-regulation of demand.

Input: Demand properties p_1^u, \dots, p_l^u , the forecasted state s_{l+1}^u and the threshold θ .

Training phase:

1. Compute the demand states by clustering p_1^u, \dots, p_l^u as in (4).
2. Compute the transition probability $\mathbb{T}_{a \rightarrow b}^u$ for all possible transitions.
3. Compute the transitions from step 2 that satisfy the DR objective.
4. Compute the regulated state \hat{s}_{l+1}^u from the transitions of step 3 with maximum $\mathbb{T}_{a \rightarrow b}^u < \theta$.

Testing phase:

if $\mathbf{a}_l^u = \{s_l^u, s_{l+1}^u\}$ satisfies the DR objective **then**

5. No demand regulation is required.

else

6. Change from forecasted state s_{l+1}^u to the regulated \hat{s}_{l+1}^u .
7. Compute efficiency: AVG reduction, RSD reduction or increase in LF.

end if

Algorithm 1 illustrates the local heuristic that realizes the online self-regulation model. The heuristic is executed by each

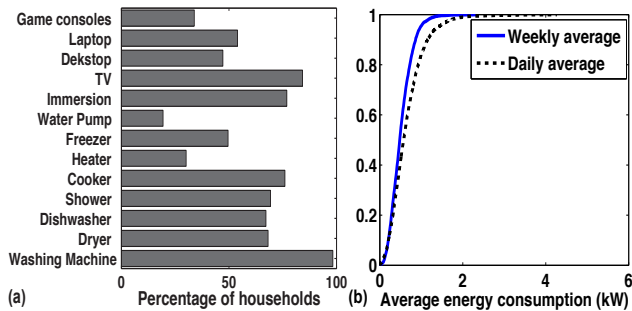


Fig. 2: Dataset characteristics.

household. It consists of a training and testing phase. In the training phase, all possible demand adjustments that satisfy the DR objectives are computed and ranked according to the transition probability metric. The training phase completes with the computation of the regulated state, in case the constraint for a maximum $T_{a \rightarrow b}^u < \theta$ is satisfied. The testing phase checks if the adjustment from the current demand state s_t^u to the forecasted demand state s_{t+1}^u satisfies the DR objective. If the objective is satisfied, no regulation is required otherwise the forecasted state is adjusted to the regulated state. Each household is assumed to be equipped with an information system that can translate the forecasted demand state to the regulated demand state [11]. Based on this adjustment, the efficiency of the heuristic can be computed by measuring the AVG reduction, RSD reduction or LF increase, depending on the selected quality feature.

The self-regulation model is *online* and the training model proposed is adaptive, wherein the temporal metrics are updated after each time period. To regulate the demand, households have to only identify the current and forecasted demand states at each time period.

V. EXPERIMENTAL EVALUATION

This section illustrates the experimental evaluation by employing a dataset [12] of 4,232 residential households to identify target consumers for the DR programs. The performance of the proposed online self-regulation of demand is evaluated empirically. Furthermore, the proposed methodology can be applied to any Smart Grid dataset without modifying the algorithm. The experimental evaluation is repeated with the REFIT dataset [13] confirming the findings illustrated in this paper and the results are available in [14].

A. Real-world smart meter data

CER dataset [12] collected during a smart metering trial in Ireland is used for empirical evaluation. The dataset contains energy consumption measurements from 4,232 households every 30 minutes between July 2009 and December 2010 (75 weeks in total). The objective of the trial was to investigate the effect of feedback on household electricity consumption. Each participating household fills out a questionnaire before and after the trial. The questionnaire contains questions about

the socio-economic status of the residential consumer, appliance stock, properties of the dwelling, and the consumption behavior of the occupants. Fig. 2(a) shows the distribution of loads across households in the dataset. The x-axis represents the percentage of the households having an appliance. Furthermore, Fig. 2(b) shows the distribution of daily and weekly average energy consumption across all households.

B. Cluster computation and evaluation

Expectation Maximization (EM) clustering [10], [15] is employed to determine number of clusters based on the demand properties. One of the major limitations with clustering algorithms such as *k*-means clustering is its requirement of prior knowledge on the number of clusters, *k*. EM clustering iteratively refines an initial clustering model to fit the data based on the principle of maximum likelihood estimation.

The number of clusters found for all the households is 7 and 5 for daily and weekly AVG features, respectively. Similarly, 5 and 4 clusters are found for the RSD feature and, 5 and 5 clusters are found for the LF feature with daily and weekly granularity, respectively. Members of Cluster 1, for the AVG feature indicate households with low average energy consumption. Similarly, members of Cluster 1 for the RSD feature indicate households with low demand variation and members of Cluster 5 for the LF feature indicate households with low demand peaks.

The number of clusters computed with the unsupervised EM approach is validated with two well-known cluster evaluation metrics [16]: Davies-Bouldin Index (DBI) and Silhouette. The cluster evaluation metrics verify the number of clusters and confidence of EM method. More details on the cluster evaluation metrics can be found in [14].

Summary: *Clustering identifies the demand state of the households. Cluster evaluation metrics such as DBI and silhouette verify the accuracy of cluster formation.*

C. Temporal dynamics of demand

Fig. 3 shows the average transition probability of all households for the AVG, RSD, and LF features. The higher the gradient, the higher is the probability of transition from one demand state to another. Households in a certain demand state have higher probability to remain in the same state than transiting to other demand states. This can be seen in Fig. 3(a) and 3(b), where a household has a high probability to remain in the same demand state, indicating a constant average demand. However, for the RSD and LF features the transitions are more rapid indicating the variations in demand and sudden peaks, respectively. The transition probability from a high RSD state to a low RSD state is low, indicating not so drastic variation in the demand as seen in Fig. 3(c) and 3(d). Hence, DR programs employed by utilities should consider step-wise reduction matching the variations instead of immediate reduction in demand variation. Fig. 3(e) and 3(f) show the transition probabilities for the LF feature. The households change their load factor quite often as depicted by the transitions in low demand states.

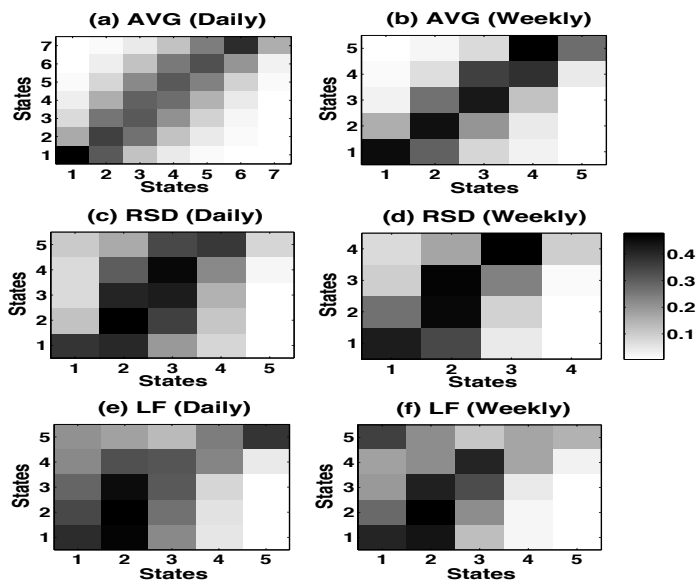


Fig. 3: Transition probabilities for the quality features – AVG, RSD and LF.

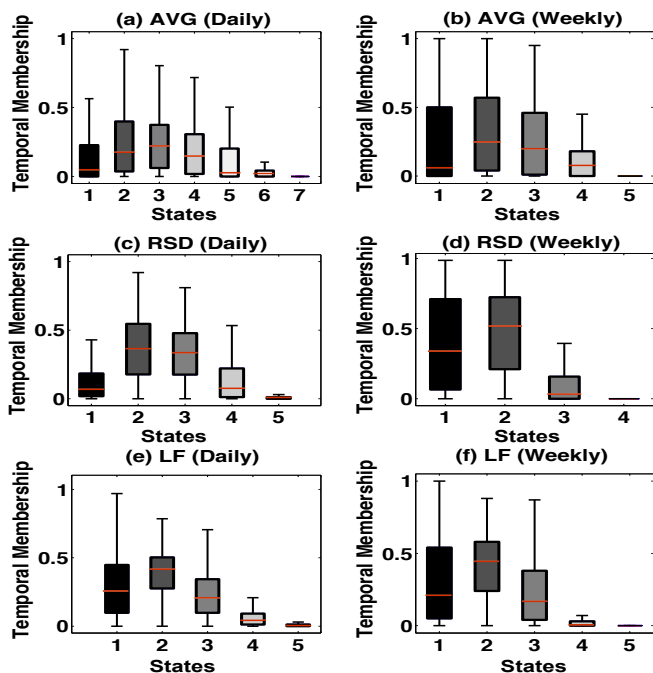


Fig. 4: Temporal membership for all households.

Summary: The transition probability illustrates the temporal adjustments of demand. The results show that for the AVG feature transitions are more fixed than the ones of RSD and LF features, where households change their states frequently.

Fig. 4 shows the average temporal membership of all households. The box plots describe the distribution of households for each demand state membership. The lowest line segment indicates the minimum temporal membership value of a household and the top line segment indicates the maximum temporal membership value of a household. The rectangular box indicates the distribution of temporal membership values

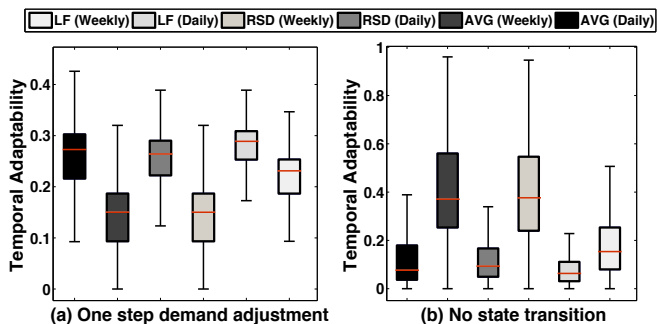


Fig. 5: Temporal adaptability for all households.

for different households with the red line segment indicating the median. The majority of the households belong to the intermediate demand states (States 2, 3 and 4) as seen in Fig. 4(a) and 4(b) for the AVG daily and weekly properties. Indeed, less than 10% of the households belong to low and high AVG state. Temporal membership reveals the most favorable demand state of a household. Utilities can use this information in order to provide tailored recommendations.

Fig. 4(c) shows that around 40% of households have high membership probability in state 2 and 3 indicating the majority of the households having moderate demand variations. However, Fig. 4(d) shows that around 80% of households have high membership probability in demand state 1 and 2 indicating a low variation in weekly demand. Thus, weekly demand variation of households is more stable compared to the daily variation, which increases the membership probability associated with the weekly properties. Hence, varying the granularity level provides insights on how household demand properties change over time. Fig. 4(e) and 4(f) show the membership of households for the LF feature. The majority of the households are distributed over low demand states, indicating high demand peaks.

Summary: Temporal membership reveals the most probable demand state of a household. With respect to the AVG feature, only 10% of the households belongs to low demand states indicating that the majority of the households are either moderate or high energy consumers.

Fig. 5(a) show the average temporal adaptability of all households for different quality features with transitions that aim to reduce average energy demand, demand variation and demand peaks. This work considers, (i) one step demand adjustment – transition from one state to another (consecutive or non-consecutive states); (ii) two step demand adjustment – two consecutive transitions from one state to another; and (iii) no transition – self-transitions to the same demand state. An adjustment from a high demand state to a low demand state for the AVG and RSD features indicates the reduction in average demand and variation (e.g., transitions from State 5 to 1 (one step) or State 5 to 3 and then to 1 (two step)). Similarly for the LF feature, demand adjustments from a low LF state to a high LF state indicates reduction in demand peaks. Fig. 5(a) shows around 30% of the households can reduce AVG daily demand with one step demand adjustment.

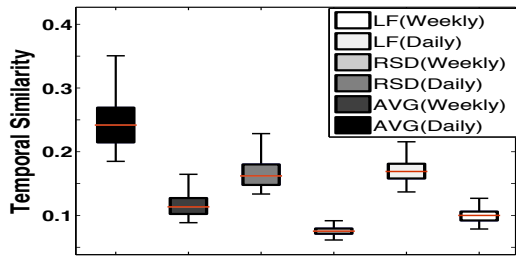


Fig. 6: Average temporal similarity for all households.

For the RSD and LF features, 30% of the households have transitions that can result in reduction of demand variation and demand peaks. The total number of households adaptable for weekly granularity is around 15% for all the quality features. This observation is due to stabilization of demand properties over a week. Fig. 5(b) shows the number of households having state transitions to the same state (for example, transition from State 2 to itself). The households containing no/self-transition, indicate the consumers who are not adaptable towards demand regulations. Hence, utilities can use temporal adaptability to identify households that can participate in the DR programs.

Summary: Temporal adaptability identifies households that are potential target consumers for the DR programs. The results show that around 30% and 15% of the households can participate in the DR programs for daily and weekly granularity, respectively.

Fig. 6 show the average temporal similarity of all households for the AVG, RSD and LF features. Around 25% and 10% of the households have similar demand state transitions for the AVG feature with daily and weekly granularity respectively. Similarly, for the RSD and LF features around 16% and 18% of the households have same transitions for the daily demand. DR programs can use temporal similarity to determine potential households, which have similar demand variation for peak reduction and peak shifting.

Summary: The results show that, around 25%, 16% and 18% of households have similar demand state transitions among the 4,232 households for daily AVG, RSD and LF features, respectively.

D. Online self-regulation of demand

The online self-regulation model considers over a year of energy consumption data for the training phase. Since the proposed model is adaptive and online, the duration of training data can be varied. Fig. 7 shows the demand regulation for each quality feature with both daily and weekly demand properties. The x-axis represents the threshold value θ indicating the probability of having a demand adjustment that satisfies the DR objective. The y-axis indicates the demand regulation in percentage. The figure also illustrates the number of households participating in the demand regulation.

Fig. 7(a) and 7(b) show the total energy reduction by all households for daily and weekly AVG demand properties. Each day around 3000 households have demand adjustments that can support energy reduction, resulting in 33% daily average energy reduction (this corresponds to 3.5kW of power)

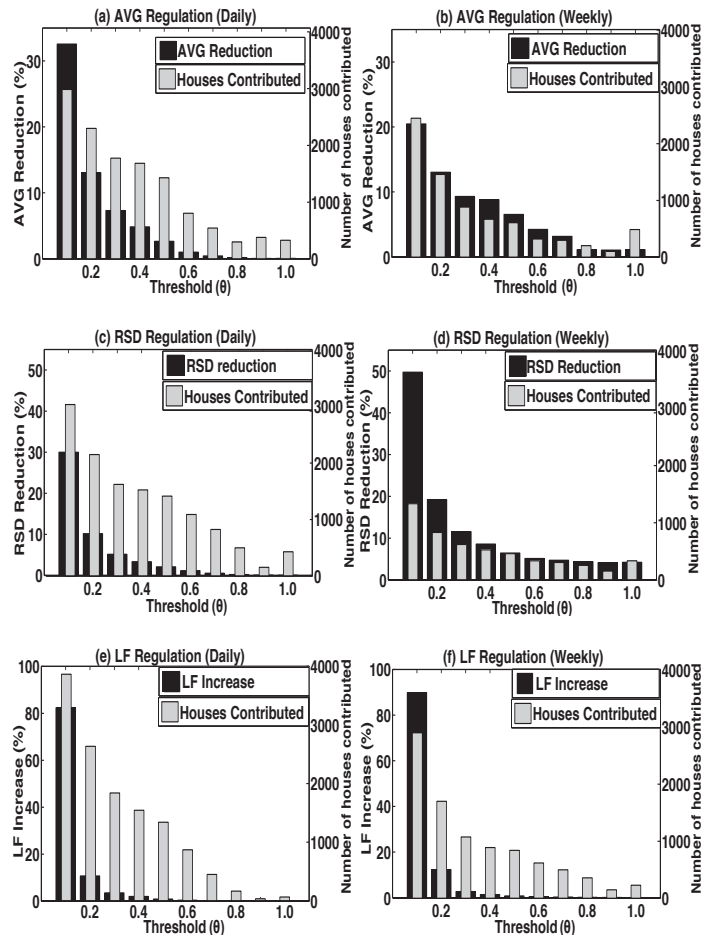


Fig. 7: AVG, RSD and LF regulation for daily and weekly demand properties.

for threshold $\theta = 0.1$. With the increase in θ , the number of households participating in demand reduction decreases. This means that not every household has a demand adjustment with high probability, which can regulate the demand. Moreover, the percentage of energy reduction decreases with the increase in θ . For example, when $\theta > 0.9$, even though around 400 households have state transitions that can regulate the demand, the average energy reduction per day is low. This is because, most of these households selected for $\theta > 0.9$ have low energy consumption. Hence regulating the demand of these households results in low demand reduction. The θ value can be used to select the households, which can participate in demand reduction. Utilities can set a low θ value during the peak period to select more households for demand regulation and a high θ value during the off-peak period. Fig. 7(b) shows the demand reduction for weekly demand properties and it follows a similar trend like daily reduction. For all θ values, reduction of 10% is achieved for daily AVG demand properties.

Fig. 7(c) and 7(d) show the RSD regulation by all households for daily and weekly demand properties. Demand variation is reduced by 30% and 50% for daily and weekly RSD feature when $\theta = 0.1$. The RSD regulation is higher for the

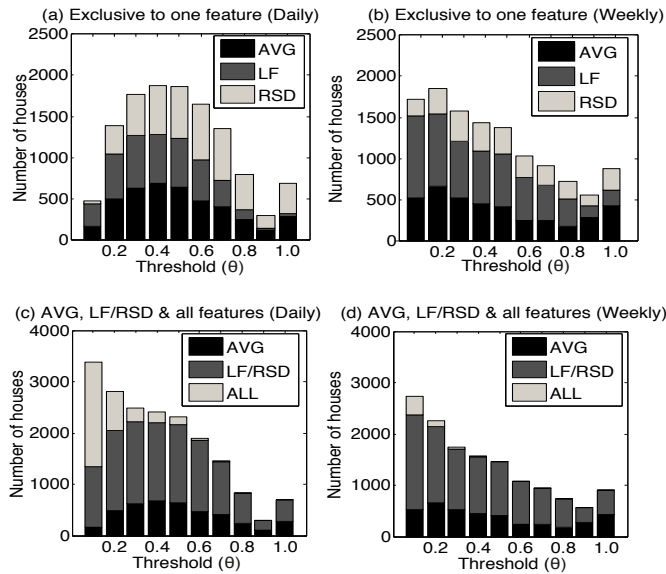


Fig. 8: Number of consumers participating in DR programs.

weekly demand than the daily demand. This indicates that households prefer to adjust their demand properties during the week as compared to the daily regulation. For all θ values, the demand variation is reduced by 15% for weekly RSD properties.

Fig. 7(e) and 7(f) show LF regulation by all households for daily and weekly demand properties. Households regulate the LF by reducing the peak demand. Load factor is increased by 80% for both daily and weekly demand properties when $\theta = 0.1$. The number of households participating in demand peak shaving gradually decreases, with the increase in θ . For all θ values, LF increase of 15% is achieved for both daily and weekly demand properties.

The results from the self-regulation model can be used to identify the households that participate in different DR programs. Furthermore, recommendations can be provided to the utilities regarding their DR programs. For example, utilities are encouraged to choose daily AVG demand properties over weekly AVG demand properties for effective demand reduction program. Similarly, for an effective reduction in demand variation, utilities need to select the weekly RSD demand properties over daily RSD demand properties. Utilities can either select daily or weekly LF demand properties for the demand peak shaving as they result in similar LF improvement.

Summary: *The online demand regulation model enables average reduction of 10% in daily average energy demand, 15% in weekly demand variation and 15% in daily demand peak shaving for all θ values.*

Fig. 8 illustrates the distribution of households participating in a DR program for all θ values. Fig. 8(a) and 8(b) show the households that participate either towards (i) reduction in demand (AVG) or (ii) reduction in demand variation (RSD) or (iii) reduction in demand peak (LF). The number of households participating towards demand reduction (AVG) for θ between 0.3 and 0.5 is comparatively higher than for other θ values. This indicates that these households have frequent

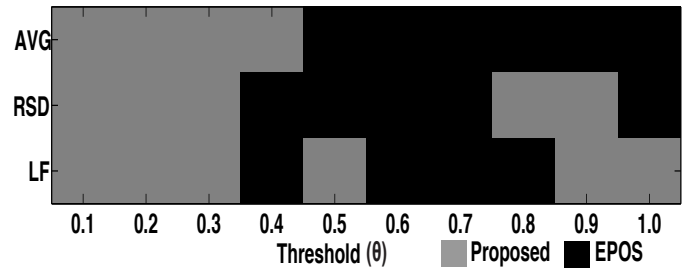


Fig. 9: Comparison of the proposed model against EPOS for θ values in the range 0 and 1.

demand adjustments that regulate the average demand. A large number of households participate in reduction of demand variation when θ is greater than 0.4, indicating that these households have frequent state transitions from low RSD demand state to high RSD demand state. In contrast, more number of households participate in demand peak shaving when θ is lower than 0.5.

Fig. 8(c) and 8(d) show the households that participate either towards (i) demand reduction (only AVG feature) or (ii) demand variation and demand peak shaving (LF and/or RSD features) or (iii) all the three DR objectives. The number of households participating to all the three features reduce as θ increases and is maximum when $\theta = 0.1$. Utilities can use these insights to choose the appropriate θ value for the selection of households towards the DR program. For example, incentives to consumers with lower values of θ can increase the likelihood of their participation in DR programs. In [12], consumers are incentivized to participate in DR program either based on (i) time of use tariffs, (ii) weekend tariffs and (iii) behavioral change in energy consumption.

Summary: *The results show that the selection of θ plays a crucial role in identifying the target consumers for the DR programs.*

Evaluating the proposed distributed methodology with other related methodologies is a challenge and requires an equivalent context, same dataset and experimental settings. However, this paper contributes a constructive empirical comparison with EPOS, the Energy Plan Overlay Self-stabilization system [4]. EPOS is a fully decentralized mechanism for planning and optimizing demand, and employs the same CER [12] dataset for its evaluation. The experimental evaluation settings of EPOS are replicated¹ and compared with the proposed model for different values of θ .

Fig. 9 shows the performance of the proposed model against EPOS for θ values in the range 0 and 1. Each colored block indicates the model with the highest performance for the corresponding θ value. The consumer associated with regulation can be managed with a relevant choice of θ . This means that the selection of this parameter is a trade-off and can make the proposed methodology perform higher or lower than other methodologies.

¹Three selection functions of EPOS viz., MIN-Demand, MIN-Relative-Deviations and MAX-Load-Factor are used for comparison. These three functions corresponds to the proposed AVG, RSD and LF regulation.

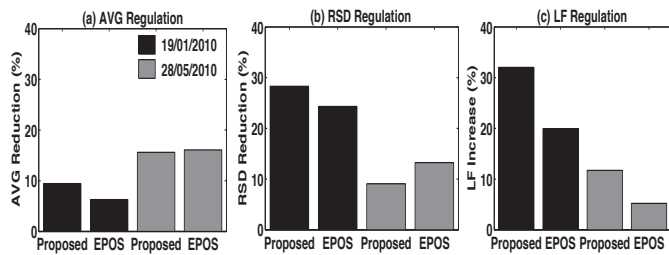


Fig. 10: Comparison of the proposed model with EPOS.

Fig. 10 shows the comparison of the proposed online model against EPOS on two specific days viz., 19/01/2010 and 28/05/2010 for all features. Threshold value of $\theta \leq 0.5$ is used to obtain the regulation results. The proposed online self-regulation model has a higher performance than EPOS across all quality features. Demand regulation can be further improved by allowing higher θ values that implies higher consumer tolerance in discomfort. EPOS studies a scenario in which all households participate in the process of demand regulation. In contrast, the online self-regulation model identify the households for DR program based on the temporal characteristics of the demand. Consequently, only households that have a valid transitions satisfying the DR objective is selected.

E. Validation with survey data

The experimental findings derived are validated with the survey data collected from the trial [12]. Each participant is asked questions regarding the collection of energy data, their attitude towards energy reduction, environment, etc. The objective of this validation is to quantify how close the data-driven analysis is to the survey data. Specifically, the survey responses of consumers are compared with the demand regulation results. The following questions are selected from the survey questionnaire:

- Q1: *I/we am/are interested in changing the way I/we use electricity if it reduces the bill.*
- Q2: *It is too inconvenient to reduce our electricity usage.*
- Q3: *I/we am/are interested in changing the way I/we use electricity if it helps the environment.*

The answers to the above question is in the range [1, 5], where 1 stands for a strong agreement and 5 stands for a strong disagreement. Questions Q1 and Q2 are used to compare the results obtained for the AVG feature and Question Q3 is used to compare results obtained for the RSD and LF features. Consumer survey responses are grouped into two categories, (i) households which agree towards reduction (survey response: 1,2,3,4) and (ii) households which do not agree towards reduction (survey response: 5). The hypothesis here is that a survey response of strong disagreement (i.e., response 5) means the consumer has no interest towards DR programs. Hence, any other response indicates the willingness towards the DR program. Grouping of consumer survey responses with different combinations is also evaluated, viz., (i) households with survey response (1,2,3) and households with survey response (4,5); and (ii) households with survey

TABLE II: Comparison of survey data with analysis result.

Questions	TP	FP	FN	TN	F1-score
Q1	0.68	0.12	0.02	0.17	0.90
Q2	0.63	0.20	0.02	0.15	0.85
Q3-RSD	0.69	0.13	0.09	0.09	0.86
Q3-LF	0.68	0.15	0.09	0.08	0.85

response (1,2) and households with survey response (3,4,5). Jaccard similarity coefficient is used to compare the results from the data analysis to the survey results. It is defined as,

$$J(R_s, R_a) = \frac{|R_s \cap R_a|}{|R_s \cup R_a|}, \quad (10)$$

where, $J(\cdot)$ is the Jaccard similarity coefficient [17], R_s and R_a are the set of households obtained based on the outcome of the survey response and data analysis respectively. The similarity coefficient is the ratio of intersection and union of these two sets and takes a value [0, 1]. The output of $J(R_s, R_a)$ indicates the percentage of households, which are found both in the survey and analysis results. To evaluate the similarity, the following statistical measures are derived:

- True Positive (TP): The number of households that are present both in survey and analysis. This is similar to Jaccard similarity coefficient.
- False Positive (FP): The number of households that are present in the survey, but are not present in the analysis.
- False Negative (FN): The number of households that are not present in the survey, but are present in the analysis.
- True Negative (TN): The number of households that are not present in both survey and analysis.
- F1-score: The measure of accuracy and is obtained by calculating the harmonic mean of precision and recall [18].

Table II shows the TP, FP, FN, TN and F1-score for all questions when compared to the data analysis results. The analysis correctly identifies 70% of the consumers in the survey data, who agree with the reduction. FP shows the percentage of consumers who responded positively towards reduction but are not found in the analysis. This observation can be explained by the fact that survey response collected is from only one occupant of a household and this opinion may be different from the other occupants in the household. Similarly, FN indicates the consumers who do not agree towards reduction but are found participating in the analysis. These households could be the potential new target consumers for the utilities. The FN in the dataset for the AVG feature is around 2% (85 households) and for the RSD/LF feature it is around 9% (380 households). Overall, the analysis results are around 87% accurate (F1-score) when compared to the survey data. Due to paucity of space, results from different groupings of consumer survey responses are not shown in detail.²

Summary: *Validation results show 70% similarity among the consumers identified in the data analysis and survey. New potential target consumers close to 10% are determined for the DR programs, which are not apparent in the survey data.*

²When survey response (1,2,3) are grouped together, the TP for Q1,Q2, Q3-RSD and Q3-LF is 0.64, 0.67, 0.69 and 0.69 respectively. Furthermore, when survey response (1,2) are grouped together, the TP for Q1 is 0.65, Q2 is 0.55, Q3-RSD is 0.65 and Q3-LF is 0.63.

VI. COMPARISON WITH RELATED WORK

Numerous DR programs [2]–[6] have been proposed to motivate changes in the consumers’ power consumption. These DR programs can be broadly classified into centralized and decentralized schemes [19]. In centralized scheme, a central controller collects all the demand information from consumers for DR decisions [2], [3]. Decentralized scheme allows consumers to coordinate directly with each other to participate in DR programs [4], [7], [11].

Current investigations on DR programs do not consider temporal resolution of energy consumption of households. The growth of Internet of Things (IoT) in the recent years has enabled not only monitoring energy consumption of appliances in a household, but also understanding the behavior of consumers vis-a-vis energy [20], [21]. Temporal analysis presented in this work is used to understand demand adjustments of households. In contrast with the state-of-the-art DR algorithms, temporal modeling and metrics proposed here can be used to reason whether or when a certain demand transition is feasible by households. An online self-regulation model for the adjustment of demands is presented for various DR objectives. The model is empirically evaluated with one of the largest publicly available dataset [12]. Furthermore, the analysis results are validated against survey data collected from more than 4000 households. The proposed methodology is highly scalable and privacy-preserving as the consumer energy information is locally stored. Table III shows the comparison of the proposed decentralized demand regulation scheme with state-of-the-art techniques. Only a few DR programs take into account preferences of consumers and often they need to be specified explicitly [3], [4]. In contrast, the proposed work derives preferences and characteristics of consumer from their energy usage over time. Majority of the literature are concerned with simulation or numerical analysis compared to the real data employed here. This is one of the first ones to comprehensively analyze the temporal dynamics of demands.

Optimization-based models are designed for DR programs with various objectives. Zhu *et al.* derive optimal power consumption, by taking into account loads that can shift or adjust their consumption in successive time periods [2]. This centralized scheme requires consumers to communicate their demand needs and usage patterns for each appliance. In contrast, the proposed analyze the temporal demand of households to derive consumer characteristics such as how often the demand pattern varies and which consumers are willing to participate in DR. Joe-wong *et al.* propose a day-ahead device-specific scheduling that is based on task schedules, which considers heterogeneity in appliance delay tolerance [3]. This centralized model employs convex optimization to derive demand schedules. However, the main problem is that it requires fine-grained appliance level energy data and also continuous real-time communication between the energy provider and the consumers. Recent work [4] shows how to manage the energy demand of households by analyzing historic aggregated energy consumption data. Pournaras *et al.* propose a decentralized approach for demand-side self-management [4], [27], where software agents represent the

TABLE III: Comparison of state-of-the-art techniques.

Work	Method	Consumer preference	Study type	Temporal analysis
[2]	Centralized	No	Simulation	No
[3]	Centralized	Partial	Simulation	No
[4]	Decentralized	Partial	Data-driven	No
[7]	Decentralized	Partial	Simulation	No
[11]	Decentralized	No	Simulation	No
Proposed	Decentralized	Yes	Data-driven	Yes

demand preferences of consumers and control their demand by selecting a plan according to the criteria defined by a selection function. The decentralized approach enforces all the consumers to select a plan that meets the DR objective set by the utility. In contrast, this work identifies the target consumers who can participate in different DR programs by analyzing the temporal dynamics of demand. Baharlouei *et al.* propose a decentralized scheme along with a fairness index to minimize total generation cost with a smart billing mechanism [7]. This approach assumes all consumers are flexible in participating towards DR. In this work, the selection of consumers and the discomfort associated with the demand regulation is governed by the threshold parameter θ . Several insights obtained from temporal analysis can be applied to develop more effective consumer-centric DR programs.

Successful implementation of DR programs rely on the identification and participation of the target consumers. The majority of previous efforts on the identification of target consumers relied on customer self-reported data [22], [23]. Large scale deployments of smart meters has paved the way to analyze real-time energy consumption to provide insights into energy usage of households [20], [24], [25]. Moss *et al.* investigate the segmentation of consumers into groups based on the similarity of energy usage [25], whereas, Chicco *et al.* study different unsupervised clustering algorithms to classify consumers, based on the load pattern shape [26]. The majority of earlier work does not study the temporal transitions for classification of households. This work analyzes the temporal dynamics of demand by considering multiple quality features such as AVG, RSD and LF.

VII. CONCLUSIONS AND FUTURE WORK

This paper concludes that a data-driven methodology for understanding and measuring the temporal dynamics of energy demand adjustments is promising. Based on this methodology, an online self-regulation model was introduced that can identify consumers who can adjust their demands to meet various DR objectives. Since the time series analysis is used, the approach could be generally applied to any application domain that deals with such data. Experimental evaluation with demand data from real-world Smart Grids shows that around 30% and 15% of the consumers can be incentivized to participate in daily and weekly DR programs. In this case, DR programs achieve 10% reduction in the average daily demand, 15% reduction in the weekly demand variations and 15% reduction in daily demand peak. The data-driven analysis was also validated with the data from the survey.

The applicability of the proposed methodology in other Smart Grid applications or even other domains is part of

the future work. A broader range of techno-socio-economic systems in which temporal dynamics play a crucial role in their regulating complexity is the future applicability of this research.

REFERENCES

- [1] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart Grid The New and Improved Power Grid: A Survey," *Communications Surveys & Tutorials, IEEE*, vol. 14, no. 4, pp. 944-980, 2012.
- [2] Z. Zhu, J. Tang, S. Lambotharan, W. H. Chin, and Z. Fan, "An integer linear programming based optimization for home demand-side management in smart grid," in *Proc. IEEE PES ISGT*, pp. 1-5 2012.
- [3] C. Joe-Wong, S. Sen, S. Ha, and M. Chiang, "Optimized day-ahead pricing for smart grids with device-specific scheduling flexibility," *IEEE J. Sel. Areas Communication*, vol. 30, no. 6, pp. 1075-1085, 2012.
- [4] E. Pournaras, M. Vasirani, R. E. Kooij, and K. Aberer, "Decentralized Planning of Energy Demand for the Management of Robustness and Discomfort," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 4, pp. 2280-2289, 2014.
- [5] R. F. Chang, and C. N. Lu, "Feeder reconfiguration for load factor improvement," *Power Engineering Society Winter Meeting, IEEE*, vol. 2, no. 7, pp. 980-984, 2002.
- [6] W. Shen, V. Babushkin, Z. Aung, and W. L. Woon, "An ensemble model for day-ahead electricity demand time series forecasting," in *Proceedings of the fourth international conference on Future energy systems, ACM e-Energy*, 2013.
- [7] Z. Baharlouei and M. Hashemi, "Efficiency-Fairness Trade-off in Privacy-Preserving Autonomous Demand Side Management," *IEEE Transactions on Smart Grid*, vol. 5, no. 2, pp. 799-808, 2014.
- [8] J. Lu, T. Sookoor, V. Srinivasan, G. Gao, B. Holben, J. Stankovic, E. Field, and K. Whitehouse, "The smart thermostat: using occupancy sensors to save energy in homes," in *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems, ACM SenSys*, pp. 211-224, 2010.
- [9] P. Berkhin, "A survey of clustering data mining techniques," *Grouping multidimensional data. Springer, Berlin Heidelberg*, pp. 25-71, 2006.
- [10] B. D. Chuong, and B. Serafim, "What is the expectation maximization algorithm?," *Nature Biotechnology*, vol. 26(8), pp. 897-899, 2008.
- [11] C. Chen, J. Wang, S. Kishore, "A Distributed Direct Load Control Approach for Large-Scale Residential Demand Response," in *IEEE Transactions on Power Systems*, vol. 29, no. 5, pp. 2219-2228, 2014.
- [12] Electricity customer behaviour trial, The commission for Energy Regulation (CER), 2012.
- [13] D. Murray, J. Liao, L. Stankovic, "A data management platform for personalised real-time energy feedback," in *Int Conf Energy Efficiency in Domestic Appliances and Lighting, EEDAL*, 2015.
- [14] S. N. A. U. Nambi, E. Pournaras and R. V. Prasad, "Clustering and Temporal Evaluation of Energy Demand," in *arXiv preprint*, 2016.
- [15] C. Frale, and A. E. Raftery, "How Many Clusters? Which Clustering Method? Answers via Model-Based Cluster Analysis," *Computer Journal*, vol. 41, pp. 578-588, 1998.
- [16] E. Rendon, I. Abundez, A. Arizmendi and E. M. Quiroz, "Internal versus external cluster validation indexes," *International Journal on Computer Communication*, vol. 5, no. 1, pp. 27-34, 2011.
- [17] P. Tan, M. Steinbach, and V. Kumar, "Introduction to Data Mining," *Addison-Wesley Longman Publishing Co., Inc.*, USA.
- [18] D. M. W. Powers, "Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation," in *Journal of Machine Learning Technologies*, pp. 37-63, 2011.
- [19] J. S. Vardakas, N. Zorba and C. V. Verikoukis, "A Survey on Demand Response Programs in Smart Grids: Pricing Methods and Optimization Algorithms," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 152-178, 2015.
- [20] S. N. A. U. Nambi, A. Reyes Lua and R. Prasad, "LocED: Location-aware energy Disaggregation Framework," in *2nd ACM International Conference on Embedded Systems For Energy-Efficient Built Environments, ACM BuildSys*, 2015.
- [21] C. Sarkar, S. N. A. U. Nambi, R. V. Prasad, A. Rahim, R. Neisse and G. Baldini, "DIAT: A Scalable Distributed Architecture for IoT," *IEEE Internet of Things Journal*, vol. 2, no. 3, pp. 230-239, 2015.
- [22] T. F. Sanquist, H. Orr, B. Shui, and C. Bittner, "Lifestyle factors in U.S. residential electricity consumption," *Energy Policy*, vol. 42, 2012.
- [23] W. Abrahamse, and L. Steg, "How do socio-demographic and psychological factors relate to households direct and indirect energy use and savings?," *Journal of Economic Psychology*, vol. 30, October 2009.
- [24] S. N. A. U. Nambi, T. G. Papaioannou, D. Chakraborty, and K. Aberer, "Sustainable energy consumption monitoring in residential settings," in *proceedings of INFOCOM, IEEE*, 2013.
- [25] S. J. Moss, "Market segmentation and energy efficiency program design," *CIEE Behavior and Energy Program*, 2008.
- [26] G. Chicco, R. Napoli, and F. Piglion, "Comparisons among clustering techniques for electricity customer classification," *IEEE Transactions on Power Systems*, vol. 21, no. 2, pp. 933-940, 2006.
- [27] E. Pournaras, M. Vasirani, R. E. Kooij, and K. Aberer, "Measuring and Controlling Unfairness in Decentralized Planning of Energy Demand," in *the proceedings of the IEEE International Energy Conference-EnergyCon 2014, Dubrovnik, Croatia, 2014*.



Akshay Uttama Nambi S. N. received the B.E. degree in computer science and engineering from Visvesvaraya Technological University, Belgaum, India, and is currently pursuing the Ph.D. degree at Delft University of Technology, Delft, The Netherlands. He is currently with the Embedded Software Group, Delft University of Technology. Prior to this, he was a visiting researcher with the Ecole Polytechnique Federale de Lausanne (EPFL), Lausanne, Switzerland, and the Indian Institute of Science (IISc), Bangalore, India.

His research interests are in the areas of data mining, machine learning, Internet of Things, cyber physical systems, smart grids and energy harvesting.



Evangelos Pournaras received the B.Sc. degree in technology education and digital systems from the University of Piraeus, Piraeus, Greece; the M.Sc. degree in internet computing from the University of Surrey, Guildford, Surrey, U.K.; and the Ph.D. degree from the Delft University of Technology, Delft, The Netherlands, in 2013.

He is currently a postdoctoral researcher in the Computational Social Science group, at ETH Zurich, Zurich, Switzerland. Evangelos has also been a visiting researcher at EPFL in Switzerland and has industry experience at IBM T.J. Watson Research Center in the USA. He serves the editorial board and the program committees of several international conferences and journals. He has several publications in high-impact journals and conferences in the area of distributed systems, including a best journal paper award. Evangelos is currently working on the Nervousnet project focusing on fully decentralized and privacy-preserving services for social sensing and mining.



Venkatesha Prasad R. received the B.E. degree in electronics and communication engineering and M.Tech. degree in industrial electronics from the University of Mysore, Mysore, India, in 1991 and 1994, respectively, and the Ph.D. degree from the Indian Institute of Science, Bangalore, India, in 2003. He has been a Consultant with the ERNET Laboratory, Indian Institute of Science (IISc). From 1999 to 2003, he was also a Consultant with CEDT, IISc, involved with VoIP application developments. In 2003, he was a Team Leader with Esquebe Communication Solutions Pvt Ltd., Bangalore, India, where he was involved in the development of various real-time networking applications.

From 2005 to 2012, he was a Senior Researcher, and since 2012, he has been an Assistant Professor with the Delft University of Technology, Delft, The Netherlands. The work at TUDelft has resulted in 180+ publications. Dr. Prasad is a Senior Member of the ACM and IEEE.