2012 Seventh International Conference on Knowledge, Information and Creativity Support Systems

Demand response architectures and load management algorithms for energy-efficient power grids: a survey

Yee Wei Law* Tansu Alpcan* Vincent C.-S. Lee† Anthony Lo‡ Slaven Marusic* Marimuthu Palaniswami*
*Dept. of Electrical & Electronic Engineering, The University of Melbourne, Parkville, VIC 3010, Australia
†Clayton School of Information Technology, Monash University, Clayton, VIC 3800, Australia
‡Delft University of Technology, Mekelweg 4, 2628CD Delft, The Netherlands

***Abstract*—A power grid has four segments: generation, transmission, distribution and demand. Until now, utilities have been focusing on streamlining their generation, transmission and distribution operations for energy efficiency. While loads have traditionally been a passive part of a grid, with rapid advances in ICT, demand-side technologies now play an increasingly important role in the energy efficiency of power grids. This paper starts by introducing the key concepts of demand-side management and demand-side load management. Classical demand-side management defines six load shape objectives, of which "peak clipping" and "load shifting" are most widely applicable and most relevant to energy efficiency. At present, the predominant demand-side management activity is demand response (DR). This paper surveys DR architectures, which are ICT architectures for enabling DR programs as well as load management. This paper also surveys load management solutions for responding to DR programs, in the form of load reduction and load shifting algorithms. A taxonomy for "group load shifting" is proposed. Research challenges and opportunities are identified and linked to ambient intelligence, wireless sensor networks, nonintrusive load monitoring, virtual power plants, etc.**

***Index Terms*—Smart grid, energy efficiency, demand-side load management, energy management, load shifting**

I. Introduction

It is estimated that if power grids were 5% more efficient, the energy savings would equate permanently eliminating the fuel and greenhouse gas emissions from 53 million cars [1]. Consequently, demand-side energy efficiency has been receiving a lot of attention in addition to efficiency in generation, transmission and distribution. For example, Australia's technical energy-efficiency improvement between 1990 and 2004 was only about a third of the average amongst OECD countries. The below-average performance has motivated Australia to set a National Energy Efficiency Target to achieve world-class saving by 2015.

Let us start by analysing from "first principle" how a power grid can be made more energy-efficient from the demand side. Line loss is proportional to the current squared, so it is easily understandable that a grid is more energy-efficient with low demand, and therefore load reduction is an obvious way of improving a grid's energy efficiency. Utilities may not have incentives to reduce overall demand, but peak demand reduction helps preventing grid instability.

Flattening the demand curve is another way of making a grid more energy-efficient. To understand this, consider a load that draws a current of $2i$ for half of the day, but no current for the rest of the day, and thereby incurring a line loss that is proportional to $(2i)^2 \times \frac{1}{2}$day. Consider another load that draws a current of i throughout the day, and thereby incurring a line loss that is proportional to $i^2 \times 1$day. The latter load which represents a flat demand incurs half as much line loss. Therefore, a flat demand curve is better for energy efficiency, and also better for infrastructure utilization. Naturally, a grid will also be more energy-efficient if more consumers use energy-efficient appliances.

While utilities have no direct control over their customers' loads, a utility can perform *demand-side management*, i.e., "to plan, implement and monitor activities designed to influence customer uses of electricity in ways that will produce desired changes in the utility's load shape" [2]. Out of the six so-called *load shape objectives* associated with classical demand-side management, three are relevant to our energy efficiency goal: (i) *peak clipping*: reduction of peak load; (ii) *load shifting*: shifting of load from peak to off-peak periods; (iii) *strategic conservation*: reduction of sales. Peak clipping and load shifting coincide with the energy efficiency strategies we have identified earlier. Peak clipping and load shifting reduce network volatility by shaving local demand peaks, thereby assisting constrained networks to cope with summer and winter demand peaks, and reducing the need for investment in grid-infrastructure reinforcement. Strategic conservation is more relevant to the situation where energy resources are scarce and will not be further pursued here.

Currently, the predominant demand-side management activity is *demand response* (DR). A DR program is "a tariff or program established to motivate changes in electric use by end-use customers in response to changes in the price of electricity over time, or to give incentive payments designed to induce lower electricity use at times of high market prices or when grid reliability is jeopardized" [3]. The term *demand-side load management* (equivalently, demand-side energy management, or load management in short) refers loosely to the "adjustment of demand to match supply" [4], and can be understood as a client's response to demand-side management, represented

978-0-7695-4861-6/12 $26.00 © 2012 IEEE
DOI 10.1109/KICSS.2012.45

primarily by DR programs. **Essentially, this paper surveys DR architectures, and load management solutions for meeting the peak clipping and load shifting objectives of demand-side management, in the form of load reduction and load shifting algorithms**. The relations between the concepts just discussed are visualized in Fig. 1.

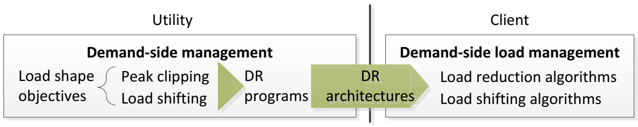

Fig. 1. Relations between demand-side management and demand-side load management concepts.

Our main contribution is that we clarify and make connections between essential concepts of demand-side management and demand-side load management. Unlike recent surveys on smart grid communications [5]–[9], we focus heavily on DR architectures and algorithms in detail and with precision. Our survey of load management algorithms is more extensive and technically in-depth than [10], [11]. Additionally, our taxonomy of group load shifting is new.

The paper is organized as follows. Table I lists frequently used acronyms. Section II introduces DR programs, and explains how they influence the design of DR architectures and load management algorithms. Discussing architectures or algorithms without the other tends to leave an incomplete picture. Therefore, Section III discusses DR architectures, before Sections IV and V discuss load reduction and load shifting algorithms respectively, together with the research challenges and opportunities in these areas. Section VI concludes.

TABLE I
FREQUENTLY USED ACRONYMS IN ALPHABETICAL ORDER

AMI	Advanced metering infrastructure	GLS	Group load shifting
CHP	Combined heat and power	HAN	Home area network
DR	Demand response	ISO	Independent system operator
EMS	Energy management system	NILM	Nonintrusive load monitoring
		VPP	Virtual power plant

II. Demand response programs

DR programs are either price-based or incentive-based.

Price-based DR programs are programs where the tariff fluctuates according to the real-time cost of electricity. Examples are *critical peak pricing* and *time-of-use* pricing. In critical peak pricing (aka dynamic peak pricing), customers are notified in advance of critical peak times – limited to several days per year – during which the tariffs will be much higher than average. In time-of-use pricing, the tariff varies with different time blocks of the day.

The downside of these schemes is the potential grid-destabilizing *rebound effect* [12]–[14]. As observed during the Californian pilot study of time-of-use and critical peak pricing [15], a local demand peak, called the *rebound peak*, arises at the end of a critical period when a large number of loads are re-connected to the grid at roughly the same time. Load management algorithm should be designed to avoid this effect.

Incentive-based DR programs are programs where a utility rewards its customers for their participation. Examples include *peak-time rebate* and *direct load control*. A peak-time rebate program offers a credit or rebate to customers who reduce usage during critical peak hours; the value of this peak-time reduction is monetized in the wholesale market and returned to retail customers by the DR provider (utility most likely) [16]. The difference between peak-time rebate and dynamic pricing programs such as time-of-use and critical peak pricing is that the former rewards the customers if they reduce their peak-time usage, but does not punish them for not changing their usage. Direct load control is a program by which the program operator remotely shuts down or cycles its customers' appliances (e.g., electric water heaters) on short notice [3]. Peak-time rebate and direct load control programs are so far the most widely implemented incentive-based DR programs.

III. Demand response architectures

DR would not be possible without an *advanced metering infrastructure* (AMI), i.e., a two-way communication infrastructure between a utility's enterprise network and its smart meters, whose purpose, besides *automatic meter reading*, is providing up-to-date tariff information to customers. Multiple AMI vendors designate the types of networks that constitute an AMI as *Neighborhood Area Network* and *Field Area Network*. In a Neighborhood Area Network, nodes called *collectors* collect meter data from downstream smart meters and forward the data upstream toward the backhaul network. These data are encoded according to the IEEE standard P1377 (equivalently, ANSI C12.19) and messaged using the IEEE standard P1703 (equivalently, ANSI C12.22). A Field Area Network is for connecting field devices to a utility's SCADA master.

Downstream, an AMI is connected to *Home Area Networks* (HANs), *Building Area Networks* and *Industrial Area Networks* via smart meters. As the name implies, these networks are for homes, buildings and industrial complexes respectively. HAN devices are envisioned to communicate securely with the smart meter, so that (i) devices like in-home displays can securely retrieve and display tariffs, usage and other data; (ii) the smart meter can send load control instructions to HAN devices; (iii) the smart meter can measure the output of microgeneration sources (e.g., solar panels and small wind turbines). While there may be residential customers who prefer to manually monitor and adjust their appliances, automated DR is more cost-effective, especially for business and industrial customers. *Automated DR algorithms are essentially load management algorithms. A DR architecture is an ICT architecture that supports the activation or deployment of load management algorithms by a utility in a client's designated appliances as determined by the DR program the client subscribes to.* Existing DR architectures are surveyed below.

OpenADR (Open Automated DR) [17] is an open specification of communications data models designed to provide interoperable DR signals to building and industrial control systems (e.g., BACnet) that are pre-programmed to take automatic action based on the signals. It is also an architecture (see Fig. 2(a)). The *DR Automation Server* (DRAS) is the central component. Utilities access the server's *Utility or ISO Operator Interface* to manage DR programs and event, perform automated bidding in the electricity market, manage participant accounts, etc. Participants access the server's *Participant Operator Interface* to opt out of DR events, submit feedback, perform automated bidding, etc.

Whirlpool Smart Device Network [18] is a HAN-centric architecture (see Fig. 2(b)). The *Whirlpool Integrated Service Environment* is a collection of web services providing interfaces – some of which similar to OpenADR's DR Automation Server – to utilities and consumers. The *Smart Device Controller* hosts a set of proprietary load management algorithms called the *in-home energy management system*, which can be modified by the Whirlpool Integrated Service Environment.

The **Australian HAN guideline** [19] specifies that a HAN can have two partitions: a *Utility Private HAN* and a *Customer HAN*, bridged by a *Premise EMS* (Energy Management System) (see Fig. 2(c)). The Utility Private HAN includes the smart meter and HAN devices registered with the utility. The Customer HAN includes HAN devices that do not have secure connections with the smart meter. The Premise EMS plays a similar role to Whirlpool's Smart Device Controller.

PowerMatcher [20] is a multi-agent architecture explicitly designed for *supply and demand matching* (see Fig. 2(d)). An agent residing in every device bids and buys (or sells if the device is a producer) in the electricity market. The Home Energy Management Box implements a local energy management strategy based on the user's preference. The Box and the Exchange Agent in effect act as two levels of "supply and demand matchers" in a tree structure (with Boxes constituting the leaves, Exchange Agent being the root), attempting to match demand coming from below to supply available above.

In all the architectures above, the smart meter plays a passive role in automated DR, understandably because it lacks upgradability and the bandwidth necessary to support fine-grained load management. This renders LeMay et al.'s smart meter-centric Meter Gateway Architecture [12] dated. The division of load management tasks is such that the equivalents of Smart Device Controller or Premise EMS are primarily responsible for load shifting, while the loads themselves are responsible for load reduction (energy efficiency).

As soon as AMIs unleash the possibilities of load management on a per-household basis, experts realize that similar architectures should be in place to support cooperative load management among a group of households, which has the potential to mitigate the rebound effect. The IEEE standard 1888 serves just this purpose. The standard defines the *ubiquitous green community control network* (UGCCNet) architecture (see Fig. 3), consisting of four major logical components:

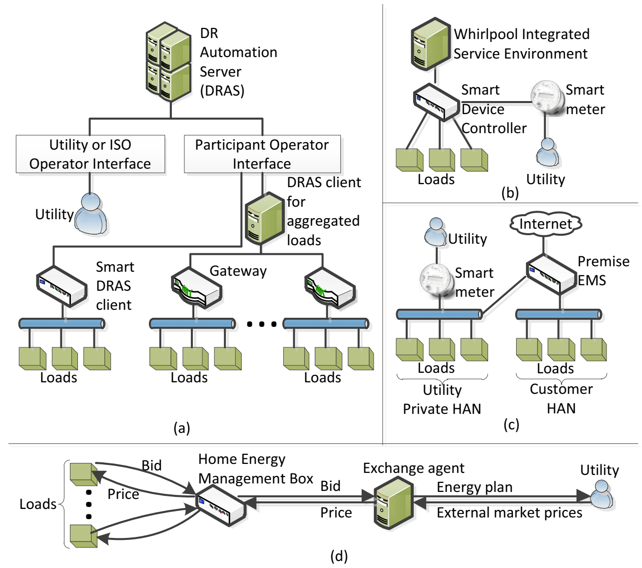

Fig. 2. DR architectures: (a) OpenADR, (b) Whirlpool Smart Device Network, (c) Australian HAN guideline, (d) PowerMatcher.

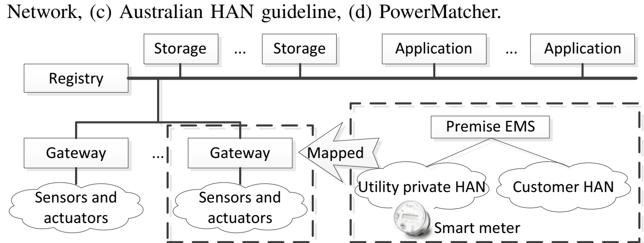

Fig. 3. The UGCCnet architecture; and mapping of an Australian HAN network to an IEEE 1888 sensor and actuator sub-network.

Registry, Gateway, Storage, and Application. The Registry serves as a broker for other components. A Gateway provides access to physical sensors and actuators in its sub-network. The Storage components archive data, whereas the Application components provide user interfaces, access sensors and actuators via Gateways, as well as request data from or store data in Storage components. Fig. 3 shows a potential mapping of an Australian HAN to an IEEE 1888 sensor and actuator sub-network. In this mapping, a Premise EMS acts as a Gateway and hosts a copy of a load management Application designed to optimize energy usage cooperatively. A *building automation and control network* implementing BACnet (ASHRAE/ANSI standard 135-1995, ISO standard 16484-5) can also be extended to serve as an IEEE 1888 sensor and actuator sub-network.

IV. Load reduction algorithms

The DR architectures discussed in the previous section serve as platforms for load management algorithms. It is conceivable that consumers often use more energy than they really need to, e.g., indoor air-conditioning is often colder/warmer than necessary in summer/winter. While meeting the peak clipping objective entails only load reduction at demand peaks, it is in the consumers' interest to match usage closely to requirement at all times, to reduce overall energy bills. For this, load reduction algorithms are used and are discussed in this section.

Load reduction is especially important for buildings because in the US and EU [21], buildings account for 40% of energy consumption; whereas in Australia, the energy used by buildings accounts for approximately 20% of Australia's greenhouse gas emissions, split roughly evenly between homes and commercial buildings.

For businesses in Australia, heating, ventilation and air conditioning (HVAC) typically accounts for up to 40% of energy bills, and is the biggest electricity consumer. Lighting is the second largest electricity consumer. Lighting control can be quite simple, and thanks to the proliferation of ZigBee-based wireless sensors, can be low-cost too. For example, magnetic reed switches on doors, passive infrared motion sensors, and a simple heuristic are sufficient to detect room occupancy (see [22] and [23] for sample testbeds). Upon detection of room occupancy, ambient light sensors can be used to check if current illumination is below 500 lux, in which case the system should activate additional lighting following safety and health regulations [24]. Requiring every occupant to wear an electronic badge to ease detection may not be practical but does enable occupants' preferences for illumination level and so on to be taken into account [25].

HVAC control is more challenging. HVAC mainly conditions temperature and CO_2. Conventional HVAC control strategies (such as [26]) focus on maintaining the temperature level using as little energy as possible, but (i) waste energy on unoccupied rooms, and (ii) do not optimize the ventilation rate. While we can improve these strategies by sensing room occupancy, optimizing CO_2 ventilation rate requires knowledge of the number of occupants. Standard 62.1 of the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) specifies the minimum ventilation rate of a breathing zone to be

$$R_p P_z + R_a A_z,$$

where R_p is the minimum airflow rate per person, P_z is the number of people in the zone, R_a is the minimum airflow rate per zone area, and A_z is the zone area [27]. Optimal HVAC control thus requires knowledge of the number of occupants in a room. Measuring CO_2 level is problematic because CO_2 sensors are slow to detect CO_2 buildup [28]. An array of passive infrared sensors deployed above a doorway can be used for counting people passing through the doorway [29]. However, this method requires sensors to be deployed above every doorway and is only applicable to confined spaces. Worse, a miscount will result in a permanent error. Most other methods for people counting are based on machine vision.

Vision-based human detection and tracking has been studied for decades, but it is only recently that lightweight techniques using low-cost *wireless camera sensor networks* start appearing. So-called *address-event imagers* selectively extract and output a small number of features of interest such as motion, direction of motion, etc. from every image, instead of the image itself [30]. Processing these features not only requires less computational power, storage and bandwidth than processing the original images, but is also less privacy-infringing. Teixeira and Savvides' solution [31] is an example of using address-event imagers for people counting. Using their solution, camera sensor nodes with wide-angle lenses are mounted on the ceiling, in enough number so that an entire room can be captured. The cameras detect moving occupants by frame differencing: after subtracting the previous frame from the current one, the resulting pixels that are above a certain intensity threshold represent objects that have moved, among which objects of a certain pre-determined size are then classified as humans. Due to noise, simple frame differencing may result in "smeared" or "disconnected" blobs; furthermore, frame differencing by definition filters out static and slow-moving occupants. To overcome these problems, Teixeira and Savvides' solution consists of algorithms for making frame differencing robust, and for tracking moving objects. Overall, their solution performs well for less than 5 occupants. See [32], [33] for description of similar systems for people counting, but the detection rates of these systems have not been published.

A. Research challenges and opportunities

The algorithms just discussed fall easily under the framework of intelligent buildings, the research on which actually dates back to decades ago. In the advent of wireless sensor networks, building intelligence evolved into ambient intelligence, and the energy efficiency aspect started getting more attention following the push on smart grids. Ambient intelligence algorithms are generally based on the multi-agent paradigm, and are primarily designed for learning, predicting and supporting user lifestyles in a single-user environment [34]. These algorithms need to be extended to support multi-user environments, with energy efficiency being one of the optimization objectives. Effective sensing is a fundamental challenge. In particular, sensing and estimating occupancy in order to match ventilation rate to the number of people in an area for maximum HVAC energy efficiency is an open problem. Wireless camera sensor networks provide a low-cost mean for people counting, but the best implementation known cannot correctly count more than 5 occupants. More efficient address-event imagers, heterogeneous sensing, and sensor fusion are expected to provide improvement.

Although HVAC is known to be the biggest energy consumer, *energy audits* are useful for identifying other potential sources of wastage in a building. These audits often require single-point measurements to be *disaggregated* into individual appliance's power consumption figures, using a procedure called *nonintrusive load monitoring* (NILM) [35]. However, NILM products that are cost-effective (requires cheap hardware and installation) and easy to use (requires little calibration) have yet to appear. Survey [36] discusses the latest research challenges and opportunities in this area.

V. Load shifting algorithms

When consumption cannot be reduced, cost savings can be achieved by shifting the load to a cheaper billing period. Load shifting assumes the possibility of *remotely* scheduling an appliance, as opposed to *locally* scheduling an appliance by (say) setting the timer on the appliance. A day is divided into slots (e.g., [37], [38] use 1-hour slots, whereas [39] uses 6-minute slots), during each of which electricity price remains constant, and load shifting is determining which of these time slots to turn on a particular appliance. For discussion of load shifting, we use the notation below:

τ	denotes the length of a time slot;
m	denotes the number of time slots per day;
$\mathcal{U}$	denotes the set of users;
$\mathcal{A}_u$	denotes user u's set of appliances;
$p_{u,a}(t)$	denotes the power / load profile of user u's appliance a, where $t = 0$ is the time when a is turned on, and $p_{u,a}(t) = 0, \forall t < 0$;
$x_{u,a}$	denotes the time slot when user u's appliance a is turned on;
$l_{u,a}$	denotes the number of time slots during which user u's appliance a is required to be turned on;
C_i	denotes the price per unit energy at the ith time slot;
$e_{u,a,i}$	denotes the energy consumed by user u's appliance a for the ith time slot;
$E_i^{\max}$	denotes the maximum energy consumed per household for the ith time slot.

By "user", we mean an entity, such as an entire household or business unit, controlling a set of appliances. We can express $e_{u,a,i}$ as a function of $x_{u,a}$, i.e.,

$$e_{u,a,i}(x_{u,a}) = \int_{t=i\tau}^{(i+1)\tau} \mathsf{u}[(x_{u,a} + l_{u,a})\tau - t] p_{u,a}(t - x_{u,a}\tau) dt, \tag{1}$$

where $\mathsf{u}()$ is the Heaviside step function. Notice the integration in (1) is "terminated" at time $t = (x_{u,a} + l_{u,a})\tau$, because

$$\mathsf{u}[(x_{u,a} + l_{u,a})\tau - t] = \begin{cases} 1, & t \le (x_{u,a} + l_{u,a})\tau, \\ 0, & t > (x_{u,a} + l_{u,a})\tau. \end{cases}$$

For user u, the objective is solving the optimization problem:

$$\begin{aligned} \min & \sum_{a \in \mathcal{A}_u} \sum_{i=0}^{m-1} C_i \cdot e_{u,a,i}(x_{u,a}), \\ \text{s.t.} & \sum_{a \in \mathcal{A}_u} e_{u,a,i}(x_{u,a}) \le E_i^{\max}, \forall i = 0, \ldots, m-1. \end{aligned} \tag{2}$$

The inequality constraint in (2) is used to prevent all users from scheduling all their appliances at the same time (when C_i is lowest). This inequality constraint is absent in Kishore and Snyder's first load shifting scheme [40, (1)], which unsurprisingly is plagued by the rebound peak problem.

Many load shifting schemes (e.g., [37], [40], [41]) employ an artificial discomfort / inconvenience / waiting cost to model the cost incurred for the delay between the instance when a user turns on an appliance and the instance when the appliance is turned on according to schedule. Due to the lack of an objective way of quantifying inconvenience and compelling scenarios where there is an economic need to quantify inconvenience caused by load shifting, (2) does not include such an artificial cost.

Instead of inconvenience cost, Pedrasa et al. [42] introduce the cost metric "perceived benefit" $b_{u,a,i}(e_{u,a,i})$, which measures the price u is willing to pay for the service level that a will provide if it consumes $e_{u,a,i}$ energy during the ith time slot. The optimization objective becomes

$$\min \sum_{a \in \mathcal{A}_u} \sum_{i=0}^{m-1} \left(C_i e_{u,a,i} - b_{u,a,i}(e_{u,a,i}) \right). \tag{3}$$

In (3), $e_{u,a,i}$ is not a function of $x_{u,a}$ as in (2), but is itself the optimization variable. When $e_{u,a,i}$ fluctuates from slot to slot, so does the power, in a way that deviates arbitrarily from the load profile. While this may be acceptable for energy-dependent appliances, power-dependent appliances may perform sub-optimally, or not at all, or even be damaged as a result. Another problem with (3) is that there is no universally accepted way of quantifying perceived benefits in monetary values.

Mohsenian-Rad et al. [37] assign an exact, predicted total usage amount E_a to every appliance a, besides capping the usage for each time slot i at $E_i^{\max}$. Their formulation, omitting any waiting cost, is as follows:

$$\min \sum_{a \in \mathcal{A}_u} \sum_{i=0}^{m-1} C_i \left(\sum_{a \in \mathcal{A}_u} e_{u,a,i} \right) \cdot e_{u,a,i}, \tag{4}$$

$$\begin{aligned} \text{s.t. } & \alpha_{u,a} \le e_{u,a,i} \le \beta_{u,a}, \forall i = 0, ..., m-1, \forall a \in \mathcal{A}_u, \\ & \sum_{i=0}^{m-1} e_{u,a,i} = E_a, \forall a \in \mathcal{A}_u, \\ & \sum_{a \in \mathcal{A}_u} e_{u,a,i} \le E_i^{\max}, \forall i = 0, \ldots, m-1. \end{aligned} \tag{5}$$

The differences between (4) and (2) are:

- C_i in (4) is a function that increases with the total energy consumed by the appliances during time slot i. This is to simulate the *inclining block rates* pricing scheme.
- $e_{u,a,i}$ in (4) is not a function of $x_{u,a}$, but is itself the optimization variable constrained by (5), where $\alpha_{u,a}$ and $\beta_{u,a}$ represent the lower and upper energy limits per time slot characterizing appliance a respectively. The caveat of using this optimization variable has been discussed.

Erol-Kantarci et al. [43] propose to schedule appliances sequentially, but this temporal constraint may actually lead to higher electricity bills.

Xiao et al. [44] propose a heuristic based on the Longest Processing Time algorithm that tries to minimize the maximum bill for each time slot, rather than the daily bill. The effect is that although the user's daily bill is not minimized, the *user's* demand curve is effectively flattened.

All the load shifting schemes discussed so far are individual-oriented. When users in a neighborhood collaborate to determine the optimal energy allocation for each time slot, the *system's* demand curve can be flattened more effectively. We call this kind of scheme *group load shifting* (GLS). GLS can be *centralized*, in which case the utility alone dictates the load schedule for each user, by considering all users in a neighborhood as a group. Some vendors are already providing centralized GLS solutions although they are not necessarily labelled as such (e.g., GE's Grid IQ Demand Optimization System DR1000). GLS can also be *distributed*, in which case each user coordinates with one another directly and does not require any more than a pricing signal from the utility. GLS can also be *hierarchical*, in which case each user coordinates with a "global planner" to derive the optimal schedule based on input from all the users in the neighborhood. The following subsections survey distributed and hierarchical GLS schemes.

A. Distributed group load shifting schemes

Kishore and Snyder [40] propose a distributed "neighborhood-level load scheduling" protocol, where users in a neighborhood contend for energy from a finite energy resource for every time slot. The protocol is heuristic and assumes the "energy management controllers" in a neighborhood are one hop away from each other, which is a severe limitation. Other issues include: packet collisions are not handled; no countermeasures against selfish controllers (e.g., controllers that do not wait for a random delay before requesting for energy).

Mohsenian-Rad et al. [38], [45] propose a distributed scheme where each user u in group $\mathcal{U}$ aims to minimize the daily bill of $\mathcal{U}$, by changing their single-user formulation (4) to:

$$\min \sum_{u\in\mathcal{U}} \sum_{a\in\mathcal{A}_u} \sum_{i=0}^{m-1} C_i \left(\sum_{u\in\mathcal{U}} \sum_{a\in\mathcal{A}_u} E_a \right) \cdot e_{u,a,i}, \tag{6}$$

$$\text{s.t. } \alpha_{u,a} \le e_{u,a,i} \le \beta_{u,a}, \forall i = 0, ..., m-1, \forall a \in \mathcal{A}_u, \forall u \in \mathcal{U}, \tag{7}$$

$$\sum_{i=0}^{m-1} e_{u,a,i} = E_a, \forall a \in \mathcal{A}_u, \forall u \in \mathcal{U}. \tag{8}$$

C_i in (6) is a function that increases with the total energy consumed by the group during time slot i, such that the cost function in (6) is strictly convex. Real-world pricing models however rarely lead to a strictly convex cost function, e.g., BC Hydro's pricing model [45]. (8) means that every appliance is assigned an exact, pre-determined total usage amount, which implies that every user u has a pre-determined usage amount, which further implies that his/her daily bill is a *constant* fraction of the group's daily bill. Thus, by minimizing the group's daily bill, every user essentially minimizes his/her own daily bill. According to n-person game theory, a unique Nash equilibrium exists where all users adhere to the strategy of minimizing the group's daily bill.

For communication efficiency, these schemes can be implemented on top of UGCCNet (see Section III) or comparable architectures. Distributed GLS schemes have the advantage that users do not need to surrender control of their appliances to their utilities, but do expose the users to security and privacy risks. *None* of the above schemes have been designed with security in mind. Defending these schemes against malicious users will significantly increase their complexity.

B. Hierarchical group load shifting schemes

Due to the limited information flow between users, hierarchical GLS is potentially more secure than distributed GLS. However, the viability of both distributed and hierarchical GLS understandably hinges on how much savings users can make from using these technologies. In this respect, hierarchical GLS schemes is also potentially more advantageous, due to the possibility of realizing the *virtual power plant / producer* (VPP) concept using these schemes. There are several definitions of VPP [46] but basically, a VPP is a group of energy producers and consumers acting as though they are a single energy producer with *stable*, defined hourly output that is tradable on the *spinning reserve* market. The idea of a fleet of distributed but interconnected micro CHP (combined heat and power) generators and storages appearing on the energy market as a single VPP has been around for a while. Recently, several hierarchical GLS schemes have been proposed for managing micro CHP generators alongside energy-consuming appliances, so that participating users can collectively operate as a VPP [20].

To this end, Molderink et al. [39], [47] propose a hierarchical GLS scheme consisting of three steps:

Step 1 – Local prediction: Each user predicts his/her heat demand for the next day using neural networks, using data from preceding weeks and local outdoor temperatures as input. Based on the heat demand and the required micro CHP output in fulfillment of the user's role in the VPP, a local electricity production plan is derived and sent to the global planner.

Step 2 – Global planning: The global planner has to solve the micro CHP scheduling problem: given m time slots, n micro CHPs with their own production plans for each time slot, which micro CHPs should be turned on at each time slot to meet the VPP production goal for every time slot?

The problem is non-trivial because each micro CHP has its own constraints, e.g., its production plan, its maximum output power, the minimum durations it must keep running for each run and stay off after each run to minimize wearing, the startup and shutdown delays of its micro CHP, etc. In fact, (a simplified version of) the problem is shown to be NP-complete in the strong sense [48].

For this reason, a heuristic is used to adjust the local plans by minimizing the mismatch between the local plans with the global (VPP production) plan. An earlier version of the heuristic [49] applies different electricity prices to different houses, which is not realistic.

Step 3 – Local scheduling: An appliance's load profile is modeled to consist of *power ranges*, i.e., user u's appliance a's power consumption takes a value between $\alpha_{u,a}^{\text{PR }1}$ and $\beta_{u,a}^{\text{PR }1}$, or between $\alpha_{u,a}^{\text{PR }2}$ and $\beta_{u,a}^{\text{PR }2}$, and so on; where the superscript

"PR r" represents the rth power range, and $\alpha_{u,a}^{\text{PR}\ r} \leq \beta_{u,a}^{\text{PR}\ r}$. Note the case $\alpha_{u,a}^{\text{PR}\ r} = \beta_{u,a}^{\text{PR}\ r}$, and the case $\alpha_{u,a}^{\text{PR}\ r} \leq \beta_{u,a}^{\text{PR}\ r} \leq 0$ (when a actually generates electricity) are possible. Local scheduling for each time slot is expressed as the optimization problem (9)–(12):

$$\min \sum_r \text{Cost}_{\text{running}} \cdot p_{u,a}^{\text{PR}\ r} + \text{Cost}_{\text{startup}} \cdot x_{u,a}^{\text{PR}\ r}, \tag{9}$$

$$\text{s.t. } x_{u,a}^{\text{PR}\ r} \in \{0,1\}, \sum_r x_{u,a}^{\text{PR}\ r} = 1, \tag{10}$$

$$x_{u,a}^{\text{PR}\ r} \alpha_{u,a}^{\text{PR}\ r} \leq p_{u,a}^{\text{PR}\ r} \leq x_{u,a}^{\text{PR}\ r} \beta_{u,a}^{\text{PR}\ r}, \forall r, \tag{11}$$

$$\tau \sum_{a \in \mathcal{A}_u} \sum_r p_{u,a}^{\text{PR}\ r} = \text{local heat, electricity demand} + \text{demand from global plan.} \tag{12}$$

The idea of this formulation is to select one power range for each appliance (constraint (10)), and within the selected range select a power value (constraint (11)), that satisfies the heat and electricity demand (constraint (12)) and gives the lowest cost. The cost has two components: a running cost that is proportional to $p_{u,a}^{\text{PR}\ r}$, and a constant startup cost associated with the power range r.

C. Research challenges and opportunities

Modeling: Accurate modeling of load profiles is desired. As a result of extensive metering, it is now understood that appliances can be modeled as either finite state machines (FSMs) or continuously variable appliances [35]. FSMs operate on discrete power levels, e.g., washing machines. Continuously variable appliances operate on a continuous range of power levels, e.g., dimmer lights. Some appliances that can be modeled as FSMs, e.g., dishwashers, behave in part like continuously variable appliances. Yet some other appliances that can be modeled as FSMs, e.g., plasma TVs, fluctuate quite a bit in power consumption depending on their workload and user activities. No existing schemes have endeavored to model appliances accurately. Results from the NILM literature should provide input to this modeling challenge. Additionally, NILM measurements provide training data for the prediction of local demand, as required by many load shifting schemes.

Security: Distributed GLS schemes involve exchanging of power usage information among the participants. The IEEE standard 1888 has identified some broad security requirements, but has not discussed potential attacks and countermeasures in detail. Privacy invasion and false data injection are among the threats that need to be addressed in depth. Key to these issues is the establishment of a suitable access control structure and associated key management scheme.

Reliability: So that a VPP can provide spinning reserve services, the reliability of the relevant electrical as well as communication networks needs to be ensured. In this regard, the global planner should leverage the redundancy inherent in a large fleet of microgenerators, by monitoring and tracking individual microgenerators, and devising the global plan correspondingly. Less reliable microgenerators should be assigned a lower target and therefore a lower share of the profit. Opportunities exist for the design of a real-time control scheme to monitor for and counterbalance possible undersupply or oversupply by any of the microgenerators. A significant challenge lies in the communication efficiency necessary to achieve real-time response.

Economics: Missing from existing hierarchical GLS schemes is a specification of the global planner's market strategy for maximizing the VPP's profit. Together with microgrid, VPP has been identified as one of two main strategies for *integrating* distributed generation into a power grid. Creation of market models and negotiation strategies for microgrids and VPPs is thus an important issue (see review [50]). Also missing from existing hierarchical GLS schemes is a specification of the users' demand response strategy when the VPP suffers a supply deficit. Both market models for VPPs and demand response can be designed and analyzed using game theory.

VI. Conclusion

Driven by sustainability initiatives and advances in ICT infrastructures, demand-side management is supplanting supply-side management (increasing generators to meet demand). Classical demand-side management defines six load shape objectives, of which "peak clipping" and "load shifting" are most widely applicable and most relevant to energy efficiency. At present, the predominant demand-side management activity is demand response (DR). This paper surveys DR architectures, which are ICT architectures for enabling DR programs as well as load management. This paper also surveys load management solutions for responding to DR programs, in the form of load reduction and load shifting algorithms. For load reduction, multi-agent ambient intelligence technologies, wireless camera sensor networks, heterogeneous sensing, and sensor fusion are expected to provide significant energy efficiency improvements to conventional HVAC control. For load shifting, this paper proposes a taxonomy for group load shifting, whose objective is to flatten the system's demand curve more effectively than traditional load shifting. Research challenges and opportunities for load shifting are identified to be in the areas of appliance modeling, economic modeling, security, reliability, and communications. To the development of both load reduction and load shifting algorithms, nonintrusive load monitoring can provide useful input.

References

[1] *The Smart Grid: An Introduction*, US Department of Energy, 2008.

[2] C. Gellings, "The concept of demand-side management for electric utilities," *Proc. IEEE*, vol. 73, no. 10, pp. 1468–1470, Oct. 1985.

[3] *Benefits of Demand Response in Electricity Markets and Recommendations for Achieving them*, US Department of Energy, Feb. 2006, report to the United States Congress.

[4] K. Wacks, "Utility load management using home automation," *IEEE Trans. Consum. Electron.*, vol. 37, no. 2, pp. 168–174, May 1991.

[5] V. Gungor and F. Lambert, "A survey on communication networks for electric system automation," *Computer Networks*, vol. 50, no. 7, pp. 877–897, 2006.

[6] W. Wang, Y. Xu, and M. Khanna, "A survey on the communication architectures in smart grid," *Computer Networks*, vol. 55, no. 15, pp. 3604–3629, 2011.

[7] S. K. Tan, M. Sooriyabandara, and Z. Fan, "M2M Communications in the Smart Grid: Applications, Standards, Enabling Technologies, and Research Challenges," *International Journal of Digital Multimedia Broadcasting*, vol. 2011, 2011, article ID 289015, 8 pages.

[8] J. Gao, Y. Xiao, J. Liu, W. Liang, and C. P. Chen, "A survey of communication/networking in Smart Grids," *Future Gener. Comput. Syst.*, vol. 28, no. 2, pp. 391–404, Feb. 2012.
[9] Z. Fan, P. Kulkarni, S. Gormus, C. Efthymiou, G. Kalogridis, M. Sooriyabandara, Z. Zhu, S. Lambotharan, and W. Chin, "Smart grid communications: Overview of research challenges, solutions, and standardization activities," *IEEE Communications Surveys Tutorials*, vol. PP, no. 99, pp. 1–18, 2012.
[10] M. Erol-Kantarci and H. T. Mouftah, *Energy Management Systems*. InTech, 2011, ch. 13: Demand Management and Wireless Sensor Networks in the Smart Grid.
[11] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart grid – the new and improved power grid: A survey," *IEEE Communications Surveys Tutorials*, vol. PP, no. 99, pp. 1–37, 2011.
[12] M. LeMay, R. Nelli, G. Gross, and C. A. Gunter, "An integrated architecture for demand response communications and control," in *Hawaii International Conference on System Sciences*. Los Alamitos, CA, USA: IEEE Computer Society, 2008, pp. 174–183.
[13] M. A. Piette, S. Kiliccote, and G. Ghatikar, "Design and implementation of an open, interoperable automated demand response infrastructure," Lawrence Berkeley National Laboratory, LBNL Paper LBNL-63665, 2008.
[14] J. W. Black and R. Tyagi, "Potential problems with large scale differential pricing programs," in *2010 IEEE PES Transmission and Distribution Conference and Exposition*, Apr. 2010, pp. 1–5.
[15] P. McAuliffe and A. Rosenfeld, "Response of residential customers to critical peak pricing and time-of-use rates during the summer of 2003," California Energy Commission, Tech. Rep., Sep. 2004.
[16] B. R. Alexander, "Dynamic Pricing? Not So Fast! A Residential Consumer Perspective," *The Electricity Journal*, vol. 23, no. 6, pp. 39–49, 2010.
[17] M. A. Piette, G. Ghatikar, S. Kiliccote, E. Koch, D. Hennage, P. Palensky, and C. McParland, "Open automated demand response communications specification (version 1.0)," California Energy Commission, PIER Program CEC5002009063, Apr. 2009.
[18] T. Lui, W. Stirling, and H. Marcy, "Get smart," *IEEE Power and Energy Magazine*, vol. 8, no. 3, pp. 66–78, May-June 2010.
[19] *Advanced Metering Infrastructure Home Area Network (HAN) Functionary Guideline*, Dept. of Primary Industries, Nov. 2008, ver. 0.5.
[20] J. K. Kok, C. J. Warmer, and I. G. Kamphuis, "PowerMatcher: multiagent control in the electricity infrastructure," in *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*, ser. AAMAS '05. ACM, 2005, pp. 75–82.
[21] S. Baden, P. Fairey, P. Waide, P. de T'serclaes, and J. Laustsen, "Hurdling Financial Barriers to Low Energy Buildings: Experiences from the USA and Europe on Financial Incentives and Monetizing Building Energy Savings in Private Investment Decisions," in *Proc. 2006 ACEEE Summer Study on Energy Efficiency in Buildings*. American Council for an Energy-Efficient Economy, 2006.
[22] A. Marchiori and Q. Han, "Distributed wireless control for building energy management?" in *Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building*, ser. BuildSys '10. ACM, 2010, pp. 37–42.
[23] Y. Agarwal, B. Balaji, R. Gupta, J. Lyles, M. Wei, and T. Weng, "Occupancy-driven energy management for smart building automation," in *Proc. 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building*, ser. BuildSys '10. ACM, 2010, pp. 1–6.
[24] D. T. Delaney, G. M. P. O'Hare, and A. G. Ruzzelli, "Evaluation of energy-efficiency in lighting systems using sensor networks," in *Proc. First ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings*, ser. BuildSys '09. ACM, 2009, pp. 61–66.
[25] L.-W. Yeh, Y.-C. Wang, and Y.-C. Tseng, "iPower: an energy conservation system for intelligent buildings by wireless sensor networks," *International Journal of Sensor Networks*, vol. 5, no. 1, pp. 1–10, 2009.
[26] A. Kusiak and G. Xu, "Modeling and optimization of hvac systems using a dynamic neural network," *Energy*, no. 0, 2012.
[27] V. L. Erickson and A. E. Cerpa, "Occupancy based demand response HVAC control strategy," in *Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building*, ser. BuildSys '10. ACM, 2010, pp. 7–12.
[28] S. Wang and X. Jin, "CO_2-based occupancy detection for on-line outdoor air flow control," *Indoor and Built Environment*, vol. 7, no. 3, pp. 165–181, 1998.
[29] K. Hashimoto, K. Morinaka, N. Yoshiike, C. Kawaguchi, and S. Matsueda, "People count system using multi-sensing application," in *1997 International Conference on Solid State Sensors and Actuators (TRANSDUCERS '97)*, vol. 2, Jun. 1997, pp. 1291–1294.
[30] T. Teixeira, E. Culurciello, J. Park, D. Lymberopoulos, A. Barton-Sweeney, and A. Savvides, "Address-event imagers for sensor networks: evaluation and modeling," in *The Fifth International Conference on Information Processing in Sensor Networks (IPSN)*, 2006, pp. 458–466.
[31] T. Teixeira and A. Savvides, "Lightweight People Counting and Localizing for Easily Deployable Indoors WSNs," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 4, pp. 493–502, Aug. 2008.
[32] A. Kamthe, L. Jiang, M. Dudys, and A. Cerpa, "SCOPES: Smart Cameras Object Position Estimation System," in *Wireless Sensor Networks*, ser. Lecture Notes in Computer Science, U. Roedig and C. Sreenan, Eds. Springer Berlin / Heidelberg, 2009, vol. 5432, pp. 279–295.
[33] L. Gasparini, R. Manduchi, and M. Gottardi, "An ultra-low-power contrast-based integrated camera node and its application as a people counter," in *2010 Seventh IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, Sep. 2010, pp. 547–554.
[34] D. J. Cook, J. C. Augusto, and V. R. Jakkula, "Ambient intelligence: Technologies, applications, and opportunities," *Pervasive and Mobile Computing*, vol. 5, no. 4, pp. 277–298, 2009.
[35] G. Hart, "Nonintrusive appliance load monitoring," *Proc. IEEE*, vol. 80, no. 12, pp. 1870–1891, Dec. 1992.
[36] M. Zeifman and K. Roth, "Nonintrusive appliance load monitoring: Review and outlook," *IEEE Trans. Consum. Electron.*, vol. 57, no. 1, pp. 76–84, Feb. 2011.
[37] A.-H. Mohsenian-Rad and A. Leon-Garcia, "Optimal residential load control with price prediction in real-time electricity pricing environments," *IEEE Trans. on Smart Grid*, vol. 1, no. 2, pp. 120–133, 2010.
[38] A.-H. Mohsenian-Rad, V. Wong, J. Jatskevich, and R. Schober, "Optimal and autonomous incentive-based energy consumption scheduling algorithm for smart grid," in *Innovative Smart Grid Technologies (ISGT)*, Jan. 2010, pp. 1–6.
[39] A. Molderink, V. Bakker, M. Bosman, J. Hurink, and G. Smit, "Management and control of domestic smart grid technology," *IEEE Trans. Smart Grid*, vol. 1, no. 2, pp. 109–119, Sep. 2010.
[40] S. Kishore and L. Snyder, "Control mechanisms for residential electricity demand in smartgrids," in *First IEEE International Conference on Smart Grid Communications (SmartGridComm 2010)*, Oct. 2010, pp. 443–448.
[41] L. D. Ha, S. Ploix, E. Zamai, and M. Jacomino, "Tabu search for the optimization of household energy consumption," in *2006 IEEE International Conference on Information Reuse and Integration*, Sep. 2006, pp. 86–92.
[42] M. Pedrasa, T. Spooner, and I. MacGill, "Coordinated scheduling of residential distributed energy resources to optimize smart home energy services," *IEEE Trans. Smart Grid*, vol. 1, no. 2, pp. 134–143, 2010.
[43] M. Erol-Kantarci and H. Mouftah, "Wireless sensor networks for cost-efficient residential energy management in the smart grid," *IEEE Trans. Smart Grid*, vol. 2, no. 2, pp. 314–325, Jun. 2011.
[44] J. Xiao, J. Y. Chung, J. Li, R. Boutaba, and J.-K. Hong, "Near optimal demand-side energy management under real-time demand-response pricing," in *2010 International Conference on Network and Service Management (CNSM)*, Oct. 2010, pp. 527–532.
[45] A. Mohsenian-Rad, V. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia, "Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid," *IEEE Trans. Smart Grid*, vol. 1, no. 3, pp. 320–331, Dec. 2010.
[46] E. A. Setiawan, *Concept and Controllability of Virtual Power Plant*. Kassel University Press GmbH, 2007.
[47] A. Molderink, V. Bakker, M. Bosman, J. Hurink, and G. Smit, "A three-step methodology to improve domestic energy efficiency," in *Innovative Smart Grid Technologies (ISGT), 2010*, Jan. 2010, pp. 1–8.
[48] M. Bosman, V. Bakker, A. Molderink, J. Hurink, and G. Smit, "On the microCHP scheduling problem," in *3rd Global Conference on Power Control and Optimization (PCO)*. Australia: PCO, Feb. 2010.
[49] V. Bakker, M. Bosman, A. Molderink, J. Hurink, and G. Smit, "Demand side load management using a three step optimization methodology," in *First IEEE International Conference on Smart Grid Communications (SmartGridComm)*, Oct. 2010, pp. 431–436.
[50] E. Mashhour and S. Moghaddas-Tafreshi, "A review on operation of micro grids and virtual power plants in the power markets," in *2nd International Conference on Adaptive Science Technology (ICAST 2009)*, Jan. 2009, pp. 273–277.