iDR: Consumer and Grid Friendly Demand Response System*

Vikas Chandan[#1], Tanuja Ganu[#1], Tri Kurniawan Wijaya[#2]†, Marilena Minou[#3],
George Stamoulis[#3], George Thanos[#3], Deva P. Seetharam[#1]
[#1]IBM Research India, [#2]EPFL, Switzerland, [#3]Athens University of Economics and Business, Greece

ABSTRACT

Peak demand is a major challenge for power utilities across the world. Demand Response (DR) is considered to be effective in addressing peak demand by altering consumption of end consumers, so as to match supply capability. However, an efficient DR system needs to respect end consumer convenience and understand their propensity of participating in a particular DR event, while altering the consumer demand. Understanding such preferences is non-trivial due to the large-scale and variability of consumers and the infrastructure changes required for collecting essential (smart meter and/or appliance specific) data.

In this paper, we propose an inclusive DR system, *iDR*, that helps an electricity provider to design an effective demand response event by analyzing its consumers' house-level consumption (smart meter) data and external context (weather conditions, seasonality etc.) data. *iDR* combines analytics and optimization to determine optimal power consumption schedules that satisfy an electricity provider's DR objectives - such as reduction in peak load - while minimizing the inconvenience caused to consumers associated with alteration in their consumption patterns. *iDR* uses a novel context-specific approach for determining end consumer baseline consumptions and user convenience models. Using these consumer specific models and past DR experience, *iDR* optimization engine identifies -(i) when to execute a DR event, (ii) who are the consumers to be targeted for the DR, and (iii) what signals to be sent. Some of *iDR*'s capabilities are demonstrated using real-world house-level as well as appliance-level data.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous

General Terms

User preferences, Human Factors, Optimal Scheduling

Keywords

Demand Response, Energy Consumption Scheduling, User Preferences, Smart Grids, Smart Meters

1. INTRODUCTION

Demand Response (DR) programs aim to alleviate the peak demand problem and provide higher system reliability by altering consumer demand in response to the power grid's supply and economic conditions. Various evaluation studies indicate that DR can be an effective mechanism for addressing the challenges of growing energy demand and related supply-demand imbalances [1–3]. A staff report by Federal Energy Regulatory Commission [4] estimates that the feasibility of peak demand reduction in the United States can be up to 20% by using DR technologies with full participation. As per another study [5], DR programs alone could achieve up to half of the European Union's 2020 targets concerning energy savings and CO2 emissions. Additionally, many electricity suppliers around the world are deploying smart meters to gather fine-grained spatio-temporal consumption data, and to provide a bi-directional communication mechanism [6]. This infrastructural enhancement is a significant enabler for bringing DR vision to reality.

However, the success of DR programs essentially depends upon the end consumers' participation. Various evaluation studies [7, 8] based on results of DR pilots and surveys of participants indicate that *discomfort/inconvenience* caused during a DR event is a key factor that adversely affects DR participation. Additionally, it was found that demographic attributes such as family size, income level, participants' activity structure, etc. are relevant factors for DR participation. Hence, any effective DR system needs to understand the end consumers' electricity usage preferences and their demographic attributes while choosing the right set of consumers for a DR contract and/or for a particular DR event.

Energy suppliers/aggregators offer various types of DR contracts to the end consumers, such as Time of Use (ToU) Rates, Capacity Biding Programs (CBP), Demand Biding Programs (DBP) and Peak Day Pricing (PDP) [9]. Any particular DR program needs to confer *fairness* amongst consumers under the same DR contract. Additionally, it should also take into account adverse side-effects of DR, such as rebound effect [10]. Considering above mentioned challenges, it is non-trivial to determine - (i) the *time window* for executing a DR event, considering supply conditions and

predicted consumer demand, (ii) the *right set of consumers to be targeted* for a particular DR event considering factors such as end consumers' preferences, DR contracts and DR participation history, (iii) *target reduction per consumer* and *expected DR outcome.*

To address the afore-mentioned challenges, in this paper, we present *iDR*, an inclusive DR planning system, that helps energy suppliers design effective DR events, while taking into account consumers' preferences and fairness of the system with respect to the DR contracts. The key contributions of this paper are:

1. A novel context-based method to determine baseline consumption and quantify inconvenience caused to consumers (in the form of utility functions), using smart meter data.

2. A novel and simple context-based approach to determine inconvenience (in the form of utility functions) using appliance-level data, if available.

3. A stochastic optimization framework that uses the above utility functions to determine when to execute DR events, which consumers to target, and what signals to send.

4. A mechanism to ensure fairness amongst multiple consumers with same type of DR contract while planning DR events.

The remainder of this paper is organized as follows. In Section II, we provide an overview of the related work. In Section III, we describe our context based approaches for determining end consumers' preferences of electricity usage. In Section IV, we define an optimization framework for optimal and fair DR scheduling. We showcase the evaluation of *iDR* on a real-world dataset in Section V, and finally, in Section VI, we conclude our work and discuss avenues for future work.

2. RELATED WORK

Determining consumer preferences for individual appliance usage and using these preferences for optimal appliance scheduling is well studied in literature. In [11], the authors consider optimal household appliance scheduling for dynamic pricing. They propose a system, *Yupik* that determines preferred time of use for individual appliances and generates appliance usage schedules to minimize both a household's energy costs and potential lifestyle disruptions. [12] presents a methodology to schedule appliances taking into account cost, scheduling preferences and climactic comfort requirements. In [13], authors propose an approach for determining consumer preferences for appliance use by categorizing household appliances into four types and deriving different utility functions for each of these types. Further, they propose an optimal DR strategy based on utility maximization, where dual objectives are formulated for cost minimization from the utility provider's perspective, and social welfare maximization from the consumers' perspective. [14] extends this work for computation time reduction by proposing load consolidation and a LP framework.

Though the existing literature provides various approaches for modeling and determining consumer preferences and optimal scheduling for demand response events or dynamic pricing, their practical usefulness is somewhat limited because of the underlying complexity associated with the use of appliance level utility models only. For example, a utility provider might not have access to appliance level consumption information for all its consumers. This motivates the need for a DR planning methodology which is not dependent on fine-grained appliance-level information, but works well with house-level (smart meter) consumption data only. In this paper, we present a novel approach for determining consumer preferences using smart meter data only, and/or appliance level data, if available. Additionally, the existing work mainly considers automatic demand response scenarios and hence proposes the direct control or scheduling of individual home appliances. On the contrary, we consider a more practical and widely prevalent demand response scenario wherein the utility provider sends DR signals to consumers, and thereafter consumers can accept or decline the signals and determine appliance schedules as per their convenience. In this context, this paper presents a *stochastic optimization* approach to determine when to send DR signal (optimal DR timeslots), whom to send it (selecting the right set of consumers) and how much reduction to target.

There is also an increasing interest in obtaining useful DR related insights from smart meter data such as application specific consumer segmentation or baseline load forecasting for individual consumers. For example, smart meter data has been used for consumer segmentation in [15] and [16]. A body of literature also exists on baseline consumption estimation. KEMA Inc. [17] and EnerNOC [18] provide excellent overviews of the baseline estimation methods employed by utilities in the United States. The above-mentioned references serve as enablers or background for our approach of optimal DR using smart meter data. However, existing work has not used smart meter data for understanding consumer preferences and subsequent DR planning. The proposed work addresses these gaps by providing a methodology to plan and schedule DR events based on smart meter data alone as well as combining appliance-level data, if available.

3. DETERMINING CONSUMER PREFERENCES

In this section, we present methodologies to model the utility (benefit) derived by a consumer as a function of her consumption. These models form the basis of the optimization framework presented in section 4. We propose two utility modeling frameworks - one at the aggregate house level (using smart meter data), and another which uses appliance level consumption data if such high resolution sensing infrastructure is available. The common nomenclature used in this paper is shown in Table 1.

3.1 Preference mining using smart meter data

Analysis of historical consumption data for a consumer over time can provide information on her consumption preferences. We use two quantities - baseline and utility - to quantify such preferences. Baseline is defined as an estimate of the electrical load drawn by a consumer in the absence of any DR related curtailment actions. Utility is defined as the 'benefit' obtained by a consumer corresponding to a given amount of consumption. The objective of preference mining in this work using smart meter data is to determine baselines and utilities for each consumer. The underlying methods are explained below.

Symbol	Description
d	General notation to represent a day
i	General notation to represent a consumer
t	General notation to represent a timeslot
T	Number of timeslots in a day d
N	Number of consumers served by a utility company
$Q_{i,t,d}$	Consumption in kWh by consumer i during timelot t in day d
$Q^b_{i,t,d}$	Baseline consumption in kWh by consumer i during timeslot t in day d
$U_{i,t,d}(.)$	Utility function associated with consumer i during timeslot t in day d
$Q^s_{t,d}$	Maximum amount of energy in kWh that the utility can supply during timelot t in day d
n_t	Maximum number of consumers that the utility decides to target for DR during timeslot t
η^t_{max}	Maximum reduction in consumption as a fraction of baseline consumption
$\nu_{i,t,d}$	Boolean variable representing whether consumer i should be targeted for DR during timeslot t in day d
$\Delta Q_{i,t,d}$	Reduction in consumption (kWh) from baseline by consumer i during timelot t in day d

Table 1: Nomenclature of common symbols

3.1.1 Baseline estimation

Existing methods for baseline estimation involve averaging [17], regression [19] and time series analysis [20]. In this work, we obtain baselines by first identifying appropriate consumer contexts. A context is defined as a combination of external and internal factors that influence a consumer's consumption decisions. For instance, at a month level, influencing contexts could be season of a year (summer, autumn, winter or spring) or at a finer resolution month of the year (January, February, etc.). Similarly, at a day level, possible contexts are weekday/weekend or day of the week(Monday, Tuesday, etc.).

Multiple contexts may be associated with a particular day. As an example, October 17th, 2013 can be represented by contexts such as {Autumn}, {Autumn, weekday}, {October, Thursday}, etc. To choose the most appropriate context, and obtain the corresponding baseline consumptions, we proceed as follows:

1. We denote the day for which baselines are to be estimated by d. We assume that this day is divided into T discrete timeslots. We first identify the set $C(d)$ consisting of all contexts and their combinations that explains the day d.

2. Based on the historical data, we find the context $C^* \in C(d)$, whose consumption has lowest dispersion. Here dispersion can be quantified using statistical measures such as standard deviation or inter-quartile range.

3. The baseline consumption for each timeslot $t \in \{1, 2, ..., T\}$ is then determined by a measure of central tendency - such as mean or median - applied to historical consumption data at timeslot t on previous days in the context C^*.

In obtaining the baseline consumption as described above, historical data can be used in several different ways. The simplest method is to use a static time window, where consumption data from all previous days in the context are used,

irrespective of how far in the past these days are from the day d. On the other hand, a moving window approach only uses consumption data from N_{prev} most recent days. An exponentially moving average approach is a variation of the moving window approach, where a weighted average is performed with more importance (higher weights) being given to the more recent days.

Figure 1 provides an evaluation of the proposed context based baseline estimation method with respect to existing approaches based on average mean absolute error (MAE) across consumers for residential consumer [21] and commercial consumer [22] datasets. Our proposed approach performs better than existing approaches on both datasets. For a more detailed comparative analysis of various methods for baseline prediction, the user is directed to the study [17].
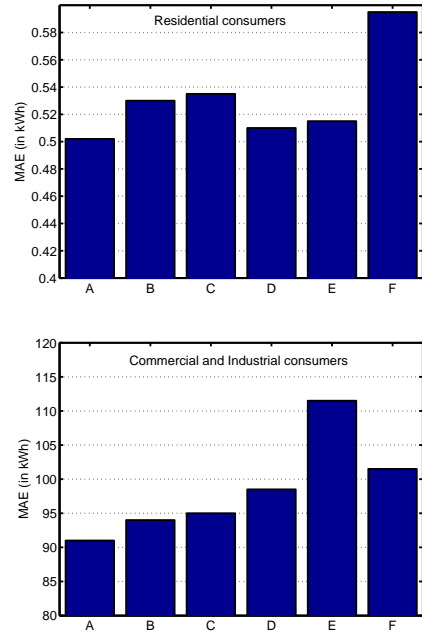

Figure 1: Evaluation of the proposed baseline estimation method with respect to existing approaches based on average mean absolute error (MAE) across all consumers for Residential consumers [21] and Commercial Consumers [22] datasets. Approaches A and B present our proposed context-based baselining method with (A) using Enlarged window with exponentially moving average and (B) using Enlarged window. Approaches C to F present existing methods [17] - (C) CAISO, (D) ISONE, (E) Low5Of10 and (F) NYISO.

3.1.2 Utility quantification

To express the utility derived by a consumer as a function of her consumption, we make the following assumptions:

1. A consumer derives maximum utility when her consumption equals baseline, since by definition the baseline represents the preferred level of consumption.

2. Any deviation from baseline - either above or below - results in a reduction of utility.

In order to satisfy the above assumptions, we propose the use of probability distribution of consumption during the

specific timeslot (on the past consumption data filtered by optimal context C^*) to quantify utilities. However, it should be noted that the use of probability distribution reflects a specific choice, among several candidate methods for quantifying utilities that can potentially satisfy the above assumptions.

The probability associated with a consumption of q kWh, reflects the consumer's preference for consuming an amount of energy q KWh during the specified slot and under the context C^*. To make this setting more formal, we introduce the concept of utility functions.

The utility function is defined as follows:
Consider a consumer i, and a timeslot t in the day, where $i \in \{1, 2, ..., N\}$, and $t \in \{1, 2, ..., T\}$ and d is the day in consideration for DR planning. Let C^* be the optimal context for the triplet $\{i, t, d\}$ based on the method discussed in Section 3.1.1 and $Q_{i,t,d}^b$ be the corresponding estimated baseline consumption. For any particular consumption $Q_{i,t,d}$ by the consumer at this time slot of the day, associated utility function $U_{i,t,d}(.)$ is expressed as:

$$U_{i,t,d}(Q_{i,t,d}) = \frac{p^{C^*}(Q_{i,t,d})}{p^{C^*}(Q_{i,t,d}^b)} \qquad (1)$$

where $p^{C^*}(q)$ gives the probability of consumption q in the past data under optimal context C^*. The consumer preferences - captured via baselines and utility functions as described above - are inputs that are used to set up the optimization frameworks presented in Section 4.

3.2 Preference mining using appliance data

So far, we have presented an approach for quantifying consumer preferences using aggregated, household level, consumption data. Nevertheless, assuming that appliance level data is available, for example where AMIs are in place, there is an evident potential to leverage this information in order to form a more detailed and accurate depiction of user's preferences. This section presents a methodology for achieving that. The proposed methodology is based on mining utility (benefit) derived from each monitored appliance. More specifically, it estimates the importance that each consumer places on a specific appliance via calculating a corresponding "weight". These weights constitute a measure of how flexible (elastic) the consumers are in altering the typical use of each appliance, for example in the case of a DR event.

Our methodology follows three distinct phases. At first, by using statistical analysis, we derive the utility values of each appliance in order to formulate the consumption profile of each household. Next, we fit the derived utility values from phase 1 to those resulting from the utility functions of [13], thus also validating the results from our approach against a more theoretic one that assumes the knowledge of utility functions. Finally, exploiting the fitted utility functions we mine the weights that reflect consumer's preferences regarding the usage of appliances, under specific realistic assumptions.

As a prerequisite, our work adopts the categorisation of household devices that is proposed in the widely cited work [13]. According to this reference, the typical household appliances can be classified into four types. The first type includes appliances such as refrigerator and air-conditioner which control temperature. The second type includes appliances such as washing machine where the consumer only cares about whether a task is completed before a certain

time. The third category includes appliances such as lighting that must be ON for a certain period of time. The fourth type includes appliances such as TV or computer that the consumer uses for entertainment. Each type is characterized by a utility function $U_{i,\alpha}(q_{i,\alpha})$ that models how much consumer i values the consumption vector $q_{i,\alpha}$ and a set of linear constraints on $q_{i,\alpha}$.

To experimentally evaluate our approach, we use real consumption data from 6 households in India. The available data comprises of sensor readings at a granularity of ten seconds for four different appliances: air-conditioner, fridge, washing machine and television (TV). These readings are extracted for a given context C, that is weekdays in July and correspond to three of the four appliance categories proposed by [13] (Type 1, 2 and 4), as the fridge and the air-conditioner both belong to the same, Type 1 category. This is because sensor readings for Type 3 appliances (e.g. lighting) were not available.

For conciseness, we present application details of our proposed methodology to only one of the three categories of devices (Type 4). However, the same approach was used for the other three devices.

3.2.1 Phase 1: Derivation of utility values from appliance level data

In Phase 1, the utility values for the respective devices are empirically estimated using statistical analysis. We now explain this process for Type 4 devices. Type 4 category includes appliances such as TV, video games, and computers that a consumer uses for entertainment. In this case, the consumer's utility depends on two factors: how much power is consumed at each time she wants to use them, and how much total power is consumed over the entire day. At each time, $t \in \{1, 2, ..., T\}$, we assume that the consumer i attains a utility $U_{i,t,\alpha}(q_{i,t,\alpha})$ from energy consumption $q_{i,t,a}$ on appliance α.
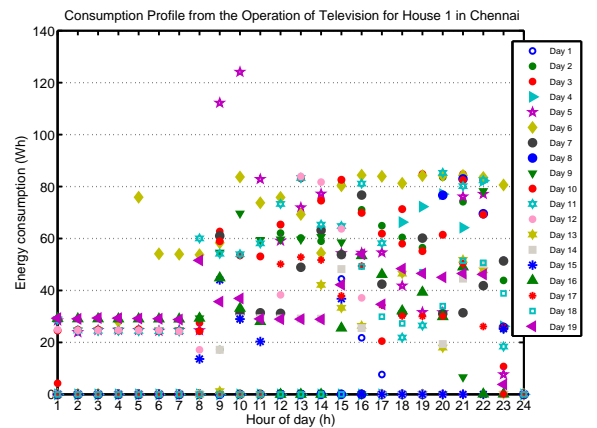

Figure 2: Hourly consumption of TV operation of House 1 during weekdays in July in Chennai, India. The figure shows the behavior of a specific consumer (House 1) over a specific appliance under specific context (weekdays in July).

By constructing the histograms of the consumption vector $Q_{i,d,\alpha} := \{q_{i,t,\alpha}, t \in \{1, 2, ..., T\}\}$, over all timeslots, as shown in Figure 2, we can analyse the consumer behavior i.e. the timeslots when their appliances are operated and the associated consumption level. For this type of appliances we can assume that higher the consumption, the greater the sat-

isfaction experienced by the consumers, which in economic theory is expressed as a "utility value". At this point, as we do not have any knowledge of the function that expresses the relationship between the consumption values and the obtained utility values, the latter can be approximated by the empirical cumulative distribution function for given consumer i, appliance α, timeslot t and context, as follows:

$$U_{i,t,\alpha}(q_{i,t,\alpha}) = \begin{cases} 0, & \text{if } q_{i,t,\alpha} < q_{i,t,\alpha}^{min} \\ eCDF_{i,t,\alpha}(q_{i,t,\alpha}), & \text{if } q_{i,t,\alpha} \in [q_{i,t,\alpha}^{min}, q_{i,t,\alpha}^{max}] \\ 1, & \text{if } q_{i,t,\alpha} > q_{i,t,\alpha}^{max} \end{cases}$$

(2)

Here, $q_{i,t,\alpha}^{min}$ and $q_{i,t,\alpha}^{max}$ are minimum and maximum consumption observed in timeslot t, respectively in the given context, and $eCDF_{i,t,\alpha}(x)$ represents the value of the empirical cumulative distribution function for consumer i when operating appliance α at power level x at timeslot t. In particular, given the range of observed consumption values $[q_{i,t,\alpha}^{min}, q_{i,t,\alpha}^{max}]$, and by assuming that the utility is an increasing function of consumption and thus monotonic, we define the utility value associated with a consumption x as the cumulative frequency of consumption values $\leq x$ in this range.

Note that, when calculating the probability of a consumption value to be less than a certain value, the eCDF takes into consideration situations of zero consumption, i.e. when no utility is gained. Therefore, we do not explicitly take such situations into account in deriving the utility values.

Next, to define the utility value $U_{i,d,\alpha}(Q_{i,d,\alpha})$ for a specific day d, a given consumer i, and a given appliance α, where

$$Q_{i,d,\alpha} = [q_{i,1,\alpha}, q_{i,2,\alpha}, ..., q_{i,T,\alpha}]^T,$$

we estimate it as the median - weighted sum:

$$U_{i,d,\alpha}(Q_{i,d,\alpha}) = \frac{\sum_{t=1}^{t=T} (median_{i,t,\alpha} \times U_{i,t,\alpha}(q_{i,t,\alpha}))}{\sum_{t=1}^{t=T} median_{i,t,\alpha}}, \quad (3)$$

where, $median_{i,t,\alpha}$ is the median of consumption data of appliance α in timeslot t for a consumer i and given context, and it is used in order to avoid the analysis to be skewed in favour of very high or very low consumption values.

3.2.2 Phase 2: Fitting the estimated utility values to the utility functions of [13]

In the next step, our approach uses the utility functions of the different types of appliances from [13] represented as continuously differentiable concave functions of the total consumption in timeslot t and day d. Our objective is to approximate these coefficients such that the resulting utility values match those statistically estimated in Phase 1. This can be achieved by minimizing the distance of the curve formed from the utility values, i.e. the eCDF curve at the points of evaluation, from the curve of the utility functions in [13].

We adopt the utility functions as well as the initialization data defined in the experimental section of [13]. In particular, for Type 4 appliances, the utility function takes the form of:

$$U_{i,t,\alpha}(q_{i,t,\alpha}) = C_\alpha - (b_\alpha - \frac{q_{i,t,\alpha}}{\bar{q}_\alpha})^{-1.5}$$

(4)

where $C_\alpha \geq 0$, $b_\alpha \geq 0$ and $\bar{q}_\alpha > 0$ are the variables to be fitted. In this case, the fitting is executed at each timeslot t due to the nature of the appliance and corresponds to a non linear regression problem.
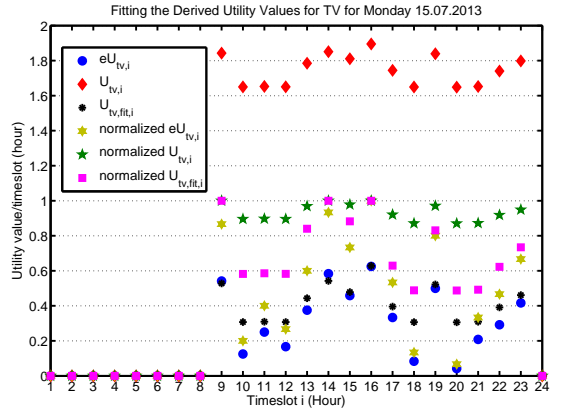

Figure 3: TV - Fitting the derived utility values to utility values based on [13] for Monday - Jul 15, 2013

Figure 3 presents the results for a particular consumer and a particular day, while the notation follows the nomenclature in Table 2. The day was divided into 4 timeslots, that is $T = 4$. The coefficients obtained using least squares fit are $C_{TV} = 2.2953$, $b_{TV} = 0.5786$ and $\bar{q}_{TV} = 7.5397 \times 10^3$. From Fig. 3, we observe that the normalized utility values ($U_{TV,fit,t}$), after fitting are close to the normalized values of $U_{TV,t}$ from the utility functions.

Symbol	Description
$eU_{\alpha,t}$	Estimated utility values derived in Phase 1 for appliance α during timeslot t
$U_{\alpha,t}$	Utility values resulting from utility functions in [13] for appliance α during timeslot t
$U_{\alpha,fit,t}$	Utility values resulting from applying fitted coefficient values to the utility functions for appliance α during timeslot t
normalized $eU_{\alpha,t}$	is calculated as $\frac{eU_{\alpha,t}}{\max(eU_{\alpha,t})}$
normalized $U_{\alpha,t}$	is calculated as $\frac{U_{\alpha,t}}{\max(U_{\alpha,t})}$
normalized $eU_{\alpha,fit,t}$	is calculated as $\frac{U_{\alpha,fit,t}}{\max(U_{\alpha,fit,t})}$

Table 2: Nomenclature of symbols used in Figure 3

In general, the results from the fitting methodology for all appliances imply that our approach of statistically deriving the consumers preferences, in terms of utility values, with use of detailed consumption data and without any knowledge about the utility functions has small deviations from the case where full knowledge is available, as in [13].

3.2.3 Phase 3: Mining the weights of each appliance

To derive the weights of each appliance that reflect the consumer's preferences, we use the assumption that given two different days of the same context, e.g. two weekdays, the consumption profile might vary, but the total Net Benefit (NB) obtained (i.e. the utility gained minus the monetary cost) is the same. This implies that for both days if the energy tariff scheme is known - a valid assumption in our case - each consumer chooses that consumption, which maximizes her total NB. To apply this assumption to our methodology, we calculate one respective weight per timeslot for the TV (type 4 appliance), fridge (type 1 appliance) and air-conditioner (type 1 appliance) and one per day for

the washing machine (type 2 appliance). We use only one weight for the washing machine as we are interested only in the total consumption of the appliance after each operation, since the respective load (in one washing cycle) can be interrupted by the consumer and shifted to subsequent timeslots, which cannot be the case, for example, in the use of an air-conditioning device.

Thereby, given a set of appliances $\alpha \in A, A := \{AC, FR, WM, TV\}$, the weights $w_{i,\alpha}, \alpha \in A$ and the same context for each day $d, d \in C$, the NB is of the following form:

$$NB_{i,d} = \sum_{\alpha \in A} w_{i,\alpha} U_{i,d,\alpha}(Q_{i,d,\alpha}) - C_{i,d}(\sum_{\alpha \in A} Q_{i,d,\alpha}), \quad (5)$$

where $Q_{i,d,\alpha} = [q_{i,1,\alpha}, q_{i,2,\alpha}, q_{i,3,\alpha}, q_{i,4,\alpha}]$, and $C_{i,d}(\sum_{\alpha \in A} Q_{i,d,\alpha})$ are the cost functions based on which the consumer is charged for electricity consumption. Note that in this phase we consider all type 1 appliances of consumer i as a single appliance for simplicity. Thus, by equating the NB formulas for all the combinations of days from the available set of days C and applying linear least squares optimization, we derive the weights presented in Table 3.

Timeslot	$w_{i,Type4}$	$w_{i,Type1}$
1	0.0004	29.7085
2	0.0013	0.4750
3	4.6355	0.6577
4	0.5947	0.0007
$w_{i,Type2}$		
7.0732		

Table 3: Values of appliances' weights using NB

The results indicate that consumers are rather inelastic to changes in the consumption of type 1 appliances (i.e. their corresponding weights exhibit the greatest values amongst all), which is rational due to the fact that the use of such appliances (for e.g. cooling devices) in the given context of July weekdays is of paramount importance to them. We also observe that the weights that correspond to type 4 appliances can take mostly smaller values than those of type 1 and type 2 appliances respectively. This implies that this type of load can be curtailed at a specific time period or shifted to another time period, without causing significant inconvenience to consumers.

On the other hand, we observe that type 2 appliances are more elastic in terms of the time of their operation. This means that their load cannot be curtailed or shifted before a specific task is finished but they can operate at anytime during the day, which explains why the resulting value shown in Table 3 is greater than the values of type 4. Generally, the conclusions one can deduce about the importance of the appliances in the consumers' everyday life are consistent with the respective bibliography about the categorization of appliances in terms of shiftable and curtailment load. An energy provider can therefore leverage our methodology to quantify each consumer's personal preferences, and use them in order to predict their behaviour in the case of a DR signal being sent to them.

To conclude, although we have presented two frameworks for understanding and quantifying consumers' preferences, in the remainder of this paper we use the framework based on smart meter data alone (Section 3.1) for further development of iDR. This is because the rapid penetration of smart meters in today's grids have allowed utility companies access to consumers' aggregate consumption data at the household level. Therefore, we believe that at the current state of metering infrastructure, the framework based on smart meter data is more universally applicable than that based on appliance level data. However, we recommend the use of appliance level preference mining, wherever possible, as we believe it is a more accurate representation of consumers' consumption patterns.

4. OPTIMIZATION FRAMEWORK

In this section, we present a methodology that can be used by the utility provider for planning DR events. In particular, for a given day, we seek to determine the following in advance: (i) when should DR events be conducted, (ii) which consumers should be targeted, and (iii) what DR signals should be sent to each such consumer. Our approach involves solving an optimization problem, which uses consumers' preferences quantified in Section 3.1. The goal of the optimization is to plan DR response events such that the inconvenience caused to consumers is minimized while the utility provider's targets with regard to reduction in consumption are achieved. In the presentation of this methodology, we do not make any specific assumption on the form of the utility function. Therefore, it can be applied to any candidate utility function deemed appropriate by the reader. An example of such a utility function was provided in (1).

Before a formal presentation of the framework, we first define the underlying nomenclature. The number of consumers served by the utility is denoted by N. The day for which DR planning is to be performed is denoted by d. It is divided into T timeslots. Consider a consumer i, and a timeslot t in the day, where $i \in \{1, 2, ..., N\}$, and $t \in \{1, 2, ..., T\}$. For the triplet $\{i, t, d\}$, the best matching context C^* can be obtained using the methodology described in Section 2. Using this context, a baseline consumption $Q_{i,t,d}^b$ and the corresponding utility function $U_{i,t,d}(.)$ are determined.

We assume that for each timeslot t in the day d under consideration, the utility provider determines, in advance, the maximum amount of energy, $Q_{t,d}^s$ to be supplied. Such a determination can be made based on a combination of the provider's power procurement schedule from power generating companies for that day, and economics associated with power purchase. For instance, if it is known that buying electricity during certain time periods (e.g. peak demand periods) would be more expensive than during other periods, the utility provider may choose to limit the amount of energy procured during those time periods.

Given the above setting, we first determine the set T_{DR} of timeslots when DR events should be conducted. This set consists of timeslots when the expected (baseline) consumption, added across consumers exceeds the desired supply, and is more formally expressed as:

$$T_{DR} = \left\{ t : \sum_{i=1}^{N} Q_{i,t,d}^b \geq Q_{t,d}^s, t \in \{1, 2, ...T\} \right\}. \quad (6)$$

Next, to determine the consumers to be targeted for DR and the DR signals to be sent, we propose the following optimization frameworks. The first framework assumes that the targeted consumers would always participate in the DR event. The second framework assumes that these consumers would participate only with some pre-determined probability. Two design variables - n_t and η_{max}^t - are used in these

frameworks. The former represents the maximum number of consumers that the utility would like to target during timeslot t for participation in a DR event, whereas the latter represents the maximum reduction in consumption, expressed as a fraction of the baseline consumption $Q_{i,t,d}^b$ for each consumer, that the utility provider can request during a DR event at timeslot t. The use of n_t as a design variable is motivated by the need to reduce the computational requirements associated with the optimization. For example, in a realistic scenario, a utility provider can cater to millions of consumers, but might like to involve only a few thousand consumers at any given time for DR.

4.1 Deterministic framework

For all $t \in T_{DR}$, the following optimization problem is solved:

$$\left\{ \nu_{i,t,d}^*, \Delta Q_{i,t,d}^* \right\} =$$

$$\underset{\left\{ \nu_{i,t,d}, \Delta Q_{i,t,d} \right\}}{\operatorname{argmax}} J_d := \sum_{i=1}^{N} U_{i,t,d} \left(Q_{i,t,d}^b - \nu_{i,t,d} \Delta Q_{i,t,d} \right) \quad (7)$$

Subject to:

$$\sum_{i=1}^{N} \left(Q_{i,t,d}^b - \nu_{i,t,d} \Delta Q_{i,t,d} \right) \leq Q_{t,d}^s, \quad (8)$$

$$\nu_{i,t} \in \{0,1\} \quad \text{for all} \quad i \in \{1,2,...,N\}, \quad (9)$$

$$\sum_{i=1}^{N} \nu_{i,t} \leq n_t, \text{ and} \quad (10)$$

$$0 \leq \Delta Q_{i,t,d} \leq \eta_{max}^t Q_{i,t,d}^b \quad \text{for all} \quad i \in \{1,2,...,N\}. \quad (11)$$

The objective of the above optimization, expressed by (7) is to maximize the aggregated utility obtained by the consumers during a DR event. In other words, we seek to minimize the inconvenience caused to the consumers resulting from reduction in consumption. The decision variables in the framework are as follows: (i) for each $i \in \{1,2,...,N\}$, the Boolean variable $\nu_{i,t,d}$ represents whether consumer i should be targeted for DR during timeslot t, indicated by a value of 1, and 0 otherwise; (ii) for each $i \in \{1,2,...,N\}$, the variable $\Delta Q_{i,t,d} \geq 0$ represents the target reduction in consumption for consumer i during the DR timeslot t. The constraint (8) restricts the aggregate consumption during each DR timeslot $t \in T_{DR}$ to be within the consumption bound $Q_{t,d}^s$ imposed by the utility provider. The DR design variables n_t and η_{max}^t, explained earlier, appear in constraints (10) and (11), respectively.

4.2 Stochastic framework

We denote the probability of user i responding to a DR event at timeslot t in day d by $p_{i,t,d}$. This probability can be determined, for instance, by analyzing the user's response history to previous DR signals sent to her on days and timeslots which correspond to the context C^*.

Next, we introduce a random boolean variable $\alpha_{i,t,d}$, which indicates the ith consumer's decision to participate, when the utility provider targets that consumer for a DR event. This setting can be mathematically expressed using the following equations:

$$\alpha_{i,t,d} \in \{0,1\} \quad \text{for all} \quad i \in \{1,2,...,N\}, \quad \text{and} \quad (12)$$

$$P\left(\alpha_{i,t,d} = 1 | \nu_{i,t,d} = 1 \right) = p_{i,t,d}. \quad (13)$$

A modified objective function J_s is defined as shown below, where $E(.)$ is the expected value operator.

$$J_s := E\left(\sum_{i=1}^{N} U_{i,t,d} \left(Q_{i,t,d}^b - \alpha_{i,t,d} \nu_{i,t,d} \Delta Q_{i,t,d} \right) \right). \quad (14)$$

The stochastic version of the constraint (8) is given by

$$E\left(\sum_{i=1}^{N} \left(Q_{i,t,d}^b - \alpha_{i,t,d} \nu_{i,t,d} \Delta Q_{i,t,d} \right) \right) \leq Q_{t,d}^s \quad (15)$$

The stochastic optimization framework, therefore seeks to maximize J_s given by (14), subject to the constraints (9) to (11), and the constraints (12), (13) and (15). Hence it represents a chance-constrained mixed integer program. The expected values appearing in equations (14) and (15) can be evaluated using (13), resulting in deterministic versions of these equations given by (16) and (17), respectively. It should be noted that by setting $p_{i,t,d}$ to 1, we recover the deterministic framework presented in Section 4.1.

$$J_s = \sum_{i=1}^{N} \left[U_{i,t,d} \left(Q_{i,t,d}^b \right) - \nu_{i,t,d} p_{i,t,d} X \right]$$

$$\text{where } X = \left\{ U_{i,t,d} \left(Q_{i,t,d}^b \right) - U_{i,t,d} \left(Q_{i,t,d}^b - \Delta Q_{i,t,d} \right) \right\}. \quad (16)$$

$$\sum_{i=1}^{N} \left(Q_{i,t,d}^b - \nu_{i,t,d} p_{i,t,d} \Delta Q_{i,t,d} \right) \leq Q_{t,d}^s. \quad (17)$$

The use of a stochastic framework can also allow the utility provider to ensure fairness on the basis of DR contract. For example, if the contract specifies a maximum number of DR signals $N_{DR,i}^{max}$ to be sent to a consumer i in a given time slot t, the probabilities $p_{i,t,d}$ can be iteratively adjusted (e.g. once every day) as shown below:

$$p_{i,t,d} = 1 - \frac{N_{DR,i}^{curr}}{N_{DR,i}^{max}} \quad (18)$$

Here, $N_{DR,i}^{curr}$ denotes the total number of times the consumer i has been selected at time slot t since the beginning of the contract upto the current iteration. We observe that (18) results in probabilities of participation which are monotonically decreasing functions of the number of times the consumer has been targeted for DR. When the consumer has been selected $N_{DR,i}^{max}$ number of times, $p_{i,t,d}$ becomes zero, which ensures that she would not be targeted for DR any further.

4.3 Feasibility condition

In this section, we prove a necessary and sufficient condition for feasibility of the afore-mentioned optimization framework, and determine a feasible point that can be used as a candidate initial point to solve it.

Theorem 1: *Without loss of generality, we assume that for a given $t \in T_{DR}$, the consumers indexed by $i \in \{1,2,...,N\}$ are arranged in the increasing order of $p_{i,t,d} Q_{i,t,d}^b$. The problem of maximizing J_s given by (16) subject to the constraints (17) and (9) to (11) admits a feasible solution if and only if the following inequality is satisfied:*

$$\sum_{i=1}^{N} Q_{i,t,d}^b - Q_{t,d}^s \leq \eta_{max}^t \sum_{i=N-n_t+1}^{N} p_{i,t,d} Q_{i,t,d}^b. \quad (19)$$

If the condition (19) *is satisfied, decision variables obtained by the assignments* (20) *and* (21) *are feasible.*

If $i \in \{1, 2, ..., N - n_t\}$, $\nu_{i,t,d} = 0$, $\Delta Q_{i,t,d} = 0$, and $\quad$ (20)

If $i \in \{N - n_t + 1, ..., N\}$, $\nu_{i,t,d} = 1$, $\Delta Q_{i,t,d} = \eta^t_{max} Q^b_{i,t,d}$. $\quad$ (21)

PROOF. We first prove the 'necesssay' part. If for all $i \in \{1, 2, ..., N\}$, $\nu_{i,t,d}$ and $\Delta Q_{i,t,d}$ are such that they satisfy (17) and (11), they also satisfy:

$$\sum_{i=1}^{N} Q^b_{i,t,d} - Q^s_{t,d} \leq \sum_{i=1}^{N} \nu_{i,t,d} p_{i,t,d} \Delta Q_{i,t,d}$$

$$\leq \eta^t_{max} \sum_{i=1}^{N} \nu_{i,t,d} p_{i,t,d} Q^b_{i,t,d}. \quad (22)$$

Since the set $\{1, 2, ..., N\}$ is sorted in the increasing order of $p_{i,t,d} Q^b_{i,t,d}$, the expression in the RHS of (22) attains a maximum value of $\eta^t_{max} \sum_{i=N-n_t+1}^{N} p_{i,t,d} Q^b_{i,t,d}$ due to the remaining constraints (9) and (10). Therefore, (22) can be satisfied only if: $\sum_{i=1}^{N} Q^b_{i,t,d} - Q^s_{t,d} \leq \eta^t_{max} \sum_{i=N-n_t+1}^{N} p_{i,t,d} Q^b_{i,t,d}$.

To prove the 'sufficient' part, we assume that the condition (19) is satisfied. It can be easily verified that the decision variables obtained by assignments (20) and (21) satisfy the following relationship for all $i \in \{1, 2, ..., N\}$.

$$\nu_{i,t,d} p_{i,t,d} \Delta Q_{i,t,d} = \eta^t_{max} \nu_{i,t,d} p_{i,t,d} Q^b_{i,t,d}. \quad (23)$$

Using (19) and (23), we establish that the above chosen decision variables also satisfy the constraint (17). Since these decision variables also satisfy (9), (10) and (11), we conclude that this choice of decision variables is feasible, that is it satisfies all underlying constraints. Hence the condition (19) leads to a feasible choice of decision variables, given by (20) and (21). $\quad \square$

Note that if the condition (19) is not satisfied for a particular choice of design parameters, either η^t_{max} or n_t or both might have to be increased until it is satisfied.

5. EXPERIMENTAL EVALUATION

In this section, we perform simulation experiments using a real world consumption data set to evaluate the DR planning methodology presented in Section 4. We compare and evaluate the performance of three approaches - (i) deterministic optimization (Section 4.1), (ii) stochastic optimization (Section 4.2), and (iii) a rule based approach presented later in Section 5.5. The goal of these experiments is to understand and explain the results of the proposed optimization frameworks using intuition, and to quantify the efficacy of these frameworks by comparing with the rule-based approach. Each of these experiments is performed in two steps. In the first step, only 10 consumers are included to facilitate an easy understanding of the results. In the second step, the optimization is repeated on the complete set of consumers to demonstrate the scalability of *iDR*. Lastly, we also include an experiment, where we investigate the impact of the fairness strategy proposed in (18).

5.1 Data collection and pre-processing

We use consumption data available in the CER Ireland dataset [21]. This dataset contains approximately 5000 consumers, both residential and small and medium enterprises.

Measurements were obtained using smart meters for a period of 1.5 years, from July 2009 to December 2010. Since the dataset was collected as a part of a dynamic pricing trial, we select consumers who are in a single control group. In addition, we choose residential consumers with no missing data, thereby resulting in $N = 500$ consumers. For each day, and for each consumer, the recorded data is used to obtain consumption in kWh for each one hour time slot ($t \in \{1, 2, ..., 24\}$) during the day. In this way, we obtain the consumption $Q_{i,t,d}$ corresponding to each triplet $\{i, t, d\}$, where i is the consumer index ($i \in \{1, 2, ..., N\}$), t represents a slot in the day ($t \in \{1, 2, ..., 24\}$), and d refers to a day in the dataset.

5.2 Parameters for optimization

In the simulation experiments reported in this section, we choose the day d_{DR} for DR planning as April 1, 2011, which represents the first day of summer in 2011. We select the optimal context C^*, such that $d_{DR} \in C^*$, for each consumer as described in Section 3.1.1. Let d_C be the set of days belonging to context C^* for the given consumer. For simplicity and computational efficiency, we assume Gaussian distribution of consumption over the set d_C to calculate utility functions using (1). Thus, on day d_{DR}, for each pair $\{i, t\}$, where $i \in \{1, 2, ..., N\}$ and $t \in \{1, 2, ..., 24\}$, we determine the baseline consumption $Q^b_{i,t,d_{DR}}$ and the utility function $U_{i,t,d_{DR}}(.)$ using equations (24) and (25).

$$Q^b_{i,t,d_{DR}} = \mu_{i,t,d_C}, \quad (24)$$

$$U_{i,t,d_{DR}}(Q_{i,t,d_{DR}}) = \exp\left\{ -\frac{\left(Q_{i,t,d_{DR}} - Q^b_{i,t,d_{DR}}\right)^2}{2\sigma_{i,t,d_C}} \right\}. \quad (25)$$

Here, μ_{i,t,d_C} and σ_{i,t,d_C} denote the mean and standard deviations, respectively, of the set $\{Q_{i,t,d} : d \in d_C\}$. Equation (24) sets the baseline consumption for a consumer in any timeslot to her mean consumption in that timeslot observed on days which lie in the appropriate context C^*. Equation (25) uses a Gaussian function to describe a utility value between 0 and 1 for consumption in each timeslot, ensuring that the maximum utility of 1 is obtained when consumption equals baseline consumption. Note that (25) satisfies the assumptions in Section 3.1.2.

As the first step in DR planning, we determine the set of timeslots in the day, T_{DR} to be targeted for DR. The sum $\sum_{1=1}^{N} Q^b_{i,t,d_{DR}}$ of baseline consumptions for all consumers, corresponding to each hourly timeslot in the day is plotted in Figure 4. The utility provider's assumed target energy delivery schedule is also shown in Figure 4. From a comparison of these plots, we determine, using (6) that the set of timeslots to be targeted for DR is $T_{DR} = \{13, 22\}$. To solve the mixed integer nonlinear optimization problems (MINLP) appearing in Sections 4.1 and 4.2, we use the NOMAD algorithm [23] implemented in the OPTI toolbox in MATLAB [24]. NOMAD is a mesh-adaptive direct search algorithm with the potential to find global solutions to MINLPs. It should be noted that the objective of this work is not to determine the best available solution method for the MINLP considered. Hence our use of NOMAD should be treated as one of the choices among several possible solvers, and the interested reader is free to use any other solver to perform these experiments. To analyze and understand the results of the optimization, the baseline consumptions $Q^b_{i,t,d_{DR}}$ and the standard deviations σ_{i,t,d_C} for each consumer for the DR

timeslots $t = 13$ and $t = 22$ are shown in Tables 4 and 5, respectively. Given the aforementioned setup and parameters, we now describe various simulation experiments to evaluate the proposed DR planning methodology.
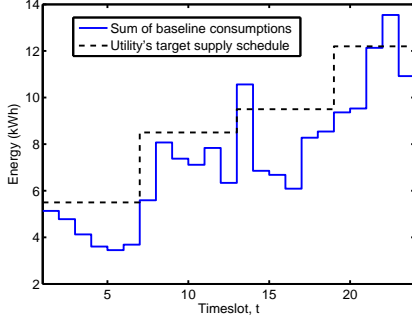

Figure 4: Comparison of the sum of baseline consumptions and the utility provider's power delivery schedule over the day d_{DR}.

Consumer number	$Q^b_{i,t,d_{DR}}$ (kWh)	σ_{i,t,d_c} (kWh)	$p_{i,t,d_{DR}}$ (for experiment 2)
1	3.143	2.685	0.9
2	0.393	0.373	0.9
3	0.305	0.287	0.9
4	0.417	0.449	0.9
5	1.077	1.268	0.9
6	1.928	1.752	0.1
7	0.164	0.251	0.1
8	1.140	1.324	0.1
9	1.440	1.495	0.1
10	0.680	0.982	0.1

Table 4: Baseline consumptions ($Q^b_{i,t,d_{DR}}$), standard deviations (σ_{i,t,d_c}) and probabilities of participation ($p_{i,t,d_{DR}}$) chosen for timeslot $t = 13$. These parameters are used to run the simulation experiments.

5.3 Approach 1: Deterministic case

In the first aprroach, we assume that each consumer's probability of participation in a DR event is unity, i.e. for each $t \in T_{DR}$, and for each $i \in \{1, 2, ...N\}$, $p_{i,t,d_{DR}} = 1$. The DR design parameters n_t and η^t_{max} which are required by the optimization framework presented in Section 4.1 are set to 3 and 0.25 respectively. It can be easily verified that this choice of these parameters satisfies the feasibility criterion (19) obtained earlier.

5.4 Approach 2: Stochastic case

In the second approach, we assign non-unity probabilities of participation $p_{i,t,d_{DR}}$ in a DR event for each consumer. Since DR participation information was not reported in the data set being used, we assign these probabilities as shown in Tables 4 and 5. With this choice of probabilities, the feasibility condition (19) is not satisfied. To address this, we increase n_t from 3 to 4.

5.5 Approach 3: Rule based strategy

In the third approach, we investigate a rule based strategy (Algorithm 1) for determination of which consumers should

Consumer number	$Q^b_{i,t,d_{DR}}$ (kWh)	σ_{i,t,d_c} (kWh)	$p_{i,t,d_{DR}}$ (for experiment 2)
1	0.675	0.357	0.9
2	0.988	0.626	0.9
3	0.729	0.491	0.9
4	1.060	0.962	0.9
5	3.116	2.261	0.9
6	1.079	1.062	0.1
7	0.388	0.594	0.1
8	1.701	1.304	0.1
9	1.723	0.966	0.1
10	1.275	1.153	0.1

Table 5: Baseline consumptions ($Q^b_{i,t,d_{DR}}$), standard deviations (σ_{i,t,d_c}) and probabilities of participation ($p_{i,t,d_{DR}}$) chosen for timeslot $t = 22$. These parameters are used to run the simulation experiments

be targeted for DR and what DR signals should be sent. In the absence of appropriate ground truth information, we use results obtained from this rule based scheme as the reference to test the value added by the DR planning methodology presented in Section 4. It identifies a set of n_t consumers which achieve the targeted reduction in demand by navigating through the set of consumers in the increasing order of their reduction in utility (steps 2, 3 and 4). Once a set of such consumers is obtained, it then assigns the demand reduction target for each of those consumers in such a way that their percentage reduction over their baseline consumption is the same (step 5). Such an assignment is performed to ensure a 'fair' DR strategy across the target consumers.

Algorithm 1: Rule based strategy for DR planning at timeslot $t \in T_{DR}$

STEP 1: For each consumer i, determine the reduction in utility $\Delta U^{max}_{i,t,d_{DR}}$ and the absolute expected reduction in consumption $\Delta Q^{max}_{i,t,d_{DR}} := p_{i,t,d_{DR}} \eta^t_{max} Q^b_{i,t,d_{DR}}$ from baseline consumption corresponding to a percentage reduction of η^t_{max}.

STEP 2: Sort consumers in ascending order of $\Delta U^{max}_{i,t,d_{DR}}$. We refer to this sorted set by S_t. Set m = 1.

STEP 3: Set the 'current decision set' D_t to consumers $\{m, m + 1, ..., m + n_t - 1\}$ in S_t. Set m = m + 1.

STEP 4: Repeat step 3 until the total expected reduction in consumption corresponding to D_t, $\sum_{i \in D_t} \Delta Q^{max}_{i,t,d_{DR}}$ is greater than or equal to the targeted reduction $\sum_{i=1}^{N} Q^b_{i,t,d} - Q^s_{t,d}$.

STEP 5: The final set of consumers to be targeted is given by D_t. The target reduction for each consumer $i \in D_t$, is then set to: $\Delta Q_{i,t,d_{DR}} = \dfrac{\left(\sum_{j=1}^{N} Q^b_{j,t,d} - Q^s_{t,d} \right) Q^b_{i,t,d_{DR}}}{\sum_{j \in D_t} p_{j,t,d_{DR}} Q^b_{j,t,d_{DR}}}$.

5.6 Results

The decision variables obtained using the three approaches described above are presented and compared in Figures 5 and 6. The consumers selected for DR during timeslots $t = 13$ and $t = 22$, are shown in Figure 5. The corresponding DR signals (targeted consumptions in kWh) are shown in Figures 6.

Due to the choice of the utility function in (25), for the same reduction over baseline consumption, a consumer with
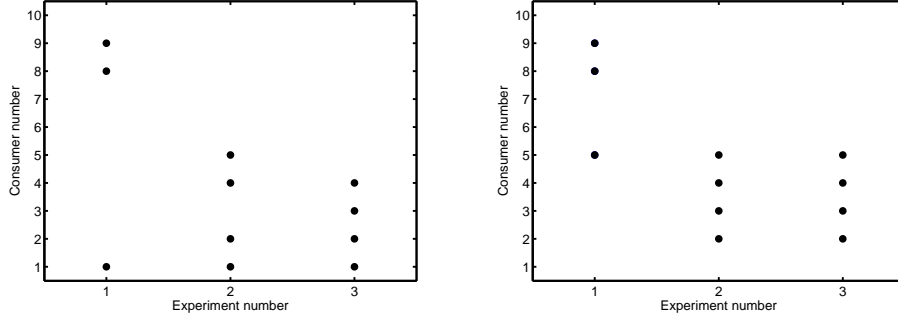
Figure 5: Comparison of consumers selected for DR using three approaches during time slots $t = 13$ (left) and $t = 22$ (right). Note that at $t = 13$, the stochastic case results in certain non-intuitive choice of consumers (e.g. 2 and 4 which have small baselines, but higher participation probabilities.) Similarly, at $t = 22$, the stochastic case results in certain non-intuitive choice of consumers (e.g. 2 and 3 which have small standard deviations)
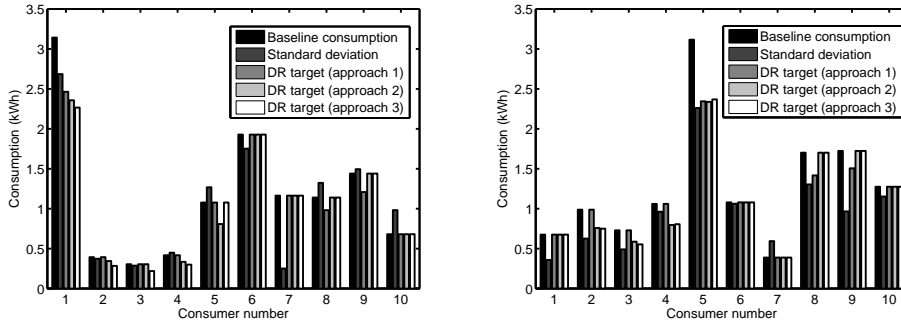

Figure 6: Comparison of DR target consumptions for consumers during time slots $t = 13$ (left) and $t = 22$ (right) using the first three optimization approaches. The baseline and standard deviation based on historical data are also shown for reference. Note that at $t = 13$, while consumers with small baseline and standard deviation (e.g. 2 and 4) are never targeted in the deterministic case (approach 1), they can be targeted in the stochastic case (approach 2) due to their higher participation probabilities.

larger standard deviation, σ_{i,t,d_C} would be subjected to a smaller reduction in utility than a consumer with smaller standard deviation. Also, for the same percentage reduction over baseline consumption, consumers with higher baseline consumption provide higher curtailment of demand than consumers with lower baseline consumption. As a result of these observations, we expect that consumers with large standard deviation and high baseline consumption should be chosen for DR in approach 1. From Figure 5, it can be verified that the optimal choice of consumers for DR indeed conforms to the above intuition.

The intuition stated above also holds for approach 2, after taking into account the probabilities of participation: we expect that consumers with large standard deviation, high baseline consumption and high probability of participation should be targeted for DR. It is difficult to apply this intuition directly, since a set of n_t consumers satisfying all the above 3 criterion does not exist. Therefore, in Figure 5, while we observe that some consumers satisfying these criterion (consumers 1 and 5 at $t = 13$) are selected, the selection of remaining consumers (2 and 4) is non-trivial. This supports the use of a systematic optimization framework - such as the one proposed in this work - since intuition might be limited and inadequate.

Next, we compare the results of approach 1 and approach 2 to the results of the the rule based strategy (approach 3). We first compare the expected reduction in aggregate con-

sumption $\left(\Delta Q_{\text{expected}}^t \right)$ obtained using these approaches, which is computed using the following equation:

$$\Delta Q_{\text{expected}}^t = \sum_{i=1}^{N} p_{i,t,d_{DR}} \Delta Q_{i,t,d_{DR}} \qquad (26)$$

The expected reduction in consumption in each of these approaches during timeslot $t = 13$ is 1.069 kW which exactly matches the desired reduction $\left(\sum_{1=1}^{N} Q_{i,t,d_{DR}}^b - Q_{t,d}^s \right)$. The same observation was made for DR at timeslot $t = 22$. Therefore, we conclude that the consumpting mitigation objective of DR is met in all these approaches. We also quantify the inconvenience caused to consumers in each of these approaches. This is done by computing the expected reduction in utility from the maximum value of 1 for each consumer chosen to participate in DR, and adding it over all such consumers. The inconvenience caused to consumers for each of these approaches and during each DR timeslot, as obtained using this methodology, is shown in Figure 7. We observe that for DR during timeslot $t = 13$, approach 3 results in almost twice as much inconvenience as that caused in approach 2. For timeslot $t = 22$, the inconvenience caused is about 9% more in approach 3 than approach 2. From these results, we conclude that the rule based strategy is at best suboptimal and reinforces the importance of using a systematic optimization framework, such as the one pro-
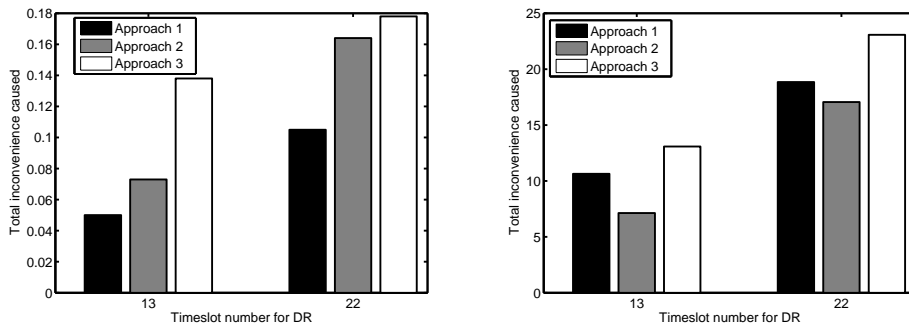
Figure 7: Comparison of inconvenience caused (total reduction in utility over all consumers) using the three approaches, when number of consumers is 10 (left) and when number of consumers is 500 (right). Note that for DR at timeslot $t = 13$, the inconvenience caused using heuristic rule-based DR strategy (approach 3) is almost twice that of iDR (approach 2).

posed, over heuristic rule based strategies for DR planning. Note that we do not perform a comparison of approaches 1 and 3 because they represent two different setups, the latter corresponds to a stochastic setting whereas the former is deterministic.

Lastly, we re-run the afore-mentioned experiments for the larger set of $N = 500$ consumers. The probabilities $p_{i,t,d_{DR}}$ were assigned arbitrarily, and we assume that the timeslots corresponding to DR are still $t = 13$ and $t = 22$, when the utility's desired supply power $Q^s_{t,d_{DR}}$ falls short of the expected demand $\sum_{i=1}^{N} Q^b_{i,t,DR}$ by 10%. The DR design variables η^t_{max} and n_t are set to 0.25 and 167 respectively. The observations made were similar to those reported above for 10 consumers. The targeted DR reduction was achieved in all these approaches and the inconvenience comparison is presented in Figure 7.

5.7 Fairness assessment

The objective of this experiment is to evaluate the fairness strategy proposed in (18). For simplicity, we create 10 identical clones of consumer 1 used in the previous three experiments, resulting in a group of consumers with similar consumption preferences. The parameters n_t in (10) and $N^{max}_{DR,i}$ in (18) are both set to 5, and the timeslot $t = 13$ is chosen for simulations.

Figure 8 shows the number of times each consumer is targeted, as DR iterations progress. It is observed that after the end of a sufficiently large number of iterations (in this case 10), all consumers have been chosen an equal number of times. This verifies that the proposed DR planning strategy is 'fair' across all consumers in the group.

6. DISCUSSIONS AND FUTURE WORK

In this paper, we presented iDR, a planning methodology that helps electricity providers in designing effective demand response events. In summary, iDR helps answer three important questions: (i) when to plan DR events, (ii) which consumers to target, and (iii) what signals to send. The proposed approach involved the estimation of baseline consumptions using historical consumption data obtained from smart meters, and then the development of utility functions using smart meter data to gauge the inconvenience associated with consumers shifting their demand from baselines. These utility functions were then used in the solution of a chance constrained, mixed integer nonlinear program. The approach was tested in simulation using data from a public
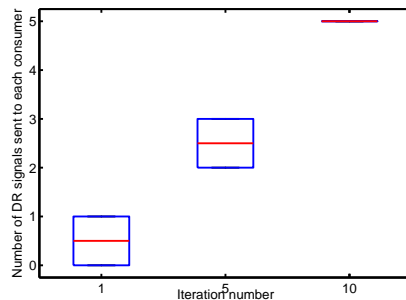

Figure 8: Progression of consumer selection as a function of DR iterations, shown as a box plot. Note that each consumer has been selected equal number of times at the end of 10 iterations, ensuring fairness

data source. Results indicated that the DR objectives were achieved, and iDR resulted in lower inconvenience than a heuristic rule based approach.

Based on the work presented in this paper, we identify the following avenues of future research involving iDR. The optimization framework presented in Section 4 and evaluated in Section 5 corresponds to the utility mining framework based on aggregate consumption (Section 3.1). An important area of investigation in future is the development of a DR planning methodology, analogous to that presented in Section 4, but based on appliance level utility functions (Section 3.2), if appliance level consumption information is available. Comparison of the results of these two optimization frameworks is expected to provide deeper insights towards establishing optimal consumption schedules for effective DR planning. A Gaussian utility function was used in the simulation experiments reported in Section 5. As an alternative, empirically derived utility functions from the consumption dataset can also be implemented. However, the potentially additional computational overhead introduced would need to be balanced against the value added by the use of such an empirical, purely data driven approach. The relationship between DR participation probabilities and incentives also needs detailed investigation. This can enable an augmentation of the existing capabilities of iDR. For example, it can enable the development of appropriate incentives in order to alter the participation probabilities of a set of consumers, such as those with poor DR participation history. In general, the

use of participation probabilities in the framework provides a handle for the electricity providers to achieve other DR objectives besides demand mitigation, such as the development of effective DR contracts. The mitigation of unwanted outcomes such as rebound effects [10] can be included in the optimization framework to increase its practical value. Also, extensions of the proposed approach to plan DR for consumer groups instead of individual consumers can result in potential advantages such as a reduction in computational overhead. Previous work in the area of consumer segmentation such as [16] can be used in this context. Experimental results were not reported in this paper. However, we identify the real-world testing of *iDR* through pilot studies to gauge its true potential as a DR planning tool as an important final step in its development. Such pilot studies are being planned as a part of the European Union funded WATTA-LYST project [25], which would involve user studies to test the accuracy of the utility models developed in Section 3, effect of consumer participation in DR execution, as well as the efficacy of the DR planning methodology proposed in Section 4. Additionally, we plan to test *iDR* in simulation on a larger dataset than the one currently used, for a detailed scalability analysis.

7. REFERENCES

[1] N. Armaroli and V. Balzani, "The future of energy supply: challenges and opportunities," *Angewandte Chemie International Edition*, vol. 46, no. 1-2, pp. 52–66, 2007.

[2] U.S. Department of Energy, "Benefits of demand response in electricity markets and recommendations for achieving them," http://eetd.lbl.gov/ea/emp/reports/congress-1252d.pdf, 2006.

[3] K. Spees and L. Lave, "Impacts of responsive load in pjm: Load shifting and real time pricing," *The Energy Journal*, vol. 29, no. 2, pp. 101–122, 2008.

[4] Federal Energy Regulatory Commission, "A National Assessment of Demand Response Potential," http://www.ferc.gov/legal/staff-reports/06-09-demand-response.pdf, 2009.

[5] A. Chardon, O. Almen, P. E. Lewis, J. Stromback, and B. Chateau, "Demand Response: a decisive breakthrough for Europe. How Europe could save Gigawatts, Billions of Euros and Millions of tons of CO2." Capgemini, Tech. Rep., June 2008.

[6] Smart Metering Project in the United Kingdom, "Smart Metering Projects Map," http://bit.ly/GSs1o5, 2013.

[7] Peak Load Management Alliance (PLMA), "Demand Response: Design Principles for Creating Customer and Market Value," http://www.peaklma.com, November 2002.

[8] US Department of Energy, "Benefits of demand response in electricity markets and recommendations for achieving them," Feb. 2006.

[9] F. Yoo and C. Riker, "PG&E Automated Demand Response Program: IDSM Lessons Learned," http://energyefficiencysmartgrid.org/presentations.

[10] D. P. Seetharam, A. Agrawal, T. Ganu, J. Hazra, V. Rajaraman, and R. Kunnath, "Hidden costs of power cuts and battery backups," in *Proceedings of the fourth international conference on Future energy systems.* ACM, 2013, pp. 39–50.

[11] T. Bapat, N. Sengupta, S. K. Ghai, V. Arya, Y. B. Shrinivasan, and D. Seetharam, "User-sensitive scheduling of home appliances," in *Proceedings of the 2nd ACM SIGCOMM workshop on Green networking.* ACM, 2011, pp. 43–48.

[12] A. Agnetis, G. de Pascale, P. Detti, and A. Vicino, "Load scheduling for household energy consumption optimization," *IEEE Transactions on Smart Grid*, vol. 4, no. 4, pp. 2364–2373, Dec 2013.

[13] N. Li, L. Chen, and S. H. Low, "Optimal demand response based on utility maximization in power networks," in *2011 Power and Energy Society General Meeting.* IEEE, 2011, pp. 1–8.

[14] M. Kanakakis, M. Minou, C. Courcoubetis, G. D. Stamoulis, and G. Thanos, "A practical iterative price-based approach for optimal demand-response," *in proceedings of The Third International Conference on Smart Grids, Green Communications and IT Energy-aware Technologies*, 2013.

[15] A. Albert, R. Rajagopal, and R. Sevlian, "Segmenting consumers using smart meter data," in *Proceedings of the Third ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings.* ACM, 2011, pp. 49–50.

[16] T. K. Wijaya, T. Ganu, D. Chakraborty, K. Aberer, and D. P. Seetharam, "Consumer segmentation and knowledge extraction from smart meter and survey data," in *SIAM International Conference on Data Mining (SDM14)*, 2014.

[17] K. Inc., "Pjm empirical analysis of demand response baseline methods," *White Paper*, 2011.

[18] EnerNOC, "The demand response baseline," *White Paper*, 2011.

[19] W. Charytoniuk, M. Chen, and P. Van Olinda, "Nonparametric regression based short-term load forecasting," *IEEE Transactions on Power Systems*, vol. 13, no. 3, pp. 725–730, 1998.

[20] 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, 2013, pp. 51–62.

[21] CER, "Electricity customer behaviour trial," http://www.ucd.ie/issda/data/commissionforenergyregulationcer/, 2012.

[22] "Enernoc: Open data to facilitate research and development," http://open.enernoc.com/data/, 2013.

[23] S. Le Digabel, "Algorithm 909: Nomad: Nonlinear optimization with the mads algorithm," *ACM Transactions on Mathematical Software (TOMS)*, vol. 37, no. 4, p. 44, 2011.

[24] Opti Toolbox - A free MATLAB toolbox for optimization, http://www.i2c2.aut.ac.nz/Wiki/OPTI/.

[25] "WATTALYST," http://www.wattalyst.org.