Information Infrastructure for Cellular Load Management in Green Power Delivery Systems

Mahnoosh Alizadeh*, Anna Scaglione*, Robert J. Thomas† and Duncan Callaway‡
*University of California, Davis †Cornell University ‡University of California, Berkeley

Abstract—**In this paper, we outline a specific communication framework to support Direct Load Control programs on the edge of the power distribution network. We suggest a model where the arrival process of the smart appliances is made visible to control centers, forming a cellular Microgrid infrastructure, through Home Energy Management Systems (HEMS). These appliances then wait to receive an authorization message before starting to function. The cell control center uses this information to choose an optimal departure process that serves the waiting appliances while minimizing its operational costs. The described information exchange strategy gives the cell the ability to better match the load to the available green energy supply and its day ahead energy bid. We show that this model will allow to increase the integration of intermittent resources in the power grid, with modest communication rate requirements.**[1]

I. Introduction

Currently, power system operators have little control over the dynamic behaviour of the load. In most areas, the only possible load modification is load shedding, i.e. an intentional power outage so that power can be kept in balance in the grid. In monitoring and predicting the load, the current approach is sensing the bundled request and accruing statistics to forecast the future evolution of the demand. With the addition of appliances like Plug-in (Hybrid) Electric Vehicles (PHEV), Smart cooling/heating systems, and the introduction of active loads like discharging EVs or Community Energy Storage (CES), serving as a buffer or a source of energy, volatility in the power distribution network can be greatly increased [1] and thus, traditional prediction techniques will degrade considerably in performance. Utilities have also been aware of this fact and they have tried to find ways of matching consumption with generation by making the load observable, thanks to Smart meters, whose deployment is referred to as the Advanced Metering Infrastructure (AMI) [2].

Even with full observability of this volatile load, the current tradition of having no control over the load's behaviour is likely to increase the need for operating reserves for the grid to function reliably and will increase the price of electricity. What is envisioned in the so called Smart Grid project is that certain *smart* devices like electric vehicles can become flexible agents distributed across the grid whose demand can be modified, either in a centralized or a distributed manner, to integrate the volatile local renewable generation sources in the system on the edge of the distribution network. One approach that has been proposed to handle this problem is using Direct Load Control programs [3], [4], which aim to control part of the load due to smart appliances directly from a central unit. For a control center to be able to remotely modify the load, the smart appliances should be able to communicate with the control center. But how? What is the data that needs to be communicated? What is a suitable structure to formulate the Load Management problem?

We address these questions in this paper. Specifically, the paper is organized as follows: We will first give an overview of the related works on load management. In Section II, we provide a system model that is the basis to determine the information that is required to reconstruct a portion of the load in a control center. We investigate a data representation methodology that can support this infrastructure in Section III. After we calculate the total amount of delay experienced by customers participating in a DLC program in a given interval, in Section, IV-B, we will give a general definition for the type of optimization required to solve the centralized scheduling problem. We will discuss possible data gathering and distribution strategies in section V and finally, in section VI, we will give numerical results on how a DLC program can help match a volatile load to the available generation supply.

A. Previous Work

Previous work on load management techniques can be mainly categorized into one of the following:

- Price-based load control
- Load control through curtailment and scheduling

Price-based load control strategies include time-of use (TOU) or real-time (RT) pricing techniques. In TOU pricing strategies, the price data is usually decided months or years before the actual time of use. There have been several studies on determining these rates, which requires a dynamic model for the price response of customers, typically derived based on experiments [5], [6]. On the other hand, in real time pricing scenarios, the price data is provided only hours before consumption. TOU and RT pricing have been made possible through the deployment of the AMI. In both cases, each customer decides on their energy use pattern individually given the pricing data. In the case of RT pricing, the need for an automated system to help the customer in making these decisions seems apparent. There is, in fact, an extensive literature emerging on Home Energy Management Systems (HEMS) [7] [8]. For these systems to operate, the price data is delivered to the customer and an automation device installed inside the home will plan the use of appliances given their

[1]This work was funded by the UIUC TCIPG Project sponsored by DOE under the Award DE-OE0000097.

power consumption, price data and job deadlines. While these strategies can alleviate the situation, we think that central load control may better solve the pressing problem of managing resources , because of the economy of scale and increasing flexibility that pooling together resources can warrant.

The second group of load control strategies are usually applied directly by a control center but require customer participation. They first emerged in the 90's and are currently employed through the so-called *Interruptible Load* programs where, upon receiving a notice, the customers turn off some of their appliances for a pre-determined amount of time (15-30 minutes). In these programs, a constrained optimization problem is solved to determine a load curtailment strategy [9], [10]. In recent years, the deployment of advanced two-way communication links between the customers and the utility has attracted attention towards *Direct Load Control* (DLC) programs that allow a control center to directly control end-use appliances. For example, a few papers have attempted in the past 2-3 years to deal with the problem of centralized electric vehicle scheduling [11]. This focus on EVs [12], [13] may be due to the fact that unlike many appliances already used in a commercial scale by end-use customers, electric vehicles are yet in their introductory stages and do not yet have fixed standards. On the other hand, when commercialized, these appliances will play a huge role on the demand side of the market and if managed optimally, they can be put to great use due to the inherent flexibility in their time of use.

Here, we propose a traffic management model for centralized load control in a Smart Grid. The smart loads communicate with their HEMS and their information is forwarded to the Community Energy Management System (CEMS), which takes on the tasks of scheduling them. A graphical representation of the communication infrastructure, the decision blocks and the load reconstruction is given in Fig. 1.

II. Modeling the Traffic

We assume that H different types of appliances arrive in the system following non-stationary Poisson arrival processes

$$a_j(t) = \sum_{i=1}^{\infty} u(t - t^a_{i,j}), \quad j = 1, \dots, H \tag{II.1}$$

with $u(t)$ the unit step, $t^a_{i,j}$ the arrival time the i^{th} appliance of type j, and an arrival rate of $\lambda_j(t)$ for each type of appliance. Each arrival event has an associated dimensionless parameter set $C_{i,j}$ that describes its energy request. These parameters are assumed to be i.i.d. random variables with a known stationary distribution $f^j_C(c)$, independent of the arrival times. Each arrival event is modeled by a tuple $(j, t^a_{i,j}, C_{i,j})$. The uplink information to the cell control is this tuple; upon its reception, the cell control can map the appliance type and energy request code $(j, C_{i,j})$ one to one with a complex load phasor $g_j(t; C_{i,j})$ (in volt-ampere), representing the evolution of active and reactive power for a type-j load turned on at time zero. For appliances like EVs $C_{i,j}$ can be a scalar representing the number of time units the battery charge lasts, and can be used to scale the argument of a real pulse, which approximates well the car charging profile, $g(t; C_{i,j}) \approx g(t/C_{i,j})$; but, in general, $C_{i,j}$ can be quantized vector of Fourier or Wavelet coefficients, used as a basis expansion to represent the the known load evolution after activation. To explain more important concepts, we look at the simple case of EVs, and a single scalar parameter $C_{i,j}$ in $g(t/C_{i,j})$.

The ability to schedule the loads means that the control center has the authority to delay the starting time for each smart appliance turning on. When the appliance is authorized to turn on, it is said to *depart*.[2] Hence, these appliances form a *departure process*, where $t^d_{i,j}$ is the time instant when the i^{th} appliance of the j^{th} type is authorized to start functioning:

$$d_j(t) = \sum_{i=1}^{\infty} u(t - t^d_{i,j}), \quad t^d_{i,j} \geq t^a_{i,j}. \tag{II.2}$$

Thus, the future load to satisfy can then be decomposed as:

$$L(t) = L^N(t) + L^S(t) \tag{II.3}$$

where $L^N(t)$ represents the base load which we presumably have no control over, while $L^S(t)$ is the *controllable part of the load* due to the smart appliances. For the rest of the paper we consider $L^N(t)$ highly predictable, using standard load forecasting techniques (see e.g. [14], [15]). In addition, it is convenient to include in $L^N(t)$ the load due to previously scheduled smart appliances, since we assume that the functioning of an appliance cannot be interrupted once started.

The smart load, $L^S(t)$ can now be written as the following function of the departure process

$$L^S(t) = \sum_{j=1}^{H} \sum_{i \in I_j} g_j \left(\frac{t - t^d_{i,j}}{C_{i,j}} \right) \quad , t^d_{i,j} \geq t^a_{i,j} \tag{II.4}$$

where the set I_j includes all appliances that require scheduling in the j^{th} arrival process, i.e. it includes the appliances of the j^{th} type that have already arrived in the system at time t but have not yet been authorized to function or the ones that are expected to arrive in the future.

III. Load Reconstruction

We saw that the information required to model the future load of the system in the control center includes visibility of the arrival process ($t^a_{i,j}$), the type of the arriving appliances, and their energy request code, $C_{i,j}$. To collect this information, we will need a data gathering methodology. To find a suitable one, we first have to look at how we are going to use this communicated information.

To avoid large storage and computational costs, we quantize the charging codes $C_{i,j}$ of appliances, which for scalar codes, like the battery charge for EVs, will correspond to

$$Q(C_{i,j}) \in \{C_{q,j}, q = 1, \dots, Q_j\} \tag{III.1}$$

The quantization is also a tool to separate the appliances in classes of service; we assume that arriving appliances of

[2]The actual finishing time of the job is a one to to one mapping from this departure time, depending on the queue to which the appliance belongs.

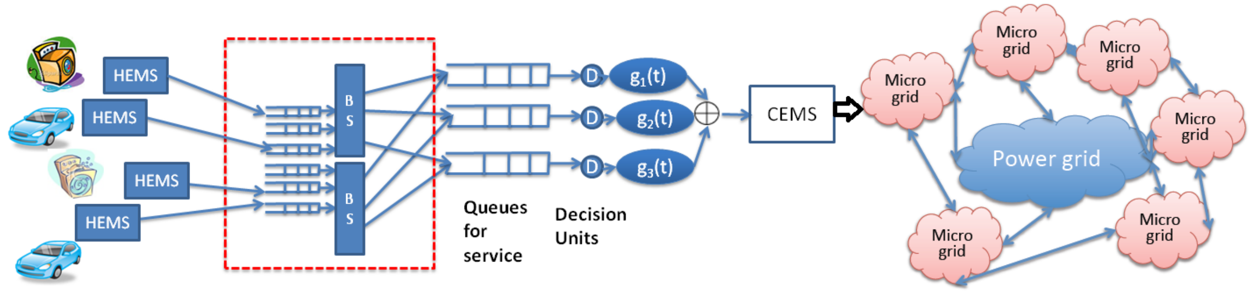

Fig. 1. Communication and Control Architecture

the j^{th} type are put in Q_j separate queues depending on their charging code. The statistics of each queue (rate of arrival) can be obtained from $\lambda_j(t)$ and $f_C^j(c)$. Assuming that we round up to the next allowable quantized value and $C_{0,j} = 0$:

$$\lambda_{q,j}(t) = [f_C^j(C_{q,j}) - f_C^j(C_{q-1,j})]\lambda_j(t), \quad q = 1, \dots, Q_j$$

From this point on, we use the variable $C_{q,j}$ to denote the job duration associated with the q-th queue of the the type-j appliances. The number of quantization levels Q_j can be chosen to satisfy one of these two performance metrics:

•Minimize the Q_j's while satisfying a desired maximum distortion W in load reconstruction, i.e.

$$\min_{Q_j} \quad \sum_{j=1}^{H} Q_j, \quad \text{s.t.} \ \sum_{j=1}^{H} \ E\left(W_j\right) < W, \qquad \text{(III.2)}$$

where $E(W_j)$ is the largest expected distortion bt the j-th type

$$\sum_{q=1}^{Q_j} \lambda_{q,j}^{\max} \int_{t=0}^{\infty} \int_{x=C_{q-1,j}}^{C_{q,j}} \left| g\left(\frac{t}{x}\right) - g\left(\frac{t}{C_{q,j}}\right) \right| f_C^j(x) dt dx$$

due to quantization errors. $\lambda_{q,j}^{\max}$ denotes the maximum of the quantity over time. The $C_{q,j}$'s will be determined uniquely from Q_j and the range of $C_{i,j}$ if a uniform quantizer that rounds up to the nearest level is assumed.[3]

•Under a maximum bit rate constraint, minimize the distortion,

$$\min_{Q_j} \sum_{j=1}^{H} E[W_j(t)]$$

$$\text{s.t.} \ \max_{\lambda} R_{HEMS\text{-}MAC} < R_1^{max}, \ \max_{\lambda} R_{BS\text{-}CEMS} < R_2^{max}$$

where the functions $R_{HEMS\text{-}MAC}$ and $R_{BS\text{-}CEMS}$ are calculated in section V. It is clear that the optimum in both cases is reached when the constraints are tight.

Since arrivals in different queues are independent processes, the arrival process $a_j(t)$ can be divided into Q_j separate processes whose state can be represented by a vector $\bar{a}_j(t)$ of length Q_j with the property that $\|\bar{a}_j(t)\|_1 = a_j(t)$. The corresponding departure processes from each queue can also be represented by a vector $\bar{d}_j(t)$, also satisfying $\|\bar{d}_j(t)\|_1 = d_j(t)$. We also know that $\bar{a}_j(t) \succeq \bar{d}_j(t)$, where $\succeq$ represents element by element inequality. This is due to the simple fact that the number of departures from each queue can never be larger than the number of arrivals.

[3]For the moment we assume a uniform quantization, but in future work we will examine the impact that different quantization can have on the optimization.

The quantization of the charge durations in Q_j levels allows a simple system representation of the relationship between the individual queue departure processes and the total smart load. In fact, since the $C_{q,j}$'s are discrete values, we can rewrite the load due smart appliances as

$$L^S(t) = \sum_{j=1}^{H} \sum_{q=1}^{Q_j} \frac{\partial}{\partial t} [\bar{d}_j(t)]_q \star g_j\left(\frac{t}{C_{q,j}}\right) \qquad \text{(III.3)}$$

where the time derivative $\frac{\partial}{\partial t}[\bar{d}_j(t)]_q$ will produce a Dirac delta each time an arrival occurs in the q^{th} queue of the j^{th} appliance type and $\star$ represents the convolution operation.

Combinatorial complexity is an inherent feature of scheduling problems. To be able to work online, a real time solution methodology should reduce the problem size. Since we are unable to communicate the arrival time with infinite precision, it is reasonable to use a predetermined set of discrete time decision intervals to reduce this complexity,

$$t_{i,j}^d \in \{l\triangle | l \in \mathrm{N}\} \qquad \text{(III.4)}$$

Thus, the departure process $\bar{d}(t)$ can only have increments at these discrete set of time instants. Due to this, an alternate representation of the smart load for $t \geq l_0\triangle = \lceil \frac{t}{\triangle} \rceil \triangle$ is,

$$L^S(t) = \sum_{j=1}^{H} \sum_{q=1}^{Q_j} \sum_{l=l_0}^{\infty} [\bar{d}_j(l\triangle) - \bar{d}_j((l-1)\triangle)]_q g_j\left(\frac{t - l\triangle}{C_{q,j}}\right) \qquad \text{(III.5)}$$

Assuming that the charging pulse $g(t)$ can be considered constant during intervals of length $\triangle$, we can write the load as a function of a set of discrete variables. Hence, from this point on, we discretize the arrival and departure processes and replace the previously defined quantities with their discrete counterparts. For example,

$$\begin{array}{ll} [\bar{a}_j(l\triangle)]_q \to a_{q,j}(l), & [\bar{d}_j(l\triangle)]_q \to d_{q,j}(l) \\ g_j(\frac{l\triangle}{C_{q,j}}) \to g_{q,j}(l), & L^S(u\triangle) \to L^S(u) \end{array} \qquad \text{(III.6)}$$

Which leads to the following representation of the load in terms of the decisions $d_{q,j}(l)$ and samples of the pulses $g_j(t)$

$$L^S(u) = \sum_{j=1}^{H} \sum_{q=1}^{Q_j} \sum_{l=l_0}^{\infty} [d_{q,j}(l) - d_{q,j}(l-1)] g_{q,j}(u-l) \quad , u \geq l_0 \qquad \text{(III.7)}$$

Before further clarifying the communication support infrastructure in Section V, we discuss the general formulation of the cellular energy control optimization in the next section.

IV. Optimization Problem

A. Modeling the Delay

To formulate an optimization problem that schedules smart appliances, a quantity that requires modeling is the delay experienced by these appliances. It is a well-known result in traffic flow theory that the total delay experienced by the customers in a queue, i.e. $\sum_i(t_i^d - t_i^a)$ is equal to the area of the queue polygon, which is a function that represents the state of the queue $s(t)$ versus time. It is obtained by superimposing the departure and arrival profiles, $a(t)$ and $d(t)$:

$$s(t) = a(t) - d(t) \tag{IV.1}$$

The delay experienced in the past by the appliances currently present in the system is not amendable and so, it is not of any interest in the formulation of the optimization. As a result, for the purpose of optimal scheduling in the future, one can replace the total delay cost with the total delay experienced by all the customers in the future, i.e. what we call the *Delay Cost Increment* (DCI) for the appliances of the j^{th} type:

$$\mathrm{DCI}_j(t) \triangleq C_D^j \int_t^\infty s_j(\tau)d\tau = C_D^j \int_t^\infty (a_j(\tau) - d_j(\tau))d\tau$$

where C_D^j is the cost per unit of time delay for the j^{th} type. If the delay cost C_D^j is a function of time, we can write

$$\mathrm{DCI}_j(t) = \int_t^\infty C_D^j(\tau)(a_j(\tau) - d_j(\tau))d\tau \tag{IV.2}$$

After quantizing the $C_{i,j}$'s , the DCI defined in (IV.2) can be rewritten in following alternative ways,

$$C_D^j \int_t^\infty (\|\bar{a_j}(\tau)\|_1 - \|\bar{d_j}(\tau)\|_1)d\tau$$

One should note that when using a discrete set of decision epochs, we cannot change the delay experienced by the vehicles before $l_0\triangle$. Applying this fact to (IV.3), we can replace the delay cost increment at time t with

$$\begin{aligned}\mathrm{DCI}_j(l_0\triangle) &= C_D^j \int_{l_0\triangle}^\infty \|\bar{a_j}(\tau) - \bar{d_j}(\tau)\|_1 d\tau \qquad \text{(IV.3)}\\ &= C_D^j \left\| \int_{l_0\triangle}^\infty [\bar{a_j}(\tau) - \bar{d_j}(\tau)]d\tau \right\|_1 \\ &= C_D^j \left\| \sum_{l=l_0}^\infty \int_{l\triangle}^{(l+1)\triangle} \bar{a_j}(\tau)d\tau - \bar{d_j}(l\triangle)\triangle \right\|_1\end{aligned}$$

Since the delay experienced by the customers between their arrival time until the next possible decision epoch is also not amendable, we can also remove the associated penalty from our cost and thus, reformulate the delay cost as,

$$\mathrm{DCI}_j(l_0\triangle) = C_D^j \left\| \sum_{l=l_0}^\infty [\bar{a_j}(l\triangle)\triangle - \bar{d_j}(l\triangle)\triangle] \right\|_1 \tag{IV.4}$$

Note that the interchangeability of integration (or sum) and the $\mathcal{L}_1$ norm is possible since $\bar{a_j}(t) - \bar{d_j}(t) \succeq 0$.

Now, we will look into ways of using this new information exchange strategy to help the grid operate more reliably.

B. General Formulation of Scheduling Problem

Assuming that consumers participate in a DLC program by subscribing to the cellular load management service, we now want to formulate a general scheduling problem for these entities. The operational costs of these control center include:

1) Wholesale market day-ahead bidding cost: by looking on the day ahead forecasts of its local generation units and its load pattern, the control center participates in the day ahead wholesale market and purchases a certain amount of power for every hour of the next day so that it can safely serve all of its load reliably.
2) Wholesale market real-time bidding cost: if the control center cannot serve its customers with its local generation and its day ahead bid, it buys more energy from the central grid in a spot market.

Other costs include the cost of purchasing energy from the local market for distributed generation and the inconvenience cost paid to customers as an incentive for participating in a DLC program, which we assume to be proportional to the amount of delay they experience. On the other hand, the utilities of the control center are from selling electricity to customers, carbon taxing utility and the wholesale market ancillary service utility, if any such services are provided.

Thus, the control center minimizes the following objective function to find the optimum schedule for smart appliances,

$$\begin{aligned}&\min_D \; E[\text{Cost of retail entity in real time}] = \\ &\min_D \; E\{\sum_{l=l_0}^{l_0+T}[C(L^N(l)+L^S(l),l) - U(L^N(l)+L^S(l),l)] \\ &\qquad + \sum_{j=1}^{H} \mathrm{DCI}_j(l_0)\} \\ &\text{s.t.} \quad d_{q,j}(l-1) \le d_{q,j}(l) \le a_{q,j}(l) \\ &\qquad d_{q,j}(l) \in \mathbb{Z}_+ \\ &\qquad d_{q,j}(T) = a_{q,j}(T) \end{aligned} \tag{IV.5}$$

Where D is the entire decision space, i.e. $D = \{d_{q,j}(l), l \geq l_0\}$ and T is the look-ahead horizon. U is the utility function associated with selling energy equal to $L^N(l) + L^S(l)$ in the retail electricity market to end-use customers. If dynamic retail pricing strategies are implemented, U will be a function of time. C is the cost that the retail utility incurs in at time l when purchasing electricity equal to $L^N(l) + L^S(l)$ from a central wholesale market or the local intermittent resources. DCI(l_0) is the delay cost increment defined in section IV-A, calculated in a finite horizon T. The first and second constraints are due to causality and the third constraint requires that no arriving appliance is delayed beyond time $l_0 + T$.

V. Information Gathering and Distribution

A. The Cell Uplink

In the previous sections, we saw that only the number of customers arriving in each queue between each two decision epochs is what is required to determine an optimal departure process from these queues and, optimally modify the smart part of the load $L^S(t)$. So, the control center has to have an information gathering strategy that avoids unnecessary communication. It only needs to gather the value of the vectors $\bar{a}_j(l\Delta)$ that aggregate the arrivals in the queues of the j^{th} appliance type during the interval $[l-1,l)\Delta$, for real-time applications. Each HEMS can locally compute the accrued inconvenience cost and communicate it offline to track the quality of service delivered to each individual home, which means that the data communicated for real time scheduling can be both anonymized, as well as aggregated as they flow towards the CEMS. Hence, the identification of the HEMS system, and its authentication are not functions that need to be performed in real-time and do not therefore correspond to the traffic with delay constraints. Assuming that there are H different types of appliances, the time sensitive bits correspond to the digital communication of the tuple $(j, t^a_{i,j}, Q(C_{i,j}))$, which requires clearly $\log(H)$ to specify the type j and $\log Q_j$ bits for the charging class. The arrival time $t^a_{i,j}$ here represents an index that multiplies a discrete time interval equal to the scheduling resolution Δ.

Let $t^n_{i,j} > t^a_{i,j}$ be the *notification* time index of the arrival, i.e. the time interval $k\Delta$ when the event is first recorded and added to the corresponding queue $a_{i,j}(t^a_{i,j}\Delta)$. Assuming that

$$t^n_{i,j} - t^a_{i,j} < D,$$

where D is a maximum network delay, the arrival time can be encoded modulo D, using the notification time as side information, and computing the arrival time as $t^a_{i,j} = t^n_{i,j} - [t^n_{i,j}]_D + [t^a_{i,j}]_D$. In this case, clearly, encoding $t^a_{i,j}$ only requires $\log_2 D$ bits. Hence, considering that $\lambda_j(l\Delta)$ is the traffic of type-j loads arriving in the system, the HEMS access channel needs to support an aggregate traffic of

$$R_{HEMS-MAC}(l) = \frac{1}{\Delta}\sum_{j=1}^{H}\lambda_j(l\Delta)\log_2(HDQ_j). \quad \text{(V.1)}$$

As discussed before, the traffic can be aggregated at the first network relay, acting as a base station (BS); for example a BS could map one to one with each area transformer or to a ISP node, coalescing the arrival times into information about the arrival vector $\bar{a}_j(l\Delta)$. Assuming that each component can be approximated with a Poisson r.v., the communication rate of the aggregate arrival vector is bounded by:

$$R_{BS-CEMS}(l) = \sum_{j=1}^{H} Q_j \frac{1}{2}\log(2\pi e\lambda_j(l\Delta)). \quad \text{(V.2)}$$

Next, we discuss a messaging strategy for the downlink feedback.

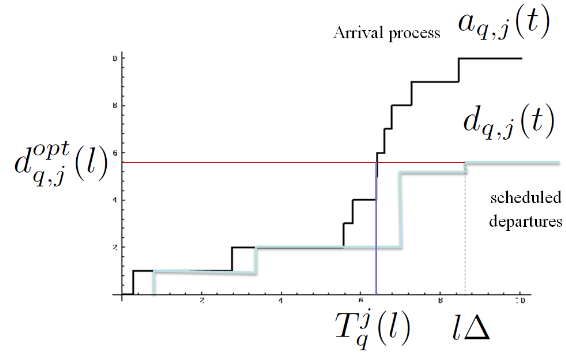

Fig. 2. Mapping decisions into feedback messages

B. Communicating the decisions back

Consistently with our uplink model, we envision a downlink message structure that preserves the anonymity of the scheduled user. Once the CEMS decides the optimum schedule, it sends a record with H feedback messages, one for each type, to let the vector $\bar{d}^{opt}_j(l)$ of appliances in: the j-th feedback consists of a $Q_j \times 1$ vector $\bar{T}^j(l)$, which alerts all appliances of type j in the corresponding classes $q = 1, \ldots, Q_j$ that arrived before time $T^j_q(l)$, to enter the system.

The calculation of these vectors is performed as indicated in Fig. 2 and summarized in the following equation:

$$T^j_q(l) = \max\{\tau \le l : a_{q,j}(\tau) \le d^{opt}_{q,j}(l)\} \quad \text{(V.3)}$$

This system makes sure that the departures match the desired value, while guaranteeing anonymity of the access requests. Also, it is not necessary to transmit absolute times $T^j_q(l)$. Considering that the delay is an explicit cost for the optimization, and that the optimum decisions are correlated since the queue states are correlated, $T^j_q(l)$ can be differentially encoded with a relatively modest rate requirement. More specifically, assuming that ρ_j is the minimum correlation coefficient among the decisions for appliances of type j in any class and at any time, and that the variance of the delay for appliances in class j is bounded by χ_j, then each feedback vector can be encoded in $Q_j \frac{1}{2}\log_2(2\pi e(1-\rho_j^2)\chi_j)$ bits every Δ.

VI. Numerical Analysis

A. Traffic Estimation in a Real Scenario

In today's development of the AMI, the BS can be associated with the gateways installed in small neighborhoods. The communication links between these gateways and the CEMS support very high rates (T1 lines, DSL, Wimax, etc.) and thus, the communication of the arrival and decision messages is well provisioned. The links that may not have a bandwidth as large as these links are the ones connecting the smart meters to the gateways, or the Local area network, which is based on either RF or PLC. In this section, we calculate the number of bits each arrival message may contain in a real deployment of the DLC technique introduced in this paper. To choose a suitable Q for the EV case, the constraint in III.2 can be simplified if we assume that g is a rectangular pulse. Then, each EV will cause a false load in the optimization for a maximum time duration of $C_{i,j}^{\max}/Q_j$. Now, let us consider a realistic scenario: Assume that a Microgrid is handling the charge scheduling of 500 vehicles and they all arrive in a span of 5 hours (4-9 pm) uniformly. If the CEMS wants to make sure no more than 10 EVs are falsely assumed to be charging at any time, then the resolution required to quantize $C_{i,j}$ is around 5 minutes. If we assume that $C_{i,j}^{\max} = 8$ hours, then $Q_j = 96$, which requires 6 bits to communicate. Assuming that the delay of the AMI network is 15 minutes, then the arrival time will also require 4 bits to be communicated. Also, if 4 types of smart appliances are present in the optimization, 2 more bits are needed to communicate the type. This sums up to 12 bits per arrival message, which is a very modest rate. This shows that the main concern is therefore coverage and reliability rather than high speed connectivity.

B. Load scheduling

In this section, we will present numerical results for a simple test situation to provide insight on how direct load management of appliances can improve the ability of a control center to match its available generation supply. Due to lack of space, we are unable to present the underlying reasoning that leads from the operational costs in Section IV-B to the selected forms for functions U and C in IV.5, which are the following:

$$\begin{aligned} C(L^N(l) + L^S(l), l) &= C_{dv}(l)|P(l) - (L^N(l) + L^S(l))| \\ U(L^N(l) + L^S(l), l) &= C_{rt}(l)(L^N(l) + L^S(l)) \end{aligned} \tag{VI.1}$$

where we will assume to have perfect knowledge of the available generation $P(l)$, the retail prices $C_{rt}(l)$ and the deviation penalties $C_{dv}(l)$ in the look-ahead horizon.

The simulated scenario is as follows: we assume that only a single type of smart appliance is present in the system with a minimum and maximum job size of 1 and 4 units of time. The appliances arrive in the system following a Poisson arrival process with a mean arrival rate $\lambda = 3$ for each job size, assumed to be constant for simplicity and also, known a-priori. The optimizer solves the problem with a one-step lookahead rollout on a certainty equivalent controller that uses linear programming to determine the best scheduling strategy. After each step, the scheduled loads are added to the uncontrollable load term for future epochs. Fig. 3 compares the uncontrolled load profile with the profile resulting from the DLC strategy. The results show a 40 percent reduction in deviation from the generation profile $P(l)$ in a horizon of 50 units of time. The look-ahead horizon in these simulations is assumed to be 5 units of time and to be fair, no appliance is allowed to be delayed beyond t = 50. So, the number of appliances receiving service is equal in both load profiles. The results also show that the average delay experienced by customers in this program is less than one unit of time.

VII. Conclusions

In this paper, we first introduced a traffic management model required for direct load control in smart power systems. Also, we gave a general formulation for the optimization problem that needs to be solved by a control center in charge of scheduling the appliances. A more in depth discussion of this optimization will be provided in future works.

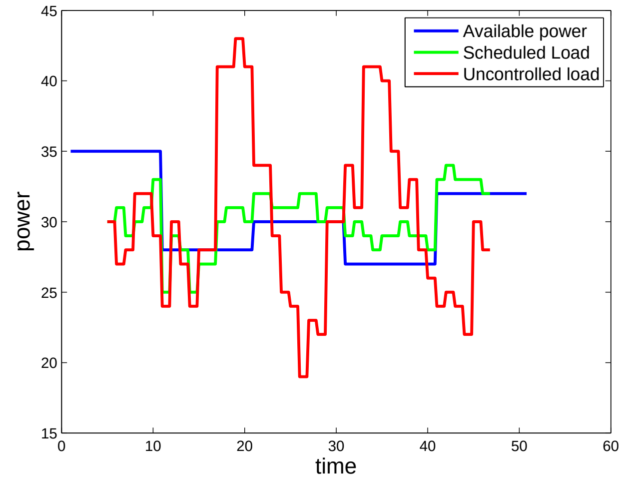

Fig. 3. Simulation Results

References

[1] J. Lopes, F. Soares, and P. Almeida, "Integration of electric vehicles in the electric power system," *Proceedings of the IEEE*, vol. 99, no. 1, pp. 168 –183, jan 2011.

[2] R. Brown, "Impact of smart grid on distribution system design," in *Power and Energy Society General Meeting , 2008 IEEE*, july 2008, pp. 1 –4.

[3] Y.-Y. Hsu and C.-C. Su, "Dispatch of direct load control using dynamic programming," *Power Systems, IEEE Transactions on*, vol. 6, no. 3, pp. 1056 –1061, Aug. 1991.

[4] K. Ng and G. Sheble, "Direct load control-a profit-based load management using linear programming," *Power Systems, IEEE Transactions on*, vol. 13, no. 2, pp. 688 –694, may 1998.

[5] S. Zeng, J. Li, and Y. Ren, "Research of time-of-use electricity pricing models in china: A survey," in *IEEM 2008*, dec 2008, pp. 2191 –2195.

[6] Y. Jia-hai, "Customer response under time-of-use electricity pricing policy based on multi-agent system simulation," in *Power Systems Conference and Exposition, IEEE PES*, nov 2006, pp. 814 –818.

[7] J. Paradiso et al., "Guest editors' introduction: Smart energy systems," *Pervasive Computing, IEEE*, vol. 10, no. 1, pp. 11 –12, jan 2011.

[8] D.-M. Han and J.-H. Lim, "Design and implementation of smart home energy management systems based on zigbee," *Consumer Electronics, IEEE Transactions on*, vol. 56, no. 3, pp. 1417 –1425, 2010.

[9] K. Y. Huang, "Demand subscription services-an iterative dynamic programming for the substation suffering from capacity shortage," *Power Systems, IEEE Transactions on*, vol. 18, no. 2, pp. 947 – 953, may 2003.

[10] M. Pedrasa, T. Spooner, and I. MacGill, "Scheduling of demand side resources using binary particle swarm optimization," *Power Systems, IEEE Transactions on*, vol. 24, no. 3, pp. 1173 –1181, aug 2009.

[11] M. C. Caramanis and J. Foster, "Management of electric vehicle charging to mitigate renewable generation intermittency and distribution network congestion," *48th IEEE CDC*, pp. 4717–4722, dec 2009.

[12] K. Mets et al., "Optimizing smart energy control strategies for plug-in hybrid electric vehicle charging," in *Network Operations and Management Symposium Workshops*, 2010, pp. 293 –299.

[13] O. Sundstrom and C. Bindin, "Optimization methods to plan the charging of electric vehicle fleets," in *in Proceedings of the International Coriference on Control, Communication and Power Engineering*, 2010.

[14] G. Gross and F. Galiana, "Short-term load forecasting," *Proceedings of the IEEE*, vol. 75, no. 12, pp. 1558 – 1573, dec 1987.

[15] K.-H. Kim, J.-K. Park, K.-J. Hwang, and S.-H. Kim, "Implementation of hybrid short-term load forecasting system using artificial neural networks and fuzzy expert systems," *Power Systems, IEEE Transactions on*, vol. 10, no. 3, pp. 1534 –1539, aug 1995.