To appear in CAMSAP2011. Author Copy - do not redistribute.

Demand Side Management Trends in the Power Grid

Mahnoosh Alizadeh, Zhifang Wang and Anna Scaglione
University of California, Davis
Email: {malizadeh,zfwang,ascaglione}@ucdavis.edu

***Abstract*—The roadmap of the Smart Grid project includes exploiting the intrinsic elasticity of electricity demand in the future, to make it responsive to the near term cost of supplying generation. This would curb costly peaks of demand and allow for a greater penetration of renewable energy sources. The mechanisms serving this purpose are referred to as Demand Side Management (DSM) and Demand-Response (DR) programs. While it is clear that DSM and DR will be indispensable to loosen the control over the generation and decrease the reserve requirements, there is much debate on what is the right architecture for DSM and DR programs. In this paper we discuss current trends that are being considered as candidates for DSM and DR and critically compare them, outlining research directions that should be pursued in the future to overcome this dilemma.**[1]

I. Introduction

Demand Side Management (DSM) systems have evolved over the past three decades through systematic activities of the Power Utilities, as well as Government policies, designed to change the amount and timing of electricity consumption. Such DSM measures have been implemented for load management, for increasing energy efficiency, and for electrification (i.e. the strategic increase of electricity use) [1].

Initially DSM programs mainly comprised reliability-driven load management measures, used occasionally to manage emergency situations. Recently, more sophisticated and rapid forms of DSM have emerged, extending the level of consumer interaction for this services, through appropriate incentives [2]. One of the great promises of Smart Grid is to foster advanced forms of DSM that continuously control the load for potentially all consumers. The most notable idea emerging today is the inclusion of programs that enable direct price responsiveness, even of individual loads [1]. In a wider definition, DSM may also include items such as renewable energy systems, combined heat and power systems, independent power purchase, and all instruments that allow to meet the customer demand with the highest efficiency [3].

In this paper we discuss the two most popular trends of demand management, which lie at the two opposite sides of the control spectrum. At one end are Real Time Pricing strategies, discussed in Section II and, at the other end, are Direct Load Control (DLC) strategies, presented in Section III. As we argue in Section I-A there is a pressing need for effective and reliable DSM technology. With numerous legislations being enacted around the world requiring that the consumption of renewable resources (like wind and solar energy) be increased to very high targets soon, and considering the degree of uncertainty that will be introduced on the generation side of the power grid as a result of this, it will be economically infeasible to maintain today's levels of grid reliability without creating incentives for the customers to follow the variable generation supply, rather than the other way around. DSM advances, and RTP in particular, have been made possible by technological progress in data processing and in communications. But the network and processing infrastructure for DSM is clearly best defined by knowing what the right DSM application is. As this paper outlines, this is still a subject of heated debate and research. In Section IV, we provide a new model which can strike the right balance between the options considered so far.

[1]This work was funded under TCIPG and CERTS Projects sponsored by DOE.

A. Why is DSM technology important

The most important rule in the operation of power systems is that there has to be a continuous balance between demand and generation. Currently, there are two very opposite views on how green generation can be integrated into the wholesale market: **1)** *Generation Reserves*: volatile and non-dispatchable plants like wind and solar should be backed by clean but controllable resources like hydro or natural gas units that will start generating energy whenever the intermittent resources are not able to meet their scheduled generation requirements. This is the approach followed today. But the main problem with this solution is limited availability of clean but dispatchable plants. DOE reports that today 10% of all generation assets and 25% of distribution infrastructure are used less than 400 hours per year (5% of the time). To triple the penetration of renewable, these numbers will more than triple. **2)** *Demand Side Management*: the volatility introduced by these intermittent resources on the generation side will be compensated by price responsive loads, such as RTP, or DLC programs. As we will see, DLC programs are mostly designed for emergency situations while RTP programs, although very appealing, face the very challenging problem of determining what these price signals should be.

II. Real-time Pricing

One of the most serious contenders in the DSM research arena is RTP. The concept of real-time prices has been around for about three decades now [4]: instead of shielding the customer completely from fluctuations of energy costs in the spot market (as it is done when using flat rates and time-of-use tariffs) in RTP, *price signals* delivered to the customers will provide incentives to modify their demand and alleviate the pressure on the grid, with the reward of lowering their bill. What is not obvious is how should the utility calculate and post price signals in order to attain a stable operating point for the system, balancing generation and demand.

Theoretically, the real-time price of electricity at each node is the shadow price obtained from solving the optimization for the spot market security constrained generation dispatch [5], [6]. In a system with N generator and L load buses, each generating G_i or consuming D_i units of power respectively, and W transmission lines, this optimization is given by

$$\begin{aligned} G_i^* = \underset{G_i}{\text{argmin}} & \sum_{i=1}^{N} C_i(G_i) \qquad \text{(II.1)} \\ \text{s.t.} & \sum_{i=1}^{N} G_i - \sum_{i=1}^{L} D_i - \text{Loss} = 0 \\ & \left| H[G_1 \dots G_N \ D_1 \dots D_L]^T \right| \leq \bar{F}^{\max} \\ & G_i^{\min} \leq G_i \leq G_i^{\max} \end{aligned}$$

where $C_i(\cdot)$ is the generation cost function and H is a matrix that relates power flow on transmission lines to nodal power inputs. The first constraint ensures power balance; the second ensures that flows on all the transmission lines lie within their limit, given by the vector $\bar{F}^{\max}$, and the third one defines generation capacity limits. The Lagrangian of (II.1) is:

$$\begin{aligned} \mathcal{L} = & \sum_{i=1}^{N} C_i(G_i) - \theta \left(\sum_{i=1}^{N} G_i - \sum_{i=1}^{L} D_i - \text{Loss} \right) \\ & - \bar{\mu}^T \left(H[G_1 \dots G_N \ D_1 \dots D_L]^T - \bar{F}^{\max} \right) \\ & + \sum_{i=1}^{N} [\nu_i^{\max}(G_i - G_i^{\max}) - \nu_i^{\min}(G_i - G_i^{\min})], \qquad \text{(II.2)} \end{aligned}$$

where θ, $\bar{\mu} = [\mu_1, \cdots, \mu_W]^T$, $\nu_i^{\max}$ and $\nu_i^{\min}$ are the Lagrange multipliers at the optimal solution. The Locational Marginal Price (LMP) for load bus i is,

$$\lambda_i = \frac{\partial \mathcal{L}}{\partial D_i} |_{G_i = G_i^*}, \qquad \text{(II.3)}$$

which represents the marginal cost of providing one additional unit of power at that bus under the optimal generation dispatch. Simplifying assumptions or additional constraints can be added to (II.1), such as adopting linear or quadratic generation costs and including additional contingency constraints [7]. LMPs are the true costs of serving loads, which include costs of generation, grid losses and congestion on transmission lines. Theoretically, one could calculate these prices before real-time and by releasing them the system should converge to its optimal operating point. This will happen if:

- Perfect forecasts of demand values D_i's are available

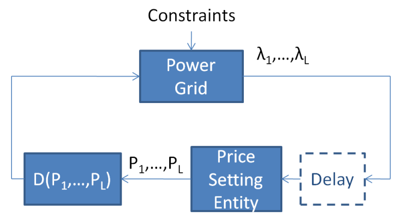

Fig. 1. System model with RTP

- The customers are shielded from this true cost and do not react to fluctuations of the their associated LMP (like what happens in most of the power grid today)
- Most importantly, all generators should be *price taking rational agents*, so that posting the calculated LMPs will make them choose the dipatched value in (II.1).

If, however, the customer are exposed to even limited information about these true costs, like what all real-time pricing techniques aim to do, there will be an extra feedback loop added to the equation, i.e. the demand at the i-th load bus, D_i will be a function of λ_i and thus, (II.3) has to be modified to

$$\lambda_i = \frac{\partial \mathcal{L}(\lambda_1, \dots, \lambda_i, \dots, \lambda_L)}{\partial D_i} \qquad \text{(II.4)}$$

A diagram of the system is shown in Fig. 1. Calculating the true cost of serving the demand if the customers react to real-time market fluctuations will require perfect knowledge of customer behavior in reacting to price signals for every different load bus in the system at the time the LMPs are being calculated for. [8] provides a dynamic model to analyse the dynamics of supply, demand, and clearing prices in a power grid with real-time retail pricing and information asymmetry. It shows that a power market accomodating common RTP techniques may possibly experience volatile prices and demand, or even lose its stability. Although proven with major simplifying assumptions, the important point made is that stability should be taken into account when designing a load control system. The authors conclude that more intricate models for demand in each area, knowledge about consumer behavior in response to dynamic prices received, and a thorough understanding of the implications of different market mechanisms and system architectures are needed before real-time pricing techniques can be implemented in large-scale.

Currently, power system operators have two major approaches to calculate LMPs for the real-time market. Ex-anté prices indicate the value of the LMPs before the true value of the demand is released, using predicted values of the stochastic variables like the demand and intermittent renewable resources. Ex-post values, on the other hand, are calculated after the load is served and with deterministic knowledge of all the values for demand and generation.

Since real-time price signals need to be delivered to customers beforehand to allow some planning time, they should be of the same nature as ex-anté LMPs. Also, it is very unlikely that ex-post adjustments will be allowed to affect how customers are billed, since this would expose the public to unacceptable risks. The same approach just described, is adopted by most RTP researchers in the literature. The real-time price sent to customers is derived either from a direct ex-anté analysis of wholesale market prices, or by adding some modifications to account for consumer satisfaction. To do so, either a term representing the benefit of customers from consuming electricity is added to the cost function in (II.1) or, a cap on the variations of the price signal in enforced. For example, in [9], the authors maximize the social benefit, which includes known cost functions $C_i(.)$ representing the cost of production of energy and known benefit functions $B_i(.)$ representing the consumers

$$\max \sum_{i=1}^{L} B_i(D_i) - \sum_{i=1}^{N} C_i(G_i) \tag{II.5}$$

To calculate the price, the authors assume that both generation units and customers are price taking agents with known limits on their consumption and generation values and declare the price at each bus as the marginal value of objective function in this maximization problem (II.5) under various constraints. [10]–[12] follow a similar approach. The gap caused by the difference between ex-anté and ex-post prices should be compensated by the utility, similar to what is done today for the gap between flat or TOU rates and ex-post LMPs. Some of the problems that can arise from this approach are as follows

- Perfect knowledge of the utility of the customers is assumed at every load bus (usually simple analytic functions), which is unrealistic, at least in the current situation
- If the price is set using an incorrect prediction the behavior of customers, prices that are posted may lead to system instability
- Demand is only assumed to be dependent on the current price. This assumption is not valid for the demand for electricity since electricity is not delivered instantly in packets and appliances need time to finish their jobs.
- Generation owners may try to arbitrage the market.
- At least the major part of generation assets should have a fully deterministic and controllable nature in order to let the utility calculate valid prices in advance. This assumption may not hold with the addition of a considerable amount of intermittent resources to the grid.

To conclude this section, it is important to remark that in addition to concerns on market stability, any power imbalance resulting from volatile demand, if not compensated correctly and timely, may accumulate and cause the grid itself to experience oscillation or even lose stability. This is why the socio-economic feedback of the energy market, with all the uncertainty it carries, cannot be closed in real time and some type of direct control may prove necessary, blurring the boundary between RTP and Direct Load Control (DLC).

III. Direct Load Control

Unlike RTP, Direct Load Control, or *Interruptible Load* programs, have been widely and successfully practiced for over a decade now. During peak load hours, utilities have the option to curtail the load due to certain appliances like air conditioners or water heater belonging to participating customers for a predetermined duration of time (usually, 15-30 minutes). This is done by sending a curtailment *signal* to the target appliance from a central dispatch center.

Ever since the 80's, several researchers have worked on finding an optimum curtailment schedule that will maximize the benefits of the utility while avoiding unacceptable dissatisfaction for participating customers. Typically, an optimization problem is defined that minimizes a certain cost function of the load during a look-ahead horizon. Common constraints are maximum curtailment time and a minimum pay-back period between two successive curtailments, during which the appliance is allowed to function without interruption, to catch up on its duty. This type of control is effective only for certain appliances. Many researchers then use Dynamic Programming methods to solve this optimization problem [14], [15], while others use some form of simplification to turn it into a Linear Program, which is less computationally intensive. DLC solutions usually try to have a long enough look-ahead horizon in order to avoid rebound peaks due to payback periods, since the modified load $L^{'}(t)$ at time t is given by,

$$\begin{aligned} L^{'}(t) = L(t) &- \text{Curtailed Load from DLC} \\ &+ \text{Payback Load from Previous Curtailments} \end{aligned}$$

Note that DLC schedules are sometimes solved for in coordination with the unit commitment problem (II.1) since they will affect the demand values D_i [16].

While interrupting certain loads can help alleviate the problems with high peak demand when facing generation shortage, most DLC solutions are designed merely for emergency situations. With the addition of a remarkable amount of unpredictable renewable resources like wind and solar energy to the grid, the frequency of these *emergency* situations will increase substantially and forced curtailment of the load, even if it is backed by customer participation, will no longer be a sufficient measure to match volatile and unpredictable generation with inelastic demand. Also, deciding how the customers should be paid for these interruptions will become increasingly unclear, bringing about similar complexity as RTP.

IV. Discussion

In section II, we saw that the problems with determining appropriate price signals are mainly due to the uncertainty in how customers respond to variable real-time prices, i.e. an unknown feedback behaviour. If, however, this feedback loop is somehow opened or its responses are based on settings that can be learned by a control center, the stated problems will no longer exist (or will be mitigated). What we propose is to add a certain intermediate mechanism to help reshape the original demand subject to a mutually agreed level of QoS.

A. *Remove the feedback loop with cellular scheduling*

If instead of making individual decisions, customer use electric loads that are capable of specifying their request for energy to a control center, releasing the control of their starting time to their associated neighborhood scheduler, the problematic feedback loop will be removed. The neighborhood scheduler will operate as a local energy retailer, which determines the activation time of the requesting smart loads such that it can shape the load profile to be as close as possible to the available generation supply (the day-ahead bid + the available local intermittent resources), avoiding the rebound peak problems that arise in DLC. We call this scheme Digital Direct Load Scheduling (DDLS). Detailed description of how the required transactions happen can be found in [17], [18].

B. Pricing based on known customer tailored response

If decisions that are made in response to price signals are based on simple strategies that can be announced in advance to a local control center, the uncertainty due to customer behaviour will be reduced. Moreover, if the control center is notified when each smart appliance is turned on, in a fashion similar to the DDLS strategy described in Section IV-A, the control center will be fully aware of the outcome of posting a price signal. However, the degrees of freedom in shaping the load are greatly reduced from what is possible in the DDLS, since customers in a single neighborhood will most likely receive the same real-time prices.

This strategy can lead to a more sophisticated DLC program and a more crude form of RTP. When a customer turns on his air conditioning unit, a simple message is communicated to the control center announcing the arrival of the smart AC device in the grid and its simple control strategy, like a table that will determine the functioning status of the AC depending on the current price signal. If all the AC units work under similar control strategies that can be described in a simple table, the control unit can distribute price signals based on its deterministic knowledge of customer responses and execute a more sophisticated DLC program, on a day-to-day basis.

C. Research opportunities

Prospective DSM technologies serve as bridging interfaces that coordinate the fluctuations of demand with opportunistically available energy supplies. On the demand side, real time information and statistical signal processing techniques are of paramount importance in both capturing the dynamics of the load, and accruing information efficiently from networks of distributed resources. On the generation side, the problem is not only that of unreliable predictions for volatile generation, but also that of the risks associated with this generation. These aspects have not sufficiently been stressed, as most of the pricing models we discussed do not make use of statistics other than the expected value of the random load/generation. Since DSM solutions are far from being settled on, it is clear that the best approach to understand what information support the problem needs is to think of scalable solutions to solve it.

V. Conclusion

In this paper, we gave an overview of envisioned load management programs for the Smart Grid. We concluded that, to match the strict reliability standards of the power grid, load control programs should address optimally the uncertainties due to customer behavior, volatile generation and market as well as grid stability. This brings new opportunities for statistical signal processing and optimization, which will be key to the success of DSM programs.

References

[1] "Primer on demand-side management with an emphasis on price-responsive programs," Charles River Associates, Tech. Rep., 2005. [Online]. Available: http://siteresources.worldbank.org/INTENERGY/Resources/PrimeronDemand-SideManagement.pdf

[2] J. Eto, C. Goldman, G. Heffner, B. Kirby, J. Kueck, M. Kintner-Meyer, J. Dagle, T. Mount, W. Schultze, R. Thomas, and R. Zimmerman, "Innovative developments in load as a reliability resource," in *Power Engineering Society Winter Meeting, 2002. IEEE*, vol. 2, 2002, pp. 1002 – 1004 vol.2.

[3] "Demand side management," Ministry of Power Government of India, Tech. Rep., 2011. [Online]. Available: http://www.powermin.nic.in/distribution/demand_side_management.htm

[4] M. Caramanis, R. Bohn, and F. Schweppe, "Optimal spot pricing: Practice and theory," *Power Apparatus and Systems, IEEE Transactions on*, vol. PAS-101, no. 9, pp. 3234 –3245, sept. 1982.

[5] W. Hogan, E. Read, and B. Ring, "Using mathematical programming for electricity spot pricing," *International Transactions in Operational Research*, vol. 3, no. 3-4, pp. 209–221, 1996. [Online]. Available: http://dx.doi.org/10.1111/j.1475-3995.1996.tb00048.x

[6] E. Litvinov, "Design and operation of the locational marginal prices-based electricity markets," *Generation, Transmission Distribution, IET*, vol. 4, no. 2, pp. 315 –323, february 2010.

[7] M. Huneault and F. D. Galiana, "A survey of the optimal power flow literature," *Power Systems, IEEE Transactions on*, vol. 6, pp. 762–770, may 1991.

[8] M. Roozbehani, M. A. Dahleh, and S. K. Mitter, "Volatility of power grids under real-time pricing," arXiv.org, Quantitative Finance Papers 1106.1401, Jun. 2011. [Online]. Available: http://ideas.repec.org/p/arx/papers/1106.1401.html

[9] J. Y. Choi, S.-H. Rim, and J.-K. Park, "Optimal real time pricing of real and reactive powers," *Power Systems, IEEE Transactions on*, vol. 13, no. 4, pp. 1226 –1231, nov 1998.

[10] M. Roozbehani, M. Dahleh, and S. Mitter, "Dynamic pricing and stabilization of supply and demand in modern electric power grids," in *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*, oct. 2010, pp. 543 –548.

[11] R. Sioshansi, "Evaluating the impacts of real-time pricing on the cost and value of wind generation," *Power Systems, IEEE Transactions on*, vol. 25, no. 2, pp. 741 –748, may 2010.

[12] P. Samadi, A. Mohsenian-Rad, R. Schober, V. Wong, and J. Jatskevich, "Optimal real-time pricing algorithm based on utility maximization for smart grid," in *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*, oct. 2010, pp. 415 –420.

[13] J. D. Glover and M. S. Sarma, *Power System Analysis and Design*. Brooks/Cole, 2001.

[14] A. Cohen and C. Wang, "An optimization method for load management scheduling," *Power Systems, IEEE Transactions on*, vol. 3, no. 2, pp. 612 –618, may 1988.

[15] K.-Y. Huang and Y.-C. Huang, "Integrating direct load control with interruptible load management to provide instantaneous reserves for ancillary services," *Power Systems, IEEE Transactions on*, vol. 19, no. 3, pp. 1626 – 1634, aug. 2004.

[16] 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.

[17] M. Alizadeh, A. Scaglione, and R. Thomas, "Direct load management of electric vehicles," in *The 36th International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011, pp. 5964–5967.

[18] M. Alizadeh, A. Scaglione, R. Thomas, and D. Callaway, "Information infrastructure for cellular load management in green power delivery systems," in *IEEE SmartGridComm 2011, to appear*.