

# Optimizations of Power Consumption and Supply in the Smart Grid: Analysis of the Impact of Data Communication Reliability

Dusit Niyato, *Member, IEEE*, Qiumin Dong, Ping Wang, *Member, IEEE*, and Ekram Hossain, *Senior Member, IEEE*

**Abstract**—Data communications infrastructure will play an important role to transfer various information in the smart grid. In this paper, we consider the reliability of the smart grid data communications infrastructure and its impact on the power consumption and supply optimizations. For optimizing the power consumption, we consider a deferrable load scheduling method which is modeled by using a constrained Markov decision process (CMDP) model, taking into account the unavailability of the home area network (HAN) and neighborhood area network (NAN) gateways. For optimizing the power supply, we consider an economic dispatch method which is modeled by using a stochastic programming (SP) model, taking into account the unavailability of the exact power demand and supply information. The power consumption and supply costs are analyzed. In addition, we show how these costs can be reduced through the deployment of component redundancy in the smart grid data communications infrastructure.

**Index Terms**—Smart grid data communications infrastructure, meter data management system (MDMS), deferrable load scheduling, economic dispatch, reliability analysis.

## I. INTRODUCTION

With advanced data communications technologies, the smart grid is emerging as the future power system which can operate in a cooperative, responsive, economical, and organic manner. One of the most important features of the smart grid is the meter data management system (MDMS) which collects, exchanges, and processes the meter data with an objective of minimizing the power consumption and supply costs. The MDMS can implement a demand response (DR) program through real-time pricing mechanism. The time-varying price can motivate the consumers to defer their power consumption to the off-peak period. Also, the MDMS can utilize power demand and supply information transferred from the consumers and distributed energy resource (DER) for the economic dispatch (ED) [1]. The power consumer and public utility can perform the optimizations of the demand response and economic dispatch, respectively. The power consumption and supply optimizations can achieve the minimum costs under the assumption that the MDMS has complete information, which is provided by the smart grid data communications infrastructure.

The smart grid data communications infrastructure is a special purpose data network to support various smart grid applications. The smart grid data communications infrastructure is generally composed of the home area network (HAN) and neighborhood area network (NAN) connected with each other in a hierarchical structure to transfer the smart grid-related data. The reliability of the smart grid data communications infrastructure is important for the smart grid applications to operate optimally. Although a few work in the literature analyzed the reliability of such a network, they did not analyze the impact of reliability on the power consumption and supply optimizations.

This paper considers the use of the smart grid data communications infrastructure to support power consumption and supply optimizations in the MDMS. Specifically, optimization models are developed for demand response by deferrable load scheduling and economic dispatch (i.e., power supply purchasing) considering the reliability of HAN and NAN. These optimization models enable us to analyze the impact of the reliability of smart grid data communications infrastructure on the costs of the power consumption (for the consumers) and power supply optimization (for the public utility). The contributions of this paper can be summarized as follows:

- *Optimization of power consumption:* The deferrable load scheduling problem is formulated as a constrained Markov decision process (CMDP) model. The CMDP model considers connectivity (i.e., reliability of communications) between the deferrable load through HAN and NAN gateways and the MDMS server which determines the availability of the price information in making scheduling decision. The optimal scheduling policies are obtained under complete and incomplete information cases.
- *Optimization of power supply:* The economic dispatch problem is formulated by using a stochastic programming (SP) model. The SP model considers two forms of information about the demand (e.g., from deferrable loads) and supply (e.g., from DERs). If the smart grid data communications infrastructure is available, the exact demand and the DER supply information can be used in the SP model. Otherwise, only demand and DER supply distributions, which can be estimated from the historical data, can be used. The optimal power purchasing strategy is obtained to minimize the power supply cost of the

D. Niyato, D. Qiumin, and P. Wang are with the School of Computer Engineering, Nanyang Technological University (NTU), Singapore (dniyato.S090052, wangping@ntu.edu.sg).

E. Hossain is with the Department of Electrical and Computer Engineering at University of Manitoba, Canada (ekram@ee.umanitoba.ca).

public utility.

The proposed optimization models follow the concept of *value of information analysis*. By having complete information, the optimal scheduling policy for the deferrable load and optimal power purchasing strategy can achieve the minimum costs for the consumer and the public utility, respectively. However, with incomplete information (e.g., due to unavailability of the smart grid data communications infrastructure), the optimal solutions may not be obtained. Therefore, the costs incurred due to the unreliability of the communications infrastructure can be quantified, which is the major usefulness of the analysis presented in this paper. Note that the problem of pricing of a public utility is not considered in this paper. The pricing is a different issue which should be investigated in a separate work.

From the energy efficiency perspective, the objective of the proposed scheme can be considered as to maximize the energy efficiency (i.e., maximize the utilization of power consumption and power generation). This objective can be achieved through the deferrable load scheduling and power supply optimization under uncertainty, in the power consumption and generation sides, respectively. For the consumption side, the energy efficiency is achieved in the sense that the power consumption will be delayed if the power price is high. The power price will be high when there is an insufficient supply of cheap power sources (e.g., renewable power sources with nearly zero variable cost for the fuel). For the supply side, the energy efficiency is achieved in the sense that the only necessary power will be generated given the unpredictable demand. This will reduce the cost of running expensive power generator (e.g., diesel power generator).

The rest of this paper is organized as follows. Section II presents a review of the related work. Section III presents the system model of MDMS and the smart grid data communications infrastructure. Section IV proposes the optimization formulation of the deferrable load scheduling as the CMDP model. The power supply optimization formulated as the SP model is presented in Section V. Section VI presents the performance evaluation. Section VII states the conclusion of the paper.

## II. RELATED WORK

### A. Power Consumption Optimization

Power consumption optimization (e.g., demand response) can be implemented to encourage users to shift or defer their power consumption to the off-peak period using an incentive from pricing. In [2], a Markov decision process (MDP) model was formulated to solve the scheduling problem of the deferrable load under an uncertain and time-varying power price. The objective is to minimize the expected cost at the consumer side. The state is the amount of power consumption units, while the actions are to wait or to start the load. The MDP can be solved using a backward induction method. In addition, by using the information about past and current prices, an algorithm based on a finite time horizon to find the decision threshold to operate the non-interruptible and interruptible loads with deadline constraint was proposed.

In [3], the multi-timescale power consumption scheduling problem was considered. A scheduling algorithm for traditional and opportunistic energy users was presented. The power demand from traditional energy users is uncertain and is assumed to be the function of a power price. The opportunistic energy users can compare the current power price with the acceptable price level, and opportunistic energy users wait if the power price is not at the desirable level. The scheduling algorithm was designed in two timescales (i.e., day-ahead conventional energy procurement and real-time retail market) to maximize the profit of the public utility. A multi-timescale Markov decision process (MMDP) was adopted to obtain the solution of the scheduling algorithm.

[4] formulated a binary particle swarm optimization (BPSO) model to schedule the interruptible loads. The objective is to minimize not only the cost but also the frequency of interruptions. The problem was found to be a nonlinear and discrete optimization problem. However, with the BPSO model, the near-optimal solution can be obtained with low computational complexity. In [5], an optimization model was proposed to schedule the power consumption of various loads given the power prices. The objective is to maximize the utility of the users. The utility is a function of the power consumption. The constraints of the model are the minimum daily energy consumption level, maximum and minimum hourly load levels, and ramping limits on such load levels. A robust optimization model was applied to address the price uncertainty. The model is transformed into a linear program which can be solved efficiently.

However, the above work did not consider the reliability of the smart grid data communications infrastructure to transfer price and other information. It was assumed that the necessary information is always available.

### B. Power Supply Optimization

Power supply optimization is a classical problem in a power system. A number of work in the literature studied this problem by using different optimization techniques. For example, [7] considered the optimal reactive power dispatch problem as a nonlinear multimodal and mixed-variable optimization problem. The objective is to minimize the active power loss in a transmission network with the constraint on meeting power demand. The seeker optimization algorithm (SOA) was proposed to obtain the optimal solution. A stochastic programming model was formulated for the unit commitment problem in [8]. The objective is to minimize the cost due to power system reliability taking the uncertainties of random outage of generators, transmission lines, and load forecasting inaccuracy into account. In addition, the loss-of-load-expectation (LOLE) was considered in the model as a constraint and used for calculating the cost of supplying the reserve.

[9] considered the cost minimization problem for the power supply from different generators to meet the power demand. The wind energy conversion system (WECS) was also included in the system under consideration. The wind power estimation was introduced in which the power costs due to overestimation and underestimation are to be minimized.

TABLE I: Notations

$\bar{\alpha}$	Average throughput of a deferrable load
$\mathbf{C}$	Transition matrix of “available” and “repair” states of smart grid communications infrastructure
$\mathcal{C}(\cdot, \cdot)$	Immediate cost function
$d$	Average delay of a job of a deferrable load
$D_{\max}$	Threshold of maximum delay of a job of a deferrable load
$\mathcal{D}(\cdot, \cdot)$	Immediate delay function
$\Delta$	Action space of a deferrable load
$E_{\text{det}}$	Deterministic DER supply from the connected DERs
$\epsilon$	Unavailability
$f_{\text{dem}}(\omega), f_{\text{sup}}(\gamma)$	Probabilities of demand and DER supply scenarios (i.e., $\omega$ and $\gamma$ , respectively)
$F_{\max}$	Threshold of maximum probability of job queue of a deferrable load to be full
$\mathcal{F}(\cdot, \cdot)$	Immediate probability function of full queue of a deferrable load
$\gamma, \Gamma$	Scenario of the aggregated DER supply of disconnected DERs, and DER supply scenario space, respectively
$\mathcal{J}_C, \mathcal{J}_D, \mathcal{J}_F$	Average cost of the deferrable load, average delay of a job, and probability of full queue
$\kappa$	Probability that the deferrable load will consume the power by taking action “run”
$L_{\text{det}}$	Aggregated deterministic demand from the connected loads
$MTBF, MTTR$	Mean time between failure and mean time to repair, respectively
$\mu$	Cost of deferrable load
$\omega, \Omega$	Scenario of power demand, and scenario space of demand, respectively
$p_{\max}, \bar{p}$	Maximum price state and average power price, respectively
$\mathbf{P}$	Price transition probability matrix
$\pi^*(l, a)$	Scheduling policy (i.e., probability of taking action $a$ when the current state is $l$ )
$\Psi_g, \Psi_{\text{pen}}$	Price of power purchased from generator $g$ , and penalty cost of not meeting power demand, respectively
$\phi(l, a)$	Stationary probability of state $l$ and action $a$
$q_{\max}$	Maximum size of a job queue
$\bar{q}$	Average number of jobs in the deferrable load
$\mathbf{Q}(\cdot)$	Transition probability matrix for all state of a deferrable load
$s_0, s_{\max}$	Initial and maximum operating stages, respectively
$\mathbf{S}^u(\cdot), \mathbf{S}^f(\cdot)$	Probability matrices for the transition of operating stage and wait states when the deferrable load is not at and is at the last operating stage, respectively
$\mathbf{T}^u(\cdot), \mathbf{T}^f(\cdot)$	Transition matrices for the connection, price, operating stage, and wait states
$\tau$	Power supply cost
$\Theta$	State space of the deferrable load
$\mathcal{U}$	Set of power consumption units
$w_{\max}$	Maximum wait state
$x_g$	Amount of power purchased from generator $g$
$y_{\omega, \gamma}$	Amount of under-supplied power and subject to penalty in scenarios $\omega$ and $\gamma$ for demand and DER supply, respectively

Similarly, [10] considered the wind power in the economic dispatch. However, the objective is to minimize the atmospheric emission (e.g., carbon oxides, oxides of nitrogen [NOx], and sulfur oxides [SOx]) from conventional generators. The optimization problem was formulated as a nonlinear problem.

### C. Reliability Analysis of Data Communications System for Power Systems

Reliability analysis, which is used to evaluate the capability of a system to operate normally, was performed for telecommunications systems including universal mobile telecommunications system (UMTS) [11], IEEE 802.11 [12], and sensor networks [13]. In [14], a reliability analysis of the supervisory control and data acquisition (SCADA) system for the power systems was presented. [15] analyzed the shortcomings of traditional communication and information systems used in a power grid and focused on the reliability and performance of real-time monitoring, operating control system, communication system, and delayed restoration which contributes to the power system failure. In [16], a reliability analysis for the wide-area measurement system (WAMS) based on the IEEE 14-bus system was presented. This WAMS is composed of a phasor measurement unit (PMU) as a sensor to measure the characteristics of the power delivery, a phasor data concentrator (PDC), and a control center (CC). In [16], the availability of

the connectivity from the PMU to CC was analyzed. A Markov chain model was proposed to capture the different failure and repair events of the optical-fiber system (OFS). In [17], a similar reliability analysis was performed for the WAMS with a centralized distributed model, where the measurements are collected from PMUs, aggregated by PDCs, and summarized and sent to a monitor center of the WAMS. A Markov chain was used in the analysis.

However, to the best of our knowledge, there is no reliability analysis which considers the power consumption and supply optimizations. The analysis of costs due to lack of complete price information and exact demand and supply information was not performed in the literature before. The relationship of the proposed scheme with existing methods can be summarized as follows. The proposed scheme extends the existing methods by considering the uncertainty and unavailability of the data/information used to optimize the power consumption and supply. The impact of such uncertainty and unavailability of the data communication infrastructure can be quantified as the cost of suboptimal solution (i.e., obtained with incomplete information) compared with the optimal solution (i.e., obtained with complete information).

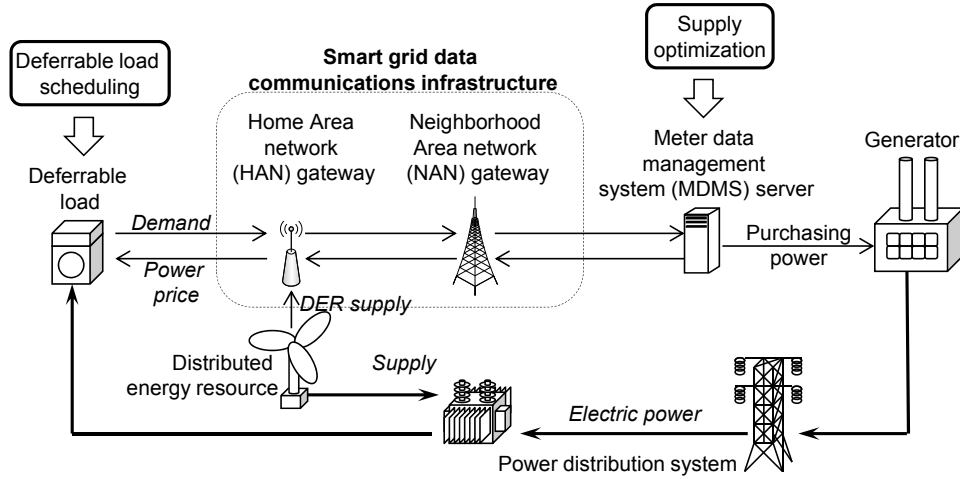


Fig. 1: System model of meter data management system (MDMS).

### III. SYSTEM MODEL: COMMUNICATIONS INFRASTRUCTURE, POWER CONSUMPTION, AND POWER SUPPLY IN SMART GRID

#### A. Meter Data Management System (MDMS)

We consider the model of the meter data management system (MDMS) shown in Fig. 1 [18]. The MDMS belongs to a public utility. The public utility purchases electric power from generators and delivers it to the consumers or users<sup>1</sup> through the power distribution system. The MDMS has a smart grid data communications infrastructure to facilitate the power delivery among the public utility, consumers, and generators. The main components in the system model under consideration are as follows.

- **Power Generator:** The public utility purchases the electric power from the power generators. Different power generators can offer power supply with different prices and can have different capacities.
- **Power Distribution System:** The public utility operates a power distribution system which provides a facility (e.g., transmission lines and power distributions substations) to transfer electric power from the generators to the consumers.
- **Distributed Energy Resource (DER):** The DER is a small power generator (e.g., wind turbine or solar panel) to provide an alternative electricity supply to the users through the power distribution system. The DER can be installed by the users or the public utility.
- **MDMS Server:** The MDMS server is a centralized controller and provides data storage for the MDMS. In the system model under consideration, the MDMS server sends the power price information to the consumers and maintains the demand information from the consumers. Based on the demand information, the MDMS server performs power supply optimization for purchasing power from the power generators.

<sup>1</sup>In the rest of this paper, the terms “consumer” and “user” are used interchangeably.

- **Smart Grid Data Communications Infrastructure:** The public utility uses a smart grid data communications infrastructure with the purpose to inform the power price information to the consumers and also to collect the power demand from consumers and the power supply from DERs. Machine-to-machine (M2M) communications can be used for the protocol stack [19] [20].
- **Deferrable Load:** The deferrable load is a special power consumption unit (e.g., dishwasher, washing machine, and computer) whose operation is divided into stages. The deferrable load does not need to operate continuously (i.e., interruptible) and it can be put in the wait state between two operating stages. A scheduling mechanism can be implemented such that the deferrable load can be operated with the minimum power consumption.

#### B. Reliability of the Smart Grid Communications Infrastructure

We consider a smart grid communications infrastructure (Fig. 1). A service area of the public utility is divided into communities. Each community has a neighborhood area network (NAN) gateway (e.g., cellular base station) to provide two-way data communications between the MDMS server and home area network (HAN) gateways [21]. The HAN gateway provides a communication channel among power consumption units (e.g., through a smart meter) and DER in a local area (e.g., a house). On the uplink direction, the HAN gateway can collect the power demand information from the power consumption units, power usage information from the smart meter, and power supply information from the DER. The power demand and supply information is reported to the MDMS server through NAN gateway. On the downlink direction, the MDMS server sends the power price to the consumers (i.e., power consumption units), through the NAN gateway and subsequently the HAN gateway. The smart grid data communications infrastructure is part of an advanced metering infrastructure (AMI) [22].

We consider the reliability of the smart grid communications infrastructure. We can define a Markov chain to model the state

(i.e., “available” and “repair”) of the HAN and NAN gateways. The transition probability matrix of the Markov chain can be defined as follows:

$$\mathbf{C} = \begin{bmatrix} 1 - 1/MTBF & 1/MTBF \\ 1/MTTR & 1 - 1/MTTR \end{bmatrix} \begin{array}{l} \rightarrow \text{“available”} \\ \rightarrow \text{“repair”} \end{array} \quad (1)$$

where  $MTBF$  and  $MTTR$  are the mean time between failure (MTBF) and mean time to repair (MTTR) [23], respectively. The mean time between failure (MTBF) is a basic measure of the reliability of the components and systems. The MTBF represents the average time period between failures of a component and a system during operation. The mean time to repair (MTTR) is a measure of the maintainability of the repairable components and systems. The MTTR represents the average time period needed to repair a failed component or system. The first and second rows of matrix  $\mathbf{C}$  correspond to the “available” and “repair” states, respectively. Let  $\mathbf{C}_{\text{NAN}}$  and  $\mathbf{C}_{\text{HAN}}$  denote the transition probability matrices of NAN and HAN gateways, respectively. The transition probability matrix of the connection from the HAN gateway and the NAN gateway to the MDMS server can be obtained from

$$\mathbf{C}_{\text{con}} = \mathbf{C}_{\text{NAN}} \otimes \mathbf{C}_{\text{HAN}} \quad (2)$$

where  $\otimes$  is the Kronecker product. The first row of  $\mathbf{C}_{\text{con}}$  corresponds to the state “connected” while the rest are “disconnected” (i.e., the first row corresponds to the case when both the HAN and NAN gateways are in the “available” states). Note that, in the rest of this paper, the term *connection* refers to the data transfer channel of a HAN gateway and a NAN gateway to the MDMS server. The connection could be either in “connected” or “disconnected” state.

We also consider the case that redundant NAN gateways are deployed. In this case, there will be two NAN gateways to provide the connection between HAN gateway and MDMS server. The transition probability matrix of the connection becomes

$$\mathbf{C}_{\text{con}} = \mathbf{C}_{\text{NAN}} \otimes \mathbf{C}_{\text{NAN}} \otimes \mathbf{C}_{\text{HAN}}. \quad (3)$$

In this case, the first, third, and fifth rows correspond to the state “connected” while the rest are “disconnected”. Specifically, the first, third, and fifth rows of the matrix  $\mathbf{C}_{\text{con}}$  defined in (3) are the cases that one of the NAN gateways is available and a HAN gateway is available.

We assume that a HAN gateway can be easily repaired when a failure occurs. Its redundancy is not considered in this paper. Nevertheless, if redundant HAN gateways exist, a model similar to that of the redundant NAN gateways can be applied.

### C. Power Consumption Optimization (Deferrable Load Scheduling Problem)

For the consumption side, the deferrable load can be scheduled based on a different state (e.g., power price). This is also known as the deferrable load scheduling problem. For example, the deferrable load can run and consume power when the price is low, and defer the operation when the price is high. It is assumed that real-time pricing (RTP) [24], [25] is used

by the public utility for the demand side management (DSM). That is, the electrical price can be changed dynamically, which motivates the users to avoid consuming power during the peak period [2]. The deferrable load scheduler (i.e., a component which runs the scheduling mechanism of a deferrable load) takes advantage of the price information transmitted through the smart grid communications infrastructure to optimize the scheduling policy. This optimal scheduling policy can be obtained by formulating and solving a constrained Markov decision process (CMDP) model. However, if the price information is not available to the deferrable load, the optimal scheduling policy will be affected. This case will be considered in the CMDP model as well. The details of the CMDP model will be presented in Section IV.

### D. Power Supply Optimization (Economic Dispatch)

For the supply side, the public utility can obtain the power supply from DERs or purchase electric power from generators. It is assumed that all the power from DERs will be supplied to the consumers. However, if the power from the DERs is not enough (i.e., the net demand in a community is more than the net DERs’ supply in the same community), the public utility purchases power from generators. This is referred to as the economic dispatch mechanism. In the economic dispatch, the public utility has to make contracts with the generators periodically (e.g., hour ahead market). However, if the power supply generated from DERs and purchased from the generators is not enough, the public utility has to pay a penalty according to the under-supplied power. Note that the penalty could be the fine charged by the government or the regulator if the public utility fails to meet the power delivery requirement or it could be the extra cost to run a special quick-start generator to meet the instantaneous power demand.

The MDMS server can obtain the exact demand information from the power consumption units (i.e., loads) and exact DER supply information from DERs through the smart grid communications infrastructure. With this exact demand and DER supply information, the public utility can purchase power supply from generators to meet the demand. However, if the exact demand and DER supply information is not available (e.g., due to failure of a NAN gateway), the public utility can use the historical demand data and DER supply data stored in the MDMS server to estimate the demand and DER supply distributions (i.e., probabilities of demand and DER supply). The distributions of demand and DER supply can be used to optimize the procurement of power from the power generators. In this case, a stochastic programming (SP) model is formulated with an objective of minimizing the cost (i.e., power cost and penalty). The details of the SP model will be presented in Section V.

### E. Relationship between Power Consumption and Power Supply Optimizations

The relationship between power consumption and supply optimizations is as follows. The MDMS server sends the price information to the deferrable load. The deferrable load optimizes the scheduling policy (i.e., consumption optimization).

Given this policy, the deferrable load sends the demand information back to the MDMS server. Also, the DER measures and sends the current capacity of the power supply to the MDMS server (Fig. 1). The demand and DER supply information is used for power supply optimization by the public utility to purchase power from the generators. However, the demand information does not need to be from the deferrable load only. Other loads can also send the demand information to the MDMS server. For example, a user can manually program the household appliance to operate only when the price is low.

#### IV. DEFERRABLE LOAD SCHEDULING: MODELING AND ANALYSIS UNDER COMMUNICATION UNRELIABILITY

In this section, we first consider the power consumption optimization, referred to as the deferrable load scheduling. The details of the deferrable load and the state and action spaces of the constrained Markov decision process (CMDP) model are described. The analysis under communication unreliability is performed by incorporating the availability and unavailability of the smart grid communications infrastructure into the state space of CMDP. As a result, the CMDP is solved for the optimal policy with the assumption that the power price is not perfectly known (i.e., when the communications infrastructure is unavailable). The cost of unavailable smart grid communications infrastructure is analyzed by applying the optimal policy of the deferrable load scheduling to the CMDP whose power price state is assumed to be perfectly known.

##### A. State and Action Spaces

We consider a deferrable load whose operation can be divided into  $s_{\max}$  operating stages. The operation of each operating stage can be deferred for the maximum of  $w_{\max}$  time periods [2]. The deferrable load may have a queue to keep the jobs from the consumer before running the jobs. The scheduler of the deferrable load can observe various states (e.g., power price information transmitted by the MDMS server through the NAN and HAN gateways) and make decision to run or to defer the deferrable load in different operating stages.

The state space of the deferrable load is defined as follows:

$$\Theta = \left\{ (S, \mathcal{W}, \mathcal{P}, \mathcal{C}, Q); S = \{s_0, \dots, s_{\max}\}, \quad (4) \right. \\ \left. \mathcal{W} = \{w_0, \dots, w_{\max}\}, \mathcal{P} = \{p_1, \dots, p_{\max}\}, \right. \\ \left. \mathcal{C} = \{\text{connected}, \text{disconnected}\}, Q = \{0, \dots, q_{\max}\} \right\}$$

where  $S$  is the operating stage of the deferrable load and  $S = s_0$  is the initial operating stage when the deferrable load retrieves a job from a queue and is ready to run.  $\{s_0, \dots, s_{\max}\}$  is the set of all operating stages.  $\mathcal{W}$  is the wait state (i.e., the number of time periods that the deferrable load has already deferred in the current operating stage).  $\{w_0, \dots, w_{\max}\}$  is the set of all wait states.  $\mathcal{P}$  is the price state and  $\{p_1, \dots, p_{\max}\}$  is the set of all price states.  $\mathcal{C}$  is the connection state (i.e., “connected” or “disconnected” with the MDMS server to obtain the price information).  $Q$  is the number of jobs in the deferrable load including the running job.  $q_{\max}$  is the maximum number of jobs in the deferrable load.

The general action space of the scheduler is  $\Delta = \{\text{run}, \text{defer}\}$ . Alternatively, the action space can be defined based on a state  $l \in \Theta$  as follows:

$$\Delta(l) = \begin{cases} \{\text{run}\}, & (w = w_{\max}) \text{ AND } (q > 0) \\ \{\text{defer}\}, & (q = 0) \\ \{\text{run}, \text{defer}\}, & \text{otherwise} \end{cases} \quad (5)$$

where a composite state  $l$  is defined as  $l = (s, w, p, c, q)$ . The state variables  $s, w, p, c,$  and  $q$  correspond to the states  $S, \mathcal{W}, \mathcal{P}, \mathcal{C},$  and  $Q$ , defined in (4), respectively. The first condition in (5) indicates that the deferrable load needs to be run when its maximum wait state is reached and there is a job in the deferrable load. The second condition indicates that if there is no job in the deferrable load, the only action is to defer the load. Otherwise, the scheduler has two actions, i.e., to run or to defer an operation of the deferrable load. An example of a state transition diagram of the deferrable load excluding queue and connection state transitions is shown in Fig. 2.

##### B. Transition Probability Matrix

The transition probability matrix of the deferrable load can be derived based on the action.

1) *Action “run”*: Let  $\mathbf{S}^u(\text{run})$  and  $\mathbf{S}^f(\text{run})$  denote the probability matrices for the transition of operating stage and wait states when the deferrable load is not at and is at the last operating stage, respectively. The rows of these matrices correspond to the states  $(s_0, w_0), (s_0, w_1), \dots, (s_0, w_{\max}), (s_1, w_0), \dots, (s_{\max}, w_{\max})$ . The elements of these matrices can be obtained as in (6)-(7), where  $S_{(s_i, w_j), (s_{i'}, w_{j'})}^u(\text{run})$  and  $S_{(s_i, w_j), (s_{i'}, w_{j'})}^f(\text{run})$  are the transition probabilities from state  $(s_i, w_j)$  to state  $(s_{i'}, w_{j'})$  when the deferrable load is not at and is at the last operating stage, respectively.  $\theta(s_i, w_j)$  is the probability that the deferrable load finishes operating stage  $s_i$  in a time period. The probabilities defined in (6) and (7) indicate transition to the next operating stage and starts at wait state  $w_0$ . When the deferrable load finishes all operating stages, the scheduler returns to the original state  $(s_0, w_0)$ .

Then, we combine the transition of connection and price states. Let  $\mathbf{P}$  denote the price transition probability matrix [6]. The transition matrices for the connection, price, operating stage, and wait states are obtained as follows:

$$\mathbf{T}^u(\text{run}) = \begin{bmatrix} \cdots & [\mathbf{C}_{\text{con}}]_{\text{connected}} \mathbf{P} \otimes \mathbf{S}^u(\text{run}) & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & [\mathbf{C}_{\text{con}}]_{\text{disconnected}} \mathbf{I} \otimes \mathbf{S}^u(\text{run}) & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & [\mathbf{C}_{\text{con}}]_{\text{disconnected}} \mathbf{I} \otimes \mathbf{S}^u(\text{run}) & \cdots \end{bmatrix} \quad (8)$$

$$\mathbf{T}^f(\text{run}) = \begin{bmatrix} \cdots & [\mathbf{C}_{\text{con}}]_{\text{connected}} \mathbf{P} \otimes \mathbf{S}^f(\text{run}) & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & [\mathbf{C}_{\text{con}}]_{\text{disconnected}} \mathbf{I} \otimes \mathbf{S}^f(\text{run}) & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & [\mathbf{C}_{\text{con}}]_{\text{disconnected}} \mathbf{I} \otimes \mathbf{S}^f(\text{run}) & \cdots \end{bmatrix} \quad (9)$$



$$S_{(s_i, w_j), (s_{i'}, w_{j'})}^u(\text{defer}) = \begin{cases} 1, & ((s_{i'} = s_i) \text{ AND } (w_{j'} = w_{j+1}) \text{ AND } (s_i \neq s_{\max}) \text{ AND } (w_j \neq w_{\max})) \text{ OR} \\ & ((s_{i'} = s_{i+1}) \text{ AND } (w_{j'} = w_0) \text{ AND } (s_i \neq s_{\max}) \text{ AND } (w_j \neq w_{\max})) \\ 0, & \text{otherwise} \end{cases} \quad (17)$$

$$S_{(s_i, w_j), (s_{i'}, w_{j'})}^f(\text{defer}) = \begin{cases} 1, & (s_i = s_{\max}) \text{ AND } (w_j = w_{\max}) \text{ AND } (s_{i'} = s_0) \text{ AND } (w_{j'} = w_0) \\ 0, & \text{otherwise.} \end{cases} \quad (18)$$

the complete transition probability matrix  $\mathbf{Q}(\text{defer})$  similar to that in (10) is obtained. In summary, the difference between the actions “run” and “defer” is at the operating stage and wait state transitions (i.e.,  $\mathbf{S}^u(\cdot)$  and  $\mathbf{S}^f(\cdot)$ ).

### C. Optimization Formulation

A stochastic optimization problem (i.e., CMDP model) can be formulated to obtain an optimal scheduling policy. This policy is to achieve the minimum long-term average cost of the deferrable load given that the average delay of a job and probability that the deferrable load has the maximum number of jobs (i.e., a queue is full) are maintained at the thresholds (i.e.,  $D_{\max}$  and  $F_{\max}$ , respectively). The average cost of the deferrable load  $\mathcal{J}_C$ , average delay of a job  $\mathcal{J}_D$ , and probability of full queue  $\mathcal{J}_F$  are defined as follows:

$$\mathcal{J}_C = \limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{t'=1}^t \mathbb{E}(\mathcal{C}(\mathcal{L}_{t'}, \mathcal{A}_{t'})) \quad (19)$$

$$\mathcal{J}_D = \limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{t'=1}^t \mathbb{E}(\mathcal{D}(\mathcal{L}_{t'}, \mathcal{A}_{t'})) \quad (20)$$

$$\mathcal{J}_F = \limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{t'=1}^t \mathbb{E}(\mathcal{F}(\mathcal{L}_{t'}, \mathcal{A}_{t'})) \quad (21)$$

where  $\mathcal{L}_{t'} \in \Theta$  and  $\mathcal{A}_{t'} \in \Delta$  are the state and action variables, respectively, for the deferrable load at time  $t'$ .  $\mathbb{E}(\cdot)$  denotes the expectation.  $\mathcal{C}(l, a)$ ,  $\mathcal{D}(l, a)$ , and  $\mathcal{F}(l, a)$ , where  $l \in \Theta$  and  $a \in \Delta$ , denote the immediate cost function, immediate delay function, and immediate probability function of full queue, respectively.

The immediate cost function of the deferrable load is due to the power consumption and can be defined based on the state  $l = (s, w, p, c, q)$  and action  $a$  as in (22) where  $\beta_{\text{run}}(s)$  is the power consumption when the deferrable load runs for operating stage  $s$ .  $\beta_{\text{wait}}(s, w)$  is the power consumption when the deferrable load waits at operating stage  $s$  and the wait state is  $w$ .  $p$  is the power price, which is known if the connection state is “connected”. However, if the connection state is “disconnected”, the average price  $\bar{p}$  is assumed by the deferrable load regardless of the price state. Let  $\vec{\eta}$  denote the stationary probability vector of the price state.  $\vec{\eta}$  can be obtained by solving  $\vec{\eta}^T \mathbf{P} = \vec{\eta}^T$  and  $\vec{\eta}^T \mathbf{1} = 1$ , where  $\mathbf{1}$  is a vector (with an appropriate size) of ones. The average price can be obtained from  $\bar{p} = \vec{\eta}^T [p_1, \dots, p_{\max}]^T$ . In (22), the deferrable load consumes zero power if there is no waiting job.

The immediate delay function can be expressed as follows:

$$\mathcal{D}(l, a) = q/\alpha. \quad (23)$$

In the long term (i.e.,  $t \rightarrow \infty$ ), the numerator of the delay function  $\mathcal{D}(l, a)$  becomes the average number of jobs in the deferrable load. Therefore, the long term average of the immediate delay function is given by the Little’s law.

The immediate probability of full queue is obtained from

$$\mathcal{F}(l, a) = \begin{cases} 1, & q = q_{\max} \\ 0, & \text{otherwise.} \end{cases} \quad (24)$$

Next, the CMDP model is solved to obtain the scheduling policy denoted by  $\pi^*(l, a)$ , which is the probability of taking action  $a$  when the current state is  $l$  (i.e., randomized policy). Note that this scheduling policy is obtained under an incomplete information case, since when the connection state of the deferrable load to the MDMS server is “disconnected”, the price state cannot be observed. Therefore, this scheduling policy  $\pi^*(l, a)$  is optimal in the sense of incomplete information only. The details of this incomplete information case will be elaborated in Section IV-E.

The CMDP model can be expressed as follows:

$$\min_{\pi} \mathcal{J}_C(\pi) \quad (25)$$

$$\text{s.t. } \mathcal{J}_D(\pi) \leq D_{\max} \quad (26)$$

$$\mathcal{J}_F(\pi) \leq F_{\max} \quad (27)$$

where the average cost of the deferrable load, average delay of a job, and probability of full queue are defined as functions of the policy  $\pi$ , i.e.,  $\mathcal{J}_C(\pi)$ ,  $\mathcal{J}_D(\pi)$ , and  $\mathcal{J}_F(\pi)$ , respectively. To obtain the scheduling policy, the CMDP model can be transformed into an equivalent linear programming (LP) model [26]. In this case, there is a one-to-one mapping between the solution denoted by  $\phi^*(l, a)$  of the LP model and the scheduling policy denoted by  $\pi^*(l, a)$  of the CMDP model. Let  $\phi(l, a)$  denote the stationary probability of state  $l$  and action  $a$ . The equivalent LP model can be expressed as follows:

$$\mu = \min_{\phi(l, a)} \sum_{l \in \Theta} \sum_{a \in \Delta} \phi(l, a) \mathcal{C}(l, a) \quad (28)$$

$$\text{s.t. } \sum_{l \in \Theta} \sum_{a \in \Delta} \phi(l, a) \mathcal{D}(l, a) \leq D_{\max} \quad (29)$$

$$\sum_{l \in \Theta} \sum_{a \in \Delta} \phi(l, a) \mathcal{F}(l, a) \leq F_{\max} \quad (30)$$

$$\phi(l, a = \text{defer}) = 0, \quad (w = w_{\max}) \text{ AND } q > 0 \quad (31)$$

$$\phi(l, a = \text{run}) = 0, \quad q = 0 \quad (32)$$

$$\sum_{a \in \Delta} \phi(l', a) = \sum_{l \in \Theta} \sum_{a \in \Delta} \phi(l, a) Q(l'|l, a), \quad l' \in \Theta \quad (33)$$

$$\sum_{l \in \Theta} \sum_{a \in \Delta} \phi(l, a) = 1, \quad \phi(l, a) \geq 0 \quad (34)$$

$$\mathcal{C}(l, a) = \begin{cases} p\beta_{\text{run}}(s), & ((a = \text{run}) \text{ OR } (w = w_{\text{max}})) \text{ AND } (c = \text{connected}) \\ p\beta_{\text{wait}}(s, w), & (a = \text{defer}) \text{ AND } (w \neq w_{\text{max}}) \text{ AND } (c = \text{connected}) \\ \bar{p}\beta_{\text{run}}(s), & ((a = \text{run}) \text{ OR } (w = w_{\text{max}})) \text{ AND } (c = \text{disconnected}) \\ \bar{p}\beta_{\text{wait}}(s, w), & (a = \text{defer}) \text{ AND } (w \neq w_{\text{max}}) \text{ AND } (c = \text{disconnected}) \\ 0, & q = 0. \end{cases} \quad (22)$$

where  $Q(l'|l, a)$  denotes the probability that action  $a$  is taken and the state changes from  $l$  to  $l'$ . This probability is the element of matrix  $\mathbf{Q}(\cdot)$  (e.g., as obtained from (10) for the action “run”). The objective in (28) is to minimize the average cost of the deferrable load. The constraints in (29) and (30) are to maintain the delay and probability of full queue at the thresholds, respectively. The constraint in (31) indicates that the action “defer” is not feasible if the maximum wait state is reached when there is a job in the deferrable load. The constraint in (32) indicates that the action “run” is not feasible if there is no job. The constraint in (33) satisfies the Chapman-Kolmogorov equation.

The scheduling policy, which is the probability of taking a particular action at a particular state, can be obtained from the solution  $\phi^*(l, a)$  of the above LP model. The scheduling policy for the deferrable load is obtained from

$$\pi^*(l, a) = \frac{\phi^*(l, a)}{\sum_{a' \in \Delta} \phi^*(l, a')}, \text{ for } l \in \Theta \text{ and } \sum_{a' \in \Delta} \phi^*(l, a') > 0. \quad (35)$$

If  $\sum_{a' \in \Delta} \phi^*(l, a') = 0$ , the specific action “defer” is taken. Note again that this scheduling policy is optimal for the incomplete information case.

#### D. Performance Measures

Various performance measures can be obtained for the deferrable load given the optimal scheduling policy. The average number of jobs in the deferrable load can be obtained from

$$\bar{q} = \sum_{l \in \Theta} \sum_{a \in \Delta} \phi^*(l, a) q \quad (36)$$

where  $l = (s, w, p, c, q)$ . The throughput of the deferrable load (i.e., the average number of finished jobs per time period) can be obtained from

$$\bar{\alpha} = \alpha \left( 1 - \sum_{q=q_{\text{max}}} \sum_{a \in \Delta} \phi^*(l, a) \right) \quad (37)$$

where  $\sum_{q=q_{\text{max}}} \sum_{a \in \Delta} \phi^*(l, a)$  is the probability that the queue of the deferrable load is full. The average delay of a job can be obtained from the Little’s law as follows:

$$\bar{d} = \bar{q} / \bar{\alpha}. \quad (38)$$

The probability that the deferrable load will consume power by taking action “run” is obtained from

$$\kappa = \sum_{l \in \Theta} \sum_{a=\text{run}} \phi^*(l, a). \quad (39)$$

This probability can be used as the statistical information for the MDMS server to optimize the power supply. The details of the power supply optimization will be presented in Section V.

#### E. Cost of the Unavailability of the Smart Grid Communications Infrastructure

The optimal scheduling policy  $\pi^*(l, a)$  is obtained from the fact that the connection between the deferrable load and the MDMS server may not be available. As a result, the price state transition is not taken into account when the connection state is “disconnected”. This is referred to as the incomplete information case. To analyze the cost incurred due to the unavailability of the smart grid data communications infrastructure, we start by deriving the optimal cost of the deferrable load when the price information is available (i.e., complete information case). To obtain the optimal cost when a connection state is always “connected”, the immediate cost function defined in (22) becomes

$$\mathcal{C}^*(l, a) = \begin{cases} p\beta_{\text{run}}(s), & (a = \text{run}) \text{ OR } (w = w_{\text{max}}) \\ p\beta_{\text{wait}}(s, w), & (a = \text{defer}) \text{ AND } (w \neq w_{\text{max}}) \\ 0, & q = 0. \end{cases} \quad (40)$$

Also, the transition probability matrix  $\mathbf{Q}(\cdot)$  has to be modified. Specifically, the elements of matrices  $\mathbf{T}^u(\cdot)$  and  $\mathbf{T}^f(\cdot)$  obtained in (8) and (9), respectively, will be  $[\mathbf{C}_{\text{con}}]_{\text{connected}} \mathbf{P} \otimes \mathbf{S}^u$ ,  $[\mathbf{C}_{\text{con}}]_{\text{disconnected}} \mathbf{P} \otimes \mathbf{S}^u$ ,  $[\mathbf{C}_{\text{con}}]_{\text{connected}} \mathbf{P} \otimes \mathbf{S}^f$ , and  $[\mathbf{C}_{\text{con}}]_{\text{disconnected}} \mathbf{P} \otimes \mathbf{S}^f$  (i.e., price state can be always transmitted and  $\mathbf{I}$  is not used). Let the modified transition probability matrix of  $\mathbf{Q}(\cdot)$  be denoted by  $\tilde{\mathbf{Q}}(\cdot)$ . The elements of this matrix  $\tilde{\mathbf{Q}}(\cdot)$  are denoted by  $\tilde{Q}(l'|l, a)$  (i.e., probability that the action  $a$  is taken and the state changes from  $l$  to  $l'$  in the complete information case). The optimal scheduling policy can be obtained from the optimal solution of the LP problem defined as follows:

$$\begin{aligned} \mu^* &= \min_{\phi(l, a)} \sum_{l \in \Theta} \sum_{a \in \Delta} \phi(l, a) \mathcal{C}^*(l, a) \\ &\text{s.t.} \quad (29) - (34) \end{aligned} \quad (41)$$

where the immediate cost function as defined in (40) is applied. The solution of this LP problem is denoted as  $\phi^*(l, a)$  (i.e., stationary probability). The optimal scheduling policy  $\pi^*(l, a)$  can be obtained similar to that in (35).  $\mu^*$  is the optimal cost in the complete information case.

Next, we want to evaluate the cost of the scheduling policy  $\pi^*(l, a)$  obtained from the incomplete information case in the complete information environment. In other words, we want to quantify the actual cost due to using the scheduling policy  $\pi^*(l, a)$ . This cost will be suboptimal due to random failure of the HAN and NAN gateways (i.e., the unavailability of the smart grid data communications infrastructure). First, the stationary probability denoted by  $\delta(l)$  of state  $l \in \Theta$  is

obtained by solving the following equations:

$$\delta(l') = \sum_{l \in \Theta} \delta(l) \sum_{a \in \Delta} \tilde{Q}(l'|l, a) \pi^*(l, a) \text{ and } \sum_{l \in \Theta} \delta(l) = 1, \forall l' \in \Theta. \quad (42)$$

The actual cost of the scheduling policy  $\pi^*(l, a)$  is obtained from

$$\mu^\dagger = \sum_{l \in \Theta} \sum_{a \in \Delta} \pi^*(l, a) \delta(l) \mathcal{C}^*(l, a). \quad (43)$$

In summary, if the connection state between the deferrable load and the MDMS server is always “connected”, the cost of the deferrable load will be  $\mu^*$  (i.e., optimal cost in a complete information case). However, if the connection state is “disconnected” due to the failure of HAN and NAN gateways, the actual cost will be  $\mu^\dagger$ . Therefore, in the power consumption optimization, the difference between the costs in the incomplete and complete information cases can be obtained from

$$\hat{\mu} = \mu^\dagger - \mu^*. \quad (44)$$

## V. POWER SUPPLY OPTIMIZATION: MODELING AND ANALYSIS UNDER COMMUNICATION UNRELIABILITY

In this section, we analyze the power supply optimization for the MDMS of the public utility. First, a stochastic programming (SP) formulation [27] is presented. The solution of the SP formulation is obtained under the demand uncertainty. Then, the impact of the unavailability of the smart grid communications infrastructure on the power supply optimization is analyzed in terms of cost incurred to the utility company. The cost is obtained by applying the solution of the SP formulation to when the demand information cannot perfectly observed.

### A. Stochastic Programming Formulation

We assume that the public utility has two sources of power supply, i.e., from distributed energy resources (DERs) and generators. Without loss of generality, the power supply from a DER will be transmitted to the power distribution system before delivered to the power consumption units (i.e., loads). In the economic dispatch, the public utility optimizes the power purchased from generators such that the total cost is minimized (i.e., power supply optimization). However, during the economic dispatch, there are stochastic parameters that the public utility has to take into account.

- *Power supply*: The power supply from a DER may not be exactly known by the MDMS server if the communications infrastructure between DER and the MDMS server is not available (e.g., due to failure of HAN and NAN gateways).
- *Power demand*: The power demand may not be exactly known by the MDMS server due to two factors. First, the load can be random (i.e., random load). The operation of random load is unpredictable (e.g., users can turn on lights anytime). Second, the load (e.g., deferrable load) fails to report its demand information to the MDMS server, again due to the unavailability of the smart grid data communications infrastructure between load and the MDMS server.

Therefore, the public utility can use the SP model to cope up with the uncertainties to achieve the minimum power supply cost.

The SP model of the economic dispatch problem for the public utility can be formulated as follows:

$$\tau = \min_{x_g, y} \Psi_g x_g + \mathbb{E}(\Psi_{\text{pen}} y) \quad (45)$$

$$\text{s.t. } x_g + y \geq (L_{\text{det}} + \mathcal{L}) - (E_{\text{det}} + \mathcal{E}) \quad (46)$$

$$x_g \leq G_g \quad (47)$$

$$x_g, y \geq 0 \quad (48)$$

where  $\Psi_g$  and  $x_g$  are the price and amount of power purchased from generator  $g$ , respectively.  $G_g$  is the capacity of generator  $g$ .  $\Psi_{\text{pen}}$  is the penalty cost.  $y$  is the amount of under-supplied power which is subject to penalty.  $L_{\text{det}}$  is the aggregated deterministic demand from the connected loads (e.g., deferrable load is able to report the exact demand information to the MDMS server).  $\mathcal{L}$  is a random parameter for the aggregated demand from the disconnected loads and random loads.  $E_{\text{det}}$  is the deterministic DER supply from the connected DERs.  $\mathcal{E}$  is the random parameter for the aggregated DER supply from the disconnected DERs. The objective in (45) is to minimize the expected total power supply cost of the public utility. The constraint in (46) represents the balance of power demand and supply.

It is important to ensure that the constraint in (46) is always satisfied given the random parameters  $\mathcal{L}$  and  $\mathcal{E}$ . Therefore, in the SP model, the scenarios of the random parameters (i.e., realizations of the random parameters) are considered. The scenario of the aggregated demand of disconnected loads and random loads is denoted by  $\omega \in \Omega$  (i.e., for random parameter  $\mathcal{L}$ ) where  $\Omega$  is the demand scenario space. Likewise, the scenario of the aggregated DER supply of disconnected DERs is denoted by  $\gamma \in \Gamma$  (i.e., for random parameter  $\mathcal{E}$ ) where  $\Gamma$  is the DER supply scenario space. The probabilities of demand and DER supply scenarios  $\omega$  and  $\gamma$  are denoted by  $f_{\text{dem}}(\omega)$  and  $f_{\text{sup}}(\gamma)$ , respectively. To meet the constraint defined in (46), the decision variable  $y$  is associated with the scenarios of demand and DER supply denoted by  $y_{\omega, \gamma}$ . In particular,  $y_{\omega, \gamma}$  is the amount of under-supplied power and subject to penalty in scenarios  $\omega$  and  $\gamma$  for demand and DER supply, respectively. Note that the probability of demand for the deferrable load can be obtained from (39). For example,  $f_{\text{dem}}(\omega_1) = \kappa$  and  $f_{\text{dem}}(\omega_0) = 1 - \kappa$  are the probability that the deferrable load consumes and does not consume power, which correspond to scenarios  $\omega_1$  and  $\omega_0$ , respectively.

Let  $\mathbb{U}$  denote the set of power consumption units (i.e., loads) whose exact demand information is not known by the MDMS server, where  $\mathbb{U} = \mathbb{U}_{\text{dis}} \cup \mathbb{U}_{\text{ran}}$ . Specifically, the set of loads whose exact demand information is not known is composed of the set of disconnected load  $\mathbb{U}_{\text{dis}}$  (e.g., due to unavailability of the data communications infrastructure) and the set of random loads  $\mathbb{U}_{\text{ran}}$ . Let  $\mathbb{R}$  denote the set of DERs whose exact supply information is not known by the MDMS server.

Now, we can transform the SP model defined in (45)-(48) to the deterministic equivalent as follows:  $\tau^\dagger =$

$$\begin{aligned} \min_{x_g, y_{\omega, \gamma}} \quad & \Psi_g x_g + \sum_{\omega \in \Omega(\mathbb{U})} \sum_{\gamma \in \Gamma(\mathbb{R})} f_{\text{dem}}(\omega) f_{\text{sup}}(\gamma) \Psi_{\text{pen}} y_{\omega, \gamma} \\ \text{s.t.} \quad & x_g + y_{\omega, \gamma} \geq (L_{\text{det}} + L_{\omega}) - (E_{\text{det}} + E_{\gamma}) \quad (49) \\ & x_g \leq G_g \quad (50) \\ & x_g, y_{\omega, \gamma} \geq 0. \quad (51) \end{aligned}$$

Note here the scenario space of the demand from disconnected loads and random load, and the scenario space of disconnected DER supply are defined as the corresponding sets of loads and DERs whose exact demand and DER supply information, i.e.,  $\Omega(\mathbb{U})$  and  $\Gamma(\mathbb{R})$ , respectively, are not known.

The scenario space of demand of disconnected loads and random loads is defined as  $\Omega = \bigcup_{u \in \mathbb{U}} \Omega_u$ , where  $u \in \mathbb{U}$  is the load whose exact demand information is not known by the MDMS server and  $\Omega_u$  is the scenario space of such a load. Similarly, the scenario space of DER supply of disconnected DERs is defined as  $\Gamma = \bigcup_{r \in \mathbb{R}} \Gamma_r$ , where  $r \in \mathbb{R}$  is the DER whose exact supply information is not known by the MDMS server and  $\Gamma_r$  is the scenario space of such a DER. The probability distribution of the aggregated demand of disconnected loads and random loads can be estimated (e.g., using methods presented in [28], [29]) and derived as follows:

$$f_{\text{dem}}(\omega) = \prod_{u \in \mathbb{U}} f_{\text{dem}}(\omega_u), \quad \forall \omega \in \Omega \quad (52)$$

where  $f_{\text{dem}}(\omega_u)$  is the probability of demand scenario  $\omega_u$  of the disconnected or random load  $u$ . The aggregated demand is obtained from  $L_{\omega} = \sum_{u \in \mathbb{U}} L_{\omega_u}$  for all  $\omega \in \Omega$ , where  $L_{\omega_u}$  is the demand of a scenario of the disconnected or random load  $u$ . Similarly, the probability of the aggregated supply of disconnected DER can be obtained from

$$f_{\text{sup}}(\gamma) = \prod_{r \in \mathbb{R}} f_{\text{sup}}(\gamma_r), \quad \forall \gamma \in \Gamma \quad (53)$$

where  $f_{\text{sup}}(\gamma_r)$  is the probability of a supply scenario  $\gamma_r$  of the disconnected DER  $r$ . The aggregated supply is obtained from  $E_{\gamma} = \sum_{r \in \mathbb{R}} E_{\gamma_r}$   $\forall \gamma \in \Gamma$ , where  $E_{\gamma_r}$  is the supply of scenario of the disconnected DER  $r$ .

### B. Cost of the Unavailability of the Smart Grid Communications Infrastructure

Without a failure of HAN and NAN gateways, the random parameter in the economic dispatch problem is just the demand of random loads. In this case, the MDMS server has exact demand and supply information from loads and DERs, respectively. The optimal power supply strategy of the public utility can be obtained by solving the SP model defined as follows:

$$\min_{x_g, y_{\omega}} \quad \Psi_g x_g + \sum_{\omega \in \Omega(\mathbb{U}_{\text{ran}})} f_{\text{dem}}(\omega) \Psi_{\text{pen}} y_{\omega} \quad (54)$$

$$\text{s.t.} \quad x_g + y_{\omega} \geq (L_{\text{det}} + L_{\omega}) - E_{\text{det}} \quad (55) \\ (50) - (51).$$

In the long term, the optimal power supply cost in the complete information case (i.e., if the MDMS server always

knows the exact demand information from all loads and exact supply information from all DERs) can be obtained from

$$\tau^* = \min_{x_g, y} \quad \Psi_g x_g + \Psi_{\text{pen}} y \quad (56)$$

$$\text{s.t.} \quad x_g + y \geq \mathbb{E}(\mathcal{L}_{\text{tot}}) - \mathbb{E}(\mathcal{E}_{\text{tot}}) \quad (57) \\ \text{and (47) - (48)}$$

where  $\mathbb{E}(\mathcal{L}_{\text{tot}})$  and  $\mathbb{E}(\mathcal{E}_{\text{tot}})$  are the expected total demand and expected DER supply, respectively. In particular  $\tau^*$  is the expected optimal power supply cost considering the complete information case.

The difference between costs of the public utility when the MDMS server does not have and when the MDMS server does have the exact demand and DER supply information can be obtained from

$$\hat{\tau} = \left( \tau^\dagger \epsilon + \tau^* (1 - \epsilon) \right) - \tau^* \quad (58)$$

where  $\epsilon$  is the unavailability of the smart grid data communications infrastructure. The unavailability can be obtained from

$$\epsilon = \sum [\vec{v}]_{\text{disconnected}} \quad (59)$$

where  $[\vec{v}]_{\text{disconnected}}$  is the element of vector  $\vec{v}$  associated with ‘‘disconnected’’ state. The vector  $\vec{v}$  can be obtained by solving  $\vec{v}^T \mathbf{C}_{\text{con}} = \vec{v}^T$  and  $\vec{v}^T \mathbf{1} = 1$ .  $\mathbf{C}_{\text{con}}$  is the connection state transition probability matrix of the smart grid data communications infrastructure (e.g., as shown in (2) and (3)).

## VI. PERFORMANCE EVALUATION

### A. Numerical Results

1) *Consumer side – Deferrable load scheduling*: For the deferrable load scheduling, we consider a deferrable load with 4 operating stages which can wait for 3 time periods. The length of a time period is 15 minutes. The deferrable load can have 4 jobs waiting in a queue. The job is generated with probability 0.1 in each time period. The probability threshold for the deferrable load to have maximum 4 jobs waiting in a queue is 0.01. The delay threshold of a job is 170 minutes. When the deferrable load operates, it consumes power of 1 kWh in all operating stages. On the other hand, when the deferrable load defers an operation, it consumes 10% of that in the run state. The probability of the deferrable load to finish running stage  $s$  in a time period is one. For the public utility, there are 3 price levels, i.e., 5, 10, and 15 cents per kWh. We consider the uniform price transition (i.e., probabilities of price 5, 10, and 15 cents per kWh are equal, i.e., about 0.333). The mean time between failure (MTBF) and the mean time to repair (MTTR) of a HAN gateway are 15,000 hours (i.e., 1 failure per two years) and 1 hours, respectively. The MTBF and MTTR of a NAN gateway are 8,780 hours (i.e., 1 failure per year) and 48 hours, respectively.

We compare different scheduling policies for the deferrable load, i.e., ‘‘optimal’’, ‘‘always defer’’, and ‘‘always run’’. The optimal policies are obtained from the CMDP model, in which we consider two cases, with and without price information. In the ‘‘always run’’ policy, the scheduler always runs the

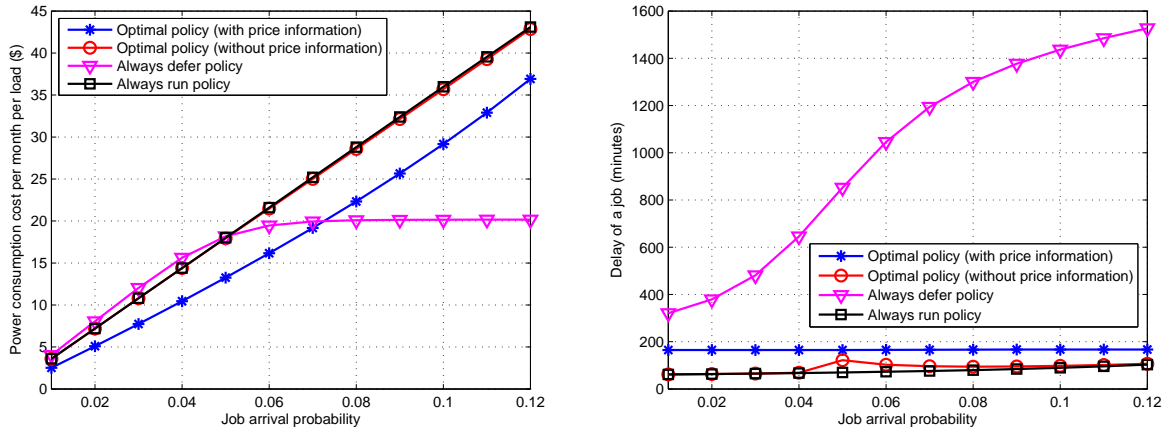


Fig. 3: (a) Power consumption cost of a deferrable load and (b) delay of a job under varied job arrival probabilities and different scheduling policies.

deferrable load. In the “always defer” policy, the scheduler always defers the deferrable load, except when the last wait state is reached. Fig. 3(a) shows the power consumption cost of the deferrable load. When the job arrival probability is small (i.e., less than 0.07), clearly, the optimal policy with price information achieves the lowest power consumption cost, since the scheduler can defer running a deferrable load to the time periods when the price is known to be low. For a low job arrival probability, the “always run” policy achieves a lower power consumption cost than that of the “always defer” policy. The reason is that the load is deferred, it has to consume some energy during a wait state. On the other hand, if the load is run immediately, there is no energy consumption during a wait state. However, when the job arrival probability is large (i.e., more than 0.08), the “always defer” policy achieves the lowest cost, which is even lower than that of the optimal policy with price information. However, the “always defer” policy suffers from the long delay of a job (Fig. 3(b)). Also, since the probability of full queue for the “always defer” policy is high, many jobs will not be accepted by the deferrable load, which mainly contributes towards the lowest power consumption. However, these behaviors of the “always defer” policy are undesirable from the user’s perspective. The optimal policy with price information can maintain the delay of a job at the target level (i.e., 170 minutes). Although the “always run” policy achieves a small delay, the “always run” policy consumes more power than that of the optimal policy with price information. We observe that without price information, the optimal policy acts similar to that of the “always run” policy, whose power consumption cost is much higher than that of optimal policy with price information.

In summary, with the power consumption optimization, the optimal scheduling policy of a deferrable load can be obtained such that the power consumption cost is minimized while successfully maintaining the delay at the acceptable level. However, the optimal scheduling policy will be effective only when the price information is available. If the price information is not available (e.g., due to the failure of a

NAN gateway), the optimization of deferrable load scheduling becomes suboptimal.

Now, we evaluate the availability performance metric and its impact on the power consumption cost of a deferrable load. Fig. 4(a) shows the availability of a NAN gateway (i.e., probability that a NAN gateway will be in an “available” state) for different mean time between failure (MTBF) of a NAN gateway. Without redundancy, if a NAN gateway fails, the connection will become unavailable (i.e., a deferrable load will be disconnected and unable to obtain the price information). On the other hand, with redundancy, if one NAN gateway fails, the deferrable load can still obtain the price information through the redundant NAN gateway. As expected, when the MTBF of a NAN gateway increases (given that the failure happens randomly), the availability increases. Without redundancy, the availability increases noticeably. With redundancy, the availability is almost one. Therefore, as the MTBF increases, the increase in the availability and the decrease in the cost are marginal. This result shows that increasing the redundancy can have a more significant effect on improving the availability of the smart grid data communications infrastructure when compared to that due to increasing the MTBF. The availability is important for the power consumption cost. Without price information, the deferrable load cannot be optimally scheduled. As a result, the power consumption cost increases as the MTBF decreases (Fig. 4(b)).

2) *Public utility side – Power supply optimization:* Next, we consider the power supply optimization of the public utility. The public utility can obtain the power demand information from the power consumption units through a connection of HAN and NAN gateways. We consider 100 DERs and the capacity of each DER is 1kWh. We assume that the DERs belong to the public utility. The probability of generating power is 0.7. The penalty of under-supplied power is 100 cents per kWh. There is one generator selling power with price 8 cents per kWh. There are 200 deferrable loads and the probability for each deferrable load to consume power of 1 kWh and 0 kWh (e.g., when the scheduler of a deferrable load

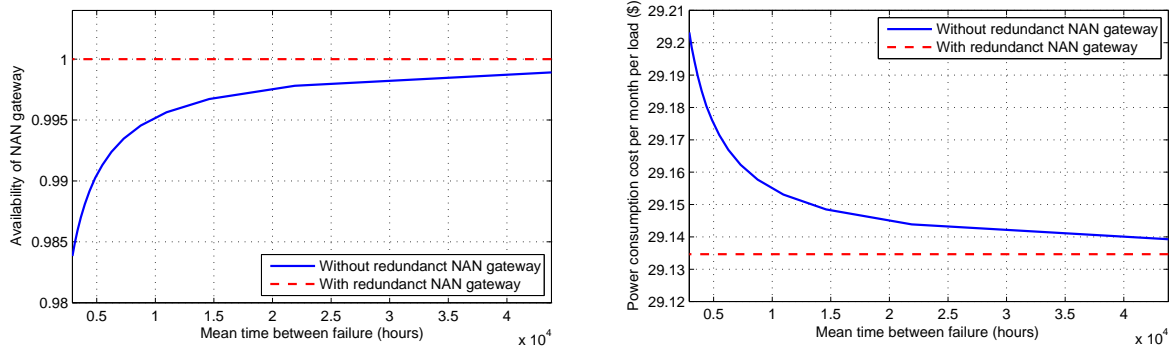


Fig. 4: (a) Availability of a NAN gateway, and (b) power consumption cost of a deferrable load under different mean time between failure (MTBF) of a NAN gateway.

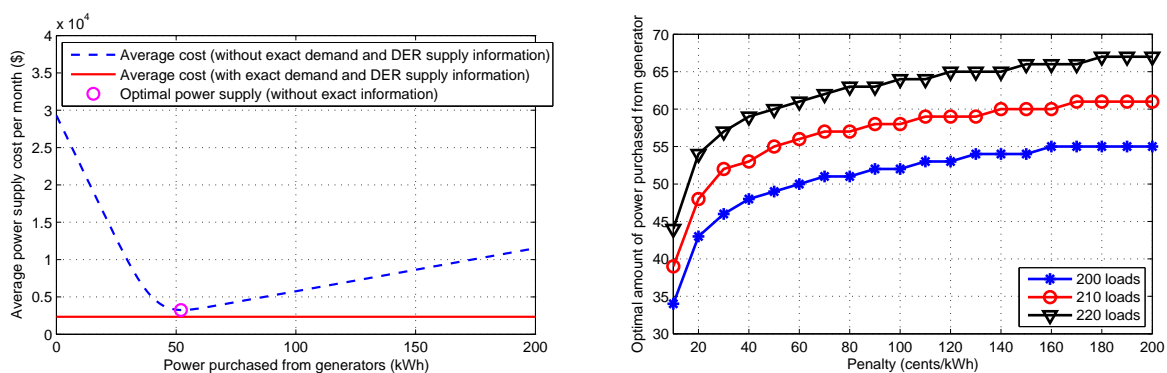


Fig. 5: (a) Average power supply cost of a public utility under amount of power purchased from generators, and (b) optimal solution of the stochastic programming model under different penalty of under-supplied power.

takes actions “run” and “defer”) is 0.45 and 0.55, respectively.

Fig. 5(a) shows the average power supply cost of the public utility in two cases, i.e., with and without exact demand and DER supply information. With demand and DER supply information, the public utility can purchase power from the generator to meet the demand. As a result, this is the optimal solution and the average power supply cost is minimized. Note here that the average power cost obtained with exact demand and DER supply information is shown in Fig. 5(a) as a reference point and it does not vary with the value of power purchased from the generators (shown in the X-axis). Without demand and DER supply information, the public utility estimates the demand and DER supply distributions. Fig. 5(a) also shows the power supply cost when the amount of power purchased from the generator is varied. First, the power supply cost decreases since the cost due to penalty of under-supplied power decreases. However, at a certain point, the power supply cost increases since the power purchased from the generator is not fully utilized. We observe that there is an optimal point of the power purchased from the generator such that the power supply cost of the public utility is minimized. Given the demand and DER supply distributions, this optimal power can be obtained by solving the stochastic programming (SP) model. Note that, the gap between the optimal power supply cost obtained without exact demand and DER supply

information and the cost obtained with such information, is due to the unavailability of the smart grid data communications infrastructure.

Fig. 5(b) shows the optimal amount of power purchased from the generator without exact demand and DER supply information. The optimal amount of power is obtained by solving the SP model. The penalty of under-supplied power is varied. As expected, when the penalty increases, the public utility tends to purchase more power from the generator to minimize the cost. However, as the penalty increases, the amount of power purchased from the generator becomes saturated due to the fact that the purchased power may not be utilized, resulting in the over-supplied power.

Fig. 6 shows the power supply cost of the public utility under mean time between failure (MTBF) of a NAN gateway. Similar to the results for power consumption cost shown in Fig. 4(b), the power supply cost decreases as the MTBF increases (i.e., higher availability). With a redundant NAN gateway, the communication reliability among loads and DERs increases significantly and hence the power supply cost reduces significantly.

Fig. 7(a) shows the power supply cost of the public utility under different number of available DERs. Clearly, with more DERs, the power supply of the public utility decreases. Also, with exact demand and DER supply information, the power

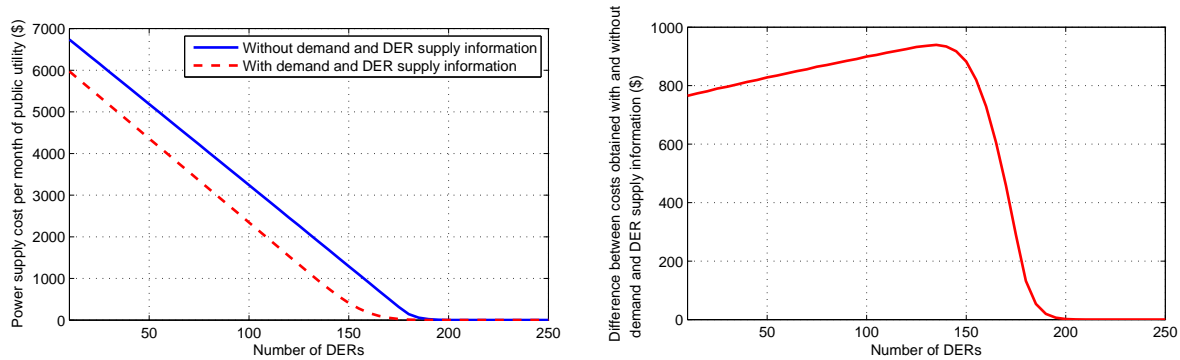


Fig. 7: (a) Power supply cost of a public utility with and without exact demand and DER supply information and (b) their difference, when the number of DERs is varied.

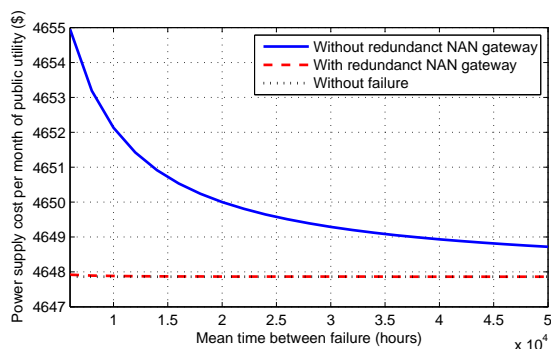


Fig. 6: Power supply cost of a public utility under different mean time between failure of a NAN gateway.

supply cost is smaller than that without such information. Fig. 7(b) shows the difference between the power supply costs with and without exact demand and DER supply information. Interestingly, when the number of DERs increases, the difference first increases and then it decreases. This is due to the fact that when there are more DERs, the exact demand and DER supply information can help the public utility to better purchase the power. As a result, the power supply cost decreases at a faster rate than that without exact demand and DER supply information. However, when the power supply from DERs is enough to accommodate the demand, the power supplies of both cases approach zero, and the difference also decreases to zero.

## VII. CONCLUSION

We have analyzed the impact of the reliability of the smart grid data communications infrastructure on the optimizations of power consumption and supply. For the power consumption optimization, the deferrable load scheduling requires price information from the meter data management system (MDMS) server. If the connection of home area network (HAN) and neighborhood area network (NAN) gateways from deferrable load to the MDMS server is unavailable, the scheduling policy will be suboptimal and it will result in a higher power consumption cost. Likewise, for the power supply optimization, the economic dispatch of the public utility requires

exact demand and distributed energy resource (DER) supply information to purchase the power from generators. If the connection of HAN and NAN gateways is unavailable, the public utility has to use only the statistical information to optimize the power purchasing, which results in a higher power supply cost.

Performance evaluation has been performed based on the analysis of the power consumption and supply optimizations with complete and incomplete information. Some interesting results have been discussed. For example, the difference between power supply costs obtained with and without exact demand and DER supply information may not be monotonic to the number of DERs. In addition, the redundancy of the NAN gateway has been evaluated. With the redundant NAN gateways, the power consumption and supply costs incurred to the consumer and public utility can be significantly reduced.

In our future work, the reliability analysis of the smart grid data communications infrastructure will be extended to include the security attacks, and their impacts on the cost will be analyzed.

## REFERENCES

- [1] S. Stoft, *Power System Economics: Designing Markets for Electricity*, Wiley-IEEE Press, May 2002.
- [2] T. T. Kim and H. V. Poor, "Scheduling power consumption with price uncertainty," *IEEE Transactions on Smart Grid*, vol. 2, no. 3, pp. 519-527, September 2011.
- [3] M. He, S. Murugesan, and J. Zhang, "A Markov decision process approach to multi-timescale scheduling and pricing in smart grids with integrated wind generation," in *Proceedings of IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pp. 125-128, December 2011.
- [4] M. Pedrasa, T. D. Spooner, and I. F. MacGill, "Scheduling of demand side resources using binary particle swarm optimization," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1173-1181, Aug. 2009.
- [5] A. J. Conejo, J. M. Morales, and L. Baringo, "Real-time demand response model," *IEEE Transactions on Smart Grid*, vol. 1, no. 3, pp. 236-242, Dec. 2010.
- [6] W. Shi and V. W. S. Wong, "Real-time vehicle-to-grid control algorithm under price uncertainty," in *Proceedings of IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pp. 261-266, October 2011.
- [7] C. Dai, W. Chen, Y. Zhu, and X. Zhang, "Seeker optimization algorithm for optimal reactive power dispatch," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1218-1231, Aug. 2009.
- [8] L. Wu, M. Shahidehpour, and T. Li, "Cost of reliability analysis based on stochastic unit commitment," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1364-1374, Aug. 2008.

- [9] J. Hetzer, D. C. Yu, and K. Bhattacharai, "An economic dispatch model incorporating wind power," *IEEE Transactions on Energy Conversion*, vol. 23, no. 2, pp. 603-611, June 2008.
- [10] X. Liu and W. Xu, "Minimum emission dispatch constrained by stochastic wind power availability and cost," *IEEE Transactions on Power Systems*, vol. 25, no. 3, pp. 1705-1713, Aug. 2010.
- [11] S. Dharmaraja, V. Jindal, and U. Varshney, "Reliability and survivability analysis for UMTS networks: An analytical approach," *IEEE Transactions on Network and Service Management*, vol. 5, no. 3, pp. 132-142, September 2008.
- [12] W. Hneiti and N. Ajlouni, "Dependability analysis of wireless local area networks," in *Proceedings of International Conference on Information and Communication Technologies (ICTTA)*, pp. 2416-2422, April 2006.
- [13] H. Wen, C. Lin, F. Ren, Y. Yue, and X. Huang, "Retransmission or redundancy: Transmission reliability in wireless sensor networks," in *Proceedings of IEEE International Conference on Mobile Adhoc and Sensor Systems (MASS)*, Oct. 2007.
- [14] A. G. Bruce, "Reliability analysis of electric utility SCADA systems," *IEEE Transactions on Power Systems*, vol. 13, no. 3, pp. 844-849, August 1998.
- [15] Z. Xie, G. Manimaran, V. Vittal, A. G. Phadke, and V. Centeno, "An information architecture for future power systems and its reliability analysis," *IEEE Transactions on Power Systems*, vol. 17, no. 3, pp. 857-863, August 2002.
- [16] Y. Wang, W. Li, and J. Lu, "Reliability analysis of wide-area measurement system," *IEEE Transactions on Power Delivery*, vol. 25, no. 3, pp. 1483-1491, July 2010.
- [17] L. Wen-xia, L. Nian, F. Yong-feng, Z. Li-xin, and Z. Xin, "Reliability analysis of wide area measurement system based on the centralized distributed model," in *Proceedings of Power Systems Conference and Exposition (PSCE)*, pp. 1-6, March 2009.
- [18] J. Zhou, R. Hu, and Y. Qian, "Scalable distributed communication architectures to support advanced metering infrastructure in smart grid," *IEEE Transactions on Parallel and Distributed Systems*, in press.
- [19] Z. Md. Fadlullah, M. Fouda, N. Kato, A. Takeuchi, N. Iwasaki, and Y. Nozaki, "Towards intelligent machine-to-machine communications in smart grid," *IEEE Communications Magazine*, vol. 49, no. 4, pp. 60-65, Apr. 2011.
- [20] M. Fouda, Z. Md. Fadlullah, N. Kato, R. Lu, and X. Shen, "A lightweight message authentication scheme for smart grid communications," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 675-685, Dec. 2011.
- [21] Y. Yan, Y. Qian, H. Sharif, and D. Tipper, "A survey on smart grid communication infrastructures: Motivations, requirements and challenges," *IEEE Communications Surveys & Tutorials*, in press.
- [22] F. Bouhaf, M. Mackay, and M. Merabti, "Links to the future: Communication requirements and challenges in the smart grid," *IEEE Power and Energy Magazine*, vol. 10, no. 1, pp. 24-32, Jan.-Feb. 2012.
- [23] E. E. Lewis, *Introduction to Reliability Engineering*, Wiley, November 1995.
- [24] S. Borenstein, "The long-run efficiency of real-time electricity pricing," *The Energy Journal*, vol. 26, no. 3, pp. 93-116, 2005.
- [25] S. P. Holland and E. T. Mansur, "The short-run effects of time-varying prices in competitive electricity markets," *The Energy Journal*, vol. 27, no. 4, pp. 127-155, 2006.
- [26] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, Wiley-Interscience, April 1994.
- [27] J. R. Birge and F. Louveaux, *Introduction to Stochastic Programming*, Springer, June 2011.
- [28] W. Charytoniuk, M. S. Chen, P. Kotas, and P. Van Olinda, "Demand forecasting in power distribution systems using nonparametric probability density estimation," *IEEE Transactions on Power Systems*, vol. 14, no. 4, pp. 1200-1206, Nov 1999.
- [29] S. Minasiewicz, G. S. Hope, and O. P. Malik, "Probability distributions of active power demand a design approach," *IEEE Transactions on Power Delivery*, vol. 1, no. 3, pp. 197-202, July 1986.

#### ACKNOWLEDGMENT

This work was supported by a grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada. This work was done in the Centre for Multimedia and Network Technology (CeMNet) of the School of Computer Engineering, Nanyang Technological University, Singapore.



**Dusit Niyato** (M'09) is currently an assistant professor in the School of Computer Engineering, at the Nanyang Technological University, Singapore. He received B.E. from King Mongkut's Institute of Technology Ladkrabang (KMITL) in 1999. He obtained his Ph.D. in Electrical and Computer Engineering from the University of Manitoba, Canada in 2008. His research interests are in the area of radio resource management in cognitive radio networks and broadband wireless access networks, and smart grid communications and networking.



**Qiumin Dong** is a Ph.D. student in the School of Computer Engineering, Nanyang Technological University (NTU), Singapore. He received the B.Eng. degree in Electronic and Information Engineering from Huazhong University of Science and Technology in 2008. His research interests are in the area of smart grid communications and networking, and dynamic spectrum access in cognitive radio networks.



**Ping Wang** received the BE and ME degrees from Huazhong University of Science and Technology, Wuhan, China, in 1994 and 1997, respectively, and the Ph.D. degree from University of Waterloo, Ontario, Canada, in 2008, all in electrical engineering. Since June 2008, she has been an assistant professor in the School of Computer Engineering, Nanyang Technological University, Singapore. Her current research interests include QoS provisioning and resource allocation in multimedia wireless communications. She was a co-recipient of the Best Paper Award from *IEEE Wireless Communications and Networking Conference (WCNC) 2012* and *IEEE International Conference on Communications (ICC) 2007*. She is a member of the IEEE. She is an Editor of *IEEE Transactions on Wireless Communications*, *EURASIP Journal on Wireless Communications and Networking*, and *International Journal of Ultra Wideband Communications and Systems*.



**Ekram Hossain** (S'98-M'01-SM'06) is a Full Professor in the Department of Electrical and Computer Engineering at University of Manitoba, Winnipeg, Canada. He received his Ph.D. in Electrical Engineering from University of Victoria, Canada, in 2001. Dr. Hossain's current research interests include design, analysis, and optimization of wireless/mobile communications networks, cognitive radio systems, smart grid communications and networking, and network economics. He has authored/edited several books in these areas (<http://www.ee.umanitoba.ca/ekram>). Dr. Hossain serves as the Editor-in-Chief of the *IEEE Communications Surveys and Tutorials* (for the term 2012-2013) and an Editor of *IEEE Wireless Communications*. Previously, he served as the Area Editor of the *IEEE Transactions on Wireless Communications* in the area of "Resource Management and Multiple Access" from 2009-2011 and an Editor of the *IEEE Transactions on Mobile Computing* from 2007-2012. Dr. Hossain has won several research awards including the University of Manitoba Merit Award in 2010 (for Research and Scholarly Activities), the 2011 IEEE Communications Society Fred Ellersick Prize Paper Award, and the *IEEE Wireless Communications and Networking Conference 2012 (WCNC'12)* Best Paper Award. He is a registered Professional Engineer in the province of Manitoba, Canada.